

**ANL312**

**Text Mining and Applied Project Formulation**

**End-of-Course Assessment**

**Jul 2024**

|  |  |
| --- | --- |
| Name | B HARIPPRIYA |
| PI Number | K2170295 |
| T-Group | T03 |
| Date Submitted | 01/11/2024 |

Table of Contents

[Chapter 1: Introduction 2](#_Toc181228160)

[Chapter 2: Literature Review 4](#_Toc181228161)

[Chapter 3: CRISP-DM Process 6](#_Toc181228162)

[Chapter 4: Summary 24](#_Toc181228163)

[References 25](#_Toc181228164)

# **Chapter 1: Introduction**

Project Background

The topic that I have chosen is applications of text mining on customer satisfaction of a specific company. The reason for choosing this topic is based on the famous quote by Harry Gordon Selfridge which is “The customer is always right” (OSMOND, 2023). The importance of customer satisfaction is highlighted in this single sentence as it emphasizes on the business’s need to listen to customers and improve on its shortcomings to reduce customer churn and increase customer’s loyalty and engagement to the business (Indeed Editorial Team, 2024). If businesses do not pay importance to customer satisfaction, customers would take their business elsewhere and its reputation would be damaged through the customer’s negative word of mouth leading to loss of revenue (Donner , 2024).

Customer satisfaction is a measurement that helps to determine how well a business is performing and it can be measured through customer feedback forms/surveys, reviews on Google and social media (Szyndlar, 2024). To analyze large amounts of customer feedback manually is impossible. To help us, text mining is engaged to identify trends and pinpoint common themes being highlighted in reviews which in turn allows businesses to make informed business decisions to enhance overall customer satisfaction (Linkedin Community, 2023).

For this topic, the company that I have decided to perform text mining on is Air India. Air India is an airline which is ranked at 90 at the world’s top 100 airlines in 2024 announced in the world airline awards (Skytrax, 2024). This airline generally receives criticism regarding its punctuality, engineering and maintenance issues (Sinha, 2024). Customers vent their frustration and satisfaction across different platforms like Google reviews, TripAdvisor, social media, etc.

Business analytics concepts involved

The key components of business analytics are Business, Data and Technology. The business component refers to the ability to come up with a problem which can be solved using business analytics. The Data component refers to the ability to collect relevant data that will be useful in solving the business problem at hand. The Technology component refers to the ability to identify suitable techniques that can be used to solve the business problem.

The business analytics concepts involved are data management and descriptive analysis. Data management is to collect, clean and store data for analysis (J., 2023). In text mining, text parsing is done by extracting, cleaning and creating a collection of words using NLP and the steps in text parsing are tokenization, stemming, spelling normalization, stop words, synonyms, etc. (Linkedin Community, 2023). Descriptive analysis is analyzing past data to understand the issue at hand and why it happened (J., 2023). For this project, Air India customer reviews from previous years are extracted to understand the factors that contribute to customer dissatisfaction and to work on those areas so that high customer satisfaction can be achieved.

Project objectives and role of text mining

The project objective is to identify the key factors that impact customer satisfaction towards Air India and compare it with the common issues identified above to see if the same complaints are arising or are there any new criticism and use these insights gathered to come up with data-driven decisions that will improve services, improving customer experience, leading to which leads to high customer satisfaction and churn.

In order to satisfy the project objectives text mining can be performed. Text mining is identified as the process of transforming unstructured data to structured data to identify patterns and insights and it uses Natural Language processing (NLP) which allows computers to understand the human way of text. Text mining consists of different techniques that can be performed to slowly uncover the meanings hidden in unstructured data (Zhang, 2024).

For example, topic modelling can be used to group similar words together to identify different point of topics that is widely brought up by customers. Sentiment analysis can also be used to identify the sentiments of the text either positive or negative (Zhang, 2024).

# **Chapter 2: Literature Review**

There are 2 articles that I have found where text mining has been applied on customer satisfaction that I have found relevant of respective companies. The 2 companies in focus are Airbnb and Honda.

The first article was a study conducted by Santos, M., Ribeiro, R., Batista, F. & Correia, A. where they extracted reviews from Airbnb properties in Lisbon, Portugal from 2010 to 2019 to find out what are the key factors that lead to customer satisfaction. The dataset consisted of 590,070 reviews and the researchers used topic modelling and sentiment analysis in their text mining process. Topic modelling was used to identify the main topics that were discussed among customers. Latent Dirichlet Allocation (LDA) was used to extract out main topics by creating a list of topics based on the probability of the word appearing in the document. Sentiment analysis was also used to gauge the emotional tone of the reviews. VADER was used to detect sentiments in reviews and the higher the rating of a word, the more positive the word is (Santos et al. (2010)).

The results of the study showed that the main topics extracted from topic modelling were location, physical aspects of the property and the service of the host. Out of these topics, the mentions of service of the host decreased and mentions of physical aspects of the property increased. This could mean that the quality of service provided by the host may have declined and the property was maintained well which could have lead to the dissatisfaction of the customers. In terms of sentiment analysis, it was revealed that about 98% of the reviews were positive but this number had also decreased over the years as Airbnb reduced human contact from its services. Thus, human interaction is a major factor in fostering customer satisfaction, according to the researchers' conclusions (Santos et al. (2010)).

The second article was a study conducted by Princess Eunike where the researcher extracted reviews from Google reviews from 3 car dealers in Indonesia with regards to customer satisfaction throughout the journey of buying Honda products as well as buying Honda products. The total size of the dataset consisted of 3214 reviews across 3 dealers. This researcher also uses topic modelling and sentiment analysis to extract main topics discussed from customer reviews. For topic modelling, Non-Negative Matrix Factorization (NMF) was used where 2 matrixes, W and H, were formed. Matrix W displays an approximate value for each word in a topic and Matrix H displays an approximate value of the importance of the topic in each document. For sentiment analysis, a decision tree and SenticNet5 was used to gauge the emotional tone of the reviews and a score will be given between -1 to 1 where values closer to -1 refers to negative and +1 refers to positive (Princess Eunike, 2024).

The results of the study showed that the main topics that were extracted from topic modelling across the 3 car dealers were service experience, buyer experience, waiting room, waiting time, etc. For sentiment analysis, in each topic the average importance and performance level scores were listed for each topic identified in topic modelling. It can be inferred that services has a high emotional tone score compared to service experience, online order and waiting time. The importance performance analysis matrix was also used to prioritize topics for targeted action. In conclusion, the targeted action for specific dealers were listed which are waiting room faciltity improvement, buying experience and services (Princess Eunike, 2024).

The implications of the references discussed is useful for the current topic. This is because the approach to text mining used is applicable for my analysis. The references used topic modelling and sentiment analysis to anlayse the factors contributing to customer satisfaction. By using this, Air India will be able to identify recurring themes highlighted among customers and identify the overall emotional tone of customers. Based on the results, it can focus on the improvements that can be made to enhance customer satisfaction and improve the travel experience of guests.

# **Chapter 3: CRISP-DM Process**

CRISP-DM process allows analysts and data mining professionals to follow a standard process to help generate a solution to the given business problem (Hotz, 2024). There are a total of 6 phases in the CRISP-DM process. They are business understanding, data understanding, data preparation, modelling, evaluation and deployment. The CRISP-DM process is used for this project to find a solution for Air India.

Phase 1: Business Understanding

This phase is to develop a clear understanding of the business needs and define business objectives to improve business performance. For this project, the business objective is to identify the key factors that drives customer satisfaction or dissatisfaction towards Air India services by creating a text mining model and using text categorization using SPSS Modeler.

Phase 2: Data Understanding

This phase is to collect the relevant data that is needed to solve the business problem. Since the objective is to identify the key factors that contribute towards customer’s sentiments regarding Air India, it is ideal to collect data where customers convey their feelings regarding their experience flying with Air India. I have found a dataset on Kaggle where the dataset consists of customer reviews from 2018 to 2024 that were extracted from TripAdvisor and it talks about the personal experiences of customers regarding Air India airlines. The dataset consists of 6 columns which are the date and time of the review being published, the rating given from 1 to 5, the type which is review, helpful\_votes which is the number indicating the helpfulness of the review to others, the title of the review and the review itself (Karunarathna, 2024).

The text column is the important column of the dataset for our analysis as the sentiments of the customers are revealed in that column. The quality of the dataset is at an acceptable range where there are no missing values nor duplicate values. Therefore, this dataset is suitable for analysis.

Phase 3: Data Preparation

This phase is to transform the original dirty data into clean, formatted data which will be used for analysis later on. The purpose of data cleaning is to ensure that the data is accurate and would not corrupt the results after analysis which would lead to incorrect business decisions made that would further damage the company. The process of preparing the data is explained from the next paragraph onwards.

The first order of business is to ensure that the measurements of the fields are read correctly. An excel node is created to perform this action. In *Figure 1* below, the measurements of different fields are displayed after the values were read. In this situation, the published\_date, type, title and text fields are categorized as typeless. This is accurate as these fields do not conform to the other measurements such as nominal, ordinal, etc. The field helpful\_votes is continuous and this is accurate because they are numbers indicating the upvotes made by other customers to indicate how helpful the comment is. The field rating is read as continuous but it is actually supposed to be ordinal as ratings can be ranked and they have a meaningful order. Therefore, in *Figure 2*, I have changed the measurement to ordinal and read the value again.

A screenshot of a computer

Description automatically generated

*Figure 1: Roles of fields in csv file*

A screenshot of a computer

Description automatically generated

*Figure 2: Change rating field measurement to ordinal*

The second order of business is to check if there are any missing values, null values or outliers. A data audit node is attached to the var. file node to perform this action. In *Figure 3* below, since the only 2 fields that aren’t typeless are rating and helpful\_votes, they are the only fields that will be displayed in the data audit. Looking at the statistics below, it can be seen that there are no missing or duplicate values. However, there are outliers in the helpful\_votes field but I would not be performing any data cleaning as this field is not going to be used for the analysis later on as only the text field is needed. A table node is then attached to visualize the data to see if it is formatted correctly. The initial count of records are at 4840 as shown in *Figure 4*.

A screenshot of a computer

Description automatically generated

*Figure 3: Data audit of csv file*

A screenshot of a computer

Description automatically generated

*Figure 4: Number of records before data cleaning*

The third order of business is to clean the data to prepare for analysis. A few data quality issues have been identified and they need to be taken care of.

The first data quality issue is that there are empty values in the text column for 3 records. This means that only a summary of the review was given in the title with no elaboration on the review given. These records were extracted by adding a select node and using the formula “length(trim(text)) > 0” which discards any records that does not have a length which is more than 0 as shown in *Figure 5*. The reason for choosing discard option is to extract the records to visualize the number of records. To clean the data, Include option will be selected to include the records that have a length more than 0 and discard the rest. The total number of records after discarding comes down to 4837 as shown in *Figure 6*.

A screenshot of a computer

Description automatically generated

*Figure 5: select node formula to view empty text records*

A screenshot of a computer

Description automatically generated

*Figure 6: Number of records after removing empty text records*

The second data quality issue as seen in *Figure 6* is that in the text column there are words that are uppercase and some that are in lowercase. The appropriate treatment is to change all the words in the text column from upper to lower as only lowercase words are accepted in text mining. To perform this treatment, a filler node is added to the select node and the condition is to check if the text column contains strings by using “is\_string(text)” and use “uppertolower(text)” command to change all uppercase to lowercase words as shown in *Figure 7*. *Figure 8* shows the table node after the words have been changed to lowercase in the text column.

A screenshot of a computer

Description automatically generated

*Figure 7: Filler node settings with formula to change words from upper to lower case*

A screenshot of a computer

Description automatically generated

*Figure 8: records after words are changed to lowercase*

The last data quality issue is to identify and remove duplicate records if any. To perform this action, distinct node is selected. The field needed for grouping is the text field as only this field is going to be used for text mining and no duplicates should be in the field as shown in *Figure 9*. *Figure 10* shows that the count of records has dropped by 1 to 4836 which indicates that a single duplicate record was removed.

This concludes the data preparation phase and the next phase which is modelling can be performed.

A screenshot of a computer

Description automatically generated

*Figure 9: Distinct node to remove duplicates*

A screenshot of a computer

Description automatically generated

*Figure 10: Records after removing duplicates*

Phase 4: Modelling

This phase is when a model is created to help analyse and identify areas of concern so that appropriate data driven business decisions can be made. For this project, a text mining model will be created. The reason why text mining model is suitable for this project is because it helps analyse and transform unstructured data which is text reviews into structured data which is standardized into a table format. This will help in identifying any trends or relationships in the reviews.

In order to perform text mining, a text mining node needs to be added into the stream. Once the node is added, under text field, “text” should be selected as this is the field that contains all of the detailed reviews which needs to be dissected to find any insights as shown in *Figure 11*.

A screenshot of a computer

Description automatically generated

*Figure 11: Text mining node to choose the column*

Next, a suitable resource template needs to be chosen. In this situation, since the reviews from customers who flew Air India are regarding customer satisfaction, the appropriate resource template is Customer Satisfaction Opinions (English) as seen in *Figure 12*.

A screenshot of a computer

Description automatically generated

*Figure 12: Text mining node to choose the resource template*

Lastly, the option “Accommodate spelling for a minimum root character limit of 5” should be selected as shown in *Figure 13*. This will make sure that all of the words that are commonly spelled wrongly or grouped together. Now, the text mining node is ready to run.

A screenshot of a computer

Description automatically generated

*Figure 13: Text mining node to accommodate spelling errors*

Once the text mining node starts to run, it will first extract out all of the concepts and assign some of them into types while some concepts are left unknown. In this case, 23,076 concepts were extracted as shown in *Figure 14* and about roughly 13,000 concepts were unknown as shown in *Figure 15*. There are 61 different types including “Unknown” for the concepts as shown in Figure 16. The next step is to create some new types so that some of the concepts can be converted from “Unknown”.

A screenshot of a computer

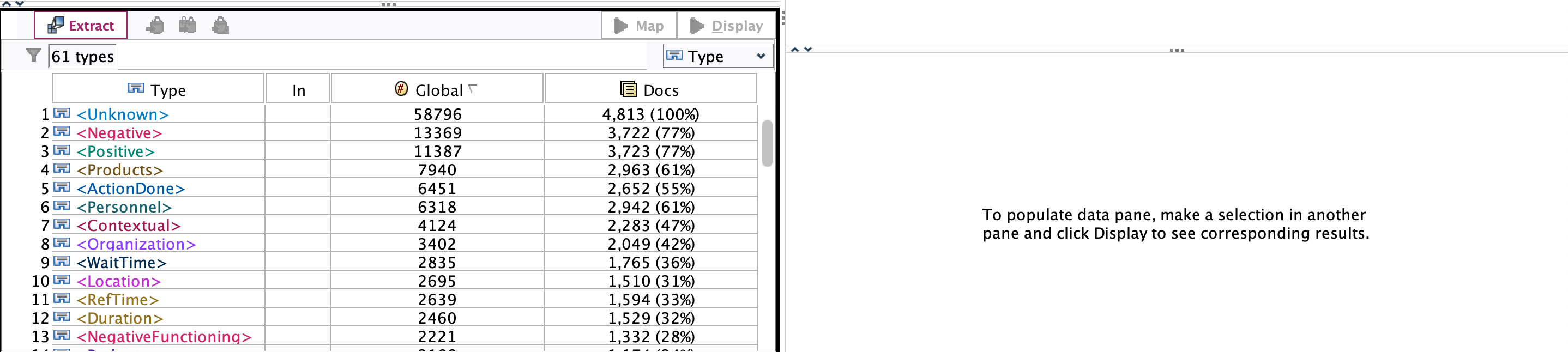
Description automatically generated

*Figure 14: Original result after concepts are extracted*

A screenshot of a computer

Description automatically generated

*Figure 15: The number of concepts that have “Unknown” types*



*Figure 16: The number of types extracted from the dataset*

In order to create categories, 2 activities needs to be performed which are to retype concepts and to assign synonyms to common targets that are misspelled. There might be instances where customers write reviews in a hurry or if customers get angry they will usually write reviews in frustration which cause them to misspell words and leads to loss of meaning of the review. Therefore, with these 2 activities being done, it helps with easy analysis when creating categories.

The first activity is to assign unknown concepts with a new type. Since the number of unknown concepts is too large, I created a sample size of 54 concepts to be retyped. *Table 1* below shows the 54 concepts to be retyped, the default type which is unknown and the appropriate type that the concept was categorized under. All concepts are retyped and the new types are stored under Local library.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **S/N** | **Concept to be retyped** | **Default Type** | **More Appropriate Type** | **Library** |
| **1** | flight time wait | Unknown | FlightExperience | Local Library |
| **2** | flight handling | Unknown | FlightExperience | Local Library |
| **3** | Flight delays | Unknown | FlightExperience | Local Library |
| **4** | Seat stank | Unknown | PlaneHygiene | Local Library |
| **5** | Feces | Unknown | PlaneHygiene | Local Library |
| **6** | Stench | Unknown | PlaneHygiene | Local Library |
| **7** | Sanitization | Unknown | PlaneHygiene | Local Library |
| **8** | State of the plane | Unknown | PlaneHygiene | Local Library |
| **9** | Unhygienic flights | Unknown | PlaneHygiene | Local Library |
| **10** | blood | Unknown | PlaneHygiene | Local Library |
| **11** | Cabin conditions | Unknown | PlaneComfort | Local Library |
| **12** | Charging port | Unknown | PlaneComfort | Local Library |
| **13** | Bed | Unknown | PlaneComfort | Local Library |
| **14** | Seat jutted | Unknown | PlaneComfort | Local Library |
| **15** | Extra leg room front | Unknown | PlaneComfort | Local Library |
| **16** | Watch movies | Unknown | PlaneComfort | Local Library |
| **17** | Head phone slot | Unknown | PlaneComfort | Local library |
| **18** | Headrest | Unknown | PlaneComfort | Local library |
| **19** | Seat handrests | Unknown | PlaneComfort | Local Library |
| **20** | Leg room | Unknown | PlaneComfort | Local Library |
| **21** | Chicago to delhi to Bangalore | Unknown | FlightPath | Local Library |
| **22** | Canada to india | Unknown | FlightPath | Local Library |
| **23** | Jodhpur - Mumbai | Unknown | FlightPath | Local Library |
| **24** | Jfk-ndl flight | Unknown | FlightPath | Local Library |
| **25** | London to india | Unknown | FlightPath | Local Library |
| **26** | Hyderabad-mumbai-shirdi flight | Unknown | FlightPath | Local Library |
| **27** | Anger | Unknown | Emotions | Local Library |
| **28** | Plight | Unknown | Emotions | Local Library |
| **29** | soulles | Unknown | Emotions | Local Library |
| **30** | Vomiting | Unknown | PlaneFood | Local Library |
| **32** | Biryani | Unknown | PlaneFood | Local Library |
| **33** | Croissant | Unknown | PlaneFood | Local Library |
| **34** | Inflight cuisine | Unknown | PlaneFood | Local Library |
| **35** | Scanty breakfast | Unknown | PlaneFood | Local Library |
| **36** | Inflight breakfast | Unknown | PlaneFood | Local Library |
| **37** | Non-vegetarian | Unknown | PlaneFood | Local Library |
| **38** | Expired meals | Unknown | PlaneFood | Local Library |
| **39** | Mashed potato affair | Unknown | PlaneFood | Local Library |
| **40** | Gulab jamun | Unknown | PlaneFood | Local Library |
| **41** | Maiyonnaise | Unknown | PlaneFood | Local Library |
| **42** | Uththapam | Unknown | PlaneFood | Local Library |
| **43** | Tin | Unknown | PlaneFood | Local Library |
| **44** | Roti | Unknown | PlaneFood | Local Library |
| **45** | Salt | Unknown | PlaneFood | Local Library |
| **46** | Paratha | Unknown | PlaneFood | Local Library |
| **47** | Indian wine | Unknown | PlaneFood | Local Library |
| **48** | Snack meal | Unknown | PlaneFood | Local Library |
| **49** | Manners | Unknown | AirIndiaCrewService | Local Library |
| **50** | Warmth | Unknown | AirIndiaCrewService | Local Library |
| **51** | Helpful | Unknown | AirIndiaCrewService | Local Library |
| **52** | Mixed attitudes | Unknown | AirIndiaCrewService | Local Library |
| **53** | Shout | Unknown | AirIndiaCrewService | Local Library |
| **54** | Air hostess courtesy | Unknown | AirIndiaCrewService | Local Library |

*Table 1: Sample concepts to have types retyped from “Unknown”*

The second activity is to create targets for common group of words or misspelled words that have the same meaning. Since the concepts size is too large, I have created a sample size of 10 targets with synonyms added to the local library as shown in *Table 2* below.

|  |  |  |  |
| --- | --- | --- | --- |
| **S/N** | **Targets** | **Synonyms** | **Library** |
| **1** | Cabin crew | Flight steward, stewardess, air hostess, male air host | Local Library |
| **2** | government | Govt, govt. , gov | Local Library |
| **3** | Air india | ai, air indis, aair india | Local Library |
| **4** | flew | flue | Local Library |
| **5** | Ticket | tkt | Local Library |
| **6** | strongly | strngly | Local Library |
| **7** | flight | Fhight, flih | Local Library |
| **8** | Common sense | Coman senesce | Local Library |
| **9** | about | bout | Local Library |
| **10** | continuosly | contnously | Local Library |

*Table 2: Synonyms created for a target word*

Once these activities are completed, the categories can be created. *Figure 17* below shows the final text mining model with the categories and respective subcategories. The four main categories that I have created which I found relevant to the business problem is Food Quality, Aircraft maintenance, Crew Service and overall customer satisfaction.

A screenshot of a computer

Description automatically generated

*Figure 17: Final text mining model with categories and sub-categories*

Phase 5: Evaluation

In this phase, the text mining model results and the readiness of the text mining model for deployment are to be evaluated. The categories here have more negative elements than positive. Although the doc counts are very low for positive sub-categories, the top 2 positive sub-categories are Positive in Service of Crew category with 8 doc counts and Positive in Maintenance of Aircraft category with 7 doc counts.

The top 2 negative sub-categories are Negative in Maintenance of Aircraft category with 35 doc counts and Negative in Overall customer satisfaction with 31 doc counts. *Figures 18 to 21* shows the sample docs of the top 2 positive and negative sub-categories.

**Positive sub-categories**

A screenshot of a computer

Description automatically generated

*Figure 18: sample docs of Positive sub-category of Service of Crew category*

A screenshot of a computer

Description automatically generated

*Figure 19: sample docs of Positive sub-category of Maintenance of Aircraft category*

**Negative sub-categories**

A screenshot of a computer

Description automatically generated

*Figure 20: Negative sub-category of Maintenance of Aircraft category*

A screenshot of a computer

Description automatically generated

*Figure 21: Negative sub-category of Overall Customer Satisfaction category*

Phase 6: Deployment

I feel that the model is ready for deployment based on the categories. It has effectively highlighted the words that portrays the sentiments of customers and are categorized correctly. Therefore, it will be useful for senior management to have a high overview on the important aspects that affect customer satisfaction scores. This model can be deployed for public use and more categories can be created depending on the business problem.

# **Chapter 4: Summary**

The project objective that was addressed in Phase 1 of the CRISP-DM process was to identify the key factors that were contributing to the low satisfaction among customers who fly with Air India. Based on the text mining model, it can be seen that Aircraft maintenance, Service of crew as well as Food Quality plays an important part in customer satisfaction. In terms of aircraft maintenance, the main highlights were regarding faulty armrests and dirty toilets. In terms of Food quality, the main highlights were that the food was a health hazard, undercooked and led to diarrhea. In terms of Service of the crew, the main highlights were that they did not take any accountability, no helping tendency, expressionless, etc.

In conclusion, the senior management of Air India are able to witness these conditions clearly from the text mining model which were categorized from the thousands of customer reviews on the TripAdvisor website and have a clear idea on what actions needs to be done in order to regain trust and confidence of customers to fly with Air India again.

**Word Count: 3730 words (excluding cover page, table of content, references, words under figures and tables, in-text citations)**

# **References**

 Donner , M. E. (2024, March 27). *The decline of customer service: Why companies are failing to delight and what needs to change*. LinkedIn. <https://www.linkedin.com/pulse/decline-customer-service-why-companies-failing-what-vjojf#:~:text=Loss%20of%20Customer%20Loyalty%3A%20When,after%20just%20one%20bad%20experience>.

Hotz, N. (2024, April 28). *What is CRISP DM?*. Data Science Process Alliance. <https://www.datascience-pm.com/crisp-dm-2/>

Indeed Editorial Team. (2024, August 17). Why is customer satisfaction important? 7 common benefits | Indeed.com UK. <https://uk.indeed.com/career-advice/career-development/why-customer-satisfaction-important>

J., P. K. (2023, March 6). *What is business analytics? definition, importance & examples*. LinkedIn. <https://www.linkedin.com/pulse/what-business-analytics-definition-importance-examples-jha>

Karunarathna, K. (2024, March 29). *Air India Tripadvisor reviews*. Kaggle. <https://www.kaggle.com/datasets/kanchana1990/air-india-tripadvisor-reviews>

Linkedin Community. (2023, July 9). *How do you analyze customer feedback with text mining?*. How to Analyze Customer Feedback with Text Mining. <https://www.linkedin.com/advice/0/how-do-you-analyze-customer-feedback#:~:text=Through%20text%20mining%2C%20you%20can,or%20categories%20of%20the%20text>.

OSMOND, C. (2023, September 2). *The customer is always right - origin, meaning & explanation*. GRAMMARIST. <https://grammarist.com/phrase/the-customer-is-always-right/>

Princess Eunike (2024). Implementation Importance-Performance Analysis Method to Increase Customer Satisfaction of Honda Motorcycle Dealer using Text Mining. International Journal of Current Science Research and Review, 7(2), 1094-1100

Santos, M., Ribeiro, R., Batista, F. & Correia, A. (2022). Examining Airbnb guest satisfaction tendencies: A text mining approach. Current Issues in Tourism. 25 (22), 3607-3622

Sinha, S. (2024, September 22). *Expanding Air India’s punctuality woes leave flyers frustrated: India News - Times of India*. The Times of India. <https://timesofindia.indiatimes.com/india/expanding-air-indias-punctuality-woes-leave-flyers-frustrated/articleshow/113562007.cms#>

Skytrax. (2024, July 8). *World’s top 100 airlines 2024*. SKYTRAX. <https://www.worldairlineawards.com/worlds-top-100-airlines-2024/>

Szyndlar, M. (2024, October 17). *Customer satisfaction: Why it’s important in 2024*. Survicate. <https://survicate.com/customer-satisfaction/importance-customer-satisfaction/#:~:text=In%20simple%20words%2C%20customer%20satisfaction,predict%20business%20growth%20and%20revenue>.

Zhang, Y. (2024). *ANL312 Text Mining and applied project formulation*. Singapore University of Social Sciences