FYP Project Plan

Framework: PyTorch

**Approach 1**

Dataset:

* Clean and preprocess the dataset by removing unnecessary columns and invalid data.
* Split data and labels into train, test and validation
* Tokenise the data column (text or sentences) into vectors and pad the sequences to the maximum length decided and feed into the model.

Model:

* Design the model in a way such that the sarcasm detection layers and sentiment analysis layers have their own neural network layers (e.g. transformer, lstm, rnn)
* The model will output both logits that belong to the sarcasm detection layers and the sentiment analysis layers.
* Experiment with different types of neural network layers for both sarcasm detection and sentiment analysis to see which type of layers are most suitable for this task.

Training and Evaluation:

* Select suitable loss function to calculate the loss for sentiment analysis layer and sarcasm detection layer.
* Calculate weighted sum of both losses for backpropagation and update the model parameters using an optimiser.
* After training the model with the chosen number of epochs, evaluate the model’s accuracy on sarcasm detection and sentiment analysis.

Conclusion:

* To show whether sarcasm detection benefits sentiment analysis or not, the model’s result for sarcasm detection and sentiment analysis would be combined in a way that for example if a data is deemed to have a positive sentiment but also sarcastic, it will be concluded to be having a negative sentiment rather than positive.
* Results from this model would then be compared to another model that is only trained on sentiment analysis to draw conclusions.

**Approach 2**

Dataset:

* Clean and preprocess the dataset by removing unnecessary columns and invalid data.
* Split data and labels into train, test and validation
* Tokenise the data column (text or sentences) into vectors and pad the sequences to the maximum length decided and feed into the model.

Model:

* Design a sarcasm detection only model as the first step by experimenting with different types of neural network layers.
* After the sarcasm detection model is trained, it will be used as a feature extractor and extract relevant features from the intermediate layers of the model while freezing the parameters.
* After which, new layers will be added on top of the frozen layers for sentiment analysis.
* When training and fine tuning this model’s parameters, only the parameters of the new layers will be adjusted.

Training and Evaluation:

* Select suitable loss function to calculate the loss for the initial sarcasm detection only model and another loss function for the subsequent full model with sentiment analysis layers added onto it.
* Calculate weighted sum of both losses for backpropagation and update the model parameters using an optimiser (only the parameters of the newly added sentiment analysis layers will be updated in the full model).
* After training the model with the chosen number of epochs, evaluate the model’s accuracy on sentiment analysis built on top of sarcasm detection.

Conclusion:

* To show whether sarcasm detection benefits sentiment analysis or not, this model’s result would be compared to another model whose model layers are purely designed for sentiment analysis to draw conclusions.