**Sentiment Analysis on Yelp Reviews using LSTM and DistilBERT**

This project performs multiclass sentiment analysis (positive, neutral, negative) on Yelp reviews using two deep learning models: \*\*LSTM\*\* and \*\*DistilBERT\*\*.

It includes training, hyperparameter optimization (HPO), and testing notebooks for both models.

All required dependencies are listed within each respective notebook.

## **General Pipeline**

1. **Data Preparation**
   * Load yelp\_dataset\_train.csv and yelp\_dataset\_test.csv
2. **Hyperparameter Optimization**
   * Run the correct HPO.ipynb on each model to find optimal training parameters using a 10% sample of the training set
3. **Training**
   * Use the selected hyperparameters to train the models on the full training set in their respective training.ipynb
4. **Evaluation**
   * Run the respective testing.ipynb to test the model, generate metrics, confusion matrices, and plots
5. **Analysis**
   * Run additional\_testing.ipynb (included in each model folder) for:
     + LIME and vocabulary visualization (LSTM & DistilBERT)
     + Attention weight heatmaps (DistilBERT only)