VIT | Vidyalankar Institute of Technology
ACCREDITED A+ BY NAAC

RESUME CLASSIFIER

Project Guide - Prof. Suja Jayachandran

1. Bhumi Avhad – 20102B2002
2. Sapana Survase – 20102B2005
3. Pratik Haldankar – 20102A2006
4. Dipesh Bedmutha – 20102A2009

Problem Statement

Design a system using Tensorflow that

Classifies a resume based on its potential fit for different job categories

Our Motivation

- Create a faster system for classifying resumes
- Reduce human effort for recruiters
- Help applicants improve their resumes before applying to jobs
- Improve current resumes classification systems, which have limitations.
  - Can only be used by companies, not available to applicants to get feedback
  - Many don’t use machine learning, which makes making changes, like adding more jobs categories and retraining the classifier,difficult to do

Overview of Solution

- Classify resumes into categories - java Developer, Business Analyst, Data Scientist, etc.
- Used a dataset consisting of many categories
- Solution included a Machine Learning model for classification
  - Model takes the text in the resume and classifies it into one of the categories
  - Depicts confidence in the category choice (as a percentage)
- Display top five categories and confidence percentages
- Deploy trained model on https server
- Test on teammate’s resume

LITERATURE SURVEY

- Senthil Kumaran et al. [1] used an intelligent tool for ontology called EXPERT mapping-based candidate screening to create an automated system for intelligent screening of prospects for recruitment, enhancing the precision with which candidates are matched to the requirements of the job.
-

LITERATURE SURVEY

- Jagan Mohan Reddy D et al. [2] suggested joining efficient candidates before resume selection, so that the entire process can be completed in a timely and cost-effective manner. Some characteristics, such as age and salary hike, cannot be used directly for classification due to substantial variations in values that must be transformed into bins.

Methodology

- Preprocessing

The resume’s provided as input would be shortlisted in this procedure to remove any special or garbage characters from the resumes.

NLTK tokenizers are used to break the dataset into tokens.

Stop word removal, lemmatization and vectorization are among the preprocessing operations performed on the tokenized dataset.

- Removing Stop Words:

  Stop words such as and, the, was, and others appear very often in words and limit the process which determines prediction, thus they are removed.

- Lemmatization:

  Lemmatization reduces derived phrases to make entirely sure that the underlying word is accurately associated with the language.

  The routine phases of lemmatization are as follows: • Convert the text corpus into a list of words. • Make a corpus concordance, which includes all of the word list entries as they appear in the corpus. • Based on the concordance, link the word-forms to their lemmas.

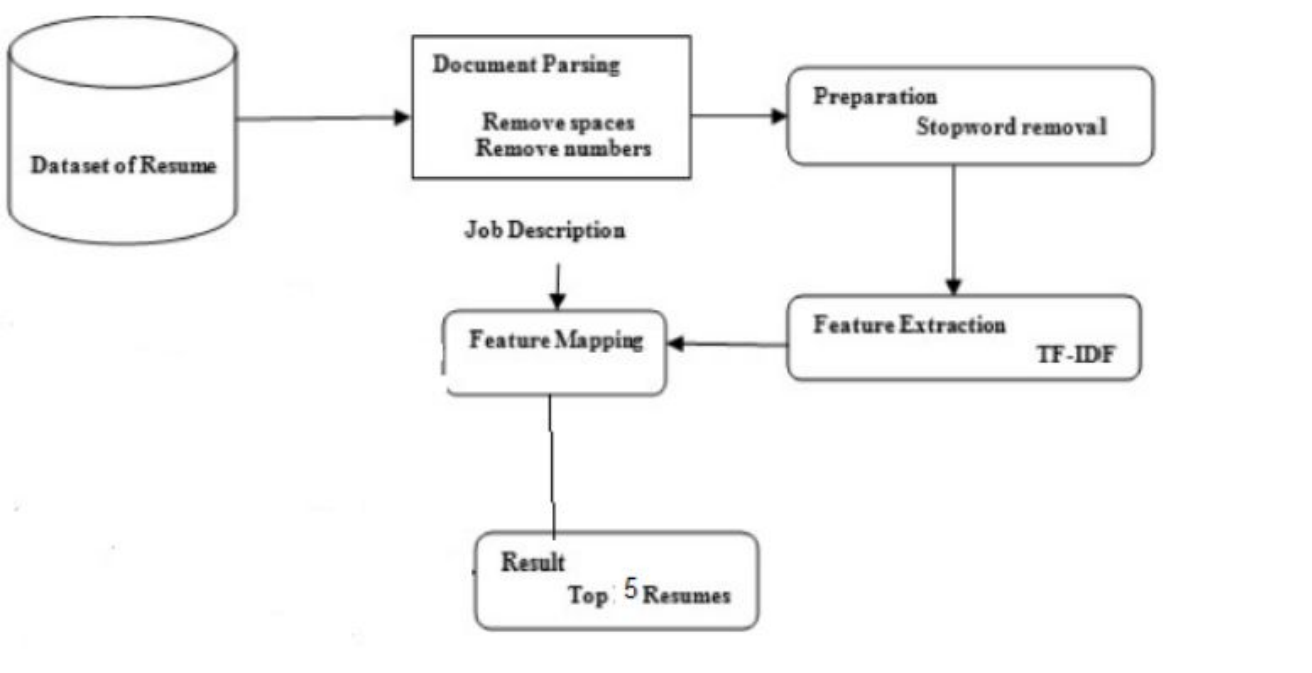
- The k-NN model :

  Is used in this model to find the resumes that are closest to the specified job description.

Requirements

- Language implemented : python
- Hardware : Windows 8/10/11. 32/64 bit
- Libraries Imported: nltk, numpy,pandas,re,

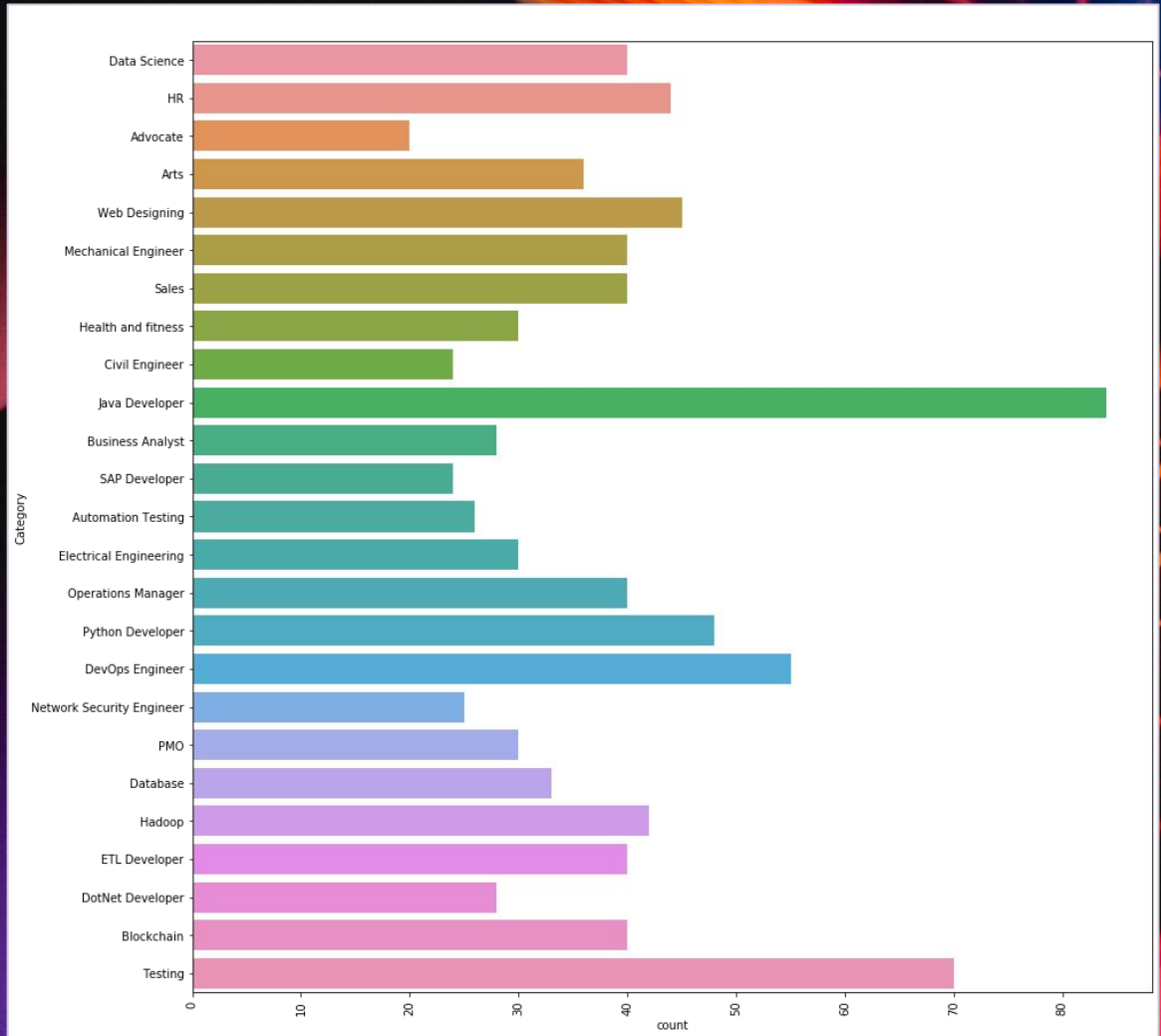
Proposed System



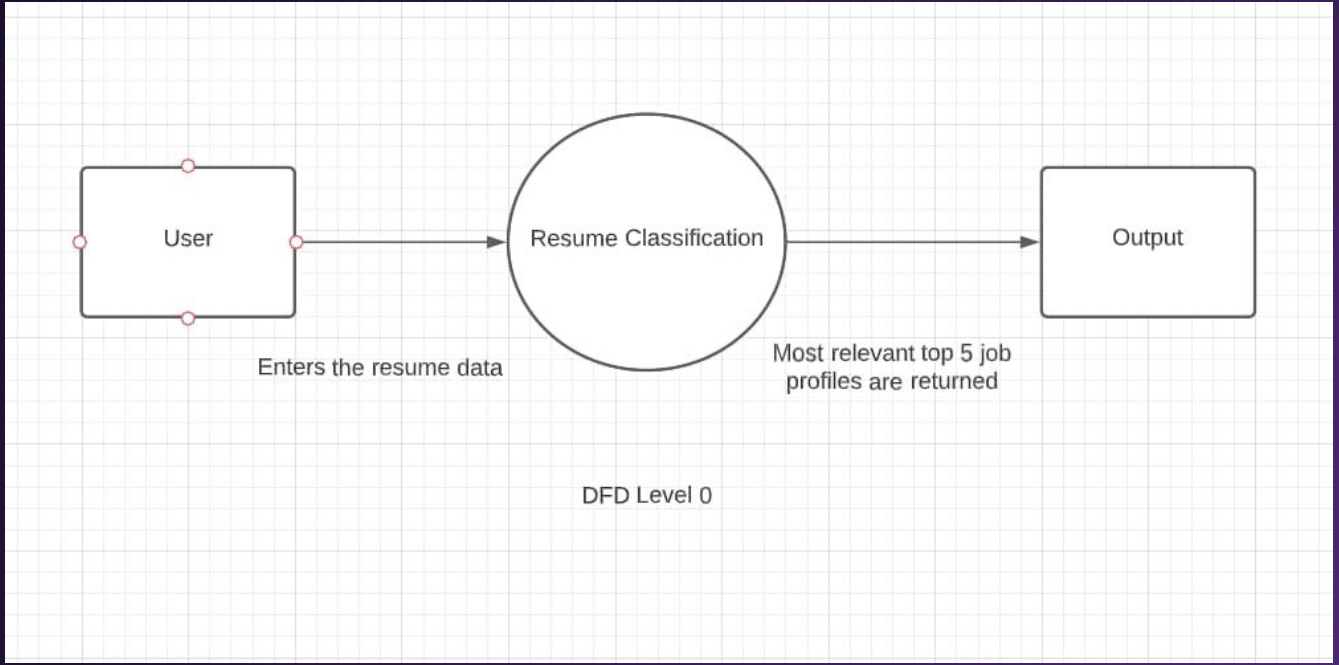
Categories in Resume Dataset

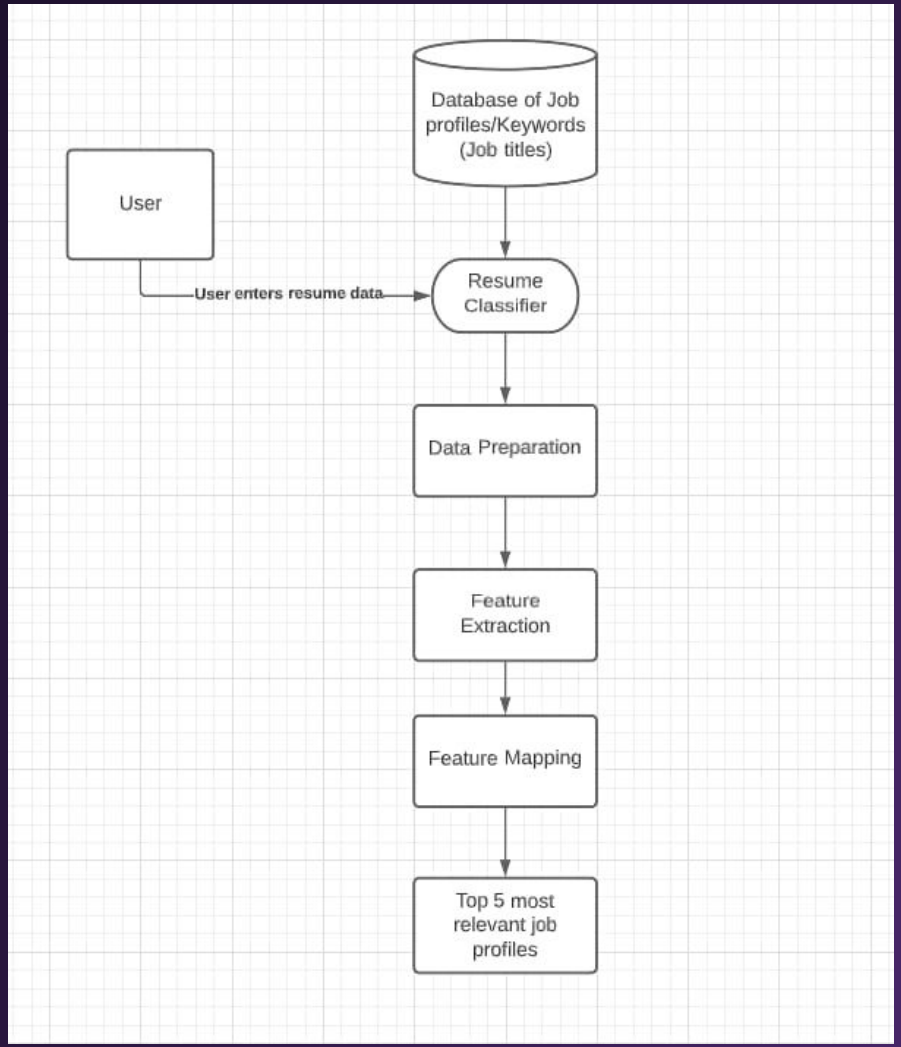
Category	# in Dataset
Java Developer	84
Testing	70
DevOps Engineer	55
Python Developer	48
Web Designing	45
HR	44
Hadoop	42
Sales	40
Blockchain	40
Mechanical Engineer	40
ETL Developer	40
Operations Manager	40
Data Science	40
Arts	36
Database	33
PMO	30
Electrical Engineering	30
Health and fitness	30
DotNet Developer	28
Business Analyst	28
Automation Testing	26
Network Security Engineer	25
Civil Engineer	24
SAP Developer	24
Advocate	20

Analyzing Dataset

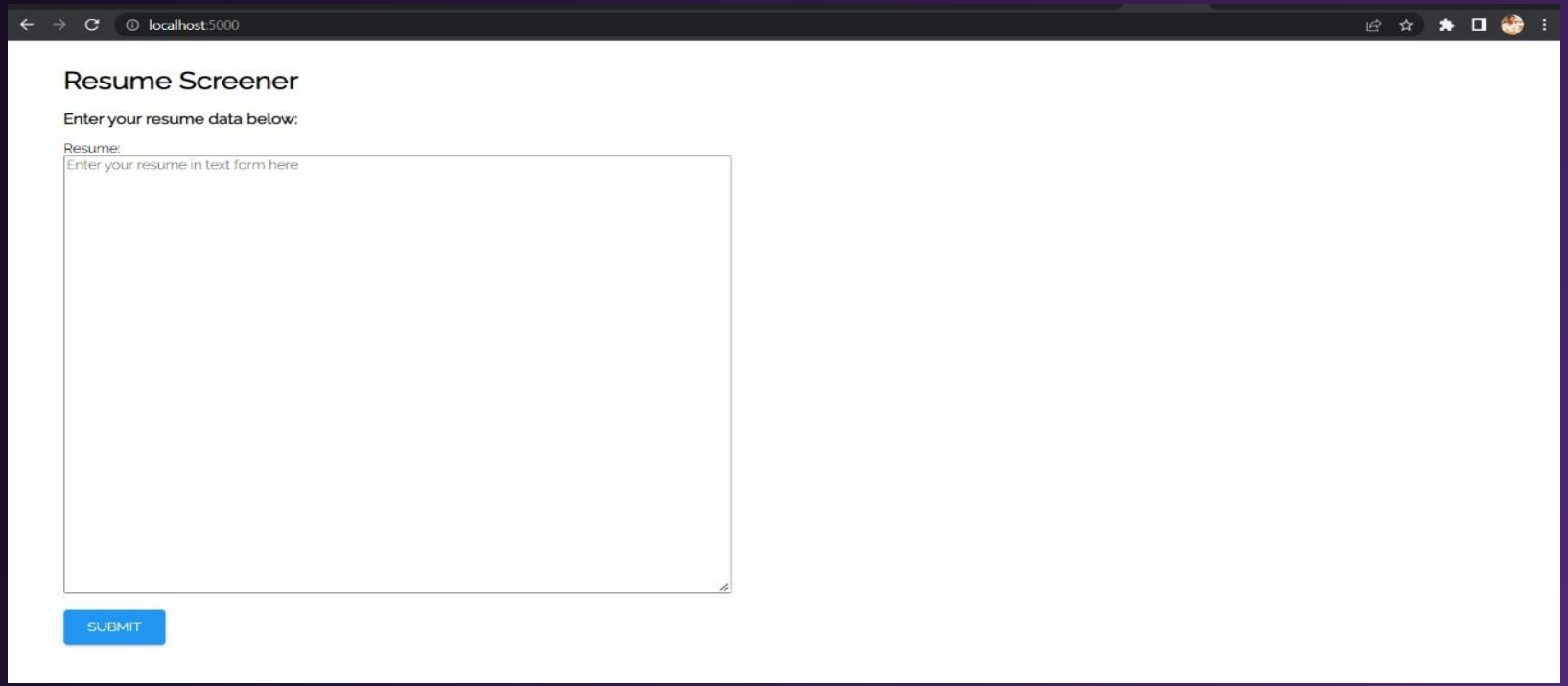


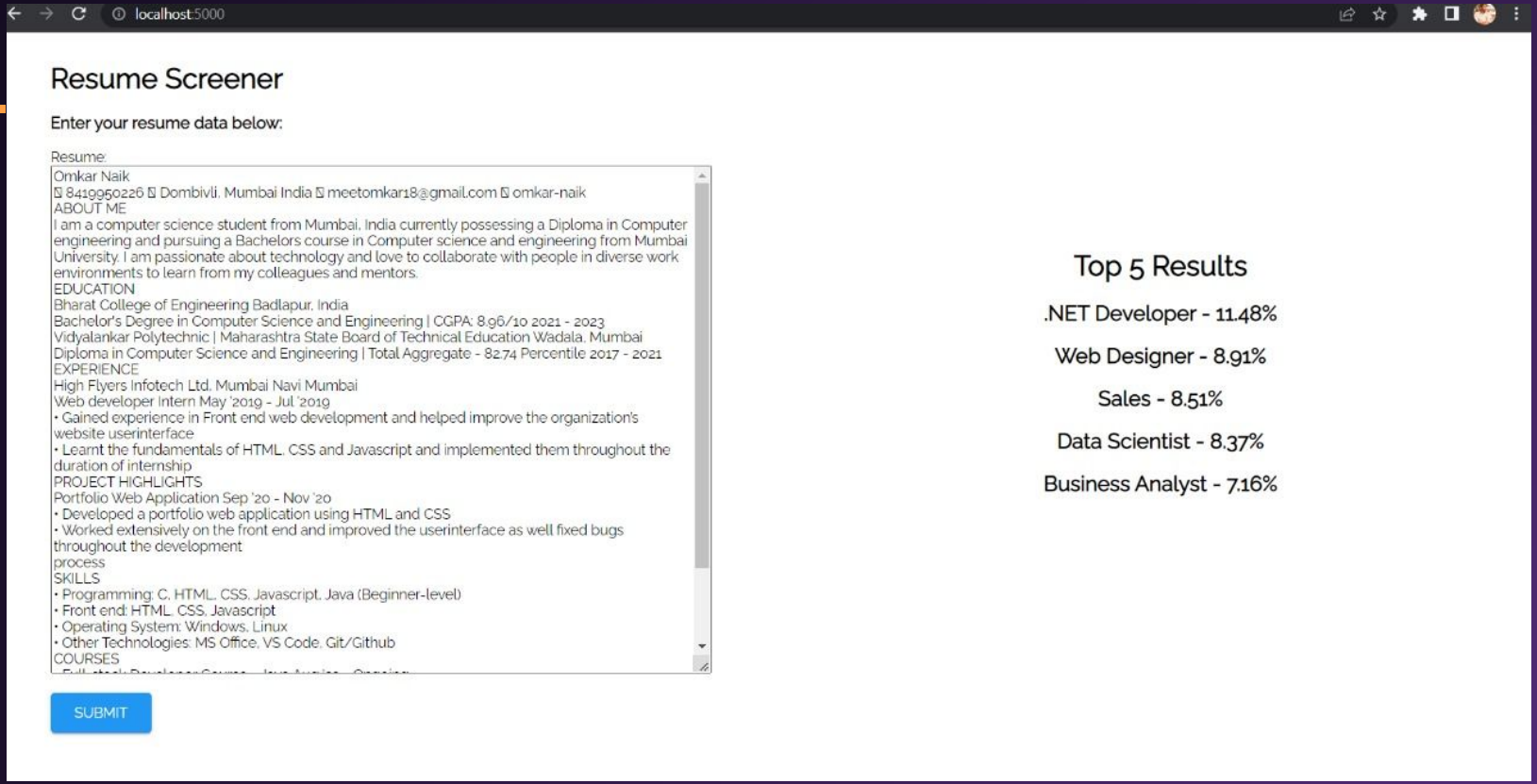
DFD


User
Resume Classification
Output
Enters the resume data
Most relevant top 5 job profiles are returned
DFD Level 0


Database of Job profiles/Keywords (Job titles)
User
User enters resume data
Resume Classifier
Data Preparation
Feature Extraction
Feature Mapping
Top 5 most relevant job profiles

Result


localhost:5000
Resume Screener
Enter your resume data below:
Resume:
Enter your resume in text form here
SUBMIT


localhost:5000
Resume Screener
Enter your resume data below:
Resume:
Omkar Naik
8419950226 Dombivli, Mumbai India meetomkar18@gmail.com omkar-naik
ABOUT ME
I am a computer science student from Mumbai, India currently possessing a Diploma in Computer engineering and pursuing a Bachelors course in Computer science and engineering from Mumbai University. I am passionate about technology and love to collaborate with people in diverse work environments to learn from my colleagues and mentors.
EDUCATION
Bharat College of Engineering Badlapur, India
Bachelor's Degree in Computer Science and Engineering | CGPA: 8.96/10 2021 - 2023
Vidyalankar Polytechnic | Maharashtra State Board of Technical Education Wadala, Mumbai
Diploma in Computer Science and Engineering | Total Aggregate - 82.74 Percentile 2017 - 2021
EXPERIENCE
High Flyers Infotech Ltd. Mumbai Navi Mumbai
Web developer Intern May '2019 - Jul '2019
• Gained experience in Front end web development and helped improve the organization's website userinterface
• Learnt the fundamentals of HTML, CSS and Javascript and implemented them throughout the duration of internship
PROJECT HIGHLIGHTS
Portfolio Web Application Sep '20 - Nov '20
• Developed a portfolio web application using HTML and CSS
• Worked extensively on the front end and improved the userinterface as well fixed bugs throughout the development
process
SKILLS
• Programming: C, HTML, CSS, Javascript, Java (Beginner-level)
• Front end: HTML, CSS, Javascript
• Operating System: Windows, Linux
• Other Technologies: MS Office, VS Code, Git/Github
COURSES
SUBMIT
Top 5 Results
.NET Developer - 11.48%
Web Designer - 8.91%
Sales - 8.51%
Data Scientist - 8.37%
Business Analyst - 7.16%

Future Scope

- Investigate more complex neural network architectures.
- Expand dataset with more resume categories.
- Develop models that can give feedbacks to users to improve their resume.
- Develop model that analyzes resume in detail
- Upscale the existing project to show the job profiles on the company career portals which match the specific keywords in the candidate's resume. List the most relevant profiles first, so that the chances of a candidate getting selected is increased significantly.
- Suggest resumes that are similar to the submitted resume

References

- Sankar, A. (2013). “Towards an automated system for intelligent screening of candidates for recruitment using ontology mapping (EXPERT)”. International Journal of Metadata, Semantics and Ontologies, 8(1), 56.
- https://doi.org/10.1504/ijmso.2013.054184 [2] Jagan Mohan Reddy D, Sirisha Regella., “Recruitment Prediction using Machine Learning”, IEEE Xplore, 2020.
- https://www.ijert.org/research/resume-classification-and-ranking-using-knn-and-cosine-similarity-IJERTV10IS080057.pdf[3]International Journal of Engineering Research & Technology (IJERT)

Thank you !!