© April 2020 | IJIRT | Volume 6 Issue 11 | ISSN: 2349-6002

Investigating Airline Reviews for Improved Customer Services using Artificial Intelligence

Srishti Kohli[1], Rahul Bajaj[2]
[1]*Indira Gandhi Delhi Technical University for Women, Delhi*
[2]*Vellore Institute of Technology, Chennai*

Abstract- **The aim of this paper is to deduce the customers' assessment of their experience of flight service using the customer reviews and ratings. Analysis of the feedback from customers is of prime importance for any industry. This analysis can help the industry to provide better services and customer satisfaction. We are leveraging the power of deep learning to mine relevant information from raw review data. We undertake sentiment analysis of airline reviews and clustering techniques to identify customer distribution. The research is done in order to understand customer behavior and their perception about the flight services with respect to various airlines.**

I.INTRODUCTION

To ensure the growth of an organisation, analysing its customers' feedback and experience is a major concern. Sentiment analysis is the task of analysing a piece of text to extract the emotions expressed in it and identifying whether it is positive, negative or neutral. In this research, we identify the sentiment or opinion a customer holds about his experience on a flight from his reviews. The automated task of sentiment analysis helps categorizing reviews into positive or negative, otherwise done manually. This enables faster bifurcation of reviews from customers for easier study of the same.

Moreover, we use unsupervised learning to group similar reviews to obtain flights with similar properties. This helps in finding closest alternatives to a flight.

II. DATA EXPLORATORY STUDY

Data used for this research belongs to the public domain and is curated from airlinequality.com. The dataset comprises 70,285 instances. The number of airlines reviewed is 443. For each airline the number of reviews ranges from a minimum of 1 to a maximum of 2355 with mean lying around 157. The publishing dates of reviews range between 06-01-2002 and 17-12-2017. Each review consists of the customer's perception of the airline on the following attributes- Seat comfort, Inflight entertainment, cabin staff service, food and beverages, value for money. All the mentioned attributes are rated between 0 and 10. The customers are also further divided into 4 classes: Business Class, First Class, Economy class, Premium Economy. We perform Clustering on the mentioned attributes for all the classes of customers.

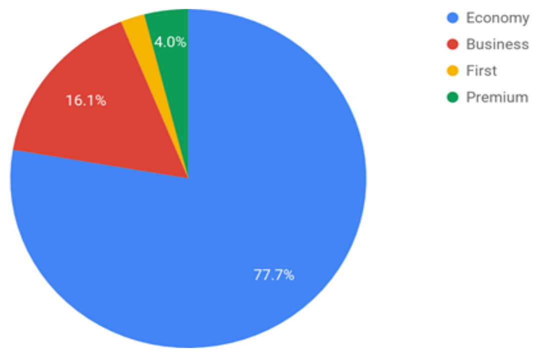

Fig 1. Customer Distribution

Furthermore, each review contains a textual description of the customer's experience on the flight. We perform sentiment analysis on this textual data using Natural Language Processing and various Deep Learning Techniques. Each review is labelled with positive or negative recommendations for the airline.

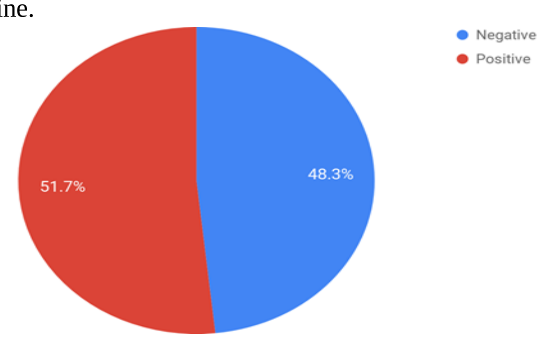

Fig 2. Review Distribution

III. RELATED WORK

Earlier studies have been performed analysing traveller satisfaction based only on statistical data collected from Skytrax. Our research makes a comprehensive study analysing text reviews eliciting customer behaviour from the statistical distribution and correlation of the ratings provided. Similar cluster based studies have been performed on a smaller and different dataset from Skytrax. Our research spans a dataset of about 15 years covering a wider span of reviews and airlines.

IV. METHODOLOGY

This research involves different machine learning, natural language processing and deep learning methodologies.

A. Clustering

In the process of identifying customer distribution we have used the K-means algorithm. K-Means works on the principle of dividing data points into K clusters where K is being determined using elbow method. In this process K was calculated as 6. The clustering is done for each cabin flown, Economy, Premium Economy and Business class. For the sake of convenience we have changed first class to business class for effective cluster formation. With these 6 clusters, we computed the parameter values to identify which values are more affecting the customer distribution. The clustering helps to relate a new case and find alternatives to similar choices with respect to the specified service fields in the data. This clustering will enable us to understand Customer behavior better.

B. Sentiment Analysis

For the task of sentiment analysis, the textual data is first preprocessed and each word is converted to its embedding. We use both pre trained and custom trained embedding. This is followed by feeding the processed data into various neural networks for classification.

a. Pre Processing Pipeline

The raw textual review data is preprocessed using various Natural Language processing techniques. Firstly, the data is tokenized. This is followed by removal of stop words and removal of names of geographical locations using Named Entity Recognition. Lastly, lemmatization of all words is done.

b. Word Embeddings

We convert all the words to their respective embeddings before feeding the data into the network. In this research, we follow three approaches for the same.

In the first approach, we use Pre-trained GloVe vectors which provide 50 dimensional embeddings for 6B tokens.

The second approach uses the Word2Vec model for building word embeddings. This model is a two-layer deep neural network that is trained to reconstruct linguistic contexts of words. The words with common contexts are positioned close to each other in the vector space. We construct the Word2Vec Model on 80% of the data in our corpus to produce 50 dimensional embeddings for each word.

In the third approach, we use the Keras Embedding layer, to produce 100 dimensional embeddings for each word.

c. Model Used

The model takes input of size 200 X 50 (200 words per review and 50 dimensional embedding per word). The reviews longer than 200 are truncated and those shorter are padded to obtain a constant size of 200 words per review.

We use the following networks in our model:

Convolutional Neural Network: Convolutional Neural Network comprises a network of nodes/neurons, to learn the weights and biases, employing convolution as the linear operation. It can be understood as replacing matrix multiplication in a neural network with convolution in at least one of the layers. Here we use 1D Convolution layer to learn features from the text to perform Sentiment Analysis.

Recurrent Neural Network: Recurrent Neural Networks enable sequential processing of data with respect to time. Output generation takes the state of the network into account, such that it depends on the previous computations. In addition to weights in a traditional deep neural network, it contains a state component to take previous results into consideration. We also employed LSTM to overcome the vanishing gradient problem.

The architectures used are discussed below:

Layer	Output shape	Parameters
Convolution1D	200 X 64	9664
Convolution1D	200 X 128	24704
Convolution1D	200 X 256	98560
Convolution1D	200 X 512	393728
LSTM	200 X 256	787456
LSTM	256	525312
Dense	256	65792
Dense	128	32896
Dense	64	8256
Dense	32	2080
Dense	1	33

Table: Layers in Architecture 1

Layer	Output shape	Parameters
Keras Embedding	100	5160200
Simple RNN	64	10560
Simple RNN	64	8256
Simple RNN	64	8256
Dense	1	65

Table: Layers in Architecture 2

Layer	Output shape	Parameters
Keras Embedding	100	5160200
Simple RNN	64	10560
Dense	1	65

Table: Layers in Architecture 3

V. RESULTS

A. Clustering

Cluster Distribution Values for each parameter:

1. Value for Money
2. Seat Comfort
3. Cabin Staff Service
4. Food and Beverages
5. Inflight Entertainment

	X0	X1	X2	X3	X4	X5
Sub-class		I	II	III	IV	V
Business	-0.2417	0.9623	0.1116	0.3807	0.4075	-0.0131
Economy	-0.2944	0.8756	0.1080	0.6068	0.823	0.0962
Premium Economy	-0.2531	1.0740	-0.4098	0.7002	0.2345	0.2811

Table: Cluster distribution

B. Sentiment Analysis

The results are obtained with 70%, 10% and 20% train, validation and test split.

	Model	Testing Accuracy (%)
A	Pre trained Glove Vector and Architecture 1	87.01
B	Custom Word2Vec and Architecture 1	88.18
C	Architecture 2	87.38
D	Architecture 3	84.64

Table - Sentiment Analysis results

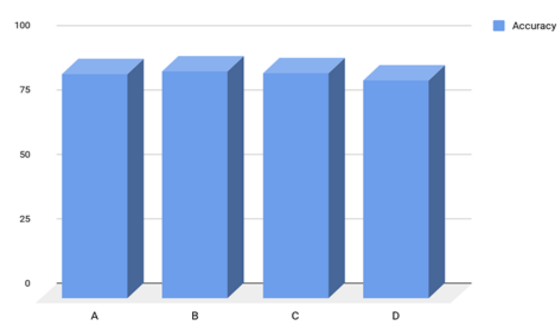

Fig : Model and Accuracy

VI. CONCLUSION

The results for sentiment analysis were comparable for various approaches but the best results were obtained from Word2Vec.
Investigation on airline reviews, consisting of Sentiment Analysis and Clustering is successful.

REFERENCES

[1] Skytrax- airlinequality.com
[2] I. Yakut, T. Turkoglu, and F. Yakut. Understanding customers' evaluations through mining airline reviews. International Journal of Data Mining & Knowledge Management Process, 2015
[3] chollet 2015 keras, title=Keras, author=Chollet, Fraçois and others, year=2015
[4] en.wikipedia.org/wiki/Word2vec
[5] Jeffrey Pennington, Richard Socher, and Christopher D. Manning. 2014. GloVe: Global Vectors for Word Representation
[6] arXiv:1301.3781