Team 16

Statistical Software Project

Fact-checking on Scientific claims

Amrutha Karuturi, ak2508
Taniya Kulkarni, tk740

Problem Statement

- Spread of misinformation
- Complications of scientific jargon
- Lack of verification tools
- Public health risk

Literature Review

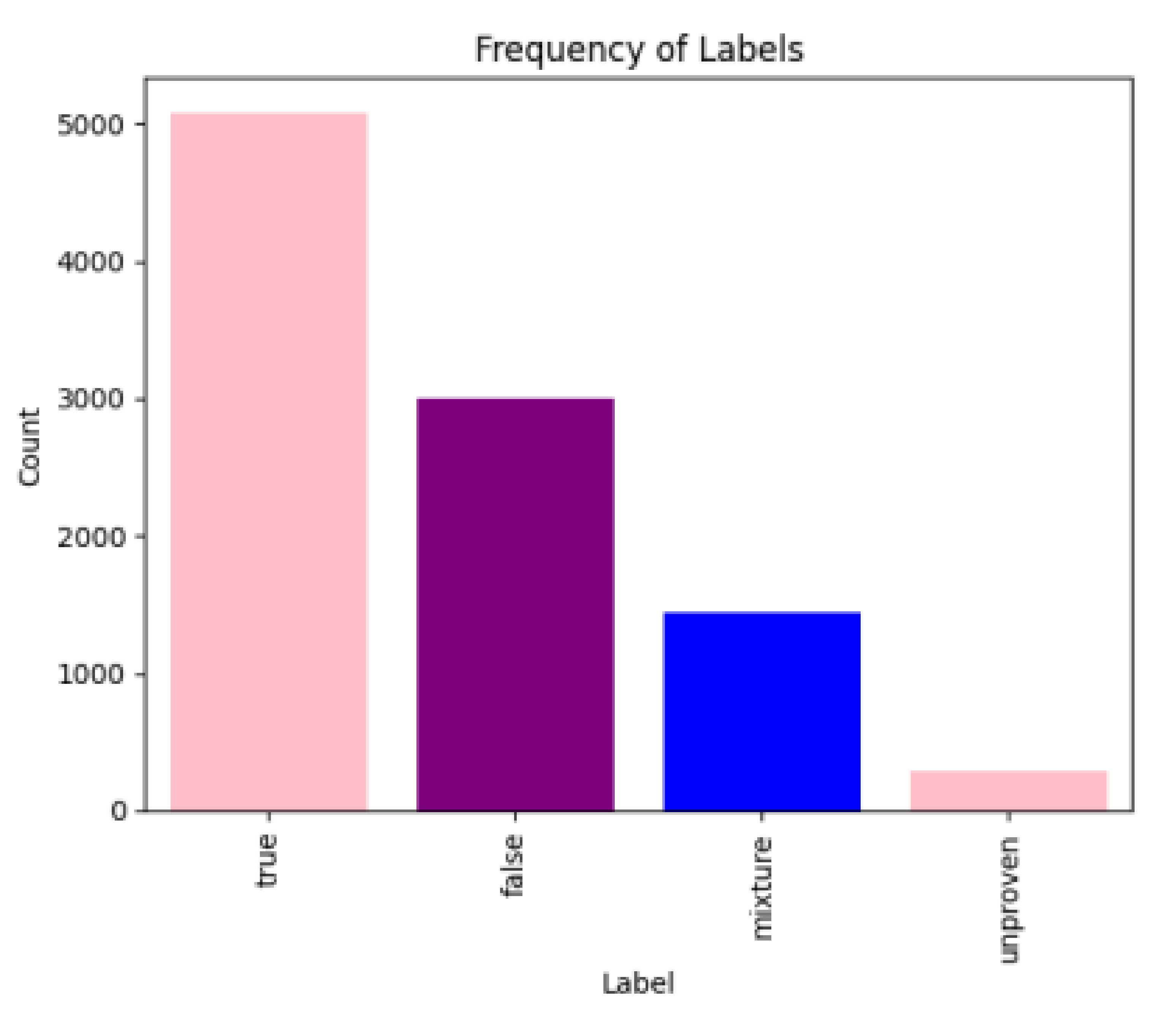
- **Paper**: Explainable Automated Fact-Checking for Public Health Claims (Kotonya, Neema et. al):
  - Novel Dataset for fact-checking
  - Framework for Veracity Predictions
  - Evaluation Through Coherence Properties

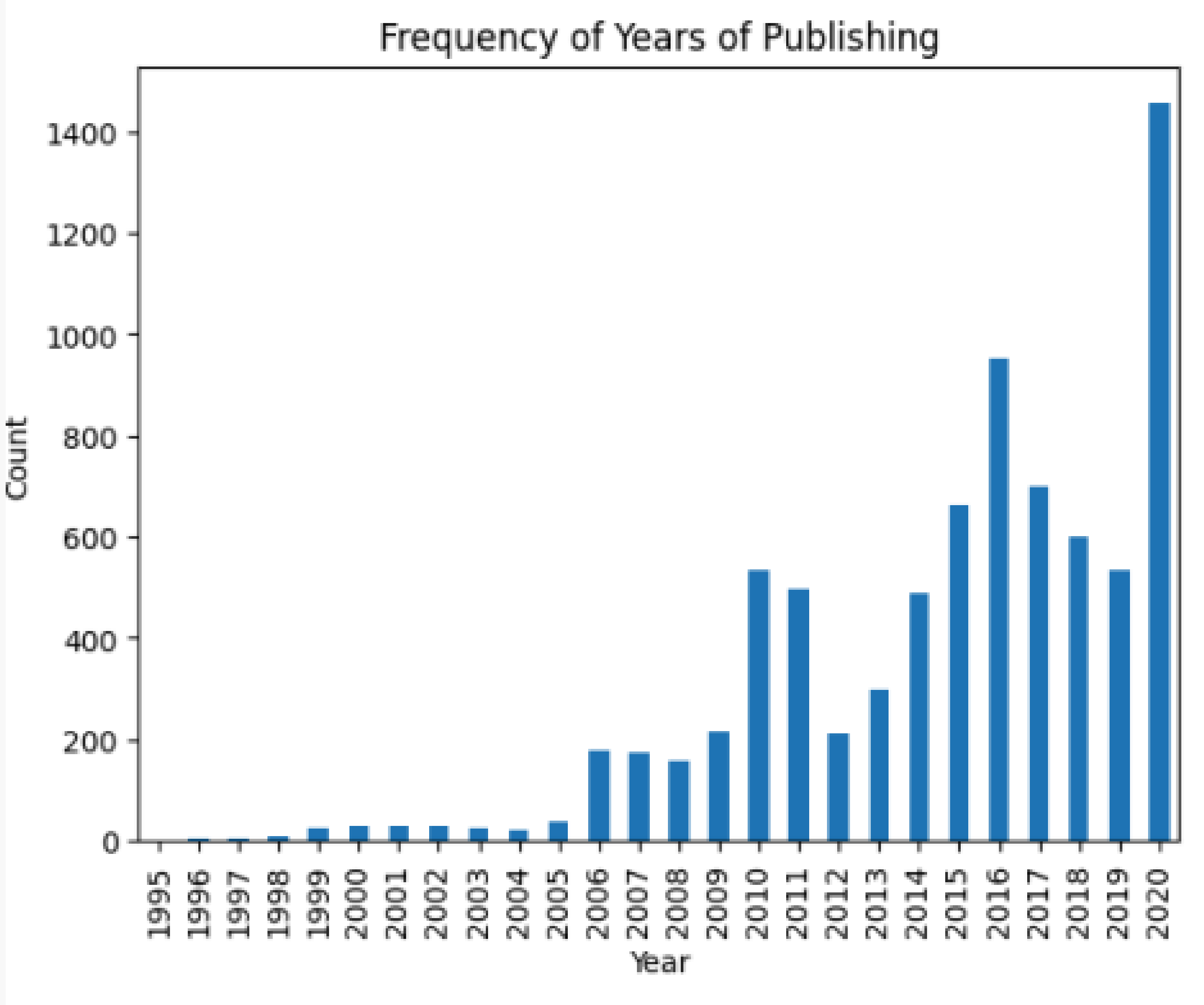
Dataset

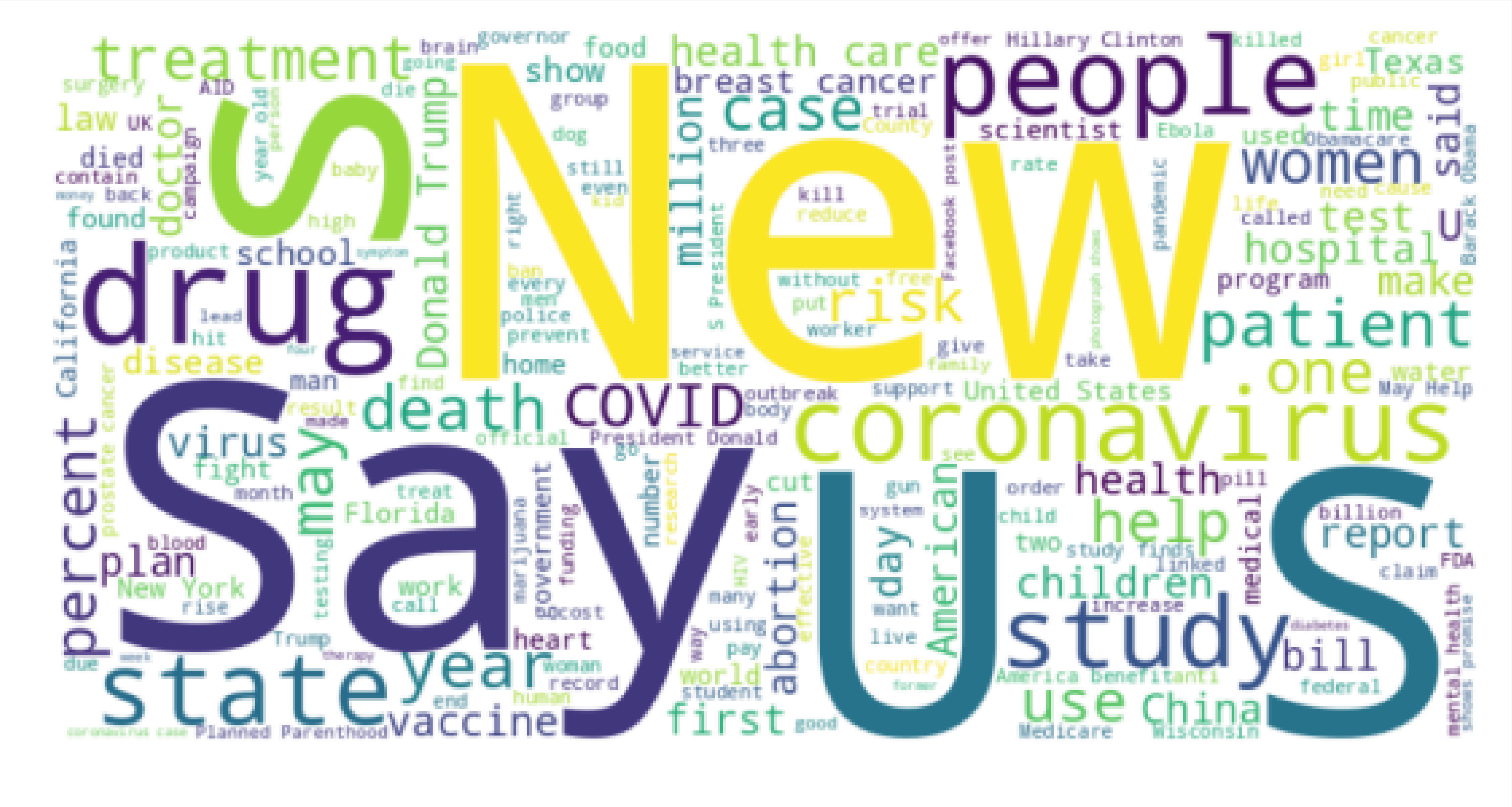
- PUBHEALTH fact-checking dataset by Neema Kotonya
- Consists of nearly 12000 rows
- Has 10 columns, out of which we make use of claim, main_text and label

	claim_id	claim	date_published	explanation	fact_checkers	main_text	sources	label	subjects
0	15661	"The money the Clinton Foundation took from fr...	April 26, 2015	"Gingrich said the Clinton Foundation ""took m...	Katie Sanders	"Hillary Clinton is in the political crosshair...	https://www.wsj.com/articles/clinton-foundatio...	false	Foreign Policy, PunditFact, Newt Gingrich,
1	9893	Annual Mammograms May Have More False-Positives	October 18, 2011	This article reports on the results of a study...		While the financial costs of screening mammogr...		mixture	Screening,WebMD,women's health

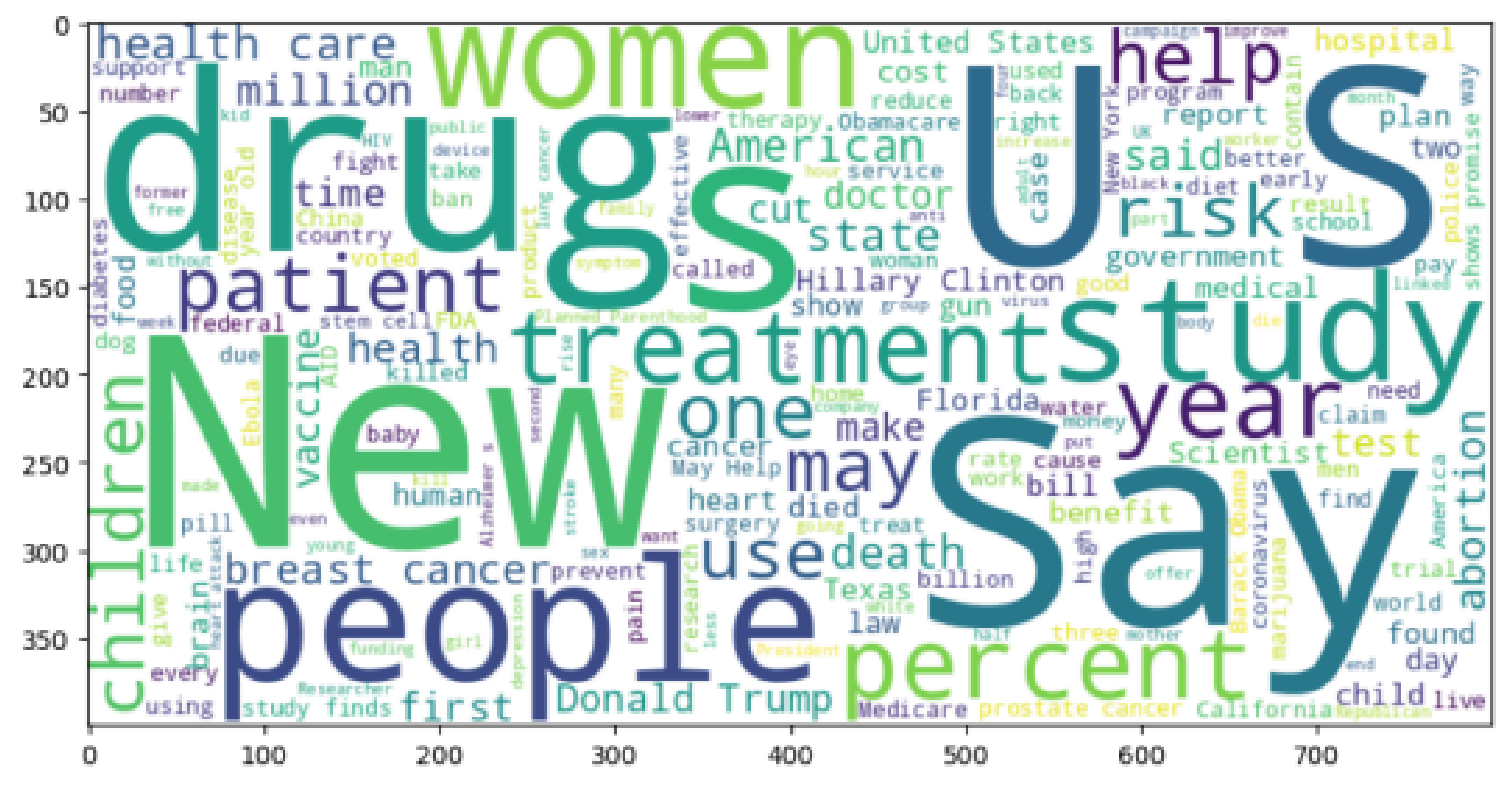
Exploratory Data Analysis



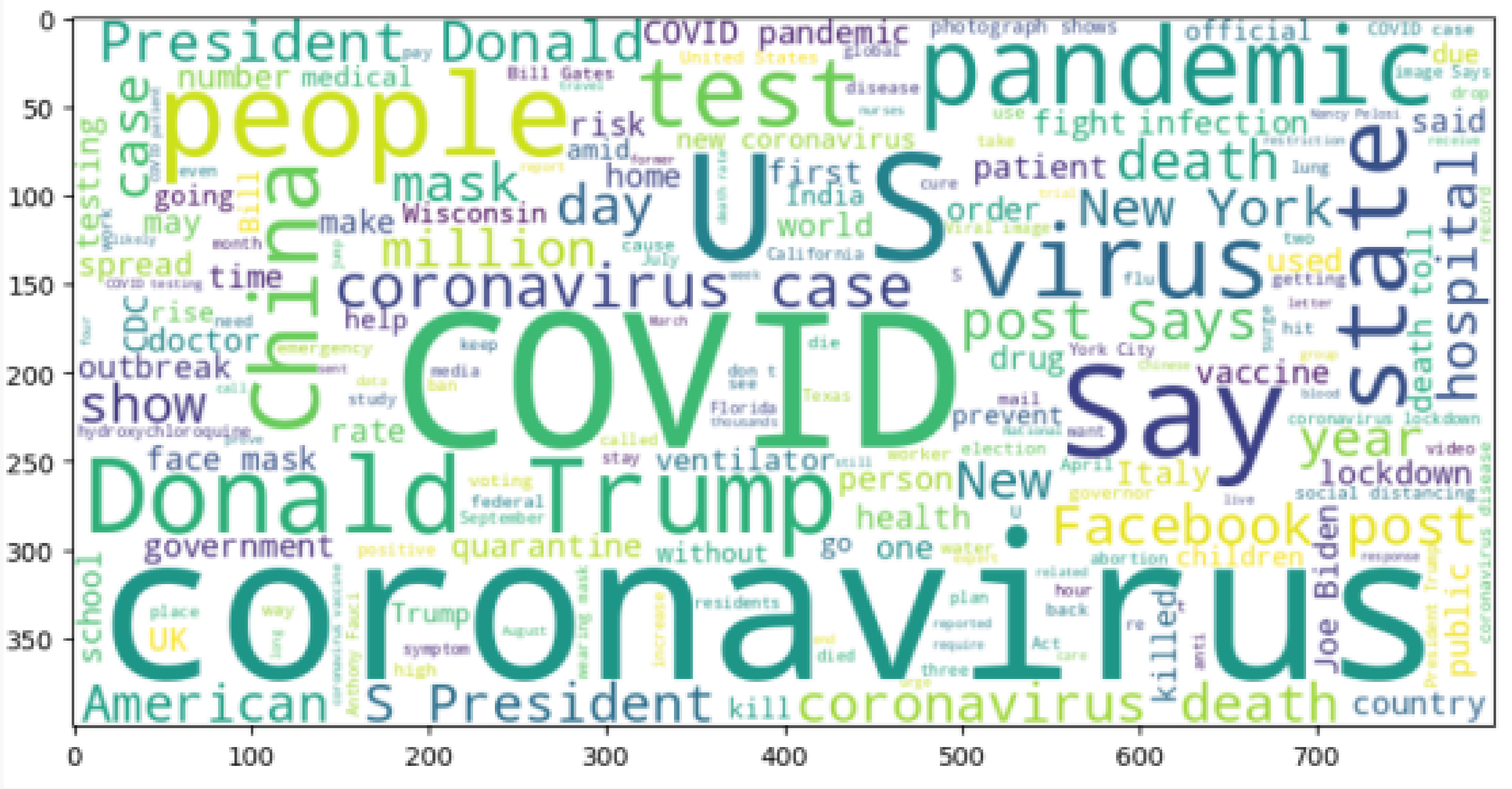




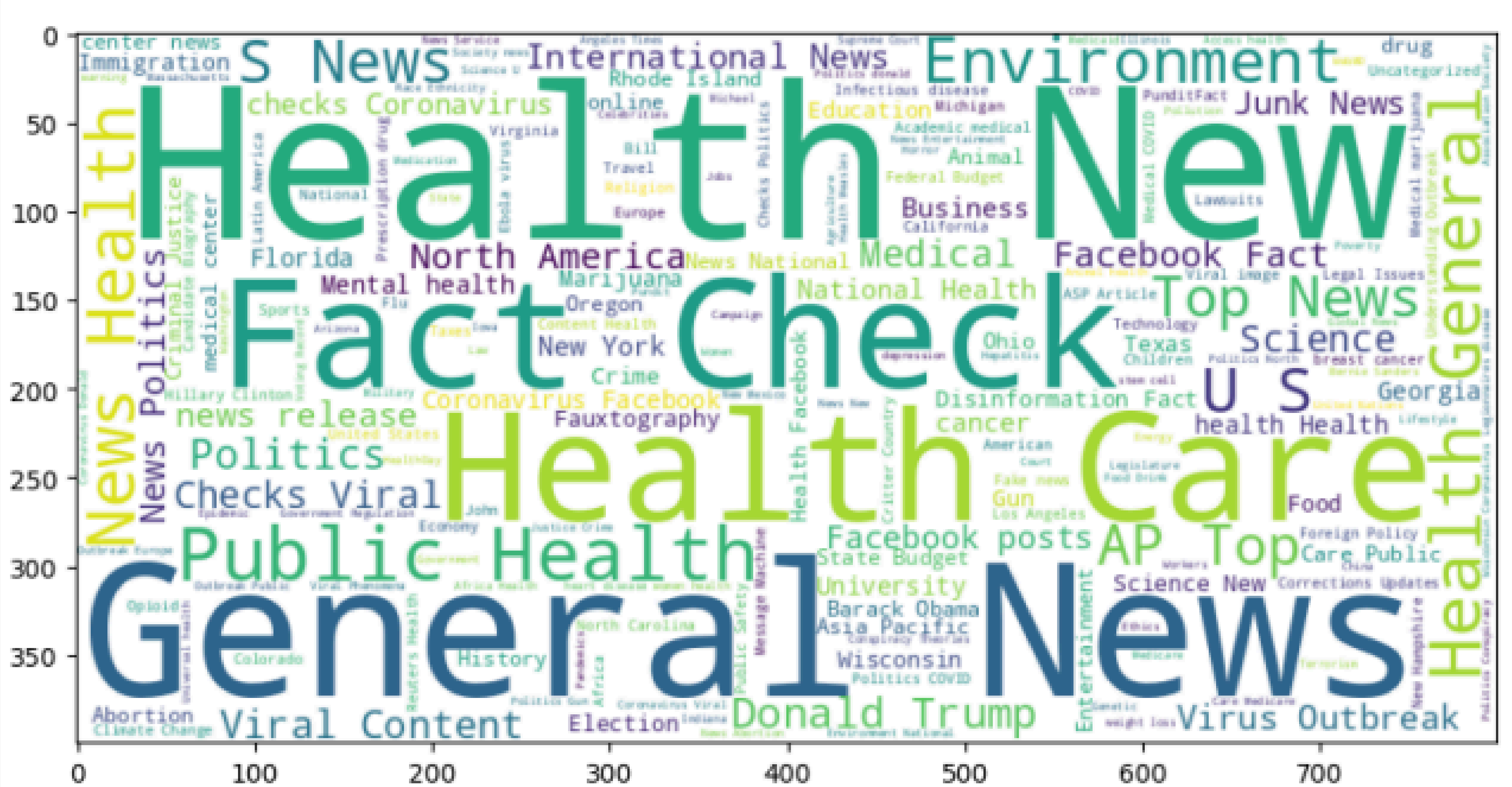
Word cloud for ‘Claim’



Word cloud for ‘Claim’ pre-pandemic



Word cloud for ‘Claim’ post-pandemic



Word cloud for ‘subject’

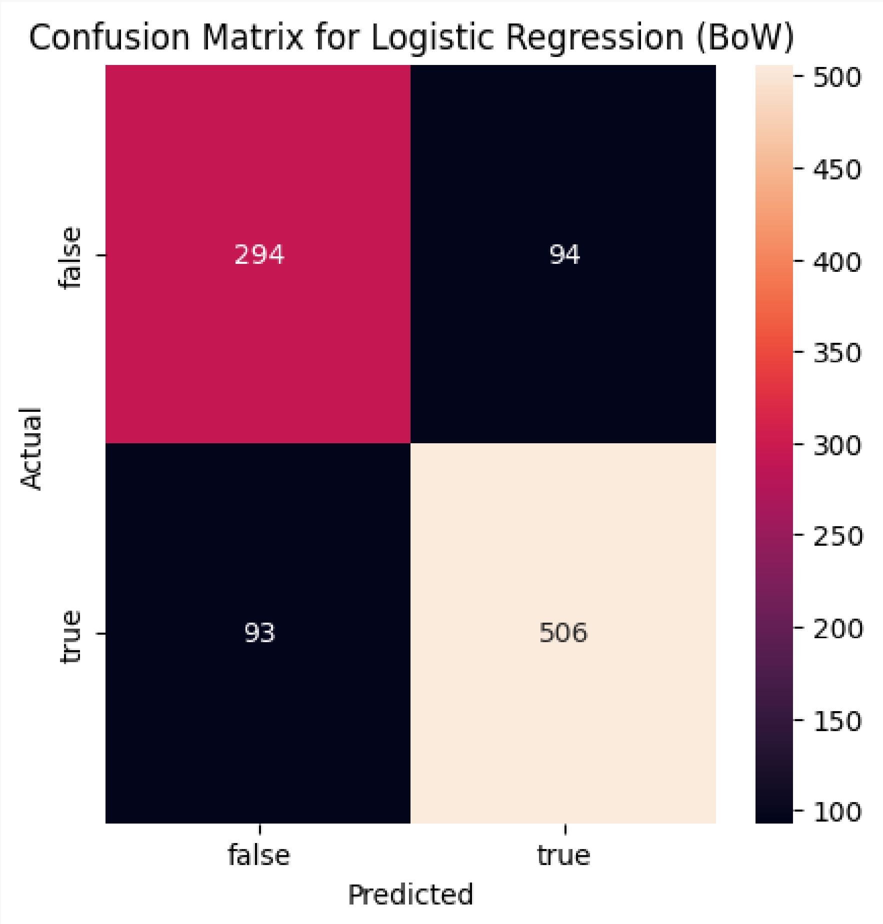
Methodology

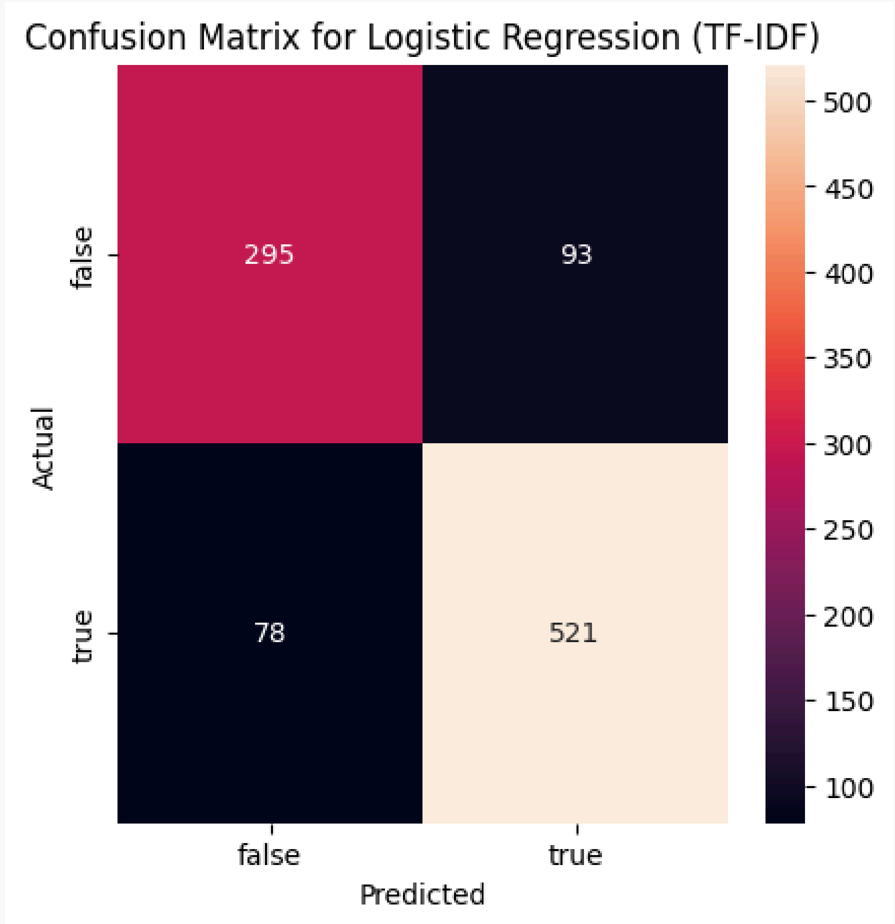
- **Text tokenizers used:**
  - Bag of Words
  - TF-IDF
  - Word2Vec
  - DistilBERT tokenizer
- **Models used:**
  - Logistic Regression
  - Simple Neural Network
  - DistilBert For Sequence Classification
  - SciBERT with top-k sentence retrieval

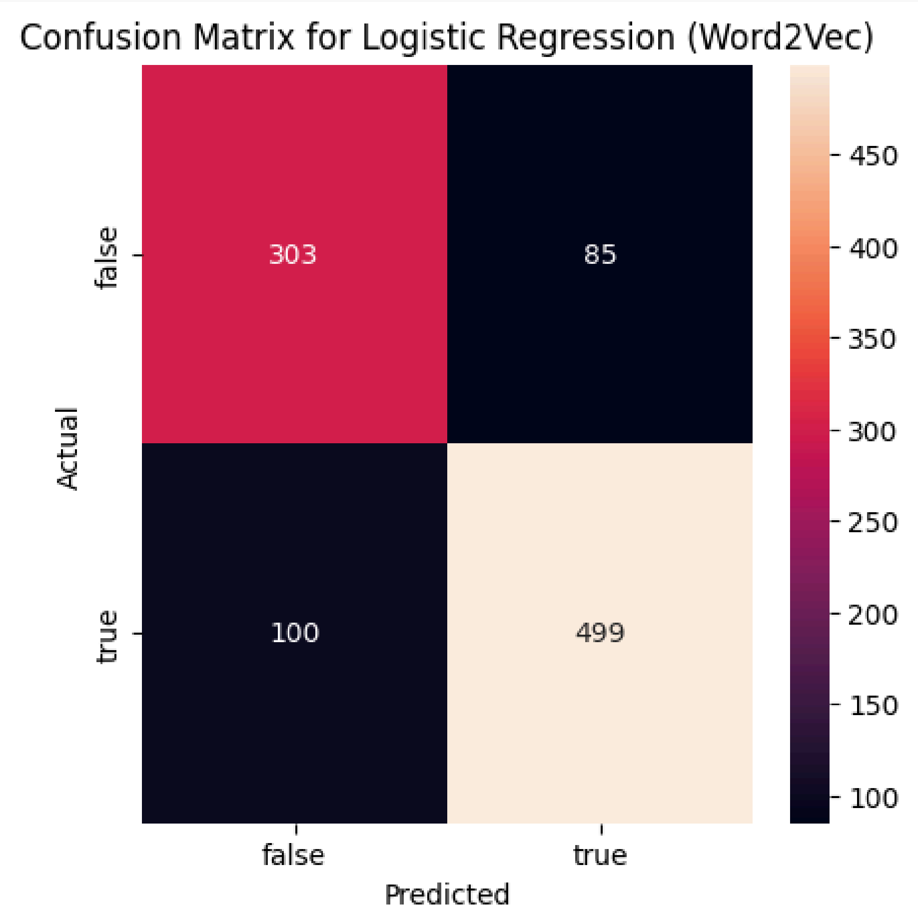
Performance metrics for all labels

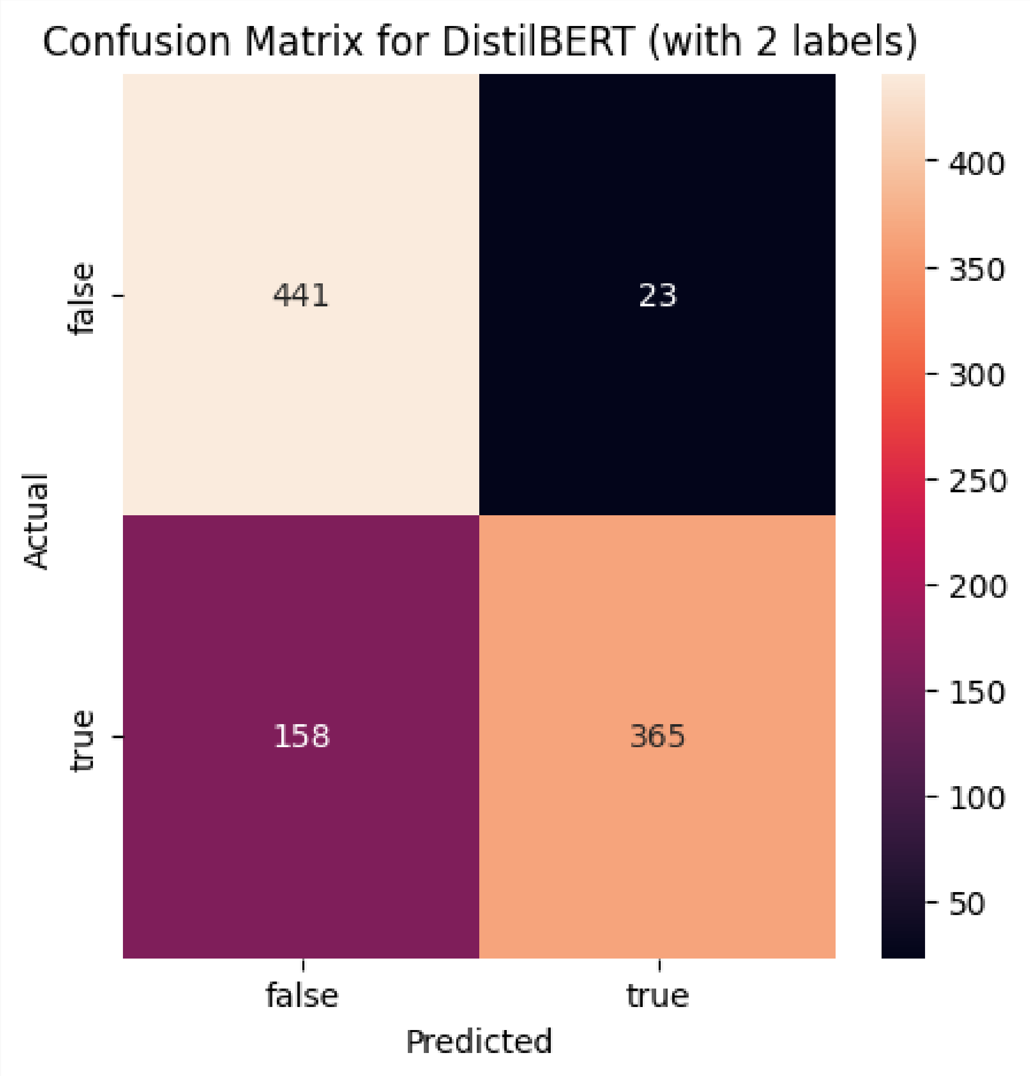
Model	Accuracy	Macro precision	Macro recall	F1 score
BOWs Logistic	0.1517	0.4663	0.3199	0.1916
TF-IDF Logistic	0.6407	0.5702	0.5063	0.6489
Word2vec Logistic	0.5531	0.4749	0.5068	0.5803
BOWs NN	0.442	0.4534	0.4705	0.5051
TF-IDF NN	0.2879	0.4629	0.4055	0.3526
Word2vec NN	0.485	0.4395	0.4215	0.5028
DistilBERT embeddings - DistilBERT sequence classifier	0.6586	0.5238	0.5314	0.6621

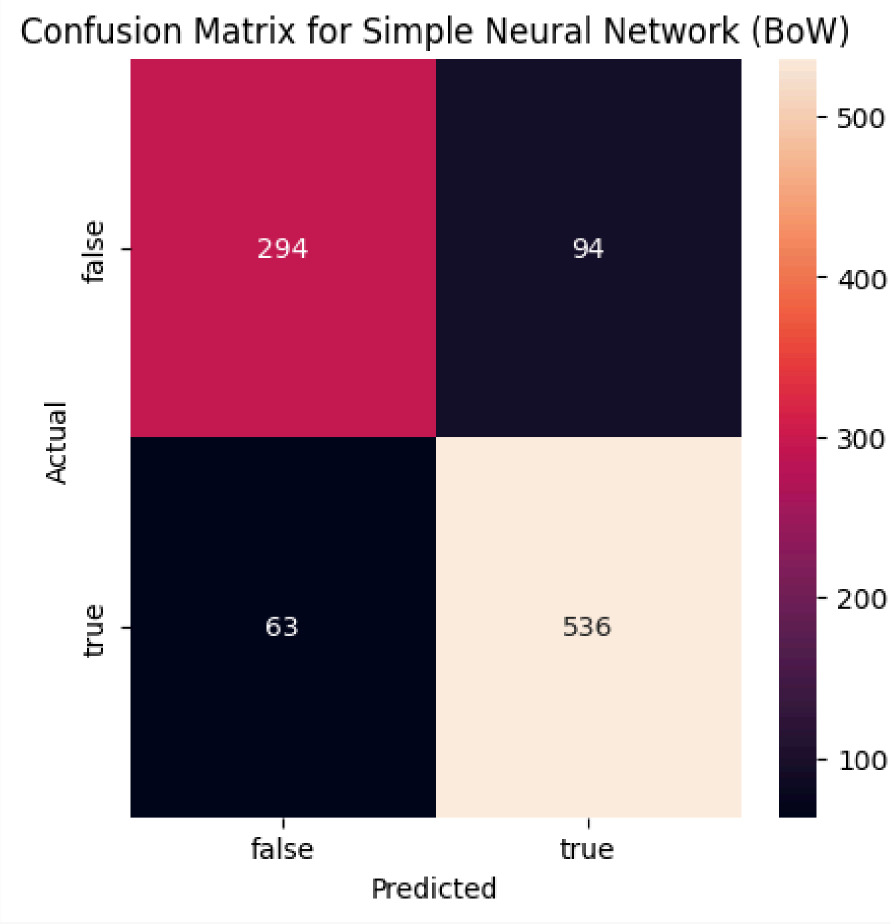
For 2 labels (true & false)

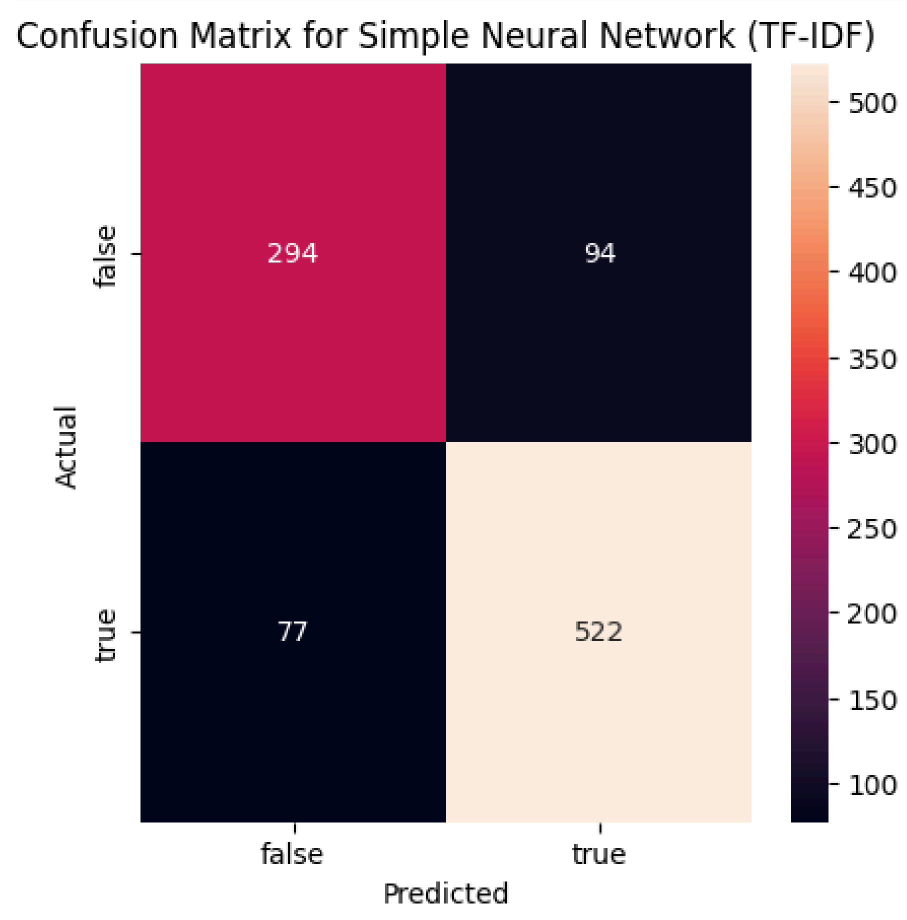


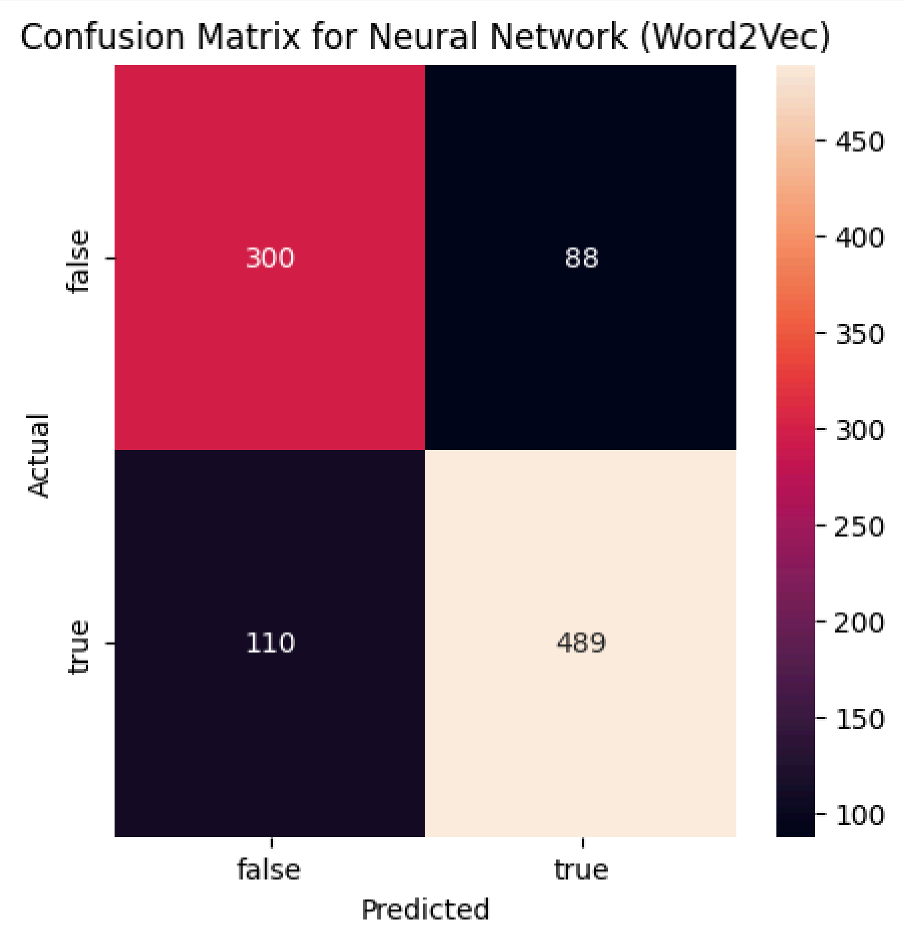








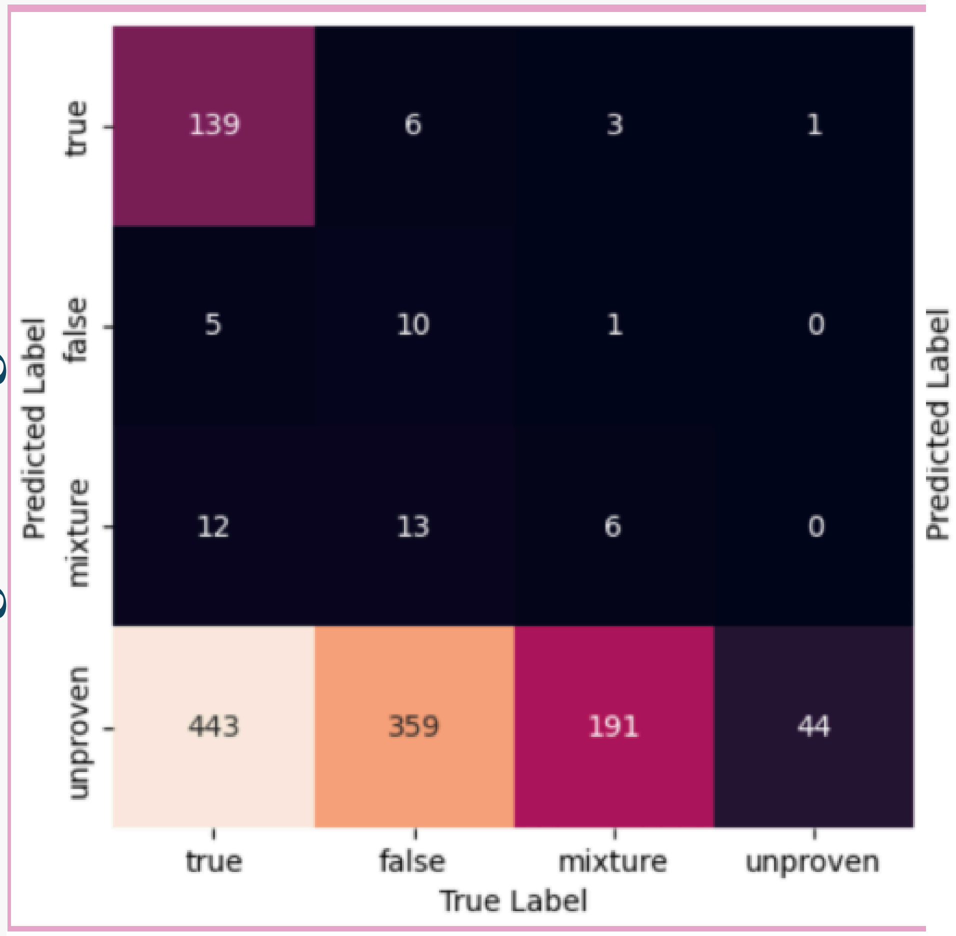


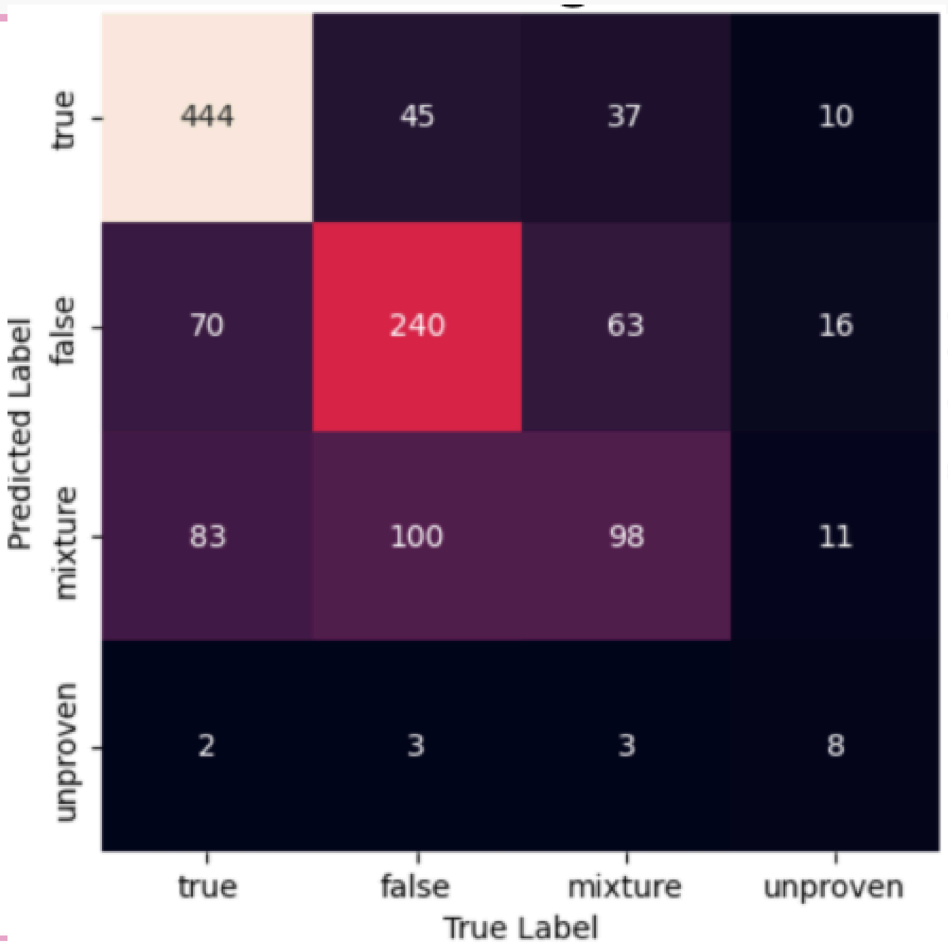


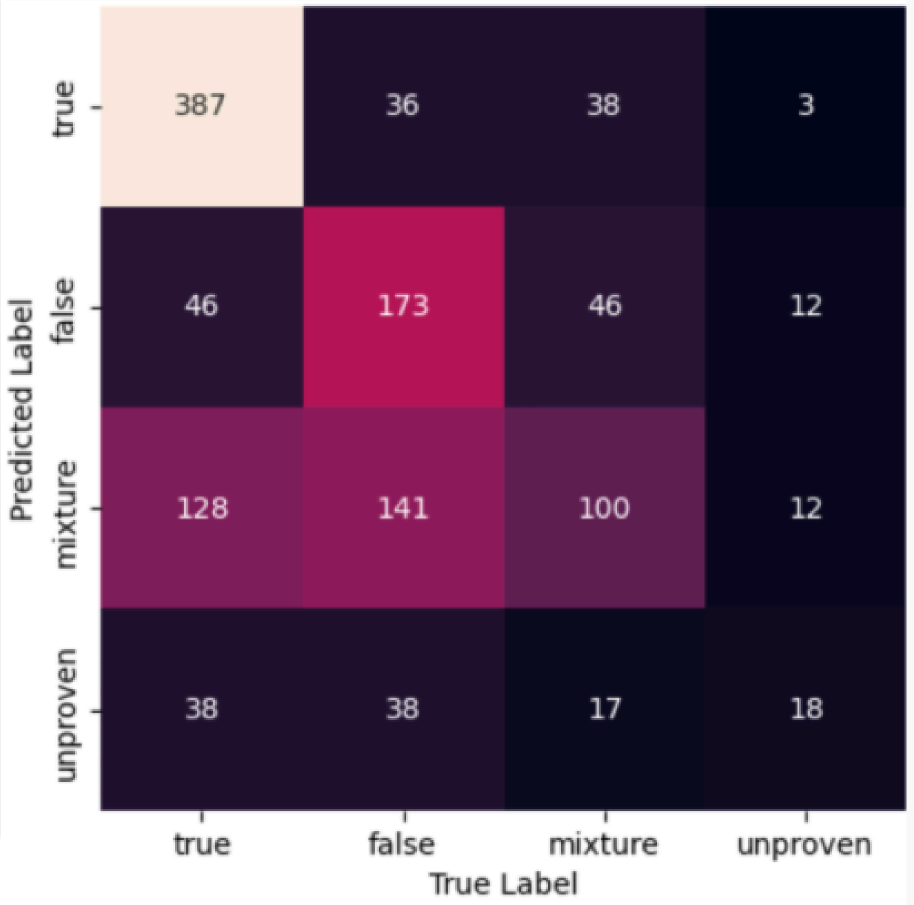
For all labels

BOWs | *TF-IDF* | *Word2vec*

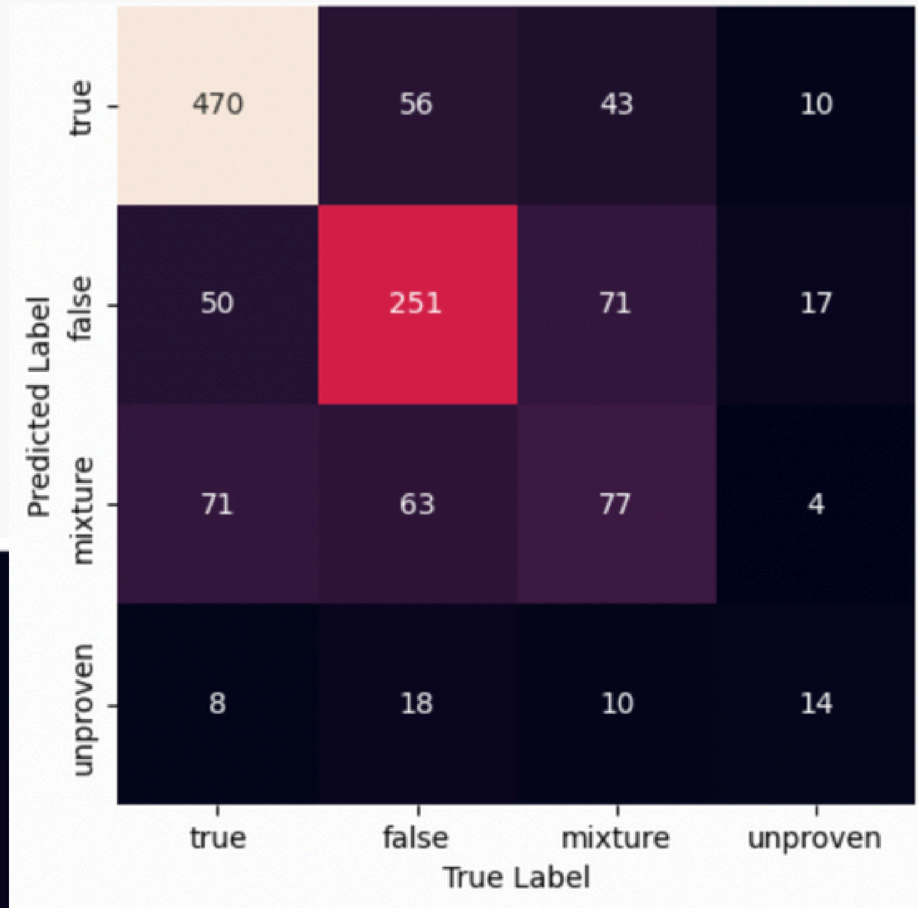
Logistic regression



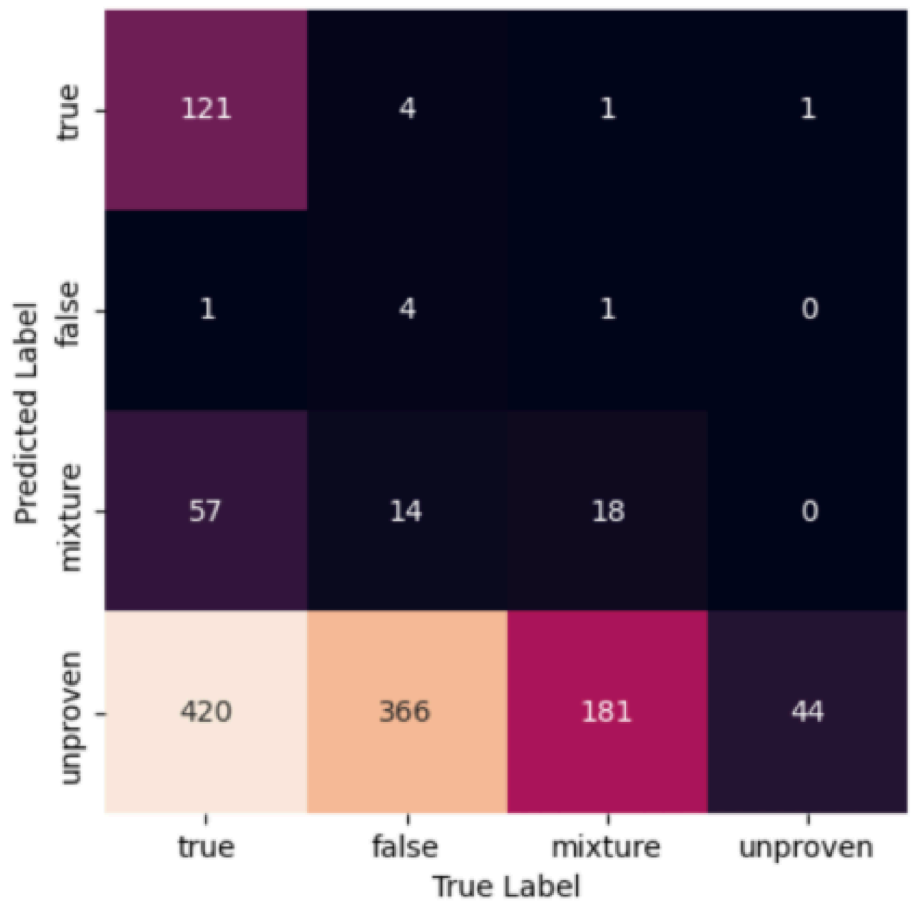


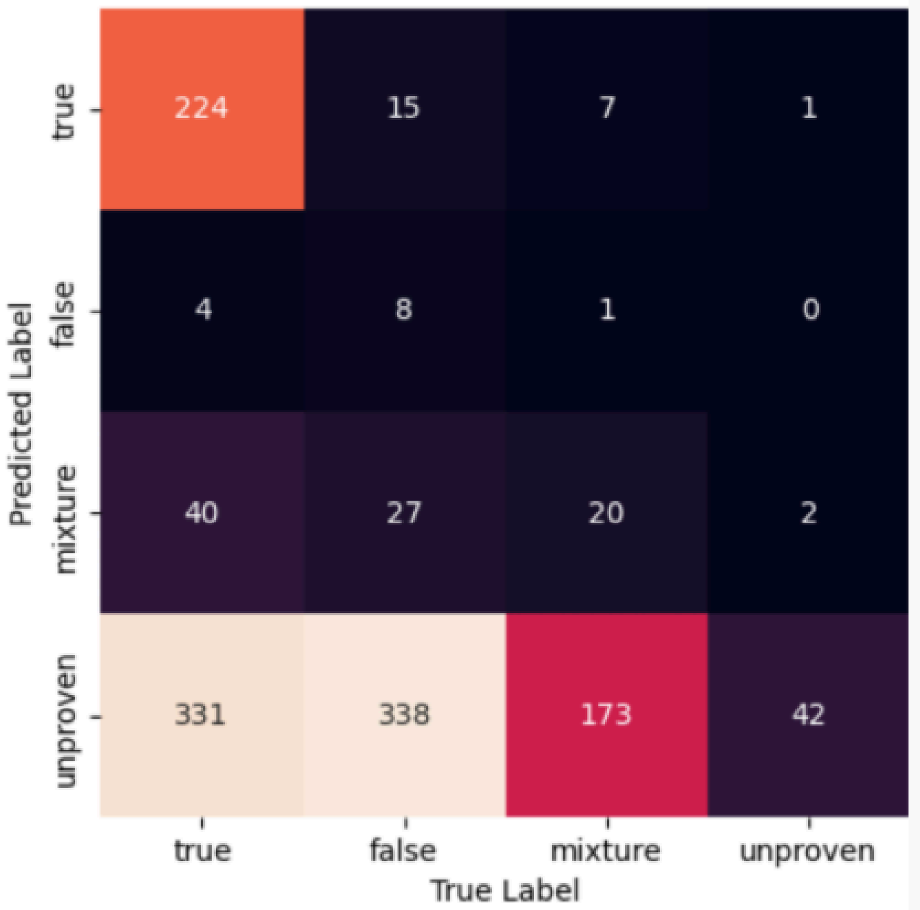


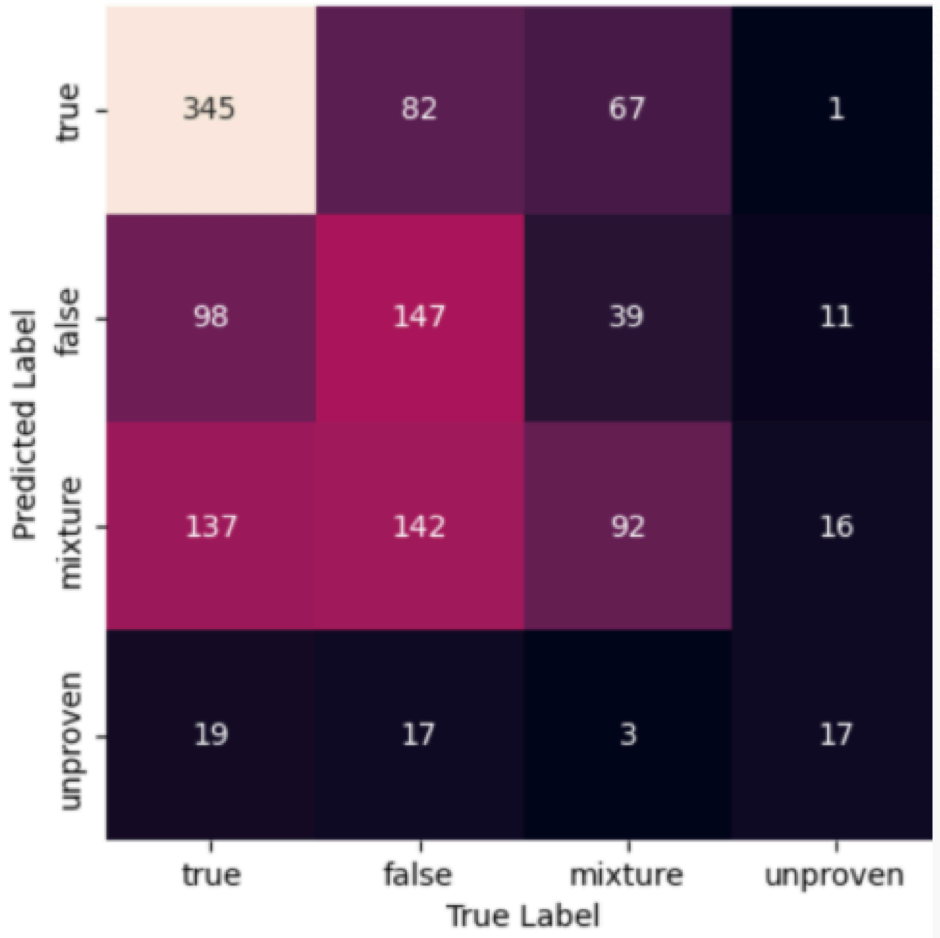
Distil-Bert



Nueral Network

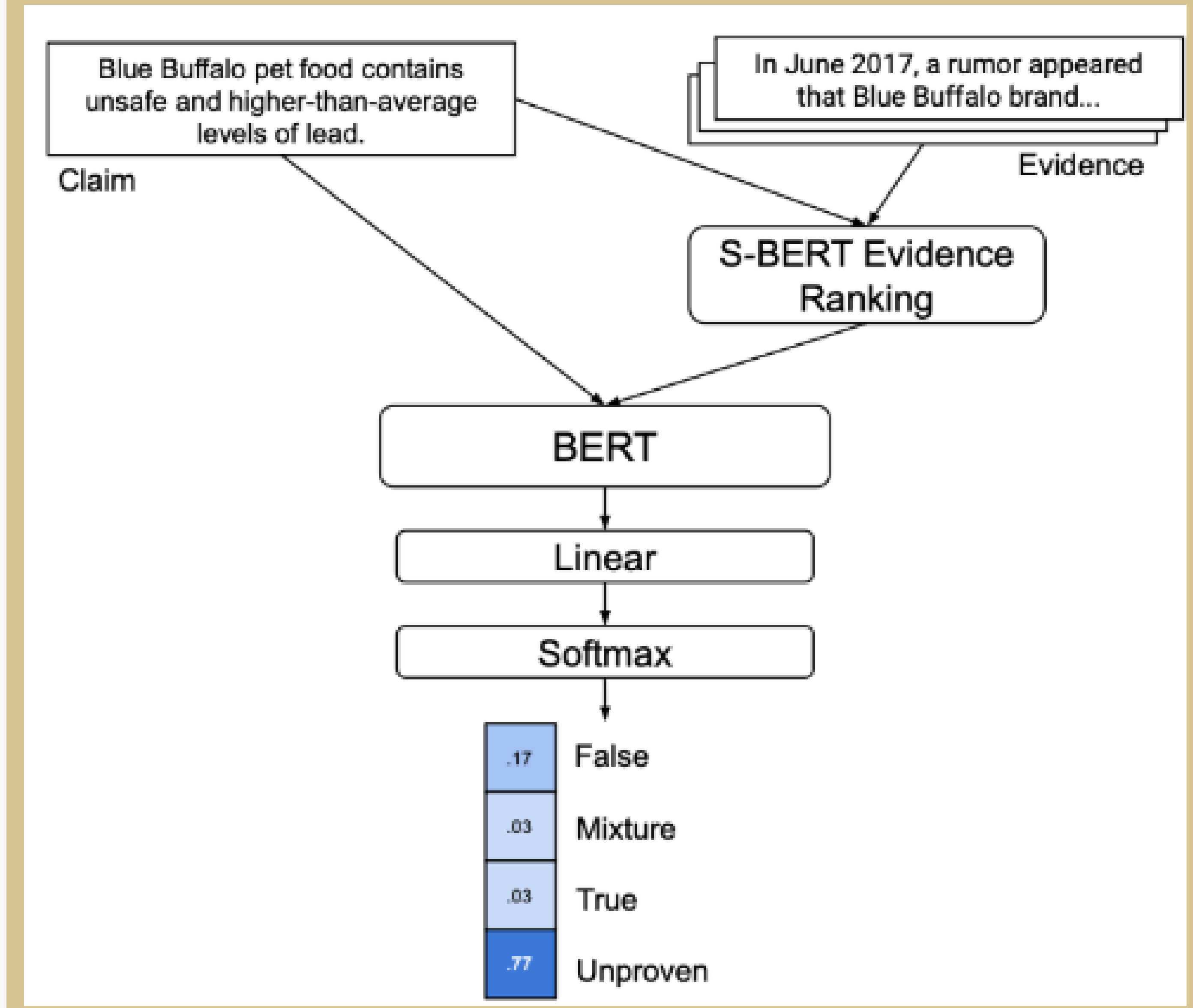




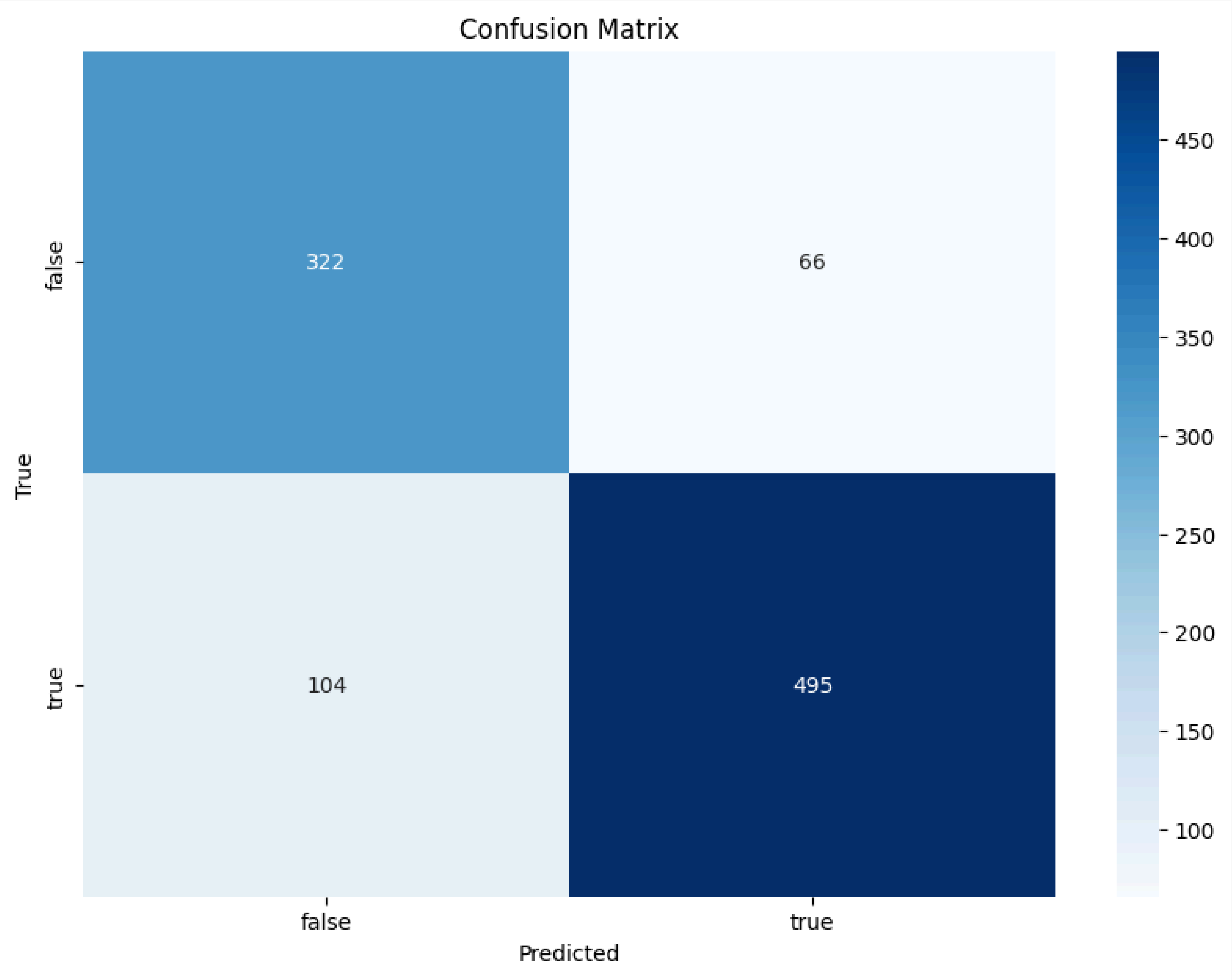


SciBERT Implementation

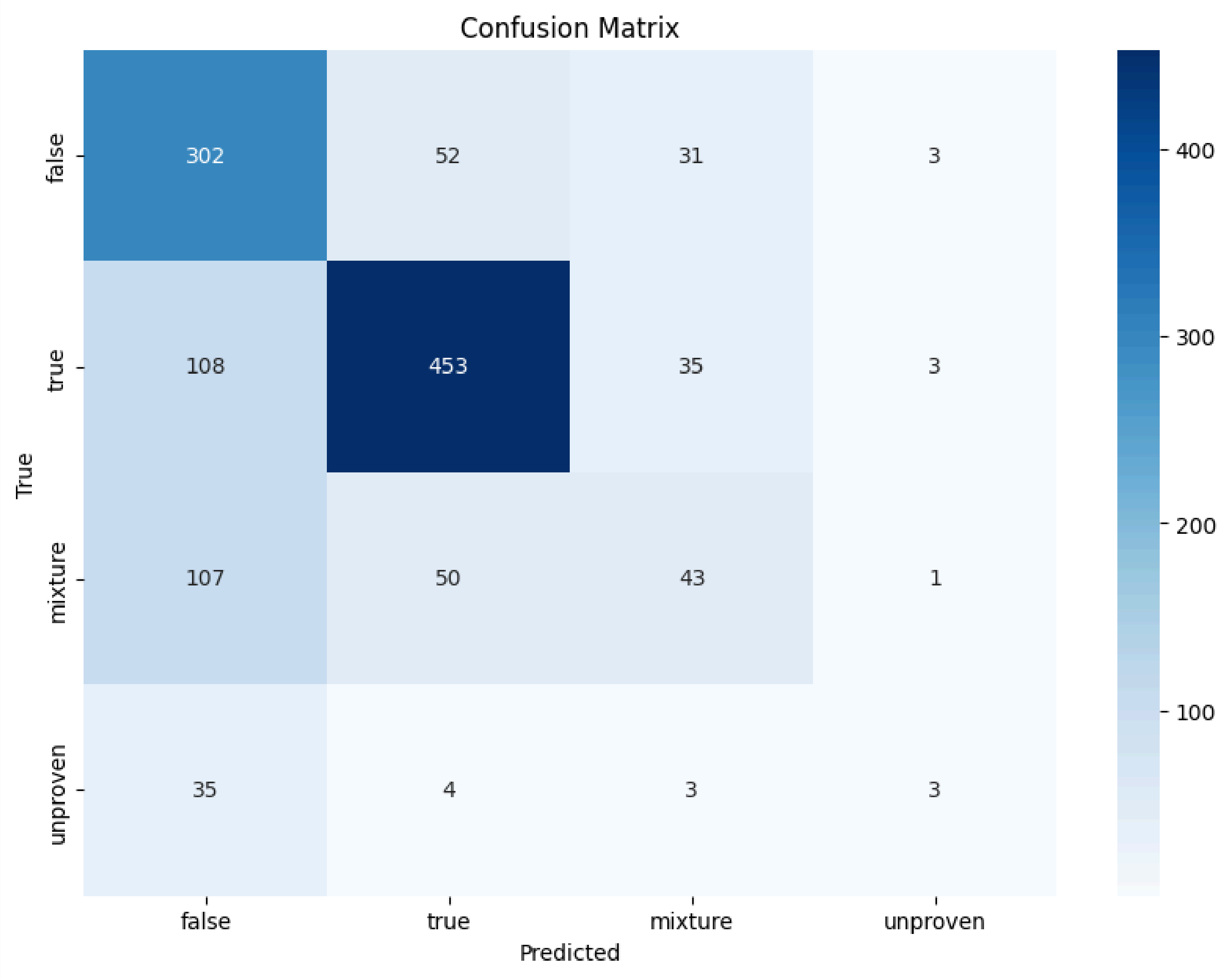
1. **Sentence retrieval:** Implement sentence retrieval to select the top k relevant sentences, preserving the main context by filtering out irrelevant lines.
2. **Sentence-Bert(SBERT):** Employ SBERT, based on BERT, to encode contextualized representations of evidence sentences and rank them by cosine similarity to the claim.
3. **Top k-selection:** Select the top k-ranked sentences for veracity prediction.
4. **For classification:** Consider models like SciBERT (which are more specialized in health data).



SciBERT with top-k sentence retrieval



For 2 labels



For all labels

Result

- SciBERT with top-k sentence retrieval performs with an accuracy of **67%** for all labels and **82%** for 2 labels
- DistilBERT performs with an accuracy of **65%** for all labels and **80%** for 2 labels
- Model tuned for binary classification task performs much better than multi-class classification task model

THANK YOU!