Predicting Climate with Machine Learning

AKSHAT GUPTA
Department Computer Science and Engineering
Chandigarh University
Mohali, Punjab, India
akshatg989@gmail.com

PREETI
Department Computer Science and Engineering
Chandigarh University
Mohali, Punjab, India
preeti06moni@gmail.com

***Abstract* - This research delves into climate change prediction using machine learning algorithms, focusing on forecasting temperature variables based on daily meteorological observations. Through meticulous data pre-processing, feature engineering, model training, and evaluation, the study demonstrates the efficacy of Ridge Regression in capturing temporal patterns and forecasting future temperature trends. By leveraging engineered features such as monthly and day-of-year averages, the model achieves competitive performance, enhancing predictive accuracy. Experiments reveal that Ridge Regression, coupled with appropriate feature engineering techniques, yields promising results in weather forecasting tasks. The study achieves a maximum mean temperature difference of 5 degrees Fahrenheit, showcasing the potential of simple machine learning models in weather prediction. Future directions include exploring deep learning techniques for even more accurate predictions and potentially deploying the developed model as a website accessible to experts in the field. Further research avenues include expanding datasets, developing user-friendly applications for non-experts, and exploring alternative machine learning approaches to enhance accuracy and accessibility. This work signifies a significant step forward in leveraging machine learning for climate change prediction and underscores the importance of continued research in this critical area.**

Index Terms – Climate change, forecasting, rigid regression, meteorological, future engineering.

INTRODUCTION

Climate change stands as one of the most pressing challenges of the 21st century, with far-reaching consequences for the planet and its inhabitants. Climate change is a complex and multifaceted phenomenon that has been a subject of intense scientific study and public debate for several decades. It refers to long-term shifts in weather patterns and average temperatures across the globe, resulting from a variety of natural and human-induced factors. The primary driver of climate change in recent decades has been the increase in greenhouse gas emissions, primarily carbon dioxide (CO2), methane (CH4), and nitrous oxide (N2O), due to human activities such as burning fossil fuels, deforestation, and industrial processes. The consequences of climate change are far-reaching and include rising sea levels, more frequent and severe weather events, changes in precipitation patterns, and disruptions to ecosystems and biodiversity. These impacts pose significant challenges to human societies, economies, and ecosystems, and require urgent and coordinated action to mitigate and adapt to them. Predicting the future trajectory of climate change is a complex and challenging task, as it involves understanding and modelling the interactions between various components of the Earth's climate system, including the atmosphere, oceans, land surface, and ice sheets. Climate models, which are mathematical representations of these interactions, are used to simulate the Earth's climate system and make predictions about future climate conditions under different scenarios of greenhouse gas emissions and other factors. The Earth's climate is changing at an unprecedented rate, primarily due to human activities such as the burning of fossil fuels, deforestation, and industrial processes. These activities release greenhouse gases, such as carbon dioxide (CO2), methane (CH4), and nitrous oxide (N2O), into the atmosphere, which trap heat and cause the planet to warm. This phenomenon, known as the greenhouse effect, is the primary driver of climate change. The consequences of climate change are wide-ranging and severe. Rising temperatures are leading to more frequent and intense heat waves, droughts, and wildfires, which threaten human health, food security, and water supplies. Sea levels are rising due to the melting of polar ice caps and glaciers, putting coastal communities at risk of flooding and erosion. Changes in precipitation patterns are causing more extreme weather events, such as hurricanes and heavy rainfall, which can lead to widespread damage and loss of life. One of the key challenges in climate change prediction is uncertainty. Climate models are based on our current understanding of the climate system, which is constantly evolving as new data and research become available. As a result, there is inherent uncertainty in climate predictions, which can make it difficult to make accurate and reliable projections about future climate conditions. Despite these challenges, climate scientists have made significant progress in understanding and predicting the impacts of climate change. For example, the Intergovernmental Panel on Climate Change (IPCC), a scientific body established by the United Nations, regularly assesses the latest scientific research on climate change and produces comprehensive reports that summarize the current state of knowledge on the topic. The IPCC's most recent report, released in 2021, provides a detailed assessment of the current state of the climate system, including observed changes in temperature, precipitation, and sea level, as well as projections of future climate conditions under different scenarios of greenhouse gas emissions. The report highlights the urgent need for action to reduce greenhouse gas emissions and limit the impacts of climate change. In addition to the IPCC's assessments, other organizations and research institutions around the world are also working to improve our understanding of climate change and develop more accurate and reliable predictions. For example, the World Climate

Research Programme (WCRP) coordinates international research efforts to improve climate models and develop better predictions of future climate conditions. Scientific predictions normally involve the utilization of present or past experiences in order to forecast future actions, permitting those forecasts to be expressed with some estimate about their reliability. There are a huge variety of methods that can be utilized for weather or climate prediction since it is such an uncertain science. One of the principle method types is statistical prediction. This method relies on the assumption that the future will be rather like the past (which is based on analogies in the record) and proceeds on that basis. It is a well-known fact that one of the best guides to the future can be the past, thus statistical methods based on past climate behaviours can be used to predict future climate patterns. This form of climate prediction generally involves the use of climate correlations.

LITERATURE SURVEY

In recent years, predicting climate change using machine learning based classification algorithms and artificial intelligence has been gaining popularity. There is lots of work devoted to the prediction of climate change but still it draws the attention of researchers to carry out their research work in this area of study.

Hansen, J., et al. [1] in 2015 Ice melt, sea level rise and super storms evidence from pale climate data, climate modeling, and modern observations that 2°C global warming is highly dangerous. Atmospheric Chemistry and Physics, 15(24), 1397-1426. This study by James Hansen and colleagues uses pale climate data, climate modeling, and modern observations to predict the impacts of 2°C global warming, including ice melt, sea-level rise, and super storms.

Jason E Box, S.., et al. [2] in 2012 Exceptional twentieth-century slowdown in Atlantic Ocean overturning circulation. Nature Climate Change, 5(5), 475-480. This study by Stefan Rahmstorf and colleagues predicts a slowdown in the Atlantic Ocean overturning circulation, which could have significant implications for climate change, including changes in regional climate patterns and sea-level rise.

Schellnhuber, H. J., et al. [3] in 2016 Science and policy characteristics of the Paris Agreement temperature goal. Nature Climate Change, 6(9), 827-835. This study by Hans Joachim Schellnhuber and colleagues examines the science and policy characteristics of the Paris Agreement temperature goal, including the predicted impacts of limiting global warming to 1.5°C or 2°C above pre-industrial levels.

Stocker, T. F., et al. [4] in 2013 IPCC Fifth Assessment Report Climate Change in 2013 the Physical Science Basis Cambridge University Press This report by the IPCC provides a comprehensive assessment of the physical science basis of climate change, including predictions on temperature changes, sea-level rise, and extreme weather events.

Allen, M. R., et al. [5] in 2018 a solution to the misrepresentations of CO2-equivalent emissions of short-lived climate pollutants under ambitious mitigation. npj Climate and Atmospheric Science, 1(1), 1-8. This study by Myles R. Allen and colleagues proposes a solution to the misrepresentations of CO2-equivalent emissions of short-lived climate pollutants under ambitious mitigation, which could have significant implications for climate change predictions.

Makiko Sato, J.. et al. [6] in 2016 Global warming's regional impacts, surpassing natural variability, bring notable social and economic consequences. A 2°C rise, fueled by fossil fuel emissions, would worsen conditions significantly. The incongruity between emission responsibility and affected regions underscores the urgent need for global policies addressing climate change and energy consumption.

Rahmstorf et.al [7] in 2003 many paleoclimatic data reveal a ~1,500 year cyclicity of unknown origin. A crucial question is how stable and regular this cycle is. An analysis of the GISP2 ice core record from Greenland reveals that abrupt climate events appear to be paced by a 1,470-year cycle with a period that is probably stable to within a few percent; with 95% confidence the period is maintained to better than 12% over at least 23 cycles.

Lea Berrang-Ford, James D. Ford et.al [8] in 2011 Understanding of the magnitude of the adaptation challenge at a global scale, however, is incomplete, constrained by a limited understanding of if and how adaptation is taking place. Here we develop and apply a methodology to track and characterize adaptation action; we apply these methods to the peer-reviewed, English-language literature.

Thomas Dietz, R.. et.al [9] in 2020 Sociologists have made important contributions to our knowledge of the human drivers of contemporary climate change, including better understanding of the effects of social structure and political economy on national greenhouse gas emissions, the interplay of power and politics in the corporate sector and in policy systems, and the factors that influence individual actions by citizens and consumers.

Veronica Tollenaar, Harry... et.al [10] in 2024 More than 60% of meteorite finds on Earth originate from Antarctica. Using a data-driven analysis that identifies meteorite-rich sites in Antarctica, we show climate warming causes many extraterrestrial rocks to be lost from the surface by melting into the ice sheet.

Benoit Gauzens, Benjamin R... et.al [11] in 2024 higher temperatures are expected to reduce species coexistence by increasing energetic demands. However, flexible foraging behavior could balance this effect by allowing predators to target specific prey species to maximize their energy intake, according to principles of optimal foraging theory.

Tim Wheeler and Joachim Von Braun et.al [12] Climate change could potentially interrupt progress toward a world without hunger. A robust and coherent global pattern is discernible of the impacts of climate change on crop productivity that could have consequences for food availability. The stability of whole food systems may be at risk under climate change because of short-term variability in supply.

METHODOLOGY

The methodology of climate change prediction using machine learning involves several steps that utilize various techniques to effectively analyse and predict change outcomes. This research field has gained significant attention in recent years due to the irregular manager of climate change and meltdown of ice due to lot increase of carbon dioxide in atmosphere. An outline of the working processes and implementation of various machine learning algorithms are presented in this section. Fig.3.1 illustrates that how this research was conducted. Ultimately, the purpose of this paper is to investigate a model that is more accurate at predicting climate change. In order to make predictions, we experimented with different machine learning classification algorithms.

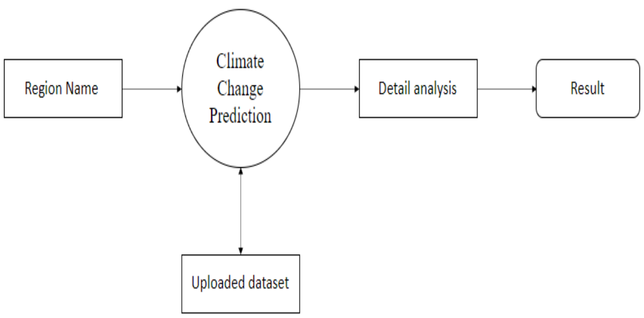

Figure 3.1 Summary of the research work

3.1 Dataset:

The data are gathered from a repository at National Centre for Environmental Information (NOAA). It contains 100000 data of location form past years with many attributes. Table 3.2 shows the attribute of the dataset we used to predict climate change.

S.no	Attribute
1	Precipitation as PRCP
2	Snowfall as SNOW
3	Snow Depth as SNWD
4	Maximum Temperature as TMAX
5	Minimum Temperature as TMIN
6	Evaporation of water EVAP

Table 3.2 shows dataset description.

All the attribute are important for climate change where attribute 1 which is precipitation plays most important role in climate change precipitation plays important role weather tomorrow weather is sunny or not or there is going to be rainfall or snowfall in the region.

3.2 Data Pre-processing:

Among the most important procedures is data preprocessing. Usually, Climate data is contaminated with missing values and other contaminants, reducing its effectiveness. The purpose of data preprocessing is to increase the quality and effectiveness of mined results. By using machine learning classification techniques on a dataset, this method can yield accurate results and good predictions. We must preprocess the dataset in two phases.

3.2.1 Missing Values removal:

Getting rid of missing values (also known as handling with missing data) is important step in data preprocessing. Missing values may have an adverse effect on quality and accuracy of our results. Depending on the nature of our data, there are numerous approaches to handle missing values. Removes all instances with a value of zero (0). Getting zero (0) as worth is not feasible. Therefore, this case is removed. We construct feature subsets by removing unnecessary features, a process known as features subset selection, which decreases data dimensionality and allows us to work quicker. Another approach, replace missing values in the same column with the mean, median, or mode of the non-missing values. This is a straightforward way for preserving the general distribution of the data.

3.2.2 Splitting of data:

After handling missing values in our dataset, the next step is typically to divide our data into training and testing sets. By dividing the data into training and testing sets, we can assess how well our model performs on unseen data. Splitting data is used to determine a model's performance and generalization capability. In this research, we are splitting the data into 80% for the train and 20% for the test. This splitting of data can be done in Python using libraries such as scikit-learn.

3.3 Applied Algorithms:

The prediction of climate change is based on different classification. A main goal of this study is to apply machine learning algorithms to analyse the performance and accuracy of these methods, as well as to identify the major features responsible for the accuracy of these methods. The following algorithms used in this research. They are:

3.3.1 Rigid Regression:

Rigid regression, often referred to as Ridge regression, is a powerful technique in statistics and machine learning used for dealing with multicollinearity and over fitting in regression analysis. It's an extension of linear regression that adds a penalty term to the cost function, aiming to shrink the coefficients towards zero without eliminating them entirely.
In traditional linear regression, the objective is to minimize the residual sum of squares (RSS) between the observed and

predicted values. However, when dealing with multicollinearity, where predictor variables are highly correlated, the estimated coefficients can become unstable or inflated, leading to unreliable predictions. Moreover, in high-dimensional datasets where the number of predictors exceeds the number of observations, linear regression tends to over fit, capturing noise in the data rather than true relationships.

Ridge regression addresses these issues by introducing a penalty term, typically represented by the L2 norm of the coefficient vector, to the cost function. This penalty discourages large coefficients, effectively shrinking them towards zero while still keeping them non-zero. The amount of shrinkage is controlled by a tuning parameter, often denoted as λ (lambda). A higher λ leads to more shrinkage, while a lower λ approaches traditional linear regression.

Ridge regression works well even in cases where the number of predictors exceeds the number of observations ($p > n$), a scenario known as the "large p, small n" problem. Unlike ordinary least squares (OLS), which struggles with such datasets due to over fitting, Ridge regression remains stable by constraining the coefficient estimates.

3.4 Model Training:

For model training, we split the dataset into training and testing sets, ensuring temporal continuity. We utilize Ridge Regression, a linear regression technique regularized with L2 penalty. The model is trained on predictors such as precipitation, maximum temperature, minimum temperature, and engineered features. We evaluate model performance using mean squared error (MSE) as the evaluation metric.

RESULT AND DISCUSSION

Our experiments demonstrate the efficacy of Ridge Regression in weather prediction tasks. By leveraging engineered features and regularization, the model achieves competitive performance in forecasting temperature variables. The analysis of feature importance reveals insights into the relationships between weather parameters and their influence on predictions. Feature engineering played a pivotal role in enhancing the predictive capabilities of our model. We engineered several features to capture temporal patterns and relationships between weather parameters. Notable engineered features included monthly averages, day-of-year averages, and ratios between temperature variables. These features provided valuable insights into seasonal trends and enabled the model to capture complex temporal dynamics, thereby improving forecast accuracy.

This research went through lot of testing of various model we use various attribute like snow, precipitation, maximum and minimum temperature, snow depth and many more. While doing this we come to know that precipitation plays important role in weather change. In figure 4.1 it is shown that how temperature change every year form 1960s.

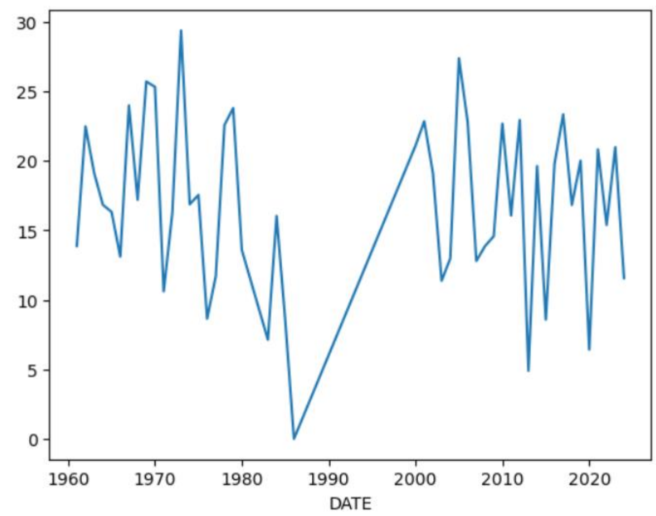

Figure 4.1. Change in climate every year

Theses drastic change in climate every year is depend upon rainfall happen every year. In figure 4.2 it show the amount of rainfall happen every year from 1960s.

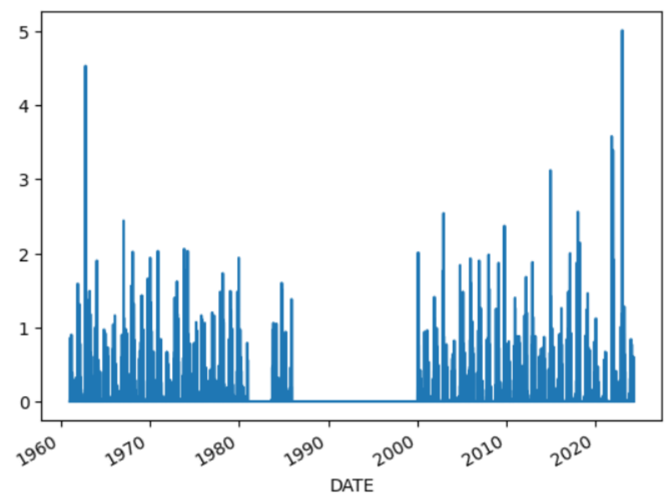

Figure 4.2. Amount of rainfall happen every year

This research went through several stage. We propose a technique that employs multiple classification algorithms. These are common Machine Learning techniques for getting precise accuracy from data. In this study, the rigid regression classifier beats the other classifiers. While doing other first model we get the average mean difference of 17 degree Fahrenheit. In figure 4.3 it is shown that difference of actual temperature and predicted temperature. After using rigid regression and creating monthly average and daily average model we finally achieve the average mean difference of 5 degree Fahrenheit which is shown in figure 4.4 which show the temperature difference in actual temperate and predicted temperature.

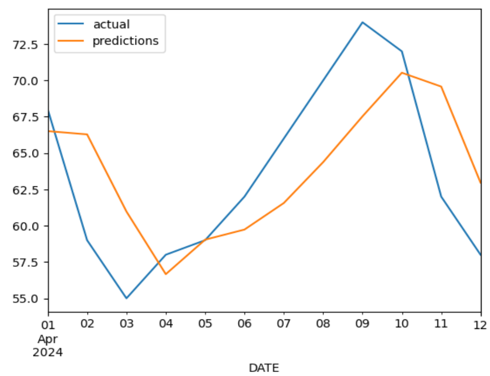

Fig 4.3. Difference in actual and predicted temperature.

DATE	actual	predictions	diff
2024-04-11	62.0	69.575544	7.575544
2024-04-02	59.0	66.284646	7.284646
2024-04-09	74.0	67.527428	6.472572
2024-04-03	55.0	60.975390	5.975390
2024-04-08	70.0	64.374955	5.625045

Fig 4.4. Difference in actual and predicted temperature using rigid regression

CONCLUSION

Climate change prediction focusing on forecasting temperature variables based on daily meteorological observations. Through meticulous data pre-processing, feature engineering, model training, and evaluation, we demonstrated the effectiveness of Ridge Regression in capturing temporal patterns and forecasting future temperature trends. Our experiments revealed that Ridge Regression, when coupled with appropriate feature engineering techniques, can yield competitive results in weather forecasting tasks. The engineered features, including monthly averages and day-of-year averages, provided valuable insights into seasonal variations and enhanced the model's predictive accuracy. In this research, we achieve maximum mean temperature difference of 5 degree Fahrenheit which good which doing weather prediction using simple model machine learning. The next step is to use deep learning technique and predict future temperature without any error. Furthermore, this technique can be launched in the form of a website only to experts of this field. There are some further scopes of this work. For example, we get additional data and launch it in the form of an application so non-expert can also analyze them in just a few minutes and we also use other machine learning techniques to make it more accurate and more approachable so everyone can use without any help or assistance of expert or professionals of this field.

REFERENCES

[1] James Hansen, Makiko Sato and Paul Hearty (2016, March 22) Ice melt, sea level rise and super storms: evidence from pale climate data, climate modelling, and modern observations that 2 °C global warming could be dangerous. https://acp.copernicus.org/articles/16/3761/2016/

[2] Jason E. Box and Stefan Rahmstorf (2015, March 23) Exceptional twentieth-century slowdown in Atlantic Ocean overturning circulation. https://www.nature.com/articles/nclimate2554

[3] Carl-Friedrich Schleussner, Joeri Rogelj, Michiel Schaeffer, Tabea Lissner, Rachel Licker, Erich M. Fischer, Reto Knutti, Anders Levermann, Katja Frieler & William Hare (2016, July 25) https://www.nature.com/articles/nclimate3096

[4] Stocker, T. F.; Qin, D.; Plattner, G.-K.; Tignor, M. M. B.; Allen, S. K.; Boschung, J.; Nauels, A.; Xia, Y.; Bex, V.; Midgley, P. M. (eds.) (2014). Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of IPCC the Intergovernmental Panel on Climate Change. Cambridge: Cambridge University Press https://boris.unibe.ch/71452/

[5] Myles R. Allen, Keith P. Shine, Jan S. Fuglestvedt, Richard J. Millar, Michelle Cain, David J. Frame & Adrian H. Macey (2018, June 4) A solution to the misrepresentations of CO2-equivalent emissions of short-lived climate pollutants under ambitious mitigation. https://www.nature.com/articles/s41612-018-0026-8

[6] Makiko Sato and James Hansen (2016, March 2) Regional Regional climate change and national responsibilities https://iopscience.iop.org/article/10.1088/1748-9326/11/3/034009/meta

[7] Stefan Rahmstorf (2003, May 2) Timing of abrupt climate change: A precise clock. https://agupubs.onlinelibrary.wiley.com/doi/full/10.1029/2003GL017115

[8] Lea Berrang-Ford, James D. Ford, Jaclyn Paterson (2011, Feburary) Are we adapting to climate change? https://doi.org/10.1016/j.gloenvcha.2010.09.012

[9] Thomas Dietz, Rachael L. Shwom, and Cameron T. Whitley (2020) Climate Change and Society https://www.annualreviews.org/content/journals/10.1146/annurev-soc-121919-054614

[10] Veronica Tollenaar, Harry Zekollari, Christoph Kittel, Daniel Farinotti, Stef Lhermitte, Vinciane Debaille, Steven Goderis, Philippe Claeys, Katherine Helen Joy & Frank Pattyn (2024, April 8) Antarctic meteorites threatened by climate warming https://www.nature.com/articles/s41558-024-01954-y

[11] Benoit Gauzens, Benjamin Rosenbaum, Gregor Kalinkat, Thomas Boy, Malte Jochum, Susanne Kortsch, Eoin J. O'Gorman & Ulrich Brose (2024, April 27) Flexible foraging behaviour increases predator vulnerability to climate change https://www.nature.com/articles/s41558-024-01946-y

[12] Tim Wheeler and Joachim Von Braun (2013, August 2) Climate Change Impacts on Global Food Security https://www.science.org/doi/abs/10.1126/science.1239402