

Project Dependencies and Requirements

Machine Learning Text Classification Project

March 27, 2025

1 BERT Classification Script (`bert_classification.py`)

1.1 Python Dependencies

Library	Minimum Version
torch	1.7.0
transformers	4.10.0
pandas	1.3.0
numpy	1.21.0
scikit-learn	0.24.0
re	Standard Library
random	Standard Library

1.2 Hardware Requirements

- Recommended: CUDA-enabled GPU with at least 8GB VRAM
- Minimum: CPU with multi-core processing
- Recommended RAM: 16GB or more

2 Naive Bayes Classification Script (`br_classification_optimized.py`)

2.1 Python Dependencies

Library	Minimum Version
pandas	1.3.0
numpy	1.21.0
scikit-learn	0.24.0
nltk	3.6.0
imbalanced-learn	0.8.0
re	Standard Library
string	Standard Library

2.2 Hardware Requirements

- Minimum: Multi-core CPU
- Recommended RAM: 8GB or more

3 Installation Instructions

```
# Recommended: Create a virtual environment
python -m venv ml_textclassification_env
source ml_textclassification_env/bin/activate # On Windows: ml_textclassification_env\Scripts\activ

# Install dependencies
pip install torch transformers pandas numpy scikit-learn nltk imbalanced-learn
python -m nltk.downloader stopwords wordnet
```