



SENTIMENT ANALYSIS WITH MNB

Benjamin Jaron

OVERVIEW

We build a model for SXSW that enables them to predict positive sentiment with a 73% recall score. This will help SXSW fortify their strengths.

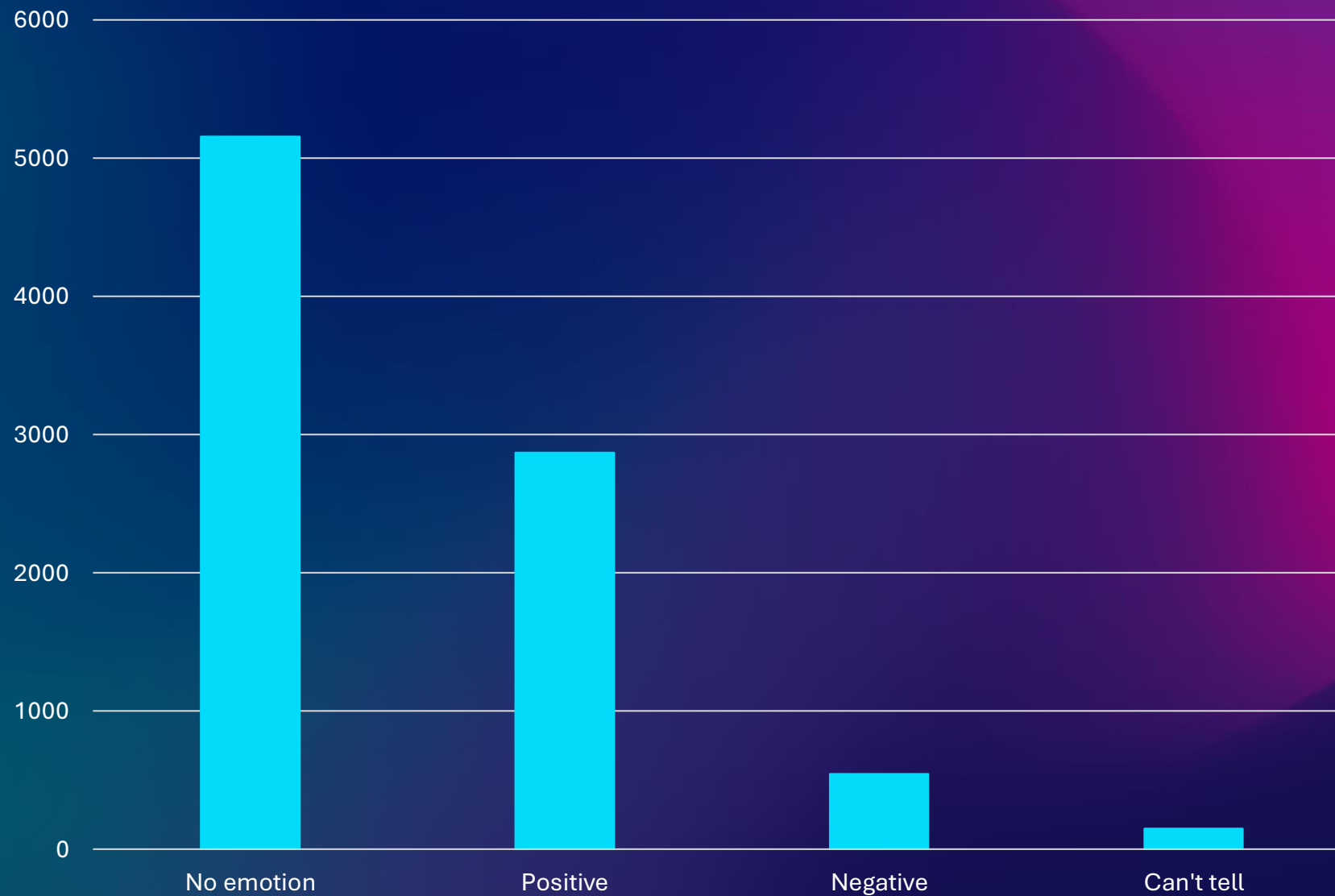




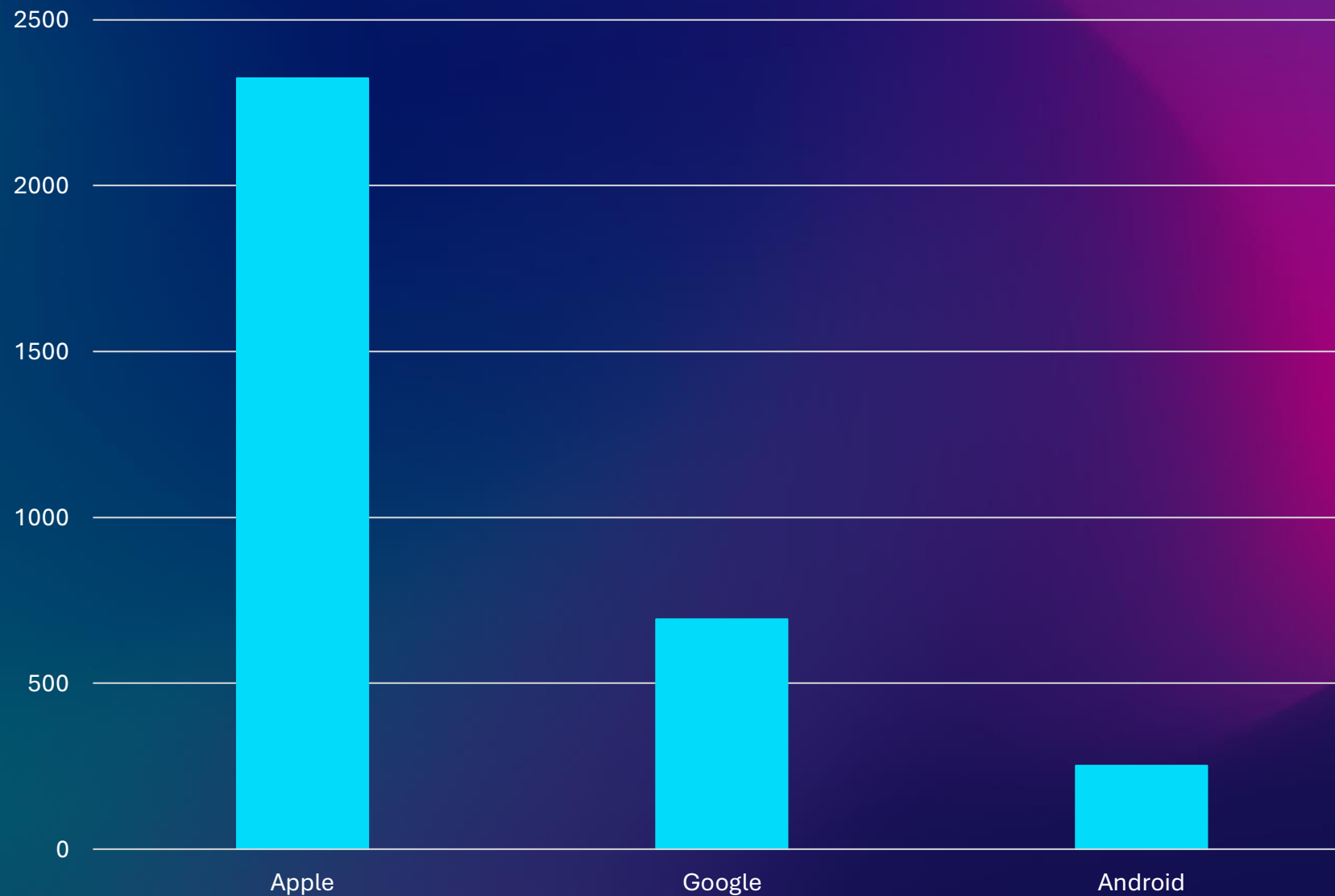
BUSINESS PROBLEM

- Data set of 8,721 tweets from the SXSW tech conference
- Need to provide actionable insights from positive tweets
- Provide model to classify positive tweets

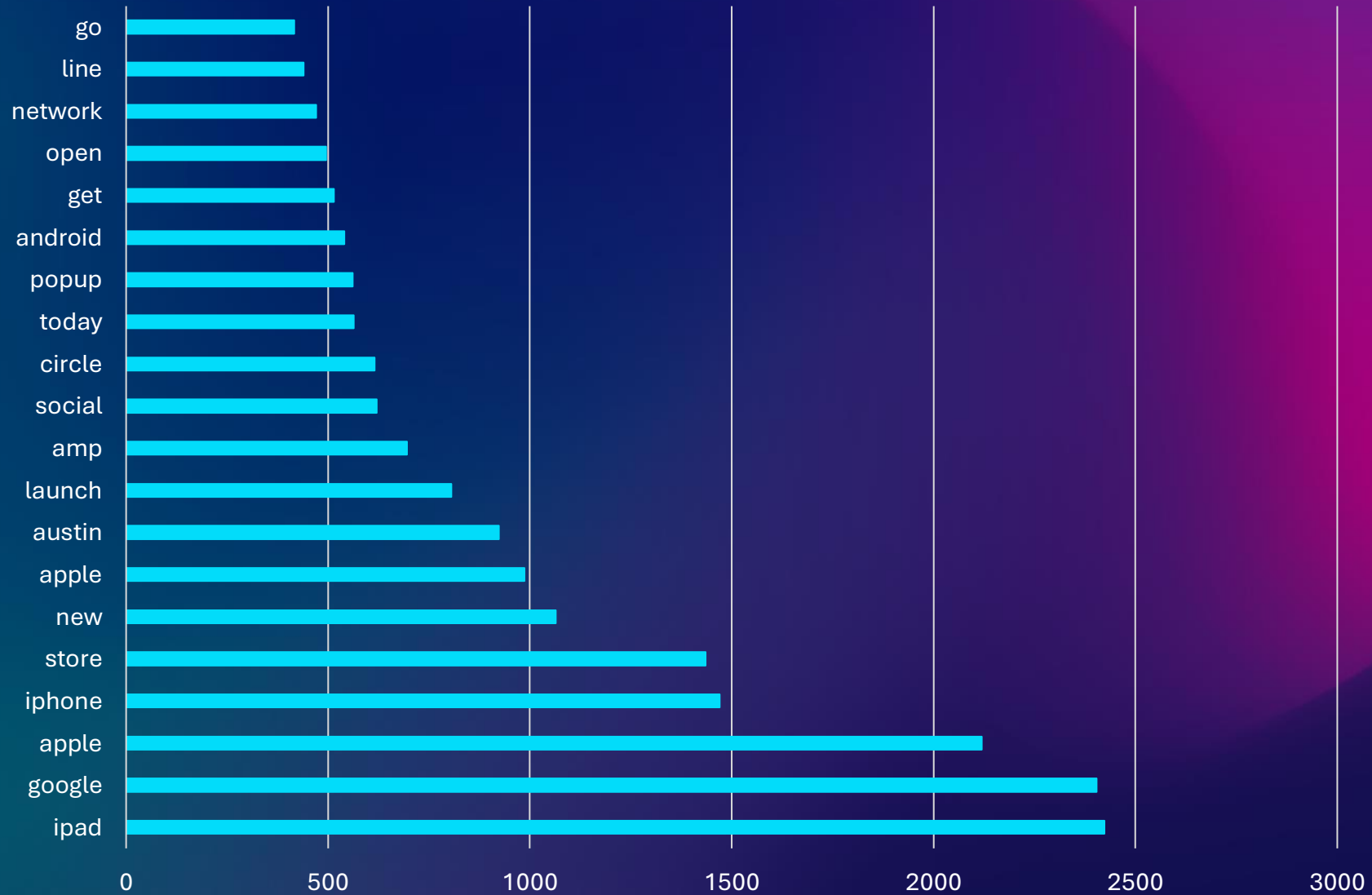
DATA IMBALANCE



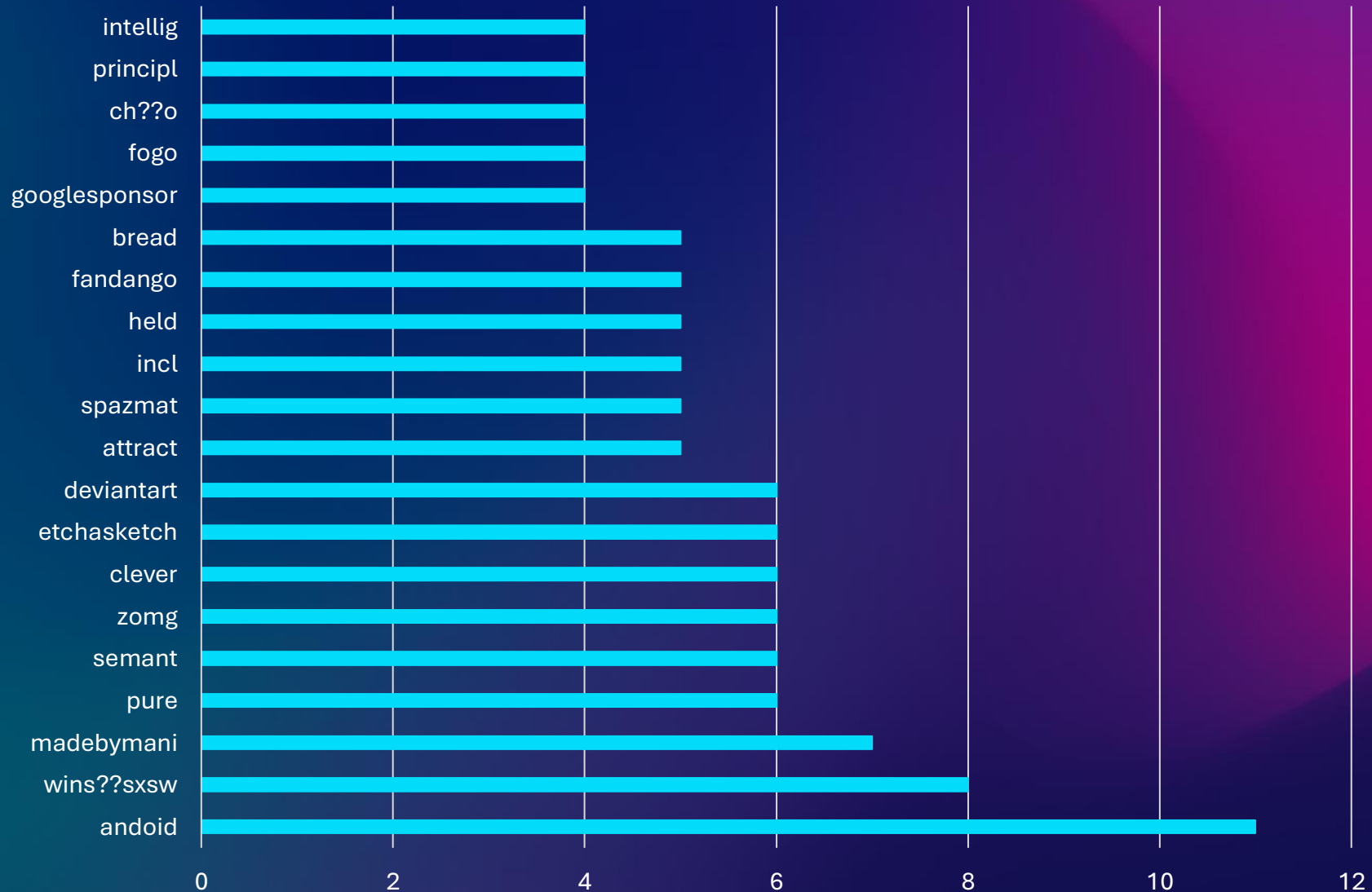
WHERE ARE MOST TWEETS DIRECTED?



MOST COMMON WORDS



MOST COMMON UNIQUE POSITIVE WORDS



RECALL: TOTAL AMOUNT OF TRUE POSITIVES

True Positive Our model correctly predicts a positive tweet	False Positive Our model predicts that a tweet is positive, but it really is not
False Negative Our model predicts that a tweet is not positive, but it really is positive	True Negative Our model correctly predicts that a tweet is not positive

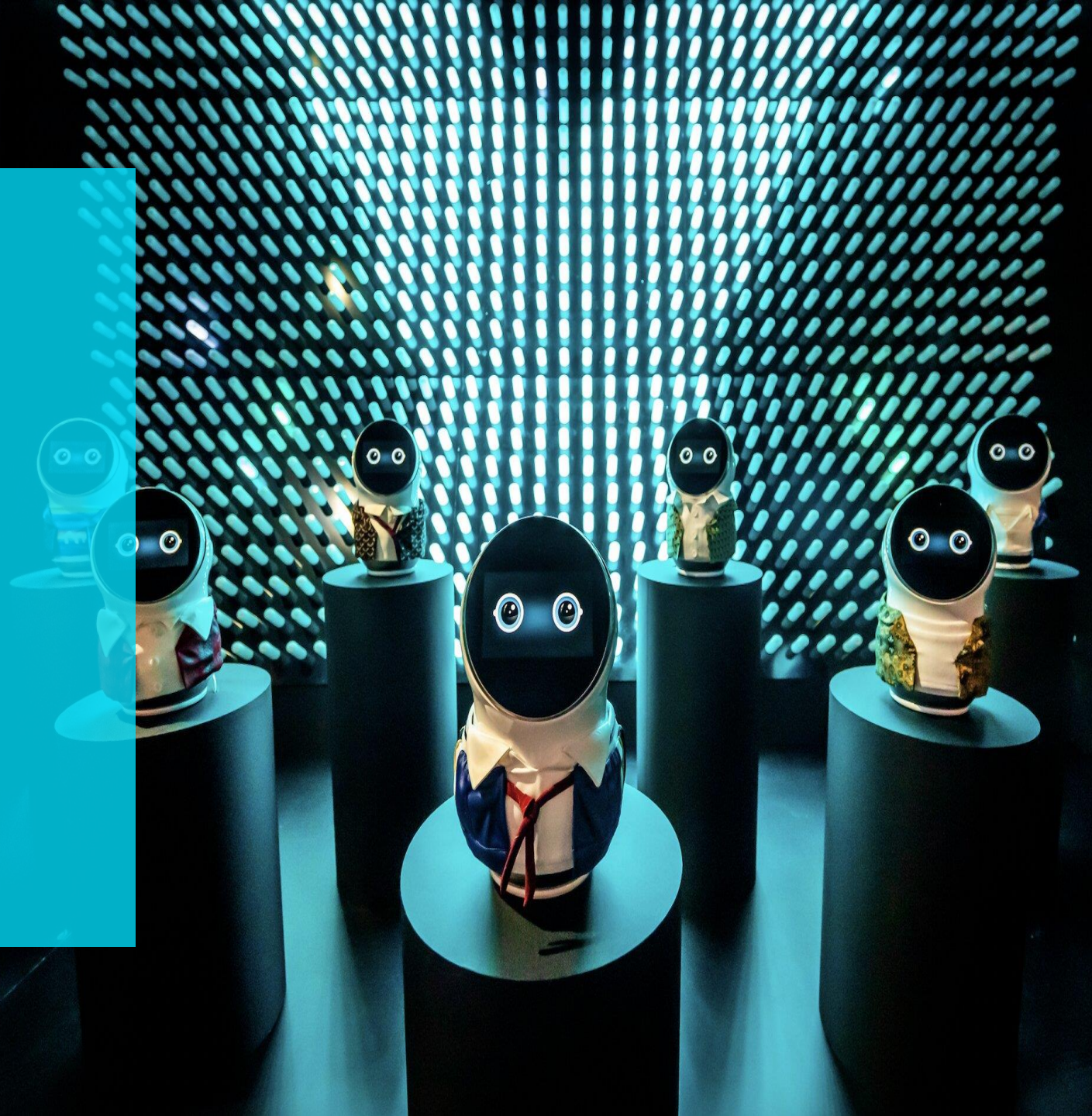
RECALL: TOTAL AMOUNT OF TRUE POSITIVES

True Positive Our model correctly predicts a positive tweet	False Positive Our model predicts that a tweet is positive, but it really is not
False Negative Our model predicts that a tweet is not positive, but it is positive	True Negative Our model correctly predicts that a tweet is not positive

OUR MODEL

73%

We can predict recall with a
score of 73%



THANK YOU

Benjamin Jaron

bmjaron@gmail.com