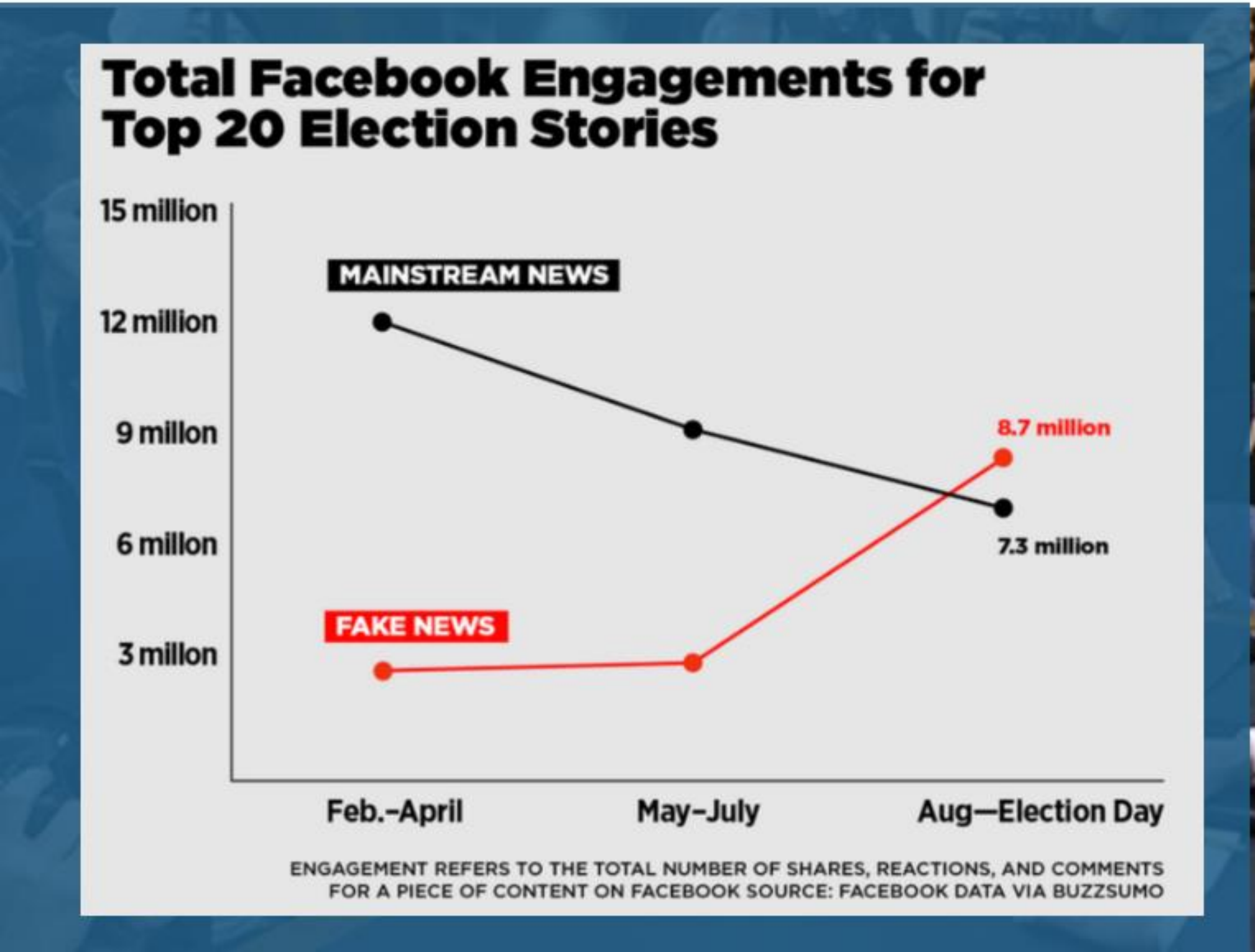
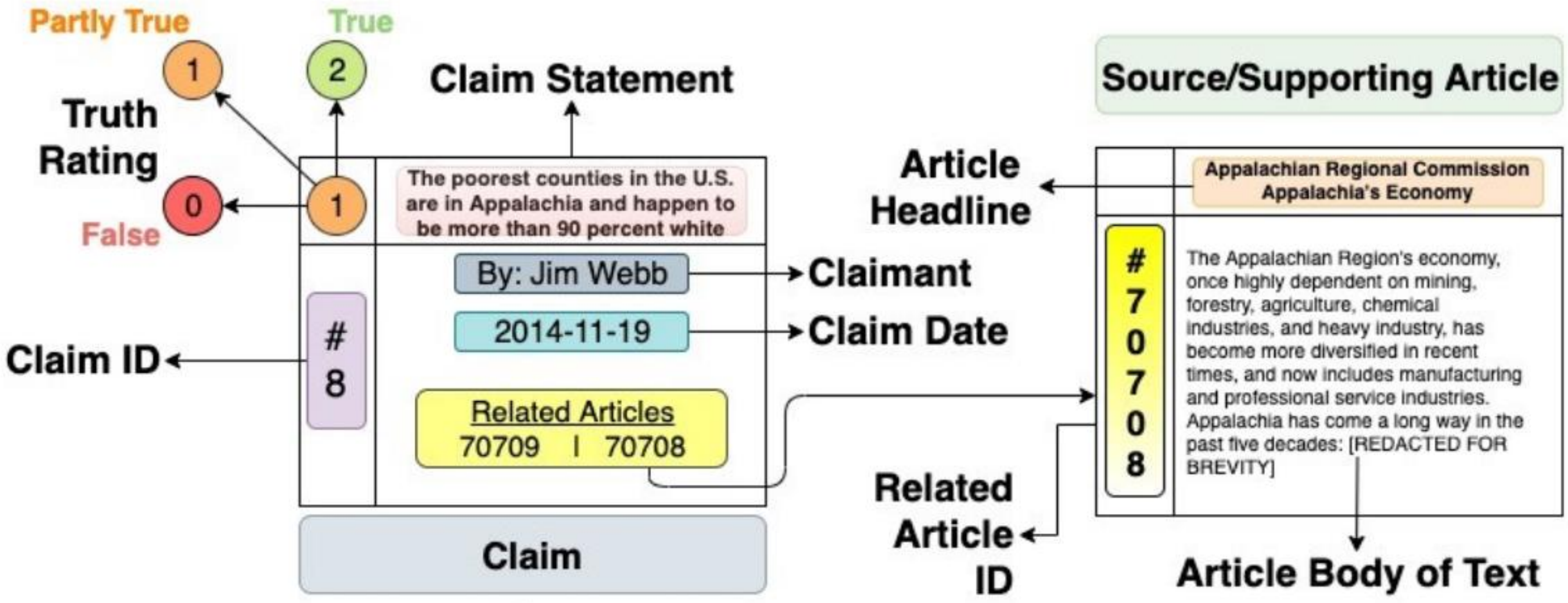
Fact or Fake News: Developing a Fact Checking Algorithm

Prepared by Group #5

Hijun Seo, Jeong Cheol Seok, Meji Lee, Mickole Mulano, Negar Balaghi

INTRODUCTION





Leaders Prize
Creating the next generation of technology innovators

PROBLEM IDENTIFICATION

1

Effectiveness

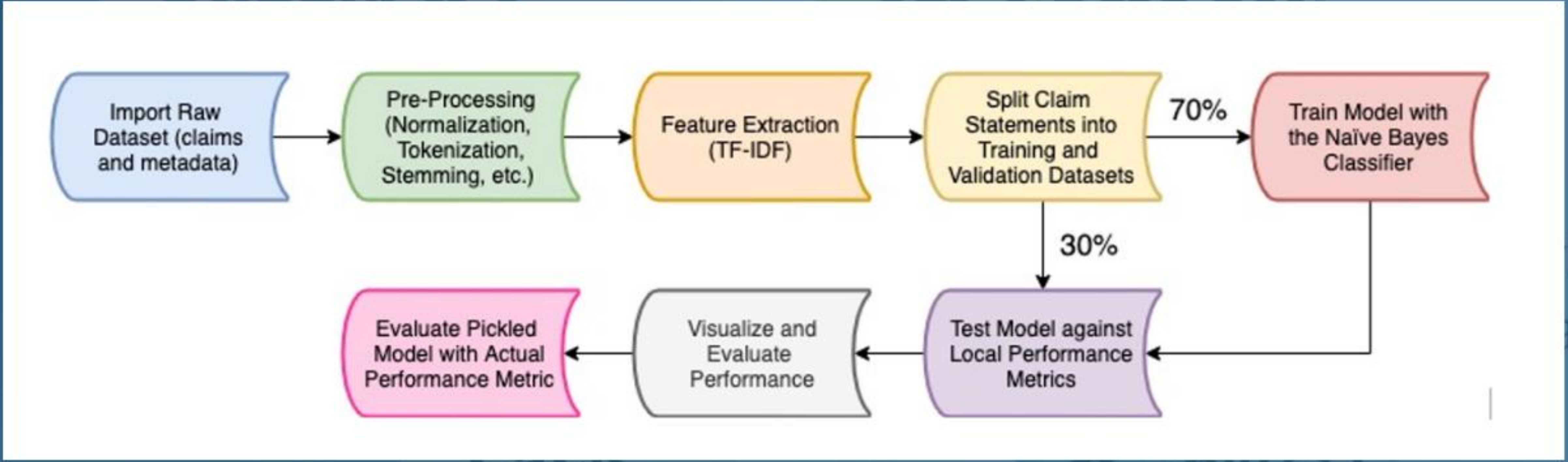
Robust and high accuracy as measured by the weighted average of the precision and sensitivity of the predictions of truth ratings for each claim

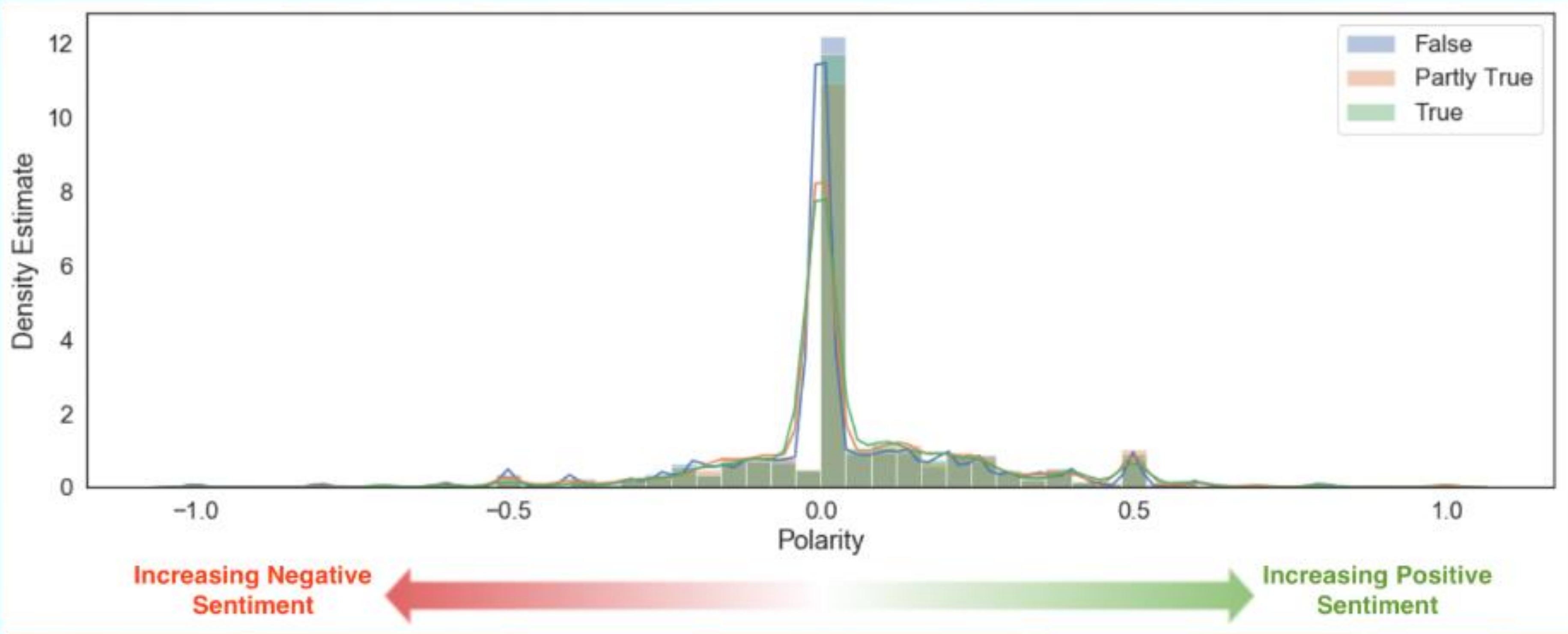
2

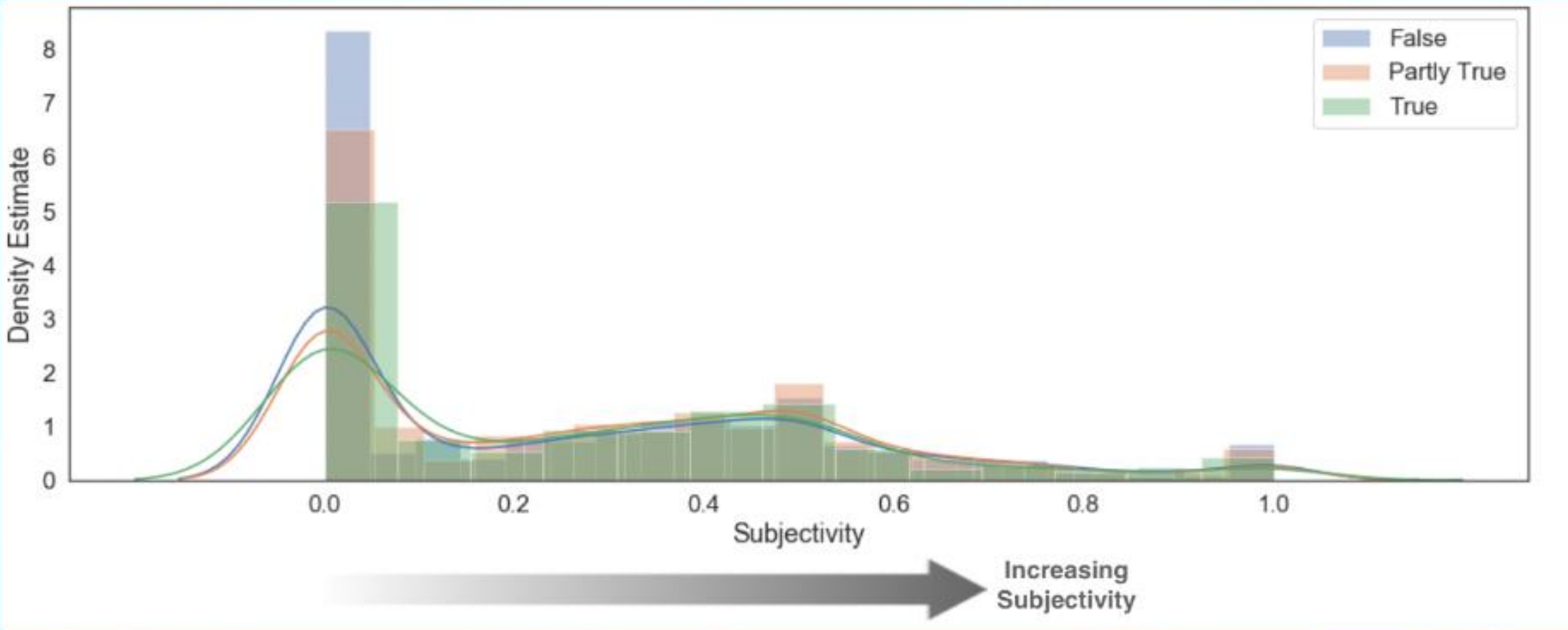
Efficiency

Must provide a robust output of the predictions of truth ratings for each claim in the shortest amount of time as possible.

DATA ANALYSIS



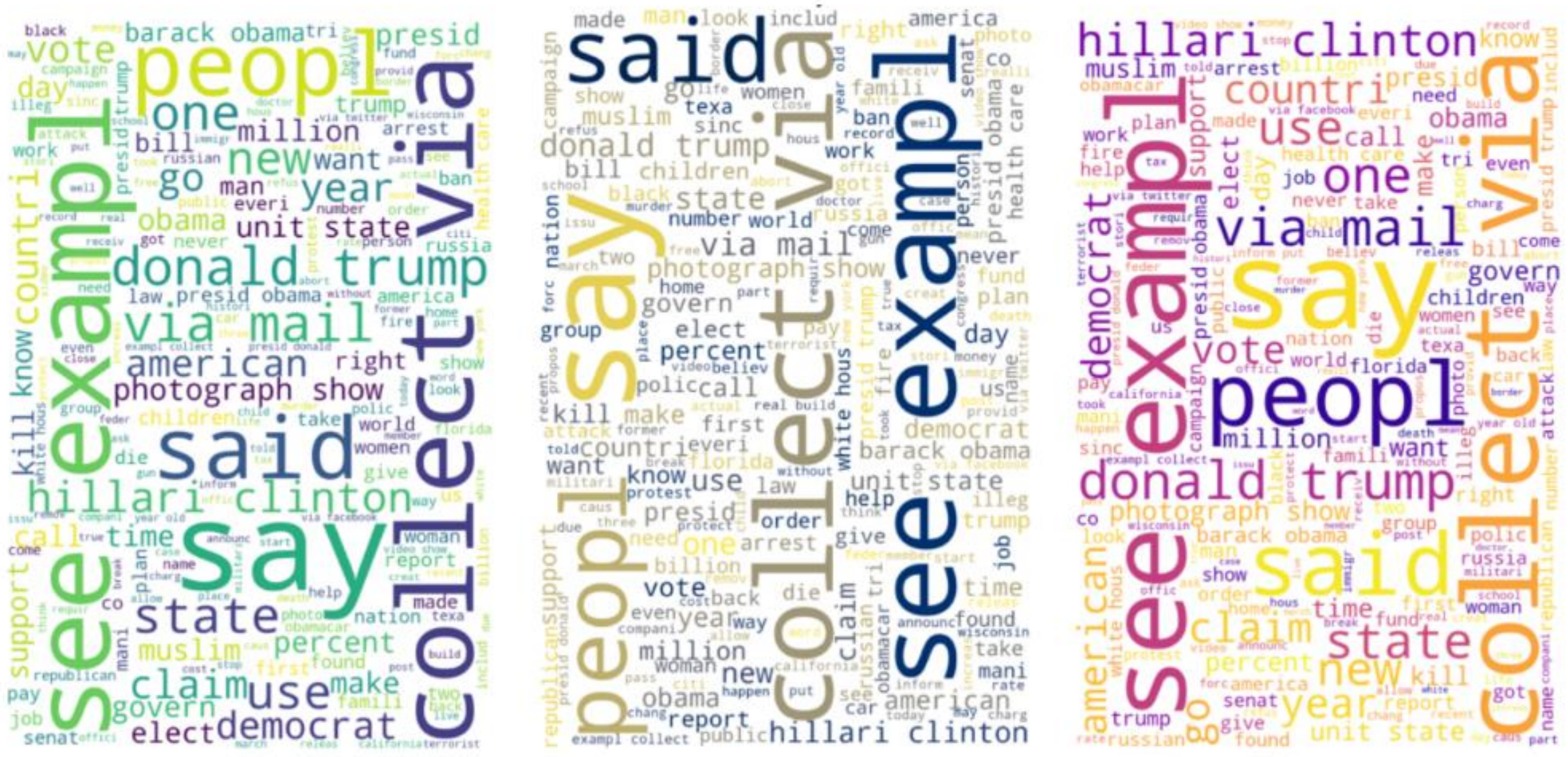




DATA VISUALIZATION

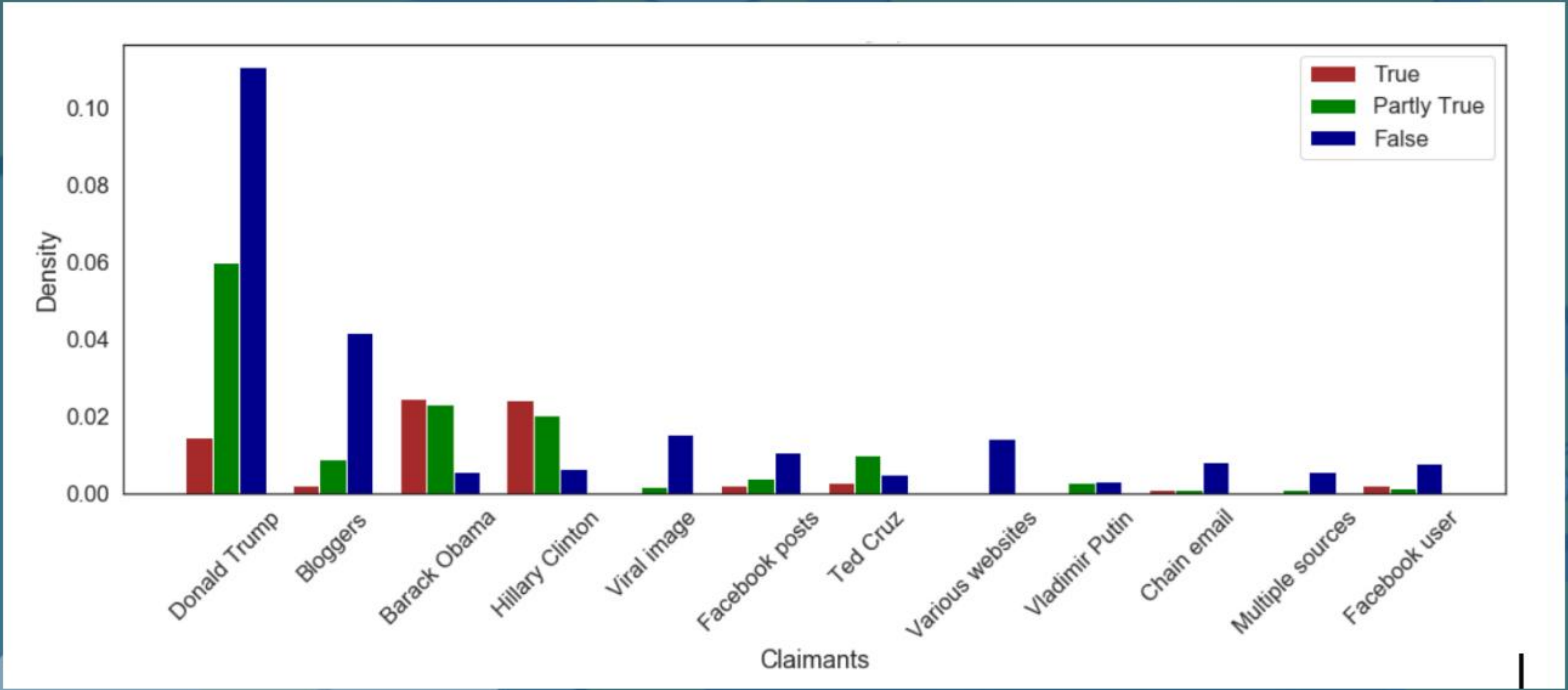
Application of current industry solutions for predicting polarity or subjectivity (eg. TextBlob) resulted in poor correlation with claim labels

DATA VISUALIZATION



(a) True Claims **(b)** Partly True Claims **(c)** False Claims

The most frequently used words within each of the three truth rating categories are largely similar, suggesting that use of only word frequencies to train an ML classifier will result in poor classification of the claims



DATA VISUALIZATION

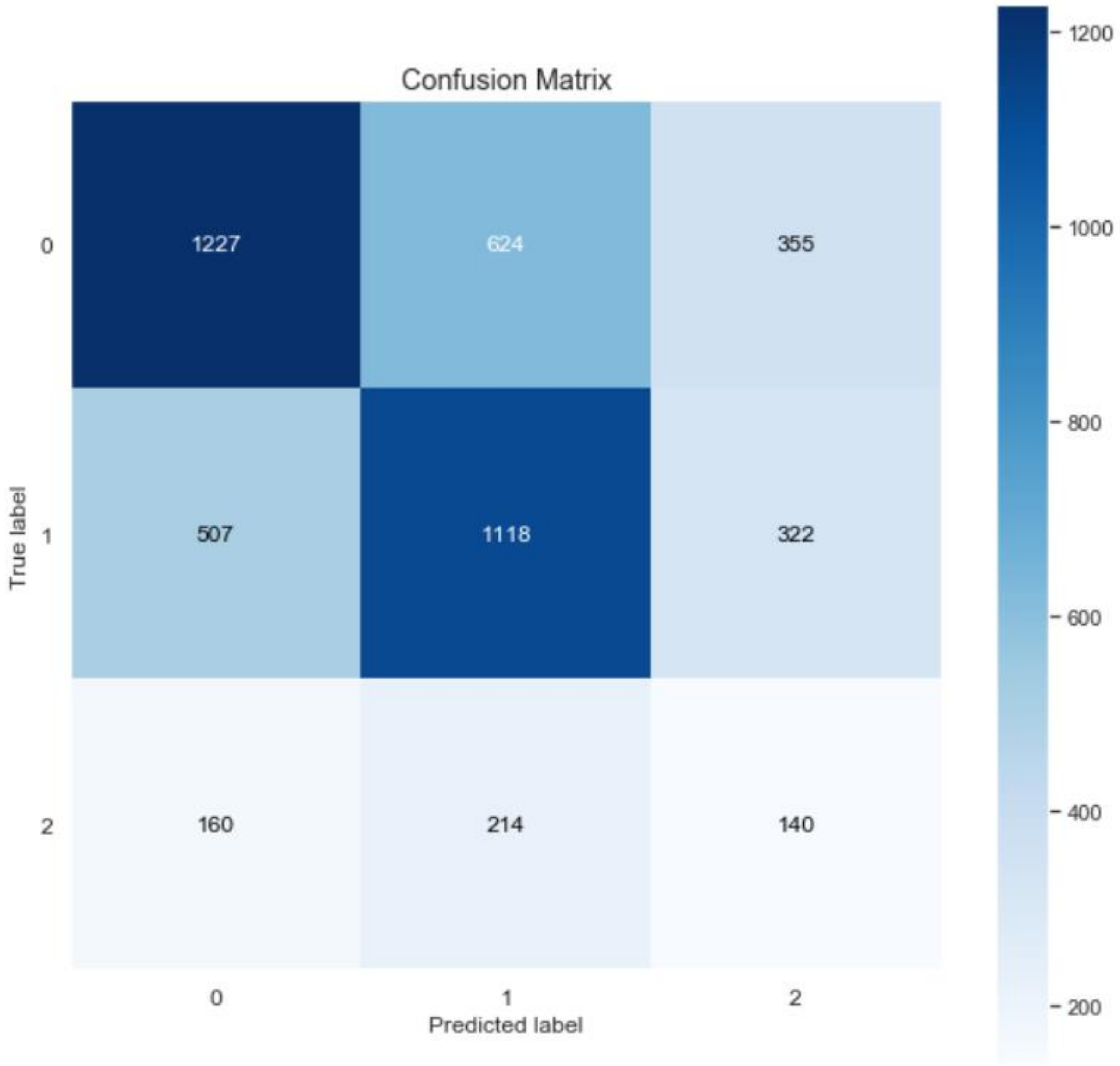
Correlation of the claimants data and the claim truth labels displays clear differences between the three types of claims, to the point of achieving near deterministic categorization for some of the claimants

ALGORITHM APPLICATION

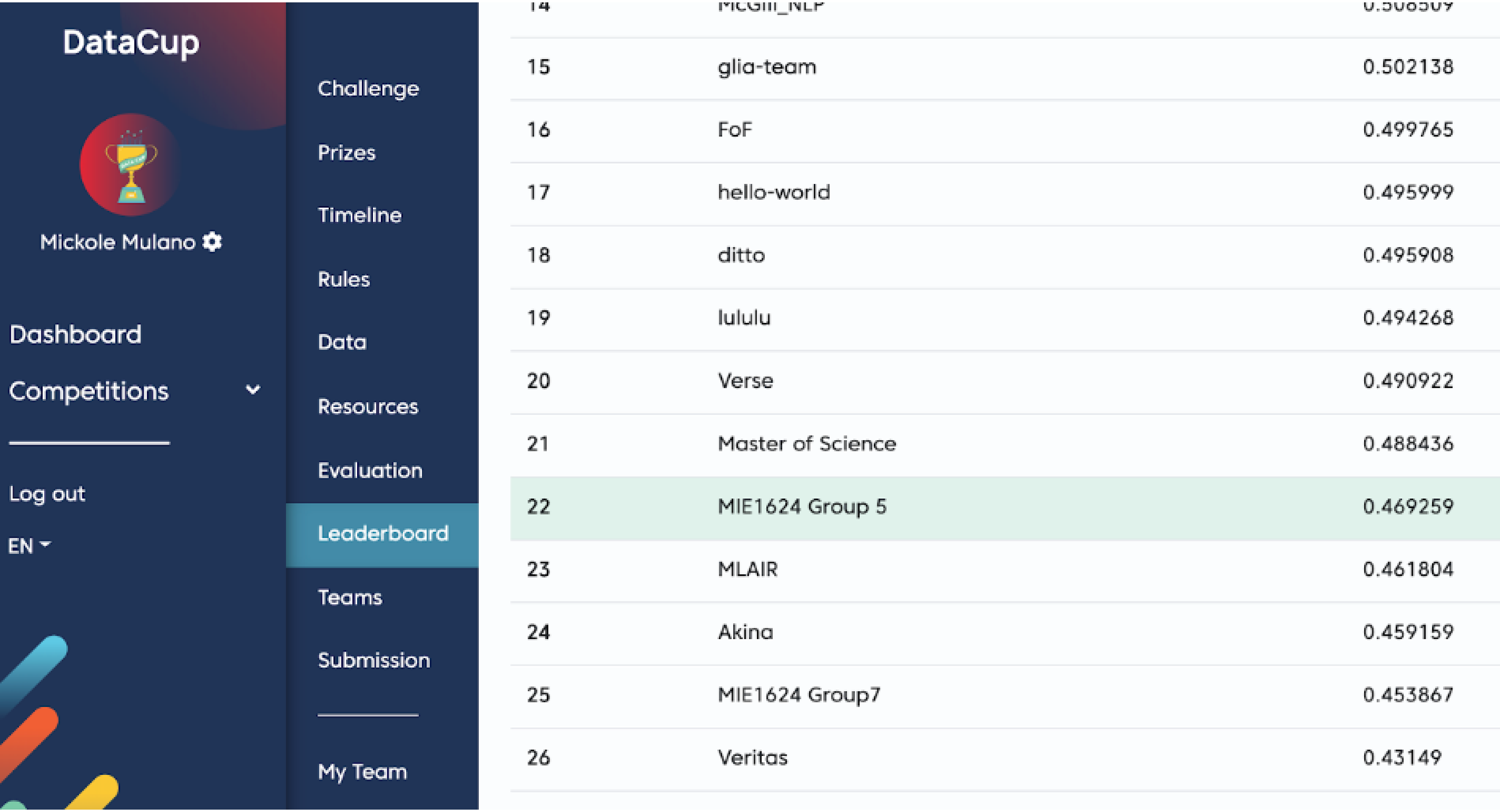
Eight models were trained with different ML classifiers. The performance of each algorithm was evaluated using the F1-score. Bernoulli Naïve Bayes classifier was selected as the final model, as it yielded the best performance amongst those tested.

Table 1: Comparison of model performance against local performance metrics

ML Classifier	`Training F1-Score	Validation F1-Score
Logistic Regression	60.66%	45.82%
Linear Discriminant Analysis	61.89%	46.63%
Multinomial Naïve Bayes	55.03%	46.06%
Bernoulli Naïve Bayes	54.29%	47.36%
Support Vector Machine (linear kernel)	59.42%	45.35%
XGBoost	44.34%	43.98%
Random Forest	97.17%	43.96%
Decision Trees	99.88%	41.99%



ALGORITHM EVALUATION



ALGORITHM EVALUATION

DISCUSSION

Using TF-IDF to create the feature set

Improves model efficiency may result in reduced model accuracy. Could use a more controlled but manual method such as Latent dirichlet allocation (LDA) for clustering

Use of related articles

Related articles would be critical in the cases where the training and validation datasets are from different general topics

Use of the stemming function to reduce word inflection

Does not take the meaning of the words into account. Lemmatization could be a better option

In case we have another

Lorem ipsum dolor sit amet, consectetur adipiscing elit.

DISCUSSION

Debunking false news and neutralizing false claims

Removal or insertion of claims that correct false claims

"Pre-bunking" to prevent false news

True information is preemptively injected into social media before false information is even created

Early detection of false news

First step and our objective


TRUE
Thank You!
FALSE