



Microsoft Innovation Lab

Summer Internship 2019

CritiQali - Customer Feedback Analysis

Team Members

Abhigna L Shastry

Karan Kumar G

Sumukh Aithal K

Mentor

Sumanth V Rao



Problem Statement

To help companies across Hotel and Airline Industries improve their sales and service by enabling them to understand the customer's expectations and issues. Design a "Customer Feedback Model" to analyze the customer review data of that company, and identify the most critical issues based on their severity and impact on customers as well as management.





Domains

1

Natural Language Processing: Natural language processing (NLP) is the ability of a computer program to understand human language as it is spoken and written. NLP is a component of Machine learning(ML) and Artificial Intelligence (AI). It deals with how to program computers to process and analyze large amounts of natural language data

2

Full stack development: A combination of both Front-end and Back-end web development to design and host fully functional websites and web applications



Why use our product?

- Customer satisfaction is the top priority for any company
- Large companies receive thousands of critical remarks across online platforms on a daily basis.
- Assigning human labour to manually read through such a huge quantity of reviews is inefficient.
- Our product aims at providing a automated approach to the same issue
- With this advanced method, companies can provide better and faster solutions to customer issues, become more popular and customer-friendly

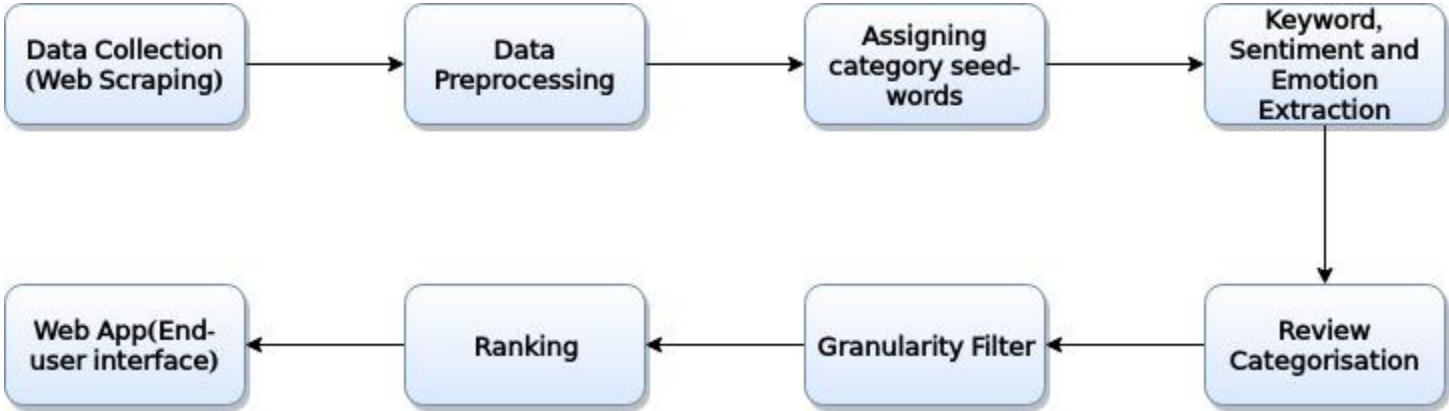


Approach

1. **Web Scraping** : to extract reviews from online reviews platform(TripAdvisor)
2. **Data Preprocessing** : to clean raw data and convert it to suitable format to run NLP processes
3. **Keyword, Sentiment and Emotion extraction** : using IBM's Natural Language Understanding(NLU) API
4. **Review Categorisation** : categorise each of the reviews based on a similarity score
5. **Granularity Filter** : to retain reviews which provide specific information, and discard the more general ones
6. **Ranking** : to rank the reviews based on relevance and severity and obtain the most critical issues
7. **Web App** : Online interface for any company to use our product



Workflow





Weekly Progress

Literature Survey, Web
Scraping

Review Categorisation



Data Preprocessing and
Aspect Keyword Generation

Aspect-Opinion
Extraction and Sentiment
Analysis



Weekly Progress

Keyword generation using
IBM NLU, Granularity
Filtering

Week 5

Week 6

Ranking, Emotion score
generation

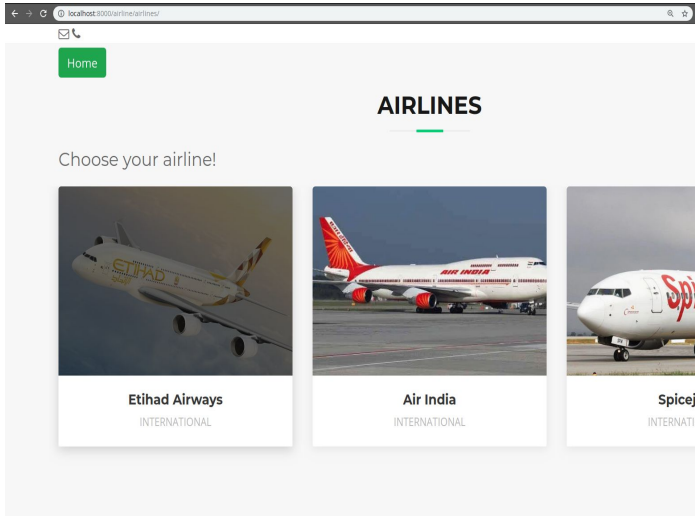
Web app dev,
Containerisation and
Modularisation

Week 7

Week 8

Extended to Airline
industries, Fine Tuning
and Report writing

Results - Airlines



AMBIENCE



Issue No.	Review	Date	From	To	Class	Source
1	Seats were warn, foot rests and walls with many scuff marks, toilets can be best described as having a lick and promise to clean.	04-2019	New Delhi	London	Business Class	
2	The enetertainment system as broken, the seat divider was flimsy and filthy.	03-2018	New Delhi	Chicago	Business Class	
3	Seat to Delhi were a bit dirty but fine on our way back.	02-2019	Singapore	New Delhi	Economy	
4	After I had booked the flight I read several comments about how dirty the AI seats were and how poor the service was.	03-2018	London	New Delhi	Business Class	
5	The seat was so dirty that when the stewardess came around with the hot towels, I used it to was the complete headrest and arms.	04-2019	New Delhi	Singapore	Business Class	

✕ Close Window

Results - Airlines

localhost:8000/categories/airline/Air%20India/

Airlines

AIR INDIA

Most Critical Issues



Results - Hotel

Hotels

GOLDEN PALMS HOTEL & SPA

Most Critical Issues



FOOD



Issue No.	Review	Date	Source
1	The staff were very unapproachable, the taste of food in restaurant was pathetic while the prices were of 5 star rates.	Nov-18	
2	The food menu for the room service is too expensive (Better to go to the restaurant and have food to save some bucks)	Feb-18	
3	A dead fly was stuck on the bread with many other flying around (0 for hygiene) the infrastructure looks old and slightly run down.	Jul-17	
4	Food is above average, room service is extremely delayed, many fountains have stopped functioning.	Apr-18	
5	The food was bad and priced at Rs 1250 per head..too expensive for the kind of food served..should have been priced at no more than Rs 500/ head.	May-17	

✕ Close Window

Results - Hotel

CLEANLINESS



Issue No.	Review	Date	Source
1	The rooms had a strange musty smell and the toilets were just not clean as one would expect in a property like this.	Feb-17	
2	when we checked into our room, it was stinking with alcohol & cigarette smell from the previous occupant.	May-18	
3	as it was stinking it seems their housekeeping staff has not visited that floor from 15 days.	Sep-17	
4	The hallway stinks and inside the room foul smell of some food.	Mar-19	
5	(no towels in the room, dirty bedsheets.....	Jan-19	

× Close Window



Challenges

1. Data was unlabeled and hence unsupervised learning techniques had to be used which performed poorly.
2. Recognizing emotion and sentiment based on aspect keyword was challenging as the state of the art techniques were all based on supervised learning.
3. Grammatically incorrect reviews were not classified and ranked correctly.
4. Ranking issues based on granularity and frequency. The more specific issues were not frequent and hence identifying them was a challenge.



Future Work

1. Build a robust model in place of the IBM Natural Language Understanding API, which returns emotions and sentiment based on aspect keywords
2. Generalize the product across industries such as Education, Restaurants, etc.
3. A more optimised ranking algorithm
4. Dynamically input details from the companies and return the critical issues real-time.



THANK YOU!

