05.20.25

CANVA CORPORATION

Natural Language Processing

Tweet Analysis

Agenda

01 Introduction

02 Data

03 Modeling

04 Recommendations and Conclusions

05 Next Steps

01

Introduction

CANVA CORPORATION

Tweet Analysis

KBO Marketing Company is a marketing firm that wants a model that can predict sentiment for technology products such as smartphones. The C-suite executives task their in-house data scientists to develop a prototype

Modeling……..

Completed the following:

- Leveraged tweets targeted at Apple and Google
- Two categories: *Not Positive* and *Positive*

02

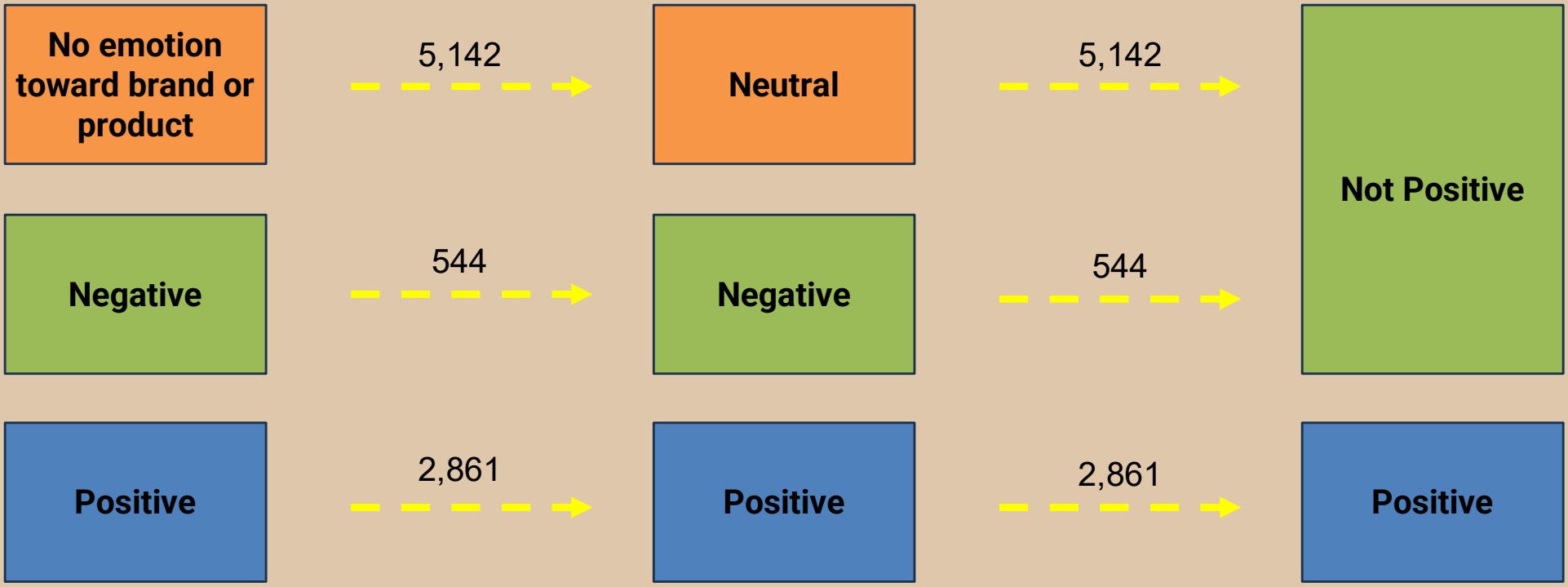
Data

CANVA CORPORATION

Data Description

- csv file
  - Initially 8,721 Observations, or Tweets
  - 3 Columns
- Columns are the following:
  - Tweet
  - Apple / Google Brand or Product (i.e. – *iPad, iPhone App, Android*)
  - Emotion

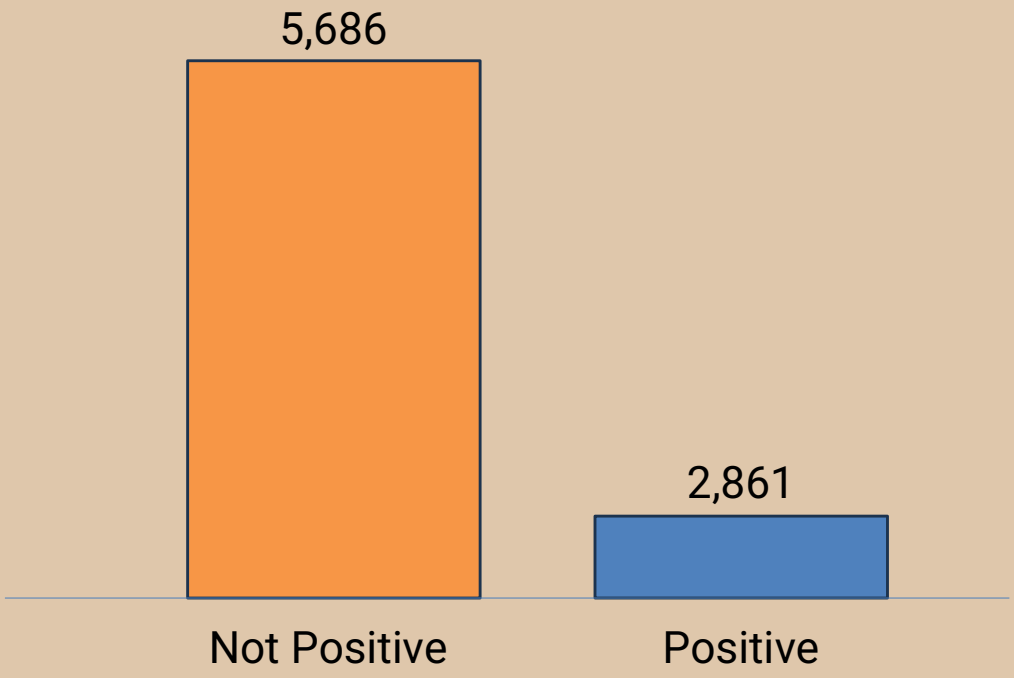
Data Description



CANVA CORPORATION

Data Description

- Class Imbalance



03

Modeling

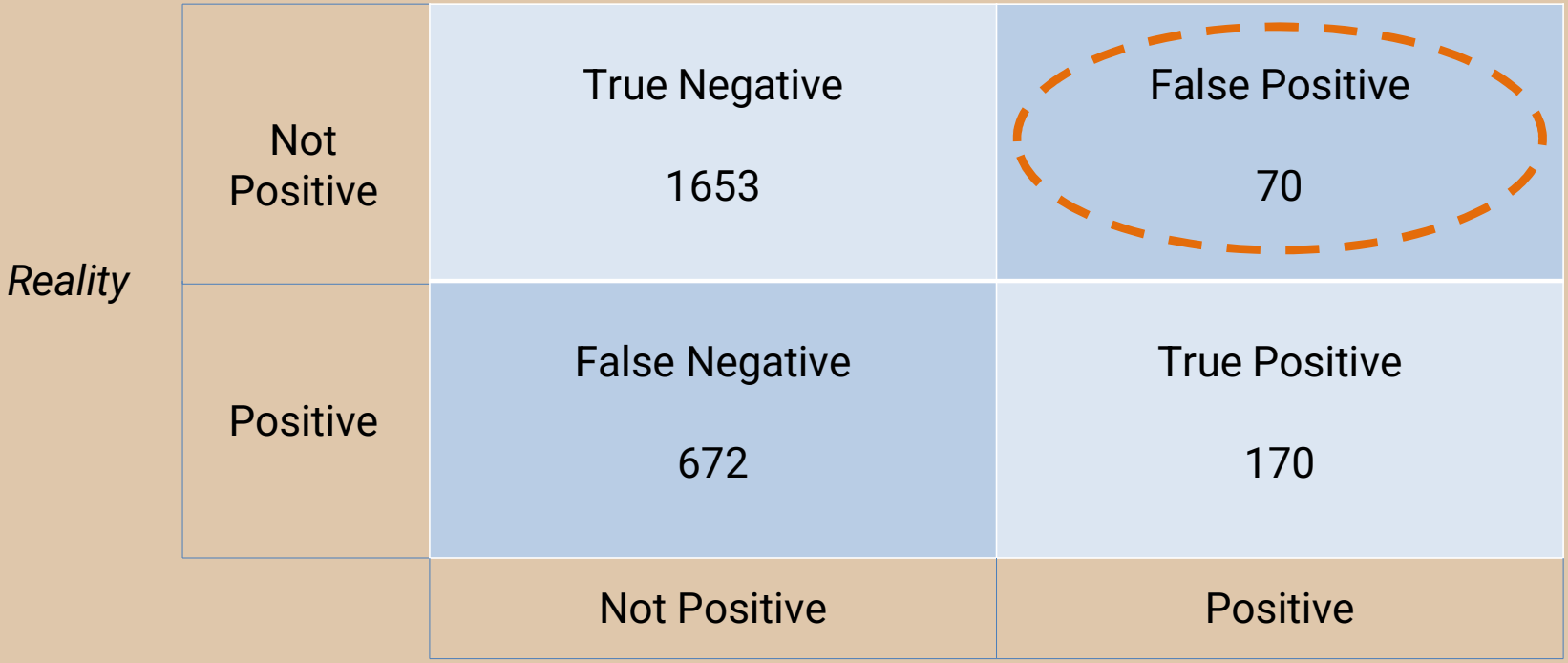
CANVA CORPORATION

Naive Bayes

- What is Naïve Bayes?
  - Algorithm based on Bayes Theorem

CANVA CORPORATION

Confusion Matrix



04

Recommendations and Conclusions

CANVA CORPORATION

Recommendations and Conclusions

- Current Model
  - ≈70% Precision
  - Based on 8,547 Customers

05

Next Steps

CANVA CORPORATION

Next Steps

1 – Public Online Communities

2 – Other Smartphone and Tech Companies

3 – Private Online Communities

06

Appendix

CANVA CORPORATION

Appendix

- Serrano, Luis (2022). “Using Probability to its Maximum: The naïve Bayes model”
- Liang, David (2024). “Intro – Probability in Python: Coin Toss, Dice, and Poker Explained”

Model Precision

- ≈70.8%
- How is Precision Calculated?

$$= \frac{\text{Number of True Positives}}{\text{Number of Positives Predicted by the Model}}$$

$$= \frac{170}{240}$$