

SOFTWARE DEVELOPMENT FOR ARTIFICIAL INTELLIGENCE

FALL 2022 – Department of Computer Science, UNT

PROJECT NAME: Resume Score Calculator using NLP

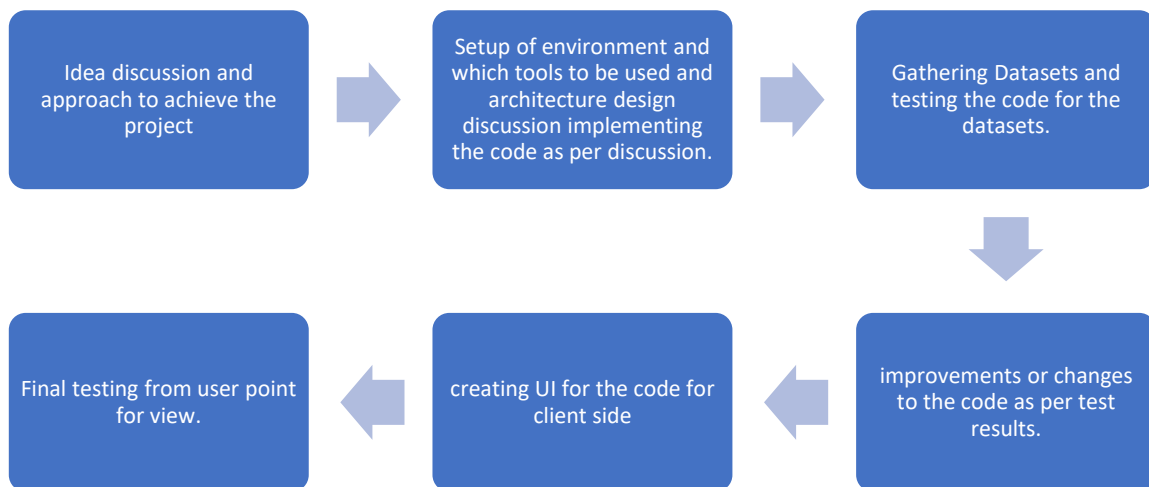
PARTICIPANTS:

Team Size - 4

- 1) Yash Zauwar - 11439929
- 2) Goutham Pallapothu - 11519845
- 3) Veeranjanyulu Muppalla - 11532514
- 4) Yashwanth Gajula – 11524980

Workflow:

- Idea discussion and approach to achieve the project.
- Setup of environment and which tools to be used and architecture design discussion implementing the code as per discussion.
- Each team member will contribute towards some part of the code, gather the dataset, testing, documenting the project progress, creating UI, setting it up as a web project.
- Connecting over meeting and work towards the goal of the project.



Weekly meetings: Every Wednesday 2:00 PM to 4:00 PM

Location: Zoom Meetings or in Person at Library.

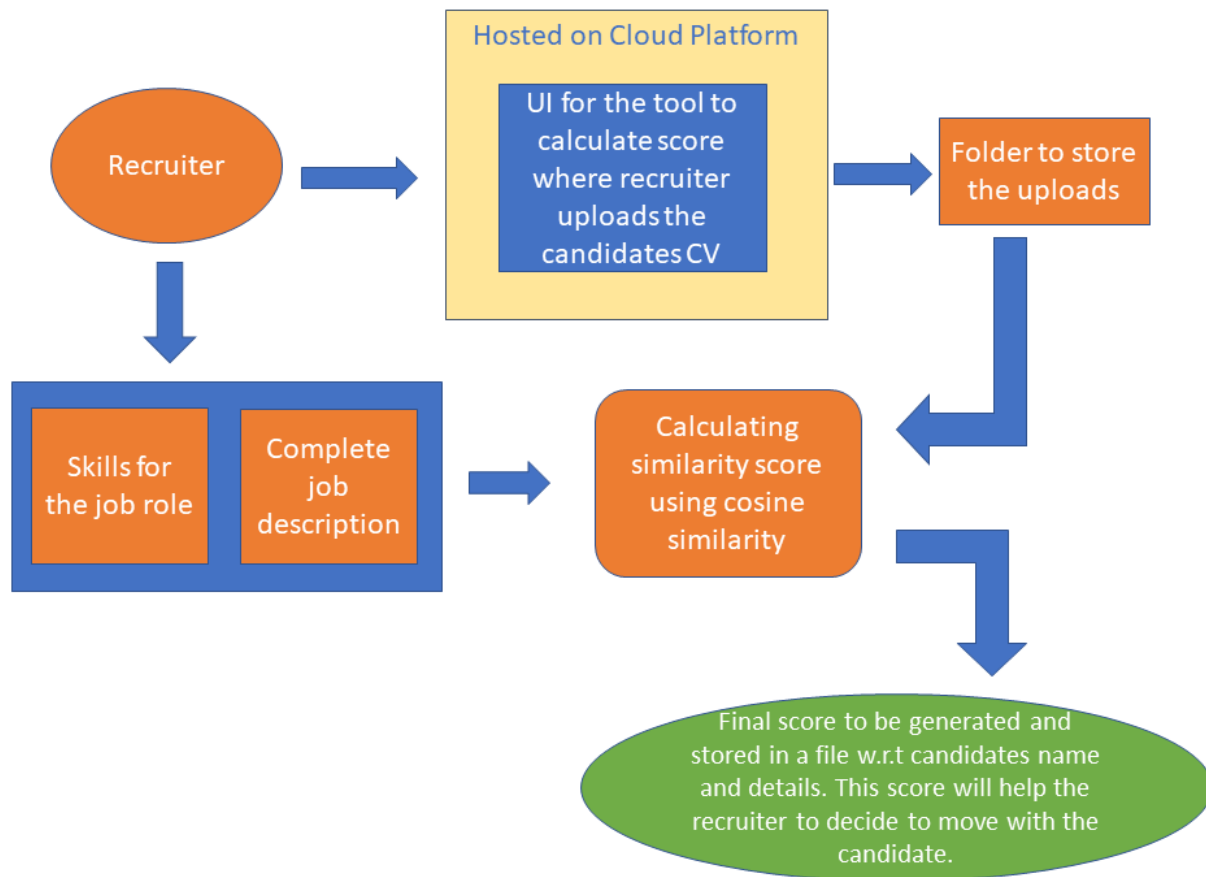
Link to Project: <https://github.com/YashZauwar/SoftwareDevelopmentforAI>

Project Abstract: Resume screening is the process of determining if a candidate meets the qualifications for a position by comparing the information on their resumes—such as their education, skills, certifications, experience, and achievements—with the needs of the position. Particularly in situations where there are a lot of applications, resume screening is essential to determining whether an applicant advances to the next step of the recruiting process or not.

To find the best candidate for a job opportunity based on thresholds, specific criteria, or scores, this repository employs text mining and natural language processing algorithms for evaluating thousands of resumes objectively and without bias in a matter of minutes.

- Cloud Platform: Microsoft Azure
- Dataset: Resumes from various sources on the internet.
- Deployment: Heroku or other python web app deployer.
- Stretch Goal: To deploy a web app which will show the score of a candidate for the recruiter and make it easy to proceed with hiring steps.

Project Design and Milestones:



Project Milestones:

- 1) Gathering Datasets (Resumes, skills, job descriptions), project environment setup,
- 2) Developing code, applying cosine similarity methods in python,
- 3) Developing UI,
- 4) Deploying the application to cloud,
- 5) Testing,
- 6) Release.

Programming Languages used: Python

Modules: python modules like texthero, numpy, cleantext, pandas, pdf2image etc.,

UI: User interface will be developed using HTML, CSS, php, etc.,

Code will be deployed in: Azure Cloud, Flask programming.

Resources and Related Projects:

1: <https://blog.apilayer.com/build-your-own-resume-parser-using-python-and-nlp/>

Resumes are a great example of unstructured data. Since there is no widely accepted resume layout; each resume has its own style of formatting, different text blocks, or even category titles do vary a lot. I don't even need to mention how big a challenge it is to parse multilingual resumes.

2: <https://www.geeksforgeeks.org/how-to-calculate-cosine-similarity-in-python/>

We calculate the **Cosine Similarity** between the two non-zero vectors. A vector is a single-dimensional signal **NumPy array**. Cosine similarity is a measure of similarity, often used to measure document similarity in text analysis. We use the below formula to compute the cosine similarity.

We will be using some techniques from the above URLs and few internet resources as we progress in the project which will be referenced in respective project increments.

With the help of cosine similarity, similarity score will be calculated for the skill set of the candidate and from the sheet and job description provided by the recruiter.
