# Sentiment Analysis Of Airline Tweets

Le-Anne Lee | DSI 12

# Overview

- Introduction
- Problem Statement
- Understanding Airline Customer Tweets
- Findings from Modelling
- Future Implementation

# Impact of Social Media

Average Daily Time Spent on Social Networking Worldwide

144

minutes per day

More than half of the world's population use social media

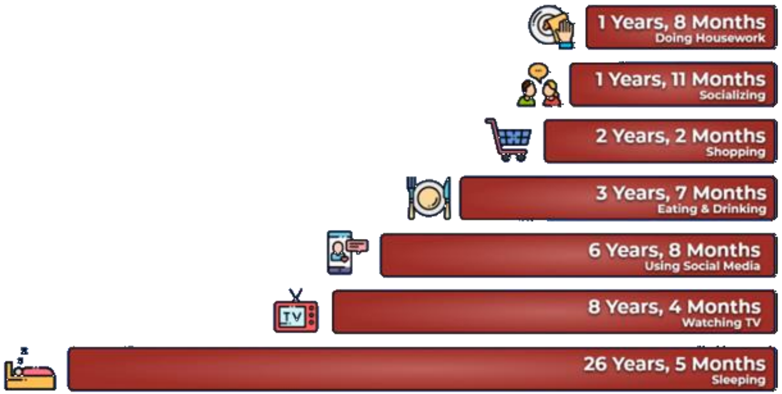
3.8bil

people in 2020

Source: Broadbandsearch

# Impact of Social Media

## Average Time Spent in a Lifetime





Source: Broadbandsearch

Source: Bureau of Labour Statistics

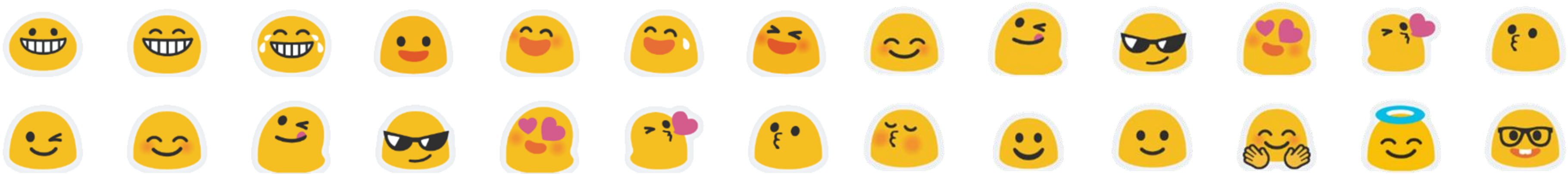
# Sentiment Analysis

## What it does

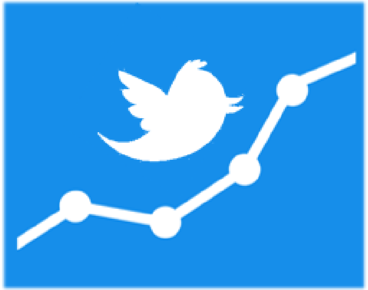
- Interpretation and classification of emotions within text
- Time-consuming to sieve through large amounts of data

## Business Outcome

- Identify customer sentiment towards brands / services
- Automatically analyse customer feedback to tailor services to meet their needs



# Problem Statement



How can airlines exploit twitter data to better respond to customer's needs?

# About the Data

- Airline tweets gathered in Feb 2015
- 14640 tweets

- Contains:
  - tweet id
  - airline sentiment
  - airline sentiment confidence
  - airline
  - name
  - retweet count
  - text
  - tweet coordinate
  - tweet created
  - tweet location
  - user timezone
  - negative reason

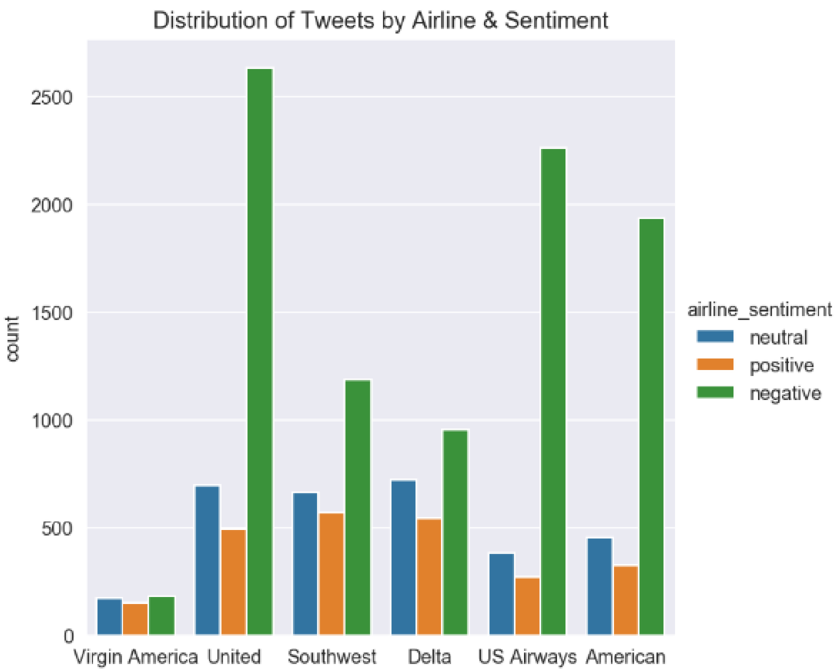
# Understanding your customers

# Data Insights

01

**Distribution of tweets by Airline & Sentiment**

Implications:
People tend to tweet more when they have something negative to say



# Data Insights

02

**Incorrect Tagging of Original Dataset**

- All 2222 observations tagged as Delta posts were referring to @JetBlue

| | text | airline |
|---|---|---|
| 6746 | @JetBlue Yesterday on my way from EWR to FLL j... | Delta |
| 6747 | @JetBlue I hope so because I fly very often an... | Delta |
| 6748 | @JetBlue flight 1041 to Savannah, GA | Delta |
| 6749 | @JetBlue They weren't on any flight, they just... | Delta |
| 6750 | @JetBlue everyone is here but our pilots are n... | Delta |

# Types of Customer Tweets

## 03 #CUSTOMER_SERVICE_HOTLINE

8949
@JetBlue I have a internal bleed in my foot, and I am flying next Tuesday, what should I. Do :( these. Leeds come randomly. I
8945
@JetBlue no, I am fine to fly! Haha, they come at random, with the hemophilia, but what if i need extra assistance
8570
@JetBlue is it true there's a new A320 livery and a new website...?!
8569
@JetBlue BEST SEAT ON A E190 to board early. READY. SET. GO!
8568
@JetBlue I can't pay 30 bucks xD
8566
@JetBlue my mom wanted me to change her seat along with my sister, but their two different reservations and idk
8565
@JetBlue can't change it. True blue points and I can't get to a phone
8558
@JetBlue can you DM?
8409
@JetBlue is flight 51 on 4/24/15 moved back? When I booked it said we arrive 11:31 but now it says 12:08 😢
8297
@JetBlue I can't do that flight. I need a Late Flightr one! I need you to change my flight. You guys changed it and now I can't do that time!
8296
@JetBlue I can't. I don't have acces to a phone rn. My iPhone broke. :/ would rather change it now then Late Flightr.
8295
@JetBlue should I check in my awesome bag on my flight or carry it on... Decisions...
8294
@JetBlue okie doke! Knowing you, you will fix this ;)
8097
@JetBlue thanks to the gent on the phone who fixed my BOS-MCO flight and the fee waiver! A320 now :) #flyfi ! I forget her name :(

# Types of Customer Tweets

## 03 #THE_SPAMMER

@USAirways YOU ARE THE BEST AIRWAYS!!!!!!!!!! FOLLOW ME BACK, PLEASE 🙏🙏🙏🙏🙏🙏🙏🙏🙏🙏🙏😢😢😢😢😢😢😢😢🙏🙏🙏🙏
🙏🙏🙏🙏🙏🙏🙏🙏🙏🙏🙏🙏🙏🙏🙏🙏🙏
11706
@USAirways YOU ARE AMAZING!!! YOU ARE THE BEST!!! FOLLOW ME PLEASE 🙏🙏🙏😢😢😢🙏🙏🙏
11714
@USAirways YOU ARE THE BEST AIRWAYS! Follow me please!!!!!🙏🙏🙏🙏🙏🙏🙏🙏🙏🙏🙏🙏🙏🙏🙏🙏🙏🙏🙏🙏🙏🙏🙏🙏🙏🙏🙏
11720
@USAirways YOU ARE AMAZING!!! YOU ARE THE BEST!!! FOLLOW ME PLEASE AND I FOLLOW YOU BACK;)🙏🙏🙏✌😉)
12313

# Types of Customer Tweets

## 03 #NEWS_SOURCE

'@JetBlue Fliers to Gain Access to WSJ Content - Analyst Blog - Nasdaq http://t.co/dWEse7Xidr'

'@JetBlue Airways Stock Rating Lowered by Vetr Inc. (JBLU) - Dakota Financial News http://t.co/QW2eBEEMVg'

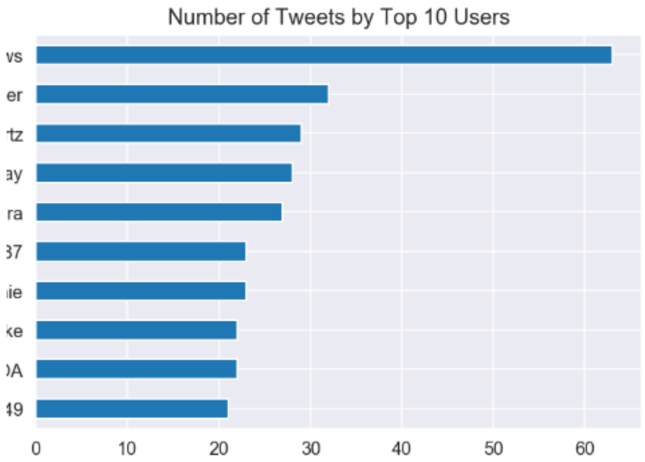
"@JetBlue's new CEO Robin Hayes battles to appease passengers and Wall Street - Business In Savannah http://t.co/KKAY8XaPs1"

# Types of Customer Tweets

03

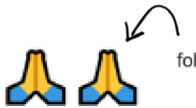
**Highest number of tweets by the same user :**
**60** out of **14640** over tweets.

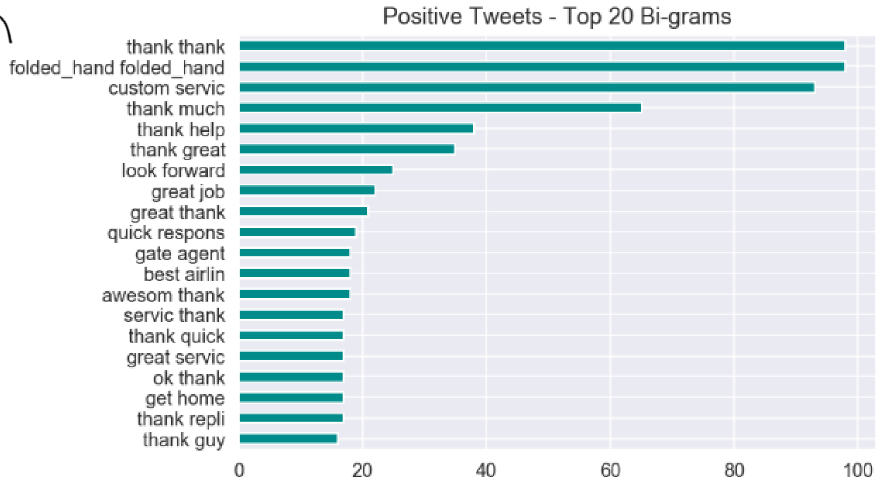
**Implications:**
No concern that the data is overly skewed by one user



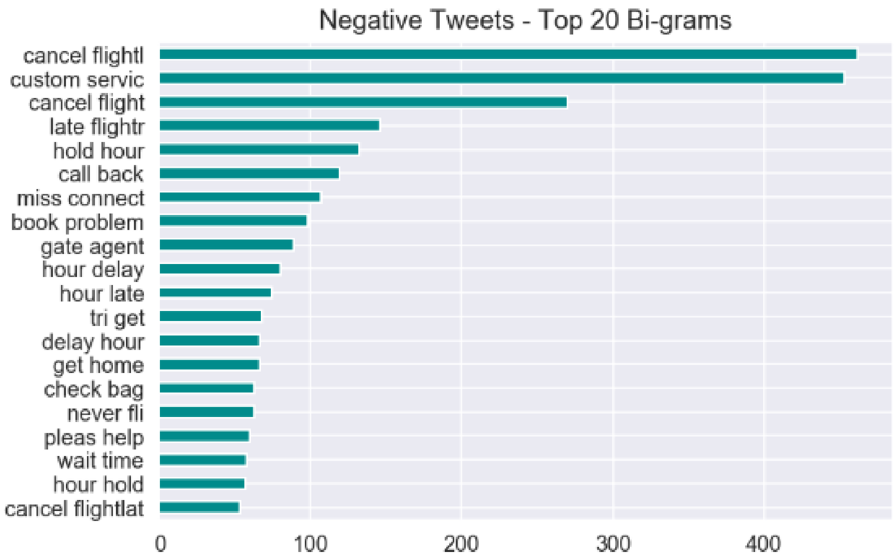
# Common Words

#POSITIVE





# Common Themes #NEGATIVE



# Data Preprocessing

- Retain emojis
- Retain stopwords
- Spelling errors
- Short forms

# Modelling Results

| Model | Accuracy Score |
|---|---|
| Logistic Regression | 0.801 |
| Naive Bayes | 0.748 |
| Decision Tree | 0.681 |
| Random Forest Classifier | 0.761 |
| Ada Boost | 0.755 |
| Gradient Boost | 0.751 |
| XG Boost | 0.774 |

| Model | Accuracy Score |
|---|---|
| Vanilla RNN Model | 0.781 |
| RNN with LSTM | 0.750 |
| RNN with LSTM & Dropout | 0.778 |

# Summary of Findings

## Model Performance

- Performed best on negative tweets
- Precision for Neutral tweets was only 70%

## Areas of Improvement

- POS Tagging
- Topic Modeling
- Word2Vec
- Include more positive and neutral posts

# Model Deployment


Removal of Spam Tweets


Sentiment Analysis


Negative Tweets to be Addressed by Customer Service


Topic Modelling

# Conclusion

- Model was able to accurately classify 80.1% of tweets
- Pre-processing of features is the most important
- Future work includes:
  - classification of spam tweets
  - topic modelling

# Thank you.