Report on Detection of Persuasive Techniques in Texts and Images

Akshayaa Udayakumar, 2021-22 AIRo, Matricula: 1988599

Abstract

The goal of this project is to study the ways in which memes can be used to spread propaganda and manipulate public opinion. By analyzing both the textual and visual content of memes, the project seeks to understand the various techniques used in these memes to persuade people to believe certain claims. To do this, deep learning and machine learning algorithms are utilized to classify the memes based on their use of these techniques.

The data used for training and testing the models has been provided, with the used techniques already defined. The models are developed using the training data and then validated using the test data. This project is divided into three subtasks, with the third task being the main focus of this project. In this task, both the textual and visual content of memes are analyzed, making it a multilabel classification task. The goal is to determine not just what the meme is saying but also how it is saying it. By combining the results of both the textual and visual analysis, a more complete picture of the techniques used in these memes can be obtained.

Task 3

The present task considers not only the text of the data but also the image of the meme, revealing the substantial impact of images on the techniques employed. In the first sample, only one technique was identified based solely on the text, but with the addition of the image, a second technique was identified. Similarly, in the second sample, the text was inadequate to determine the technique, but the image made it possible to identify the technique used. In contrast, the third sample demonstrates that some images used have no significance and do not contribute to the classification. These findings indicate that images can either positively or negatively affect the classification of techniques.

Algorithm

i. Preprocessing

The dataset of images is processed using the vision transformer.

ii. Feature Extraction

The models extract the features themselves.

iii. Apply Deep Learning Model

Both methods are used to classify the images in required techniques. These models contain many layers to process the image and classify it.

iv. Apply Machine Learning Model

For multilevel classification one vs rest classifier is used using the methods of logistic regression and random forest classifier.

v. Validation

The model is then validated on the testing dataset and training accuracy, validation accuracy, f1 score.

Vision Transformer	Training Accuracy	Val Accuracy	F1 Score
Random Forests	0.83406113537117	0.015	0.243678432291
Logistic Regression	0.98689956331877	0.02	0.305544680174