Journal of Climate Change, Vol. 10, No. 3 (2024), pp. 23-30.
DOI 10.3233/JCC240020

Exploring Climate Interactions: A Fuzzy Rough Set Theory Approach

Leena Sharma[1], Hemlata Saxena[2] and Minakshi Panchal[1,2]*
[1]Department of Applied Science and Humanities, Pimpri Chinchwad Collge of Engineering, Pune
[2]School of Applied Sciences, Career Point University, Kota, Rajasthan, India
✉ mins.29feb@gmail.com

Received June 20, 2024; revised and accepted July 28, 2024

Abstract: Analysing climate change is challenging due to climate data's intricate and dynamic nature. The primary issue is starting with high dimensionality. High dimensionality impacts the model's performance, computation time, cost, and accuracy. Feature selection can be employed as a strategy to address the issue of dimensionality reduction, resulting in more precise insights and the identification of more explicit patterns. Various techniques are used for feature selection. Still, there is scope for progress in this field.

This study uses fuzzy rough set theory (FRST) to perform feature selection in the analysis of climatic data. The dataset in the present study, obtained from Kaggle, is an authentic climate change dataset in the real world. FRST effectively addresses uncertainty and vagueness in climate data by identifying the most relevant temperature parameters and treating them as the deciding attribute.

We identified 25 reducts from the original dataset using FRST. Compared to the original dataset, the best reducts had good classification accuracy. It indicates that FRST reducts preserve the essential features of the original climate data, assuring the reduced dataset's integrity and relevance. FRST was more accurate than usual climate data analysis methods, proving its efficacy.

Keywords: Fuzzy rough set theory; Feature selection; Climate change, Reducts, Greenhouse gases.

Introduction

Climate change is one of the most significant global issues, affecting economies, cultures, and the environment. Global warming and climate change provide significant challenges with potentially terrible consequences, as observed globally. We must implement stringent policies and strictly adhere to safety guidelines to prevent the extensive loss of the ecosystem agriculture (Malhi et al., 2021), water resources, human health (Charlson et al., 2022; Valentová & Bostik, 2021), ecosystems (Malhi et al., 2020), and the economy (Wade & Jennings, 2016). Climate models and computer simulations conducted by climate experts show that the Earth's average temperature could increase by 2–9.7°F (1.1–5.4°C) by the year 2100 (Lindsey & Dahlman, 2021). The primary cause of this increase in temperature is mainly the emission of Green House Gases like water vapours, carbon dioxide (CO_2), methane (CH_4), nitrous oxide (N_2O), ozone (O_3) and chlorofluorocarbon (CFC), etc. CO_2 emission is the leading cause of global warming in the second half of the 20th century (Marland et al., 2003). The 2018 intergovernmental study on climate change predicts severe consequences if global greenhouse gas (GHG) emissions are not stopped within 30 years (Rosenberg, 2010). Climate

*Corresponding Author

research datasets are massive and complicated, requiring advanced data analytics to understand and prevent their consequences. However, climate data's high complexity and uncertainty make analysis difficult (Hassani et al., 2019).

Machine learning algorithms are crucial in climate change data analysis because they can efficiently handle intricate data and identify patterns and linkages that conventional methods may overlook. ML finds high-impact issues, such as overcoming gaps in intelligent grids and disaster management (Rolnick et al., 2023). Support vector machine is used for classification and support vector regression for prediction tasks in climate change analysis (Chen et al., 2011; Khan et al., 2021; Tripathi et al., 2006). The neural network is another machine learning tool used widely for prediction and classification, and it is an effective tool for image processing (Maqsood et al., 2022; Bône et al., 2023). K-means clustering is a popular unsupervised machine learning approach for feature similarity-based dataset clustering. K-Means clustering can group comparable climate patterns, identify regional climate zones, and assess climate variable temporal trends in climate change data (Doan et al., 2023; Huang & Jane, 2009; Sadeghi et al., 2022; Gupta & Jain, 2018). By combining machine-learning approaches with statistical methods or domain-specific models, hybrid machine-learning models can improve climate change data analysis predictions, classification, and insights (Anaraki et al., 2021; Kumar, 2023; Yasodha & Ananthanarayanan, 2015).

In order to use machine learning algorithms, data should be precise and clean. Climate datasets often have many unnecessary or redundant variables, complicating analysis and resulting in less accurate conclusions. Due to errors in measurement, missing values, and environmental dynamics, climate data could be imprecise, compounding the issue of high dimensionality (Papadopoulos & Balta, 2022).

Support vector machine, Regression, Neural network, and K-NN are some methods used for feature selection climate data analysis. Regression is mostly used to find more impacting factors. Different regression models are used as per the necessity, like meta-analysis regression (MRA) (Chaikumbung, 2023), spurious Regression, and regressions involving non-stationary variables (Cummins et al., 2022). logistic regression machine learning algorithms to identify greenhouse gas emissions data pre-processed using min-max normalization forecast well (Adnan et al., 2023). Multi-collinearity and a high number of variables may make OLS Regression impractical, but LASSO regression, Ridge regression, and Elastic net Regression can overcome OLS's shortcomings (Yamaka et al., 2021). Regression has specific limitations: it cannot handle data with multi-collinearity (Dormann et al., 2013), data quality issues like missing or imprecise data (Thorne et al., 2011), data-infused with outliers (Hawkins, 2004). Recent advancements in artificial intelligence, including deep learning, have shown promise, but their application in climate studies is still evolving.

Fuzzy Rough Set Theory (FRST) is a mathematical framework that combines fuzzy logic and rough set theory (Dubois & Prade, 1990), providing an adaptable approach for handling uncertainty and imprecision in data processing, feature extraction, Rule induction, and decision-making (Acharjya & Rathi, 2022; Bhatt & Gopal, 2005; Ewees et al., 2020; Lasisi et al., 2016; Pamucar et al., 2023; Tsang et al., 2008; Zhao et al., 2010). It enhances the current rough set theory by incorporating partial membership, enabling a more nuanced depiction of uncertain and indistinct information (Jensen & Shen, 2009). Fuzzy Rough Set Theory provides a flexible framework for addressing uncertainties in predicting the impact of greenhouse gases. It enhances interpretation using a nuanced approach to data imprecision and ambiguity (Zadeh, 1965).

Fuzzy rough set theory has shown effectiveness for feature selection for complex and uncertain datasets in recent years. Fuzzy rough set theory handles imprecision and uncertainty, making it ideal for climate data. Fuzzy rough set theory can improve traditional data analytics by identifying relevant features and removing redundancy, providing more accurate and meaningful insights into climatic trends and interactions. This study applies fuzzy rough set theory to analyse climate data and uncover essential features causing temperature change. Doing so presents a sophisticated method for managing uncertainty and improving our understanding of the complex nature of climate interactions. Also, we applied regression analysis to identify essential features and compare the performance of both methods for the prediction of temperature change.

Research Methodology

Fuzzy Rough Set Theory (FRST) is an expanded version of Rough Set Theory (RST) that incorporates fuzzy set concepts to handle uncertainty and vagueness. The fundamental mathematical terminology utilized in FRST includes the following:

Fuzzy Sets and Fuzzy Relations

- Fuzzy Set $\overline{A}$ on $U \times X$:
  $\overline{A}$ is a fuzzy set defined on the Cartesian product $U \times X$ where U is the set of Universe, and X is the set of attributes. $A(\mu, x)$ Represents the membership degree μ of x in the fuzzy set $\overline{A}$
  Mathematically, $\overline{A}: U \times X \rightarrow [0, 1]$
- **Fuzzy Relation $\overline{R}_A$ on $U \times U$:**
  R_A is the fuzzy relation resulting from the fuzzy set $\overline{A}$, and it shows associations between elements in the universe U. Fuzzy relation $\overline{R}_A\,(s, t)$ for the elements s, t, $\in U$, represents the degree between u and v based on the fuzzy set $\overline{A}$.
  Mathematically, R_A: $U \times U \rightarrow [0, 1]$, where
  $R_A(s, t) = \min x \in X \{A(s, x), A(t, x)\}$)

Lower and Upper Fuzzy Approximations:

- **Fuzzy Lower Approximation ($\underline{A}(B)$):**
  The set consists of features that certainly belong to the fuzzy set $\overline{A}$, which is a subset of $\overline{B}$. The minimum function assures the inclusion of only things with high assurance.The fuzzy lower approximation of $\overline{A}$ is the union of the lower membership degrees of $\overline{A}$ for each element in $\overline{B}$ where subset $\overline{A} \subseteq U$,
  Mathematically, $\underline{A}(B) = \cup u\ B_A(u)$, where $A(u) = \min_{x \in X} A(u, x)$.
- **Fuzzy Upper Approximation ($\overline{A}(B)$):**
  This includes features about the fuzzy set A within the subset B. The maximal function ensures the inclusion of components that have partial membership.The fuzzy upper approximation of $\overline{A}$ is the union of the upper membership degrees of A for each element in $\overline{\boldsymbol{B}}$ where subset $B \subseteq U$,
  Mathematically,

$$\overline{A}(B) = \cup u \in BA(u), \text{ where } A(u) = \max_{x \in X} A(u, x).$$

Core and Boundary Regions

- **Fuzzy Core ($\mathit{Core}\ (\overline{A})$):** The fuzzy core of set A consists of items for which the lower and upper approximations are identical. This indicates that these components indeed belong to $\overline{A}$. The fuzzy core of $\overline{A}$ denotes the set of features for which the lower and upper approximations overlap entirely.
  Mathematically, Core $(A) = \{u \in U |\ A(u) = A(u)\}$.
- **Fuzzy Boundary (∂A):** The fuzzy boundary of $\overline{\boldsymbol{A}}$ represents the set of features for which the lower and upper approximations do not overlap.

$$\text{Mathematically, } \partial A = U \backslash \mathrm{Core}(A).$$

Fuzzy Reduct ($\overline{A}$) and Decision Rules

A fuzzy reduct is a minor collection of attributes that preserves the discernibility properties in the fuzzy context. Identifying fuzzy reducts involves finding a subset of attributes $B \subseteq X$ such that $\overline{A}(B) = \underline{A}(B)$.
Mathematically, Reduct $(A) = \{x \in X \mid \underline{A}(x) = \overline{A}(x)\}$.

- **Fuzzy Decision Rule ($B \rightarrow D$):**
  A fuzzy decision rule is derived from a fuzzy reduct B and decision attribute D. It represents a relationship between the subset of attributes B and the decision attribute D in the fuzzy context.

$$\text{Mathematically, } B \rightarrow D \text{ holds if } A(B \cup \{D\}) = A(B \cup \{D\}).$$

Fuzzy Rough Set Theory combines fuzzy and rough sets to handle imprecision and uncertainty. The mathematical framework involves defining fuzzy relations based on fuzzy sets, computing fuzzy lower and upper approximations, identifying fuzzy core and boundary regions, and determining fuzzy reducts to extract meaningful information in uncertainty.

Climate data often originate from diverse sources such as satellite observations, ground-based measurements, and climate models. Sometimes, it is susceptible to missing or incomplete observations. Fuzzy Rough Set Theory is uniquely suited for analysing complex climate-related data because it can handle uncertainties, integrate multisource data, extract meaningful patterns, and provide decision support in fuzzy environments.

Case Study

In this study, we used climate data from May 1983 to December 2008 taken from the Kaggle website. The available data include a total of 9 attributes and 308 observations. Details about the data are given in Table 1.

Pre-Processing of the Data

Fuzzy Rough Set Theory (FRST) is a powerful method for detecting and managing outliers, missing values, and erroneous data in datasets. It achieves this by employing fuzzy membership functions and approximations. The algorithm assigns a membership degree to each data point, enabling the identification of outliers that significantly depart from the usual data ranges. FRST establishes border regions by utilising lower and higher approximations, thereby identifying outliers as data points that fall outside these bounds. FRST addresses missing and inaccurate data by incorporating uncertainty via fuzzy sets and imputing values based on related data

Table 1: Data details

Sr. No.	*Attribute*	*Details*	*Source*
1	Temp	The deviation, measured in degrees Celsius, between the average global temperature during a specified period and a reference value	Climatic Research Unit at the University of East Anglia.
2	CO_2	Carbon Dioxide	ESRL/NOAA Global Monitoring Division
3	N_2O	Nitrous Oxide	
4	CH_4	Methane	
5	CFC-11	Tri Chloro-Fluoro methane	
6	CFC-12	Di Chloro-Difluoro methane	
7	Aerosols	The mean stratospheric aerosol optical depth at 550 nm	Godard Institute for Space Studies at NASA
8	TSI	The total solar irradiance	SOLARIS-HEPPA project website.
9	MEI	multivariate El Nino Southern Oscillation index	ESRL/NOAA Physical Sciences Division

points. This approach ensures that the analysis remains resilient even in the presence of incomplete or erroneous information. This methodology is highly suitable for intricate datasets, accurately identifying outliers.

Moreover, it offers flexibility in adjusting to evolving data patterns.

Results and Discussion

Decisional Table

The data is organised into conditional and decisional attributes in FRST and is referred to as a decisional table. In the current data, 'Temp' is the only decisional attribute, whereas the remaining eight variables are conditional attributes. We used R software to reduce the FRST attribute. The data was initially normalised, and then we used R's 'RoughSetPackage' to generate a set of reducts (a reduced collection of attributes).

As previously noted, climate data is collected from various sources, including satellite observations, ground-based measurements, and climate models. Climate change is a global phenomenon with interconnected systems. So, it is observed that the model needs help to capture the broader impacts and interactions. We obtained a total of 25 reducts. The performance of reducts is given in Table 2. The performance of each reduct is assessed using a Random Forest model, employing a cross-validation testing technique with ten folds and three repetitions. This combination comprehensively evaluates the model's accuracy and ability to generalize. Table 3 gives Reducts and their performance.

Interpretation from the FRST Reducts

RMSE, Rsquared, and MAE are metrics used to check the accuracy of the reducts for prediction. RMSE penalizes more significant errors to reveal the model's prediction accuracy. R squared measures the model's predictive ability by measuring the proportion of variance explained. MAE is easy to read and measures average error prediction; these measures would be combined to evaluate model performance.

• High performing reducts

Based on their frequent inclusion in top reducts and agreement with the scientific understanding of climate change factors, CO_2, N_2O, CFC_12, TSI, MEI, and Aerosols are the essential criteria for temperature prediction.

• Low performing reducts

CH_4 and CFC.11 attributes perform poorly, implying that they alone cannot explain temperature variance and that additional anthropogenic causes must be considered. These aspects are less critical in temperature modeling but should be addressed in complete climate assessments.

• Overall Conclusion for the performance of reducts

The original data's RMSE of 0.0697, R-squared of 0.8561, and MAE of 0.0531 provide a solid temperature prediction baseline. However, the reducts generated by FRST, with fewer features, also performed comparable and, in some cases, better accuracy metrics than the original dataset for temperature prediction. Let us compare the metric performance given in Table 3.

The reducts obtained by FRST exhibit a high level of accuracy, comparable to the complete dataset. This indicates that these selected features effectively represent the crucial factors influencing temperature

Table 2: Reducts and their performance

Reduct	*RMSE*	*Rsquared*	*MAE*
CO_2, CFC.12, TSI	0.07297839	0.8385665	0.05617610
CFC.11, TSI	0.1017114	0.6914656	0.07805482
MEI, CO_2, N_2O, CFC.12	0.07402929	0.8337500	0.05718342
MEI, CO_2, N_2O, CFC.12, TSI, Aerosols	0.07072951	0.8496601	0.05438029
N_2O, TSI	0.08201193	0.7945125	0.06340683
N_2O, CFC.11	0.07483289	0.8301461	0.05746496
MEI, CO_2, N_2O, CFC.12, TSI, Aerosols	0.09669404	0.7158756	0.07461719
CH_4, TSI	0.0959312	0.7167543	0.07375469
CH_4, CFC.12	0.08075624	0.8002665	0.06284183
CH_4, CFC.11	0.08342491	0.7871987	0.06374292
CH_4, N_2O	0.08545164	0.7797217	0.06536471
CO_2, Aerosol	0.06869305	0.7982172	0.05767735
CO_2, TSI	0.01010565	0.7900788	0.06850357
CO_2, CFC12	0.07487109	0.8327623	0.05766616
CO_2, N_2O	0.05268504	0.8136582	0.0531278
CO_2, CH_4	0.1094267	0.634997	0.08648479
CO_2, CFC11	0.0784023	0.8125331	0.06036547
MEI, CFC.11	0.1011338	0.691984	0.0782230
MEI, CH_4	0.1165169	0.5820787	0.08816766
MEI , CO_2, N_2O, CFC_12, TSI, Aerosols	0.07876883	0.8135444	0.05929967
MEI, N_2O	0.07286242	0.8392453	0.05586151
MEI, CFC.12	0.079821	0.8102288	0.06082631
MEI, TSI	0.1814416	0.0964548	0.1449752
MEI, Aerosols	0.09638725	0.7186111	0.07241071
MEI , CO_2, CH_4, TSI , Aerosols	0.08077717	0.8053545	0.06187857

Table 3: Performance of original data and high-performing reducts

	RMSE	*R-Squared*	*MAE*
Original data	0.0697	0.8561	0.0531
MEI, CO_2, N_2O, CFC.12, TSI, Aerosols	0.070729	0.8496	0.05438

changes. This suggests that by concentrating on these essential elements, targeted interventions and analysis can be employed to attain comparable predicted results, hence improving the efficiency and interpretability of the model.

Interpretation from Regression Analysis

The regression model will help to find the relationship between temperature change and all other factors. Here, temperature change will be considered as the dependent variable. All other factors, i.e. (MEI), Carbon Dioxide (CO_2), Methane (CH_4), Nitrous Oxide (N_2O), Chlorofluorocarbons (CFC.11 and CFC.12), Total Solar Irradiance (TSI), and Aerosols will be treated as dependent variables. We used R-programming to get the results for the regression model. Data is pre-processed using R to deal with outliers and multi-co-linearity. The statistical summary of the model is given in Table 4

Table 4: Results from regression model

Variable	*Estimate*	*Std. Error*	*t value*	*P-value*
(Intercept)	-127.7	19.19	-6.654	1.36E-10
MEI	0.06632	0.006186	10.722	< 2e-16
CO_2	0.005207	0.002192	2.375	0.0182
CH_4	6.371E-05	0.000498	0.128	0.8982
N_2O	-0.01693	0.007835	-2.161	0.0315
CFC.11	-0.007278	0.001461	-4.98	1.07E-06
CFC.12	0.004272	0.000876	4.875	1.77E-06
TSI	0.09586	0.01401	6.844	4.38E-11
Aerosols	-1.582	0.2099	-7.535	5.86E-13

The model clarifies the intricate interactions that cause temperature changes, making it useful for climate change research and policymaking. According to the regression model, MEI, CO_2, TSI, and aerosols significantly impact temperature. The unexpected N_2O results suggest more research into its role and interactions in the climate system.

Comparisons of Results from FRST Reducts and Regression

The regression model and Fuzzy Rough Set Theory (FRST) analysis reveal key factors influencing temperature variations. The regression model performs well, with an R-squared of 0.744, explaining 74.4% of temperature variance. This model identifies MEI, CO_2, CFC.11, CFC.12, TSI, and Aerosols as significant contributions, as expected. The regression also shows surprising results, such as the negative coefficient for N_2O, suggesting complicated interactions or confounding factors not captured by the model.

However, FRST reducts are also accurate, with numerous combinations approaching the original dataset. The best features MEI, CO_2, N_2O, CFC.12, TSI, and Aerosols have an R-squared of 0.8496, almost matching the original data's 0.8561. These features capture the most temperature variance, validating the regression model.

Both techniques consistently identify CO_2 as an essential component, supporting their global warming functions. CFC_12 and TSI are also substantial, showing anthropogenic emissions and natural fluctuation. Although CH_4 is less successful in the regression model, it occurs in effective reducts in the FRST analysis, demonstrating its importance in certain settings. MEI's consistency among reducts underscores its importance in recording ENSO-related temperature fluctuations.

Conclusion

This work by FRST introduces a powerful approach for selecting features that can handle non-linear interactions and the complexity of climatic variables more effectively, resulting in more accurate and realistic conclusions. FRST's high-performing reducts effectively capture data variance comparable to the regression model, thereby accurately capturing real-world climate dynamics. FRST may provide a more accurate representation of reality in complex, uncertain settings such as climate science. Both FRST and Regression have their advantages and downsides. FRST is more computationally intensive, whereas regression is more straightforward to implement. On the other hand, FRST is more adept at handling outliers than regression. In the future, a hybrid technique can be employed, utilising FRST to identify crucial features and regression analysis to determine the weightage of each feature.

Table 5: Performance of the regression model

Metric	*Value*	*Interpretation*
Residual Standard Error	0.09182	RSE 0.09182 is low which indicates a good fit between the model's predictions and observed data.
Multiple R-squared	0.744	74.4% prediction accuracy indicates a strong correlation between independent and dependent variables.
Adjusted R-squared	0.7371	The adjusted R-squared provides a more precise estimate of model performance when several predictors exist. When predictors are used, 73.71% of the variation is explained.
F-statistic	108.6	The F-statistic assesses regression model significance. Here, a high F-statistic indicates a better data fit than a model without independent variables.
P-value (F-statistic)	< 2.2e-16	A p-value of < 0.05 shows a significant regression model, indicating that the independent factors collectively affect the dependent variable (temperature).

References

Acharjya, D.P. and Rathi, R., 2022. An integrated fuzzy rough set and real coded genetic algorithm approach for crop identification in smart agriculture. *Multimedia Tools and Applications*, **81(24):** 35117-35142. https://doi.org/10.1007/s11042-021-10518-7

Adnan, A., Yolanda, A.M., Erda, G., Goldameir, N.E. and Indra, Z., 2023. The comparison of accuracy on classification climate change data with logistic regression. *Sinkron*, **8(1):** 56-61. https://doi.org/10.33395/sinkron.v8i1.11914

Anaraki, M.V., Farzin, S., Mousavi, S.F. and Karami, H., 2021. Uncertainty analysis of climate change impacts on flood frequency by using hybrid machine learning methods. *Water Resources Management*, **35(1):** 199-223. https://doi.org/10.1007/s11269-020-02719-w

Bhatt, R.B. and Gopal, M., 2005. On fuzzy-rough sets approach to feature selection. *Pattern Recognition Letters*, **26(7):** 965-975. https://doi.org/10.1016/j.patrec.2004.09.044

Bône, C., Gastineau, G., Thiria, S., Gallinari, P. and Mejia, C., 2023. Detection and attribution of climate change using a neural network. *Journal of Advances in Modeling Earth Systems*, **15(10):** e2022MS003475. https://doi.org/10.1029/2022MS003475

Chaikumbung, M., 2023. The effects of institutions and cultures on people's willingness to pay for climate change policies: A meta-regression analysis. *Energy Policy*, **177:**113513. https://doi.org/10.1016/j.enpol.2023.113513

Charlson, F., Ali, S., Augustinavicius, J., Benmarhnia, T., Birch, S., Clayton, S., Fielding, K., Jones, L., Juma, D., Snider, L., Ugo, V., Zeitz, L., Jayawardana, D., La Nauze, A. and Massazza, A., 2022. Global priorities for climate change and mental health research. *Environment International*, **158:** 106984. https://doi.org/10.1016/j.envint.2021.106984

Chen, H.L., Yang, B., Liu, J. and Liu, D.Y., 2011. A support vector machine classifier with rough set-based feature selection for breast cancer diagnosis. *Expert Systems with Applications*, **38(7):** 9014-9022. https://doi.org/10.1016/j.eswa.2011.01.120

Cummins, D.P., Stephenson, D.B. and Stott, P.A., 2022. Could detection and attribution of climate change trends be spurious regression? *Climate Dynamics*, **59(9-10):** 2785-2799. https://doi.org/10.1007/s00382-022-06242-z

Doan, Q. Van, Amagasa, T., Pham, T.H., Sato, T., Chen, F. and Kusaka, H., 2023. Structural k-means (S k-means) and clustering uncertainty evaluation framework (CUEF) for mining climate data. *Geoscientific Model Development*, **16(8):** 2215-2233. https://doi.org/10.5194/gmd-16-2215-2023

Dormann, C.F., Elith, J., Bacher, S., Buchmann, C., Carl, G., Carré, G., Marquéz, J. R. G., Gruber, B., Lafourcade, B., Leitão, P.J., Münkemüller, T., et al., 2009. Collinearity: A review of methods to deal with it and a simulation study evaluating their performance. *Ecography*, **36(1)**. https://doi.org/10.1111/j.1600-0587.2012.07348.x

Dubois, D. and Prade, H., 1990. Rough fuzzy sets and fuzzy rough sets. *International Journal of General Systems*, **17(2-3):** 191–209. https://doi.org/10.1080/03081079008935107

Ewees, A.A., Elaziz, M.A., Arafa, R.M. and Ghoniem, R.M., 2020. Improved approach based on fuzzy rough set and sine-cosine algorithm: A case study on prediction of osteoporosis. *IEEE Access*, **8:** 203190–203202. https://doi.org/10.1109/ACCESS.2020.3036831

Gupta, V. and Jain, M.K., 2018. Investigation of multi-model spatiotemporal mesoscale drought projections over India under climate change scenario. *Journal of Hydrology*, **567**. https://doi.org/10.1016/j.jhydrol.2018.10.012

Hassani, H., Huang, X. and Silva, A.E., 2019. Big data and climate change. *In:* Big Data and Cognitive Computing, **3(1)**. https://doi.org/10.3390/bdcc3010012

Hawkins, D. M. (2004). The Problem of Overfitting. *In: Journal of Chemical Information and Computer Sciences*, **44(1)**. https://doi.org/10.1021/ci0342472

Huang, K.Y. and Jane, C.J., 2009. A hybrid model for stock market forecasting and portfolio selection based on ARX, grey system and RS theories. *Expert Systems with Applications*, **36(3 PART 1)**. https://doi.org/10.1016/j.eswa.2008.06.103

Jensen, R. and Shen, Q., 2009. New approaches to fuzzy-rough feature selection. *IEEE Transactions on Fuzzy Systems*, **17(4)**. https://doi.org/10.1109/TFUZZ.2008.924209

Khan, N., Shahid, S., Ismail, Bin, T. and Behlil, F., 2021. Prediction of heat waves over Pakistan using support vector machine algorithm in the context of climate change. *Stochastic Environmental Research and Risk Assessment*, **35(7)**. https://doi.org/10.1007/s00477-020-01963-1

Kumar, S. (2023). A novel hybrid machine learning model for prediction of CO_2 using socio-economic and energy attributes for climate change monitoring and mitigation policies. *Ecological Informatics*, *77*. https://doi.org/10.1016/j.ecoinf.2023.102253

Lasisi, A., Ghazali, R., Deris, M. M., Herawan, T. and Lasisi, F., 2016. Extracting information in agricultural data using fuzzy-rough sets hybridization and clonal selection theory inspired algorithms. *International Journal of Pattern Recognition and Artificial Intelligence*, **30(9)**. https://doi.org/10.1142/S0218001416600089

Lindsey, R. and Dahlman, L., 2021. Climate change: global temperature projections. National Oceanic and Atmospheric Administration (NOAA).

Malhi, G.S., Kaur, M. and Kaushik, P., 2021. Impact of climate change on agriculture and its mitigation strategies: A review. *In:* Sustainability (Switzerland), **13(3)**. https://doi.org/10.3390/su13031318

Malhi, Y., Franklin, J., Seddon, N., Solan, M., Turner, M.G., Field, C.B. and Knowlton, N., 2020. Climate change and ecosystems: Threats, opportunities and solutions.

In: Philosophical Transactions of the Royal Society B: Biological Sciences (Vol. 375, Issue 1794). https://doi.org/10.1098/rstb.2019.0104

Maqsood, J., Farooque, A. A., Abbas, F., Esau, T., Wang, X., Acharya, B. and Afzaal, H., 2022. Application of artificial neural networks to project reference evapotranspiration under climate change scenarios. *Water Resources Management*, **36(3)**. https://doi.org/10.1007/s11269-021-02997-y

Marland, G., Pielke, R.A., Apps, M., Avissar, R., Betts, R.A., Davis, K.J., Frumhoff, P.C., Jackson, S.T., et al., 2003. The climatic impacts of land surface change and carbon management, and the implications for climate-change mitigation policy. *Climate Policy*, **3(2):** 149-157. https://doi.org/10.1016/S1469-3062(03)00028-7

Pamucar, D., Torkayesh, A.E. and Biswas, S., 2023. Supplier selection in healthcare supply chain management during the COVID-19 pandemic: A novel fuzzy rough decision-making approach. *Annals of Operations Research*, **328(1).** https://doi.org/10.1007/s10479-022-04529-2

Papadopoulos, T. and Balta, M.E., 2022. Climate change and big data analytics: Challenges and opportunities. *International Journal of Information Management*, **63:** 102448. https://doi.org/10.1016/j.ijinfomgt.2021.102448

Rolnick, D., Donti, P.L., Kaack, L.H., Kochanski, K., Lacoste, A., et al., 2023. Tackling climate change with machine learning. *In:* ACM Computing Surveys, **55(2):** 1-96. https://doi.org/10.1145/3485128

Rosenberg, C.. 2010. Intergovernmental Panel on Climate Change. *In:* Encyclopedia of the U.S. Government and the Environment: History, Policy, and Politics [2 volumes] (Vol. 1). https://doi.org/10.4324/9781315270326-109

Sadeghi, M., Naghedi, R., Behzadian, K., Shamshirgaran, A., Tabrizi, M.R. and Maknoon, R., 2022. Customisation of green buildings assessment tools based on climatic zoning and experts judgement using K-means clustering and fuzzy AHP. *Building and Environment*, **223:** 109473. https://doi.org/10.1016/j.buildenv.2022.109473

Thorne, P.W., Willett, K.M., Allan, R.J., Bojinskski, S., et al., 2011. Guiding the creation of a comprehensive surface temperature resource for twenty-first-century climate science. *Bulletin of the American Meteorological Society*, **92(11):** ES40-ES47. https://doi.org/10.1175/2011BAMS3124.1

Tripathi, S., Srinivas, V.V. and Nanjundiah, R.S., 2006. Downscaling of precipitation for climate change scenarios: A support vector machine approach. *Journal of Hydrology*, **330(3-4):** 621-640. https://doi.org/10.1016/j.jhydrol.2006.04.030

Tsang, E.C.C., Chen, D., Yeung, D.S., Wang, X.Z. and Lee, J.W.T., 2008. Attributes reduction using fuzzy rough sets. *IEEE Transactions on Fuzzy Systems*, **16(5)**. https://doi.org/10.1109/TFUZZ.2006.889960

Valentová, A. and Bostik, V., 2021. Climate change and human health. *In:* Military Medical Science Letters (Vojenske Zdravotnicke Listy), **90(2)**. https://doi.org/10.31482/mmsl.2021.010

Wade, K. and Jennings, M., 2016. The impact of climate change on the global economy. *Schroders*.

Yamaka, W., Phadkantha, R. and Rakpho, P., 2021. Economic and energy impacts on greenhouse gas emissions: A case study of China and the USA. *Energy Reports*, **7:** 240-247. https://doi.org/10.1016/j.egyr.2021.06.040

Yasodha, P. and Ananthanarayanan, N.R., 2015. Analysing big data to build knowledge based system for early detection of ovarian cancer. *Indian Journal of Science and Technology*, **8(14)**. https://doi.org/10.17485/ijst/2015/v8i14/65745

Zadeh, L.A., 1965. Fuzzy sets. *In:* Information and Control. **8:** 338-33.

Zhao, S., Tsang, E.C.C., Chen, D. and Wang, X., 2010. Building a rule-based classifiera fuzzy-rough set approach. *IEEE Transactions on Knowledge and Data Engineering*, **22(5):** 624-638. https://doi.org/10.1109/TKDE.2009.118