# Aspect Based Sentiment Analysis of Airline Tweets

# Who are we – We are the 🤗

Jose Bordon
Data Scientist

Ryan Kirby
Control Manager

Shruti Korada
Business Analyst

Zachary Stockmal
Software Engineer

# The AI Canvas

## What task/decision are you examining?

We are examining tweets and deciding how to analyze the aspect or topic of that tweet as well as the customer's sentiment towards it.
We want to give the company an evaluation of the topic, so they can task out whether it requires remediation.

### Prediction

Predict whether a tweet towards an airline has positive or negative sentiment, and what aspect caused that.

### Judgment

Identifying a truly negative tweet can give the company insight on where their efforts should be placed and quickly resolved. Falsely identifying a negative tweet as positive could leave the company with more unsatisfied customers.

### Action

The airline can then identify the source of many of their negative tweets and work to resolve it, as well as improve their brand's standing on social media (restart a service, address a delay, etc.)

### Outcome

When an airline acts on the aspect presented that was negative, we expect to see if in a different timeframe the negative sentiment towards the aspect go down. (example: 100 people hate the food during 2021, they change the food in 2022, then we only see 20 people hate the food)

### Training

Our sentiment model will stay the same, but our aspect identifying model will stay trained on a historical record of the tweets before hand with generic clusters setup. As the prediction machine is deployed, we will receive more data from the newer tweets that come in, and have it adjusted to the whims of the public over time.

### Input

A steady stream of tweets about the airline will generate our aspects and sentiment analysis for the model to work from. Gaps in knowledge from the tweets can be supplemented with flight records and other public data relating to the time period. (holidays, other social media, etc.)

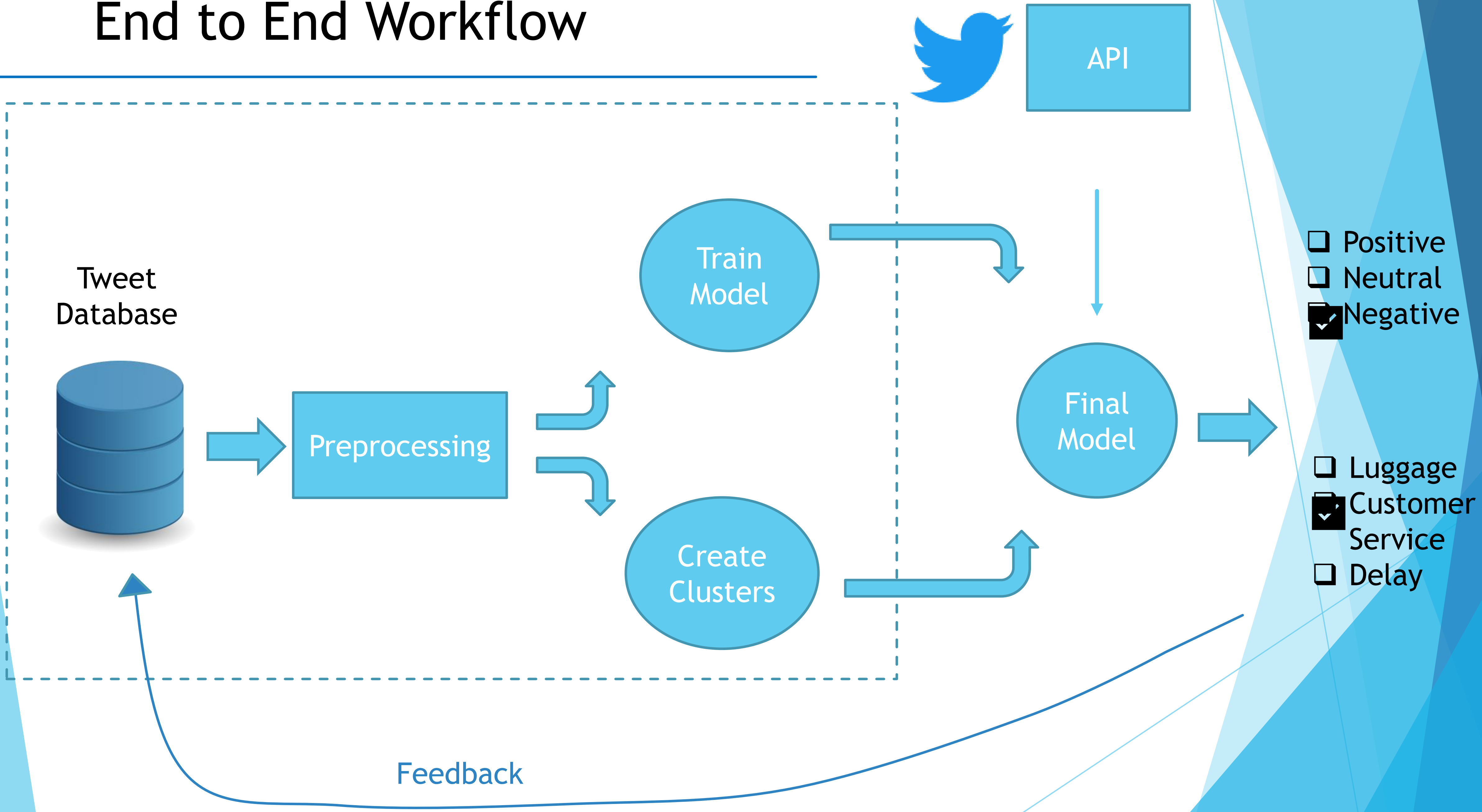
### Feedback

On a monthly cycle we will package all the tweets run through our model with their respective sentiment prediction and aspect identified and outsource the validity of models predictions. If we are happy with our positive rate we will let the model continue as is, if not we will have to retune our model with the new data

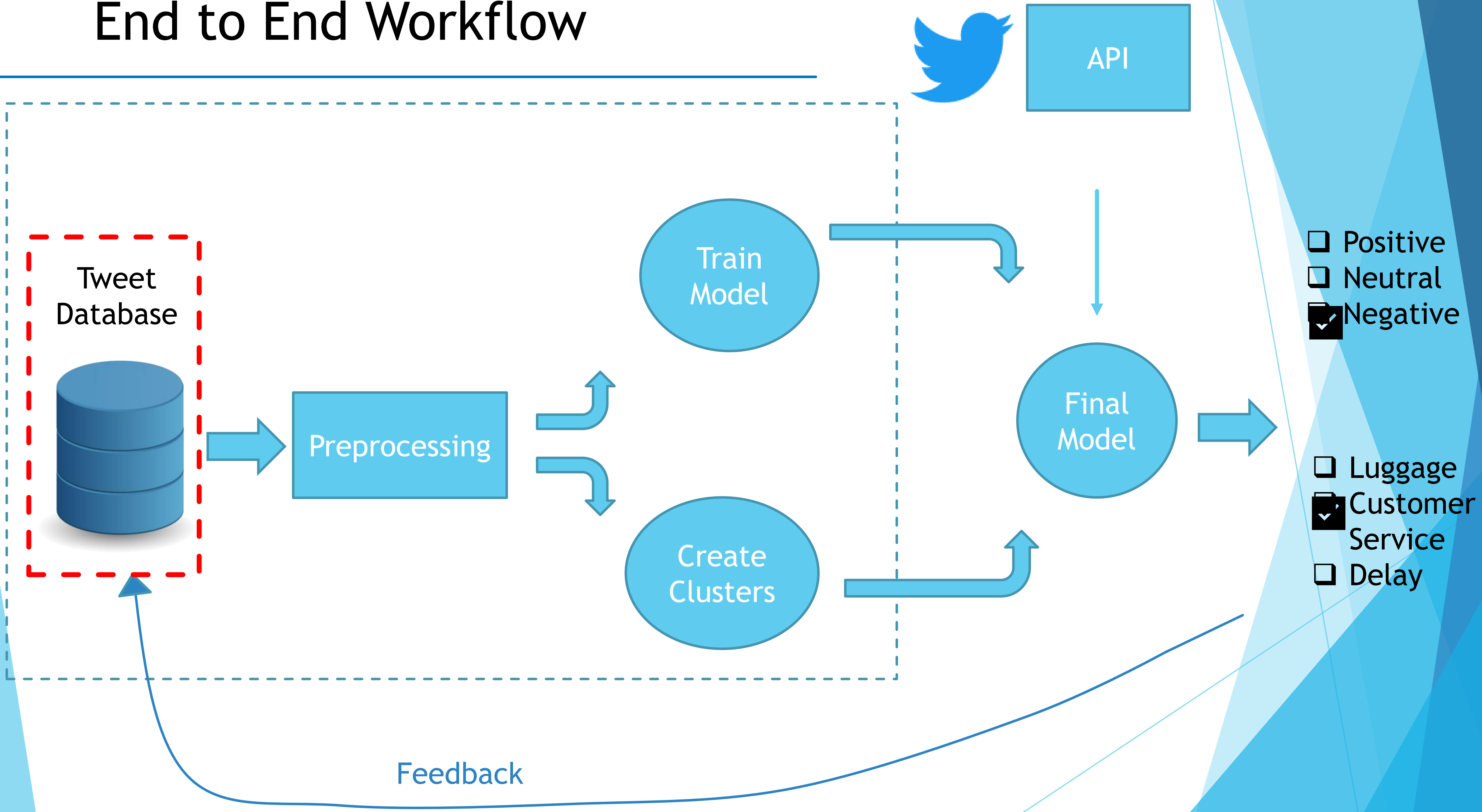
## How will this AI impact on the overall workflow?

The goal of this AI is reducing the workload of social listening for a company. Customer service representatives will need to be retrained to identify the aspects and which ones require escalation or remediation. This AI will give airline companies a more efficient and effective method of finding the issues their company is causing for the customers. Usually customer surveys go unanswered, but tweets are more easily accessible making it a better identifier of a company's standing on social media, and what issues need their attention.

# End to End Workflow



# End to End Workflow



# Our Data

Twitter is a social network where users can post “tweets”, tweets are short post of up 140 characters.

The data set that we used consistent of approximately 15,000 tweets from users that have an @airline in their text. Airlines consist of United, Southwest, Delta, US Airways, American, and Virgin America

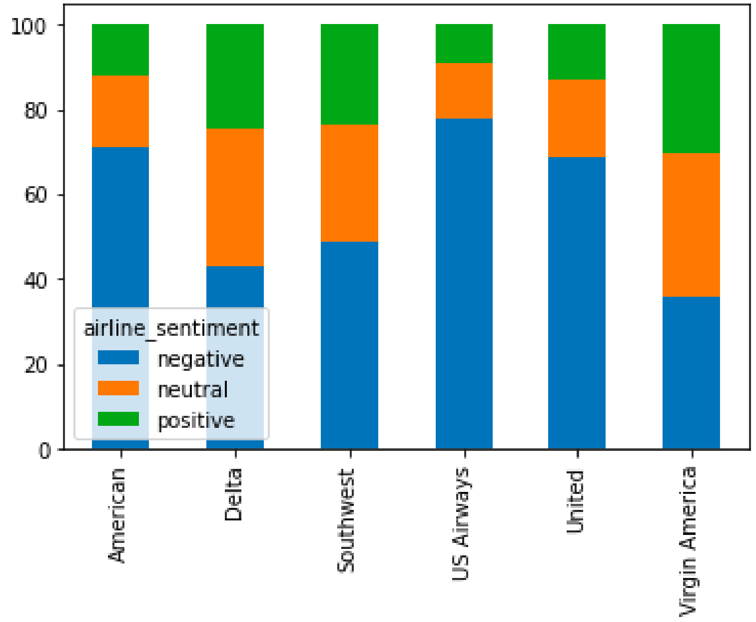
## Tweet


[redacted] · 1h ···
@UnitedAirlines_ direct flight booked was changed no notification with lay overs. I purchased seats, not honored. Now in the last seat on both legs. NOT what I purchased!!
After an hour on hold a
they fixed it, only to find when I arrived they did'nt! #frustrated #unitedsucks

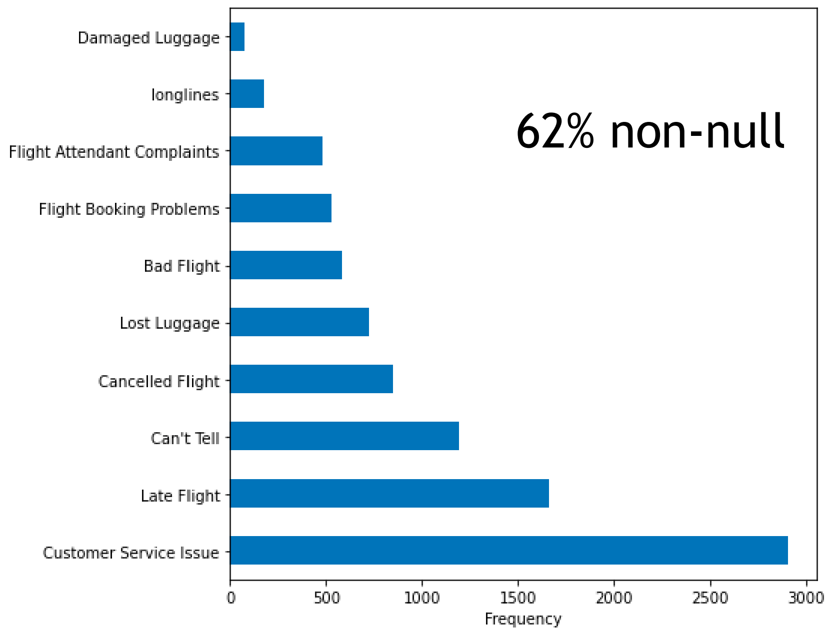
## Datapoint

| Airline | Tweet | Sentiment | Aspect | Other Cols. But not used |
|---|---|---|---|---|
| United | @UnitedAirlines_Direct flight booked was changed no notification with lay over. I purchases seats, not h…….s | Negative | Customer Service | ……… |

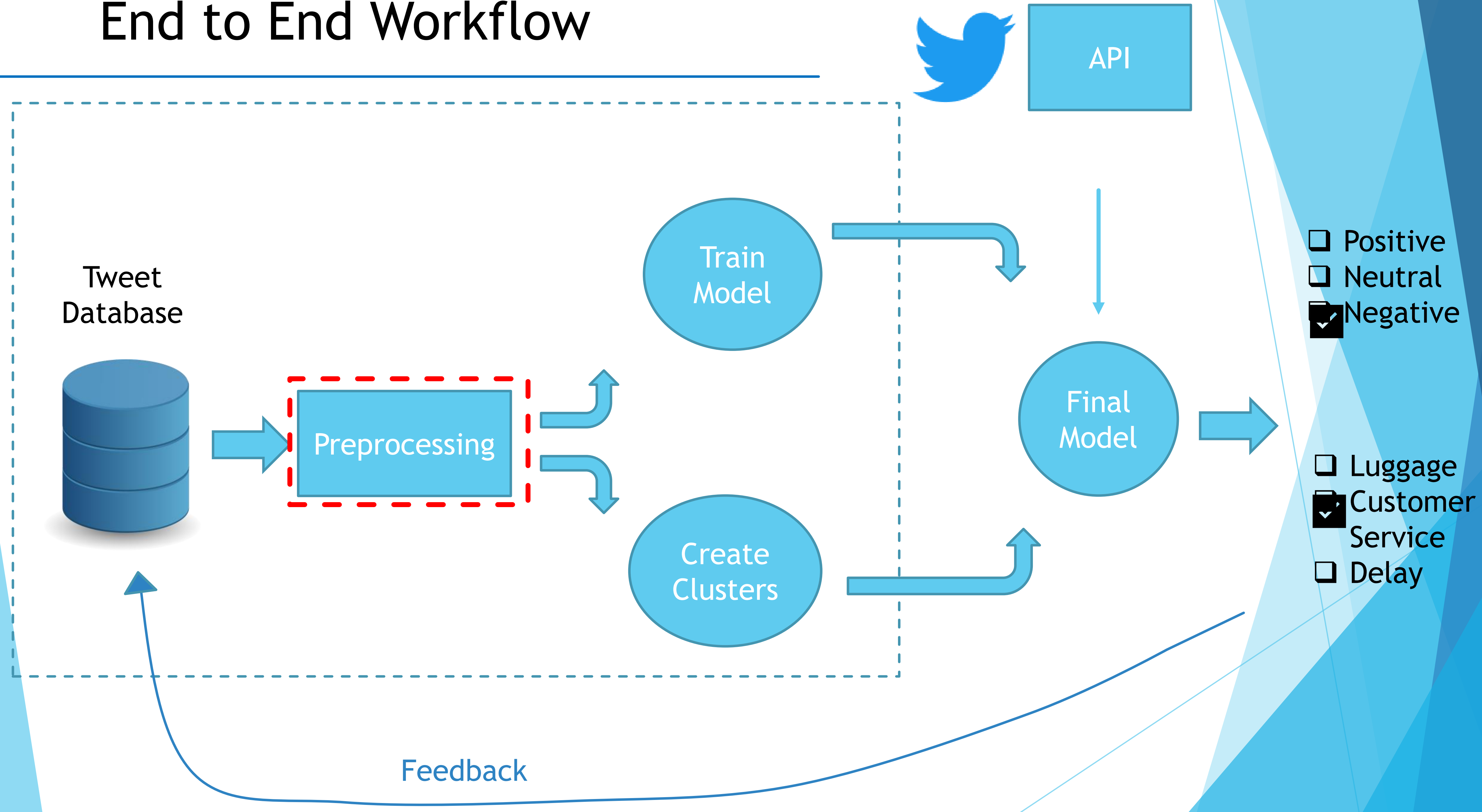
## Airline Sentiment




## Aspect



# End to End Workflow



# Data Scrubbing

**Tweets contain portions of text that aren't related to the sentiment or aspect and have to removed**

1. Remove Tagged Users
2. HTML Decoding
3. Remove any links
4. Remove any characters that aren't letters
5. Remove Stop Words
6. Stemming


[redacted]

43 minutes on hold with American Airlines..Im stuck on a loop about vouchers expiring...no pretty music...im sure they will pick up any minute🙏 #AmericanAirlines #AmericanAirlinesOnHold @americanairlnes youtu.be/Tw7HlhXBn2o @AmericanAirlines

10:31 PM · Jun 12, 2021 · Twitter for Android

**17** Likes

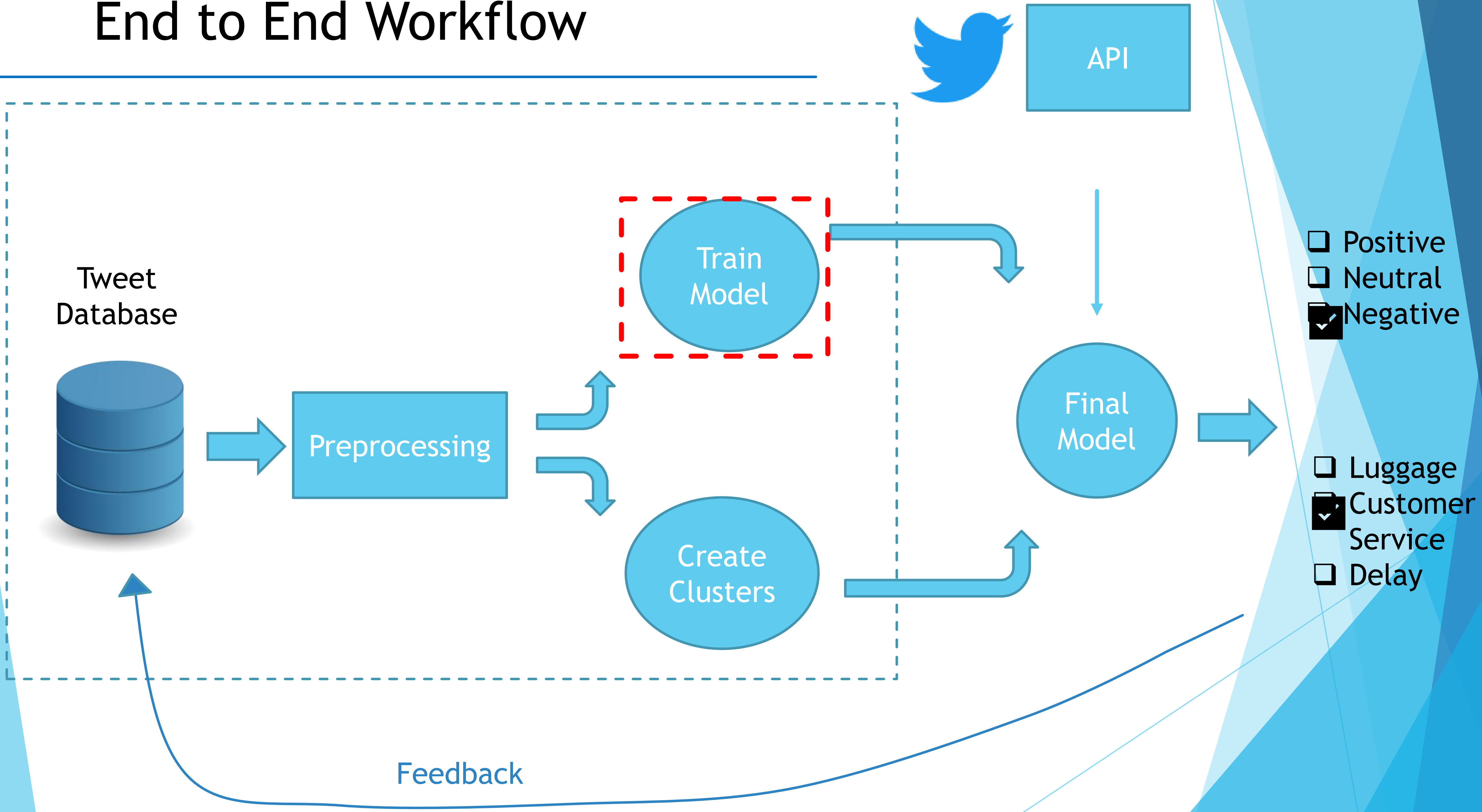
## Pre

43 minutes on hold with American Airlines..Im stuck on loop about vouchers expiring..no pretty music..im sure they will pick up any minuteA #Americanairlines#AmericanAirlinesOnHold@americanairlnes youtu.be/Tw7HlhXBn2o @AmericanAirlines

## Post

minute hold American Airlines Im stuck loop voucher expire no pretty music im sure they pick up minute Americanairline AmericanAirlinesOnHold

# End to End Workflow



# Sentiment Analysis

Baseline Analysis – Naïve Bayes - 55% Accuracy

•VADER - SentimentIntensityAnalyzer (nltk): 65%

- Precision: 0.898
- Recall: 0.504
- Accuracy: 0.653
- F1 Score: 0.646

•Hugging Face (BERT): 79%

- Precision: 0.939
- Recall: 0.711
- Accuracy: 0.790
- F1 Score: 0.809

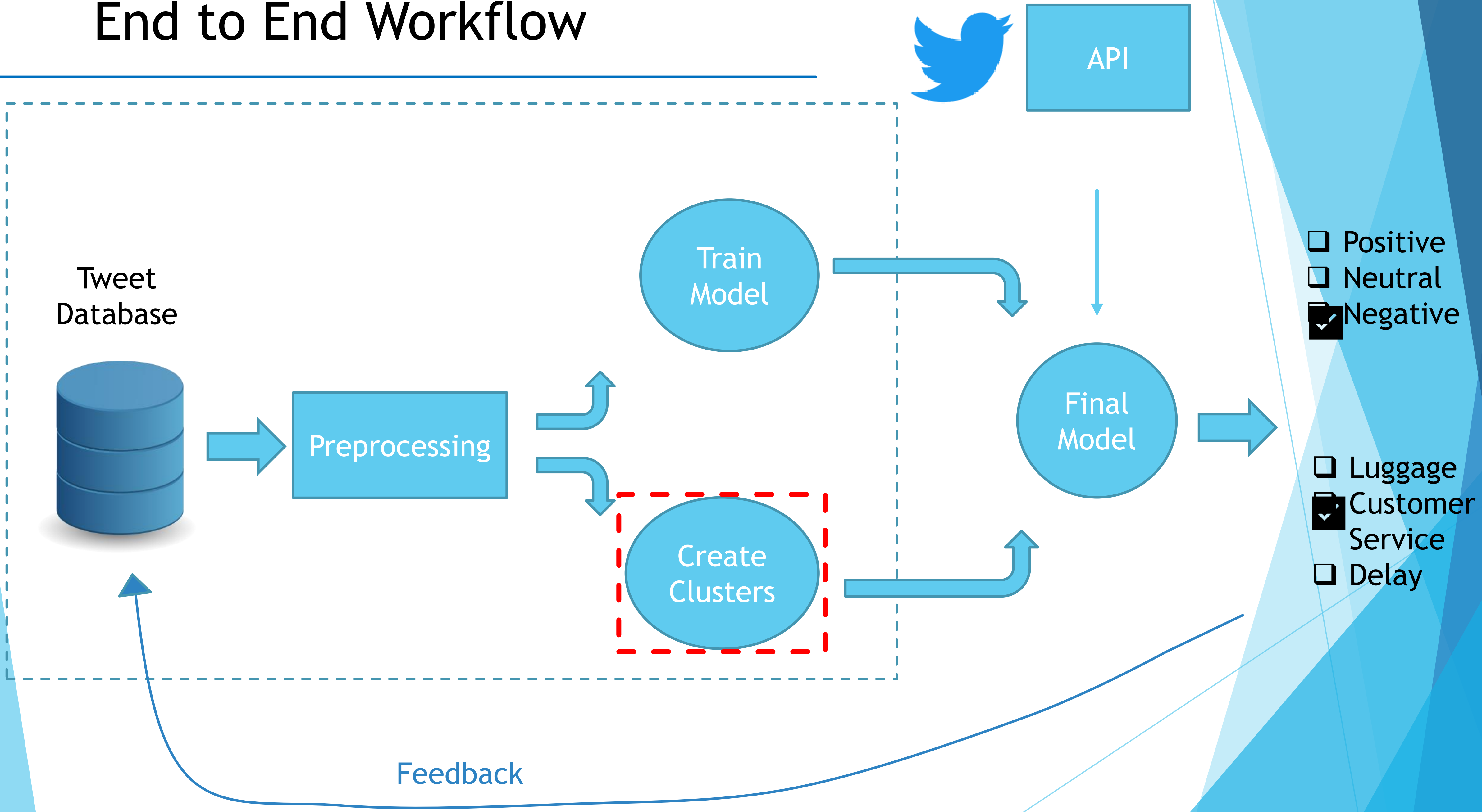
•Textblob x NaiveBayesAnalyzer (nltk): 69%

- Precision: 0.775
- Recall: 0.716
- Accuracy: 0.692
- F1 Score: 0.744

•Fine-tuned Hugging Face (BERT): 89% on the test subset.

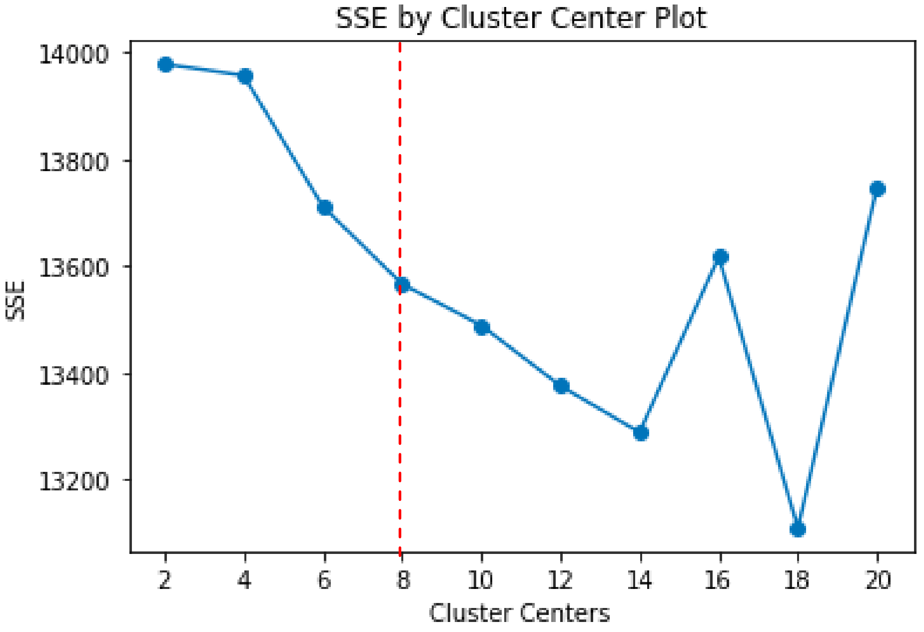
- With another airline tweets dataset:
- Precision: 0.853
- Recall: 0.738
- Accuracy: 0.791
- F1 Score: 0.791

# End to End Workflow



# Aspect Building

Due to aspects being unsupervised, we used an elbow curve to measure the inertia when adding more clusters to decide an appropriate amount




At 8 clusters we get little improvement in inertia

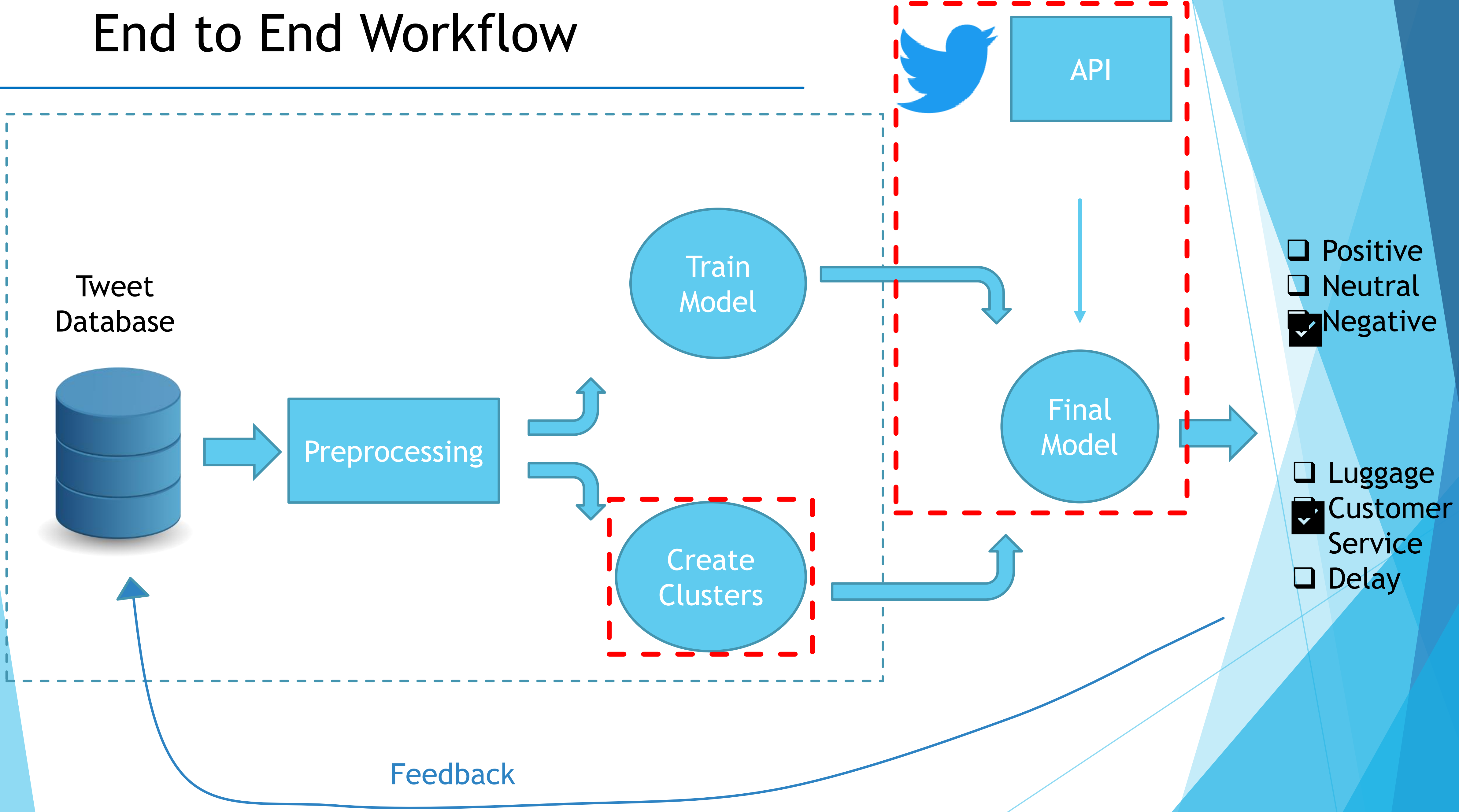
| Cluster 1 | | Cluster 2 | | Cluster 3 | Cluster 4 | |
|---|---|---|---|---|---|---|
| Plane | Left | Help | Because | Delayed | Great | Updates |
| Sit | Just | Bag | Phone | Late | Awesome | Very |
| Hour | New | Change | Want | Flightr | Appreciate | Safe |
| Gate | Passengers | Time | Thanks | Missed | Follow | Okay |
| Waited | Please | Know | | Connecting | Sent | Got |
| Why | Seats | Working | | Min | Good | Yes |
| Boarding | Stop | Booking | | hrs | | respond |

| Cluster 5 | Cluster 6 | Cluster 7 | Cluster 8 |
|---|---|---|---|
| Need | Hold | Service | Rebook |
| Trying | Cancelled | Customer | Tomorrow |
| Days | minutes | Worst | Dfw |
| Like | | Terrible | Reschedule |
| Guy | | Today | ticket |
| Check | | Poor | |
| | | line | |

Due to overlap between clusters, some clusters we're merged resulting in 4 clusters

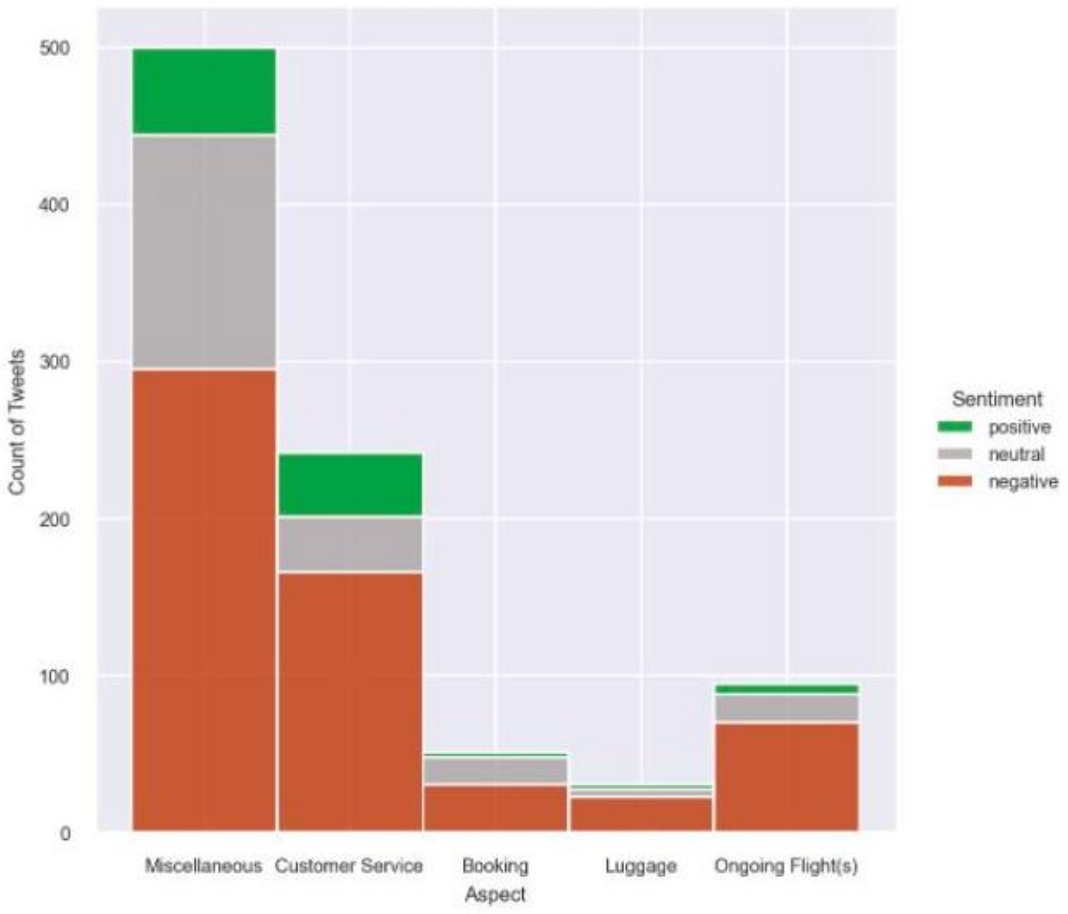
| | | | | |
|---|---|---|---|---|
| Customer Service | Ongoing Flights | Booking | Luggage | Wait time |

# End to End Workflow

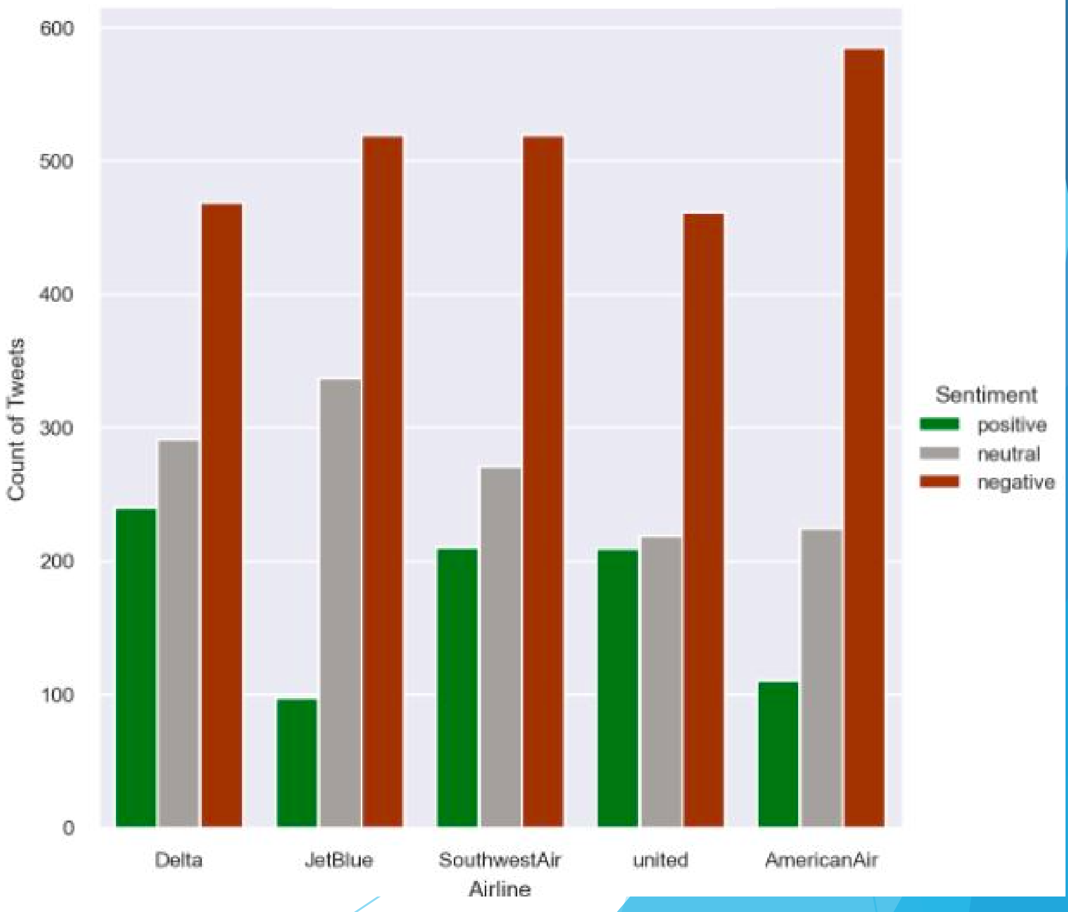


# Demo Part 1

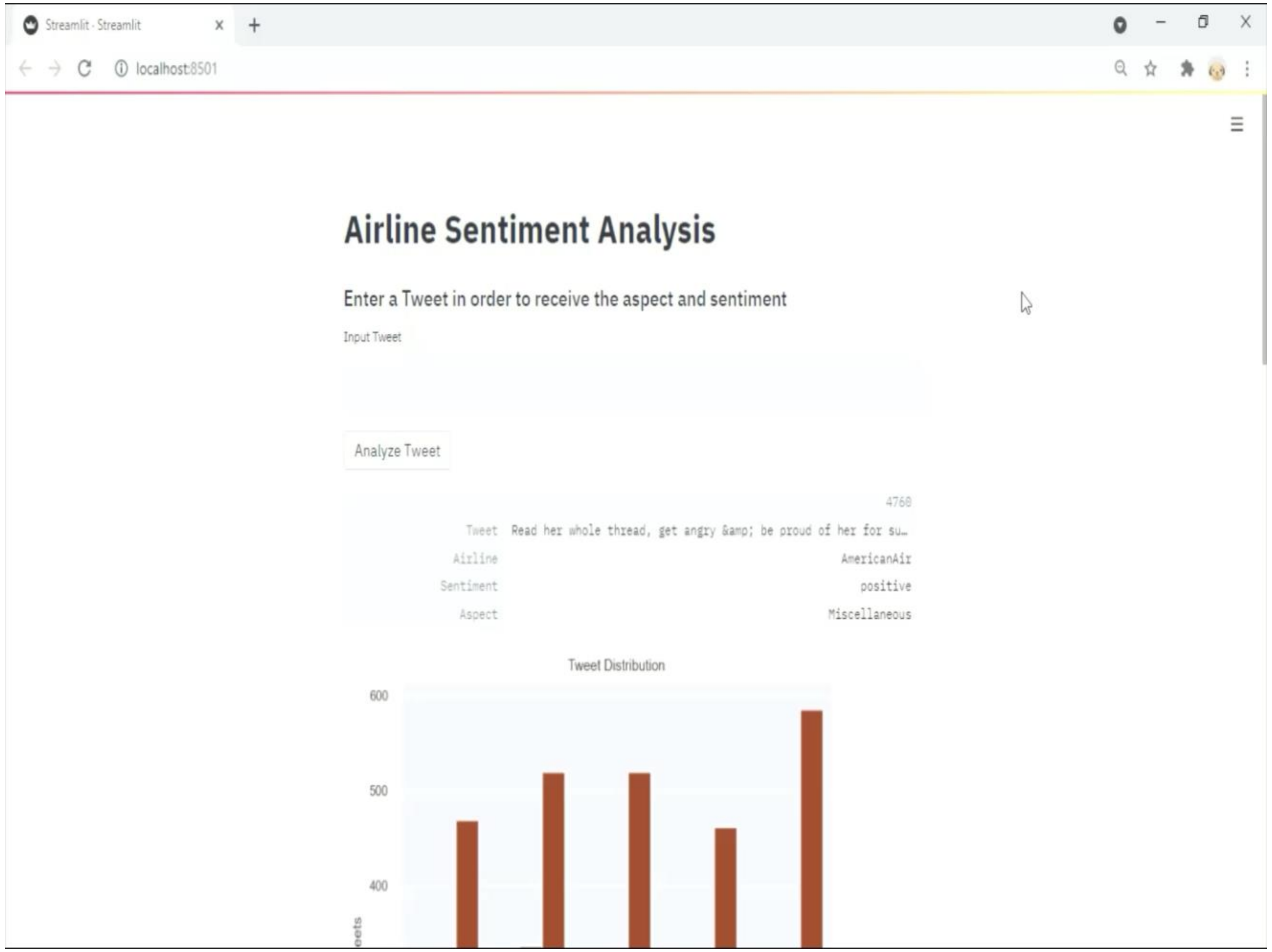
Aspect Analysis for American Air




Tweet Distribution



# Demo Part 2



Streamlit · Streamlit
localhost:8501
Airline Sentiment Analysis
Enter a Tweet in order to receive the aspect and sentiment
Input Tweet
Analyze Tweet
4760
Tweet
Read her whole thread, get angry & be proud of her for su…
Airline
AmericanAir
Sentiment
positive
Aspect
Miscellaneous
Tweet Distribution
600
500
400

# Recommendation

- Fine tuning the hugging face model showed drastic improvements in sentiment detection comparatively to the baseline. If the model we're to deteriorate with time, we are confident that minor tweaks could keep its current accuracy.
- Aspect clustering was not as straight forward and required a lot of manual intervention and as of now would not be sustainable. Could outsource this portion
- When customer service is using our model they can continually add input whether a tweet is correctly classified, adding to our training data over time
- Lastly, we would love to expand our original training dataset to a more current timeframe. With Covid-19 changing travel so drastically there may be new things that aren't covered with our training set.

# What else could we do

- Each tweet being put to more than one cluster, multi-aspect analysis
- When more data is added, incorporate a seasonality affect
- Dynamic Clusters

# Conclusion

## Build a Social Listening Tool

- Inform customer service on customer complaints are trending
- discover new pockets of negative sentiment, or new topics (e.g. masks)
- benchmarking against other companies (i.e. where do we perform better, what should we advertise?