

Abstract

The outbreak of the Novel Coronavirus or the COVID-19 in various parts of the world has affected the world as a whole and caused millions of deaths. This remains an ominous warning to public health and will be marked as one of the greatest pandemics in world history. This aims to provide a better understanding of how various Machine Learning models can be implemented in real-world situations. Apart from the analysis done on the world figures, this paper also analyzes the current trend or pattern of Covid-19 transmission in India. With the help of datasets from the Ministry of Health and Family Welfare of India, this study puts forward various trends and patterns experienced in different parts of the world. The data to be studied has been obtained for 154 days i.e. from January 22, 2020, till June 24, 2020. For future references, the data can be further analyzed, and more results can be obtained.

TABLE OF CONTENTS

Chapter.No	TITLE	Page No.
	ABSTRACT	5
	LIST OF FIGURES	7
1	INTRODUCTION	9
	1.1 MACHINE LEARNING	10
	1.2 MACHINE LEARNING STRATEGIES	10
	1.2.1. SUPERVISED LEARNING	11
	1.2.2. UNSUPERVISED LEARNING	11
2	LITERATURE SURVEY	13
	2.1. RELATED WORK	13
3	METHODOLOGY	18
	3.1. EXISTING SYSTEM	18
	3.2. PROPOSED SYSTEM	18
	3.3. SOFTWARE ENVIRONMENT	19
	3.4. SYSTEM REQUIRMENTS	29
	3.4.1. HARDWARE REQUIRMENTS	29
	3.4.2. SYSTEM REQUIRMENTS	29
	3.5. SYSTYEM ARCHITECTURE	30
	3.6. MODULES	30
	3.7. PROPOSED ALGORITHMMS	32
	3.8. SYSTEM DESIGN AND TESTING PLAN	35
4	RESULTS AND DISCUSSION	40

5	CONCLUSION AND FUTURE WORK	43
	5.1. CONCLUSION	43
	5.2. FUTURE WORK	43
	REFERENCES	45
	APPENDICES	49
	A. SOURCE CPDE	49
	B. SCREENSHOTS	53
	C. PUBLICATION	55

LIST OF FIGURES

Figure No.	FIGURE NAME	Page no.
1.1	MACHINE LEARNING CLASSIFICATION	12
1.2	MACHINE LEARNING TASK	12
3.5	SYSTEM ARCHITECTURE	30
3.8	DATA FLOW DIAGRAM	39

CHAPTER 1

INTRODUCTION

MACHINE learning (ML) has proved itself as a prominent field of study over the last decade by solving many very complex and sophisticated real-world problems. The application areas included almost all the real-world domains such as healthcare, autonomous vehicle (AV), business applications, natural language processing (NLP), intelligent robots, gaming, climate modeling, voice, and image processing. ML algorithms' learning is typically based on trial and error method quite opposite of conventional algorithms, which follows the programming instructions based on decision statements like if-else. One of the most significant areas of ML is forecasting, numerous standard ML algorithms have been used in this area to guide the future course of actions needed in many application areas including weather forecasting, disease forecasting, stock market forecasting as well as disease prognosis. Various regression and neural network models have wide applicability in predicting the conditions of patients in the future with a specific disease. There are lots of studies performed for the prediction of different diseases using machine learning techniques such as coronary artery disease, cardiovascular disease prediction, and breast cancer prediction. In particular, the study is focused on live forecasting of COVID-19 confirmed cases and study is also focused on the forecast of COVID-19 outbreak and early response. These prediction systems can be very helpful in decision making to handle the present scenario to guide early interventions to manage these diseases very effectively. The most challenging aspect of its spread is that a person can possess the virus for many days without showing symptoms. The causes of its spread and considering its danger, almost all the countries have declared either partial or strict lockdowns throughout the affected regions and cities. Medical researchers throughout the globe are currently involved to discover an appropriate vaccine and medications for the disease. Since there is no approved medication till now for killing the virus so the governments of all countries are focusing on the precautions which can stop the spread. Out of all precautions, "be informed" about all the aspects of COVID-19 is considered extremely important. To

contribute to this aspect of information, numerous researchers are studying the different dimensions of the pandemic and produce the results to help humanity.

1.1 MACHINE LEARNING

Machine learning could be a subfield of computer science (AI). The goal of machine learning typically is to know the structure information of knowledge of information and match that data into models which will be understood and used by folks. Although machine learning could be a field inside technology, it differs from ancient process approaches.

In ancient computing, algorithms are sets of expressly programmed directions employed by computers to calculate or downside solve. Machine learning algorithms instead give computers to coach on knowledge inputs and use applied math analysis so as to output values that fall inside a particular vary. thanks to this, machine learning facilitates computers in building models from sample knowledge so as to modify decision-making processes supported knowledge inputs.

1.2 Machine Learning Strategies

In machine learning, tasks square measure typically classified into broad classes. These classes square measure supported however learning is received or however feedback on the educational is given to the system developed. Two of the foremost wide adopted machine learning strategies square measure supervised learning that trains algorithms supported example input and output information that's tagged by humans, and unattended learning that provides the algorithmic program with no tagged information so as to permit it to search out structure at intervals its computer file.

1.2.1 Supervised Learning

In supervised learning, the pc is given example inputs that square measure labelled with their desired outputs. The aim of this technique is for the algorithmic program to be ready to “learn” by comparison its actual output with the “taught” outputs to search out errors, and modify the model consequently. Supervised learning thus uses patterns to predict label values on extra unlabelled information. For example, with supervised learning, an algorithm may be fed data with images of sharks labelled as fish and images of oceans labelled as water. By being trained on this data, the supervised learning algorithm should be able to later identify unlabelled shark images as fish and unlabelled ocean images as water.

A common use case of supervised learning is to use historical information to predict statistically probably future events. It's going to use historical stock exchange info to anticipate approaching fluctuations, or be used to filter spam

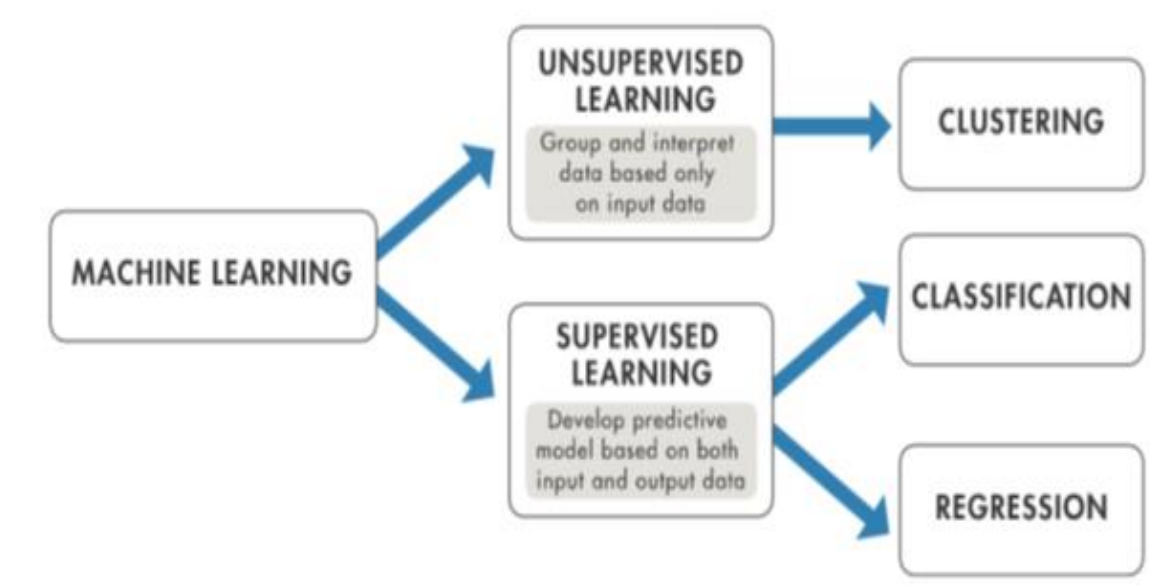
emails. In supervised learning, labeled photos of dogs are often used as input file to classify unlabeled photos of dogs.

1.2.2 UnSupervised Learning

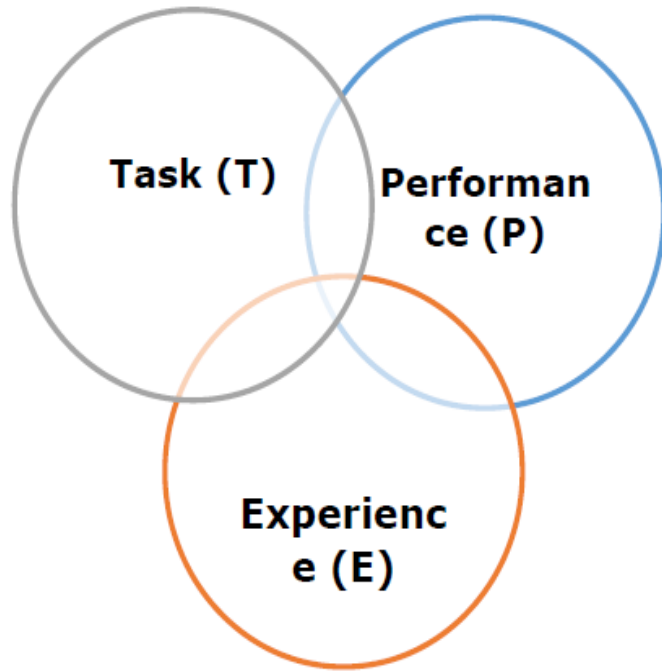
In unattended learning, information is unlabeled, that the learning rule is left to seek out commonalities among its input file. The goal of unattended learning is also as easy as discovering hidden patterns at intervals a dataset, however it should even have a goal of feature learning, that permits the procedure machine to mechanically discover the representations that square measure required to classify data.

Unsupervised learning is usually used for transactional information. You will have an oversized dataset of consumers and their purchases, however as a person's you'll probably not be able to add up of what similar attributes will be drawn from client profiles and their styles of purchases.

With this information fed into Associate in Nursing unattended learning rule, it should be determined that ladies of a definite age vary UN agency obtain unscented soaps square measure probably to be pregnant, and so a promoting campaign associated with physiological condition and baby will be merchandised.



1.1 MACHINE LEARNING



CLASSIFICATION

1.2 MACHINE LEARNING

CHAPTER 2

LITERATURE REVIEW

Literature survey is the most important step in software development process. Before developing the tool it is necessary to determine the time factor, economy and company strength. Once these things are satisfied, then the next step is to determine which operating system and language can be used for developing the tool. Once the programmers start building the tool the programmers need lot of external support. This support can be obtained from senior programmers, from book or from websites. Before building the system the above consideration are taken into account for developing the proposed system.

The major part of the project development sector considers and fully survey all the required needs for developing the project. For every project Literature survey is the most important sector in software development process. Before developing the tools and the associated designing it is necessary to determine and survey the time factor, resource requirement, man power, economy, and company strength. Once these things are satisfied and fully surveyed, then the next step is to determine about the software specifications in the respective system such as what type of operating system the project would require, and what are all the necessary software are needed to proceed with the next step such as developing the tools, and the associated operations.

2.1 RELATED WORK

COVID-19 Epidemic Analysis using Machine Learning and Deep Learning Algorithms

The catastrophic outbreak of Severe Acute Respiratory Syndrome - Coronavirus (SARS-CoV-2) also known as COVID-2019 has brought the worldwide threat to the living society. The whole world is putting incredible efforts to fight against the spread of this deadly disease in terms of infrastructure, finance, data sources, protective gears, life-risk treatments and several other resources. The artificial intelligence

researchers are focusing their expertise knowledge to develop mathematical models for analyzing this epidemic situation using nationwide shared data. To contribute towards the well-being of living society, this article proposes to utilize the machine learning and deep learning models with the aim for understanding its everyday exponential behaviour along with the prediction of future reachability of the COVID-2019 across the nations by utilizing the real-time information from the Johns Hopkins dashboard.

COVID-19 in India: State-wise Analysis and Prediction

The highly infectious coronavirus disease (COVID-19) was first detected in Wuhan, China in December 2019 and subsequently spread to 212 countries and territories around the world, infecting millions of people. In India, a large country of about 1.3 billion people, the disease was first detected on January 30, 2020, in a student returning from Wuhan. The total number of confirmed infections in India as of May 3, 2020, is more than 37,000 and is currently growing fast.

COVID-19 Outbreak Prediction with Machine Learning

Several outbreak prediction models for COVID-19 are being used by officials around the world to make informed-decisions and enforce relevant control measures. Among the standard models for COVID-19 global pandemic prediction, simple epidemiological and statistical models have received more attention by authorities, and they are popular in the media. Due to a high level of uncertainty and lack of essential data, standard models have shown low accuracy for long-term prediction. Although the literature includes several attempts to address this issue, the essential generalization and robustness abilities of existing models needs to be improved. This paper presents a comparative analysis of machine learning and soft computing models to predict the COVID-19 outbreak. Among a wide range of machine learning

CHAPTER 3

METHODOLOGY

3.1 EXISTING SYSTEM

- Information from news articles played an important role in empowering citizens during an epidemic.
- However, our analysis found more negative articles than positive articles which is a concern (2548 vs. 561).
- The word clouds corresponding to the negative and positive statements displayed similar words.
- A content analysis of text mined Ebola on news articles and scientific publications showed similar findings

3.2 PROPOSED SYSTEM

- The also presented a comprehensive study of the spread of the virus outbreak situation in India which will further help in taking necessary steps to manage the huge population of India.
- For the same, two Machine Learning models were used but in future Deep Learning models or hybrid two or models can be used to forecast the further spread of the virus.
- Real-time data has been used for this analysis and not the conventional data because the data validity of real time data highly depends on the time and its response time requirements come from the external world.
- Although the real time data can be relaxed in a few cases unlike the conventional data which needs to satisfy every case, real time data is approximately correct with performance metrics calculated as a number of transactions missing their deadlines per unit time.