



ELEVENTH INTERNATIONAL CONFERENCE ON
BUSINESS ANALYTICS AND INTELLIGENCE

9 – 11 DECEMBER, 2024

BOOK OF
ABSTRACTS

**11TH INTERNATIONAL CONFERENCE ON BUSINESS ANALYTICS AND INTELLIGENCE
9-11 DECEMBER, 2024**

Table of Contents

Sl.No	ABSTRACT TITLE	Page No
1	LLM Hallucination Detection with Guardrails to RAG Pipeline	2
2	Unveiling the Private Labels Transition in India's FMCG Market: A Multi-Stage Analysis with Switch Regression	3
3	AI Spends - An approach for organizations to identify and optimize their AI Spends	4
4	Customer Churn Prediction in Telecom Business	6
5	Intelligent analytical framework to reduce customer retention efforts in the SaaS industry	7
6	Open World Classification using Projected Gradient Descent (PGD)	8
7	Improving CNN Classification using Fractal Geometry Based Network and Residual Connections	9
8	Predicting the Stochastic Volatility and Spillover effect of Foreign Currencies. An integrated approach using the LSTM Attention mechanism and BEKK GARCH	10
9	Hyper-Personalized Enterprise-Level Product Recommendations for HPE's eCommerce Store	11
10	Refining Marketing Mix Modeling: The Power of Hierarchical Mixed-Effects Models in Retail Chains	12
11	Adaptive Contrastive Search for Text Generation	13
12	Enhancing Intelligent Document Processing & RAG with Ant Colony Optimization and Gen AI Synergy	14
13	Optimizing Medical Equipment's Information Retrieval: The Role of Knowledge Graphs and Advanced Re-rankers	15
14	Video RAG: Leveraging Semantic Chunking and Retrieval-Augmented Generation for Enhanced Information Extraction	16
15	Analytics for Enhancing Employability of MBA Graduates Using Predictive Modeling	17
16	Sentence Level Correction of Common Spoken Grammatical Errors with Deep Neural Networks	18
17	Addressing Emerging Risks in LEVs: A Novel Framework for Insurance Pricing Using Transfer Learning and Telematics	19
18	An approach towards handling LLMs Hallucinations Using LLM ensemble	20
19	SaaS Cloud License Management - Predictive Cost optimization	21
20	A Study on Predicting Interest Rate Movement of Banks	22
21	Detection of Protective Gear Compliance in Mine and Factory Settings Using Convolutional Vision Transformers	23
22	A Prescriptive Model for the Depot Locations and Allocation of City Buses to Depots	24
23	Enhancing On-Time Loan Repayment and Customer Retention in Microfinance Loans through Machine Learning Algorithms	25

Sl.No	ABSTRACT TITLE	Page No
24	Steering E-commerce Success through Next Basket Recommendation System	26
25	Leveraging Machine Learning to Enhance Last-Mile Delivery Efficiency: The Impact of Agent Experience, Environmental Factors, and Geographical Distance	27
26	Impact of shared account usage on personalisation algorithms	28
27	Making meeting sense	29
28	Automotive Sales and Service Recommendation with Singular Value Decomposition and Multi-Label Classification Models	30
29	Aspect-Based call transcript Summarization: Generation & Validation	31
30	Democratizing Data Insights: Empowering Business Analytics through Generative AI	32
31	Analyzing and Mitigating Prompt Injection Techniques in AI Language Models	33
32	Transforming Video Content into Dynamic Multimodal SOPs	34
33	A Deep Dive into RAG Paradigms: Selecting the Right Approach for Specific Business Applications	35
34	Predicting Dynamic Mobile Phone Price using Reinforcement Learning: Helping Customers Under Price Uncertainty	36
35	From Decision Trees to Boosting: A Study of Classification Algorithms for BNPL Credit Risk	37
36	Transforming Airline Service Quality: Harnessing User-Generated Content and Data Mining for Deep Customer Insights	38
37	Performance Evaluation using Data Envelopment Analysis	39
38	A Fusion of AI and Human Expertise for Enhanced Code Generation and Debugging	40
39	Revolutionizing Contract Management with Generative AI: Uncovering Loopholes with Precision and Speed	41
40	Key Phrase Extraction From Chairmans Speech using Transformers	42
41	Benchmarking Advanced Video Generation Models: Evaluating Effectiveness in Creative Media and Corporate Communication	43
42	Enhancing Analysis of Reviews with Deep Learning and Explainable AI for Selecting Hotels on Specific Aspects	44
43	Leveraging Workplace It and Behavioural Analytics for Customized L&D Plans	45
44	Audio2Action: A Framework for Automated Task & Insight Generation from Speech	46
45	Federated Learning for Complaint Classification: Navigating Non-IID Textual Data Challenges and Ensuring Privacy and Explainability	48
46	Integration of Chatbot Based Non-Medical Services in Hospitals	50
47	Enhancing Customer Service With Real-Time Feedback and Verification System	51
48	Assessing Bayesian and Multi-Fidelity Hyperparameter Optimization Strategies for Several Machine Learning Models	53
49	Efficient PII Removal with Zero-Shot NER and Dual Encoders	54
50	Survey of Large Language Models : Capability, Efficiency and Security	55
51	Evaluation and Prediction of Sustainable Livelihood Capital - A Case Study	56

Sl.No	ABSTRACT TITLE	Page No
52	Pedagogical Effectiveness impacting Learning Outcome - An Analysis with Formative Assessment	57
53	Business analytics and firm performance: A hybrid review of existing literature	58
54	Analysing Barriers to AI and Automation Adoption in Retail: A Grey-DEMATEL Approach	59
55	Generating Actionable Insights from Ticket Data for IT Operations	60
56	Profiling Customers using RFM Modelling for Credit Card holders using k-means clustering	62
57	Evaluating Consistency: Sentiment In Annual Reports VS Financial Performance	63
58	Revolutionizing Cybersecurity with Generative AI: Adaptive Systems for Non-Compliance & Threat Detection	64
59	Conformal Prediction Approaches for Accurate Carbon Footprint Forecasting	65
60	Predicting the Performance of Sites for Solar Power Plant	66
61	Factors Influencing the NIRF Score of Indian Management Institutions	67
62	Spillover Network Analysis in the APAC region using Minimum Spanning Tree	68
63	Asset Growth and Investor Preferences Across Firm Life Cycles: Indian Evidence	69
64	Transformative AI for OPD - Enhancing efficiency and productivity	70
65	Classifying and diagnosing ASD through Deep learning techniques using brain EEG signals and MRI images	71
66	Consumer Adoption Intention Towards Electric Vehicles: Insights from India	73
67	Role Of Prospect Theory In Insurance Company's Investment: A Behavioural Finance Study in Karnataka	74
68	The emergence of ratings of Amazon's reviews: From pooling to separating equilibrium	75
69	Enhancing Dispute Validation in Customer Service through Fine-Tuned Large Language Models and Advanced Embedding Techniques	76
70	Optimizing Multimodal Ai Performance: A Comparative Analysis Of Prompting Techniques Across GPT-4.0, Gemini, and Claude	77
71	Does Brand Equity Exist in Banking Sector? A Discriminant Analysis	78
72	Smart Analytics For Energy Savings in Telecom Network	79
73	A Novel Approach for Evaluating Performance Consistency in Time Series Forecasting Models	80
74	Transforming Highway Monitoring: A Study of traditional methods of monitoring large- scale infrastructure like highways and the possibility of application of emerging technologies like drones and artificial intelligence in conjunction with GIS approaches for a Drone Analytics Management System (DAMS)	81
75	Large Language Models for PL/SQL Code Explanations: A Comprehensive Exploration and Evaluation	83
76	Adapting Agricultural Water Management Practises using Artificial Intelligence for Climate Resilience and Sustainability of Virudhunagar District of Tamilnadu	84

Sl.No	ABSTRACT TITLE	Page No
77	Aligning CSR with SDGs: A Comprehensive Sustainability Reporting Tool for Indian Enterprises	85
78	Comparative Evaluation of Advanced AI Models and Services for Complex Tabular Data Extraction from Documents	87
79	Stock Forecasting and Sector-Specific Recommendation using Deep Learning	88
80	Automated Content Moderation Using AI: An Agentic Approach to Process Text, Audio, Video, and Images	89
81	Advancing Sentiment Analysis with Long Short-Term Memory (LSTM) Networks	90
82	Optimizing Fare Structures Using Shortest Path Algorithms and Real-time Adjustments in Urban Metro Networks	91
83	Approach, application and implementation strategy for a 3D Digital Twin, an Artificial Intelligence-based transformative tool	92
84	RoundTable: Leveraging Dynamic Schema and Contextual Autocomplete for Enhanced Query Precision in Tabular Question Answering	94
85	Exploring Ethical Paradigms in Digital Marketing: A Bibliometric Approach	95
86	Automated Extraction of Startup Performance Data from Pitch Decks	96
87	Analysing YouTube Advertising of Indian Companies with Focus on Celebrity Endorsements	97
88	Retail Analytics: Artificial Intelligence based Propensity Models	98
89	Advancing Engineering Drawing Interpretation: A Comparative Study of Prompting Techniques	99
90	Sustainable Distribution Planning- An Optimization Approach	100
91	Cloud Threat Hunting	101
92	Conversational AI for Enhanced Customer Journeys	102
93	Enhancing Airline Operations using ARIMA: Passenger Load & OTP Evaluation	103
94	ChatGPT's Influence on Student Learning: An Exploratory Study Using Sentiment Analysis	104
95	Harnessing Data Labs for the Evolution and Growth of Metro Rail Systems	105
96	Volatility forecasting of BSE Private Bank Index using ARIMA and GARCH models	106
97	Exploring Thematic Patterns in 35 News Articles for Coffee badging: A Detailed Analysis using Python	107
98	An Exploratory Study of Stock Price Movements from Earnings Call Transcripts: A case of Indian Capital Market	108
99	Building Resilience through ESG: Evaluating Indian PSUs in the Context of Global Emerging Markets	109
100	User Generated Content Influencing homestay business in Karnataka	110
101	Dynamic cricket match predictions through LSTM based sequential data modeling	111
102	Implementation of a Unified Data Lake Platform: An End-to-End Framework for Data, BI and Visualization Practice enabling Data-Driven Decision Making	112

Sl.No	ABSTRACT TITLE	Page No
103	Data Driven Decision Making Based Operation Research in Supply Chain Management	113
104	Measuring the Influence of Message and Content on NBRx Using Multi-Touch Attribution: A Healthcare Analytics Approach	114
105	Machine Learning-Based Prediction of Gross Enrolment Ratio in Higher Education Institutes in India	115
106	Decoding Economic Success in Indian Cinema: A Predictive Explainable AI Approach	116
107	Detecting Scams in Real-Time: A Machine Learning Approach to Identify Fraudulent Calls and Messages	117
108	Image-Based Content Generation with Cultural Variation: A Dual Attention and Fuzzy Logic Framework	118
109	Integrated Traffic Management with Incident Prevention and Detection in Real-Time	119
110	Comparative Analysis of Machine Learning Models for Optimized Detection and Personalization in Mental Health	120
111	Efficient Deep Learning Techniques for Text Categorization in Resource-Constrained Environments	121
112	Optimizing Zomato Deliveries: A Comparative Analysis of Linear Regression, Decision Trees, and Random Forest	122
113	Optichanneling Omnichannel Retail Strategies by Integrating A IoT	123
114	Harnessing Predictive Analytics and Consumer Behavior Data for Co-Creation in Digital Commerce: A Path to Building Stronger Brand Identity	124
115	Analyzing Customer Preference and Satisfaction - A Study on Key Attributes Influencing Apple iPhone Purchase	125
116	Impact of Telehealth Adoption on Hospital Demand Forecasting	126
117	Smart and Sustainable Packaging for Fruit Logistics	127
118	The emergence of ratings of Amazon's reviews: From pooling to separating equilibrium	128
119	Hyper-Personalized Enterprise-Level Product Recommendations for HPE's eCommerce Store	129
120	Impact of Macroeconomic Determinants on the Performance of the Indian Stock Market after COVID-19: A Multivariate Time Series Approach	130



ELEVENTH INTERNATIONAL CONFERENCE ON
BUSINESS ANALYTICS AND INTELLIGENCE

9 – 11 DECEMBER, 2024

BOOK OF
ABSTRACTS

LLM HALLUCINATION DETECTION WITH GUARDRAILS TO RAG PIPELINE

Vaibhav Mishra

Data Scientist

Genpact

Bengaluru, India

Vaibhav.mishra@genpact.com

Vipul Sharma

Data Scientist

Genpact

Bengaluru, India

Vipul.sharma2@genpact.com

Arjun Seth

Data Scientist

Genpact

Bengaluru, India

arjun.seth1@genpact.com

Tavva Prudhvith

Data Scientist

Genpact

Bengaluru, India

tavva.prudhvith@genpact.com

Prakash Selvakumar

Assistant Vice President

Genpact

Bengaluru, India

prakash.selvakumar@genpact.com

Large Language Models (LLMs) have become increasingly more powerful and capable with time. By generating human-like text and understanding complex language patterns, they assist with tasks ranging from chatbots to content creation, revolutionizing how we interact with information. While they are very capable, they also face a significant challenge: hallucination. Hallucination refers to situations where LLMs generate factually inaccurate or misleading responses, veering off-the-rails from the truth. This research paper presents a comprehensive framework to address this issue, ensuring LLMs provide reliable and accurate information.

The solution is twofold: implementing guardrails to keep LLMs on track and developing effective methods for hallucination detection. Two approaches are developed for the guard railing process and two more approaches are explored for the hallucination detection task. Each module will be explored in this research paper.

By combining the above methods, the framework ensures that responses are factually correct and thoroughly vetted through hallucination detection checks. Making it unique, robust and novel, increasing the reliability of LLM-powered applications.

Keywords: Large Language Model (LLM), Hallucination, guardrails, Retrieval Augmented Generation (RAG), Finetuning, Knowledge Graphs.

UNVEILING THE PRIVATE LABELS TRANSITION IN INDIA'S FMCG MARKET: A MULTI-STAGE ANALYSIS WITH SWITCH REGRESSION

Baba Gnanakumar P

School of Management

Kristu Jayanti College

Bengaluru, India

gnanakumar12000@yahoo.com

ORCID IDM 0000-0002-6618-1532

Despite their global dominance, private labels hold a surprisingly modest share of the Indian fast-moving consumer goods (FMCG) market. This discrepancy, with a global average share of 16.3% compared to India's less than 1%, presents a fascinating research opportunity. This study delves into the reasons behind this low penetration and investigates the potential for future growth of private labels in the Indian FMCG landscape. The research has a two-pronged approach: understanding consumer loyalty drivers and the interplay between price sensitivity and quality perception. Firstly, it explores the various loyalty-building tactics employed by private label companies. This initial stage aims to identify key factors influencing consumer decisions through factor analysis. The commitment analysis utilizes a dataset encompassing 12 variables related to store performance and customer behaviour, gleaned from a survey of 502 validated customer responses. By employing factor analysis, the research seeks to identify underlying relationships among these variables and categorize them into distinct factors. The Kaiser-Meyer-Olkin (KMO) Measure of Sampling Adequacy test and Bartlett's Test of Sphericity were conducted to ensure the data's suitability for this technique. The second stage delves into the "Benjamin Franklin effect" in the context of private labels. This effect explores how retailers can foster a transition from cognitive dissonance (psychological discomfort) to cognitive consonance (psychological comfort) in customers. By leveraging this effect, retailers can strengthen customer relationships with private labels and cultivate long-term loyalty. Finally, the research employs switch regression analysis. This method is particularly useful when the relationship between variables undergoes distinct changes across different conditions. In this context, switch regression is used to examine the nuanced interplay between price sensitivity and the four key factors identified in the obviously first stage. By analyzing switching behaviour between private labels and national brands, the research sheds light on how price sensitivity interacts with customer behaviour and store performance. The research findings offer valuable insights for private label companies in the Indian market. The analysis of loyalty drivers highlights the importance of fostering loyalty value through targeted strategies and emphasizes the need for continuous product innovation, especially for price-sensitive customers. Additionally, the research underscores the significance of developing effective customer retention strategies and maintaining a focus on competitor monitoring and adapting pricing tactics. Furthermore, the price sensitivity analysis, conducted using switch regression, reveals the need for nuanced pricing strategies that go beyond just cost considerations. It also reinforces the enduring power of loyalty programs in influencing consumer behaviour. By implementing these insights, private label companies can cultivate stronger customer loyalty, optimize pricing strategies, and ultimately achieve long-term success in the Indian FMCG market. This research paves the way for a more comprehensive understanding of consumer behaviour towards private labels and equips companies with the knowledge required to navigate the dynamic Indian retail landscape.

Keywords: Private labels, Customer loyalty, Switch Regression

AI SPENDS - AN APPROACH FOR ORGANIZATIONS TO IDENTIFY AND OPTIMIZE THEIR AI SPENDS

Ravi Krishna

AI CoE, Engineering
Flexera Software Solutions LLP
Bengaluru, India
rbettegowda@flexera.com

Amit Sharma

AI CoE, Engineering
Flexera Software Solutions LLP
Bengaluru, India
asharma@flexera.com

Barath Ravichander

AI CoE, Engineering
Flexera Software Solutions LLP
Bengaluru, India
bravichander@flexera.com

Priya Srivatsav

FinOps, Product Management
Flexera Software Solutions LLP
Bengaluru, India
psrivastav@flexera.com

Sunil Kumar

FinOps, Engineering
Flexera Software Solutions LLP
Bengaluru, India
skumar2@flexera.com

Niladri Ray

VP, Engineering
Flexera Software Solutions LLP
Bengaluru, India
nray@flexera.com

As organizations increasingly invest in Artificial Intelligence (AI) technologies, the need to effectively manage and optimize AI expenditures has become a pressing concern. AI Spends is a novel approach that addresses this challenge by providing a comprehensive and structured methodology for analyzing, categorizing, and evaluating AI investments. Our framework empowers organizations to identify areas of inefficiency, measure the effectiveness of AI initiatives, and make informed decisions about future investments. By leveraging AI Spends, organizations can:

- * Gain visibility into AI spending patterns across departments and teams
- * Identify and eliminate duplicated efforts and waste
- * Optimize resource allocation and reduce costs
- * Improve AI adoption rates and ROI
- * Enhance the overall effectiveness of AI initiatives inside an organization

Our approach is grounded in a data-driven methodology that incorporates machine learning, data analytics, and expert knowledge. We demonstrate the efficacy of AI Spends has been successfully applied in various industries, resulting in significant cost savings and improved AI adoption rates. This paper presents the AI Spends framework, its implementation, and case studies demonstrating its effectiveness in optimizing AI investments.

Keywords: AI investments, AI optimization, data-driven approach, machine learning, cost savings, ROI improvement.

CUSTOMER CHURN PREDICTION IN TELECOM BUSINESS

PraveenaSri Perini

Department of Economics

ICFAI School of Social Sciences

The ICFAI Foundation for Higher Education (IFHE)

Hyderabad, India

praveena.sriperini@ibsindia.org

The prospect of predicting churn in the B2B environment remains unresolved despite the burgeoning availability of electronic customer traces. Nevertheless, its unique features such as a minimum number of customers with absurdly higher transactional values symbolize the relevance of identifying potentially churning customers. This research study aims to investigate the pertinent factor variables that influence the churn rate of Telco Company using a real-world data set with a sample size of 7021 subscriptions. 12 ML Algorithm models such as Logistic Regression, Support Vector, Gradient Boosting, K-Nearest Neighbor, Decision Tree, Bagging Classifier, Random Forest, AdaBoost, Gaussian Naïve Bayes, Extra Trees, and Multi-Layer Perceptron Classifier are arrayed to a huge dataset to recognize the varying patterns of churn rates thereby leveraging the learning curve of the Business. To assess the justification of the archetypal, the F1 and AUC were espoused where a score of 91 percent and 85 percent were achieved. These business analytical models aided in effective managerial decision-making regarding customer retention strategies.

Keywords: Customer Churn, Telco Company, B2BEnvironment, Feature Selection

INTELLIGENT ANALYTICAL FRAMEWORK TO REDUCE CUSTOMER RETENTION EFFORTS IN THE SAAS INDUSTRY

V Sasikiran Angara and KS Manu

Software as a Service, or SaaS, is a business model that uses the internet to offer software applications as a service instead of installing them locally on devices. In the software as a service (SaaS) sector, churn is a crucial indicator as it directly affects a business's earnings, prospects for expansion, and viability over time. Because SaaS companies mostly rely on recurring income from subscriptions, high churn rates can be detrimental to their operations. Customer retention is crucial for SaaS companies as it is frequently more profitable and cost-effective than bringing on new customers. Retention expenses and efforts can be decreased by focusing on an appropriate set of customers. This study focuses on an intelligent analytical framework that uses machine learning and artificial intelligence techniques to find the ideal group of customers for a SaaS-based organization to retain. In the proposed methodology, feature selection models will be used to identify the most significant drivers of churn, the data will be split into train and test sets and the most popular predictive models like logistic regression, random forest, support vector machine, and neural networks will be applied to the train set. The confusion matrix will be used to assess the outcomes of the test set. The classification models calculate the likelihood of churn. The Kaplan-Meier estimate will be used for survival analysis to determine the odds of survival based on tenure. Lastly, the prediction models' interpretability will be enhanced by using explainable AI models like SHAP (SHapley Additive exPlanations) and LIME (Local interpretable model-agnostic explanations). This analytical framework provides the probability of churn and the likelihood of survival considering the time factor. It also describes the most important features that are considered by the model to arrive at the probability. This enables the company to segment the data based on the probabilities of churn and survival, and the feature importance and respective retention strategies can be planned for each segment. By implementing the suggested analytical methodology, the business may determine which customers are most important to target with customer retention strategies.

Keywords: SaaS, Churn analysis, Machine Learning Predictive Models, Survival Analysis, Explainable AI

OPEN WORLD CLASSIFICATION USING PROJECTED GRADIENT DESCENT (PGD)

Lakshay

Data Science & Insights

Genpact

Bangalore, India

lakshay.lakshay@genpact.com

Shalini Mittal

Data Science & Insights

Genpact

Bangalore, India

shalini.mittal@genpact.com

Amol Kumar

Data Science & Insights

Genpact

Bangalore, India

amol.kumar1@genpact.com

Shrey Bhardwaj

Data Science & Insights

Genpact

Bangalore, India

shrey.bhardwaj@genpact.com

Muskan Wadhwa

Data Science & Insights

Genpact

Bangalore, India

muskan.wadhwa@genpact.com

Classification in open world scenarios poses significant challenges for machine learning models due to the diverse nature of data and the potential for encountering previously unseen classes. To encounter this problem, we require negative samples which can act as our unseen class samples for the training process. Generating quality negative samples is the main key for an accurate open-world classification. In this paper, we perform open world classification using Projected Gradient Descent (PGD) negative sampling. This method uses the gradient alteration process to generate negative samples from the known samples. This paper evaluates the accuracy of open-world classification when we use this negative sampling process and compares the classification results while using different types of classifiers and their impact on the open-world classification process. In real world application, the model can be exposed to never seen before data which leads to inaccurate classification and increase in error. These issues can be resolved, and the accuracy of the results can be more precise and accurate using PGD negative sampling.

Keywords: Open world classification, convolutional neural networks, Projected gradient descent, negative sampling

IMPROVING CNN CLASSIFICATION USING FRACTAL GEOMETRY BASED NETWORK AND RESIDUAL CONNECTIONS

Shalini Mittal

Data Science & Insights
Genpact
Bangalore, India

shalini.mittal@genpact.com

Lakshay

Data Science & Insights
Genpact
Bangalore, India

lakshay.lakshay@genpact.com

Amol Kumar

Data Science & Insights
Genpact
Bangalore, India

amol.kumar1@genpact.com

Shrey Bhardwaj

Data Science & Insights
Genpact
Bangalore, India

shrey.bhardwaj@genpact.com

Arjun Seth

Data Science & Insights
Genpact
Bangalore, India

arjun.seth1@genpact.com

In this paper, we introduce a hybrid image classification model inspired by FractalNet [1] and ResNet [2], leveraging fractal patterns to create sequential convolution blocks of user-defined depth with residual connections between alternate fractal blocks. The network can be as shallow or deep as required, balancing the trade-off between response time and accuracy. The fractal block employs self-similar patterns alongside parallel convolutions at different scales to enhance image pattern recognition. Skip connections, inspired by ResNet [2], facilitate deeper network architectures and improve the convergence rate on the CIFAR-10 dataset. By combining the strengths of FractalNet's [1] multi-scale architecture and ResNet's [2] effective training of very deep networks, our model aims to achieve superior performance in image classification tasks. Additionally, extensive data augmentation and hyperparameter tuning are employed to optimize the model's performance, ensuring robustness and generalization. Experimental results demonstrate that our hybrid model outperforms certain traditional architectures on CIFAR-10, providing a promising approach for advanced image recognition systems.

Keywords: Fractal Geometry, ResNet, Convolution Neural Networks, Google-Net

PREDICTING THE STOCHASTIC VOLATILITY AND SPILLOVER EFFECT OF FOREIGN CURRENCIES. AN INTEGRATED APPROACH USING THE LSTM ATTENTION MECHANISM AND BEKK GARCH

Dr T Viswanathan

Associate Professor

Welingkar Institute of Management Development and Research
Bengaluru

Viswanathan.t@welingkar.org

Abhirup Chakraborty

Student Research Assistant

Symbiosis Institute of Business Management, Bengaluru
Bengaluru, India

abhirup.chakraborty25qf@sibm.edu.in

This paper proposes a novel approach for predicting the stochastic volatility and spillover effect of foreign currency pairs. For the purpose of analysis, we utilize a sample of a sample of daily Open, High, Low and Closing (OHLC) prices of 11 pairs of foreign currencies between 2004 and 2024. We employ a hybrid model integrating the BEKK (Baba, Engle, Kraft, and Kroner) GARCH with long short term memory (LSTM) enhanced by a self-attention mechanism. The performance of the LSTM model is tested using metrics such as MSE and RMSE. This integrated approach helped us to predict exchange rates with minimum RMSE and observe the transmission of volatility from lower to higher-indexed markets. The research study offers a profound methodology to the financial institutions, corporates and other stakeholders to forecast currency volatility and spillover effects, thereby effectively managing foreign currency risk.

Keywords: Stochastic Volatility, Volatility spillover, LSTM, Attention Mechanism, BEKK GARCH

HYPER-PERSONALIZED ENTERPRISE-LEVEL PRODUCT RECOMMENDATIONS FOR HPE'S ECOMMERCE STORE

Pankaj Kumar

Global Marketing Hewlett Packard Enterprise
India
pankajk@hpe.com

Dies P Varghese

Global Marketing Hewlett Packard Enterprise
Canada
dies.varghese@hpe.com

Jyoti Khare

Global Marketing Hewlett Packard Enterprise
India
jyoti.khare@hpe.com

Market research highlights the growth of B2C eCommerce as a strategic channel, valued for enhancing customer satisfaction and enabling upsell/cross-sell opportunities. HPE B2C eCommerce is rapidly becoming the preferred choice for buyers of various technology products. As eCommerce grows and buyers increasingly make larger, more complex purchases online, delivering an exceptional digital experience is no longer optional—it's essential. HPE has a clear vision to transition its volume products business online, particularly targeting the SMB segment. The HPE eCommerce Business team has proactively recognized this significant opportunity and is already taking steps to revamp its online store, 'Elite'. The goal is to offer a hassle-free, frictionless, and personalized shopping experience, thereby driving incremental revenue and higher conversion rates. This strategic shift positions HPE to meet the evolving needs of its customers in a rapidly changing digital marketplace.

HPE's extensive and complex product portfolio, including both off-the-shelf (BTO) and customizable (CTO) products, has traditionally relied on business acumen and sales intuition to build customer-specific Elite ecommerce stores. However, this approach has made it challenging to optimize product portfolio size and ensure comprehensive coverage across all product categories. To address this, a scientific, algorithm-driven approach is essential for better product portfolio management. Different sections of the Elite store require tailored, contextual, and cognitive recommendations to enhance the customer experience. For instance, on the homepage, it is crucial not only to recommend products the customer previously viewed but also to incorporate **transaction-based** and **similar user-base** recommendations. By using an **ensemble of algorithms**, we can deliver more accurate and effective product suggestions. When a customer navigates to a product details page (such as a configurator page), **content-based** recommendations become vital. For example, if a customer is viewing a DL server, the recommendation engine should suggest compatible products across relevant categories—such as memory, processors, power supplies, services etc. These recommendations should help the customer choose the right products for optimal performance, cost-efficiency, and balance.

By analyzing customer digital behavior for specific product interests, a ML model, known as the Digital Visitor Intent (DVI) model, predicts a visitor's product intent score, indicating their level of interest in that product. This DVI model and ensemble-based approach delivers **hyper-personalized recommendations** at HPE's B2C Elite eCommerce store, guiding customers to the best product bundles tailored to their needs. Implementing such advanced recommendation algorithms will transform how HPE's Elite store serves its customers, making it a more powerful tool for both optimizing the product portfolio and enhancing the overall buying experience. The proposed solution focuses on customer interest, product compatibility, and product refresh to deliver enterprise-level product recommendations.

Keywords: Machine Learning, Market Basket Analysis, Collaborative Filtering, Data Mining, XGBoost, Digital Visitor Intelligence, Product Recommendation

REFINING MARKETING MIX MODELING: THE POWER OF HIERARCHICAL MIXED-EFFECTS MODELS IN RETAIL CHAINS

Chiranjiv Das

Senior Data Scientist

Epsilon India

Bengaluru, India

chiranjiv.das@epsilon.com

Pramod A

Lead Data Scientist

Epsilon India

Bengaluru, India

pramod.a@epsilon.com

When developing Market Mix Models (MMM) for a major retail chain focusing on specific Stock Keeping Units (SKUs), we encountered several substantial challenges due to data pooling practices. Pooling in linear models involves combining data from different groups or sources to create a single model, which can lead to problems, especially when there are significant differences between the groups. The main issues included omitted variable bias, heterogeneity of variance, masking of interaction effects, ecological fallacy, and loss of detailed insights. To address these challenges, we employed multilevel (hierarchical) models to account for group-specific effects and varying relationships across different groups.

Keywords: market mixed modelling, hierarchical mixed effect modelling, return on investment, random effects model, carryover and saturation effects

CONTEXT-ENHANCED CONTRASTIVE SEARCH FOR IMPROVED LLM TEXT GENERATION

Jaydip Sen

Department of Data Science
Praxis Business School
Kolkata, INDIA
email: jaydip.sen@acm.org

Rohit Pandey

Department of Data Science
Praxis Business School
Kolkata, INDIA
email: rohit.pandeyecon@gmail.com

Hetvi Waghela

Department of Data Science
Praxis Business School
Kolkata, INDIA
email: waghelah@acm.org

Recently, Large Language Models (LLMs) have demonstrated remarkable advancements in Natural Language Processing (NLP). However, generating high-quality text that balances coherence, diversity, and relevance remains challenging. Traditional decoding methods, such as beam search and top-k sampling, often struggle with either repetitive or incoherent outputs, particularly in tasks that require long-form text generation. To address these limitations, the paper proposes a novel enhancement of the well-known Contrastive Search algorithm, Context-Enhanced Contrastive Search (CECS) with contextual calibration. The proposed scheme introduces several novelties including dynamic contextual importance weighting, multi-level Contrastive Search, and adaptive temperature control, to optimize the balance between fluency, creativity, and precision. The performance of CECS is evaluated using several standard metrics such as BLEU, ROUGE, and semantic similarity. Experimental results demonstrate significant improvements in both coherence and relevance of the generated texts by CECS outperforming the existing Contrastive Search techniques. The proposed algorithm has several potential applications in the real world including legal document drafting, customer service chatbots, and content marketing.

Keywords: Large Language Models (LLMs), Contrastive Search, Text Generation, Contextual Calibration, Coherence Optimization, Adaptive Decoding, Creative Content Generation, Natural Language Processing (NLP).

ENHANCING INTELLIGENT DOCUMENT PROCESSING & RAG WITH ANT COLONY OPTIMIZATION AND GEN AI SYNERGY

Divyanshi Yadav

Data Scientist

Data Science & Insights

Genpact

Bengaluru, India

divyanshi.yadav1@genpact.com

Vipul Sharma

Data Scientist

Data Science & Insights

Genpact

Bengaluru, India

vipul.sharma2@genpact.com

Prakash Selvakumar

Assistant Vice President

Data Science & Insights

Genpact

Bengaluru, India

prakash.selvakumar@genpact.com

As organizations continue to expand their digital presence, sifting through vast amounts of documents for essential information becomes a monumental task. This is where the traditional Retrieval-Augmented Generation (RAG) systems are typically employed. While they provide a reasonable solution, they are not without their drawbacks. The chunking and retrieval process followed in the typical RAG solution is inefficient and with the modern digital landscape, the surge for more advanced methods capable of efficiently retrieving relevant information keeps on increasing. Traditional strategies and straightforward algorithmic solutions are proving inadequate in meeting the demands for speed, precision, and cost-effectiveness. This research paper introduces a novel method that enhances Document Processing by combining Ant Colony Optimization (ACO) with Generative Artificial Intelligence (Gen AI) techniques, with a particular focus on the RAG system. The objective is to overcome the shortcomings of existing document processing techniques by utilizing the capabilities of Gen AI for better document clustering, keyword extraction, and layout analysis. The proposed approach utilizes a combination of regular expressions and Ant Colony Parsing Optimization (ACPO) for context association with headers, sub headers, and page numbers. This methodology ensures that when a query is posed, the system intelligently identifies the respective header, sub header, and context to answer the question more efficiently, thereby saving time and reducing operational costs. This research employs a quantitative methodology, utilizing a dataset of diverse documents to test and compare the efficacy of the solution against traditional document processing methods. We employ multiple techniques to analyse the performance metrics, including the accuracy of document clustering, the relevance of keyword extraction, and the effectiveness of layout analysis in facilitating improved RAG use cases. We anticipate the proposed approach will demonstrate significant improvements in the accuracy and efficiency of IDP & RAG tasks compared to traditional methods. By harnessing the synergistic potential of Generative AI, and ACO, this work aims to set a new benchmark for intelligent document processing that can adapt to the evolving needs of various industries. This study provides practical insights for businesses seeking to enhance their document management systems with cutting-edge technologies.

Keywords: Intelligent Document Processing (IDP), Ant Colony Optimization (ACO), Generative Artificial Intelligence, Document Clustering, Keyword Extraction, Document Summarization, Layout Analysis, Business Analytics, Retrieval Augmented Generation (RAG) System, Textual Data Management

OPTIMIZING MEDICAL EQUIPMENT'S INFORMATION RETRIEVAL: THE ROLE OF KNOWLEDGE GRAPHS AND ADVANCED RE-RANKER

Ayush Sinha

Data Science & Insights

Genpact

Bangalore, India

ayush.sinha@genpact.com

Ish Chaudhary

Data Science & Insights

Genpact

Bangalore, India

ish.chaudhary@genpact.com

In the realm of healthcare, managing and staying up to date with the vast amount of information required for the proper use and maintenance of medical devices and equipment presents a significant challenge. For every piece of medical equipment, there exists a plethora of Standard Operating Procedures (SOPs) and resource materials that detail its operation, maintenance, and safety guidelines. These resources are essential for ensuring that medical devices are used efficiently, safely, and in compliance with regulatory standards. The problem lies in the fact that, with each update or upgrade, healthcare professional must frequently update their knowledge, a process prone to error and time-consuming. Additionally, the vast array of SOPs for these devices, stored in various formats and locations, complicates access and training, posing a logistical challenge and risking patient safety. There is requirement of a more streamlined, efficient, and error-proof method of managing, updating, and disseminating the vast amount of critical information associated with the operation and maintenance of medical devices. Our solution introduces a groundbreaking approach to revolutionize medical information retrieval through the strategic application of Knowledge Graphs (KGs), leveraging the innovative Relation Extraction by end-to-end Language generation (REBEL) model, enhanced by an advanced Reranker that prioritizes the most pertinent information nodes. This work not only advances the field of natural language processing and information retrieval but also offers a scalable and practical solution that has the potential to significantly improve the way individuals' access and utilize medical information.

Keywords: Knowledge Graphs, Natural Language Processing, Information Retrieval, Relation Extraction, REBEL Model, Medical Device Management, Standard Operating Procedures

VIDEO RAG: LEVERAGING SEMANTIC CHUNKING AND RETRIEVAL AUGMENTED GENERATION FOR ENHANCED INFORMATION EXTRACTION

Raj Kumar Yadav

Senior Data Scientist

Genpact

Bengaluru, India

rajkumar.yadav2@genpact.com

Gonuguntla Harichandana

Data Scientist

Genpact

Bengaluru, India

ORCID: 0009-0004-8223-7817

Koundinya N V S S

Senior Data Scientist

Genpact

Hyderabad, India

ORCID: 0000-0003-1663-3112

Vaibhav Misra

Data Scientist

Genpact

Bengaluru, India

vaibhav.mishra@genpact.com

Ankit Sati

Senior Data Scientist

Genpact

Bengaluru, India

ankit.sati@genpact.com

Chirag Jain

Principal Data Scientist

Genpact

Bengaluru, India

chirag.jain4@genpact.com

Video has become an ubiquitous source of information, with about 90% of global internet traffic expected to be video by 2024 (Cisco). This massive volume of video content presents a significant opportunity, but also a major challenge: extracting relevant information from videos remains a complex task. Traditional question answering systems, often limited to text or images, struggle to handle the multimodal nature of video.

Keywords: Video Question Answering, Multimodal Understanding, Semantic Chunking, Video Retrieval-Augmented Generation, AI, Video Analytics, Information Extraction.

ANALYTICS FOR ENHANCING EMPLOYABILITY OF MBA GRADUATES USING PREDICTIVE MODELLING

Dr. Sushma V

Assistant Professor

Dept. of Business Administration

Sahyadri College of Engineering & Management, Mangaluru

Email ID: sushma.mba@sahyadri.edu.in

Phone No.: +91 9632053663

Dr. Vishal Samartha

Professor and Director-MBA Program

Dept. of Business Administration,

Sahyadri College of Engineering & Management, Mangaluru

Email ID: director.mba@sahyadri.edu.in

Phone No.: +91 9886184413

Dr. Suhan M

Additional Professor

Dept. of Commerce

MAHE, Manipal University, Manipal

Email ID: suhan.mendon@manipal.edu

Phone No.: +91 9986103367

Employers in developing countries report that a lack of skilled workers is a major and increasing bottleneck for their operations, affecting their capacity to innovate. With reference to the India Skill Report (ISR) 2023, MBA graduates are only at 60.1% employable talents among the youth. Karnataka has never been in top 3 among the states with maximum supply of employable talent in the last 9 years as per the ISR 2023. Management institutions have shouldered the responsibility of enabling MBA graduates to some extent, yet the challenges on employability still exists. The purpose of this study is to develop and empirically test a theoretical model that explains the effect of career-related social support (CRSS) and employability opportunities/activities (EOQ) on perceived employability (PE) by bringing in literacies for life (LFL) as a mediator. The model is conceptualised in the backdrop of social cognitive theory. Using a cross-sectional research, 1,117 final year MBA students of 2020-22 batch from Visvesvaraya Technological University (VTU), Belagavi, Karnataka, India affiliated MBA autonomous and non-autonomous colleges (both rural and urban) in Karnataka, India were surveyed. The colleges were further grouped under 4 VTU regions – Belagavi, Bengaluru, Kalaburgi and Mysuru. In order to carry out the study, questionnaire method was chosen as a data collection instrument. Questions related to the four constructs were driven through the questionnaire. In addition to this, certain generic information was collected in the same questionnaire in order to understand the background of the respondents. A total of 7,018 was the population size from 36 Colleges. Out of which 1,117 was derived as the sample size considering 95% confidence level and 5% margin of error. This has been proportionately divided among the selected colleges based on the population. Structured Equation Modelling (SEM) was conducted using SMART PLS for data analysis. Measurement Model was tested for reliability and validity. Structural Model was checked for the Model Fitness. Radar chart highlights the skill that require focus. A systematic approach to skill assessment of MBA students to enhance employability has been proposed incorporating the factors taken for the study.

Keywords: Higher Education, MBA, Skills, Perceived Employability, Career social support, Employability opportunities

SENTENCE LEVEL CORRECTION OF COMMON SPOKEN GRAMMATICAL ERRORS WITH DEEP NEURAL NETWORKS

Malay Shikhar

Department of Computer Science and Engineering
PES University
Bengaluru, India
malayshikharsahay@gmail.com

Sai Nikhila Kuchimanchi

Department of Computer Science and Engineering
PES University
Bengaluru, India
nikhila.kuchimanchi@gmail.com

Maitreyi Kar

Department of Computer Science and Engineering
PES University
Bengaluru, India
maitreyikar13@gmail.com

Liwa Al Haq Muhammad Kunhi

Department of Computer Science and Engineering
PES University
Bengaluru, India
liwaalhaq@gmail.com

R Bharathi

Department of Computer Science and Engineering
PES University
Bengaluru, India
rbharathi@pes.edu

Grammatical errors disrupt conversation clarity and effectiveness, even for fluent speakers. Such errors can make even those fluent in a language oblivious to what is being conveyed. This study presents a novel approach, using natural language processing and transformer models to both identify and provide feedback on common grammatical errors spanning 10 types. These include adjective and adverb misuse, determiner selection, preposition selection, conjunction selection, verb form and tense, subject-verb agreement, noun and pronoun errors. The system operates as a cascaded pipeline of a binary sentence classifier and a grammar corrector. Sentences are classified as “correct” or “incorrect” by the classifier while Google’s T5 model, fine-tuned on C4-200M for grammatical error correction (GEC), corrects errors in ungrammatical sentences. ERRANT annotates the errors present in the input sentence, concluding the evaluation. Model validation performed on various datasets shows the grammar corrector performs best on error-filtered JFLEG with BLEU 0.89 and ROUGE 0.95. The binary sentence classifier also achieved an accuracy score of 0.78 on a custom dataset. While it is effective, its applicability is currently limited to English, with plans for multilingual support and real-time applications exploring the expansion of the model’s capabilities.

Keywords: Grammar error correction, deep learning, spoken errors, sentence classification, grammar error detection

ADDRESSING EMERGING RISKS IN LEVS: A NOVEL FRAMEWORK FOR INSURANCE PRICING USING TRANSFER LEARNING AND TELEMATICS

Sudaman Thoppan Mohanchandralal

In the dynamic landscape of insurance, staying ahead requires the ability to adapt to emerging risks swiftly. With the advent of new technologies, changing consumer behaviors, and evolving market dynamics, insurers face the imperative to assess and price new risks effectively. Due to the emergence of light electrical vehicles (LEVs) and their increasing popularity among consumers, I choose to assess and price the associated risks. This paper details the first results of the concept of applying a novel framework innovated for evaluating and pricing the risks associated with insuring LEVs. The first efforts achieved is through the exploration and usage of external data sources along with internal data sources where available. Two varieties of transfer learning was considered while one of the best fit approach is chosen and this is detailed in a separate section. I firmly believe that this initiative will extend the academic literature that already exists in the area of risk management and dynamic pricing for SME's emerging risks and is also a potent value driver that can significantly enhance businesses revenue, operational excellence and profitability while securing SME from emerging risks. The approach I took is to start with desk search followed by interviews with subject matter experts (underwriters). Based on the outcome of these two steps, I started to gather data. I started with internal data, followed by external data for three countries vis a vis Netherlands, Belgium and France. Based on the data collected, I started to build the risk features and then proceeded to build the first model. Sections below explain these steps in greater detail.

AN APPROACH TOWARDS HANDLING LLMS HALLUCINATIONS USING LLM ENSEMBLE

Jagan P S

AI CoE, Engineering
Flexera Software Solutions LLP
Bengaluru, India
Jsrilakshmi@flexera.com

Amit Sharma

AI CoE, Engineering
Flexera Software Solutions LLP
Bengaluru, India
asharma@flexera.com

Barath Ravichander

AI CoE, Engineering
Flexera Software Solutions LLP
Bengaluru, India
bravichander@flexera.com

Punyakeerthi BL

AI CoE, Engineering
Flexera Software Solutions LLP
Bengaluru, India
pbl@flexera.com

Mallikarjuna S

AI CoE, Engineering
Flexera Software Solutions LLP
Bengaluru, India
msubramani@flexera.com

Ensuring factual accuracy and reliability in the outputs of Large Language Models (LLMs) remains a key challenge due to the frequent occurrence of hallucinations, where models generate incorrect or nonsensical information. This paper proposes a novel multi-LLM approach that mitigates hallucinations by combining and cross-verifying responses from multiple LLMs. A ranking system evaluates each response based on factual accuracy, coherence, and alignment with verified data sources, while an additional mechanism penalizes hallucinatory content. This approach enhances the reliability of LLM-generated text, making it suitable for critical applications such as medical advice, legal analysis, and content creation.

Keywords: large language models, hallucinations, factual accuracy, multi-LLM, response reliability

SAAS CLOUD LICENSE MANAGEMENT - PREDICTIVE COST OPTIMIZATION

Drupad SS

AI CoE, Engineering
Flexera Software Solutions LLP
Bengaluru, India
dsatish@flexera.com

Amit Sharma

AI CoE, Engineering
Flexera Software Solutions LLP
Bengaluru, India
asharma@flexera.com

Barath Ravichander

AI CoE, Engineering
Flexera Software Solutions LLP
Bengaluru, India
bravichander@flexera.com

Aamer Sharif

Product Development
Flexera Software Solutions LLP
Melbourne, Australia
asharif@flexera.com

Tim Ferguson

Architecture
Flexera Software Solutions LLP
Melbourne, Australia
tferguson@flexera.com

Efficient management of SaaS cloud licenses is crucial for organizations to maximize ROI and ensure compliance. However, current research lacks exploration of AI- driven approaches to this problem. This presents a significant opportunity for research and development within the field. This research proposes an AI-driven solution to automate and optimize SaaS cloud license management. By leveraging advanced machine learning techniques, our system analyses vast amounts of historical usage data to identify inefficiencies, underutilized licenses, and potential cost savings. Key features include:

- Analyzing existing license usage: We need to assess whether the current licensing type is appropriate for their usage.
- Predictive License Usage Forecasting: Accurately predicting future usage patterns to optimize procurement and renewal decisions
- Intelligent License Recommendation: Recommending optimal license configurations based on user behavior and organizational needs
- Automated License Discovery and Classification: Automatically identifying and categorizing SaaS licenses across various cloud platforms.

Through the implementation of this AI-powered solution, organizations can significantly reduce costs, improve operational efficiency, and enhance overall cloud resource management.

A STUDY ON PREDICTING INTEREST RATE MOVEMENT OF BANKS

Dr. Indumathi R and Nithin Kumar P

Interest rates are a cornerstone of economic stability and financial market efficiency, influencing borrowing costs, savings returns, and the broader economic environment. In India, public sector bank, plays a pivotal role in the financial system, with its interest rate movements significantly impacting the banking industry and the economy. This study aims to develop a predictive model for interest rate movements by analyzing historical data, macroeconomic variables, and competitive positioning. By understanding the determinants of interest rate changes, the research seeks to provide a framework for accurate predictions. These insights are crucial for policymakers, financial institutions, businesses, and individual investors to optimize investment strategies and enhance the effectiveness of monetary policy. The findings of this study have significant implications for strategic planning, risk management, and financial decision-making.

Keywords: Interest rate Prediction, Macroeconomic Variables, Financial Market Efficiency, Monetary Policy

DETECTION OF PROTECTIVE GEAR COMPLIANCE USING CONVOLUTIONAL VISION TRANSFORMERS

Neeraja Vaidya

Principal Data Scientist
Neoware Technology Solutions
Mumbai, India
neeraja@neoware.ai

Rishibalaji Srinivasan

Dept. of Chemical Engineering
Indian Institute of Technology
Guwahati, India
s.rishibalaji@iitg.ac.in

Aparna Yegnaraman

Data Scientist
Neoware Technology Solutions
Chennai, India
aparna@neoware.ai
ORCID: 0000-0002-5759-7851

Swaminath Venkataraman

Customer Success Officer
Neoware Technology Solutions
Chennai, India
swami@neoware.ai

In mines and factories, workers must wear protective equipment such as goggles, suits, shoes, masks, gloves, and helmets. Companies enforce the use of such mandatory gear to prevent workplace injuries and fatalities, as well as to avoid financial and legal repercussions. However, continuously inspecting workers can be time-consuming and prone to errors. To address this, we have developed a solution that detects workers without protective gear in images and videos from mining and factory environments, triggering warnings when violations are identified. This solution leverages the latest advancements in deep learning, specifically convolutional vision transformers. Our model is trained on a diverse set of images from various mining and factory settings, including those captured in low light and from a distance. It can detect violations in real-time, enabling companies to take immediate corrective actions, avoid legal issues and financial losses, and most importantly, protect human lives.

Keywords: protective gear compliance, deep learning, computer vision, construction safety, convolutional vision transformers.

A PRESCRIPTIVE MODEL FOR THE DEPOT LOCATIONS AND ALLOCATION OF CITY BUSES TO DEPOTS

Rupkatha Ghosh and M. Mathirajan

Department of Management Studies

Indian Institute of Science

Bangalore, India

rupkathag@iisc.ac.in ; msdmathi@iisc.ac.in

In India, rapid urbanization and dramatic growth in private vehicle ownership has created a great pressure on the Urban Road Transport Organizations (URTO) to augment mass transit capacity and to fulfill the expanding population's needs. The city buses run by these URTO are cost-effective due to the government's obligation to serve all sections of society and connect most of the parts of the city on subsidized rates. Further, as the city size expands due to migration of people, the number of commuters increases, for which a good number of new additional buses need to be introduced. To accommodate an increased number of buses, the existing depot capacity might not be sufficient, bringing the need to open new depots. For opening a new depot(s), the organizations incur a fixed cost (FC). Further, every URTO creates depots on the outskirts of a city. Over the period, a few of the locations of the depots become the inner part of city as city size increases, creating menace to the public leading to the closure of a few existing depots. The closing of existing depot(s) causes salvage cost (SC) to URTO. Thus, every URTO strategically does to open a few new depots and close a few existing depots. This situation happens once in 2/3/4 years' time. Thus, the decision problem on depot location (DL) while opening new depot(s) and closing existing depot(s), a strategic-level decision problem, is one of the classical facility location problems. Depot location also impacts the non-revenue travel distance covered by each of the buses. In the early morning, there is a non-revenue travel distance from depot to starting terminus of route of a bus and in the late-night, there is a non-revenue travel distance from last terminus of the route back to depot. This travel distance is called the dead kilometers in literature. The cost associated with dead kilometer is called Dead Kilometer Cost (DKC). Though DKC cannot be eliminated, it is possible to minimize total DKC by optimal allocation of city buses to depots. This decision problem is called Allocation of City Buses to Depot (ACBD). This decision problem: ACBD problem is one of the tactical-level decision problems, as the frequency of deciding on the ACBD might occur whenever (a) a considerable number of new buses are added, (b) one-way routes within a city change etc. Further, the decision on ACBD is also influenced by depot operating cost (DOC). Both the decision problems on depot location (DL) and ACBD problem associated with URTO are inseparable and this integrated problem of URTO is termed as DL-ACBD problem. Based on review of existing literature, only one study has considered FC, SC, DKC and DOC while taking optimal decisions on DL-ACBD problem, but not in Indian context. Further, an important cost component: environmental cost (EC) is not considered while optimizing the DL-ACBD problem in general, particularly in Indian context. Considering this research gap, this study has proposed a prescriptive model for DL-ACBD problem of URTO by simultaneously minimizing total DKC, FC, SC, DOC, and EC. Further, the workability of the proposed prescriptive model is demonstrated through a numerical example developed for representing the DL-ACBD problem.

Keywords: depot location, allocation of city buses to depots, dead kilometer cost, fixed cost, salvage cost, depot operating cost, environmental cost, MILP model

ENHANCING ON-TIME LOAN REPAYMENT AND CUSTOMER RETENTION IN MICROFINANCE LOANS THROUGH MACHINE LEARNING ALGORITHMS

Mayapremil S

II Year PGDM

Thiagarajar School of Business,

Madurai, India

mayapremil_2313093@tsm.ac.in

Dr. Nataraj Balasubramanian

Faculty Supervisor,

Assistant Professor, Chairperson MDP,

Thiagarajar School of Management,

Madurai, India

This research aims to improve On-Time Repayment (OTR) rates and customer retention at a microfinance institution by identifying key factors influencing repayment behavior and customer loyalty. The study provides actionable insights to optimize loan approval processes, reduce delinquency risks, and enhance customer relationships. Conducted in two phases the first phase involved analyzing loan disbursement data using regression analysis, Random Forest classification, and decision tree algorithms (CHAID and CART) to identify predictors of OTR behavior, including borrower demographics, loan characteristics, and repayment history. The second part focused on customer retention among OTR borrowers through decision tree algorithms to discern primary factors influencing loyalty. The findings revealed critical variables that significantly impact OTR rates, offering insights into improving loan eligibility criteria and risk assessment models. Key factors influencing customer satisfaction and retention were also identified, informing the design of targeted retention programs. Collectively, these findings provide a strategic framework for enhancing both loan repayment performance and customer loyalty within the financial services sector. However, the study's geographic limitation to the Kelambakkam branch of Belstar may affect the generalizability of the results. Future research should broaden the scope and include additional variables, such as CIBIL scores, to enhance prediction accuracy. The study's practical implications include providing a foundation for risk scoring models and tailored retention strategies that can reduce delinquency rates, boost profitability, and foster customer loyalty.

Keywords: On-Time Repayment (OTR), Loan Disbursement, Machine Learning, Predictive Analytics, Risk Assessment, CHAID, CART

STEERING E-COMMERCE SUCCESS THROUGH NEXT BASKET RECOMMENDATION SYSTEM

Ashis Kumar Panda

Senior Data Scientist

Epsilon India Data and Digital Technology Solutions LLP

Gaurav Sarda

Data Scientist 2

Epsilon India Data and Digital Technology Solutions LLP

The proliferation of e-commerce platforms has revolutionized shopping behaviors, necessitating advanced recommendation systems to enhance user experience and drive sales. Traditional recommendation systems have revolutionized e-commerce by harnessing data-driven personalization to meet the rising demands of consumers, where UX plays a pivotal role in influencing shopping behaviors and choices. This has amplified upselling, cross-selling, and customer lifetime value, while fueling customer retention, revenue growth, and a robust competitive edge. Among these, Next basket recommendation has emerged as a pivotal innovation, leveraging sophisticated algorithms that takes personalization to the next level by predicting entire sets of items that a user is likely to purchase in their next shopping session, based on their past buying patterns and the buying patterns of similar users. This method focuses on understanding the sequence and timing of purchases, enabling e-commerce platforms to offer highly relevant product suggestions at the right moment, offering a holistic and personalized shopping experience. We propose a novel approach to developing a Next Basket Recommendation system by combining collaborative filtering technique with personalized item frequency analysis to enhance the accuracy and relevance of recommendations by integrating both user interaction data and overall shopping patterns. We then analyze and assess the performance of the next basket recommendation (NBR) method on multiple metrics. The major implications of this research extend to dynamically influencing consumer purchasing decisions, driving customer loyalty, boosting sales revenue, through targeted promotions and personalized recommendations. This study shows the transformative potential of NBR in the e-commerce sector, providing a foundation for further advancements in predictive analytics and customer relationship management.

Keywords: Next Basket Recommendation, Personalized Item Frequency, Collaborative Filtering

Leveraging Machine Learning to Enhance Last-Mile Delivery Efficiency: The Impact of Agent Experience, Environmental Factors, and Geographical Distance

Prof. Nikita Ramrakhiani

Vivekanand Education
Society's Business School
Mumbai, India
nikita.ramrakhiani@ves.ac.in

Yash Ramesh Tambe

Vivekanand Education
Society's Business School
Mumbai, India
yash.tambe2325p@ves.ac.in

Indraneel Deshmukh

Vivekanand Education
Society's Business School
Mumbai, India
indraneel.deshmukh2325p@ves.ac.in

The intricate tapestry of logistics operations, encompassing transportation, warehousing, inventory, and packaging, must be meticulously orchestrated to ensure the seamless and timely flow of goods and services within a supply chain, but the growing volume of e-commerce orders has strained traditional delivery optimization methods. From freight delivery routes to personal travel itineraries, travel time is a key factor in urban transportation. By leveraging machine learning models, delivery companies can improve route planning, resource allocation, and customer communication, ultimately enhancing overall efficiency and customer satisfaction. Through a rigorous exploration of delivery data, this research aims to uncover the underlying relationships between delivery time, traffic congestion, weather variability, and agent competency, providing valuable insights for optimizing delivery processes. For the purpose of this research, we examined the data of Amazon in India to investigate the factors influencing last-mile delivery efficiency and develop random forest models to predict delivery times accurately. The advanced system proves to be a valuable asset in optimizing the distribution chain, resulting in substantial improvement in ratings from consumers. A sophisticated algorithmic framework was designed, incorporating a series of interconnected steps that leverage various machine learning and data mining techniques as the initial phase of a project to improve delivery route efficiency.

Keywords: Last-mile delivery, random forest, machine learning, Data Visualization, Supply chain

IMPACT OF SHARED ACCOUNT USAGE ON PERSONALISATION ALGORITHMS

Aditi Kolekar

Indian Statistical Institute

Tezpur, India

kolekaraditi22@gmail.com

Sudarshan Muralidharan

Department of Management Studies

NALSAR University of Law

Hyderabad, India

sudarshanmuralidharanmba202325@doms.nalsar.ac.in

Personalisation algorithms are crucial in optimising user experiences on digital platforms by delivering recommendations tailored to individual preferences. However, their efficacy can be significantly reduced when accounts are shared among multiple users, leading to skewed recommendations and diminished user satisfaction. This research investigates the impact of shared account usage on the effectiveness of personalisation algorithms, addressing a notable gap in current understanding. We analyse how shared usage scenarios disrupt the accuracy and relevance of personalised recommendations and identify the factors contributing to these distortions. Our study introduces and evaluates innovative approaches to counteract these effects, focusing on the development of adaptive recommendation systems that can dynamically adjust to multiple user profiles within a single account. Through empirical analysis and system modeling, we propose strategies to enhance the precision of personalisation algorithms in shared environments, ultimately aiming to improve the overall user experience for all involved parties.

Keywords: Personalisation Algorithms, User Experience, Recommendations, Shared Accounts, Skewed Recommendations, User Satisfaction

UNLOCKING THE POTENTIAL OF MEETING NOTES

Mousumi Acharya

Data Scientist

Fidelity Investments

Bangalore, India

Mousumi.Acharya@fmr.com

Nitin Kumar

Data Scientist

Fidelity Investments

Bangalore, India

Nitin.Kumar@fmr.com

Jino Mathews

Director

Fidelity Investments

Bangalore, India

Jino.Mathews@fmr.com

Meeting notes play a key role in documenting the essence of interactions during meetings. They serve as organizational memory and asset while also ensuring effective communication and clear accountability among each of the concerned stakeholders. With increasing frequency, modes of interaction and complexity of business meetings, it is increasingly becoming difficult to keep up with the immense amount of information exchanged during the meetings in absence of a well-established approach to record, store and access the same. With text mining operations on meeting notes, organizations can unlock the capability of processing the meeting notes further such as enhancing productivity by saving time in future meeting preparations since concise summary of notes and broader topics discussed can be quickly referenced to drive future discussions, tracking progress as well as maintaining business continuity during times of change. They also foster a sense of accountability by clearly outlining the action items for individual stakeholders along with respective timelines, that can help improve follow-up and ensure that projects remain on track. In this paper, we are proposing an approach to use different Large Language Models (LLMs) to achieve 3 downstream information extraction tasks and evaluate the performance of same in respective areas. The three tasks are "extracting Key-phrase" to map/structure the data, "generating concise summary" and "extracting Action items" from the meeting notes.

Keywords: NLP, AI, Generative AI, Topics, Themes, Summary

AUTOMOTIVE SALES AND SERVICE RECOMMENDATION WITH SINGULAR VALUE DECOMPOSITION AND MULTI-LABEL CLASSIFICATION MODELS

Vishal Kailaswar

Auto Analytics

Epsilon India Data and Digital Technology Solutions LLP

Bengaluru , India

kailaswar.vishal@epsilon.com

Abhiraj Singh

Auto Analytics

Epsilon India Data and Digital Technology Solutions LLP

Bengaluru , India

abhiraj.singh@epsilon.com

The global automotive market is expanding with plethora of feature packed car models being launched by all the automakers. Enabling customers to make right choice of car models based on their needs and thereafter ensuring periodic services to build loyalty is key for the growth of automotive business. In this study, we present sales and service recommendation models tailored for the automotive industry. The solution includes deploying advanced AI algorithm and model to enable the customer acquisition and retention through personalized recommendations. Initially, to enable customer acquisition strategy implemented sales recommendation algorithm using Singular Value Decomposition (SVD) algorithm to predict the top two car models that a customer is likely to purchase based on historical purchase patterns. The SVD algorithm further was incorporated to suggest the top three “complementary services” at the time of car purchase and achieved goal of generating additional upfront revenue. Post-purchase, a Multi-Label Classification (MLC) model was used to predicts multiple vehicle services (brakes, shocks, tire, oil) needs within the next 60 days for each customer. Robustness of model was ensured by utilizing Optuna for hyperparameter tuning to optimize the precision and recall and minimized errors. Performance metrics indicate that targeting the top 30% of customers captured approximately 70% of the service demand. One of our global automaker clients incorporated model based personalized recommendations strategies and that enabled their business decisions with significant improvements in customer acquisition and retention. Adopting the strategies also increased their operational efficiency by aligning with the digital transformation and improving customer retention through personalized periodic vehicle services leading to higher repair order (RO) and loyalty.

Keywords: Single Value Decomposition, Multi-Label Classification, Repair Order.

ASPECT-BASED CALL TRANSCRIPTS SUMMARIZATION

Nitin Kumar

Data Scientist
Fidelity Investments
Bangalore, India
Nitin.Kumar@fmr.com

Talamanchi Sai Yaswanth Reddy

Director, Data Science
Fidelity Investments
Bangalore, India
Talamanchi.Yaswanthreddy@fmr.com

Bidhan Roy

VP, Data Science
Fidelity Investments
Bangalore, India
Bidhan.Roy@fmr.com

Call Transcripts summarization has gained significant attention recently, thanks to the power of recent Large Language Models (LLMs). However, the generated summaries often fall short due to redundant or inaccurate content, primarily attributed to the unstructured and complex nature of human-human interactions. Generic summarization, which takes a broad approach to summarizing call transcripts can be achieved using LLMs. However, this approach doesn't emphasize enough on specific aspects like queries or goals. In this paper, we are solving two major problems in this area. Firstly, we aim to generate coherent and informative summaries while avoiding redundancy. We are proposing a framework for aspect-based summarization, encompassing Customer Problem Understanding, Agent Assessment, Friction Point Analysis, and Multi-level Intent Extraction. Secondly, we propose guidelines for utilizing LLMs for evaluating generated responses, employing another LLM for this purpose.

Keywords: LLMs, NLP, Summarization, Artificial intelligence, Machine Learning

DEMOCRATIZING DATA INSIGHTS – EMPOWERING BUSINESS ANALYTICS THROUGH GENERATIVE AI

Divyanshi Yadav

Business Analyst
Data Science & Insights
Genpact
Bengaluru, India
divyanshi.yadav1@genpact.com

Prakash Selvakumar

Assistant Vice President
Data Science & Insights
Genpact
Bengaluru, India
prakash.selvakumar@genpact.com

In the rapidly evolving business landscape, the ability to quickly and accurately analyze large volumes of data for decision-making is paramount. Traditional methods of querying knowledge graphs (KGs) for business analytics often require technical expertise, leading to delays and dependency on specialized personnel. This paper presents an innovative solution leveraging generative artificial intelligence (AI) and advanced natural language processing (NLP) techniques to address these challenges. The research builds upon efforts to enhance business analytics by enabling non-technical users to query massive databases using natural language, democratizing access to critical insights and speeding up decision-making. The core of the proposed approach involves using generative AI, specifically Large Language Models (LLMs), equipped with curated prompts to autonomously generate database queries from user-input natural language questions. This work focuses on optimizing this process and mitigating inaccuracies, commonly referred to as "LLM hallucination," by employing techniques such as reflection methods for error checking and a fallback mechanism for query correction. Additional tested approaches include chain of thought and pathways parsing to enhance understanding and accuracy, along with a voting method to further reduce the likelihood of hallucination. These mechanisms ensure users can obtain comprehensive, accurate, and easily understandable insights from complex databases. This novel pipeline facilitates translating layman's terms into precise database queries and provides natural language narratives and visual representations of the data as part of the response. This technique has proven significantly innovative for organizations needing frequent, in-depth database analysis for decision-making purposes, including system and departmental lineage, business strategy, and more. The paper details the experimental setup, methods, and performance evaluation, demonstrating the approach's effectiveness on multiple factors. Notably, this system offers reliable recommendations and enables the visualization of data subgraphs, contributing to a more intuitive understanding of the underlying information. This study contributes to the field by providing a comprehensive solution to a prevalent problem in business analytics, showcasing the potential of generative AI and NLP in transforming data querying processes. The findings underscore the importance of integrating advanced AI techniques to facilitate more accessible, accurate, and efficient information retrieval and analysis, ultimately empowering businesses to make better-informed decisions with minimal technical dependency.

Keywords: Business Analytics, Generative AI, Knowledge Graphs, Natural Language Processing, Large Language Models, Reflection Technique, Chain of Thought, Pathways Parsing, Information Retrieval, Decision Making.

ANALYZING AND MITIGATING PROMPT INJECTION TECHNIQUES IN AI LANGUAGE MODELS

Meghan Challa

Data Science Intern

Genpact

Hyderabad, India

meghan.challa@genpact.com

Shashank Kumar

Data Science Intern

Genpact

Hyderabad, India

shashank.kumar4@genpact.com

Kolli Nethre Sai

Senior Data Scientist

Genpact

Hyderabad, India

kollinethre.sai1@genpact.com

Koundinya N V S S

Senior Data Scientist

Genpact

Hyderabad, India

Mohit Pawar

Senior Data Scientist

Genpact

Gurgaon, India

mohit.panwar@genpact.com

Chirag Jain

Principal Data Scientist

Genpact

Bengaluru, India

chirag.jain4@genpact.com

Prompt injection techniques exploit vulnerabilities in AI language models, manipulating their outputs through specifically crafted prompts. As natural language processing (NLP) becomes increasingly integrated into various applications, understanding and mitigating these risks is essential. This paper presents a comprehensive comparative analysis of different prompt injection methods, assessing their effectiveness, robustness, and ethical implications. Organizations using AI models face significant risks, with the average cost of a data breach reaching \$4.35 million in 2022, highlighting the financial and operational stakes involved. Our methodology systematically addresses these vulnerabilities by employing adversarial inputs to stress-test model performance and utilizing role-based prompts to evaluate adaptability. We investigated techniques such as jailbreaking, token smuggling, and payload splitting, identifying gaps in existing safety protocols. Key findings reveal that while direct injection techniques are effective, they often lack robustness, and ethical concerns arise from adversarial methods. Based on these insights, we propose best practices for managing prompt injection risks and emphasize the need for future research to develop automated tools for detection and mitigation, enhancing the overall security of AI systems.

Keywords: Prompt injection, AI language models, Vulnerabilities, Mitigation, Adversarial inputs, Ethical implications

TRANSFORMING VIDEO CONTENT INTO DYNAMIC MULTIMODAL SOPS

Gonuguntla Harichandana

Data Scientist
Genpact
Bengaluru, India

Koundinya N V S S

Senior Data Scientist
Genpact
Hyderabad, India

Raj Kumar Yadav

Senior Data Scientist
Genpact
Bengaluru, India

rajkumar.yadav2@genpact.com

Pritam Varma

Data Scientist Intern
Genpact
Hyderabad, India

pritam.varma@genpact.com

Prudhvith Tavva

Senior Data Scientist
Genpact
Bengaluru, India

tavva.prudhvith@genpact.com

Prakash Selvakumar

Principal Data Scientist
Genpact
Bengaluru, India

prakash.selvakumar@genpact.com

Chirag Jain

Principal Data Scientist
Genpact
Bengaluru, India

chirag.jain4@genpact.com

The ubiquity of video content in today's digital landscape presents a compelling opportunity to revolutionize training and documentation practices across industries. This paper presents an AI-driven framework for transforming video content into structured, multimodal Standard Operating Procedures (SOPs). Our methodology integrates advanced multimodal processing, encompassing image recognition, audio transcription, and contextual summarization, to produce SOPs enriched with text, visuals, and flowcharts. By leveraging state of the art multimodal AI models, our approach captures the full spectrum of instructional content, accommodating diverse learning preferences. Experimental results demonstrate the system's effectiveness in converting varied video formats into detailed, actionable SOPs, underscoring its potential to enhance organizational efficiency and training outcomes.

Keywords: Video Content Transformation, Multimodal Standard Operating Procedures (SOPs), Artificial Intelligence (AI), Computer Vision, Multimodal AI Models, Vision Agents.

A DEEP DIVE INTO RAG PARADIGMS - SELECTING THE RIGHT APPROACH FOR SPECIFIC BUSINESS APPLICATIONS

Ashrita Naidu Vaka

Data Science Intern

Genpact

Hyderabad, India

ashritanaidu.vaka@genpact.com

Yadma Meghna Reddy

Data Science Intern

Genpact

Hyderabad, India

yadma.meghnareddy@genpact.com

Koundinya N V S S

Senior Data Scientist

Genpact

Hyderabad, India

ORCID: 0000-0003-1663-3112

Kolli Nethre Sai

Senior Data Scientist

Genpact

Hyderabad, India

kollinethre.sai1@genpact.com

Ish Chaudhary

Senior Data Scientist

Genpact

Bengaluru, India

ish.chaudhary@genpact.com

Chirag Jain

Principal Data Scientist

Genpact

Bengaluru, India

chirag.jain4@genpact.com

This paper examines Retrieval-Augmented Generation (RAG) paradigms—Naive RAG, Graph RAG, Agentic RAG, and Multimodal RAG—across healthcare, finance, knowledge management, and customer support. As businesses face challenges like data integration inefficiencies and fraud, advanced RAG models offer solutions for real-time insights and actionable recommendations. Each paradigm is evaluated using metrics such as Mean Reciprocal Rank (MRR), F1 score, context precision, and context recall, leveraging datasets tailored to sector-specific issues. This paper reveals trends indicating that Multimodal RAG is particularly effective in healthcare, Agentic RAG performs well in finance, Graph RAG is advantageous for knowledge management, and Naive RAG is suitable for customer support. Future research could investigate hybrid RAG systems to enhance adaptability and address complex business challenges more effectively.

PREDICTING MOBILE PHONE PRICE CHANGES USING REINFORCEMENT LEARNING: HELPING CUSTOMERS UNDER PRICE UNCERTAINTY

Soumak Banerjee

FPM, Research Scholar
International School of Management Excellence
Bangalore, India
soumakb.fpm22@isme.in

Dr. Shampa Nandi

Professor and Principal
International School of Management Excellence
Bangalore, India
<https://orcid.org/0000-0002-6270-288x>

For mobile phone consumers with device prices below 20,000 INR, understanding the price fluctuations is crucial for consumers if the offer price of the product changes frequently and thus makes it unpredictable. There are numerous price prediction models available under uncertainty and volatility, still, there remains a gap in forecasting the dynamic price trends in e-commerce environments. The main objective of this study is to bridge this gap by applying Machine Learning techniques called Reinforcement Learning to predict the future price of the mobile phone device. Over 30 days, daily offer prices of 44 mobile models were collected and analyzed. The data shows that approx. 30% of the mobile models do not change their price while most of the products show low to medium price change during the period. The methodology involves the development of an algorithm that integrates reinforcement learning by training the model with the data being collected to recognize the pattern of future prices and thus forecast price trends. As a key finding, research reveals the model's effectiveness in capturing market dynamics and providing a predicted price with some degree of accuracy. the pricing model. As bound by rationality, the consumer's decision-making process can be enhanced using the model. Using machine learning with behavioral theories like reference price theory, the study provides a new approach to price prediction. The study also helps a deeper understanding of prospect theory and its impact on the pricing model.

Keywords: Reinforcement Learning, ecommerce, dynamic price.

FROM DECISION TREES TO BOOSTING: A STUDY OF CLASSIFICATION ALGORITHMS FOR BNPL CREDIT RISK

Dr. Sunitha K

Dept of finance, accounts and
economics, VIT Business School
Vellore Institute of Technology,
Vellore, India
Email: sunitha.k@vit.ac.in

Dr Dilip D

Dept of operations management,
VIT Business School
Vellore Institute of Technology, Vellore,
India
Email: dilip.d@vit.ac.in

Buy Now, Pay Later (BNPL) services have become a popular financial tool empowering consumers and boosting sales in different sectors. These services allow consumers to delay payments or split them into smaller instalments, providing flexibility and better access to goods and services. The increasing adoption of BNPL underscores the need for effective credit risk assessment and management. Accurately assessing credit risk in BNPL transactions is crucial for minimizing default rates and ensuring the financial stability of lending institutions. Proper risk assessment helps predict customer repayment behavior, set credit limits and reduce financial losses. The study examines the performance of various classification models on a BNPL dataset to determine their effectiveness in predicting creditworthiness and managing associated risks. The dataset includes features such as employment status, merchant category, BNPL usage frequency, transaction amounts, and tenure. We tested several classification algorithms, including Logistic Regression, Decision Tree, Random Forest, Support Vector Machine (SVM), K-Nearest Neighbors (KNN), XGBoost and CatBoost. Each model underwent thorough hyperparameter tuning using Grid Search to optimize performance. Evaluation metrics such as accuracy, precision, recall and AUC provided a comprehensive assessment of each model's predictive abilities. Our analysis shows that SVM emerged as the most improved model post tuning with the highest accuracy and F1 score (0.9083). With an accuracy of 0.867, CatBoost remains very competitive and may be preferred due to lower computational requirements and faster training times on a large dataset. This research provides valuable insights into the effectiveness of various classification models for BNPL credit risk assessment. It offers practical guidance for financial institutions looking to enhance risk management strategies and improve decision-making processes in BNPL lending.

Keywords: Credit risk assessment, AI, ML, Predictive Models, BNPL credit risk

TRANSFORMING AIRLINE SERVICE QUALITY: HARNESSING USER-GENERATED CONTENT

Ayaskanta Panigrahi

Research Scholar

Indian Institute of Management Tiruchirappalli

Tiruchirappalli, India

ayaskanta.ed2305006@iimtrichy.ac.in

Vinu Cheruvil Thomas

Associate Professor

Indian Institute of Management Tiruchirappalli

Tiruchirappalli, India

vinu@iimtrichy.ac.in

The airline industry has traditionally utilized questionnaires and surveys to discern customer needs for enhancing service quality. However, adapting to the evolving customer perceptions of airline service quality continues to pose a significant challenge. Recently, User-Generated Content (UGC) has emerged as a promising alternative for identifying customer needs, capturing personal opinions, experiences, and preferences related to services or products. Despite this, significant research has mainly concentrated on quantitative ratings found on websites, resulting in inconsistent and incomplete assessments of airline service quality. Traditional methods such as questionnaires, surveys, and quantitative ratings fail to fully capture the complexities of airline service quality. To tackle this challenge, we aim to precisely map the service quality dimensions by extracting additional information embedded in unstructured online reviews using Latent Dirichlet Allocation (LDA). Our proposed method integrates both numeric ratings and extensive textual reviews to analyse sentiments across various airline service quality dimensions identified in the initial mapping phase. This paper illustrates how topic modelling and sentiment analysis on user-generated content can be applied to comprehend airline service quality and customer sentiments concerning quality attributes. The expected findings include identifying crucial attributes responsible for airline service quality and clarifying the sentiments linked to different service quality dimensions. These insights will enable service quality managers to make informed, data-driven decisions for continuous improvement. Moreover, this method will empower marketing managers to effectively manage large-scale data in interactive, action-oriented marketing research.

Keywords: User Generated Content, Service Quality, Latent Dirichlet Allocation, Sentiment Analysis

OPTIMIZING OPERATIONAL EFFICIENCY IN LAP CREDIT DEPARTMENT: AN EMPIRICAL DEA APPROACH

Deepika Bisht*

Student

Jaipuria Institute of Management
Noida, India

Deepika.bisht.25n@jaipuria.ac.in

(* Corresponding Author)

Aayushi Kalra

Assistant Professor

Jaipuria Institute of Management
Noida, India

Aayushi.kalra@jaipuria.ac.in

Pragya Gupta

Associate Professor

Jaipuria Institute of Management
Noida, India

pragya.g@jaipuria.ac.in

In the competitive banking financial sector, optimizing operational efficiency is critical for maintaining high service standards and achieving organizational goals. This study focuses on evaluating the performance of the Loan against Property (LAP) credit department at one of the biggest Private Sector Commercial Bank in India using DEA. Despite their potential to enhance profitability and efficiency, DEA remains underutilized in the industry due to its complexity. This study aims to bridge gap between DEA's theoretical potential and its practical application to make it more accessible to managers. The research involves collecting data on the performance of credit managers for a month. The Variable Returns to Scale (VRS) model of DEA was used for the analysis. VRS accounts for non- proportional relationship between inputs and outputs. The findings highlight the best practices and areas needing improvements, guiding the implementation of targeted strategies to enhance overall efficiency, service delivery and a competitive edge. The study provides insights into the efficiency of LAP Credit team at bank but also demonstrates the application of DEA in a complex organisational setting.

Keywords: Data Envelopment Analysis, Returns to Scale, Efficiency, Banks

A FUSION OF AI AND HUMAN EXPERTISE FOR ENHANCED CODE GENERATION AND DEBUGGING

Harshitha H M

AI/ML Data Science

Genpact

Bengaluru, India

Harshitha.M@genpact.com

Prudhvith Tavva

AI/ML Data Science

Genpact

Bengaluru, India

tavva.prudhvith@genpact.com

Prakash Selvakumar

AI/ML Data Science

Genpact

Bengaluru, India

Prakash.Selvakumar@genpact.com

This paper presents a novel semi-automated approach to code generation and bug fixing, uniquely blending artificial intelligence (AI), human expertise, and adaptive learning. Addressing the escalating complexity of software development, our approach combines AI-driven code planning, generation, and optimization with crucial human oversight. Unlike traditional reinforcement learning, our system employs a unique prompt-based approach, enabling a large language model (LLM) to directly learn and apply human feedback. Our system distinguishes itself by leveraging Azure OpenAI for seamless integration of GPT model with the incorporation of prompt engineering techniques to refine AI responses. A comparative analysis against AutoGen and CrewAI, using a diverse dataset ranging from simple to complex internet-sourced and NLP-related problem statements, demonstrates a 30% reduction in bug-fixing time and a >40% improvement in user satisfaction. This superior performance

stems from our system's ability to adapt to human input in real-time, facilitated by prompt engineering and enhanced by Explainable AI (XAI) techniques such as SHAP, which provide transparency into the LLM's decision-making process. A single, multi-functional LLM handles planning, code generation, library identification, documentation, and code optimization, leveraging a codebase cache for efficient auto-completion and Q&A. A dedicated code modifier function ensures adaptability to evolving user requirements. The system's iterative refinement loop, triggered by inaccuracies or user intervention, further refines code quality. Our hybrid model, combining AI capabilities with human oversight, offers a significantly more efficient and effective software development methodology. Future work will focus on expanding adaptive learning techniques to further refine our system's real-time responsiveness and adaptability.

Keywords: Code generation, Bug fixing, Human-in-the-loop, Reinforcement learning, Explainability.

REVOLUTIONIZING CONTRACT MANAGEMENT WITH GENERATIVE AI - UNCOVERING LOOPHOLES WITH PRECISION AND SPEED

Koundinya N V S S

Senior Data Scientist

Genpact

Hyderabad, India

ORCID: 0000-0003-1663-3112

Meghana Kurra

Data Scientist

Genpact

Hyderabad, India

kurra.meghana@genpact.com

B Lakshit

Data Scientist

Genpact

Hyderabad, India

b.lakshit@genpact.com

Upamanyu Deepak Wable

Data Scientist

Genpact

Hyderabad, India

upamanyu.deepakwable@genpact.com

Kolli Nethre Sai

Senior Data Scientist

Genpact

Hyderabad, India

kollinethre.sai1@genpact.com

Chirag Jain

Principal Data Scientist

Genpact

Bengaluru, India

chirag.jain4@genpact.com

Contracts are vital to business operations yet managing them is labor-intensive and prone to errors. This paper explores how generative AI can revolutionize Contract Lifecycle Management by automating loophole detection. We evaluated six prompting techniques across three language models—GPT-4o, Gemini-1.5 Flash, and Llama-3.2—using a dataset of 50 contracts from healthcare, banking, and insurance sectors. Advanced prompting significantly improved the model's performance to identify ambiguous language, missing clauses, unenforceable terms, and lack of consideration. GPT-4o achieved the highest F1-score of 90% with Tree-of-Thought Prompting. However, sophisticated prompts increased cost as well as processing time, highlighting a trade-off between accuracy and efficiency. Precision, recall, and F1-score demonstrated the models' effectiveness. Results suggest that integrating AI with optimized prompting can reduce review times, lower error rates, and mitigate contractual risks. Thus, this research addresses a gap in applying AI and prompting techniques to legal contract analysis, offering practical insights for enhancing contract management systems.

Keywords: Contract Lifecycle Management, Prompting Techniques, Contractual Loopholes, Language Models

KEY PHRASE EXTRACTION FROM CHAIRMANS SPEECH USIN TRANSFORMERS

Subhasis Dasgupta

Dept of Data Science

Praxis Business School

Kolkata, India

subhasisdasgupta1@acm.org

Jaydip Sen

Dept of Data Science

Praxis Business School

Kolkata, India

jaydip.sen@acm.org

Chairman's speech is an important section of the annual report of a company. In this section the chairman tries to highlight the key areas where the company investigated and some key aspects which the company tried to address in the previous financial year. That is why, chairman's report and director's report are considered important for analyzing the performance of the company. The key words or phrases used in this report are also relevant for structured analysis. The current work focuses on the extraction of these keywords in an automated manner by using the state-of-the-art transformers models. The models hyperparameters were tuned for getting the best results. The dataset was manually created by going through about 105 annual reports and custom models were fine tuned for the task in hand.

Keywords: Text annotation, DistilBERT, INCEPTION, Tokenization, Embedding, Token classification.

BENCHMARKING ADVANCED VIDEO GENERATION MODELS: EVALUATING EFFECTIVENESS IN CREATIVE MEDIA AND CORPORATE COMMUNICATION

Upamanyu Deepak Wable

Data Scientist
Genpact
Hyderabad, India
upamanyu.deepakwable@genpact.com

B Lakshit

Data Scientist
Genpact
Hyderabad, India
b.lakshit@genpact.com

Meghana Kurra

Data Scientist
Genpact
Hyderabad, India
kurra.meghana@genpact.com

Kolli Nethre Sai

Senior Data Scientist
Genpact
Hyderabad, India
kollinethre.sai1@genpact.com

Koundinya N V S S

Senior Data Scientist
Genpact
Hyderabad, India
ORCID: 0000-0003-1663-3112

Chirag Jain

Principal Data Scientist
Genpact
Bengaluru, India
chirag.jain4@genpact.com

This study explores the capabilities and limitations of recent open-source video generation models for creative media and corporate applications. Although text-to-video models have advanced significantly in recent years, challenges persist in achieving realistic video outputs across dimensions such as physical accuracy, smooth transitions, and consistent object representation. We selected three open-source models— CogVideoX-5B, Pyramid-flow-sd3, and Zeroscope-V2—based on their potential for practical applications and evaluated them using a set of prompts designed to assess physical realism, transition quality, object depiction, and suitability for creative use. Findings reveal that while CogVideoX and Pyramid-flow show promise for short-form applications like training animations and advertisements, there remain limitations, particularly in generating longer, coherent scenes. This study highlights the progress in open-source video generation while identifying areas where further improvement is needed.

Keywords: text-to-video generation, creative media application, generative AI, animated training creation

ENHANCING ANALYSIS OF REVIEWS WITH DEEP LEARNING AN EXPLAINABLE AI FOR SELECTING HOTELS ON SPECIFIC ASPECTS

Shaukat Ali Shahee

Assistant Professor,
Information Technology and Systems,
IIM Kashipur, Kashipur, India

Dillip Rout

Assistant Professor,
Department of Computer Science and Engineering,
C.V. Raman Global University, Bhubaneswar,
India

NO ABSTRACT IN MANUSCRIPT

LEVERAGING WORKPLACE IT AND BEHAVIOURAL ANALYTICS FOR CUSTOMISED L&D PLANS

Ms. Amrita Shil

PGDM Business Design 2023-25

Welingkar Institute Of Management Development & Research, Mumbai
Mumbai, India

BD2023-amrita.shil@welingkar.org

Employee performance management is essential for organizational success, especially in today's dynamic global business environment. One of the key challenges in workforce performance management is the need to provide Learning and Development (L&D) plans that are tailored to the specific needs of individual employees. In the era of behavioral analytics, traditional data-driven approaches are becoming outdated, as behavioral analytics offers deeper insights into employee performance. This paper introduces a system architecture for a behavioral analytics system that leverages workplace IT infrastructure and integrates with Learning Management Systems (LMS) and Performance Management System (PMS) to deliver real-time, customized L&D plans. Additionally, the paper examines the business implications and potential challenges in implementing such a system, providing insights into its transformative potential in human resource management.

Keywords: Behavioural analytics, Customized L&D plans, Learning management system, Performance management system, Workplace IT ecosystem

AUDIO2ACTION: A FRAMEWORK FOR AUTOMATED TASK & INSIGHT GENERATION FROM SPEECH

Sukanya Bag

Data Science and Insights
Genpact
Bengaluru, India
[sukanya.bag1 @genpact.com](mailto:sukanya.bag1@genpact.com)

Subhadip Mandal

Data Science and Insights
Genpact
Bengaluru, India
subhadip.mandal@genpact.com

Ankit Sati

Data Science and Insights
Genpact
Bengaluru, India
ankit.sati@genpact.com

Chirag Jain

Data Science and Insights
Genpact
Bengaluru, India
chirag.jain4@genpact.com

Business Challenge: In today's fast-paced business environment, audio content, including meeting recordings, webinars, and knowledge-sharing sessions, is essential for facilitating information exchange and supporting decision-making processes across various sectors such as business, academia, and industry. A universal challenge emerges when individuals face the labor-intensive and time-consuming task of going through lengthy audio content to stay informed and aligned with ongoing projects or educational curricula. This affects not only businesses on a large scale, but also academic institutions, healthcare providers, and various other organizations that rely on the efficient transfer of knowledge followed by decision-making, thus creating a significant productivity bottleneck. **Proposed Solution:** Our end-to-end plug-and-play framework addresses this challenge head-on and is capable of not only processing meeting recordings and extracting information out of them but also a wide array of informational audio content. Our system adeptly generates precise audio summaries, assigns, and generates tasks discussed in organizational team meetups, and identifies key highlights and insights from a variety of organizational and educational audio content, which becomes handy while generating documentation and operational procedures for future use. **Methodology:** Leveraging state-of-the-art transcription and diarization techniques powered with Large Language Models (LLMs), our application transforms audio-video recordings into concise audio digests. The application integrates advanced audio processing, followed by transcription, summarization, and audio generation. The key feature of the system is the plug-and-play of the components, which allows developers to add and remove tools based on requirements. The application is built following popular software design principles and follows a CI/CD workflow for easy integration and testing of new features seamlessly. It also has a set of defined tests to ensure the proper working of the application after major changes. The application's broad compatibility with a variety of audio/video content ensures it can be adopted widely, by businesses to boost employee productivity, decision-making and operational workflow/procedure creation with high efficiency. **Experiments and Results:** One of the critical components of our framework is the integration of Machine Learning Operations (MLOps) practices, which ensures the continuous integration and delivery of the integrated tools. This allows our system to adapt quickly to new data inputs, improving the accuracy and efficiency of audio/ video processing tasks. The use of containerization technologies like Docker our application to scale horizontally, handling multiple users and large audio files in parallel without compromising performance. This scalability is crucial for large organizations with extensive audio archives and high-frequency recordings. Extensive experimentation and testing on the real-world scenarios discussed, have shown promising results on our

framework, and is seen to significantly streamline the process of information extraction and sharing, leading to enhanced collaboration and productivity in any context where information sharing is vital. The application is designed to be easily scalable and handle multiple users parallelly. Next Steps: As we advance, our future objective is to integrate the application directly with popular asynchronous work and project management tools to automatically assign tasks to users and update existing tasks based on the updates from recordings.

Keywords: Speech2Text, Audio Transformers, LLM, MLOps

FEDERATED LEARNING FOR COMPLAINT CLASSIFICATION: NAVIGATING NON-IID TEXTUAL DATA CHALLENGES AND ENSURING PRIVACY AND EXPLAINABILITY

Honey Yadav

Doctoral Scholar

Indian Institute of Technology Delhi

New Delhi, India

yadav.honey19933@gmail.com

Ekta Srivastava

Postdoctoral Researcher

Indian Institute of Technology Delhi

New Delhi, India

ektasri@iitd.ac.in

Sandeep Kumar

Assistant Professor

Indian Institute of Technology Delhi

New Delhi, India

ksandeep@ee.iitd.ac.in

Federated learning (FL) architecture presents a promising solution to complaint classification problems by enabling collaborative model training across multiple institutions without sharing raw data, potentially enhancing classification accuracy while maintaining data privacy. However, the challenge of Non-IID (Non-Independent and Identically Distributed) data remains a significant concern in achieving optimal model convergence and accuracy in federated environments. This study explores the application of FL in financial complaint classification, focusing on the effects of Non-IID data distribution across multiple clients. The dataset contains financial complaints related to transaction failures, annotated with attributes such as complaint labels, severity levels, sentiments, and emotions. The primary aim of this study is to investigate and compare how uneven and even data distribution affects the performance of a federated learning model in a classification task. Using Latent Dirichlet Allocation (LDA), we simulate uneven data distribution among five clients to mimic real-world scenarios where each client possesses different proportions of complaint types, sentiments, and emotions. LDA-based data distribution enables us to create clients with data that vary significantly in characteristics such as sentiment polarity (e.g., negative vs. neutral) and emotion (e.g., sadness, anger, surprise). We employ BERT (Bidirectional Encoder Representations from Transformers) for the sequence classification task. Each client trains a local BERT model using its assigned data, and these models are periodically aggregated into a global model using three distinct aggregation methods: simple averaging, weighted averaging based on client data sizes, Krum aggregation, and federated stochastic gradient descent (FedSGD). The models are evaluated based on accuracy over multiple rounds of federated learning to track how the global model adapts to Non-IID client data. Our method includes preprocessing the text data by tokenizing it with a BERT tokenizer and converting the texts to TF-IDF (Term Frequency-Inverse Document Frequency) representations for LDA-based client assignments. The training process involves fine-tuning BERT models for five epochs per client, followed by global model aggregation after each round of training. The findings highlight the importance of the aggregation method in federated learning setups, particularly in scenarios where data distribution among clients is heterogeneous. Krum aggregation, with its resilience to outliers, provided a more consistent improvement in the global model's accuracy. These results support the dilemma among organizations with uneven/inadequate data availability that can achieve reliable global model performance with high privacy and explainability. The experimental results demonstrate that Krum aggregation outperforms simple and weighted averaging in achieving a stable and robust global model. Furthermore, findings have yet to be evaluated, with an increase in the number of rounds embedded with the differential privacy approach and an increase in interpretation using SHAP to establish the arguments by building explainable systems. Moreover, even data distribution among clients showed more stable and robust performance over 30 rounds with improved reliability. At round 0, the model aggregated accuracy (simple averaging) is 96.34%, and weighted aggregated accuracy is 96%, which represents better baseline performance as compared to uneven data distribution. These results conclude that

even data distribution among clients has better performance and room for improvement in the FL process with the increased number of rounds.

Keywords: Federated Learning, Latent Dirichlet Allocation (LDA), Topic Modeling, Complaint Classification, Explainable Systems

INTEGRATION OF CHATBOT BASED NON-MEDICAL SERVICES IN HOSPITALS

Abhilash V S

CET School of Management
College of Engineering Trivandrum

Raju G

Professor
Xavier Institute of Management and Entrepreneurship

DEEPIKA D NAIR

CET School of Management CET School of Management
College of Engineering Trivandrum

The integration of chatbot technology across various industries has transformed customer service practices in the recent past. The effectiveness and implications of chatbot-based customer support in hospital environments are examined in this study. This study aims to identify the complex factors influencing customer satisfaction, engagement, and healthcare outcomes by looking at the use of chatbots for customer interactions in hospitals. The purpose of this study is to explain the key elements influencing adoption and efficiency of chatbot-based customer service in improving hospital operations by analyzing chatbot functionality, user experiences, and feedback systems. The study used multiple regression analysis on a sample of 121 individuals who were using hospital services, and the results showed a substantial relationship between the chatbot's performance, the decrease in waiting times, customer perception, and trust in the chatbot's knowledge.

eywords: Chatbot and AI, nonmedical services, customer satisfaction

ENHANCING CUSTOMER SERVICE WITH REAL-TIME FEEDBACK AND VERIFICATION SYSTEM

Subhadip Mandal

Data Science and Insights

Genpact

Bengaluru, India

subhadip.mandal@genpact.com

Sukanya Bag

Data Science and Insights

Genpact

Bengaluru, India

sukanya.bag1 @genpact.com

Chirag Jain

Data Science and Insights

Genpact

Bengaluru, India

chirag.jain4@genpact.com

Ankit Sati

Data Science and Insights

Genpact

Bengaluru, India

ankit.sati@genpact.com

Business Challenges: Current customer service systems face several critical issues that hinder their ability to provide efficient customer satisfaction. Prolonged response times often result from the need for service agents to manually search and verify customer information, leading to delays in addressing customer queries and requests. High operational costs are another significant challenge, as manual processes require more time and resources. Manual verification errors further intensify these issues, potentially leading to security concerns and customer dissatisfaction. Additionally, inconsistent service quality is a common problem, as different agents may have varying levels of skill and experience, resulting in uneven customer experiences. These challenges collectively contribute to reduced customer satisfaction and operational inefficiency. **Proposed Solution:** This paper discusses a system designed to improve customer service interactions through the integration of advanced audio processing, transcription, sentiment analysis, actionable feedback for service agents, and a real-time automatic customer verification feature. The proposed methodology consists of five key stages: uploading and preprocessing audio files, converting speech to text, analyzing sentiment, generating feedback and communication suggestions, and fetching user information from a database automatically in real-time to help the agent verify the customer.

This system supports popular audio file formats and integrate seamlessly with existing service workflows to enhance agent efficiency and customer satisfaction. By addressing the mentioned business challenges, this system eliminates the need for manual searches and verifications, thereby reducing response times, lowering operational costs, minimizing verification errors, and ensuring more consistent service quality. The system's capability to provide feedback and suggestions to service agents further helps improve their communication with customers, ultimately enhancing overall customer satisfaction and operational efficiency.

Methodology: The system supports popular audio file formats, which are pre-processed for compatibility with transcription services such as Azure Real-time Diarizer and OpenAI's Whisper. A significant feature of this system is its use of Large Language Models (LLMs) for sentiment analysis and

feedback generation. Additionally, the system employs SQL agents to securely and efficiently verify customer identity in real-time by retrieving user information from a database based on the conversation. This not only improves security but also streamlines the service process.

Experiments and Results: Extensive testing and experimentation were conducted to evaluate the system's performance in real-world scenarios. The integration of transcription services, sentiment analysis and feedback tools demonstrated high accuracy in converting speech to text, identifying customer sentiments and providing feedback to the agent based on the customer sentiments. The real-time customer verification feature proved effective in reducing the time service agents spent on manual verification processes, leading to quicker resolution of customer queries and enhanced overall service efficiency.

Next Steps: This research aims to show how Generative AI can facilitate better customer service outcomes, emphasizing the importance of integrating such systems for the advancement of the industry. Future work will focus on refining the system's sentiment analysis capabilities and further enhancing the security and efficiency of the real-time customer verification process. Additionally, exploring the integration of more advanced AI models and expanding the system's adaptability to various customer service environments will be key areas of development.

Keywords: Audio, Sentiment Analysis, LLM, MLOps

ASSESSING BAYESIAN AND MULTI-FIDELITY HYPER PARAMETER OPTIMIZATION STRATEGIES FOR SEVERAL MACHINE LEARNING MODELS

Shailendra Kadre

School of Business and Management (SBM)

Christ University

Bangalore, India

shailendrakadre@gmail.com

Hyper parameter Optimization (HPO) has been a hot research area amongst machine learning researchers. Grid Search (GS) and Random Search (RS) HPO methods are prevalent amongst machine learning (ML) practitioners across the industry. These methods require more resources and iterations than Bayesian Optimization (BO) and Multi-fidelity (MF) HPO methods. Still, the BO and MF methods adoption is not as widespread (in the industry) as GS and RS.

Several research efforts have evaluated various HPO methods on standard datasets. Still, very few researchers have assessed these HPO methods specifically for the datasets in the domain of banking and finance. These datasets often have a high-level class imbalance ratio. The current research uses five finance domain datasets to evaluate BO and MF families of HPO methods. Multiple model evaluation criteria are used, such as ROC AUC (Receiver Operating Characteristic Area Under the Curve) and CPU time to measure the performance of these HPO methods.

Our two hundred and forty independent experiments with eight optimizers and six machine learning (ML) algorithms show that it is possible to get the same results with any optimizers provided you use Random Forests (RF) or XGBoost (XGB) algorithms. Experimental results across the five datasets show that RF with BOHB optimizer provides the best results and the best program runtimes. Similar trend was observed for all the five datasets regarding all the experiments. It is sufficient to support the consistency and authenticity of the conclusions made in this research project. To the advantage of practitioners, it is adequate that they go with only one experiment of running RF with BOHB to get the best optimization results. Logistic Regression (LR) may still be used to get average results if a baseline is required.

The actionable conclusions made in this project can save considerable effort, time, and resources for the industry practitioners and the ML researchers attempting ML model optimization for business and scientific research problems.

Keywords: Machine learning, algorithms, ANN, optimization, model parameters & hyper parameters, hyper parameters tuning, Bayesian optimizers, multi- Fidelity optimizers

EFFICIENT PII REMOVAL WITH ZERO-SHOT NER AND DUAL ENCODERS

Nitin Agarwal

Microsoft

Bangalore, India

mnitin3@gmail.com

Rutvik Acharya

Atlassian

Bangalore, India

rutvik25@gmail.com

Personally Identifiable Information (PII) removal is a critical task in data privacy and security, requiring the identification and redaction of sensitive entities such as names, addresses, and social security numbers from unstructured text. Traditional Named Entity Recognition (NER) models used for PII removal are limited to predefined entity types, necessitating retraining for each new PII category. This paper presents zero-shot NER architectures that enable the efficient removal of any type of PII without extensive retraining.

We leverage two advanced architectures for zero-shot NER in the context of PII removal: bi-encoder and poly-encoder models. The bi-encoder architecture separates the encoding of input text and PII entity types into distinct transformer models, allowing for efficient and scalable processing. PII entity type encodings can be pre-computed and reused across different input texts, reducing computational overhead. The poly-encoder architecture enhances the bi-encoder approach by incorporating a post-fusion step to model interactions between input text and PII entity representations explicitly, addressing the lack of interentity understanding in standalone bi-encoder models.

To evaluate the effectiveness of these architectures for PII removal, we conduct experiments using a diverse, high-quality dataset containing various types of PII. We compare the performance of our proposed models with existing zero-shot NER approaches, such as GLiNER, in terms of precision, recall, and F1 score. The results demonstrate that our bi-encoder model outperforms GLiNER in identifying and removing PII entities, setting a new benchmark for zero-shot NER in the context of data privacy and security.

These architectures offer several advantages for PII removal, including the ability to recognize an unlimited number of PII entities simultaneously, faster inference with preprocessed PII entity embeddings, and better generalization to unseen PII categories. These advancements enable the development of efficient and scalable PII removal systems capable of handling diverse and evolving PII requirements, ensuring compliance with data privacy regulations and protecting sensitive information.

In this paper, we present an adaptive approach to PII detection that dynamically selects between GLiNER and Presidio models based on contextual analysis. Our methodology first analyzes input text for regional markers, script patterns, and format variations to determine the most suitable model for PII detection. GLiNER is prioritized for Western contexts and standardized formats, while Presidio handles region-specific and non-standard patterns. This context-aware selection is complemented by a robust validation framework that includes both primary and secondary validation layers, confidence scoring, and enhanced processing for ambiguous cases. Experimental results demonstrate a 12%-14% improvement in overall accuracy compared to single-model approaches, with particularly strong performance in handling diverse regional formats and multi-script environments, while maintaining acceptable processing overhead.

Keywords: Data privacy, Personally Identifiable Information, Deep Learning, Named Entity Recognition, Poly-encoder architecture

SURVEY OF LARGE LANGUAGE MODELS : CAPABILITY, EFFICIENCY AND SECURITY

Yug Makhecha and Dhaval Mehta

School of Engineering and Technology

Navrachana University

Vadodara, India

dhruvin.83@gmail.com

Large language models (LLMs) are a category of AI based models which are trained on an enormous amount of data, making them capable to interact with humans using Natural Language Processing and making them capable of understanding, summarizing, generating and predicting new content in the form Text, Audio, Video. These models are based on transformer which learns context and meaning by tracking relationships in the sequential data. Sometimes It uses Recurrent Neural Networks(RNN), Generative Adversarial Networks (GAN), Autoregressive Transformers, and diffusion models. IT Companies like OpenAI, Meta, Google, Microsoft, and Mistral have introduced their version of LLM for the enterprises and users. Text generation and embedding models manage text in small pieces which are called tokens. Tokens represent commonly occurring sequences of characters which is nearly upto 4 characters. The efficiency here is heavily dependent upon the Token Size for the Models. LLM use cases for text-based content can be classified into : Generation, Summarization, Translation, Classification, Sentiment analysis and Chatbots. Additionally, companies have introduced Large Language Vision Models(LLVM) especially for generating, analysing the Images and Videos. These models are capable enough to manage visual quality and relevance to the user's prompt. Some of these are open source while some are charging usage wise(According to Token Size used for Input and Output). Even these models have some important features and these features are differentiating them with their competitors. In this paper we have compared, analysed the functionalities of various LLM and LLVM. In addition to that we have also compared them by considering the various criteria like Pre-Training, Prompt Engineering, Retrieval augmented generation (RAG), Fine-Tuning and Hallucination by applying them to solve the particular tasks.

Keywords: Artificial Intelligence, Natural Language Processing, Generative AI, Large Language Models, ChatGPT

EVALUATION AND PREDICTION OF SUSTAINABLE LIVELIHOOD CAPITAL - A CASE STUDY

Vasantha Rao Sappati

Ophthalmic, Asst,

Department of Medical and health,

Andra Pradesh

vasanthophthalmicofficer@gmail.com

Narayana Rao Kandukuri

Principal, GPT, Visakhapatnam

Visakhapatnam, India

kandukuri67@gmail.com

The concept of sustainable livelihood capital, comprising human, social, natural, financial, and physical assets, is essential for understanding community resilience and development in rural and vulnerable populations. This study aims to identify and validate measurement items for assessing sustainable livelihood capital using by engaging diverse stakeholders, including subject matter experts, community leaders, development practitioners, and local beneficiaries.

This study employs the Combined Compromise Solution (CoCoSo) method, a Multi-Criteria Decision-Making (MCDM) technique, to assess the sustainable livelihood capital of a specific region. Discriminant analysis is applied to build a model that classifies the observations into distinct groups based on their livelihood capital levels. A case study of a Village in Srikakulam is presented to identify the category of sustainable livelihood capital that the village belongs to. The Results of the Discriminant analysis are compared by Artificial Neural network Analysis.

Keywords: Sustainable Livelihood, MCDM, Discriminant Analysis, Livelihood Capitals, Evaluation, Prediction

PEDAGOGICAL EFFECTIVENESS IMPACTING LEARNING OUTCOME - AN ANALYSIS WITH FORMATIVE ASSESSMENT

Maitreyee Das

Alliance School of Business
Alliance University
Bangalore, India
m.das@live.com

Heena Bhardwaj

Alliance School of Business
Alliance University
Bangalore, India
89heenabhardwaj@gmail.com

Abhishel Dutta

Department of Environmental Science
Chulalongkorn University
Bangkok, Thailand
duttabob@gmail.com

Gautam Dutta

Centre for Northeast Studies
Indian Institute of Foreign Trade
Kolkata, India
gautamiift@gmail.com

Effectiveness of pedagogical tools has always remained an area of research interest for academicians. The movement towards outcome-based education, technological advancement and subsequent changes in the behavior and the expectations in the learners community has forced the policy makers and educationists to come up with more innovative ways of course delivery to make the classroom experience more enriching. Formative assessment is a way of ongoing evaluation and feedback process which entails more interaction between the instructors and the learners resulting in an enhanced learning outcome. Using partial least square structured equation modelling this case study was conducted to evaluate the effectiveness of formative assessment for the postgraduate management students in an Indian business school. The result of our study showed that effectiveness of formative assessment has a positive correlation with students' level of satisfaction, and it also positively impacts their intention in actively engage in the class which in turn results in the enhanced learning outcome.

Keywords: Outcome Based Education, Learning Outcome, Formative Assessment, Students' Satisfaction, Pedagogy, Continuous Evaluation and feedback

BUSINESS ANALYTICS AND FIRM PERFORMANCE: A HYBRID REVIEW OF EXISTING LITERATURE

Hina Siddiqui and Ayesha Farooq

Despite the growing interest in the relationship between business analytics and firm performance, the literature is very limited. No attempt has been made to systematically synthesise the literature on the relationship between business analytics and firm performance. The objective of this study is to significantly enhance the understanding of the key facets of the relationship between business analytics and firm performance by conducting hybrid review of the concerned literature and propose agenda for future research. We have used Scientific Procedures and Rationales for Systematic Literature Reviews (SPAR 4 SLR) protocol proposed by Paul et al.,2021, a framework based systematic literature review technique to expand the body of knowledge by providing detailed insight into the theoretical bases, contexts and methods. Additionally, using bibliometric mapping to identify the most productive authors, sources, countries and affiliations. In addition, conceptual structure of literature is identified using co-occurrence network, intellectual structure is presented using co-citation network and social structure is presented using collaboration network and countries' collaboration map. We have used thematic maps to identify key themes. Further, we have provided future research directions using theories, contexts and methods framework. The article concludes with the limitations of the study. This hybrid study makes a significant contribution to the business analytics and firm performance research by highlighting the development of literature and focussing on most critical research fronts that were not assessed by previous literature assessments.

Keywords: business analytics; firm performance; SPAR-4 SLR; bibliometric review; systematic literature review

Analysing Barriers to AI and Automation Adoption in Retail: A Grey-DEMATEL Approach

Preeti Sagar and Madhvendra Misra

Retail is an essential part of our lives; we encounter it every day, often without even realizing it. As the world rapidly embraces technological advancements, the retail industry, too, has had to evolve. Retailers, striving to maintain their relevance, have increasingly adopted new technologies to enhance customer shopping experiences. These innovations have undeniably made life easier and more exciting. However, as with everything, there are two sides to the coin. It has also disrupted established shopping behaviours in customers' minds, causing some to resist the use of these new technologies or alter their revisit intentions. In this proposed study, we will examine how the adoption of labour-saving innovations, such as artificial intelligence (AI) and automation, shopping apps, is transforming the retail industry. However, the integration of these technologies often encounters various barriers that can significantly impact established shopping behaviours. Most research primarily focuses on the impact of adopting technological innovations on customer satisfaction. This proposed study employs both grey system theory and the Decision-Making Trial and Evaluation Laboratory (DEMATEL) approach, a multi-criteria decision-making (MCDM) method to analyse the interrelationships among the barriers (identified from the literature) to adopting innovation in retail for the 'Gems and Jewellery' industry. The findings reveal that certain barriers, such as consumer resistance to change, technological barriers, product and performance barriers, time barriers, and financial considerations, play a pivotal role in shaping the adoption process. Understanding these factors is crucial for retailers aiming to leverage AI and automation to enhance customer experiences while minimizing disruptions to traditional shopping patterns. The study contributes to the literature by offering a structured framework for assessing and mitigating barriers to innovation adoption, thereby facilitating a smoother transition towards automated retailing solutions.

Keyword: labour saving innovations, artificial intelligence & automation, grey theory, dematel.

GENERATING ACTIONABLE INSIGHTS FROM TICKET DATA FOR IT OPERATIONS

L. Janardhan Rajan

Advanced Analytics
TechM Consulting
Tech Mahindra
Bengaluru, India

Rajan.Janardhan@TechMahindra.com

Dr Satyajit Acharya

Advanced Analytics
TechM Consulting
Tech Mahindra
Bhubhaneswar, India

Satyajit.Acharya@TechMahindra.com

Nisha Nirmal Gangwani

Advanced Analytics
TechM Consulting
Tech Mahindra
Bengaluru, India

NG00769202@techmahindra.com

Meghna Zutshi

Advanced Analytics
TechM Consulting
Tech Mahindra
Bengaluru, India

MZ00614727@techmahindra.com

Sagareeka Soni

Advanced Analytics
TechM Consulting
Tech Mahindra
Bengaluru, India

SS00347214@TechMahindra.com

IT Operations (including Service Management teams), across enterprises and carriers, often face a plethora of challenges like negative customer perception, process bottlenecks and inefficiencies, SLA breaches, data quality issues, process non-compliance, inadequate staffing, ascertain the effectiveness and impact of Service Improvement Plan implementations. Signatures of all such issues are found in ticket data through prudent application of Statistics, Natural Language Processing, Process Mining and other custom algorithms. Ticket extracts from the Service Desk when subjected to a slew of statistical techniques and algorithms can reveal existing issues in an IT Operations setup. Comprehensive improvement plans can be initiated based on such insights.

The authors of this paper are a team of ITIL [1] Practitioners and Data Scientists who have successfully developed an award-winning open-source product called Snoopy - a comprehensive IT Service Analytics solution, designed for rigorous analysis of ticket data and built using R. The product has found acceptance across multiple fortune 500 customers. Furthermore, services of the team are frequently called upon to support large deals, transformational projects, and operational assessments.

This paper discusses different categories of insights generated and value delivered based on the analysis done using Snoopy.

Keywords: IT Service Analytics, Snoopy, Service Improvement, Process Mining, Sentiment Analysis

PROFILING CUSTOMERS USING RFM MODELLING FOR CREDIT CARD HOLDERS USING K-MEANS CLUSTERING

Dr.J.Deepa

Assistant Professor
Department of Management Studies
Coimbatore Institute of Technology
Coimbatore, India
deepaj@cit.edu.in

Dr.S.Thilaga

Assistant Professor
Department of Management Studies
Coimbatore Institute of Technology
Coimbatore, India
thilaga.s@cit.edu.in

Dr.D.Kavitha

Assistant Professor
Department of Management Studies
Coimbatore Institute of Technology
Coimbatore, India
kavithad@cit.edu.in

Dr.K.Swathine

Assistant Professor
Department of Management Studies
Coimbatore Institute of Technology
Coimbatore, India
swathinek@cit.edu.in

Purpose: As competition heats up, the banking industry is fostering customer behavior to improve their top lines on customer segmentation, targeting, and personalized marketing strategies for credit card holders. Consequently, estimating the Customer Life

Value (CLV) of credit card customers is a potential value that allows the bank to tailor their services, rewards, and retention.

Study design/methodology/approach: This paper describes how extended K-means clustering is performed on the credit card consumers to identify the target cluster group. Using a proposed approach, cluster red represents a high spending group, and also addresses the issue of the small class.

Findings: Banks can more effectively target customer segmentation based on purchase behavior, thereby increasing revenue, but also reducing unnecessary costs and also improving the efficiency of marketing efforts to existing customers.

Originality/value: Due to the fast development of online platforms, RFM (Recency, Frequency, & Monetary) is one of the techniques to profile the credit card customer into various recent transactions, frequency, amount, spending pattern and utilization.

Keywords: Credit Card Customer, Customer Life Value, RFM.

EVALUATING NARRATIVE SENTIMENT AGAINST PERFORMANCE IN INDIAN FINANCIAL SERVICES

Mohit Pareek

MSE Business School
Madras School of Economics
Chennai, India
phd21mohit@mse.ac.in

Sudharsan S P

MSE Business School
Madras School of Economics
Chennai, India
pgdm22sudharsan@mse.ac.in

Rakesh Nigam

MSE Business School
Madras School of Economics
Chennai, India
rakesh@mse.ac.in

This study investigates the effectiveness of annual reports as communication tools by examining the relationship between narrative sentiment in Management Discussion and Analysis (MD&A) sections and financial performance in the Indian financial services sector. Using a state-space framework combined with sentiment analysis, we examine 450 annual reports from 30 companies over 15 years (from year 2009-10 to year 2023-24), focusing on communication patterns during significant economic events such as demonetization, Goods and Services Tax (GST) implementation, and the COVID-19 pandemic. Our findings reveal that while institutions generally maintain alignment between sentiment and performance, they demonstrate strategic flexibility during economic disruptions. Traditional banks show more stable communication patterns, while non-banking financial companies (NBFCs) demonstrate greater adaptability, particularly in recent years. These insights contribute to our understanding of financial communication strategies in emerging markets and may guide future practices in corporate reporting for more effective stakeholder communication.

Keywords: Sentiment Analysis, State-Space Framework, Financial Performance, Economic Events, Indian Financial Services

REVOLUTIONIZING CYBERSECURITY WITH GENERATIVE AI: ADAPTIVE SYSTEMS FOR NON-COMPLIANCE & THREAT DETECTION

Devesh Srivastava

Business Leader & Gen AI Architect
Cisco Systems
Boston, USA
devesriv@cisco.com

Sambuddha Das

Customer Delivery Software Architect
Cisco Systems
Bangalore, India
sambudas@cisco.com

The growth of health technology and increased data sharing between healthcare providers and patients have introduced significant compliance and security challenges, particularly around HIPAA (Health Insurance Portability and Accountability Act) compliance. This project aims to develop an AI-driven compliance and security breach detection system for healthcare industries, a health-monitoring platform, by using a multi-layered approach. The system is designed to identify potential regulatory violations, unauthorized access patterns, and emerging security threats within physiological data, ensuring HIPAA adherence and data security.

CONFORMAL PREDICTION APPROACHES FOR ACCURATE CARBON FOOTPRINT FORECASTING

Nitin Agarwal

Microsoft

Bangalore, India

mnitin3@gmail.com

Rutvik Acharya

Atlassian

Bangalore, India

rutvik25@gmail.com

Sanjana Das

GSK

Bangalore, India

dassanjana102@gmail.com

Climate change represents one of the most significant global challenges of our time, with carbon dioxide (CO₂) emissions being a primary contributor to the greenhouse effect and subsequent global warming. As the urgency to address climate change grows, the need for accurate and reliable estimation of CO₂ emissions becomes increasingly critical. Such forecasts are essential for policymakers, environmental planners, and global organizations as they develop and implement strategies to mitigate the adverse impacts of climate change.

This paper presents a comprehensive study on the application of conformal prediction to the task of forecasting CO₂ emissions. Conformal prediction offers a robust framework that goes beyond traditional point estimates, providing probabilistic predictions accompanied by well-defined confidence intervals. This approach allows for a more nuanced understanding of the potential future values of CO₂ emissions, capturing the inherent uncertainty and variability that often characterize these emissions.

Traditional forecasting methodologies, which typically generate single-point predictions, are limited in their ability to accommodate the complex, multi-faceted nature of CO₂ emissions. These methods often fail to account for the wide range of factors—such as economic activity, technological advancements, policy changes, and natural events—that can influence emissions. Consequently, they may produce forecasts that are either overly optimistic or pessimistic, leading to suboptimal policy decisions.

In contrast, the conformal prediction-based approach leveraged in this paper provides a range of possible outcomes, rather than a single deterministic value. This range-based forecasting technique is particularly valuable in the context of CO₂ emissions, where uncertainty is inherent due to the diverse and often interdependent factors at play. By offering predictions that include a spectrum of potential scenarios, our approach supports the development of more adaptive and flexible policies that are better suited to addressing the unpredictable nature of climate change.

In conclusion, the application of conformal prediction to CO₂ emission forecasting represents a significant advancement in the field of environmental forecasting. This method enhances the reliability of predictions and provides critical insights that can inform more effective and responsive climate policies.

Keywords: Machine Learning, Forecasting, Conformal Prediction, Confidence Calibration, Climate Impact Prediction

PREDICTING THE PERFORMANCE OF SITES FOR SOLAR POWER PLANT

Raman SVV

Govt Polytechnic

Vijayawada, India

svvramana369@gmail.com

Deva Kumar MIS

JNTU, Ananthapur

Ananthapur, India

devakumar.jntua@gmail.com

Site for solar power plant prediction is a multifaceted problem that involves analyzing numerous variables such as solar irradiance, temperature, humidity, and plant-specific parameters. Traditional methods often fall short in handling the complexity and variability inherent in these data sets. In this paper, performance of sites for solar power plants is determined through ratio and non-ratio-based methods of MCDM. Aggregate ranks of these solar power plants are obtained through ensembling rank method and derived from the categories of sites of solar power plants. The categories obtained are considered as dependent variable and twenty criteria for selection of solar powerplants are considered as independent factors. By implementing categorical regression analysis, category of sites of alternate solar power plants are predicted. The study will be useful to identify the best location among the alternates to establish the solar power plant.

Keywords: Solar Power Plant, Regression Analysis, MCDM

FACTORS INFLUENCING THE NIRF SCORE OF INDIAN MANAGEMENT INSTITUTIONS

Aishwarya Mahendra Kapase

Student
PES University
Bengaluru, India
mkaishwarya2730@gmail.com

Biplab Sarkar

Professor
PES University
Bengaluru, India
biplabsarkar@pes.edu

Higher education institute rankings play a crucial role in shaping the perceptions of students, parents, employers, and policymakers. They provide a benchmark for comparing the quality of institutions based on various criteria such as teaching, research, and international outlook. Rankings help students make informed decisions about where to study, ensuring they choose institutions that align with their academic and career goals. For universities, high rankings enhance their reputation, attract top talent, and foster international collaborations.

Additionally, rankings can influence funding and policy decisions, driving institutions to improve their standards and innovate in their educational offerings. In India, higher education institutes are ranked by the National Institutional Ranking Framework (NIRF), established by the Ministry of Human Resource Development in September 2015. Rankings have been conducted nine times for Indian higher education institutes from 2016 to 2024. Institutes are ranked based on five parameters: teaching, learning, and resources (TLR); research and professional practice (RP); graduation outcomes (GO); outreach and inclusivity (OI); and peer perception (PP). In 2024, Indian higher education institutes were ranked in 16 categories: overall, universities, colleges, research institutions, engineering, management, pharmacy, medical, dental, law, architecture and planning, agriculture and allied sectors, innovation, open university, skill university, and state public university. This research paper aims to investigate the factors influencing the NIRF score of Indian management institutions category, considering the data submitted by the institutions on the NIRF portal. The study uses a step-wise multiple linear regression (MLR) model on a dataset of 100 management institutes. It seeks to identify the key determinants that contribute to a B-school's NIRF score. The analysis explores the variation in NIRF scores based on multiple independent variables, such as total student intake, number of students placed, and median salary of placed students, PhD student intake, total amount of sponsored research, amount received from consultancy project, annual earnings from executive development programs (EDP) and management development programs (MDP), and number of faculty members.

In our model, the adjusted R-square was 0.763, with the probability of F-statistics below 0.05, which is fairly good. This means that 76.3 percent of the variation in NIRF scores is significantly explained by the explanatory variables considered in the model. Among these variables, the median salary of placed students, annual earnings from EDP/MDP Programs and the total amount of sponsored research play significant roles in determining the NIRF score. The findings of this research provide valuable insights for educational institutions to understand the critical areas that need to be prioritized for improvement to enhance their NIRF rankings.

Keywords: NIRF score, management institutions, step-wise MLR

SPILOVER NETWORK ANALYSIS IN THE APAC REGION USING MINIMUM SPANNING TREE

Kartikey Kashyape

Equity Research Analyst
IndSec Securities and Finance Limited
Mumbai, India
kartikkashyape@gmail.com

Maneesh Gupta

Research Scholar
IIM Mumbai
Mumbai, India
maneesh.gupta.2018@iimmumbai.ac.in / 0009-0001-4696-9832

This study analyses spillover networks across six global economic and financial shocks on thirteen different Asia-Pacific countries' stock market indices using Minimum Spanning Tree. Prim's algorithm is used for making the Minimum Spanning Tree network diagram. The main contribution of the study is to help understand the variability in the connectedness structure of the stock indices in the Asia-Pacific region across six crisis periods. After analysing the six events, Singapore's Straits Time Index appeared to be the central hub of contagion as the indices in the Asia-Pacific region are very closely connected and form clusters during economic shocks. The reason for Singapore emerging as the central contagion hub can be associated with its geographical location in the Pacific Ocean. The country connects the western countries through the Singapore Strait to China, South Korea and Japan. Furthermore, the country's favourable corporate laws and regulation make it an ideal destination. The study provides valuable insights for a variety of stakeholders within the financial ecosystem. Investors can benefit significantly from acquiring a more profound comprehension of the complex interrelations among Asia-Pacific stock indices and the impact of various economic shocks. By identifying the Straits Time Index of Singapore as a central contagion point and acknowledging the limited effect of the SSE Composite Index (China) despite its economic importance, investors can enhance their decision-making processes regarding portfolio diversification and risk management strategies through understanding potential opportunities and risks in the region.

Keywords: minimum spanning tree, contagion, crisis events, APAC countries, equity markets

ASSET GROWTH AND INVESTOR PREFERENCES ACROSS FIRM LIFE CYCLES: INDIAN EVIDENCE

Gopikumar V

Senior Assistant Professor
Xavier Institute of Management and Entrepreneurship
Kochi – 683 503, Kerala, India
gopikumarv@xime.org

Smitha Nair

Assistant Professor (Sr. Grade)
Amrita School of Business
Kochi – 682 041, Kerala, India
n_smitha@asb.kochi.amrita.edu

Do all investors celebrate when companies grow quickly? Not necessarily. Their preferences differ based on their access to information about the investments, their ability to forecast the impact of asset expansions on firm value, and potential agency conflicts. Our study examines how five distinct investor types – insiders (promoters), mutual funds, foreign institutional investors (FII), corporates, and retail investors, respond to the asset growth decisions of firms across different life cycle stages. We find that domestic mutual funds and promoters, tend to avoid growth and maturity stage companies with aggressive asset expansions. In contrast, foreign investors do the opposite, indicating a higher risk-seeking behavior. Additionally, mature firms with asset growth tend to attract foreign and retail investors plausibly due to perceived stability and the availability of public information about such investments. Our results indicate that mutual funds and promoters are more adept at identifying the empire-building efforts of managers, whereas foreign institutional investors, corporations, and retail investors fail to do so. Further, our study underscores the significance of considering the firm life cycle as a conditioning variable in understanding diverse investor preferences, thereby offering valuable insights for investment strategy and decision-making.

Keywords: Firm life cycle; Asset growth; Investor behavior; Agency costs; Emerging market

TRANSFORMATIVE AI FOR OPD - ENHANCING EFFICIENCY AND PRODUCTIVITY

Dhananjaya Bedkani Linganaik

NEC Corporation India

[dhananjaya.bl@india.nec.com](mailto:ghananjaya.bl@india.nec.com)

Sanjeeth Veigas

NEC Corporation India

sanjeeth.veigas@india.nec.com

Vishnu Kumar Jakhoria

NEC Corporation India

vishnu.jakhoria@india.nec.com

In developing countries like India, the strained doctor-to-population ratio places significant pressure on healthcare providers. With the increasing adoption of Electronic Health Records (EHRs), doctors spend an additional 10-15% of their time on form-filling and 15-20% on data retrieval, adversely affecting care quality and availability. Many practitioners resort to capturing information in natural language, which limits the data's utility for analytics and retrieval.

While recent advances in deep learning, such as handwriting recognition and voice-to-text technologies, promise to alleviate this burden, they face practical limitations. Handwriting recognition remains nascent, with accuracy concerns hindering clinical adoption. Voice-to-text solutions struggle with noisy environments, code-switching, multi-party interactions, and cultural nuances, such as indirect communication of diagnoses in India. Furthermore, these technologies primarily address data entry without transforming unstructured text into a structured EHR-compatible format.

Our approach leverages advanced Natural Language Processing (NLP) techniques, Large Language Models (LLMs), and Retrieval-Augmented Generation (RAG) to automate HER form-filling, facilitate access to historical records, and enhance clinical data for analytics. By streamlining these processes, we aim to improve doctor-patient interaction time and overall care quality.

Keywords: Medical NLP, Medical LLM, RAG, Summarization, Question Answering, EHR Automation.

CLASSIFYING AND DIAGNOSING ASD THROUGH DEEP LEARNING TECHNIQUES USING BRAIN EEG SIGNALS AND MRI IMAGES

Mrinmoy Roy

PhD Scholar, Dept. of CSE, MIT,
MAHE, Manipal, Mangalore, India
mroy2612@gmail.com,
0000-0002-4646-9301

Dr Srikanth Prabhu

Professor, Dept. of CSE, MIT, MAHE,
Manipal
Mangalore, India

Dr Venkatesh Bhandage

Assistant Professor, Dept. of CSE, MIT,
MAHE, Manipal
Mangalore, India

Dr Manjunath Asuti

Associate Professor, Dept. of ECE,
BNMIT, Bangalore
India

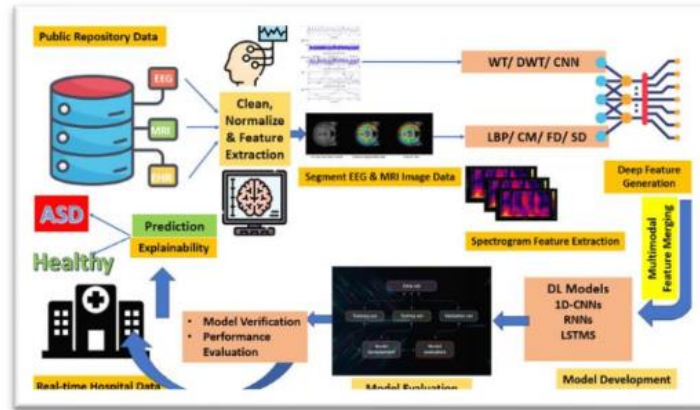
Dr Andrew J

Associate Professor, Dept. of CSE,
MIT, MAHE, Manipal
Mangalore, India

This study investigates the use of Magnetic Resonance Imaging (MRI) data and Electroencephalography (EEG) signals to improve the classification and diagnosis of Autism Spectrum Disorder (ASD) through a multi-modal deep learning approach. The study creates robust models for the diagnosis of ASD by utilizing convolutional neural networks (CNNs) for MRI analysis and recurrent architectures such as Long Short- Term Memory (LSTM) networks for EEG time-series data.

The research aims to enhance the generalizability of the model by combining real-time clinical data from hospital collaborations with public datasets, such as the Autism Brain Imaging Data Exchange (ABIDE), to address the limitations of conventional diagnostic methods. By filtering noise and extracting pertinent features, sophisticated preprocessing methods guarantee that input data is of the highest quality. The study accentuates the interpretability and explainability of the model by utilizing visualization techniques, thereby facilitating its clinical adoption. The findings indicate that the diagnostic accuracy, precision, and recall of a multimodal fusion of EEG and MRI data are substantially enhanced in comparison to those of individual modalities. This research provides valuable insights into the neural correlates of ASD and establishes a pathway for the development of personalized interventions, thereby bridging the gap between artificial intelligence and healthcare to enhance patient outcomes.

Graphical Abstract



Keywords: Electroencephalography (EEG), Feature Extraction, Autism Spectrum Disorder (ASD), Magnetic Resonance Imaging (MRI)

CONSUMER ADOPTION INTENTION TOWARDS ELECTRIC VEHICLES: INSIGHTS FROM INDIA

Suraj Ghosh

Research Scholar

PES University

Bengaluru, India

surajghosh@pesu.pes.edu

Biplab Sarkar

Professor

PES University

Bengaluru, India

biplabsarkar@pes.edu

Electric Vehicles (EVs) are increasingly recognized as a crucial solution for reducing carbon emissions from the transportation sector, which is vital for advancing environmental sustainability. Despite the introduction of various policy measures aimed at promoting EV adoption, the uptake of EVs in India remains limited. Understanding the factors influencing consumer adoption intentions is essential for developing effective strategies to enhance EV uptake, particularly in the context of a developing country like India. This study aims to explore the influence of several key factors - namely financial incentives, information provision, and perceived economic benefits - on the adoption intentions of Indian consumers towards EVs. Additionally, it examines the moderating role of environmental concern on these relationships. A cross-sectional survey was conducted involving 493 young and educated individuals residing in Bengaluru, India. The survey specifically targeted university students, who are considered potential early adopters of EV technology due to their openness to new innovations and their role in shaping future consumer trends. Data were collected using a structured online questionnaire, designed to capture various dimensions of consumer attitudes towards EVs and their adoption intentions. The collected data were analyzed using Partial Least Squares Structural Equation Modelling (PLS-SEM) to test the proposed hypotheses and examine the relationships between the variables. The analysis revealed that financial incentives and perceived economic benefits significantly and positively impact EV adoption intentions. Perceived economic benefits of using EVs was found to be a strong driver of consumer decisions regarding EVs compared with financial incentives. This underlines the importance of perceived economic benefits on adoption intentions. However, information provision measures did not show a significant effect on adoption intentions. Furthermore, the study found that financial incentives partially mediate the relationship between perceived economic benefits and adoption intention. This highlights the interdependence of economic benefits and perceived economic benefits in influencing consumer intentions. Environmental concern, although positively associated with adoption intention, did not significantly moderate the impact of policy measures such as financial incentives and information provision. This finding suggests a complex relationship between environmental values and economic incentives, where higher environmental concern does not necessarily enhance the effectiveness of financial incentives or informational strategies. In summary, the study underlines the critical role of financial incentives and perceived value in shaping EV adoption intentions in India. It also highlights the need for a nuanced approach in policy formulation, considering the impacts of different measures and the intricate relationship between environmental concern and economic incentives. These insights are valuable for policymakers and stakeholders aiming more effective strategies to promote EV adoption, ultimately contributing to the achievement of environmental sustainability goals.

Keywords: electric vehicles, consumer adoption, environmental sustainability, PLS-SEM, policy

ROLE OF PROSPECT THEORY IN INSURANCE COMPANY'S INVESTMENT: A BEHAVIOURAL FINANCE STUDY IN KARNATAKA

Vanishree K

Assistant Professor
Faculty of Management and Commerce
Ramaiah University of Applied Sciences, Bangalore

Dr. Deepa T M

Assistant Professor
Department of Collegiate Education,
Government First Grade College, Kushalnagara

Prospect Theory, a seminal theory in behavioral economics and finance, was developed by psychologists Daniel Kahneman and Amos Tversky in 1979. This theory serves as a fundamental framework for understanding decision-making under conditions of risk and uncertainty, a central focus of this study regarding investment decision-making in insurance companies.

THE EMERGENCE OF RATINGS OF AMAZON'S REVIEWS: FROM POOLING TO SEPARATING EQUILIBRIUM

Vallurupalli Vamsi, Prachi Sinha, Deepak K Sinha

Fake product reviews are a bane of online marketing. Their presence often forces even producers of quality products to buy such reviews, resulting in a pooling equilibrium and loss of consumers' value. We show that with the advent of product review ratings by FakeSpot, ReviewMeta, ReviewIndex, and others, a separating equilibrium emerges in which producers of high-quality products do not buy fake reviews. The lack of fake reviews reported by these ratings and Amazon's original ratings signal high quality to consumers for these products. We numerically establish the validity of this result over a wide range of parameter values. Data from FakeSpot, ReviewMeta, and amazon.com support our hypotheses.

Keywords: Retail Analytics; Fraud Analytics; Web Analytics

ENHANCING DISPUTE VALIDATION IN CUSTOMER SERVICE THROUGH FINE-TUNED LARGE LANGUAGE MODELS AND ADVANCED EMBEDDING TECHNIQUES

Kolli Nethre Sai

Senior Data Scientist
Genpact India Pvt. Ltd
Hyderabad, India

kollinethre.sai1@genpact.com

Ish Chaudhary

Data Scientist
Genpact India Pvt. Ltd
Bengaluru, India

ish.chaudhary@genpact.com

Varsha Rani

Lead Data Scientist
Genpact India Pvt. Ltd
Gurgaon, India

varsha.rani@genpact.com

In the realm of customer service, resolving disputes efficiently and accurately is paramount, yet traditional manual relying on manual processing and rule-based methods often fall short due to their labor-intensive nature, susceptibility to human error, and inability to scale effectively. This paper introduces an advanced solution that leverages the capabilities of fine-tuned Large Language Models (LLMs) and advanced embedding techniques to automate and enhance the dispute validation process. Our proposed framework excels at extracting relevant information utilizes the latest natural language processing capabilities of LLMs to extract relevant information from dispute documents with unprecedented from diverse dispute documents with high precision, overcoming the limitations inherent in conventional methods. Additionally, it addresses the critical challenge of product identification within extensive SKU databases by employing dense vector of identifying products within extensive SKU databases. The research underscores disputes, thereby enhancing operational efficiency, reducing resolution times, and facilitating strategic business decisions. This research highlights the potential of LLM-based solutions in transforming the landscape of retail dispute resolution, offering a scalable, efficient, and accurate alternative the transformative potential of LLM-based solutions in streamlining dispute validation processes within the retail sector and suggests avenues for future exploration, to traditional methods. Future work will explore the framework's applicability across different markets and product categories and the integration of additional advanced including the expansion to other markets and product categories and the integration of additional advanced technologies.

Keywords: Dispute Validation, Natural Language Processing, Finetuning LLM, Key-value Pair, SKU Databases, Operational Efficiency, Advanced Embedding Techniques.

OPTIMIZING MULTIMODAL AI PERFORMANCE: A COMPARATIVE ANALYSIS OF PROMPTING TECHNIQUES ACROSS GPT-4.0, GEMINI, AND CLAUDE

Yadma Meghna Reddy

Data Science Intern

Genpact

Hyderabad, India

yadma.meghnareddy@genpact.com

Ashrita Naidu Vaka

Data Science Intern

Genpact

Hyderabad, India

ashritanaidu.vaka@genpact.com

Pritam Varma

Data Science Intern

Genpact

Hyderabad, India

pritam.varma@genpact.com

Koundinya N V S S

Senior Data Scientist

Genpact

Hyderabad, India

ORCID: 0000-0003-1663-3112

Kolli Nethre Sai

Senior Data Scientist

Genpact

Hyderabad, India

kollinethre.sai1@genpact.com

Chirag Jain

Principal Data Scientist

Genpact

Bengaluru, India

chirag.jain4@genpact.com

As multimodal AI advances, selecting effective prompting technique is crucial for optimizing model performance. This study aims to compare seven prompting techniques: Zero-shot, Few-shot, Chain-of-Thoughts (CoT), Self-consistency, Prompt chaining, Step-back prompting, and Meta Prompting across three leading AI models: GPT-4o, Gemini, Claude, evaluating their effectiveness on various tasks. Various multimodal tasks have been chosen including named entity recognition, image captioning, video summarization and audio transcription. These tasks cover multiple data types-text, image, video and audio allowing for evaluation on a broader context. The evaluation criteria include accuracy, coherence, and contextual relevance. By analyzing the models performance, this paper provides insights for selecting and implementing the most effective prompting strategies in various applications.

Keywords: multimodal AI, prompting techniques, model performance, named entity recognition, image captioning

DOES BRAND EQUITY EXIST IN BANKING SECTOR? A DISCRIMINANT ANALYSIS

Meena Rani N

Department of PGDM

Xavier Institute of Management and Entrepreneurship
Bangalore, India

Orcid: 0000-0001 7149 9089

Shivani Kapoor

Department of MBA

Jain Online (Deemed to University)
Bangalore, India

Orcid: 0000-0002-9627 6524

Purpose: This paper attempts to explore the various sources of brand equity in the banking sector and seeks to apprehend the role of brand loyalty in the brand equity framework. Additionally, it tries to decipher how these sources predict the likelihood of the consumer to continue to patronage the bank.

Design/Methodology/Approach: Primary data were collected from 489 customers of different commercial banks in the Bangalore area. This study collected both qualitative and quantitative data and used the same for analysis and drawing inferences. Structural Equation Modelling and Discriminant analysis were undertaken to validate the results.

Findings: Statistical analysis showed validity and reliability of the instrument, the Structural Equation Modelling output shows 5 constructs - awareness, association, perceived quality, ethics and transparency, and environment-friendly practices significantly contributing to brand loyalty, which is the core of brand equity. To examine the factor that contributes most to predicting the future intentions of the consumers regarding continued patronage of the bank, the discriminant analysis was executed which showed awareness and perceived quality are the significant differentiators in determining continued patronage of the bank, and the model predicted with 97 per cent accuracy the future intentions of the consumers regarding continued patronage.

Originality: This paper broadens the understanding of Brand equity sources, alongside conventional elements of brand awareness, association, and perceived quality, the rapidly emerging elements such as ethics and transparency, and sustainability also proved to be significant contributors to brand equity in the banking sector. Additionally, Awareness, Perceived Quality, Ethics and Transparency, and Environmental Friendliness are the most significant components determining continued patronage of the customers. The study also sheds light on brand loyalty which serves as a proxy for brand equity.

Keywords: Brand equity, banking sector, predictive model, ethics and transparency, environmental friendliness, patronage and discriminant analysis.

Harnessing the Power of Analytics for Energy Savings in Telecom Networks

Rohit Agarwal

Lead Product Manager

HCLSoftware

Bengaluru, India

rohit.agarwal@hcl-software.com

Murali Krishnan K H

Senior Software Engineer

HCLSoftware

Bengaluru, India

murali.krishnankh@hcl-software.com

With the global surge in mobile data demand, Mobile Network Operators (MNOs) are increasingly challenged to align operational efficiency with environmental responsibility. This paper presents a novel solution: HCLSoftware's Intelligent Dynamic Energy Savings (IDES), powered by HCLSoftware Smart Analytics, engineered to enhance energy efficiency across telecom networks while preserving superior service quality.

By harnessing advanced artificial intelligence, HCLSoftware IDES dynamically adjusts energy consumption at the site level, achieving substantial reductions in both operational costs and carbon emissions. Through integrated Business Intelligence (BI)-driven analytics, MNOs gain real-time insights to monitor, optimize, and document energy use effectively.

This proactive approach not only contributes to environmental sustainability but also drives compliance with emerging green standards in telecommunications. Findings from deployments demonstrate that the Analytics driven energy saving solution from HCLSoftware enables MNOs to achieve around a 10% reduction in energy consumption leading to an estimated CO2 reduction of 9000 Metric Tons and an annual savings exceeding \$2.4 Million for a network of roughly 10,000 radio sites, setting a benchmark for sustainable operations in the industry.

Keywords: AI-driven Energy Savings, Mobile Network Sustainability, Energy Efficiency in Telecom, BI in Telecom, HCL Software, IDES, Smart Analytics, Carbon Footprint Reduction, Mobile Network Optimization.

A NOVEL APPROACH FOR EVALUATING PERFORMANCE CONSISTENCY IN TIME SERIES FORECASTING MODELS

Deeksha Gopalan

Data Scientist
Genpact
Bengaluru, India
deeksha.gopalan@genpact.com

Komal Sandhu

Principal Data Scientist
Genpact
Bengaluru, India
komal.sandhu@genpact.com

Koundinya N V S S

Senior Data Scientist
Genpact
Hyderabad, India
ORCID: 0000-0003-1663-3112

Rajeev Ranjan

Principal Data Scientist
Genpact
Bengaluru, India
rajeev.ranjann@genpact.com

B Lakshit

Data Scientist
Genpact
Hyderabad, India
b.lakshit@genpact.com

Chirag Jain

Principal Data Scientist
Genpact
Bengaluru, India
chirag.jain4@genpact.com

This paper addresses the critical challenge of evaluating performance consistency in time series forecasting models, particularly in the context of limited data availability. Traditional evaluation methods often lead to issues of overfitting and underfitting, as they fail to account for the temporal dependencies inherent in time series data. To tackle this problem, we propose a novel methodology that employs a hybrid expanding-sliding window approach for training and a comprehensive testing strategy utilizing all unseen future data thus maximizing data utilization. We introduce the Genpact Evaluation Metric (GEM) to evaluate both the magnitude and consistency of forecasting errors across multiple validations. Our methodology has been validated on real-world transportation and P&L datasets, revealing that traditional models, particularly exponential smoothing and PROPHET, often overfit during training and cross-validation, resulting in poor performance in actual forecasting scenarios. In contrast, our approach consistently selects models that demonstrate stability across different datasets. Thus, our research enhances the reliability of evaluation techniques in time series forecasting, underscoring the need for improved methodologies in the field.

Keywords: Time Series Forecasting, Performance Consistency, Data Scarcity, Performance Evaluation.

TRANSFORMING HIGHWAY MONITORING DRONE ANALYTICS MONITORING SYSTEM (DAMS)

Utkarsh Singh

CEO & Founder
DronaMaps Pvt Ltd.
New Delhi, India
usingh@dronamaps.com

Poonam Gupta

Chief Business Officer
DronaMaps Pvt Ltd.
New Delhi, India
pgupta@dronamaps.com

Ayushi Mishra

COO & Co Founder
DronaMaps Pvt Ltd.
New Delhi, India
amishra@dronamaps.com

Ronak Khandelwal

Sr. Data Science Engineer
DronaMaps Pvt Ltd.
New Delhi, India
rkhandelwal@dronamaps.com

Rahul Kumar

Associate Director,
DronaMaps Pvt Ltd.
New Delhi, India
rkumar@dronamaps.com

Hardik Jain

Data Science Engineer
DronaMaps Pvt Ltd.
New Delhi, India
hardik@dronamaps.com

Everything that humankind has built either gets captured in the form of numbers in a table, or two dimensional maps. The problem with both these modalities is that the medium of capture is not sufficient to enable accurate measurement or visual representation of physical infrastructure as it exists in reality. The easiest way to solve this problem is to create a solution that creates a digital replica of the world and uses algorithms to measure it. The traditional methods of monitoring highway construction and maintenance are mostly based on manual data collection augmented with drone video collection or Lidar based survey vehicles. They have been inefficient, prone to errors, or extremely costly. This research study presents a new approach to highway monitoring using drone based maps and artificial intelligence, called the Drone Analytics Monitoring System (DAMS).

DAMS leverages high-resolution drone images of infrastructure to create 3D drone maps which are then analyzed using AI algorithms to identify construction progress, road conditions, plantations, and safety hazards. The system provides insights through user-friendly analytics suites, enabling better decision-making. As a company, DronaMaps specialises in reconstruction of drone based maps and AI based analytics on it. We are deploying a system for the National Highways Authority of India and Ministry of Road Transport and Highways. Therefore, this paper would focus more on the usage of drones.

The study outlines the development and testing of DAMS, showcasing its potential to revolutionize highway monitoring. The system's ability to improve efficiency, safety, and compliance makes it a valuable tool for infrastructure development. Additionally, the study explores the integrated management ability of highway construction and maintenance, providing a more efficient platform for department coordination. The DAMS platform establishes a timely, stable, and unified information-based collaborative working environment for all parties involved in the construction, maintenance, and operation stages. This enables relevant departments to share information, work together, and coordinate business on a unified platform, forming a scientific system of knowledge accumulation.

Furthermore, the study emphasizes the intangible assets in the highway management system, as the DAMS platform will accumulate valuable intangible assets and information wealth in the form of 3D maps for highway construction and management. This will realize the value-added efficiency of highway assets and provide reference for future highway construction and management.

Finally, the study discusses the interface for public participation in the highway construction and maintenance system. The DAMS platform services for the public, such as reporting of public safety hazards, health of plantations, or damage or decay of the highway furniture with geotagged photographs.

Keywords: highways, infrastructure, drones, artificial intelligence, road condition monitoring; pavement distress evaluation; pavement monitoring; smart sensors; AI; deep learning; machine learning; Drone Analytics Monitoring System (DAMS), high-resolution drone imagery, 3D drone maps, AI algorithms, highway monitoring, construction progress, road conditions, safety hazards, plantation health, infrastructure development, National Highways Authority of India (NHAI), Ministry of Road Transport and Highways (MoRTH), manual data collection, Lidar survey vehicles, intangible assets in highway management, geotagged photographs, public safety hazards, collaborative platform, digital twin, and asset management efficiency.

LARGE LANGUAGE MODELS FOR PL/SQL CODE EXPLANATIONS: A COMPREHENSIVE EXPLORATION AND EVALUATION

Vivek Tak

Data Scientist
Fidelity Investments
Bangalore, India
vivek.tak@fmr.com

Radha N

Senior Data Scientist
Fidelity Investments
Bangalore, India
radha.n@fmr.com

Nandeesh Umesha

Data Scientist
Fidelity Investments
Bangalore, India
nandeesh.umesha@fmr.com

Amol Jadhav

Director
Fidelity Investments
Bangalore, India
amol.jadhav@fmr.com

Advancements in technology necessitate replacing legacy systems for improved usability and development of business applications. Updating requirement documents for legacy systems is challenging due to the complexity of reverse engineering and understanding system requirements from legacy code. Clear explanations of legacy code are vital for requirement analysts to bridge gaps and missing requirements. Extensive research and development have been taking place in this area, harnessing natural language processing techniques, particularly with the use of large language models. In this paper, we present our observations and methodology for code explanations. This study focuses on the step-by-step code explanation of PL/SQL queries using Large Language models. We have developed a chunking method that aims to break down the code into smaller portions. The key idea behind our methodology is to split the code into meaningful sections while preserving the logical flow. This methodology effectively breaks down the code into digestible portions, facilitating code explanation. The evaluation criteria have been defined based on the overall completeness, correctness, and conciseness of the explanations. The findings demonstrate that incorporating system instructions in zero-shot prompt setting significantly improves the performance. The method of parsing and chunking enhances the completeness, correctness of the explanations by minimizing the omission of main facts and conditions in the translated text.

Keywords: Large Language Model, Natural Language Processing, PL/SQL Code Explanation

ADAPTING AGRICULTURAL WATER MANAGEMENT PRACTISES USING ARTIFICIAL INTELLIGENCE FOR CLIMATE RESILIENCE AND SUSTAINABILITY OF VIRUDHUNAGAR DISTRICT OF TAMILNADU

Dr.G.VENKATESHWARAN*, Mrs.S. KASTHURI, and Mr.P.SHREE SHRAWAN KUMAR

Assistant Professors, Department of Management Studies,
Velammal College of Engineering and Technology (Autonomous), Madurai
(Affiliated to Anna University, Chennai) Tamil Nadu, India

This study examines the adaptation of agricultural water practices for enhanced climate resilience and sustainability, focusing on the Virudhunagar District of Tamil Nadu. A biodiversity hotspot and critical water source, faces increasing pressure from climate change, which exacerbates water scarcity and impacts agricultural productivity using AI. By assessing traditional and modern water management techniques, this research identifies strategies that can mitigate climate impacts and promote sustainable agriculture. Key practices explored include rainwater harvesting, drip irrigation, and watershed management. The study integrates field surveys, hydrological data analysis, and stakeholder interviews to provide a comprehensive understanding of effective water management in this unique region. Findings emphasize the importance of community involvement, policy support, and technological innovation in building resilient agricultural systems. This research aims to contribute to the broader discourse on sustainable water management, offering actionable insights for policymakers, farmers, and environmentalists in Tamil Nadu and similar agro-climatic regions. Overall, this research aims to create a resilient agricultural framework that not only addresses immediate water management challenges but also builds long-term sustainability and climate adaptability using AI in Virudhunagar District of Tamil Nadu.

Keywords: Agriculture, Climate Change, Artificial Intelligence, Sustainability.

MAPPING CSR WITH SDGS FOR SUSTAINABILITY REPORTING OF INDIA ENTERPRISES

Gulshan Kumar Yadav

Rajendra Mishra School of Engineering Entrepreneurship
Indian Institute of Technology Kharagpur
Kharagpur, India
gulshan.kyadav.777@gmail.com

Dr. Titas Bhattacharjee

Rajendra Mishra School of Engineering Entrepreneurship
Indian Institute of Technology Kharagpur
Kharagpur, India
titas@see.iitkgp.ac.in

Dr. Atasi Mohanty

Rekhi Centre of Excellence for the Science of Happiness
Indian Institute of Technology Kharagpur
Kharagpur, India
atasim@cet.iitkgp.ac.in

The alignment of Corporate Social Responsibility (CSR) interventions with the Sustainable Development Goals (SDGs) is crucial for achieving sustainability in developing countries like India. This research aims to develop a comprehensive CSR-SDG mapping tool to effectively align CSR activities with SDG targets. This tool will enable the CSR practicing Indian enterprises to report effective impacts, build reputation, trust, and attract investment while promoting long-term sustainability and regulatory compliance.

The research methodology involves a detailed analysis of the SDG framework and CSR provisions in India. The SDG framework document, downloaded from the United Nations website, provides descriptions for the 17 SDGs and 169 associated targets. CSR provisions are obtained from the Ministry of Corporate Affairs website, detailing the domains and activities recognized as CSR under the Companies Act, 2013. These documents serve as inputs for mapping the similarity between CSR interventions and SDG goals.

To achieve the objective, the study employs Natural Language Processing (NLP) techniques to convert human language into machine-readable numeric characteristics. The texts are pre-processed to remove any special characters and then vectorized using the pre-trained model ‘MiniLM layer 6 version 2’ from the sentence-transformers library in Python. This model maps sentences and paragraphs to a dense vector space, enabling tasks like clustering and semantic search. Cosine similarity is used to measure the similarity between sentences, providing a quantitative analysis of the content.

The analysis reveals that aligning CSR interventions with SDGs and their respective target indicators can enhance the reporting of sustainability goals by the government. Also, corporates can use this framework to strategically plan CSR activities and report sustainable actions aligned with UN SDGs, enhancing legitimacy, stakeholder trust, and investor interest. The study finds that 14 out of 17 SDGs are addressed by CSR interventions, with varying degrees of association. For instance, CSR intervention which focuses on improving health and sanitation, shows a strong association with SDG 6 (Clean Water and Sanitation). Other CSR interventions, such as those related to education, livelihood, and environmental conservation, also demonstrate significant alignment with multiple SDGs.

The results indicate that while some CSR interventions are strongly related to specific SDGs, others show moderate or weak associations. This highlights the need for a more strategic approach to CSR activities to maximize their impact

on sustainability goals. The developed CSR-SDG mapping tool provides a robust framework for corporates to plan their CSR activities strategically and report their contributions to sustainability more effectively.

In conclusion, the alignment of CSR interventions with SDG targets offers a comprehensive mechanism for achieving sustainability goals. The CSR-SDG mapping tool developed in this research can serve as a valuable resource for enterprises and policymakers in India, enhancing the overall sustainability reporting and contributing to the nation's progress towards the SDGs.

Keywords: CSR-SDG Alignment, Natural Language Processing (NLP), Sustainability Reporting Introduction (Heading 1)

COMPARATIVE EVALUATION OF ADVANCED AI MODELS AND SERVICES FOR COMPLEX TABULAR DATA EXTRACTION FROM DOCUMENTS

Shashank Kumar

Data Science Intern

Genpact

Hyderabad, India

shashank.kumar4@genpact.com

Koundinya N V S S

Senior Data Scientist

Genpact

Hyderabad, India

ORCID: 0000-0003-1663-3112

Meghan Challa

Data Science Intern

Genpact

Hyderabad, India

meghan.challa@genpact.com

Divya Jyothi Mutyala

Machine Learning Engineer

Genpact

Hyderabad, India

divyajyoti.mutyala@genpact.com

Kolli Nethre Sai

Senior Data Scientist

Genpact

Hyderabad, India

kollinethre.sai1@genpact.com

Chirag Jain

Principal Data Scientist

Genpact

Bengaluru, India

chirag.jain4@genpact.com

Extracting complex tabular data from documents is challenging due to diverse document layouts and non-standard tables. This paper evaluates five advanced AI models for tabular data extraction: Google Document AI (Vision AI), Amazon Textract, Azure Document AI, and LayoutLMv3. Each model employs distinct techniques, from OCR to transformer-based architectures, with varying strengths in processing structured and unstructured documents. Through experiments on complex datasets with multi-page tables and irregular cell formats, we assess each model's accuracy, speed, scalability, and cost-effectiveness. Results indicate that transformer-based models, such as and LayoutLMv3, excel at understanding complex table structures, while OCR-based services like Amazon Textract and Google Document AI perform well in general document parsing. This study offers a practical comparison to inform AI solution selection based on specific document intelligence needs.

Keywords: Tabular data extraction, document parsing, artificial intelligence, OCR, transformer models, Google Document AI, Amazon Textract, Azure Document AI, LayoutLMv3, document intelligence, deep learning, complex tables, multi-page tables.

STOCK FORECASTING AND SECTOR-SPECIFIC RECOMMENDATION USING DEEP LEARNING

Aravilli Atchuta Ram, Aryan Wadhwa, Avani Dhagam, Jay Mintu Shah, R Bharathi

Department of Computer Science and Engineering
PES University, Bengaluru, India
pes2202101182@pesu.pes.edu

The stock market is extremely dynamic and uncertain by nature. Several studies have applied various statistical, machine learning and deep learning models for stock forecasting and recommendation. However sector-specific (e.g., healthcare, technology etc.) stock recommendations remain unexplored in the literature. Hence, a unified model was implemented for stock forecasting and sector-specific recommendation using deep learning. A hybrid forecasting model is proposed combining LSTM with ARIMAX to predict the closing price of the stock, incorporating exogenous variables across different sectors. Additionally, explainable AI techniques were employed to enhance model transparency and increase investor trust. The performance of the proposed hybrid model and individual LSTM and ARIMAX models was evaluated using standard forecasting metrics. Furthermore, a sector-specific recommendation model was proposed that takes into consideration various factors and recommends the stock price movement indicator for each stock. The recommendation model was then evaluated using classification metrics. This study contributes to developments in the finance sector and opens opportunities for more personalized investment strategies and offers a comprehensive approach to enhance investment decisions.

Keywords: Recommender System, Stock Price Forecasting, Deep Learning, ARIMAX, Explainable AI

AUTOMATED CONTENT MODERATION USING AI: AN AGENTIC APPROACH TO PROCESS TEXT, AUDIO, VIDEO, AND IMAGES

Kolli Nethre Sai

Senior Data Scientist

Genpact India Pvt. Ltd

Hyderabad, India

kollinethre.sai1@genpact.com

Upamanyu Deepak Wable

Data Scientist

Genpact India Pvt. Ltd

Hyderabad, India

upamanyu.deepakwable@genpact.com

Koundinya N V S S

Senior Data Scientist

Genpact India Pvt. Ltd

Hyderabad, India

ORCID: 0000-0003-1663-3112

Divya Jyothi Mutyala

ML Engineer

Genpact India Pvt. Ltd

Hyderabad, India

divyajyoti.mutyala@genpact.com

Chirag Jain

Senior Data Scientist

Genpact India Pvt. Ltd

Bengaluru, India

chirag.jain4@genpact.com

Meghana Kurra

Data Scientist

Genpact India Pvt. Ltd

Hyderabad, India

kurra.meghana@genpact.com

The vast and unregulated flow of content on the internet presents significant challenges in maintaining safe and respectful digital environments. Conventional content moderation methods, which depend heavily on human moderators, face significant challenges related to scalability, high costs, slower response rates, and the emotional strain placed on reviewers. In response to these limitations, this paper proposes an automated content moderation solution utilizing advanced Artificial Intelligence (AI) techniques, including Large Language Models (LLMs) and Vision AI agents. The solution processes text, audio, video, and image-based content, ensuring that inappropriate material is flagged and addressed efficiently. This automated pipeline leverages Whisper for audio transcription and uses LLMs for textual analysis, while Vision AI agents handle video and image moderation. Our approach significantly reduces human involvement, improves accuracy, and ensures scalability, achieving an 80-85% accuracy rate across various content types. This solution contributes to a safer online environment and more effective content management practices for organizations.

Keywords: Vision AI Agents, Large Language Models (LLMs), Content Moderation, Artificial Intelligence

ADVANCING SENTIMENT ANALYSIS WITH LONG SHORTTERM MEMORY (LSTM) NETWORKS

Nidhi Umashankar

Dept. of AI&ML,
BMS Institute of Technology and Management
Bangalore,India
nidhiumashankar03@gmail.com

K Sai Geethanjali

Dept. of AI&ML,
BMS Institute of Technology and Management
Bangalore,India
sgak2014@gmail.com

Sentiment analysis is the ability to detect emotional states in text, an area of very high importance in NLP and very relevant for many of its applications, such as monitoring social media, product reviews, and the investigation of customer feedback. It presents a sophisticated technique for the usage of LSTM networks-an advanced version of Recurrent Neural Networks (RNNs). Such approaches, however, go way beyond simple traditional machine learning techniques, Naïve Bayes and logistic regression, which can easily fail for complex sequential patterns in very long texts or vague texts. Our experiment was based on the Yelp polarity dataset, which contained many user reviews with positive and negative tags. The pre-processing of the full text included lowercasing, punctuation elimination, and stop words removed from it. The model bidirectional LSTM efficiently controlled sentiment prediction through internal memory and gating functions, which easily handled both short-term and long-term dependencies. The proposed method used performance evaluation metrics including accuracy, precision, recall, F1- score, and AUC-ROC, where this method does significantly better than conventional approaches as it achieves an ensemble accuracy of 95.83%, F1-score of 95.85%, precision of 95.33%, and recall of 96.39%. This work embraces the effectiveness of LSTM networks in performing tasks related to complex and context-dependent sentiment prediction and potential applications in e-commerce and social media analysis, where the concept of having a correct sentiment captured from huge volumes of text is very relevant.

Keywords: Sentiment Analysis, Long Short-Term Memory (LSTM), Convolutional Neural Networks (CNN), Ensemble Learning, Yelp Dataset, Deep Learning, Text Classification.

OPTIMIZING FARE STRUCTURES USING SHORTEST PATH ALGORITHMS AND REAL-TIME ADJUSTMENTS IN URBAN METRO NETWORKS

Dr. K. P. Mani

Field Specialist

Automatic Fare Collection System

Chennai Metro Rail Limited

Chennai, Tamil Nadu, India

vkpmany@gmail.com; mani.kp@cmrl.in;

Mr. K. Thiruvalluvan

Consultant

Automatic Fare Collection System

Chennai Metro Rail Limited

Chennai, Tamil Nadu, India

thiruvalluvan.k@cmrl.in;

Mr. Kumar Satyam

Assistant Manager

Automatic Fare Collection System

Chennai Metro Rail Limited

Chennai, Tamil Nadu, India

kumar.satyam@cmrl.in;

Mr. S. Packiriraj

Senior AFC Engineer, General Consultant

Automatic Fare Collection System

Chennai Metro Rail Limited

Chennai, Tamil Nadu, India

packiriraj.s@nkab.cmrl.in;

This study presents a dynamic fare optimization model for urban metro networks, integrating shortest-path algorithms and real-time data to enhance operational efficiency and pricing fairness. By evaluating algorithms such as A*, Dijkstra, and K-Shortest Path (KSP), the research assesses optimal route determination, fare calculation based on travel distance and interchanges, and real-time pricing adjustments. The model incorporates live data streams for dynamic fare adjustments and employs evolutionary algorithms for optimizing revenue and passenger satisfaction. Simulations on real-world metro networks validate the model's scalability, efficiency, and potential for broader adoption. The proposed framework improves fare accuracy, load balancing, and revenue optimization, making it a versatile solution for modern metro systems.

Keywords: Shortest-path algorithm; real-time adjustments; optimal routes; multi-corridor journeys; dynamic pricing model; fare computation; real-time integration; optimization; fare accuracy; operational efficiency.

APPROACH, APPLICATION AND IMPLEMENTATION STRATEGY FOR A 3D DIGITAL TWIN, AN ARTIFICIAL INTELLIGENCE-BASED TRANSFORMATIVE TOOL

Utsav Mathur

Founder / CEO

GMetri

Bangalore, India

utsav@gmetri.com

Piyush Kumar Mathur

Head, dg2n (3D Digital-Twin Division)

GMetri

Bangalore, India

piyush@gmetri.com

Sahil Ahuja

Founder / CTO

GMetri

Bangalore, India

sahil@gmetri.com

Operational efficiency is an urgent priority for every industry, yet critical data often remains hidden in the murky depths of databases, accessible to only a few. This data, originally formatted to suit machines rather than humans, hinders understanding and slows decision-making where clarity and speed are paramount. Designed for machine-readability, databases and systems have created a world where information is siloed, opaque, and complex. As a result, the skills needed to interpret machine-centric data have outpaced those aimed at making this data accessible, comprehensive, and actionable for human decision-makers.

3D Digital Twin technology stands as a revolutionary solution, positioning real-time and historical data as a single source of truth. By providing a centralized, visually immersive 3D model, digital twins closely mimic how humans perceive the world, making data not only accessible but intuitively understandable. Through dynamic 3D visualizations, digital twins enable organizations to optimize operations, enhance safety, and drive rapid, data-informed insights across sectors. This paper will demonstrate the versatility of 3D Digital Twin solutions through applications in warehousing and oil and gas, highlighting the universal impact of technology that brings critical data into vivid, actionable view.

While the promise of digital twins has been discussed for years, successful implementation has often remained elusive. Without the right foundations—robust data integration, scalable infrastructure, and clear operational goals—digital twin projects can stall or fail to deliver meaningful value. This paper introduces “dg2n”, a structured framework designed to guide organizations through each stage of implementing a 3D Digital Twin effectively and scalably. By following the dg2n roadmap, companies can establish a comprehensive, future-proof solution that maximizes the impact of digital twin technology across their operations, bringing the full promise of digital twins within reach.

A critical challenge in digital twin implementation is ensuring the veracity of data, as discrepancies frequently arise between data observed in the field and what is recorded in systems. These errors often stem from system or database design limitations, leading to data that may not accurately reflect real-world conditions. Bridging this gap is a primary objective of dg2n, achieved through a suite of mobile-based tools and processes that continuously validate and update data. By enabling a bi-directional data flow and seamless integration into daily operations, dg2n allows data to move fluidly between systems and the field, ensuring that digital twins remain true to their physical counterparts and provide reliable insights.

Keywords: 3D Digital Twin, dg2n Framework, Data Veracity, Operational Efficiency, Real-time Data Integration, Predictive Maintenance, Bi-directional Data Flow, Mobile Data Tools, Data-driven Decision Making, Process Optimization, Industrial Automation, Visualization Technology, Warehouse Optimization, Oil and Gas Operations, Scalable Digital Twin Solutions.

ROUNDTABLE: LEVERAGING DYNAMIC SCHEMA AND CONTEXTUAL AUTOCOMPLETE FOR ENHANCED QUERY PRECISION IN TABULAR QUESTION ANSWERING

Pratyush Kumar

West Pharmaceutical Services, Inc.
Bengaluru, India
pratyushk2011@gmail.com

Aman Chadha

Stanford University
Amazon GenAI
Cupertino, USA
hi@aman.ai

Kuber Vijaykumar Bellad

West Pharmaceutical Services, Inc.
Bengaluru, India
kuber.bellad4@gmail.com

Bharat Vadlamudi

West Pharmaceutical Services, Inc.
Bengaluru, India
bharat0355@gmail.com

With advancements in Large Language Models (LLMs), a major use case that has emerged is querying databases in plain English, translating user questions into executable database queries, which has improved significantly. However, real-world datasets often feature a vast array of attributes and complex values, complicating the LLM's task of accurately identifying relevant columns or values from natural language queries. Traditional methods cannot fully relay the dataset's size and complexity to the LLM. To address these challenges, we propose a novel framework that leverages Full Text Search (FTS) on the input table. This approach not only enables precise detection of specific values and columns but also narrows the search space for language models, thereby enhancing query accuracy. Additionally, it supports a custom auto-complete feature that suggests queries based on the data in the table. This integration significantly refines the interaction between the user and complex datasets, offering a sophisticated solution to the limitations faced by current table querying capabilities. This work is accompanied by an application 1 for both Mac and Windows platforms, which readers can try out themselves on their own data.

Keywords: LLMs, Queries, Databases, Querying, Datasets, Attributes, Values, Columns, Search, Accuracy, Autocomplete, Interaction, Refinement, Complexity, Table, Application. AI, GenAI, Data

EXPLORING ETHICAL PARADIGMS IN DIGITAL MARKETING: A BIBLIOMETRIC APPROACH

Smita M Gaikwad

Faculty of Management Science
CMS B-School JAIN (Deemed to- be-Univesity) Bangalore, India
dr.smita_gaikwad@cms.ac.in

Umesh Chandra

Faculty of Management Science
CMS B-School JAIN (Deemed to- be-Univesity) Bangalore, India
dr.umesh_chandra@cms.ac.in

Krishna Koppa

Faculty of Management Science
CMS B-School JAIN (Deemed to- be-Univesity) Bangalore, India
krishnakoppa@cms.ac.in

Ajith Kumar M

Faculty of Management Science
CMS B-School JAIN (Deemed to- be-Univesity) Bangalore, India
ajithkumar_m@cms.ac.in

Vishal Soodan

Faculty of Management Science
CMS B-school JAIN (Deemed to- be-Univesity) Bangalore, India
dr.vishalsoodan@cms.ac.in

Over the last 20 years, there has been considerable progress in digital marketing and consumer ethics but we can see that there is limited research on specific areas of digital marketing ethics. Despite the development of digital technologies and their impact on advertising practice as a general matter, particularly concerning internet advertising, this limitation is maintained. The sub-discipline of digital marketing ethics has only seen a handful of investigations with many limitations. This work therefore aims to review the current literature on digital marketing ethics, employing bibliometric methods. In this research paper, we have conducted our reviews using bibliometric techniques: 'co-citation authorship, co- occurrence, bibliographic coupling document analysis', and we have highlighted the field's origins and current themes. We then mapped our findings using free science visualization software (VOS viewer). We have examined the implications of this study and found new variables to carry out future research.

Keywords: Digital Marketing ethics, bibliometric, VOS Viewer, Advertising practices.

AUTOMATED EXTRACTION OF STARTUP PERFORMANCE DATA FROM PITCH DECKS

Sruthi Kannan

Rajendra Mishra School of Engineering
Entrepreneurship
IIT Kharagpur
India
sruthisresearch@gmail.com

Divya Eshwar

Department of Computer Science &
Engineering
PES University
India
divyaeshwar.22@gmail.com

Guhan Chandrasekaran

Engineering Research & Development
Flyna Technologies
India
cguhan@gmail.com

Bhaskar Bhowmick

Rajendra Mishra School of Engineering
Entrepreneurship
IIT Kharagpur
India
bhaskar@see.iitkgp.ac.in

CS Kumar

Department of Mechanical Engineering
IIT Kharagpur
India
kumar@mech.iitkgp.ac.in

This research explores the application of Generative AI and Large Language Models (LLMs) to automate the extraction of key growth metrics—such as revenue and profit/loss—from startup pitch decks, thereby streamlining the evaluation process. By utilizing OpenAI's embedding models and Facebook AI Similarity Search (FAISS) for similarity search, the framework efficiently captures critical performance data from pitch decks containing unstructured data. This automated approach overcomes the inefficiencies of manual data extraction while enhancing both the accuracy and scalability of the analysis. Additionally, the system incorporates impact metrics, aligning the extracted data with relevant UN Sustainable Development Goals (SDGs) to ensure a comprehensive evaluation of the startups' impact on sustainability. Our study delves into data from 94 India-based startups focused on sustainability. Using Python and LlamaIndex, unstructured text is extracted and converted into embeddings via OpenAI's API. FAISS facilitates efficient document similarity searches, enabling the retrieval of key financial metrics such as revenue, profit/loss, and impact. This RAG (Retrieval-Augmented Generation) based implementation showcases the potential for scalable, automated analysis of pitch decks, enhancing accuracy and efficiency in financial data extraction. Thus, this study aims to enhance transparency and fairness in startup evaluations, benefiting both investors and entrepreneurs through data-driven insights and consistent evaluations.

Keywords: Generative AI, Large Language Models, Startup Pitch Decks, Financial Decision-Making, OpenAI Embeddings, FAISS, LlamaIndex, Automated Financial Analysis, Venture Capital, Investment Analysis, Document Similarity Search, NLP for Finance, Startup Evaluation, Automated Data Extraction.

THE CELEBRITY SALES GAME: WORTH THE INVESTMENT?

Harshavardhan G

B.Sc Data Science
SVKM'S NMIMS

Akhil Khorana

B.Sc Data Science
SVKM'S NMIMS

This project examines YouTube advertising activity of Indian companies between 2021 and 2024, with a specific focus on identifying ads featuring celebrity endorsements. As digital marketing becomes increasingly central to brand promotion, YouTube serves as a key platform for companies to reach consumers. Celebrity endorsements, known for boosting visibility and brand credibility, play a significant role in these campaigns. In this regard, the project utilizes the YouTube Data API, which retrieves data from the official YouTube channels of each selected company. The following information is extracted during the analysis: the company name, date of publication, and links to the advertisements. Special attention is given to filtering unofficial or user-generated content by only taking into consideration advertisements published by the companies themselves. The system also uses keywords and other methods to identify possible celebrity endorsements in the ads. This study aims to draw from insight on how Indian companies have positioned YouTube for advertising in these three years, with the purpose of ascertaining the role that celebrity endorsement has played in their digital marketing strategies. This will allow one to understand if the celebrity-driven campaigns are a common feature and whether they stayed influential in brand recall between 2021 and 2024.

RETAIL ANALYTICS: ARTIFICIAL INTELLIGENCE BASED PROPENSITY MODELS

Shreya Chhabra

Central Analytics
Tata Motors
Mumbai, India

Shreya.Chhabra@tatamotors.com

Shilpi Chakraverti

Central Analytics
Tata Motors
Pune, India

Shilpi.Chakraverti@tatamotors.com

Gautam Nath

Central Analytics
Tata Motors
Pune, India

Gautam.nath@tatamotors.com

Pratik Bagellu

Central Analytics
Tata Motors
Pune, India

PVB5248888@tatamotors.com

Abhinav Shukla

Central Analytics
Tata Motors
Pune, India

ASS529946@tatamotors.com

Amit Maheshwari

Central Analytics
Tata Motors
Pune, India

amit.maheshwari@tatamotors.com

The automotive industry is undergoing a rapid transformation driven by emerging technologies, with Tata Motors Ltd. (TML) at the forefront. Leveraging artificial intelligence (AI) and machine learning (ML), TML is innovating across Engineering, Manufacturing, Sales, Marketing, and Customer Experience. This commitment to AI/ML enables TML to maintain a competitive edge across all business units, particularly in the sales function, where increasing lead conversions is a key objective. Predictive analytics has become central to enhancing sales performance by identifying high-potential prospects—both new and existing customers—who are likely to engage and purchase vehicles. This paper explores TML's application of propensity models to optimize lead conversion across the pipeline. Various machine learning algorithms, including Random Forest, XGBoost, Logistic Regression, and LightGBM, were assessed to predict customer propensity. XGBoost achieved the highest accuracy for prospect ranking at 67%, while CATBoost delivered an 87% accuracy in pipeline prioritization. For predicting exchange propensity, Random Forest achieved a 59% accuracy rate. These model outputs were integrated into field-level applications, enabling more targeted outreach by sales teams. The deployment of these models has led to higher conversion rates and more efficient sales efforts, underscoring TML's commitment to leveraging AI/ML for actionable customer insights and reinforcing its leadership in digital innovation within the automotive industry.

ADVANCING ENGINEERING DRAWING INTERPRETATION – A COMPARATIVE STUDY OF PROMPTING TECHNIQUES

Pritam Varma

Data Science Intern

Genpact

Hyderabad, India

pritam.varma@genpact.com

Kolli Nethre Sai

Senior Data Scientist

Genpact

Hyderabad, India

kollinethre.sai1@genpact.com

B Lakshit

Data Scientist

Genpact

Hyderabad, India

b.lakshit@genpact.com

Divya Jyothi Mutyala

Machine Learning Engineer

Genpact

Hyderabad, India

divyajyoti.mutyala@genpact.com

Koundinya N V S S

Senior Data Scientist

Genpact

Hyderabad, India

ORCID: 0000-0003-1663-3112

Chirag Jain

Principal Data Scientist

Genpact

Bengaluru, India

chirag.jain4@genpact.com

This paper explores the automation of information extraction from engineering drawings, which are challenging to interpret due to their visual complexity. We propose a multi-faceted approach that employs advanced prompting techniques—such as Chain-of-Thought, Tree-of-Thought, and Reflexion prompting—within multimodal models to improve component detection and dimension extraction. Our research includes a custom dataset categorized into simple, medium, and complex engineering drawings, enabling robust model evaluation. Results show that advanced prompting significantly enhances extraction accuracy. Additionally, we leverage Vision Agents for automation of information extraction and analysis tasks in tandem with prompting techniques. This work underscores the potential of automated processes to streamline engineering design, reducing the workload on design engineers and improving the usability of engineering drawings.

Keywords: Information Extraction, Engineering Drawings, Automation, Multimodal Models, Prompting Techniques, Component Detection, Dimension Extraction, Vision Agents.

SUSTAINABLE DISTRIBUTION PLANNING- AN OPTIMIZATION APPROACH

Dr. Anuj Prakash, Premanand Bhupathy Raju, Manjula Saravanan

anujpra@gmail.com, premanand_braju@yahoo.com, manjula_honey@yahoo.co.in,
Tata Consultancy Services Limited, Bangalore- 560100, India

In recent years, the retail industry is going into a very tough situation due to competitiveness in the market, more demanding customers, hyper-personalization of the products, increasing rates of transportation, and pressurization for less carbon emission. The retail companies are looking for the distribution solution to meet the customers' demand at a lower transportation cost with less carbon emission. The distribution of goods, like sending goods to the stores from RDCs (Regional Distribution Centers) or CDCs (Centralized Distribution Centers), is having high fuel consumption resulting in high carbon emission due to long distances and higher frequencies of trips.

Based on the above discussion, the major point is to select the trucks in such a manner that the carbon emission should be minimized along with cost minimization, which is the main aim of any retail organization, while the service level should be maintained. In the present paper, an optimization approach has been introduced for distribution planning, where decisions should be made for the number of trucks required to transport the goods for fulfilling the orders. Genetic Algorithm (GA) has been implemented for solving the optimization problem, which provides us the near-optimal solution. The algorithm has been tested on the numerical example to show the efficacy of the algorithm.

Keywords: Distribution, Truck Optimization, Sustainability, Carbon Emission, Genetic Algorithms

HUNTING CLOUD EXFILTRATION

Meena Vyas and Craig Chamberlain

There have been four large-scale cases of data exfiltration and several smaller ones. Cloud exfiltration is not always detected as it happens and is sometimes uncovered later during incident response and investigation. Cloud intrusion sets often consist of credentialed access—someone’s credentials are compromised—which also can be resistant to detection. Hence, it becomes difficult to differentiate between a threat actor and normal activity. Globally distributed cloud systems can also pose challenges to security teams organized around conventional endpoint or network instrumentation and tooling. In many cases, incidents of credentialed access do not originate from an endpoint, or even a network, that is under the control of the victim organization, in which case EDR (Endpoint Detection and Response) and NSM (Network Security Monitoring) tooling cannot see it. In order to detect credentialed access, we need cloud service logs in some sort of analysis platform that can perform anomaly detection in addition to specification-based detection using static rules.

Why anomaly detection? In cloud systems, most log events involve things that are dual-use—the same methods to discover and exfiltrate data are used routinely to do authorized work. Many of these methods are used heavily enough to produce log entries numbering in the tens of millions. S3 audit events, when S3 is used as a data layer at scale, can number in the billions, in many log sets. Conventional rule-based alerting may find some instances of credentialed access but will also produce an unacceptably high number of false positives at this scale.

One of the most frequent requests heard is for some freely available code or tools that can be picked up and used to apply machine learning to cloud threat hunting. We put together a sort of starter kit with Python code and a small example CloudTrail dataset enriched with MITRE ATT&CK tactics. The CloudTrail data came from a real-world account and has been sanitized and tokenized so that it can be released.

CONVERSATIONAL AI FOR ENHANCED CUSTOMER JOURNEYS

Abhishek Kumar

Leader, Data Science

Cisco Systems

The incorporation of conversational AI has become a transformative force in the quickly changing field of customer service and support, greatly improving customer journeys. This study examines the creation and implementation of an advanced conversational bot intended to handle the different difficulties engineers and support staff have when providing intricate technical answers. The future of customer service and interaction is embodied by our conversational AI system, which stands out for its use of sophisticated large language models (LLMs) and a strong data ingestion infrastructure. Specifically created to optimize service engineers' processes, this AI-driven method successfully tackles the difficulties involved in organizing and carrying out intricate deliverables.

The solution offers a thorough and contextually appropriate assistance experience by combining unstructured data from several technical publications with structured data from databases. We consume enormous volumes of technical data, such as product roadmaps, configuration and design manuals, datasheets, and customer-specific data. To guarantee quick and precise information retrieval, this data is painstakingly processed and saved in a hybrid retrieval system that blends full-text search features with vector search. A Knowledge Graph (RAG) serves as the foundation for the system's design, enabling the smooth integration and contextualization of various data sources. Our conversational AI makes use of cutting-edge LLMs, such as GPT-4-O-mini, which have been optimized to comprehend and react to intricate technical inquiries. To produce accurate and useful insights, these models are trained in specialized domains. Real-time data fetching is ensured by the RAG's capacity to create queries and communicate with several APIs. For engineers who require current knowledge to make wise judgments, this capability is helpful. By offering thorough summaries and insights, the AI lessens the cognitive load on engineers, freeing them up to concentrate on making strategic decisions and solving problems. The system also includes sophisticated features like entity recognition, sentiment analysis, and context validation, which guarantee correct and pertinent responses.

To sum up, our conversational AI technology makes a substantial contribution to service engineering and customer support. We have developed a tool that improves the overall customer journey and increases engineers' productivity and effectiveness by utilizing LLMs in conjunction with an advanced data ingestion and retrieval framework. The technical foundations, real-world uses, and revolutionary effects on customer care and support will all be covered in this paper.

Keywords: Conversational AI, Data Ingestion Framework, Domain-Specific Training, Contextualization, RAG.

ENHANCING AIRLINES OPERATIONS USING ARDL: PASSENGER LOAD AND OTP EVALUATION

Anurag Sethi

Operations Control Center
Interglobe Aviation
Gurugram, India

anuragsethi1718@gmail.com

Shivank Singh

Operations and Planning
Celebi airport services India Pvt. Ltd.
New Delhi, India

shivanksingh680@gmail.com

Dr. Soumen Raj

School of Business
UPES
Dehradun, India

soumen9.11.1991@gmail.com

Airline profitability depends on efficient fleet utilization and precise scheduling. As passenger loads and flight delays are difficult to predict, airlines face serious operational issues. This study seeks to address these issues by employing the ARDL (Auto Regressive Distributed Lag) framework for forecasting passenger loads and on-time performance (OTP). In the study, a comparative analysis of historical data using ARDL provides relevant forecasts that airlines may utilize to optimize fleet allocation and aircraft scheduling. Through the study, we show that ARDL provides convincing forecasts that perform well in terms of accuracy and reliability, resulting in better planning and less operational disruption. Overall, this study concludes that when ARDL forecasts are integrated into airline operations, decision-making can become more reliable and innovative, and provide a competitive edge for fleet management, scheduling, and resource allocation.

Keywords: Airlines, Forecasting, Aviation, Operations, ARDL.

CHATGPT'S INFLUENCE ON STUDENT LEARNING: AN EXPLORATORY STUDY USING SENTIMENT ANALYSIS

Raghavendra Rao G N

Associate Dean, School of Management
ICFAI Foundation for Higher Education,
Off Campus Center Bangalore (Deemed to be University u/s 3 of UGC Act 1956)
Bangalore, India
raghavendraraog.n@ibsindia.org

Manisha Singh

Dean, School of Management
ICFAI Foundation for Higher Education,
Off Campus Center Bangalore (Deemed to be University u/s 3 of UGC Act 1956)
Bangalore, India
manisha@ibsindia.org

Muralidhara G V

Professor of Practice & Dy.
Director - SCMS Case Research Centre School of Commerce & Management Studies,
Dayananda Sagar University Bangalore, India
muralidharagv@dsu.edu.in

Large language models like ChatGPT are transforming education by improving students' knowledge, problem-solving, and analytical skills. This research examined the effects of ChatGPT on student behavior in various contexts. This study focuses on Indian graduate students and uses surveys and sentiment analysis to capture their perceptions and experiences of ChatGPT in academic settings. The goal is to evaluate ChatGPT's role as an educational tool while addressing dependency issues. The results will contribute to the discussion about AI in education and highlight the need for a balanced approach that promotes critical and creative thinking. The study argues for collaboration between educators, policymakers, and AI developers to ensure ethical and effective AI integration into education.

Keywords: ChatGPT, Large Language Models, Educational Impact, AI in Education, Student Perceptions, VADER Model, Sentiment Analysis, Topic Modeling, Problem-Solving Skills, Analytical Thinking.

HARNESSING DATA LABS FOR THE EVOLUTION AND GROWTH OF METRO RAIL SYSTEMS

Mr. Ashok Shankar. S

Manager – IT and AFC

Automatic Fare Collection System

Chennai Metro Rail Limited

Chennai, Tamil Nadu

India

ashokshankar.s@cmrl.in;

Dr. K. P. Mani

Field Specialist

Automatic Fare Collection System

Chennai Metro Rail Limited

Chennai, Tamil Nadu

India

vkpmany@gmail.com; mani.kp@cmrl.in

This study explores the transformative potential of Data Labs in enhancing metro rail systems' operational efficiency, passenger satisfaction, sustainability, and ridership growth. Leveraging synthesized and real-world data from multiple sources, four hypotheses were tested to assess the impact of data-driven decision-making on metro operations. The results indicate a significant improvement in operational efficiency and passenger satisfaction post-implementation of Data Labs. Furthermore, increased metro ridership contributed to reduced private vehicle dependency and traffic congestion, supporting urban sustainability goals. The findings provide valuable insights for metro operators and urban planners, highlighting the strategic role of Data Labs in optimizing metro performance and promoting sustainable urban mobility.

Keywords: Data Labs, Metro Rail Systems, Operational Efficiency, Passenger Satisfaction, Sustainability, Ridership Growth, Data-Driven Decision-Making, Urban Mobility, Transportation Analytics, Public Transit Optimization.

VOLATILITY FORECASTING OF BSE PRIVATE BANK INDEX USING ARIMA AND GARCH MODELS

(JEL Code: G170)

Pinak Dusane and Shubhangi Chatterjee

Prin. L. N. Welingkar Institute of Management Development & Research
(WeSchool), Mumbai

This paper evaluates the efficiencies of the two major time series models, ARIMA and GARCH, in volatility prediction of the BSE Private Bank Index as a critical benchmark for performance evaluation of privately-owned banks in India. With accurate volatility predictions crucial for risk management, investment strategies, and stability in general, the paper will draw a comparison of the forecasting capabilities of ARIMA and GARCH models by using historical data from the BSE Private Bank Index. Evidence is given that, even though the ARIMA model fits linear trends in the series well, it essentially lacks robustness in capturing volatility clustering as well as the time-varying volatility characteristics of financial markets. By contrast, GARCH(1,1) represents the dynamics of volatility correctly and is very close to reality in the market, showing great capability in modeling the volatility cluster as well as changes in the market. These results suggest that the GARCH model delivers a more reliable framework for modeling future volatility in the private banking sector and should be used at least in risk management, portfolio optimization, and the pricing of derivatives. It indicates further research on how GARCH models can be combined with other advanced machine learning methods to improve predictive results in various financial scenarios.

Keywords: Financial Forecast, Forecasting Returns, Risk Analysis, Stock Market Forecasting, Stock Market Simulation, Volatility Forecasting, Time Series Analysis

EXPLORING THEMATIC PATTERNS IN 35 NEWS ARTICLES: A DETAILED ANALYSIS USING PYTHON

Dr. Jeeta Sarkar, Aayush Garg, Gourab Chatterjee
Jaipuria Institute of Management, Noida

The growth of social data on the web is increasing rapidly nowadays. This leads to researchers to access the data and information for the research purpose. During the global COVID-19 outbreak, many individuals as well as organizations and government agencies are posting their viewpoints regarding the coronavirus. The study focuses on the sentiment analysis of news of the Google news using Python programming language with TextBlob library. The news articles have been collected, pre-processed and then used for text mining and sentiment analysis using Google Colab. The graphical representation has been provided on the data after sentiment analysis based on two specified hashtags.

Keywords: #coffee badging and #coronavirus.

AN EXPLORATORY STUDY OF STOCK PRICE MOVEMENTS FROM EARNINGS CALLS: A CASE OF INDIAN STOCK MARKETS

Ram Sundar G

Dept of CSE
Amrita School of Computing,
(Amrita Vishwa Vidyapeetham)
line 4: Bengaluru, India
grams1979@gmail.com

Peeta Basa Pati

Dept of CSE
Amrita School of Computing,
(Amrita Vishwa Vidyapeetham)
line 4: Bengaluru, India
bp_peeta@blr.amrita.edu

Radha D

Dept of CSE
Amrita School of Computing,
(Amrita Vishwa Vidyapeetham)
line 4: Bengaluru, India
d_radha@blr.amrita.edu

The primary focus of financial market research has been on identifying patterns in stock price, accounting, as well as other numerical "hard" data displayed on P&L statements or earnings per share reports. But it's common knowledge that decision-makers frequently rely on "soft" written documents to comprehend hard data it present. Current developments on computer techniques for large-scale analysis of unstructured also soft text-based data provide opportunities to comprehend the behavior of financial markets, which may enhance market equity and investment opportunities. Earnings calls are a crucial and common type of soft data. CEOs typically use earnings calls, which are recurring (often quarterly) remarks, to try and sway investors' perceptions of company's performance in the past as well as future. Here, we examine statistical connections among analyst recommendations, stock performance, corporate sales, and earnings calls. Our analysis spans ten years, from January 2015 to December 2019, and includes around 1000 transcripts of earnings calls of NSE 50 companies. We provide three discoveries in this study. First, there is little relationship among stock price fluctuations following earnings call also recommendations made by analysts before the earnings call. Second, we accurately and consistently forecast changes in stock prices across four key economic sectors by processing the semantic aspects of earnings calls using our graph neural network-based approach. Third, majority of situations, the semantic properties of transcripts that is, conventional hard data—are better indicators of stock price fluctuations than sales then earnings per share.

Keywords: stock price movement, Earnings call, natural language processing

BUILDING RESILIENCE THROUGH ESG: EVALUATING INDIAN PSUS IN THE CONTEXT OF GLOBAL EMERGING MARKETS

Shreyansh Kumar Sharma

Department of Mathematics

Indian Institute of Technology Kharagpur, Kharagpur, India

shreyansh.tebi@kgpian.iitkgp.ac.in

Moulik Kumar

Department of Physics

Indian Institute of Technology Kharagpur, Kharagpur, India

moulik@kgpian.iitkgp.ac.in

Abhinav Thakur

Department of Humanities and Social Sciences

Indian Institute of Technology Kharagpur, Kharagpur, India

abhinavthakur@kgpian.iitkgp.ac.in

Aumogh Sengupta

Department of Mechanical Engineering

Indian Institute of Technology Kharagpur, Kharagpur, India

aumogh.sengupta@kgpian.iitkgp.ac.in

Manas Jha

Department of Mathematics

Indian Institute of Technology Kharagpur, Kharagpur, India

manas.jha@kgpian.iitkgp.ac.in

Public Sector Undertakings (PSUs) play a critical role in India's economic development and sustainability efforts. With the increasing global focus on Environmental, Social, and Governance (ESG) factors, there is a growing need to evaluate the ESG performance of Indian PSUs in alignment with the country's sustainability goals, such as the National Action Plan on Climate Change (NAPCC) and the United Nations Sustainable Development Goals (SDGs). ESG performance is becoming a key measure of corporate responsibility and resilience, particularly in the face of economic downturns and crises. This research seeks to assess how effectively Indian PSUs adhere to ESG principles and integrate sustainability into their operations. The primary objective is to evaluate the corporate performance of PSUs using ESG indicators while considering the sustainability targets set by the Indian government. In addition to assessing the resilience of Indian PSUs during times of economic stress, the research will also compare their corporate performance with that of PSUs in other emerging economies to provide a global context to the findings. By using machine learning techniques, this study aims to develop advanced performance metrics based on ESG indicators. These metrics will be derived from both traditional financial data and alternative data sources, such as environmental records, corporate sustainability disclosures, and government policy reports. A critical component of the research methodology includes stress testing and adversarial scenario analysis, which will simulate how PSUs perform under different economic conditions, including downturns and crises. These tests will reveal potential vulnerabilities in the ESG strategies of PSUs, helping to identify areas for improvement in their corporate governance and sustainability practices. The comparative analysis with PSUs from other emerging markets will provide insights into how Indian public enterprises rank globally in terms of sustainability and governance resilience. The expected findings from this research are aimed at providing data-driven insights that can guide the development of more robust ESG frameworks for PSUs, not only in India but also in other emerging markets. The study will produce actionable recommendations for policymakers and corporate leaders, with a focus on enhancing the sustainability, resilience, and governance of public sector entities.

USER GENERATED CONTENT INFLUENCING HOMESTAY BUSINESS IN KARNATAKA

Mrs. Smitha Nadiga

Assistant Professor

Jyothy Institute of commerce and Management, Bengaluru University
Bengaluru, India

smitha.nadiga@jyothyicm.org

Dr. Roopa J

Assistant Professor

PES University
Bengaluru, India

roopaj@pes.edu

This study focusses on the homestay industry in Karnataka, India, located at an experimental stage of UGC literature trying to conduct a full business analysis for a growing sector. With the increasing popularity of experiential tourism, homestays are now being considered as a major alternative accommodation choice across culturally rich regions such as Karnataka.

The study employs a unique mixed-method approach that integrates secondary analysis of User-Generated Content (UGC) from various online platforms with primary data collected through surveys and interviews of homestay owners and guests. A total of 300 visitors and 150 homestay owners from different districts in Karnataka filled out standardized questionnaires to contribute to the primary data. 20 key stakeholders that are tourism officials and professionals in the industry were interviewed. Secondary data about homestays in Karnataka had been collected from popular travel websites, social media platforms, and online review sites which covered more than 10k user reviews/ratings of these homestays.

UGC content analysis found four important dimensions: authenticity of experience, hospitality, cleanliness and value for money as significant drivers leading to overall guest satisfaction. An examination of the finances of homestays reveals differing degrees of profitability based mainly on their location and the effectiveness of their marketing strategies. Additionally, it highlights the necessity for proficiency, in digital marketing the challenges posed by hotel competition and the impact of regulations on homestay operators.

The present research would fill the literature gap by conducting an empirical investigation into factors that influence user generated content and its consequences for business. Precisely, the homestay industry in Karnataka is of interest in this study.

These findings make suggestions on how the tourism boards, legislators, and homestay operators might enhance this industry's sustainability and competitiveness. The useful suggestions and further direction of study can be made for allowing us to understand issues deeper and ways we can integrate UGC into a homestay business plan.

Keywords: Homestay Tourism, User-Generated Content, Karnataka Tourism, Experiential Accommodation, Digital Marketing in Hospitality

DYNAMIC CRICKET MATCH PREDICTIONS THROUGH LSTM BASED SEQUENTIAL DATA MODELING

Debjyoti Sadhu

Senior Quantitative Analytics Specialist

Wells Fargo

Bengaluru, India

debjyotisadhu@gmail.com

Supriyo Sarkar

Quantitative Analytics Specialist

Wells Fargo

Bengaluru, India

supriyo.stat@gmail.com

Saurav Kuamr

Quantitative Researcher

Franklin Templeton

Hyderabad, India

sauravk82@gmail.com

The intricate dynamics of sports match outcomes pose significant challenges for predictive modeling, compelling sports analysts and data scientists to seek advanced methodologies. This paper presents a novel application of Long Short-Term Memory (LSTM) networks—an advanced class of recurrent neural networks—tailored to forecast cricket match results with heightened accuracy. We propose a sophisticated dual-model framework that constructs detailed player and team profiles from extensive historical data, serving as robust input features for two parallel LSTM architectures. By concurrently processing insights from both models, our approach adeptly captures the multifaceted interplay between individual brilliance and team dynamics. Comprehensive experimentation across domestic, international, and IPL datasets for the IPL 2022 season reveals that our dual-model LSTM framework consistently outperforms conventional Machine Learning models and standalone predictive models. The implications of this research extend beyond cricket, offering foundational insights for predictive analytics in various sports, enhancing strategic decision-making for teams, and refining odds calculation in sports betting environments.

Keywords: Deep Learning, RNN, LSTM, parallel models, sports predictions, IPL

IMPLEMENTATION OF A UNIFIED DATA LAKE PLATFORM: AN END-TO-END FRAMEWORK FOR DATA, BI AND VISUALIZATION PRACTICE ENABLING DATA-DRIVEN DECISION MAKING

Sougata Mazumder

Data and Analytics, Advisory
PwC India
Kolkata, India
sougata.mazumder@pwc.com

Tuhin Majumder

Data and Analytics, Advisory
PwC India
Kolkata, India
tuhin.m.majumder@pwc.com

Arghya Roy

Data and Analytics, Advisory
PwC India
Kolkata, India
arghya.r.roy@pwc.com

In today's world of data driven decision making systems, it is utmost important to make an enormous amount of data (aka Big Data) readily available to the consumption side i.e. reporting layer. At the core, Big Data comprises of 4 'V's, i.e., Velocity, Volume, Variability and Value. While the first three are dependent on Cloud Service providers (i.e. AWS, Azure, GCP, etc.), increasingly it is being realized that it is the 'Value' that creates the highest impact on any Data Driven Decision Making Process. In this paper, we will concentrate on unlocking the Value by discussing the implementation of a Unified Data Lake Platform (UDLP) versatile across industries. Adopting data-driven decision-making culture in an organization necessitates the need for a single source of truth in the form of a UDLP. In this paper we walk through one such UDLP that ensures seamless integration with various data sources (like SAP, CRM, SFTP, Third-Party APIs, etc.), clean the data using predefined and customized data quality check (DQC) process dynamically and finally transform and model those data using standard Industry practices. So essentially this data platform or the framework aims to create and maintain standard data models that are Sector and Technology Agnostic. In other words, this data platform would usher benefits to all types of Manufacturing Industries (except Service) by unifying, cleaning and modelling data from myriads of source systems. This data platform can be considered as a plug-and-play accelerator. The UDLP solution has been implemented across various industries and this robust framework is proven in paving the way for any organization into its cloud and data journey. As a by-product, it also gears up an organization to be data-ready for leveraging the latest GenAI capabilities as well.

Keywords: Data Lake; Data Architecture; UDLP; ETL; Cloud Platform; Big Data; Data Modeling

DATA DRIVEN DECISION MAKING BASED OPERATION RESEARCH IN SUPPLY CHAIN MANAGEMENT

Debanshee Datta

Faculty of Management
SRMIST, Kattankulathur
Chennai, India
dd2982@srmist.edu.in

Dr. V.M. Shenbagaraman

Faculty of Management
SRMIST, Kattankulathur
Chennai, India
shenbagv1@srmist.edu.in

Dr. Arunfred

Faculty of Management
SRMIST, Kattankulathur
Chennai, India
arunfren@srmist.edu.in

Recent trends on machine learning technology provides an improvised data analysis method that supports the entire enterprise process of procurement management. Hence it provides an improved management effectiveness of the operations in the supply chain so called as the smart supply chain operations. It analyses the upstream and downstream industry market status data in the supply chain and various primary data present in the enterprise management activities. The backbone of this smart supply chain management is based on multi-criteria decision making technique that operates on various products attributed by various criteria such as price, stock availability, lead times, order quantities, shipping times and shipping cost. Overall three different types of market product data have been selected for decision making analysis. This kind of analysis aimed at investigating the effect of supply chain management of product marketing decisions for a company. The data reduction method based on adaptive statistics was the most effective in revealing the market status and promoting the efficient operation decision-making based on supply chain management. This research work is focused towards the implementation of two multi-criteria decision making methods namely sum of additive weighting (SAW), Weighted Product Method (WPM) and technique of order preference of similarity ideal solution (TOPSIS) for selecting the best product for supplying from manufacturing company or industry to customer or to select the best supplier of the selected best product. Our experiment is based on three products such as 'Haircare', 'Skincare' and 'Cosmetics'. The reason for selecting TOPSIS, WPM and SAW methods of multi-criteria decision making is to find a comparative analysis via performance evaluation of both the methods in decision making based operation research in the domain of supply chain management.

MEASURING THE INFLUENCE OF MESSAGE AND CONTENT ON NBRX USING MULTI-TOUCH ATTRIBUTION: A HEALTHCARE ANALYTICS APPROACH

Chandran Venkatesan,
Senior Data Scientist, Mathco
chandran.v@mathco.com

Manoj Gundreddy,
Data Scientist Manager, Mathco
manoj.gundreddy@mathco.com

Pharmaceutical marketing has increasingly adopted omnichannel strategies, engaging healthcare providers (HCPs) across multiple digital and traditional touchpoints. Multi-Touch Attribution (MTA) is commonly used to assess the impact of these strategies on New-to-Brand prescriptions (NBRx). However, traditional MTA models tend to focus on attributing impact at the tactic level (e.g., email campaigns, webinars) without adequately differentiating between variations in messaging and content, which may lead to distinct outcomes among HCPs. This lack of detail limits the generation of actionable insights needed to optimize personalized marketing approaches. This paper introduces a novel MTA framework that leverages modified SHAP to improve attribution accuracy by analysing the granular impact of messaging and content within an omnichannel marketing context. SHAP values provide data-driven insights into how personalized messages influence NBRx across distinct HCP cohorts. The single-model approach captures the interactions between marketing channels, tactics, and creative assets, optimizing both overall strategies and individual components. Model Cross-validation ensures robust attribution estimates, while SHAP enhances model transparency by quantifying the contributions of specific messages. To mitigate increased variance at granular levels, data preprocessing consolidates features and reorganizes message and content based on activity volume, ensuring statistical stability and actionable insights. This enables pharmaceutical marketing teams to align performance metrics more precisely with marketing effectiveness, both at the tactic and granular levels. The approach also enhances targeting accuracy and optimizes marketing strategies to drive better NBRx outcomes. The paper includes large-scale validation across multiple drug brands, including those in immunology, oncology, and neurology, demonstrating how integrating SHAP into MTA models significantly improves campaign performance and marketing efficiency within the pharmaceutical sector.

Keywords: Multi-Touch Attribution, SHapley Additive exPlanations (SHAP), Omnichannel pharmaceutical marketing, New-to-Brand prescriptions, message and content analysis

MACHINE LEARNING-BASED PREDICTION OF GROSS ENROLMENT RATIO IN HIGHER EDUCATION INSTITUTES IN INDIA

Ms. Shalini Jayakumar

Research Scholar

SRM Institute of Science and Technology
Kattankulathur, Chennai, Tamil Nadu,
India

sj5103@srmist.edu.in

Prof. V M Shenbagaraman

Chair-Information Systems and Technology Management
SRM Institute of Science and Technology
Kattankulathur, Chennai, Tamil Nadu,
India

shenbagv1@srmist.edu.in

Dr. P. Saravanan

Associate Professor

SRM Institute of Science and Technology
Kattankulathur, Chennai, Tamil Nadu,
India

saravamp2@srmist.edu.in

The Gross Enrolment Ratio (GER) in Indian Higher Education Institutes (HEIs) is a crucial indicator for understanding the nation's educational and academic performance. Analyzing GER trends remains challenging due to various factors, including demographic, university-related features, and socioeconomic factors. In this study, enrolment was predicted to use various machine learning (ML) algorithms, including Linear Regression (LR), Decision Tree (DT), and Random Forest (RF). The dataset used to develop ML models included types and specializations of universities, enrolment in various programs, and social group data. The pre-processed data was split with an 80:20 ratio for training and testing, and a 5-fold cross-validation was performed. The ML results indicated that RF with hyperparameter was performed better than the other algorithms, with accuracies of 0.95, and 0.96 respectively. Feature importance analysis showed a strong correlation (0.77) between enrolment in scheduled caste and type of universities. Similarly, enrolment in state public institutions and institutes of national importance correlated more strongly (>0.75) compared to central universities (0.52). Among social groups, Scheduled caste enrolment showed a correlation of 0.82, compared to 0.43 for scheduled Tribe enrolment. This study predicts the enrolment ratio using ML algorithms for Indian HEIs and identifies the most influential factors.

Keywords: Gross Enrolment Ratio, Machine Learning, Higher Education Institute, Predictions, Feature Importance

DECODING ECONOMIC SUCCESS IN INDIAN CINEMA: A PREDICTIVE AND EXPLAINABLE AI APPROACH

Chitra Gautam

Indian Institute of Management Lucknow
Lucknow, India
phd23020@iiml.ac.in

Rohan Kumar Sinha

Indian Institute of Management Lucknow
Lucknow, India
phd22022@iiml.ac.in

Pradeep kumar

Indian Institute of Management Lucknow
Lucknow, India
pradeepkumar@iiml.ac.in

The Indian film industry, being the second most prominent in the world, has experienced substantial growth over recent decades and contributes significantly to the national economy. Despite its scale, only a fraction of released films achieves profitability due to a high-volume, low-value strategy that limits global revenue share. This study aims to address the gap in research focused on the Indian film market by investigating the major drivers of economic success using modern AI and explainable AI techniques. Drawing on an original dataset of 100 Indian films from IMDb Pro, spanning 2010 to 2024, we applied multiple regression models including linear, ridge, decision tree, random forest, and neural network regression. The target variable is taken as the gross worldwide revenue, and transformations were conducted on collected data to enhance model accuracy. The analysis utilized LIME (Local Interpretable Model-agnostic Explanations) for explaining predictions, offering transparent insights into significant determinants such as budget, star power, director influence, genre, and public awareness. This research contributes by identifying key economic drivers in the Indian film industry and providing practical insights for industry stakeholders to optimize decision-making.

Keywords: Predictive Analytics, Movies, Explainable AI, Artificial Neural network

DETECTING SCAMS IN REAL-TIME: A MACHINE LEARNING APPROACH TO IDENTIFY FRAUDULENT CALLS AND MESSAGES

Naveen Kumar A

Department of MBA,
PES University Bengaluru, India
being.naveen.a@gmail.com

Assistant Professor Deepika K S

Department of MBA,
PES University Bengaluru, India
ksdeepika@pes.edu

With the rise of digital communication, online scams/frauds have become increasingly easy and unavoidable, resulting in significant financial losses for innocent individuals, which is a threat to the society, individuals' financial security, resulting in economic losses and erosion of trust in online transactions. This research proposes a machine learning approach to detect scams in real-time, focusing on calls and messages (text and social media). Utilizing a dataset of fraudulent cases in Bangalore in 2024, we develop a natural language processing (NLP) model, along with key machine learning, decision making techniques such as ID3 and Cosine similarity index, text analytics, etc. to classify text based interactions/conversations as scam or legitimate. Our approach uses text analysis, and intent detection to identify threats, flags indicative of fraudulent activity. The proposed system aims to provide an accurate solution for individuals and organizations to mitigate the risk of scams, promoting a safer digital environment. This research aims to contribute to the field of NLP, Digital safety, and machine learning by exploring the application of these technologies in combating cybercrime. This study also contributes to the development of effective scam detection systems, providing a practical solution for mitigating financial losses and promoting a secure digital environment. The proposed system can be integrated into existing communication platforms of enterprises or personal apps to prevent fraud and protect individuals from financial exploitation, ultimately enhancing the safety and reliability of digital communications and finally the safety of society.

Keywords: NLP, ID3, Fraud detection, Financial Fraud, Online Safety

IMAGE-BASED CONTENT GENERATION WITH CULTURAL VARIATION: A DUAL ATTENTION AND FUZZY LOGIC FRAMEWORK

Ranjith Gnana Suthakar A

Department of Computer Science,
PES University Bengaluru, Karnataka - 560100, India.
Email ID: ranjith.gs@pes.edu

Sandesh B J

Department of Computer Science
PES University Bengaluru, Karnataka - 560100, India.
Email ID: sandesh_bj@pes.edu

Content Generation (CG), which provides meaningful descriptions to aid people in easily understanding visual content, is the process of creating descriptive text for images. Existing works have failed to investigate festive moments and individual culture during CG, leading to inaccuracies in differentiating cultural contexts for effective understanding. To address these issues, the proposed work introduces the Dual Attention-based MinTheil'sTsallis-LeL2Cun Sin-Shifted Convolved Gated Recurrent Unit (DAM (TL)²C²S-CGRU) with multi-data for CG enhancement. Initially, images are gathered and pre-processed. Subsequently, objects in the image are detected, and human images are extracted to identify emotions, enabling a clearer understanding of the image. Finally, the preprocessed image, detected objects, symbols, birds/animals, recognized emotions, and features such as dress, color, and facial characteristics are input into the DAM (TL)²C²S-CGRU classifier. This classifier generates culturally varied enhanced content with a remarkable accuracy of 99.67%. The proposed approach outperformed conventional methods, demonstrating its superiority.

Keywords: Canonical Renyi Correlation-based Median Filter (CRC-MF), Contrast Limited Adaptive Histogram Equalization (CLAHE), Eta-Kendall's Tau-based RetinaNet (EKT-RN), SCramér'sV Fuzzy (SCV-Fuzzy), Generating Captions, Text Conversion, Emotion Recognition (ER).

INTEGRATED TRAFFIC MANAGEMENT WITH INCIDENT PREVENTION AND DETECTION IN REAL-TIME

Muhammad Bilal S

Department of Artificial Intelligence and Data Science
Sri Sairam Institute of Technology,
Chennai, Tamil Nadu, India
sit22ad119@sairamtap.edu.in

Dr. Su Suganthi

Department of Artificial Intelligence and Data Science
Sri Sairam Institute of Technology,
Chennai, Tamil Nadu, India
hodai@sairamit.edu.in

Sathiya A

Department of Artificial Intelligence and Data Science
Sri Sairam Institute of Technology,
Chennai, Tamil Nadu, India
sathya.ai@sairamit.edu.in

Kaviya Sanjay K

Department of Artificial Intelligence and Data Science
Sri Sairam Institute of Technology,
Chennai, Tamil Nadu, India
sit22ad072@sairamtap.edu.in

Hari S K

Department of Artificial Intelligence and Data Science
Sri Sairam Institute of Technology,
Chennai, Tamil Nadu, India
sit22ad072@sairamtap.edu.in

In an era where road safety has become increasingly critical, innovative methods for managing traffic incidents and expediting emergency response times are essential. Traditional traffic management systems often rely on manual reporting, which introduces delays and inefficiencies that can worsen traffic hazards and compromise public safety. To overcome these challenges, this study proposes a real-time system integrating cameras, sensors, and artificial intelligence (AI) algorithms to transform traffic incident detection and management. This advanced system automates emergency service alerts, significantly reducing response times and enhancing traffic safety. The framework incorporates real-time updates via smartphone applications, AI-driven incident detection, continuous monitoring, and connected vehicle technologies. Furthermore, it includes a computer vision-based stress level detection feature, which identifies risks associated with driver fatigue and stress, enabling early warnings to mitigate accidents. By addressing the limitations of manual systems, this technology sets a new benchmark for proactive traffic control and incident prevention, fostering responsible driving practices and efficient emergency management.

Keywords: Road Safety, Traffic Management, Emergency Response, Real-Time System, Artificial Intelligence (AI), Traffic Incident Detection, Cameras and Sensors, Automated Alerts, Incident Prevention, Connected Vehicle Technologies, Driver Fatigue, Stress Level Detection, Proactive Traffic Control, Efficient Emergency Response, Responsible Driving Practices

COMPARATIVE ANALYSIS OF MACHINE LEARNING MODELS FOR OPTIMIZED DETECTION AND PERSONALIZATION IN MENTAL HEALTH

Ritik Samanta

PGDM Business Analytics
Vivekanand Education Society's Business School
Mumbai, India
ritik.samanta2325p@ves.ac.in

Prof. Nikita Ramrakhiani

PGDM Business Analytics
Vivekanand Education Society's Business School
Mumbai, India
nikita.ramrakhiani@ves.ac.in

Mrudula Geddam

PGDM Business Analytics
Vivekanand Education Society's Business School
Mumbai, India
geddam.mrudula2325p@ves.ac.in

The growing concern for mental health underscores an urgent need for innovative approaches to address treatment gaps caused by stigma and limited access to care. Disorders such as major depression and anxiety, which affect over a tenth of the population, are among the leading global health burdens, particularly in India, as reported by the World Health Organization. The post-pandemic era has exacerbated stress and social isolation, contributing to the rising prevalence of mental health disorders. This study explores advanced, data-driven methodologies to enhance the detection and intervention of mental health conditions.

A comparative analysis of predictive models—including Logistic Regression, Decision Tree, and Random Forest—evaluates their accuracy, precision, and effectiveness in identifying mental health conditions. These models are benchmarked against traditional diagnostic methods to determine the most valid and timely approaches for identifying mental health disorders. By highlighting the strengths and limitations of each model, the study provides guidance for selecting appropriate diagnostic and intervention strategies.

The research incorporates data visualization tools such as Power BI and Tableau, which enable dynamic and actionable presentations of complex data. These visual dashboards facilitate real-time monitoring of mental health trends, allowing for better assessment of risks and the development of targeted intervention plans. Demographic and lifestyle variables, such as gender, state of residence, family living status, education, sleep duration, income, and exercise, are analyzed for their correlations with psychological outcomes.

The study anticipates that integrating advanced machine learning techniques with interactive data visualization will significantly improve the diagnosis, prevention, and management of mental health disorders, optimizing treatment strategies and promoting quality care. This approach represents a crucial advancement in mental health management for India and the global community.

Keywords: Mental Health, Machine Learning, Data Visualization, Predictive Modeling, Logistic Regression, Decision Tree, Random Forest, Gradient Boosting Machines, Intervention Strategies, Post-Pandemic Trends

OPTIMIZING LARGE LANGUAGE MODELS FOR RESOURCE-CONSTRAINED ENVIRONMENTS: A PARAMETER-EFFICIENT APPROACH USING QLoRA AND PROMPT TUNING

Shivay Shakti

Senior Data Scientist, Comscore
New Delhi, India
shivayshakti2025@gmail.com

Drishti Hajong

MBA, Indian Institute of Technology (IIT)
New Delhi, India
drishti.hajong26@dmsiitd.org

Dr. Priyanshi Dubey

Medical Researcher,
Government Medical College (GMC)
Ratlam, India
priyanshidubeyjbp@gmail.com

Mudit Jain

BTech, Maharaja Agrasen
Institute of Technology New Delhi, India
mudit.jain2303@gmail.com

Apoorva Bhatnagar

MS, Nanyang Technological University (NTU)
Singapore
APOORVA006@e.ntu.edu.sg

The rapid growth of AI solutions, especially in resource-constrained environments, demands efficient and cost-effective deployment methods. Large Language Models (LLMs) pose substantial computational challenges, often making their implementation infeasible for many real-world applications. This study investigates parameter-efficient fine-tuning approaches, focusing on Quantized Low-Rank Adaptation (QLoRA) and Prompt Tuning, in conjunction with DistilBERT, to address these limitations. The proposed approach achieved a 36.2% reduction in memory usage and a 50% reduction in inference costs while preserving an accuracy of 87.75% compared to baseline models. These findings demonstrate that combining these techniques offers multiplicative benefits in resource optimization without significant performance loss, making them highly practical for real-world, resource-constrained AI deployments.

Keywords: Parameter-Efficient Fine-Tuning, Large Language Models, QLoRA, Prompt Tuning, Resource-Constrained Environments, NLP, Memory Optimization, Deployment Cost Reduction, Text Classification, Quantization, Low-Rank Adaptation

A OPTIMIZING DELIVERY: A MULTIFACETED ANALYSIS OF ZOMATO'S DELIVERY PERFORMANCE

Prof. Nikita Ramrakhiani

Vivekanand Education Society's Business School
Mumbai, India
nikita.ramrakhiani@ves.ac.in

Prerna Joshi

Vivekanand Education Society's Business School
Mumbai, India
prernajoshi.joshi2325p@ves.ac.in

Tanmay Sinpure

Vivekanand Education Society's Business School
Mumbai, India
tanmay.sinpure2325p@ves.ac.in

As of 2023, the Indian food delivery market was valued at approximately USD 36.3 billion (INR 2,997.9 billion) and is projected to experience significant growth driven by urbanization, rising disposable incomes, and evolving consumer preferences. This study examines the online food delivery (OFD) domain using data analytics and machine learning algorithms. Our research adopts a multifaceted approach, incorporating exploratory data analysis with visualization tools such as Power BI and comparative modeling using Linear Regression, Decision Tree, and Random Forest techniques to predict delivery times for Zomato delivery agents.

Zomato's delivery data was analyzed to uncover patterns, trends, and relationships between key factors like order volume, travel distance, traffic conditions, and agent performance. Additional variables, such as traffic congestion, peak hours, and weather conditions, were included to improve prediction accuracy. This study offers actionable insights into optimizing Zomato's delivery performance through recommendations related to route optimization, peak-hour management, agent training and assignment, and weather contingency planning.

The findings provide a comparative evaluation of machine learning algorithms in predicting delivery times and offer a roadmap for Zomato to enhance customer satisfaction, lower operational costs, and maintain its competitive advantage in the fast-growing food delivery industry.

Keywords: Zomato, Machine Learning, Data Visualization, Online Food Delivery, Regression, Decision Tree, Random Forest

OPTICHANNELING OMNICHANNEL RETAIL STRATEGIES BY INTEGRATING AIoT

Sinduja L

School of business and management
Christ (Deemed to be university)
Bangalore, India
Sinduja.l@res.christuniversity.in

Vedha Balaji

School of business and management
Christ (Deemed to be university)
Bangalore, India
vedha.balaji@christuniversity.in

Kerena Anand

School of business and management
Christ (Deemed to be university)
Bangalore, India
kerena.anand@christuniversity.in

Elangovan N

School of business and management
Christ (Deemed to be university)
Bangalore, India
elangovan.n@christuniversity.in

The Indian retail industry is undergoing a rapid digital transformation driven by advancements in Artificial Intelligence of Things (AIoT), opening new avenues for delivering highly personalized customer experiences. This study examines the emerging retail strategy of Optichanneling, where retailers optimize consumer interactions based on preferences and behaviors rather than merely integrating all channels into a traditional omnichannel framework.

By employing AIoT technologies such as dynamic pricing, predictive analytics, and in-store automation, retailers aim to create tailored shopping experiences. However, these advancements bring to the forefront critical concerns regarding consumer data privacy and trust. This paper applies the Institutional-Based Trust Theory, the Technology Acceptance Model (TAM), and the Privacy Calculus Theory to analyze how AIoT can be utilized for personalized optichannel strategies while ensuring data protection.

Using an exploratory inductive research approach, this study identifies key dimensions of optichanneling and proposes a conceptual model for AIoT-enabled personalized retail strategies. It also examines the strategic impact of these approaches on consumer data privacy and trust. The discussion provides insights into the challenges and opportunities facing the Indian retail sector as it seeks to adopt AIoT technologies and build trust-based, personalized optichannel solutions to enhance customer experiences.

Keywords: AIoT, Optichannel, Data Privacy, Personalization, Shopping Experience

HARNESSING PREDICTIVE ANALYTICS AND CONSUMER BEHAVIOR DATA FOR CO-CREATION IN DIGITAL COMMERCE: A PATH TO BUILDING STRONGER BRAND IDENTITY

Dr Sandesh Ramakant Bhat

0009-0005-1626-2886

MLA Academy of Higher Learning,
Bengaluru India

Ms. Poornima Nair

0000-0001-9259-291X

Alliance Ascent College, Alliance University
Bengaluru, India

Pragathi Prakash

0009-0007-2729-7915

School of Commerce and Management , Dayanand Sagar University
Bengaluru, India

The integration of predictive analytics and consumer behavior data within the evolving digital marketplace offers businesses a transformative opportunity to drive co-creation with consumers, thereby enhancing brand identity and loyalty. This study explores the role of predictive analytics in improving the understanding and forecasting of consumer behavior to facilitate more collaborative and personalized digital commerce experiences.

The research investigates how data-driven insights, focusing on key variables such as purchase frequency, customer lifetime value, product preferences, and consumer engagement, contribute to fostering co-creation projects. It further examines the influence of consumer behavior factors, including price sensitivity, feedback, and browsing patterns, on co-creation processes. These insights are pivotal in building emotional connections with consumers, strengthening brand loyalty, and fostering an enduring relationship between businesses and their customers.

By leveraging predictive analytics, organizations can identify actionable opportunities to personalize the consumer experience, enhancing engagement and ensuring their strategies align with consumer preferences. This study underscores the critical importance of integrating analytics into consumer engagement strategies, offering a roadmap for businesses aiming to navigate the dynamic digital commerce landscape effectively.

Keywords: Predictive Analytics, Consumer Behavior, Co-Creation, Digital Commerce, Brand Identity, Customer Engagement, Purchase Likelihood, Consumer Preferences, Data-Driven Insights, Brand Loyalty

ANALYZING CUSTOMER PREFERENCE AND SATISFACTION – A STUDY ON KEY ATTRIBUTES INFLUENCING APPLE IPHONE PURCHASE

Srihari Choori

Student, MBA

PES University Bangalore, India

sriharichoori13@gmail.com

Kiran Kumar

Student, MBA

PES University Bangalore, India

kirankumarr853@gmail.com

Aishwarya Mahendra Kapase

Student, MBA

PES University Bangalore, India

mkaishwarya2730@gmail.com

Technology has significantly impacted every facet of life, and the smartphone industry is no exception. Among the many players in this sector, Apple has distinguished itself as a game-changer with its cutting-edge features, including AI-powered cameras, seamless integration with augmented reality (AR) applications, enhanced privacy controls, and the ultra-fast A-series chips. These innovations have elevated the iPhone beyond a mere communication device. Additionally, the introduction of unique features such as satellite connectivity for emergency situations further strengthens Apple's position in the premium market.

Understanding customer preferences regarding these new features is crucial for both industry players and researchers. The competitive advantage lies not only in offering superior hardware but also in creating a robust ecosystem, continuous updates, and a high-quality after-purchase experience that meets the evolving expectations of tech-savvy consumers. This study aims to evaluate the factors influencing overall satisfaction and customer preferences in the purchase decision of the Apple iPhone.

A primary survey was conducted, consisting of 10 questions, to gather insights into the various factors that shape customer satisfaction and purchase intent. The findings of this study will provide valuable insights for industry players to refine their strategies and strengthen their market position, while also contributing to a better understanding of consumer behavior in the context of high-tech smartphones.

Keywords: Apple, iPhone, Customer Preferences, Purchase Decision, Technology, AI, Augmented Reality, Privacy, Consumer Satisfaction, Smartphone Industry

IMPACT OF TELEHEALTH ADOPTION ON HOSPITAL DEMAND FORECASTING

Preethish T

MBA Student

PES University Bangalore, India

Preethishjb007@gmail.com

The paper titled "Impact of Telehealth Adoption on Hospital Demand Forecasting" explores the implications of integrating telehealth technologies into healthcare systems, particularly focusing on changes in patient volume and resource allocation. The research investigates two primary perspectives based on a review of existing literature. One viewpoint suggests that the continued expansion of telehealth services, including teleconsulting, alleviates pressure on healthcare establishments by enhancing patient access and efficiency. Conversely, another perspective argues that the impact of telehealth on hospital demand and patient throughput may not be as positive as initially anticipated, instead reshaping how healthcare resources are utilized.

The study participants were doctors from various departments across multiple hospitals, with data collected via questionnaires to understand the challenges of integrating telehealth into hospital demand forecasting models. The collected data was analyzed using Logistic Regression through the JMP software. The findings highlight that while telehealth offers advantages such as flexible scheduling and cost savings, it also presents challenges like limitations in patient examinations and concerns over the quality of service. Telehealth has proven beneficial in managing chronic diseases, but its full potential as a healthcare delivery tool remains under discussion.

The paper discusses how the introduction of telehealth affects hospital demand forecasts, often replacing traditional methods. It also addresses organizational and policy barriers to integrating telehealth into existing healthcare systems. Despite these challenges, the study concludes that the opportunities for using telehealth in hospital demand forecasting are significant. However, achieving widespread adoption and ensuring high-quality service delivery will require collaboration among stakeholders and further technological advancements.

Keywords: Telehealth adoption, Hospital demand forecasting, Patient volume, Teleconsulting, Healthcare resources

SMART AND SUSTAINABLE PACKAGING FOR FRUIT LOGISTICS

Poongodi B

KCT Business School
Kumaraguru College of Technology
Coimbatore, India
poongodi@kctbs.ac.in

Mukundan Seshadri

GRNTECH Packaging & Storage
Solutions
Chennai, India
mukundan@grntec.co.in

Thyanidhi Dinesh E & Kalaranjani V

KCT Business School
Kumaraguru College of Technology
Coimbatore, India
thyanithidinesh.22mba@kct.ac.in,
kalaranjani.23mba@kct.ac.in

Biodegradable packaging solutions are gaining traction as sustainability becomes a key priority in various industries. In the context of the fresh fruit sector, the study "Sustainable and Smart Packaging for High-Value Fruits" explores retailers' perceptions and preferences regarding biodegradable and smart packaging technologies. According to Ghaffar et al. (2023), retailers are under increasing pressure to adopt sustainable practices, including eco-friendly packaging, in response to the rising demand from environmentally conscious consumers. Millennials, in particular, are inclined toward sustainable consumption, driven by the availability of sustainable products and packaging.

Previous research highlights the growing use of bio composites in smart packaging production, using techniques like extrusion and injection molding to create packaging that not only preserves food quality but also contributes to a circular economy by utilizing renewable resources. Poongodi et al. have demonstrated shelf-life extension using non-woven technology and bioingredients, while Ganeson et al. have studied intelligent packaging, which provides real-time data to enhance food safety and reduce waste. These innovations help foster better interactions between consumers and products, promoting informed purchasing decisions and reducing food loss.

A systematic review of recent literature using the Scispace AI research tool has identified key variables such as customer satisfaction, environmental impact assessment, food waste reduction, and sustainability initiatives. The study addresses gaps in the literature regarding market dynamics and logistical aspects, particularly in the context of fruit retailers. The research examines the environmental, economic, and retailer perspectives on incorporating sustainable packaging solutions in the high-end retail fruit industry.

The findings reveal that retailers generally hold a positive view of biodegradable packaging, recognizing its advantages such as improved shelf life, enhanced product quality, and increased brand recognition. The study underscores the importance of prioritizing sustainable and smart packaging to meet consumer expectations, improve corporate image, and maintain environmental responsibility. The integration of such packaging is crucial as consumer demand for eco-friendly products continues to rise, aligning with sustainability goals and addressing food waste and environmental concerns.

Keywords: Sustainable packaging, Smart packaging, Environmental impact, Retailer perception, Technology acceptance

THE EMERGENCE OF RATINGS OF AMAZON'S REVIEWS: FROM POOLING TO SEPARATING EQUILIBRIUM

Vallurupalli Vamsi, Prachi Sinha, Deepak K Sinha

Fake product reviews pose a significant challenge in online marketing, leading to a phenomenon where even producers of high-quality products feel compelled to purchase fake reviews, creating a pooling equilibrium that diminishes consumer trust and value. This study demonstrates that the introduction of product review rating platforms such as FakeSpot, ReviewMeta, and ReviewIndex helps establish a separating equilibrium. In this new equilibrium, high-quality product producers refrain from buying fake reviews. The absence of fake reviews, as reported by these rating systems, along with Amazon's original ratings, signals product quality to consumers.

Our research numerically validates this result over a wide range of parameter values, using data from FakeSpot, ReviewMeta, and Amazon to support our hypotheses. The findings suggest that these rating systems play a crucial role in fostering a market environment where genuine products are more easily distinguishable from fraudulent ones, enhancing consumer trust and driving better purchasing decisions.

Keywords: Retail Analytics; Fraud Analytics; Web Analytics

HYPER-PERSONALIZED ENTERPRISE-LEVEL PRODUCT RECOMMENDATIONS FOR HPE'S ECOMMERCE STORE

Pankaj Kumar

Global Marketing Hewlett Packard Enterprise India

pankajk@hpe.com

Dies P Varghese

Global Marketing Hewlett Packard Enterprise Canada

dies.varghese@hpe.com

Jyoti Khare

Global Marketing Hewlett Packard Enterprise India

jyoti.khare@hpe.com

Market research highlights the growth of B2C eCommerce as a strategic channel, valued for enhancing customer satisfaction and enabling upsell/cross-sell opportunities. HPE B2C eCommerce is rapidly becoming the preferred choice for buyers of various technology products. As eCommerce grows and buyers increasingly make larger, more complex purchases online, delivering an exceptional digital experience is no longer optional—it's essential. HPE has a clear vision to transition its volume products business online, particularly targeting the SMB segment. The HPE eCommerce Business team has proactively recognized this significant opportunity and is already taking steps to revamp its online store, 'Elite'. The goal is to offer a hassle-free, frictionless, and personalized shopping experience, thereby driving incremental revenue and higher conversion rates. This strategic shift positions HPE to meet the evolving needs of its customers in a rapidly changing digital marketplace.

HPE's extensive and complex product portfolio, including both off-the-shelf (BTO) and customizable (CTO) products, has traditionally relied on business acumen and sales intuition to build customer-specific Elite ecommerce stores. However, this approach has made it challenging to optimize product portfolio size and ensure comprehensive coverage across all product categories. To address this, a scientific, algorithm-driven approach is essential for better product portfolio management. Different sections of the Elite store require tailored, contextual, and cognitive recommendations to enhance the customer experience. For instance, on the homepage, it is crucial not only to recommend products the customer previously viewed but also to incorporate transaction-based and similar user-base recommendations. By using an ensemble of algorithms, we can deliver more accurate and effective product suggestions. When a customer navigates to a product details page (such as a configurator page), content-based recommendations become vital. For example, if a customer is viewing a DL server, the recommendation engine should suggest compatible products across relevant categories—such as memory, processors, power supplies, services etc. These recommendations should help the customer choose the right products for optimal performance, cost-efficiency, and balance.

By analyzing customer digital behavior for specific product interests, a ML model, known as the Digital Visitor Intent (DVI) model, predicts a visitor's product intent score, indicating their level of interest in that product. This DVI model and ensemble-based approach delivers hyper-personalized recommendations at HPE's B2C Elite eCommerce store, guiding customers to the best product bundles tailored to their needs. Implementing such advanced recommendation algorithms will transform how HPE's Elite store serves its customers, making it a more powerful tool for both optimizing the product portfolio and enhancing the overall buying experience. The proposed solution focuses on customer interest, product compatibility, and product refresh to deliver enterprise-level product recommendations.

Keywords: Machine Learning, Market Basket Analysis, Collaborative Filtering, Data Mining, XGBoost, Digital Visitor Intelligence, Product Recommendation

IMPACT OF MACROECONOMIC DETERMINANTS ON THE PERFORMANCE OF THE INDIAN STOCK MARKET AFTER COVID-19: A MULTIVARIATE TIME SERIES APPROACH

Jagdish Changani

Research Scholar

Department of Business Intelligence,
Gujarat University, Ahmedabad.

changani.j.g@gmail.com

Dr. Ravi Gor

Department of Mathematics

Gujarat University, Ahmedabad

ravigor@hotmail.com

The COVID-19 pandemic created unprecedented economic challenges worldwide, profoundly affecting the Indian stock market. This study investigates the influence of key macroeconomic determinants—interest rates, inflation (CPI), exchange rates, commodity prices (gold and oil), industrial production, and forex reserves—on the performance of the Indian stock market, with a focus on the Nifty 50 index during the post-COVID-19 recovery period. Using monthly data from January 2020 to August 2024, advanced econometric techniques, including Vector Autoregression (VAR), Johansen's Cointegration Test, and Granger Causality Analysis, reveal significant long-term relationships and short-term causalities among these variables.

Key findings highlight the pivotal role of interest rates, forex reserves, and oil prices in shaping industrial output and market stability, while gold emerges as a safe-haven asset with limited short-term predictive capacity. Stationarity tests confirm the suitability of most variables for time-series modeling, ensuring robust and reliable analysis. The results underscore the interconnectedness of macroeconomic factors and their implications for policy, investment, and market behavior in a volatile economic environment.

These insights offer practical applications for policymakers to design effective fiscal and monetary strategies, for investors to optimize portfolio management, and for businesses to align strategies with economic trends. This research bridges gaps in understanding the evolving dynamics of India's stock market post-COVID-19 and provides a foundation for future studies.

Keywords: Stock Market Dynamics, Post-Pandemic Economic Impact, Financial Market Recovery, Macroeconomic Indicators, Emerging Markets Analysis, Volatility and Risk Management

THANK YOU