Requirements

```
Requirements for Bug Report Classification Tool
Last Updated: August 2023
Python Version: 3.8+

1. Core Python Packages
Package Version    Purpose
scikit-learn    1.0.2  Model training, evaluation, and metrics.
nltk    3.7    Text preprocessing, tokenization, and VADER sentiment
analysis.
gensim  4.2.0  Word2Vec embeddings.
imbalanced-learn    0.9.0  SMOTE for class balancing.
pandas  1.4.3  Data loading, preprocessing, and analysis.
numpy   1.23.1 Numerical operations.
scipy   1.9.1  Statistical tests (Wilcoxon, Cliff's Delta).
2. NLP & Text Processing
Package Version    Notes
spacy   3.4.1  Lemmatization and advanced tokenization.
en-core-web-sm  3.4.0  spaCy's small English model. Install via: python -m
spacy download en_core_web_sm
tqdm    4.64.1 Progress bars for data processing.
3. Visualization
Package Version    Purpose
matplotlib  3.5.3  Generating plots (precision-recall curves).
seaborn 0.11.2 Enhanced visualizations (metric comparisons).
4. Workflow & Reproducibility
Package Version    Purpose
jupyter 1.0.0  Running analysis notebooks.
requests    2.28.1 Fetching data via GitHub API.
ipykernel   6.19.2 Kernel support for Jupyter notebooks.
5. Critical Version-Specific Notes
imbalanced-learn==0.9.0:

Newer versions (≥0.10+) may alter SMOTE's sampling behavior, causing
reproducibility issues.

gensim==4.2.0:

Compatibility fix for training Word2Vec on Python 3.8+.

scikit-learn==1.0.2:

Required for class_weight="balanced" in Random Forests.

6. Installation
Create a virtual environment:
python -m venv venv
source venv/bin/activate  # Linux/Mac
venv\Scripts\activate     # Windows

Install packages:
pip install -r requirements.txt

Download spaCy model:
python -m spacy download en_core_web_sm
```