Project Proposal Team 149

Tanima Joshi, Mitchell Miller, Vrinda Naik, Prakshi Payal, Utkarsh Saxena, Mansi Saxena

March 28, 2025

1. Introduction [2%]
Currently, job market forecasting relies on historical employment data, economic models, and time-series analysis to predict job demand, salary trends, and required skills. Government reports [3] and industry studies provide foundational insights into labor market trends [1] [2]. Machine learning and NLP techniques enable the extraction of structured insights from unstructured job postings, predicting skill adjacencies and emerging roles [11] [8] [14]. Tools and frameworks offer robust methodologies for time series forecasting and workforce planning [15] [18]. Studies assess automation risks and workforce shifts based on past trends [7] [6] and explore how external shocks like the COVID-19 pandemic accelerate demand for specific skills [9]. Our project intends to target below limitations of the current practices. Most forecasts provide only national-level trends, failing to address localized labor market dynamics [3][2]. Studies often analyze only specific sectors, limiting their applicability to broader workforce planning [5] [16]. Many studies rely on data that is 2–3 years old, resulting in stale and insufficient insights for real-time decision-making [1] [7]. Insights from reports are often broad and do not provide clear, actionable guidance for career planning or skill acquisition [2] [3].

2. Problem Definition [3%]
Our aim is to forecast employment opportunities, trends in salary fluctuations, and the requisite skills for these positions within the United States. This project will empower job seekers by highlighting where opportunities exist and the knowledge they must acquire. It will aid companies in strategizing for the workforce they will require and determining compensation in the future. We aim to create an interactive job market map that displays job availability and regional salary projections for the next 2-3 years.

3. Literature Survey [5%]
Job market forecasting relies on historical employment data, economic models, and time-series analysis to predict job demand, salary trends, and required skills. Government reports [3] and industry studies provide foundational insights into labor market trends [1] [2]. Machine learning and NLP techniques enable the extraction of structured insights from unstructured job postings, predicting skill adjacencies and emerging roles [11] [8] [14]. Tools and frameworks offer robust methodologies for time series forecasting and workforce planning [15] [18]. Studies assess automation risks and workforce shifts based on past trends [7] [6] and explore how external shocks like the COVID-19 pandemic accelerate demand for specific skills [9]. Collectively, these approaches provide a broad understanding of the labor market but lacks actionable insights for stakeholders. Current practices lack of regional insights, provide narrow industry focus, use outdated data usage and provide no direct guidance for job seekers.

4. Proposed method
The project will be executed in phases: first, we will gather and clean employment data from sources like the National Labor Exchange and US DOL. Then, we will analyze regional job trends and use machine learning models to forecast future employment patterns. These insights will be visualized on an interactive map, allowing users to explore job data by region, industry, and time frame.

We take a holistic approach to tackling the problem by giving the user a complete picture of the employment market. This includes **forecasting job demand**, **analyzing salary trends**, and **skill requirements clustering** across location and occupation. We tackle each of these three parts with a different approach, that gives us an overview of the job market and demand in United States.
For job demand and salary trends, we created a forecasting framework using data comprising 10 years of occupational employment data with over 4 million records. Our preprocessing involved cleaning null values and removing duplicates. Our goal is to predict the job demand and salary trend for the next two years in the dataset (2024-2026).

We used a backtesting approach that trained models on nine years of data (2014-2022) and tested against 2023 figures. To forecast the next two years of job demand and salary trends in the US by state, we tested Exponential Smoothing, ARIMA, Prophet, Linear Regression and Extreme Gradient Boosting (XGB) models. We performed a rolling backtest of the above models to make robust, reliable and informed decisions on model selection. Rolling backtest avoids the pitfalls of simply fitting a model to the entire dataset and assuming it will maintain that level of accuracy in the future. Some models might perform exceptionally well on certain segments of the data but poorly on others. A rolling backtest reveals these inconsistencies. Below is a summary of each of the models tested:

Exponential smoothing works by assigning exponentially decreasing weights to past observations. This means that more recent data points have a greater influence on the forecast than older data points. **ARIMA** uses past values of the time series to predict future values, differences the raw observations to make the time series stationary (i.e., to remove trends and seasonality) and produces a moving average. **Prophet** is developed by Facebook's Core Data Science team. It's designed for time series data with strong seasonal patterns and handles missing data and outliers well. **Linear regression** is used to model the relationship between a dependent variable and one or more independent variables by fitting a linear equation to observed data to find the best-fitting line (or plane in multiple regression) that minimizes the difference between the predicted values and the actual values of the dependent variable. **XGB** is an implementation of gradient boosting, a machine learning technique that builds an ensemble of decision trees. It sequentially adds trees, where each tree corrects the errors of the previous ones.

Forecasting Job Demand – We completed backtesting for linear regression, random forest and Prophet to predict employment trends across five major occupational categories (Office/Administrative Support, Food Preparation, Sales, Transportation, and Healthcare Practitioners). In the future, we plan to improve our forecasting framework by adding ARIMA, Exponential Smoothing, and Gradient Boosting techniques for approximately 20 job categories instead of just 5 using the same data.

Analyzing Salary Trends – We completed all the five forecasting techniques on the salary by location data. We plan to choose the model with the lowest percentage error for the testing dataset.

Skills Requirements Clustering – This method outperforms state-of-the-art techniques by offering state-specific, granular analysis of job postings. It uses **Word2Vec** for context-aware job grouping, ensuring more accurate clustering based on actual job descriptions. Combining K-Means and DBSCAN allows flexible and dynamic clustering, with hyperparameter tuning to find optimal cluster sizes. Sub-clustering further refines job groups, providing detailed insights into job prospects. This approach identifies top skills more effectively and offers actionable, localized insights for policymakers and businesses, making it highly scalable and adaptable to evolving job market trends. With thousands of job postings scattered across various sources, we have designed an approach to group similar jobs together and, based on these groupings, identify the top skills required. We perform this analysis for each state, providing insights into the most in-demand skills and corresponding job titles within a specific state. To achieve this, we begin by cleaning the text data, including the Job Title and Job Description, to remove any stop words. We use the Word2Vec encoding model to assign semantic meaning to the descriptions. Afterward, we apply several clustering algorithms, such as **K-Means** and **DBSCAN**, to group similar jobs together. We also perform hyperparameter tuning to determine the optimal value of K and use dendrograms to assist in determining the appropriate clusters for DBSCAN. Once we have grouped similar jobs, we identify the top jobs in each state. To further refine our clusters, we divide them into sub-clusters to represent different job prospects within each state. For instance, if a cluster (Cluster A) contains 200 data points, we split it into smaller sub-clusters to better categorize the job roles. As a result, if Cluster A represents five distinct job prospects, we divide it into five sub-clusters, each corresponding to a unique job title. These titles are then assigned to the state associated with Cluster A, providing a more granular view of the job market within that state.

5. Evaluation

For forecasting job demand and salary trends, our evaluation testbed consists of historical employment data from 2014-2023 with 2023 reserved for testing model accuracy. Similarly, our testbed for clustering consists of a dataset of job postings from various sources, which includes fields such as State, Cluster, Sub-Cluster, Job Title, and Job Description. Questions addressed by the experiment include:

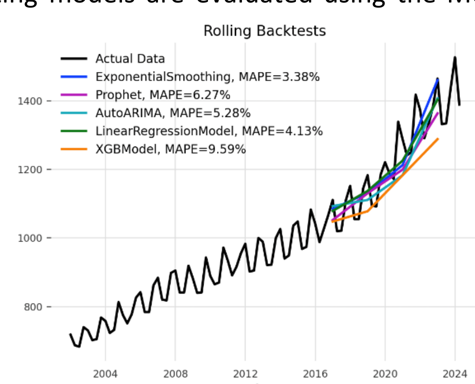
1. What are the most in-demand job titles and skills within each state?
2. How can we cluster similar job postings based on job descriptions, ensuring meaningful groupings?
3. What are the optimal clustering configurations (e.g., value of K for K-Means, DBSCAN settings)?
4. How can sub-clusters help refine job prospects and skill identification within large clusters?
5. How can dynamic and semantic clustering improve the accuracy of job market analysis compared to traditional methods?

Observations:

Forecasting job demand – We found significant differences in model accuracy across occupations, with Random Forest achieving exceptional accuracy for Office/Administrative positions (0.46% MAPE) and Sales occupations (1.30% MAPE), while Linear Regression performed best for Transportation occupations (0.30% MAPE) and Prophet worked best for Healthcare Practitioners (0.45% MAPE). Despite having the best overall MAPE (3.19%), Random Forest produced identical predictions for all future years, making it unsuitable for long-term forecasting. The employment data showed substantial disruption around 2020 for most occupations, with Office/Administrative Support positions showing a sharp decline (from 153M in 2018 to 137M in 2019) that never fully recovered. Healthcare Practitioners showed the most stable employment trajectory with consistent growth, while Transportation occupations displayed a dramatic growth pattern, increasing from 70M in 2017 to 85M in 2019. Our five-year Prophet forecasts revealed divergent trajectories: declining trends for Office/Administrative Support, Food Preparation, and Sales occupations, while Transportation and Healthcare Practitioners showed continued growth. Interestingly, we found that model performance varied significantly by occupation - Linear Regression had poor performance for Food Preparation (10.79% MAPE) and Sales (8.16% MAPE) but excellent results for Transportation (0.30% MAPE), suggesting that occupation-specific model selection is more effective than a one-size-fits-all approach.

Job	Linear Regression MAPE	Random Forests MAPE	Prophet MAPE
Office and Administrative Support Occupations	5.99	0.46	7.10
Food Preparation and Serving Related Occupations	10.79	7.86	10.67
Sales and Related Occupations	8.16	9.89	9.89
Transportation and Material Moving Occupations	0.30	2.27	1.86
Healthcare Practitioners and Technical Occupation	2.77	3.59	0.45

Salary Trends – As shown in the image below, the forecasting models are evaluated using the Mean Absolute Percentage Error (MAPE) where lower MAPE values indicate better forecasting accuracy. The data spans from approximately 2004 to 2024. The models are used to forecast the period from roughly 2018 to 2024. The Exponential Smoothing model has the lowest MAPE of 3.38%, making it the most accurate model among those tested and the XGBModel has the highest MAPE of 9.59%, indicating the poorest performance. The other models fall in between, with the Linear Regression Model offering a reasonably good balance of simplicity and accuracy. With



this result, as we move on to the next steps of our project, we conclude that we will proceed with the Exponential Smoothing model to predict salary trends in the US.

Skills Requirements Clustering – Preprocessing improved the quality of the text, making it easier to capture semantic meaning and group similar job postings. Word2Vec successfully captured the semantic meaning of job descriptions, allowing for more accurate job grouping than traditional keyword matching. Certain job titles that were originally labeled differently (e.g., "Retail Manager" and "Store Manager") were grouped together due to their shared semantic content. K-Means was effective in grouping similar jobs but required fine-tuning to find the optimal K. Used **dendrograms** to help determine the optimal parameters for DBSCAN (e.g., epsilon and min_samples). Sub-clustering helped distinguish different job roles within a large cluster (e.g., distinguishing between "Sales Manager" and "Retail Supervisor" in the same initial cluster.

6. Conclusions and Discussion [5%]

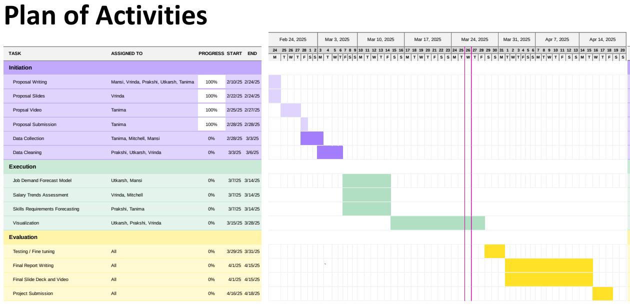
All team members have contributed a similar amount of effort. Below are the key results and takeaways for the three parts of our project:

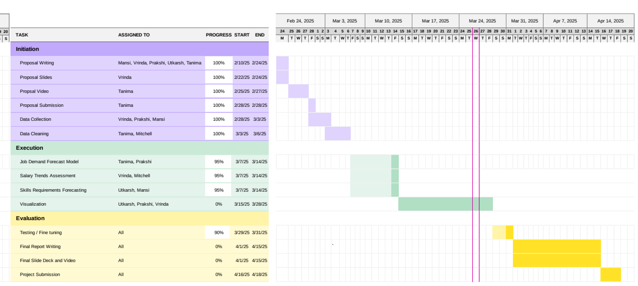
Forecasting Job Demand – We applied 3 forecasting techniques to the job demand data. Model accuracy for employment forecasting varies significantly by occupation, with Random Forest, Linear Regression, and Prophet each excelling in different sectors. Employment data showed major disruptions around 2020, and Prophet forecasts predict divergent future trends across occupations. We will pursue occupation-specific model selection for accurate forecasting, as a one-size-fits-all approach is ineffective.

Salary Trends – We applied 5 forecasting techniques to the job demand data. Exponential Smoothing had the lowest MAPE (3.38%) and was deemed the most accurate model for forecasting salary trends from 2018-2024. XGBModel showed the poorest performance with the highest MAPE (9.59%). Based on these results, Exponential Smoothing was chosen for further salary trend predictions.

Skills Requirements Clustering – Text preprocessing and Word2Vec enhanced job posting clustering by accurately capturing semantic meaning, grouping similar roles like "Retail Manager" and "Store Manager." K-Means and DBSCAN, with parameter tuning via dendrograms, effectively grouped and sub-clustered jobs. Sub-clustering refined large clusters, distinguishing between similar but distinct roles like "Sales Manager" and "Retail Supervisor."

Plan of Activities





Bibliography

1. Carvalho, João. "Preparing for the Future of Work: Trends in Job Market 2030." *GFoundry*, November 25, 2023. https://gfoundry.com/future-of-jobs-2030/.
2. Center for Education and the Workforce, Georgetown University. "After Everything: Projections of Jobs, Education, and Training Requirements through 2031." Accessed February 25, 2025. https://cew.georgetown.edu/cew-reports/projections2031/.
3. DeZarn, Nicholas, Stanislava Ilic-Godfrey, and Emily Krutsch. "Occupational Projections Overview, 2021–31." *Monthly Labor Review*. U.S. Bureau of Labor Statistics, April 2023. https://www.bls.gov/opub/mlr/2023/article/occupational-projections-overview-2021-31.htm.
4. Ellis, Lindsay. "Even Harvard M.B.A.s Are Struggling to Land Jobs." *The Wall Street Journal*, January 15, 2025. https://www.wsj.com/lifestyle/careers/harvard-mba-employment-rate-job-hunt-difficulty-addfc3ec.
5. Fukui, Sadaaki, Wesley Wu, and Michelle P. Salyers. "Applying Machine Learning to Human Resources Data: Predicting Job Turnover among Community Mental Health Center Employees." *The Journal of Mental Health Policy and Economics* 26, no. 2 (June 2023): 63. https://pmc.ncbi.nlm.nih.gov/articles/PMC10424701/.
6. Handel, Michael. "Growth Trends for Selected Occupations Considered at Risk from Automation." *Monthly Labor Review*. U.S. Bureau of Labor Statistics, July 2022. https://www.bls.gov/opub/mlr/2022/article/growth-trends-for-selected-occupations-considered-at-risk-from-automation.htm.
7. Holzer, Harry. "The US Labor Market in 2050: Supply, Demand and Policies to Improve Outcomes." IZA Policy Paper No. 148, 2019.
8. Howison, Mark, Chelsea Lacewell, Morgan Olson, Joosung Lee, Erik Brynjolfsson, Morgan R. Frank, and David Autor. "Extracting Structured Labor Market Information from Job Postings with Generative AI." *Amazon Science*, 2024. https://www.amazon.science/publications/extracting-structured-labor-market-information-from-job-postings-with-generative-ai.
9. Krutsch, Emily, Sara Hylton, and Lindsey Ice. "What the Long-Term Impacts of the COVID-19 Pandemic Could Mean for the Future of IT Jobs." *Beyond the Numbers*. U.S. Bureau of Labor Statistics, February 2022. https://www.bls.gov/opub/btn/volume-11/what-the-long-term-impacts-of-the-covid-19-pandemic-could-mean-for-the-future-of-it-jobs.htm.
10. Liang, Zhang, Xiangmin Liu, and Yitong Hu. "Degrees of Return: Estimating Internal Rates of Return for College Majors Using Quantile Regression." *American Educational Research Journal*, 2024. https://journals.sagepub.com/stoken/default+domain/W9MBX7ZGK8FKK86JZ3Z9/full.
11. Müller, Oliver, and Christop Düsing. "Forecasting the Impact of Artificial Intelligence on the Labor Market through Natural Language Processing of Job Posting." *Social Science Research Network*, January 1, 2024. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4710299.
12. National Bureau of Economic Research. "Projecting Changes in the US Labor Force." 2025. https://www.nber.org/brd-20221/projecting-changes-us-labor-force.
13. Petrongolo, Barbara, Regis Barnichon, and Paula Garda. "The Ins and Outs of Forecasting Unemployment: Using Labor Force Flows to Forecast the Labor Market." *Brookings*, September 2012. https://www.brookings.edu/articles/the-ins-and-outs-of-forecasting-unemployment-using-labor-force-flows-to-forecast-the-labor-market/.
14. Senthurvelautham, S., and N. Senanayake. "A Machine Learning-Based Job Forecasting and Trend Analysis System to Predict Future Job Markets Using Historical Data." In *Proceedings of the 8th IEEE International Conference for Convergence in Technology 2023 (I2CT 2023)*, April 7-9, 2023. IEEE. https://doi.org/10.1109/I2CT57861.2023.10126233.

15. Taylor, Sean J., and Benjamin Letham. "Forecasting at Scale." *PeerJ Preprints*, September 27, 2017. https://peerj.com/preprints/3190/.
16. "Evaluation of the Trends in Jobs and Skill-Sets Using Data Analytics: A Case Study." *Journal of Big Data*. n.d.
17. "On the Radar: Predicting Near-Future Surges in Skills' Hiring Demand to Provide Early Warning to Educators." PMC. n.d.
18. "Optimizing Workforce Efficiency: Leveraging Integrated Business Analytics and Machine Learning for Enhanced Performance Prediction." N.d.
19. Koskinen, L., Salonen, J., & Nummi, T. (2007). Modeling and Predicting Individual Salaries. Presented at the 2nd PBSS Colloquium, Helsinki, Finland. https://www.actuaries.org/pbss/colloquia/helsinki/papers/koskinen.pdf

Appendix

LITERATURE SURVEY (EXTENDED):

Name	**Paper Title**	**Article Link**	**Main Idea**	**Why It Will Be Useful**	**Potential Shortcomings**
Vrinda	Projecting Changes in the US Labor Face (link)	https://www.nber.org/brd-20221/projecting-changes-us-labor-force	The study examines the impact of population aging on labor force participation in the U.S. and Germany, projecting workforce trends up to 2060. Using a dynamic microsimulation model, researchers find that while the U.S. labor force is expected to grow by 16.2% (25.2 million workers) due to population growth, education, and pension reforms, Germany's labor force is projected to decline by 10.7% (4.4 million workers) due to demographic shifts.	By contrasting U.S. projections with those of Germany—where the labor force is expected to decline by 10.7% (4.4 million workers) between 2020 and 2060—the report offers a broader context for understanding how different demographic and policy environments can influence labor force trends.	The projections assume current policies remain unchanged. Future alterations in immigration laws, retirement benefits, or education funding could significantly impact labor force dynamics, potentially rendering these projections less accurate.
Vrinda	Preparing for the Future of Work: Trends in Job Market 2030 (link)	https://gfoundry.com/future-	The study examines the evolving labor market influenced by technological advancements,	The report notes the rapid expansion of the gig economy, significantly	The report provides a global overview but lacks detailed regional or country-specific data, which

		of-jobs-2030/	demographic shifts, and changing consumer preferences. It highlights the dual impact of automation and artificial intelligence (AI), which are expected to automate 42% of tasks by 2027, potentially displacing jobs in routine and repetitive sectors such as data entry and manufacturing. Conversely, these technologies are projected to create approximately 97 million new roles in fields like data science, AI, cybersecurity, and healthcare.	impacting the nature of work and employment patterns, this means it could be helpful to see the effects of other macroeconomic factors on employment.	may limit the applicability of its findings to localized projections.
Vrinda	After Everything: Projections of Jobs, Education, and Training Requirements through 2031 (link)	https://cew.georgetown.edu/cew-reports/projections2031/	The study forecasts that by 2031, 72% of jobs in the U.S. will necessitate postsecondary education or training. This marks an increase from 68% in 2021. On average, there will be 18.5 million job openings annually between 2021 and 2031, with approximately 12.5 million of these requiring at least some college education.	The study delineates a bifurcated economy into managerial and professional economy and blue-collar and skilled-trades economy. This helps further alienate our project into more focused sectors and allows us to take into consideration differences in macro sectors and their needs.	The report, published in March 2024, relies on data projections that may not fully account for unforeseen economic shifts, technological advancements, or global events occurring after its release.

Utkarsh	AI’s Labor Market Impact via NLP of Job Postings	https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4710299	The paper proposes a novel methodology to assess AI's impact on labor markets by analyzing unstructured job postings using NLP and deep learning. Traditional approaches rely on structured databases like O*NET, which have limitations in timeliness and coverage. The authors develop a two-stage model: 1. Task Extraction: A BERT-based model identifies specific tasks from job descriptions with 87.6% accuracy. 2. Feasibility Prediction: A classifier predicts whether each task can be performed by AI within 0-5 years (short-term), 5-10 years (medium-term), or 10+ years (long-term), achieving 73.4% accuracy. Empirical analysis of 12,000 business operations jobs suggests 25% of tasks could be AI-driven within 5 years, rising to 69% within a decade.	Repurposing the paper’s regression analysis to link salary data with skill requirements will be very helpful in the group project. Adapting the BERT-based NLP pipeline to identify skill clusters using SHAP explanations.	The potential shortcomings of the paper include - 1. Narrow occupational scope: The paper focuses exclusively on "business operations" jobs from Monster.com. This should be expanded to multiple sectors (healthcare, tech, manufacturing) using diverse sources like LinkedIn, Indeed, and government databases. 2. Temporal Resolution issues: In light of the rapidly evolving tech landscape, using 5-10 year horizons without real-time adaptability causes problems. This should be changed to tiered forecasting like 0-2 years, 3-5 years and 5-10+ years.

Utkarsh	Generative AI for Labor Market Extraction	https://www.amazon.science/publications/extracting-structured-labor-market-information-from-job-postings-with-generative-ai	The paper discusses application of generative AI to extract structured labor market insights from unstructured online job postings, addressing gaps in traditional survey-based data collection. By analyzing 6,800 job postings divided across 68 occupational groups, the authors developed a scalable method using Amazon Bedrock's foundation models to parse details such as educational requirements, remote-work flexibility, and benefits offerings with high accuracy.	The paper provides actionable insights that directly align with our project's objectives. 1. The GenAI framework extracts structured data from unstructured job postings, enabling analysis of emerging roles and skill adjacencies faster than traditional surveys. 2. Aggregating data at a hyperlocal level and providing forecasts will be a massive value proposition for the project.	There are some key shortcomings in the current research. 1. Skill Demand Analysis Gap: Focuses on high-level attributes (education, remote work) but lacks granular skill extraction (eg. Python vs R proficiency etc.) We can improve upon this by applying NLP to map skill adjacencies (e.g., "TensorFlow → machine learning engineer") and forecast demand surges. 2. Aggregates data nationally, missing metro-specific trends (eg. AI engineering demand in Austin vs Boston). We can improve upon this by getting geographic data from BLS to localize forecasts.
Utkarsh	A Machine Learning-Based Job Forecasting System	https://rgu-repository.worktribe.com/OutputFile/1993469	The paper proposes a machine learning-based system to forecast job market trends, specifically targeting the software industry, to aid job seekers, organizations, and academic institutions in adapting to evolving employment demands. It tries to address	The paper provides a few foundational ideas which can be used in our project. 1. Data Collection Strategy: The paper's use of web scraping to gather 522,180 job listings validates this approach for employment trend analysis.	The paper has the following limitations, which can be improved upon: 1. Skill extraction method: It uses TF-IDF for keyword extraction, which lacks nuance in detecting emerging or context-specific skills. This can be improved by applying BERT or

			gaps in prior research by incorporating real-time data from job boards and combining quantitative analysis (e.g., job posting frequency) with qualitative aspects like skill requirements.	2. Skill Extraction: While the paper uses TF-IDF for keyword extraction, we can improve semantic analysis by applying BERT/GPT-4 embeddings to detect emerging skills.	GPT-4 embeddings for semantic skill clustering. 2. Forecasting approach: The paper uses Bi-LSTM which focuses on historical patterns without validating against unforeseen disruptions. This can be improved by combining quantitative models (ARIMA, Prophet) with qualitative inputs (like Delphi method).
Mansi	Labor Market 2050: Automation & Policy	https://docs.iza.org/pp148.pdf	The main idea of the paper, *The US Labor Market in 2050: Supply, Demand and Policies to Improve Outcomes* by Harry J. Holzer, is an analysis of the anticipated changes in the U.S. labor market by 2050. The paper explores the effects of demographic shifts, automation, and changing employment practices on labor supply and demand. Key points including demographic changes, automation , employment trends etc.	Use real-time job market data (e.g., LinkedIn, Glassdoor, Bureau of Labor Statistics) to refine and validate the paper's projections. Develop predictive models for employment and salary trends using machine learning or statistical methods. Provide industry-specific insights rather than broad economic trends. Incorporate regional labor market variations to make your findings more actionable for job seekers and employers.	There are several potential shortcomings in the paper like Limited Use of Data-Driven Forecasting , lack of granular, industry-specific forecasts, regional analysis and company specific forecasts.

Mansi	ML for Workforce Planning in Mental Health	https://pmc.ncbi.nlm.nih.gov/articles/PMC10424701/	The paper titled "Applying Machine Learning to Human Resources Data: Predicting Job Turnover among Community Mental Health Center Employees" explores the use of machine learning techniques to analyze HR data for predicting employee turnover in community mental health centers. By identifying patterns and factors associated with staff departures, the study aims to enhance retention strategies and reduce turnover rates in these critical healthcare setting	We can apply similar predictive modeling techniques to analyze broader employment trends, such as job demand across industries or salary progression over time. We can expand the variable set to include skills, industry trends, automation impact, and regional economic factors to enhance forecasting accuracy.	The paper focuses on a single sector (mental health centers), making its findings less generalizable to broader employment trends across multiple industries. The study primarily examines job turnover, which is a subset of workforce trends but does not address broader employment opportunities, salary trends, or future skill requirements. The study does not explore how salaries fluctuate over time or what factors influence salary growth within an industry.
Mansi	Forecasting Unemployment Using Labor Force Flows	https://www.brookings.edu/articles/the-ins-and-outs-of-forecasting-unemployment-using-labor-force-flows-to-forecast-the-	This paper presents a model that leverages labor force flow data to predict unemployment rates. By analyzing transitions between employment, unemployment, and non-participation, the model offers real-time forecasts that outperform traditional methods, such as the Survey of Professional Forecasters and the	These insights can help predict job openings and workforce stability by understanding how workers transition between jobs and how industries experience employment fluctuations. We can integrate real-time job postings, wage data, and economic indicators to	The paper primarily forecasts unemployment rates, rather than providing a broader analysis of job opportunities, salary trends, and skill requirements. The paper treats the labor market as a whole and does not differentiate between industries or occupations. The paper does not analyze which skills will be in demand or how workers can

		labor-market/	Federal Reserve Board's Greenbook. This approach provides a more dynamic and accurate tool for anticipating labor market conditions	create up-to-date employment forecasts.	adapt to labor market changes.
Prakshi	Optimizing Workforce Efficiency: Leveraging Integrated Business Analytics and Machine Learning for Enhanced Performance Prediction	(PDF) Optimizing Workforce Efficiency: Leveraging Integrated Business Analytics and Machine Learning for Enhanced Performance Prediction	The paper proposes a framework using business analytics and machine learning to predict employee performance, optimize workforce efficiency, and enable proactive decision-making through systematic data collection, preprocessing, modeling, and validation	Its methodology of leveraging historical and real-time data to forecast trends aligns with your goal of identifying employment opportunities, salary fluctuations, and requisite skills for job seekers and companies.	Reliance on high-quality data, lack of granularity for specific roles/regions, ethical concerns, and potential resistance to adoption may limit its effectiveness despite its robust methodology.
Prakshi	Evaluation of the trends in jobs and skill-sets using data analytics:	Evaluation of the trends in jobs and skill-sets using data analytics: a case study | Journal of Big	The study uses data analytics (LSI, LDA, FA, NMF) to analyze job market trends, skill demands, and education-industry mismatches, focusing on the oil and gas sector.	It provides a data-driven framework to forecast employment opportunities, salary trends, and required skills, aiding job seekers and companies in workforce planning.	It lacks regional specificity and does not address emerging industries, remote work trends, or demographic shifts, limiting its broader applicability.

		Data | Full Text			
Prakshi	On the radar: Predicting near-future surges in skills' hiring demand to provide early warning to educators	On the radar: Predicting near-future surges in skills' hiring demand to provide early warning to educators - PMC	The paper proposes an AI-driven methodology to predict emerging skills in the labor market by analyzing job ad trends, focusing on skills with sudden surges in hiring demand. This approach aims to provide early warnings to educators and training providers, enabling them to adapt curricula and training programs to meet future workforce needs.	This methodology aligns with the goal of forecasting employment opportunities, salary trends, and requisite skills. By identifying emerging skills early, it can help job seekers understand where opportunities lie and what skills to acquire, while assisting companies in workforce planning and compensation strategies. The use of job ad data and predictive analytics offers a data-driven approach to understanding labor market dynamics.	The methodology relies heavily on job ad data, which may not capture all labor market trends, especially in emerging industries or regions with limited job ad availability. It also focuses on granular skill predictions but may lack broader insights into macroeconomic factors, such as recessions or geopolitical shifts, that could impact employment trends. Additionally, the approach may struggle with predicting skills that emerge very rapidly or are not well-represented in job ads.
Tanima	Occupational projections overview, 2021-31	https://www.bls.gov/opub/mlr/2023/article/occupational-projections-overview-2021-31.htm	The paper provides a comprehensive overview of occupational projections in the United States from 2021 to 2031, as developed by the Employment Projections program of the U.S. Bureau of Labor Statistics. It highlights anticipated changes in employment across 24	The paper directly aligns with forecasting job demand by projecting growth trends across different occupational groups and identifying key drivers of employment changes. It provides detailed insights into salary trends by reporting on	While the paper offers a broad national overview, it lacks regional specificity, which could be crucial for understanding localized job market trends. Furthermore, the analysis primarily focuses on occupational groups and does not delve deeply into granular predictions for individual

			occupational groups, including detailed data on growth trends, numeric and percentage changes, factors driving these changes, median annual wages, and educational or training requirements for occupational entry. The paper's primary contribution is serving as a resource for job seekers, career counselors, researchers, and policymakers, offering insights into future labor market demands.	median wages for each occupational group and specific occupations. Additionally, the discussion of typical education and training requirements for occupational entry supports skills analysis, helping individuals and organizations understand the competencies needed to meet future labor market demands.	occupations beyond some examples. Additionally, it does not address the potential impact of emerging industries, remote work trends, or international factors on the U.S. job market. The paper also does not explicitly consider how demographic shifts or socioeconomic disparities might affect workforce opportunities across different groups.
Tanima	Growth trends for selected occupations considered at risk from automation	https://www.bls.gov/opub/mlr/2022/article/growth-trends-for-selected-occupations-considered-at-risk-from-automation.htm	The article "Growth trends for selected occupations considered at risk from automation," published in the Monthly Labor Review in July 2022, examines the employment growth trends of occupations identified as susceptible to automation. It critically analyzes previous studies that have predicted significant job displacement due to automation, such as the claim that 47 percent of U.S. jobs	By analyzing historical employment data of occupations deemed at risk from automation, the article offers valuable insights into job demand trends. Understanding which occupations have experienced growth or decline despite automation concerns can inform future job demand forecasts. While the article primarily focuses on employment trends, it indirectly	The paper does not provide granular predictions about future employment trends or consider emerging technologies' potential impact on these occupations.

			were at risk between 2010 and 2030. The paper contributes to the discourse by providing empirical data on how these at-risk occupations have fared in terms of employment growth, offering a nuanced perspective on the actual impact of automation on the labor market.	informs salary trends and skills analysis by identifying which occupations remain resilient, suggesting a continued or evolving demand for specific skill sets.	
Tanima	What the long-term impacts of the COVID-19 pandemic could mean for the future of IT jobs	https://www.bls.gov/opub/btn/volume-11/what-the-long-term-impacts-of-the-covid-19-pandemic-could-mean-for-the-future-of-it-jobs.htm	This article examines how the COVID-19 pandemic has influenced employment projections for Information Technology (IT) occupations. It highlights that, prior to the pandemic, IT jobs were already expected to see significant growth; however, the pandemic has further amplified the importance of IT professionals, leading to increased employment projections in this sector. The article provides insights into how the accelerated adoption of digital technologies and remote work has heightened the demand for IT expertise.	The article directly relates to forecasting job demand by presenting updated employment projections for IT occupations in the context of the pandemic. It suggests that the surge in remote work and digital transformation has led to a heightened demand for IT professionals, potentially influencing salary trends upward due to increased competition for skilled workers. Additionally, the article implies a shift in the requisite skills for IT positions, emphasizing the need for expertise	The analysis does not delve into the granularity of predictions for individual IT roles, nor does it address potential disparities in employment opportunities across different regions or industries. Furthermore, the article does not explore the long-term sustainability of the increased demand for IT professionals or how emerging technologies, such as artificial intelligence, might impact future job prospects and skill requirements within the IT sector.

				in areas such as cybersecurity, cloud computing, and IT infrastructure to support remote operations.	
Mitchell	Degrees of Return: Estimating Internal Rates of Return for College Majors Using Quantile Regression	https://www.aera.net/Newsroom/Degrees-of-Return-Estimating-Internal-Rates-of-Return-for-College-Majors-using-Quantile-Regression	Going to college is still valuable but it depends on the degree	Answers the question about whether college is worth it	Assumes that the majors valued in the past will continue to be valued in the future.
Mitchell	Even Harvard M.B.A.s Are Struggling to Land Jobs	https://www.wsj.com/lifestyle/careers/harvard-mba-employment-rate-job-hunt-difficulty-addfc3ec	A MBA may not be as valuable as it once was	Signals that the job market may be valuing different skills	The desired skills for the job market will change over time and this looked at initial job placement not necessary career growth.
Mitchell	Forecasting at Scale (Prophet)	https://peerj.com/prepr	Prophet can be a strong time series forcasting method	Justifies why we may want to use prophet to forecast the job	Relies on historical patterns being persist.

		ints/3190/		market and salary growth	
Mitchell	Modeling and Predicting Individual Salaries: A Study of Finland's Unique Dataset	https://www.actuaries.org/pbss/colloquia/helsinki/papers/koskinen.pdf	Wages for the middle quartile earners can be predicted more accurately than the lower or upper quartiles.	Wages can be more accurately predicted during times of economic growth.	The data is from Finland and may not be representative for the world