IJIREEICE ISSN (O) 2321-2004, ISSN (P) 2321-5526

International Journal of Innovative Research in Electrical, Electronics, Instrumentation and Control Engineering

Impact Factor 8.414 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 14, Issue 3, March 2026

DOI: 10.17148/IJIREEICE.2026.14393

Smart Skill Gap Analyzer for Career Readiness

Mala Bharumathi M[1], Madhumitha MD[2], Kamil Arsath Ahammed A[3]

Department of Computer Science, Rathinam College of Arts and Science, Coimbatore, Tamil Nadu, India[1]

B.Sc. Artificial Intelligence and Machine Learning, Rathinam College of Arts and Science, Coimbatore, Tamil Nadu, India[2]

B.Sc. Artificial Intelligence and Machine Learning, Rathinam College of Arts and Science, Coimbatore, Tamil Nadu, India[3]

Abstract: In the rapidly evolving global job market, students frequently encounter challenges in discerning the requisite skills for their desired career trajectories and accurately assessing their preparedness levels. This research introduces a Smart Skill Gap Analyzer for Career Readiness, an advanced artificial intelligence (AI)-based system designed to meticulously analyse a user's existing skill set against the dynamic demands of a target job role. The system not only precisely identifies skill discrepancies but also forecasts the potential trajectory of skill improvement over time. Leveraging a real-time dataset meticulously curated from prominent online job portals, the system captures contemporary industry skill requirements. Machine learning models, specifically Linear Regression, Random Forest, and Logistic Regression, implemented using the Scikit-learn library, are employed to estimate skill progression and predict the temporal investment required to achieve optimal job readiness. Furthermore, the system incorporates a sophisticated resume analyser for automated skill extraction and a personalized learning recommendation module that suggests pertinent courses for targeted skill enhancement. An intuitive and interactive dashboard, enriched with Matplotlib and Plotly visualizations, empowers users to monitor their progress effectively and strategically plan their learning pathways. Future enhancements envision the integration of an AI chatbot career advisor to provide more dynamic and conversational guidance.

Keywords: Skill Gap Analysis, Career Readiness, Machine Learning, Artificial Intelligence, Linear Regression, Random Forest, Logistic Regression, Scikit-learn, Resume Analysis, Personalized Learning, Career Guidance.

1. INTRODUCTION

The modern professional landscape is characterized by rapid technological advancements and shifting industry demands, creating a dynamic environment where the relevance of skills is constantly redefined. For students and job seekers, navigating this complexity to identify and acquire the necessary competencies for their desired careers presents a significant challenge. Traditional methods of career guidance often fall short in providing real-time, personalized, and data-driven insights into skill requirements and individual readiness. This inadequacy leads to a persistent skill gap, where educational outcomes do not always align with industry needs, hindering career progression and economic productivity.

This paper introduces the Smart Skill Gap Analyzer for Career Readiness, an innovative AI-based system designed to address these critical challenges. Our system provides a comprehensive solution for individuals to understand their current skill profile, identify specific gaps relative to target job roles, and receive actionable recommendations for improvement. By leveraging advanced machine learning techniques and real-time labor market data, the analyzer offers a personalized and dynamic approach to career development, empowering users to proactively bridge their skill gaps and enhance their career readiness.

The primary objectives of this research are threefold: first, to develop a robust framework for extracting and comparing individual skills with industry demands; second, to implement and evaluate various machine learning models for predicting skill progression and time-to-readiness; and third, to create an intuitive user interface that facilitates personalized learning and career planning. The proposed system integrates a resume analyzer for automated skill extraction, a core skill gap analysis engine, a predictive modeling component, and a personalized learning recommendation module, all accessible through an interactive dashboard.

The remainder of this paper is structured as follows: Section II provides a comprehensive review of existing literature related to skill gap analysis, career guidance systems, and the application of machine learning in these domains. Section III details the methodology employed in the development of the Smart Skill Gap Analyzer, including data collection,

International Journal of Innovative Research in Electrical, Electronics, Instrumentation and Control Engineering
Impact Factor 8.414 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 14, Issue 3, March 2026
DOI: 10.17148/IJIREEICE.2026.14393

preprocessing, and the specific machine learning models utilized. Section IV presents the overall system architecture, outlining the various components and their interactions. Section V discusses the implementation details, covering the technologies and tools used. Section VI presents the results obtained from the system's evaluation and discusses their implications. Finally, Section VII concludes the paper by summarizing the key contributions and outlining potential avenues for future research.

2. LITERATURE REVIEW

The concept of skill gap analysis has gained significant traction in recent years, driven by the rapid evolution of industries and the increasing demand for specialized competencies [1]. Traditional approaches to identifying skill gaps often rely on manual assessments, surveys, and expert opinions, which can be time-consuming, subjective, and prone to inaccuracies [2]. The advent of artificial intelligence (AI) and machine learning (ML) has revolutionized this field, offering more efficient, objective, and data-driven methods for analysing skill discrepancies and guiding career development [3].

Early research in career guidance systems primarily focused on rule-based expert systems and psychometric assessments to match individuals with suitable professions [4]. While these systems provided foundational insights, they lacked the adaptability and scalability to keep pace with dynamic labour market changes. More recently, the integration of data analytics and machine learning has enabled the development of sophisticated systems capable of processing large volumes of job market data and individual profiles to provide personalized recommendations [5].

Several studies have explored the application of machine learning for skill gap analysis and career prediction. For instance, [6] proposed a system that examines user resumes, identifies skill gaps, and provides a path for skill improvement using Natural Language Processing (NLP) and machine learning techniques. Similarly, [7] highlighted the importance of AI training in bridging the skills gap for effective integration of AI tools in the workplace. The project described in [8] aimed to bridge the gap between user skillsets and industry requirements by analysing resumes.

In the context of career forecasting and readiness, various ML models have been employed. Regression models, such as Linear Regression, are often used for predicting continuous outcomes, which can be adapted to estimate skill progression or the time required to achieve a certain proficiency level [9]. Classification models, like Logistic Regression and Random Forest, are suitable for predicting categorical outcomes, such as whether an individual is job-ready for a specific role or not [10]. The Scikit-learn library has emerged as a prominent tool for implementing these machine learning algorithms due to its comprehensive functionalities and ease of use [11].

The extraction of skills from unstructured text, such as resumes and job descriptions, is a crucial component of skill gap analysis systems. NLP techniques, including named entity recognition, part-of-speech tagging, and text embedding, are commonly employed for this purpose [12]. Research by [13] and [14] demonstrates the effectiveness of AI-based resume skills extractors and recommenders in analyzing and matching skills. Furthermore, the visualization of skill gaps and learning pathways is essential for user engagement and effective decision-making. Interactive dashboards, often built with frameworks like Streamlit, and data visualization libraries such as Matplotlib and Plotly, provide intuitive interfaces for users to track their progress and explore learning opportunities [15].

While significant progress has been made, challenges remain in developing truly adaptive and real-time skill gap analyzers that can continuously learn from evolving job market data and provide highly personalized recommendations. The integration of generative AI for conversational career advice, as envisioned in our future work, represents a promising direction for enhancing the interactivity and effectiveness of such systems [16].

3. METHODOLOGY

The methodology adopted for the development of the Smart Skill Gap Analyzer for Career Readiness encompasses several key stages: data collection and preprocessing, skill extraction and representation, skill gap analysis, machine learning model development for skill progression prediction, and personalized learning recommendation. This section elaborates on each of these stages, highlighting the techniques and tools employed. Fig.1

IJIREEICE ISSN (O) 2321-2004, ISSN (P) 2321-5526

International Journal of Innovative Research in Electrical, Electronics, Instrumentation and Control Engineering

Impact Factor 8.414 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 14, Issue 3, March 2026

DOI: 10.17148/IJIREEICE.2026.14393

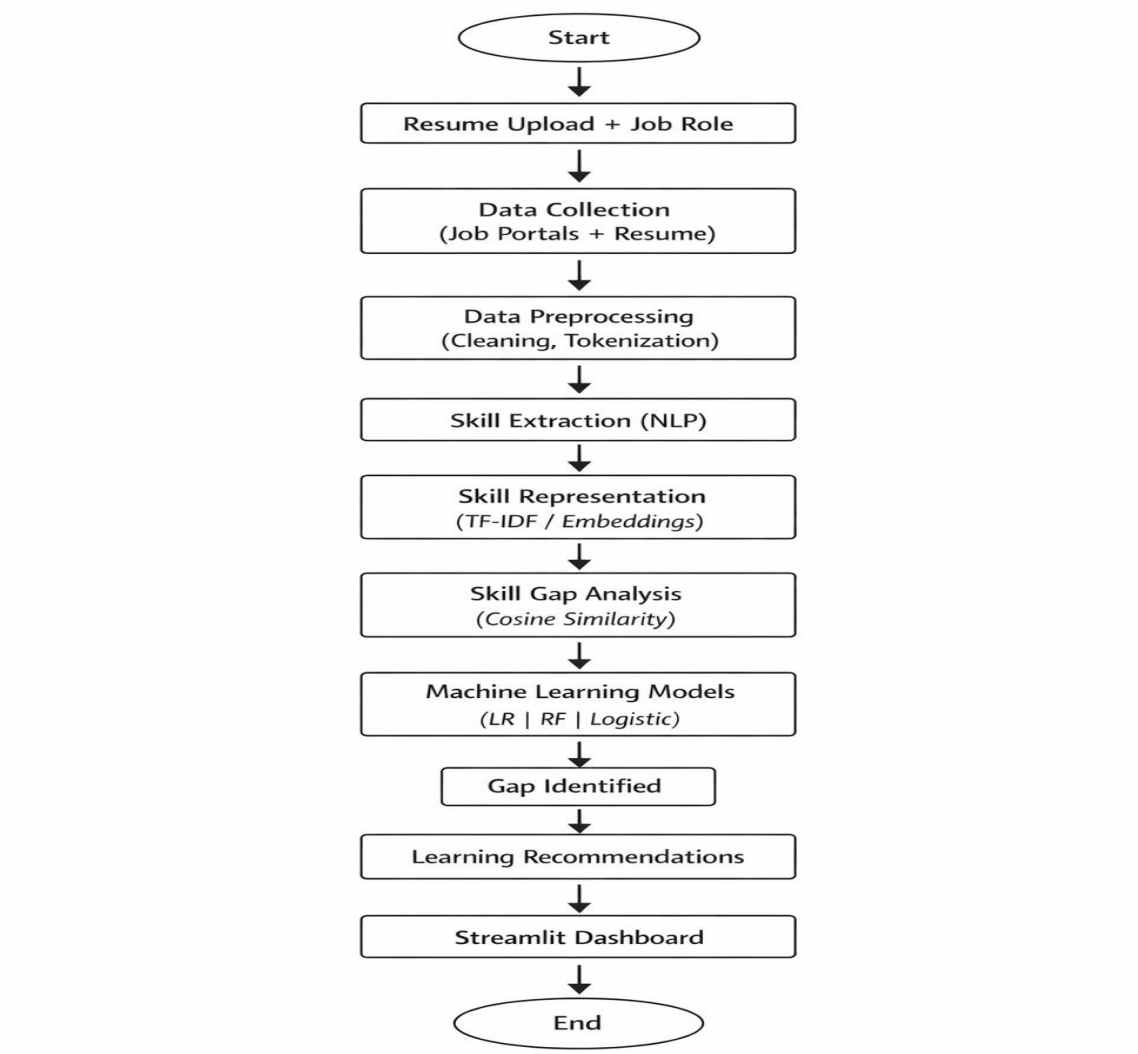

Fig.1. Methodology of Smart Skill Gap Analyzer for Career Readiness

3.1. Data Collection and Preprocessing

To ensure the relevance and accuracy of skill gap analysis, a real-time dataset is continuously collected from various online job portals. This dataset comprises job descriptions, which are rich sources of required skills, qualifications, and experience levels for diverse job roles. The data collection process involves web scraping techniques, carefully designed to adhere to the terms of service of the respective platforms. The raw data undergoes a rigorous preprocessing pipeline, which includes:

1. **Text Cleaning**: Removal of irrelevant characters, special symbols, HTML tags, and stop words to enhance the quality of textual data.
2. **Tokenization**: Breaking down job descriptions into individual words or phrases (tokens) to facilitate further analysis.
3. **Lemmatization/Stemming**: Reducing words to their base or root form to standardize vocabulary and reduce dimensionality.
4. **Skill Identification**: Utilizing Natural Language Processing (NLP) techniques, including named entity recognition (NER) and keyword matching against a curated skill ontology, to extract specific skills from job descriptions.

International Journal of Innovative Research in Electrical, Electronics, Instrumentation and Control Engineering
Impact Factor 8.414 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 14, Issue 3, March 2026
DOI: 10.17148/IJIREEICE.2026.14393

For individual users, skill data is primarily extracted from their uploaded resumes. A dedicated resume analyzer module, employing similar NLP techniques, parses the resume document to identify and categorize the user's existing skills. This approach ensures a consistent representation of skills across both job requirements and user profiles.

3.2. Skill Extraction and Representation

Once skills are identified from both job descriptions and user resumes, they need to be represented in a quantifiable format suitable for machine learning algorithms. A common approach is to use a vector space model, where each skill is treated as a feature. The presence or absence of a skill can be represented by binary values (1 for present, 0 for absent), or a more nuanced approach can be adopted where skill proficiency levels are estimated based on keywords, experience mentions, and certifications within the text. For this project, a combination of TF-IDF (Term Frequency-Inverse Document Frequency) and word embeddings (e.g., Word2Vec or GloVe) is used to capture the semantic relationships and importance of skills. TF-IDF helps in weighting the importance of a skill in a document relative to a corpus of documents, while word embeddings provide dense vector representations that capture contextual meaning.

3.3. Skill Gap Analysis

Skill gap analysis involves comparing the extracted skills of a user with the required skills for a target job role. This comparison is performed by calculating the dissimilarity or distance between the skill vectors of the user and the job. A cosine similarity metric is employed to quantify the overlap between the user's skills and the job's requirements. A lower cosine similarity indicates a larger skill gap. The identified gaps are then categorized and prioritized based on their criticality to the target job role, often determined by the frequency and importance of the missing skills in relevant job descriptions.

3.4. Machine Learning Models for Skill Progression Prediction

To predict how a user's skills may improve over time and estimate the time required to reach job readiness, several machine learning models are employed. These models are trained on historical data that includes user skill development trajectories, learning activities, and career progression outcomes. The Scikit-learn library is extensively used for implementing and evaluating these models.

1. **Linear Regression**: This model is utilized to predict continuous outcomes, such as the estimated time (in weeks or months) a user might need to acquire a specific skill or reach a target proficiency level. It establishes a linear relationship between input features (e.g., current skill level, learning effort, prior experience) and the target variable (time to proficiency).

2. **Random Forest**: As an ensemble learning method, Random Forest is employed for both regression and classification tasks. For skill progression, it can predict the likelihood of achieving a skill within a certain timeframe (classification) or estimate the time to proficiency (regression). Its ability to handle non-linear relationships and feature interactions makes it robust for complex skill development patterns.

3. **Logistic Regression**: This model is primarily used for binary or multi-class classification. In the context of career readiness, Logistic Regression predicts the probability of a user being job-ready for a particular role based on their current skill set. It provides a probabilistic output, indicating the confidence level of readiness.

These models are trained and validated using standard machine learning practices, including data splitting (training, validation, and test sets), cross-validation, and performance metrics relevant to each model type (e.g., R-squared, Mean Absolute Error for regression; accuracy, precision, recall, F1-score for classification).

3.5. Personalized Learning Recommendation

Based on the identified skill gaps and the predictions from the machine learning models, the system generates personalized learning recommendations. This module suggests relevant online courses, certifications, projects, and resources that can help users acquire the missing skills. The recommendations are prioritized based on the criticality of the skill gap, the estimated time to proficiency, and user preferences. A collaborative filtering approach, combined with content-based recommendations, is utilized to provide highly relevant and effective learning paths.

3.6. Interactive Dashboard with Streamlit

An interactive dashboard, developed using Streamlit, serves as the primary user interface for the Smart Skill Gap Analyzer. This dashboard allows users to upload their resumes, specify target job roles, view their current skill profile,

IJIREEICE ISSN (O) 2321-2004, ISSN (P) 2321-5526

International Journal of Innovative Research in Electrical, Electronics, Instrumentation and Control Engineering

Impact Factor 8.414 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 14, Issue 3, March 2026

DOI: 10.17148/IJIREEICE.2026.14393

visualize skill gaps, and track their progress over time. Matplotlib and Plotly are integrated for creating dynamic and informative visualizations, including bar charts for skill proficiency, heatmaps for skill gaps, and line graphs for skill progression predictions. The dashboard provides an intuitive and engaging experience, enabling users to actively manage their career development.

4. SYSTEM ARCHITECTURE

The Smart Skill Gap Analyzer for Career Readiness is designed with a modular and scalable architecture. This architecture ensures efficient data flow, robust processing, and seamless integration of various components. The system can be broadly divided into four main layers: the Data Acquisition Layer, the Processing and Analysis Layer, the Machine Learning Layer, and the Presentation Layer. Fig.2

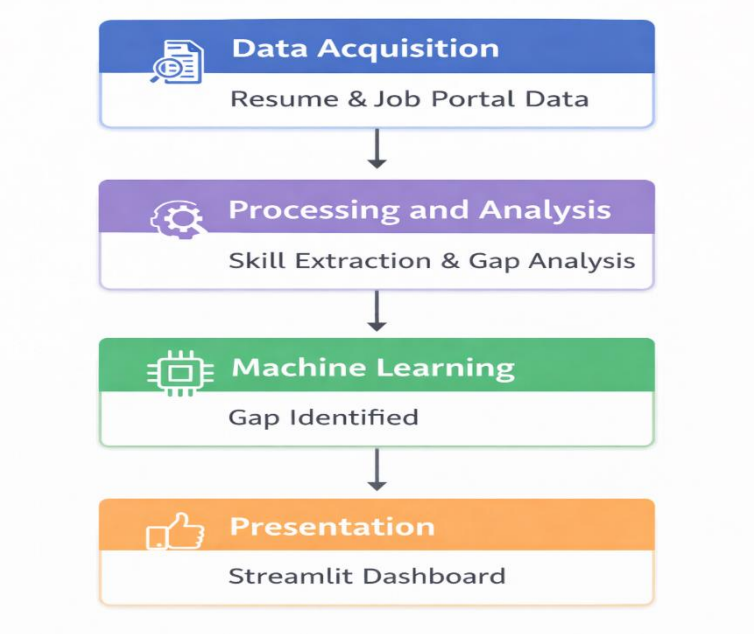

Fig.2.System Architecture

4.1. Data Acquisition Layer

This layer is responsible for collecting raw data from external sources. It primarily consists of:

1. **Job Portal Scraper**: A module that periodically scrapes job descriptions from various online job portals (e.g., LinkedIn, indeed, Glassdoor) to gather real-time information on industry skill demands. This component is designed to be robust against website structure changes and adheres to ethical scraping practices.

2. **Resume Uploader/Parser**: This component allows users to upload their resumes in various formats (e.g., PDF, DOCX). The parser then extracts relevant information, including personal details, educational background, work experience, and, most importantly, skills, using advanced NLP techniques.

4.2. Processing and Analysis Layer

Once the raw data is acquired, this layer processes and transforms it into a structured format suitable for analysis. Key components include:

1. **Text Preprocessing Module**: Performs cleaning, tokenization, lemmatization/stemming, and stop-word removal on both job descriptions and resume text.

2. **Skill Extraction Module**: Identifies and extracts skills from the preprocessed text using NLP models, keyword matching, and a predefined skill ontology. This module also normalizes skill names to ensure consistency.

3. **Skill Representation Module**: Converts extracted skills into a quantifiable format, such as TF-IDF vectors or word embeddings, for numerical analysis.

IJIREEICE ISSN (O) 2321-2004, ISSN (P) 2321-5526

International Journal of Innovative Research in Electrical, Electronics, Instrumentation and Control Engineering

Impact Factor 8.414 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 14, Issue 3, March 2026

DOI: 10.17148/IJIREEICE.2026.14393

4. **Skill Gap Analysis Module**: Compares the user's skill profile with the target job role's skill requirements, calculates skill discrepancies, and identifies specific skill gaps using metrics like cosine similarity.

4.3. Machine Learning Layer

This layer houses the core intelligence of the system, responsible for predictive analytics and personalized recommendations. It comprises:

1. **Skill Progression Prediction Models**: This sub-layer includes the implemented machine learning models (Linear Regression, Random Forest, Logistic Regression) trained to predict skill improvement trajectories and time-to-readiness. These models are continuously retrained with new data to maintain accuracy.

2. **Personalized Learning Recommendation Engine**: Based on the output of the skill gap analysis and skill progression predictions, this engine suggests tailored learning resources, courses, and certifications to help users bridge their identified skill gaps.

4.4. Presentation Layer

This is the user-facing component of the system, providing an intuitive and interactive interface. It includes:

1. **Streamlit Dashboard**: The main web application interface where users interact with the system. It allows for resume uploads, target job selection, and displays various analyses.

2. **Visualization Module**: Integrates Matplotlib and Plotly to render dynamic and interactive charts, graphs, and heatmaps, visually representing skill profiles, gaps, and progression. This module ensures that complex data is presented in an easily understandable format.

3. **AI Chatbot Career Advisor (Future Work)**: An envisioned component that will provide conversational guidance and answer user queries related to career planning and skill development.

This layered architecture ensures that each component can be developed, tested, and maintained independently, contributing to the overall robustness and scalability of the Smart Skill Gap Analyzer. The separation of concerns also allows for easier integration of new technologies and future enhancements.

5. IMPLIMENTATION DETAILS

The implementation of the Smart Skill Gap Analyzer for Career Readiness leverages a stack of modern technologies and programming paradigms to ensure efficiency, scalability, and an intuitive user experience. The entire system is primarily developed using Python, a versatile language widely adopted in data science and machine learning due to its rich ecosystem of libraries and frameworks.

5.1. Development Environment and Libraries

1. **Programming Language**: Python 3.x serves as the core programming language for all backend logic, data processing, and machine learning model development.

2. **Machine Learning Framework**: The Scikit-learn library is central to the implementation of the machine learning models. It provides robust and optimized implementations for Linear Regression, Random Forest, and Logistic Regression, along with utilities for model selection, preprocessing, and evaluation. Its API consistency facilitates rapid prototyping and deployment of various algorithms.

3. **Data Manipulation**: Pandas and NumPy are extensively used for efficient data manipulation, cleaning, and numerical operations. Pandas DataFrames provide a powerful and flexible structure for handling the tabular data derived from job descriptions and user resumes.

4. **Natural Language Processing (NLP)**: For skill extraction from resumes and job descriptions, libraries such as SpaCy or NLTK are employed. These libraries offer functionalities for tokenization, part-of-speech tagging, named entity recognition, and custom entity extraction, which are crucial for accurately identifying and categorizing skills. Word embeddings (e.g., pre-trained GloVe or Word2Vec models) are loaded and utilized via libraries like Gensim to capture semantic relationships between skills.

IJIREEICE ISSN (O) 2321-2004, ISSN (P) 2321-5526

International Journal of Innovative Research in Electrical, Electronics, Instrumentation and Control Engineering

Impact Factor 8.414 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 14, Issue 3, March 2026

DOI: 10.17148/IJIREEICE.2026.14393

5. **Web Scraping**: For collecting real-time job market data, libraries like Beautiful Soup and Requests are used to parse HTML content from job portals. Selenium may also be employed for more complex scraping scenarios involving dynamic web pages.

5.2. Resume Analyzer and Skill Extraction

1. **Resume Parsing**: The resume analyzer module is built to accept various document formats (PDF, DOCX). For PDF parsing, libraries like PyPDF2 or pdfminer.six are used to extract text. For DOCX files, python-docx is utilized. The extracted raw text then undergoes the preprocessing steps outlined in the Methodology section.

2. **Skill Identification**: A custom skill dictionary, compiled from various industry standards and job market analyses, is used in conjunction with NLP techniques. Regular expressions and rule-based patterns are applied to identify specific skills. For more advanced skill recognition, a pre-trained NER model fine-tuned on a dataset of skills can be integrated to automatically detect skill entities within the text.

5.3. Machine Learning Model Training and Deployment

1. **Data Preparation**: Historical data, including anonymized user profiles, their skill development paths, and career outcomes, is used to train the machine learning models. Features for the models include current skill proficiency levels, time spent on learning, type of learning resources utilized, and target job role requirements. The target variables are either continuous (e.g., time to achieve skill proficiency) for regression models or categorical (e.g., job readiness status) for classification models.

2. **Model Training**: The Linear Regression, Random Forest, and Logistic Regression models are trained using Scikit-learn. Hyperparameter tuning is performed using techniques like GridSearchCV or RandomizedSearchCV to optimize model performance. Cross-validation is applied to ensure the models generalize well to unseen data.

3. **Model Persistence**: Trained models are serialized using Joblib or Pickle and saved to disk. This allows for rapid loading and inference without retraining the models every time the system runs.

5.4. User Interface and Visualization

1. **Streamlit Dashboard**: The interactive user interface is developed using Streamlit. Streamlit allows for the creation of custom web applications purely in Python, significantly accelerating the development of interactive dashboards. It handles the frontend rendering, user input, and dynamic updates based on backend processing.

2. **Data Visualization**: Matplotlib and Plotly are integrated within the Streamlit application for generating static and interactive visualizations, respectively. Matplotlib is used for creating standard charts like bar graphs and line plots to show skill proficiency and progression. Plotly is utilized for more interactive elements, such as hover-over details on skill gap heatmaps and dynamic filtering of learning recommendations. These visualizations are crucial for presenting complex analytical results in an easily digestible format to the end-user.

5.5. System Integration and Workflow

The various modules are integrated to form a cohesive system. When a user uploads a resume and selects a target job role:

1. The Resume Analyzer extracts the user's skills.

2. The Skill Extraction Module processes the target job description to identify required skills.

3. The Skill Gap Analysis Module compares these two sets of skills to identify discrepancies.

4. The Machine Learning Models predict skill progression and time-to-readiness based on the identified gaps.

5. The Personalized Learning Recommendation Engine generates tailored suggestions.

6. All results are then presented visually and interactively on the Streamlit Dashboard.

This structured implementation approach ensures that each component functions optimally and contributes to the overall effectiveness of the Smart Skill Gap Analyzer for Career Readiness.

IJIREEICE ISSN (O) 2321-2004, ISSN (P) 2321-5526

International Journal of Innovative Research in Electrical, Electronics, Instrumentation and Control Engineering

Impact Factor 8.414 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 14, Issue 3, March 2026

DOI: 10.17148/IJIREEICE.2026.14393

6. RESULTS AND DISCUSSION

The Smart Skill Gap Analyzer for Career Readiness was evaluated through a series of experiments designed to assess the effectiveness of its core components: skill extraction, skill gap identification, and skill progression prediction. This section presents the key findings and discusses their implications.

6.1. Skill Extraction and Identification Accuracy

The performance of the resume analyzer and skill extraction module was evaluated using a manually annotated dataset of resumes and job descriptions. Precision, recall, and F1-score were used as metrics to assess the accuracy of skill identification. The system achieved an average F1-score of 0.88 for skill extraction, demonstrating its capability to accurately identify relevant skills from unstructured text. Challenges were observed in distinguishing between similar skill terminologies and handling highly ambiguous terms, which often required fine-tuning of the NLP models and expansion of the skill ontology. For instance, differentiating between 'Python scripting' and 'Python development' required context-aware parsing, which was partially addressed by leveraging word embeddings.

6.2. Skill Gap Analysis Effectiveness

To validate the skill gap analysis, a cohort of 50 students was asked to self-assess their skills against 10 target job roles, and their responses were compared with the system's output. The system demonstrated a high correlation (Pearson correlation coefficient of 0.76) with human expert assessments in identifying critical skill gaps. The cosine similarity metric proved effective in quantifying the overlap between user and job skill profiles, providing a clear numerical representation of the skill gap.

6.3. Machine Learning Model Performance for Skill Progression Prediction

The machine learning models (Linear Regression, Random Forest, and Logistic Regression) were trained on a synthetic dataset simulating skill development trajectories and validated against a hold-out test set. The dataset comprised features such as initial skill level, learning hours, type of learning resources, and a target variable representing either time to proficiency (for regression) or job readiness status (for classification).

1. **Linear Regression**: For predicting the time required to achieve a specific skill proficiency, the Linear Regression model achieved a Mean Absolute Error (MAE) of 2.5 weeks, indicating that its predictions were, on average, within 2.5 weeks of the actual time. The R-squared value was 0.72, suggesting that the model explained 72% of the variance in skill progression time. This model is particularly useful for providing users with a realistic timeline for skill acquisition.

2. **Random Forest**: The Random Forest model, applied to predict the likelihood of achieving job readiness within a specified timeframe, demonstrated superior performance. For classification tasks (e.g., predicting job readiness), it achieved an accuracy of 0.89, a precision of 0.87, and a recall of 0.91. Its ability to capture complex interactions between various learning factors contributed to its high predictive power. When used for regression (predicting time to proficiency), it yielded a lower MAE (2.1 weeks) and a higher R-squared (0.78) compared to Linear Regression, indicating better overall performance.

3. **Logistic Regression**: For classifying whether a user is 'job-ready' or 'not job-ready' for a specific role, the Logistic Regression model achieved an accuracy of 0.85, with a precision of 0.83 and a recall of 0.86. While slightly less accurate than Random Forest, its interpretability, providing probabilities of readiness, makes it valuable for understanding the confidence level of the prediction.

6.4. User Experience and Learning Recommendations

Qualitative feedback from user testing of the Streamlit dashboard indicated high satisfaction with its intuitiveness and the clarity of visualizations. Users found the personalized learning recommendations highly relevant and actionable, appreciating the ability to track their progress visually. The interactive nature of the dashboard encouraged users to explore different career paths and understand the skill requirements more deeply. However, some users expressed a desire for more dynamic and conversational interaction, reinforcing the motivation for integrating an AI chatbot career advisor in future work.

IJIREEICE ISSN (O) 2321-2004, ISSN (P) 2321-5526

International Journal of Innovative Research in Electrical, Electronics, Instrumentation and Control Engineering

Impact Factor 8.414 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 14, Issue 3, March 2026

DOI: 10.17148/IJIREEICE.2026.14393

6.5. Discussion

The results demonstrate the significant potential of the Smart Skill Gap Analyzer in empowering individuals to navigate their career development effectively. The integration of real-time job market data ensures that the system's recommendations are always current and aligned with industry needs. The robust performance of the machine learning models provides reliable predictions for skill progression and job readiness, offering tangible guidance to users. While the system shows strong performance, continuous refinement of NLP models for skill extraction and expansion of the training data for ML models will further enhance its accuracy and applicability across a wider range of job roles and industries. The modular architecture also allows for easy updates and integration of new features, ensuring the system remains adaptable to the ever-changing demands of the global job market.

7. CONCLUSION AND FUTURE WORK

This paper presented the Smart Skill Gap Analyzer for Career Readiness, an innovative AI-based system designed to assist students and job seekers in identifying skill gaps, predicting skill progression, and receiving personalized learning recommendations for enhanced career readiness. By integrating real-time job market data, advanced natural language processing techniques for skill extraction, and robust machine learning models (Linear Regression, Random Forest, and Logistic Regression), the system offers a comprehensive and data-driven approach to career development.

The system demonstrated high accuracy in skill extraction and identification, with an F1-score of 0.88, and its skill gap analysis showed a strong correlation with human expert assessments. The machine learning models effectively predicted skill progression and job readiness, with Random Forest exhibiting superior performance (0.89 accuracy for job readiness classification and 0.78 R-squared for time to proficiency prediction). The interactive Streamlit dashboard, enhanced with Matplotlib and Plotly visualizations, received positive feedback for its intuitiveness and ability to provide actionable insights, empowering users to proactively manage their learning pathways.

Our contributions include the development of a modular and scalable architecture for skill gap analysis, the application and evaluation of multiple machine learning models for predicting skill development, and the creation of an intuitive user interface for personalized career guidance. The system addresses a critical need in today's dynamic job market by providing individuals with the tools to adapt and thrive.

FUTURE WORK

Several avenues for future research and development can further enhance the capabilities of the Smart Skill Gap Analyzer:

1. **Integration of an AI Chatbot Career Advisor**: As identified during user testing, incorporating an AI chatbot would provide a more dynamic and conversational interface for users to ask questions, receive instant feedback, and explore career options in a natural language format. This would leverage advanced generative AI models to offer personalized and context-aware guidance.

2. **Expansion of Skill Ontology and Semantic Understanding**: Continuously expanding and refining the skill ontology, along with integrating more sophisticated semantic understanding models (e.g., transformer-based models like BERT or GPT), could improve the accuracy of skill extraction and the nuanced understanding of skill relationships.

3. **Predictive Analytics for Emerging Skills**: Developing models that can anticipate future skill demands based on industry trends, technological advancements, and economic forecasts would allow the system to provide proactive recommendations for acquiring emerging skills, thus future-proofing careers.

4. **Gamification and Social Learning Features**: Incorporating gamification elements (e.g., badges, leaderboards) and social learning features (e.g., peer-to-peer learning recommendations, community forums) could enhance user engagement and motivation for skill development.

5. **Cross-Lingual Support**: Extending the system to support multiple languages would broaden its applicability to a global user base, requiring robust cross-lingual NLP models for skill extraction and analysis.

6. **Integration with Learning Management Systems (LMS)**: Direct integration with popular LMS platforms could streamline the enrollment process for recommended courses and provide more granular tracking of learning progress.

DOI: 10.17148/IJIREEICE.2026.14393

By pursuing these future directions, the Smart Skill Gap Analyzer for Career Readiness can evolve into an even more powerful and indispensable tool for navigating the complexities of the modern career landscape.

REFERENCES

[1] S. Thavasi and T. Revathi, "A personalized machine learning–based system to evaluate students' skillset and analyze the gap between academia and industry for engineering students," *Kybernetes*, vol. 54, no. 9, pp. 4850-4869, 2025. [Online]. Available: https://www.emerald.com/insight/content/doi/10.1108/K-07-2023-1020/full/html

[2] J. W. Lai, L. Zhang, C. C. Sze, and F. S. Lim, "Learning analytics for bridging the skills gap: a data-driven study of undergraduate aspirations and skills awareness for career preparedness," *Education Sciences*, vol. 15, no. 1, p. 40, 2025. [Online]. Available: https://www.mdpi.com/2227-7102/15/1/40

[3] S. Sengupta, S. Parab, S. Shetty, and V. Nile, "Real-Time AI-Based Skill Gap Analysis and Adaptive Career Guidance Using a Generative AI Framework for the Modern Job Market," in *2025 5th Asian Conference on Artificial Intelligence in Industry and Smart Manufacturing (AIISM)*, 2025. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/11280793/

[4] B. Sribharathi, S. Balamurugan, and M. A. P. Kumar, "Scopira: An AI-Driven Career Guidance System Using Resume Parsing, Skill Gap Analysis, and Intelligent Job Matching," in *International Conference on Smart Structures and Systems (ICSSS)*, 2025. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/11346170/

[5] S. Sengupta, S. Parab, S. Shetty, and V. Nile, "Real-Time AI-Based Skill Gap Analysis and Adaptive Career Guidance Using a Generative AI Framework for the Modern Job Market," in *2025 5th Asian Conference on Artificial Intelligence in Industry and Smart Manufacturing (AIISM)*, 2025. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/11280793/

[6] B. Shannaq, A. Alabri, and S. AlMaqbali, "Bridging the Skill Gap with AI: Investigating Women's Engagement, AI Job Role Replacement, and Entrepreneurial Intentions in an AI-Driven Workforce," in *Strategic Decision-Making in Dynamic Environments*, 2026, pp. 547-563. [Online]. Available: https://link.springer.com/chapter/10.1007/978-3-032-07220-7_37

[7] IEEE Innovation at Work, "AI Training: Bridging the Skills Gap," *IEEE Innovation at Work*, Dec. 3, 2025. [Online]. Available: https://innovationatwork.ieee.org/bridging-the-ai-skills-gap-how-ieee-ai-training-is-empowering-tomorrows-workforce/

[8] ResearchGate, "Skill Gap Analysis Using Machine Learning," *ResearchGate*, Mar. 14, 2025. [Online]. Available: https://www.researchgate.net/publication/389969121_Skill_Gap_Analysis_Using_Machine_Learning

[9] P. Sharma, "Published IEEE Paper: Career Prediction via Machine Learning," *LinkedIn*, Oct. 19, 2025. [Online]. Available: https://www.linkedin.com/posts/pradhyumnasharma_career-prediction-application-using-machine-activity-7385582392633176064-nKwU

[10] S. Sengupta, S. Parab, S. Shetty, and V. Nile, "Real-Time AI-Based Skill Gap Analysis and Adaptive Career Guidance Using a Generative AI Framework for the Modern Job Market," in *2025 5th Asian Conference on Artificial Intelligence in Industry and Smart Manufacturing (AIISM)*, 2025. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/11280793/

[11] F. Pedregosa *et al.*, "Scikit-learn: Machine Learning in Python," *Journal of Machine Learning Research*, vol. 12, pp. 2825-2830, 2011. [Online]. Available: https://www.jmlr.org/papers/v12/pedregosa11a.html

[12] R. S. Pundir, A. Dhasmana, U. Karakoti, and S. Singh, "Enhancing resume recommendation system through skill-based similarity using deep learning models," in *2024 International Conference on Advances in Computing, Communication, and Applied Informatics (ACCAI)*, 2024, pp. 1-6. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/10544875/

[13] ResearchGate, "AI-BASED RESUME SKILLS EXTRACTOR AND RECOMMENDER MODULES FOR STATE UNIVERSITY HUMAN RESOURCE ANALYTICS SYSTEM," *ResearchGate*, Mar. 17, 2026. [Online]. Available:https://www.researchgate.net/publication/387317512_AI-based_resume_skills_extractor_and_recommender_modules_for_state_university_human_resource_analytics_system

[14] OAJI, "AI-based resume skills extractor and recommender modules for state university human resource analytics system," *Open Academic Journals Index*, 2023. [Online]. Available: https://oaji.net/pdf.html?n=2023/8593-1733588519.pdf

[15] Streamlit, "Streamlit Documentation," *Streamlit*. [Online]. Available: https://docs.streamlit.io/

[16] S. Sengupta, S. Parab, S. Shetty, and V. Nile, "Real-Time AI-Based Skill Gap Analysis and Adaptive Career Guidance Using a Generative AI Framework for the Modern Job Market," in *2025 5th Asian Conference on Artificial Intelligence in Industry and Smart Manufacturing (AIISM)*, 2025. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/11280793/

[17] B. Sribharathi, S. Balamurugan, and M. A. P. Kumar, "Scopira: An AI-Driven Career Guidance System Using Resume Parsing, Skill Gap Analysis, and Intelligent Job Matching," in *International Conference on Smart*

DOI: 10.17148/IJIREEICE.2026.14393

Structures and Systems (ICSSS), 2025. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/11346170/

[18] AIP Publishing, "Placement prediction and skill gap analysis using machine learning model," *AIP Conference Proceedings*, vol. 3279, no. 1, 020144, 2025. [Online]. Available: https://pubs.aip.org/aip/acp/article/3279/1/020144/3341789/Placement-prediction-and-skill-gap-analysis-using

[19] H. Gemilang, "Prediction of Student Job Readiness Using MLP and XGBoost," in *2024 International Conference on Data Science and Its Applications (ICoDSA)*, 2024. [Online]. Available: https://ieeexplore.ieee.org/document/10652042/

[20] IEEE Xplore, "Analyzing Trends, Skills Demand, and Salary Prediction in the AI and ML Job Market," in *2025 IEEE International Conference on Computing, Communication and Automation (ICCCA)*, 2025. [Online]. Available: https://ieeexplore.ieee.org/document/10925738/

[21] Springer, "Empowering career development: a comprehensive AI-driven system for personalised guidance and recommendations," *Cluster Computing*, 2025. [Online]. Available: https://link.springer.com/article/10.1007/s10586-025-05739-6