SOCIAL MEDIA OPINION MINING FOR ISKCON TEMPLE USING NLP

Vinodhini Ranganathan*
St John's Research Institute, Bangalore

Vigneshwaran SR
Indian Institute of Science, Bangalore

* Email: vinodhiniranganathan14@gmail.com, M: 9488317475

Abstract

People visit temples to pray for their own physical, mental and material well being. Very few go to temples for indulging themselves in some charitable social work. In today's world, temples can be categorized into two types: traditional temples that emphasize & preach "bhakti" and contemporary temples that preach & practice "social responsibility with bhakti". One notable TEMPLE (Tranquility Education Medication Purification Love Engagement) that falls in the second category is "ISKCON" (The International Society for Krishna Consciousness). Temples like ISKCON are not only meant to be a place for religious worship, but also as a charitable society that provides free food, free education etc. Such societies see rapid growth in a short span, thereby making it difficult for the management to monitor people's perceptions and identify factors that denigrate their brand equity. With the growing availability and popularity of opinion-rich resources such as online review portals and social media, new opportunities and challenges arise as people can now actively use social mediums to not only share their opinions but also read others' opinions. Therefore, monitoring social platforms is very essential for a temple's feedback management system. By extracting and analyzing the comments from social media platforms, such as Twitter and Facebook, the management can easily identify visitor's pain points and take immediate action to rectify them. In this paper, we develop a systematic approach to extract users' opinions about ISKCON Bangalore from social media and then perform sentiment and emotion analysis to understand visitor's concerns.

1. Introduction

In the recent times, social media has become an indispensable channel for building brand awareness, engagement, and marketing. Social media data is a treasure trove of information compared to traditional surveys because of the volume of user-generated content available online. Another notable advantage with social media data is accessibility - the data is usually openly available and easy to extract.

Sentiment analysis (SA) sometimes alternatively mentioned as opinion mining, is a research area which aims to analyze people's sentiments or opinions toward entities such as topics, events, individuals, issues, services, products, organizations, and their attributes (Liu, 2012). In today's generation, people depend a lot on online sentiment reviews. In fact, 90% of customer's decisions such as visiting a restaurant or booking movie tickets depend on online reviews (Ling et al., 2014). Moreover, SA has become an essential entity for brands in decision-making (Tawunrat and Jeremy, 2015; Matthew et al., 2015) and it is heavily used across many fields such as consumer information, marketing, books, websites, and social. However, SA for temples and charitable societies is still unheard of. Quite often, temples are seen as a structure reserved for religious or spiritual rituals and activities such as prayer and sacrifice. One might wonder why temples want to gauge their brand equity through SA. In this study, we have attempted to analyse the brand equity of ISKCON through social media opinions/reviews.

2. Objective

The main objective of this study can be boiled down to the below aspects

- Identify the facilities/factors that creates a negative impact on people when they visit the temple from social media reviews and resolve/improve these factors.
- Reduce the manual work spent on the current analysis by using Machine learning (ML) and automate the whole process by creating a dashboard. This dashboard would have to be plug-and-play, which automatically gives the emotions, sentiment splits and few descriptive plots when the data is fed into it.

3. Review of Literature

Research on Sentiment Analysis (SA) has seen rapid growth and changes in terms of content over the years. Prior to the availability of massive online content such studies mainly relied on survey-based methods and were focused on public or expert opinions rather than users or customers' opinions. The first paper that matched our search on SA was published by Stanger (1940) and it was titled ''The Cross-Out Technique as a Method in Public Opinion Analysis''. In 1945 and 1947, three papers appeared that addressed measuring public opinions in post WWII countries that had suffered during the war (Japan, Italy, and Czechoslovakia) and they were all published in the journal Public Opinion Quarterly (Knutson, 1945; Fegiz, 1947; Adamec, 1947).

In the mid-90s, once computer-based systems started to appear, a paper titled ''Elicitation, Assessment, and Pooling of Expert Judgments Using Possibility Theory'' was published by Sandri et.al (1995) where a computer system was used for expert opinion analysis for analyzing industrial safety by pooling the opinions. Another branch of work that was highly influential and eventually led to the birth of modern sentiment analysis carried out by the Association for Computational Linguistics (ACL) founded in 1962. ACL seeded the concept of computer-based sentiment and Wiebe (1997) proposed methods to detected subjective sentences from a narrative in 1990 and later proposed a gold standard for this. (Wiebe et.al, 1999).

Being precise of what has been the first paper of modern sentiment analysis is hard as early years used fluctuating terminology. Using citation counts from Scopus and Google Scholar, we can pinpoint many influential papers started from the turn of 21st century. Hatzivassiloglou and McKeown (1997) published a paper titled ''Predicting the Semantic Orientation of Adjectives'' where they used data from 1987 Wall Street Journal corpus and built a list of positive and negative adjectives and predicted whether conjoined adjectives are of the same or different orientation. ''Thumbs up? Sentiment Classification using Machine Learning Techniques'' Pang et.al (2002) used movie review data and found out that machine learning classification outperformed human-produced baselines. This kick started research on modern day SA. In the same year, (Turney, 2002) published a paper with a very similar title: ''Thumbs Up or Thumbs Down? Semantic Orientation Applied to Unsupervised Classification of Reviews'' where he used online reviews of automobiles, banks, movies, and travel destinations and achieved on average accuracy of 74% for recommendations. More groundbreaking papers were published in 2003. Turney and Litmann (2003) in a paper titled ''Measuring Praise and Criticism: Inference of Semantic Orientation from Association'' proposed a method to automatically infer the semantic orientation of a word from a statistical context. Another notable work titled, ''Mining the peanut gallery: Opinion extraction and semantic classification of product reviews'' positioned SA in the context of product reviews available in the Web. The authors argued that automated analysis of such reviews is needed due to the high volumes (Dave et.al, 2003).

As of 2016, the number of papers on SA counts to 6996. In the same way sentiment analysis has been done using various methodologies and for very diverse applications. Their detailed use of methodologies and applications can be found in the survey papers (Lin et.al, 2018; Doaa, 2016; Kumar, 2015; Walaa et.al, 2014). However, the SA methodology proposed in this paper is shown in Figure 1.

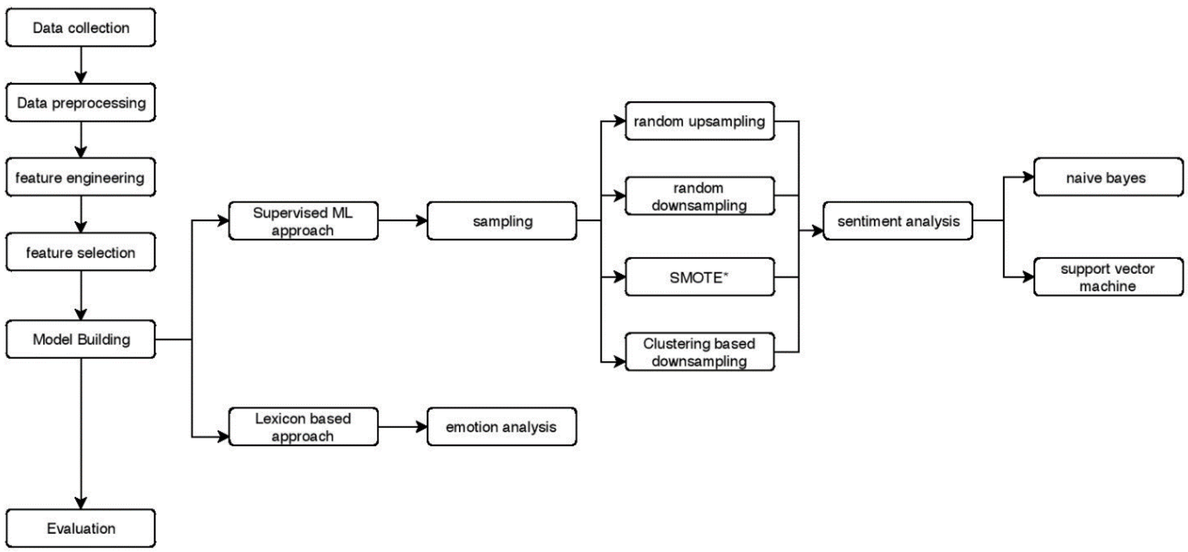

Figure 1: Proposed Sentiment Analysis (SA) Methodology

4. Data Collection and Preprocessing

4.1 Data collection

Data was manually gathered in "CSV" files by the ISKCON IT team from three social mediums: TripAdvisor (TA), Facebook (FB) and Google plus (G+). The data had 5685 reviews/comments collected from February 2013 to August 2017, out of which 66% of the comments were from TA, 25% from G+ and 9 % from FB. Each comment was classified into one of the four different categories: positive, negative, neutral and mixed. A sample of the raw data is given in Figure 2.

REVIEW ID	SOURCE	REVIEW BY	REVIEW DATE	REVIEW SUBJECT	REVIEW COMMENT	REVIEW RATING	REVIEW TYPE
1	Facebook	Vinod Kumar	10/1/2015		So Dedicated and Divotional Temple Lord Venkateswara Swamy in Bangalore. I Like it	5	POSITIVE
2	Google + HK HILL	kamalyadav 1987	3/3/2015		Place is great but they made it business on the name of God. But this is not the only worship place which is doing so?	5	MIXED
3	Trip Advisor	er_mansury	1/9/2014	Its a mall rather than a templ	They have build this tample to earn money. Everywhere they created stalls and the path is around those stalls and you ave to apss though each stall. No peace at all. Delhi Isckon temple is better than this,	1	NEGATIVE
4	Trip Advisor	bisbani23	19/7/2016	ISKON temple Bengalore	It is situated near Sandal soap factory. It is a holy temple for Hindu. God Radha Krishna stay here. Surrounding is picturised.	4	NEUTRAL

Figure 2: Sample Data

4.2 Data Preprocessing

As with any data analysis, data cleaning was done with due diligence since there were many duplicates and missing data because of the manual data collection process. Other issues with the data was that the comments had other Indian languages written in English. There were also a few comments in French and German that required translation for further analysis.

Problems with the data

- Duplicates - Same comment posted twice by 2 different members of the IT team
- Duplicates - Same comment posted twice with different dates by the same person on social media.
- Missing data - Comments were not copied fully from the social mediums. The comments had missing lines and broken sentences.
- Data error - Same comment classified into 2 types(positive and negative) by 2 different members
- Data Error - Same comment posted in 2 different social mediums(TA and FB) by the same person
- Comments in multiple languages that require translation to English
- Informal, slang ,short words
- Spell errors and creative spellings

The total number of reviews came down to 4940(18% reduction) after removing duplicates. The remaining data was cleaned further by removing punctuations, stop words and whitespaces. The data was normalized with stemming and lemmatization and by converting all the words to lowercase. Creative spelling and slangs were handled by using slang dictionaries and replacing slang words with English dictionary words. Spell errors for words with one or two missing or misplaced characters was corrected using spell dictionaries.

5. Exploratory Data analysis

After cleaning the data, next step was to perform Exploratory Data Analysis (EDA) to understand and summarize the data. The first step was to plot a few histograms to identify the proportions between social mediums and the review type. Most of the reviews were from TripAdvisor as shown in Figure 3. It was also observed that the data was imbalanced, since most of the reviews fell into the positive class as shown in Figure 3 . A detailed approach of handling imbalanced data can be found in Section 7.

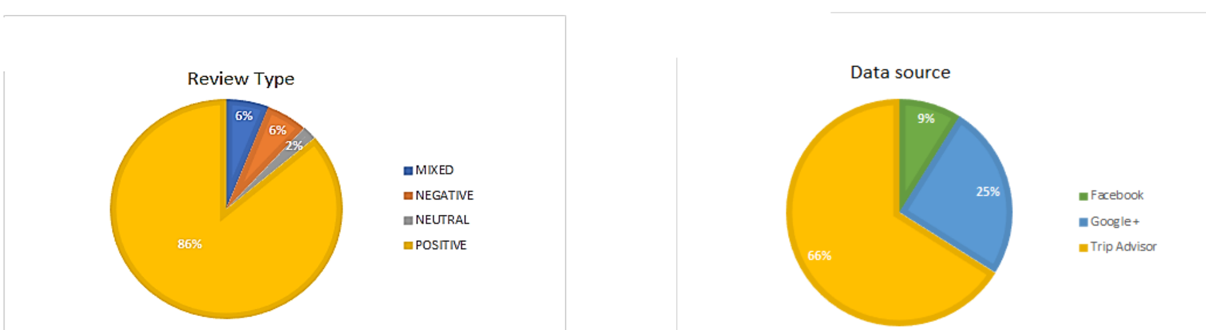

Figure 3: Proportion of Data Review Type and Data Source

Second step of the analysis was to plotting some charts to see if there were any trends in the data. EDA brought interesting insights and the behavioral aspects of how people spent time online and in temples. A summary of the some interesting trends is listed below

- According to ISKCON, months of June and July see the highest number of visitors because of festivals and foreign tourists. This trend was reflected in the number of reviews posted online (Figure 4).

- According to ISKCON they receive the maximum number of visitors on weekends. However, this trend was not reflected in the number of reviews online. While ISKCON saw crowds on weekends, most reviews were posted on Monday. One could reason that people do not review a place immediately after visit but take their time since they spend most their time outdoors on weekends and most of the times the TA e-mail review notifications come only after a couple of days after the visit.(Figure 4)

- In continuation with this finding, 43% of all negative reviews posted on Monday had the keywords “crowded”, “long queue” while only 18% of all positive reviews had the keyword “queue management”. This clearly reiterates the point that people had visited the temple during weekends.

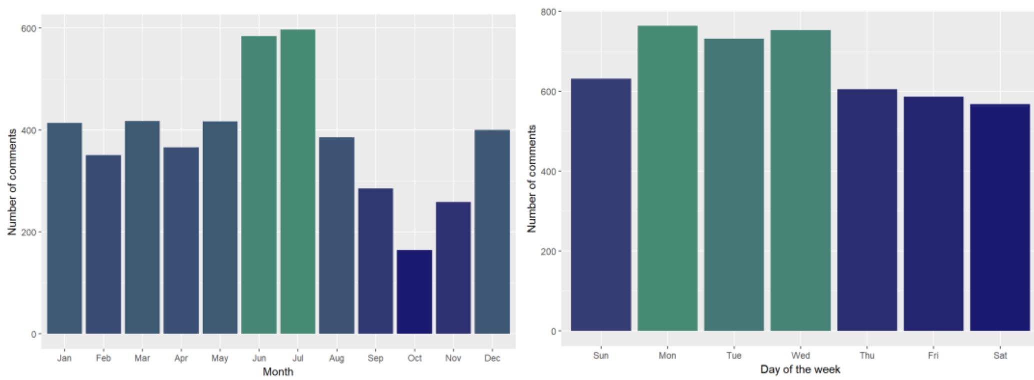

Figure 4: Trend Analysis - Month and Days of the week

- The highest number of negative comments was received between Jun and Jul 2017.On further inquiry, it was found that ISKCON had included many souvenir shops on the only exit pathway of the temple. 90% of all negative comments in this period has the keyword “commercial”.

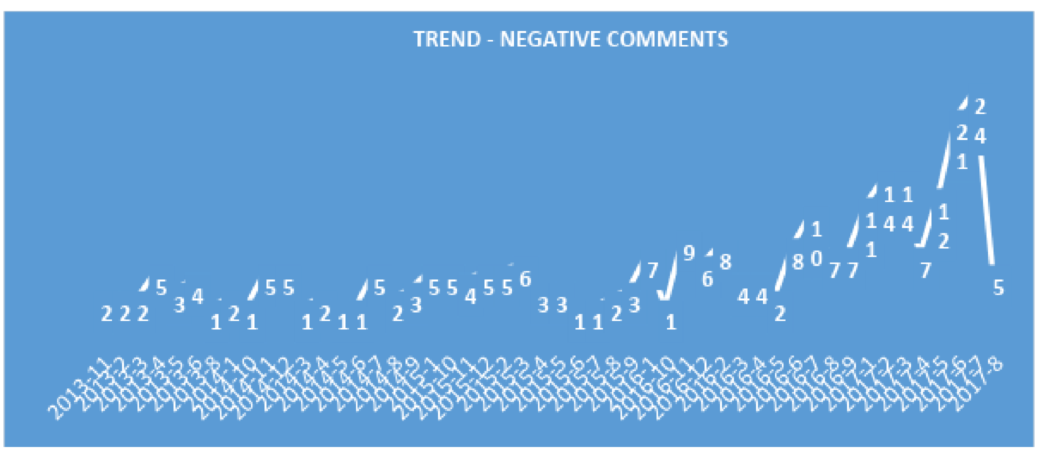

Figure 5: Month wise Trend Analysis for Negative Comments

6. Feature Engineering and Selection

6.1 Feature Engineering

Feature engineering (FE) is the process of transforming the raw data into features that act as inputs for the ML model. Even basic algorithms can yield good accuracies when feature engineering is done in the right way. In the context of text classification words would act as features for the model. Feature engineering was done by performing n-gram tokenization. For purposes of improving accuracy, a combination unigram (UG), bigram (BG) and trigram (TG) tokens were used. Feature vectors were built for each of them and the model was tested with each combination. As supervised ML approach was used or sentiment analysis, the features (UG, BG, TG) tokens were built only from the review comments. For emotion analysis, FE was done by using a lexicon-based approach. A dictionary of seed words for each emotion was built and used for further analysis.

6.2 Feature Selection

Feature selection (FS) is the process of selecting a subset of features from the previous step. FS is important since it aids in improving the classifier accuracy by remove the noise features. In addition, FS reduces the size of the feature dictionary fed to classifiers, thereby reducing the run-time of the classifiers. FS was done by using three different approaches: Chi-Squared metric (CS), Term Frequency-Inverse Document Frequency (TF-IDF) and Bag-Of-Words (BOW). BOW gave the least accuracy, while CS and TF-IDF gave close results. However, BG - CS gave the highest accuracy across all approaches. A notable inference was that while the accuracy of CS dipped after BG, the accuracy of TF-IDF kept improving from UG to TG. Result of feature selection and feature engineering process is furnished in Figure 6.

Feature engineering	Feature selection	Model Accuracy
UG	BOW	0.64
UG	TF-IDF	0.67
UG	CS	0.7
BG	BOW	0.68
BG	TF-IDF	0.78
BG	CS	0.82
TG	BOW	0.72
TG	TF-IDF	0.8
TG	CS	0.79

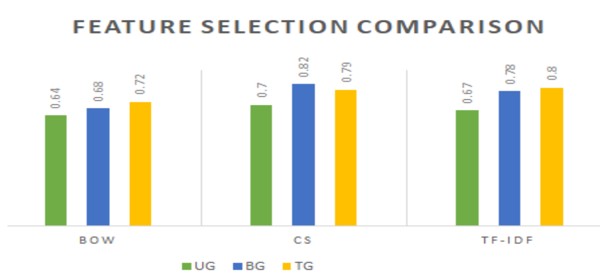

Figure 6: Results of Feature Engineering and Feature Selection

7. Building ML model for Sentiment Analysis with Imbalanced data

For building ML models, Naïve Bayes (NB) and Support Vector Machine (SVM) classifiers were used. The data was split into 80% training data and 20% test data. Upon careful analysis it was found that the dataset was imbalanced i.e. the 4 classification categories (positive, negative, neutral and mixed were not approximately equally represented. The number of positive comments outnumbered the other three classes - negative, neutral and mixed (Figure. 7).Classifiers give poor misleading results on such highly imbalanced data. Hence, various sampling strategies were tested to balance the data before feeding it to the classifier. Summary of the sampling techniques used are given below

- Random up sampling(RUS) - Randomly select samples from each of the minority classes i.e negative, neutral and mixed comments
- Random down sampling(RDS) - Randomly select samples from the majority class i.e. positive comments
- Synthetic Minority Oversampling Technique(SMOTE) - Minority classes are up sampled by creating ``synthetic'' examples rather than using random up sampling with replacement
- Clustering based downsampling (CDS) – Cluster the majority class using KNN, merge the comments (down sample), and balance the outputs with the other minority classes.

The balanced dataset from sampling was then fed into two classifier algorithms: NB and SVM. A comparison of accuracies of the various models for each classes is given in Figure 7. It can be seen that NB and SVM gave similar results with Random downsampling. All methods gave consistently good accuracies for the positive class whereas many deviations were observed for the other classes. Neutral classes produced low accuracies for almost all methods except SMOTE and clustered downsampling. Comparing all methods and their run times, SMOTE-NB gave highest accuracy and low run-time.

NB				SVM			
Sampling Technique	**Classes**	**Sensitivity**	**Specificity**	**Sampling Technique**	**Classes**	**Sensitivity**	**Specificity**
Random Upsampling	Positive	0.97	0.91	Random Upsampling	Positive	0.95	0.97
Random Upsampling	Negative	0.6	0.26	Random Upsampling	Negative	0.35	0.21
Random Upsampling	Neutral	0.5	0.12	Random Upsampling	Neutral	0.45	0.09
Random Upsampling	Mixed	0.48	0.33	Random Upsampling	Mixed	0.26	0.37
Random Downsampling	Positive	0.94	0.89	Random Downsampling	Positive	0.95	0.92
Random Downsampling	Negative	0.44	0.45	Random Downsampling	Negative	0.45	0.4
Random Downsampling	Neutral	0.35	0.35	Random Downsampling	Neutral	0.35	0.57
Random Downsampling	Mixed	0.67	0.67	Random Downsampling	Mixed	0.65	0.73
SMOTE	Positive	0.99	0.95	SMOTE	Positive	0.95	0.96
SMOTE	Negative	0.71	0.88	SMOTE	Negative	0.75	0.66
SMOTE	Neutral	0.84	0.75	SMOTE	Neutral	0.85	0.71
SMOTE	Mixed	0.79	0.83	SMOTE	Mixed	0.75	0.63
Clustered Downsampling	Positive	0.96	0.89	Clustered Downsampling	Positive	0.95	0.9
Clustered Downsampling	Negative	0.68	0.76	Clustered Downsampling	Negative	0.65	0.49
Clustered Downsampling	Neutral	0.66	0.65	Clustered Downsampling	Neutral	0.65	0.34
Clustered Downsampling	Mixed	0.85	0.75	Clustered Downsampling	Mixed	0.85	0.18

Figure 7: Comparison of Accuracies

8. Emotion Analysis (EA) using lexicon based approach

Emotions run high when people visit temples. In this paper, we have attempted to perform emotion analysis by using a Dictionary based approach. The objective of this analysis was to mine eight emotions (anger, fear, anticipation, trust, surprise, sadness, joy, and disgust) from the reviews. A dictionary mapping [8] was done for every word in a comment and a frequency score was calculated based on the presence/absence of the word. Each comment was then classified into one of the eight emotions based on the highest score. Emotion analysis (Figure. 8) revealed that visitors had positive emotions while at ISKCON. Negative emotions such as fear, sadness and disgust almost 3 times less than the positive emotions. Of all positive emotions, trust and joy were scored equally, followed by anticipation.

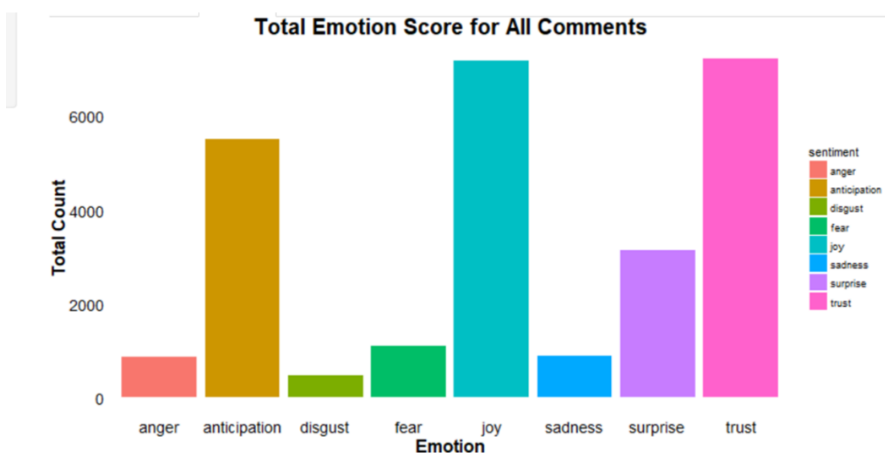

Figure 8: Emotion Analysis

9. Conclusion

The whole study can be categorized into three sections namely Trend Analysis, Sentiment-Emotion Analysis and Deployment. Trend Analysis helps us to unearth hidden patterns in the data which could otherwise go unnoticed. For example, we can understand when and how often people use social media to post reviews. One important insight we found was that people take time, usually 2-3 days after their visit, to post reviews online. Another insight was that visitors did not like the commercial shops on the exit pathway and a clear negative trend was seen once these shops were installed. As a result of this analysis, we derived a lot of other real-time insights that helped ISKCON base their decisions.

In the second section, we have proposed a Sentiment Analysis approach that can be used to identify the overall sentiment polarity and emotion of the people visiting ISKCON. Firstly, we have tried various methods for feature selection and found that Chi-square metric with bigram tokens gave the best results for the given dataset. Secondly, for identifying the overall sentiment polarity we have employed a supervised ML approach and handled the imbalanced data using various sampling techniques, which yielded us good accuracies. It has to be noted that Naive Bayes has given better results compared to SVM in many combinations. However, SVM performed better for certain minority classes. Out of all sampling-algorithm methods, SMOTE-NB combination gave the highest accuracy and this model can be improvised further when the data volume increases. Thirdly, we have also tried doing emotion analysis, since temple is a place where people come in with one mindset and leave with another. The results showed that the overall emotion was largely positive and negative emotions were minimal.

Finally, the deployment plan was to build an automated system that would reduce the manual efforts. To achieve this, we have provisioned and automated the whole process from data collection to deriving results by building an R Shiny app (refer appendix) that would directly give the sentiment and emotions as plots. These plots were made interactive in the sense that the top 10 topics for each class (positive, negative, neutral and mixed) would be listed based on user selections. When the user clicks on each topic all comments of that topic would be displayed. Going one level deeper, when the user clicks on each of these comments the app would list the demographics (date, name, location and social medium) of the person who has posted the comment.

10. References

1. Lane, Peter C.r., et al. “On Developing Robust Models for Favourability Analysis: Model Choice, Feature Sets and Imbalanced Data.” *Decision Support Systems*, vol. 53, no. 4, 2012, pp. 712–718., doi:10.1016/j.dss.2012.05.028.
2. Li, Shoushan, et al. “Imbalanced Sentiment Classification.” *Proceedings of the 20th ACM International Conference on Information and Knowledge Management - CIKM '11*, 2011, doi:10.1145/2063576.2063994.
3. Lin, Weitong, et al. ”A survey of sentiment analysis in social media.” *Knowledge and Information Systems*, 2018, doi:10.1007/s10115-018-1236-4.
4. Liu, Bing. “Opinion Mining and Sentiment Analysis.” *Web Data Mining*, 2011, pp. 459–526., doi:10.1007/978-3-642-19460-3_11.
5. Medhat, Walaa, et al. “Sentiment Analysis Algorithms and Applications: A Survey.” *Ain Shams Engineering Journal*, vol. 5, no. 4, 2014, pp. 1093–1113., doi:10.1016/j.asej.2014.04.011.

6. Mountassir, Asmaa, et al. "An Empirical Study to Address the Problem of Unbalanced Data Sets in Sentiment Classification." *2012 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, 2012, doi:10.1109/icsmc.2012.6378300.
7. Xia, Rui, et al. "Ensemble of Feature Sets and Classification Algorithms for Sentiment Classification." *Information Sciences*, vol. 181, no. 6, 2011, pp. 1138–1152., doi:10.1016/j.ins.2010.11.023.
8. Zhu, Xiaodan, et al. "NRC-Canada-2014: Recent Improvements in the Sentiment Analysis of Tweets." *Proceedings of the 8th International Workshop on Semantic Evaluation (SemEval 2014)*, 2014, doi:10.3115/v1/s14-2077.

Appendix

Deployed R Shiny App