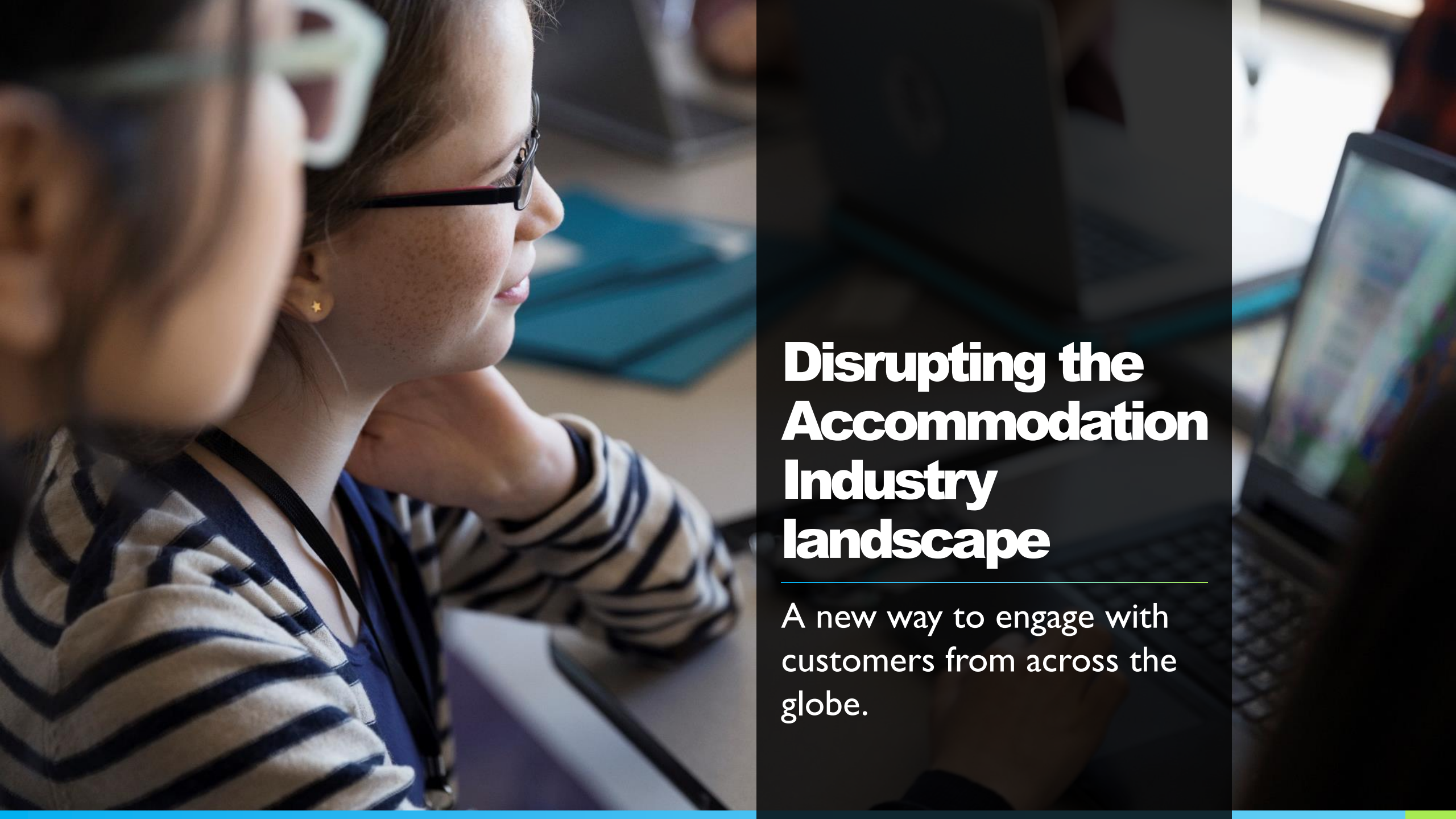


Disrupting the Accommodation Industry landscape

A new way to engage with customers from across the globe.

Sentiment Analysis of Accommodation Industry Reviews

…using Intelligent Feature Extraction.

Project proposed by

Nikhil Miranda - 1BY14EC046

Under the Guidance of

Dr. Hanumantharaju M.C. & Prof. Laxmisagar H.S.

Abstract

- The problem of **Sentimental Analysis** mostly deals with the **classifying** the sentiments into majorly **3** categories (Positive | Neutral | Negative)

- The **traditional sentiment classification** involves treating the entire sentence as a text document and then classifying the sentiments based on all the words **individually**.

- In this project, the data will be **pre-processed** first and then **cleansed** for training the **ML model**, which will be then used for **feature extraction**.

Introduction

- Over the past **TWO** years **social media** has become an emerging phenomenon due to the huge and **rapid advances in information technology** and its services.

- People are using social media on a **daily basis** to communicate their opinions with each other about a wide variety of products and services, which has made it a rich resource for **text mining and sentiment analysis**.

- In this project a standard approach will be presented using open source tools which will take care of the possible errors which occur during the **manual analysis of the reviews**.

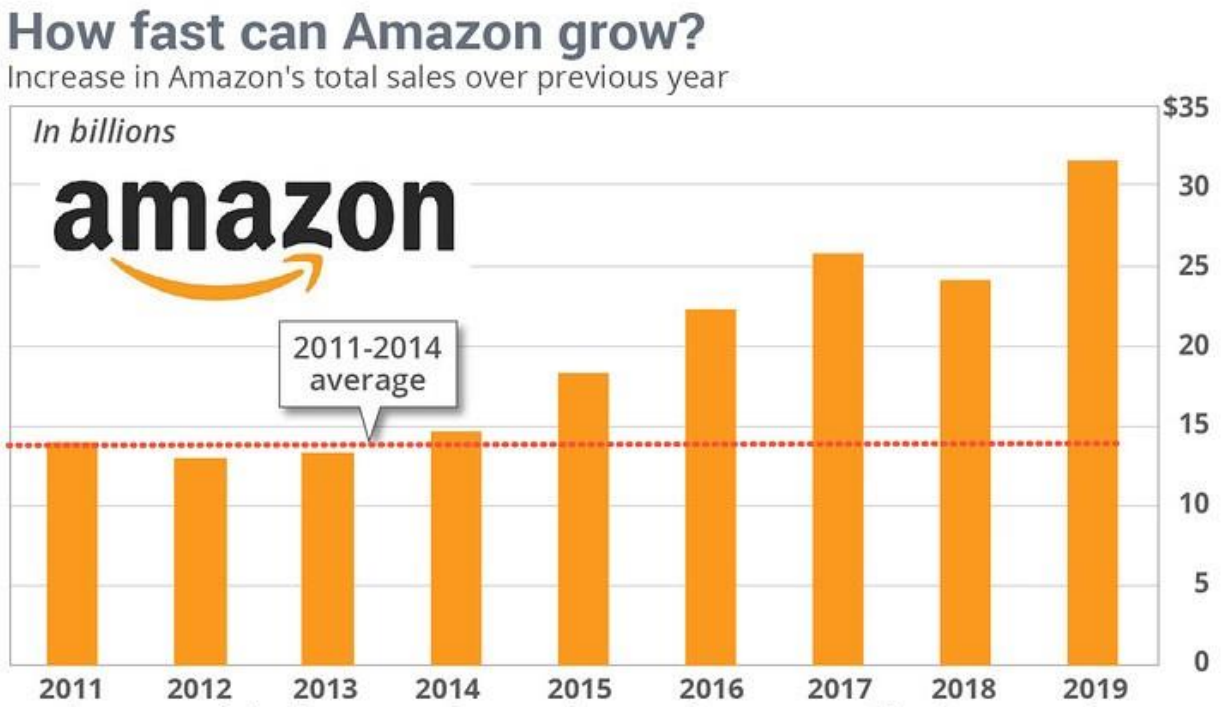
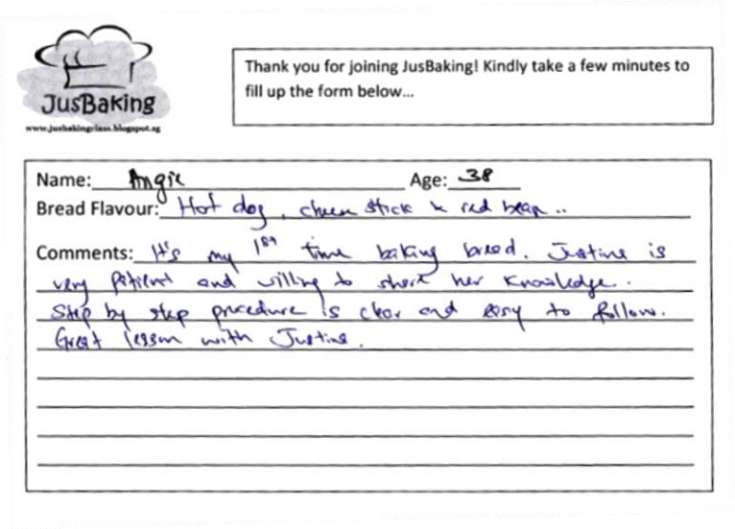


Image taken from forbes.com depicting Amazon's Prime Customer Growth

Motivation

- The Review System was introduced in order to inform the **manufacturer** of a specific product about **the faults** in it so that the user can use this information know what faults he need **to rectify in his next upcoming product**.


JusBaking

Thank you for joining JusBaking! Kindly take a few minutes to fill up the form below...

Name: Angie Age: 38

Bread Flavour: Hot dog, cheese stick & red bean..

Comments: It's my 1st time baking bread. Justine is very patient and willing to share her knowledge. Step by step procedure is clear and easy to follow. Great lesson with Justine.

Hand Written Reviews

Sloppy Hand Writing
And
Time Consuming


Reviewed 1 week ago via mobile

Believe the 5 star rating you won't be disappointed.

...The rooftop bar provides great music, some interesting and colorful drinks, wonderful staff and fabulous views into the centre of the hotel and the city. We also had massages at the spa that were relaxing and comfortable the staff bent over backwards to accommodate our... More

123DLP456
Toronto, Canada
27 13

Thank 123DLP456

Response from Kornél M, Music Director at Aria Hotel Budapest by Library Hotel Collection

Responded 5 days ago

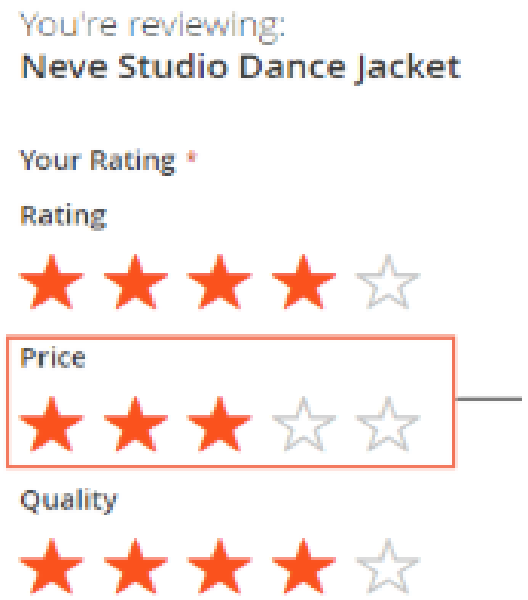
Dear 123DLP456, Many thanks for your nice review. You have just set the perfect time for your visit: Budapest has changed a lot during the last couple of years, and it has gained back its old glamour which made it a paradise for artists, musicians... More

Online Text Review

Difficult to Analyse
And
Time Consuming

Motivation

- But lately the consumers have been using the Review System **to judge the product of its quality** by deducing **their own inferences** from other consumer's reviews.



Star Rating Reviews

Less Expressive
And
Limited Review

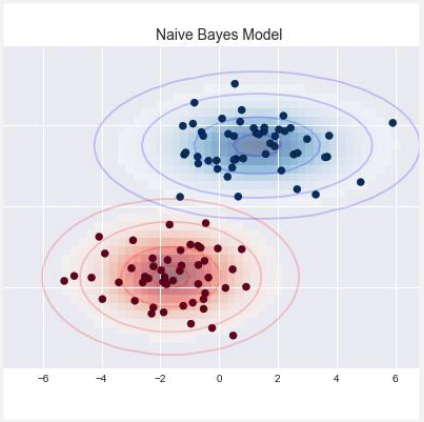
Present Review System

Difficult to Analyse
And
Time Consuming

Problem Statement

Given a feedback review, **classify** whether the message is of **positive, negative, or neutral sentiment**. For messages conveying both positive and negative sentiment, whichever is the **stronger** sentiment should be chosen.

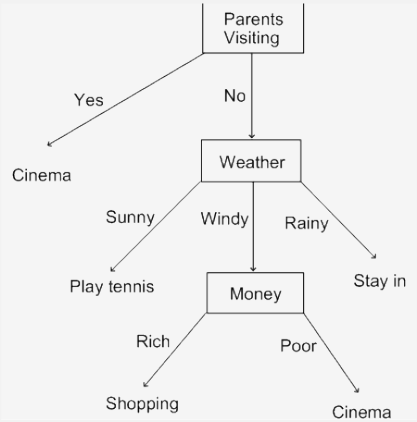
Literature Survey



Naïve Bayes

Accuracy : **53.99%**

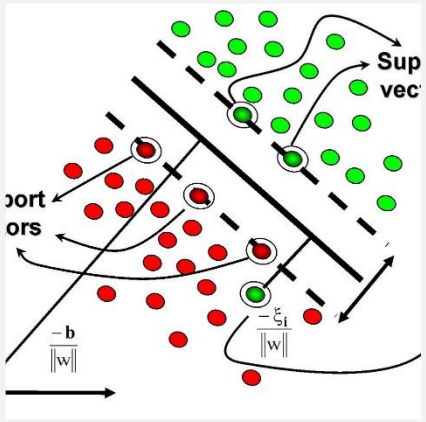
- ✓ Highly Scalable
- ✓ Fast for Small Dataset
- ✓ Good for Text
- ▪ Considers only Independent Features



Decision Trees

Accuracy : **53.97%**

- ✓ Highly Scalable
- ✓ Good for Text
- ✓ Fast for Small Dataset
- ▪ Over Fitting
- ▪ Unusable for Continuous Variables



Support Vector Machine

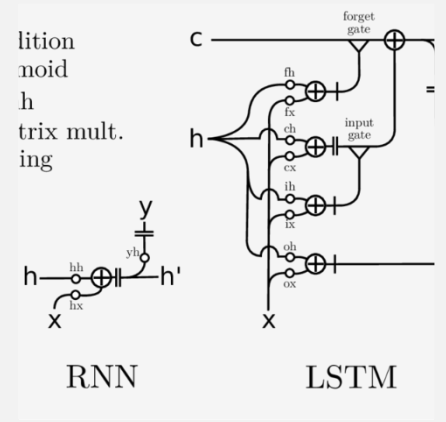
Accuracy : **60.11%**

- ✓ Good for Unknown data
- ✓ Works with unstructured data
- ✓ Highly Scalable
- ▪ Difficult to understand Final Model
- ▪ Long time for Training

XGBoost

Accuracy : **65.11%**

- ✓ Highly Scalable
- ✓ More Expressive
- ✓ Low Memory Usage
- ▪ Long time for Training



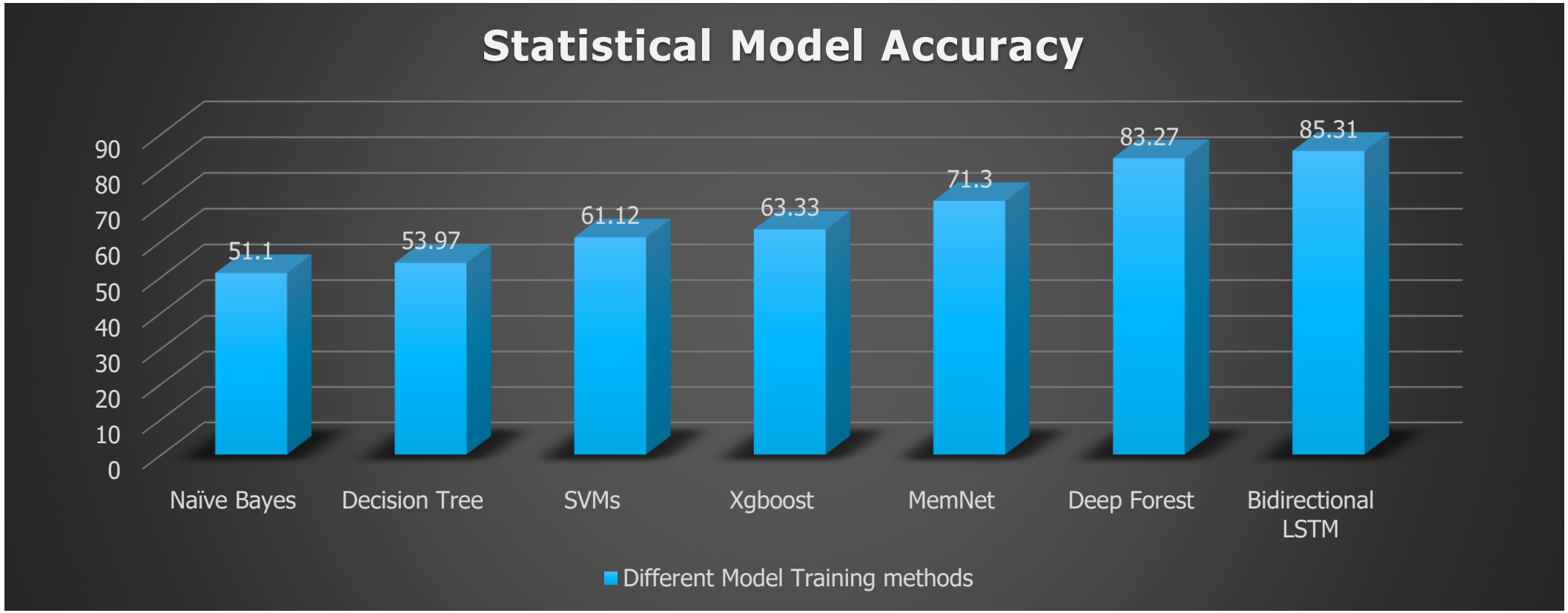
Bi-directional LSTM

Accuracy : **85.31%**

- ✓ Fit for Sequential Data
- ✓ Learn Long Sequences
- ▪ Difficult to Train
- ▪ Gradient Vanishing
- ▪ Exploding Gradient

Literature Review

Among the various **ML Algorithms** that are popularly recognized for sentiment analysis **Naive Bayes, SVM, Deep Forest & Bidirectional LSTM** have shown promising results in Product Review classification and in recent Sentiment Analysis researches.

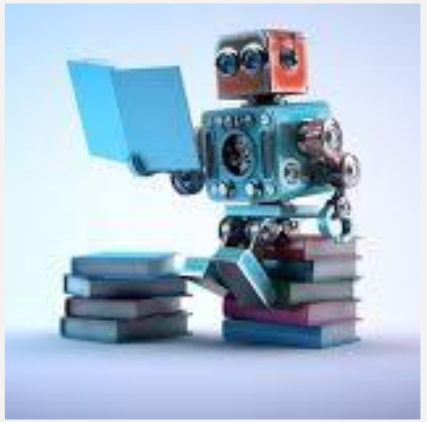


Limitations of Existing System

There are majorly **FIVE** types of problem which are difficult to handle when it comes to sentiment analysis.

1. **Sarcasm** in a sentence.
2. **Emoticon** Detection. (☺,☹,etc.)
3. **Fake** Review Detection.
4. **Spelling** mistakes.
5. **Vocabulary** mistakes due to linguistic constraints.

Proposed Methodology



Data Accumulation

Web Scraping Framework

- scrapy.io
- import.io
- spinn3r.io

Data Cleaning

Data Pre-Processing

- Removing stop words
- Removing spaces
- Removing punctuation

Data Warehousing

Data Analytics

- Amazon RedShift
- TeraData
- Oracle
- MATLAB

Statistical Modelling

Machine Learning Models

- Naive Bayes
- Support Vector Machines
- Deep Forest
- Bidirectional LSTM

Visualisation

Graphs & Plots

- Tableau
- R studio
- MATLAB
- Infogr.am
- SmartDraw

System Requirements

Processor

- Intel i3-7100 3.9 GHz
- AMD Ryzen 5 1600
- And above.

RAM

- 8 GB DDR4 2666MHZ
- And above.

Graphic Processor

- Nvidia GeForce 1060 Ti
- AMD RX470 4GB
- And above.

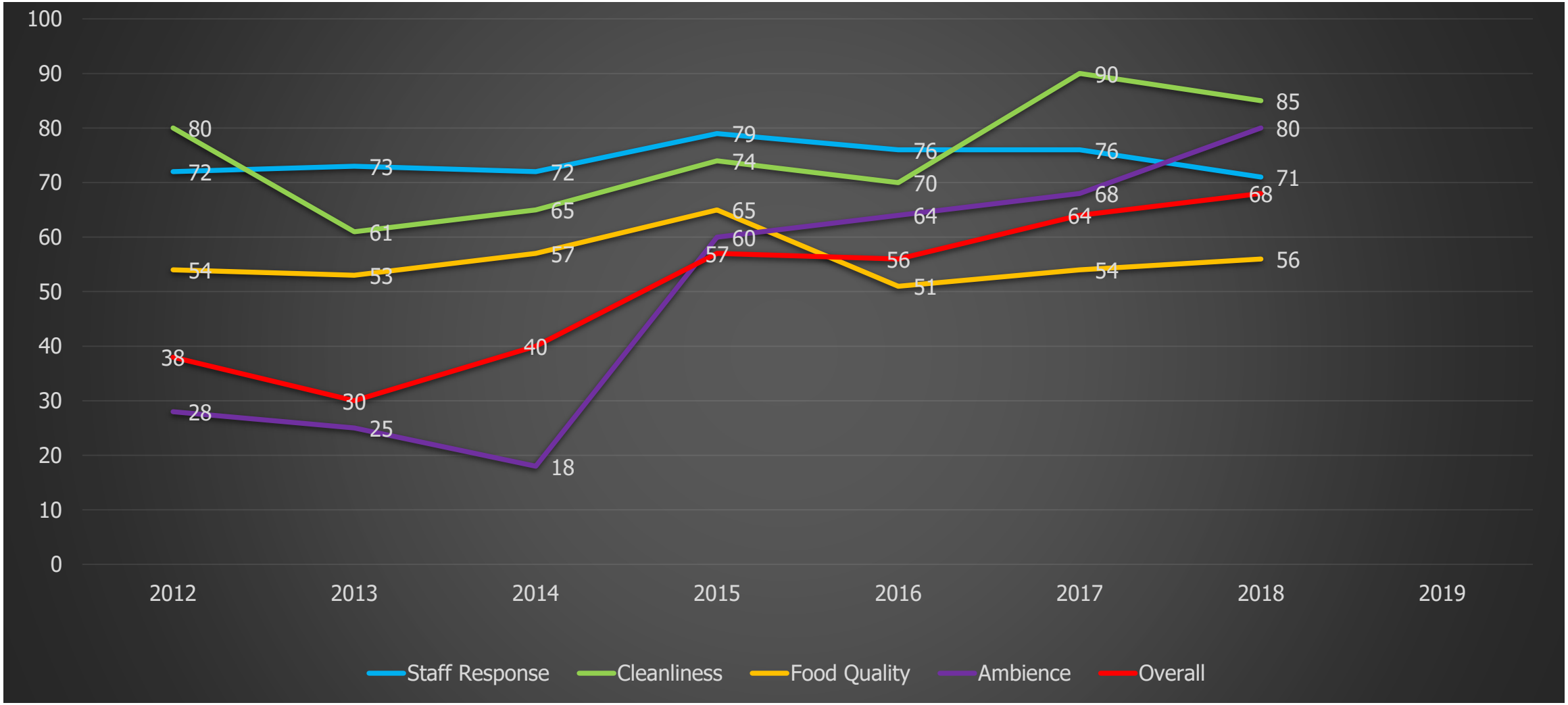
Hard Drive

- 50GB Free Space
- And above.

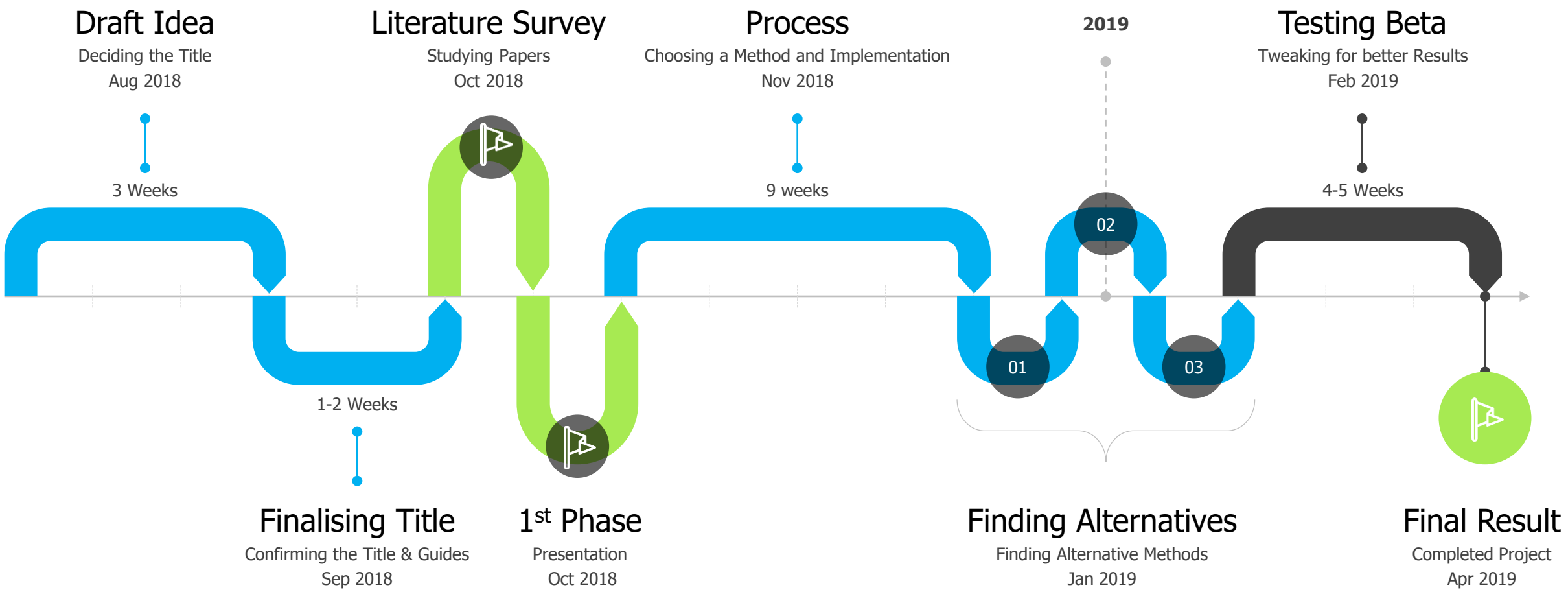
Cloud Services

- 60 GB Data (S3)
- 4.0 GHz (EC2)
- 16 GB RAM (EC2)
- eg1.2xlarge (EC2)
- Price : $0.80/hr

Expected Outcome



Project Roadmap



Team Members

Nikhil Miranda

STUDENT | BMSIT&M

Dr. Hanumantharaju M.C.

HOD, ECE | BMSIT&M

Prof. Laxmisagar H.S.

Asst. Prof. ECE | BMSIT&M

Thank You

Nikhil Miranda
https://nikhilmiranda.github.io