**Report on Pre-Interview Task - Multi-label classification using transformer and BERT**

**Introduction**

The pre-interview task involved building a multi-label classification model using BERT, a transformer-based neural network, to predict the sentiment of movie reviews. The task included preprocessing the dataset, building and training the model, and making predictions.

**Preprocessing**

The dataset was preprocessed by removing segment duplication using SentenceId, initializing empty arrays to store tokenized text, and one-hot encoding the sentiment. A tf.data.Dataset object was built using input and label tensors, transformed into the correct format for the model, and batched and shuffled. The data was split into 90-10 for training and validation.

**Building the Model**

The BERT base model was initialized and a frame was built around it using typical tf.keras layers. The frame included two input layers, one for input indices and another for attention masks, a post-BERT dropout layer to reduce the likelihood of overfitting and improve generalization, a max pooling layer to convert the 3D tensor output by BERT to a 2D tensor, and final output activations using softmax to obtain the winning class prediction. The model was initialized and input and output layers were fed into the initialization function.

**Training**

The training process was initialized with suitable training parameters, including Adam optimizer with weight decay, categorical cross-entropy loss function, and categorical accuracy. The model was compiled and trained on 1 epoch for less training time. The model was saved after training.

**Predictions**

The saved model was loaded and the data was prepared. A string was taken as input and tokenized using a BertTokenizer, returning token tensors as a dictionary with 'input\_ids' and 'attention\_mask' key-value pairs. The input string was formatted into the correct dictionary-tensors format and passed into the model.predict method, which returned a Numpy array of probabilities across each output label.

**Conclusion**

The pre-interview task involved building a multi-label classification model using BERT, a transformer-based neural network, to predict the sentiment of movie reviews. The task required preprocessing the dataset, building and training the model, and making predictions. The task provided a good understanding of BERT and its applications in natural language processing.