



# Detecting Context Misinformation in Text and Images

Deepika SIVASANKARAN, Mohd AAMIR

SUPERVISORS: Raphael Troncy, Paolo Papotti, Youri Peskine

27, June 2022



# Introduction

- Why Fake news is a problem?
  - Fake news brings **panic and misunderstanding against truth** among the people.
  - Has an adverse effect on the society and **manipulates public opinion.**

# Objective

Verification of any information found online needs to be done to avoid the aftermath occurring after believing online news which are fake.

Thus a machine learning model is created to analyze such news in relation to image contexts and classify them to the various levels of misclassifications.

# Example of a misleading image



Maryland drive  
gets probation for  
Delaware crash  
that killed 5 NJ  
family members

**True**



New 'Natural  
Feeding' trend has  
parents puking on  
babies

**Satire/Paroday**



Neuroscience  
Says Doing This 1  
Thing Makes You  
Just as Happy as  
Eating 2,000  
Chocolate Bars

**Misleading  
Content**



My plane hit an  
orca right after  
takeoff

**Manipulated  
Content**



Bowl of mussels

**False  
Connection**



I just thought that  
was sitting in the  
deli

**Imposter  
Content**

# Dataset

- a small part of Fakeddit dataset which contains text and images along with 3 types of labels for each sample.

- for **2-way**: classified into **True, False**

- for **3 way**: classified into **True, False, Partially True/False**

And **6-way**: classified into **True, Satire, Misleading, Manipulated, False Connection, Imposter**

# Dataset

- The subset of data contains 17622 rows and 16 columns given in the following distribution:

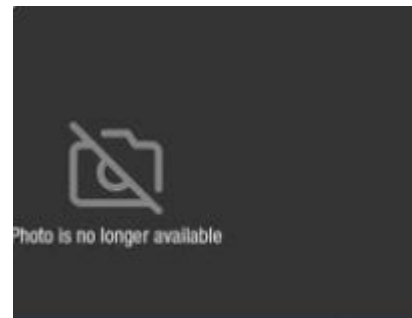
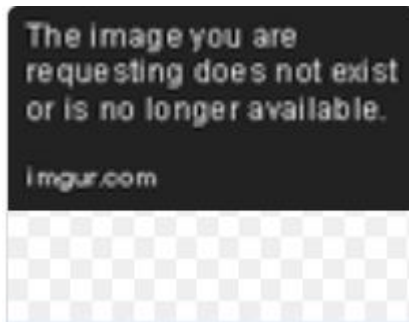
author	16096
clean_title	17622
created_utc	17622
domain	16750
hasImage	17622
id	17622
image_url	17603
linked_submission_id	872
num_comments	16750
score	17622
subreddit	17622
title	17622
upvote_ratio	16750
2_way_label	17622
3_way_label	17622
6_way_label	17622
dtype:	int64

# Problem With the Dataset

The dataset contains no null values in both sentences and images.

Although the dataset contains no null values or missing images, in the case of images, there are around 300 to 400 images which does not contain any meaningful information.

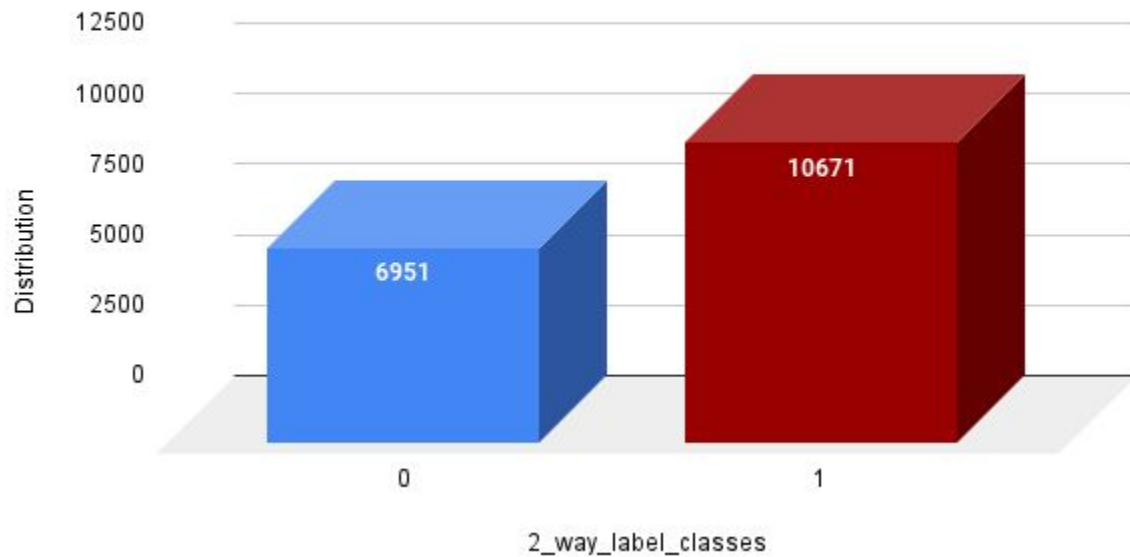
An example of one is shown:



# Dataset distribution

2\_way\_label Data Distribution

Total: 17622

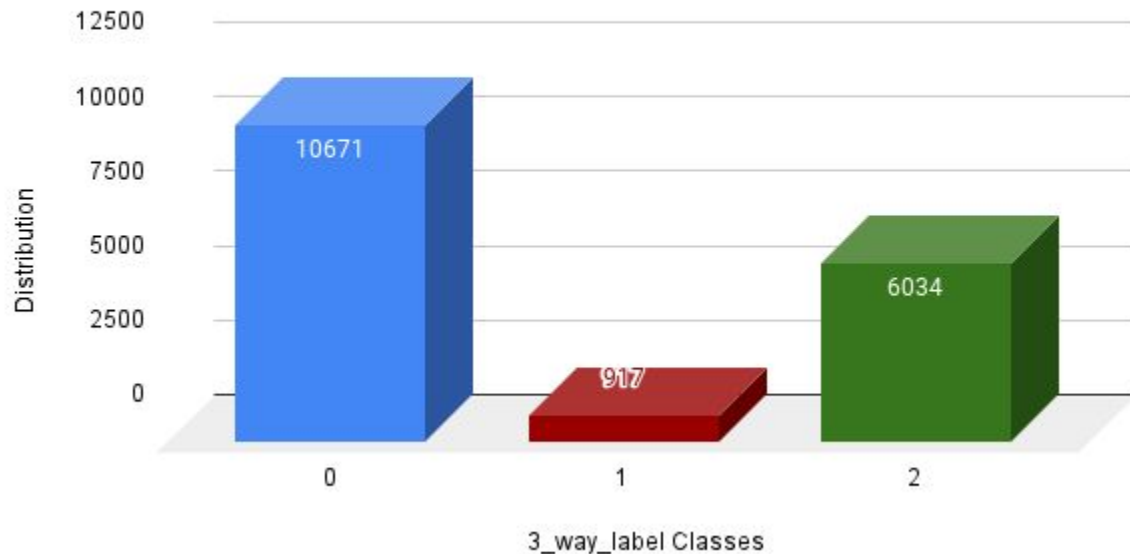




# Dataset distribution

3\_way\_label Data Distribution

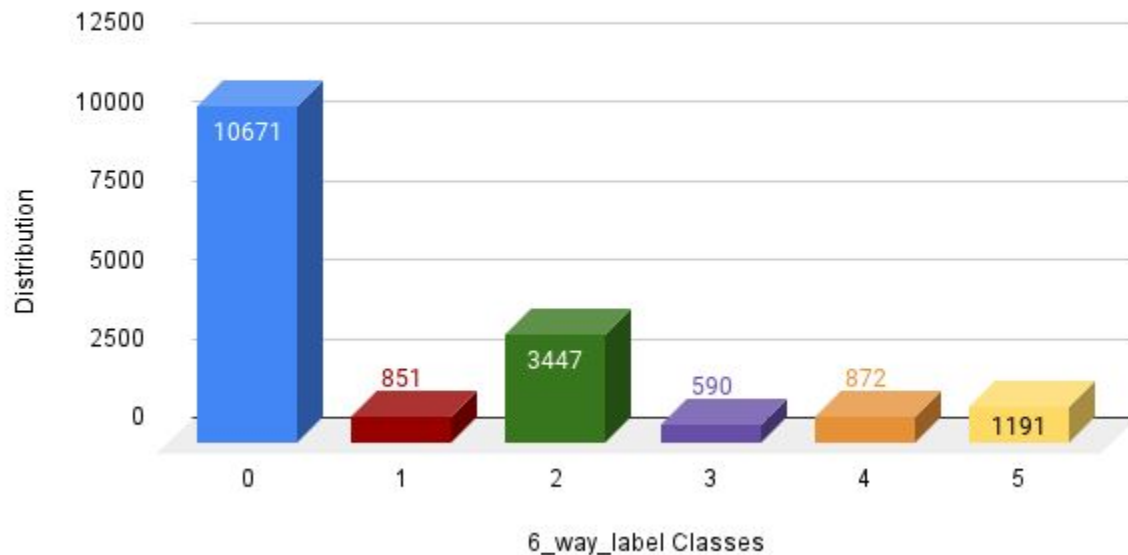
Total: 17622



# Dataset distribution

6\_way\_label Data Distribution

Total: 17622



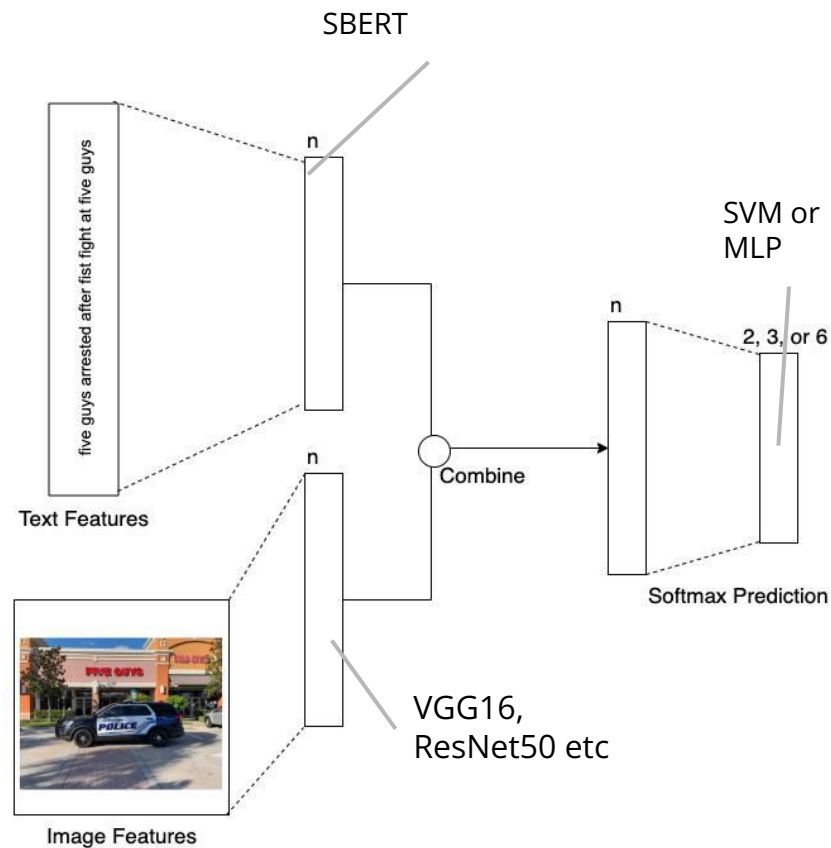
# Experiment Setup

- For **text**, S-BERT was used to generate text embeddings
- For **images**, ResNet was used to extract the features

Preprocessing: Images are resized to 224x224 since the dimensions differ in both height and width

# Experimental Setup

Fig. Model Architecture



# Experimental Setup

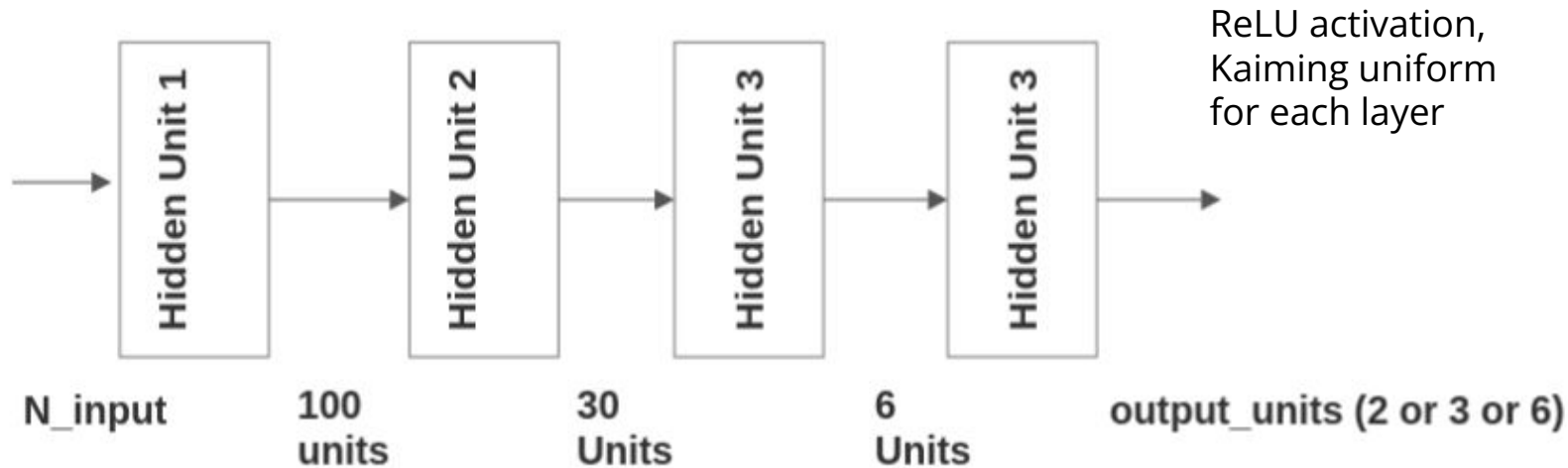
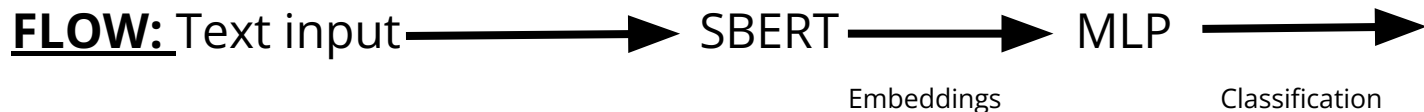


Fig: MLP Architecture

# For TEXT Only



For each of the classification, 2 models of SBERT were used.

1. paraphrase-MiniLM-L6-v2
2. all-mpnet-base-v2

Model	2_way_label	3_way_label	6_way_label
S-BERT (paraphrase-MiniLM-L6-v2)	0.80	0.79	0.73
S-BERT (all-mpnet-base-v2)	0.84	0.82	0.75

# For IMAGE only

**FLOW:** Image input → Pretrained Model → MLP → Classification  
Embeddings

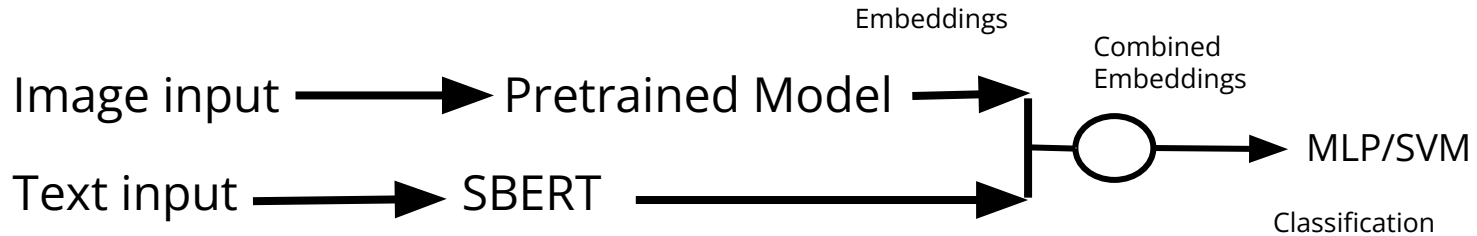
For each of the classification, 5 pretrained models were used.

Pretrained models: ResNet18, ResNet50, ResNet101, VGG16, EfficientNet\_B0

Model	2_way_label	3_way_label	6_way_label
ResNet18	0.75	0.76	0.76
ResNet50	0.80	0.80	0.81
ResNet101	0.77	0.61	0.60
VGG16	0.78	-	-
EfficientNet_b0	0.76	0.80	0.75

# Text + Image

## FLOW:



The best performing models for text and images are used to extract features.

The embeddings are combined using minpooling, maxpooling, append and average pooling

Classifiers such as MLP and SVM are used.

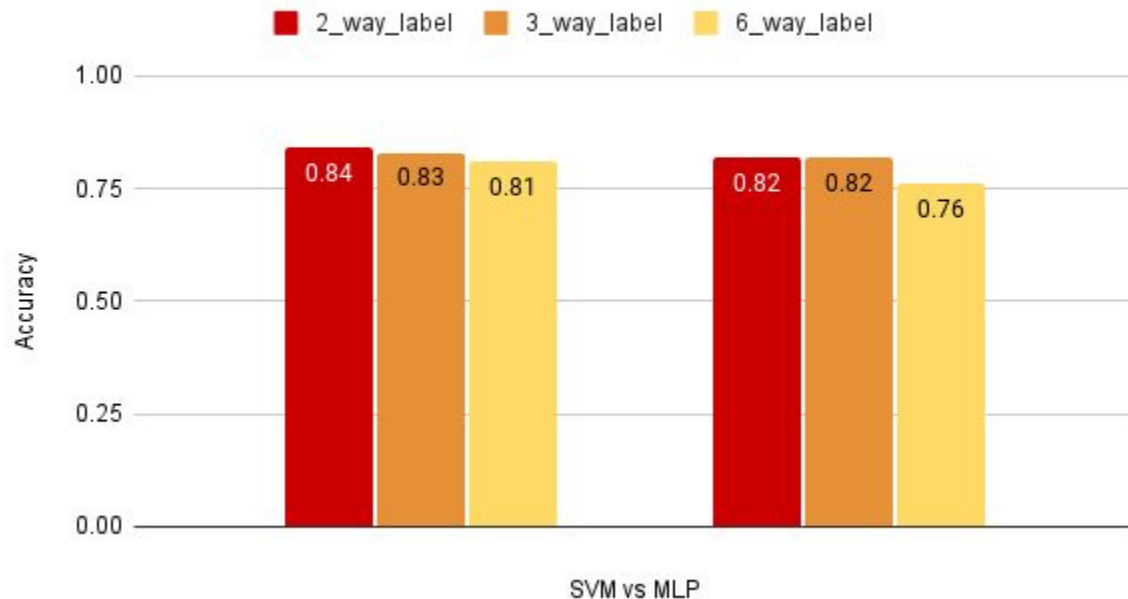


# Text + Image

Model	2_way_label	3_way_label	6_way_label
SBERT+ResNet by concat	0.82	0.79	0.61
SBERT+ResNet by maxpool	0.80	0.79	0.62
SBERT+ResNet by Minpool	0.82	0.82	0.76
SBERT+ResNet by avgpool	0.80	0.80	0.61
SBERT+ResNet by minpool on SVM	0.84	0.83	0.81

# Graph Comparison of Best models

Comparison of MLP vs SVM



# SVM better than MLP?

We observe the SVM is better than the MLP in terms of accuracy since it adjusts its own hyperplane to establish the boundaries between classes

# RESULT

Best combination: SBERT (all-mpnet-base-v2) + ResNet50 on SVM

Areas of improvement / Drawbacks in the system:

- Accuracy not reliable
- Dataset is imbalanced
- Biased model

# Further Improvements

- Train with more data to address imbalance in distribution.
- Extend project to verify accuracy of generated text in image and video captioning.
- Using comments upvotes, downvotes the dataset could be further made more informative to train better models.
- Since the dataset is imbalanced, a penalty based point system can be included to overcome this.

THE END!

Thank you

