SENTIMENT ANALYSIS OF STARBUCKS REVIEWS

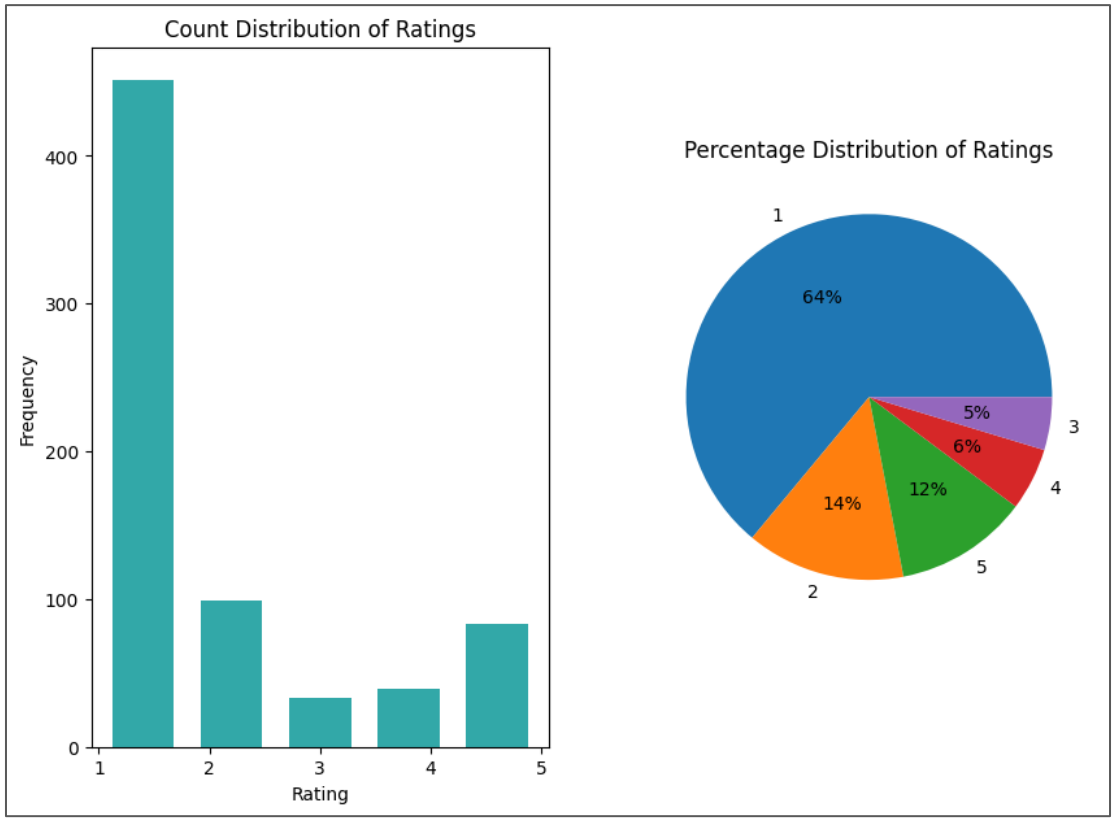
TEMILOLA FAMAKINWA

EXECUTIVE SUMMARY

- Brian Niccol became **Starbucks' CEO** in Sept 2024
- Tracking **customer sentiment** in the first 12 months can give an **indication** of **impact of Niccol's strategy and vision**.
- Developed a model for predicting customer sentiment from reviews.
- Achieved **81% accuracy** in prediction of customer sentiment using a recurrent neural network model.
- Potential areas of improvement: **customer service**, **wait times**, **payment methods and drive thru** service.

DATA

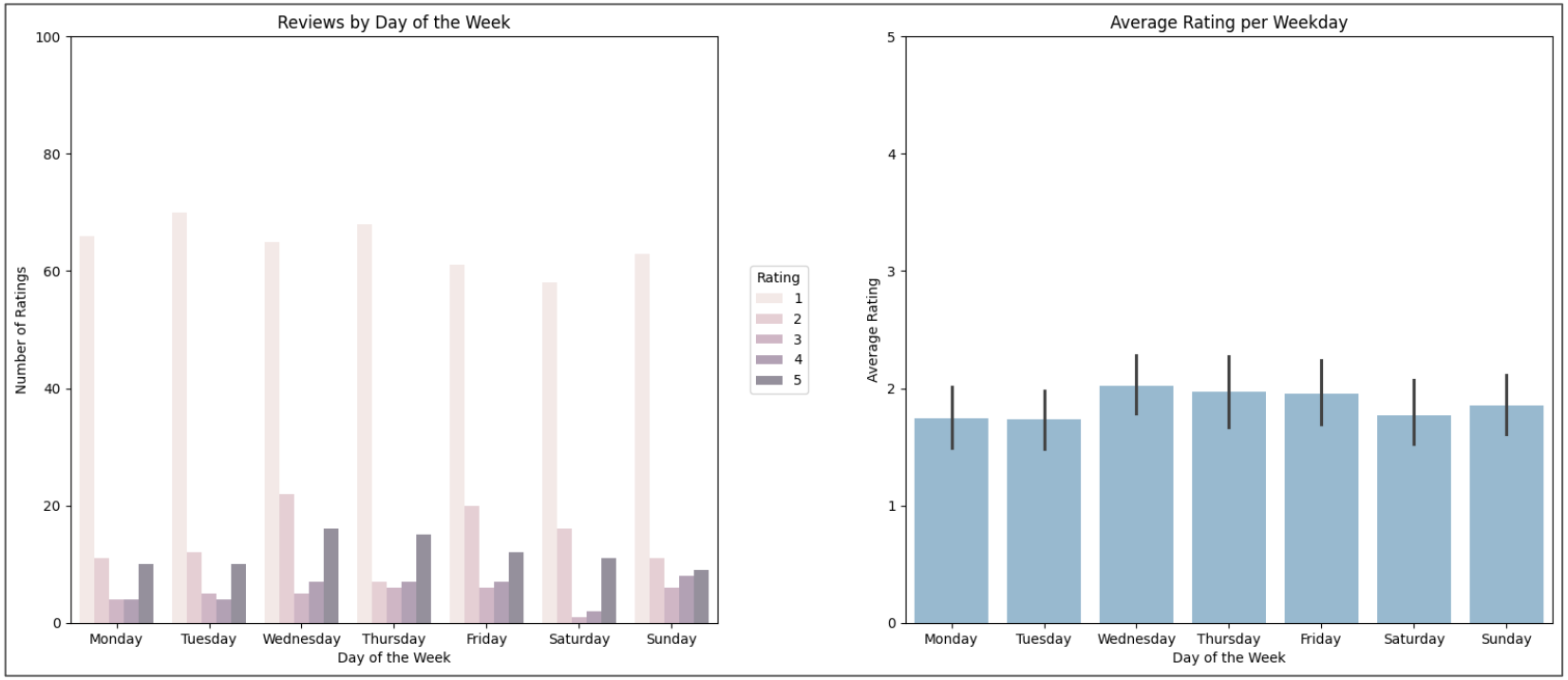
- Starbucks Yelp Review (2009 – 2023)[1]
- 850 entries
- CSV Format
- **Imbalance**: 78% of reviews are negative
  - address during modelling.



[1] https://www.kaggle.com/datasets/harshalhonde/starbucks-reviews-dataset

CUSTOMER RATINGS BY WEEKDAY

Customer ratings are **lowest** at the **beginning** of the week and **highest in the middle to end** of the week.

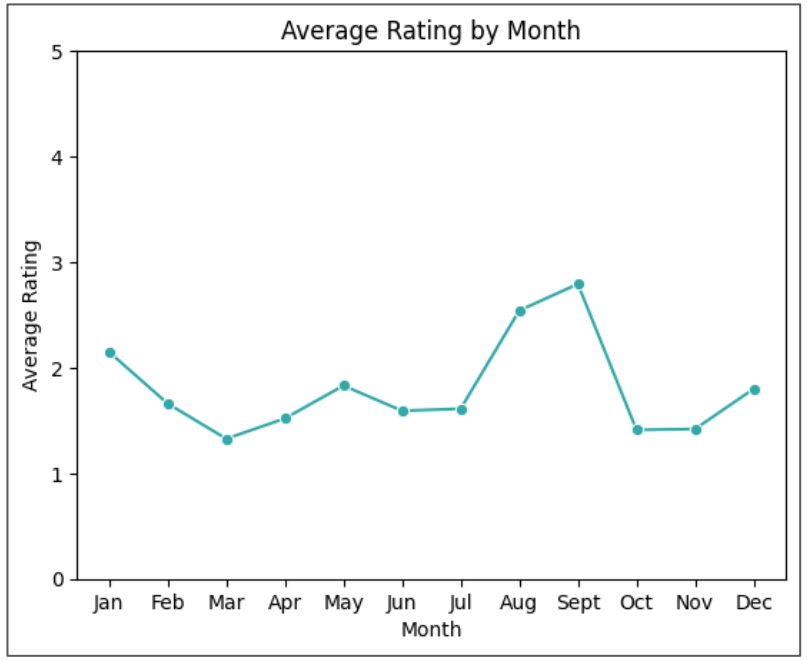


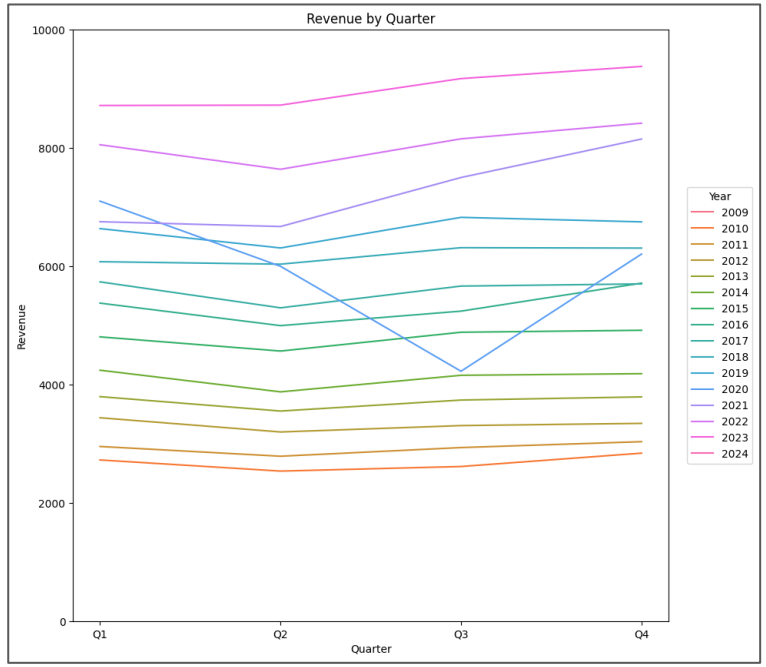
Do customers have a less optimistic outlook at the beginning of workweek?

CUSTOMER RATINGS BY MONTH

On average, **customer ratings peak** between **August** and **September** during the year.

Revenue[1] shows **uptick during Q3 - Q4** (fall/holiday season) annual.



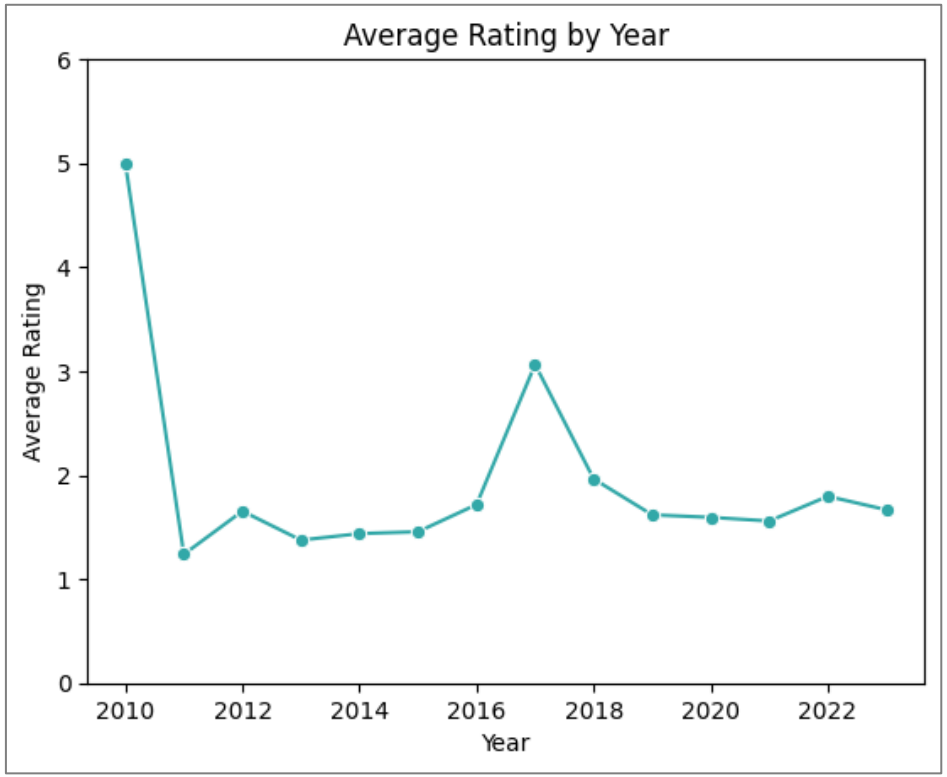


Are ratings in August-September driven by positive outlook associated with the start of the fall/holiday season?

[1]https://www.netcials.com/financial-revenue-growth-usa/829224-STARBUCKS-CORP/

CUSTOMER RATINGS BY YEAR

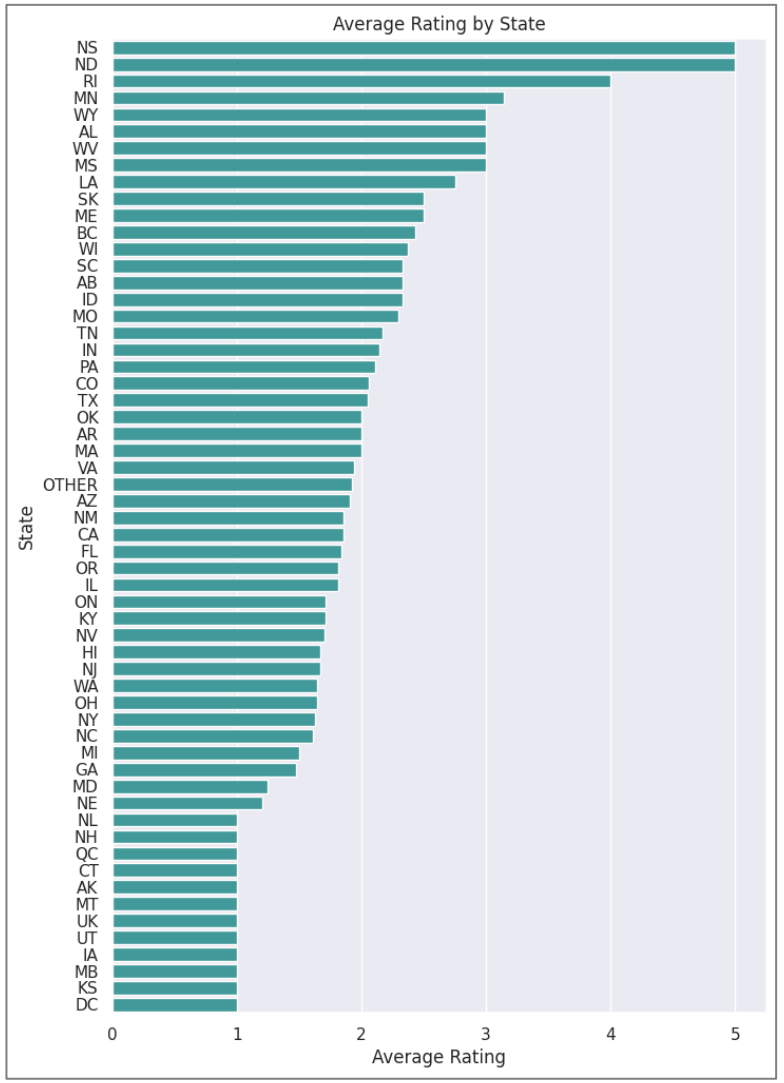
Customer ratings had a noticeable **increase in 2017** compared to other years



What programs and initiatives did Starbucks have during 2017?

CUSTOMER RATINGS BY US/CANADIAN STATES & PROVINCES

Highest ratings were recorded in **Nova Scotia**, **North Dakota,** and **Rhode Island**.



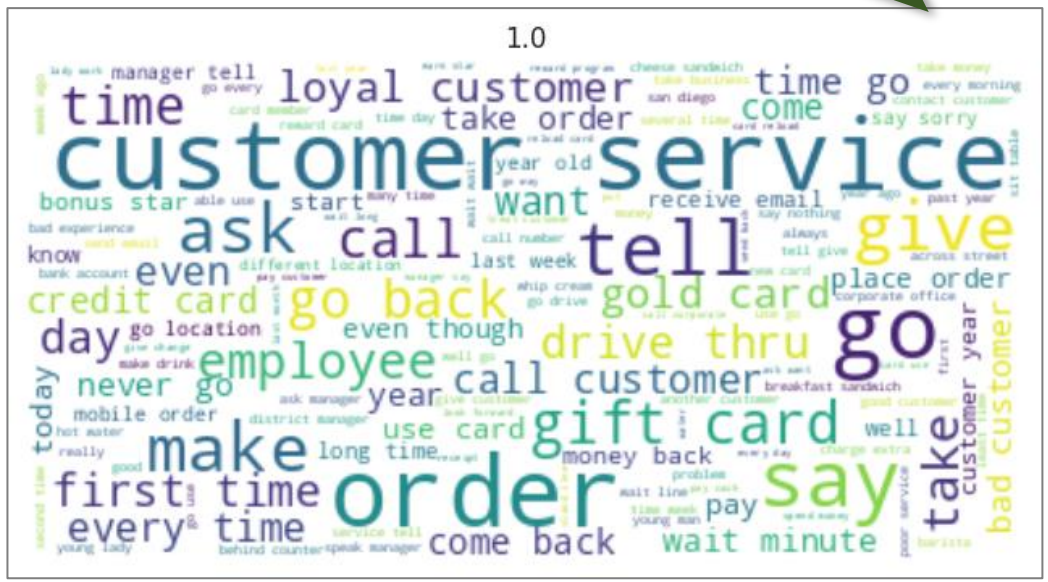
Could these ratings be correlated to regional groupings?

When things go right…

- Good, ‘Professional’ ‘customer service’
- Drink ‘variety’, ‘Good food’
- ‘Atmosphere’, ‘Feels like home’
- ‘Allergy’ (well taken care of?)

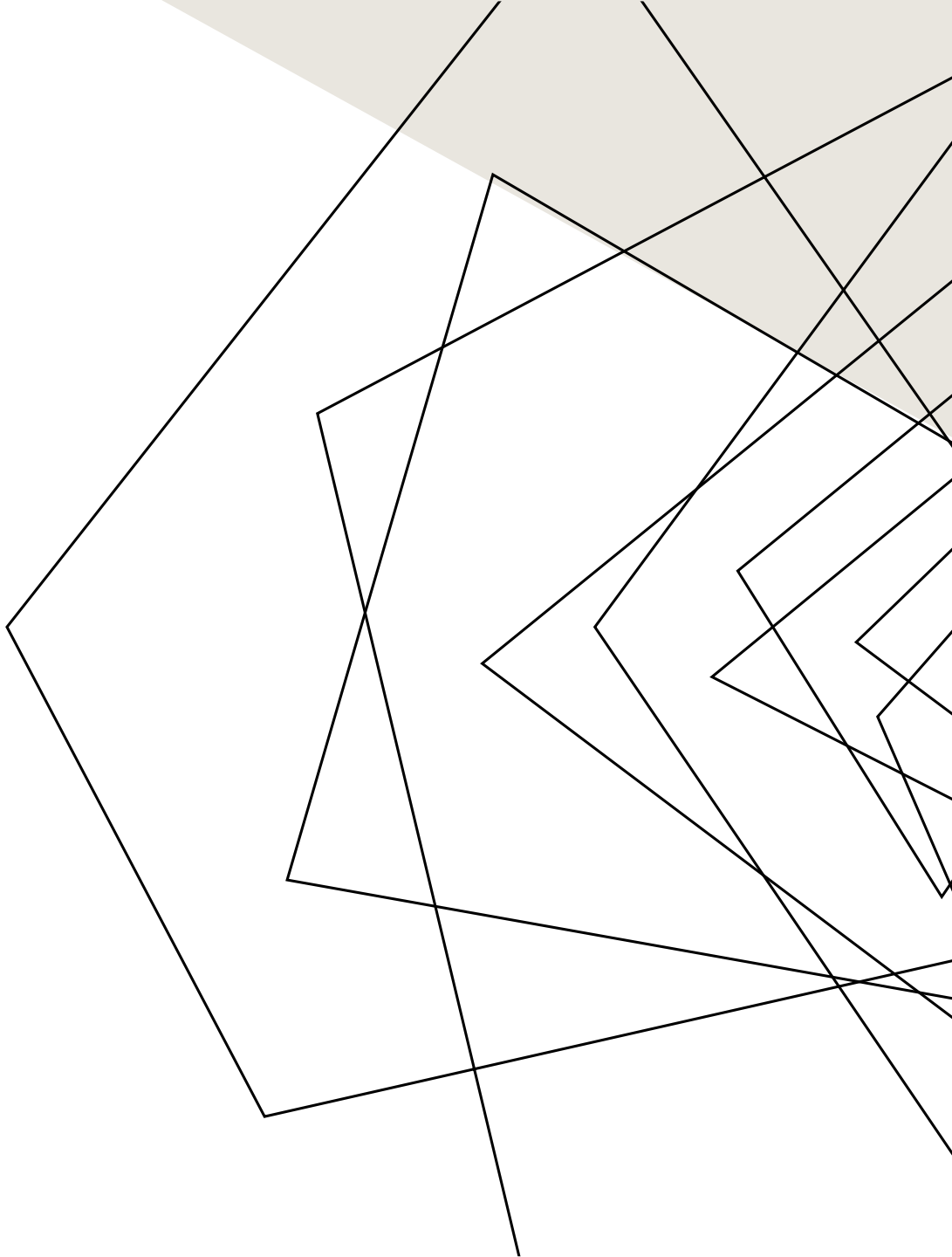
When improvements need to be made…

- ‘Poor’ ‘customer service’
- ‘Long’ ‘ Wait’ ‘Line’ ‘Time’
- ‘Bank account’ ‘ Credit Card’ ‘ Gift card’ ‘Mobile order’ ‘Pay’ ‘Reward card’ ‘Gold card’
- ‘Drive thru’



MODELLING

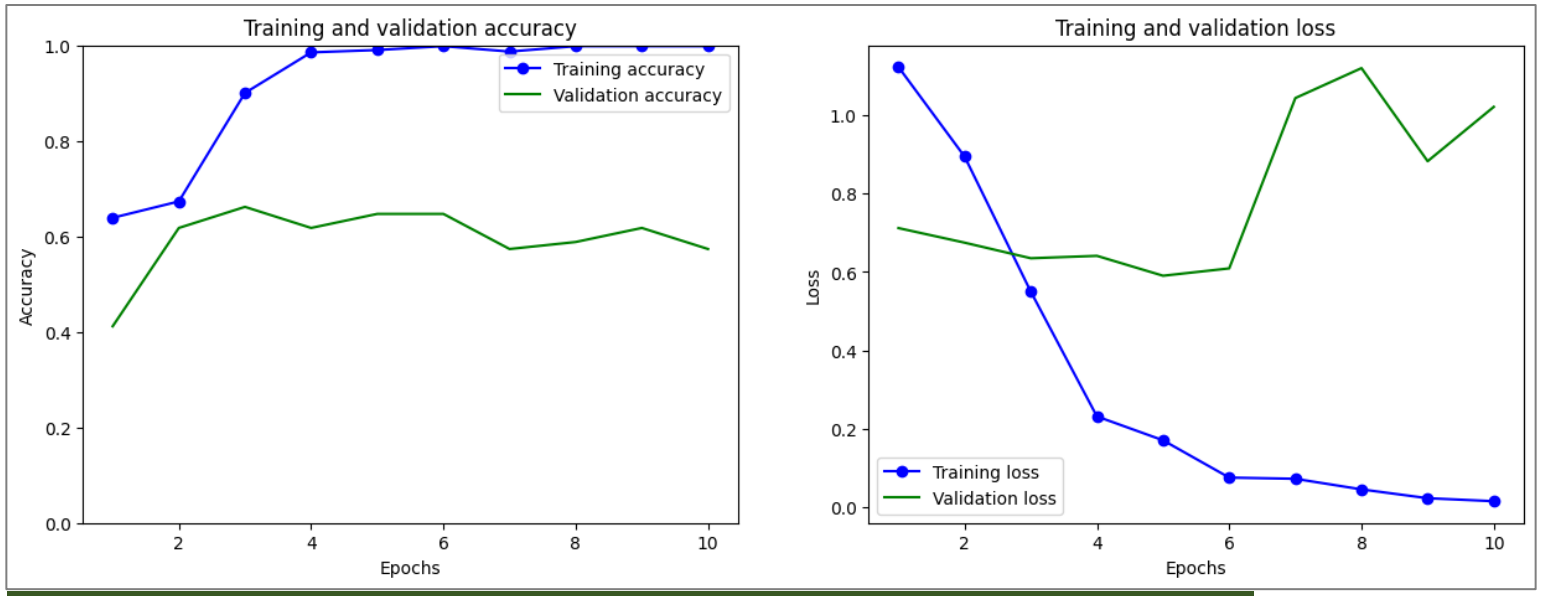
- Recurrent neural networks (RNNs) which are designed to deal with sequential data were used.
    - Simple RNN
    - Long Short-Term Memory (LSTM)
    - Transformer based model (BERT) for transfer learning
- Review ratings: 1-2 = Negative (0), 3-5 = Positive (1)
- Class weights were set during the model fitting step to address class in-balance.
- Hyperparameter tuning was conducted on the best model.

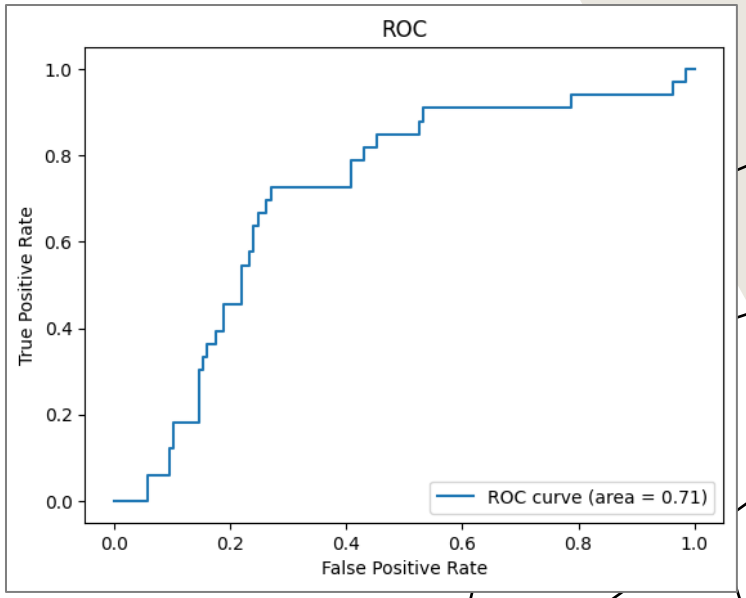


SIMPLE RNN MODEL

- **Training accuracy:** 99%
- **Test accuracy:** 66 %
- **Area under the Receiver Operator Characteristic Curve (ROC) (AUC):** 0.71

```
Model: "Simple_RNN"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 embedding_2 (Embedding)     (None, None, 32)          143232

 simple_rnn_1 (SimpleRNN)    (None, 64)                6208

 dropout_2 (Dropout)         (None, 64)                0

 dense_2 (Dense)             (None, 1)                 65

=================================================================
Total params: 149505 (584.00 KB)
Trainable params: 149505 (584.00 KB)
Non-trainable params: 0 (0.00 Byte)
```



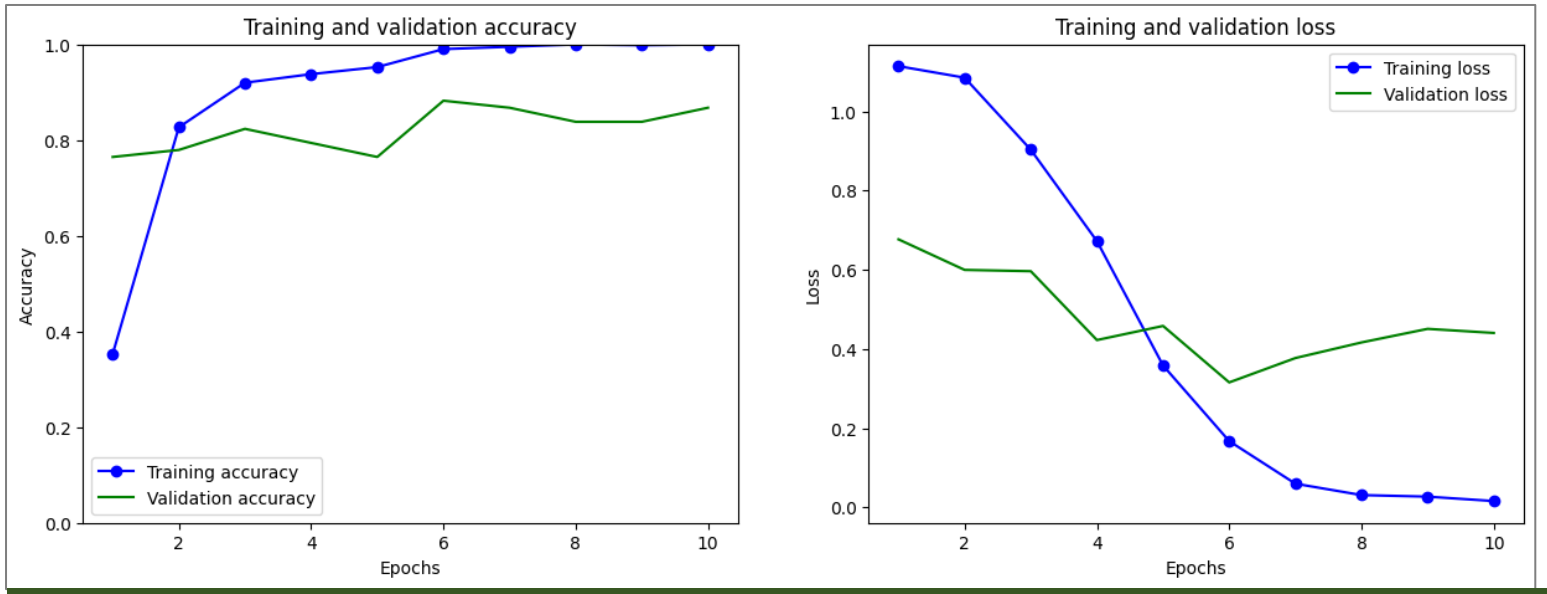


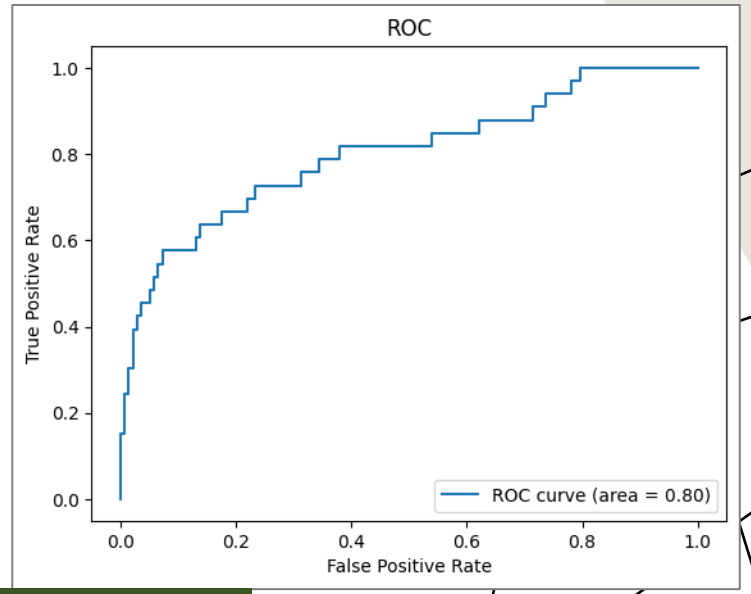
Model seems to be overfitting the training data.

LSTM MODEL

- **Training accuracy:** 100%
- **Test accuracy:** 0.85 %
- **Area under the Receiver Operator Characteristic Curve (ROC) (AUC):** 0.80

```
Model: "LSTM"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 embedding_3 (Embedding)     (None, None, 32)          143232

 bidirectional_1 (Bidirecti  (None, 128)               49664
 onal)

 dropout_3 (Dropout)         (None, 128)               0

 dense_3 (Dense)             (None, 1)                 129

=================================================================
Total params: 193025 (754.00 KB)
Trainable params: 193025 (754.00 KB)
Non-trainable params: 0 (0.00 Byte)
```



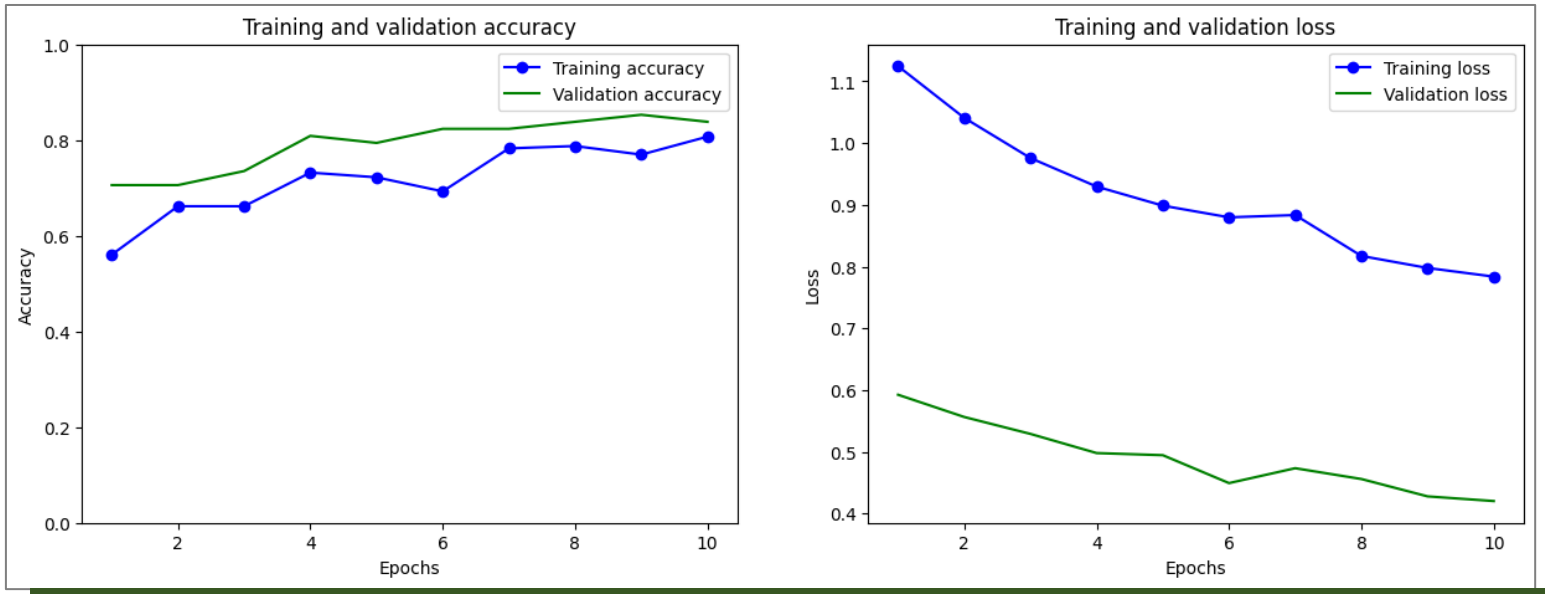


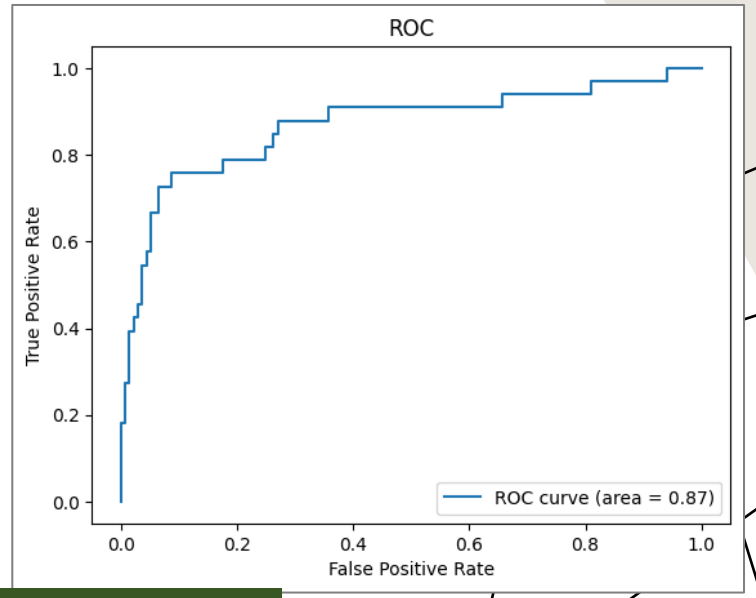
Model is an improvement over Simple RNN but still shows overfitting behavior.

TRANSFER LEARNING MODEL

- **Training accuracy:** 81%
- **Test accuracy:** 85 %
- **Area under the Receiver Operator Characteristic Curve (ROC) (AUC):** 0.87

```
Model: "model"
__________________________________________________________________________________________________
 Layer (type)                Output Shape              Param #   Connected to
==================================================================================================
 review_input (InputLayer)   [(None,)]                 0         []

 preprocessing (KerasLayer)  {'input_mask': (None, 128)  0        ['review_input[0][0]']
                             , 'input_type_ids': (None,
                              128),
                              'input_word_ids': (None,
                             128)}

 BERT_encoder (KerasLayer)   {'encoder_outputs': [(None  28763649  ['preprocessing[0][0]',
                             , 128, 512),                           'preprocessing[0][1]',
                              (None, 128, 512),                     'preprocessing[0][2]']
                              (None, 128, 512),
                              (None, 128, 512)],
                              'default': (None, 512),
                              'sequence_output': (None,
                             128, 512),
                              'pooled_output': (None, 5
                             12)}

 dropout (Dropout)           (None, 512)               0         ['BERT_encoder[0][5]']

 classifier (Dense)          (None, 1)                 513       ['dropout[0][0]']

==================================================================================================
Total params: 28764162 (109.73 MB)
Trainable params: 513 (2.00 KB)
Non-trainable params: 28763649 (109.72 MB)
```





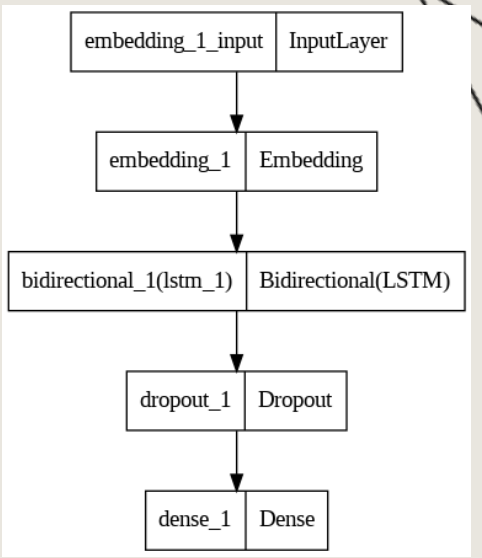
Model seems to be showing less overfitting than both LSTM and Simple RNN.

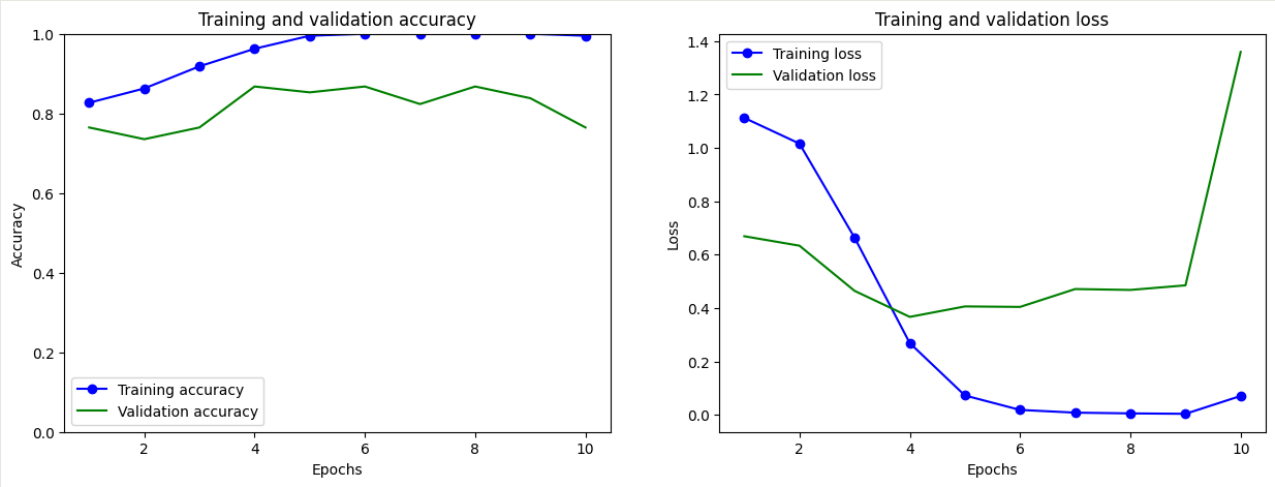
HYPERPARAMETER TUNING

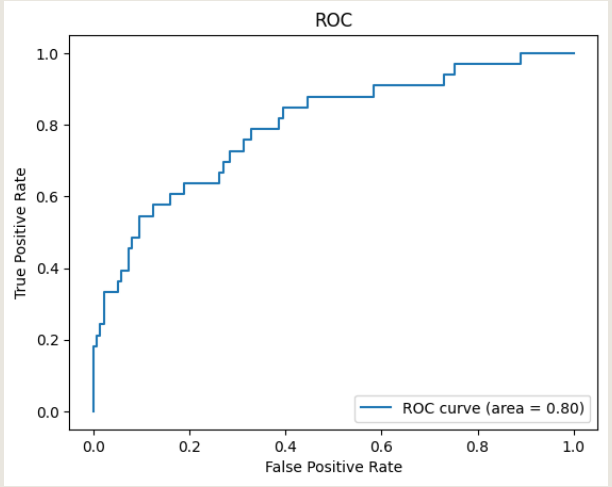
- LSTM model was selected as the best model for further analysis because it:
  - achieved above 80% accuracy target
  - is a less complex and smaller model (in size) than the BERT model
- Hyperparameters:
  - Dropout rate
  - Learning rate

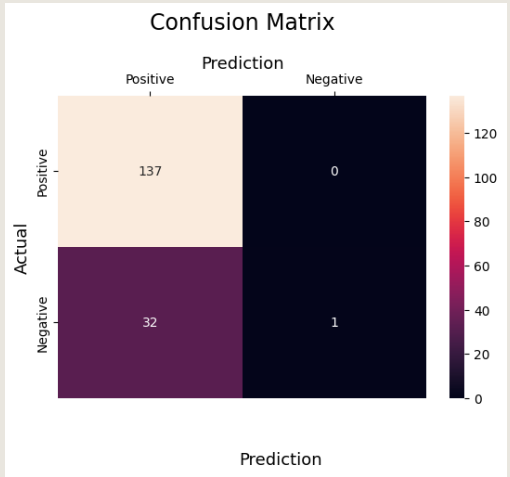
HYPERPARAMETER TUNING RESULTS

- **Best parameters:** Dropout rate: 25%, Learning rate: 0.002
- **Training Accuracy:** 99.5%
- **Test Accuracy:** 81.1%





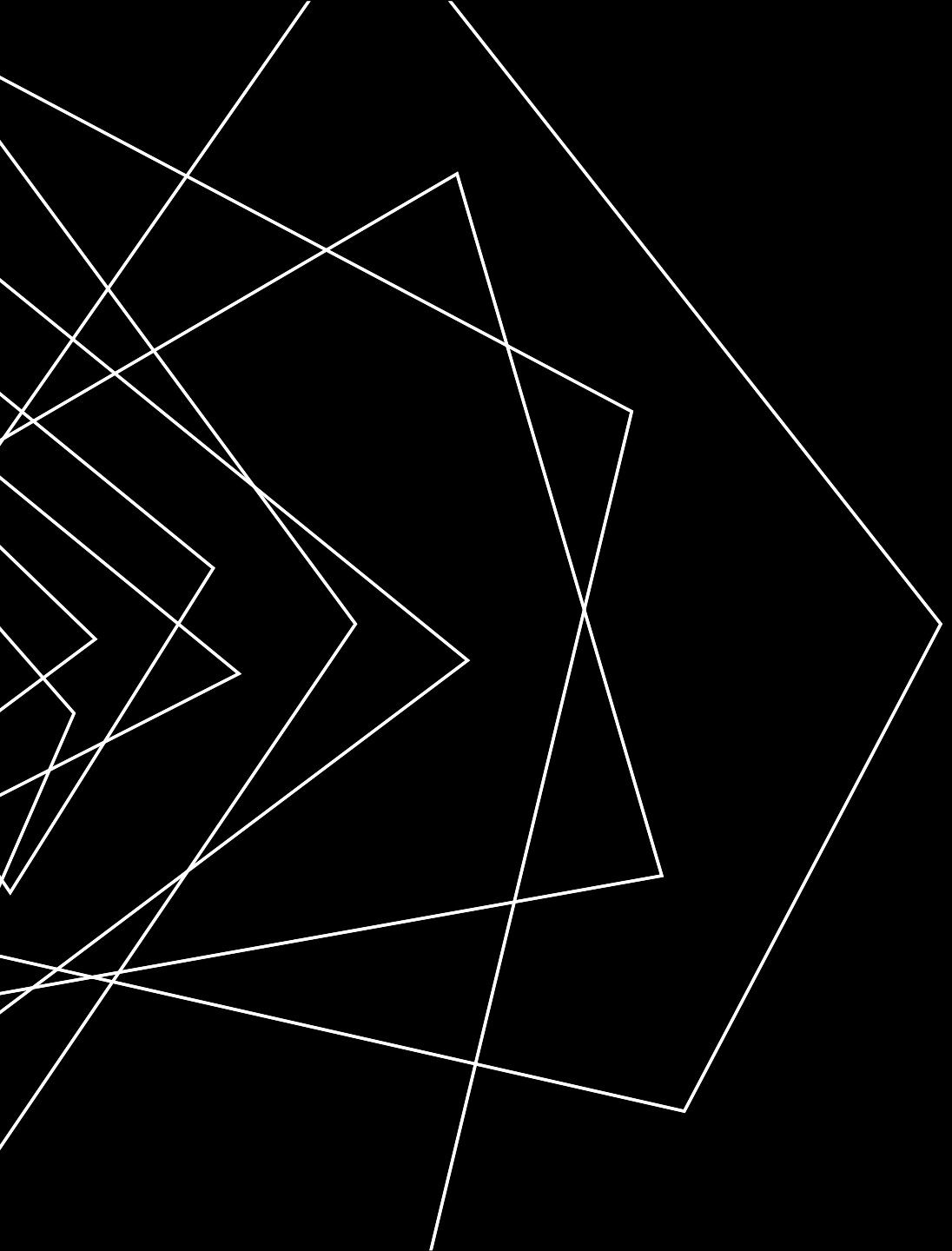




Model seems to show signs of overfitting. Rate of predictions for the positive and negative classes are 3% and 100%, respectively. Imbalance should be addressed during hyperparameter tuning.

CONCLUSION

- Business recommendations:
  - Investigate if marketing campaign addressing start-of-week blues could improve ratings, sales at Starbucks.
  - Design fall/holiday focused menus that leverage general optimism of the seasonal change.
  - Dig into campaigns/events that were run in 2017 to see if they truly drove increased sales; consider reusing them.
  - Create surveys for store employees and customers to collect feedback to better understand if there are opportunities for improvement in **customer service**, **wait times**, **payment methods and drive thru** service.
- The LSTM model predicted customer sentiment with **81% accuracy**.
  - It is likely overfit.
  - Continue to monitor how it performs on new review data.
  - Consider using the **transfer learning model** instead to improve performance on unseen data.



THANK YOU

Temilola Famakinwa
tefamakinwa@gmail.com
www.linkedin.com/in/famakinwatemilola/