Dependencies

To run the provided code successfully, the following Python libraries and frameworks are required. Below is a list of dependencies along with their installation instructions.

3.1. Core Dependencies

1. Python 3.x

- The code is written in Python, so a compatible version of Python (3.7 or higher) is required.

2. PyTorch

- PyTorch is used for deep learning model training and dataset handling.

- Install PyTorch using the following command (choose the appropriate version for your system):

Command:pip install torch

3. Pandas

- Pandas is used for data manipulation and handling CSV files.

- Install Pandas using:

pip install pandas

4. NumPy

- NumPy is used for numerical computations and is often required by other libraries.

- Install NumPy using:

5. Scikit-learn

- Scikit-learn is used for label encoding and metrics calculation.

- Install Scikit-learn using:

pip install scikit-learn

3.2. Transformers Library

6. Hugging Face Transformers

- The `transformers` library is used for pre-trained models, tokenizers, and training utilities.

- Install the library using:

pip install transformers

3.3. Additional Dependencies

Datasets

- The `datasets` library is used for dataset handling and preprocessing.

- Install the library using:

pip install datasets

8. SHAP (SHapley Additive exPlanations)

- SHAP is used for model explainability and interpretability.

- Install SHAP using:

pip install shap

9. tqdm

- tqdm is used for progress bars during training (optional but recommended).

- Install tqdm using:

pip install tqdm

3.4. Optional Dependencies

10. Matplotlib/Seaborn

- These libraries are used for visualizing metrics or data (optional).

- Install them using:

pip install matplotlib seaborn

3.5. Installation Command

To install all the required dependencies in one go, use the following command:

pip install torch pandas numpy scikit-learn transformers datasets shap tqdm

3.6. Summary of Dependencies

| Library | Purpose |

|-------------------------|-----------------------------------------------------------------------------|

| Python 3.x | Programming language. |

| PyTorch | Deep learning framework. |

| Pandas | Data manipulation and CSV handling. |

| NumPy | Numerical computations. |

| Scikit-learn | Label encoding and metrics calculation. |

| Transformers | Pre-trained models, tokenizers, and training utilities. |

| Datasets | Dataset handling and preprocessing. |

| SHAP | Model explainability and interpretability. |

| tqdm | Progress bars during training. |

| Matplotlib/Seaborn | Visualization of metrics or data (optional). |