RESEARCH PAPER

A Systematic Review of Skill Monitoring and Career Roadmap Prediction Using Computational Intelligence

Sumit Mittu[1], Lovi Raj Gupta[2], Gulshan Kumar[3]
[1]School of Computer Science and Engineering, Lovely Professional University, India
Email: sumit.mittu@gmail.com
[2]School of Computer Science and Engineering, Lovely Professional University, India
Email: lovi.raj@lpu.co.in
[3]School of Computer Science and Engineering, Lovely Professional University, India
Email: gulshan.kumar@lpu.co.in
Received: 25th Aug, 2026 | Revised: 1st Sep, 2026 | Accepted: 5th Sep, 2026 | Available Online: 7th Sep, 2026

How to cite this article: Mittu S, Gupta LR, Kumar G., A Systematic Review of Skill Monitoring and Career Roadmap Prediction Using Computational Intelligence. Int J Drug Deliv Technol. 2026;16(82s): 226-237; DOI: 10.25258/ijddt.16.82s.24

Abstract—The fast-paced evolution of job markets and the dynamic changes in the skill requirements by the employers have urged a need for intelligent systems that can support skill monitoring, predicting career roadmap, and job recommendation. To address these challenges, diverse computational intelligence techniques (including machine learning, natural language processing, and recommendation systems) have been proposed over the last decade. In this paper, a systematic review of the existing approaches in the areas such as skill-assessment, matching of resume with the job description, career prediction, and personalized recommendation system is being presented. The studies reviewed have been classified on the basis of data sources, skill representation approaches, modeling techniques, and evaluation metrics. A comparative analysis highlights the strengths and limitations of current methods. Finally, open research challenges and future research directions are identified to guide the development of scalable, adaptive, and explainable career guidance systems.

Index Terms—Skill Monitoring, Career Roadmap Prediction, Job Recommendation, Computational Intelligence

I. Introduction

Technological advancements and rapidly changing industry requirements have significantly transformed the global job market, especially after the COVID-19. According to the Reserve Bank of India (RBI) data , for the number of graduates passing out, the number of jobs created, and the number of vacancies filled in India for the recent years, it is observed that while the number of graduates passing out annually has been increasing slightly, the number of jobs created has been growing at a faster rate. However, the number of vacancies filled has not kept pace with the number of jobs created, indicating a gap between job creation and actual employment [1], [2]. As per these reports, In India, during the year 2021, 1.07 Cr graduates passed-out and 4.66 Cr jobs were created of which only 2.00 Cr vacancies were filled. The numbers slightly rose to 1.08 Cr graduates passed-out with 4.70 Cr job created and 2.05 Cr vacancies filled, in 2022. Similar trends were observed for the year 2023 and for year 2024, 1.10 Cr graduates passed-out, 4.80 Cr jobs created while only 2.15 Cr vacancies were filled [1], [2].

While job creation has increased, the employability gaps still persist due to mismatches between acquired skills and industry demands, affected by the changing landscape of education as well as careers due to the advent of Gen AI. Traditional career guidance systems lack the ability to dynamically assess skills, predict career trajectories, and provide personalized recommendations. As per the India Skills Report 2024, several reasons are reported in [3], [4] for the gap between job creation and the actually filled positions, in India. The prominent reasons are stated :

1) Skills Mismatch: Many graduates lack the skills required by employers, leading to a mismatch between the supply of labor and the demand for specific skills. This is often referred to as the ”skills deficit”.
2) Quality of Jobs: A significant portion of the jobs created are low-skilled and low-paying, which may not be attractive to many job seekers. This results in vacancies remaining unfilled despite job creation.
3) Geographic Mismatch: Jobs may be created in urban areas, while job seekers are located in rural areas, leading to a geographic mismatch.
4) Informal Sector: The informal sector employing large part of the Indian workforce, often lacks job security and other employee benefits, making formal sector jobs more desirable choice for candidates.
5) Economic Slowdown: Economic slowdowns and uncertainties can lead to companies being cautious about hiring, even when vacancies exist.
6) Automation and Technology: The increased automation & technology has led to displacement of traditional jobs, creating a gap between the number of jobs created and those filled.

To answer such concerns, a multi-fold approach is required, that includes improving education and training programs, promoting job creation in diverse sectors, and ensuring better job quality. Recent research has leveraged computational intelligence techniques such as machine learning, natural language processing, and recommender systems to address these limitations but these widely fragmented solutions are evaluated under non-homogeneous environments, and often lack in terms of scalability and explainability.In this paper, a systematic review of the skill monitoring frameworks and the frameworks for career roadmap prediction is presented and consolidated to identify the gaps in existing research.

This review contributes by focusing on the following aspects:

- Skill-monitoring Taxonomy
- Approaches for Career Roadmap prediction
- Comparatively analyzing the datasets, modeling techniques, and evaluation approaches.
- Identify open challenges or research gaps and explore future research perspectives.

II. Review Methodology

This systematic literature review analyzes relevant studies published during the past decade (from year 2015 to 2024). The focus of this systematic review was to explore the evolution of recruitment systems (from manual systems to automated systems) that transition from the basic keyword-based approaches to the advanced semantic analysis technique that make use of deep learning frameworks.

A. Data Sources and Search Strategy

The search strategy targeted several high-impact publication repositories in computer science, artificial intelligence and machine learning, and human resources (HR), etc. so that a comprehensive coverage of the domain can be be ensured. Existing research papers were collected from primary digital libraries (like IEEE Xplore, ACM Digital Library, SpringerLink, and Elsevier (ScienceDirect)) and other sources that included quality journal publications, conference proceedings, preprints, arXiv articles and press/media reports issued by government or private organizations, etc. The search queries involved the relevant domain and technology related keywords like:

- Resume screening and skill-set identification
- Matching resume against job description
- Skill taxonomy and profile representation
- Career path/roadmap prediction,
- Machine learning techniques and evaluation methods, etc.

B. Inclusion Criteria

The criteria listed below were followed to filter the studies ensuring relevance and quality. Preference was given to select the peer-reviewed journal articles, conference proceedings, and high-quality preprints that provided empirical results or proposed novel approaches and frameworks. Purely conceptual papers without any algorithmic implementation or those missing any data analysis, were avoided. Care was taken that the selected papers relate to skill based taxonomy, or automation in some stage of the recruitment process, e.g. focus on skill assessment, resume–job matching, prediction of career roadmap, or analysis of labor market. Papers that proposed or evaluated Computational Intelligence techniques like machine Learning (ML), Natural Language Processing (NLP), Deep Learning (LSTM, BERT), etc. were also studied. These studies proposed techniques for handling tasks such as classification, text extraction, semantic interpretation, or Ontology-based reasoning, etc.

C. Classification Strategy

A four-fold strategy (considering, data sources, skills representation, modeling techniques and evaluation metrics) has been used to classify the selected works, analyze them, and identify trends & gaps in the existing state-of-the-art: The segregation of the reviewed works was done on the basis of their data origin so as to differentiate them between the datasets that are – publicly available or proprietary datasets acquired through from academic institutions, government/labor departments, employers or private recruitment agencies. A set of papers considered used taxonomies from standard frameworks like International Standard Classification of Occupations (ISCO), European Skills, Competences, Qualifications and Occupations (ESCO), or Occupational Information Network (O*NET) for mapping unstructured text to structured terms. Multiple works considered used word embeddings (e.g., Word2Vec, FastText, RoBERTa) for representing the skills as vectors in a semantic space.

A set of papers were considered that used modeling techniques based on traditional machine learning classifiers like K-Nearest Neighbors (KNN), Support Vector Machines (SVM), Naïve Bayes, and Random Forest etc. for categorizing and ranking the resumes. Papers were also considered that included modeling techniques based on advanced/deep learning neural architectures like Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs) LSTM, etc. to capture sequential dependencies in career paths. Another set of papers were considered that included systems which integrate Knowledge Graphs and Ontologies for explicitly modeling the relationships or associations between candidates' skills, and job descriptions.

Quantitative Metrics like Accuracy, Precision, Recall, and F1-Score have been prominently used to determine performance of the proposed models in the selected works. Qualitative validation methods like Human-in-the-loop (evaluation by HR experts) have also been observed in some of the selected works.

III. Background and Preliminaries

This section briefly discusses the initial basic concepts namely, skill taxonomies (like ESCO, DISCO, O*NET), representation of resume and features, and commonly used computational intelligence techniques in career guidance systems.

A. Skill Taxonomies

Standardised taxonomies are essential to inter-connect the terminology used by the job seekers in their resumes and those used by the employers in the job descriptions. Such systems can facilitate a common jargon to define occupations, skills, and qualifications, helping labour market dynamics analysis in a well-defined manner [5], [6].

International and European Standards: The International Labour Organization (ILO) developed the International Standard Classification of Occupations (ISCO). The ISCO is as a widely adopted four-tier hierarchical structure that organises globally available professions or occupations. [6], [7]. While the basline is defined by the ISCO, European organizations extend the ISCO framework for addressing fine granularities. The European Skills, Competences, Qualifications and Occupations (ESCO) classification adds the taxonomy of skills and competences to ISCO, and also incorporated a multi-lingual for 27 languages [5]. ESCO, thus, plots skills to particular occupations, permitting a thorough cross-link to figuring out the skill gaps. Likewise, the European Dictionary of Skills and Competences (DISCO) was curated as an early multilingual thesaurus with standardized skill descriptions across European borders. [5], [8]

National and Regional Standards: The Occupational Information Network (O*NET) standard facilitates occupational information in the United States [9]. Acting as an ontology, the ONET describes an extensive set of occupation descriptors along with the information of skills, capabilities, and work activities applicable to an occupation [5]. This ontology has been frequently considered as a baseline to train the machine learning models for job vacancies' classification [6]. A few of the specific frameworks have also been developed by other countries. For instance, the Thailand Standard Classification of Occupations (TSCO) and the Thailand Standard Industrial Classification (TSIC) classify the occupations based on requirements of the industry [10].

Ontologies and Dynamic Taxonomies: Beyond static lists, research has moved towards ontology-based approaches. For instance, the Computer Science Ontology (CSO) is a large-scale ontology of research areas generated automatically to represent technical competencies [9]. Furthermore, private sector entities and researchers have developed data-driven ontologies using Natural Language Processing (NLP) on large datasets of job advertisements to create dynamic classifications that adapt to market changes more rapidly than curated standards [5].

B. Resume and Job Description Representations

To apply computational techniques to recruitment, unstructured textual data from resumes and job descriptions (JDs) must be converted into machine-readable representations.

A fundamental approach involves the Vector Space Model (VSM), where jobs and candidates are represented as vectors in an n-dimensional space [12]. In this model, text is often pre-processed into a 'Bag of Words' (BoW), where documents are represented by the frequency of their constituent terms [6]. More advanced techniques utilise word embeddings, such as Word2Vec and FastText, to represent words as real-number vectors that capture semantic meaning and context [26], [27]. These embeddings allow for the calculation of similarity between resumes and JDs based on the distance between their vector representations [19].

Resumes are often treated as structured documents composed of distinct sections such as education, work experience, and skills [9], [15]. Advanced parsing systems segment these documents to extract specific attributes (e.g. age, gender, specific technical skills) [11]. Recent frameworks have proposed hierarchical representations that disentangle latent skill factors, modelling the competency level of candidates and the difficulty of job requirements at different granularity levels [28].

Graph-based representations model recruitment data as networks. For example, knowledge graphs can represent skills and

TABLE I
OVERVIEW OF STANDARD SKILL TAXONOMIES AND OCCUPATIONAL FRAMEWORKS

Acronym	Full Name	Region/ Scope	Description and Structure
ISCO	International Standard Classification of Occupations	Global (ILO)	A four-level hierarchical structure organising occupations into 10 major groups. It classifies jobs based on skill levels mapped to formal education (ISCO-88, ISCO-08).
ESCO	European Skills, Competences, Qualifications and Occupations	European Union	A multilingual classification (27 languages) extending ISCO. It segments concepts into three pillars: occupations, skills/competences, and qualifications to facilitate cross-border matching.
ONET	Occupational Information Network	United States	The primary source of US occupational information. It provides comprehensive descriptors including knowledge, skills, abilities, and work activities, often used for training ML models.
DISCO	European Dictionary of Skills and Competences	Europe	A structured thesaurus and vocabulary designed to describe skills and competences for education and labor markets, supporting semantic matching.
ROME	Repertoire Operationnel des Metiers et des Emplois	France	A professional mobility tool organizing jobs into inventories of common titles, activities, and skills based on proximity principles.
TSCO	Thailand Standard Classification of Occupations	Thailand	Modeled after ISCO, it categorizes occupations into 9 major groups based on skill levels, ranging from managers to elementary occupations.
CSO	Computer Science Ontology	Domain-Specific (Global)	A large-scale, automatically generated ontology of research areas in computer science, utilizing semantic relations like super-area and contributes-to.
NACE	Statistical Classification of Economic Activities	European Community	A standard classification system for economic sectors/industries, often used in conjunction with ISCO for labor market intelligence.

TABLE II
FEATURE REPRESENTATIONS OF RESUMES AND JOB DESCRIPTIONS IN SELECTED STUDIES

Study	Resume (Candidate) Features	Job Description (Vacancy) Features	Representation / Technique
Rodriguez & Chavez [11]	Demographics: Age, Sex, Civil Status. Quals: Education, Work Exp., Skills.	Requirements: Job Title, Age, Sex, Civil Status, Education, Exp., Years of Exp.	Profile Matching: Attributes extracted to build profiles for clustering/matching.
Almalis et al. (FoDRA) [12]	Measurable Skills: Education level, Experience, Language, Programming, Network/Tech support knowledge.	Weighted Attributes: 8 dimensions (Exact, Lower limit, Upper limit, Range).	Vector Space: 8-dimensional vectors mapped to scoring table (Minkowski distance).
Mishra et al., [9]	Sections: Contact, Experience (Gap behavior, Role type), Academic (Grades, Projects), Soft Skills.	Requirements: Degree, Skills (Technical & Cultural), Domain concepts.	Ontology/Graph: Skill graphs generated; matched using ontologies and Word2Vec.
Koh & Chew [13]	Standard Parameters: Qualification, Job choices, Salary, Experience, Knowledge, Location.	Vacancy Parameters: Domain, Title, Position, Location, Salary, Required Knowledge.	Standard Template: Unstructured data parsed into templates; missing data inferred via ontology.
G.M. & Suganthi [14]	Clusters: Primary/Secondary Skills, Adjectives (Quality), Adverbs. Fields: Education, Certificates.	Clusters: Primary/Secondary Skills, Adjectives, Adverbs. Fields: Responsibilities, Pref. Skills.	Clustering/Jaccard: Similarity computed between specific word clusters.
Li et al. [15]	Sections: Profile, Education, Work Experience, Activities, Skills, Others.	Description Text: Full JD text used to extract semantic meaning.	Section Encoding: Segmented sections encoded via Transformer models (BERT).
Lamba et al. (SRS) [16]	Tokens: Skills, Experience, College Name (tier ranking), Extra-curriculars.	Constraints: Occupational Category, Skill Arrays, Experience requirements.	Scoring Model: Tokenized and scored based on weighted sections (e.g., College Score).
Lee et al. [17]	Preference Vector: Skill level (binary) in specific subjects (e.g., Inventory Mgmt, Quality Mgmt).	Requirement Vector: Need (binary) for specific subjects.	Vector Matching: Preference vectors derived from subject lists.
Martinez-Gil et al. [8]	Profile Set: Items defined in knowledge bases (e.g., DISCO, ESCO) covering Education, Skills, Languages.	Job Offer Set: Required items defined in knowledge bases.	Transformation Cost: Edit distance operations (insert, delete, replace) between sets.
Jain et al. [18]	Summarized Text: Lemmatized tokens from parsed PDF CVs.	Query Document: Job description treated as a query vector.	LSA / LDA: Latent Semantic Analysis and Topic Modeling on TF-IDF matrices.
Wowczko [19]	N/A (Focus on Vacancy Analysis).	Terms: Job Title, Job Description (full text converted to N-grams).	Text Mining/KNN: Term-Document Matrices with TF-IDF weighting.
Ali et al. [20]	Tokens: Cleaned text (stopwords removed, stemmed) from resumes.	Categories: Target labels (e.g., Java Developer, Testing) derived from JD clusters.	TF-IDF / SVM: Term Frequency-Inverse Document Frequency features fed to ML classifiers.
Satheesh et al. [21]	Entities: Extracted via NER (Name, Skills, Designation, Organization).	Keywords: Technical requirements (e.g., ”Python”, ”Machine Learning”).	SpaCy NER / Cosine: Named Entity Recognition followed by similarity scoring.
Kinge et al. [22]	Attributes: Skills (Sr), Education (Er), Experience (Xr).	Description: Full text for TF-IDF vectorization.	Weighted Formula: Hybrid score using weighted attributes and Cosine Similarity.
Kumari et al. [23]	Homogeneous Text: Converted from heterogeneous formats (DOC, PDF).	Criteria: Search engine queries (Criteria 1, 2, 3...).	Naïve Bayes: Classifier used to interact with storage and match candidates.
Tejaswini et al. [24]	Summarized Text: Generated using Genism library.	Description: Scaled and summarized text.	KNN / Cosine: Content-based filtering using similarity measures.
Ikudo et al. [25]	Admin Data: Job Title, Age, Wage, Transitions (Preceding/Succeeding titles).	N/A (Focus on Occupational Classification).	Random Forest: Classification based on administrative HR records and titles.

occupations as nodes, with edges representing relationships such as ”belongs to” or ”requires” [29]. This structure supports complex reasoning, such as identifying skill adjacencies or visualising a candidate’s fit within a specific domain [9].

C. ML and NLP Techniques in Career Guidance Systems

A blend of Machine Learning (ML) and Natural Language Processing (NLP) approaches is significantly contributing towards the development of automated career guidance and recruitment systems – ML for classification matching, and NLP for extracting information. NLP can be effectively used to parse unstructured text from resumes/CVs and job descriptions (JDs) using techniques like tokenization, stop-word removal, Named Entity Recognition (NER) and Term Frequency-Inverse Document Frequency (TF-IDF), etc. Tokenization and Stop-word Removal is the basic pre-processing technique used to clean the text by removing the non-informative words from it [14], [24]. NER is used to identify and extract specific entities such as candidate names, organisation names, locations, and skills etc. from raw text [20], [26]. NER can be supported by libraries like SpaCy and Stanford CoreNLP to accomplish this task [9], [30]. The TF-IDF statistic is widely used to evaluate the importance of a word in a document relative to a corpus and hence it serves as a feature extraction method for downstream classification [20], [31]. A variety of supervised learning algorithms are employed to classify resumes into job categories or predict candidate suitability. Traditional classifiers based on Supervised machine learning techniques like Support Vector Machines (SVM), Naïve Bayes (Multinomial and Bernoulli) Classifier, K-Nearest Neighbours (KNN), and Logistic Regression are frequently adopted benchmarks for resume classification tasks [19], [20]. Techniques such as the Random Forest classifiers have also been found effective in predicting the occupational categories [25]. Ensemble methods such as XGBoost and AdaBoost have also been used to predict candidate suitability classes (e.g., ’Most Suitable’, ’Not Suitable’) with high accuracy [14].

Recent systems have adopted deep learning architectures to capture complex semantic relationships. Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs), including Long Short-Term Memory (LSTM) networks, are used for sequence modelling and text-classification [18], [20]. The state-of-the-art transformer models such as the Bidirectional Encoder Representations from Transformers (BERT) and RoBERTa can be used to generate context-aware embeddings and can significantly improve the accuracy of match between job seekers and the job description [32]. After generating the representations, these intelligent systems compute the matching scores using different metrics – like Euclidean Distance, Jaccard Similarity and Cosine Similarity – that can be used to rank order the candidates against job description [14], [18], [24]. The computational techniques can be categorized by their particular role/function within the recruitment process – (e.g., Pre-processing, Representation, Classification, Matching) – as presented in Table III appropriately citing the relevant studies.

IV. Taxonomy of Existing Approaches

Existing studies have been classified into the following categories based on the review:

A. Skill Assessment and Profiling Systems

These approaches focus on extracting and quantifying skills from resumes, assessments, or online profiles to create structured representations of candidates. Resume Parsing and Entity Extraction: A fundamental step in profiling involves converting unstructured text (PDF, DOCX) into structured formats. Approaches utilize Natural Language Processing (NLP) techniques, such as Named Entity Recognition (NER) using models like SpaCy, to automatically extract entities like candidate names, organizations, and specific skills [21]. Other systems employ regular expressions and pattern matching to parse details including education and contact information [30].

Competence Level Prediction: Apart from simple extraction, the more sophisticated systems are seeked to identify the skill level of a person. Recent studies have made use of context-based transformers for identifying skills at different levels (Clinical Research Coordinator levels 1-4) by categorizing the document sections, such as ”Work Experience” and ”Education.” [15].

Ontology-Based Profiling: To standardize skill representation, systems often map extracted terms to established taxonomies such as the Computer Science Ontology (CSO) or ESCO [9]. Some approaches generate skill graphs where nodes represent competencies and edges represent relationships, allowing for a more semantic understanding of a candidate’s profile [9].

Cognitive Diagnosis: New approaches have brought forward theories of cognitive diagnoses to profiling where the hiring process is seen as an exercise problem solving process. The system endeavors to untangle latent skill elements in order to develop a clear model of expertise relative to job demands. [28].

B. Career Path Prediction Models

These prediction approaches make use of supervised and unsupervised machine learning methods in predicting career paths through the use of historical data.

Using historical data from administrative and HR information, researchers can extract job transition and concurrent job title data in order to define career paths. Classifiers such as the random forest classifier have been used to predict occupational classifications based on transition of jobs, hence enriching career path data [25]. Some of the career prediction frameworks use Case Based Reasoning (CBR) methods to identify the gap between a candidate’s ”acquired” skill set against the market’s ”required” skill set. This model can help predict the next course of action to help candidates acquire the missing skills [26]. The platform called ”Create Your Own Future” (CYOF) uses AI toolkits to enable workers to configure their own career paths. In the case of this system, users are enabled to provide rat- ings of their competencies (e.g., from ”undeveloped” to ”very competent”), which are then matched against job families [33]. Methods based on unsupervised learning, including clustering, have been used to study student preferences by grouping together individuals with similar characteristics. This approach is also effective in predicting an appropriate career for individuals in line with labor market preferences [17].

C. Resume–Job Matching and Recommendation Systems

The systems apply various methods for matching candidate’s information against job descriptions based on similarity, embeddings, or hybrid recommendation algorithms. Most systems employ mathematical distance metrics to measure how well a candidate matches to a specific job. This includes the application of Cosine Similarity on TF-IDF vectors [22], [24], JJaccard similarity for skill clusters [14], and expanded Minkowski distances to account for range-based qualifications as part of the FoDRA algorithm [12]. Some novel systems also apply ”transformation costs” to determine the ”edit distance” needed to transform a candidate’s profile information into a description of the job’s requirements [8]. To account for keyword mismatch, some systems apply Latent Semantic Analysis (LSA) and Latent Dirichlet Allocation (LDA). These systems convert candidate’s resumes and job descriptions to

TABLE III
SUMMARY OF MACHINE LEARNING AND NLP TECHNIQUES IN CAREER GUIDANCE SYSTEMS

Category	Technique / Algorithm	Application in Recruitment Domain	Representative Studies
Text Representation & Embeddings	TF-IDF	Weighting keywords (skills, education) based on frequency to filter non-informative terms from resumes.	Kinge et al. [22], Jain et al. [18], G.M. & Suganthi [14]
	Word2Vec / Doc2Vec	Mapping words to low-dimensional vector spaces to capture semantic relationships between job titles and skills.	Boselli et al. [27], [29], Aljohani et al. [26]
	FastText	Extension of Word2Vec used to handle out-of-vocabulary terms in job descriptions.	Aljohani et al. [26]
	Transformers (BERT, RoBERTa)	Context-aware embedding generation for deep semantic understanding of competence levels and sentence matching.	Li et al. [15], Aleisa et al. [32]
Information Extraction (NLP)	Named Entity Recognition (NER)	Extracting specific entities (Candidate Name, Organization, Skills, Dates) from unstructured resume text.	Satheesh et al. [21], Kinge et al. [22]
	Lemmatization & Tokenization	Pre-processing steps to reduce words to root forms (e.g., "managing" to "manage") for consistent matching.	Jain et al. [18], G.M. & Suganthi [14]
	TextRank	Summarizing long resumes or job descriptions into concise key sentences.	Jain et al. [18]
Supervised Classification (ML)	Support Vector Machines (SVM)	Classifying resumes into occupational categories (e.g., Java Developer, Tester) and text classification of vacancies.	Kinge et al. [22], Boselli et al. [27]
	Random Forest	Decision-tree based ensemble method used for occupational classification based on administrative HR data.	Ikudo et al. [25], Boselli et al. [27]
	Naïve Bayes	Probabilistic classifier used for resume screening based on feature independence assumptions.	Kinge et al. [22], Kumari et al. [23]
Unsupervised Learning	Clustering (K-Means, BIRCH)	Grouping similar job profiles or skills to identify labor market taxonomies without labeled data.	Aleisa et al. [32], Rodriguez & Chavez [11]
	Latent Dirichlet Allocation (LDA)	Topic modeling to discover hidden themes (e.g., specific skill sets) within large corpora of resumes.	Jain et al. [18]
	Latent Semantic Analysis (LSA)	Reducing dimensionality of text data to find latent correlations between terms and documents.	Jain et al. [18]
Similarity & Matching Metrics	Cosine Similarity	Measuring the angle between candidate and job vectors to determine ranking/fit.	Tejaswini et al. [24], Aljohani et al. [26]
	Jaccard Similarity	Calculating the intersection over union of skill sets (e.g., matching keywords between CV and JD).	G.M. & Suganthi [14]
	Extended Minkowski Distance	Measuring distance between candidate attributes and job requirements involving range-based values.	Almalis et al. (FoDRA) [12]
Deep Learning Architectures	CNN, LSTM, GNN	Capturing sequential dependencies in career paths and modeling complex graph-based skill interactions.	Boselli et al. [27], [29], DISCO [28]

the semantic space where topics are discovered and matching precision is improved [18]. Graph edit distance and ontological matching enable the comparison of the structural and semantic associations among the skills listed in the resume and those required in a job description, thus providing an advanced match compared to keyword analysis [9]. Supervised learning methods like KNN, SVM and Stochastic Gradient Descent (SGD) are commonly used in resume classification or ranking by suitability to certain jobs [19], [20], [24]. Some systems utilize "hiring history" algorithms for considering previous employment [10]. Sophisticated systems use various models together. For example, "Intelligent Job Matching" employs a self-learning engine that automatically populates the empty fields of a resume through ontology matching before matching process [13]. Others combine collaborative filtering with content-based filtering to enhance recommendation accuracy [12].

D. Integrated Skill Monitoring Frameworks

There is an increasing trend for combining assessment, prediction, and recommendation into a single platform, using big data for Labor Market Intelligence (LMI). Models such as WoLMIS harvest and categorize millions of online job openings that serve as a basis for high-level decision making. These models employ machine learning techniques to categorize job openings into standard taxonomies (such as ISCO) and extract skills that can help to monitor labor market dynamics in real-time [6], [29]. AI Recruiting Model (AIRM) introduces a holistic AI architecture with a Data Lake, which includes an Initital Screening layer (based on BIRCH clustering), a Mapping layer (based on RoBERTa embeddings and FAISS ranking), and a Preferences layer. Such a model helps to reduce delays and offer personalized recommendations to recruiters and applicants [32]. Programs such as OpenSKIMR and Europass use skill matching together with labor market analytics. The portals leverage standard taxonomies (ESCO, DISCO) for cross-border skill matching and identifying skill gaps on a regional/national scale [5]. Modern approaches place emphasis on building centralized data repositories (Data Lakes), where raw data is collected from a wide range of sources (data from the education sector, labor market). This enables applying deep learning models to extract patterns that can be useful in career counseling and policy-making [32].

V. Comparative Analysis

The technological evolution of automated recruitment and career guidance systems can be distinguished by the transition from keyword-based heuristic approaches to semantic-oriented deep learning techniques. Traditional methodologies have primarily been based on the use of statistical metrics such as the TF-IDF technique combined with classical classification algorithms such as Naive Bayes, K-nearest neighbors (KNN), and Support Vector Machines (SVM) classifiers [19], [20], [23]. Although computationally efficient and interpretable, such approaches often suffer from semantic inconsistencies where it is difficult for a model to understand that certain terms (coding vs. programming) denote the same skill [18]. Therefore, recent advancements in research have focused on applying deep learning and natural language processing (NLP) techniques to develop models that utilize advanced techniques such as Word2Vec, LSA/LDA, and Transformer-based models (BERT, RoBERTa) to create meaningful embeddings of the underlying skill semantics [27], [32], [34]. Although accurate, such models require considerable computational power and extensive data sets that can still pose a challenge in the area. Another rising trend in research includes combining data-driven models with structured knowledge bases (ESCO, O*NET, CSO) to increase the accuracy and interpretability of results [5], [9].

Table IV summarizes representative studies, highlighting techniques and limitations.

A. Key Insights from the representative studies

1) Technological Shift to Semantics: There seems to be a definite trend away from syntactic matching towards semantic matching. The use of Latent Semantic Analysis (LSA), word embedding techniques such as Word2Vec and GloVe enables the system to comprehend synonyms and skills hierarchies, thereby increasing the accuracy of matching. [18], [19], [32].
2) Trade-off Between Accuracy and Interpretability: Although deep learning models (such as LSTM, CNN, and BERT) often yield high accuracy levels (which can be higher than 90%), they serve as ”black box” approaches. On the other hand, ontology-based and rule-based approaches have poorer performance rates yet provide transparent reasoning behind their decisions, which is essential in avoiding hiring biases [20], [36], [37].
3) Data Heterogeneity and Parsing Challenges: A universal bottleneck that exists in all studies is the absence of structure in the resumes (PDF, DOCX). For effective systems, there must be good pre-processing procedures that incorporate Named Entity Recognition (NER) and text segmentation to convert heterogeneous documents into structured formats before the analysis can take place [21], [30], [35]. The absence of standardized format for the resumes within the industry still poses a problem in scalable parsing. [13].
4) Domain Specificity vs. Generalization: There have been many suggested models that tend to be very specific, designed to learn from datasets specific to the Information Technology/Software industry or specific regions such as Saudi Arabia or China. The issue is that such models do not exhibit good transfer learning properties for blue-collar jobs or other languages, highlighting a need for more diverse, multi-domain training corpora [10], [26], [32].
5) Metric Standardization Gap: There is no common metric for measuring success. Classification-based architectures use Precision, Recall, and F1-Score metrics, whereas recommendation-based models use Cosine Similarity, NDCG, and Mean Reciprocal Rank (MRR) metrics. This creates an obstacle in benchmarking between diverse models since there are differences in metrics. [16], [20], [24].

B. Trade-offs between accuracy, computational cost, and matching quality across different algorithms

Comparative Accuracy of Machine Learning Classifiers: Fig. 1 highlights that while most modern classifiers perform well, Support Vector Machines (specifically Linear SVC) and Stochastic Gradient Descent (SGD) outperform probabilistic models like Naïve Bayes (Bernoulli) in resume classification tasks.

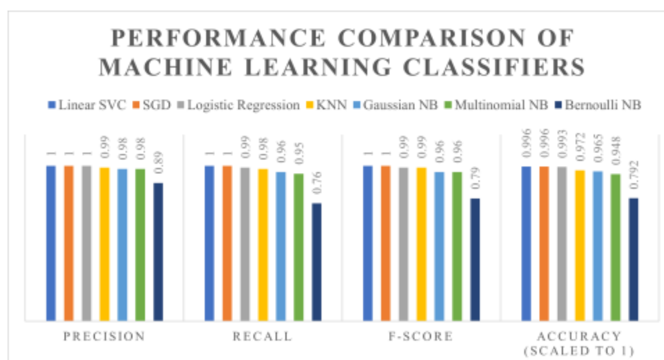

Fig. 1. Comparison of Machine Learning Classifiers

Matching Quality vs. Algorithm Type: This comparison distinguishes between statistical fitness models (like Least Squares or Pearson Correlation) and heuristic/hybrid models like the proposed Multiple Preference Matching Algorithm (MPMA) [17]. Fig. 2 demonstrates that statistical models suffer from high ”Miss-matching” rates (placing candidates in

TABLE IV
COMPARISON OF SKILL MONITORING AND CAREER RECOMMENDATION APPROACHES

Study	Technique	Evaluation Metrics	Limitations
Yu et al. [28]	DISCO Framework (Hierarchical Disentangled Cognitive Diagnosis)	AUC, HR@K, NDCG@K	Relies on interaction data availability; complexity in disentangling hierarchical skill factors.
Aleisa et al. [32]	AI Recruiting Model (BIRCH, RoBERTa, FAISS)	Accuracy (Human-in-the-loop), Execution Time	Limited geographic scope (Saudi Arabia); fails to recognize transferable skills across sectors.
Li et al. [15]	Context-aware Transformers (BERT) with Section Encoding	Accuracy, Fleiss Kappa	Misclassification frequent between adjacent competence levels; high computational resource requirement.
Martinez-Gil [8]	Adaptive Matching with Transformation Costs	Spearman Rank Correlation	Relies on expert knowledge consistency; excludes multimedia/personal features.
Jain et al. [18]	LSA and LDA with TextRank Summarization	Precision (True/False Positives)	LDA showed significantly lower precision than LSA; sensitive to noisy data.
Tejaswini et al. [24]	TF-IDF, Cosine Similarity, KNN	Parsing Accuracy, Ranking Accuracy	Implicit text compression via summarization leads to information loss; cannot serve as sole selection criterion.
Wowczko [19]	KNN Classification, Text Mining (R)	Accuracy	Focused solely on IT sector in Ireland; manual verification needed for ambiguous job titles.
Almalis et al. [12]	FoDRA (Extended Minkowski Distance)	Suitability Value (Distance-based)	Sensitivity to the choice of the p variable; requires testing on diverse field datasets.
Sridevi & Suganthi [14]	Jaccard Similarity on Clusters + XGBoost	Classification Rate (Accuracy)	Jaccard similarity is sensitive to data size; relies on quality of keyword clustering; Does not currently integrate social media features or soft skill validation.
Rodriguez & Chavez [11]	Profile Matching Model (WEKA)	Similarity Measures	Dataset limited to specific recruitment agency records (2,283 profiles).
Koh & Chew [13]	Inference Engine (Forward Chaining) with Ontology	Distance Score Calculation	Relies on domain experts for defining weighting factors and threshold values.
Lee et al. [17]	Multi-Phase Matching Algorithm (MPMA)	Matching Accuracy (Perfect/Second Perfect)	Statistical fitness models produced mis-matching; requires specific preference vectors.
Boselli et al. [27]	SVM (Linear/RBF) on ISCO Taxonomy	Precision, Recall, F1-Score	Title features alone often insufficient; data biased by seasonal effects (e.g., teaching jobs).
Ikudo et al. [25]	Random Forest on Administrative Data	Predictive Accuracy, Gini Impurity	Ambiguity in specific titles (e.g., ”Scientist”, ”Fellow”) leads to misclassification.
Mishra et al. [9]	Ontology (ESCO) + Word2Vec + Hay Criteria	Matching Score (MR), Accuracy	Limited analysis axes; requires user tuning for mandatory filtering fields.
Lamba et al. [16]	Rule-based Dictionary Matching (SRS)	Accuracy, F1-Score (Confusion Matrix)	Requires exact keyword matching; relies on manual inputs for experience/college.
Satheesh et al. [21]	SpaCy NER + Cosine Similarity	Parsing Accuracy, Ranking Accuracy	Model failed to recognize startup company names; limited to specific file formats.
Amin et al. [35]	TF-IDF + Weighted Cosine Similarity	Similarity Score (0-1)	Manual input required for years of experience; dictionary creation required for new companies.
Mittal et al. [30]	Rule-based Information Extraction (JSON)	Precision, Recall, F1 Score	Difficulty filtering suitable candidates due to volume; reliance on specific domain fields.
Aljohani et al. [26]	CBR + BM25 Matching	Weighted Scores	Requires extensive data cleaning (HTML tags, phone numbers); specific to Saudi market.

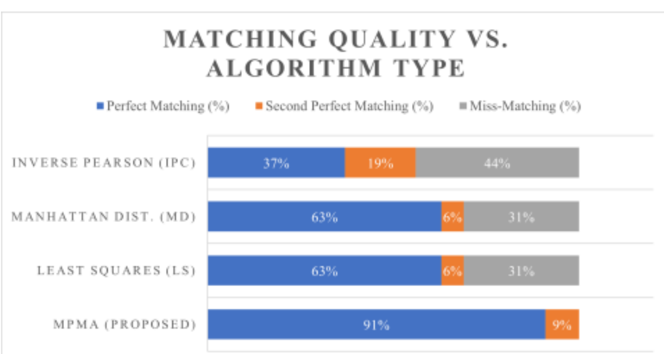

Fig. 2. Matching Quality vs. Algorithm Type

wrong jobs), whereas the proposed MPMA model eliminates miss-matching.

Computational Cost vs. Performance High accuracy often comes at a computational cost. Fig. 3 and Fig. 4 compare the performance of of different algorithms to execute versus their success rate.

Impact of Pre-processing on Precision Fig. 5 illustrates the ”Technological Shift” mentioned in this study and shows that Latent Semantic Analysis (LSA) generally outperforms Latent Dirichlet Allocation (LDA), but both are significantly improved when text summarization (TextRank) is applied as a pre-processing step.

Comparison of Deep Recommendation Models presented in Table V highlights the state-of-the-art in implicit feedback systems.

VI. OPEN RESEARCH CHALLENGES

Despite notable progress in the development of automated career guidance and recruitment systems, several significant challenges hinder their widespread adoption and effectiveness.

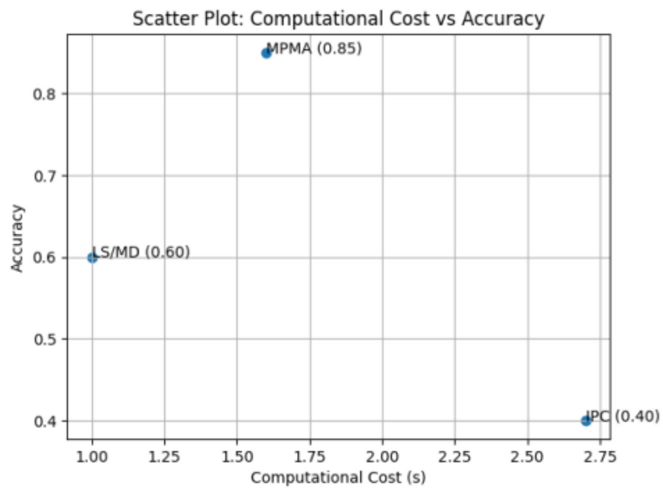

Fig. 3. Computational Cost vs. Accuracy

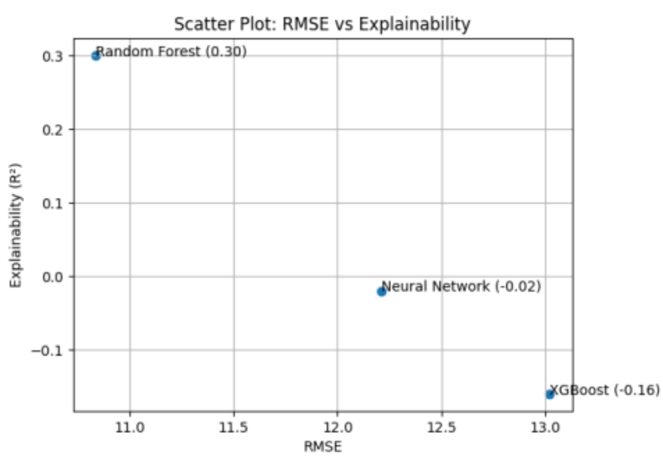

Fig. 4. RMSE vs Explainability (R^2)

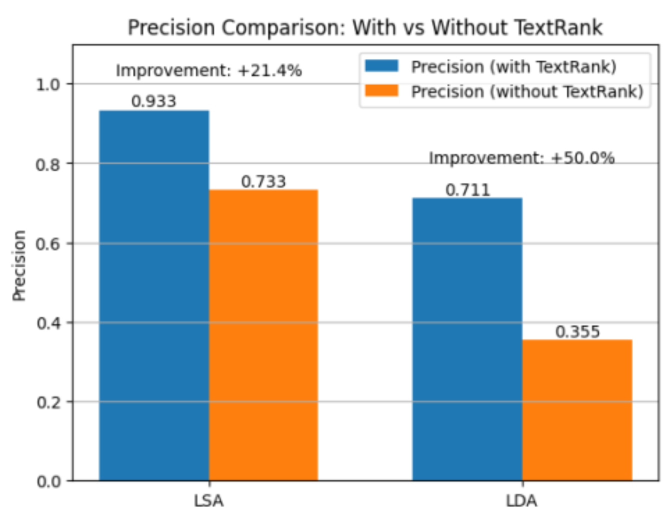

Fig. 5. Comparative Precision: With vs. Without Pre-processing

The complexity of labor markets, combined with the limitations of current AI methodologies, presents the following open research problems:

1) Lack of continuous and dynamic skill monitoring.
2) Limited personalization in skill development recommendations.
3) Poor integration with educational ecosystems.
4) Absence of standardized benchmarks and datasets.

A. Lack of Continuous and Dynamic Skill Monitoring

The rapid development in technologies has significantly reduced the life span of professional skills, where the half-life period of skills is less than five years and as low as two and a half years in tech industries [38]. Conventional approaches of tracking labor market demands through surveys have inherent drawbacks of significant delays, where obtaining survey data takes as long as one year making them unfit for instant decisions [6], [29]. Though there have been attempts at using data-driven methods such as web vacancies to identify emerging skills in an instantaneous manner [39], conventional ontologies (e.g., ISCO, ESCO) are often too slow in updating themselves to reflect emerging micro-trends as well as software adjacent to the demands of the industry [5]. Moreover, current systems lack an understanding of how the skills of a candidate have evolved over the years, where the models classify candidates having more experience as beginners because they fail to credit the progression of skills across different job roles [32]. There is a dire necessity for dynamic skill monitoring framework where ontology can be regularly updated and skills of a person can be monitored over time.

B. Limited Personalization in Skill Development Recommendations

Although present AI systems are quite accurate when predicting someone's performance or recommending jobs for him/her, they are often not personalized enough to provide valuable recommendations for career growth. The rationale behind the most common artificial intelligences is usually based on statistical averages and does not take into account unique features of each student/job applicant. [36]. Therefore, recommendations can be generic or unrealistic; for instance, algorithmic recommendations could include the suggestion that a student minimize their interaction with course content based on false correlations within the general population [36]. Evidence suggests that "one-size-fits-all" training programs

TABLE V
PERFORMANCE COMPARISON OF DEEP GRAPH LEARNING MODELS (HR@10)

Base Model	Standard	With DISCO	Improvement
Matrix Factorization	0.5835	0.6337	+8.6%
NGCF	0.6144	0.7145	+16.3%
LightGCN	0.6364	0.7070	+11.1%
DPGNN	0.7018	0.7029	+0.15%

are not effective because adult learners tend to favor learning in practice and personalized paths [38]. Additionally, current methods cannot accurately recognize the transferable skills of individuals across various industries, and thus fail to recommend plausible career changes for individuals who have experience managing operations within one particular field [32]. Achieving true personalization requires "instance-level" explainability that can identify exactly which variables a specific user needs to manipulate to achieve their unique professional goals [36].

C. Poor Integration with Educational Ecosystems

A huge gap still exists between what skills are sought for in the labor market and what skills are taught in educational establishments. There is a lack of a central data repository or a data infrastructure that integrates both education datasets and labor market datasets and, hence, the impossibility of matching supply and demand [32]. Whereas job ads are being increasingly annotated using standardized skill classifications, educational offers (curricula, syllabi, and training programs) seldom follow standardized metadata schemas such as ESCO [5]. This situation leads to the inability to automatically assess the "skills supply," i.e., to carry out accurate gap analysis between educational achievements and labor market needs. [5], [26]. Moreover, a self-perpetuating vicious circle is formed: SMEs need skilled employees but do not have sufficient means for their adequate training while graduates lack the required skill set [3], [26]. Future models should resolve this problem through parameterization of the educational process and tight integration of curricula with labor market intelligence [3], [26].

D. Absence of Standardized Benchmarks and Datasets

Automated recruitment research lacks standardized benchmarking techniques and therefore has limited comparability of model performance. Research is often done using private or ad hoc datasets collected from a particular region (e.g., Saudi Arabia, China, UK), platform, or source, limiting the applicability of results [20], [26]. There is no consensus on a "ground truth" for occupation categorization due to inherent inconsistency in titles between organizations or HR information systems [25]. For instance, ambiguous occupational titles such as "Specialist" or "Analyst" can describe an entirely different occupation in another domain, causing noise and thus reducing the efficiency of supervised models [19], [32]. Moreover, the absence of a common language or taxonomy between nationalities poses further challenges to cross-linguistic matching and mobility analysis [6], [27]. Creating freely accessible multi-lingual benchmark data sets is critical to achieving state-of-the-art in algorithmic hiring fairness and transparency [32].

VII. Future Research Directions

The review of the current landscape indicates that while automated recruitment systems have achieved high accuracy in specific tasks, they often lack the adaptability, transparency, and holistic understanding required for complex real-world deployment. Future work could focus on, but not limited to, the following key areas:

A. Adaptive and Longitudinal Skill Monitoring Models

Present frameworks tend to see candidate profile as a snapshot. Yet, the fast "half-life" of professional knowledge—less than five years in technical domains [9] demands constant surveillance. Future studies have to move from taxonomic mapping to longitudinal study, tracing how the competence map of a given employee changes over time [12]. Studies should target "ideographic" studies that track individual career trajectories and not just statistical patterns, so as to trace the decline in skills and identify the need for upgrading measures to counteract obsolescence [9], [12]. Frameworks should be designed as "career configuration" platforms (like Create Your Own Future), which perceive employability as a lifelong process and not a one-off matching exercise, and update skill graphs constantly through learning and employment conditions [9].

B. Explainable and Ethical AI (XAI) for Career Guidance

As deep learning models grow increasingly sophisticated, the process of making decisions becomes even more opaque, raising questions of bias and fairness. Studies on the interpretability of recruiting recommendation engines still remain vastly underexplored [9]. Future research designs will require analyzing cases of "misprediction," or cases when algorithms malfunction, to uncover what is causing these errors, such as biased predictions against individuals who have followed non-linear career trajectories, thereby improving overall algorithmic fairness [12]. Another research avenue to explore includes developing "cognitive diagnostic" models that allow for separating different skill sets in order to provide comprehensible explanations of what made a particular individual a good match (or mismatch) for a position [9]. As a way to avoid bias and subjectivity associated with hiring processes, future systems must incorporate human feedback loops during training to ensure that AI predictions match human expert opinions [9].

C. Integration of Generative AI for Skill Validation and Assessment

The advent of Large Language Models (LLMs) and Generative AI presents potential paths towards overcoming keyword matching. Future research should focus on exploring the potential of generative AI entities (such as AI mentors in video form) that are capable of performing sentiment analysis as well as engaging in conversations in order to determine the underlying barriers, motivations, and "vocational personality" of the candidate that is usually not revealed by static CVs [9]. The research could focus on using pre-trained language models (such as BERT, RoBERTa) for generating contextualized interview questions or personalized learning pathways aligned with skill deficiencies [9].

D. Scalable Frameworks Validated in Real-World Educational Settings

Most current research uses datasets collected offline (such as those found on Kaggle) and not deployed in live environments. It is necessary to have frameworks that help fill the gap between labor market intelligence (LMI) and the academic curriculum. Future systems should be deployed and tested in real-world educational environments to ensure that they can efficiently help students find ”thriving economy” opportunities [9]. Research must address the latency and scalability issues associated with processing national-level datasets (Big Data), moving from Minimum Viable Products (MVPs) to robust architectures capable of handling millions of job seeker queries in real-time [9].

E. Multi-modal Recruitment Analysis

Existing approaches make extensive use of textual information. Future work needs to venture into multi-modal processing to get a more holistic representation of the candidates. Combining technologies like facial emotion analysis and speech to text analysis while conducting a video interview can add more informative data about soft skills, which cannot be quantified by texts alone [9]. Training models capable of concurrently processing multi-modal data sets (text, visual layout of the CV, audio) can improve resume comprehension [9].

F. Blockchain for Trust and Credential Verification

One problem with the automated screening process is that the self-reports of skill cannot be easily verified. The future architecture must consider exploring Blockchain technology in building an immutable and authenticated form of ”Smart CVs”. This way, algorithms will be able to match applicants based on credible information instead of claims, greatly increasing trust [9].

G. Cross-Lingual and Cross-Domain Generalization

Most existing systems are developed using English language datasets tailored specifically to the IT field. Future models have to be tested on different language datasets (such as Arabic and European languages) to establish their cross-border viability [9]. Future research should concentrate on spotting capability adjacency to understand how skills in one field (e.g. manufacturing) can help in another (e.g. technology) [9].

VIII. Conclusion

This paper comprehensively reviewed the application of computational intelligence techniques in monitoring skills and predicting career roadmaps. The objective of this study is therefore to provide information on the current state of affairs as well as highlight areas that need further research. The findings clearly show a paradigm shift in this area of study with the shift away from the use of keyword-oriented algorithms like Boolean matching and Term Frequency – Inverse Document Frequency (TF-IDF) toward an architecture that utilizes Deep Learning and NLP technologies like transformer models (BERT, RoBERTa) and sequential models (CNN, LSTM) to bridge the gap between unstructured resume data and structured job descriptions, resulting in a classification and matching accuracy above 90% [15], [30]. Moreover, the combination of standardized knowledge bases such as ESCO, O*NET, and the Computer Science Ontology (CSO) has become crucial to normalize data heterogeneity and facilitate labor market mobility across borders [5], [6].

However, still several obstacles exist within this field. The ”black-box” problem of sophisticated neural networks has raised issues about explainability and bias, requiring more emphasis on the concept of Explainable Artificial Intelligence (XAI) [36]. Moreover, a gap also exists between the supply of education and demand in the labor market; current mechanisms are efficient in matching current candidates but fall short when it comes to offering lifelong guidance for upskilling in light of rapid developments in the industry [26], [33]. Looking forward, the next generation of career guidance systems must move beyond static profile matching to become dynamic ”career configuration” tools [33]. Future research should prioritize the development of dynamic Labor Market Intelligence (LMI) systems that update skill taxonomies in real-time [6], the incorporation of multi-modal data (video, audio) for holistic candidate assessment [31], and the utilization of Generative AI to create personalized, interactive career coaching agents [26]. In the end, the effective combination of the AI-driven recruitment process and the education planning system becomes crucial in solving the problem of the skills shortage at the global level.

References

[1] V. Taparia and V. Taparia, “India created 4.66 cr jobs in FY24; employment growth rate rises to 6

[2] H. Times, “RBI data shows India created far more jobs than private surveys depict— Business News,” 7 2024. [Online]. Available: https://www.hindustantimes.com/business/rbi-data-shows-india-created-far-more-jobs-than-private-surveys-101720446996720.html

[3] ONLYIAS, “India skills report 2024 and skill development in india,” December 2023, accessed: 2025-10-10. [Online]. Available: https://pwonlyias.com/editorial-analysis/india-skills-report-2024/

[4] “Employment trends in india: Addressing challenges and solutions for job creation,” ONLYIAS, April 2024, accessed: 2025-10-10. [Online]. Available: https://pwonlyias.com/editorial-analysis/india-skills-report-2024/

[5] R. Rentzsch and M. Staneva, “Skills-matching and skills intelligence through curated and data-driven ontologies,” in Proceedings of the DELFI Workshops 2020, Heidelberg, Germany, 2020.

[6] R. Boselli, M. Cesarini, S. Marrara, F. Mercorio, M. Mezzanzanica, G. Pasi, and M. Viviani, “WoLMIS: a labor market intelligence system for classifying web job vacancies,” Journal of Intelligent Information Systems, vol. 51, no. 3, pp. 477–502, 9 2017. [Online]. Available: https://doi.org/10.1007/s10844-017-0488-x

[7] P. Elias, “Occupational classification (isco-88): Concepts, methods, reliability, validity and cross-national comparability,” OECD Labour Market and Social Policy Occasional Papers, Tech. Rep. 20, 1997.

[8] J. Martinez-Gil, A. L. Paoletti, and M. Pichler, “A novel approach for learning how to automatically match job offers and candidate profiles,” Information Systems Frontiers, vol. 22, no. 4, pp. 1265–1274, 2020.

[9] R. Mishra, R. Rodriguez, and V. Portillo, “An ai based talent acquisition and benchmarking for job,” arXiv preprint arXiv:2009.09088, 2020.

[10] S. Kitichalermkiat, V. Singh, S. Chaowiang, S. Tangprasert, P. Chiraprawattrakun, and N. Leeprechanon, “Job matching using artificial intelligence,” International Journal on Soft Computing, Artificial Intelligence and Applications (IJSCAI), vol. 11, no. 3, pp. 1–12, 2022.

[11] L. G. Rodriguez and E. P. Chavez, "Feature Selection for Job Matching Application using Profile Matching Model," 2019 IEEE 4th International Conference on Computer and Communication Systems (ICCCS), pp. 263–266, 2 2019. [Online]. Available: https://doi.org/10.1109/ccoms.2019.8821682

[12] N. D. Almalis, G. A. Tsihrintzis, N. Karagiannis, and A. D. Strati, "Fodra - a new content-based job recommendation algorithm for job seeking and recruiting," in 2015 6th International Conference on Information, Intelligence, Systems and Applications (IISA). IEEE, 2015, pp. 1–10.

[13] M. F. Koh and Y. C. Chew, "Intelligent job matching with self-learning recommendation engine," Procedia Manufacturing, vol. 3, pp. 1959–1965, 2015.

[14] G. M. Sridevi and S. K. Suganthi, "Ai based suitability measurement and prediction between job description and job seeker profiles," International Journal of Information Management Data Insights, vol. 2, no. 2, p. 100109, 2022.

[15] C. Li, E. Fisher, R. Thomas, S. Pittard, V. Hertzberg, and J. D. Choi, "Competence-level prediction and resume & job description matching using context-aware transformer models," arXiv preprint arXiv:2011.02998, 2020.

[16] D. Lamba, S. Goyal, V. Chitresh, and N. Gupta, "An integrated system for occupational category classification based on resume and job matching," in International Conference on Innovative Computing and Communication (ICICC 2020), 2020, available at SSRN 3607282.

[17] D. Lee, M. Kim, and I. Na, "Artificial intelligence based career matching," Journal of Intelligent & Fuzzy Systems, vol. 35, no. 6, pp. 6061–6070, 2018.

[18] L. Jain, M. A. H. Vardhan, G. Kathiresan, and A. Narayan, "Optimizing people sourcing through semantic matching of job description documents and candidate profile using improved topic modelling techniques," in Advances in Artificial Intelligence and Data Engineering (AIDE 2019). Springer, 2021, pp. 899–908.

[19] I. A. Wowczko, "Skills and vacancy analysis with data mining techniques," Informatics, vol. 2, no. 4, pp. 31–49, 2015.

[20] I. Ali, N. Mughal, Z. H. Khand, J. Ahmed, and G. Mujtaba, "Resume classification system using natural language processing and machine learning techniques," Mehran University Research Journal of Engineering and Technology, vol. 41, no. 1, pp. 65–79, 2022.

[21] K. Satheesh, A. Jahnavi, L. Iswarya, K. Ayesha, G. Bhanusekhar, and K. Hanisha, "Resume ranking based on job description using spacy ner model," International Research Journal of Engineering and Technology (IRJET), vol. 7, no. 5, pp. 74–77, 2020.

[22] B. Kinge, S. Mandhare, P. Chavan, and S. M. Chaware, "Resume screening using machine learning and nlp : A proposed system," International Journal of Scientific Research in Computer Science, Engineering and Information Technology (IJSRCSEIT), vol. 8, no. 2, pp. 253–258, 2022.

[23] S. Kumari, P. Giri, S. Choudhury, and S. R. Patil, "Automated resume extraction and candidate selection system," International Journal of Research in Engineering and Technology (IJRET), vol. 3, no. 1, pp. 206–208, 2014.

[24] K. Tejaswini, V. Umadevi, S. M. Kadiwal, and S. Revanna, "Design and development of machine learning based resume ranking system," Global Transitions Proceedings, vol. 3, no. 1, pp. 371–375, 2022.

[25] A. Ikudo, J. Lane, J. Staudt, and B. A. Weinberg, "Occupational classifications: A machine learning approach," SSRN Electronic Journal, 2018.

[26] N. R. Aljohani, M. A. Aslam, A. O. Khadidos, and S.-U. Hassan, "A methodological framework to predict future market needs for sustainable skills management using ai and big data technologies," Applied Sciences, vol. 12, no. 14, p. 6898, 2022.

[27] R. Boselli, M. Cesarini, F. Mercorio, and M. Mezzanzanica, "Using machine learning for labour market intelligence," in Machine Learning and Knowledge Discovery in Databases (ECML PKDD 2017). Springer, 2017, pp. 330–342.

[28] X. Yu, C. Qin, Q. Zhang, C. Zhu, H. Ma, X. Zhang, and H. Zhu, "Disco: A hierarchical disentangled cognitive diagnosis framework for interpretable job recommendation," arXiv preprint arXiv:2410.07671, 2024.

[29] R. Boselli, M. Cesarini, F. Mercorio, and M. Mezzanzanica, "Labour market intelligence for supporting decision making," in Proceedings of the DELFI Workshops 2020, Heidelberg, Germany, 2020.

[30] V. Mittal, P. Mehta, D. Relan, and G. Gabrani, "Methodology for resume parsing and job domain prediction," Journal of Statistics and Management Systems, vol. 23, no. 6, 2020.

[31] R. Pal, S. Shaikh, S. Satpute, and S. Bhagwat, "Resume classification using various machine learning algorithms," in ITM Web of Conferences, vol. 44. EDP Sciences, 2022, p. 03011.

[32] M. A. Aleisa, N. Beloff, and M. White, "Implementing airm: a new ai recruiting model for the saudi arabia labour market," Journal of Innovation and Entrepreneurship, vol. 12, p. 59, 2023.

[33] C. Grosso, N. Sazen, and R. Boselli, "Ai-implemented toolkit to assist users with career "configuration": The case of create your own future," in Proceedings of 26th ACM International Systems and Software Product Line Conference (SPLC'22). ACM, 2022, pp. 1–8.

[34] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "Bert: Pre-training of deep bidirectional transformers for language understanding," in Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers), 2019, pp. 4171–4186.

[35] S. Amin, N. Jayakar, M. Kiruthika, and A. Gurjar, "Best fit resume predictor," International Research Journal of Engineering and Technology (IRJET), vol. 6, no. 8, pp. 813–820, 2019.

[36] M. Saqr and S. López-Pernas, "Why explainable ai may not be enough: predictions and mispredictions in decision making in education," Smart Learning Environments, vol. 11, no. 1, p. 52, 2024.

[37] S. Bharadwaj, R. Varun, P. S. Aditya, M. Nikhil, and G. C. Babu, "Resume screening using nlp and lstm," in Proceedings of the Fifth International Conference on Inventive Computation Technologies (ICICT 2022). IEEE, 2022.

[38] J. Tamayo, L. Doumi, S. Goel, O. Kovács-Ondrejkovic, and R. Sadun, "Reskilling in the age of ai," Harvard Business Review, 2023.

[39] E. Colombo, F. Mercorio, and M. Mezzanzanica, "Applying machine learning tools on web vacancies for labour market and skill analysis," Università Cattolica del Sacro Cuore and CRISP, Tech. Rep., 2018.