Final Project Proposal - Social Media Misinformation Detection System

In today’s day and age rampant misinformation can be found in a variety of different avenues on the internet and social media specifically. Some of those avenues can be seen on social media as links to articles, videos, social media posts, pictures, along with other methods. The point of the matter is that misinformation is all around us on the internet and it would be nice to detect it before digesting it as trustworthy information.

In my project I plan on utilizing BERT (Bidirectional Encoder Representations from Transformers) model developed by Google to detect false information. By training Bert on the labeled data set FakeNewsNet, the hope is that examples of false and valid information can be reliably distinguished after the model has been trained and, from there, conclude how accurate the transformer-based model is or is not and why the accuracy of results being seen are what they are.

To analyze the detection performance of the model a combination of recall, precision, accuracy, and F1-scoring will be utilized. On top of this a detailed analysis will be conducted on the cases that led to the model producing incorrect results compared to the testing data. The hope is that this project will showcase how effective (or potentially ineffective) transformer models can be distinguishing between accurate information and false information and produce a case for why they should or shouldn’t be used for misinformation detection on social media sites.

Feedback:

Dear Carlos,

This proposal is excellent.

Here is a short summary of your proposal: You aim to develop a Social Media Misinformation Detection System using the BERT model. By training BERT on the FakeNewsNet dataset, you intend to distinguish between false and valid information effectively. Your evaluation will involve metrics such as recall, precision, accuracy, and F1-score, and you'll analyze cases where the model produces incorrect results.

Here are some things that are really strong about your proposal:  
- Clear identification of the problem of misinformation on social media and the need for detection.  
- Thoughtful choice of using BERT, a well-regarded transformer model, for the task.  
- Comprehensive evaluation plan using multiple metrics and error analysis.

Here are some things that could be improved or expanded on:  
- Consider specifying any pre-processing steps you might take with the FakeNewsNet dataset.  
- You might want to explore the impact of different hyperparameters on BERT's performance.

Here are some suggested corpora to look at:  
- FakeNewsNet is a great choice, and you might also consider the LIAR dataset available on Kaggle for additional training or validation.

Here are some evaluation metrics to consider:  
- In addition to the ones you've mentioned, you might want to explore the ROC-AUC score for a more comprehensive evaluation.

Here are some search terms or paper titles of research papers to consider exploring:  
- "BERT for Fake News Detection" on Google Scholar.  
- "Transformer models for misinformation detection" on arXiv.

Here is how I used the rubric to calculate your grade:  
- NLP Algorithms/Architectures/Techniques (50 points): Clear and well-chosen use of BERT for misinformation detection.  
- Application (30 points): Well-articulated application in the context of social media misinformation.  
- Evaluation (20 points): Comprehensive evaluation plan using multiple metrics.

Total points: 100.

Approach”