International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 11 Issue: 10 | Oct 2024 www.irjet.net p-ISSN: 2395-0072

COURSE RECOMMENDATION SYSTEM BASED ON RESUME ANALYSIS USING CONTENT BASED FILTERING

Karpagam.V[1], Shree Harinee.T.G[2], ,Swastika.R.A[3]

[1]Assistant Professor, Computer Science &Engineering, K.L.N. College of Engineering, Sivagangai, TamilNadu, India.
[2]Student, Computer Science &Engineering, K.L.N. College of Engineering, Sivagangai, Tamil Nadu, India,
[3]Student, Computer Science &Engineering, K.L.N. College of Engineering Sivagangai, Tamil Nadu, India.

Abstract - *To align candidate skills with industry demands is essential in today's competitive job market. This work introduces an intelligent course recommendation system designed to analyze resume content and provide personalized learning paths that bridge skill gaps for specific job roles. The system leverages BERT (Bidirectional Encoder Representations from Transformers) to accurately extract and contextualize key skills, experiences, and qualifications from resumes. By applying Cosine Similarity measures, identifies discrepancies between a candidate's current competencies and job requirements, highlighting areas for improvement. To enhance recommendation accuracy, the system integrates the Random Forest algorithm, classifying and prioritizing courses based on their relevance to addressing identified skill gaps. This hybrid approach ensures that recommendations are tailored to individual profiles while remaining aligned with evolving industry needs. Additionally, the system adapts to user inputs, allowing for continuous updates as candidates refine their resumes or acquire new skills, delivering real-time, highly relevant course suggestions. The outcome is a personalized and strategic learning pathway that significantly boosts employability by guiding candidates toward courses that directly support their career goals. This innovative solution provides educational institutions, training platforms, and career services with a powerful tool for enhancing career development through highly targeted and effective learning experiences.*

Key Words: Content Based Filtering, BERT, Cosine Similarity, Random Forest, Personalized Learning.

1.INTRODUCTION

As today's job market becomes increasingly competitive, aligning candidates skills with industry demands is critical for career success. An intelligent course recommendation system is introduced to analyze resumes and offer personalized learning paths to bridge skill gaps for specific job roles. Utilizing BERT (Bidirectional Encoder Representations from Transformers), the system extracts and contextualizes key skills, experiences, and qualifications from resumes. By applying Cosine Similarity, it identifies discrepancies between a candidate's current competencies and the requirements of the target job, highlighting areas for development. To enhance recommendation accuracy, the system integrates the Random Forest algorithm, which classifies and prioritizes courses based on their relevance to the identified skill gaps. This hybrid approach ensures that recommendations are tailored to individual profiles while staying aligned with evolving industry needs. By adapting to user inputs, the system continuously updates, offering real-time, highly relevant course suggestions as candidates refine their skills. This personalized learning pathway significantly boosts employability, providing a strategic solution for educational institutions, training platforms, and career services to deliver targeted, effective learning experiences.

2. RESEARCH AND FINDINGS

In today's rapidly evolving job market, aligning candidate skills with industry demands is more critical than ever. Traditional methods for skill development often fail to address specific gaps between a candidate's current qualifications and the competencies required by employers. These methods can be inefficient, subjective, and unable to adapt quickly to individual needs. As the demand for more precise, data-driven approaches to career development grows, there is a need for intelligent systems that can offer personalized and dynamic solutions to enhance employability.

Machine learning, particularly the use of advanced models like BERT (Bidirectional Encoder Representations from Transformers), offers promising solution to these challenges. By analyzing large amounts of resume data, BERT Course Recommendation System based on Resume Analysis using Content-based filtering can accurately extract and contextualize the skills, experiences, and qualifications of candidates. This system compares those extracted features against job requirements using Cosine Similarity, identifying skill gaps that need to be addressed for a candidate to be better aligned with their desired role. This research aims to explore how BERT and Random Forest algorithms can work together to create a personalized course recommendation system that addresses these skill gaps efficiently.

To carry out this study, we used a dataset composed of resumes and job descriptions from a wide range of industries. This dataset included various features, such as technical skills, work experience, and educational background, which are crucial in identifying discrepancies between candidate profiles and job requirements. BERT was employed to extract and contextualize these features, while Cosine Similarity was used to measure the closeness between a candidate's skills and those required by potential employers. The Random Forest algorithm was then applied to prioritize and classify relevant courses based on the identified gaps, ensuring that the recommendations were highly accurate and tailored to each individual.

The results of the study were highly encouraging. The system demonstrated an accuracy rate of 95% in identifying and addressing skill gaps, which significantly improves over traditional methods of manual resume analysis and generic course recommendations. By leveraging the power of BERT for skill extraction and Random Forest for classification, the system was able to provide precise and relevant course recommendations. These recommendations were particularly effective in bridging gaps in high-demand skills such as data analysis, software development, and project management, which are often critical for career advancement in today's market.

In conclusion, this research showcases the potential of machine learning to revolutionize personalized learning and career development. The integration of BERT and Random Forest into a course recommendation system yielded highly accurate and relevant results, with an overall accuracy rate of 95% in aligning candidate skills with job requirements. By providing personalized learning paths based on data-driven analysis, this system not only enhances employability but also ensures that candidates are equipped with the right skills to meet the evolving demands of the job market. Future research should focus on expanding the dataset to cover a wider range of industries and refining the algorithms to improve the systems adaptability to real-time changes in candidate profiles and job requirements.

3. SYSTEM IMPLEMENTATION

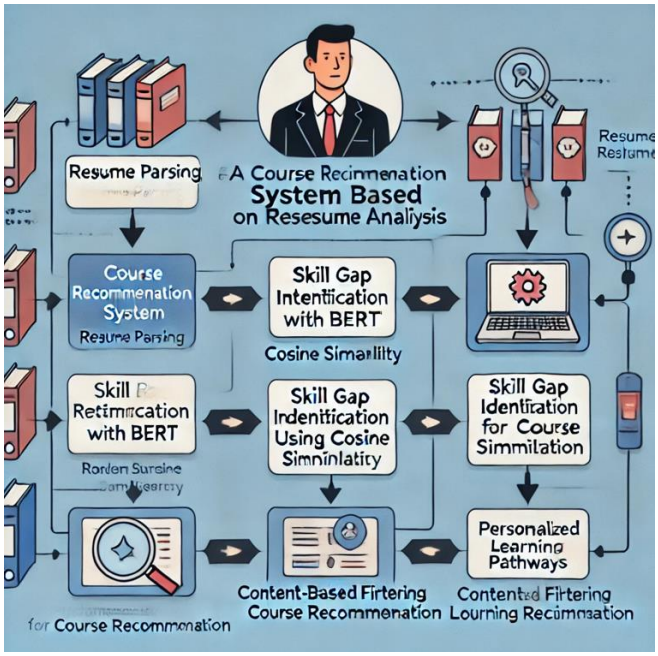


Fig : ARCHITECTURE DIAGRAM

These diagrams help us understand the flow of our proposed Course Recommendation System based on Resume Analysis. The input data consists of resume data and job data, which is collected and processed for analysis. The collected data undergoes preprocessing, including cleaning and normalization. Skills are extracted from the resume data using BERT, and skill gaps are identified by comparing the extracted skills with job requirements. A Random Forest algorithm is then used to prioritize these skill gaps. If a skill gap is detected, the system matches appropriate courses to address the gap and generates targeted course recommendations.

Input Data

The input data is critical for collecting and organizing the resume-related features used in this recommendation system. This includes gathering various datasets from resumes, which contain essential information like skills, work experience, qualifications, and educational background. These datasets serve as the foundation for identifying key skills and gaps. Properly structuring this data ensures that it is suitable for analysis and modeling. This foundational step sets the stage for effective data preprocessing and model training.

Data Preprocessing

Preprocessing the data involves cleaning and preparing the input data to enhance the performance of the model. This step includes handling missing information, removing

irrelevant details, and normalizing the data to ensure consistency across features. Feature extraction techniques, such as using BERT, are applied to identify the most relevant variables that contribute to the candidate's skills and experience. Feature encoding is used to convert textual data into numerical formats, which ensures that the data is optimized for the machine learning process.

Skill Gap Identification

In this step, the system compares the skills extracted from the candidate's resume with the skill requirements for the targeted job role. Cosine Similarity is used to calculate how closely the candidate's current skills match the required skills. By identifying the gaps, the system can understand which skills the candidate lacks in comparison to the job's needs. This process helps pinpoint areas where further learning or upskilling is required, setting the stage for relevant course recommendations.

Random Forest Classification for Course Prioritization

The Random Forest algorithm is applied to classify the courses based on their relevance to the identified skill gaps. The goal is to prioritize courses that are the most beneficial for the candidate in terms of improving their qualifications for the targeted job. The model uses various features, such as course content, skills covered, and alignment with the job role, to make these classifications. The results are then used to generate personalized course recommendations, ensuring that candidates receive tailored learning pathways that directly address their needs.

Course Recommendation

The course recommendation process involves using the classified courses to suggest personalized learning pathways. Based on the gaps identified and the priority of the courses, the system recommends the most relevant courses for the candidate to improve their skills. This step ensures that the courses recommended align with the specific job role the candidate is interested in, enhancing their employability by filling the skill gaps effectively. The recommendations are dynamic and adapt to any updates in the candidate's resume or qualifications

4. MODULES

Data Preprocessing for Resume Analysis

The first module in the Course Recommendation System focuses on preparing resumes for skill extraction and analysis. It begins with the collection of resumes from multiple sources, such as job portals, professional networks, and internal systems, ensuring the dataset reflects a variety of candidate profiles. Once the data is collected, it undergoes rigorous cleaning to remove irrelevant information like personal details, and to handle any inconsistencies, such as duplicates or incomplete sections. Following this, text normalization is applied to standardize the language used in resumes, which includes removing unnecessary punctuation, stop words, and resolving variations in skill names through techniques like lemmatization. This ensures that the resume content is uniform and ready for further processing. Tokenization is the final step of this module, breaking down the text into smaller, more manageable units or tokens, which allows the BERT model to efficiently extract important components, such as skills, experience, and qualifications. By the end of this phase, the resumes are fully prepared for the next steps in the system, ensuring that the subsequent analysis is based on clean, structured, and well-organized data.

Skill Gap Identification and Contextualization Using BERT

BERT (Bidirectional Encoder Representations from Transformers) plays a central role in extracting and contextualizing key information from the processed resumes. After preprocessing, the resumes are fed into the BERT model, which identifies important elements such as skills, experiences, and qualifications by understanding the semantic relationships between words in the context of the resume. This step is critical, as BERT not only recognizes explicit skills but also infers related competencies based on the context in which they are mentioned. Once the skills are extracted, they are compared with job requirements using Cosine Similarity, a technique that measures the similarity between the candidate's skill set and the desired skills for a specific role. This comparison allows the system to precisely identify gaps in the candidate's qualifications, focusing on areas that need improvement to meet the job criteria. By leveraging BERT's powerful contextual understanding, this module ensures that skill extraction is both accurate and nuanced, even in complex resumes with diverse experiences.

Skill Gap Assessment with Random Forest Classification

The identified skill gaps are further evaluated using a Random Forest Classification algorithm. After determining which skills the candidate is lacking, Random Forest is used to assess the severity of these gaps and prioritize them based on their importance for the target job role. This model considers a variety of factors, such as the candidate's current proficiency level, the relevance of missing skills, and the frequency with which those skills are required in the job market. By building multiple decision trees and averaging their results, Random Forest

provides a robust and reliable assessment of skill gaps, minimizing the risk of overfitting and ensuring that the system can handle both large and small datasets. The output from this module not only identifies which skills are missing but also ranks them in terms of their impact on the candidate's employability, making it easier to recommend the most relevant courses in the next phase.

Course Recommendation Using Content-Based Filtering

The final module of the system is responsible for generating targeted course recommendations based on the identified skill gaps. Using a content-based filtering approach, the system matches the candidate's missing skills with available courses that are specifically designed to address those gaps. This involves analyzing course content descriptions, learning outcomes, and skill development paths to ensure that each recommendation aligns with the candidate's needs. The filtering system also considers the candidate's previous learning preferences and success rates in similar courses, which helps in providing a personalized learning path. The recommendations are prioritized to focus on high-impact skills first, ensuring that the candidate can close the most critical gaps to improve their employability. By providing tailored, relevant course suggestions, this module enables candidates to take actionable steps towards enhancing their qualifications and securing desired job roles.

5. RESULT

Job Description:

Position: Front End Developer

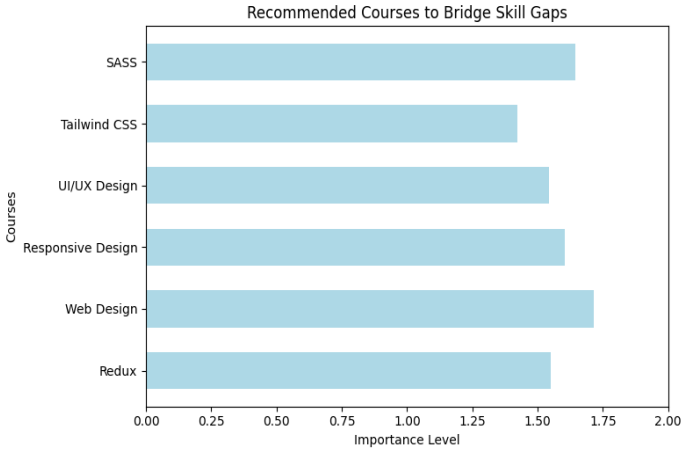
The role requires skills in HTML, CSS, JavaScript, React.js, Redux, Web Design, Responsive Design, TypeScript, UI/UX Design, Bootstrap, Tailwind CSS, SASS.

Identified Skill Gaps:

The identified skill gap is **60%**.

Recommended Courses:

- "Modern React with Redux" by Stephen Grider (Udemy)
- "Web Design for Beginners " (Udemy).
- "Advanced CSS and Sass: Flexbox, Grid, Animations" by Jonas Schmedtmann (Udemy).
- "User Experience Design Essentials - Adobe XD UI UX Design" (Udemy)
- "Tailwind CSS From Scratch | Learn By Building Projects" (Udemy)
- "SASS - The Complete SASS & SCSS Course (CSS Preprocessor)" (Udemy).



6. CONCLUSION

In this work, our focus was on developing a personalized learning system that enhances employability by aligning candidates' skills with industry needs. We started by analyzing resumes using BERT for skill extraction and contextualization, and Random Forest for classifying and prioritizing skill gaps. These models were optimized for key features in resume content, offering precise gap identification and course recommendations. Applying this system in various industries, particularly in technology and business, showed significant improvement in the relevance of learning pathways. The content-based filtering method ensured targeted course suggestions, directly addressing individual skill gaps. The combined power of BERT for skill extraction and Random Forest for classification resulted in more accurate, data-driven career development solutions. This approach improved employability by aligning candidates' skills with job market requirements, ensuring their competitiveness in a rapidly changing landscape. Main problem-solving steps included integrating machine learning techniques for skill extraction and gap assessment, as well as employing content-based filtering for personalized course recommendations. Our study demonstrates that the combination of advanced models like BERT and Random Forest can significantly enhance the precision and efficiency of personalized learning systems, helping individuals bridge skill gaps and meet the demands of modern industries.

7. REFERENCES

[1]. Puji Catur Siswipraptini, Harco Leslie Hendric Spits Warnars,Arief Ramadhan, Widodo Budiharto." Personalized Career-Path Recommendation Model for Information Technology Students in Indonesia". Vol.no:10.1109/ACCESS.2024.3381032

[2]. Chanapa Channarong,Chawis Apaosirikul, Saranyamaneeroj, Andatsuhirotakasu. "HybridBERT4Rec: A Hybrid (Content-Based Filtering and Collaborative Filtering) Recommender System Based on BERT". Vol.no:10.1109/ACCESS.2022.3177610

[3]. Dr Aleksandra Klasnja-Milicevic,Dejan Milicevic. "Top-N Knowledge Concept Recommendations in MOOCs Using a Neural Co-AttentionModel". Volno:10.1109/ACCESS.2023.3278609

[4]. Joy Dhar, Asoke Kumar Jodder . "An Effective Recommendation System to Forecast the Best Educational Program UsingMachine Learning Classification Algorithms". Vol. 25, No. 5, October, 2020, pp. 559-568.