

# Efficient Weather Prediction Model using Relevant Machine Learning Approach

Prof Abhishek Pandey, Harsha Khanna

Takshshila College of Engineering and Technology, Jabalpur, Madhya Pradesh, India

## ABSTRACT

### Article Info

Volume 7 Issue 5

Page Number: 124-130

Publication Issue :

September-October-2020

### Article History

Accepted : 15 Sep 2020

Published : 23 Sep 2020

The alternate in worldwide temperatures, current beyond three years natural disasters, rising sea levels, lowering polar regions may be causing the problem of expertise and predicting these weather phenomena. Prediction is a top importance and that they may be run and simulated as laptop simulations to expect climate variables temperature, precipitation, rainfall and and so forth. Temperature prediction is the most crucial venture for predicting early prediction of rainfall may additionally helps to peasant's in addition to for the people because most people in india may be depends upon the agriculture. This dissertation explains approximately a couple of linear regression approach for the temperature estimation or prediction. It may helps to farmers for taking appropriate decisions on crop yielding. As normally at the equal time there can be a scope to research the prevalence of floods or droughts. The a couple of linear regression evaluation technique applied on the dataset of noaa climate records. The test and our a couple of linear regression method take advantage of the ideal consequences for the temperature than simple linear regression methodology and lasso regression, ridge regression.

**Keywords** : Multinomial Linear Regression based classification, judgment, machine learning algorithms, non compensatory, Lasso Regression, Ridge Regression.

## I. INTRODUCTION

Machine Learning is the study and construction of sets of computer instructions that can gain (understanding of deep things) from sample dataset and make data-driven (statements about possible future events) or decisions on new data. Tom M. Mitchell gave/given a formal definition: "A computer program is said to learn from experience E with respect to some class of tasks T and performance measure P if its performance at tasks in T, as

measured by P, improves with experience E"[1]. It involves development of computer programs which changes or learns when exposed to new data which is like data mining. Both systems search through data to look for patterns. However, data mining extracts data for human understanding whereas machine learning uses that data to detect patterns in data and (change to make better/change to fit new conditions) program actions in the same way/in that way.

Machine learning is done always based on (instances of watching, noticing, or making statements) or data, direct experience, or instruction. So, in general, machine learning is about learning to do better in the future based on what was experienced in the past. The goal is to figure out learning sets of computer instructions that do the learning automatically without human (action that helps a bad situation) or help. The machine learning way of thinking can be viewed as "programming by example." Often we have a particular job in mind, such as spam filtering. But rather than programming the computer to solve the job directly, in machine learning, we look (for) methods by which the computer will come up with its own program based on examples that we provide. Machine learning is a core subarea of (not made by nature/fake) intelligence. It is very unlikely that we will be able to build any kind of smart system capable of (doing/completing) complex tasks such as language or vision, without using learning to get there. These tasks are otherwise simply very hard to solve.

The major advantage of machine learning over static programming is the results are often more (very close to the truth or true number) with machine learning than static programming results because the machine learning sets of computer instructions are data driven, and can examine large amounts of data. On the other hand, a human expert who writes static programs is likely to be guided by sloppy/unclear impressions or perhaps an examination.

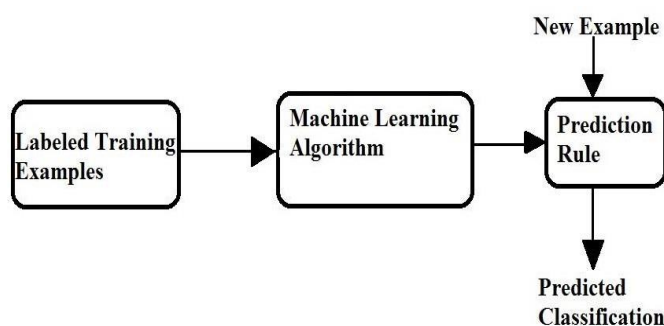


Figure 1. Diagram of a general Machine Learning Process

a relatively small number of examples or data. Figure 1 show the general process involved in a typical

machine learning model.

For instance, it is easy for humans to label images of letters by the character represented, but we would have trouble in explaining how we did it in precise terms. Another reason to study machine learning is the hope that it will provide insights into the general phenomenon of learning. Some of the details we might learn are the intrinsic properties of a given learning problem that makes it hard or easy to solve and know ahead of time about what is being learned to be able to learn it effectively. In this report, we are interested in designing machine learning algorithms, but we also hope to analyze them mathematically to understand their efficiency. Through theory, we hope to understand the intrinsic difficulty of a given learning problem and we attempt to explain phenomena observed in actual experiments with learning algorithms.

## II. RELATED WORK

This Area will start with a light about key walking idea using this guess look for after down after with the aid of way of a move chart which clears up step sharp (solid basic structure on which bigger things can be built) covered. To treatment the guess, we would clear a case for strengthening results from a dating net net page and the growing python content applied from run it. The python substance might be visibly lit up maximum evacuated aspect watchful inside the walking with subsection checked for after with the useful thing/valuable supply of the results were given from that code.

The focal (things you constantly think about) within the usage of this figuring are its ridiculously overdone (high) quality, wild to amazing and interesting events and no (not having enough of something) about data. It has its personal top notch lovely setting tangles in like way which is probably its needed things for part of memory and very confusing check. This guess works with every number-based and clean homes.

Ling chen, xu lai (2020) [1] in comparison the experimental results received/got through (not made by nature/fake) nerve-related/brain-related community (ann) and autoregressive incorporated moving common (arima) in forecasting the hourly wind speed. On comparison, ann model produces a higher result while compared to arima model.

Jyoti clear jellywal, renuka nagpal et al., (2019) [2] has completed crime analysis the use of ok-method clustering at the crime dataset. This version is developed the usage of quick miner device. The grouped-together effects are carefully studied via way of plotting the values through the years. The model this way ends/decides from the (process of figuring out the worth, amount, or quality of something) that the range of murders decreases from 2018 to 2019.

Shiju sathyadevan, devan m. S et al., (2020) [3] expected/looked ahead to the areas that have too much/too many opportunity for crime number and saw (in your mind) crime inclined areas. The authors labeled the facts the use of the childlike (because of a lack of understanding) bayes classifiers set of rules that is a supervised learning (in almost the same way) to a (related to studying numbers) approach for class and has given 90% (quality of being very close to the truth or true number).

Lawrence mcclendon and natarajan meghanathan (2020) [4] used many (statement about a possible future event) sets of computer instructions along with linear moving backward, (serving to add something) moving backward, and selection stump sets of computer instructions the usage of the equal set of input (functions), on the communities and crime dataset. Standard, the linear moving backward set of rules gave the pleasant results in comparison to the three decided/figured out on sets of computer instructions. The main advantage of linear moving backward set of computer instructions is, it could deal with randomness inside the test data to a positive

extent (without getting/causing too much 15 of (statement about a possible future event) mistakes).

Rasoul kiani, siamak mahdavi et al., (2020) [5] proposed a (solid basic structure on which bigger things can be built) for (describing a possible future event) the crimes by means of the use of clustering sets of computer instructions. That is carried out the use of fastminer tool. If you want to boom the performance of (statement about a possible future event), ga ((related to tiny chemical assembly instructions inside of living things) set of rules) is used for detecting The victim is expected by the model. This gadget is as a result expected/looked ahead to to ease the heavy load of the police branch in based totally mostly on these sets of computer instructions the solving the murder events. (things that aren't part of the main group) in the records. This model has produced a (quality of being very close to the truth or true number) of 91.

The rate of crimes that takes area because of (related to the meaning of words) social engineering attacks and explores the (ability to actually be done) of (describing a possible future event) (related to people who use a product or service) likelihood of being harmed or influenced by lie/lying-based totally completely attacks. The authors have (described a possible future event) the use of logistic moving backward and a random (natural area with trees) (statement about a possible future event) model, with the (quality of being very close to the truth or true number) costs of .68 and .71, (match up each pair of items in order).

S. Sivaranjani, s. Sivakumari et al., (2019) [7] used diverse clustering processes just like the ok-approach clustering, agglomerative clustering and density based totally spatial clustering with noise (dbscan) algorithms are used to cluster crime activities in tamil nadu. The overall overall performance of every clustering algorithms is evaluated the usage of the metrics which include precision, recollect and f-

diploma, and the results are in comparison. Primarily based at the above metrics, dbSCAN set of rules gave the splendid results as compared to the alternative selected algorithms.

Chirag kansara, rakhi gupta et al., (2019) [8] proposed a version which examine the sentiments of the people in twitter and predicts whether or not they're able to grow to be hazard to unique individual or society. This version is applied the usage of naive bayes classifier which classifies the humans by manner of sentiment assessment.

### III. PROPOSED WORK AND RESULTS

#### MODEL DESCRIPTION:

Linear regression set of rules with a couple of variables predicting the final results of the aim variable. This is also known as a couple of linear regression. Simple linear regression model has a non-stop outcome and one predictor, whereas a a couple of linear regression version has a non-forestall outcome and more than one predictors (non-stop or specific). A easy linear regression version would have the form:

$$y = \alpha + x\beta + \epsilon$$

A multivariable or multiple linear regression model would take the form:

$$y = \alpha + x_1\beta_1 + x_2\beta_2 + \dots + x_k\beta_k + \epsilon$$

where Y is a continuous established variable, x is a single predictor in the simple regression model, and x1, x2, ..., xk are the predictors within the more than one regression model.

In Facts, the imply squared mistakes (mse) or suggest squared deviation (msd) of an estimator (of a manner for estimating an unobserved amount) measures the not unusual of the squares of the errors — this is, the average squared distinction among the predicted values and what is definitely predicted.

Multiple linear regression can model more complex courting which comes from numerous capabilities together. They should be utilized in instances wherein one specific variable isn't evident sufficient to map the connection between the independent and the mounted variable.

#### 2.Lasso regression

Mathematics behind lasso regression is quiet similar to that of ridge only difference being instead of adding squares of theta, we will add absolute value of  $\Theta$ .

$$\min \left( \|Y - X\theta\|_2^2 + \lambda \|\theta\|_1 \right)$$

Here too,  $\lambda$  is the hypermeter, whose value is equal to the alpha in the Lasso function.

#### 3. Ridge regression

So, now you have an idea how to implement it but let us take a look at the mathematics side also. Till now

```

poly_degree = 2, interaction_only = False
Model: linear_model.Ridge

Alpha (Regularization strength): 3

X_train.shape = (100284, 78)
y_train.shape = (100284, 1)
X_test.shape = (8457, 78)
y_test.shape = (8457, 1)
For training set:
Mean squared error (train): 4.773

Variance score (train): 0.847

model_runtime (training) = 0.092 (seconds)

model_runtime (predict train set) = 0.007 (seconds)

For test set:
Mean squared error (test): 4.675

Variance score (test): 0.837

model_runtime (predict test set) = 0.001 (seconds)

Correct format and time range!
Prediction start from 2016-10-05 00:00:00

Prediction end at 2016-10-15 00:00:00

Detail in df_plot.csv:
    
```

our idea was to basically minimize the cost function, such that values predicted are much closer to the desired result.

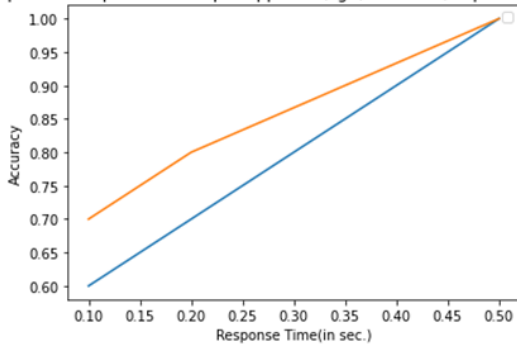
Now take a look back again at the cost function for regression line.

$$\min \left( \|Y - X(\theta)\|_2^2 + \lambda \|\theta\|_2^2 \right)$$

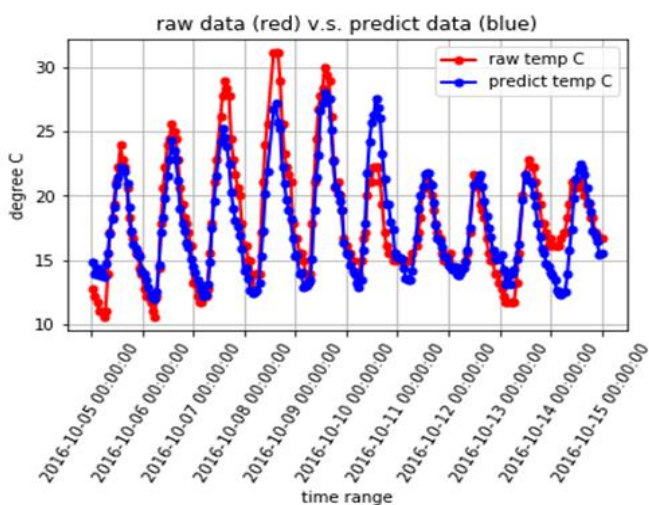
Here if you notice, we come across an extra term, which is known as the penalty term.  $\lambda$  given here, is actually denoted by alpha parameter in the ridge function. So by changing the values of alpha, we are basically controlling the penalty term. Higher the values of alpha, bigger is the penalty and therefore the magnitude of coefficients are reduced.

#### IV. RESULTS

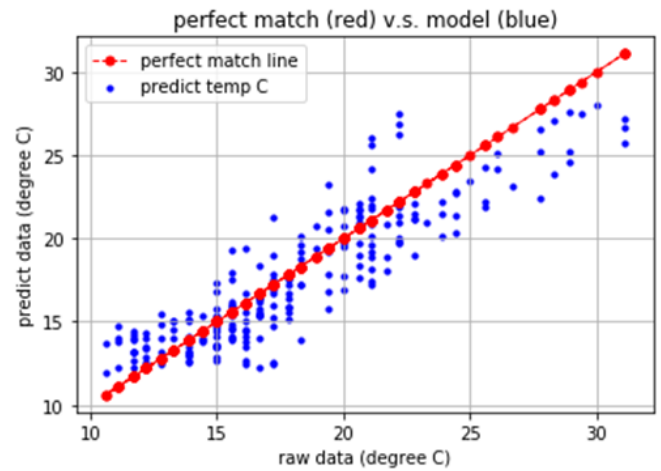
Comparison Graph of Base Paper Approach(LgR) And MLR(Proposed Approach)!



(a) This represent comparison of temperature in centigrade and time with blue line in graph represent predicted value of temperature and red line actual value of temperature of weather dataset.



(b) This represent comparison of predicted data and raw data with blue dots in graph represent predicted value of temperature and red line perfect line of temperature of weather dataset in Multinomial linear



(c) **Comparison Graph** : This graph represents comparison of Multinomial linear regression (orange line) and LgR(Blue line) algorithm with respect to response time of model and accuracy of both the models.

#### V. CONCLUSION

A Maximum of the existing structures are implemented the use of statistical procedures, the above module turned into applied using multi-linear regression and made the system extra correct than previous prediction techniques. The hassle with the existing modules lies of their inability to keep in mind the impact of every value of every parameter, on the relation. In other phrases, the effect of every single cost isn't always considered and a more generalized equation is produced as opposed to a unique relation. For example, inside the svm technique, a plane that carries the facts is generated and the equation of this plane is used as a relation for prediction. This poses a trouble as information points that are a ways apart from every different in terms of significance grow to be at the identical aircraft and their person effect receives unnoticed. The usage of multi linear regression deals with this problem and optimizes the results. Ancient records used for this module has impartial attributes like temperature, wind pace, wind direction, humidity, and ecosystem stress. They may be used to calculate the quantity of rainfall. This

module offers an accurate prediction 88% of the time.

## VI. FUTURE WORK

In order that we will expect rain in the destiny 12 months by using knowing weather factors which is very useful for farmers for their agricultural work. That is the handiest prediction regarding rain but no longer correct due to weather elements. As we education and organisation personnel making prepared in some time.

## VII. REFERENCES

- [1]. L. Miller and X. Xu, "Ambient PM2.5 human health effects—Findings in China and research directions," *Atmosphere*, vol. 9, no. 11, p. 424, 2020.
- [2]. Y. Hao and Y.-M. Liu, "The influential factors of urban PM2.5 concentrations in China: A spatial econometric analysis," *J. Cleaner Prod.*, vol. 112, pp. 1443–1453, Jan. 2019.
- [3]. J. Chen, Y. Song, G. Wang, and Q. Liu, "WSN-aided haze pollution governance: Modelling public willingness based on structural equations," *Int. J. Sensor Netw.*, vol. 29, no. 2, pp. 111–120, 2019.
- [4]. T. Qiu, J. Liu, W. Si, and D. O. Wu, "Robustness optimization scheme with multi-population co-evolution for scale-free wireless sensor networks," *IEEE/ACM Trans. Netw.*, vol. 27, no. 3, pp. 1028–1042, Jun. 2019.
- [5]. T. Qiu, B. Li, W. Qu, E. Ahmed, and X. Wang, "TOSG: A topology optimization scheme with global-small-world for industrial heterogeneous Internet of Things," *IEEE Trans. Ind. Informat.*, vol. 15, no. 6, pp. 3174–3184, Jun. 2019.
- [6]. J. Chen, J. Lu, J. C. Avise, J. A. DaMassa, M. J. Kleeman, and A. P. Kaduwela, "Seasonal modeling of PM2.5 in California's San Joaquin Valley," *Atmos. Environ.*, vol. 92, pp. 182–190, Aug. 2014.
- [7]. P. E. Saide, G. R. Carmichael, S. N. Spak, L. Gallardo, A. E. Osses, M. A. Mena-Carrasco, M. Pagowski, "Forecasting urban PM10 and PM2.5 pollution episodes in very stable nocturnal conditions and complex terrain using WRF–Chem CO tracer model," *Atmos. Environ.*, vol. 45, no. 16, pp. 2769–2780, May 2011.
- [8]. X. Yang, Z. Zhang, Z. Zhang, L. Sun, C. Xu, and L. Yu, "A long-term prediction model of beijing haze episodes using time series analysis," *Comput. Intell. Neurosci.*, vol. 2016, pp. 1–7, Jul. 2016.
- [9]. J. J. Carbajal-Hernández, L. P. Sánchez-Fernández, J. A. Carrasco-Ochoa, and J. F. Martínez-Trinidad, "Assessment and prediction of air quality using fuzzy logic and autoregressive models," *Atmos. Environ.*, vol. 60, pp. 37–50, Dec. 2012.
- [10]. P. J. G. Nietoo, E. Combarro, J. Del Coz Díaz, and E. Montañés, "A SVM-based regression model to study the air quality at local scale in Oviedo urban area (Northern Spain): A case study," *Appl. Math. Comput.*, vol. 219, no. 17, pp. 8923–8937, May 2013.
- [11]. L. Wu and H. Zhao, "Using FGM(1,1) model to predict the number of the lightly polluted day in Jing-Jin-Ji region of China," *Atmos. Pollut. Res.*, vol. 10, no. 2, pp. 552–555, Mar. 2019.
- [12]. A. Challoner, F. Pilla, and L. Gill, "Prediction of indoor air exposure from outdoor air quality using an artificial neural network model for inner city commercial buildings," *Int. J. Environ. Res. Public Health*, vol. 12, no. 12, pp. 15233–15253, Dec. 2015.
- [13]. B. Liu, S. Yan, J. Li, G. Qu, Y. Li, J. Lang, and R. Gu, "A sequence-to-sequence air quality predictor based on the n-step recurrent prediction," *IEEE Access*, vol. 7, pp. 43331–43345, 2019.
- [14]. Y. Bai, Y. Li, X. Wang, J. Xie, and C. Li, "Air pollutants concentrations forecasting using back propagation neural network based on wavelet decomposition with meteorological

conditions," *Atmos. Pollut. Res.*, vol. 7, no. 3, pp. 557–566, May 2016.

- [15]. Q. Wu and H. Lin, "A novel optimal-hybrid model for daily air quality index prediction considering air pollutant factors," *Sci. Total Environ.*, vol. 683, pp. 808–821, Sep. 2019.
- [16]. Y. Zhan, Y. Luo, X. Deng, H. Chen, M. L. Grieneisen, X. Shen, L. Zhu, and M. Zhang, "Spatiotemporal prediction of continuous daily PM<sub>2.5</sub> concentrations across China using a spatially explicit machine learning algorithm," *Atmos. Environ.*, vol. 155, pp. 129–139, Apr. 2017.

**Cite this article as :**

Sh