भारतीय प्रबंध संस्थान बेंगलूर
INDIAN INSTITUTE OF MANAGEMENT
BANGALORE

DCAL
DATA CENTRE & ANALYTICS LAB

PROJECT ABSTRACTS

Project Title: Optimisation of Operation Theatre Utilisation and Keeping Abreast with Surgery Trends and Technologies at Apollo Hospitals

Abstract:

While the rising levels of income and wealth in India have resulted in better standards of living, the changed lifestyles have also led to rising instances of related diseases and disorders. Armed with the latest information, the patient now demands world class medical treatment and the best and latest surgical technologies. Keeping the changed scenario is mind healthcare service providers such as Apollo Hospitals have to plan well in advance to satisfy the resultant growing demand of speciality surgeries. This project involves discovering and discerning the trends in various surgical specialities over the last few years so as to forecast future trends and to facilitate optimal planning and preparedness in terms of surgeons, healthcare staff and operating equipment and infrastructure.

Project Title: Data Driven Analysis of Non-Revenue Water for BWSSB

Abstract:

In Bengaluru, nearly 48% of water is lost due to various physical and commercial factors which include leaks in distribution networks, losses from pilferage, non-functional meters, unauthorised connections, theft, etc. Bengaluru receives water primarily from Cauvery which is nearly 100 km away and the water is pumped nearly 500 m high. With rapidly depleting ground water resources due to overdraft, it is imperative that concerted efforts are undertaken to conserve water to ensure sustainability of the city. BWSSB intends to bring down the non-revenue water (NRW) levels to more acceptable levels and has embarked on a slew of initiatives.

This project aims to analyse real and apparent losses incurred by the application of relevant analytical techniques. Real losses comprise leakages, overflows at the utility's storage tanks etc. Apparent losses comprise meter under registration, reporting issues, theft of water, illegal connections, etc.

The study will analyse NRW data for the Srinagar – DMA (District Metering Area) and will include relevant data from other DMAs.

Project Title: Analysing social media debate and market sentiment around Maggi fiasco

Abstract:

Social media plays an important role in how consumers discover, research and share information about brands and products. Social media data combined with advance analytics techniques offers businesses ability to identify patterns in customer sentiment, any reputational risk and gauge marketing effectiveness.

This project is an attempt to demonstrate these capabilities by presenting the story of Maggi India as it goes through its worst turmoil – reports of high level of lead and mislabelling of the product followed by negative press, product recall, government ban and the negative effects of all of this amplified on social media. Nestle has meticulously won the legal battle but was their emotional come-back campaign equally efficient in regaining consumers trust and loyalty? How pervasive was the controversy – how did it impact the sales of Maggi and rival instant noodle brands? How did the market react to the ban? How did institutional investors perceive the ban? Has Nestle been able to regain the market share and recover its number one position in the instant noodles market?

Project Title: Analysis of Brand Switching Trend in Cookies Category and Product Strategy Based on Market Basket Analysis at Reliance Retail

Abstract:

The Indian hypermarkets and supermarkets form an extremely competitive market space. In order to address issues pertaining to the inventory, discount and profit margin, companies generally try and launch their own private labels or try to understand the brand switching within the category, so as to come up with innovative solutions to address the inventory and discount problems. Cookies is one such category, which is facing a similar set of problems and in order to address the issue, the project is divided into two major sub-projects: (a) Brand switching within the cookies category and (b) Affinity analysis for cookies. In order to study the brand switching problem, the Markov chains model is used. This will help in understanding the switching within the various cookie brands and to understand and estimate the percentage share of each cookie brand after *T* periods within Reliance retail.In order to study the affinity between cookies and other products, the market basket analysis is used so as to try and understand other products which are generally purchased along with the cookies. This will help in arriving at the various cross-selling schemes and also help in shelf space planning within the stores.

Project Title: iD. Special – Demand Forecasting for Home-Made Fresh Food

Abstract: iD Fresh Food (India) Private Ltd., is a leading ready-to-cook and eat packaged food company. The company is known for its popular product such as idly and dosa batter, Malabar parotas, wheat parotas, chapattis and chutneys. This project is an investigation of iD's supply–demand forecasting process and how it can be improved. The current forecasting technique follows almost the Naive method, in that it uses recent supply data to project supply for the current day. For the future state, an attempt has been made to recommend an alternative analytical technique that forecasts demand at a SKU-level per beat. One of iD's main operating regions, Mumbai, where it supplies about 8 different SKUs using 25+ beats was selected for the study and model build.

The resulting models which are analysed in this report are forecasts that predict the volume per SKU to be loaded into the van used for supply on a particular beat for a particular day. Multiple forecasting techniques are tested to determine the best performing options.

Project Title: Optimal Gate Assignment at a Major Metro Airport in India

Abstract:

For all major international airport operators, resource capacity utilisation and optimisation is the key to achieving its objective of hassle-free passenger transit, improved cost of operations and enabling increased airline and flight connections. Gate scheduling involves assigning gates to flights and is a key activity of airport operations. It is important for an airport operator to utilise the available gates in an optimal manner to maintain its service-level agreements with airline operators, while improving passenger service delivery. The model implemented uses a mixed integer programming approach to generate the schedule while adhering to the assignment policy. A further study of the variations in the fight arrival/departure times was accommodated to make the schedule more robust. The Metro Airport Operator's international terminal has seen rapid growth in airline connectivity and the proposed dynamic real-time scheduling model will help it to achieve operational excellence.

Project Title: Package Component Analysis for Manipal Hospitals: Designing "Standardized Packages" based on Consumption Pattern of Various Package Components

Abstract:

Be it hospitality, entertainment or any other service industry – customer satisfaction is one of the key success measures, which would make or break the existence of the brand. Health care is one such commodity, which is no longer termed as an institute that helps people walk out of ailments. Instead, people expect quality service with a high sensitivity to price. Consumers prefer to go for a package-priced surgery/procedure considering it to be financially beneficial; in return, the hospital attracts more customers. Package pricing is an act of placing several products or services together in a single package and selling for a lower price than would be charged if the items were sold separately. The package usually includes one big-ticket product and at least one complementary good. The intention of this study is to provide recommendations to the hospital, on what items are rightly justified to be under inclusions and exclusions bucket and in turn reduce the variance between net amount and the package amount.

Project Title: Customer Analytics for a Non-Banking Financial Company (NBFC) Segmentation and Campaign Strategies

Abstract:

Markets are driven by sentiments but none matches the uniqueness of the gold loan industry being driven by sentiments. The inherent social challenges of the industry is that demand is driven by unforeseen contingencies requiring immediate capital infusion with short to medium term maturity and the attached social stigma of having to pawn family jewellery, restricts sharing of good experiences of the engagements. The entry of banks into the gold loan business poses an immediate threat as these institutions can leverage the low cost of funds to offer a lower rate of interest. It has become imperative for the NBFC to start looking at the big question – "from where will the next million customers come?" This paper attempts to identify and classify the customer base of the particular NBFC and propose "Below-the-Line (BLT)" campaign strategies to attract more customers and thereby augment the market share. The analysis will also undertake the discovery and unveiling of insights from the data.

Project Title: Ensuring Safer Banking at ATMs Using Real-Time Intelligent Video Analytics

Abstract:

The paper proposes an analytical framework for detecting crime which is useful for the current security systems of banking institutions. The intelligent video analytics solution can be integrated with the existing security systems of banks to provide near real-time predictions of crimes and illegal activities to make banking through ATMs safe and secure. The proposed solution to detect and prevent crime and illegal activities at ATMs is divided into three phases. The first phase includes capturing live video and audio feeds from cameras deployed in ATMs and processing frames from the video feeds. In the second phase, processed frames or images are analysed using different image processing and analytical models to extract different features and identify any anomalies. Once the features are identified, the final phase involves classifying the events using different classification techniques and creating an alert if the event is suspicious. The study uses open source deep learning neural network models to detect facial features for mask disguise and emotion detection. The solution also makes use of face detectors such as OpenCV Haar Cascades and DLib for identifying faces in a frame which is a pre-processing activity for using the deep learning framework.

Project Title: Sentiment Analysis of BJP's Digital Media Campaign in Maharashtra

Abstract:

India's 16th general election was held in nine phases from April to May 2014. The 13th Legislative Assembly election for Maharashtra was held in a single phase on 15 October 2014. The state of Maharashtra presents a very interesting case to conduct a detailed analysis of the election results to understand the voting behaviour and patterns between the general election and the legislative election. A comparative study has also been carried out using 2004 and 2009 legislative assembly election results for the state to understand the voting shift. The entire state was segmented into 6 regions and assembly-wise analysis was performed to understand margin of victories, native strong and weak territories for each alliance and the potential impact of a combined alliance. The secondary aim of this project is to understand the impact of BJP's digital media campaign, particularly their "Whatsapp" campaign *Bhau Cha Dhakka* (Brother's Push) and *Aghadi Cha Ghotala* (scams by former government). The analysis looks at the impact of these two campaigns on the outcome of the election results in the targeted regions. The project uses exploratory data analysis techniques and advanced visualisation to understand the outcome of the election results.

Project Title: Marketing Campaign Effectiveness for Manipal Hospitals

Abstract:

Every year, corporations spend huge amounts of money on marketing campaigns. It is important to understand where/how the money is being spent and if the money is being spent wisely. The project aims to study the way in which the marketing campaigns are conducted in the healthcare domain and observe patterns in terms of departments/locations/seasons and try to correlate which of these if any offer returns in terms of increased footfalls or revenue for the hospital.

Project Title: Out of Order - Predicting ATM Failure.

Abstract:

This project aims to predict ATM failures in advance so that preventive maintenance schedule can be adopted by ATM operations team rather than the reactive maintenance model. This will help to reduce the down time of the ATMs. The project also aims to optimise the cost of maintenance by leveraging the cash replenishment visit for preventive maintenance repairs. The solution will evaluate clustering techniques to identify anomalous/similar switch log messages to identify failures. Sequence based prediction techniques will be used to predict future failure.

Project Title: Supply Chain Analytics – Developing a model by which sourcing costs can be optimised across Unilever Factories

Abstract

Companies need to keep generating fuel for growth in terms of cash to invest in brand building and brand development. A significant part of costs is incurred in supply chain by way of materials, factory operations and logistics costs. The dynamics of the markets cause swings such that there is always new scope to optimise costs by either cutting costs or moving sourcing to cost-effective locations. Unilever has more than 500 manufacturing units across the globe manufacturing more than 10,000 SKUs. Several of these units manufacture similar products to cater to markets nearby. This scenario allows the company to explore and benchmark units to arrive at an optimal sourcing decision. The project aims at developing a model that looks into the forecasts prepared by the company to do the benchmarking and identify the most optimal sourcing mix from a business category point of view.

Project Title: Crystal Ball – Developing a Volume Forecasting Model (Top Departments) for Manipal Hospitals

Abstract

Providers usually keep a close watch on key performance indicators (indicative not exhaustive) that include a mix of key service metrics such as patient satisfaction scores, waiting times and financial metrics such as revenue generation, bed utilisation and staff–patient ratio apart from many clinical quality parameters. Healthcare providers have traditionally relied on historical forecasts for predicting patient demand. Estimating the right demand would help the hospitals in better tactical and strategic planning/hiring for staff and bed availability. Forecasts also help in making key infrastructure decisions across the various departments, and better manage quality of care and patient experience. As a pioneer in health care, Manipal Hospitals is among the largest hospital networks in India, serving over 2 million patients annually, with a network of over 15 hospitals, managing and operating over 5,200 beds across the group. The research aims at developing a volume forecasting model for specified departments.

Project Title: Design with Feedback in Mind: Predict Net Promoter Score (NPS) for Analytics Training Programs

Abstract:

Net Promoter Score (NPS) is based on the fundamental perspective that every company's customers can be divided into three categories namely the promoters, passives and detractors. These three groups are determined by asking customers the following question: "How likely is it that you would recommend this company to a friend or colleague?" Using an 11-point scale, promoters score 9 or 10; passives score 7 or 8, and detractors score 0–6. The NPS is calculated by subtracting the percentage of detractors from the percentage of promoters. The score, which is an indicator of participants' satisfaction, is a strong predictor of the training's ability to drive sustainable learning and knowledge. NPS analysis can help measuring the satisfactory outcome of the trainings conducted, identifying the key drivers influencing the satisfaction of participants, determining the patterns. The insights generated through NPS analysis can bring positive changes into the system to improve efficiency and effectiveness of the training function. This paper captures the essence of the objective stated above by analysing the available data for a series of trainings conducted in 2015 across different expertise areas.

Project title: TweetFolio Investment Advisor

Abstract:

Social media is a place where users present themselves to the world, revealing personal details and providing insights into their lives. This study aims to understand how some of this information can be utilised to recommend an individual with optimum investment strategies based on his/her social media activity. The study presents a method by which a user's financial personality and investment needs can be predicted through the publicly available information on his/her Twitter profile. Further, the recommender algorithm will select the best combination of asset classes for investment, based on the determined financial personality (risk-taking ability, longevity, etc.) and demographic profile (age, gender, education, etc.).

Project Title: PROJECT – UNIBIC: Developing a Framework for Optimising the Logistics Costs

Abstract

UNIBIC Foods India Pvt. Ltd. manufactures and markets premium cookies. UNIBIC has a major challenge in reducing the logistics cost involving the distribution of cookies across 15 primary distributors and 35 secondary distributions, which includes private labels, exports, and institutions.

The objective of the project is to optimise the logistics costs by using various techniques of data analytics and provide potential solutions to the stakeholders and process owner and predicting the optimal number of trips for the transportation and distribution without affecting the demand requirements.

Project Title: Making farming profitable by predicting market prices of crops based on mobile applications

Abstract:

Contribution of agriculture to India's GDP decreased from 23% in 2000 to 18% in 2014 [1], due to many factors including climate change, lower productivity and increased cost of farming. There is a constant need to improve productivity and make agriculture more profitable to provide livelihood to the vast workforce depending on agriculture sector and ensure food security. Jayalaxmi Agro Tech Private Limited (JA Tech), a social entrepreneurship start-up firm, developed several crop specific mobile applications for farmers in English and regional languages of India. Farmers who click on these applications contribute to data collect. Important details like location, crop details and number of clicks are available which help indicate important aspects. For example, clicks on a particular disease icon by many farmers in a specific area indicates possible outbreak of a disease. If the disease is not contained, there is a high probability that it impacts yield from the crop. The quality of produce and yield along with market parameters such as supply, demand and exports significantly influence market price of the crop.

This paper proposes an advanced analytics model to predict market price of crops by using crop specific mobile usage data, along with secondary data such as geographical, weather, crop cultivation, yield, production and market inputs.

Project Title: Value Analytics in Motor Insurance

Abstract:

Motor insurance accounted for 39.41% of the gross direct premiums till September 2015 in the Rs.78,000 crore premium per annum non-life insurance industry in India. A majority of the Indian insurance players are struggling to keep the motor portfolio operationally profitable. The project sponsor wants to optimise their operational profitability, thereby garnering a competitive edge in the market. With the increasing Internet usage across various customer segments in India, the project sponsor is looking to optimise the E-business channel with two key objectives: (1) Design an optimum media plan to improve effectiveness of the marketing spend, with a higher ROI and reduced cost of customer acquisition and (2) Identify loss-making segments by evaluating Customer Lifetime Value, and thereby providing recommendations for an optimal customer mix. The data analysis leads to a linear optimisation model on basis of key parameters from vehicle and policy data to reduce the cost of acquisition; whereas the customer life time value analysis leads to identification of vehicle segments where the company needs to be aggressive (in New Business and Retention Business) and the segments where they need to be conservative in the long-term point to increase the portfolio's Top Line while maintaining or improving the bottom line.

Project Title: RecoE – A Recommender Engine for Content Websites

Abstract:

Website publishers have realised that is not only important to attract visitors to the site through innovative marketing campaigns, but also to convert them into loyal customers. The RecoE project's objective is to put together a data-driven system to recommend pages that would be of interest to visitors, based on the content of the page they are currently reading. The project involves the following approach. Key content features of pages published on a website are identified, followed by summarising this information into a matrix which can be used to compute similarity scores and make recommendations. First, relevant information is scraped from the website and pre-processed to eliminate noise. Then, the content is tokenised and the term frequency – inverse document frequency (TF-IDF) scores are computed. Next, the similarity scores between different webpages are computed by using metrics such as Euclidian distance and cosine similarity. Finally, the models are validated and the best algorithm that can be coded into a workflow for deployment on the live site is arrived at.

Project title: Sentiment Analysis of Bollywood Movies – *Queen* and *Gulab Gang*

Abstract:

Sentiment analysis is the task of extraction of sentiment, the positive or negative orientation that a writer expresses towards some object. This project investigates sentiments about the films – *Gulaab Gang* and *Queen* – at the time of trailer launch and movie release to understand how close is social chatter to field survey. Using movie reviews from websites such as IMDB, Mouthshut, YouTube, etc. as data, the project applies three machine learning methods (Naive Bayes, maximum entropy classification, and support vector machines) to determine the sentiment of each review. The study concludes by examining factors that make the sentiment classification problem more challenging. Finally, an attempt is made to provide suggestions to improving the performance of the models using advanced language processing techniques.

Project Title: Prediction of Customer Cancellations and Monthly Collections

Abstract:

VBHC Value Homes Private Limited (VBHC) applies contemporary building technology, computer-aided design methods, a lean manufacturing process and ingenious industrial engineering processes to deliver affordable housing. To ensure better planning and cost control, VBHC wants to address two significant problems: (1) Since lower sale price makes the cost of acquisition very high in the affordable housing segment, identifying possible cancellations at earlier stage is very helpful in controlling the sales and marketing costs. How can VBHC predict unit cancellations that are likely to happen? (2) From a cash flow perspective, accurate prediction on monthly collection is essential for future project planning and investments. How can VBHC predict monthly collections that are due from customers and can the customers who are more likely to default be identified as soon as a bill is generated?

Project Title: Unified Social Media Analytics Report and Dashboard Using Advanced Analytics to Help the Largest Online Automotive Portal of India

Abstract:

CarDekho, a fast-growing company, an online search venture for new cars and used cars, has been very active in social media marketing. CarDekho has strived to increase performance and ROI through all possible channels and has delivered effective results every time. Social media is one of those channels which has acquired many leads and converted them into consumers. Hence, it is essential to track the conversion of such leads into consumers and social media metrics play an important role in tracking the impact of social media activity on gaining leads. This project aims at creating a report which helps to understand CarDekho's performance with the help of key performance indicators (KPIs) with respect to its competitors and understand how its impact is growing. The study resulted in successful building and deployment of a dashboard for CarDekho which is being used to better understand the effectiveness of advertising campaigns and work out the ROI. Also, CarDekho is using a monthly unified report which was prepared to understand and track its competitors and plan its social marketing efforts.

Project Title: Real time Audience Engagement metrics using Video Analytics

Abstract:

Many institutions today highly rely on the engagement of audience in an event and solicit feedback from them by means of survey or feedback forms. Several participants hesitate to provide feedback due to time constraints or other reasons thereby bringing down the response percentage to merely 10%. Feedback from the audience is an extremely important criterion needed to assess the engagement quotient and there is a need for a reliable solution with a quick turnaround time to assess the same.

This paper aims at how we can get beyond the means of survey and devise an intelligent feedback gathering through means of video technology and Analytics. The method uses features like audience's head pose (yaw and pitch angles) and face detection program to detect engagement score along with the audience mood metrics (by means of happiness emotion index). The solution is robust enough to be agnostic of speaker movement around the room as it uses the concept of collective intelligence and unsupervised clustering algorithm to derive outliers (disengaged audience) within each group. The solution being real-time, helps a speaker identify quadrants or groups of audience where the audience is disengaged and takes actions in real-time to improve the overall engagement.