### Model Card: Sentiment Analysis on Earnings Call Transcripts

#### 1. **Model Overview**

* **Task**: Multiclass classification for sentiment analysis on earnings call transcripts, to predict the impact of earning call on stock price: positive, neutral, negative
* **Models Used**:
  + **Logistic Regression**
  + **Support Vector Machine (SVM)**
  + **Feedforward Neural Network (FNN)**
  + **Gated Graph Neural Network (GNN)**

#### 2. **Datasets**

* **Input Features**: Derived from text embeddings generated using a pre-trained language model (RoBERTa).
* **Labels**: negative, neutral, positive.

#### 3. **Model Architectures**

* **Logistic Regression**:
  + Simple linear model used as a baseline.
* **Support Vector Machine (SVM)**:
  + Linear kernel used for classification.
* **Feedforward Neural Network (FNN)**:
  + **Layers**: 1 hidden layer with 128 units.
  + **Activation**: ReLU.
  + **Loss Function**: Cross-entropy loss.
* **Gated Graph Neural Network (GNN)**:
  + **Graph Layer**: Graph Convolutional Network (GCN).
  + **Recurrent Layer**: GRU (Gated Recurrent Unit).
  + **Dropout**: Applied after each layer to prevent overfitting.

#### 4. **Fine-Tuning and Optimization**

* **Optuna**:
  + Used for hyperparameter optimization for both FNN and GNN models.
  + Parameters tuned: hidden\_dim, learning\_rate, num\_epochs
* **Metrics**:
  + **F1 Score**: Used alongside accuracy to ensure balanced performance across all classes.
* **Handling Class Imbalance**:
  + Introduced dropout 50% to GNN to mitigate overfitting and improve generalization.

#### 5. **Results**

* **FNN and GNN Performance**:
  + FNN with below set up had the best result: with F1 score of 38.27% on test set  
    hidden\_dim: 128

learning\_rate: 0.00256106251118606

num\_epochs: 300

* + Best GNN model from Optuna achieved F1 Score of 37% on train data, but the F1 score dropped to 13% on test data, indicating potential overfitting. We thus introduced dropout of 50% to get superior generalization on the test set, with a F1 score of 37% on test set.

|  |  |  |
| --- | --- | --- |
| **Model** | **F1 Score** | **Fine tuning method** |
| FNN | 38.27 | Optuna |
| GNN | 36.98 | Dropout |
| SVM | 35.72 | Grid Search |
| Logistic Regression | 37.09 | No optimisation |

#### 6. **Challenges and Mitigation**

* **Model Convergence Issues**:
  + For logistic regression, convergence warnings were addressed by increasing max\_iter, applying feature scaling, and adjusting the solver and regularization parameters.
* **Class Imbalance**:
  + GNN models struggled with class imbalance, leading to poor F1 scores on the minority class. Class weights and dropout were introduced to address this.

#### 7. **Model Save and Reuse**

* **Best Model Saving**:
  + The best model found through Optuna was saved using torch.save() to ensure consistent results upon re-running.
* **Re-evaluation**:
  + The saved model was reloaded and re-evaluated on the test set to confirm performance metrics.