Quest Journals
Journal of Architecture and Civil Engineering
Volume 7 ~ Issue 11 (2022) pp: 55-64
ISSN(Online) : 2321-8193
www.questjournals.org

Research Paper

Forecasting of Cyclone Precipitation Using Different Machine Learning Techniques

P. Jayaram Prakash [1] ,P. Pavani[2], P. Gayatri[3], P. Akhil[4]
Department of civil engineering, GMRIT.
Rajam,532127.

ABSTRACT: *In today's world, technology has been increasing with many software applications such as Machine learning ruling the world. By using this machine learning, we can predict the Extreme weather events such as Cyclone precipitation, storms, tropical cyclones, etc. The machine learning models contain different algorithms like support vector regression (SVR), decision tree (DT), logistic regression (LR) and the neural networks like convolutional neural network (CNN), recurrent neural networks (RNN) and artificial neural networks (ANN). By using these algorithms, we can forecast the future data by using the previous data. The factors affecting these extreme weather events are wind speed, temperature gradient, sea surface temperature (SST), precipitation, potential intensity, etc. Based on this factors and algorithmic models, we can forecast the tropical cyclone intensity with high accuracy. Comparing the algorithm models, support vector regression gives the correct prediction with high accuracy. This model generates the equation to forecast the extreme weather events. It is a simple technique to forecast extreme weather events when compared to manual calculation and prediction. It is the day improve the effectiveness of the forecast by using the machine learning. It is all about the representative algorithm which effectively work on the forecasting the extreme weather events. The event will be taken as the main function and factors used as a variable to predict weather event when it happens in the future. Instead of manual prediction, we use the machine learning algorithms to forecast the future data. Through summarizing and analyzing the challenges of extreme weather events, the machine learning model successful in recent years.*
KEYWORDS: *Tropical cyclone, cyclone precipitation, vertical wind shear, temperature gradient.*

R*eceived 01 Nov., 2022;* ***R****evised 09 Nov., 2022;* ***A****ccepted 12 Nov., 2022*
P*ublished with open access at* ***www.questjournals.org***

I. INTRODUCTION

Many meteorologists research and world meteorologists' society (WMS) says that it is difficult forecast the extreme weather events by manual predictors and it's time to use the Machine learning models and algorithms to forecast this type of extreme weather events like Cyclone precipitation, Tropical cyclone, Storm surges etc. Before going to apply the Machine learning (ML) models we have to know about these events, cyclone precipitation is a cyclone in the atmosphere that has large low pressure having circular wind motion. It is an extreme weather event which happens at irregular intervals of time. However, many problems are solved by predictive skills but it is difficult to find these events by using the statical and non-linear methods, this reduces the performance of the model. These statical models are unable to deal with the complex and non-linear relationship between the factors affecting these extreme weather events. As Increase in the SST there is probability of increase in the occurring of the TC, they are directly proportional each other. Also decrease in the vertical wind shear then the intensity of the tropical cyclone also increases. Generally, the most popular forecast models are inaccurate due to installation. In order to increase the accuracy of the prediction we introduce the Machine learning (ML), it is a type of Artificial intelligence for forecast the future data by using the previous data. Based on the applications of ML divided into three categories: *feature selection, clustering and regression***.** Feature selection algorithm can eliminate the irrelevant data and increase the effectiveness of the task and accuracy , while coming to the Clustering algorithm used in pattern recognition and data driving to convert the group of data into parts which makes the simplest to understand about that event and finally coming to the Regression algorithm create the relationship between the independent and dependent variables , In this one representative algorithm is Support Vector Machine (SVM) and another is Support vector Regression (SVR) , SVM have been introduced by *vapnik's in 1995* for solving the pattern recognition problems. In this method higher dimensional data as input space and create the hyperplane in the space. It creates a solved quadratic equation based on the factors we take. Polynomials, splines, radial basis and multi-layer perceptron are applied on this SVP, Now coming to the Support Vector Regression which can effectively deal with non-linear problems by defining the Kernel Functions. SVR is the main machine learning algorithm which gives the high accuracy about the results. In addition to these algorithms, Decision Tree (DT) algorithm gives the classification rules for factors affecting the extreme weather events with high accuracy. Some of the tasks can be solve by using the neural networks such as Convolutional neural networks (CNN), Artificial neural networks (ANN) and Recurrent neural networks (RNN), these networks used for the mapping the non-linear variables and relates the function and variables. By Taking the pictures of extreme weather events it can create a one map to relates the all factors which affects the extreme weather events. These neural networks excellent in feature learning, due to this mapping difficulty level of the problem is avoided. Deep learning algorithms more suitable for the complex applications, we have to use the different machine learning algorithms for different type of the problems which gives the effectiveness and accuracy in the problem. These CNN and RNN includes that capturing the pictures of extreme weather events through hyperspectral, anomalies and Synthetic Aperture Radar (SAR).

Many difficulties faced in forecasting the cyclone precipitation due to insufficient understanding of the mechanism and less interaction with these extreme weather events. Machine learning provides a good way to forecast the cyclone precipitation to improve the accuracy about the model. There are some tools to use the machine algorithm model in that tool to create a quadratic equation which gives the result about the short-term forecasting and long-term forecasting. Tools such as *Rapid miner, SVM light, Weka, LIBSVM and KNIME* etc. These tools take the high input data as factors affecting the cyclone precipitation and main function as Cyclone precipitation. Except this algorithm there are many algorithms used in the machine learning model to forecast the cyclone precipitation genesis, Multiple linear regression (MLR) and Linear regression (LR) which can used to forecast the tropical cyclone intensity and storm conditions. We have to use this algorithm models for forecasting the cyclone precipitation with high accuracy.

Long term forecasting using the problem data forecast for long time period of time while coming to the short-term forecasting uses the problem data and forecast future data for short period of time. Long-term forecasting uses the algorithms like MR, MLP and SVR etc. Long-term forecasting uses the algorithms like SVR, LR, ANN, CNN and DT etc. SVR is one main representative algorithm which gives the high accuracy comparing to the all-algorithms models regarding machine learning.

II. METHODOLOGY

1.MACHINE LEARNING

Machinelearning is defined as the use and development of advanced capable of learning and accommodate without specific instructions by analyzing and draw inference from the data patterns using algorithms and statical models, which focuses on the use of data and algorithms to imitate the way that humans learn, gradually improving its accuracy.Building mathematical models is the primary function of machine learning, which is a collection of computer algorithms, making conclusions from samples using statistical models. In the case of a model with predefined parameters, Computer programs that employ practice data or experiences to improve performance are said to be learning when they are run models' parameters. The model can forecast the situation of the world in the future or explain the information gleaned from data. Machine learning algorithms can also be divided into several categories based on the learning tasks, such as prediction, feature selection, and dimensionality reduction. Because

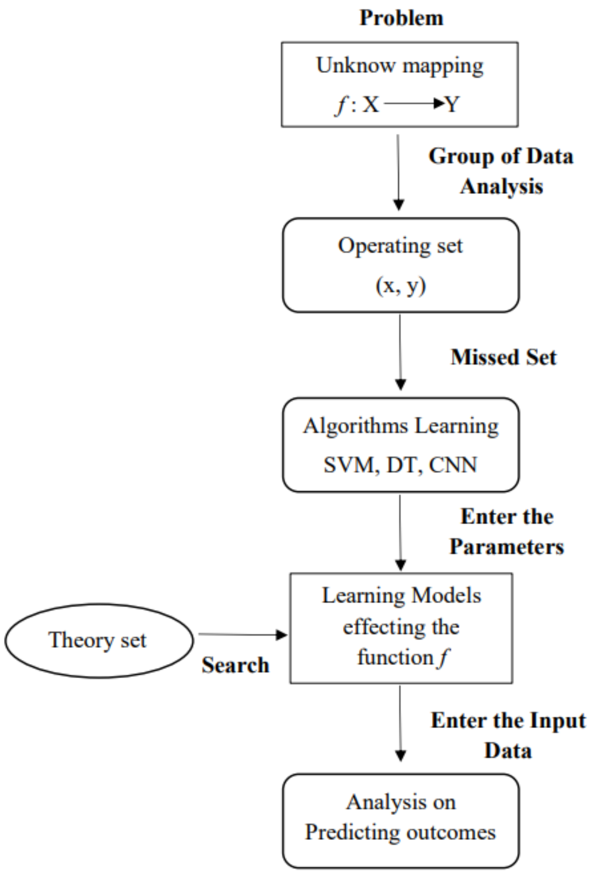


Diagram1. Workflow of Machine learning

this review focuses on Only predictive algorithms for TC forecasting will be discussed here. In general, if the model's goal is to predict discrete values, this type of learning task is known as "classification," while predicting continuous values is known as "regression." Furthermore, depending on whether the training data is labelled or not, learning tasks can be broadly classified into "supervised learning" and "unsupervised learning," with classification and regression representing the former and clustering representing the later. Algorithm works in this machine learning effectively with the variables. It takes the input space as X and output space as Y with a function fthen we represent relation. between X and Y as f :X to Y, which function depends on the Nonlinear parameters which belongs to our extreme weather event.

This flow chart explains the process of machine learning using different models and algorithms to forecast the extreme weather event. It takes the high input data of the function variables and establish the relation between the variables and gives a outcome quadratic equation which is useful for the short term and long-term forecasting ofthe extreme weather events.

2. **Representative Algorithm:**

Support Vector Regression:

Learn to build and analyze a Support Vector Regression (SVR) model in Machine Learning. The Machine Learning programmed is used to create Linear and Polynomial Regression models in a tool. In this article, we will walk through the programmed for creating a non-linear Support Vector Regression model.

Overview of SVR:

Regression with Support Vectors Understand how to build and analyze a Support Vector Regression (SVR) model in Machine Learning. The Machine Learning tool for creating Linear and Polynomial Multiple regression. In this paper, we will know the process of creating a non-linear Support Vector Regression model.

The SVR aims to fit the best line within a target value (Distance between hyperplane and boundary line), a, as opposed to other regression models that aim to minimize the error between the real and projected value. As a result, we may state that the SVR model tries to satisfy the criteria $-a<y-wx+b< a$. It forecast the value using the points along this boundary.

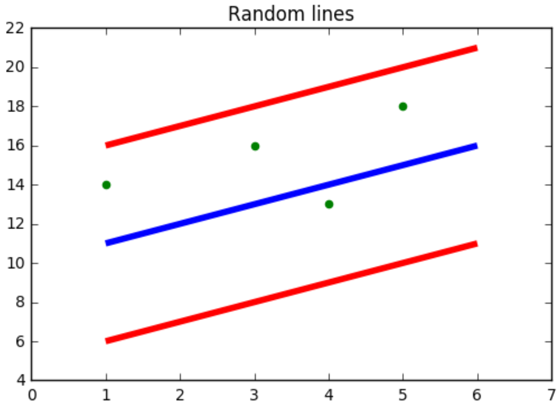
SOURCE: For a non-linear regression, the kernel function transforms the data to a higher dimension and performs the linear separation. Here we will use the rbfkernel.In this example, we will go through the implementation of support vector regression (SVR), in which we will predict extreme weather events by using the sources.

PROBLEM ANALYSIS:

In this data, we have one independent and one dependent variable, Marks. In this problem, we have to train a SVR model with this data to understand the correlation between the function and variables and be able to predict the student's mark based on their number hours dedicated studies.

WORKING OF SVR:

Now that we are clear on SVR terminology, let's examine how the algorithm functions. For instance, in a classification issue, we must distinguish between red and blue data points. The occurring events are the dots in the picture, and the random lines separate the range of a progressive event, which defines the value of the event. Based on these graphs, we create an algorithm for prediction and analysis of the problem. It gives the **98%** accuracy about model.



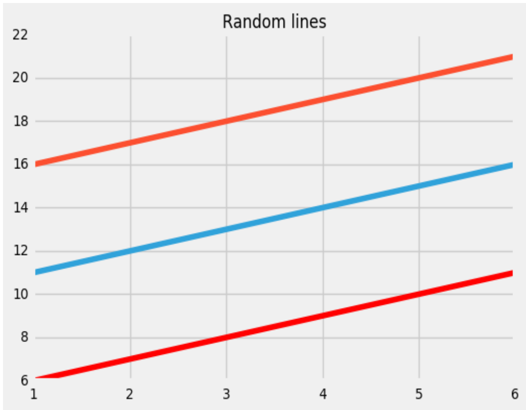


Diagram 2. separated lines

Diagram3.separated medium

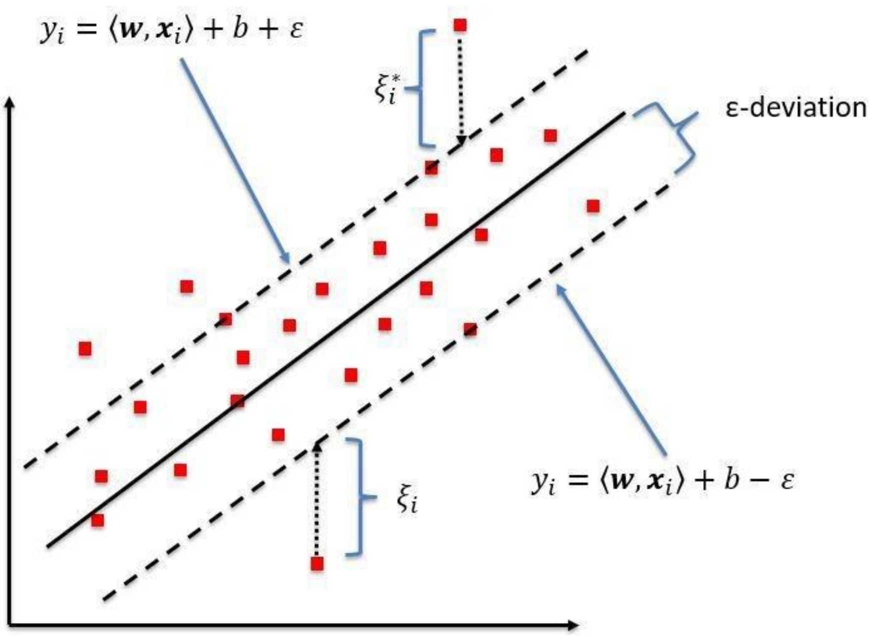


Diagram 4. Graphical representation of SVR

Decision tree:

It is a tool with uses in many different sectors. Regression and classification issues can both be solved using decision trees. The term itself implies that it uses a tree-like flowchart to display the forecasts resulting from a sequence of feature-based splits. Root branches serve as the starting point, while leaves make the final decision. A non-parametric deep learning technique for classification and regression is called a decision tree (DT). The objective is to learn simple decision rules derived from the data features in order to build a model that predicts the value of a target variable. An example of an equation's approximation is a tree.

Using a set of this if decision rules, the decision tree in the example below learns from data to approximate a linear model. The complexity of the decision rules and the model's fit increase with the depth of the tree.

Graphical representation:

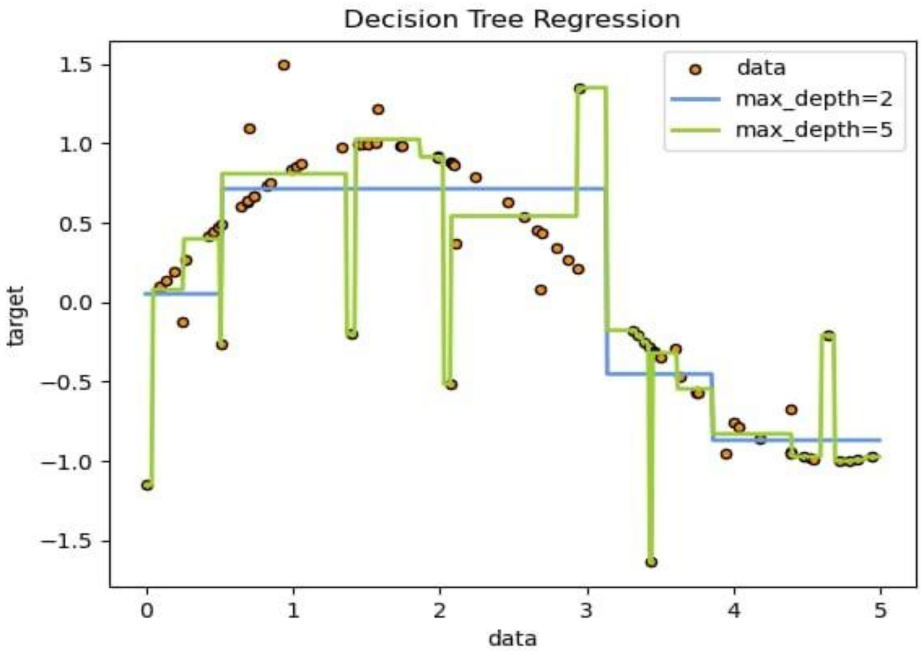


Diagram 5. Branches of tree process

https://scikit-learn.org/stable/auto_examples/tree/plot_tree_regression.html

Decision tree algorithm gives the accuracy **91%** comparing the all the models in the machine learning, which is easy to get through the branches. The selection to make important divisions major effect on a tree's accuracy. Regression and classification trees have different decision criteria. To decide whether to divide a node into two or more sub-nodes, decision trees employ a variety of techniques. The homogeneity of newly formed sub-nodes is increased by sub-node creation. In other words, we can claim that the branches purity improves in relation to the desired variable. The decision tree divides the nodes based on all variables that are accessible before choosing the split that produces the most homogeneous sub-branches.

Random forest:

Using both classification and regression, the supervised learning method random forest is applied. However, classification issues are its primary usage. A forest is made up of trees, and a forest with more trees is a forest with greater stability. In a similar manner, the random forest method builds decision trees on data samples, obtains predictions from each one, and then votes to determine which is the best answer. It is an ensemble method, which reduces over-fitting by averaging the outcome, making it superior than a single decision tree. It has two methods, they are *Bagging*, *Boosting*. The classification of decision trees that the random forest algorithm creates while growing the trees provides additional randomness to the forest. A better model is produced as a result of the algorithm's search for the best features from the random subset of features while splitting a node. As a result, only a random subset of the features is used when splitting a node. We can also utilize random thresholds for each feature to create additional random trees instead of looking for the best target level for each feature.

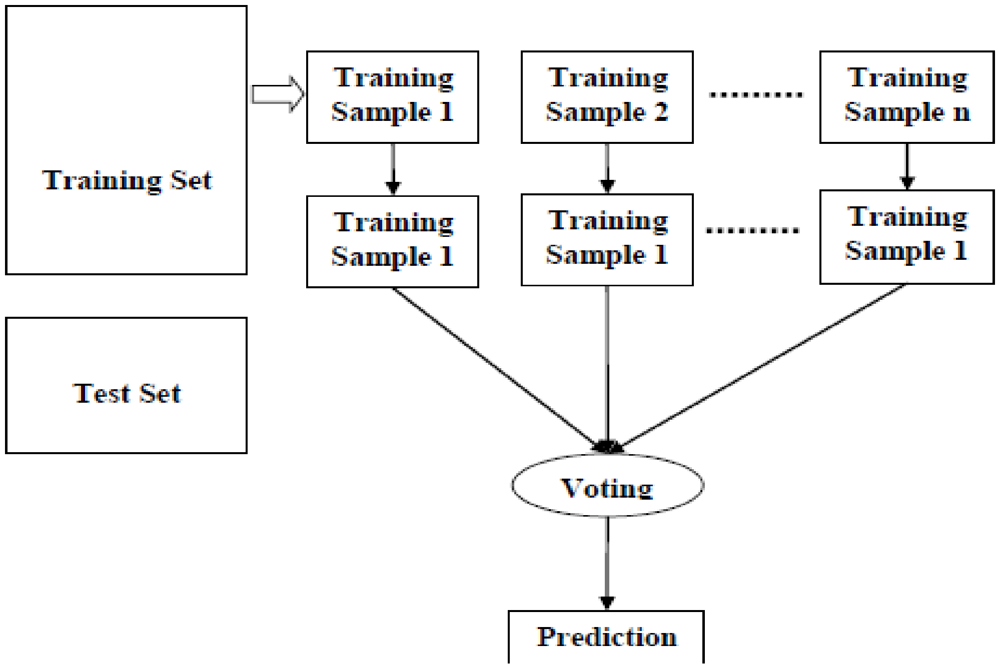
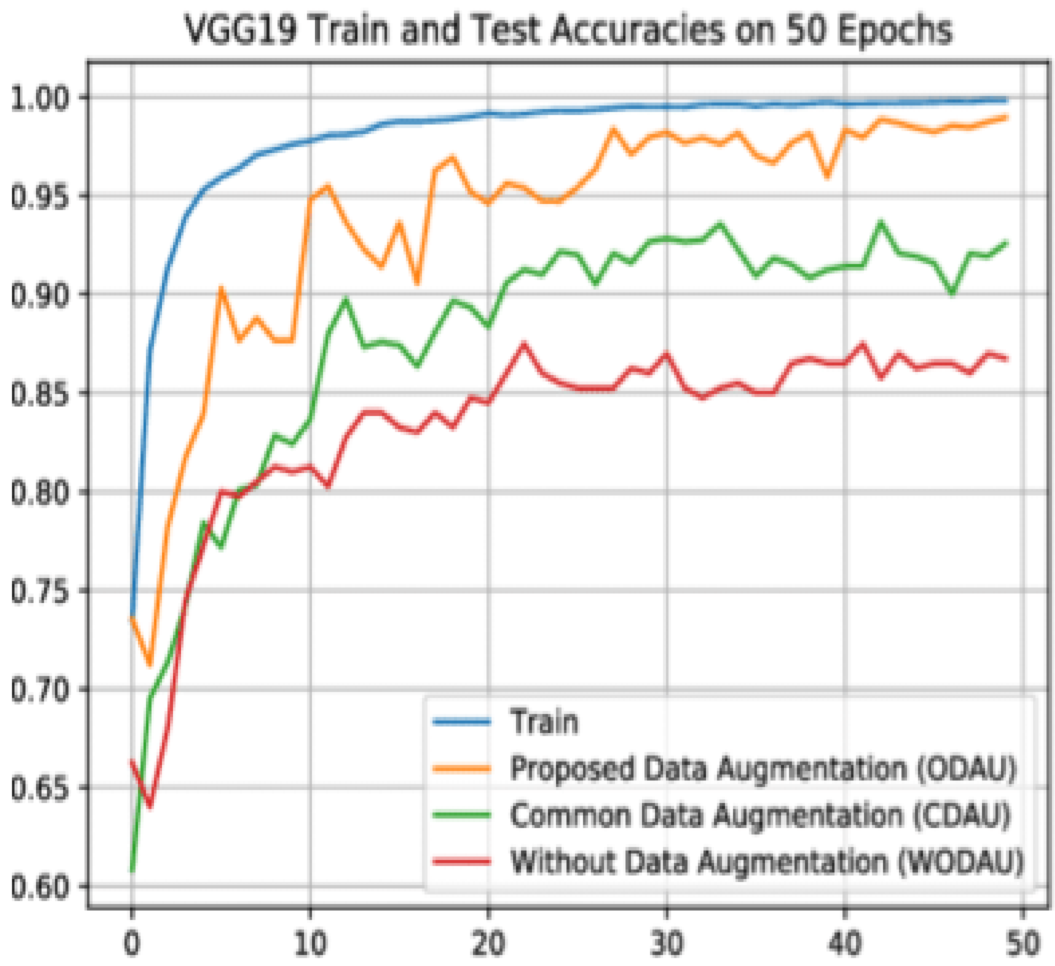


Diagram 6. Training set process of Random Forest

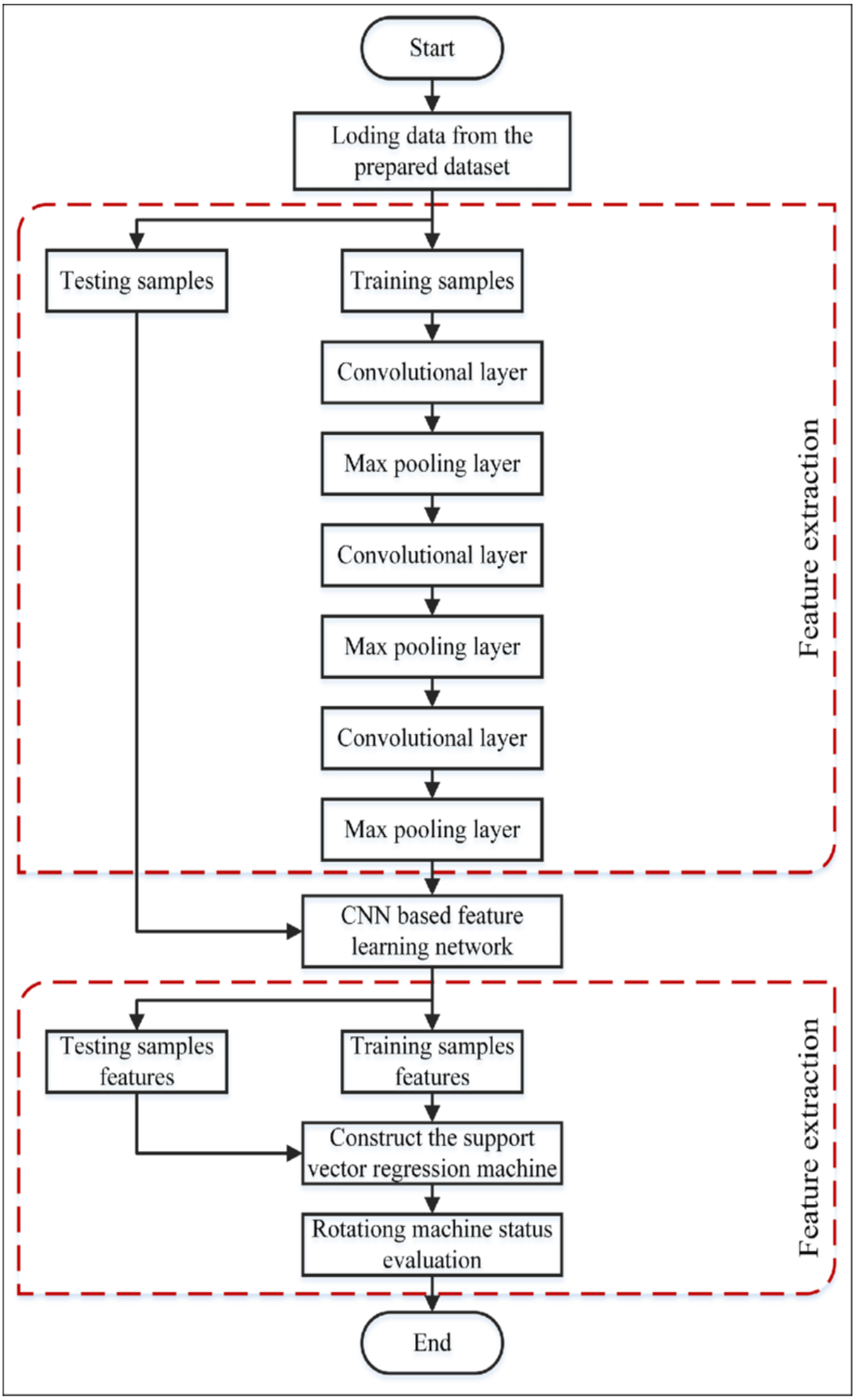
The random forest algorithm creates a forest in the shape of a group of decision trees, increasing randomness as the trees grow. The algorithm search for the greatest features from the random subset of features when splitting a node, adding more diversity and improving the model. Therefore, only a random subset of the features is considered when splitting a tree. We may also build more random trees by using random thresholds for each feature rather than attempting to find the best possible standard. The accuracy of these Random Forest is **87.87%.**

Convolutional neural networks:

Multiple layers of a CNN are possible, and each layer trains the CNN to detect the many aspects of an input image. Each image is given a filter or kernel to create an output that gets better and more detailed with each layer. The filters may begin as basic characteristics in the lower layers. In order to check and identify features that particularly indicate the input item, the difficulty of the filters increases with each additional layer. As a result, the partially recognized image from each layer's output, or convolved image, serves as the input for the layer. The CNN knows the image or object it represents in the final layer, which is a layer. The input image is filtered through a number of different filters during convolution. Each filter performs its function by turning on specific aspects of the image, after which it sends its output to the filter in the following layer. The operations are repeated for dozens, hundreds, or even thousands of layers as each layer learns to classify various features. Finally, the CNN is able to recognize the full object after processing all the picture data through its many layers.



https://www.researchgate.net/figure/Accuracy-and-loss-graphs-for-the-CNN-algorithms-Each-figure-consists-of-accuracy-



1.4.1. Flow map of CNN

https://www.researchgate.net/figure/Algorithm-flowchart-of-the-CNN-

These convolutional neural networks give the accuracy of **70%,** which form a network based on the factors affecting the function.

Tools used for Machine learning algorithms:

1.SVM Light

2.Rapid Miner
3.Weka
4.LIBSVM
5.KNIME

1.1 COMPARASION TABLE:

Algorithm model	Accuracy
SVM (Support Vector Regression)	98%
DT (Decision Tree)	91%
RF (Random Forest)	87.87%
CNN (Convolutional Neural Networks)	70%

III. RESULTS

Analyzing all of these studies has directed us to the assumption that SVR algorithm model are also the only machine learning model that consistently perform well. SVM has the highest level of forecasting accuracy when compared to all-Machine Learning models for extreme weather events. This SVM is part of the regression that links the dependent and independent variables in space and operates in a dimensionality-based space. SVR is mostly used to anticipate severe weather conditions like tropical cyclones and cyclone precipitation. With only the variables that have an impact on the cyclone precipitation as functions, this approach generates a quadratic equation. When comparing the DF, CNN, RF, and SVR, it shows that they are all 98% accurate. Model using the decision tree method yields 91%, Due to the lack of the dimensionality space needed to anticipate that extreme weather occurrence, the decision tree algorithm model only provides 91% of the process. When it comes to algorithm models, the SVR is far more effective than Random Forest and convolutional neural networks. In additionto forecasting extreme weather, air quality index forecasting is also done. Many forecasting methods claim that SVR is the best algorithm for applying machine learning to predict events with high accuracy. In place of manual forecasting methods, machine learning techniques are now crucial for predicting extreme weather events. Despite the fact that this SVR algorithm model functions with the majority of forecasting methods, it is integrated with other machine learning models, and it provides the high accuracy. Therefore, we chose the SVR algorithm model as the most crucial machine learning model for forecasting extreme weather occurrences.

This SVR algorithm model gives the accuracy of 98% among all the algorithm models in the machine learning. so, we prefer the SVR instead of other models in the machine learning to forecast the extreme weather events.

IV. CONCLUSION

Based on these Machine learning models and algorithms we conclude that SVR algorithm model gives the high accuracy other than all models. By forecasting the extreme weather events we mostly used this SVR model. Decision tree algorithm model also use to forecast the rare weather events with high accuracy with respective to SVR algorithm model. In this model, it relates the dependent and independent variables in the input and output space. It takes the all type of the input data and gives the output with the high accuracy; this model not depend on the Dimensionality space and instead of SVR we use the straight line continues values in the space. We use this model to forecast the extreme weather events rather than any algorithm model in the machine learning models. The minimum two support vectors is used to forecast the event, on the basis of all other machine learning models we use this SVR models to refer more to forecast. On our study we give more preference to the support vector regression to predict the cyclone precipitation. So, we apply the machine learning to forecast the cyclone precipitation instead of manual forecasting. This SVR gives the equation in terms of the factors affecting the cyclone precipitation. We know the factor values to get the predicted value in Machine learning process.

REFERENCES

[1]. Machine Learning in Tropical Cyclone Forecast Modeling: A Review [Rui Chen, Weimin Zhang, and Xiang Wang] [Published: 27June 2020.]
[2]. Tropical cyclones and climate change [Thomas R. Knutson, John L. McBride, Johnny chan, Kerry Emanuel, Greg Holland, Chris Land Sea, Isaac Held, James P. Kossin, A. K. Srivastava, Masato sugi] [Published online: 21 February 2010] [Doi: 10.1038/NGE0779]
[3]. Response of tropical sea surface temperature, precipitation, and tropical cyclone-related variables to changes in global and local forcing [Kerry Emanuel, Adam Sobel] [published 20 June 2013] [doi:10.1002/jame.20032, 2013]
[4]. Homogeneous record of Atlantic hurricane surge threat since 1923 [Aslak Grinsted, John C. Moore, Svetlana Jevrejeva]
[5]. Machine Learning Applied to Weather Forecasting [Mark Holmstrom, Dylan Liu, Christopher Vo] [Dated: December 15, 2016]
[6]. A Novel Nonlinear Combination Model Based on Support Vector Machine forRainfall Prediction [Keshen Lu, Lingzhi Wang] [DOI 10.1109/CSO.2011.50]

[7]. Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation [Christopher B. Field, Vicente Barros, Thomas F. Stocker, Qin Dahe, David Jon Dokken, Gian-Kasper Plattner, Kristie L. Ebi, Simon K. Allen, Michael D. Mastrandrea, Melinda Tignor, Katharine J. Mach, Pauline M. Midgley] [published 2012]

[8]. Hurricane track variability and secular potential intensity trends [James P. Kossin · Suzana J. Camargo] [Published online: 21 October 2009] [DOI 10.1007/s10584-009-9748-2]

[9]. Is climate change increasing the frequency of hazardous events? [Pascal Peduzz]

[10]. Attribution of climate variations and trends to human influences and natural variability [Kevin E. Trenberth] [publishe:2011] [doi: 10.1002/wcc.142]

[11]. SVM-Based Model for Short-Term Rainfall Forecasts at a Local Scale in the Mumbai Urban Area, India [Vinay Nikam, Kapil Gupta] [published:2014] [DOI: 10.1061/(ASCE)HE.1943-5584.0000875]

[12]. Forecasting daily potential evapotranspiration using machine learning and limited climatic data [Alfonso F. Torres, Wynn R. Walker, Mac McKee] [published :15 December 2010]

[13]. Forced and unforced ocean temperature changes in Atlantic and Pacific tropical cyclogenesis regions [B. D. Santer, T. M. L. Wigley, P. J. Gleckler, C. Bonfils, M. F. Wehne, K. AchutaRao, T. P. Barnett, J. S. Boyle, W. Bru¨ ggemann, M. Fiorino, N. Gillett, J. E. Hansen, P. D. Jones, S. A. Klein, G. A. Meehl, S. C. B. Raper, R. W. Reynolds, K. E. Taylor, W. M. Washington] [published: April 7, 2006] [doi:10.1073/pnas.0602861103]

[14]. Climate Change 2013 The Physical Science Basis [Thomas F. Stocker, Dahe Qin, Gian-Kasper Plattner, Melinda M.B. Tignor, Simon K. Allen, Judith Boschung, Alexander Nauels, Yu Xia, Vincent Bex, Pauline M. Midgley] [published 2013]

[15]. Bender, M.A.; Ginis, I.; Kurihara, Y.J. Numerical simulations of tropical cyclone-ocean interaction with a high-resolution coupled model. J. Geophys. Res. Atmos. [1993, 98, 23245–23263].

[16]. Wang, Y.; Rao, Y.; Tan, Z.M.; Schönemann, D. A Statistical Analysis of the Effects of Vertical Wind Shear on Tropical Cyclone Intensity Change over the Western North Pacific. Mon. Weather Rev. 2015, 143, 3434–3453.

[17]. Suykens, J.A.; Vandewalle, J. Least squares support vector machine classifiers. Neural Process. Lett. 1999,9, 293–300.

[18]. Drucker, H.; Burges, C.J.; Kaufman, L.; Smola, A.J.; Vapnik, V. Support vector regression machines. In Proceedings of the 9th International Conference on Neural Information Processing Systems; MIT Press: Denver, CO, USA, 1997; pp. 155–161.

[19]. Lawrence, S.; Giles, C.L.; Tsoi, A.C.; Back, A.D. Face recognition: A convolutional neural-network approach. IEEE Trans. Neural Netw. 1997, 8, 98–113.

[20]. Wijnands, J.S.; Qian, G.; Kuleshov, Y. Variable selection for tropical cyclogenesis predictive modeling. Mon. Weather Rev. 2016, 144, 4605–4619.

[21]. Richman, M.B.; Leslie, L.M.; Ramsay, H.A.; Klotzbach, P.J. Reducing Tropical Cyclone Prediction Errors Using Machine Learning Approaches. Procedia Compute. Sci. 2017, 114, 314–323.