International Research Journal of Innovations in Engineering and Technology (IRJIET)

ISSN (online): 2581-3048

Volume 6, Issue 4, pp 1-8, April-2022

https://doi.org/10.47001/IRJIET/2022.604001

Applying ARIMA Model to Predict Future Jobs in the Saudi Labor Market

[1]Mashael Alyahya, [2]Mohammed Hadwan

[1,2]Department of Information Technology, College of Computer, Qassim University, Buraydah, Saudi Arabia

[2]Department of Computer Science, College of Applied Sciences, Taiz University, Taiz, Yemen

***Abstract* - The labor force is one of the most critical components of advanced countries, as an investment in the labor force leads to the country's economic development. The development of the labor force and human resources requires the development of education to obtain outputs that comply with the labor market requirements. In recent years, technology has become the dominant factor in the development of countries as strength began measured by technical progress. Therefore, this paper focused on predicting future technical jobs for the KSA, especially the government sector. The used dataset is named Government job advertisements (GJA) and owned by the Ministry of Human Resources and Social Development (MHRSD). The Autoregressive Integrated Moving Average (ARIMA) was developed to obtain the results of future jobs. Root Mean Square Error (RMSE), Mean Absolute Error (MAE), Mean Squared Error (MSE), and Mean Absolute Percentage Error (MAPE) are the elevation metrics that will be used to assess the performance of the ARIMA. Results demonstrated that the ARIMA achieves 5.45 of RMSE, 2.19 MAE, 5.44 MSE, and 4581076540077221.0 MAPE.**

Keywords: ARIMA, Machine Learning, Predicting jobs, KSA, Saudi Arabia.

I. INTRODUCTION

Future job prediction has been considered one of the trend subjects in recent years. In general, job prediction Influenced by advances in technology since the 1980s, technology has shaped the labor market changes and made small to medium forecast [1]. Employability plays a significant role in individuals' life as it is known as the ability of the graduate to get a satisfying job [2]. In other words, employability is defined as the ability to gain and retain fulling work [3]. Nevertheless, there are two sides to employability: individuals' skills (hard or soft skills) and the labor market demand to be fitted. On the other side, unemployment is widespread in many countries, where individuals are directly affected by it, known as "persons above a specified age not being in paid employment or self-employment but currently available for work during the reference period" [4]. Based on the General Authority for Statistics (GAS), figure 1 shows the unemployment rate for Saudi and non-Saudi individuals from 2016-2021, where the unemployment rate is defined as the ratio of "unemployed people to the labor force (employed and unemployed) at working-age population (15+ years)" [5]. The unemployment rate of Saudi males has increased, while the female unemployment rate has decreased through the years. Global phenomena also significantly affect the level of unemployment, such as the spread of the Coronavirus (Covid-19) affected the employment of individuals in medium to large companies, as 27% of companies reported that they would reduce the employment of graduates this year. As the technology changing the global workforce landscape, the merging technology in predicting future jobs facilitates the process of finding the right job for the right candidate [6]. Many researchers use Machine Learning (ML) techniques to predict long-term labor market needs, which can be defined as "predicting future employment opportunities in the labor market by predicting the factors that affect the demand for skills or competencies, as imposed by the company" [7].

To be in line with the objectives of the Ministry of Human Resources and Social Development(MHRSD)which improve skill's level that meets labor market, this paper aims to predict future jobs of Saudi government epically in Information Technology (IT) sector [8]. After reviewing previous studies and various predictions and results reached by researchers, the Autoregressive Integrated Moving Average (ARIMA) model will be developed. ARIMA has demonstrated the high and efficient prediction performance due to the various studies.

	Unemployment Rates					
	Saudi			Non Saudi		
	Male	Female	Total	Male	Female	Total
2016 Q2	5.4	33.7	11.6	0.4	2.3	0.6
2016 Q3	5.7	34.5	12.1	0.6	2.9	0.8
2016 Q4	5.9	34.5	12.3	0.4	1.6	0.5
2017 Q1	7.2	33.0	12.7	0.5	2.1	0.7
2017 Q2	7.4	33.1	12.8	0.7	3.3	0.9
2017 Q3	7.4	32.7	12.8	0.4	1.6	0.5
2017 Q4	7.5	31.0	12.8	0.5	2.5	0.7
2018 Q1	7.6	30.9	12.9	0.7	2.6	0.9
2018 Q2	7.6	31.1	12.9	0.5	2.5	0.7
2018 Q3	7.5	30.9	12.8	0.6	3.1	0.9
2018 Q4	6.6	32.5	12.7	0.6	4.4	1.0
2019 Q1	6.6	31.7	12.5	0.4	2.5	0.6
2019 Q2	6.0	31.1	12.3	0.2	0.9	0.3
2019 Q3	5.8	30.8	12.0	0.2	1.0	0.3
2019 Q4	4.9	30.8	12.0	0.3	1.3	0.4
2020 Q1	5.6	28.2	11.8	0.4	2.0	0.5
2020 Q2	8.1	31.4	15.4	2.3	9.5	3.1
2020 Q3	7.9	30.2	14.9	1.9	9.1	2.7
2020 Q4	7.1	24.4	12.6	1.7	9.1	2.6
2021 Q1	7.2	21.2	11.7	1.3	5.5	1.9

Figure 1: The unemployment rate for Saudi and non-Saudi individuals from 2016-2021

International Research Journal of Innovations in Engineering and Technology (IRJIET)

ISSN (online): 2581-3048

Volume 6, Issue 4, pp 1-8, April-2022

https://doi.org/10.47001/IRJIET/2022.604001

The main contribution of this paper is using the government job datasets to predict the future of the labor market of KSA. Also, the use of the ARIMA model is an excellent contribution to this paper, as no previous studies are using it with the government jobs dataset.

II. LITRATURE REVIEW

This section will focus on recent research papers related to predicting future jobs collected from various trusted scientific journals and research gates. Many research paper focusing on different aspect of kingdom of Saudi Arabia as in [9]–[13]. This research paper focused on predicting labor market in Saudi Arabia. The research papers will be presented in ascending order, based on publication, from 2018 to 2022. Moving to the Automotive Job Market in Morocco, this research paper [14] focused on predicting the needs of the Automotive Job Market and the reasonable salary by using job offers. The adopted ML methods were Decision tree (DT), Gradient boosting, and Random Forest (RF). Researchers collected the dataset from Moroccan job portals from January 2017 until January 2018 and gained 8027 vacancies after deleting the redundant data. As a result, RF outperforms other ML methods and achieves 83p% accuracy.

In some cases, the Appearance of inaccurate job matching depends on the semi-structured resumes. In this research paper [15], the researchers proposed an approach to effectively matching semi-structured resumes and job descriptions. They used six commonly used models: Word n-grams, TF-IDF, bag-of-words (BOW), Bag-of-Means, Doc2Vec, and CNN. The dataset contained 1314 resumes 3809 job descriptions (JDs) from different domains. The used evaluation metrics were Accuracy, Precision, Recall, and f1-score. In the end, results indicated that TF-IDF was better than BOW and Word n-gram.

Traditional methods take a long time to find the right student for a job based on the required skills. Hence in 2018 [16], researchers used NB, DT, Neural Network, and KNN with the Weka -3.4.9 tool to find the likelihood of finding a suitable job in a company. They compared the techniques by different evaluation metrics and the time taken for the classification process. According to the researchers' analysis results, the KNN outperformed the other algorithms with the lowest error rate. NB achieved the least time to analyze the data while DT took the longest processing time compared to all different algorithms.

In 2018, [17] proposed a CNN model for matching job seekers to suitable jobs. The proposed model was named Person-Job Fit Neural Network (PJFNN), learned from a Chinese company's previous job application records. The experiments showed the efficiency of the proposed model based on different evaluation metrics Area Under Curve (AUC) and P-value.

The main goal [20] was to propose a Deep Neural Network model (DNN) to predict future job specifications such as payroll and position name. The researchers compared the proposed model with SVM, RF, and Extreme Gradient Boosting (XGB). They applied ML models on a real-world dataset that contained 70,000 Chinese resumes. By evaluated the different ML models, the experiments proved that the proposed DNN model achieves high accuracy.

In present days ML techniques are widely used for prediction tasks. The researchers of [21] used ML models to predict the job's salary based on specific points to solve the salary specification challenges. The used dataset contained 244,768 data points. The researcher used DT and RF models and evaluated them based on Accuracy, Mean absolute error (MAE), Mean squared error (MSE), f1_score (macro), and f1_score (weighted). Lately, they found that the RF classifier achieved high accuracy more than the DT classifier.

The industry's success mainly depends on the workers' skills. One of the main issues facing any industry is finding a proper candidate to fill the gap of another skilled worker leaving the job. [22] solved the issues by ML models and compared them based on Accuracy. The ML models: AdaBoost (AB), DT, Gradient Boost (GB), KNN, Linear Discriminant Analysis (LDA), LR, Multi-Layer Perceptron (MLP), NB, RF, Stochastic Gradient Descent Classifier (SGD), and SVM. The used dataset consisted of 53 characteristics of 17,588 players and was collected from the Kaggle website [23]. Results showed that LDA outperformed the other models.

Due to the volatile employment structure, many workers leave their job unexpectedly which lead a problem to the company stability. [24] researchers designed a model to predict the chances of a worker to either leave or stay in the position. They used RF, XGBoost, CatBoost, and LightGBM models and applied them to the analyticsvidhya.com dataset [25]. What is more, results proved that CatBoost had the highest accuracy among other models.

Job prediction models can improve the computing environment, defined as attaching a network of computers in various locations to make a task. Furthermore, job prediction models can predict the actual run time of the jobs to improve their work. This research paper [26] focused on predicting the run time of employment using different prediction models: linear regression, weight least square, multi-layer perceptron, and the proposed model. The evaluation metrics to test the models were prediction rate, Sum Square Error (SSE), and Root Mean Square Error (RMSE). The results showed that

ISSN (online): 2581-3048
Volume 6, Issue 4, pp 1-8, April-2022
https://doi.org/10.47001/IRJIET/2022.604001

using linear regression with a smoothing spline fitting model achieved high efficiency.

This research paper [27] focused on analyzing the job logs where researchers collected for many years. The researchers compared different ML algorithms, multivariate linear regression models, multivariate polynomial models, linear neural networks, and BP neural networks. They used a dataset that includes 1,241,163 entries of job logs. The experimental results showed that the researchers obtained a reasonable accuracy using different ML algorithms.

Later in 2019, the researchers of [28] used different Supervised ML methods to predict employee turnover within an organization. They used DT, RF, Gradient Boosting Trees, XGB, LR, SVM, Neural Networks, Linear Discriminant Analysis, NB, and KNN methods. They used two datasets; the first dataset collected from the United States of America bank contained 14,322 employee entries. The second dataset created by IBM Watson Analytics contained 1,470 employee entries. The results showed that they differ according to the size and complexity of the data, but the XGB method proved to be the best for complex and large data.

By focusing on the Saudi labor market, this research paper [29] extracted data from online job portals named LinkedIn and Bayt.com to examine the demand of recruitment specialty for Information Technology students [30][31]. It focused on recommending the highest need of IT skills to improve the outcomes of IT students and solve the unemployment problem. The researchers used different ML algorithms, which are SVM, NB, and KNN.

One of the challenges a data scientist faces is the high dimensionality of datasets and how to extract valuable knowledge. In this research paper [32], the researchers proposed an efficient predictive system suited to the high dimensional dataset by improving scalable random forest (SRF) algorithm on a platform named Apache Spark. They used the proposed model with two techniques of data reduction techniques, Principal Component Analysis (PCA) and Information Gain (IG). In addition, they applied the proposed model to a real-world dataset from Parallel Workload Archives (PWA). The experimental results showed that the proposed model with the dimension reduction techniques achieved high performance compared with the RF algorithm.

This German research paper [18], researchers collected data from job advertisements from 1950 to 2014. The bidirectional Long Short-Term Memories (BiLSTMs) model was used and proved its worth with 89.8% accuracy. Recently, advertising jobs in web pages have played a significant role in the labor market and provide an opportunity to control and monitor the labor market. In this research paper [19], researchers proposed an approach to classify online job advertisements through ML techniques in text classification. First, they evaluated different ML techniques SVM Linear, SVM RBF Kernel, RF, and Artificial Neural Networks (ANNs) on a dataset contained 75 546 job vacancies Italian. After evaluating the classifiers, they found that SVM Linear outperforms all other ML techniques, which obtained 0.93 of Precision, Recall, and F1-score.

Most job seekers' information such as gender, age, education, and company scale can predict future job details. The researchers of [33] proposed a CNN model with five layers: one input layer, convolution layer with a ReLU activation, max pooling and dropout, Dense layer, and softmax layer to predict future job information. They verified the model on a real-world dataset contained 70,000 resumes.

Job involvement plays an essential role in organizations. Therefore, [34] developed a prediction model able to predict employees' job involvement. The developed model depended on a generalized linear model (GLM), including LR and binomial classification. And the dataset was available at international business machines (IBM). [34] Researchers used Accuracy, Area Under the Curve (AUC), Precision, Recall, and F-measure to evaluate the models' performance. ML models achieved the best results in terms of accuracy.

Most graduated students confine themselves to a particular domain; therefore, the competition in the specific field will be high. Such a problem was the primary goal of this research paper [35]. The researchers tried to find a suitable solution to prevent and avoid searching for jobs in attractive domains. The researchers used job category and the individuals' location to predict the job openings applying the LSTM model. The results of the LSTM model compared with the Simple Recurrent Neural Network (SRNN) Model and indicated that the LSTM model was 96% effective.

Recently, [36] focused on determining a suitable job for job seekers based on job details and required skills. They used various Deep Neural Network models TextCNN, Bi-GRU-LSTM-CNN, and Bi-GRU-CNN, where the input was the IT job details and the output was the predicted job. The different models were evaluated based on four evaluation metrics: Accuracy, Precision, Recall, and F1-score. The results indicated that the Bi-GRU-CNN model proved its efficiency and achieved the highest marks.Many job portals published in recent years post job openings that benefit learners, job seekers, and education administrators. For example, in Vietnam, [37] proposed a system that collected and integrated a vast jobs dataset to predict job opportunities and recommend the required courses for a specific job. They used time-series

ISSN (online): 2581-3048
Volume 6, Issue 4, pp 1-8, April-2022
https://doi.org/10.47001/IRJIET/2022.604001

forecasting models ARIMA and LSTM to predict future employment and applied them to a dataset contained more than 700.000 job advertisements. The results showed that ARIMA outperforms LSTM.

Resume matching systems provide different advantages: ease the decision-making process, explore the suitable candidate, and reduce time-consuming. Regarding a large number of resumes, these systems should classify the users to the correct categories carefully. [38] Proposed a system divided into two steps classification and recommendation. Through the classification process, the researchers compared four models: RF, Multinomial NB, LR, and Linear Support Vector Classifier (Linear SVM). What is more, the recommendation models were Content-Based Recommendation using Cosine Similarity and KNN. Dataset collected from an online job portal and Kaggle website. The evaluation metric was Accuracy, and the results presented that Linear SVM performed best for other ML classifiers with 78.53% of Accuracy.

In a similar context, the researchers of [39] examined the future jobs in light of the Neom project in Saudi Arabia. They convened with ten experts in different fields such as Technology, Science, and Human Resources Planning and used the Delphi method to achieve their goal. The future jobs in the labor market are affected by many changes: the development of technology, the spread of the Coronavirus, the expansion of patents. Although patents directly affect the labor market, there is no way to link job information with patents. To consider this problem, the researchers [40] used the classification code to classify similar patents and match them to similar jobs. The used dataset was collected from the O∗ NET website and the classification codes from the website of CPC [41][42]. The researchers calculated the similarities between job representation and patent classification codes to gain a result. The results indicated that the approach is reasonable but limited for some cases.

III. METHODS

3.1 Data collection

This paper uses the GJA dataset [43] for the years (1437-1441), which means from 2016 - 2020, as shown in table 1.

Table 1: The years in Hijri and Gregorian

Hijri Years	Gregorian Years
1437	2016
1438	2017
1439	2018
1440	2019
1441	2020

The GJA dataset is freely public through the Open Data web portal [44]. The dataset published and owned by the Ministry of Human Resources and Social Development (MHRSD) was created on November 8, 2020, and the last update was on November 24, 2020.

3.2 Data Preprocessing

Most datasets require processing and cleaning which means removing inconsistencies from data to improve data quality [45]. Regarding the GJA dataset, preprocessing is crucial before using it in the prediction models. Listed below are the preprocessing steps that have been applied to the GJA dataset:

- Remove unimportant columns:

The dataset contains unnecessary columns such as the type and date of the advertisement and the advertisement sector. For the cleaning processes, the removed columns are: advertisement date and type, headquarter, headquarter area, the sector of advertisement, gender, rank.

- Remove irrelevant data:

Since the main objective of this paper is to predict the IT future jobs, the other jobs are out of the research scope. The dataset contains multiple jobs such as teacher, writer, legal researcher, photographer, etc. therefore, all jobs unrelated to the IT sector were removed. After removing 67,430 irrelevant records, the remaining IT jobs are 1487 records.

- Convert the dataset:

The original dataset is published in Arabic language; therefore, the researchers convert all records into English to use the prediction models ideally.

3.3 ARIMA model

In order to achieve the main goals of this paper, the researchers used the ARIMA model as it is considered as one of the linear models in time series forecasting.

3.3.1 Applying the ARIMA model on a single timeline

In order to apply the ARIMA model, the researchers initially used a single timeline to ensure the quality of the outputs and results. First, the researchers uploaded the GJA dataset. After that, the dataset was divided into three columns: the job type, the date after it was modified, the number of each job repetition. The dataset would be equipped according to the requirements of the ARIMA model, as it deals with the time series dataset. Various Python libraries were called to train the ARIMA model, such as the Stats models, the open-source library for statistical analysis in Python [46]. The ARIMA

International Research Journal of Innovations in Engineering and Technology (IRJIET)

ISSN (online): 2581-3048

Volume 6, Issue 4, pp 1-8, April-2022

https://doi.org/10.47001/IRJIET/2022.604001

model requires training and testing datasets; the training dataset consisted of 85 samples, whereas the testing dataset consisted of 40 samples. Then the Pandas tool was used to normalize the data within 0 and 1 range.

3.3.2 Applying the ARIMA model on multiple timelines

In the next phase of the ARIMA model application, the model was applied on multiple timelines covering the whole dataset. The ARIMA model applied to each job type through the range of years. Next figure 2 presents the results of using ARIMA for Computer engineer where the RMSE = 0.09, MAE = 0.22, MSE = 0.084, and MAPE = 352734435743071.1. More to that, figure 3 shows the plot of the computer engineer prediction result.

```
Results calculation:
ARIMA Test Score for computer engineer : 0.09 RMSE
MAE: 0.22166294924779323
MSE: 0.08453728725056071
MAPE: 352734435743071.1
Date by months 1: 0
Date by months 2: 3
Date by months 3: 0
Date by months 4: 0
Date by months 5: 0
Date by months 6: 0
Date by months 7: 0
```

Figure 2: Applying the ARIMA on the Computer engineer job

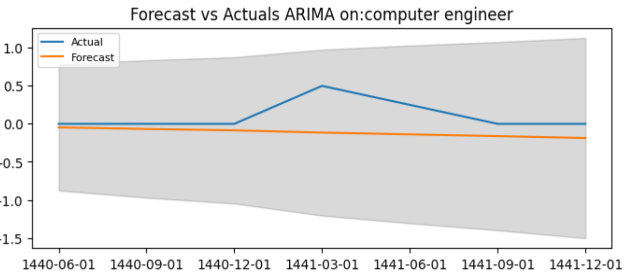

Figure 3: Plot of the Computer engineer prediction result

3.4 Accuracy Metrics for Time-series Prediction

The primary goal of measuring the performance of ML models is to improve the performance and reach a logical result [47]. Different accuracy metrics for time-series predictions are used to show the model performance:

- Root-Mean-Square- Error (RMSE): is always employed for evaluation goals. It is used as a standard for many types of research: air quality and climate.
- Mean Absolute Error (MAE): is a simple evaluation metric that sums the absolute values of the error to get the total errors and divides the total errors by n[48].
- Mean Square Error (MSE): compared prediction techniques for accuracy under quadratic loss[49].
- Mean Absolute Percentage Error (MAPE): usually used when different observations have few series with large values[50].

IV. RESULTS AND DICUSSION

To discover the significant findings, the ARIMA model was assessed using a variety of metrics: RMSE, MAE, MSE, and MAPE. The RMSE low value indicated that the ARIMA obtained quite good performances. The MAE values are always close to the RMSE, which means the errors are of the same magnitude. If the values have a significant difference, that means high errors of the samples. The MSE and MAPE have small values; the model has good performance.

The ARIMA performances are as follow: 5.45 of RMSE, where the MAE= 2.19, 5.44= MSE, and 4581076540077221.0 MAPE. The reason behind the MAPE value of the ARIMA model is that the undefined values “zero values” appear in the actual dataset. As a result, the ARIMA model performs good.

As this paper aimed to predict the future jobs in the KSA; the computer teacher of high school, senior software developer, system analyst, and software developer were the most desirable jobs in the future of KSA.

V. CONCLUSION

In this paper, recent research papers related to predicting future jobs were collected. Each research paper specified different objectives and covered various jobs types. As this paper focused on the KSA government jobs, the GJA dataset was collected and prepared to apply the ARIMA prediction model. The dataset consisted of five years with 1478 IT jobs. The scope was confined to IT jobs, regarding the developments of technology in the whole world and the initiative of the KSA to digital transformation. The RMSE, MAE, MSE, and MAPE were used to measure the models' performance for experimental results. In particular, the ARIMA model achieves5.45RMSE, 2.19MAE, 5.44MSE, and 4581076540077221.0 MAPE. This paper contributed to predicting the future jobs in the KSA, which directly affects individuals and many different sectors: the education and labor sectors. The computer teacher of high school, senior software developer, system analyst, and software developer were the most desirable jobs in the future of KSA.

VI. FUTURE WORK

This research can be extended and developed through:

- Compare the ARIMA model with different ML methods.
- Integrate the government and private sector jobs.
- Expand the scope and used different jobs rather than IT-related jobs.

ISSN (online): 2581-3048
Volume 6, Issue 4, pp 1-8, April-2022
https://doi.org/10.47001/IRJIET/2022.604001

REFERENCES

[1] A. Adams, “Technology and the labour market: The assessment,” *Oxford Review of Economic Policy*, vol. 34, no. 3, pp. 349–361, 2018, doi: 10.1093/oxrep/gry010.

[2] L. Harvey, “Defining and Measuring Employability,” vol. 7, no. August 2012, pp. 37–41, 2010, doi: 10.1080/1353832012005999.

[3] J. Hillage and E. Pollard, “Developing a framework for policy analysis. Research brief, 85,” *Institute for Employment Studies*, no. December, 1998.

[4] T. I. Conference, L. Statisticians, T. Thirteenth, I. Conference, L. Statisticians, E. Conference, R. Iii, E. Conference, E. Conference, R. Iii, and E. Conference, “Resolution concerning statistics of the economically active population, employment, unemployment and underemployment, adopted by the Thirteenth International Conference of Labour Statisticians,” *The Thirteenth International Conference of Labour Statisticians.*, no. i, 1982.

[5] S. A. General Authority for Statistics, “Labor Force,” *Labor market statistics Q1 2021*, 2021. https://www.stats.gov.sa/sites/default/files/LM_Q1 2021 %28Press release_EN%29_0.pdf (accessed May 20, 2021).

[6] R. Miller, “Jobs for the future,” *Industrial Heating*, vol. 74, no. 1, p. 12, 2015.

[7] N. Mezhoudi, R. Alghamdi, R. Aljunaid, G. Krichna, and D. Düştegör, “Employability prediction: a survey of current approaches, research challenges and applications,” *Journal of Ambient Intelligence and Humanized Computing*, 2021, doi: 10.1007/s12652-021-03276-9.

[8] “Ministry of Human Resources and Social Development (MHRSD).” https://www.bing.com/search?q=Ministry+of+Human+Resources+and+Social+Development+objectives&cvid=f0d0879e019b4b4fb90308666a26e231&aqs=edge..69i57.6745j0j1&pglt=43&FORM=ANNTA1&PC=EDGEDB (accessed Oct. 01, 2021).

[9] S. S. Alrumiah and M. Hadwan, “Implementing big data analytics in e-commerce: Vendor and customer view,” *IEEE Access*, vol. 9. 2021. doi: 10.1109/ACCESS.2021.3063615.

[10] N. Alajlan, M. Hadwan, and D. M. Ibrahim, “EFFECTIVENESS OF ADOPTING CLOUD-BASED E-LEARNING AT QASSIM UNIVERSITY,” *Computer journal*, pp. 1–9, 2020.

[11] A. Albelaihi, M. Alyahya, N. Aldieef, N. Alghasham, N. Alajlan, R. Alharbi, and M. Hadwan, “Social Commerce in Saudi Arabia: A literature review,” *International Journal of Engineering Research and Technology*, vol. 12, no. 12, pp. 3018–3026, 2019.

[12] S. S. Alfaozan, H. A. Alghofaily, M. A. Al-Hagery, and Mohammed A. Hadwan, “Knowledge Extraction Methods as a Measurement Tool of Depression Discovery in Saudi Society,” *International Journal of Information Technology and Computer Science*, vol. 12, no. 4, pp. 1–10, 2020.

[13] M. Hadwan, R. U. Khan, and K. I. M. Abuzanouneh, “Towards a smart campus for Qassim university : an investigation of Indoor navigation system .,” *Advances in Science, Technology and Engineering Systems Journal*, vol. 5, no. 6, pp. 831–837, 2019.

[14] I. Makdoun, G. Mezzour, K. M. Carley, and I. Kassou, “Analyzing the needs of the automotive job market in Morocco,” *13th International Conference on Computer Science and Education, ICCSE 2018*, no. Iccse, pp. 210–215, 2018, doi: 10.1109/ICCSE.2018.8468854.

[15] S. Maheshwary and H. Misra, “Matching Resumes to Jobs via Deep Siamese Network,” *The Web Conference 2018 - Companion of the World Wide Web Conference, WWW 2018*, no. April, pp. 87–88, 2018, doi: 10.1145/3184558.3186942.

[16] T. Ravi Kumar, P. Yasaswini, G. Rafi, and D. V. Krishna, “Comparative analysis on job prediction of students based on resume using data mining techniques,” *International Journal of Engineering and Technology(UAE)*, vol. 7, no. 2, pp. 1100–1103, 2018, doi: 10.14419/ijet.v7i2.7.12236.

[17] C. Zhu, H. Zhu, H. Xiong, C. Ma, F. Xie, P. Ding, and P. Li, “Person-job fit: Adapting the right talent for the right job with joint representation learning,” *arXiv*, vol. 1, no. 1, pp. 1–16, 2018.

[18] A. S. Gnehm, “Text zoning for job advertisements with bidirectional LSTMs,” *CEUR Workshop Proceedings*, vol. 2226, no. June, pp. 66–74, 2018.

[19] R. Boselli, M. Cesarini, F. Mercorio, and M. Mezzanzanica, “Classifying online Job Advertisements through Machine Learning,” *Future Generation Computer Systems*, vol. 86, pp. 319–328, 2018, doi: 10.1016/j.future.2018.03.035.

[20] Y. Deng, H. Lei, X. Li, and Y. Lin, “An improved deep neural network model for job matching,” *2018 International Conference on Artificial Intelligence and Big Data, ICAIBD 2018*, pp. 106–112, 2018, doi: 10.1109/ICAIBD.2018.8396176.

[21] S. Dutta, A. Halder, and K. Dasgupta, “Design of a novel prediction engine for predicting suitable salary for a job,” *Proceedings - 2018 4th IEEE International Conference on Research in Computational Intelligence and Communication Networks, ICRCICN 2018*, pp. 275–279, 2018, doi: 10.1109/ICRCICN.2018.8718711.

[22] V. Mathew, A. M. Chacko, and A. Udhayakumar,

ISSN (online): 2581-3048
Volume 6, Issue 4, pp 1-8, April-2022
https://doi.org/10.47001/IRJIET/2022.604001

"Prediction of suitable human resource for replacement in skilled job positions using Supervised Machine Learning," *Proceedings of the 2018 8th International Symposium on Embedded Computing and System Design, ISED 2018*, pp. 37–41, 2018, doi: 10.1109/ISED.2018.8704120.

[23] "kaggle2."

[24] P. Pulicherla, T. Kumar, N. Abbaraju, and H. Khatri, "Job Shifting Prediction and Analysis Using Machine Learning," *Journal of Physics: Conference Series*, vol. 1228, no. 1, 2019, doi: 10.1088/1742-6596/1228/1/012056.

[25] "analyticsvidhya."

[26] H. Al-Najjar, S. S. N. Alhady, and J. M. Saleh, *Improving a run time job prediction model for distributed computing based on two level predictions*, vol. 547. Springer Singapore, 2019. doi: 10.1007/978-981-13-6447-1_5.

[27] Z. Hou, S. Zhao, C. Yin, Y. Wang, J. Gu, and X. Zhou, "Machine learning based performance analysis and prediction of jobs on a HPC cluster," *Proceedings - 2019 20th International Conference on Parallel and Distributed Computing, Applications and Technologies, PDCAT 2019*, pp. 247–252, 2019, doi: 10.1109/PDCAT46702.2019.00053.

[28] Y. Zhao, M. K. Hryniewicki, F. Cheng, B. Fu, and X. Zhu, *Employee turnover prediction with machine learning: A reliable approach*, vol. 869, no. January. Springer International Publishing, 2019. doi: 10.1007/978-3-030-01057-7_56.

[29] M. Alghamlas and R. Alabduljabbar, "Predicting the Suitability of IT Students' Skills for the Recruitment in Saudi Labor Market," *2nd International Conference on Computer Applications and Information Security, ICCAIS 2019*, pp. 1–5, 2019, doi: 10.1109/CAIS.2019.8769577.

[30] "Linked in." https://www.linkedin.com/ (accessed Jul. 03, 2021).

[31] "Bayt." https://www.bayt.com/ (accessed May 15, 2021).

[32] M. C. M. Oo and T. Thein, "An efficient predictive analytics system for high dimensional big data," *Journal of King Saud University - Computer and Information Sciences*, no. xxxx, 2019, doi: 10.1016/j.jksuci.2019.09.001.

[33] M. He, D. Shen, Y. Zhu, R. He, T. Wang, and Z. Zhang, "Career Trajectory Prediction based on CNN," *Proceedings - IEEE International Conference on Service Operations and Logistics, and Informatics 2019, SOLI 2019*, pp. 22–26, 2019, doi: 10.1109/SOLI48380.2019.8955009.

[34] Y. Choi and J. W. Choi, "A study of job involvement prediction using machine learning technique," *International Journal of Organizational Analysis*, vol. 29, no. 3, pp. 788–800, 2020, doi: 10.1108/IJOA-05-2020-2222.

[35] R. S. Kumar, N. Prakash, and S. Anbuchelian, "Prediction of Job Openings in IT Sector using Long Short -Term Memory Model," *Proceedings of the 4th International Conference on IoT in Social, Mobile, Analytics and Cloud, ISMAC 2020*, pp. 945–953, 2020, doi: 10.1109/I-SMAC49090.2020.9243483.

[36] T. van Huynh, K. van Nguyen, N. L. T. Nguyen, and A. G. T. Nguyen, "Job prediction: From deep neural network models to applications," 2020.

[37] V. A. Ngo, T. T. N. Doan, T. T. Le, T. H. Tran, and B. L. Do, "Exploration and Integration of Job Portals in Vietnam," *Proceedings - 2020 RIVF International Conference on Computing and Communication Technologies, RIVF 2020*, 2020, doi: 10.1109/RIVF48685.2020.9140732.

[38] P. K. Roy, S. S. Chowdhary, and R. Bhatia, "A Machine Learning approach for automation of Resume Recommendation system," *Procedia Computer Science*, vol. 167, no. 2019, pp. 2318–2327, 2020, doi: 10.1016/j.procs.2020.03.284.

[39] G. O. Alghamdi and A. M. Alghamdi, "Towards Building Academic Entrepreneurial Programs at Saudi Universities: Predicting Future Jobs in Light of the NEOM Project," *World Journal of Education*, vol. 10, no. 4, p. 60, 2020, doi: 10.5430/wje.v10n4p60.

[40] T. Ha, M. Lee, B. Yun, and B. Y. Coh, "Job Forecasting Based on the Patent Information: A Word Embedding-Based Approach," *IEEE Access*, vol. 10, pp. 7223–7233, 2022, doi: 10.1109/ACCESS.2022.3141910.

[41] "O*NET."

[42] "cooperativepatentclassification."

[43] G. open data for S. Government, "Government job Dataset." data.gov.sa/Data/en/dataset (accessed Nov. 01, 2021).

[44] Ministry of Labor and Social Development, "Open Data." https://www.data.gov.sa/Data/en/dataset?_organization_limit=0&organization=ministry_of_labor_and_social_development_-_ministry_of_planning_and_development-information_center (accessed May 03, 2021).

[45] H. H. Do, A. Broder, J. Dean, M. R. Henzinger, K. Deshmukh, and S. Sarawagi, "Data Cleaning: Problems and Current Approaches," *IEEE Data Engineering Bulletin*, vol. 23, no. 4, 2000.

[46] M. Y. Mahan, C. R. Chorn, and A. P. Georgopoulos, "White Noise Test : detecting autocorrelation and

ISSN (online): 2581-3048
Volume 6, Issue 4, pp 1-8, April-2022
https://doi.org/10.47001/IRJIET/2022.604001

nonstationarities in long time series after ARIMA modeling," no. June 2015, 2020, doi: 10.25080/Majora-7b98e3ed-00f.

[47] "important evaultaion."

[48] C. J. Willmott and K. Matsuura, "Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance," *Climate Research*, vol. 30, no. 1, pp. 79–82, 2005, doi: 10.3354/cr030079.

[49] P. A. Thompson, "An MSE statistic for comparing forecast accuracy across series," *International Journal of Forecasting*, vol. 6, no. 2, pp. 219–227, 1990, doi: 10.1016/0169-2070(90)90007-X.

[50] A. N. Koval and A. B. Severny, "On the asymmetry of the symmetric MAPE," *international Journal of Forecasting*, vol. 4, no. 2, pp. 276–284, 1990, doi: 10.1007/BF00155226.

AUTHORS BIOGRAPHY

Mashael Alyahya, is a master student at Information Technology college at Qassim University. She is currently pursuing the M.Sc. degree in information technology. Her research interests include brain–computer interface, big data analytics, artificial intelligence, natural language processing, and the Internet of Things.

Mohammed Hadwan, received the B.Sc. degree in computer science from National University, Taiz, Yemen, in 2003,the M.Sc. degree in computer science from the School of Computer Sciences, University Science Malaysia (USM), Penang, Malaysia, in 2006, and the Ph.D. degree in computer science (artificial intelligence) from the Faculty of Information Science and Technology, National University of Malaysia (UKM), in 2012. From March 2012 to December 2015, he was the Dean of the Engineering and Information Technology College, Al Saeed University, Taiz. From May 2014 to March 2016, he was the Head of the Department computer Science, Faculty of Applied Sciences, Taiz University. He is currently a Staff Member with the Information Technology Department, College of Computer, Qassim University. He is also the Head of the Accreditation Committee and the Program Manager of the B.Sc. degree with the Information Technology Department. His research interests include AI algorithms and applications in general, optimization, big data, the Internet of Things, speech recognition, E-business, and E-commerce. He is also a member of the Data Mining and Optimization (DMO) Group and the Intelligent and Analytics Group (IAG). He is also an active researcher and a reviewer of several international and local journals.

Citation of this Article:

Mashael Alyahya, Mohammed Hadwan, "Applying ARIMA Model to Predict Future Jobs in the Saudi Labor Market" Published in *International Research Journal of Innovations in Engineering and Technology - IRJIET*, Volume 6, Issue 4, pp 1-8, April 2022. Article DOI https://doi.org/10.47001/IRJIET/2022.604001

****

