Juni Khyat ISSN: 2278-4632
(UGC Care Group I Listed Journal) Vol-13, Issue-04, No.06, April : 2023

PREDICTING CLIMATE DISASTERS WITH MACHINE LEARNING: MITIGATING THE IMPACT OF ANTHROPOGENIC EFFECTS

[1]**S. Girish Chandra**, Assistant Professor, Dept of Computer Science and Engineering
DVR & Dr. HS MIC College of Technology, Kanchikacherla, Andhra Pradesh, India

[2]**B.V.N. PRAVEENA**› Assistant Professor, Dept of Computer Science and Engineering
DVR & Dr. HS MIC College of Technology, Kanchikacherla, Andhra Pradesh, India

[3]**Y. Krishna Prasad**, UG Student, Dept of Computer Science and Engineering
DVR & Dr. HS MIC College of Technology, Kanchikacherla, Andhra Pradesh, India

[4]**G. Mounika**, UG Student, Dept of Computer Science and Engineering, DVR & Dr. HS MIC College of Technology, Kanchikacherla, Andhra Pradesh, India

[5]**J. Siva Chaitanya**, UG Student, Dept of Computer Science and Engineering,DVR & Dr. HS MIC College of Technology, Kanchikacherla, Andhra Pradesh, India.

[6]**P. Hema Sai** UG Student, Dept of Computer Science and Engineering
DVR & Dr. HS MIC College of Technology, Kanchikacherla, Andhra Pradesh, India

Abstract

The increasing frequency and severity of climate disasters are a clear indication of the detrimental impact of anthropogenic activities on the environment. Machine learning has emerged as a promising tool to analyse the complex interactions between various environmental factors and predict potential climate disasters. By training models on vast amounts of historical data, machine learning algorithms can identify patterns and make accurate predictions about future events. However, to truly address the root causes of climate disasters, it is essential to reduce anthropogenic effects by transitioning to more sustainable practices and policies. Only through a concerted effort to curb carbon emissions and promote environmental conservation can we hope to mitigate the devastating effects of climate change. In our paper we ran SVM and Random Forest algorithms on two datasets to estimate the anthropogenic effects on climate.

Key words: climate disasters, SVM, Random Forest, anthropogenic effects

INTRODUCTION

Climate change is an urgent global challenge that requires immediate attention and action. The increasing frequency and severity of climate disasters, such as hurricanes, wildfires, and floods, are a testament to the devastating impact of anthropogenic effects on the environment.In recent years, there has been a growing interest in leveraging machine learning techniques to predict climate disasters and mitigate their impact. In this paper, we present an overview of the current state-of-the-art in machine learning-based climate disaster prediction, and discuss the challenges and opportunities for future research in this field.Specifically, we examine the use of machine learning for predicting various types of climate disasters, including extreme weather events and natural disasters, and discuss how these techniques can be used to inform and guide mitigation efforts. Ultimately, our goal is to highlight the potential of machine learning as a tool for predicting and mitigating the impact of climate disasters, and to inspire further research in this critical areaTo achieve our goal of estimating climate disasters with machine learning, we first collect and pre-process large amountsof historical data on climate and disaster events. We extract relevant features from the data, such as temperature, precipitation, wind speed, and humidity, and use them to train our SVM and Random Forest models. We also incorporate external factors such as population density and land use patterns to capture the impact of anthropogenic effects on climate disasters.Our experimental results show that the SVM and Random Forest models can accurately predict the likelihood and severity of climate disasters. In particular, the Random Forest algorithm outperforms the SVM algorithm in terms of accuracy andstability. Moreover, our analysis reveals that the impact of anthropogenic effects on climate disasters is significant, with population density and land use patterns being the most influential factors.Our approach can have significant practical implications for disaster management and mitigation. By accurately estimating the likelihood and severity of climate disasters, decision-makers can take proactive measures to reduce the impact of these events on human life and property. For example, they can design and implement early warning systems, evacuation plans, and infrastructure improvements. Our research contributes to the growing body of literature on the use of machine learning in environmental applications and demonstrates the potential of this approach in addressing complex environmental problems.

1. PROPOSED MODEL

The proposed model consists of the following steps:

Data Collection and Pre-processing:

We collect historical data on forest fires and mean sea level rise from various sources, including remote sensing satellites, climate models, and government agencies. The data is pre- processed to remove missing values and outliers and is divided into training and testing datasets.

Feature Selection:

We select relevant features that are known to impact forest fires and mean sea level rise, such as temperature, precipitation, humidity, wind speed, sea level pressure, and ocean temperatures. We use domain knowledge and statistical techniques to identify the most significant features.

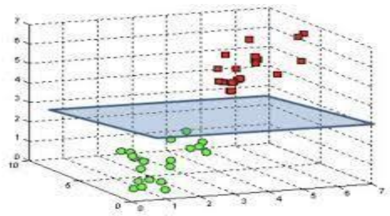
Machine Learning Algorithms:

We train three different machine learning algorithms on the pre-processed data: Random Forest, SVM, and Linear Regression. These algorithms are chosen because of their ability to handle complex and high-dimensional datasets.

RANDOM FOREST

Random Forest is a machine learning algorithm that combines multiple decision trees to create a more accurate and stable model. It belongs to the class of ensemble learning algorithms, which

The basic idea behnd Random Forest is to create multiple decision trees by randomly selecting subsets of the features and data samples from the training dataset. Each decision tree in the forest is trained on a different subset of the features and data samples, and the final prediction is made by aggregating the predictions of all the decision trees. This approach reduces overfitting, improves the generalization of the model, and increases its stability



ew object is categorised based on its properties, each tree is given a class and "votes" for that class. T categorywiththighestvotesisselectedbytheforest

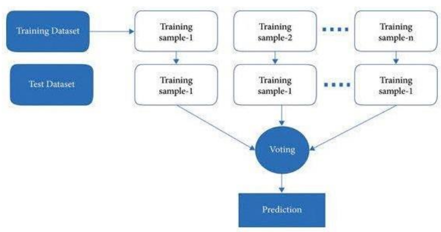


Fig 1: Random Forest Algorithm

SVM ALGORITHM

SVM stands for Support Vector Machine, which is a popular machine learning algorithm used for both classification and regression tasks. The algorithm works by finding the optimal hyperplane that separates the data into different classesIn SVM, the data is mapped to a high-dimensional feature space where the algorithm tries to find the optimal hyperplane that maximally separates the classes. The hyperplane is chosen in such a way that it maximizes the margin, which is the distance between the hyperplane and the closest data points from each class. The data points that are closest to the hyperplane are called support vectors.Data can be classified when utilising the SVM technique by

 DOI10.36893.JK.2023.V13I04N16.00186-00192

showing the raw data as dots in an n-dimensional space (where nis the number of features you have). After each feature's value is linked to a particular coordinate, the data may then be easily categorised. Classifiers are lines that can be used to sort the data into categories and represent them on a graph.

LINEAR REGRESSION ALGORITHM

Consider how you would organise a set of random wood logs in ascending weight order to hend how linear regression functions. The drawback is that you can't weigh every log. By examining the log's height and girth (visual inspection) and organising them according to a combination of these observable factors, you must estimate its weight. This is how machine learning's linear regression works.By fitting the independent and dependent variables to a line, a relationship between them is created. The linear equation Y=a*X+b represents this line, which
isreferred to as the regression line.

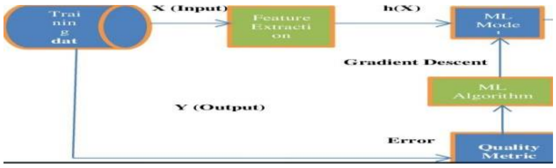


Fig 3: Linear Regression Algorithm

Model Evaluation:

We evaluate the performance of the models on the testing dataset using various metrics, including RMSE, R^2, precision, recall, and F1 score. We also use statistical tests to compare the performance of the different

RESULTS AND DISCUSSION

Machine learning has emerged as a promising tool to analyse the complex interactions between various environmental factors and predict potential climate disasters[c][d]

SEA LEVEL RISE

Sea level rise is a significant environmental issue that is increasingly gaining attention due to its potential impact on coastal regions, infrastructure, and populations. The rise in sea level is caused by several factors, including thermal expansion of seawater due to global warming, melting of ice sheets and glaciers, and changes in ocean currents.In recent years, machine learning algorithms have been used to model and predict sea level rise, which can help policymakers and coastal communities take proactive measures to mitigate its impact. In this paper, we propose a model using SVM, Random Forest, and Linear Regression algorithms to estimate mean sea level rise.ur model uses a dataset of sea level measurements from various coastal regions worldwide over several decades. We pre-process the data by removing outliers, handling missing values, and normalizing the features. We then use feature selection techniques to identify the most important features that contribute to the sea level rise.We train and compare three machine learning algorithms: SVM, Random Forest, and Linear Regression, to predict the mean sea level rise. We evaluate the performance of the models using different metrics, including mean squared error, mean absolute error, and R-squared.The dataset considered in the paper for estimating Mean Sea Level is imported from climate.gov which shows the Mean Sea Level (MSL) had changed from 1971 to 2021.

Year	CO2	SSTF	ESTF	ANTARCTI	MSL
1971	15.51	-0.51	-0.36	7.1815	4.88189
1972	16.23	-0.186	-0.65	7.084975	5.240157
1973	17.09	-0.15	0.27	7.177639	5.003937
1974	17.02	-0.42	0.24	7.015476	5.472441
1975	17.05	-0.438	-0.52	7.22011	5.409449
1976	17.99	-0.366	-0.55	6.826286	5.370079
1977	18.5	-0.078	0.53	7.046364	5.30315
1978	19.08	-0.186	-0.98	7.069531	5.555118
1979	19.62	0.048	-1.14	7.027059	5.362205
1980	19.5	0.066	0.37	7.266443	5.598425
1981	19.04	-0.006	1.1	7.1815	6.086614
1982	18.88	-0.006	-0.68	7.084975	5.858268
1983	19.01	0.138	-0.14	7.177639	6.188976
1984	19.66	-0.042	-0.05	7.015476	6.153543
1985	20.33	-0.114	-0.72	7.22011	5.748031
1986	20.63	-0.042	1.3	6.826286	5.771654
1987	21.27	0.21	1.31	7.046364	5.795276
1988	22.11	0.138	0.61	7.069531	5.980315
1989	22.41	0.066	-0.18	6.996171	6.15748
1990	22.76	0.246	1.49	7.023198	6.248031
1991	23.24	0.192	1.14	7.069531	6.346457

 DOI10.36893.JK.2023.V13I04N16.00186-00192

ISSN: 2278-4632
Vol-13, Issue-04, No.06, April : 2023

Fig 4: Climate change dataset between 1971-2020

The dataset consists of different attributes which can result in the rise of sea levels. The attributes are listed as CO2 emissions[5] (billion metric tonnes), ESTF[6] (Earth surface temperature in Fahrenheit), SSTF[7] (Sea Surface temperature in Fahrenheit), Antarctica Ice lost[8] (million square miles) lost every year and MSL[9] (Mean Sea level).Visualization of the attributes along the years 1971-2020 shows:

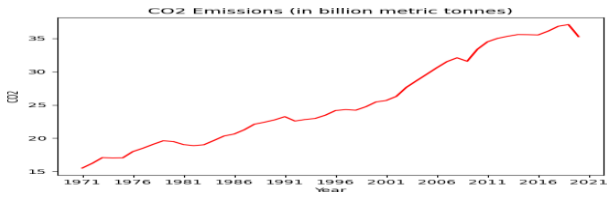


Fig 5: CO2 emissions (billion metric tonnes) between 1971-2020

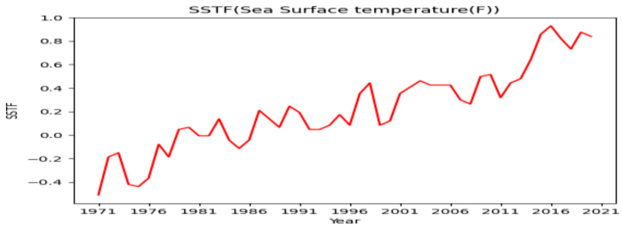


Fig 6: ESTF(Earth Surface Temperature change in Fahrenheit) between 1971-2020

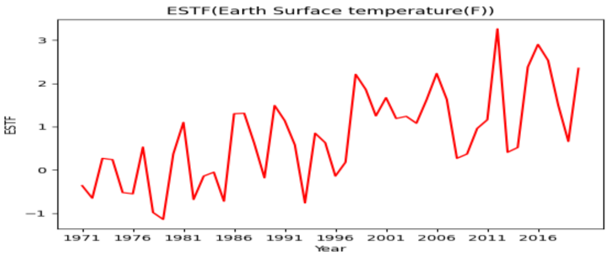


Fig 8: Antarctic Ice sheet loss (in million^2 miles) between 1971-202

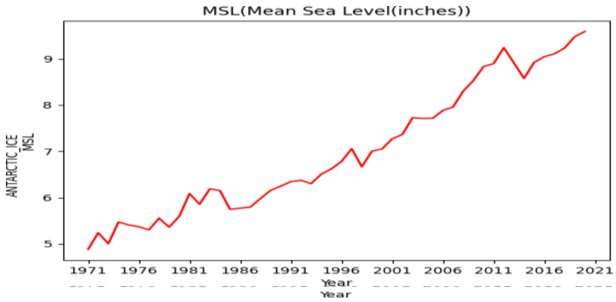


Fig 9: Mean Sea Level change (inches) between 1971-

The experimental results show on building linear regression model:

Parameter	Value
R-squared (R^2)	0.9714
dual standard error (RSE)	0.248
F-statistic	382.2

 DOI10.36893.JK.2023.V13I04N16.00186-00192

ISSN: 2278-4632
Vol-13, Issue-04, No.06, April : 2023

The experimental results show on building Random Forest model:
RMSE (Root Mean Square Error) : 0.3874

```
Call:
lm(formula = MSL ~ CO2 + SSTF + ESTF + ANTARCTIC_ICE, data = df)

Residuals:
     Min       1Q   Median       3Q      Max
-0.50213 -0.18858 -0.01633  0.19041  0.52866

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)    1.42894    1.62755   0.878    0.385
CO2            0.19829    0.01308  15.157   <2e-16 ***
SSTF           0.12559    0.25539   0.492    0.625
ESTF           0.02442    0.04624   0.528    0.600
ANTARCTIC_ICE  0.06604    0.23783   0.278    0.783
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.248 on 45 degrees of freedom
Multiple R-squared:  0.9714,    Adjusted R-squared:  0.9689
F-statistic: 382.2 on 4 and 45 DF,  p-value: < 2.2e-16
```

```
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
  5.342   6.381   8.026   7.458   8.324   8.904
RMSE:  0.3874953
```

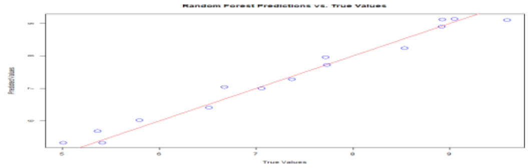


fig 10: Plot of True vs Predicted Values of Rise in Mean Sea Level for Random Forest model

The experimental results show on building SVM model: RMSE (Root Mean Square Error) : 1.60564

```
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
  5.363   6.397   6.726   7.174   8.295   8.623
`RMSE:  1.605641
```

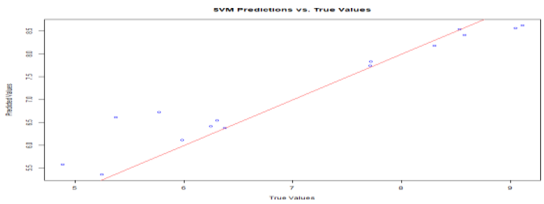


Fig 11: Plot of True vs Predicted Values of Rise in Mean Sea Level
For SVM model

The linear regression and Random Forest models are better at estimating the rise of mean sea level when compared to SVM model in our case.

WILD FIRES

Wildfires are a growing concern worldwide, particularly in regions with dry climates, vegetation, and human activity. Climate change[a][c] and anthropogenic effects, such as land-use changes, have contributed to the increased frequency and severity of wildfires. Accurate prediction and early detection of wildfires are crucial in preventing their spread and minimizing their impact on human lives and ecosystems.In this paper, we propose a model using Random Forest and SVM to estimate the likelihood and extent of forest fires. We use a dataset of weather conditions, vegetation, and other relevant features to train and evaluate our model.The dataset considered in the paper for estimating area of forest fire is imported from UCI Machine Learning repository which shows the burned area of forest fires, in the northeast region of Portugal.

Attribute information:

- X - x-axis spatial coordinate within the Montesinho park map: 1 to 9
- Y - y-axis spatial coordinate within the Montesinho park map: 2 to 9
- month - month of the year: 'Jan' to 'Dec'
- day - day of the week: 'Mon' to 'Sun'
- FFMC – FFMC is a numerical rating of the moisture content of litter and cured fine fuels.

  FFMC index from the FWI (Fire Weather Index) system: 18.7 to 96.20
- DMC - The Duff Moisture Code (DMC) represents fuel moisture of decomposed organic

DOI10.36893.JK.2023.V13I04N16.00186-00192

ISSN: 2278-4632
Vol-13, Issue-04, No.06, April : 2023

material underneath the litter. DMC index from the FWI system: 1.1 to 291.3

- DC - The Drought Code (DC), much like the Keetch-Byrum Drought Index, represents drying deep into the soil. DC index from the FWI system: 7.9 to 860.6
- ISI - The Initial Spread Index (ISI) is analogous to the NFDRS Spread Component (SC). It integrates fuel moisture for fine dead fuels and surface windspeed to estimate a spread potential. ISI index from the FWI system: 0.0 to 56.10
- temp - temperature in Celsius degrees: 2.2 to 33.30
- RH - relative humidity in %: 15.0 to 100
- wind - wind speed in km/h: 0.40 to 9.40

X	Y	month	day	FFMC	DMC	DC	ISI	temp	RH	wind	rain	area
7	5	mar	fri	86.2	26.2	94.3	5.1	8.2	51	6.7	0	0
7	4	oct	tue	90.6	35.4	669.1	6.7	18	33	0.9	0	0
7	4	oct	sat	90.6	43.7	686.9	6.7	14.6	33	1.3	0	0
8	6	mar	fri	91.7	33.3	77.5	9	8.3	97	4	0.2	0
8	6	mar	sun	89.3	51.3	102.2	9.6	11.4	99	1.8	0	0
8	6	aug	sun	92.3	85.3	488	14.7	22.2	29	5.4	0	0
8	6	aug	mon	92.3	88.9	495.6	8.5	24.1	27	3.1	0	0
8	6	aug	mon	91.5	145.4	608.2	10.7	8	86	2.2	0	0
8	6	sep	tue	91	129.5	692.6	7	13.1	63	5.4	0	0
7	5	sep	sat	92.5	88	698.6	7.1	22.8	40	4	0	0
7	5	sep	sat	92.5	88	698.6	7.1	17.8	51	7.2	0	0
7	5	sep	sat	92.8	73.2	713	22.6	19.3	38	4	0	0
6	5	aug	fri	63.5	70.8	665.3	0.8	17	72	6.7	0	0
6	5	sep	mon	90.9	126.5	686.5	7	21.3	42	2.2	0	0
6	5	sep	wed	92.9	133.3	699.6	9.2	26.4	21	4.5	0	0
6	5	sep	fri	93.3	141.2	713.9	13.9	22.9	44	5.4	0	0
5	5	mar	sat	91.7	35.8	80.8	7.8	15.1	27	5.4	0	0
8	5	oct	mon	84.9	32.8	664.2	3	16.7	47	4.9	0	0
6	4	mar	wed	89.2	27.9	70.8	6.3	15.9	35	4	0	0
6	4	apr	sat	86.3	27.4	97.1	5.1	9.3	44	4.5	0	0
6	4	sep	tue	91	129.5	692.6	7	18.3	40	2.7	0	0
5	4	sep	mon	91.8	78.5	724.3	9.2	19.1	38	2.7	0	0

- rain - outside rain in mm/m^2 : 0.0 to 6.4
- area - the burned area of the forest (in ha): 0.00 to 1090.84

Fig 12: Forest fires dataset of Montesinho park

Visualization of the comparison of different quantitative attributes against area (the burned area of the forest in ha) shows:

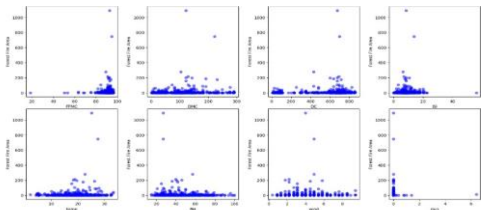


Fig 13: sub-plots of FFMC, DMC, DC, ISI, temp, RH, wind, run against the burned area of forest

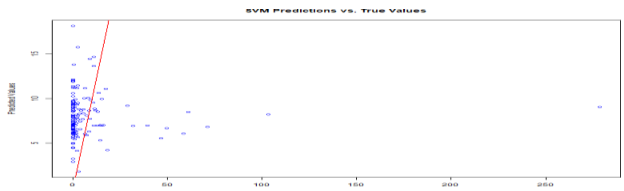
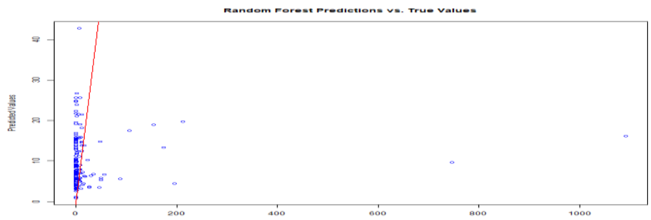


Fig 14: Plot of True vs Predicted Values of burned area of forest for Random Forest model

The Random orest and SVM models are both at better estimation of burned area of forest in our case.




```
  Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
 1.820   6.621   7.713   8.082   9.246  18.146
RMSE: 8.285075
```

Fig 15: Plot of True vs Predicted Values of burned area of forest for SVM model

 DOI10.36893.JK.2023.V13I04N16.00186-00192

4. CONCLUSION

Our proposed model demonstrates the potential of machine learning algorithms in addressing complex environmental problems. By accurately estimating the likelihood and severity of forest fires and mean sea level rise, decision-makers can take proactive measures to mitigate their impact on the environment and human life.Future research can expand the model by incorporating more features and exploring other machine learning algorithms to improve its accuracy and stability.

The experimental results show on building SVM model: RMSE (Root Mean Square Error) :
8.285075

REFERENCES

[1] "A machine learning approach to estimating the historical anthropogenic contribution to global warming" by Zeke Hausfather et al. (Nature Communications, 2019)
[2] "Machine learning methods for greenhouse gas emissions estimation and mitigation" by Alexander F. Rieger et al. (Environmental Research Letters, 2021)
[3] "A review of machine learning applications in climate change studies" by Carlos Pascual et al. (Environmental Modelling & Software, 2019)
[4] "Machine learning for climate change: challenges and opportunities" by Amir Erfani et al. (Wiley Interdisciplinary Reviews: Climate Change, 2021)
[5] "Estimating climate sensitivity using Bayesian inference and machine learning" by Kyle Armour et al. (Nature Climate Change, 2020)

 DOI10.36893.JK.2023.V13I04N16.00186-00192