

Requirements

Datasets

All the datasets from <https://github.com/ideas-labo/ISE-solution/tree/main/lab1/datasets> must be downloaded and placed into an **datasets/** directory in the project root.

Embeddings

“wiki-news-300d-1M.vec.zip” must be downloaded from <https://fasttext.cc/docs/en/english-vectors.html>, and placed into an **embeddings/** directory in the project root.

Dependencies

The following dependencies are required to run the code (these can also be found in ‘requirements.txt’). These can be installed using pip.

- **keras**
- **keras-tuner**
- **gensim**
- **numpy**
- **pandas**
- **nltk**
- **scikit-learn**
- **imbalanced-learn**
- **tensorflow**
- **scipy**

GPU Dependencies

This tool has been developed using TensorFlow as the backend to Keras.

A GPU is recommended, but not mandatory.

Generally, TensorFlow with cuda can be installed using

- **`python3 -m pip install 'tensorflow[and-cuda]'`**

Platform specific instructions can be found here: <https://www.tensorflow.org/install/pip>

Notes

This was developed and tested on Linux kernel version 6.9.3 with Python 3.10.12.