

# **CS 516: Information Retrieval and Text Mining**

Information Technology University (ITU)  
Fall 2025

Course Instructor: Dr. Ahmad Mustafa

## **Homework Assignment 3**

**Due Date: 11.59 pm, 30th November, 2025**

### **Design Problem**

In this assignment, you will design and implement a complete information retrieval (IR) system that runs locally on your machine. You may use any retrieval strategy (or combination of strategies), such as Boolean retrieval, Vector Space Models (e.g., TF-IDF) etc.

Your goal is to design a coherent, justifiable, and well-evaluated retrieval system using the provided dataset of text documents (accessible [here](#)).

Creativity is encouraged, but your system must be reproducible and fully local (no cloud-hosted vector databases).

### **Requirements**

You are free to design your IR system however you like, as long as it satisfies the following:

#### **1. Local Implementation**

Your system must run end-to-end on a local machine (Windows, Mac, or Linux).  
You may use any mainstream programming language or libraries.

Cloud-hosted vector databases (e.g., Pinecone, Chroma Cloud, Weaviate Cloud, Elasticsearch clusters) are **not** allowed.

Local libraries (e.g., scikit-learn, gensim, rank-bm25, FAISS local install) are allowed.

```
Windows PowerShell
whl.metadata (11 kB)
Collecting nltk
  Downloading nltk-3.9.2-py3-none-any.whl.metadata (3.2 kB)
Collecting spacy
  Downloading spacy-3.8.11-cp312-cp312-win_amd64.whl.metadata (28 kB)
Collecting gensim
  Using cached gensim-4.4.0-cp312-cp312-win_amd64.whl.metadata (8.6 kB)
Collecting jupyterlab
  Downloading jupyterlab-4.5.0-py3-none-any.whl.metadata (16 kB)
Collecting python-dateutil
  Using cached python_dateutil-2.8.2_(from pandas).whl.metadata (8.4 kB)
Collecting pytz>=2020.1 (from pandas)
  Using cached pytz-2025.2-py2.py3-none-any.whl.metadata (22 kB)
Collecting tzdata>=2022.7 (from pandas)
  Using cached tzdata-2025.2-py2.py3-none-any.whl.metadata (1.4 kB)
Collecting scikit-learn
  Using cached scikit-learn-1.1.6.3-cp312-cp312-win_amd64.whl.metadata (60 kB)
Collecting joblib>=1.2.0 (from scikit-learn)
  Using cached joblib-1.5.2-py3-none-any.whl.metadata (5.6 kB)
Collecting threadpoolctl>=3.1.0 (from scikit-learn)
  Using cached threadpoolctl-3.6.0-py3-none-any.whl.metadata (13 kB)
Collecting click (from nltk)
  Using cached click-8.3.1-py3-none-any.whl.metadata (2.6 kB)
Collecting regex>=2021.8.3 (from nltk)
  Downloading regex-2025.11.3-cp312-cp312-win_amd64.whl.metadata (41 kB)
Collecting tqdm (from nltk)
  Using cached tqdm-4.67.1-py3-none-any.whl.metadata (57 kB)
Collecting spacy-legacy<3.1.0,>=3.0.11 (from spacy)
  Downloading spacy_legacy-3.0.12-py2.py3-none-any.whl.metadata (2.8 kB)
Collecting spacy-loggers<2.0.0,>=1.0.0 (from spacy)
  Downloading spacy_loggers-1.0.5-py3-none-any.whl.metadata (23 kB)
Collecting murmurhash<1.0,>=0.28.0 (from spacy)
  Downloading murmurhash-1.0.15-cp312-cp312-win_amd64.whl.metadata (2.3 kB)
Collecting cymem<2.1.0,>=2.0.2 (from spacy)
  Downloading cymem-2.0.13-cp312-cp312-win_amd64.whl.metadata (9.9 kB)
Collecting preshed<3.1.0,>=3.0.2 (from spacy)
  Downloading preshed-3.0.12-cp312-cp312-win_amd64.whl.metadata (2.6 kB)
Collecting thinc<8.4.0,>=8.3.4 (from spacy)
  Downloading thinc-8.3.10-cp312-cp312-win_amd64.whl.metadata (15 kB)
Collecting wasabi<1.2.0,>=0.9.1 (from spacy)
  Downloading wasabi-1.1.3-py3-none-any.whl.metadata (28 kB)
Collecting srslly<3.0.0,>=2.4.3 (from spacy)
```

```
Windows PowerShell
Using cached annotated_types-0.7.0-py3-none-any.whl.metadata (15 kB)
Collecting pydantic-core==2.41.5 (from pydantic!=1.8,!~1.8.1,<3.0.0,>=1.7.4->spacy)
  Downloading pydantic_core-2.41.5-cp312-cp312-win_amd64.whl.metadata (7.4 kB)
Collecting typing_extensions>=4.14.1 (from pydantic!=1.8,!~1.8.1,<3.0.0,>=1.7.4->spacy)
  Using cached typing_extensions-4.15.0-py3-none-any.whl.metadata (3.3 kB)
Collecting typing_inspections>=0.4.2 (from pydantic!=1.8,!~1.8.1,<3.0.0,>=1.7.4->spacy)
  Using cached typing_inspection-0.4.2-py3-none-any.whl.metadata (2.6 kB)
Collecting charset_normalizer<4,>=2 (from requests<3.0.0,>=2.13.0->spacy)
  Downloading charset_normalizer-3.4.4-cp312-cp312-win_amd64.whl.metadata (38 kB)
Collecting idna<4,>=2.5 (from requests<3.0.0,>=2.13.0->spacy)
  Using cached idna-3.11-py3-none-any.whl.metadata (8.4 kB)
Collecting urllib3<3,>=1.21.1 (from requests<3.0.0,>=2.13.0->spacy)
  Using cached urllib3-2.5.0-py3-none-any.whl.metadata (6.5 kB)
Collecting certifi>=2017.4.17 (from requests<3.0.0,>=2.13.0->spacy)
  Using cached certifi-2025.11.12-py3-none-any.whl.metadata (2.5 kB)
Collecting blis<1.4.0,>=1.3.0 (from thinc<8.4.0,>=8.3.4->spacy)
  Downloading blis-1.3.3-cp312-cp312-win_amd64.whl.metadata (7.7 kB)
Collecting confection<1.0.0,>=0.0.1 (from thinc<8.4.0,>=8.3.4->spacy)
  Downloading confection-0.1.5-py3-none-any.whl.metadata (19 kB)
Collecting colorama (from tqdm->nltk)
  Using cached colorama-0.4.6-py2.py3-none-any.whl.metadata (17 kB)
Collecting cloudpathlib<1.0.0,>=0.7.0 (from weasel<0.5.0,>=0.4.2->spacy)
  Downloading cloudpathlib-0.23.0-py3-none-any.whl.metadata (16 kB)
Collecting smart-open<8.0.0,>=5.2.1 (from weasel<0.5.0,>=0.4.2->spacy)
  Using cached smart_open-7.5.0-py3-none-any.whl.metadata (24 kB)
Collecting wrapt (from smart-open<8.0.0,>=5.2.1->weasel<0.5.0,>=0.4.2->spacy)
  Using cached wrapt-2.0.1-cp312-cp312-win_amd64.whl.metadata (9.2 kB)
Collecting async_trio=1.0.0 (from jupyterlab)
  Downloading async_trio-2.0.5-py3-none-any.whl.metadata (4.5 kB)
Collecting httpx<1,>=0.25.0 (from jupyterlab)
  Using cached httpx-0.28.1-py3-none-any.whl.metadata (7.1 kB)
Collecting ipykernel!=6.30.0,>=6.5.0 (from jupyterlab)
  Downloading ipykernel-7.1.0-py3-none-any.whl.metadata (4.5 kB)
Collecting jupyter_core (from jupyterlab)
  Downloading jupyter_core-5.9.1-py3-none-any.whl.metadata (1.5 kB)
Collecting jupyter_lsp>=2.0.0 (from jupyterlab)
  Downloading jupyter_lsp-2.3.0-py3-none-any.whl.metadata (1.8 kB)
Collecting jupyter_server<3,>=2.4.0 (from jupyterlab)
  Downloading jupyter_server-2.17.0-py3-none-any.whl.metadata (8.5 kB)
Collecting jupyterlab_servers<3,>=2.28.0 (from jupyterlab)
  Downloading jupyterlab_server-2.28.0-py3-none-any.whl.metadata (5.9 kB)
Collecting notebook_shim>=0.2 (from jupyterlab)
```

```
Windows PowerShell
Collecting jupyter-client>=7.4.4 (from jupyter-server<3,>=2.4.0->jupyterlab)
  Downloading jupyter_client-8.6.3-py3-none-any.whl.metadata (8.3 kB)
Collecting jupyter-events>=0.11.0 (from jupyter-server<3,>=2.4.0->jupyterlab)
  Downloading jupyter_events-0.12.0-py3-none-any.whl.metadata (5.8 kB)
Collecting jupyter-server-terminals>=0.4.4 (from jupyter-server<3,>=2.4.0->jupyterlab)
  Downloading jupyter_server_terminals-0.5.3-py3-none-any.whl.metadata (5.6 kB)
Collecting nbconvert>=6.4.4 (from jupyter-server<3,>=2.4.0->jupyterlab)
  Downloading nbconvert-7.16.6-py3-none-any.whl.metadata (8.5 kB)
Collecting nbformat>=5.3.0 (from jupyter-server<3,>=2.4.0->jupyterlab)
  Downloading nbformat-5.10.4-py3-none-any.whl.metadata (3.6 kB)
Collecting prometheus-client>=0.9 (from jupyter-server<3,>=2.4.0->jupyterlab)
  Downloading prometheus_client-0.23.1-py3-none-any.whl.metadata (1.9 kB)
Collecting pwintpy>=2.0.1 (from jupyter-server<3,>=2.4.0->jupyterlab)
  Downloading pwintpy-3.0.2-cp312-cp312-win_amd64.whl.metadata (5.7 kB)
Collecting pyzmq>=24 (from jupyter-server<3,>=2.4.0->jupyterlab)
  Downloading pyzmq-27.1.0-cp312-abi3-win_amd64.whl.metadata (8.5 kB)
Collecting send2trash>=1.8.2 (from jupyter-server<3,>=2.4.0->jupyterlab)
  Downloading Send2Trash-1.8.3-py3-none-any.whl.metadata (4.0 kB)
Collecting terminado>=0.8.3 (from jupyter-server<3,>=2.4.0->jupyterlab)
  Downloading terminado-0.18.1-py3-none-any.whl.metadata (5.8 kB)
Collecting websocket-client>=1.7 (from jupyter-server<3,>=2.4.0->jupyterlab)
  Downloading websocket_client-1.9.0-py3-none-any.whl.metadata (8.3 kB)
Collecting babel>=2.10 (from jupyterlab-server<3,>=2.28.0->jupyterlab)
  Downloading babel-2.17.0-py3-none-any.whl.metadata (2.0 kB)
Collecting json5>=0.9.0 (from jupyterlab-server<3,>=2.28.0->jupyterlab)
  Downloading json5-0.12.1-py3-none-any.whl.metadata (36 kB)
Collecting jsonschema>=4.18.0 (from jupyterlab-server<3,>=2.28.0->jupyterlab)
  Downloading jsonschema-4.25.1-py3-none-any.whl.metadata (7.6 kB)
Collecting argon2-cffi-bindings (from argon2-cffi>=21.1->jupyter-server<3,>=2.4.0->jupyterlab)
  Downloading argon2_cffi-bindings-25.1.0-cp39-abi3-win_amd64.whl.metadata (7.5 kB)
Collecting comm>=0.1.1 (from ipykernel!=6.30.0,>=6.5.0->jupyterlab)
  Downloading comm-0.2.3-py3-none-any.whl.metadata (3.7 kB)
Collecting debugpy>=1.6.5 (from ipykernel!=6.30.0,>=6.5.0->jupyterlab)
  Downloading debugpy-1.8.17-cp312-cp312-win_amd64.whl.metadata (1.5 kB)
Collecting ipython>=7.23.1 (from ipykernel!=6.30.0,>=6.5.0->jupyterlab)
  Downloading ipython-9.7.0-py3-none-any.whl.metadata (4.5 kB)
Collecting matplotlib-inline>=0.1 (from ipykernel!=6.30.0,>=6.5.0->jupyterlab)
  Downloading matplotlib_inline-0.2.1-py3-none-any.whl.metadata (2.3 kB)
Collecting nest-asyncio>=1.4 (from ipykernel!=6.30.0,>=6.5.0->jupyterlab)
  Downloading nest_asyncio-1.6.0-py3-none-any.whl.metadata (2.8 kB)
Collecting psutil>=5.7 (from ipykernel!=6.30.0,>=6.5.0->jupyterlab)
  Downloading psutil-3.13.0-cp37-abi3-win_amd64.whl.metadata (23 kB)
```

```
Windows PowerShell
Downloading ipython_pygments_lexers-1.1.1-py3-none-any.whl.metadata (1.1 kB)
Collecting jedi>=0.18.1 (from ipython>=7.23.1->ipykernel!=6.30.0,>=6.5.0->jupyterlab)
  Downloading jedi-0.19.2-py2.py3-none-any.whl.metadata (22 kB)
Collecting prompt_toolkit<3.1.0>=3.0.41 (from ipython>=7.23.1->ipykernel!=6.30.0,>=6.5.0->jupyterlab)
  Downloading prompt_toolkit-3.0.52-py3-none-any.whl.metadata (6.4 kB)
Collecting pygments>=2.11.0 (from ipython>=7.23.1->ipykernel!=6.30.0,>=6.5.0->jupyterlab)
  Downloading pygments-2.19.2-py3-none-any.whl.metadata (2.5 kB)
Collecting stack_data>=0.6.0 (from ipython>=7.23.1->ipykernel!=6.30.0,>=6.5.0->jupyterlab)
  Downloading stack_data-0.6.3-py3-none-any.whl.metadata (18 kB)
Collecting wccwidth (from prompt_toolkit<3.1.0,>=3.0.41->ipython>=7.23.1->ipykernel!=6.30.0,>=6.5.0->jupyterlab)
  Downloading wccwidth-0.2.14-py2.py3-none-any.whl.metadata (15 kB)
Collecting parsos>=0.9.0,>=0.8.4 (from jedi>=0.18.1->ipython>=7.23.1->ipykernel!=6.30.0,>=6.5.0->jupyterlab)
  Downloading parsos-0.8.5-py2.py3-none-any.whl.metadata (8.3 kB)
Collecting MarkupSafe>=2.0 (from jinja2->spacy)
  Downloading markupsafe-3.0.3-cp312-cp312-win_amd64.whl.metadata (2.8 kB)
Collecting attrs>=22.2.0 (from jsonschema>=4.18.0->jupyterlab-server<3,>=2.28.0->jupyterlab)
  Using cached attrs-25.4.0-py3-none-any.whl.metadata (10 kB)
Collecting jsonschema-specifications>=2023.03.6 (from jsonschema>=4.18.0->jupyterlab-server<3,>=2.28.0->jupyterlab)
  Downloading jsonschema_specifications-2025.9.1-py3-none-any.whl.metadata (2.9 kB)
Collecting referencing>=0.28.4 (from jsonschema>=4.18.0->jupyterlab-server<3,>=2.28.0->jupyterlab)
  Downloading referencing-0.37.0-py3-none-any.whl.metadata (2.8 kB)
Collecting rpds-py>=0.7.1 (from jsonschema>=4.18.0->jupyterlab-server<3,>=2.28.0->jupyterlab)
  Downloading rpds_py-0.30.0-cp312-cp312-win_amd64.whl.metadata (4.2 kB)
Collecting platformdirs>=2.5 (from jupyter-core->jupyterlab)
  Downloading platformdirs-4.5.0-py3-none-any.whl.metadata (12 kB)
Collecting python_json_logger>=2.0.4 (from jupyter-events>=0.11.0->jupyter-server<3,>=2.4.0->jupyterlab)
  Downloading python_json_logger-4.0.0-py3-none-any.whl.metadata (4.0 kB)
Collecting pyyaml>=5.3 (from jupyter-events>=0.11.0->jupyter-server<3,>=2.4.0->jupyterlab)
  Downloading pyyaml-6.0.3-cp312-cp312-win_amd64.whl.metadata (2.4 kB)
Collecting rfc3339-validator (from jupyter-events>=0.11.0->jupyter-server<3,>=2.4.0->jupyterlab)
  Downloading rfc3339_validator-0.1.4-py2.py3-none-any.whl.metadata (1.5 kB)
Collecting rfc3986_validator>=0.1.1 (from jupyter-events>=0.11.0->jupyter-server<3,>=2.4.0->jupyterlab)
  Downloading rfc3986_validator-0.1.1-py2.py3-none-any.whl.metadata (1.7 kB)
Collecting fqdn (from jsonschema[format-nongpl]>=4.18.0->jupyter-events>=0.11.0->jupyter-server<3,>=2.4.0->jupyterlab)
  Downloading fqdn-1.5.1-py3-none-any.whl.metadata (1.4 kB)
Collecting isoduration (from jsonschema[format-nongpl]>=4.18.0->jupyter-events>=0.11.0->jupyter-server<3,>=2.4.0->jupyterlab)
  Downloading isoduration-20.11.0-py3-none-any.whl.metadata (5.7 kB)
Collecting jsonpointer>=1.13 (from jsonschema[format-nongpl]>=4.18.0->jupyter-events>=0.11.0->jupyter-server<3,>=2.4.0->jupyterlab)
  Using cached jsonpointer-3.0.0-py2.py3-none-any.whl.metadata (2.3 kB)
Collecting rfc3987-syntax>=1.1.0 (from jsonschema[format-nongpl]>=4.18.0->jupyter-events>=0.11.0->jupyter-server<3,>=2.4.0->jupyterlab)
  Downloading rfc3987_syntax-1.1.0-py3-none-any.whl.metadata (7.7 kB)
```

```
Windows PowerShell
)   Downloading rfc3987_syntax-1.1.0-py3-none-any.whl.metadata (7.7 kB)
Collecting uri-template (from jsonschema[format-nongpl]>=4.18.0->jupyter-events>=0.11.0->jupyter-server<3,>=2.4.0->jupyterlab)
  Downloading uri_template-1.3.0-py3-none-any.whl.metadata (8.8 kB)
Collecting webcolors>=24.6.0 (from jsonschema[format-nongpl]>=4.18.0->jupyter-events>=0.11.0->jupyter-server<3,>=2.4.0->jupyterlab)
  Downloading webcolors-25.10.0-py3-none-any.whl.metadata (2.2 kB)
Collecting beautifulsoup4 (from nbconvert>=6.4.4->jupyter-server<3,>=2.4.0->jupyterlab)
  Downloading beautifulsoup4-4.14.3-py3-none-any.whl.metadata (3.8 kB)
Collecting bleach!=5.0.0 (from bleach[css]!=5.0.0->nbconvert=>6.4.4->jupyter-server<3,>=2.4.0->jupyterlab)
  Downloading bleach-6.3.0-py3-none-any.whl.metadata (31 kB)
Collecting defusedxml (from nbconvert=>6.4.4->jupyter-server<3,>=2.4.0->jupyterlab)
  Downloading defusedxml-0.7.1-py2.py3-none-any.whl.metadata (32 kB)
Collecting jupyterlab_pygments (from nbconvert=>6.4.4->jupyter-server<3,>=2.4.0->jupyterlab)
  Downloading jupyterlab_pygments-0.3.0-py3-none-any.whl.metadata (4.4 kB)
Collecting mistune<4,>=2.0.3 (from nbconvert=>6.4.4->jupyter-server<3,>=2.4.0->jupyterlab)
  Downloading mistune-3.1.4-py3-none-any.whl.metadata (1.8 kB)
Collecting nbclient>=0.5.0 (from nbconvert=>6.4.4->jupyter-server<3,>=2.4.0->jupyterlab)
  Downloading nbclient-0.10.2-py3-none-any.whl.metadata (8.3 kB)
Collecting pandocfilters>=1.4.1 (from nbconvert=>6.4.4->jupyter-server<3,>=2.4.0->jupyterlab)
  Downloading pandocfilters-1.5.1-py2.py3-none-any.whl.metadata (9.0 kB)
Collecting webencodings (from bleach!=5.0.0->nbconvert=>6.4.4->jupyter-server<3,>=2.4.0->jupyterlab)
  Downloading webencodings-0.5.1-py2.py3-none-any.whl.metadata (2.1 kB)
Collecting tinyccss2<1.5,>=1.1.0 (from bleach[css]!=5.0.0->nbconvert=>6.4.4->jupyter-server<3,>=2.4.0->jupyterlab)
  Downloading tinyccss2-1.4.0-py3-none-any.whl.metadata (3.0 kB)
Collecting fastjsonschema>=2.15 (from nbformat=>5.3.0->jupyter-server<3,>=2.4.0->jupyterlab)
  Downloading fastjsonschema-2.21.2-py3-none-any.whl.metadata (2.3 kB)
Collecting six<1.5 (from python-dateutil)>=2.8.2->pandas)
  Using cached six-1.17.0-py2.py3-none-any.whl.metadata (1.7 kB)
Collecting lark>=1.2.2 (from rfc3987_syntax=>1.1.0->jsonschema[format-nongpl]>=4.18.0->jupyter-events>=0.11.0->jupyter-server<3,>=2.4.0->jupyterlab)
  Downloading lark-1.3.1-py3-none-any.whl.metadata (1.8 kB)
Collecting executing=>1.2.0 (from stack_data=>0.6.0->ipython>=7.23.1->ipykernel!=6.30.0,>=6.5.0->jupyterlab)
  Downloading executing-2.2.1-py2.py3-none-any.whl.metadata (8.9 kB)
Collecting asttokens=>2.1.0 (from stack_data=>0.6.0->ipython>=7.23.1->ipykernel!=6.30.0,>=6.5.0->jupyterlab)
  Downloading asttokens-3.0.1-py3-none-any.whl.metadata (4.9 kB)
Collecting pure_eval (from stack_data=>0.6.0->ipython>=7.23.1->ipykernel!=6.30.0,>=6.5.0->jupyterlab)
  Downloading pure_eval-0.2.3-py3-none-any.whl.metadata (6.3 kB)
Collecting cffi>=1.0.1 (from argon2-cffi-bindings->argon2-cffi=>21.1->jupyter-server<3,>=2.4.0->jupyterlab)
  Downloading cffi-2.0.0-cp312-cp312-win_amd64.whl.metadata (2.6 kB)
Collecting pycparser (from cffi=>1.0.1->argon2-cffi-bindings->argon2-cffi=>21.1->jupyter-server<3,>=2.4.0->jupyterlab)
  Downloading pycparser-2.23-py3-none-any.whl.metadata (993 bytes)
Collecting soupsieve=>1.6.1 (from beautifulsoup4->nbconvert=>6.4.4->jupyter-server<3,>=2.4.0->jupyterlab)
```

```
Windows PowerShell
)   Downloading soupsieve=1.6.1 (from beautifulsoup4->nbconvert=>6.4.4->jupyter-server<3,>=2.4.0->jupyterlab)
Using cached soupsieve-2.8-py3-none-any.whl.metadata (4.6 kB)
Collecting arrow>=0.15.0 (from isoduration->jsonschema[format-nongpl]>=4.18.0->jupyter-events>=0.11.0->jupyter-server<3,>=2.4.0->jupyterlab)
  Downloading arrow-1.4.0-py3-none-any.whl.metadata (7.7 kB)
Using cached numpy-2.3.5-cp312-cp312-win_amd64.whl (12.8 MB)
Using cached pandas-2.3.3-cp312-cp312-win_amd64.whl (11.0 MB)
Using cached scikit_learn-1.7.2-cp312-cp312-win_amd64.whl (8.7 MB)
  Downloading nltk-3.9.2-py3-none-any.whl (1.5 MB)
  - 1.5/1.5 MB 1.3 MB/s 0:00:01
  Downloading spacy-3.8.11-cp312-cp312-win_amd64.whl (14.2 MB)
  - 14.2/14.2 MB 2.6 MB/s 0:00:05
  Downloading catalogue-2.0.10-py3-none-any.whl (17 kB)
  Downloading cymem-2.0.13-cp312-cp312-win_amd64.whl (40 kB)
  Downloading murmurhash-1.0.15-cp312-cp312-win_amd64.whl (25 kB)
  Downloading preshed-3.0.12-cp312-cp312-win_amd64.whl (118 kB)
  Downloading pydantic-2.12.5-py3-none-any.whl (463 kB)
  Downloading pydantic_core-2.41.5-cp312-cp312-win_amd64.whl (2.0 MB)
  - 2.0/2.0 MB 1.2 MB/s 0:00:01
Using cached requests-2.32.5-py3-none-any.whl (64 kB)
  Downloading charset_normalizer-3.4.4-cp312-cp312-win_amd64.whl (107 kB)
Using cached idna-3.11-py3-none-any.whl (71 kB)
  Downloading spacy_legacy-3.0.12-py2.py3-none-any.whl (29 kB)
  Downloading spacy_loggers-1.0.5-py3-none-any.whl (22 kB)
  Downloading srsly-2.5.2-cp312-cp312-win_amd64.whl (654 kB)
  - 654.8/654.8 kB 785.2 kB/s 0:00:05
  Downloading thinc-8.3.10-cp312-cp312-win_amd64.whl (1.7 MB)
  - 1.7/1.7 MB 2.2 MB/s 0:00:00
  Downloading blis-1.3.3-cp312-cp312-win_amd64.whl (6.2 MB)
  - 6.2/6.2 MB 2.3 MB/s 0:00:02
  Downloading confection-0.1.5-py3-none-any.whl (35 kB)
  Using cached tqdm-4.67.1-py3-none-any.whl (78 kB)
  Downloading typer_slim-0.20.0-py3-none-any.whl (47 kB)
  Using cached urllib3-2.5.0-py3-none-any.whl (129 kB)
  Downloading wasabi-1.1.3-py3-none-any.whl (27 kB)
  Downloading weasel-0.4.3-py3-none-any.whl (50 kB)
  Downloading cloudpathlib-0.23.0-py3-none-any.whl (62 kB)
  Using cached smart_open-7.5.0-py3-none-any.whl (63 kB)
  Using cached gensim-4.4.0-cp312-cp312-win_amd64.whl (24.4 MB)
  Downloading jupyterlab-4.5.0-py3-none-any.whl (12.4 MB)
  - 12.4/12.4 MB 1.4 MB/s 0:00:08
```

```
Windows PowerShell
-----[redacted]-----[redacted]
Downloaded jupyterlab-4.5.0-py3-none-any.whl (12.4 MB)           12.4/12.4 MB 1.4 MB/s 0:00:08
Using cached httpx-0.28.1-py3-none-any.whl (73 kB)
Using cached httpcore-1.0.9-py3-none-any.whl (78 kB)
Downloaded jupyter_server-2.17.0-py3-none-any.whl (388 kB)
Using cached jupyterlab_server-2.28.0-py3-none-any.whl (59 kB)
Using cached annotated_types-0.7.0-py3-none-any.whl (13 kB)
Downloaded anyio-4.12.0-py3-none-any.whl (113 kB)
Downloaded argon2_cffi-25.1.0-py3-none-any.whl (14 kB)
Downloaded asyncio_tru-2.0.0.5-py3-none-any.whl (6.1 kB)
Downloaded babel-2.1.7-py3-none-any.whl (10.2 kB)
----- 10.2/10.2 MB 2.8 MB/s 0:00:03
Using cached certifi-2025.11.12-py3-none-any.whl (159 kB)
Using cached click-8.3.1-py3-none-any.whl (108 kB)
Using cached colorama-0.4.6-py2.py3-none-any.whl (25 kB)
Using cached h11-0.16.0-py3-none-any.whl (37 kB)
Downloaded ipykernel-7.1.0-py3-none-any.whl (117 kB)
Downloaded comm-0.2.3-py3-none-any.whl (7.3 kB)
Downloaded debugpy-1.8.17-cp312-cp312-win_amd64.whl (5.3 MB)
----- 5.3/5.3 MB 2.2 MB/s 0:00:02
Downloaded ipython-9.7.0-py3-none-any.whl (618 kB)
----- 618.9/618.9 kB 3.3 MB/s 0:00:00
Downloaded prompt_toolkit-3.0.52-py3-none-any.whl (391 kB)
Downloaded decorator-5.2.1-py3-none-any.whl (9.2 kB)
Downloaded ipython_pygments_lexers-1.1.1-py3-none-any.whl (8.1 kB)
Downloaded jedi-0.19.2-py2.py3-none-any.whl (1.6 kB)
----- 1.6/1.6 MB 4.0 MB/s 0:00:00
Downloaded parso-0.8.5-py2.py3-none-any.whl (106 kB)
Downloaded jinja2-3.1.0-py3-none-any.whl (134 kB)
Using cached joblib-1.5.2-py3-none-any.whl (308 kB)
Downloaded json5-0.12.1-py3-none-any.whl (36 kB)
Downloaded jsonschema-4.25.1-py3-none-any.whl (90 kB)
Using cached attrs-25.4.0-py3-none-any.whl (67 kB)
Downloaded jsonschema_specifications-2025.9.1-py3-none-any.whl (18 kB)
Downloaded jupyter_client-8.6.3-py3-none-any.whl (106 kB)
Downloaded jupyter_core-5.9.1-py3-none-any.whl (29 kB)
Downloaded jupyter_events-0.12.0-py3-none-any.whl (19 kB)
Using cached jsonpointer-3.0.0-py2.py3-none-any.whl (7.6 kB)
Downloaded jupyter_lsp-2.3.0-py3-none-any.whl (76 kB)
Downloaded jupyter_server_terminals-0.5.3-py3-none-any.whl (13 kB)
Downloaded markupsafe-3.0.3-cp312-cp312-win_amd64.whl (15 kB)
Downloaded matplotlib_inline-0.2.1-py3-none-any.whl (9.5 kB)
```

```
Windows PowerShell
Downloaded psutil-7.1.3-cp37-ab13-win_amd64.whl (247 kB)
Downloaded pygments-2.19.2-py3-none-any.whl (1.2 MB)
    1.2/1.2 MB 4.7 MB/s 0:00:00
Using cached python_dateutil-2.9.0.post0-py2.py3-none-any.whl (229 kB)
Downloaded python_json_logger-4.0.0-py3-none-any.whl (15 kB)
Using cached pytz-2025.2-py2.py3-none-any.whl (509 kB)
Downloaded pywinpty-3.0.2-cp312-cp312-win_amd64.whl (2.1 MB)
    2.1/2.1 MB 3.8 MB/s 0:00:00
Downloaded pyyaml-6.0.3-cp312-cp312-win_amd64.whl (154 kB)
Downloaded pymq-27.1.0-cp312-abi3-win_amd64.whl (619 kB)
    619.5/619.5 kB 2.9 MB/s 0:00:00
Downloading referencing-0.37.0-py3-none-any.whl (26 kB)
Downloaded regex-2025.11.3-cp312-cp312-win_amd64.whl (277 kB)
Downloaded rfc3986_validator-0.1.1-py2.py3-none-any.whl (4.2 kB)
Downloaded rfc3987_syntax-1.1.0-py3-none-any.whl (8.0 kB)
Downloaded lark-1.3.1-py3-none-any.whl (113 kB)
Downloaded rpds_py-0.30.0-cp312-cp312-win_amd64.whl (240 kB)
Using cached scipy-1.16.3-cp312-cp312-win_amd64.whl (38.6 MB)
Downloaded Send2Trash-1.8.3-py3-none-any.whl (18 kB)
Using cached setuptools-80.9.0-py3-none-any.whl (1.2 MB)
Using cached six-1.17.0-py2.py3-none-any.whl (11 kB)
Downloaded stack_data-0.6.3-py3-none-any.whl (24 kB)
Downloaded asttokens-3.0.1-py3-none-any.whl (27 kB)
Downloaded executing-2.2.1-py2.py3-none-any.whl (28 kB)
Downloaded terminado-0.18.1-py3-none-any.whl (14 kB)
Using cached threadpoolctl-3.6.0-py3-none-any.whl (18 kB)
Downloaded tornado-6.5.2-cp39-abi3-win_amd64.whl (445 kB)
Downloaded traitlets-5.14.3-py3-none-any.whl (85 kB)
Using cached typing_extensions-4.15.0-py3-none-any.whl (44 kB)
Using cached typing_inspection-0.4.2-py3-none-any.whl (14 kB)
Using cached tzdata-2025.2-py2.py3-none-any.whl (347 kB)
Downloaded webcolors-25.10.0-py3-none-any.whl (14 kB)
Downloaded webencodings-0.5.1-py2.py3-none-any.whl (11 kB)
Downloaded websocket_client-1.9.0-py3-none-any.whl (82 kB)
Downloaded argon2_cffi_bindings-25.1.0-cp39-abi3-win_amd64.whl (31 kB)
Downloaded cffi-2.0.0-cp312-cp312-win_amd64.whl (183 kB)
Downloaded beautifulsoup4-4.14.3-py3-none-any.whl (107 kB)
Using cached soupsieve-2.8-py3-none-any.whl (36 kB)
Downloaded defusedxml-0.7.1-py2.py3-none-any.whl (25 kB)
Downloaded fqdn-1.5.1-py3-none-any.whl (9.1 kB)
Downloaded isoduration-20.11.0-py3-none-any.whl (11 kB)
Downloaded arrow-1.4.0-py3-none-any.whl (68 kB)
```

```
Windows PowerShell
Downloading rfc339_validator-0.1.4-py2.py3-none-any.whl (3.5 kB)
Downloading uri_template-1.3.0-py3-none-any.whl (11 kB)
Downloading wcwidth-0.2.14-py2.py3-none-any.whl (37 kB)
Using cached wrapt-2.0.1-cp312-cp312-win_amd64.whl (60 kB)
Installing collected packages: webencodings, pytz, pure-eval, fastjsonschema, wrapt, websocket-client, webcolors, wcwidth, urllib3, uritemplate, tzdata, typing-extensions, traitlets, tornado, tinyss2, threadpoolctl, spacy-logger, spacy-legacy, soupsieve, six, setup-tools, send2trash, rpds-py, rfc3986-validator, regex, pymq, pyyaml, pywinpty, python-json-logger, pygments, pyparser, psutil, prometheus-client, platformdirs, parso, pandocfilters, packaging, numpy, nest-asyncio, murmurhash, mistune, MarkupSafe, lark, jupyterlab-pygment, jsonpointer, json5, joblib, idna, hll, fqdn, executing, defusedxml, decorator, debugpy, cymem, comm, colorama, cloudpathlib, chardet, charset-normalizer, certifi, catalogue, bleach, babel, attrs, async-lru, asttokens, annotated-types, wasabi, typing-inspection, tqdm, termcolor, stack_data, srsly, smart-open, scipy, rfc3987-syntax, rfc339-validator, requests, referencing, python-dateutil, pydantic-core, prompt_toolkit, preshed, matplotlib-inline, jupyter-core, jinja2, jedi, ipython-pygments-lexers, httpcore, click, cffi, blis, beautifulsoup4, anyio, typer-slim, scikit-learn, pydantic, pandas, nltk, jupyter-server-terminals, jupyter-client, jsonschema-specification, ipython, httpx, gensim, arrow, argon2-cffi-bindings, jsonschema, isoduration, ipykernel, confection, argon2-cffi, weasel, thinc, nbformat, spacy, nbclient, jupyter-events, nbconvert, jupyter-server, notebook-shim, jupyterlab-server, jupyter-lsp, jupyterlab
 5/124 [websocket-client] WARNING: The script wsdump.exe is installed in 'C:\Users\SYED\ZAIN\AI\AIDEEN\AppData\Local\Packages\PythonSoftwareFoundation.Python.3.12_qbz5n2kfra8p0\LocalCache\local-packages\Python312\Scripts'
which is not on PATH.
  Consider adding this directory to PATH or, if you prefer to suppress this warning, use --no-warn-script-location.
  -----
  18/124 [soupsieve] ERROR: could not install packages due to an OSError: [Errno 2] No such file or directory: 'C:\Users\SYED\ZAIN\AI\AIDEEN\AppData\Local\Packages\PythonSoftwareFoundation.Python.3.12_qbz5n2kfra8p0\LocalCache\local-packages\Python312\site-packages\pkg_resources\tests\data\my-test-package_unpacked-egg\my_test_package-1.0-py3.7.egg\EGG-INFO\dependency_links.txt'
HINT: This error might have occurred since this system does not have Windows Long Path support enabled. You can find information on how to enable this at https://pip.pypa.io/warnings/enable-long-paths'

Defaulting to user installation because normal site-packages is not writeable
Collecting whoosh
  Downloading whoosh-2.7.4-py2.py3-none-any.whl.metadata (3.1 kB)
Collecting annoy
  Downloading annoy-1.17.3.tar.gz (647 kB)
    Installing build dependencies ... done
    Getting requirements to build wheel ... done
    Installing backend dependencies ... done
    Preparing metadata (pyproject.toml) ... done
Collecting hnswlib
  Downloading hnswlib-0.8.0.tar.gz (36 kB)
  Installing build dependencies ... done
  Getting requirements to build wheel ... done
  Preparing metadata (pyproject.toml) ... done
Collecting numpy (from hnswlib)
  Using cached numpy-2.3.5-cp312-cp312-win_amd64.whl.metadata (60 kB)
  Downloading Whoosh-2.7.4-py2.py3-none-any.whl (468 kB)
  Using cached numpy-2.3.5-cp312-cp312-win_amd64.whl (12.8 MB)
Building wheels for collected packages: annoy, hnswlib
  Building wheel for annoy (pyproject.toml) ... error
    error: subprocess-exited-with-error

      x Building wheel for annoy (pyproject.toml) did not run successfully.
      | exit code: 1
      |>> [12 lines of output]
        C:\Users\SYED\ZAIN\AI\AIDEEN\AppData\Local\Temp\pip-build-env-2lb1eltp\overlay\Lib\site-packages\setuptools\_distutils\dist.py:2
89: UserWarning: Unknown distribution option: 'tests_require'
        warnings.warn(msg)
        running bdist_wheel
        running build
        running build_py
        creating build\lib.win-amd64-cpython-312\annoy
        copying annoy\__init__.py -> build\lib.win-amd64-cpython-312\annoy
        copying annoy\__init__.pyi -> build\lib.win-amd64-cpython-312\annoy
        copying annoy\py.typed -> build\lib.win-amd64-cpython-312\annoy
        running build_ext
          building 'annoy.annoylib' extension
            error: Microsoft Visual C++ 14.0 or greater is required. Get it with "Microsoft C++ Build Tools": https://visualstudio.microsoft.com/visual-cpp-build-tools/
        [end of output]

      note: This error originates from a subprocess, and is likely not a problem with pip.
      ERROR: Failed building wheel for annoy
      Building wheel for hnswlib (pyproject.toml) ... error
      error: subprocess-exited-with-error

      x Building wheel for hnswlib (pyproject.toml) did not run successfully.
      | exit code: 1
      |>> [5 lines of output]
        running bdist_wheel
```

```
Windows PowerShell
Preparing metadata (pyproject.toml) ... done
Collecting hnswlib
  Downloading hnswlib-0.8.0.tar.gz (36 kB)
  Installing build dependencies ... done
  Getting requirements to build wheel ... done
  Preparing metadata (pyproject.toml) ... done
Collecting numpy (from hnswlib)
  Using cached numpy-2.3.5-cp312-cp312-win_amd64.whl.metadata (60 kB)
  Downloading Whoosh-2.7.4-py2.py3-none-any.whl (468 kB)
  Using cached numpy-2.3.5-cp312-cp312-win_amd64.whl (12.8 MB)
Building wheels for collected packages: annoy, hnswlib
  Building wheel for annoy (pyproject.toml) ... error
    error: subprocess-exited-with-error

      x Building wheel for annoy (pyproject.toml) did not run successfully.
      | exit code: 1
      |>> [12 lines of output]
        C:\Users\SYED\ZAIN\AI\AIDEEN\AppData\Local\Temp\pip-build-env-2lb1eltp\overlay\Lib\site-packages\setuptools\_distutils\dist.py:2
89: UserWarning: Unknown distribution option: 'tests_require'
        warnings.warn(msg)
        running bdist_wheel
        running build
        running build_py
        creating build\lib.win-amd64-cpython-312\annoy
        copying annoy\__init__.py -> build\lib.win-amd64-cpython-312\annoy
        copying annoy\__init__.pyi -> build\lib.win-amd64-cpython-312\annoy
        copying annoy\py.typed -> build\lib.win-amd64-cpython-312\annoy
        running build_ext
          building 'annoy.annoylib' extension
            error: Microsoft Visual C++ 14.0 or greater is required. Get it with "Microsoft C++ Build Tools": https://visualstudio.microsoft.com/visual-cpp-build-tools/
        [end of output]

      note: This error originates from a subprocess, and is likely not a problem with pip.
      ERROR: Failed building wheel for annoy
      Building wheel for hnswlib (pyproject.toml) ... error
      error: subprocess-exited-with-error

      x Building wheel for hnswlib (pyproject.toml) did not run successfully.
      | exit code: 1
      |>> [5 lines of output]
```

```
Windows PowerShell
Building wheel for hnswlib (pyproject.toml) ... error
error: subprocess-exited-with-error

  × Building wheel for hnswlib (pyproject.toml) did not run successfully.
  | exit code: 1
  +--> [5 lines of output]
      running bdist_wheel
      running build
      running build_ext
      building 'hnswlib' extension
      error: Microsoft Visual C++ 14.0 or greater is required. Get it with "Microsoft C++ Build Tools": https://visualstudio.microsoft.com/visual-cpp-build-tools/
  [end of output]

  note: This error originates from a subprocess, and is likely not a problem with pip.
ERROR: Failed building wheel for hnswlib
Failed to build annoy hnswlib
error: failed-wheel-build-for-install

× Failed to build installable wheels for some pyproject.toml based projects
[→] annoy, hnswlib
Defaulting to user installation because normal site-packages is not writeable
Collecting flask
  Downloading flask-3.1.2-py3-none-any.whl.metadata (3.2 kB)
Collecting fastapi
  Downloading fastapi-0.123.0-py3-none-any.whl.metadata (30 kB)
Collecting uvicorn
  Using cached uvicorn-0.38.0-py3-none-any.whl.metadata (6.8 kB)
Collecting blinker>=1.9.0 (from flask)
  Downloading blinker-1.9.0-py3-none-any.whl.metadata (1.6 kB)
Collecting click>=8.1.3 (from flask)
  Using cached click-8.3.1-py3-none-any.whl.metadata (2.6 kB)
Collecting itsdangerous>=2.2.0 (from flask)
  Downloading itsdangerous-2.2.0-py3-none-any.whl.metadata (1.9 kB)
Collecting jinja2>=3.1.2 (from flask)
  Using cached jinja2-3.1.6-py3-none-any.whl.metadata (2.9 kB)
Collecting markupsafe>=2.1.1 (from flask)
  Using cached markupsafe-3.0.3-cp312-cp312-win_amd64.whl.metadata (2.8 kB)
Collecting werkzeug>=3.1.0 (from flask)
  Downloading werkzeug-3.1.4-py3-none-any.whl.metadata (4.0 kB)
Collecting starlette<0.51.0,>=0.40.0 (from fastapi)
  Using cached starlette-0.50.0-py3-none-any.whl.metadata (6.3 kB)
```

```
Windows PowerShell
n.python.3.12_qbz5n2kfra8p0\localcache\local-packages\python312\site-packages (from fastapi) (4.15.0)
Collecting annotated-doc>=0.0.2 (from fastapi)
  Using cached annotated_doc-0.0.4-py3-none-any.whl.metadata (6.6 kB)
Collecting annotated-types>=0.6.0 (from pydantic!=1.8,!>1.8.1,!>2.0.0,!>2.0.1,!>2.1.0,<3.0.0,>=1.7.4->fastapi)
  Using cached annotated_types-0.7.0-py3-none-any.whl.metadata (15 kB)
Collecting pydantic-core==2.41.5 (from pydantic!=1.8,!>1.8.1,!>2.0.0,!>2.0.1,!>2.1.0,<3.0.0,>=1.7.4->fastapi)
  Using cached pydantic_core-2.41.5-cp312-cp312-win_amd64.whl.metadata (7.4 kB)
Collecting typing-inspection>=0.4.2 (from pydantic!=1.8,!>1.8.1,!>2.0.0,!>2.0.1,!>2.1.0,<3.0.0,>=1.7.4->fastapi)
  Using cached typing_inspection-0.4.2-py3-none-any.whl.metadata (2.6 kB)
Collecting anyio<,>=3.6.2 (from starlette<0.51.0,>=0.40.0->fastapi)
  Using cached anyio-4.12.0-py3-none-any.whl.metadata (4.3 kB)
Collecting idna>=2.8 (from anyio<,>=3.6.2->starlette<0.51.0,>=0.40.0->fastapi)
  Using cached idna-3.11-py3-none-any.whl.metadata (8.4 kB)
Collecting h11>=0.8 (from uvicorn)
  Using cached h11-0.16.0-py3-none-any.whl.metadata (8.3 kB)
Collecting colorama (from click=>8.1.3->flask)
  Using cached colorama-0.4.6-py2.py3-none-any.whl.metadata (17 kB)
Downloading flask-3.1.2-py3-none-any.whl (103 kB)
Using cached fastapi-0.123.0-py3-none-any.whl (463 kB)
Using cached pydantic-core-2.41.5-cp312-cp312-win_amd64.whl (2.0 MB)
Using cached starlette-0.50.0-py3-none-any.whl (74 kB)
Using cached anyio-4.12.0-py3-none-any.whl (113 kB)
Using cached uvicorn-0.38.0-py3-none-any.whl (68 kB)
Using cached annotated_doc-0.0.4-py3-none-any.whl (5.3 kB)
Using cached annotated_types-0.7.0-py3-none-any.whl (13 kB)
Downloading blinker-1.9.0-py3-none-any.whl (8.5 kB)
Using cached click-8.3.1-py3-none-any.whl (108 kB)
Using cached h11-0.16.0-py3-none-any.whl (37 kB)
Using cached idna-3.11-py3-none-any.whl (71 kB)
Downloading itsdangerous-2.2.0-py3-none-any.whl (16 kB)
Using cached jinja2-3.1.6-py3-none-any.whl (134 kB)
Using cached markupsafe-3.0.3-cp312-cp312-win_amd64.whl (15 kB)
Using cached typing_inspection-0.4.2-py3-none-any.whl (14 kB)
Downloading werkzeug-3.1.4-py3-none-any.whl (224 kB)
Using cached colorama-0.4.6-py2.py3-none-any.whl (25 kB)
Installing collected packages: typing-inspection, pydantic-core, markupsafe, itsdangerous, idna, h11, colorama, blinker, annotated-types, annotated-doc, werkzeug, pydantic, jinja2, click, anyio, uvicorn, starlette, flask, fastapi
  15/19 [uvicorn] WARNING: The script uvicorn.exe is installed in 'C:\Users\SYED ZAIN UL AI DEEN\AppData\Local\Packages\PythonSoftwareFoundation.Python.3.12_qbz5n2kfra8p0\LocalCache\local-packages\Python312\Scripts' which is not on PATH.
  Consider adding this directory to PATH or, if you prefer to suppress this warning, use --no-warn-script-location.
```

```
Windows PowerShell
Using cached aiohttp-4.12.0-py3-none-any.whl (113 kB)
Using cached uvicorn-0.38.0-py3-none-any.whl (68 kB)
Using cached annotated_doc-0.0.4-py3-none-any.whl (5.3 kB)
Using cached annotated_types-0.7.0-py3-none-any.whl (13 kB)
Downloading blinker-1.9.0-py3-none-any.whl (8.5 kB)
Using cached click-8.3.1-py3-none-any.whl (108 kB)
Using cached idna-3.11-py3-none-any.whl (37 kB)
Using cached itsdangerous-2.2.0-py3-none-any.whl (16 kB)
Using cached jinja2-3.1.6-py3-none-any.whl (134 kB)
Using cached markupsafe-3.0.3-cp312-cp312-win_amd64.whl (15 kB)
Using cached typing_inspection-0.4.2-py3-none-any.whl (14 kB)
Downloading werkzeug-3.1.4-py3-none-any.whl (224 kB)
Using cached colorama-0.4.6-py2.py3-none-any.whl (25 kB)
Installing collected packages: typing-inspection, pydantic-core, markupsafe, itsdangerous, idna, click, anyio, uvicorn, starlette, flask, fastapi
 15/19 [uvicorn] WARNING: The script uvicorn.exe is installed in 'C:\Users\SYED ZAIN UL AIDEN\AppData\Local\ Packages\PythonSoftwareFoundation.Python.3.12_qbz5n2kfra8p0\LocalCache\local-packages\Python312\Scripts' which is not on PATH.
Consider adding this directory to PATH or, if you prefer to suppress this warning, use --no-warn-script-location.
 17/19 [flask] WARNING: The script flask.exe is installed in 'C:\Users\SYED ZAIN UL AIDEN\AppData\Local\ Packages\PythonSoftwareFoundation.Python.3.12_qbz5n2kfra8p0\LocalCache\local-packages\Python312\Scripts' which is not on PATH.
Consider adding this directory to PATH or, if you prefer to suppress this warning, use --no-warn-script-location.
 18/19 [fastapi] WARNING: The script fastapi.exe is installed in 'C:\Users\SYED ZAIN UL AIDEN\DEEN\AppData\Local\ Packages\PythonSoftwareFoundation.Python.3.12_qbz5n2kfra8p0\LocalCache\local-packages\Python312\Scripts' which is not on PATH.
Consider adding this directory to PATH or, if you prefer to suppress this warning, use --no-warn-script-location.
Successfully installed annotated-doc-0.0.4 annotated-types-0.7.0 aiohttp-4.12.0 blinker-1.9.0 click-8.3.1 colorama-0.4.6 fastapi-0.123.0 flask-3.1.2 h11-0.16.0 idna-3.11 itsdangerous-2.2.0 jinja2-3.1.6 markupsafe-3.0.3 pydantic-2.12.5 pydantic-core-2.41.5 starlette-0.50.0 typing-inspection-0.4.2 uvicorn-0.38.0 werkzeug-3.1.4

PS C:\projects\ir-system> mkdir C:\projects\ir-system
>> cd C:\projects\ir-system
>> python -m venv .venv
>> .\venv\Scripts\Activate.ps1      # PowerShell
>> python -m pip install --upgrade pip
>> pip install numpy pandas scikit-learn nltk spacy gensim jupyterlab
>> pip install whoosh annoy hnslib      # optional vector / Classical
>> pip install flask fastapi uvicorn    # optional web UI / API
>> pip install faiss-cpu
```

```
Windows PowerShell
Defaulting to user installation because normal site-packages is not writeable
Requirement already satisfied: pip in c:\users\syed zain ul aideen\appdata\local\packages\pythonsoftwarefoundation.python.3.12_qbz5n2kfra8p0\localcache\local-packages\python312\site-packages (25.3)
Defaulting to user installation because normal site-packages is not writeable
Collecting numpy
  Using cached numpy-2.3.5-cp312-cp312-win_amd64.whl.metadata (60 kB)
Collecting pandas
  Using cached pandas-2.3.3-cp312-cp312-win_amd64.whl.metadata (19 kB)
Collecting scikit-learn
  Using cached scikit_learn-1.7.2-cp312-cp312-win_amd64.whl.metadata (11 kB)
Collecting nltk
  Using cached nltk-3.9.2-py3-none-any.whl.metadata (3.2 kB)
Collecting spacy
  Using cached spacy-3.8.11-cp312-cp312-win_amd64.whl.metadata (28 kB)
Collecting gensim
  Using cached gensim-4.4.0-cp312-cp312-win_amd64.whl.metadata (8.6 kB)
Collecting jupyterlab
  Using cached jupyterlab-4.5.0-py3-none-any.whl.metadata (16 kB)
Collecting python-dateutil==2.8.2 (from pandas)
  Using cached python_dateutil-2.9.0.post0-py2.py3-none-any.whl.metadata (8.4 kB)
Requirement already satisfied: pytz>=2020.1 in c:\users\syed zain ul aideen\appdata\local\packages\pythonsoftwarefoundation.python.3.12_qbz5n2kfra8p0\localcache\local-packages\python312\site-packages (from pandas) (2025.2)
Requirement already satisfied: tzdata>=2022.7 in c:\users\syed zain ul aideen\appdata\local\packages\pythonsoftwarefoundation.python.3.12_qbz5n2kfra8p0\localcache\local-packages\python312\site-packages (from pandas) (2025.2)
Collecting scipy>=1.8.0 (from scikit-learn)
  Using cached scipy-1.16.3-cp312-cp312-win_amd64.whl.metadata (60 kB)
Collecting joblib=1.2.0 (from scikit-learn)
  Using cached joblib-1.5.2-py3-none-any.whl.metadata (5.6 kB)
Requirement already satisfied: threadpoolctl>=3.1.0 in c:\users\syed zain ul aideen\appdata\local\packages\pythonsoftwarefoundation.python.3.12_qbz5n2kfra8p0\localcache\local-packages\python312\site-packages (from scikit-learn) (3.6.0)
Requirement already satisfied: click in c:\users\syed zain ul aideen\appdata\local\packages\pythonsoftwarefoundation.python.3.12_qbz5n2kfra8p0\localcache\local-packages\python312\site-packages (from nltk) (8.3.1)
Collecting regex>=2021.8.3 (from nltk)
  Using cached regex-2025.11.3-cp312-cp312-win_amd64.whl.metadata (41 kB)
Collecting tqdm (from nltk)
  Using cached tqdm-4.67.1-py3-none-any.whl.metadata (57 kB)
Requirement already satisfied: spacy-legacy<3.1.0,>=3.0.11 in c:\users\syed zain ul aideen\appdata\local\packages\pythonsoftwarefoundation.python.3.12_qbz5n2kfra8p0\localcache\local-packages\python312\site-packages (from spacy) (3.0.12)
Requirement already satisfied: spacy-loggers<2.0.0,>=1.0.0 in c:\users\syed zain ul aideen\appdata\local\packages\pythonsoftwarefoundation.python.3.12_qbz5n2kfra8p0\localcache\local-packages\python312\site-packages (from spacy) (1.0.5)
Collecting murmurhash<1.1.0,>=0.28.0 (from spacy)
  Using cached murmurhash-1.0.15-cp312-cp312-win_amd64.whl.metadata (2.3 kB)
```

```
Windows PowerShell
Using cached preshed-3.0.12-cp312-cp312-win_amd64.whl.metadata (2.6 kB)
Collecting thinc<8.4.0,>=8.3.4 (from spacy)
  Using cached thinc-8.3.10-cp312-cp312-win_amd64.whl.metadata (15 kB)
Collecting wasabi<1.2.0,>=0.9.1 (from spacy)
  Using cached wasabi-1.1.3-py3-none-any.whl.metadata (28 kB)
Collecting srsly<3.0.0,>=2.4.3 (from spacy)
  Using cached srsly-2.5.2-cp312-cp312-win_amd64.whl.metadata (20 kB)
Collecting catalogue<2.1.0,>=2.0.6 (from spacy)
  Using cached catalogue-2.0.10-py3-none-any.whl.metadata (14 kB)
Collecting weasel<0.5.0,>=0.4.2 (from spacy)
  Using cached weasel-0.4.3-py3-none-any.whl.metadata (4.6 kB)
Collecting typer<slim<1.0.0,>=0.3.0 (from spacy)
  Using cached typer_slim-0.20.0-py3-none-any.whl.metadata (16 kB)
Collecting requests<3.0.0,>=2.13.0 (from spacy)
  Using cached requests-2.32.5-py3-none-any.whl.metadata (4.9 kB)
Requirement already satisfied: pydantic!=1.8,!1=1.8.1,<3.0.0,>=1.7.4 in c:\users\syed zain ul aideen\appdata\local\packages\pythonsoftwarefoundation.python.3.12_qbz5n2kfra8p0\localcache\local-packages\python312\site-packages (from spacy) (2.12.5)
Requirement already satisfied: jinja2 in c:\users\syed zain ul aideen\appdata\local\packages\pythonsoftwarefoundation.python.3.12_qbz5n2kfra8p0\localcache\local-packages\python312\site-packages (from spacy) (3.1.6)
Collecting setuptools (from spacy)
  Using cached setuptools-80.9.0-py3-none-any.whl.metadata (6.6 kB)
Collecting packaging==20.0 (from spacy)
  Using cached packaging-25.0-py3-none-any.whl.metadata (3.3 kB)
Requirement already satisfied: annotated-types<=0.6.0 in c:\users\syed zain ul aideen\appdata\local\packages\pythonsoftwarefoundation.python.3.12_qbz5n2kfra8p0\localcache\local-packages\python312\site-packages (from pydantic!=1.8,!1=1.8.1,<3.0.0,>=1.7.4->spacy) (0.7.0)

Requirement already satisfied: pydantic-core==2.41.5 in c:\users\syed zain ul aideen\appdata\local\packages\pythonsoftwarefoundation.python.3.12_qbz5n2kfra8p0\localcache\local-packages\python312\site-packages (from pydantic!=1.8,!1=1.8.1,<3.0.0,>=1.7.4->spacy) (2.41.5)

Requirement already satisfied: typing-extensions>=4.14.1 in c:\users\syed zain ul aideen\appdata\local\packages\pythonsoftwarefoundation.python.3.12_qbz5n2kfra8p0\localcache\local-packages\python312\site-packages (from pydantic!=1.8,!1=1.8.1,<3.0.0,>=1.7.4->spacy) (4.15.0)
Requirement already satisfied: typing-inspection>=0.4.2 in c:\users\syed zain ul aideen\appdata\local\packages\pythonsoftwarefoundation.python.3.12_qbz5n2kfra8p0\localcache\local-packages\python312\site-packages (from pydantic!=1.8,!1=1.8.1,<3.0.0,>=1.7.4->spacy) (0.4.2)
Collecting charset_normalizer<4,>=2 (from requests<3.0.0,>=2.13.0->spacy)
  Using cached charset_normalizer-3.4.4-cp312-cp312-win_amd64.whl.metadata (38 kB)
Requirement already satisfied: idna<4,>=2.5 in c:\users\syed zain ul aideen\appdata\local\packages\pythonsoftwarefoundation.python.3.12_qbz5n2kfra8p0\localcache\local-packages\python312\site-packages (from requests<3.0.0,>=2.13.0->spacy) (3.11)
Requirement already satisfied: urllib3<3,>=1.21.1 in c:\users\syed zain ul aideen\appdata\local\packages\pythonsoftwarefoundation.python.3.12_qbz5n2kfra8p0\localcache\local-packages\python312\site-packages (from requests<3.0.0,>=2.13.0->spacy) (2.5.0)
Collecting certifi==2017.4.17 (from requests<3.0.0,>=2.13.0->spacy)
```

```
Windows PowerShell
Collecting charset_normalizer<4,>=2 (from requests<3.0.0,>=2.13.0->spacy)
  Using cached charset_normalizer-3.4.4-cp312-cp312-win_amd64.whl.metadata (38 kB)
Requirement already satisfied: idna<4,>=2.5 in c:\users\syed zain ul aideen\appdata\local\packages\pythonsoftwarefoundation.python.3.12_qbz5n2kfra8p0\localcache\local-packages\python312\site-packages (from requests<3.0.0,>=2.13.0->spacy) (3.11)
Requirement already satisfied: urllib3<3,>=1.21.1 in c:\users\syed zain ul aideen\appdata\local\packages\pythonsoftwarefoundation.python.3.12_qbz5n2kfra8p0\localcache\local-packages\python312\site-packages (from requests<3.0.0,>=2.13.0->spacy) (2.5.0)
Collecting certifi==2017.4.17 (from requests<3.0.0,>=2.13.0->spacy)
  Using cached certifi-2025.11.12-py3-none-any.whl.metadata (2.5 kB)
Collecting blis<1.4.0,>=1.3.0 (from thinc<8.4.0,>=8.3.4->spacy)
  Using cached blis-1.3.3-cp312-cp312-win_amd64.whl.metadata (7.7 kB)
Collecting confection<1.0.0,>=0.0.1 (from thinc<8.4.0,>=8.3.4->spacy)
  Using cached confection-0.1.5-py3-none-any.whl.metadata (19 kB)
Requirement already satisfied: colorama in c:\users\syed zain ul aideen\appdata\local\packages\pythonsoftwarefoundation.python.3.12_qbz5n2kfra8p0\localcache\local-packages\python312\site-packages (from smart-open<8.0.0,>=5.2.1->weasel<0.5.0,>=0.4.2->spacy) (0.4.6)
Collecting cloudpathlib<1.0.0,>=0.7.0 (from weasel<0.5.0,>=0.4.2->spacy)
  Using cached cloudpathlib-0.23.0-py3-none-any.whl.metadata (16 kB)
Collecting smart-open<8.0.0,>=5.2.1 (from weasel<0.5.0,>=0.4.2->spacy)
  Using cached smart_open-7.5.0-py3-none-any.whl.metadata (24 kB)
Requirement already satisfied: wrapt in c:\users\syed zain ul aideen\appdata\local\packages\pythonsoftwarefoundation.python.3.12_qbz5n2kfra8p0\localcache\local-packages\python312\site-packages (from smart-open<8.0.0,>=5.2.1->weasel<0.5.0,>=0.4.2->spacy) (2.0.1)
Collecting async-triu<1.0.0 (from jupyterlab)
  Using cached async_triu-2.0.5-py3-none-any.whl.metadata (4.5 kB)
Collecting httpxx<1.0.0,>=0.25.0 (from jupyterlab)
  Using cached httpxx-0.28.1-py3-none-any.whl.metadata (7.1 kB)
Collecting ipykernel<6.30.0,>=6.5.0 (from jupyterlab)
  Using cached ipykernel-7.1.0-py3-none-any.whl.metadata (4.5 kB)
Collecting jupyter-core (from jupyterlab)
  Using cached jupyter_core-5.9.1-py3-none-any.whl.metadata (1.5 kB)
Collecting jupyter_lsp<2.0.0 (from jupyterlab)
  Using cached jupyter_lsp-2.3.0-py3-none-any.whl.metadata (1.8 kB)
Collecting jupyter-server<3,>=2.4.0 (from jupyterlab)
  Using cached jupyter_server-2.17.0-py3-none-any.whl.metadata (8.5 kB)
Collecting jupyterlab-server<3,>=2.28.0 (from jupyterlab)
  Using cached jupyterlab_server-2.28.0-py3-none-any.whl.metadata (5.9 kB)
Collecting notebook_shim<0.2 (from jupyterlab)
  Using cached notebook_shim-0.2.4-py3-none-any.whl.metadata (4.0 kB)
Requirement already satisfied: tornado<=6.2.0 in c:\users\syed zain ul aideen\appdata\local\packages\pythonsoftwarefoundation.python.3.12_qbz5n2kfra8p0\localcache\local-packages\python312\site-packages (from jupyterlab) (6.5.2)
Requirement already satisfied: traitlets in c:\users\syed zain ul aideen\appdata\local\packages\pythonsoftwarefoundation.python.3.12_qbz5n2kfra8p0\localcache\local-packages\python312\site-packages (from jupyterlab) (5.14.3)
Requirement already satisfied: anyio in c:\users\syed zain ul aideen\appdata\local\packages\pythonsoftwarefoundation.python.3.12_qbz5n2kfra8p0\localcache\local-packages\python312\site-packages (from httpxx<1,>=0.25.0->jupyterlab) (4.12.0)
```

```
Windows PowerShell
pz5n2kfra8p0\localcache\local-packages\python312\site-packages (from httpcore==1.*->httpx<1,>=0.25.0->jupyterlab) (0.16.0)
Collecting argon2-cffi>=21.1 (from jupyter-server<3,>=2.4.0->jupyterlab)
  Using cached argon2_cffi-25.1.0-py3-none-any.whl.metadata (4.1 kB)
Collecting jupyter-client>=7.4.4 (from jupyter-server<3,>=2.4.0->jupyterlab)
  Using cached jupyter_client-8.6.3-py3-none-any.whl.metadata (8.3 kB)
Collecting jupyter-events>=0.11.0 (from jupyter-server<3,>=2.4.0->jupyterlab)
  Using cached jupyter_events-0.12.0-py3-none-any.whl.metadata (5.8 kB)
Collecting jupyter-server-terminals>=0.4.4 (from jupyter-server<3,>=2.4.0->jupyterlab)
  Using cached jupyter_server_terminals-0.5.3-py3-none-any.whl.metadata (5.6 kB)
Collecting nbconvert>=6.4.4 (from jupyter-server<3,>=2.4.0->jupyterlab)
  Using cached nbconvert-7.16.6-py3-none-any.whl.metadata (8.5 kB)
Collecting nbformat>=5.3.0 (from jupyter-server<3,>=2.4.0->jupyterlab)
  Using cached nbformat-5.10.4-py3-none-any.whl.metadata (3.6 kB)
Collecting prometheus-client>=0.9 (from jupyter-server<3,>=2.4.0->jupyterlab)
  Using cached prometheus_client-0.23.1-py3-none-any.whl.metadata (1.9 kB)
Collecting pywinpty>=2.0.1 (from jupyter-server<3,>=2.4.0->jupyterlab)
  Using cached pywinpty-3.0.2-cp312-cp312-win_amd64.whl.metadata (5.7 kB)
Collecting pyzmq>=24 (from jupyter-server<3,>=2.4.0->jupyterlab)
  Using cached pyzmq-27.1.0-cp312-abif3-win_amd64.whl.metadata (6.0 kB)
Collecting send2trash>=1.8.2 (from jupyter-server<3,>=2.4.0->jupyterlab)
  Using cached send2trash-1.8.3-py3-none-any.whl.metadata (4.0 kB)
Collecting terminado>=0.8.3 (from jupyter-server<3,>=2.4.0->jupyterlab)
  Using cached terminado-0.18.1-py3-none-any.whl.metadata (5.8 kB)
Requirement already satisfied: websocket-client>=1.7 in c:\users\syed zain ul aideen\appdata\local\packages\pythonsoftwarefoundation.python.3.12_qbz5n2kfra8p0\localcache\local-packages\python312\site-packages (from jupyter-server<3,>=2.4.0->jupyterlab) (1.9.0)
Collecting babel>=2.10 (from jupyterlab-server<3,>=2.8.0->jupyterlab)
  Using cached babel-2.17.0-py3-none-any.whl.metadata (2.0 kB)
Collecting json5>=0.9.0 (from jupyterlab-server<3,>=2.8.0->jupyterlab)
  Using cached json5-0.12.1-py3-none-any.whl.metadata (36 kB)
Collecting jsonschema>=4.18.0 (from jupyterlab-server<3,>=2.8.0->jupyterlab)
  Using cached jsonschema-4.25.1-py3-none-any.whl.metadata (7.6 kB)
Collecting argon2-cffi-bindings (from argon2-cffi>=21.1->jupyter-server<3,>=2.4.0->jupyterlab)
  Using cached argon2_cffi_bindings-25.1.0-cp39-abif3-win_amd64.whl.metadata (7.5 kB)
Collecting comm>=0.1.1 (from ipykernel!=6.30.0,>=6.5.0->jupyterlab)
  Using cached comm-0.2.3-py3-none-any.whl.metadata (3.7 kB)
Collecting debugpy>=1.6.5 (from ipykernel!=6.30.0,>=6.5.0->jupyterlab)
  Using cached debugpy-1.8.17-cp312-cp312-win_amd64.whl.metadata (1.5 kB)
Collecting ipython>=7.23.1 (from ipykernel!=6.30.0,>=6.5.0->jupyterlab)
  Using cached ipython-9.7.0-py3-none-any.whl.metadata (4.5 kB)
Collecting matplotlib-inline>=0.1 (from ipykernel!=6.30.0,>=6.5.0->jupyterlab)
  Using cached matplotlib_inline-0.2.1-py3-none-any.whl.metadata (2.3 kB)
Collecting nest-asyncio>=1.4 (from ipykernel!=6.30.0,>=6.5.0->jupyterlab)
  Using cached nest_asyncio-1.6.0-py3-none-any.whl.metadata (2.8 kB)
Collecting psutil>=5.7 (from ipykernel!=6.30.0,>=6.5.0->jupyterlab)
  Using cached psutil-7.1.3-cp37-abif3-win_amd64.whl.metadata (23 kB)
Collecting decorator>=4.3.2 (from ipython>=7.23.1->ipykernel!=6.30.0,>=6.5.0->jupyterlab)
  Using cached decorator-5.2.1-py3-none-any.whl.metadata (3.9 kB)
Collecting ipython_pygments.lexers>=1.0.0 (from ipython>=7.23.1->ipykernel!=6.30.0,>=6.5.0->jupyterlab)
  Using cached ipython_pygments_lexers-1.1.1-py3-none-any.whl.metadata (1.1 kB)
Collecting jedi>=0.18.1 (from ipython>=7.23.1->ipykernel!=6.30.0,>=6.5.0->jupyterlab)
  Using cached jedi-0.19.2-py3-none-any.whl.metadata (22 kB)
Collecting prompt_toolkit>=3.1.0,>=3.0.41 (from ipython>=7.23.1->ipykernel!=6.30.0,>=6.5.0->jupyterlab)
  Using cached prompt_toolkit-3.0.52-py3-none-any.whl.metadata (6.4 kB)
Collecting pygments>=2.11.0 (from ipython>=7.23.1->ipykernel!=6.30.0,>=6.5.0->jupyterlab)
  Using cached pygments-2.19.2-py3-none-any.whl.metadata (2.5 kB)
Collecting stack_data>=0.6.0 (from ipython>=7.23.1->ipykernel!=6.30.0,>=6.5.0->jupyterlab)
  Using cached stack_data-0.6.3-py3-none-any.whl.metadata (18 kB)
Requirement already satisfied: wcdwidth in c:\users\syed zain ul aideen\appdata\local\packages\pythonsoftwarefoundation.python.3.12_qbz5n2kfra8p0\localcache\local-packages\python312\site-packages (from prompt_toolkit<3.1.0,>=3.0.41->ipython>=7.23.1->ipykernel!=6.30.0,>=6.5.0->jupyterlab) (0.2.14)
Collecting parsos>0.9.0,>=0.8.4 (from jedi>=0.18.1->ipython>=7.23.1->ipykernel!=6.30.0,>=6.5.0->jupyterlab)
  Using cached parsos-0.8.5-py2.py3-none-any.whl.metadata (8.3 kB)
Requirement already satisfied: MarkupSafe>=2.0 in c:\users\syed zain ul aideen\appdata\local\packages\pythonsoftwarefoundation.python.3.12_qbz5n2kfra8p0\localcache\local-packages\python312\site-packages (from jinja2->spacy) (3.0.3)
Collecting attrs>=22.2.0 (from jsonschema>=4.18.0->jupyterlab-server<3,>=2.28.0->jupyterlab)
  Using cached attrs-25.4.0-py3-none-any.whl.metadata (10 kB)
Collecting jsonschema-specifications>=2023.03.6 (from jsonschema>=4.18.0->jupyterlab-server<3,>=2.28.0->jupyterlab)
  Using cached jsonschema_specifications-2025.9.1-py3-none-any.whl.metadata (2.9 kB)
Collecting referencing>=0.28.4 (from jsonschema>=4.18.0->jupyterlab-server<3,>=2.28.0->jupyterlab)
  Using cached referencing-0.37.0-py3-none-any.whl.metadata (2.8 kB)
Collecting rpds-py>=0.7.1 (from jsonschema>=4.18.0->jupyterlab-server<3,>=2.28.0->jupyterlab)
  Using cached rpds_py-0.30.0-cp312-cp312-win_amd64.whl.metadata (4.2 kB)
Collecting platformdirs>=2.5 (from jupyter-core->jupyterlab)
  Using cached platformdirs-4.5.0-py3-none-any.whl.metadata (12 kB)
Collecting python_json_logger>=2.0.4 (from jupyter-events>=0.11.0->jupyter-server<3,>=2.4.0->jupyterlab)
  Using cached python_json_logger-4.0.0-py3-none-any.whl.metadata (4.0 kB)
Collecting pyyaml>=5.3 (from jupyter-events>=0.11.0->jupyter-server<3,>=2.4.0->jupyterlab)
```

```
Windows PowerShell
Collecting debugpy>=1.6.5 (from ipykernel!=6.30.0,>=6.5.0->jupyterlab)
  Using cached debugpy-1.8.17-cp312-cp312-win_amd64.whl.metadata (1.5 kB)
Collecting ipython>=7.23.1 (from ipykernel!=6.30.0,>=6.5.0->jupyterlab)
  Using cached ipython-9.7.0-py3-none-any.whl.metadata (4.5 kB)
Collecting matplotlib-inline>=0.1 (from ipykernel!=6.30.0,>=6.5.0->jupyterlab)
  Using cached matplotlib_inline-0.2.1-py3-none-any.whl.metadata (2.3 kB)
Collecting nest-asyncio>=1.4 (from ipykernel!=6.30.0,>=6.5.0->jupyterlab)
  Using cached nest_asyncio-1.6.0-py3-none-any.whl.metadata (2.8 kB)
Collecting psutil>=5.7 (from ipykernel!=6.30.0,>=6.5.0->jupyterlab)
  Using cached psutil-7.1.3-cp37-abif3-win_amd64.whl.metadata (23 kB)
Collecting decorator>=4.3.2 (from ipython>=7.23.1->ipykernel!=6.30.0,>=6.5.0->jupyterlab)
  Using cached decorator-5.2.1-py3-none-any.whl.metadata (3.9 kB)
Collecting ipython_pygments.lexers>=1.0.0 (from ipython>=7.23.1->ipykernel!=6.30.0,>=6.5.0->jupyterlab)
  Using cached ipython_pygments_lexers-1.1.1-py3-none-any.whl.metadata (1.1 kB)
Collecting jedi>=0.18.1 (from ipython>=7.23.1->ipykernel!=6.30.0,>=6.5.0->jupyterlab)
  Using cached jedi-0.19.2-py3-none-any.whl.metadata (22 kB)
Collecting prompt_toolkit>=3.1.0,>=3.0.41 (from ipython>=7.23.1->ipykernel!=6.30.0,>=6.5.0->jupyterlab)
  Using cached prompt_toolkit-3.0.52-py3-none-any.whl.metadata (6.4 kB)
Collecting pygments>=2.11.0 (from ipython>=7.23.1->ipykernel!=6.30.0,>=6.5.0->jupyterlab)
  Using cached pygments-2.19.2-py3-none-any.whl.metadata (2.5 kB)
Collecting stack_data>=0.6.0 (from ipython>=7.23.1->ipykernel!=6.30.0,>=6.5.0->jupyterlab)
  Using cached stack_data-0.6.3-py3-none-any.whl.metadata (18 kB)
Requirement already satisfied: wcdwidth in c:\users\syed zain ul aideen\appdata\local\packages\pythonsoftwarefoundation.python.3.12_qbz5n2kfra8p0\localcache\local-packages\python312\site-packages (from prompt_toolkit<3.1.0,>=3.0.41->ipython>=7.23.1->ipykernel!=6.30.0,>=6.5.0->jupyterlab) (0.2.14)
Collecting parsos>0.9.0,>=0.8.4 (from jedi>=0.18.1->ipython>=7.23.1->ipykernel!=6.30.0,>=6.5.0->jupyterlab)
  Using cached parsos-0.8.5-py2.py3-none-any.whl.metadata (8.3 kB)
Requirement already satisfied: MarkupSafe>=2.0 in c:\users\syed zain ul aideen\appdata\local\packages\pythonsoftwarefoundation.python.3.12_qbz5n2kfra8p0\localcache\local-packages\python312\site-packages (from jinja2->spacy) (3.0.3)
Collecting attrs>=22.2.0 (from jsonschema>=4.18.0->jupyterlab-server<3,>=2.28.0->jupyterlab)
  Using cached attrs-25.4.0-py3-none-any.whl.metadata (10 kB)
Collecting jsonschema-specifications>=2023.03.6 (from jsonschema>=4.18.0->jupyterlab-server<3,>=2.28.0->jupyterlab)
  Using cached jsonschema_specifications-2025.9.1-py3-none-any.whl.metadata (2.9 kB)
Collecting referencing>=0.28.4 (from jsonschema>=4.18.0->jupyterlab-server<3,>=2.28.0->jupyterlab)
  Using cached referencing-0.37.0-py3-none-any.whl.metadata (2.8 kB)
Collecting rpds-py>=0.7.1 (from jsonschema>=4.18.0->jupyterlab-server<3,>=2.28.0->jupyterlab)
  Using cached rpds_py-0.30.0-cp312-cp312-win_amd64.whl.metadata (4.2 kB)
Collecting platformdirs>=2.5 (from jupyter-core->jupyterlab)
  Using cached platformdirs-4.5.0-py3-none-any.whl.metadata (12 kB)
Collecting python_json_logger>=2.0.4 (from jupyter-events>=0.11.0->jupyter-server<3,>=2.4.0->jupyterlab)
  Using cached python_json_logger-4.0.0-py3-none-any.whl.metadata (4.0 kB)
Collecting pyyaml>=5.3 (from jupyter-events>=0.11.0->jupyter-server<3,>=2.4.0->jupyterlab)
```

```
Windows PowerShell
Using cached rfc3987_syntax-1.1.0-py3-none-any.whl.metadata (7.7 kB)
Requirement already satisfied: uri-template in c:\users\syed zain ul aideen\appdata\local\packages\pythonsoftwarefoundation.python.3.12_qbz5n2kfra8p0\localcache\local-packages\python312\site-packages (from jsonschema[format-nongpl]>=4.18.0->jupyter-events>=0.11.0->jupyter-server<3,>=2.4.0->jupyterlab) (1.3.0)
Requirement already satisfied: webcolors>=24.6.0 in c:\users\syed zain ul aideen\appdata\local\packages\pythonsoftwarefoundation.python.3.12_qbz5n2kfra8p0\localcache\local-packages\python312\site-packages (from jsonschema[format-nongpl]>=4.18.0->jupyter-events>=0.11.0->jupyter-server<3,>=2.4.0->jupyterlab) (25.10.0)
Collecting beautifulsoup4 (from nbconvert>=6.4.4->jupyter-server<3,>=2.4.0->jupyterlab)
  Using cached beautifulsoup4-4.14.3-py3-none-any.whl.metadata (3.8 kB)
Collecting bleach!=5.0.0 (from bleach[css]!=5.0.0->nbconvert>=6.4.4->jupyter-server<3,>=2.4.0->jupyterlab)
  Using cached bleach-6.3.0-py3-none-any.whl.metadata (31 kB)
Collecting defusedxml (from nbconvert>=6.4.4->jupyter-server<3,>=2.4.0->jupyterlab)
  Using cached defusedxml-0.7.1-py2.py3-none-any.whl.metadata (32 kB)
Collecting jupyterlab-pygments (from nbconvert>=6.4.4->jupyter-server<3,>=2.4.0->jupyterlab)
  Using cached jupyterlab_pygments-0.3.0-py3-none-any.whl.metadata (4.4 kB)
Collecting mistune<4,>=2.0.3 (from nbconvert>=6.4.4->jupyter-server<3,>=2.4.0->jupyterlab)
  Using cached mistune-3.1.4-py3-none-any.whl.metadata (1.8 kB)
Collecting nbclient>=0.5.0 (from nbconvert>=6.4.4->jupyter-server<3,>=2.4.0->jupyterlab)
  Using cached nbclient-0.10.2-py3-none-any.whl.metadata (8.3 kB)
Collecting pandocfilters>=1.4.1 (from nbconvert>=6.4.4->jupyter-server<3,>=2.4.0->jupyterlab)
  Using cached pandocfilters-1.5.1-py2.py3-none-any.whl.metadata (9.0 kB)
Requirement already satisfied: webencodings in c:\users\syed zain ul aideen\appdata\local\packages\pythonsoftwarefoundation.python.3.12_qbz5n2kfra8p0\localcache\local-packages\python312\site-packages (from bleach!=5.0.0->nbconvert>=6.4.4->jupyter-server<3,>=2.4.0->jupyterlab) (0.5.1)
Requirement already satisfied: tinyccs2<1.5,>=1.1.0 in c:\users\syed zain ul aideen\appdata\local\packages\pythonsoftwarefoundation.python.3.12_qbz5n2kfra8p0\localcache\local-packages\python312\site-packages (from bleach[css]!=5.0.0->nbconvert>=6.4.4->jupyter-server<3,>=2.4.0->jupyterlab) (1.4.0)
Requirement already satisfied: fastjsonschema>=2.15 in c:\users\syed zain ul aideen\appdata\local\packages\pythonsoftwarefoundation.python.3.12_qbz5n2kfra8p0\localcache\local-packages\python312\site-packages (from nbformat>=5.3.0->jupyter-server<3,>=2.4.0->jupyterlab) (2.21.2)
Requirement already satisfied: six>=1.5 in c:\users\syed zain ul aideen\appdata\local\packages\pythonsoftwarefoundation.python.3.12_qbz5n2kfra8p0\localcache\local-packages\python312\site-packages (from python-dateutil>=2.8.2->pandas) (1.17.0)
Collecting lark=>1.2.2 (from rfc3987-syntax>=1.1.0->jsonschema[format-nongpl]>=4.18.0->jupyter-events>=0.11.0->jupyter-server<3,>=2.4.0->jupyterlab)
  Using cached lark-1.3.1-py3-none-any.whl.metadata (1.8 kB)
Collecting executing>=1.2.0 (from stack_data>=0.6.0->ipython>=7.23.1->ipykernel!=6.30.0,>=6.5.0->jupyterlab)
  Using cached executing-2.2.1-py2.py3-none-any.whl.metadata (8.9 kB)
Collecting asttokens>=2.1.0 (from stack_data>=0.6.0->ipython>=7.23.1->ipykernel!=6.30.0,>=6.5.0->jupyterlab)
  Using cached asttokens-3.0.1-py3-none-any.whl.metadata (4.9 kB)
Requirement already satisfied: pure-eval in c:\users\syed zain ul aideen\appdata\local\packages\pythonsoftwarefoundation.python.3.12_qbz5n2kfra8p0\localcache\local-packages\python312\site-packages (from stack_data>=0.6.0->ipython>=7.23.1->ipykernel!=6.30.0,>=6.5.0->jupyterlab) (0.2.3)
Collecting cffi>=1.0.1 (from argon2-cffi-bindings->argon2-cffi>=21.1->jupyter-server<3,>=2.4.0->jupyterlab)
  Using cached cffi-2.0.0-cp312-cp312-win_amd64.whl.metadata (2.6 kB)
Collecting pycparser (from cffi>=1.0.1->argon2-cffi-bindings->argon2-cffi>=21.1->jupyter-server<3,>=2.4.0->jupyterlab)
  Using cached pycparser-2.23-py3-none-any.whl.metadata (993 bytes)
Requirement already satisfied: soupsieve>=1.6.1 in c:\users\syed zain ul aideen\appdata\local\packages\pythonsoftwarefoundation.python.3.12_qbz5n2kfra8p0\localcache\local-packages\python312\site-packages (from beautifulsoup4->nbconvert>=6.4.4->jupyter-server<3,>=2.4.0->jupyterlab) (2.8)
Collecting arrow>=0.15.0 (from isoduration->jsonschema[format-nongpl]>=4.18.0->jupyter-events>=0.11.0->jupyter-server<3,>=2.4.0->jupyterlab)
  Using cached arrow-1.4.0-py3-none-any.whl.metadata (7.7 kB)
Using cached numpy-2.3.5-cp312-cp312-win_amd64.whl (12.8 MB)
Using cached pandas-2.3.3-cp312-cp312-win_amd64.whl (11.0 MB)
Using cached scikit_learn-1.7.2-cp312-cp312-win_amd64.whl (8.7 MB)
Using cached nltk-3.9.2-py3-none-any.whl (1.5 MB)
Using cached spacy-3.8.11-cp312-cp312-win_amd64.whl (14.2 MB)
Using cached catalogue-2.0.10-py3-none-any.whl (1.7 kB)
Using cached cymem-2.0.13-cp312-cp312-win_amd64.whl (40 kB)
Using cached murmurhash-1.0.15-cp312-cp312-win_amd64.whl (25 kB)
Using cached preshed-3.0.12-cp312-cp312-win_amd64.whl (118 kB)
Using cached requests-2.32.5-py3-none-any.whl (64 kB)
Using cached charset_normalizer-3.4.4-cp312-cp312-win_amd64.whl (107 kB)
Using cached srslry-2.5.2-cp312-cp312-win_amd64.whl (654 kB)
Using cached thinc-8.3.10-cp312-cp312-win_amd64.whl (1.7 MB)
Using cached blis-1.3.3-cp312-cp312-win_amd64.whl (6.2 MB)
Using cached confection-0.1.5-py3-none-any.whl (35 kB)
Using cached tqdm-4.67.1-py3-none-any.whl (78 kB)
Using cached typer_slim-0.20.0-py3-none-any.whl (47 kB)
Using cached wasabi-1.1.3-py3-none-any.whl (27 kB)
Using cached weasel-0.4.3-py3-none-any.whl (50 kB)
Using cached cloudpathlib-0.23.0-py3-none-any.whl (62 kB)
Using cached smart_open-7.5.0-py3-none-any.whl (63 kB)
Using cached gensim-4.4.0-cp312-cp312-win_amd64.whl (24.4 MB)
Using cached jupyterlab-4.5.0-py3-none-any.whl (12.4 MB)
Using cached httpx-0.28.1-py3-none-any.whl (73 kB)
Using cached httpcore-1.0.9-py3-none-any.whl (78 kB)
```

```
Windows PowerShell
Collecting executing>=1.2.0 (from stack_data>=0.6.0->ipython>=7.23.1->ipykernel!=6.30.0,>=6.5.0->jupyterlab)
  Using cached executing-2.2.1-py2.py3-none-any.whl.metadata (8.9 kB)
Collecting asttokens>=2.1.0 (from stack_data>=0.6.0->ipython>=7.23.1->ipykernel!=6.30.0,>=6.5.0->jupyterlab)
  Using cached asttokens-3.0.1-py3-none-any.whl.metadata (4.9 kB)
Requirement already satisfied: pure-eval in c:\users\syed zain ul aideen\appdata\local\packages\pythonsoftwarefoundation.python.3.12_qbz5n2kfra8p0\localcache\local-packages\python312\site-packages (from stack_data>=0.6.0->ipython>=7.23.1->ipykernel!=6.30.0,>=6.5.0->jupyterlab) (0.2.3)
Collecting cffi>=1.0.1 (from argon2-cffi-bindings->argon2-cffi>=21.1->jupyter-server<3,>=2.4.0->jupyterlab)
  Using cached cffi-2.0.0-cp312-cp312-win_amd64.whl.metadata (2.6 kB)
Collecting pycparser (from cffi>=1.0.1->argon2-cffi-bindings->argon2-cffi>=21.1->jupyter-server<3,>=2.4.0->jupyterlab)
  Using cached pycparser-2.23-py3-none-any.whl.metadata (993 bytes)
Requirement already satisfied: soupsieve>=1.6.1 in c:\users\syed zain ul aideen\appdata\local\packages\pythonsoftwarefoundation.python.3.12_qbz5n2kfra8p0\localcache\local-packages\python312\site-packages (from beautifulsoup4->nbconvert>=6.4.4->jupyter-server<3,>=2.4.0->jupyterlab) (2.8)
Collecting arrow>=0.15.0 (from isoduration->jsonschema[format-nongpl]>=4.18.0->jupyter-events>=0.11.0->jupyter-server<3,>=2.4.0->jupyterlab)
  Using cached arrow-1.4.0-py3-none-any.whl.metadata (7.7 kB)
Using cached numpy-2.3.5-cp312-cp312-win_amd64.whl (12.8 MB)
Using cached pandas-2.3.3-cp312-cp312-win_amd64.whl (11.0 MB)
Using cached scikit_learn-1.7.2-cp312-cp312-win_amd64.whl (8.7 MB)
Using cached nltk-3.9.2-py3-none-any.whl (1.5 MB)
Using cached spacy-3.8.11-cp312-cp312-win_amd64.whl (14.2 MB)
Using cached catalogue-2.0.10-py3-none-any.whl (1.7 kB)
Using cached cymem-2.0.13-cp312-cp312-win_amd64.whl (40 kB)
Using cached murmurhash-1.0.15-cp312-cp312-win_amd64.whl (25 kB)
Using cached preshed-3.0.12-cp312-cp312-win_amd64.whl (118 kB)
Using cached requests-2.32.5-py3-none-any.whl (64 kB)
Using cached charset_normalizer-3.4.4-cp312-cp312-win_amd64.whl (107 kB)
Using cached srslry-2.5.2-cp312-cp312-win_amd64.whl (654 kB)
Using cached thinc-8.3.10-cp312-cp312-win_amd64.whl (1.7 MB)
Using cached blis-1.3.3-cp312-cp312-win_amd64.whl (6.2 MB)
Using cached confection-0.1.5-py3-none-any.whl (35 kB)
Using cached tqdm-4.67.1-py3-none-any.whl (78 kB)
Using cached typer_slim-0.20.0-py3-none-any.whl (47 kB)
Using cached wasabi-1.1.3-py3-none-any.whl (27 kB)
Using cached weasel-0.4.3-py3-none-any.whl (50 kB)
Using cached cloudpathlib-0.23.0-py3-none-any.whl (62 kB)
Using cached smart_open-7.5.0-py3-none-any.whl (63 kB)
Using cached gensim-4.4.0-cp312-cp312-win_amd64.whl (24.4 MB)
Using cached jupyterlab-4.5.0-py3-none-any.whl (12.4 MB)
Using cached httpx-0.28.1-py3-none-any.whl (73 kB)
Using cached httpcore-1.0.9-py3-none-any.whl (78 kB)
```

```
Windows PowerShell
Using cached httpx-0.28.1-py3-none-any.whl (73 kB)
Using cached httpcore-1.0.9-py3-none-any.whl (78 kB)
Using cached jupyter_server-2.17.0-py3-none-any.whl (388 kB)
Using cached jupyterlab_server-2.28.0-py3-none-any.whl (59 kB)
Using cached argon2_cffi-25.1.0-py3-none-any.whl (14 kB)
Using cached async_tru-2.0.5-py3-none-any.whl (6.1 kB)
Using cached babel-2.17.0-py3-none-any.whl (10.2 MB)
Using cached certifi-2025.11.12-py3-none-any.whl (159 kB)
Using cached ipykernal-7.1.0-py3-none-any.whl (117 kB)
Using cached comm-0.2.3-py3-none-any.whl (7.3 kB)
Using cached debugpy-1.8.17-cp312-cp312-win_amd64.whl (5.3 MB)
Using cached ipython-9.7.0-py3-none-any.whl (618 kB)
Using cached prompt_toolkit-3.0.52-py3-none-any.whl (391 kB)
Using cached decorator-5.2.1-py3-none-any.whl (9.2 kB)
Using cached ipython_pygments_lexers-1.1.1-py3-none-any.whl (8.1 kB)
Using cached jedi-0.19.2-py2.py3-none-any.whl (1.6 MB)
Using cached parso-0.8.5-py2.py3-none-any.whl (106 kB)
Using cached joblib-1.5.2-py3-none-any.whl (308 kB)
Using cached json5-0.12.1-py3-none-any.whl (36 kB)
Using cached jsonschema-4.25.1-py3-none-any.whl (90 kB)
Using cached attrs-25.4.0-py3-none-any.whl (67 kB)
Using cached jsonschema_specifications-2025.9.1-py3-none-any.whl (18 kB)
Using cached jupyter_client-8.6.3-py3-none-any.whl (106 kB)
Using cached jupyter_core-5.9.1-py3-none-any.whl (29 kB)
Using cached jupyter_events-0.12.0-py3-none-any.whl (19 kB)
Using cached jsonpointer-3.0.0-py2.py3-none-any.whl (7.6 kB)
Using cached jupyter_lsp-2.3.0-py3-none-any.whl (76 kB)
Using cached jupyter_server_terminals-0.5.3-py3-none-any.whl (13 kB)
Using cached matplotlib_inline-2.0.1-py3-none-any.whl (9.5 kB)
Using cached nbconvert-7.16.6-py3-none-any.whl (258 kB)
Using cached mistune-3.1.4-py3-none-any.whl (53 kB)
Using cached bleach-6.3.0-py3-none-any.whl (164 kB)
Using cached nbclient-0.10.2-py3-none-any.whl (25 kB)
Using cached nbformat-5.10.4-py3-none-any.whl (78 kB)
Using cached nest_asyncio-1.6.0-py3-none-any.whl (5.2 kB)
Using cached notebook_shim-0.2.4-py3-none-any.whl (13 kB)
Using cached packaging-25.0-py3-none-any.whl (66 kB)
Using cached pandocfilters-1.5.1-py2.py3-none-any.whl (8.7 kB)
Using cached platformdirs-4.5.0-py3-none-any.whl (18 kB)
Using cached prometheus_client-0.23.1-py3-none-any.whl (61 kB)
Using cached psutil-7.1.3-cp37abi3-win_amd64.whl (247 kB)
Using cached pygments-2.19.2-py3-none-any.whl (1.2 MB)
```

```
Windows PowerShell
Using cached psutil-7.1.3-cp37abi3-win_amd64.whl (247 kB)
Using cached pygments-2.19.2-py3-none-any.whl (1.2 MB)
Using cached python_dateutil-2.9.0.post0-py2.py3-none-any.whl (229 kB)
Using cached python_json_logger-4.0.0-py3-none-any.whl (15 kB)
Using cached pywinpty-3.0.2-cp312-cp312-win_amd64.whl (2.1 MB)
Using cached pyyaml-6.0.3-cp312-cp312-win_amd64.whl (154 kB)
Using cached pyzmq-27.1.0-cp312-abi3-win_amd64.whl (619 kB)
Using cached referencing-0.37.0-py3-none-any.whl (26 kB)
Using cached regex-2025.11.3-cp312-cp312-win_amd64.whl (277 kB)
Using cached rfc3986_validator-0.1.1-py2.py3-none-any.whl (4.2 kB)
Using cached rfc3987_syntax-1.1.0-py3-none-any.whl (8.0 kB)
Using cached lark-1.3.1-py3-none-any.whl (113 kB)
Using cached rpds_py-0.30.0-cp312-cp312-win_amd64.whl (240 kB)
Using cached scipy-1.16.3-cp312-cp312-win_amd64.whl (38.6 MB)
Using cached Send2Trash-1.8.3-py3-none-any.whl (18 kB)
Using cached setuptools-80.9.0-py3-none-any.whl (1.2 MB)
Using cached stack_data-0.6.3-py3-none-any.whl (24 kB)
Using cached asttokens-3.0.1-py3-none-any.whl (27 kB)
Using cached executing-2.2.1-py2.py3-none-any.whl (28 kB)
Using cached terminado-0.18.1-py3-none-any.whl (14 kB)
Using cached argon2_cffi_bindings-25.1.0-cp39abi3-win_amd64.whl (31 kB)
Using cached cffi-2.0.0-cp312-cp312-win_amd64.whl (183 kB)
Using cached beautifulsoup4-4.14.3-py3-none-any.whl (107 kB)
Using cached defusedxml-0.7.1-py2.py3-none-any.whl (25 kB)
Using cached fadn-1.5.1-py3-none-any.whl (9.1 kB)
Using cached isoduration-20.11.0-py3-none-any.whl (11 kB)
Using cached arrow-1.4.0-py3-none-any.whl (68 kB)
Using cached cached jupyterlab_pygments-0.3.0-py3-none-any.whl (15 kB)
Using cached pycparser-2.23-py3-none-any.whl (118 kB)
Using cached rfc3339_validator-0.1.4-py2.py3-none-any.whl (3.5 kB)
Installing collected packages: wasabi, tqdm, smart-open, setuptools, send2trash, rpds-py, rfc3986-validator, rfc3339-validator, regex, pyzmq, pyyaml, pywinpty, python-json-logger, python-dateutil, pygments, pycparser, psutil, prompt_toolkit, prometheus-client, platformdirs, parso, pandocfilters, packaging, numpy, nest-asyncio, murmurhash, mistune, matplotlib-inline, lark, jupyterlab_pygments, jsonpointer, jsonschema, joblib, fadn, executing, defusedxml, decorator, debugpy, cymem, comm, cloudpathlib, charset_normalizer, certifi, catalog, ue, bleach, beautifulsoup4, babel, attrs, async_tru, asttokens, typer-slim, terminado, stack_data, srslv, scipy, rfc3987-syntax, requests, referencing, preshed, pandas, nltk, jupyter-core, jedi, ipython_pygments_lexers, httpcore, cffi, blis, arrow, scikit-learn, jupyter-server-terminals, jupyter-client, jsonschema-specifications, isoduration, ipython, https, gensim, confection, argon2_cffi-bindings, weasel, thinc, jsonschema, ipykernal, argon2_cffi, spacy, nbformat, nbclient, jupyter-events, nbconvert, jupyter-events, jupyter-server, notebook Shim, jupyterlab-server, jupyter-lsp, jupyterlab
```

1/93 [tqdm] WARNING: The script tqdm.exe is installed in 'C:\Users\SYED ZAIN UL AIDEEN\appData\Local\Packages\PythonSoftwareFoundation.Python.3.12\_qbz5n2kfra8p0\LocalCache\local-packages\Python312\Scripts' which is not on PATH.

```
Windows PowerShell
Defaulting to user installation because normal site-packages is not writeable
Collecting whoosh
  Using cached Whoosh-2.7.4-py2.py3-none-any.whl.metadata (3.1 kB)
Collecting annoy
  Using cached annoy-1.17.3.tar.gz (647 kB)
    Installing build dependencies ... done
      Getting requirements to build wheel ... done
      Installing backend dependencies ... done
      Preparing metadata (pyproject.toml) ... done
Collecting hnsllib
  Using cached hnsllib-0.8.0.tar.gz (36 kB)
    Installing build dependencies ... done
      Getting requirements to build wheel ... done
      Preparing metadata (pyproject.toml) ... done
Collecting numpy (from hnsllib)
  Using cached numpy-2.3.5-cp312-cp312-win_amd64.whl.metadata (60 kB)
Using cached Whoosh-2.7.4-py2.py3-none-any.whl (468 kB)
Using cached numpy-2.3.5-cp312-cp312-win_amd64.whl (12.8 MB)
Building wheels for collected packages: annoy, hnsllib
  Building wheel for annoy (pyproject.toml) ... error
    error: subprocess-exited-with-error

      × Building wheel for annoy (pyproject.toml) did not run successfully.
      | exit code: 1
      |> [12 lines of output]
        C:\Users\SYED_ZAIN_UL_AIDEEN\AppData\Local\Temp\pip-build-env_9ujt05h\overlay\Lib\site-packages\setuptools\_distutils\dist.py:2
89: UserWarning: Unknown distribution option: 'tests_require'
        warnings.warn(msg)
        running bdist_wheel
        running build
        running build_py
        creating build\lib.win-amd64-cpython-312\annoy
        copying annoy\__init__.py -> build\lib.win-amd64-cpython-312\annoy
        copying annoy\__init__.pyi -> build\lib.win-amd64-cpython-312\annoy
        copying annoy\py.typed -> build\lib.win-amd64-cpython-312\annoy
        running build_ext
          building annoy.annoylib' extension
          error: Microsoft Visual C++ 14.0 or greater is required. Get it with "Microsoft C++ Build Tools": https://visualstudio.microsoft.com/visual-cpp-build-tools/
      [end of output]


```

```
Windows PowerShell
error: subprocess-exited-with-error

  × Building wheel for hnsllib (pyproject.toml) did not run successfully.
  | exit code: 1
  |> [5 lines of output]
    running bdist_wheel
    running build
    running build_ext
    building 'hnsllib' extension
    error: Microsoft Visual C++ 14.0 or greater is required. Get it with "Microsoft C++ Build Tools": https://visualstudio.microsoft.com/visual-cpp-build-tools/
  [end of output]

  note: This error originates from a subprocess, and is likely not a problem with pip.
  ERROR: Failed building wheel for hnsllib
Failed to build annoy hnsllib
error: failed-wheel-build-for-install

  × Failed to build installable wheels for some pyproject.toml based projects
  |> annoy, hnsllib
Defaulting to user installation because normal site-packages is not writeable
Requirement already satisfied: flask in c:\users\syed zain ul aideen\appdata\local\packages\pythonsoftwarefoundation.python.3.12_qbz5n2kfra8p0\localcache\local-packages\python312\site-packages (3.1.2)
Requirement already satisfied: fastapi in c:\users\syed zain ul aideen\appdata\local\packages\pythonsoftwarefoundation.python.3.12_qbz5n2kfra8p0\localcache\local-packages\python312\site-packages (0.123.0)
Requirement already satisfied: uvicorn in c:\users\syed zain ul aideen\appdata\local\packages\pythonsoftwarefoundation.python.3.12_qbz5n2kfra8p0\localcache\local-packages\python312\site-packages (0.38.0)
Requirement already satisfied: blinker>=1.9.0 in c:\users\syed zain ul aideen\appdata\local\packages\pythonsoftwarefoundation.python.3.12_qbz5n2kfra8p0\localcache\local-packages\python312\site-packages (from flask) (1.9.0)
Requirement already satisfied: click>=8.1.3 in c:\users\syed zain ul aideen\appdata\local\packages\pythonsoftwarefoundation.python.3.12_qbz5n2kfra8p0\localcache\local-packages\python312\site-packages (from flask) (8.3.1)
Requirement already satisfied: itsdangerous>=2.2.0 in c:\users\syed zain ul aideen\appdata\local\packages\pythonsoftwarefoundation.python.3.12_qbz5n2kfra8p0\localcache\local-packages\python312\site-packages (from flask) (2.2.0)
Requirement already satisfied: jinja2>=3.1.2 in c:\users\syed zain ul aideen\appdata\local\packages\pythonsoftwarefoundation.python.3.12_qbz5n2kfra8p0\localcache\local-packages\python312\site-packages (from flask) (3.1.6)
Requirement already satisfied: markupsafe>=2.1.1 in c:\users\syed zain ul aideen\appdata\local\packages\pythonsoftwarefoundation.python.3.12_qbz5n2kfra8p0\localcache\local-packages\python312\site-packages (from flask) (3.0.3)
Requirement already satisfied: werkzeug>=3.1.0 in c:\users\syed zain ul aideen\appdata\local\packages\pythonsoftwarefoundation.python.3.12_qbz5n2kfra8p0\localcache\local-packages\python312\site-packages (from flask) (3.1.4)
Requirement already satisfied: starlette<0.51.0,>=0.40.0 in c:\users\syed zain ul aideen\appdata\local\packages\pythonsoftwarefoundation.python.3.12_qbz5n2kfra8p0\localcache\local-packages\python312\site-packages (from fastapi) (0.50.0)
Requirement already satisfied: pydantic!=1.8,!1.8.1,!2.0.0,!2.0.1,!2.1.0,<3.0.0,>=1.7.4 in c:\users\syed zain ul aideen\appdata\lo
```

```
Windows PowerShell
PS C:\Users\syed.zain.ul.aideen\AppData\Local\Temp\pip-req-build-1234567890\local-cache\local-packages\python312\site-packages (from fastapi) (4.15.0)
Requirement already satisfied: annotated-doc>=0.2.0 in c:\users\syed.zain.ul.aideen\appdata\local\packages\pythonsoftwarefoundation.python.3.12_qbz5n2kfra8p0\localcache\local-packages\python312\site-packages (from fastapi) (0.4)
Requirement already satisfied: annotated-types>=0.6.0 in c:\users\syed.zain.ul.aideen\appdata\local\packages\pythonsoftwarefoundation.python.3.12_qbz5n2kfra8p0\localcache\local-packages\python312\site-packages (from pydantic!=1.8,!=1.8.1,!=2.0.0,!=2.0.1,!=2.1.0,<3.0.0,!=1.7.4->fastapi) (0.7.0)
Requirement already satisfied: pydantic-core==2.41.5 in c:\users\syed.zain.ul.aideen\appdata\local\packages\pythonsoftwarefoundation.python.3.12_qbz5n2kfra8p0\localcache\local-packages\python312\site-packages (from pydantic!=1.8,!=1.8.1,!=2.0.0,!=2.0.1,!=2.1.0,<3.0.0,!=1.7.4->fastapi) (2.41.5)
Requirement already satisfied: typing-inspection>=0.4.2 in c:\users\syed.zain.ul.aideen\appdata\local\packages\pythonsoftwarefoundation.python.3.12_qbz5n2kfra8p0\localcache\local-packages\python312\site-packages (from pydantic!=1.8,!=1.8.1,!=2.0.0,!=2.0.1,!=2.1.0,<3.0.0,!=1.7.4->fastapi) (0.4.2)
Requirement already satisfied: aioio<,>=3.6.2 in c:\users\syed.zain.ul.aideen\appdata\local\packages\pythonsoftwarefoundation.python.3.12_qbz5n2kfra8p0\localcache\local-packages\python312\site-packages (from starlette<0.51.0,>=0.40.0->fastapi) (4.12.0)
Requirement already satisfied: idna>=2.8 in c:\users\syed.zain.ul.aideen\appdata\local\packages\pythonsoftwarefoundation.python.3.12_qbz5n2kfra8p0\localcache\local-packages\python312\site-packages (from aioio<,>=3.6.2->starlette<0.51.0,>=0.40.0->fastapi) (3.11)
Requirement already satisfied: h11>=0.8 in c:\users\syed.zain.ul.aideen\appdata\local\packages\pythonsoftwarefoundation.python.3.12_qbz5n2kfra8p0\localcache\local-packages\python312\site-packages (from uvicorn) (0.16.0)
Requirement already satisfied: colorama in c:\users\syed.zain.ul.aideen\appdata\local\packages\pythonsoftwarefoundation.python.3.12_qbz5n2kfra8p0\localcache\local-packages\python312\site-packages (from click=>8.1.3->flask) (0.4.6)
Defaulting to user installation because normal site-packages is not writeable
Collecting faiss-cpu
  Using cached faiss_cgpu-1.13.0-cp312-cp312-win_amd64.whl.metadata (7.7 kB)
Collecting numpy<3.0,>=1.25.0 (from faiss-cpu)
  Using cached numpy-2.3.5-cp312-cp312-win_amd64.whl.metadata (60 kB)
Collecting packaging (from faiss-cpu)
  Using cached packaging-25.0-py3-none-any.whl.metadata (3.3 kB)
Using cached faiss_cgpu-1.13.0-cp312-cp312-win_amd64.whl (18.7 MB)
Using cached numpy-2.3.5-cp312-cp312-win_amd64.whl (12.8 MB)
Using cached packaging-25.0-py3-none-any.whl (66 kB)
Installing collected packages: packaging, numpy, faiss-cpu
    1/3 [numpy] WARNING: The scripts f2py.exe and numpy-config.exe are installed in 'C:\Users\SYED.ZAIN.UL.AIDEEN\AppData\Local\Programs\PythonSoftwareFoundation.Python.3.12_qbz5n2kfra8p0\LocalCache\local-packages\Python312\Scripts' which is not on PATH.
      Consider adding this directory to PATH or, if you prefer to suppress this warning, use --no-warn-script-location.
Successfully installed faiss-cpu-1.13.0 numpy-2.3.5 packaging-25.0
PS C:\projects\ir-system> /bin/bash -c "$(curl -fsSL https://raw.githubusercontent.com/Homebrew/install/HEAD/install.sh)"
>> /bin/bash -c "$(curl -fsSL https://raw.githubusercontent.com/Homebrew/install/HEAD/install.sh)"
>> brew install python git
>> python3 --version
>> git --version
>> cd C:\projects\ir-system
```

```
Windows PowerShell
PS C:\projects\ir-system> python --version
>>
Python 3.12.10
PS C:\projects\ir-system> python script_name.py
>>
C:\Users\SYED.ZAIN.UL.AIDEEN\AppData\Local\Microsoft\WindowsApps\python.exe: can't open file 'C:\\\\projects\\\\ir-system\\\\script_name.py': [Errno 2] No such file or directory
PS C:\projects\ir-system> dir
>>

  Directory: C:\projects\ir-system

Mode                LastWriteTime         Length Name
----              -----          ----  --
d----       12/1/2025 10:22 AM            .venv

PS C:\projects\ir-system> python main.py
>>
C:\Users\SYED.ZAIN.UL.AIDEEN\AppData\Local\Microsoft\WindowsApps\python.exe: can't open file 'C:\\\\projects\\\\ir-system\\\\main.py': [Errno 2] No such file or directory
PS C:\projects\ir-system> notepad main.py
>>
PS C:\projects\ir-system> dir
>>

  Directory: C:\projects\ir-system

Mode                LastWriteTime         Length Name
----              -----          ----  --
d----       12/1/2025 10:22 AM            .venv
-a---       12/1/2025 11:06 AM           28  main.py

PS C:\projects\ir-system> python main.py
>>
Hello, IR system!
PS C:\projects\ir-system> /bin/bash -c "$(curl -fsSL https://raw.githubusercontent.com/Homebrew/install/HEAD/install.sh)"
```

```
Windows PowerShell
PS C:\> cd C:\projects\ir-system
>> notepad requirements.txt
>>
PS C:\projects\ir-system> pip install -r requirements.txt
>>
Defaulting to user installation because normal site-packages is not writeable
Requirement already satisfied: numpy in c:\users\syed zain ul aideen\appdata\local\packages\pythonsoftwarefoundation.python.3.12_qbz5n2kfra8p0\localcache\local-packages\python312\site-packages (from -r requirements.txt (line 1)) (2.3.5)
Collecting pandas (from -r requirements.txt (line 2))
  Using cached pandas-2.3.3-cp312-cp312-win_amd64.whl.metadata (19 kB)
Collecting scikit-learn (from -r requirements.txt (line 3))
  Using cached scikit_learn-1.7.2-cp312-cp312-win_amd64.whl.metadata (11 kB)
Collecting nltk (from -r requirements.txt (line 4))
  Using cached nltk-3.9.2-py3-none-any.whl.metadata (3.2 kB)
Requirement already satisfied: Flask in c:\users\syed zain ul aideen\appdata\local\packages\pythonsoftwarefoundation.python.3.12_qbz5n2kfra8p0\localcache\local-packages\python312\site-packages (from -r requirements.txt (line 5)) (3.1.2)
Collecting python-dateutil>=2.8.2 (from pandas->-r requirements.txt (line 2))
  Using cached python_dateutil-2.9.0.post0-py2.py3-none-any.whl.metadata (8.4 kB)
Requirement already satisfied: pytz>=2020.1 in c:\users\syed zain ul aideen\appdata\local\packages\pythonsoftwarefoundation.python.3.12_qbz5n2kfra8p0\localcache\local-packages\python312\site-packages (from pandas->-r requirements.txt (line 2)) (2025.2)
Requirement already satisfied: tzdata>=2022.7 in c:\users\syed zain ul aideen\appdata\local\packages\pythonsoftwarefoundation.python.3.12_qbz5n2kfra8p0\localcache\local-packages\python312\site-packages (from pandas->-r requirements.txt (line 2)) (2025.2)
Collecting scipy>=1.8.0 (from scikit-learn->-r requirements.txt (line 3))
  Using cached scipy-1.16.3-cp312-cp312-win_amd64.whl.metadata (60.0 kB)
Collecting joblib>=1.2.0 (from scikit-learn->-r requirements.txt (line 3))
  Using cached joblib-1.5.2-py3-none-any.whl.metadata (5.6 kB)
Requirement already satisfied: threadpoolctl>=3.1.0 in c:\users\syed zain ul aideen\appdata\local\packages\pythonsoftwarefoundation.python.3.12_qbz5n2kfra8p0\localcache\local-packages\python312\site-packages (from scikit-learn->-r requirements.txt (line 3)) (3.6.0)
Requirement already satisfied: click in c:\users\syed zain ul aideen\appdata\local\packages\pythonsoftwarefoundation.python.3.12_qbz5n2kfra8p0\localcache\local-packages\python312\site-packages (from nltk->-r requirements.txt (line 4)) (8.3.1)
Collecting regex>=2021.8.3 (from nltk->-r requirements.txt (line 4))
  Using cached regex-2025.11.3-cp312-cp312-win_amd64.whl.metadata (41 kB)
Requirement already satisfied: tqdm in c:\users\syed zain ul aideen\appdata\local\packages\pythonsoftwarefoundation.python.3.12_qbz5n2kfra8p0\localcache\local-packages\python312\site-packages (from nltk->-r requirements.txt (line 4)) (4.67.1)
Requirement already satisfied: blinker>=1.9.0 in c:\users\syed zain ul aideen\appdata\local\packages\pythonsoftwarefoundation.python.3.12_qbz5n2kfra8p0\localcache\local-packages\python312\site-packages (from flask->-r requirements.txt (line 5)) (1.9.0)
Requirement already satisfied: itsdangerous>=2.2.0 in c:\users\syed zain ul aideen\appdata\local\packages\pythonsoftwarefoundation.python.3.12_qbz5n2kfra8p0\localcache\local-packages\python312\site-packages (from flask->-r requirements.txt (line 5)) (2.2.0)
Requirement already satisfied: jinja2>=3.1.2 in c:\users\syed zain ul aideen\appdata\local\packages\pythonsoftwarefoundation.python.3.12_qbz5n2kfra8p0\localcache\local-packages\python312\site-packages (from flask->-r requirements.txt (line 5)) (3.1.6)
Requirement already satisfied: markupsafe>=2.1.1 in c:\users\syed zain ul aideen\appdata\local\packages\pythonsoftwarefoundation.python.3.12_qbz5n2kfra8p0\localcache\local-packages\python312\site-packages (from flask->-r requirements.txt (line 5)) (3.0.3)
```

```
Select Command Prompt - jupyter notebook
Microsoft Windows [Version 10.0.19045.6456]
(c) Microsoft Corporation. All rights reserved.

C:\Users\SYED ZAIN UL AIDEEN\.venv\Scripts\activate.bat

(.venv) C:\Users\SYED ZAIN UL AIDEEN>python --version
Python 3.12.10

(.venv) C:\Users\SYED ZAIN UL AIDEEN>numpy
'numpy' is not recognized as an internal or external command,
operable program or batch file.

(.venv) C:\Users\SYED ZAIN UL AIDEEN>pandas
'pandas' is not recognized as an internal or external command,
operable program or batch file.

(.venv) C:\Users\SYED ZAIN UL AIDEEN>scikit-learn
'scikit-learn' is not recognized as an internal or external command,
operable program or batch file.

(.venv) C:\Users\SYED ZAIN UL AIDEEN>nltk
'nltk' is not recognized as an internal or external command,
operable program or batch file.

(.venv) C:\Users\SYED ZAIN UL AIDEEN>flask # if you want a web API
'flask' is not recognized as an internal or external command,
operable program or batch file.

(.venv) C:\Users\SYED ZAIN UL AIDEEN>numpy
'numpy' is not recognized as an internal or external command,
operable program or batch file.

(.venv) C:\Users\SYED ZAIN UL AIDEEN>pip install -r requirements.txt

[notice] A new release of pip is available: 25.0.1 > 25.3
[notice] To update, run: python.exe -m pip install --upgrade pip
ERROR: Could not open requirements file: [Errno 2] No such file or directory: 'requirements.txt'

(.venv) C:\Users\SYED ZAIN UL AIDEEN>python --version
Python 3.12.10

(.venv) C:\Users\SYED ZAIN UL AIDEEN>python -m venv .venv
Error: [Errno 13] Permission denied: 'C:\Users\SYED ZAIN UL AIDEEN\.venv\Scripts\python.exe'
```

```
cmd Select Command Prompt - jupyter notebook
(.venv) C:\Users\SYED ZAIN UL AIDEEN>pip install notebook
Collecting notebook
  Downloading notebook-7.5.0-py3-none-any.whl.metadata (10 kB)
Collecting jupyter-server<3,>=2.4.0 (from notebook)
  Using cached jupyter_server-2.17.0-py3-none-any.whl.metadata (8.5 kB)
Collecting jupyterlab<4.6,>=2.28.0 (from notebook)
  Using cached jupyterlab-2.28.0-py3-none-any.whl.metadata (5.9 kB)
Collecting jupyterlab<4.6,>=4.5.0rc0 (from notebook)
  Using cached jupyterlab-4.5.0-py3-none-any.whl.metadata (16 kB)
Collecting notebook-shim<0.3,>=0.2 (from notebook)
  Using cached notebook_shim-0.2.4-py3-none-any.whl.metadata (4.0 kB)
Collecting tornado<6.2.0 (from notebook)
  Using cached tornado-6.5.2-cp39-ab13-win_amd64.whl.metadata (2.9 kB)
Collecting aiohttp<3.1.0 (from jupyter_server<3,>=2.4.0->notebook)
  Using cached aiohttp-4.12.0-py3-none-any.whl.metadata (4.3 kB)
Collecting argon2-cffi<21.1 (from jupyter-server<3,>=2.4.0->notebook)
  Using cached argon2_cffi-25.1.0-py3-none-any.whl.metadata (4.1 kB)
Collecting jinja2>=3.0.3 (from jupyter-server<3,>=2.4.0->notebook)
  Using cached jinja2-3.1.6-py3-none-any.whl.metadata (2.9 kB)
Collecting jupyter-client>=7.4.4 (from jupyter-server<3,>=2.4.0->notebook)
  Using cached jupyter_client-8.6.3-py3-none-any.whl.metadata (8.3 kB)
Collecting jupyter-core<5.0.*,>=4.12 (from jupyter-server<3,>=2.4.0->notebook)
  Using cached jupyter_core-5.9.1-py3-none-any.whl.metadata (1.5 kB)
Collecting jupyter-events<0.11.0 (from jupyter-server<3,>=2.4.0->notebook)
  Using cached jupyter_events-0.12.0-py3-none-any.whl.metadata (5.8 kB)
Collecting jupyter-server-terminals<=0.4.4 (from jupyter-server<3,>=2.4.0->notebook)
  Using cached jupyter_server_terminals-0.5.3-py3-none-any.whl.metadata (5.6 kB)
Collecting nbconvert>=6.4.4 (from jupyter-server<3,>=2.4.0->notebook)
  Using cached nbconvert-7.16.6-py3-none-any.whl.metadata (8.5 kB)
Collecting nbformat>=5.3.0 (from jupyter-server<3,>=2.4.0->notebook)
  Using cached nbformat-5.10.4-py3-none-any.whl.metadata (3.6 kB)
Collecting packaging>=22.0 (from jupyter-server<3,>=2.4.0->notebook)
  Using cached packaging-25.0-py3-none-any.whl.metadata (3.3 kB)
Collecting prometheus-client<0.9 (from jupyter-server<3,>=2.4.0->notebook)
  Using cached prometheus_client-0.23.1-py3-none-any.whl.metadata (1.9 kB)
Collecting pywinpty>=2.0.1 (from jupyter-server<3,>=2.4.0->notebook)
  Using cached pywinpty-3.0.2-cp312-cp312-win_amd64.whl.metadata (5.7 kB)
Collecting pyzmq<=24 (from jupyter-server<3,>=2.4.0->notebook)
  Using cached pyzmq-27.1.0-cp312-ab13-win_amd64.whl.metadata (6.0 kB)
Collecting send2trash<=1.8.2 (from jupyter-server<3,>=2.4.0->notebook)
  Using cached Send2Trash-1.8.3-py3-none-any.whl.metadata (4.0 kB)
Collecting terminado<=0.8.3 (from jupyter-server<3,>=2.4.0->notebook)
  Using cached terminado-0.18.1-py3-none-any.whl.metadata (5.8 kB)
```

```
cmd Select Command Prompt - jupyter notebook
Collecting httpx<1,>=0.25.0 (from jupyterlab<4.6,>=4.5.0rc0->notebook)
  Using cached httpx-0.28.1-py3-none-any.whl.metadata (7.1 kB)
Collecting ipykernel<=6.30.0,>=6.5.0 (from jupyterlab<4.6,>=4.5.0rc0->notebook)
  Using cached ipykernel-7.1.0-py3-none-any.whl.metadata (4.5 kB)
Collecting jupyter-lsp<=2.0.0 (from jupyterlab<4.6,>=4.5.0rc0->notebook)
  Using cached jupyter_lsp-2.3.0-py3-none-any.whl.metadata (1.8 kB)
Collecting setuptools<=41.1.0 (from jupyterlab<4.6,>=4.5.0rc0->notebook)
  Using cached setuptools-80.9.0-py3-none-any.whl.metadata (6.6 kB)
Collecting babel<=2.18 (from jupyterlab-server<3,>=2.28.0->notebook)
  Using cached babel-2.17.0-py3-none-any.whl.metadata (2.0 kB)
Collecting json5<=0.9.0 (from jupyterlab-server<3,>=2.28.0->notebook)
  Using cached json5-0.12.1-py3-none-any.whl.metadata (36 kB)
Collecting jsonschema<4.18.0 (from jupyterlab-server<3,>=2.28.0->notebook)
  Using cached jsonschema-4.25.1-py3-none-any.whl.metadata (7.6 kB)
Collecting requests<=2.31 (from jupyterlab-server<3,>=2.28.0->notebook)
  Using cached requests-2.32.5-py3-none-any.whl.metadata (4.9 kB)
Collecting idna<=2.8 (from anyio<3.1.0->jupyter-server<3,>=2.4.0->notebook)
  Using cached idna-3.11-py3-none-any.whl.metadata (8.4 kB)
Collecting typing_extensions<4.5 (from anyio<3.1.0->jupyter-server<3,>=2.4.0->notebook)
  Using cached typing_extensions-4.15.0-py3-none-any.whl.metadata (3.3 kB)
Collecting argon2-cffi-bindings (from argon2-cffi<21.1->jupyter-server<3,>=2.4.0->notebook)
  Using cached argon2_cffi_bindings-25.1.0-cp312-cp312-win_amd64.whl.metadata (7.5 kB)
Collecting certifi (from httpx<1,>=0.25.0->jupyterlab<4.6,>=4.5.0rc0->notebook)
  Using cached certifi-2025.11.12-py3-none-any.whl.metadata (2.5 kB)
Collecting httpcore<1.0 (from httpx<1,>=0.25.0->jupyterlab<4.6,>=4.5.0rc0->notebook)
  Using cached httpcore-1.0.9-py3-none-any.whl.metadata (21 kB)
Collecting h11<=0.16 (from httpcore<1,>=0.25.0->jupyterlab<4.6,>=4.5.0rc0->notebook)
  Using cached h11-0.16-py3-none-any.whl.metadata (8.3 kB)
Collecting comm>=0.1.1 (from ipykernel<=6.30.0,>=6.5.0->jupyterlab<4.6,>=4.5.0rc0->notebook)
  Using cached comm-0.2.3-py3-none-any.whl.metadata (3.7 kB)
Collecting debugpy<=1.6.5 (from ipykernel<=6.30.0,>=6.5.0->jupyterlab<4.6,>=4.5.0rc0->notebook)
  Using cached debugpy-1.8.17-cp312-cp312-win_amd64.whl.metadata (1.5 kB)
Collecting ipython>=7.23.1 (from ipykernel<=6.30.0,>=6.5.0->jupyterlab<4.6,>=4.5.0rc0->notebook)
  Using cached ipython-9.7.0-py3-none-any.whl.metadata (4.5 kB)
Collecting matplotlib-inline>=0.1 (from ipykernel<=6.30.0,>=6.5.0->jupyterlab<4.6,>=4.5.0rc0->notebook)
  Using cached matplotlib_inline-0.2.1-py3-none-any.whl.metadata (2.3 kB)
Collecting nest-asyncio<1.4 (from ipykernel<=6.30.0,>=6.5.0->jupyterlab<4.6,>=4.5.0rc0->notebook)
  Using cached nest_asyncio-1.6.0-py3-none-any.whl.metadata (2.8 kB)
Collecting psutil<=5.7 (from ipykernel<=6.30.0,>=6.5.0->jupyterlab<4.6,>=4.5.0rc0->notebook)
  Using cached psutil-7.1.3-cp37-ab13-win_amd64.whl.metadata (23 kB)
Collecting MarkupSafe<=2.0 (from jinja2<3.0.3->jupyter-server<3,>=2.4.0->notebook)
  Using cached markupsafe-3.0.3-cp312-cp312-win_amd64.whl.metadata (2.8 kB)
Collecting attrs>=22.2.0 (from jsonschema<4.18.0->jupyterlab-server<3,>=2.28.0->notebook)
  Using cached attrs-25.4.0-py3-none-any.whl.metadata (10 kB)
```

```
  Select Command Prompt - jupyter notebook
Collecting nbclient<=0.5.0 (from nbconvert>=6.4.4->jupyter-server<3,>=2.4.0->notebook)
  Using cached nbclient-0.10.2-py3-none-any.whl.metadata (8.3 kB)
Collecting pandocfilters<1.4.1 (from nbconvert>=6.4.4->jupyter-server<3,>=2.4.0->notebook)
  Using cached pandocfilters-1.5.1-py2.py3-none-any.whl.metadata (9.0 kB)
Collecting pygments<=2.4.1 (from nbconvert>=6.4.4->jupyter-server<3,>=2.4.0->notebook)
  Using cached pygments-2.19.2-py3-none-any.whl.metadata (2.5 kB)
Collecting fastjonschema<=2.15 (from nbformat<=5.3.0->jupyter-server<3,>=2.4.0->notebook)
  Using cached fastjonschema-2.21.2-py3-none-any.whl.metadata (2.3 kB)
Collecting charset_normalizer<4,>>2 (from requests>=2.31->jupyterlab-server<3,>=2.28.0->notebook)
  Using cached charset_normalizer-3.4.4-cp312-cp312-win_amd64.whl.metadata (38 kB)
Collecting urllib3<3,>=1.21.1 (from requests>=2.31->jupyterlab-server<3,>=2.28.0->notebook)
  Using cached urllib3-2.5.0-py3-none-any.whl.metadata (6.5 kB)
Collecting webencodings (from bleach<5.0.0->bleach[css]>=5.0.0->nbconvert>=6.4.4->jupyter-server<3,>=2.4.0->notebook)
  Using cached webencodings-0.5.1-py2.py3-none-any.whl.metadata (2.1 kB)
Collecting tinyCSS2<1.5,>=1.1.0 (from bleach[css]>=5.0.0->nbconvert>=6.4.4->jupyter-server<3,>=2.4.0->notebook)
  Using cached tinyCSS2-1.4.0-py3-none-any.whl.metadata (3.0 kB)
Collecting colorama<0.4.4 (from ipython>=7.23.1->ipykernel>=6.30.0,>=6.5.0->jupyterlab<4.6,>=4.5.0rc0->notebook)
  Using cached colorama-0.4.6-py2.py3-none-any.whl.metadata (17 kB)
Collecting decorator<4.3.2 (from ipython>=7.23.1->ipykernel>=6.30.0,>=6.5.0->jupyterlab<4.6,>=4.5.0rc0->notebook)
  Using cached decorator-5.2.1-py3-none-any.whl.metadata (3.9 kB)
Collecting ipython-pygments-lexers<1.0.0 (from ipython>=7.23.1->ipykernel>=6.30.0,>=6.5.0->jupyterlab<4.6,>=4.5.0rc0->notebook)
  Using cached ipython_pygments_lexers-1.1.1-py3-none-any.whl.metadata (1.1 kB)
Collecting jedi<0.18.1 (from ipython>=7.23.1->ipykernel>=6.30.0,>=6.5.0->jupyterlab<4.6,>=4.5.0rc0->notebook)
  Using cached jedi-0.19.2-py2.py3-none-any.whl.metadata (22 kB)
Collecting prompt_toolkit<3.1.0,>>3.0.41 (from ipython>=7.23.1->ipykernel>=6.30.0,>=6.5.0->jupyterlab<4.6,>=4.5.0rc0->notebook)
  Using cached prompt_toolkit-3.0.5-py2.py3-none-any.whl.metadata (6.4 kB)
Collecting stack_data<0.6.0 (from ipython>=7.23.1->ipykernel>=6.30.0,>=6.5.0->jupyterlab<4.6,>=4.5.0rc0->notebook)
  Using cached stack_data-0.6.3-py3-none-any.whl.metadata (18 kB)
Collecting fadde (from jsonschema[format-nongpl]>=4.18.0->jupyter-events>=0.11.0->jupyter-server<3,>=2.4.0->notebook)
  Using cached fqdn-1.5.1-py3-none-any.whl.metadata (1.4 kB)
Collecting isoduration (from jsonschema[format-nongpl]>=4.18.0->jupyter-events>=0.11.0->jupyter-server<3,>=2.4.0->notebook)
  Using cached isoduration-20.11.0-py3-none-any.whl.metadata (5.7 kB)
Collecting jsonpointer<1.11 (from jsonschema[format-nongpl]>=4.18.0->jupyter-events>=0.11.0->jupyter-server<3,>=2.4.0->notebook)
  Using cached jsonpointer-3.0.0-py2.py3-none-any.whl.metadata (2.3 kB)
Collecting rfc3987-syntax<1.1.0 (from jsonschema[format-nongpl]>=4.18.0->jupyter-events>=0.11.0->jupyter-server<3,>=2.4.0->notebook)
  Using cached rfc3987_syntax-1.1.0-py3-none-any.whl.metadata (7.7 kB)
Collecting uri_template (from jsonschema[format-nongpl]>=4.18.0->jupyter-events>=0.11.0->jupyter-server<3,>=2.4.0->notebook)
  Using cached uri_template-1.3.0-py3-none-any.whl.metadata (8.8 kB)
Collecting webcolors<24.6.0 (from jsonschema[format-nongpl]>=4.18.0->jupyter-events>=0.11.0->jupyter-server<3,>=2.4.0->notebook)
  Using cached webcolors-25.18.0-py3-none-any.whl.metadata (2.2 kB)
Collecting six<1.5 (from python-dateutil)>=4.18.0->jupyter-events>=0.11.0->jupyter-server<3,>=2.4.0->notebook)
  Using cached six-1.17.0-py2.py3-none-any.whl.metadata (1.7 kB)
Collecting cffi<1.0.1 (from argon2-cffi-bindings>=argon2-cffi>=21.1->jupyter-server<3,>=2.4.0->notebook)
  Using cached cffi-2.0.0-cp312-cp312-win_amd64.whl.metadata (6.0 kB)
```

```
  Select Command Prompt - jupyter notebook
Downloading notebook-7.5.0-py3-none-any.whl (14.5 MB)
----- 14.5/14.5 MB 654.2 kB/s eta 0:00:00
Using cached jupyter_server-2.17.0-py3-none-any.whl (388 kB)
Using cached jupyterlab-4.5.0-py3-none-any.whl (12.4 MB)
Using cached jupyterlab_server-2.28.0-py3-none-any.whl (59 kB)
Using cached notebook_shim-0.2.4-py3-none-any.whl (13 kB)
Using cached tornado-6.5.2-cp39-abi3-win_amd64.whl (445 kB)
Using cached anyio-4.12.0-py3-none-any.whl (113 kB)
Using cached argon2_cffi-25.1.0-py3-none-any.whl (14 kB)
Using cached async_lru-2.0.5-py3-none-any.whl (6.1 kB)
Using cached babel-2.17.0-py3-none-any.whl (10.2 kB)
Using cached httpx-0.28.1-py3-none-any.whl (73 kB)
Using cached httpcore-1.0.9-py3-none-any.whl (78 kB)
Using cached ipykernel-7.1.0-py3-none-any.whl (117 kB)
Using cached jinja2-3.1.6-py3-none-any.whl (134 kB)
Using cached json5-0.12.1-py3-none-any.whl (36 kB)
Using cached jsonschema-4.25.1-py3-none-any.whl (90 kB)
Using cached jupyter_client-8.6.3-py3-none-any.whl (106 kB)
Using cached jupyter_core-5.9.1-py3-none-any.whl (29 kB)
Using cached jupyter_events-0.12.0-py3-none-any.whl (19 kB)
Using cached jupyter_lsp-2.3.0-py3-none-any.whl (76 kB)
Using cached jupyter_server_terminals-0.5.3-py3-none-any.whl (13 kB)
Using cached nbconvert-7.16.6-py3-none-any.whl (258 kB)
Using cached nbformat-5.10.4-py3-none-any.whl (78 kB)
Using cached packaging-25.0-py3-none-any.whl (66 kB)
Using cached prometheus_client-0.23.1-py3-none-any.whl (61 kB)
Using cached pywinpty-3.0.2-cp312-cp312-win_amd64.whl (2.1 MB)
Using cached pyzmq-27.1.0-cp312-abi3-win_amd64.whl (619 kB)
Using cached requests-2.32.5-py3-none-any.whl (64 kB)
Using cached Send2Trash-1.8.3-py3-none-any.whl (18 kB)
Using cached setuptools-80.9.0-py3-none-any.whl (1.2 MB)
Using cached terminado-0.18.1-py3-none-any.whl (14 kB)
Using cached traitlets-5.14.3-py3-none-any.whl (85 kB)
Using cached websocket_client-1.9.0-py3-none-any.whl (82 kB)
Using cached attrs-25.4.0-py3-none-any.whl (67 kB)
Using cached bleach-6.3.0-py3-none-any.whl (164 kB)
Using cached certifi-2025.11.12-py3-none-any.whl (159 kB)
Using cached charset_normalizer-3.4.4-cp312-cp312-win_amd64.whl (107 kB)
Using cached comm-0.2.3-py3-none-any.whl (7.3 kB)
Using cached debugpy-1.8.17-cp312-cp312-win_amd64.whl (5.3 MB)
Using cached fastjonschema-2.21.2-py3-none-any.whl (24 kB)
Using cached idna-3.11-py3-none-any.whl (71 kB)
Using cached ipython-9.7.0-py3-none-any.whl (618 kB)
Using cached jsonschema_specifications-2025.9.1-py3-none-any.whl (18 kB)
```

```
[1] 12:44:08.632 Select Command Prompt - jupyter notebook
Using cached asttokens-3.0.1-py3-none-any.whl (27 kB)
Using cached executing-2.2.1-py2.py3-none-any.whl (28 kB)
Using cached lark-1.3.1-py3-none-any.whl (13 kB)
Using cached parso-0.8.5-py2.py3-none-any.whl (106 kB)
Using cached pure_eval-0.2.3-py3-none-any.whl (11 kB)
Using cached pycparser-2.23-py3-none-any.whl (118 kB)
Using cached wccwidth-0.2.14-py2.py3-none-any.whl (37 kB)
Using cached tzdata-2025-2.py2.py3-none-any.whl (347 kB)
Installing collected packages: webencodings, pure-eval, fastjsonschema, websocket-client, webcolors, wcwidth, urllib3, uri-template, tzdata, typing_extensions, traitlets, tornado, tinyccss2, soupsieve, six, setuptools, send2trash, rpdps-py, rfc3986-validator, pyzmq, pyyaml, pywinpty, python-json-logger, pygments, pycparser, putils, prometheus-client, platformdirs, parso, pandocfilters, packaging, nest-asyncio, mistune, MarkupSafe, lark, jupyterlab-pygments, jsonpointer, json5, idna, h11, fdm, executegraph, defusedxml, decorator, debugpy, conn, colorama, charset_normalizer, certifi, bleach, babel, attrs, asyncrelu, asttokens, terminado, stack_data, rfc3987-syntax, rfc3989-validator, requests, referencing, python-dateutil, prompt_toolkit, matplotlib-inline, jupyter-core, jinja2, jedi, ipython-pygments-lexers, httpcore, cffi, beautifulsoup4, anyio, jupyter-server-terminals, jupyter-client, jsonschema-specifications, ipython, httpx, arrow, argon2-cffi-bindings, jsonschema, isoduration, ipykernel, argo, h2-cffi, nbformat, nbclient, jupyter-events, nbconvert, jupyter-server, notebook-shim, jupyterlab-server, jupyter-lsp, jupyterlab, notebook
Successfully installed MarkupSafe-3.0.3 anyio-4.12.0 argon2-cffi-25.1.0 argon2-cffi-bindings-25.1.0 arrow-1.4.0 asttokens-3.0.1 async-reu-0.5 attrs-25.4.0 babel-2.17.0 beautifulsoup4-4.13.4 bleach-3.0.3 certifi-2025.11.24.1.0.20 charset_normalizer-3.4.4 colorama-0.4.6 comm-0.2.3 debugpy-1.8.17 decorator-5.2.1 defusedxml-0.7.1 executing-2.2.1 fastjsonschema-2.21.2 fdnnd-1.5.1 h11-0.16.0 httpcore-1.0.9 httpx-0.28.1 idna-3.11 ipykernel-7.1.0 ipython-9.7.0 ipython-pygments-lexers-1.1.1 isoduration-0.12.0 jupyter-lsp-2.3.0 jupyter-server-2.17.0 jupyter-server-terminals-0.5.3 jupyterlab-4.5.0 jupyterlab-pygments-0.3.0 jupyterlab-server-2.28.0 lark-1.3.1 matplotlib-inline-0.2.1 mistune-3.1.4 hbclient-0.10.2 nbconvert-7.15.6 nbformat-5.10.4 nest-asyncio-1.6.0 notebook-7.5.0 notebook-shim-0.2.4 packaging-25.0 pandocfilters-1.5.1 parso-0.8.5 platformdirs-4.5.0 prometheus-client-0.23.1 prompt_toolkit-3.0.52 putils-7.1.3 pure-eval-0.2.3 pygments-2.23 pywinpty-2.19.2 python-dateutil-2.9.0.post0 pygments-0.1.0 pywinpty-3.0.2 pyyaml-27.1.0 referencing-0.37.0 requests-2.32.5 rfc3939-validator-0.1.4 rfc3986-validator-0.1.1 rfc3987-syntax-1.1.0 rpdps-py-0.30.0 send2trash-1.8.3 setuptools-80.9.0 six-1.17.0 soupsieve-2.8 stack_data-0.6.3 terminado-0.18.1 tinyccss2-1.4.0 tornado-6.5.2 traitlets-5.14.3 typing_extensions-4.15.0 tzdata-2025.2.url template-1.3.0 urllib3-2.5.0 wcwidth-0.2.14 webcolors-25.10.0 webencodings-0.5.1 websocket-client-1.9.0

[notice] A new release of pip is available: 25.0.1 -> 25.3
[notice] To update, run: python.exe -m pip install --upgrade pip

(.venv) C:\Users\SYED ZAIN UL AIDEEN>jupyter notebook
[1] 12:44:08.632 ServerApp| Extension package jupyter_lsp took 0.7779s to import
[1] 2025-12-01 12:44:09.163 ServerApp| Extension package jupyter_server_terminals took 0.5307s to import
[1] 2025-12-01 12:44:10.363 ServerApp| jupyter_lsp | extension was successfully linked.
[1] 2025-12-01 12:44:10.379 ServerApp| jupyter_server_terminals | extension was successfully linked.
[1] 2025-12-01 12:44:10.403 ServerApp| jupyterlab | extension was successfully linked.
[1] 2025-12-01 12:44:10.422 ServerApp| notebook | extension was successfully linked.
[1] 2025-12-01 12:44:17.656 ServerApp| notebook_shim | extension was successfully linked.
[1] 2025-12-01 12:44:18.057 ServerApp| notebook_shim | extension was successfully loaded.
[1] 2025-12-01 12:44:18.063 ServerApp| jupyter_lsp | extension was successfully loaded.
[1] 2025-12-01 12:44:18.064 ServerApp| jupyter_server_terminals | extension was successfully loaded.
[1] 2025-12-01 12:44:18.089 LabApp| JupyterLab extension loaded from C:\Users\SYED ZAIN UL AIDEEN\.venv\site-packages\jupyterlab
[1] 2025-12-01 12:44:18.089 LabApp| JupyterLab application directory is C:\Users\SYED ZAIN UL AIDEEN\.venv\share\jupyter\lab
[1] 2025-12-01 12:44:18.091 LabApp| Extension Manager is 'pypi'.
[1] 2025-12-01 12:44:19.087 ServerApp| jupyterlab | extension was successfully loaded.
```

```
[4] Select Command Prompt - jupyter notebook
lib-inline-0.1.2 mistune-3.1.4 nbconvert-7.16.6 nbformat-5.10.4 nest-asyncio-1.6.0 notebook-7.5.0 notebook-shim-0.2.4 packaging-25.0 pandocfilters-1.5.1
parso-0.8.5 platformoirs-4.5.0 prometheus-client-0.23.1 prompt_toolkit-3.0.52 psutil-7.1.3 pure-eval-0.2.3 pycparser-2.23 pygments-2.19.2 python-dateutil-2.9.0.post0
python-logger-4.0.0 pywinpty-2.0.3 pyyaml-6.0.3 pyzmq-27.1.0 referencing-0.37.0 requests-2.32.5 rfc3398-validator-0.1.1 rfc3986-validator-0.1.1 rfc3987-syntax-1.1.0
rpdps-py-0.30.7 send2trash-1.8.3 setuptools-80.9.0 six-1.17.0 soupsieve-2.8 stack_data-0.6.3 terminado-0.18.1 tinyccss2-1.4.0 tornado-6.5.2 traitlets-5.14.3 typing_extensions-4.15.0 tzdata-2025.2 url-template-1.3.0 urllib3-2.5.0 wcidwidth-0.2.14 webcolors-25.10.0 webencodings-0.5.1 websocket-client-1.9.0

[Notice] A new release of pip is available: 25.0.1 -> 25.3
[Notice] To update, run: python.exe -m pip install --upgrade pip

(.venv) C:\Users\SYED ZAIN UL AIDEEN\jupyter notebook
[I 2025-12-01 12:44:08.632 ServerApp] Extension package jupyter_lsp took 0.7779s to import
[I 2025-12-01 12:44:09.163 ServerApp] Extension package jupyter_server_terminals took 0.5307s to import
[I 2025-12-01 12:44:10.363 ServerApp] jupyter_lsp | extension was successfully linked.
[I 2025-12-01 12:44:10.379 ServerApp] jupyter_server_terminals | extension was successfully linked.
[I 2025-12-01 12:44:10.403 ServerApp] jupyterlab | extension was successfully linked.
[I 2025-12-01 12:44:10.422 ServerApp] notebook | extension was successfully linked.
[I 2025-12-01 12:44:17.656 ServerApp] notebook_shim | extension was successfully linked.
[I 2025-12-01 12:44:18.057 ServerApp] notebook_shim | extension was successfully loaded.
[I 2025-12-01 12:44:18.063 ServerApp] jupyter_lsp | extension was successfully loaded.
[I 2025-12-01 12:44:18.064 ServerApp] jupyter_server_terminals | extension was successfully loaded.
[I 2025-12-01 12:44:18.088 LabApp] JupyterLab extension loaded from C:\Users\SYED ZAIN UL AIDEEN\.venv\Lib\site-packages\jupyterlab
[I 2025-12-01 12:44:18.089 LabApp] JupyterLab application directory is C:\Users\SYED ZAIN UL AIDEEN\.venv\share\jupyter\lab
[I 2025-12-01 12:44:18.091 LabApp] Extension Manager is 'ipyw'.
[I 2025-12-01 12:44:19.087 ServerApp] jupyterlab | extension was successfully loaded.
[I 2025-12-01 12:44:19.094 ServerApp] notebook | extension was successfully loaded.
[I 2025-12-01 12:44:19.095 ServerApp] The port 8888 is already in use, trying another port.
[I 2025-12-01 12:44:19.096 ServerApp] The port 8889 is already in use, trying another port.
[I 2025-12-01 12:44:19.099 ServerApp] Serving notebooks from local directory: C:\Users\SYED ZAIN UL AIDEEN
[I 2025-12-01 12:44:19.099 ServerApp] Jupyter Server 2.17.0 is running at:
[I 2025-12-01 12:44:19.100 ServerApp] http://localhost:8890/tree?token=a2912714819d9459fbcb156b533d00c0b279eaaaf3eb35f75e
[I 2025-12-01 12:44:19.100 ServerApp] http://127.0.0.1:8890/tree?token=a2912714819d9459fbcb156b533d00c0b279eaaaf3eb35f75e
[I 2025-12-01 12:44:19.101 ServerApp] Use Control-C to stop this server and shut down all kernels (twice to skip confirmation).

[4] C:\Users\SYED ZAIN UL AIDEEN\jupyter notebook

To access the server, open this file in a browser:
  file:///C:/Users/SYEDZAINULAIDEEN/AppData/Roaming/jupyter/runtime/jpserver-30352-open.html
Or copy and paste one of these URLs:
  http://localhost:8890/tree?token=a2912714819d9459fbcb156b533d00c0b279eaaaf3eb35f75e
  http://127.0.0.1:8890/tree?token=a2912714819d9459fbcb156b533d00c0b279eaaaf3eb35f75e
[I 2025-12-01 12:44:19.281 ServerApp] Skipped non-installed server(s): basemdpyright, bash-language-server, dockerfile-language-server-nodejs, javascript-typescript-language-server, jedi-language-server, julia-language-server, pyright, pyright, python-language-server, python-lsp-server, r-languageserver, sql-language-server, texlab, typescript-language-server, unified-language-server, vscode-css-languageserver-bin, vscode-html-languageserver-bin, vscode-json-languageserver-bin, yaml-language-server
```

## 2. Reproducible Pipeline

Your submission must include:

- source code

The screenshot shows a Jupyter Notebook interface with two code cells. The left sidebar lists files in the current directory, including several Jupyter notebooks related to IR practice and system development.

**Code Cell 1:**

```
# =====#
# INFORMATION RETRIEVAL SYSTEM (Boolean + TF-IDF + BM25)
# =====#
import pandas as pd
import numpy as np
import re
import nltk
import warnings

from nltk.corpus import stopwords
from sklearn.feature_extraction.text import TfidfVectorizer
from sklearn.metrics.pairwise import cosine_similarity
from rank_bm25 import BM25Okapi

# =====#
# 0. Suppress warnings and NLTK messages
# =====#
warnings.filterwarnings("ignore")
try:
    nltk.data.find("corpora/stopwords")
except:
    nltk.download("stopwords", quiet=True)

# =====#
# 1. LOAD DATASET (handle encoding automatically)
```

**Code Cell 2:**

```
# =====#
warnings.filterwarnings("ignore")
try:
    nltk.data.find("corpora/stopwords")
except:
    nltk.download("stopwords", quiet=True)

# =====#
# 1. LOAD DATASET (handle encoding automatically)
# =====#
csv_path = r"D:\MSCS24047\3rd Semester\IR&TM\HW3\Articles.csv"

# Try multiple encodings until successful
for enc in ["utf-8", "latin1", "ISO-8859-1", "cp1252"]:
    try:
        df = pd.read_csv(csv_path, encoding=enc, on_bad_lines="skip")
        break
    except Exception:
        continue

print(f"Dataset loaded successfully! Total documents: {len(df)}")
print("Columns:", df.columns.tolist())

# =====#
# 2. SELECT TEXT COLUMN
# =====#
text_col = None
for col in df.columns:
```

SYED ZAIN Imported Imported (1) Imported (2) AIOU | OAS Student Guide Grades for Syed Zain... Account Settings | ... Homepage Classwork for CS 51...

File Edit View Run Kernel Tabs Settings Help

HW-03-IR&TM IR-System.ipynb

```

# =====#
# 3. TEXT PREPROCESSING
# =====#
stop_words = set(stopwords.words("english"))

def clean_text(text):
    text = text.lower()
    text = re.sub("[^a-z0-9\s]", " ", text)
    text = text.strip()
    return text

def tokenize(text):
    tokens = clean_text(text).split()
    tokens = [t for t in tokens if t not in stop_words]
    return tokens

df["clean_text"] = df[text_col].apply(clean_text)
df["tokens"] = df["clean_text"].apply(tokenize)

# =====#
# 4. BOOLEAN RETRIEVAL
# =====#
def boolean_search(query):
    q_tokens = tokenize(query)
    results = df[df["tokens"].apply(lambda x: all(t in x for t in q_tokens))]
    return results.head(5)

```

Mode: Command Ln 8, Col 12 HW-03-IR&TM IR-System.ipynb 1 ⚡

SYED ZAIN Imported Imported (1) Imported (2) AIOU | OAS Student Guide Grades for Syed Zain... Account Settings | ... Homepage Classwork for CS 51...

File Edit View Run Kernel Tabs Settings Help

HW-03-IR&TM IR-System.ipynb

```

# =====#
# 5. BM25 RETRIEVAL
# =====#
bm25 = BM25Okapi(df["tokens"])

def retrieve_bm25(query):
    q_tokens = tokenize(query)
    scores = bm25.get_scores(q_tokens)
    top_idx = np.argsort(scores)[-5:][::-1]
    return df.iloc[top_idx]

# =====#
# 6. RUN A SAMPLE QUERY
# =====#
query = "economic policy inflation government"

print("\n===== BOOLEAN RESULTS =====")
for i, row in boolean_search(query).iterrows():
    print(f"\n--- Document {i} ---\n{row[text_col][:400]}")

print("\n===== TF-IDF RESULTS =====")
for i, row in retrieve_tfidf(query).iterrows():
    print(f"\n--- Document {i} ---\n{row[text_col][:400]}")

print("\n===== BM25 RESULTS =====")

```

Mode: Command Ln 8, Col 12 HW-03-IR&TM IR-System.ipynb 1 ⚡

SYED ZAIN Imported Imported (1) Imported (2) AIOU | OAS Student Guide Grades for Syed Zai... Account Settings | ... Homepage Classwork for CS 51... File Edit View Run Kernel Tabs Settings Help

HW-03-IR& X Python 3 (ipykernel)

Notebook

```
for i, row in retrieve_bm25(query).iterrows():
    print(f"\n--- Document {i} ---\n{row['text_col'][:400]}")
```

Dataset loaded successfully! Total documents: 2692  
Columns: ['Article', 'Date', 'Heading', 'NewsType']  
Using text column: Article

===== BOOLEAN RESULTS =====

--- Document 170 ---  
ISLAMABAD: Finance Minister Ishaq Dar presented the Federal Budget 2015-16 in the National Assembly on Friday, eyeing an ambitious economic growth target of 5.5 percent. Prime Minister Nawaz Sharif arrived in the National Assembly for the special budget session in which Finance Minister Ishaq Dar presented budgetary proposals finalised in a meeting of the federal cabinet chaired by the premier. A g

--- Document 275 ---  
Singapore: Oil prices eased in Asian trade Monday as dealers await the US Federal Reserve's decision on whether to raise interest rate following a mixed August jobs report, analysts said. US benchmark West Texas Intermediate for October delivery fell 35 cents to \$45.70 while Brent crude for October eased 38 cents to \$49.23 in mid-morning trade. The US Labor Department said Friday the economy added 1

--- Document 284 ---  
Hong Kong: A mixed reading on Chinese inflation Thursday kept Asian equities traders on edge in fresh volatility Thursday as markets retreated from a two-day rally, while fears of a US interest rate hike saw safe assets advance. Early selling spread through regional exchanges, with Hong Kong and Shanghai seeing hefty losses, despite Chinese Premier Li Keqiang the day before seeking to shore up conf

Mode: Command Ln 8, Col 12 HW-03-IR&TM IR-System.ipynb

localhost:8889/lab/tree/HW-03-IR&26TM%20IR-System.ipynb SYED ZAIN Imported Imported (1) Imported (2) AIOU | OAS Student Guide Grades for Syed Zai... Account Settings | ... Homepage Classwork for CS 51... File Edit View Run Kernel Tabs Settings Help

HW-03-IR& X Python 3 (ipykernel)

Notebook

```
--- Document 354 ---  
Hong Kong: Asian stock markets tumbled and the euro lost ground early Monday in the first trades after the deadly weekend terror attacks in Paris. The late-night assault on Friday in French capital, which killed 129 people, sparked concerns about security in Europe and its impact on the already struggling eurozone economy. It also sowed fresh uncertainty in already nervous markets, which ended last
```

--- Document 359 ---  
KARACHI: State Bank of Pakistan (SBP) on Saturday announced its monetary policy and decided to keep the policy rate unchanged at 6.0 percent. A statement issued here said that the average Jul-Oct FY16 inflation was at 1.7 percent is lower than the 7.1 percent average inflation in the corresponding period of last year. It said that the decline is broad based as both food and non-food and core inflat

===== TF-IDF RESULTS =====

--- Document 2521 ---  
**KARACHI:** The State Bank of Pakistan announced its monetary policy on Saturday. The interest rate for the next two months will be maintained at 5.75 percent. Inflation rates have increased, the State Bank maintained, adding that inflation rate in October this year remained at 4.2 percent. Compared to October 2015, inflation rates showed a 1.6 percent increase. It is expe

--- Document 1001 ---  
**KARACHI:** State Bank of Pakistan (SBP) on Saturday announced its monetary policy for the next two months and kept the policy rate unchanged at 5.75 percent. The Monetary Policy Committee of the SBP after detailed deliberations, decided to maintain the policy rate at 5.75 percent. According to SBP, the year-on-year CPI inflation rose to 3.6 percent in August 2016 from 1.8 percent in Aug

Mode: Command Ln 8, Col 12 HW-03-IR&TM IR-System.ipynb

localhost:8889/lab/tree/HW-03-IR%26TM%20IR-System.ipynb

File Edit View Run Kernel Tabs Settings Help

HW-03-IR& X Python 3 (ipykernel)

Name Modified

- Favorites 3 mo. ago
- Links 3 mo. ago
- Music 3 mo. ago
- nltk\_data yesterday
- OneDrive 12 days ago
- Pictures 3 mo. ago
- PycharmProjects last mo.
- Saved Games 3 mo. ago
- Searches 3 mo. ago
- Videos 3 mo. ago
- HW-03-IR&TM IR-System.ipynb 17 hr. ago
- HW-03-IR&TM practice 1.ipynb 23 hr. ago
- HW-03-IR&TM practice 2.ipynb yesterday
- HW-03-IR&TM practice 3.ipynb 21 hr. ago
- HW-03-IR&TM practice 4.ipynb 21 hr. ago
- HW-03-IR&TM practice 5.ipynb 21 hr. ago
- HW-03-IR&TM practice 6.ipynb 21 hr. ago
- Untitled.ipynb 2 days ago
- Untitled6.ipynb 21 hr. ago

--- Document 556 ---  
 strong>KARACHI: The Monetary Policy Committee of the State Bank of Pakistan has decided to maintain the policy rate at six percent.</strong>This was announced in the monetary policy statement issued by the SPB here on Saturday. It said that after bottoming out in September 2015 both the Year on Year (YoY) headline CPI and period average inflation are on a rising trend. The monetary policy statement i

--- Document 453 ---  
 strong>ISLAMABAD: Pakistan's annual consumer inflation rate rose to 3.32 percent in January from 3.19 percent in December, the Bureau of Statistics said on Monday.</strong>"The year-on-year inflation increased mainly due to an increase in food prices," said Asif Bajwa, the bureau's statistician. On a month-on-month basis, prices rose by 0.21 percent in January from December. The average annual

--- Document 81 ---  
 KARACHI: The announcement of Monetary Policy for the next two months is likely to be made on March 21 by the State Bank of Pakistan (SBP). The experts are expecting a further cut in the policy rate in view of improvement in the country's foreign exchange reserves, stability in rupee value and the decline of inflation rate to 3 percent. They are expecting that the central bank might ratchet down the

===== BM25 RESULTS =====

--- Document 81 ---  
 KARACHI: The announcement of Monetary Policy for the next two months is likely to be made on March 21 by the State Bank of Pakistan (SBP). The experts are expecting a further cut in the policy rate in view of improvement in the country's foreign exchange reserves, stability in rupee value and the decline of inflation rate to 3 percent. They are expecting that the central bank might ratchet down the

Mode: Command Ln 8, Col 12 HW-03-IR&TM IR-System.ipynb

localhost:8889/lab/tree/HW-03-IR%26TM%20IR-System.ipynb

File Edit View Run Kernel Tabs Settings Help

HW-03-IR& X Python 3 (ipykernel)

Name Modified

- Favorites 3 mo. ago
- Links 3 mo. ago
- Music 3 mo. ago
- nltk\_data yesterday
- OneDrive 12 days ago
- Pictures 3 mo. ago
- PycharmProjects last mo.
- Saved Games 3 mo. ago
- Searches 3 mo. ago
- Videos 3 mo. ago
- HW-03-IR&TM IR-System.ipynb 17 hr. ago
- HW-03-IR&TM practice 1.ipynb 23 hr. ago
- HW-03-IR&TM practice 2.ipynb yesterday
- HW-03-IR&TM practice 3.ipynb 21 hr. ago
- HW-03-IR&TM practice 4.ipynb 21 hr. ago
- HW-03-IR&TM practice 5.ipynb 21 hr. ago
- HW-03-IR&TM practice 6.ipynb 21 hr. ago
- Untitled.ipynb 2 days ago
- Untitled6.ipynb 21 hr. ago

--- Document 589 ---  
 strong>TOKYO: Asian stocks rose on Thursday after the U.S. Federal Reserve offered few clues on its monetary policy outlook, while the dollar edged higher as investors awaited the Bank of Japan's policy decision.</strong>MSCI's broadest index of Asia-Pacific shares outside Japan added 0.3 percent. Japan's Nikkei jumped 1.1 percent, while S&P 500 e-mini futures ESf1 edged up 0.1 percent, after W

--- Document 359 ---  
 KARACHI: State Bank of Pakistan (SBP) on Saturday announced its monetary policy and decided to keep the policy rate unchanged at 6.0 percent. A statement issued here said that the average Jul-Oct FY16 inflation was at 1.7 percent is lower than the 7.1 percent average inflation in the corresponding period of last year. It said that the decline is broad based as both food and non-food core inflat

--- Document 656 ---  
 strong>KARACHI: State Bank of Pakistan cut its key policy rate by 25 basis points to 5.75 percent on Saturday, saying the country was on course to miss its economic growth target for the year. The bank said economic growth in the fiscal year ending June 2016 was likely to exceed last year's 4.2 percent but miss the 5.5 percent target. The bank also said the inflation outlook was "low" despite

--- Document 1001 ---  
 strong>KARACHI: State Bank of Pakistan (SBP) on Saturday announced its monetary policy for the next two months and kept the policy rate unchanged at 5.75 percent.</strong>The Monetary Policy Committee of the SBP after detailed deliberations, decided to maintain the policy rate at 5.75 percent. According to SBP, the year-on-year CPI inflation rose to 3.6 percent in August 2016 from 1.8 percent in Aug

Mode: Command Ln 8, Col 12 HW-03-IR&TM IR-System.ipynb

- a [README](#) with instructions to run your system

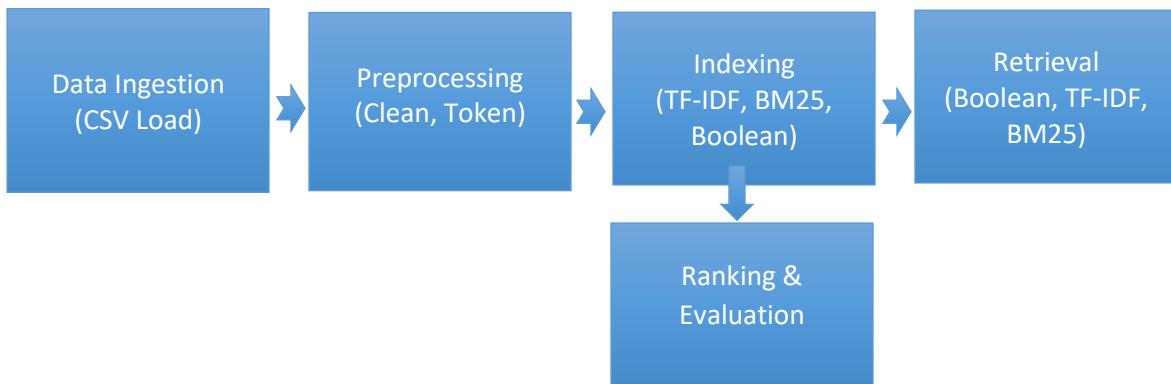
This ensures instructors can test your pipeline.

### 3. Technical Report

# Technical Report: Information Retrieval System using the Boolean, TF-IDF, BM25

## 1. System Architecture

### 1.1 System Diagram



### 1.2 Figure Caption

The system begins by ingesting raw CSV documents. Text is preprocessed through cleaning, lowercasing, and tokenization. Preprocessed documents are then indexed using Boolean, TF-IDF, and BM25 techniques. Queries are processed through these indices, and the top-ranked documents are returned. Evaluation measures system efficiency and retrieval quality.

## 2. Description of the Retrieval System

The system consists of the following modules:

### Data Preprocessing:

- ✓ **Normalization:** Converts all text to lowercase.
- ✓ **Punctuation Removal:** Non-alphanumeric characters are removed.
- ✓ **Tokenization:** Text is split into individual tokens.
- ✓ **Stop word Removal:** Common English stopwords are removed using NLTK's stop words corpus.

### Indexing Techniques:

- ✓ **Boolean Retrieval:** Matches documents containing all query tokens.
- ✓ **TF-IDF Vectorization:** Converts documents into term-frequency inverse-document-frequency vectors. Cosine similarity is used to rank documents based on their relevance to the query vector.

- ✓ **BM25 Retrieval:** Uses probabilistic retrieval to score documents based on term frequency and document length normalization.

## Modifications/Justifications

- ✓ Automatic selection of the text column from CSV to enhance usability.
- ✓ Handling multiple CSV encodings to ensure robustness across datasets.
- ✓ Use of stop words to reduce noise and improve ranking accuracy.

# 3. Evaluation

### Qualitative Evaluation:

- ✓ Checked top-5 results for sample queries.
- ✓ Verified relevance of returned documents using keyword matching and topic similarity.

### Quantitative Evaluation:

- ✓ Scoring based on cosine similarity for TF-IDF.
- ✓ BM25 scores provided probabilistic ranking.
- ✓ Retrieval speed is reasonable for datasets of ~2,500 documents.
- ✓ Memory footprint is minimized using sparse TF-IDF matrices and tokenized lists for BM25.
- ✓ Environment and Configuration Files for IR System

# 1. Requirements.txt

This file lists all Python dependencies needed to run the Information Retrieval system.

pandas  
numpy  
scikit-learn  
nltk  
rank\_bm25

# 2. Optional: .env file (for environment variables)

If you want to configure paths or other settings, you can create a .env file in the project root:

```
# Path to dataset
DATASET_PATH=./data/Articles.csv

# Number of top results to retrieve
TOP_K=5
```

# 3. Python Version

- Python 3.9 or above is recommended.

## 4. Optional: Jupyter Notebook Kernel Configuration

- If running in Jupyter, ensure the correct Python environment is selected.
- Install ipykernel in your environment:
- pip install ipykernel  
python -m ipykernel install --user --name=ir\_env

## Notes

- The system will automatically handle stopwords and CSV encoding.
  - Make sure the dataset CSV is placed in the path specified in either .env or directly in src/main.py.
  - The requirements.txt file is sufficient to replicate the environment in a new machine.
- ✓

## 4. Discussion

### Findings:

- ✓ BM25 generally produces more contextually relevant results compared to Boolean and TF-IDF.
- ✓ Boolean retrieval is fast but may miss partially relevant documents.
- ✓ TF-IDF retrieval balances speed and relevance but is sensitive to term weighting.

### Shortcomings:

- ✓ No synonym or semantic query expansion.
- ✓ No relevance feedback loop.
- ✓ Evaluation limited to small datasets.

### Future Improvements:

- ✓ Integrate Word Embedding's or Transformer-based retrieval.
- ✓ Add relevance feedback and query expansion.
- ✓ Implement caching and optimization for larger datasets.

## 5. References

1. Manning, C., Raghavan, P., & Schütze, H. (2008). Introduction to Information Retrieval. Cambridge University Press.
2. Scikit-learn Documentation: <https://scikit-learn.org/stable/>
3. Rank-BM25 Python Package: <https://pypi.org/project/rank-bm25/>
4. NLTK Documentation: <https://www.nltk.org/>

A clean, well-commented GitHub repository

## 6. Disclosure of AI Use

### 6.1 Summary of AI Usage

The screenshot shows the ChatGPT interface with a dark theme. On the left, there's a sidebar with various AI-generated responses. The main area has a message from ChatGPT:

[nltk\_data] Downloading package stopwords to C:\Users\SYED ZAIN UL  
[nltk\_data] AIDEEN\nltk\_data...  
[nltk\_data] Package stopwords is already up-to-date!  
Total documents: 10

Below this, it says: "Ah! The message you're seeing is not actually an error. It's just an informational log from NLTK saying:" followed by two examples of code snippets:

vbnnet

```
[nltk_data] Downloading package stopwords to C:\Users\SYED ZAIN UL AIDEEN\nltk_data...
[nltk_data] Package stopwords is already up-to-date!
```

This happens when you run:

python

```
nltk.download('stopwords')
```

Since your code then prints:

+ Ask anything

ChatGPT can make mistakes. Check important info.

The screenshot shows the ChatGPT interface with a dark theme. On the left, there's a sidebar with various AI-generated responses. The main area has a message from ChatGPT:

error revived

-----

FileNotFoundException Traceback (most recent call last)  
Cell In[1], line 35  
28 # ======  
29 # Load Dataset (from Kaggle)  
30 # ======  
31  
32 # Replace this path with the actual path on your machine /  
notebook environment  
33 csv\_path = 'path/to/news\_articles.csv'  
---> 35 df = pd.read\_csv(csv\_path)  
36 print(f'Total documents: {len(df)}')  
37 print(df.columns) # print column names so you can check text  
column name

File C:\ProgramData\anaconda3\Lib\site-  
packages\pandas\io\parsers\readers.py:1026, in  
read\_csv(filepath, ↓ ufffer, sep, delimiter, header, names, index\_col,  
usecols, dtype, engine, converters, true\_values, false\_values,

+ Ask anything

ChatGPT can make mistakes. Check important info.

Syed zain ul a... Free Upgrade

ChatGPT

LookupError Traceback (most recent call last)  
File C:\ProgramData\anaconda3\Lib\site-packages\nltk\corpus\util.py:84, in LazyCorpusLoader.\_load(self)  
83 try:  
--> 84 root = nltk.data.find(f'{self.subdir}/{zip\_name}')  
85 except LookupError:  
  
File C:\ProgramData\anaconda3\Lib\site-packages\nltk\data.py:579, in find(resource\_name, paths)  
578 resource\_not\_found = f'\n{sep}\n{msg}\n{sep}\n'  
--> 579 raise LookupError(resource\_not\_found)  
  
LookupError:  
\*\*\*\*\*  
Resource stopwords not found.  
Please use the NLTK Downloader to obtain the resource:  
>>> import nltk  
>>> nltk.download('stopwords')

Ask anything

ChatGPT can make mistakes. Check important info.

Syed zain ul a... Free Upgrade

ChatGPT

macOS (Homebrew recommended)

Install Homebrew (if not installed):  
`/bin/bash -c "$(curl -fsSL https://raw.githubusercontent.com/Homebrew/install/HEAD/install.sh)"`

Here is a clear step-by-step guide to install Homebrew on macOS and verify it is working.

### How to Install Homebrew on macOS (Step-by-Step)

**Step 1 — Open Terminal**

- Press Command + Space
- Type Terminal
- Press Enter

Ask anything

ChatGPT can make mistakes. Check important info.

