Informatics Institute of Technology
In Collaboration With
University of Westminster, UK

University of Westminster, Coat of Arms

Extenistic
Classification & Detecting Linguistics in Reviews

A dissertation by

Mr. Hikam Badurdeen
W1761913 / 2018419

Supervised by
Mr. Mithushan Jalangan
May 2023

This Project Proposal is submitted in partial fulfilment of the requirements for the BSc (Hons) Computer Science degree at
the University of Westminster.

Abstract

The project aims to develop a fake review detection system using the BERT (Bidirectional Encoder Representations from Transformers) model and various natural language processing techniques. The system preprocesses text data by tokenizing, removing stop words, and lemmatizing the reviews. It then uses BERT, a pre-trained transformer-based model, for sequence classification to distinguish between real and fake reviews.

The project incorporates BERT's tokenizer and sequence classification model, along with PyTorch and the Hugging Face Transformers library. The data is split into training and validation sets, and evaluation metrics such as accuracy, precision, recall, and F1 score are computed. The model is trained on the training set and evaluated on the validation set. The best model is saved for future use.

Furthermore, the project extends the fake review detection to classify different types of fake reviews, including incentive, competitor, and malicious reviews. Separate classification models are fine-tuned using the BERT model and trained on labeled data. The incentive, competitor, and malicious classification models are evaluated separately and their performance is assessed using evaluation metrics.

To enhance the analysis, the project also incorporates sentiment analysis using TextBlob and linguistic analysis using NLTK. The sentiment polarity of reviews is calculated, and linguistic features such as part-of-speech tagging are extracted.

The system provides a function to analyze and classify input text as real or fake. It outputs the predicted label, probability of being real, sentiment polarity, reason for classification, and classification results for different types of fake reviews. It also offers visualization features to display linguistic features, entity relations, and confusion matrices.

Overall, the project demonstrates the application of BERT-based models and natural language processing techniques for fake review detection and classification, providing insights into the authenticity of online reviews.