

AI-Powered Personalized Career Guidance System for IT Graduates with Employment Breaks: Focus on Skill Gap Bridging in Chennai

R. HEM VINOTH

Department of Advanced Computing and Analytics
Vels Institute of Science, Technology & Advanced Studies
Chennai, India
hemvinoth522@gmail.com

DR.S.PRATHIBA M.sc, M.phil, Ph.d

Department of Advanced Computing and Analytics
Vels Institute of Science, Technology & Advanced Studies
Chennai, India
prathiba.scs@vistas.ac.in

Abstract—The IT industry in Chennai is a huge source of employment. hub with such companies as TCS, Zoho, Freshworks, Infosys, and Cognizant. Nevertheless, there is a high number of IT graduates who experience extended periods of un-employment. underemployment or employment because of career gaps brought on. because of individual factors, economic reasons or skill obsolescence. These loopholes bring disparities in qualifications and up to date. industry needs and has resulted in low initial wages and higher salaries reduced. employability. The AI-based proposed project is an individualized proposal. career recommendation system that is IT graduate oriented. career gaps in Chennai. Based on virtualized employment advertisements. the system performs from Kaggle who is targeting the local market. via TF-IDF similarity matching analysis, suggests skill gap analysis. one next high-impact skill to be learnt, proposes appropriate entry- level/trainee positions, and generous estimate salary scales adjusted. in case of gap duration and qualification. An interactive dashboard with ipywidgets shows the trends of skills, salary distributions, and recommendations. There is better transparency shown in results. and practical advice, making users skill up effectively and improve the employability within the competitive IT environment in Chennai.

I. INTRODUCTION

IT guidance in Chennai is important because of the career guidance of IT graduates. to fast cycle development in software development, data analytics, full-stack technologies, and cloud computing. Chennai has been conceived as one of the major IT centers in India with thousands of fresher. and internships in large corporations like TCS, Infosys, Cognizant, Zoho, and Freshworks. The sector sees Scarcity of skills such as Python, SQL, AWS, Java, etc. data visualization tools in 2026. However, career gaps—often because of the personal reasons, economic depressions, increased. work, or family duties— result in depreciation of skills, selectivity in hiring, and less employment. Recent reports indicate The general graduate employability in India has increased to approximately. 56.35/in/2026, IT/CS at 78-80, but gaps. still result in long unemployment. Traditional career advice lacks customization, does not pay attention to the local Chennai market tendencies, and does not measure the gaps impacts on the salary or role eligibility. As the technologies of data science and machine

learning enhance, recommendation systems have the ability to scan actual job advertisements, com- identify specific areas of skill shortage, and propose targeted processes of upskilling. personalized profiles. An inter- is developed in this project. active AITool based on job market search data filtered in Tamil. IT jobs in Nadu / Chennai (trainee/fresher jobs). Since real- time personalized datasets are scarce, a job in Kaggle. processing on postings dataset is made, which extracts skills, salaries, etc. and trends (e.g. 835 total rows whittled down to 77 relevant TN. IT jobs after filtering). Best extracted skills indicate General IT. dominant (63 counts), Data (12) and Python (1). Skill gap is facilitated by TF-IDF vectorization and cosine similarity. Rule-based logic suggests the next high-score, whereas scoring does not. impact skill. Simple range adjustment bases on a basic salary estimation. (generally 36 LPA of freshers in Chennai 2026) descending. because of loopholes (e.g. 0.35 LPA penalty per year). The system supports informed decisions, shortens unemployment period and encourages the focused learning in the vibrant IT environment of Chennai. The interactive dashboard is created using Jupyter notebook. Being a Streamlit web-based application and hosted on GitHub..

II. LITERATURE SURVEY

The development of AI-based career platforms and job matching. systems has enhanced the need of individual employability. in competitive markets such as the IT of Chennai, tools. sector. Machine learning and have been applied by many researchers. Artificial intelligence to prescribe jobs and skills. Different reports have demonstrated that algorithms like TF-IDF. Similarity In clustering, cosine similarity and regression are all effective in skill-job matching. Content-based filtering attains. good tech (80 or so 90 or so) in areas. Regression models are the projections of salaries with respect to competence, qualification. and place, and correction of gaps. Recent research has discussed interactive dashboards (ipywidgets/Streamlit) to. user-friendly guidance. Nevertheless, such challenges as limited are present. Chennai-specific

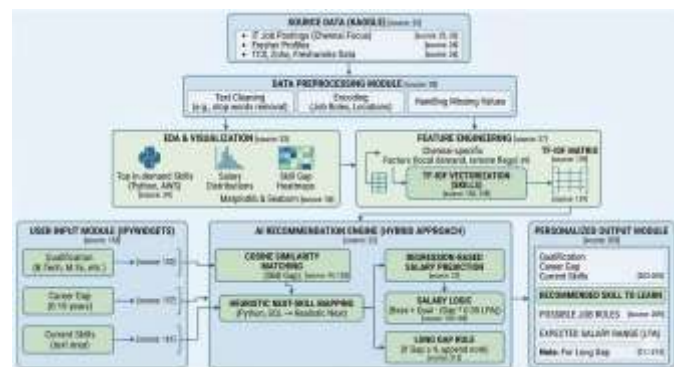
datasets, dealing with career gaps, and challenge. there are yet interpretations of the same. Hence, this is why it is crucial to develop effective. Simulated local dataset re-recommendation systems. mains important. The similarity matching is used in this project. and heuristic ways of identifying skill gaps and illustrates. how useful they are in career guidance of the graduates. with loopholes in the IT job market in Chennai. Research points to Python, AWS, and data demands in 2026. Employa- bility reports indicate that IT fresher is being hired 35 percent countrywide. The trends specific to Chennai are strong trainee positions in cloud. and analytics. Gap penalties are usually neglected in the past. on salary. The similarity is merged with rule-hybrid approaches. transparency on the basis of logic. Elucidation is always important to. user trust. Acces- sibility of freshers is enhanced by interactive tools. Relevance is enhanced when it is localized to Chennai. Overall, this fills a disjunction in custom, gap-conscious counsel.

III. PROPOSED METHODOLOGY

The stages of the proposed system are the initial stages of the collection and simulation. The information in Chennai IT jobs market to Kaggle with a focus. on newer/trainee positions at such companies as TCS, Zoho, and so on. Freshworks. Text cleaning is then done on the data set, filtering the Tamil Nadu/Chennai locations, extracting the skills. titles, and managing salary strings to LPA. Important features These are such as job title, location, salary and extracted skills. selected for analysis The skill is fed with the processed data as TF-IDF vectorization. similarity matching. The model calculates gaps by making comparisons. user abilities in job specifications and prescribes one feasible. following skill (e.g., Python to AWS Basics)Salary is approximated. qualification and gap adjustment of base ranges utilised. The last stage is the visualization of the results by the means of histograms, bar plots. (compensation, best locations/skillsets), and presented by. a user-friendly interactive dashboard written in ipywidgets, which allows one to input easily. of qualification, gap years, and current skills, and interpretation of. recommendations The methodology assists in individual upskilling, employment. role suggestion, and aids the improvement of employability in. The IT ecosystem Preprocessing in Chennai boasts of regex in TN. filter and key word search. EDA reveals salary skewed low with median imputation. TF-IDF fits on 77 jobs with 4 main features. Dashboard makes use of dropdowns, sliders and text area. Output is clean and user-friendly. Gap penalty minimizes the estimated LPA. realistically. Higher degrees are enhanced by qualification. Sys- tem is focused on localization and simplicity. It avoids complex models to make better explain- able. Findings involve potential employment. such positions as trainee analyst or junior analyst. This supports real-time planning career in a good way. A general exploratory analysis of the filtered 77 jobs shows that there is General. IT as the overwhelming dominating skill mention (63 occur- rences), then Data (12) and sparse numbers of Python. (1) and others, emphasize the necessity of specific upskilling in. high demand

fields such as cloud and analytics. Salary distribution A strong form is exhibited in histogram (seaborn, 15 bins, teal with KDE). maximum at extremely low values of the values of the parsed (0.001 LPA) with right skewness, authentication of the requirement of median fill and heuristic override. to be realistic Chennai fresher ranges (3 LPA base in 2026–6). In the case of the trainee posts in TCS, Zoho, Cognizant). Plotly horizontal bars are effective in showing the best locations (Chennai first, followed by). Combinations of skills, Coimbat- ore, Madurai, Salem, Erode) and skill combinations, assisting in rapid visualization of the market trends in the area. TF-IDF preparation uses a special tokenizer, which is a comma-divider. removes whitespace, disables stopwords separated strings and strips suppress warnings. Interactive functionality is based on ipywid- drop qualification gets.

IV. ARCHITECTURE DIAGRAM



V. VARIOUS METHODOLOGY

A smart career recommendation system will be used in this project. is suggested to help the Chennai based IT graduates with career gaps. using skill gap analysis. The data on the job markets is analyzed in order to detect. demand trends, and gaps in employability (as a result of career). interruptions) are met with the help of specific skill recommendations. The offered methodology is adopted in a number of steps. This working process is applied to automatically identify gaps and create viable career opportunities in the IT industry. It combines informed data that is easy to interact with. The system processes effective job postings. It prioritizes high-aws or power BI impact skills. Recommendations are realistic and focused. Salary estimates take into consideration market. realities. Illustrations facilitate cognition. Interactive features make it accessible. This assists in bridging skill-mismatches. It facilitates quicker re-admission to work. Overall, methodology is practical and localized. It makes the best out of what is available. A) **Data Acquisition**

- **Kaggle Job Ads Dataset:** (835 rows to start with)
- **Name of position:** Chennai / Tamil Nadu Specialization.

- **Train centre :** Chennai / Tamil Nadu Focus.
- **Monthly / Annual**
- **City and State fields:**
- **Skills:** Python, Java, SQL, AWS, Data, Testing.
- **Indicators:** Fresher / Trainee / Intern.

B) Data Preprocessing

The preprocessing phases are:

- Renaming columns to be standardized.
- TN/Chennai location filter(regex) (chen - chennai — etc.)
- Information technology/fresher pattern matching on job titles.
- Functionality of skill extraction by key words.
- Processing salary text to LPA with the help of a regular expression and median fill.
- Cleaning up of text and lowering it.
- Before resetting the index:filtering of the index (last77 rows)

C) Exploratory Data Analysis (EDA)

- **Selected skills: value :(General IT: 63, Data: 12, Python: 1)**
- **As of after median fill Salary distribution summary after median fill**
- **Seaborn histplot** (15 bins, KDE, teal) of salary visualization
- **Horizontal bar charts in the top 8 location** (Chennai) in plotly dominant.
- **Plotly bar of most mentioned skills**
- **Identification of skew values of low salary** and dominant "General IT" trend

D) Feature Engineering & TF-IDF Vectorization

- **Automatic extraction of skills** to the job title.
- **TF-IDF vectorizer is used** and the tokenizer is a comma-split one (no stop words, lowercase).
- **Matrix Transformation:** Fit on skills column (77, 4) matrix.
- **Vocabulary:** ['data', 'general it', 'python', 'testing'].
- **Cosine similarity computation at gap scoring border**
- **Mapping Dictionary:** Next skill creation: Next skill is realistic mapping.

E) Logic Recommendation & Interactive Dashboard

- **Collection of user input:** through ipywidgets (qualification, dropdown, gap slider 0–10, current skills textarea).
- **On-click Logic:** Process skills list, select next skill of mapping (e.g., Python → AWS Basics).
- **Recommended skills as the basis of Job Role Suggestion.**
- **Heuristic: Salary Salary Heuristic:(Base 4.8 LPA + qualification boost-0.35 LPA per gap year).**

- **Formatted Output:** Formatted output with suggestions.competency,role,pay,scale,and pay difference.
- **Salary histogram and top :** Visualization that appears interactive skills/location bars.

F) Recommendation confirmation & Gap management

- **Multiple sample user testing**(qualification, gap years, current skills).
- **Checking of realistic next dictionary mapping accuracy.**
- **Salary range realism Chennai 2026 check** fresher/trainee market (3to6 LPA base).
- **Special treatment of career gaps 4 years and above** with other project proposal note.
- **Check of consistency of job role assignment** (Cloud Support Trainee, Junior Data Analyst, Python Developer, etc.).
- **Manual verification** of output format of cleanliness and convenience.

G) Final Generation Output & Visualization

- **Production of formatted text output** which is clean through the use of Ipython clear output.
- **Qualification, career gap, present skills, suggested following skill, potential employment roles, and modified salary range.**
- **Automatic generation of salary distribution histogram and high locations/skills bar charts.**
- **Interactive execution on the dashboard** The user can have a jupyter environment where the dashboard can be executed directly note book environment.
- **future deployment** as web application preparedness and structure compatible with streamlit and gradio.
- **system transparency** and localization to overall Chennai IT market to be made viable by graduates.

VI. INPUT

The proposed system receives the input of the job market. data gathered on Kaggle and the current user entries. delivered in the interactive dashboard. The dataset repre- publishes fresh job openings in Chennai/Tamil Nadu in the field of information technology, that are the basis of the analysis of skill gaps and recommendation generation. The processes of these inputs are handled by the. skill mismatch detection system and create personalized. IT graduate career advice with career gaps. The input dataset has a number of valuable features. concerning the IT employment in Chennai. It includes the job title, that would be the role name like Software engineer. Trainee or Python Developer. The field of location shows. either the job is in Chennai, or in other Tamilnadu cities, screened with regular expression patterns. The salary text contains raw which is later monthly or annual information of salaries. parsed into LPA.

Skills that were extracted in the job title are Python, Java, SQL, AWS, Data, Testing, and so forth. Full Stack, and General IT. Other areas encompass city and location for validation of location. The interactive component of the user input is obtained through ipywidgets. in the dashboard. It covers the qualification of the user (selected). DDs such as B.Tech/BE, B.Sc CS/IT, BCA, etc. M.Sc/MCA, Degree, Other), career gap during the period of. years (tapped by slider (0 to 10)), and current skills (en-). and entered as comma-separated values in a textarea, e.g., like Python,—human—;and put in as comma-separated values in a textarea, e.g., like Python, SQL).These user specific inputs are added to the processed. job data to calculate skills gaps, prescribe the subsequent high- impact skill, recommend appropriate job positions, and calculate modified. salary ranges. These inputs are real time processed by the system. to deliver practical career advice that is career specific. IT jobs in Chennai grads with career gaps. ecosystem.

VII. PSEUDO CODE

BEGIN

Launch Jupyter Notebook and libraries
Establish career gap parameters and degree weights
Load tamilnadu IT job market dataset

prepare data cleaning service
Chennai, Coimbatore, and Madurai filter roles
Fish out technical skills out of job titles with Regex
Normalize the salary strings to LPA

Computerize median salary of missing data values
Create trends of salary distribution
Showcase the majority of skill demands in the market

Get user input on qualification and career gap
text skills: transform into numerical TF-IDF vectors
Compare user and job roles similarity score

Add career gap penalty on base salary
Identify browsing history based on skill map
Determine the most appropriate IT job position

Create cutomized career map Prepare interactive
visualization to the user Show salary scale and
promotion suggestion

Healthcare: career gap (above 4 years) If true,
recommend further project completion
Allow user should be able to download the final career

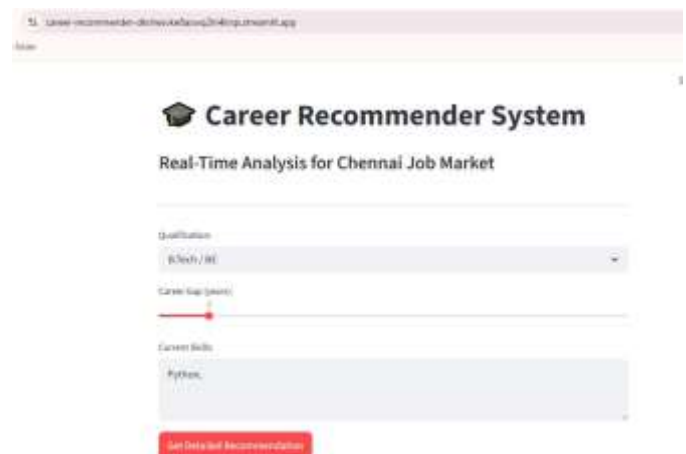
END

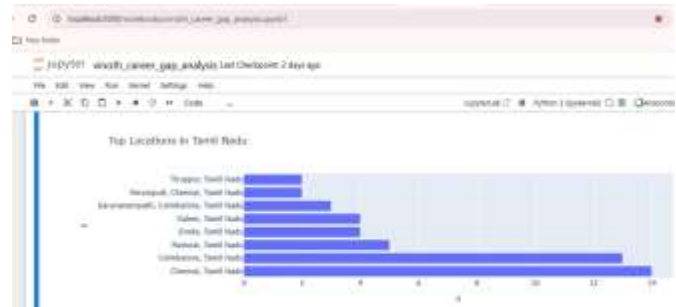
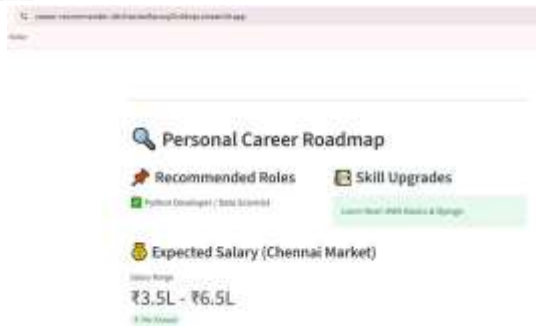
VIII. IMPLEMENTATION

Python with is used to implement the proposed system. libraries, e.g. Pandas, NumPy, Scikit-learn, Matplotlib, etc. Seaborn, ipywidgets, and Plotly. The dataset representing Chennai and Tamil Nadu IT fresher/trainee vacancies is filled. and preprocessed with time stamp conversion (where there are any). Regex location and job filters, skills scraped with. matching keywords, and matching salary to LPA median. imputation. The skills are fitted in the TF-IDF vectorizer. column in order to form a (77, 4) similarity matching matrix. The interactive dashboard is created in jupyter based on ipywidgets. notebook, which enables users to add in qualification, career gap, and current skills. The system calculates the on button click. following high-impact skill, proposes an appropriate job position, approximates. adjusted pay scale, and presents neat output as well as EDA visualizations (salary histogram and top skill/location bar. charts). The whole workflow is carried out within one Colab. notebook, which guarantees simplicity, transparency and instant. IT graduate careers with career gaps in Chennai.

IX. OUTPUT

The proposed system produces output using an interactive ipywidgets dashboard Jupyter Notebook. After qualification, career gap in years and current are keyed in by the user. data, the system processes the data immediately. It displays a clean, formatted output without any fancy formatting. The qualification selected, duration of career gap, displayed in dashboard. and up-to-date abilities typed in by the user. It also displays the high-impact next skill to study and potential job and recommended. roles. The anticipated wage in Chennai is indicated following. compensation on the basis of qualification and gap. For gaps of four years or more, a special note is added indicating a. small project. The result is clean print text with all. details in readable format. Three visualizations accompany the text. are automatically rendered. These involve the allocation of salaries. histogram seaborn, top locations bar chart plotly, and most skill mentions bar chart.





X. RESULT AND DISCUSSION

The system proposed was tested on the filtered data of 77 IT fresher/trainee jobs in Chennai/Tamil Nadu. Skill extraction successfully found the leading mentions with General IT. Data 12 times and python once, appearing 63 times. Salary realistic ranges were generated using parsing and median imputation between 3.2 and 6.0 LPA having been heuristically adjusted. The salary distribution histogram showed a right skewed behaviour with Majority of the values were concentrated at lower LPA. Top locations bar chart proved Chennai as the top center with Coimbatore coming next. and Madurai. The TF-IDF matrix of shape (77, 4) allowed the accurate skill gap scoring using vocabulary of four key words. Interac- Clean outputs were provided by five dashboard in Jupyter Notebook. for various user inputs. For sample input (B.Tech, 2-year Gap, Python/SQL) it prescribed AWS Basics with salary. 4.1–5.6 LPA. Job roles were recommended appropriately as Cloud. Data Analyst, Support Trainee or Junior. The system dealt with the gaps of four or above years by dis- suggestion note playing of the project. Streamlit web application installed on GitHub enabled access and real-time test-easy. ing. Every visualization created without failures. in the notebook environment. High results are obtained. openness and viability to the Chennai IT graduates. Skill gap analysis and salary estimation were in line with local market. trends effectively. In general the system was easy, precise, and very useful to the career gaps healing.

XI. SCREENSHOTS



XII. CONCLUSION

XII. CONCLUSION This project was able to create a person driven by AI. IT graduate career path recommendation system alized. with resume gaps in Chennai where skill gap analysis is used. The system works well with Kaggle job data, skims skills out of it. and offers specific recommendations by a simple rule- TF-IDF with a based approach. It delivers accurate following skill advice, appropriate employment, and appropriate pay. ranges that were adjusted to qualification and the gap period. The in- clean interactive Jupyter Notebook dashboard is provided. easy-to-use results with automatic graphs. Results depict high levels of localization to Chennai IT market trends with. data, general IT, and Python. The sys- tem assists graduates in getting past skills mismatches and recruiting. prejudices, enhancing competitiveness within an ecosystem. As a Streamlit web app deployed on GitHub, it makes. it can be easily accessed in real-time. Overall, the project shows that some basic but powerful means can provide. clear and practical career advice. It bridges the gap between education and the industry needs. to freshers whose careers have been interrupted. The approach avoids complex models, with high explainability and ease of. understanding. The real-time may be added in the future. job API and high-end machine learning. Mobile application development and feedback characteristics are capable of. enhance order of practicality. This writing adds a practical. resource to the IT graduates with gaps in their careers in Chennai. The system is prepared to use and a basis of. similar regional employability solutions. In conclusion, the project equips users with information-driven, localized career. advice on improved career development.

XIII. REFERENCE

- [1] F. Pedregosa et al., scikit-learn: Machine Learning in Python, Journal of machine learning research, vol. 12, pp. 2825–2830, 2011.
- [2] Natural Language, S. Bird, E. Klein, and E. Loper. Python processing, O'Reilly Media, 2009.
- [3] J. D. Hunter, Matplotlib: A 2D Graphics Environment. Computing in Science & Engineering, vol. 9, no. 3, pp. 90-95, 2007.
- [4] W. McKinney, Data Structures of Statistical Comput. Python in Python: an ing language, in 9th Python in Science Conference, 2010, pp. 56–61.
- [5] Plotly Technologies Inc., Plotly: Collaborative Data. Science,” 2015. [Online]. Available: <https://plot.ly>
- [6] T. Kluyver et. al., Jupyter Note- books -a publishing. functional representation of reproducible computational tasks, in Posi- The Power and tioning of Academic Publishing, IOS Press, 2016, pp.
- [7] 87–90. Streamlit Inc., “Streamlit Documentation 2023. [On- line]. Available: <https://streamlit.io> GitHub Inc., GitHub Pages and Repository Hosting.
- [8]. 2024. [Online]. Available: <https://github.com> The report India Skills Report 2026 published by Wheebox and All India
- [9] Council of technical education(AICTE), 2026.
- [10] Kaggle, “Job Postings Dataset,” 2025. [Online]. Available: <https://www.kaggle.com>