Offer Search App

Dependencies

- Anaconda or Miniconda
- Python = 3.10
- PyTorch = 2.1.0
- More in requirements.txt

> Ideal setup requires a Graphic Card

🛠 Installation

1. Clone repo

```
git clone https://github.com/95anantsingh/search-app.git
```

2. Create conda environment

```
conda env create -f env.yml
```

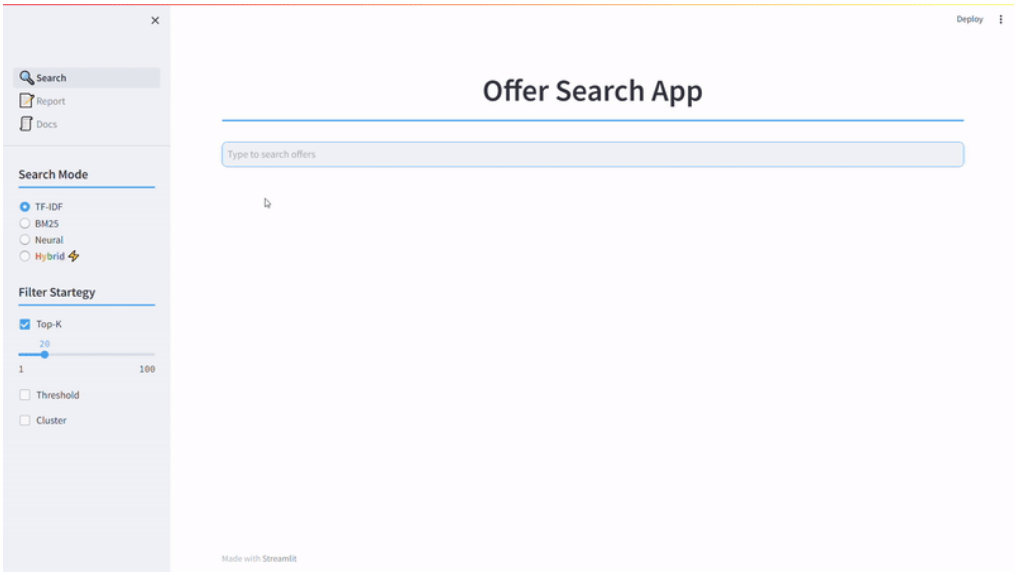
3. Download NLTK Data

```
python -m nltk.downloader punkt stopwords
```

⚡ Run App

```
cd search-app
conda activate search
streamlit run 🔍_Search.py
```

Now search for offers with options on the left to see results.



📘 Report

Project report can be found here or you can find it in the app as well.

📂 Project Structure

Folder Name	Description
.streamlit	Configuration files for Streamlit
.vscode	Visual Studio Code settings and files
core	Core application module
├── base_search.py	Base search class
├── bm25.py	BM25 search class
├── data_processor.py	Data processing code
├── hybrid.py	Hybrid search class
├── **init**.py	Initialization module
├── neural.py	Neural search class
├── offers_db.py	Offers database class
├── tfidf.py	TF-IDF search class
data	Data used by the application
├── processed	Processed data files
│ ├── database.sqlite	Offers SQLite database
│ ├── syn_queries.json	Synthetic queries
│ ├── true_scores.csv	True scores (CSV)
│ ├── true_scores_gold.csv	True scores (gold) (CSV)
│ └── true_scores_syn.csv	True scores (synthetic) (CSV)
└── raw	Raw data files
notebooks	Jupyter Notebook files
├── eval.ipynb	Evaluation notebook
├── queries.ipynb	Quer Generation notebook
└── search_exp.ipynb	Search experiment Notebook
vectors	Vector Database files
├── bm25	BM25 model files
├── neural	Neural model files
│ └── retrieval	FAISS Vector Database Files
└── tfidf	TF-IDF files
pages	Application web pages

Folder Name	Description
🔍_Search.py	Streamlit App File
env.yml	Environment configuration file
README.md	Repository README file
requirements.txt	Python package requirements

📧 Contact

If you have any question, please email `anant.singh@nyu.edu`