**DS7001**

**DS7001 Module -Course Work Plan (January 2022)**

**--------------------------------------------------------------------**

Data Science in the Age of Covid -

Review of Covid-19 impact (unemployment) on the Job Market using Machine Learning and Data Science course work & Portfolio



Contents

[ABSTRACT: 2](#_Toc92062599)

[INTRODUCTION: 2](#_Toc92062600)

[1.METHODOLOGY: 3](#_Toc92062601)

[2.MAIN ELEMENTS: 3](#_Toc92062602)

[2.1 Data Ecology in Job Market: 4](#_Toc92062603)

[*2.1.1 Data Science in Predicting the Employment Rate:* 4](#_Toc92062604)

[2.1.2 Inference: 5](#_Toc92062605)

[2.2 Data Ecology in Information Technology (IT) industry: 5](#_Toc92062606)

[2.2.1 IT Recruitment Process and AI: 5](#_Toc92062607)

[2.2.2 Inference 7](#_Toc92062608)

[2.3 Pros and Cons of the AI/Machine Learning in the Job Market: 7](#_Toc92062609)

[2.3.1 The Pros: 7](#_Toc92062610)

[2.3.2 The Cons: 8](#_Toc92062611)

[3. DISCUSSION: 10](#_Toc92062612)

[4. CONCLUSION: 10](#_Toc92062613)

[REFERENCES: 11](#_Toc92062614)

[Worksheet 1 – Online Resources 12](#_Toc92062615)

[Worksheet 2 – UK tech clusters mini-project 13](#_Toc92062616)

[Worksheet 3 – data Integration: 14](#_Toc92062617)

[Worksheet 4 – Google Trends 14](#_Toc92062618)

[Worksheet 5 – Twitter Data 19](#_Toc92062619)

|  |  |
| --- | --- |
| **Module Leader** | **Dr. Yang Li** |
| **Student Id & Name** | **2195687 – Mirnalini Gunaraj** |

# ABSTRACT:

The COVID-19 is a communicable disease which was catching all over the world. One immediate remedy that comes into rescue is the Lockdown measures. So It was implemented to stop the spread of the COVID-19 virus have had a substantial impact on many individuals in recent years. The employment market is an example of this. Since the COVID-19 crisis struck the world, the job search, job advertising on online sites, and job loss have all altered. Because of this, a large number of individuals have lost their jobs and their mental health has been adversely affected. So, this research will give us a comprehensive overview of the unemployment due to COVID and its effects to the human beings. This paper also examines the application of machine learning techniques like regression, descriptive statistics and topic modelling in the Job market ecology during a Pandemic.

# INTRODUCTION:

On January 30, 2020, WHO classified the explosion of coronavirus (COVID-19) disease as a Public Health extremity of international crisis. As of February 14, 2020, there have been 49,053 confirmed cases and 1,381 fatalities worldwide. So COVID has been considered as a pandemic disease. Greater than 31 gazillion persons have been infected with the corona virus as of September 2020, and greater than a gazillion people have lost their lives as a result of the infection.

Anxieties about disease have led several governments to put in place a wide variety of control measures. The adoption of so-called lockdown procedures has been the most extreme action to stop the spread of COVID-19. During a 'lockdown,' businesses are forced to close and people are prohibited from gathering or travelling. Nation-to-nation variations exist in the application of lockdown procedures. Restrictions were placed on sectors such as recreation (entertainment and restaurants), while enterprises such as food and agriculture and energy were not.

COVID-19 has had a significant influence on the labour market. Unemployment has soared and job listings have decreased as a result of the economic downturn. Workers' reactions to this crisis and their hunt for new occupations are an important topic. A change of pattern in job search as the result of COVID-19 might either increase or diminish the reaction of the COVID-19 on Job demand.

The economic slump has resulted in an increase in unemployment and a drop in job advertisements. We need to know how people are reacting to the crisis and whether or not they're looking for new careers. Following the COVID-19 shock, a shift in the intensity of job searches might either raise or decrease the impression of the COVID-19 agitate on labour supply. Unemployment rates throughout the major labour market sectors, as well as the consequences on individuals' mental and physical health (stress and cardiovascular disease) are examined in this research review. Also, the various machine learning techniques like Linear discriminant analysis (LDA), topic modelling, descriptive statistics are also reviewed. In addition, the various changes that has been made in the organisation to cope up with the new normal using AI has also been reviewed as it comes under the topic of Job market.

The structure of the paper is as follows: The Abstract provides a summary of the main points, and the Introduction gives the explanation of choosing the application Job market for reviewing. Methodology depicts the type of methodology which is used to the review of the data ecology elements. Main elements show the role of machine learning in finding the Unemployment rate during the COVID-19 times and various innovations implemented in the organisations using the machine learning and the AI and their various pros and cons. Discussion states the gaps in the study papers and Conclusion draws the results and future works.

# 1.METHODOLOGY:

The Methodology which is used to produce this literature review is the narrative methodology. In a narrative review methodology, it is possible to make conclusions about a topic from a variety of primary research summarised in a narrative way that incorporates the reviewers' prior knowledge, beliefs, and models. Also, it evaluates and synthesise the findings from several investigations to create new and useful information. As a result, the COVID-19 influence of data science and machine learning on the job market in different countries has been compared using this narrative method, and the findings and gaps that have been drawn from this analysis are also offered.

I have referred various Journals which have discussed the consequences of COVID-19 on the labor market of several countries like the UK, Germany, Sweden, Bangladesh & Sri Lanka as well as the use of machine learning models to predict the unemployment rate based on the Job Postings and Job search facilities of various Job sites. Also, the Comprehensive approach is used to review the impact of data science in the recruitment process.

The narrative approach is used to compare the various machine learning strategies that is applied to the Job posting data and the employment prediction rate is compared between them. We have selected the narrative methodology because it will be easy to present the results from different papers.

# 2.MAIN ELEMENTS:

The main elements of this literature review are based on how the machine learning models are used to predict the employment rate based on the Job postings by the companies in the Job sites. Also, how the Artificial Intelligence (AI) has been impacted the recruitment process for both the candidates and the Human Resource (HR) of the Organization. This Paper also deals with the Pros and Cons of the application area (Job market) and the significance of the machine learning and the data science in the Job market sector. I have reviewed the type of data, types of machine learning algorithm used and the outcome of the model to the application area or the society.

## 2.1 Data Ecology in Job Market:

The below shows the aspects of data ecology elements with respect to the entire Job market of various countries that are reviewed as part of this literature paper.

### *2.1.1 Data Science in Predicting the Employment Rate:*

The authors of the reference papers have used various machine learning models and strategies to predict the Employment rate. i.e., whether the number of Job postings are either reduced or increased sector by collecting the data from the various Job sites available in their own country. Let’s compare the data ecology elements among the countries.

#### Collection of Input Data:

[Arthur (2021)] uses the data from Reed.co.uk which is an online Job site, and the data is in the form of API (Application Program Interface) Whereas [Hayashi (2020)] in Sri Lanka and Bangladesh, the input data is from Bdjobs.com and topjobs.lk and the data is collected in form of JSON format. In Sweden, [Lena (2020)] uses data from Platsbanken.se where the dataset can be downloaded in the format of csv. In Germany [Julian (2021)], the data collection process is performed using the advanced web mining concept called robotic process automation (RPA) from the Interamt.de website.

#### Data Models:

In UK dataset, topic modelling algorithms such as **Linear Discriminant analysis (LDA) and** Doc2Vec is used to detect the topics of each advert. The usage of LDA is to show the input features in higher dimensional space so that curse of dimensionality will be rejected which results in low dimension costs. Then, decision tree classification algorithm has been applied to group the adverts into economic sectors in the Sri Lankan and Bangladesh dataset as the input data was not linear. In Sweden dataset, cluster is used to infer the user level data and Multiple Regression model has been applied on it as there were many continuous variable features. A quantitative text analysis approach has been applied on the dataset collected from the German website as they found that the data is normally distributed. So, we could see that different methodologies are applied based on the form of the dataset which is available.

#### Prediction:

In UK, the employment rate has been decreased during COVID-19. It is evident from the author Rudy Arthur (2021)’s paper that the percentage of job adverts related to full time and part time Jobs has been reduced after lockdown. Whereas in Sri Lanka and Bangladesh, Job adverts counts in the online job search websites has been reduced. The count was decreased by 88% in Bangladesh in April month of 2020 compared to the April month of 2019. At the same time the Sri Lanka’s count was and reduced around 71%. Also, in the German literature results indicate that there is a drastic increase in the Work from Home for the IT skills. So, the Job opening for IT with Work from home has been increased all over.

#### d. Model accuracies:

In the simple decision tree, model has achieved the accuracy of 0.922 and also balanced and Cohen kappa accuracies are calculated. This is quite a good score for a machine learning model.

### 2.1.2 Inference:

The machine learning model used in all the key literatures have relied only on one of the top hiring job sites of their country. It is clear that the input data is not from the difference sources. It is always good to train the models with different data instead of one. Also, the model should consider the external factors as well when trained. For Example - Brexit may be the reason for the Job deficit in some parts of United Kingdom. So, while developing the models, the external factors which is applicable to that area should also be considered.

## 2.2 Data Ecology in Information Technology (IT) industry:

The below shows the aspects of data ecology elements with respect to the IT field which are reviewed as part of this literature paper.

## 2.2.1 IT Recruitment Process and AI:

The application of Artificial Intelligence (AI) in the IT recruitment process has drastically increased nowadays (specifically after Covid). The usage is beneficial for both the candidates and the Organisation (Human Resource especially). Now, let us see the brief role of the data science in the various segments of the recruitment process.

#### AI Chatbots:

During the application Process, the AI chatbots provide immediate and quick response to the candidate questionnaires. This decreases the manpower usage during the COVID times. The chatbots are trained with the previous conversations stored in the Organisation’s database. The Input text is lemmatized, stemmed, stop words have been removed. Basically, the data cleaning is happening and then the words are converted into tokens and clustered together with the same sentiments and this data is used to train the model. So, in the future if the chatbot is facing the similar sentiment, it will be easy for model to provide the response in a good way.

#### AI Telephonic Interviews:

During the telephonic interview process, earlier manpower will be coming to the office for conducting the interviews. But, with the help of AI, the telephonic interviews are scheduled and conducted by Robots during the Covid times. Also, the scheduling process is carried out as well. The methodology behind this is as same as chatbots. The model have been trained with the earlier data and for each designation, the AI will respond appropriately. Also, the recording is done automatically while conducting the interviews. These recordings are in turn used in the future enhancements if needed.

#### CV Screening Process:

Due to the manual reading of the Curriculum Vitae (CV), there will be loads of CV’s piled up in the stack. But now, the AI is handling the CVs more efficiently. After the Pandemic breakdown, the Human resource has been piled up with large number of CVs. So, most of the top organisations has used the AI handled CV screening process. The AI uses advanced decision-making algorithm to make decisions based on the keywords available in the employee’s CV. One of the main points that needs to be considered is that there should not be any gender oriented or race-oriented differentiation while classifying CVs. The gender oriented or race-oriented differentiation is one of the biggest problems happening in recent days because of manpower doing the CV screening process. In some of the literature Papers, the models are even trained with considering the factor that the gap caused by COVID unemployment need not be considered as an factor for rejection.

#### AI Job seeking Tools:

There are many Jobs seeking artificial Intelligence tools which is helping the candidates to match their Job skill set with the Job description posted in an Organisation. This is helping the candidates to focus only on the relevant job postings which is a time saving process for both the candidates and the Organizations during the COVID times. These tools are efficient and can be used in all the sectors of the Job market. The way of creating the model involves the same algorithm which is reviewed in the AI chatbots. It has been involving the sentiment analysis tool for clustering the Skillset and the job descriptions.

#### Promotion Recommendation:

During the COVID times, the employees were given Work from home and so analysing the employee by the manager for the promotion is a no way thing. So, after the COVID, the organisations are planning to build a promotion recommender engine which is trained using the previous years promotion employee’s details and then the machine is once good with these patterns, the new employee data is provided to make it predict and say whether that employee is eligible for promotion or not. This model makes use of advanced deep learning techniques such as the Neural networks to classify the employee. In the paper [ Alexis (2021)] has used neural network because it is the great one to predict the correlations between the past events and the provided topics.

#### Employee Training Recommendation:

Also, the training to the employee can also be provided using the AI’s employee training recommendation engine. The input keywords for the engine will be the results of the assessments, their performance in the coding puzzles and algorithmic problem-solving questions. Also, the amount of time taken to resolve the problem will be fed into the model. The model is using the deep learning techniques to find the correlation and covariance factors to predict the eligibility of the employee to recommend the courses which can improve him in a great way. These methodologies can also be used in any other sectors apart from IT.

## 2.2.2 Inference

So, in the key literatures reviewed above, the models have been trained with complex algorithms to reduce the manpower during the COVID times. Also, they can be used to reduce the budget allocation given to HR to perform these activities. In one side, it is supporting to the employee and at the same time AI is taking up the opportunity of humans which in turn is affecting the employment rate of the people. Correlation is a concept which has been used in almost all the papers as it is the crucial part of building a model.

## 2.3 Pros and Cons of the AI/Machine Learning in the Job Market:

It is evident that AI and machine learning has contributed to a greater extent to the Job market employment rate analysis and recruitment process. But at the same time, there is another side of the sword that AI has caused the decline in hiring the Human Resource as well. So, let us discuss about the Pros and Cons of the application of machine learning in the Job market in the below sections:

### 2.3.1 The Pros:

The various advantages of using the AI/Machine Learning in Job market during COVID are:

1. Easy Prediction of Employment rate
2. Time Efficiency for Employee
3. Time Efficiency for Organization
4. Cost efficiency for Organization

**Easy Prediction of Employment rate:**

Using the machine learning models, the employment rate can be easily calculated as per the region wise or country wise provided the input data of Job postings or Search is provided to the model. Also, the AI helps the government to access the Economic activity as well. On analysing the Employment rate percentage, it will be easy for the government to introduce new schemes to the people who are suffering. This model can be also used in the future to check the employment activity in case of any other Virus variant emerges.

**Time Efficiency for Employee:**

In accordance with the employee, the AI has helped them to access the chatbot for Job recommendations and enquiries, which saves a lot of time for the employee due to the immediate response provided by the AI chatbots to access their performance report so that they can improve their skills as well. Also, many in person training has been cancelled and the training has been provided through AI based on their own skillset.

**Time Efficiency for Organization:**

For Organization, Conducting the interviews through telephones using AI, Chatbot implementation using AI, CV screening process, Promotion Recommendation engine and the Employee recommendation engine are saving so much of time for the Organization. Because manually reading the CVs will take more time for the Human Resource (HR) department and also there are chances that people will be discriminating them.

**Cost efficiency for Organization:**

As organisations are reducing the human, they need not pay more salary to the employees. This makes the organization to be cost effective.

### 2.3.2 The Cons:

Even though applying AI in the Job market made life easier. It has its own limitations. The various disadvantages of using the AI/Machine Learning in Job market during COVID are:

1. Relying on single source
2. Not Easy to Adapt
3. Data Centric to AI centric
4. No Privacy
5. Manpower Reduction
6. Problem in using Historical Data
7. Complexity of the Model

**a. Relying on single source:**

All our key literature models have relied on single Job sites for developing the model. So, it is not good to draw conclusion based on that. Models should have used multiple data sources before concluding the reduction in employment rate.

**b. Not Easy to Adapt:**

In a shorter span of time i.e., during the Covid times, it is not easy to adapt an AI based technology in the Job market. Because it will take more time to adapt some technology in the job market. As we could see from the literature reviews that only simple models are being created.

**c. Data Centric to AI centric:**

As we know that the Organisations are data centric by using the data lakes and the warehouses. All of a sudden, moving to an AI centric is not that easy. It takes so much of time and effort to perform that migration.

**d. No Privacy:**

In order to build an AI model, it needs to be fed with large amount of data from the previous years. Also, in our literature papers, authors are downloading from the Job sites to build a model. In that case, there is no data privacy.

**e. Manpower Reduction:**

The usage of AI will reduce the manpower soon. There is a survey that 30-40 percent of current jobs will be lost because of the artificial intelligence replacing the humans in their regular activities. There is a evidence that companies like Microsoft has completed scanning more documents about COVID to perform research about it in a shorter span of time. So, in that case companies will prefer developing more AI models rather than hiring humans.

**f. Problem in using Historical Data:**

The models used in our key literatures have used historical data so there are chances that the models are biased to existing data. So there may be chances that there will be a problem in predicting new things. For Example, In 2018, Amazon had destroyed a machine learning model which was filtering the applications because of the traditional past hiring patterns, the model did not like women.

**g. Complexity of the Model**

Applying AI in the Job market is a complex process because the employees and workers has to be trained in a proper way.

# 3. DISCUSSION:

This study has systematically reviewed the literatures which displays the use of machine learning models and AI in the Job market. It has reviewed two different kinds of literature. The first segment focused on papers which developed models to predict the employment rate during the COVID situation and the second segment of papers focussed on developing the AI models which reduces the manpower during the COVID times. On reviewing the two segments, we have the following discussion which shows the various impact on the society. Here society means the Employee and the Organisation.

* It is evident from the literature papers that the employee rate has been reduced. So, there was a negative effect on the Economy as well.
* Also, Employee were facing more stress because of unemployment and in turn cardiovascular disease as well because of this stress.
* Through these models the UK government finds out that the Job openings have reduced, and Mayor’s Academies Programme scheme has been implemented to help the unemployed.
* Employee were also happy to use the AI bots to make their life easier.
* Organisations also developed many AI models by using the COVID times as an opportunity to replace manpower.

# 4. CONCLUSION:

Finally, this literature paper has reviewed the impact of the COVID-19 in the Job market. It has shown that the COVID-19 has impacted the Job market in both good and bad way. In the Economy perspective, our key literature papers shows that there is a loss in the economy as major sectors have decrease in Job posting in the online sites and from the digitalisation perspective, our literature papers show that major enhancement has been made to make the jobs easier in day-to-day activities. Also, it has paved the way for reducing the usage of human power which is a very good way of handling a communicable disease like COVID. At the same time, there is a gap in the existing literatures that they have handled the AI usage with respect to Human resource department only and there was not much research on the technical parts. Also, the employment rate calculations have been made using the input data source from single Online job site and not from the multiple Job sites. So, it is not good to rely on the outcome fully. Also, topic modelling algorithms might have been improved more by using supervised algorithms. In addition, many of the literature papers did not calculate the accuracy value. This literature paper summarises the key points of using the AI and machine learning in the Job market industry. On reviewing we could see that the AI has contributed a lot to the Job market to bring it a good position. Also, in the future, AI can rule the Job market by doing the repetitive activities and cognitive augmentation sector.

# REFERENCES:

1. Arthur R (2021) Studying the UK job market during the COVID-19 crisis with online job ads. PLoS ONE 16(5): e0251431. <https://doi.org/10.1371/journal.pone.0251431>
2. Hayashi, Ryotaro; Matsuda, Norihiko. 2020. COVID-19 Impact on Job Postings: Real-Time Assessment Using Bangladesh and Sri Lanka Online Job Portals. © Asian Development Bank. http://hdl.handle.net/11540/12016. License: CC BY 3.0 IGO.
3. Lena Hensvik, Thomas Le Barbanchon, Roland Rathelot,Job search during the COVID-19 crisis,Journal of Public Economics,Volume 194,2021,104349,ISSN 0047-2727,https://doi.org/10.1016/j.jpubeco.2020.104349
4. Julian Koch, Ralf Plattfaut, Ingo Kregel,Looking for Talent in Times of Crisis – The Impact of the Covid-19 Pandemic on Public Sector Job Openings,International Journal of Information Management Data Insights,Volume 1, Issue 2,2021,100014,ISSN 2667-0968.
5. Paweł Nowik,Electronic personhood for artificial intelligence in the workplace,

Computer Law & Security Review,Volume 42,2021,105584,ISSN 0267-3649,

<https://doi.org/10.1016/j.clsr.2021.105584>.

1. Alexis Megan Votto, Rohit Valecha, Peyman Najafirad, H. Raghav Rao,Artificial Intelligence in Tactical Human Resource Management: A Systematic Literature Review,International Journal of Information Management Data Insights,Volume 1, Issue 2,2021,100047,ISSN 2667-0968,
2. Vijay Pereira, Elias Hadjielias, Michael Christofi, Demetris Vrontis,A systematic literature review on the impact of artificial intelligence on workplace outcomes: A multi-process perspective,Human Resource Management Review,2021,100857,ISSN 1053-4822,https://doi.org/10.1016/j.hrmr.2021.100857.