Project Proposal: ConnectGlobal

An AI-Powered Career Support for International Students in the United States of America

Problem Statement

International students face many barriers in the U.S. job search with sponsorship and OPT opportunities being uncertain and not always clear, as well as a lack of connections resulting in a weak network and less opportunities. Cold applications rarely work for anyone, even U.S. citizens, so adding these extra issues on top of that can dishearten many students, leading them to move back home after university.

Context & Motivation

Context

There are 1 million international students in the U.S. with 60%+ stating the job search to be more stressful than their coursework. [1]

Recruiters look at a resume for roughly 10 seconds or use AI filters to scan for keywords instantaneously. [2]

Motivation

I, myself, have had a hard time:

1. Finding jobs that fit my skills and experience whilst also offering OPT and H1B Visa sponsorship.

2. Networking with fellow British people working in the U.S.

Project Feasibility:

For job postings: Kaggle job posting datasets[5], LinkedIn API's

For networking: Graph ML

For extracting skills and resume match: NLP tools

Project Impact:

Societal Impact: eases student stress, promotes diversity, and builds global network

Economic Impact: saves time (20-30%), improves match quality, and boosts interviews (+10-15% callback rate)(4-6x higher hiring probability)[2]

Ethical Impact: privacy risks, bias [3]

Ethical Impact Mitigation: diverse datasets, anonymization

AI/ML Approach

Techniques

NLP (spaCy, BERT):
Extracting keywords, matching resume and jobs[4]

Supervised ML:
Rank resume-job fit

GraphML (Node2Vec):
Suggest mentors & connections

Tools

Python

Scikit-learn

Hugging Face

NetworkX

Workflow

Data Preprocessing

Job Postings
- Kaggle[5]
- LinkedIn
- Glassdoor

Resumes Samples
- Synthetic or
- Anonymized

Synthetic Networking Data
- Country
- Skills
- Industry

Resume-job similarity model

TF-IDF
- Baseline semantic similarity

BERT Embeddings[4]
- Deep semantic similarity

Model
- Supervised classification
- Outputting top-k matched jobs

Networking Recommender

Graph Construction
- Nodes = people
- Edges = shared factors

Graph Clustering
- Identifying communities of similar professionals

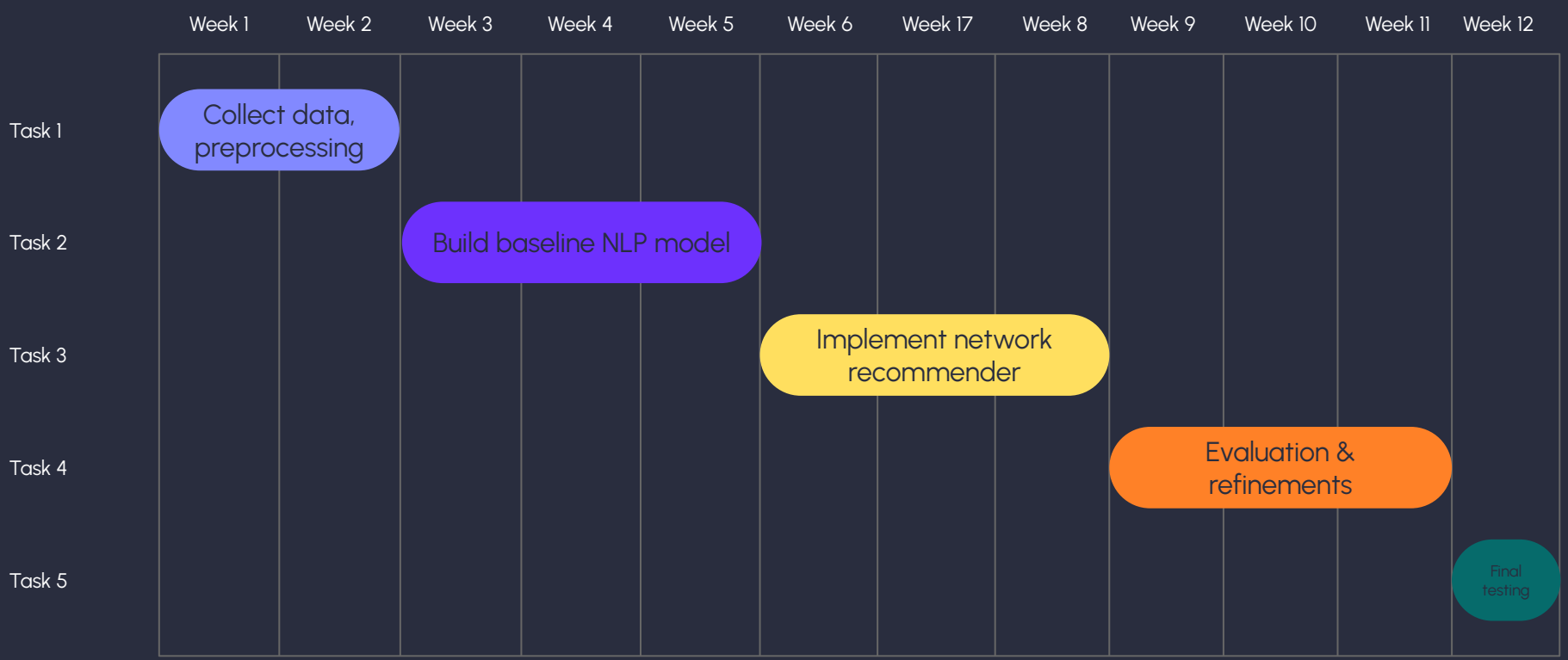
Link Prediction
- Suggest likely helpful connections

Evaluation

Job Matching
- Precision/recall
- Compare TF-IDF to BERT embeddings

Usefulness for networking
- Measured by relevance of suggested connections

Semester Timeline



Innovation vs Existing Solutions

	ConnectGlobal	LinkedIn	Indeed
Country Specific Networking	✅	❌	❌
Visa-friendly job filtering	✅	⚠️ (partial)	❌
Free AI networking assistant	✅	⚠️ (generic AI tools only)	⚠️ (generic AI tools only)
Referral likelihood insights	✅	❌	❌

Conclusion

This project has a lot of new techniques that I have not used before so I am excited to take on the challenge. The aim is to build a ML model that connects people with the right people to help further their career; and an ML model that directs people to realistic job opportunities.

I would like to add that if everything is ahead of schedule then a basic frontend website will be built to display the content in a more user friendly display.

References

- Institute of International Education (IIE). Open Doors Report 2024.[1]
- LinkedIn Economic Graph (2023). Global Talent Trends Report.[2]
- Devlin, J., et al. (2019). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. NAACL.[3]
- Angwin, J., et al. (2016). Machine Bias: Risk Assessments in Criminal Sentencing. ProPublica.[4]
- Job Postings Dataset on Kaggle (Indeed/Glassdoor data).[5]

Any questions? Ask away!

Thank you