iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 11 **Issue:** I **Month of publication:** January 2023

DOI: https://doi.org/10.22214/ijraset.2023.48899

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 11 Issue I Jan 2023- Available at www.ijraset.com

Introduction to Machine Learning and its Application

Vipul Chaudhari[1], Adyant Bajpai[2], Shruti Amrutkar[3], Sukriti Dubey[4], Dr. Poonam Gupta[5]
Department of Information Technology
[1, 2, 3, 4]UG Students, G H Raisoni College of Engineering and Management, Pune
[5]Associate Professor, G H Raisoni College of Engineering and Management, Pune

Abstract: ***This paper describes essential points of machine learning and its application. It seamlessly turns around and teach about the pros and cons of the ML. As well as it covers the real-life application where the machine learning is being used. Different types of machine learning and its algorithms. This paper is giving the detail knowledge about the different algorithms used in machine learning with their applications. There is brief explanation about the Weather Prediction application using the machine learning and also the comparison between various machine learning algorithms used by various researchers for weather prediction.***

Index Terms: Machine Learning, Application, real-life examples, Supervised learning, Unsupervised Learning, Classification, Regression, Feedback.

I. INTRODUCTION

Modern innovations like machine learning have greatly benefited both our daily lives and many commercial and professional processes. It is a subfield of artificial intelligence (AI) that concentrates on developing smart computer systems that can learn from databases by utilizing statistical methods. Machine learning, a cutting-edge field of research, enables computers to learn on their own using past data. In order to build mathematical models and generate predictions based on previously collected data or information, machine learning employs a range of methodologies. It is being used for many various purposes, such as speech recognition, recommender systems, email filtering, Facebook auto-tagging, and image identification. In the real world, there are individuals all around us who have the capacity to learn from their experiences, and we also have computers or other robots that carry out our orders. However, can a machine learn from previous knowledge or experiences in the same manner that a person does? The function of machine learning is now clear. Artificial intelligence has a subset called

Arthur Samuel first used the term "machine learning" in 1959. That is:

"With the aid of machine learning, a machine is able to predict outcomes without being explicitly programmed and automatically learn from data and improve performance over time"[1].

II. MACHINE LEARNING LIFE CYCLE

Machine learning has made it possible for computers to learn autonomously without explicit programming. To find a solution for the problem or project is the machine learning life cycle's main objective [2].

The machine learning life cycle consists of the following seven significant stages, which are mentioned below:

1) Data collection
2) Data preparation
3) Data organisation
4) Data analysis
5) Train the model
6) Testing
7) Deployment.

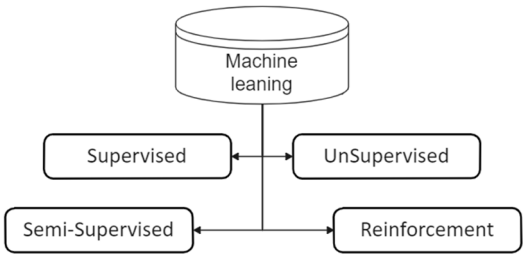
III. TYPES OF MACHINE LEARNING

Based on the methods and modalities of learning, machine learning is largely divided into four categories, which are:

1) Supervised Machine Learning
2) Unsupervised Machine Learning
3) Semi-Supervised Machine Learning
4) Reinforcement Learning

©IJRASET: All Rights are Reserved

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 11 Issue I Jan 2023- Available at www.ijraset.com



A. Supervised Machine Learning

As the name implies, supervised machine learning is based on supervision. This suggests that under the supervised learning strategy, the computers are trained using the "labelled" dataset, and after training, they predict the output. Which inputs have already been mapped to which outputs are shown here by the labelled data? To put it more clearly, we may say that we ask the machine to predict the outcome using test dataset after training it with input and related output [2].

1) Types of Supervised Machine Learning

There are two forms of supervised machine learning, and they are as follows:

a) *Classification:* Methods of classification are used to resolve problems where the output variable is categorical, such as "Yes" or "No," Male or Female, Red or Blue, etc. The classification algorithms forecast the categories that are present in the dataset. There are many classification algorithms in use today, including spam detection, email filtering, and others [3].

b) *Regression:* Regression techniques are used to resolve regression problems where the relationship between the input and output variables is linear. These are used to predict variables that have continuous outputs, such as market movements, weather forecasts, etc.

2) Advantages and Disadvantages of Supervised Learning.

a) Advantages

- We can precisely identify the classes of objects because supervised learning uses a labelled dataset.
- These algorithms are useful for forecasting the outcome based on past performance.

b) Disadvantages

- Complex problems cannot be solved by these algorithms.
- If the test data are different from the training data, it might anticipate the incorrect outcome.
- The algorithm training takes a lot of computational time.

B. Unsupervised Machine Learning

The unsupervised machine learning approach differs from the supervised machine learning method in that supervision is not required. When a system is trained on an unlabelled dataset and provides output predictions using unsupervised machine learning, no human supervision is involved. Unsupervised learning allows the models to operate independently on data that hasn't been categorised or labelled after they've been trained on it. The main objective of the unsupervised machine learning technique is to identify or classify the datasets that are not organised according to common characteristics, patterns, or differences [4].

1) Types of Unsupervised Machine Learning

Unsupervised machine learning has two components, which are as follows:

a) *Clustering:* We use the clustering approach to search for the innate groupings in the data. It is a technique for grouping the objects so that those that are most similar to one another stay together and have little to no resemblance to those in other groups. The clustering method is demonstrated by grouping customers according to their purchasing patterns [4].

b) *Association:* In a huge database, an unsupervised learning technique called association rule learning uncovers fascinating connections between variables. The main objective of this learning method is to find the connections between data points and then map the variables to maximise profit. Continuous production, web usage mining, market basket analysis, etc. are some applications of this technology.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

2) Advantages and Disadvantages of Unsupervised Learning Algorithm

a) Advantages

- These algorithms, as contrast to supervised ones, can be utilised for more challenging tasks because they operate on an unlabelled dataset.
- For a variety of jobs, unsupervised techniques are preferred since it is simpler to obtain the unlabelled dataset than the labelled dataset.

b) Disadvantages

- Due to the lack of labels in the dataset and the fact that the algorithms are not trained using the precise output in advance, the output of an unsupervised algorithm may be less accurate.
- As unsupervised learning uses a dataset that is not mapped to the output and is unlabelled, working with it is more challenging.

C. Semi-Supervised Learning

It falls between supervised learning (with labelled training data) and unsupervised learning (without labelled training data) techniques and uses a combination of labelled and unlabelled datasets during the training phase.

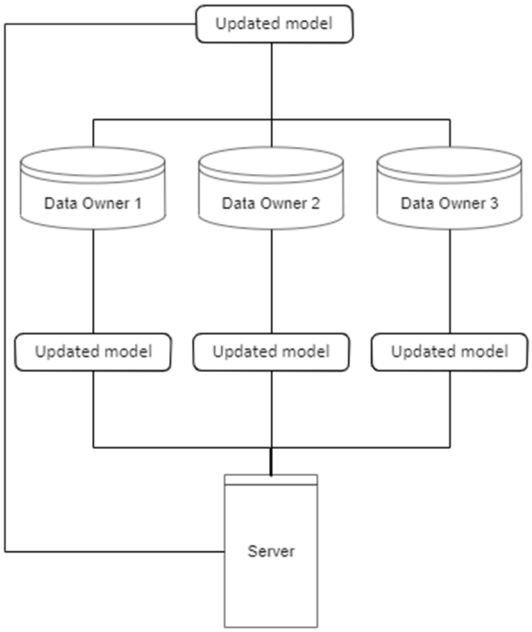
To overcome the drawbacks of supervised learning and unsupervised learning techniques, the concept of semi-supervised learning is put forth. The main objective of semi-supervised learning is to utilise all available data to the fullest extent possible rather than relying primarily on labelled data as in supervised learning.

D. Reinforcement Learning

Through striking and trailing, acting, learning from experiences, and improving performance, an AI agent can explore its environment on its own. The basis for reinforcement learning is feedback. Since a reinforcement learning agent is rewarded for every positive behaviour and punished for every negative one, its goal is to maximise rewards.

E. Federated Learning

Federated learning is a method for making data accessible for AI applications while allowing you to train AI models privately without letting anyone else see or access your data. Federated learning, a novel approach to AI education, is quickly taking over as the standard method for adhering to a slew of new regulations on the management and archiving of personal data. Federated learning also provides a mechanism to access the raw data coming from sensors on satellites, bridges, factories, and an increasing number of smart home and body devices by processing data at their source [3].



F. *Deep Learning*

Machine learning, which is simply a neural network with three or more layers, is a subset of deep learning. These neural networks aim to mimic how the human brain operates so that it can "learn" from vast volumes of data.

G. *Machine Learning*

Types of Machine Learning	Algorithm	Applications
Supervised Learning Classification	Decision trees, the Random Forest algorithm, and the Logistic Regression algorithm.	Fraud prevention, Text categorization, Spam filtering, and Image classification
Supervised Learning Regression	Simple Linear Regression Algorithm, Multivariate Regression Algorithm.	Predicting financial statements, Finance, Measure a stock's, Prediction and forecasting, Marketing, Manufacturing medicine
Unsupervised Learning Clustering	Principal component analysis, Mean-shift algorithm, DBSCAN technique, and K-Means clustering algorithm.	Scalability, Constraint-based clustering, High dimensionality, Capacity to handle noisy data
Unsupervised Learning Association	Apriori algorithm, Eclat algorithm, FP-growth algorithm	Traffic prediction, Self-driving cars and Malware Filtering, Medical Diagnosis Automatic Language Translation
Semi-Supervised learning	Consistency Regularization, Pseudo-labelling	Protein sequence classification, Speech analysis, Web content classification, and Text document classification
Reinforcement learning	Value-Based, Policy-based, Model-Based	Trading Stock Market Trading, Bid Optimization, Marketing Strategies.

IV. MACHINE LEARNING APPLIED TO WEATHER PREDICTION:

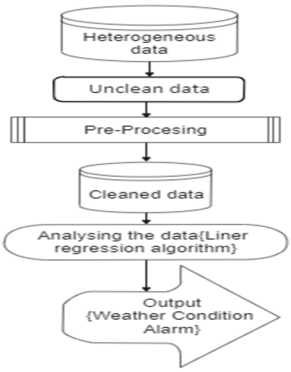
Weather forecasting is the process of predicting the climate and atmospheric conditions for a certain time. Weather conditions include rain, snow, temperature, fog, wind, and others. There are several ways for making weather predictions, including persistence forecast, climatological forecast, synoptic forecasting, statistical forecasting, computational forecast, and others [5].

A. *What is Weather Prediction?*

Weather forecasting techniques can be used to predict the weather for a specific location hours, days, or even months in advance utilising observational data, science, and technology. Weather forecasting was done through observation before computer models. The first weather forecast created by a computer was produced in 1955. In addition to being utilised for long-term forecasting, weather forecasting is used to know our everyday weather [5].

Weather forecasts are essential to us because they affect many facets of our way of life. For primary and secondary industries including farming, transportation, entertainment, and business that depend on favourable climatic conditions for their operations, as well as for making judgements on the management of severe weather, accurate weather forecasts are crucial. [6].

B. *Weather Prediction System Design*



C. Weather Prediction DATASET Available

Weather Prediction needs the training and testing of the data which is collected from various datasets and supporting web application of weather forecasting. In the Research the Datasets are from the Rp5.ru, Kaggel.com and AVIPS (Advanced Weather Information Processing System) [7], [8].

The basic parameters of the weather prediction are: Wind, Temperature, Air pressure and Moisture.

D. The Algorithm that Various Scientists use to predict the Weather

According to analysis, a variety of machine learning algorithms are employed to predict the weather.

Linear Regression: Linear regression is a straightforward method for supervised learning that may be used with both time series and structural data. The underlying presumption is that Y depends linearly on X1, X2, and Xp. The following equation represents this relationship: $Y = \beta0 + \beta1X + €$ [14].

```
x = np.array(data[features ])
y = np.array(data['Weather_
Temperature6'])
print(x.shape , y.shape)
#x_train , x_test , y_train ,
y_test = sklearn.model_selection
.train_test_split(x , y , test_size = 0.1 )
#print(x_train.shape ,y_train.shape)
#print(x_test.shape , y_test.shape)
predictions = linear.predict(x_test)
```

Random Forest classifier: A popular machine learning model with excellent performance in classification and regression applications is Random Forest. To get better outcomes, a random forest regression model can also be utilised for time series modelling and forecasting.

Authors	Algorithm Used	Parameters/ Datasets	Accuracy/ Conclusion
1) ZhanJie Wang, et al.[9]	Data mining prediction algorithm, Regression	Temperature, Rainfall, Wind speed and Historic data's	It shows good performance in term of MSE.
2) Shivam Tandon, et al.[7]	Linear Regression	Weather for 243 countries of the world, Rp5.ru	89% Accuracy
3) Munmun Biswas, et al.[8]	Naive Bayes algorithm, Chi Square algorithm	Transformed dataset.	The model's general correctness is its accuracy. (83.5%)
4) Ismaila Oshodi, et al.[6]	Gradient Boosting Classifier, Gaussian Naïve Bayes model.	Kaggle Dataset.	performing algorithm with an 84.153% prediction accuracy.
5) Dharsan G M (2022)[10]	Sector Vector Machine, Extreme learning Machine, Artificial Neural Network	Dataset used from Kaggle with basic parameters.	Its Performing accuracy is 78.3%
6) Garima Jain Bhawna Mallick (2016)[11]	ARIMA (Auto Regressive Integrated Moving Average)	Weather Property, Minimum Temperature, Maximum Temperature, Humidity and Air pressure	Accuracy is 83.24%

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

V. CONCLUSION

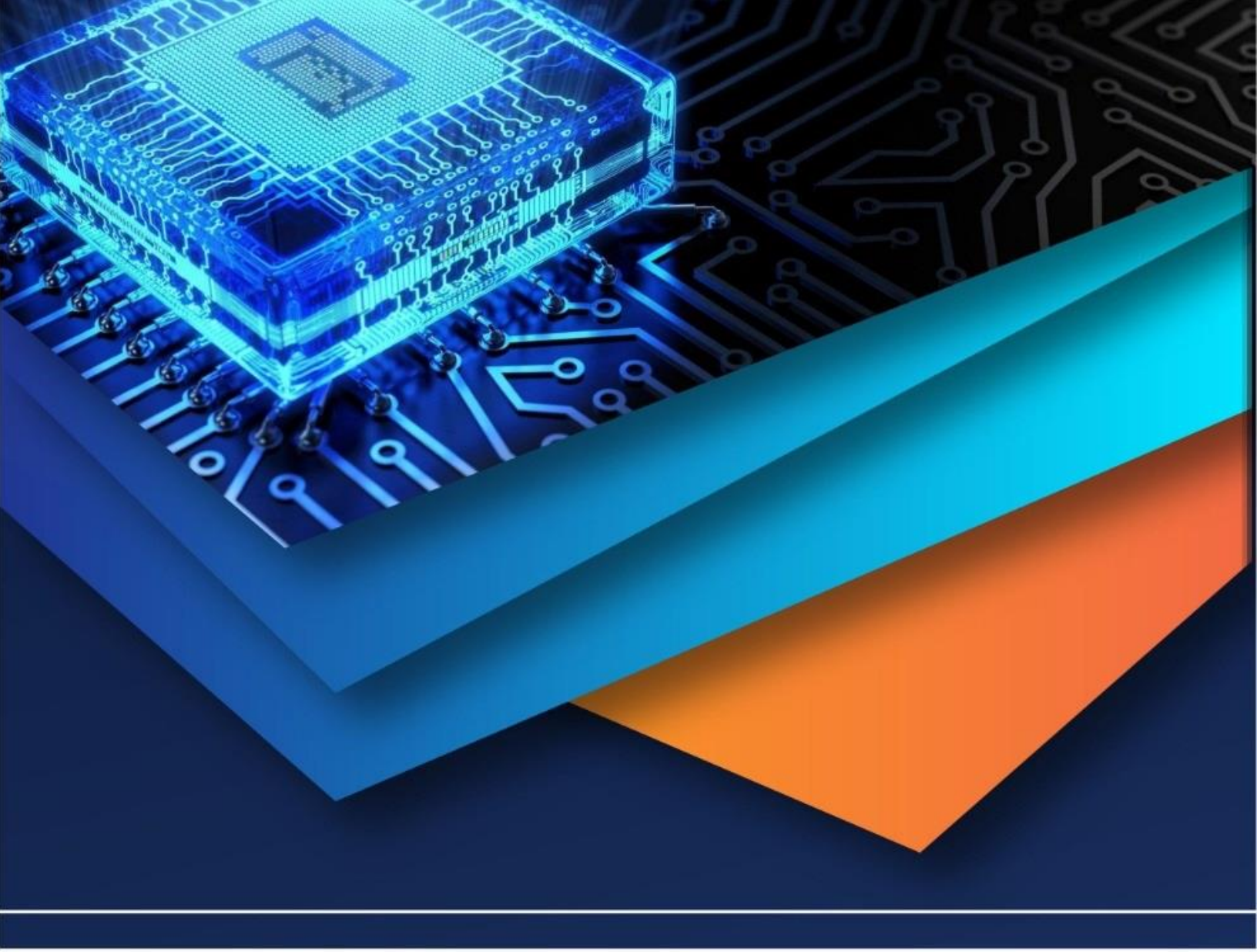
In Machine learning and its application, we learned about all the types of machine learning and the various algorithms used in it with its application. And algorithms used by various researchers for the weather prediction using machine learning are Data mining prediction algorithm, Linear Regression and etc. Got to know the accuracy of the Linear Regression algorithm which was more accurate than other algorithm used for the Weather prediction.

And we suggest to use Linear Regression algorithm for the weather forecast which gives more accuracy.

REFERENCES

[1] https://ieeexplore.ieee.org/xpl/RecentIssue.jsp?punumber=34
[2] Mark Holmstrom, Dylan Liu, Christopher Vo Stanford University "Machine Learning Applied to Weather Forecasting" (Dated: December 15, 2016)
[3] https://www.tutorialspoint.com/machine_learning/machine_learning_introduction.html
[4] https://www.javatpoint.com/machine-learning
[5] https://study.com/learn/lesson/weather-forecasting-types-tools.html
[6] Ismaila Oshodi – "Machine learning-based algorithms for weather forecasting", Preprints (www.preprints.org):2022
[7] Shivam Tandon, Pawan Kumar Singh, Abhishek Patel "Weather Prediction Using Machine Learning"- (2020)
[8] Munmun Biswas, Tanni Dhoom, Sayantanu Barua- "Weather Forecast Prediction: An Integrated Approach for Analyzing and Measuring Weather Data", December 2018 International Journal of Computer Applications 182(34):20-24
[9] ZhanJie Wang and A. B. M. Mazharul Mujib "The Weather Forecast Using Data Mining Research Based on Cloud Computing"- Journal of Physics: Conference Series 2017, J. Phys.: Conf. Ser. 910 012020
[10] Dharsan G M- "A SURVEY ON WEATHER FORECASTING AND THEIR TECHNIQUES", International Research Journal of Modernization in Engineering Technology and Science, Volume:04/Issue:02/ February -2022
[11] Garima Jain, Bhawna Mallick- "A Review on Weather Forecasting Techniques", International Journal of Advanced Research in Computer and Communication Engineering, Vol. 5, Issue 12, December 2016
[12] Anant Sharm, Ankit Yadav, Bhanu Kumar, Dr. Sunil Gupta- "Weather Forecast Application", IJIRT | Volume 6 Issue 11 | ISSN: 2349-6002
[13] Rubhi gupta-"Review on weather prediction using machine learning", Volume 8, Issue 1 | ISSN: 2321-9939 | IJEDR 2020
[14] Weather Forecasting Analysis using Linear and Logistic Regression Algorithm Tanvi Patil 1, Dr. Kamal Shah2

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)