Software Tools and Libraries Used

This project utilized several open-source libraries, frameworks, and services for natural language processing, plagiarism detection, and web application development. The following table summarizes the key tools and their purposes:

|  |  |  |
| --- | --- | --- |
| Tool / Library | Purpose | Link |
| Python 3.x | Programming language for backend, text processing, and AI. | https://www.python.org/ |
| Flask | Lightweight web framework for building REST APIs and web apps. | https://flask.palletsprojects.com/ |
| NLTK | Natural Language Toolkit for preprocessing text data. | https://www.nltk.org/ |
| sentence-transformers | Pre-trained models for semantic similarity and sentence embeddings. | https://www.sbert.net/ |
| PyPDF2 | Library to extract text from PDF files. | https://pypi.org/project/PyPDF2/ |
| python-docx | Library to extract text from DOCX files. | https://pypi.org/project/python-docx/ |
| Requests | Python library to perform HTTP requests for Google Search API. | https://docs.python-requests.org/ |
| scikit-learn | Machine learning library for similarity metrics like cosine similarity. | https://scikit-learn.org/ |
| textdistance | Library to calculate various text similarity metrics like Levenshtein distance. | https://pypi.org/project/textdistance/ |
| spaCy | Industrial-strength NLP toolkit used for dependency parsing (subject/object extraction). | https://spacy.io/ |
| Postman | API development and testing platform for testing backend routes. | https://www.postman.com/ |
| Visual Studio Code | Source code editor for Python, Flask, and HTML/CSS. | https://code.visualstudio.com/ |