J. Innovative Eng. Nat. Sci., vol.6, no. 1, pp. 1-15, 2026.
http://doi.org/10.61112/jiens.1724247

JIENS
Journal of Innovative Engineering and Natural Science
(Yenilikçi Mühendislik ve Doğa Bilimleri Dergisi)
https://dergipark.org.tr/en/pub/jiens
Original Paper

Machine learning-enabled classification of global human development using INFORM risk indicators

Merve Doğruel[a,*]

[a]*Management Information Systems, Faculty of Business and Management Sciences, Istanbul Esenyurt University, 34513, İstanbul, Türkiye.*

ARTICLE INFO

Article history:
Received June 21, 2025
Received in revised form August 8, 2025
Accepted February August 17, 2025

Available online

Keywords:

Machine learning,
Bagging,
Random forest,
Human development index,
INFORM risk

ABSTRACT

This study aims to identify the most effective machine learning model for classifying countries' Human Development Index (HDI) levels using indicators from the INFORM Risk Index. The motivation for this work lies in the growing need for data-driven methods to analyze and predict human development outcomes, particularly in the context of complex and high-dimensional socio-economic and disaster-related risk data. Traditional models often fail to capture the non-linear relationships that influence human development. To address this gap, six supervised machine learning algorithms—k-Nearest Neighbors (KNN), Linear and Nonlinear Support Vector Machines (SVM), Classification and Regression Trees (CART), Bagging, and Random Forest (RF)—were systematically evaluated. Performance was measured using weighted F1-scores on both training and testing datasets. The results reveal that while KNN, Linear SVM, and CART have limited predictive power, the Nonlinear SVM suffers from overfitting. In contrast, ensemble-based models—Bagging and RF—demonstrate superior and balanced performance, with F1-scores around 0.80 on both datasets. These methods also allow for interpretability through feature importance analysis. Socio-economic, institutional, and infrastructure-related indicators were identified as the most influential variables in predicting HDI levels. The findings highlight the strength of ensemble learning in modeling complex development-related risks and provide a robust framework for integrating machine learning into global human development analysis. This study offers valuable insights for policymakers and researchers aiming to improve forecasting, resilience planning, and development strategies.

I. INTRODUCTION AND LITERATURE REVIEW

Human development has been defined as the process of expanding people's choices. The most important concepts in this process are defined as a healthy and long life, level of education and having a good standard of living [1].

When Figure 1 is examined, it is seen that the human development index is closely related to concepts such as sustainable development, economic growth, education, income, renewable energy.

Casau et al. [2] conducted a study that evaluates the intricate relationships between environmental sustainability, welfare, and economic output. The study critiques the excessive reliance on GDP as a conventional development indicator and emphasizes the necessity of more comprehensive measures. Research conducted on alternative indicators, including the Human Development Index (HDI), Planetary Pressures Adjusted HDI (PHDI), Sustainable Development Goals (SDG) Index, and Happy Planet Index (HPI), demonstrates that economic growth does not always correspond to environmental sustainability or social welfare. The results of the study indicate that there are positive correlations between GDP and environmental degradation indicators, which underscores the necessity of sustainable economic models that enhance human welfare within environmental constraints. The study contends that policymakers should create fiscal policies that foster sustainable development and incorporate

*Corresponding author. Tel.: +90- 212-699-0990; e-mail: mervedogruel@esenyurt.edu.tr

environmental and social welfare indicators into the System of National Accounts through quantitative and qualitative analyses.

Improving fundamental aspects of human development—such as reducing poverty, expanding access to education, ensuring adequate housing, strengthening food systems, and promoting social protection—is essential to decreasing the susceptibility of individuals and communities to disaster impacts. Vulnerability to disasters arises from the interplay between physical exposure and socio-economic conditions, where underdevelopment often serves to intensify both. Inadequate development limits adaptive capacity and increases the likelihood of severe losses in the face of natural hazards. As outlined by the United Nations Office for Disaster Risk Reduction [3], disaster resilience cannot be achieved without addressing structural development gaps that shape risk exposure and coping mechanisms. Historically, risk management frameworks have paid insufficient attention to these underlying dimensions, weakening their effectiveness in protecting vulnerable populations.

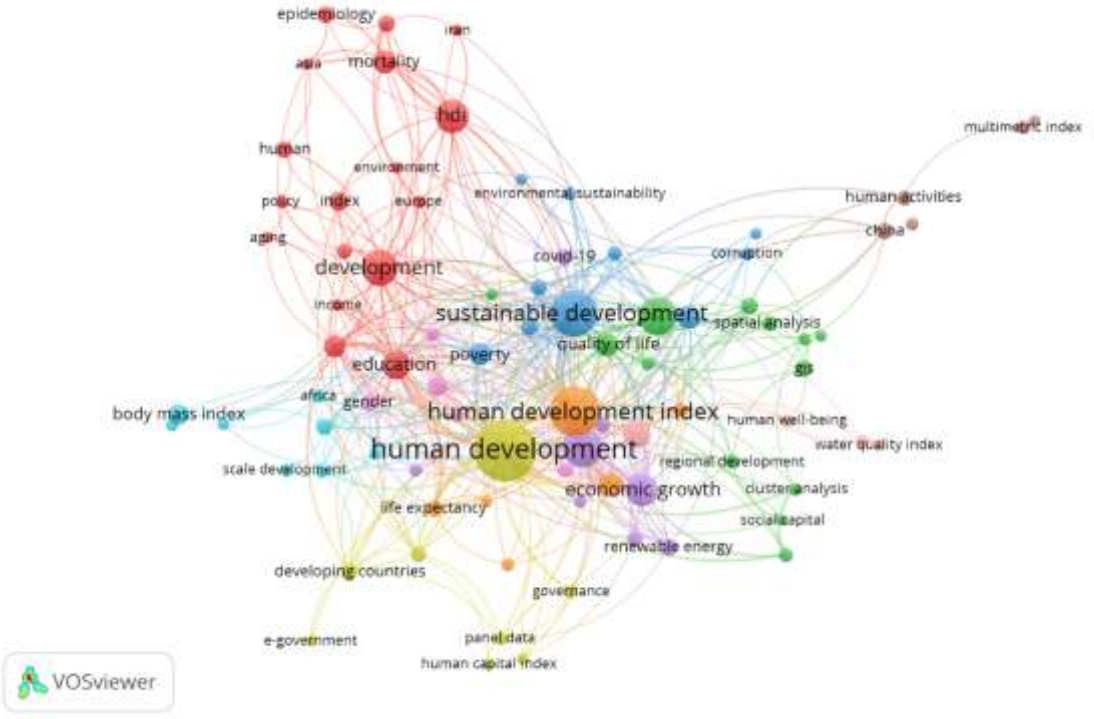

Figure 1: Keywords associated with human development index

Scholars and international organizations have argued that integrating human development processes into disaster risk reduction can reduce future disaster impacts. However, it is still unclear how to integrate disaster risk reduction with human development [4].

Oran [5] examined the disaster risk management strategies of 173 countries by utilizing data from the OECD's Human Development Index (HDI) and the INFORM Risk Index. The analysis demonstrates that disaster risk classes are correlated with the development levels of countries. It has been found that health conditions, inequality, hazards and exposure risk, and the likelihood of natural disasters are some of the factors that affect the level of

development. Consequently, it is underscored that the development levels of countries are a significant factor in the determination of disaster risk classes.

Feldmeyer et al. [6] compared the INFORM Risk Index and the World Risk Index. Both indices provide indicators that measure human vulnerability to climate change and disasters. The study analyzes the similarities and differences of these indices, revealing which regions have high human vulnerability and how development levels in these regions are related to disaster risks.

A study by Eze and Siegmund [7] applied random forest regression, spatial stratified heterogeneity, and hotspot analysis on INFORM data to examine disaster risk patterns in Africa. The results highlight increasing risk in Eastern, Southern, and Western regions, driven by exposure to floods, epidemics, and conflict, as well as by vulnerability and limited coping capacity. The study emphasizes the need for sustainability-oriented policies to reduce disaster risk.

The research conducted by Raikes et al. [4] examines the perspectives of government practitioners regarding the integration of human development and disaster risk reduction in the context of droughts and floods in Australia and Canada. This article employs a comparative case study approach, drawing on two Delphi studies and semi-structured interviews conducted with practitioners across local, provincial/state, and federal levels in Canada and Australia. The findings underscore a shared perception among participants that effective disaster risk reduction (DRR) necessitates deeper engagement with human development systems that are adaptable to local conditions. This includes the systematic integration of disaster risk data into development planning and implementation, as well as critical consideration of interconnected issues such as poverty, public health, climate resilience, social justice, equity, and human agency.

In their study, Mochizuki and Naqvi [8] aim to reformulate the Human Development Index (HDI) in a way that reflects disaster risks. It recalculates the HDI as the "Risk-Adjusted Human Development Index" (RHDI) by taking into account the economic losses of disasters. Disasters can directly affect development indicators such as health, education and infrastructure; therefore, it is argued that risks related to these effects should be included in the index.

An examination of the literature reveals a correlation between informal risk and human development. The main goal of this study is to use informatics risk indicators as predictors to rank countries by their level of human development. The most effective machine learning algorithm will be identified through comparative testing.

To address the classification of countries according to human development categories based on INFORM Risk Index indicators, this study applies a set of supervised machine learning algorithms, including K-Nearest Neighbors (KNN), Linear and Non-linear Support Vector Machines (SVM), Classification and Regression Trees (CART), Bagging, and Random Forest (RF). These models undergo training and validation using stratified k-fold cross-validation to ensure robustness and reduce sampling bias. Hyperparameters are optimized through grid search to enhance classification accuracy. The weighted F1-score serves as the primary performance metric, allowing for reliable evaluation in the presence of class imbalance. This methodological framework enables a systematic comparison of model effectiveness in mapping INFORM-based risk indicators to human development outcomes.

II. INDEXES

The Human Development Index and the INFORM Risk Index, which are the two indices employed, will be elaborated upon in this section.

2.1. INFORM Risk Index

INFORM is a forum for collaborating on quantitative assessments relevant to humanitarian crises and disasters. The Joint Research Centre of the European Commission is the scientific and technical lead for Inform, and the platform includes organisations from across the humanitarian and development sector, as well as donors and technical partners. INFORM Risk, INFORm Warning, INFORM Climate Change and INFORM Severity are INFORM products. In 2014, INFORM Partners launched an open index to assess the risk of humanitarian crises globally, and 2024 marks the tenth anniversary of Inform Risk. For 10 years, the INFORM Risk Index and other INFORM products have been a key component in the decision-making systems of many organizations worldwide and locally. Some of these organizations; World Food Programme, United Nations Office for the Coordination of Humanitarian Affairs (OCHA), European Commission Directorate-General for European Civil Protection and Humanitarian Aid Operations (DG ECHO), International Federation of Red Cross And Red Crescent Societies (IFRC), World Health Organization (WHO) [9]. The risk components encompassed by the INFORM risk index are depicted in Table 1.

Table 1. INFORM risk index

Dimensions					
Hazard and exposure		*Vulnerability*		*Lack of coping capacity*	
Categories					
Natural	*Human*	*Socio-economic*	*Vulnerable groups*	*Institutional*	*Infrastructure*
Earthquake	Current conflict intensity	Development and deprivation (50%)	Uprooted people	DRR	Communication
Tsunami	Projected conflict risk	Inequality (25%)	Other vulnerable groups	Governance	Physical infrastructure
River flood		Aid dependency (25%)			Access to health system
Coastal flood					
Tropical cyclone wind					
Drought					
Epidemic					

The INFORM risk index can be used to prioritize countries based on risk or any of its components, to reduce risk in the most accurate way, to monitor risk trends, in national and regional risk assessment, or in index adaptation studies for organizations or regions. In the 2024 report, a risk index has been calculated for 191 countries.

2.2. Human Development Index (HDI)

Human development encompasses more than merely enhancing the prosperity of the economy in which individuals reside, and it adopts a approach that aims to enhance the prosperity of human existence.

The gross domestic product is an inadequate indicator of social achievements. Gross domestic product takes into account material well-being, but does not provide information on how a countries wealth is transformed into basic needs [10]. The human development index considers not only the economic development dimension, but also the long and healthy life and education dimensions. The human development report was first published in 1990.

The dimensions and categories used in the human development index calculation for the 2024 report are shown in Table 2 [11].

Table 2: Human development index

Dimensions			
Long and health life	*Knowledge*		*A decent standart of living*
Categories			
Life expectancy at birth	Expected years of schooling	Mean years of schooling	GNI per capita (PPP $)

The human development index is calculated for 193 countries by 2024. The values for health, education, and income dimensions are standardized within a range of 0 to 1. Subsequently, the geometric meaning of these three dimensions is calculated. The value obtained is used to calculate the human development index for each country.

III. MACHINE LEARNING ALGORITHMS

In recent years, the increasing volume and complexity of data across various domains highlight the critical role of machine learning (ML) and advanced data analytics in extracting actionable insights and enabling data-driven decision making. The integration of sophisticated ML algorithms with robust data analysis frameworks revolutionizes predictive modeling, pattern recognition, and classification tasks, driving innovation in both academic research and industry applications. As data becomes more high-dimensional and heterogeneous, the demand for accurate, interpretable, and scalable classification techniques grows correspondingly. In this study, supervised classification methods such as KNN, SVM, CART, RF are focused on due to their complementary strengths in handling diverse data structures and their proven effectiveness across various classification challenges.

3.1. K-Nearest Neighbors (KNN)

The K-Nearest Neighbors (KNN) algorithm is a non-parametric, instance-based supervised learning method commonly applied to both classification and regression tasks. It determines the output for a query instance by referring to the labels (in classification) or values (in regression) of its ‘k’ nearest neighbors in the feature space, typically according to Euclidean or Manhattan distance. Its simplicity, interpretability, and flexibility-especially in modeling non-linear relationships—make KNN a widely used baseline in various domains [12].

Despite its advantages, KNN faces several notable challenges that can limit its effectiveness in practice. First, the algorithm incurs a high computational cost during inference, as it requires calculating distances between the query instance and all samples in the training dataset. This drawback becomes particularly pronounced with large-scale datasets. Second, KNN is sensitive to hyperparameters such as the choice of the number of neighbors (k), the selected distance metric, and the scaling of features, all of which significantly impact its predictive performance. Lastly, KNN’s efficacy deteriorates in high-dimensional spaces due to the “curse of dimensionality,” where the meaningfulness of distance metrics diminishes and the distinction between nearest and farthest neighbors becomes less clear, resulting in reduced classification or regression accuracy [13]. Recent literature has focused on enhancing KNN’s scalability, robustness, and applicability:

Random Kernel KNN (RK-KNN) introduces bootstrap sampling and kernel smoothing to reduce RMSE and improve generalization on large and complex datasets [13].

A comprehensive review by Halder et al. [12] catalogs modifications such as approximate search algorithms, random projection ensembles, fuzzy approaches, and power-mean variants—each designed to improve performance in high-dimensional and big data contexts. Ali et al. [14] propose a random projection ensemble (RPExNRule), which combines bootstrap sampling with low-dimensional projections, markedly improving classification stability and accuracy. These developments demonstrate that while KNN's core methodology remains unchanged, recent techniques effectively address its scalability and sensitivity limitations, making it a viable choice for large-scale data applications, fault detection, bioinformatics, and real-time analytics.

3.2. Support Vector Machine (SVM)

Support Vector Machine (SVM) is a supervised learning algorithm renowned for its effectiveness in both classification and regression tasks. It operates by finding the optimal hyperplane that maximizes the margin between support vectors—critical data points near decision boundaries—thus enhancing generalization performance. This principle, grounded in statistical learning theory and structural risk minimization, enables SVMs to achieve robust predictive accuracy in high-dimensional feature spaces [15]. In addition to its advantage of being effective in high-dimensional spaces and situations where the number of dimensions exceeds the number of samples, SVM also enhances memory efficiency by utilizing a subset of training points in the decision function, thereby maintaining strong performance [16].

For non-linearly separable data, SVMs utilize the kernel trick by transforming inputs into higher-dimensional spaces, allowing linear separation where it was not possible before. Common kernels include RBF (Radial Basis Function), polynomial, and sigmoid functions. Recent advancements, such as a novel distance-based kernel, have demonstrated significant gains in classification accuracy across diverse datasets, outperforming traditional kernels [17].

Although classical SVM is not inherently designed for big data or streaming scenarios, recent research has introduced scalable adaptations, including sample reduction techniques, parallel/distributed implementations, and online learning frameworks, to extend SVM applicability to large-scale datasets [18].

Other contemporary enhancements include robust SVM optimization, which integrates uncertainty into the optimization process to improve resilience against noisy or incomplete data, and specialized variants like p-SVM, which generalizes hinge-loss norms via p-norms to improve multiclass classification contracts and performance bounds [19].

3.3. Classification and Regression Trees (CART)

The Classification and Regression Tree (CART) algorithm, originally introduced by Breiman et al. [20], is a foundational non-parametric and non-linear supervised learning method widely employed for both classification and regression tasks. CART constructs binary decision trees by recursively splitting the dataset into subsets based on thresholds of a single predictor variable that best separates the data. For classification problems, these splits aim to maximize node purity using impurity measures such as the Gini index or cross-entropy, while regression tasks minimize variance within nodes. Each division is carried out in a rule-based, binary recursive manner, where variables can be reused across different branches of the tree depending on their predictive contribution. Owing to

its intuitive structure, interpretability, and robustness in handling both categorical and continuous variables, CART continues to be a widely adopted method across diverse application domains [21].

In recent years, ensemble learning techniques such as Random Forests and Gradient Boosted Trees have substantially elevated the predictive capacity of CART-based models. For instance, hybrid frameworks that integrate CART within metaheuristic-optimized ensembles—such as Genetic Algorithm-enhanced bagging or boosted trees—have demonstrated improved robustness and generalizability on high-dimensional and imbalanced datasets [22].

Furthermore, hyperparameter optimization has become a pivotal factor in enhancing model performance. Techniques such as Bayesian optimization are increasingly employed to systematically fine-tune critical parameters—such as max_depth, min_samples_leaf, and ccp_alpha—enabling models to achieve an optimal trade-off between predictive accuracy and overfitting [23].

3.4. Bagging

Bagging (Bootstrap Aggregating) is an ensemble learning technique designed to improve the stability and accuracy of machine learning models by reducing variance and mitigating overfitting, particularly in high-variance algorithms such as decision trees. Proposed by Breiman [24], Bagging involves generating multiple versions of a training dataset through bootstrap sampling and fitting a base learner to each. The predictions of these learners are then aggregated—typically by majority voting for classification tasks or averaging for regression—resulting in a more robust and generalizable model.

Recent developments in ensemble learning have focused on enhancing bootstrap sampling mechanisms, improving computational scalability, and adapting to challenges posed by imbalanced and high-dimensional data sets [25]. Moreover, Bagging has been advanced through the incorporation of dynamic ensemble selection strategies and online learning frameworks, enabling improved adaptability and performance in environments characterized by streaming and non-stationary data [26].

These developments affirm Bagging's continued relevance as a foundational ensemble method that enhances prediction accuracy, especially in domains where model stability and interpretability are essential. As ensemble learning continues to evolve, Bagging remains a benchmark for comparison and a building block for more complex hybrid models.

3.5. Random Forest (RF)

Random Forest is a widely used ensemble learning technique designed to improve the stability and predictive accuracy of decision trees, particularly in high-dimensional and noisy data environments. Initially introduced by Breiman [27], the method builds multiple decision trees using bootstrapped subsets of the original data and aggregates their outputs through majority voting (in classification tasks) or averaging (in regression tasks). While Random Forest shares foundational principles with bagging—most notably, the use of bootstrap aggregating to reduce variance—it introduces a critical enhancement: at each node of a decision tree, only a randomly selected subset of features is considered for splitting. This feature-level randomization increases diversity among the

individual base learners, thereby reducing correlation among trees and leading to improved generalization performance [27].

In contrast to traditional bagging, which constructs diverse learners solely by sampling data, Random Forest incorporates both data-level and feature-level randomness. This dual-randomization strategy has demonstrated superior performance across various tasks, particularly in contexts where overfitting is a concern [28]. Moreover, Random Forest is inherently suitable for handling missing values, estimating feature importance, and managing high-dimensional input spaces, making it a robust choice in domains such as bioinformatics, finance, and remote sensing.

Recent advancements in Random Forest methodologies have focused on improving computational efficiency and predictive performance through enhanced randomization techniques. In particular, Extremely Randomized Trees (Extra-Trees) introduce greater randomness by selecting split thresholds at random rather than searching for optimal splits, which accelerates training and reduces variance. This approach makes Random Forest models more scalable and better suited for large-scale datasets without compromising accuracy [29].

Additionally, efforts have been made to improve the algorithm's effectiveness on imbalanced datasets through cost-sensitive learning and class weighting techniques [25]. Furthermore, the integration of Random Forests with deep learning architectures and interpretable AI techniques—such as SHAP (SHapley Additive exPlanations)—has opened new avenues for enhancing both performance and explainability in complex predictive systems.

IV. APPLICATION

The objective of the study is to categorize the human development categories of nations based on the fundamental indicators of the Inform Risk Index by employing diverse machine learning algorithms. For this purpose, a comparison of the different machine learning algorithms will be made, and the algorithm or algorithms with the highest performance metrics will be selected and interpreted. The study's framework is depicted in Figure 2.

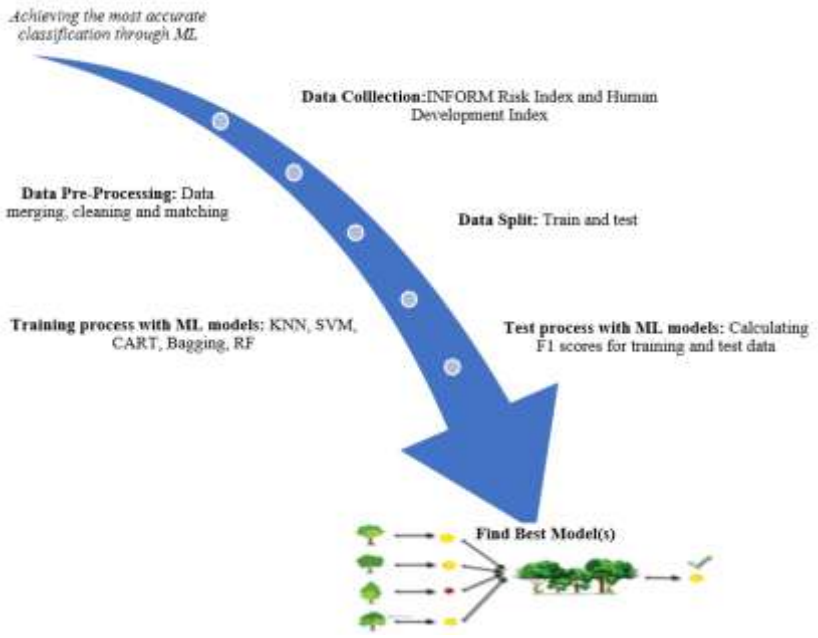

Figure 2: The study's framework

The data used for this study belong to the Inform Risk Index and HDI 2024 reports. When the common country data from the two reports is amalgamated, a dataset comprising 167 observations (countries) is obtained. The variable that each country is classified as "low", "medium", "high" and "very high" based on their IGE values has been selected as the target variable. In the classification of this target variable, the three principal dimensions of IRE, namely nature, human, socio-economic, vulnerable groups, institutional, and infrastructure characteristics, are utilized. Based on this information, it can be inferred that the dataset comprises of seven features and 167 observations. In the R program, the preparation of the data set and the application of machine learning algorithms are made. An overview of the data set used in the application is provided in Table 3.

Table 3. Dataset excerpt for illustration

No	CTRY	HDI_Category	IR_nat.	IR_human	IR_ses	IR_vuln_gro	IR_inst	IR_inf.
1	Afghanistan	4	5.8	10	8.1	6.6	7.4	6.7
2	Albania	2	5.7	0.1	2.4	2.1	5.6	2.2
3	Algeria	2	3.2	2.1	2.4	3.2	4.9	3.7
4	Angola	3	3	5	5.8	4.3	6.1	7.2
	.							
	.							
	.							
165	Yemen	4	4.4	8	8.1	8.6	8.8	6.7
166	Zambia	3	3.1	0.2	6.4	5.6	4.9	6.2
167	Zimbabwe	3	3.9	0.6	5.9	4.7	5.1	6.4

As per the fundamental principles of machine learning, 167 observation data shall be categorized into training and testing. The data set is randomly divided into 80% training and 20% test data, and in order to guarantee consistency while creating models, a 10-fold cross validation technique is employed during the training process.

Throughout the training process of the **KNN** algorithm, the k parameter is tuned over the range from 1 to 20. The best model is obtained when k is set to 14. In this configuration, the weighted F1-score is 0.7713 on the training data and 0.7082 on the test data.

In the **Linear SVM** model, the kernel parameter is set to linear, and the cost parameter is tuned over the values {0.001, 0.01, 0.1, 1, 5, 10, 100}. The best performance is obtained when the cost parameter is set to 1. The weighted F1-score of the optimal linear SVM model is 0.8781 on the training data and 0.7536 on the test data.

In the **Non-linear SVM** model, the kernel parameter is set to radial (RBF). The cost parameter is tuned over {0.01, 1, 10, 1000}, while the gamma parameter is tuned over {0.5, 1, 2, 3, 4, 5}. The optimal model is achieved with a cost parameter of 1 and a gamma value of 0.5. This best-performing non-linear SVM model yields a weighted F1-score of 0.9087 on the training data and 0.6823 on the test data.

For the **CART** algorithm, the minsplit parameter— which defines the minimum number of observations required for an internal node to be eligible for splitting— is tuned within the range of 5 to 7. Additionally, the complexity parameter (cp), which controls the overall complexity and pruning of the tree, is tuned over the values 0.03, 0.04, and 0.05. In the optimal model, the minsplit parameter is set to 5, and the cp is set to 0.03. The corresponding decision tree for this model is presented in Figure 3.

The weighted F1-score of the optimal CART model is calculated as 0.8577 for the training data, while the corresponding score for the test data is 0.7847.

In the **Bagging** model, the weighted F1-score is calculated as 1.000 for the training data, while it is 0.7954 for the test data. The plots illustrating the feature importance based on Mean Decrease Accuracy and Mean Decrease Gini are presented in Figure 4. According to the plot, the most influential variables in the prediction process appear to be socio-economic status (SES), infrastructure, and institutional features.

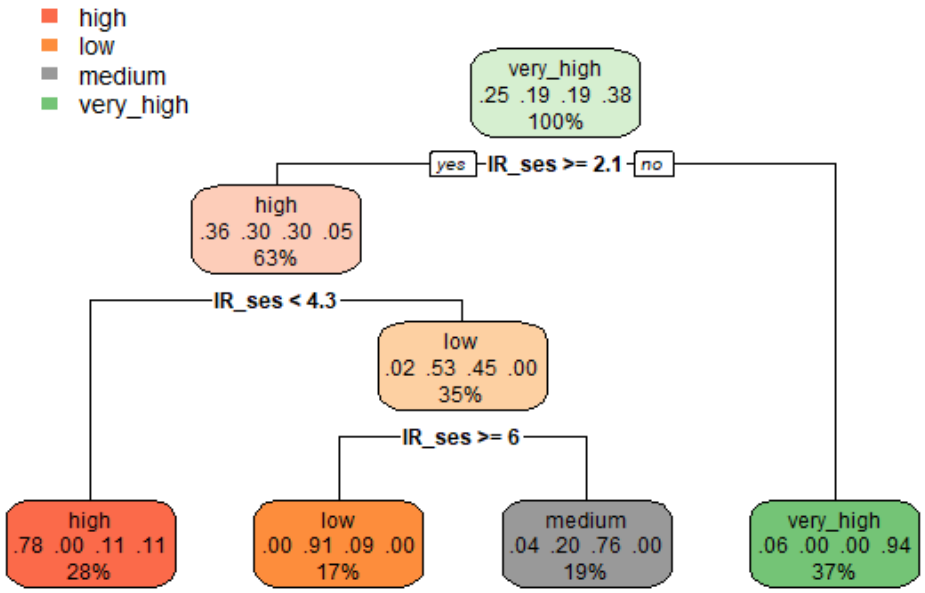

Figure 3: Tree with CART algorithm

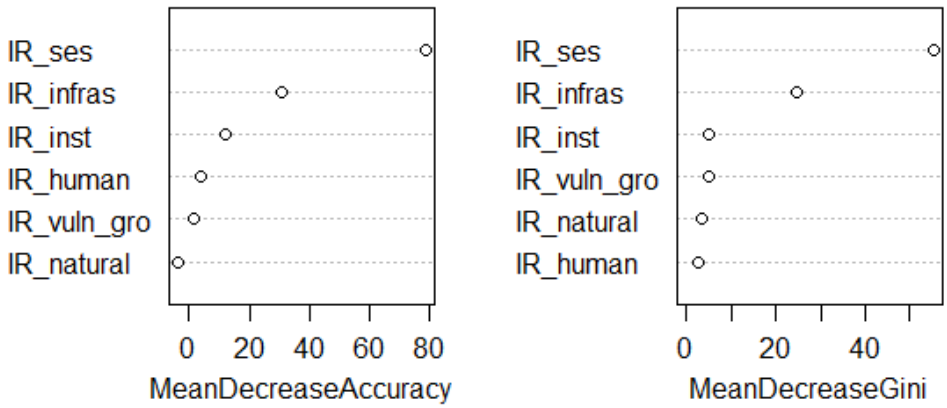

Figure 4: Features importance plot (Bagging)

In the **RF** algorithm, the mtry parameter— which determines the number of features randomly considered at each split in each tree— is tuned over the range of 2 to 6. As shown in Figure 5, the optimal value of the mtry parameter is determined to be 3.

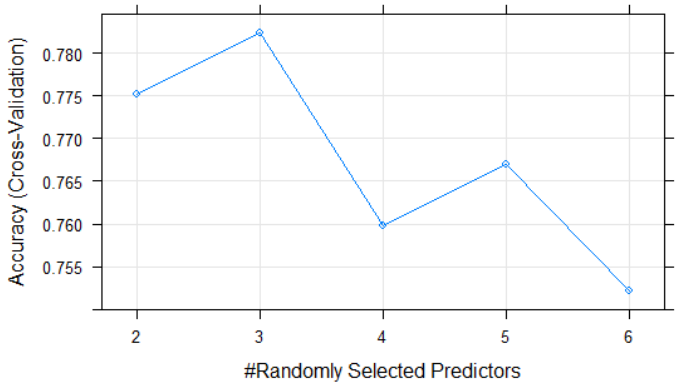

Figure 5: Randomly selected predictors

In the model constructed using three randomly selected features, the plots illustrating feature importance based on Mean Decrease Accuracy and Mean Decrease Gini are presented in Figure 6. Upon examination of Figure 6, it is observed that the results are highly similar to those obtained in the Bagging model (Figure 4).

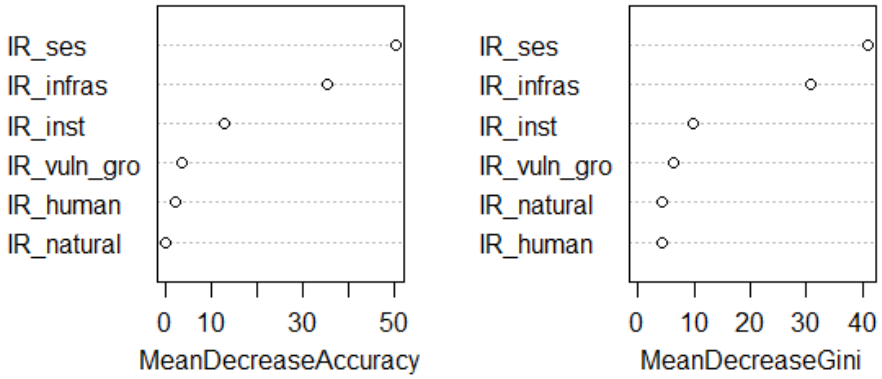

Figure 6: Features importance plot (RF)

In the Random Forest (RF) model, the weighted F1-score is calculated as 1.000 for the training data, while it is 0.7954 for the test data.

V. RESULTS AND DISCUSSION

F1 scores for the training and test datasets utilized in six machine learning algorithms are summarized in Table 4. Upon reviewing the table, it is evident that the KNN, Linear SVM, and CART algorithms exhibit relatively poor performance. A high F1 score for the training data and a significantly lower score for the test data in the context of the Nonlinear SVM algorithm suggest an overfitting issue. Conversely, the Bagging and Random Forest (RF)

algorithms generate uniform and robust F1 scores for both the training and test datasets. Therefore, the most appropriate machine learning algorithm within the scope of the study is Bagging and RF.

Table 4: F1 performance metrics for train and test data

Algorithm	KNN	SVM_Linear	SVM_Nonlinear	CART	Bagging	RF
Weighted F1_train data	0.7713	0.8781	0.9087	0.8577	1	1
Weighted F1_test data	0.7082	0.7536	0.6823	0.7847	0.7954	0.7954

Both Bagging and RF are ensemble learning methods. As evidenced by the results, both models achieved accurate classification performance without exhibiting signs of overfitting, indicating strong generalization capability. Furthermore, these algorithms provide opportunities for the interpretation of models and the selection of features.

The most influential features in the prediction process are socio-economic, institutional, and infrastructure-related variables, as demonstrated in Figures 4 and 6.

Table 1 structures the INFORM Risk Index across three principal dimensions: Hazard and Exposure, Vulnerability, and Lack of Coping Capacity. Within these, the variables pertaining to socio-economic development, institutional performance, and infrastructure robustness emerge as the most influential determinants of human development risk. These findings align with recent vulnerability models that emphasize dynamic socio-economic and institutional drivers as key to resilience [30, 31]. Moreover, recent analyses underscore that infrastructure fragility significantly amplifies hazard impacts and limits effective crisis response [32, 33]. Consequently, these structural factors—socio-economic status, institutionalism, and infrastructure —should be prioritized in risk-informed development planning to mitigate human development vulnerabilities.

The socio-economic dimension of the INFORM Risk Index is divided into three subcomponents: Development and Deprivation (50%), Inequality (25%), and Aid Dependency (25%). A society's inherent vulnerabilities and structural development are collectively evaluated by these indicators. The effectiveness of social protection mechanisms activated during crises and the degree of development and deprivation directly influence the access of individuals to essential services. Economically disadvantaged individuals typically have more limited access to shelter, healthcare, and food, which results in significant disparities in how different population groups experience the impact of crises. Aid dependency is indicative of a structural vulnerability in terms of resilience and sustainability, as it indicates a lack of internal capacity to manage shocks. In conclusion, a society's post-crisis recovery capacity is also restricted by low socio-economic development, which not only reflects current poverty levels. Therefore, this dimension is a critical determinant of human development.

Disaster Risk Reduction (DRR) and Governance comprise the institutional dimension of the INFORM Risk Index. These indicators are indicative of a country's institutional performance in terms of its ability to prevent, manage, and respond to risks. Risk planning, early warning systems, and pre-disaster preparedness are all included in DRR policies. These systems have the potential to directly reduce both human casualties and infrastructure damage. Governance encompasses factors such as accountability, transparency, and equitable access to public services. Misallocation of resources, coordination failures, and a decrease in public trust may be the consequences of inadequate governance. In summary, institutional efficacy is vital for both crisis management and a just and long-lasting recovery process following a disaster.

VI. CONCLUSIONS

This study evaluates the effectiveness of several supervised machine learning algorithms—including KNN, Linear SVM, Nonlinear SVM, CART, Bagging, and RF—in classifying countries into HDI categories using INFORM Risk Index indicators. Results show that ensemble methods, particularly Bagging and RF, outperform individual classifiers by achieving both high accuracy and strong generalization capabilities, as evidenced by balanced F1-scores on training and test datasets.

Feature importance analysis reveals that socio-economic, institutional, and infrastructure-related indicators are the most influential variables in predicting human development levels. These findings reinforce the understanding that development vulnerabilities are rooted not only in exposure to hazards but also in underlying structural factors. In particular, socio-economic inequalities, institutional capacity, and infrastructure robustness play a central role in shaping national resilience and risk outcomes. These dimensions, as defined within the INFORM Risk Index framework, offer a valuable lens for risk-informed development planning.

The study demonstrates the potential of machine learning methods as tools for extracting actionable insights from complex, multidimensional risk data. Ensemble models such as Bagging and RF offer the dual benefits of predictive power and interpretability, making them suitable for policy-relevant applications in development and disaster risk analysis.

For future research, several directions can be pursued to extend the current study. First, following Düzen et al. [34], the integration of machine learning models with multi-criteria decision-making (MCDM) techniques can enhance the transparency and interpretability of classification results, especially in high-stakes policy environments. Second, incorporating temporal components into the dataset could support trend analysis and forecasting, allowing for dynamic modeling of human development trajectories. Third, the inclusion of spatial clustering or regional segmentation may uncover geographically distinct patterns of vulnerability and resilience. Fourth, the use of advanced models—such as deep learning architectures or hybrid ensemble techniques—can be explored to better capture nonlinearities and interactions within high-dimensional risk data. Finally, applying the proposed framework to alternative indices—such as climate risk metrics, health system resilience indicators, or context-specific development benchmarks—would allow researchers to evaluate the robustness and generalizability of the approach across multiple domains.

ACKNOWLEDGMENT

We would like to thank the anonymous reviewers for their informative remarks and ideas which assisted in advancing our research.

REFERENCES

1. UNDP (United Nations Development Programme) (1990) Human Development Report 1990: Concept and Measurement of Human Development. Oxford University Press, New York.
2. Casau M, Ferreira Dias M, Leite Mota G (2024) Economics, happiness and climate change: exploring new measures of progress. Environ Dev Sustain. https://doi.org/10.1007/s10668-024-05702-2

3. UNDRR (United Nations Office for Disaster Risk Reduction) (2022) Global Assessment Report on Disaster Risk Reduction 2022: Our World at Risk – Transforming Governance for a Resilient Future. UNDRR, Geneva. https://www.undrr.org/gar2022
4. Raikes J, Smith TF, Baldwin C, Henstra D (2021) Linking disaster risk reduction and human development. Clim Risk Manag 32:100291. https://doi.org/10.1016/j.crm.2021.100291
5. Oran FÇ (2023) Afet risk yönetiminin insani gelişim endeksi çerçevesinde incelenmesi. Anadolu Univ J Econ Adm Sci 24(2):233–257. https://doi.org/10.53443/anadoluibfd.1185246
6. Feldmeyer D, Birkmann J, McMillan JM et al (2021) Global vulnerability hotspots: differences and agreement between international indicator-based assessments. Clim Change 169(12). https://doi.org/10.1007/s10584-021-03203-z
7. Eze E, Siegmund A (2024) Identifying disaster risk factors and hotspots in Africa from spatiotemporal decadal analyses using INFORM data for risk reduction and sustainable development. Sustain Dev 32(4):4020–4041. https://doi.org/10.1002/sd.2886
8. Mochizuki J, Naqvi A (2019) Reflecting disaster risk in development indicators. Sustainability 11(4):996. https://doi.org/10.3390/su11040996
9. Inter-Agency Standing Committee and the European Commission (2024) INFORM Report 2024: 10 years of INFORM. Publications Office of the European Union, Luxembourg. https://data.europa.eu/doi/10.2760/555548
10. Ricardo M (2011) Inequality and the new human development index. Appl Econ Lett 19:1–3. https://doi.org/10.1080/13504851.2011.587762
11. UNDP (United Nations Development Programme) (2024) Human Development Report 2024: Breaking the Gridlock. UNDP, USA.
12. Halder RK, Uddin MN, Uddin MA et al (2024) Enhancing K-nearest neighbor algorithm: a comprehensive review and performance analysis of modifications. J Big Data 11:113. https://doi.org/10.1186/s40537-024-00973-y
13. Srisuradetchai P, Suksrikran K (2024) Random kernel k-nearest neighbors regression. Front Big Data 7. https://doi.org/10.3389/fdata.2024.1402384
14. Ali A, Hamraz M, Khan DM, Deebani W, Khan Z (2023) A random projection k-nearest neighbours ensemble for classification via extended neighbourhood rule. arXiv preprint arXiv:2303.12210
15. Du KL, Jiang B, Lu J, Hua J, Swamy MNS (2024) Exploring kernel machines and support vector machines: principles, techniques, and future directions. Math 12. https://doi.org/10.3390/math12243935
16. Parlak B, Uysal AK (2019) On classification of abstracts obtained from medical journals. J Inf Sci 46(5):648–663. https://doi.org/10.1177/0165551519860982
17. Amaya-Tejera N, Gamarra M, Velez JI, Zurek E (2024) A distance-based kernel for classification via support vector machines. Front Artif Intell 7. https://doi.org/10.3389/frai.2024.1287875
18. Almaspoor MH, Safaei A, Salajegheh A, Minaei-Bidgoli B (2021) Support vector machines in big data classification: a systematic literature review. Preprint, Research Square. https://doi.org/10.21203/rs.3.rs-663359/v1
19. Sun H (2024) pSVM: Soft-margin SVMs with p-norm hinge loss. arXiv preprint arXiv:2408.09908
20. Breiman L, Friedman J, Olshen R, Stone C (1984) Classification and Regression Trees. Wadsworth, Belmont.
21. Doğruel M, Soner Kara S (2023) Determining the happiness class of countries with tree-based algorithms in machine learning. Acta Infologica 7(2):243–252. https://doi.org/10.26650/acin.1251650
22. Ngo G, Beard R, Chandra R (2022) Evolutionary bagging for ensemble learning. Neurocomputing 510:1–14. https://doi.org/10.1016/j.neucom.2022.08.055
23. Malhotra R, Cherukuri M (2024) A systematic review of hyperparameter tuning techniques for software quality prediction models. Intell Data Anal 28. https://doi.org/10.3233/IDA-230
24. Breiman L (1996) Bagging predictors. Mach Learn 24:123–140. https://doi.org/10.1007/BF00058655
25. Fernández A, García S, Galar M, Prati RC, Krawczyk B, Herrera F (2018) Learning from Imbalanced Data Sets. Springer Nature, Switzerland. https://doi.org/10.1007/978-3-319-98074-4
26. Wei B, Wang F et al (2024) Adaptive bagging-based dynamic ensemble selection. Expert Syst Appl 255:124860. https://doi.org/10.1016/j.eswa.2024.124860
27. Breiman L (2001) Random forests. Mach Learn 45:5–32. https://doi.org/10.1023/A:1010933404324
28. Cutler A, Cutler DR, Stevens JR (2012) Random forests. In: Cha Z, Yunqian M (eds) Ensemble Machine Learning: Methods and Applications. Springer, New York, pp 157–175. https://doi.org/10.1007/978-1-4419-9326-7_5
29. Geurts P, Ernst D, Wehenkel L (2006) Extremely randomized trees. Mach Learn 63:3–42. https://doi.org/10.1007/s10994-006-6226-1
30. Birkmann J et al (2022) Understanding human vulnerability to climate-induced disasters. Sci Total Environ 803:150065. https://doi.org/10.1016/j.scitotenv.2021.150065

31. Jamshed A et al. (2023) A bibliometric and systematic review of the Methods for the Improvement of Vulnerability Assessment in Europe (MOVE) framework: A guide for the development of further multi-hazard holistic frameworks. Jàmbá J Disaster Risk Stud 15:1486. https://doi.org/10.4102/jamba.v15i1.1486
32. Reimann L et al (2024) An empirical social vulnerability map for flood risk assessment at global scale (GlobE-SoVI). Earth's Future 12:e2023EF003895. https://doi.org/10.1029/2023EF003895
33. Verschuur J et al (2024) Quantifying climate risks to infrastructure systems: a comparative review of developments across infrastructure sectors. PLOS Clim 3(4):e0000331. https://doi.org/10.1371/journal.pclm.0000331
34. Düzen MA, Bölükbaşı İB, Çalık E (2024) How to combine ML and MCDM techniques: an extended bibliometric analysis. J Innov Eng Nat Sci 4:642–657. https://doi.org/10.61112/jiens.1475948