Resume Analyzer Using NLP

Major Project Report

Submitted in partial fulfilment of the requirements for the degree of

Bachelor of Engineering (Computer Engineering)

by:

AMRUT MANE	**TU3F2021134**
PRANIT MISAL	**TU3F2021119**
SIDDHESH NAVALE	**TU3F2021151**
ATHARVA UMAP	**TU3F2021135**

Under the Guidance of

Prof. Ramesh Shahabade.

Department of Computer Engineering

TERNA ENGINEERING COLLEGE

Nerul (W), Navi Mumbai 400706

(University of Mumbai)

(2023-2024)

TERNA ENGINEERING COLLEGE, NERUL,

NAVI MUMBAI

Department of Computer Engineering

Academic Year 2023-24

CERTIFICATE

This is to certify that the major project entitled **"Resume Analyzer Using NLP"** is a bonafide work of

AMRUT MANE	**TU3F2021134**
PRANIT MISAL	**TU3F2021119**
SIDDHESH NAVALE	**TU3F2021151**
ATHARVA UMAP	**TU3F2021135**

Submitted to the University of Mumbai in partial fulfilment of the requirement for the award of Bachelor of Engineering(Computer Engineering)

Guide	Head of Department	Principal
Prof. Ramesh Shahabade	**Dr. Kishor Sakure**	**Dr. L.K. Ragha**

Approval Sheet

Project Report Approval

This Major Project Report – an entitled "**Resume Analyzer Using NLP**" by following students is approved for the degree of ***B.E. in "Computer Engineering"***.

Submitted by:

AMRUT MANE	**TU3F2021134**
PRANIT MISAL	**TU3F2021119**
SIDDHESH NAVALE	**TU3F2021151**
ATHARVA UMAP	**TU3F2021135**

Examiners Name & Signature:

1.----------

2.----------

Date: ----------

Place: ----------

Declaration

We declare that this written submission represents our ideas in our own words and where others' ideas or words have been included, we have adequately cited and referenced the original sources. We also declare that we have adhered to all principles of academic honesty and integrity and have not misrepresented or fabricated or falsified any idea/data/fact/source in our submission. We understand that any violation of the above will be cause for disciplinary action by the Institute and can also evoke penal action from the sources which have thus not been properly cited or from whom proper permission has not been taken when needed.

AMRUT MANE	TU3F2021134	

PRANIT MISAL	TU3F2021119	

SIDDHESH NAVALE	TU3F2021151	

ATHARVA UMAP	TU3F2021135	

Date: _____

Place: _____

Acknowledgement

We would like to express our sincere gratitude towards our guide **Prof. Ramesh Shahabade**, Project Coordinators **Prof. Ramesh Shahabade** for their help, guidance and encouragement, they provided during the project development. This work would have not been possible without their valuable time, patience and motivation. We thank them for making my stint thoroughly pleasant and enriching. It was great learning and an honor being their student.

We are deeply thankful to **Dr. Kishor Sakure (H.O.D Computer Department)** and entire team in the Computer Department. They supported us with scientific guidance, advice and encouragement, they were always helpful and enthusiastic and this inspired us in our work.

We take the privilege to express our sincere thanks to **Dr. L. K. Ragha** our Principal for providing the encouragement and much support throughout our work.

AMRUT MANE	**TU3F2021134**
PRANIT MISAL	**TU3F2021119**
SIDDHESH NAVALE	**TU3F2021151**
ATHARVA UMAP	**TU3F2021135**

Date: _____

Place: _____

ABSTRACT

Many a times employees submit their resume to a company and they get rejected. Employees often keep thinking about why they got rejected while, many a times, the resume not being up to the point is the reason for the rejection. Often employees don't know why their resume got rejected, where and how to improve etc. Our resume analyzer can be used by companies as well as employees. The current resume analyzers which are present in the market are mostly company/HR oriented. We aim to make a resume analyzer which is more focused on employees. Our resume analyzer analyzes the resumes and gives the employee a thorough feedback of the resume. It makes the employee understand the positives as well as the negatives of the resume. The resume analyzer also provides a range of tips and advices to the employee regarding the resume so that they can improve their resume.

Table of Contents

List of Figures

List of Tables

Chapter 1: INTRODUCTION

1.1) Introduction:-

A resume is a critical tool in the job search process as it's often the first impression a potential employer has of a candidate. It's a concise, formal document that outlines your skills, experiences, and accomplishments, helping employers determine if you're a good fit for the job.

The importance of a resume lies in its ability to showcase your qualifications relevant to the job you're applying for. It's a chance to highlight your unique strengths, demonstrating why you're a better candidate than others. A well-crafted resume can help you stand out among numerous applicants, increasing your chances of getting an interview.

Our Resume Analyzer is a Webapp which helps candidates in improving their resumes by giving them personalized tips, and recommendations.

1.2 Organization of the project:

Chapter 1 gives a brief overview about the aim for developing this project. The introduction and the scope of the project tells us about the expected outcome of the project for the application.

Chapter 2 of the report includes the literature survey of the existing system which discusses its limitations, identifies the problem statement, and establishes the objectives of the project.

Chapter 3 includes all the activities, which help the transformation of requirement specification into implementation, implementation diagrams.

Chapter 4 describes the Methodology and procedure of implementing the project with the help of requirement specifications.

Chapter 5 describes the functionality of the system by providing the output of the system.

Chapter 6 is the conclusion. This chapter gives a summary of the entire project. It also gives the future scope for research and development in this project.

Chapter 2:-

LITERATURE SURVEY

2.1) Existing System Survey:-

Table 2.1 Literature Survey

Sr. no.	Author Name	Publish Year	Research Paper	Summary
1.	Pratibha S, Vibha Pratap, et.al.	2022	Resume Classifier and Summarizer	• Advanced machine learning technologies such as logistic regression and Support Vector machine learning are used. • This project predicts the category of the resume and provides the summarized version of the resume.
2.	R. Bharadwaj, D. Mahajan, et.al.	2023	Resume Analysis using NLP.	• NLP , an advanced technology, has been used. • Useful information is extracted from the resume and validation is done, whether or not it is inline with the company requirements. • Resume ranking is also performed by using Resume score.
3.	Preeti A. , Deepali Virmani, et.al.	2020	Resume Selector.	• Advanced technologies such as PySpark and Hadoop have been used in this project. • Hadoop is used to store terabytes of data and Pyspark is used to process those resumes parallelly in a distributed environment.(MCDM) techniques.
4.	Smiti Singhal, Dhara Ajhudia, et.al.	2020	Intelligent Recruitment System using NLP.	• This project extracts data from resumes. • Performs required analysis on data. • Converts this data into useful information for users.

Sr. no.	Author Name	Publish Year	Research Paper	Summary
5.	Nikita Phulkar, Harsha Bhute, et.al.	2021	A Framework for E-Requirement System based on Text Categorization and Semantic Analysis.	• Using NLP technology an autonomous text classification system has been made that POS tags, tokenizes and lemmatizes the data. • Phase Matcher has been used to calculate the score of resumes. • Suggest the lacking skills to users and provides the top resumes to recruiters.
6.	S. Bharadwaj, Macheria Nikhil, et.al.	2022	Resume Screening using NLP and LSTM	• This project categorizes CVs according to the skills they contain into various job options. • This project aids job seekers in evaluating what type of positions they are qualified for based on their resume skills.
7.	B Gunaseelan, Supriya Mandal, et.al	2021	Automatic extraction of segments from Resumes using Machine Learning	• This project uses multi-level classification technique • This project focuses on extracting detailed segment information like skillset, experience and education.
8.	Smiti Singhal, Dhara Ajhudia, et.al.	2023	Resume Classification using ML Techniques	• ML Techniques like Decision Tree, Random Forest, KNN, Support Vector are compared and the one with most accuracy is used. • This project aims to automatically classify resumes to their corresponding suitable positions

2.2) Problem Statement:-

- Employees often don't know how to write resumes that are effective in getting them job interviews. This can be due to a number of factors, such as lack of experience, not knowing what employers are looking for, etc.
- A resume analyzer using NLP can help employees to improve their resumes by providing them with feedback on their content and formatting.

2.3) Objectives:-

- Make a web app which parses the Resume.
- The web app should extract useful information from the resume and provide tips and recommendations to the user for making a resume.

2.4) Scope:-

- Extracts basic information of the employee: Name, Email, Contact, Resume Pages.
- Gives Experience Level of Candidate: Fresher or Experienced.
- Displays Current Skills of the Candidate.
- Shows the Job Profile that the candidate is looking for.
- On the basis of the current skills the webapp recommends new skills to the User.
- Displays Recommended Certifications.
- Gives Resume Score
- Displays Recommended YouTube Videos to the User.

Chapter 3:SOFTWARE ANALYSIS AND DESIGN

3.1 Software Model

3.1.1 Phases of the Software model:

Following the ethics of software engineering, any programmable system follows a process. Process modelling attempts to present this in an abstract form, thus representing the business process or workflow of the project.We have explored several process models for our project.

The process model we have chosen is Waterfall Model: The Waterfall Model is a traditional software development methodology that follows a linear and sequential approach. It is divided into several distinct phases, with each phase dependent on the deliverables of the previous one.

Applying the waterfall model to the development of nvolves breaking down the project into small, manageable increments and delivering functional pieces of the system in short development cycles or sprints.

Here are the typical phases in the Waterfall model:

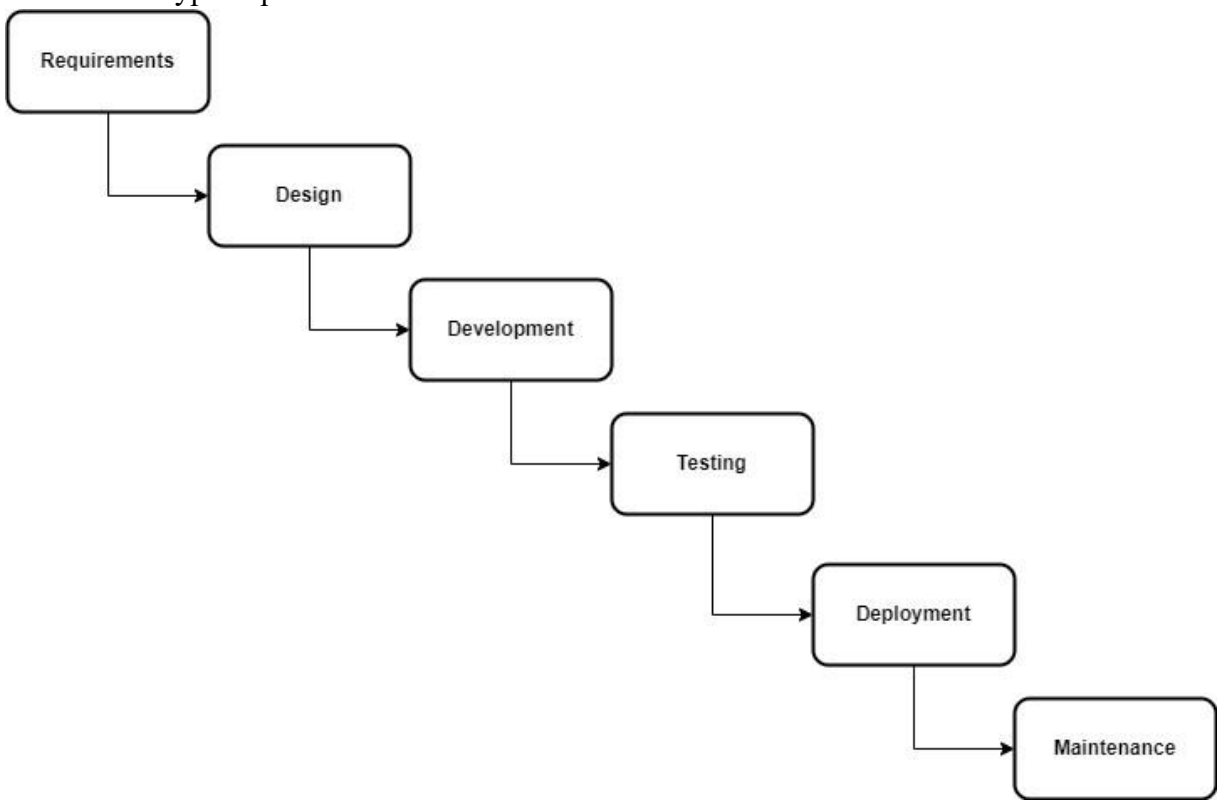

Fig 3.1 Software Model

3.2) Proposed System:-

This system leverages Natural Language Processing (NLP) to analyse resumes and provide users with personalized recommendations for improvement. Here's a breakdown of its functionalities:

1. Resume Parsing:

The system accepts resumes in pdf format.

NLP techniques extract key information like contact details, name, work experience, skills, and achievements.

2. Skill Identification and Analysis:

It identifies skills mentioned in the resume using NLP and pre-defined skill ontologies.

Skills are categorized (technical, soft skills) and compared against industry standards or specific job descriptions (if provided).

3. Personalized Recommendations:

Based on skill gaps and industry demands, the system recommends relevant skills to acquire.

It suggests online courses, certifications, or training programs to bridge identified skill gaps.

4. Resume Scoring and Feedback:

A scoring mechanism evaluates the resume based on factors like keyword density, action verbs, and formatting.

The system provides feedback on resume clarity, structure, and keyword optimization for improved Applicant Tracking System (ATS) compatibility.

5. Learning Resources:

The system recommends targeted YouTube videos or other learning resources to enhance specific skills.

6. User Interface (UI):

A user-friendly interface allows users to upload resumes, view extracted information, and access personalized recommendations.

3.3) System Requirement Specifications:-

A. Functional Requirements:-

- Data Parsing and Extraction of Important Data:

  The system must accurately parse resumes. It should extract core information like contact details, education, work experience, skills, and achievements.

- Skill Identification:

  The system should leverage NLP techniques to identify skills mentioned in the resume using a comprehensive skill ontology.

- Personalized Recommendations:

  Based on skill gaps and industry demands, the system should recommend relevant skills for users to acquire. It should integrate with online course providers or platforms to suggest suitable learning resources (courses, certifications) for identified skill gaps.

- Resume Scoring and Feedback:

  The system should implement a scoring mechanism to evaluate the resume. It should provide constructive feedback on improving resume.

- Learning Resources:

  The system should recommend targeted learning resources like YouTube videos or online courses to enhance specific skills based on the user's profile and identified gaps.

- User Interface:

  The system should provide a user-friendly interface with functionalities for:

  Uploading resumes. Viewing extracted information and identified skills. Accessing personalized recommendations (skills, courses, certifications, score, feedback, learning resources).

B. Non-functional Requirements:-

- Performance: The system should be able to provide the resume analysis without significant delays in response time.

- Security: The system should ensure the security and privacy of user data.

- Reliability: The system should have a high level of reliability, with minimal downtime and robust error handling mechanisms.

- Usability: The user interface should be intuitive and user-friendly, with clear instructions and easy navigation.

- Compatibility: The system should be compatible with a wide range of devices and web browsers to ensure accessibility for all users.

3.4) Hardware and Software Requirements:-

● **Hardware Requirement:**

1. Processor: Ryzen 5
2. Memory: 8GB
3. GPU: Nvidia
4. Internet

● **Software Requirement:**

1. Operating System: Windows 11.
2. Language: Python
3. Database: MySQL, Workbench
4. IDE: VS Code
5. Dependencies:
    - Streamlit
    - Pandas
    - Base64
    - Time
    - PDFMiner
    - Youtube DLP
    - Plotly Express
    - Nltk

3.5) Design:-

3.5.1) Gantt Chart (Timeline chart):-

A Gantt chart is a type of bar chart that illustrates a project schedule. This chart lists the tasks to be performed on the vertical axis, and time intervals on the horizontal axis. The width of the horizontal bars in the graph shows the duration of each activity. Gantt charts illustrate the start and finish dates of the terminal elements and summary elements of a project.Gantt charts are usually created initially using an early start time approach, where each task is scheduled to start immediately when its prerequisites are complete.

Fig 3.2 Timeline chart

3.5.2) Flow Chart:-

The user visits our webapp. If the user is not an admin and is a job seeking candidate then the candidate can directly upload his/her resume on the web app and get personalized recommendations.

If the user is an admin then the admin has to go through the authentication process. The admin has access to sensitive information of the users which should not be made accessible without authentication.

The admin can view details of all the job seeking candidates.

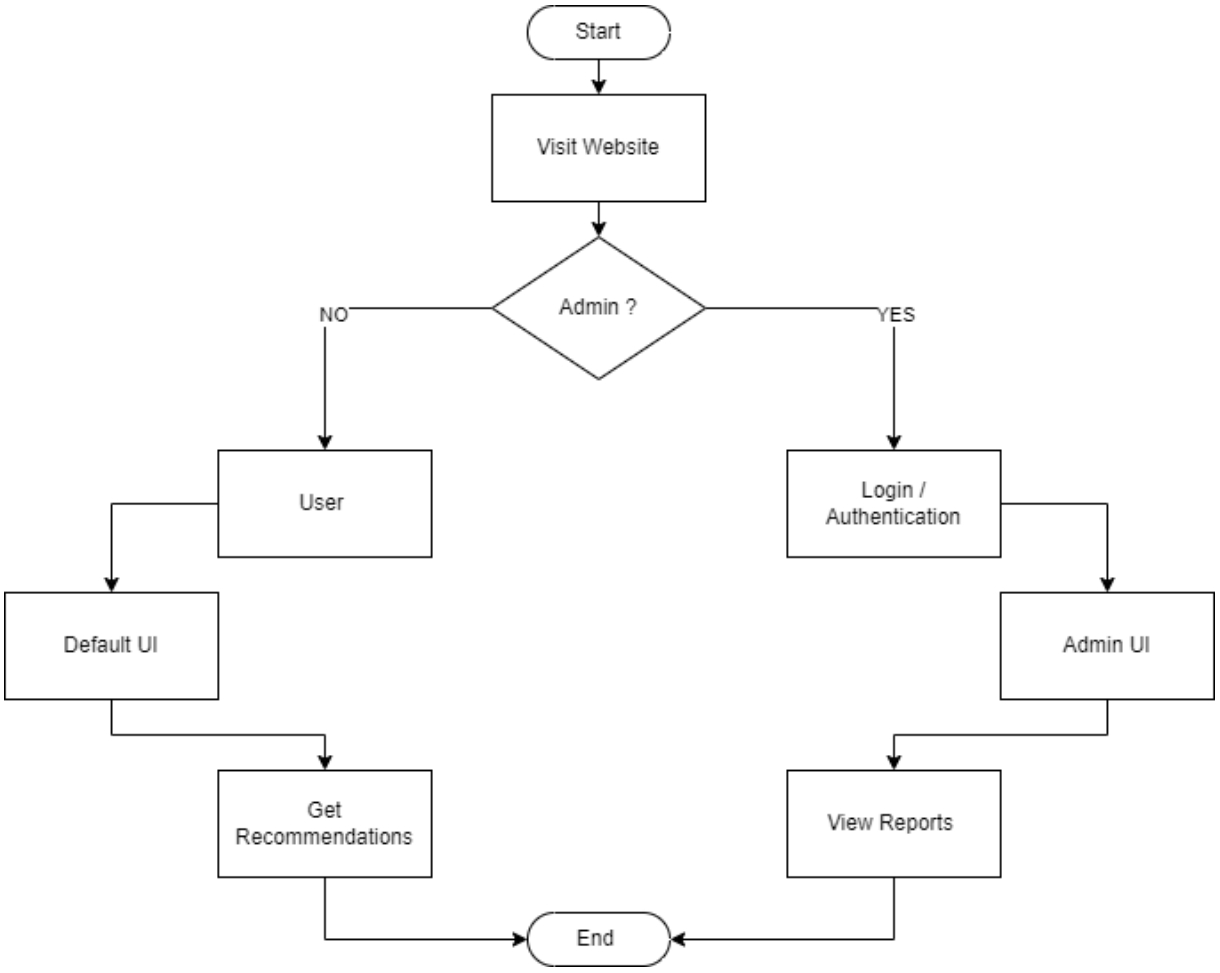

Fig 3.3 Flow Chart

3.5.3) Use Case Diagram :-

The Job Seeking Candidates can upload their resumes and view the recommendations recommended by out webapp.

The admin can view the data of all the resumes uploaded by the candidates, but it is mandatory for the admin to go through authentication process as the admin has access to personal and sensitive information of the job seeking employees.

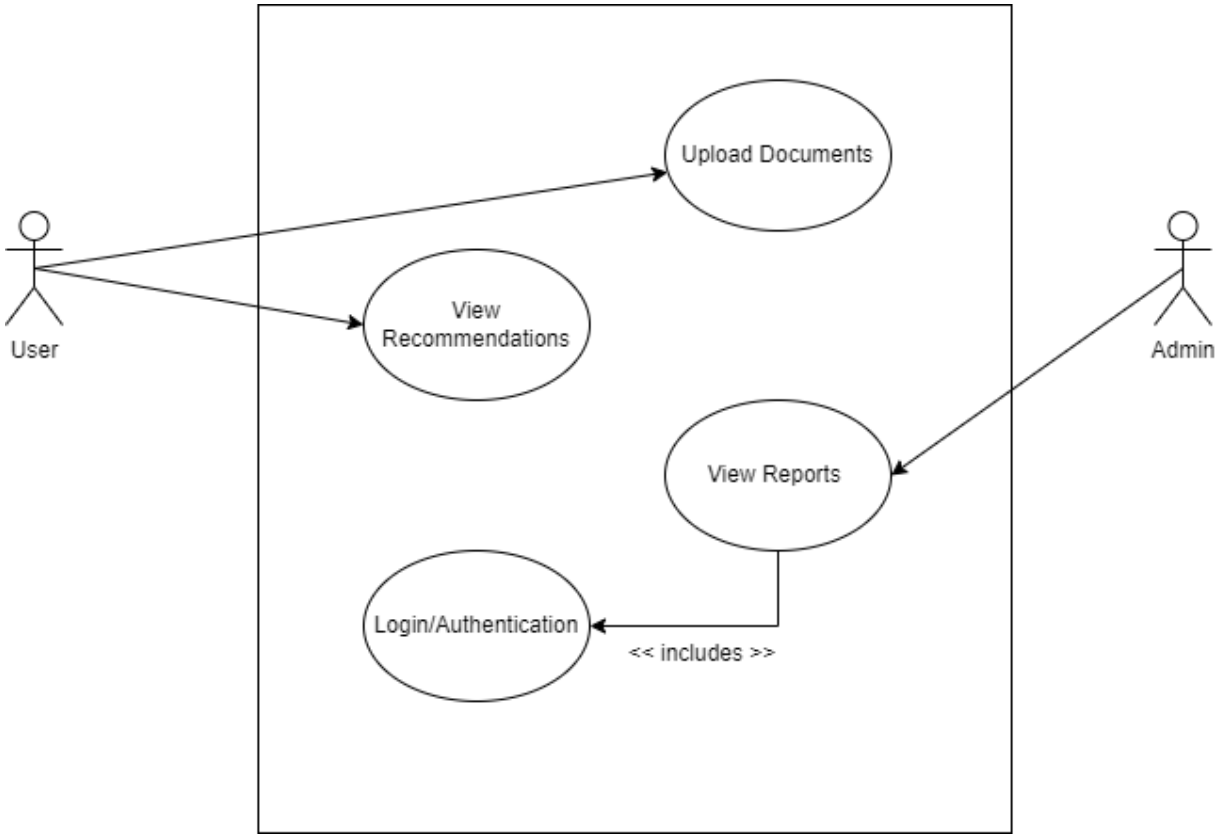

Fig 3.4 Use Case Diagram

Chapter 4:METHODOLOGY

The User visits out website.

If the user is a job seeking candidate then the users uploads his resume in the default GUI without going through the authentication process.

Using Pyresparser we parse the resume and extract useful information from the resume.

The code is written in such a way that considering the information extracted from the resume we provide personalized recommendations to the Job Seeking Candidate.

Each and every detail, recommendation that is displayed on the GUI of the User is stored in the database.

The admin has to go through authentication process.

The admin can view all the data of users that is saved in the database.

The admin can also download the excel file which contains all the User data that the database has.

4.1) METHODOLOGY:

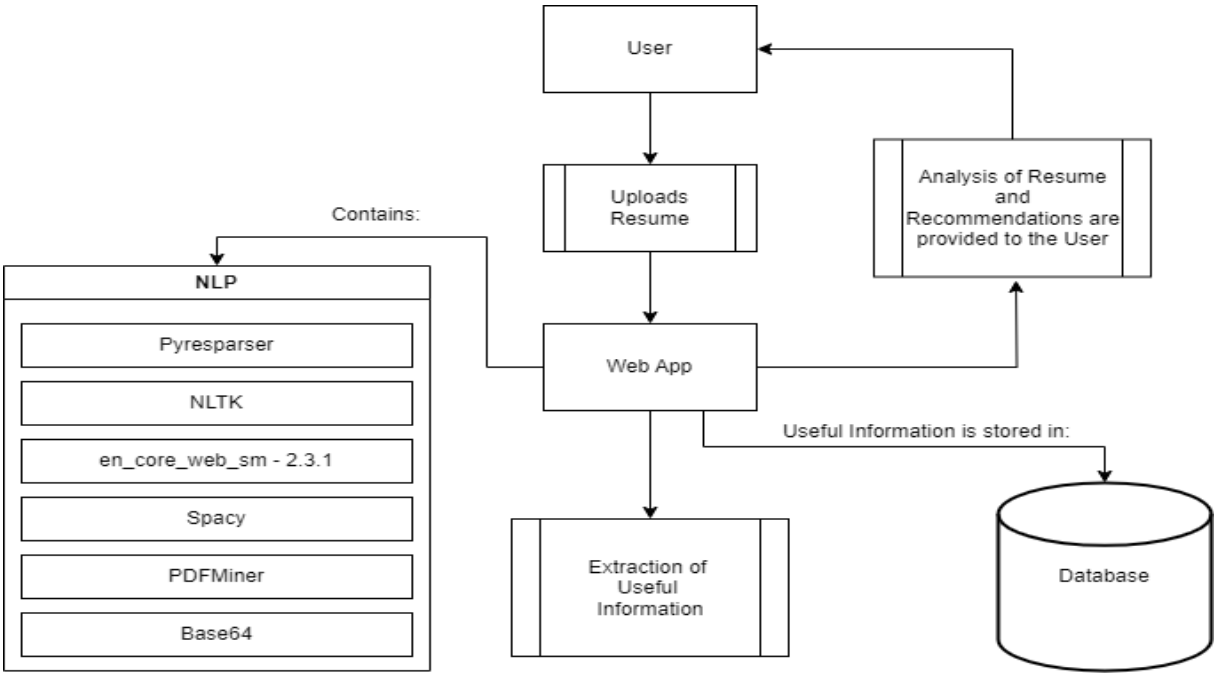

Fig 4.1 Methodology

Chapter 5: RESULTS

5.1) User :

The admin can manage all the vendors and their registered parking lots and set the pricing accordingly as shown in Figure 5.1.1 and Figure 5.1.2.

Fig 5.1.1 Default GUI

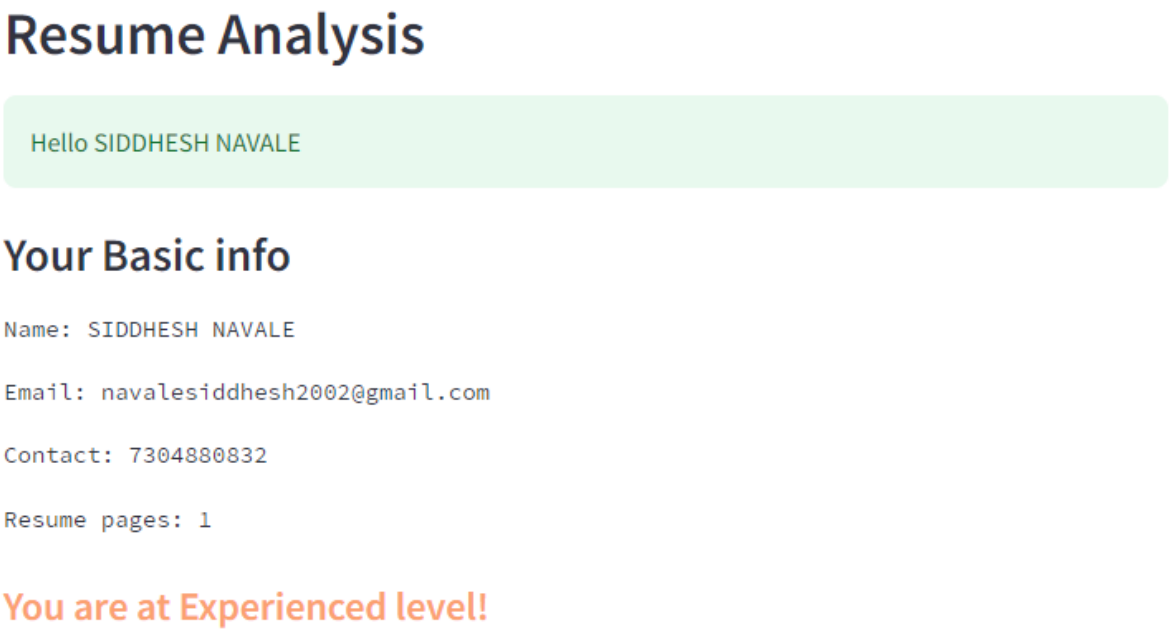

Fig 5.1.2 Basic Information and Experience Level

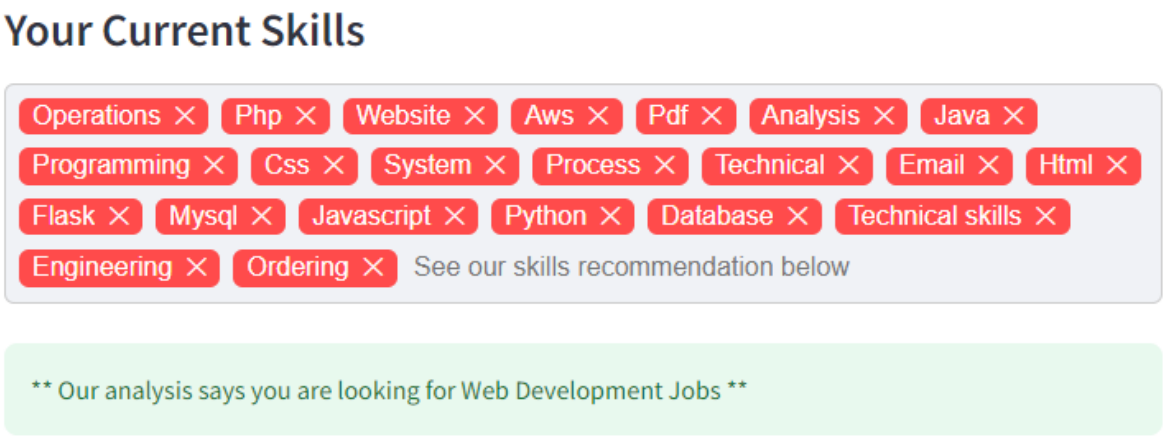

Fig 5.1.3 Your Current Skills and Job Role you are looking for.

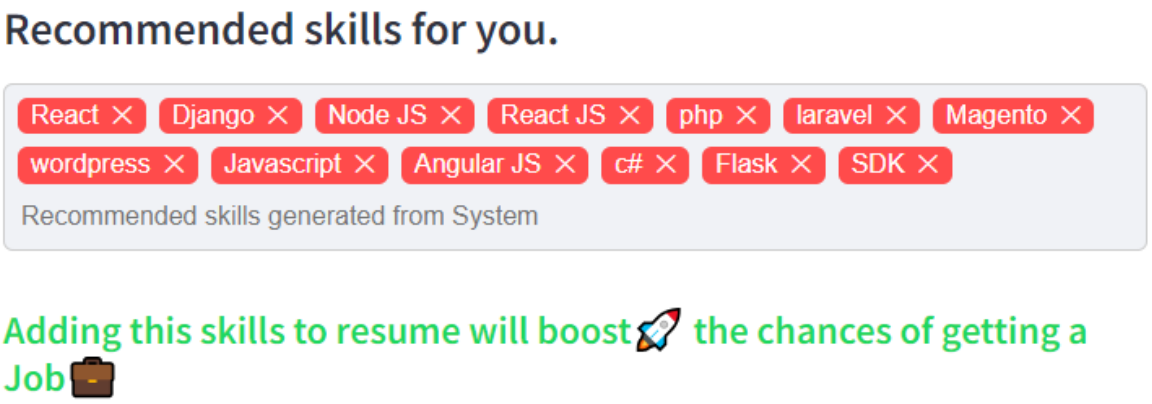

Fig 5.1.4 Recommended Skills.

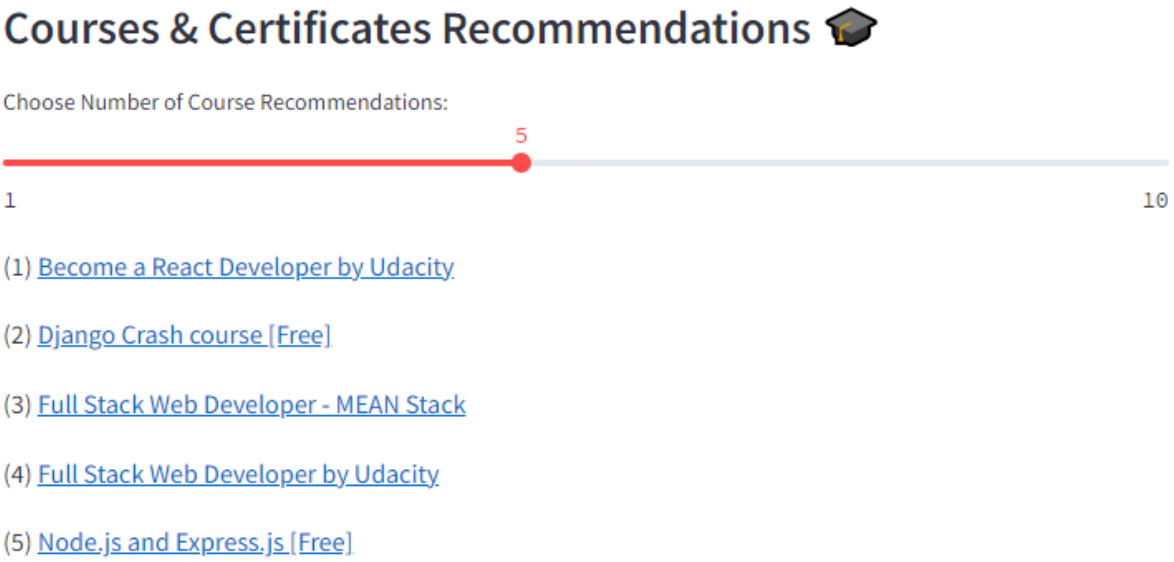

Fig 5.1.5 Course and Certificate Recommendations

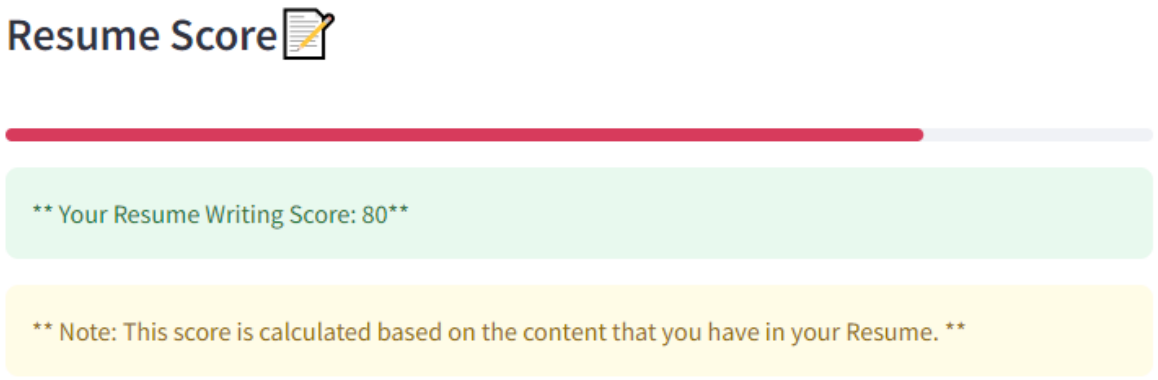

Fig 5.1.6 Resume Score.

Fig 5.1.7 Recommended Youtube Videos (Resume Enhancement Tips)

Fig 5.1.7 Recommended Youtube Videos (Interview Acing Tips)

5.2) Admin:

Fig 5.2.1 Admin Login Page

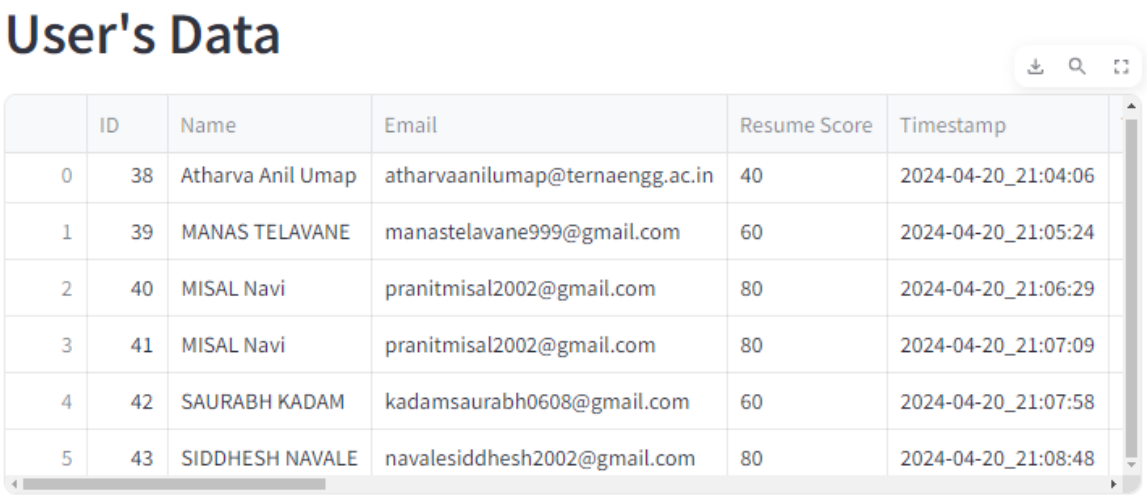

Fig 5.2.2 Uers's Data Viewable by the Admin

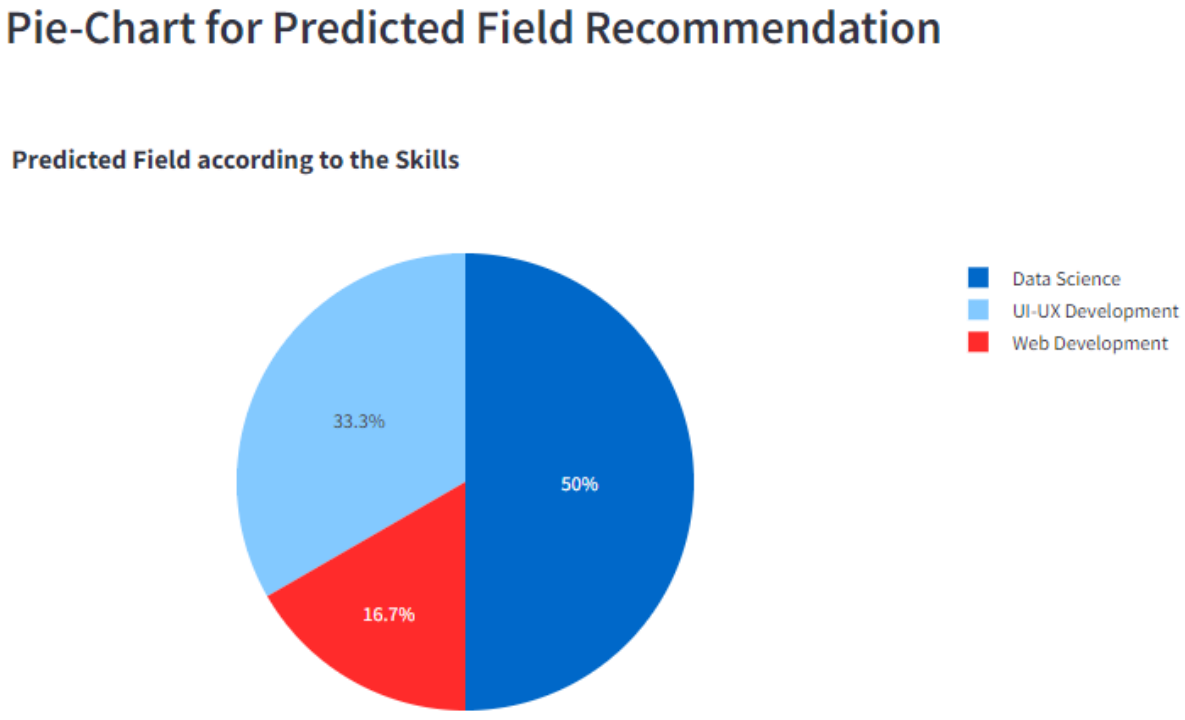

Fig 5.2.3 Pie-Chart for Predicted Field Recommendations

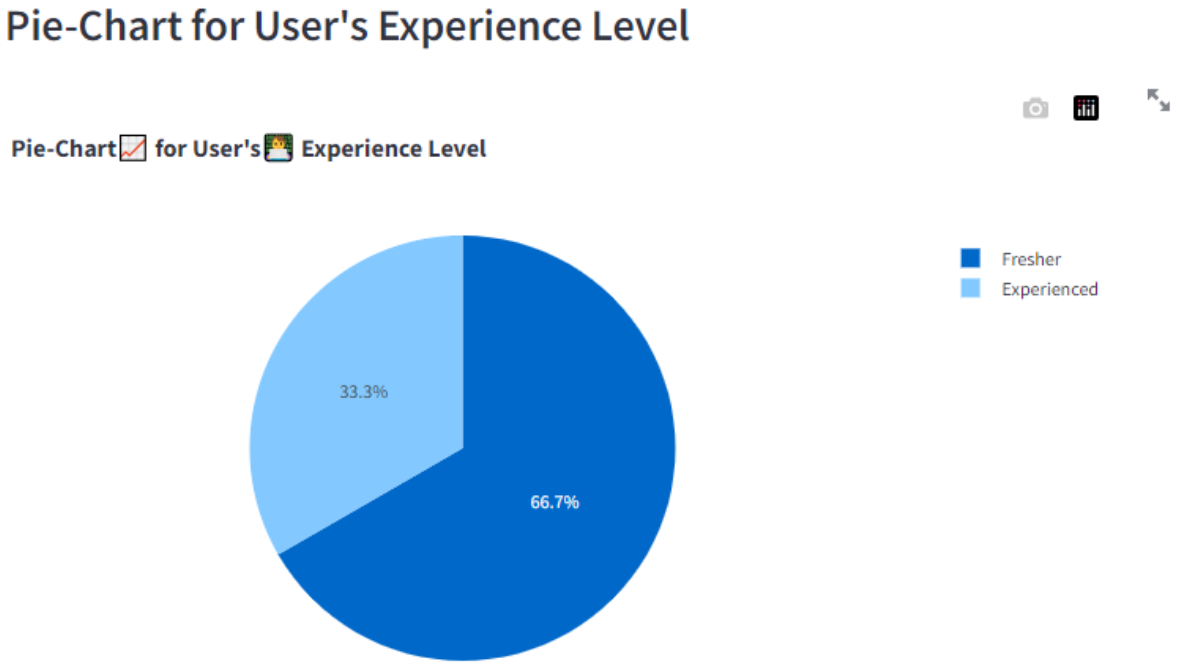

Fig 5.2.3 Pie-Chart for User's Experience Level

Chapter 6: CONCLUSION

6.1) Conclusion:-

In conclusion, this project developed a novel NLP-powered resume analyzer that transcends simple parsing. By extracting key skills and experience, the analyzer provides personalized recommendations for improvement. This includes suggested skills to acquire, relevant courses and certifications, a calculated resume score to gauge effectiveness, and targeted YouTube videos for further learning.

This comprehensive approach empowers users to actively enhance their resumes, increasing their visibility and competitiveness within the job market. This technology offers significant advantages for both employers, who can streamline the hiring process, and employees, who can tailor their resumes for maximum impact.

As NLP continues to evolve, this resume analyzer presents a foundation for further development, paving the way for a more efficient and personalized job search experience.

REFERENCES

[1] P. Swami and V. Pratap, "Resume Classifier and Summarizer," 2022 International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (COM-IT-CON), Faridabad, India, 2022, pp. 220-224, doi: 10.1109/COM-IT-CON54601.2022.9850527.

[2] Bharadwaj, R., Mahajan, D., Bharsakle, M., Meshram, K., Pujari, H. (2023). Resume Analysis Using NLP. In: Suma, V., Lorenz, P., Baig, Z. (eds) Inventive Systems and Control. Lecture Notes in Networks and Systems, vol 672. Springer, Singapore. https://doi.org/10.1007/978-981-99-1624-5_40

[3] Arora, P., Virmani, D., Jain, A., Vats, A. (2021). "RESUME SELECTOR" Using Pyspark and Hadoop. In: Pandey, V.C., Pandey, P.M., Garg, S.K. (eds) Advances in Electromechanical Technologies. Lecture Notes in Mechanical Engineering. Springer, Singapore. https://doi.org/10.1007/978-981-15-5463-6_52

[4] A. Sharma, S. Singhal and D. Ajudia, "Intelligent Recruitment System Using NLP," 2021 International Conference on Artificial Intelligence and Machine Vision (AIMV), Gandhinagar, India, 2021, pp. 1-5, doi: 10.1109/AIMV53313.2021.9670958

[5] R. Shaikh, N. Phulkar, H. Bhute, S. K. Shaikh and P. Bhapkar, "An Intelligent framework for E-Recruitment System Based on Text Categorization and Semantic Analysis," 2021 Third International Conference on Inventive Research in Computing Applications (ICIRCA), Coimbatore, India, 2021, pp. 1076-1080, doi: 10.1109/ICIRCA51532.2021.9544102.

[6] S. Bharadwaj, R. Varun, P. S., M. Nikhil and G. C. Babu, "Resume Screening using NLP and LSTM," 2022 International Conference on Inventive Computation Technologies (ICICT), Nepal, 2022, pp. 238-241, doi: 10.1109/ICICT54344.2022.9850889.

[7] B. Gunaseelan, S. Mandal and V. Rajagopalan, "Automatic Extraction of Segments from Resumes using Machine Learning," 2020 IEEE 17th India Council International Conference (INDICON), New Delhi, India, 2020, pp. 1-6, doi: 10.1109/INDICON49873.2020.9342596.

[8] B. Surendiran, T. Paturu, H. V. Chirumamilla and M. N. R. Reddy, "Resume Classification Using ML Techniques," 2023 International Conference on Signal Processing, Computation, Electronics, Power and Telecommunication (IConSCEPT), Karaikal, India, 2023, pp. 1-5, doi: 10.1109/IConSCEPT57958.2023.10169907.

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)
An International Open Access, Peer-reviewed, Refereed Journal

Ref No : IJCRT/Vol 12/ Issue 4 / 068

To,
Amrut Mane

Subject: Publication of paper at International Journal of Creative Research Thoughts.

Dear Author,

With Greetings we are informing you that your paper has been successfully published in the International Journal of Creative Research Thoughts - IJCRT (ISSN: 2320-2882). Thank you very much for your patience and cooperation during the submission of paper to final publication Process. It gives me immense pleasure to send the certificate of publication in our Journal. Following are the details regarding the published paper.

About IJCRT	:	Scholarly open access journals, Peer-reviewed, and Refereed Journals, Impact factor 7.97 (Calculate by google scholar and Semantic Scholar | AI-Powered Research Tool) , Multidisciplinary, Monthly, Indexing in all major database & Metadata, Citation Generator, Digital Object Identifier(DOI) | UGC Approved Journal No: 49023 (18)
Registration ID	:	IJCRT_256897
Paper ID	:	IJCRT24A4068
Title of Paper	:	Resume Analyzer Using NLP
Impact Factor	:	7.97 (Calculate by Google Scholar) | License by Creative Common 3.0
Publication Date	:	17-April-2024
DOI	:	
Published in	:	Volume 12 | Issue 4 | April 2024
Page No	:	j327-j330
Published URL	:	http://www.ijcrt.org/viewfull.php?&p_id=IJCRT24A4068
Authors	:	Amrut Mane, Pranit Misal, Siddhesh Navale, Atharva Umap, Ramesh Shahabade
Notification	:	UGC Approved Journal No: 49023 (18)

Thank you very much for publishing your article in IJCRT.

Editor In Chief
International Journal of Creative Research Thoughts - IJCRT
(ISSN: 2320-2882)