

Enhanced Bug Report Classifier - Requirements

This document lists all dependencies and their versions required to run the Enhanced Bug Report Classifier.

Python Version

- Python 3.6 or higher

Core Dependencies

| Package | Version | Purpose | |-----|-----|-----| | numpy | >=1.19.5 | Numerical operations and data manipulation | | pandas | >=1.1.5 | Data loading and manipulation | | scikit-learn | >=0.24.2 | Machine learning algorithms and evaluation metrics | | scipy | >=1.5.4 | Scientific computing and statistical tests | | nltk | >=3.6.2 | Natural language processing and text preprocessing | | gensim | >=4.0.1 | Word embeddings |

Visualization Dependencies

| Package | Version | Purpose | |-----|-----|-----| | matplotlib | >=3.3.4 | Basic plotting and visualization | | seaborn | >=0.11.1 | Enhanced statistical visualizations | | tabulate | >=0.8.9 | Formatted tabular data display |

Installation

You can install all dependencies using the following command:

```
pip install numpy>=1.19.5 pandas>=1.1.5 scikit-learn>=0.24.2
scipy>=1.5.4 nltk>=3.6.2 gensim>=4.0.1 matplotlib>=3.3.4
seaborn>=0.11.1 tabulate>=0.8.9
```

Or, if you prefer using a requirements.txt file:

```
pip install -r requirements.txt
```

With the following content in requirements.txt:

```
numpy>=1.19.5
pandas>=1.1.5
scikit-learn>=0.24.2
scipy>=1.5.4
nltk>=3.6.2
gensim>=4.0.1
matplotlib>=3.3.4
seaborn>=0.11.1
tabulate>=0.8.9
```

NLTK Resources

The following NLTK resources need to be downloaded:

- stopwords
- punkt
- wordnet

You can download them using the following Python code:

```
import nltk
nltk.download('stopwords')
nltk.download('punkt')
nltk.download('wordnet')
```

Word Embeddings

The classifier can optionally use pre-trained word embeddings. By default, it uses 'glove-wiki-gigaword-100' from the Gensim library, which will be automatically downloaded when first used.

Hardware Requirements

For optimal performance, especially when using word embeddings:

- RAM: 8GB minimum, 16GB or more recommended
- CPU: Multi-core processor recommended
- Disk Space: At least 2GB free space (mostly for word embeddings)

Environment Setup

It is recommended to set up a virtual environment to avoid conflicts with other Python packages:

```
# Create a virtual environment
python -m venv venv
# Activate the virtual environment
# For Windows
venv\Scripts\activate
# For Linux/macOS
source venv/bin/activate
# Install dependencies
pip install -r requirements.txt
```