# A Novel Transformer-Based Deep Learning Pipeline for Multilingual Fake News Detection

Agras, Kagan (School: METU Development Foundation Private High School)
Atay, Begum (School: METU Development Foundation Private High School)

The emergence of extensive amounts of fake news on the internet made real-time fake news detection (FND) through computational tools a necessity. However, the non-availability of annotated data and language-specific processing tools significantly undermines the application of FND in low-resource languages. To address this issue, this study introduces a transformer-based pipeline that can detect fake news and expand annotated datasets in low-resource languages. 15000 news articles were translated from English to 15 different languages using Google Translate API. mBERT (multilingual Bidirectional Encoder Representations from Transformers), which was pre-trained on a corpus of 104 languages, was fine-tuned on the combined dataset and demonstrated an accuracy rate of 97%. Through cross-lingual language understanding (XLU), our model was able to learn language-independent features and outperformed monolingual models. Experiments performed on zero-shot settings indicate that our model is capable of accurately performing FND in languages that are not included in the training set, proving its effectiveness in all languages. To increase accessibility and practicality, the model is deployed on a Chrome web extension developed with Javascript. The data inputted to the web extension and the labels assigned by the model are collected in a dataset. Impractical data are removed using a filtering model. The dataset is made publicly available, which can be beneficial for future research. Overall, this study offers a novel self-expanding pipeline for FND in all languages and the expansion of annotated data in real-time, which is a significant step towards overcoming public misinformation on a global scale.

**Awards Won:**