REST Journal on Data Analytics and Artificial Intelligence

Vol: 4(1), March 2025

REST Publisher; ISSN: 2583-5564

Website: https://restpublisher.com/journals/jdaai/

DOI: https://doi.org/10.46632/jdaai/4/1/34

SPSS For Disaster Prediction: A Machine Learning and Statistical Perspective

Jomon Jose, Jisha T

ST. Joseph's College, Devagiri, Calicut, Kerala, India.

Corresponding Author Email: jomonjose621@gmail.com

Abstract: *Disaster risk management is a critical area of study, as effective preparedness and response strategies can significantly reduce the impact of hazardous events. This research evaluates key disaster response parameters, including Risk Score, Preparedness Level, Response Time, Economic Impact, and Casualty Risk, using statistical analysis. By identifying correlations and patterns among promoting improvements disaster resilience and mitigation efforts. Research Significance: This study is significant as it provides an evidence-based method to disaster preparedness and response evaluation. By analyzing key risk factors and preparedness indicators, the findings help policymakers, emergency responders, and infrastructure planners develop more effective risk reduction strategies. The research also highlights how different environmental and infrastructural conditions influence response efficiency. Methodology: The research employs SPSS statistical analysis, utilizing methods such as, regression analysis, and ANOVA—these methods help determine the relationships between evaluation parameters and assess their collective impact on disaster response effectiveness. Alternative Influencing Factors: In addition to the primary evaluation parameters, the study considers several alternative influencing factors that may impact disaster preparedness and response, including: Disaster Type (e.g., natural vs. man-made disasters) Location Type (urban vs. rural areas) Weather Conditions (floods, storms, droughts, etc.) Ground Stability (earthquake-prone areas, landslide risks) Warning System Availability (early warning mechanisms, public alert systems) Infrastructure Readiness (emergency shelters, transportation networks, healthcare facilities) Evaluation Parameters: The key parameters assessed in this study include: Risk Score: Measures the overall level of risk associated with a disaster event. Preparedness Level: Evaluates the extent of readiness before a disaster occurs. Response Time: Assesses the efficiency of emergency response efforts. Economic Impact: Estimates the financial and infrastructural losses caused by disasters. Casualty Risk: Analyzes the potential for injuries or fatalities due to a disaster. Results: The analysis reveals a strong correlation among preparedness indicators, with Principal Component Analysis (PCA) showing that over 91% of the total variance is explained by a single component. This suggests that disaster response factors are highly interdependent. Regression and ANOVA results further confirm the statistical significance of these parameters in predicting disaster outcomes. The study emphasizes the need for enhanced early warning systems, robust infrastructure, and efficient response mechanisms to mitigate disaster risks.*

Keywords: *Disaster Risk Assessment, Preparedness Level, Response Time, Economic Impact, Casualty Risk, Principal Component Analysis (PCA), SPSS Statistical Analysis, Emergency Response Planning.*

1. INTRODUCTION

Many states in India face frequent flood disasters every year due to the lack of an effective early warning system to alert residents of affected areas. For India to progress as a smart country. A model has been developed to monitor environmental parameters with the ability to predict flood disasters. Various sensors detect parameters such as temperature, humidity, atmospheric pressure, and rainfall, and the collected data is sent to a microcontroller. Therefore, the intelligent flood disaster forecasting method proves to be more suitable for practical deployment and reliability. It provides real-time monitoring, continuous updates of environmental parameters, and accurate flood prediction, which is more effective than existing methods. The forecast chart shows that flood events were once predicted in advance by providing early warnings to at-risk communities through Think Tweet. The alerts are communicated in simple, easy-to-understand language via the internet.[1] First, the primary factors contributing to slope failure should be identified and selected. In this study, the test zone environment is examined and studied. Various environmental factors responsible for disasters are selected, evaluated, and analyzed. To enable disaster the suggested disaster forecasting support model is used to build PDAS for forecasting and monitoring. This model incorporates predictive capability techniques such as Analytical Back-propagation Neural Multivariate Statistical Analysis (MSA), Network Process (ANP), and Network Analysis (BPN) to assess and contrast the

most appropriate approach.[2] The analysis results were tabulated and used as early warning recommendations. Chen's study emphasized extensive data collection, which led to complex and time-consuming analysis processes. In contrast, the proposed fuzzy theory simplifies the assessment by using a case-inference method, providing early warnings through spreadsheets. The occurrence of debris flows is positively correlated with factors such as effective watershed size and stream length. A larger effective watershed means more water in the stream, which increases the likelihood of debris flows. Similarly, longer channel lengths extend the distance that debris flows can travel.[3] Research findings on risk and benefit assessment for hurricane disaster prevention and mitigation can contribute to improving hurricane disaster risk management technology. The first approach is a technique for estimating upper and lower limit intervals. which is applicable when there are more than three similar hurricane models based on the affected disaster area. The latter method is used when the cyclone is near, resulting in reduced uncertainty. As the cyclone approaches, the prediction range gradually decreases, leading to a continuous improvement in forecast accuracy.[4] This system forecasts the magnitude and spatial distribution of urban earthquakes using the GIS spatial analysis model, the MVC design process, and fundamental earthquake engineering concepts disasters. It helps to shift reducing urban disasters through dynamic management as opposed to static planning. In general, the contours of seismic intensity form an oval shape due to the influence of geological structures. The intensity decays very slowly in the direction of the earthquake, while it weakens very rapidly in the direction perpendicular to the fault. As distance increases, these differences gradually decrease and the boundaries become more rounded. [5] Today, many countries are using flood monitoring and disaster forecasting tools to increase disaster awareness. These forecasting tools help overcome the challenges of collecting information about flood risks by using data parameters for assessment. The primary gadget will use a water level sensor that measures in centimeters send the data wirelessly to a reporting server for analysis and evaluation.[6] With the use of this sophisticated flood disaster predicting technology, consumers and officials may get early alerts so that They are able to make well-informed choices on infrastructure reinforcement, resource distribution, and evacuation. The primary the goal aims to develop an sophisticated flood disaster prediction system for Terengganu that can deliver accurate flood forecasts and level measurements using real-time IoT-based river water data. A systematic and structured approach to creating and implementing an intelligent flood disaster forecasting system is provided by this framework. [7] The privacy and security component will put in place safeguards to preserve user anonymity, guarantee network and data security, and protect data integrity throughout the entire disaster forecasting and response. Disaster forecasting will play a key part of the suggested smart system's implementation components. The utilization of multimedia big data gathered from city sensors and OSNs would increase the precision of disaster event forecasts. Although our imaginary system has significant social appeal, this alone may not be enough to attract the necessary number of volunteer data collectors. [8] There are two primary ways that the suggested fire catastrophe management system sets itself apart. First, it provides a more thorough method by combining building information into a unified fire information breakdown system than previous studies that focused only on fire-related data. The visual representation illustrates how the existing area affects rescuers and evacuees, with darker shadows indicating increased mobility challenges. As mentioned According to the survey results in the preceding section, this system improves the existing management procedure and is a useful instrument for managing fire disasters. [9] Summary of DIAS and its use in catastrophe risk reduction and climate change analysis. According to the The Paris Agreement, the SDGs, and initiatives to slow down climate change are aligned with efforts aimed at improving resilience to climate-related disasters. Therefore, identifying and assessing overlooked disaster risks is crucial. DIAS is actively seeking opportunities to collaborate with various private sector organizations. [10] The implementation of disaster preparedness systems has been slow in the majority of African nations because of a lack of technical resources. In remote locations, keeping up globally standardized early warning systems is another significant challenge, given the unique dynamics of these communities. An all-encompassing, multidisciplinary approach to detecting, assessing, and reducing disaster-related hazards is part of disaster management. Meteorological and disaster modeling early warning systems play a key role in risk management by providing risk assessment, monitoring, forecasting, warnings, and response strategies. Making historical weather data publicly accessible encourages community engagement in climate and disaster mitigation efforts, promotes weather information sharing, ensures data quality, and supports the development of community-based weather stations. [11] Debris flows are very hazardous geological calamities that usually takes place on steep hillsides or in valleys at the base of mountains. They can also help assess the variability in the hydrological ecosystem and the possible hazards of natural disasters indices in geological catastrophe predictions. It is noteworthy that a branch of artificial intelligence called machine learning, is rapidly advancing in its use in modeling geological disasters. By creating a real-time geological catastrophe monitoring and warning system, it was able to overcome the problem of landslide warnings.[12] Despite the increased use of the technology, some management challenges persist. One of the most important issues In cloud computing, disaster management works well. This will increase the dependability of data availability and facilitate open communication between users and technical support personnel. during recuperation from disasters. server virtualization, disaster recovery, and more. There is a concern about potential data loss because the exact storage location of your data is not known with the cloud vendor. To address this, the proposed solution introduces a monitoring agent for disaster recovery, integrating prediction and monitoring at a single site using a linear regression algorithm. [13] Every disaster has warning signs, often seen as disturbances in environmental parameters such as temperature. The proposed system includes a WSN network and a web management platform. The WSN's low power consumption was accomplished by using a modified trickle routing algorithm, which consists of a WSN and a web management platform. [14] We analyze Twitter posts during based on

priority levels, distinguishing between high and low priority messages. Most disaster detection systems focus solely on determining whether a tweet is disaster-related based on its text content. Relevant Tweets are used to alert people to the need for safety precautions. It assessed the disaster risk level in Seoul by analyzing tweet text. [15] In Malaysia, the government is crucial to the distribution system's implementation following the flood disaster. Educating communities and raising awareness about disaster risk management can contribute to reducing the impact of disasters. Disaster management is divided into pre-disaster risk reduction and post-disaster recovery stages. The most important stage is the final stage of pre-disaster risk reduction. The NFIP has become One of the most established government-sponsored disaster insurance schemes worldwide. [16] Leading experts Mathematical sciences, information sciences, and disaster simulation are working together to develop the first analytics platform in history. The platform efficiently analyzes and integrates data from several observations, facilitating big data processing and real-time simulation. Understanding human behavior patterns during disasters is also important. [17] Strategies for disaster prevention and mitigation that are organized along a time axis determined by the typhoon's path, which is shaped by risk assessments of potential damage. We take into account the various institutions designated as disaster prevention offices and the responsibilities and tasks assigned to them. We aim to support disaster prevention activities in line with the chronology. This idea has led us to create a system that will identify similar cyclones when a new cyclone occurs and provide response levels in line with the timeline, including mitigation measures. Typically, disaster prevention and mitigation action plans are identified and their implementation is overseen by disaster prevention offices. Individual teams carry out the response activities assigned to them based on their assigned tasks. However, this approach rarely takes into account the path of the cyclone.[18] Since a disaster won't happen until the safe zone boundary is reached, this isn't exactly a disaster forecast is exceeded. However, it serves as a strong indicator of a high potential for disaster. This implies that an outside disruption could push a controlled system onto a path where the oracle detects an impending disaster and triggers a controlled recovery. We believe that demonstrating the validity of our disaster forecasting method for this process increases its applicability to many other processes in laboratories or industries. [19] The system improves resource utilization, threat detection, and support through disaster recovery planning, security policy management, and performance indicators. Disaster recovery and availability present significant challenges in Kubernetes environments. Maintaining business continuity across geographically distributed clusters requires thorough planning and effective management of backup and recovery processes. [20]

2. MATERIALS AND METHOD

Input parameter: 1. Disaster Type Each disaster, whether natural (e.g., floods, earthquakes, hurricanes) or man-made (e.g., industrial accidents, fires), demands a unique response strategy. For instance, early warning systems for earthquakes rely on seismic activity detection, while flood management depends on real-time water level monitoring and forecasting. 2. Region Disaster preparedness varies based on geographic and climatic conditions. Coastal regions are more prone to hurricanes and tsunamis, requiring strong embankments and evacuation plans. Mountainous areas face landslides, necessitating slope monitoring and soil stabilization efforts. 3. Season The likelihood of disasters often depends on seasonal patterns. Monsoons increase flood risks, demanding robust drainage systems, while dry seasons heighten wildfire threats, requiring strict fire control measures. Season-based disaster planning ensures timely mitigation efforts. 4. Warning System Efficient early warning systems are critical for minimizing disaster impact. Advanced alternatives include satellite-based monitoring, AI-driven predictive analytics, IoT sensor networks, and community alert systems via mobile applications or sirens. The choice of warning system depends on the disaster type and available technology. 5. Population Density: Urban areas with high population density require well-structured evacuation routes, emergency shelters, and effective communication networks. In contrast, sparsely populated rural areas may rely on community-based disaster response teams and decentralized warning systems. 6. Infrastructure: The resilience of infrastructure plays a crucial role in disaster mitigation. Alternatives include disaster-resistant building designs, underground power lines to prevent storm damage, reinforced bridges, and adaptive water management systems. Strengthening infrastructure enhances long-term disaster preparedness.

Output parameter: Evaluating the effectiveness of a disaster management system requires assessing various output parameters. These parameters determine how well the system performs in predicting, mitigating, and responding to disasters. Key output parameters include: 1. Accuracy: The accuracy of disaster prediction and response systems is crucial for minimizing false alarms and ensuring timely intervention. Advanced data analytics, AI-based models, and real-time sensor monitoring help improve accuracy by providing precise forecasts and risk assessments. A highly accurate system reduces unnecessary evacuations and optimizes resource allocation. 2. Reliability: A disaster management system must function consistently under different conditions. Reliability is determined by the system's ability to operate without failure, provide continuous monitoring, and deliver dependable alerts. Redundant communication channels, cloud-based data storage, and backup power sources enhance reliability, ensuring uninterrupted disaster response. 3. Response Time: Quick response time is essential in disaster management, as delays can lead to increased casualties and property damage. Automated alerts, AI-driven decision-making, and efficient emergency communication networks help minimize response time. A well-optimized system ensures rapid deployment of emergency services, reducing disaster impact. 4. Public Awareness: A disaster management system’s effectiveness depends not just on technology but also on the general

awareness and preparedness. Clear communication strategies, early warning alerts, and educational programs help communities understand risks and take appropriate action. Increased public awareness enhances cooperation and reduces panic during disasters. 5. Data Coverage: Comprehensive data coverage improves disaster prediction and response by integrating information from various sources, including weather satellites, IoT sensors, historical records, and social media. A wider data scope ensures better situational awareness, allowing authorities to make informed decisions and deploy resources effectively.

SPSS: SPSS provides powerful data management tools that help researchers efficiently organize, clean, and manipulate datasets. With support for multiple data formats, the software facilitates seamless data import and export across a variety of programs. In psychological research, data is often complex and heterogeneous, and SPSS streamlines the coding, labeling, and management of variables. SPSS provides advanced data visualization tools, such as histograms, scatterplots, bar charts, and box plots, that help researchers create clear and insightful graphics. These visual representations help them discover trends, relationships, and data distributions that may not be immediately apparent through numerical analysis alone. SPSS streamlines the coding, labeling, and management of variables, ensuring a well-structured dataset ready for analysis while minimizing errors and inconsistencies.[2] He begins with correlation and demonstrates how to create scatterplots, add a regression line, and create multiple scatterplots using various SPSS windows. If statistics are challenging and you are involved in second language research, including LSP studies, SPSS is worth considering. This book provides the basic background needed to perform and interpret basic statistical tests.[3] SPSS, short for allows users to enter primary and secondary data, much like Microsoft Excel. Its user-friendly menu bar facilitates easy data analysis, enabling a wide range of statistical procedures. In SPSS, correlation analysis can be used. To start this analysis, users need to go to the analysis section, choose the variables required for correlation, and select the appropriate method, such depending on the data type. Additionally, a significance test can be designed by specifying the test tail.[4] SPSS supports correlation analysis to assess relationships between quantitative variables. Depending on the data type, users can access this feature by navigating to the analysis and selecting the relevant variables. A significance test can be designed by defining the test tail. SPSS allows users to perform correlation analysis to assess relationships between quantitative variables. Depending on the data type, users can access this feature by navigating to Analysis, selecting the relevant variables, and selecting the appropriate correlation method. Additionally, they can customize the significance test by specifying the test tail. SPSS enables users to conduct correlation analysis to assess relationships between quantitative variables. Based on the data type, users can access this functionality by navigating to Analysis, selecting the relevant variables, and selecting the appropriate correlation method. They can design a significance test by defining the test tail.[5] A key A key feature of SPSS is its intuitive design, making it accessible to users without a technical background, especially in the social sciences. Users can run the software without prior knowledge of programming languages. Understanding the basic concepts of SPSS helps researchers analyze quantitative data efficiently, smoothing the process and minimizing potential challenges. SPSS requires users to define variables and input data into these variables to create cases for SPSS can perform all the essential tests needed for quantitative data analysis in the social sciences. Given its capabilities, it has become a preferred choice, but in some cases, an essential tool for social researchers to effectively analyze and present quantitative data. IBM tools among social scientists worldwide. Over the past fifty years, it has undergone many improvements to meet the evolving needs of social science researchers.[6] The macros discussed in this article provide SPSS and SAS users with an accessible command line for conducting this type of analysis. However, researchers should note that there are additional options for examining mediation in more complex models. The macro only needs to be run once when SPSS or SAS is first started, and remains active until the program is closed. Detailed instructions for using macros are provided in the appendices, and electronic copies of the macros are available for download.[7] In addition, re-arranging the data is a complex, time-consuming, and error-prone process. Dyads are considered indistinguishable when there is no consistent way to distinguish or assign a sequence to the two individuals in each pair.[8] One of the main drawbacks of SPSS is its price. Different versions offer different analytical functions and limits on Limitation on the number of use cases and variables. In addition, most licenses have an expiration date. software cannot be used unless it is renewed. The SPSS Student Suite offers comprehensive analytical tools with special features, from basic descriptive statistics to advanced general linear modeling. features that enable variable transformations in preparation for various statistical tests. software that makes it a valuable tool to master. Becoming proficient with SPSS requires some learning time, and annual license fees may also be a consideration. One of the main drawbacks of SPSS is its price. Different versions offer different analytical functions and capabilities for handling cases and variables. In addition, most licenses automatically expire after a certain period of time, after which the software cannot be used.[9] One of the main drawbacks of SPSS is its price. Different versions come with different analytical functions and capabilities for handling cases and variables. associated distance function, Euclidean distance, is consistent with our everyday perception of spatial relationships. For ease of identification, the subject number will be However, when entering multiple square matrices, an identifier is not required for each matrix. The final related subcommand relates to the content of the output from SPSS. Using the PRINT subcommand with the DATA option causes ALSCAL to display all the matrices of the original and transformed data, while the HEADER option provides a summary of all selected settings.[10] SPSS extends beyond basic statistical analysis by providing advanced modeling tools for complex research needs. It includes factor and cluster analyses to identify hidden patterns and groupings within data.

3. RESULT AND DISCUSSION

TABLE 1. Reliability Statistics

Reliability Statistics		
Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
.967	.974	5

The reliability analysis of a given scale is measured by Cronbach's alpha, which indicates exceptionally high internal consistency. A Cronbach's alpha value of 0.967 suggests that the items on the scale are highly correlated, meaning that they effectively measure the same underlying construct. This level of reliability is considered excellent because values above 0.9 generally indicate very strong internal consistency, reducing the likelihood of measurement error. In addition, the Cronbach's alpha based on standardized items is 0.974, which is slightly higher than the unstandardized value. When items are standardized, their internal consistency improves further, reinforcing the strength of the scale. The analysis is based on five items, meaning that all five contribute significantly to the overall reliability. A high Cronbach's alpha suggests that the scale is stable and reliable in measuring the intended concept. However, it is important to ensure that the items are not overly redundant, as very high values (above 0.95) may indicate that some items are very similar. Overall, this outcome measure is well-structured and suitable for research, ensuring reliable and valid data collection.

TABLE 2. Reliability Statistic individual

Item-Total Statistics	
	Cronbach's Alpha if Item Deleted
Risk Score	0.958
Preparedness Level	0.961
Response Time	0.96
Economic Impact	0.953
Casualty Risk	0.959

The item-total statistical table provides insights into the internal consistency of a questionnaire or scale by assessing how each item contributes to the overall reliability. In this case, the table provides the Cronbach's alpha value if the item is removed, which indicates what the Cronbach's alpha coefficient would be if a particular item were removed from the analysis. Higher values suggest stronger internal consistency between items. In this dataset, the overall Cronbach's alpha is high, indicating a more reliable level. The values for each item are as follows: Risk score (0.958): Removing this item would slightly reduce reliability, indicating that it is well aligned with the overall scale. Readiness level (0.961): It has the highest value, suggesting that its removal would slightly improve reliability, although it would contribute slightly less to internal consistency. Response time (0.960): Similar to readiness level, its removal could slightly increase reliability. Economic Impact (0.953): This item has a very low value, meaning it contributes well to the scale, and removing it would have a very significant negative effect on reliability. Accident Risk (0.959): This item aligns strongly with the overall scale, and removing it would not significantly improve reliability.

TABLE 3. Descriptive Statistics

	N	Range	Minimum	Maximum	Sum	Mean		Std. Deviation	Variance	Skewness		Kurtosis	
	Statistic	Statistic	Statistic	Statistic	Statistic	Statistic	Std. Error	Statistic	Statistic	Statistic	Std. Error	Statistic	Std. Error
Risk Score	30	3	2	5	109	3.63	.200	1.098	1.206	-.197	.427	-1.239	.833
Preparedness Level	30	2	3	5	126	4.20	.139	.761	.579	-.362	.427	-1.141	.833
Response Time	30	2	3	5	127	4.23	.141	.774	.599	-.441	.427	-1.160	.833
Economic Impact	30	3	2	5	110	3.67	.200	1.093	1.195	-.289	.427	-1.174	.833
Casualty Risk	30	2	3	5	123	4.10	.147	.803	.645	-.188	.427	-1.406	.833
Valid N (listwise)	30												

The descriptive statistics table provides an overview of the key statistical measures for the five variables: risk score, preparedness level, response time, economic impact, and accident risk, based on a sample size of 30 observations. These measures include range, minimum and maximum values, sum, mean, standard deviation, variance, skewness, and kurtosis. The mean of the risk score is 3.63, with scores ranging from 2 to 5. The standard deviation (1.098) indicates moderate variability. The negative skewness (-0.197) indicates a slight leftward skewness, while the kurtosis (-1.239) indicates a flatter distribution than normal. The preparedness level has a high mean of 4.20, indicating that most responses are skewed toward the higher end of the scale (3 to 5). With a low standard deviation (0.761), the data is less spread out. The skewness (-0.362) shows a slight leftward skew, and the kurtosis (-1.141) indicates a flatter distribution than the normal curve. Response time has a similar mean (4.23) and spread (standard deviation = 0.774) for readiness level, showing a trend towards higher scores. The skewness (-0.441) indicates a slight leftward skew, and the kurtosis (-1.160) indicates a somewhat flatter spread. The economic impact has an average of 3.67, with scores ranging from 2 to 5, and a standard deviation of 1.093, indicating moderate variability. Its skewness (-0.289) shows a slight leftward skew, and the kurtosis (-1.174) indicates a flatter distribution. The accident risk has an average of 4.10 and a relatively small standard deviation (0.803), meaning that the responses are fairly consistent. The skewness (-0.188) is close to zero, indicating a nearly symmetrical distribution, while the kurtosis (-1.406) indicates a flatter curve than normal.

TABLE 4. Frequencies Statistics

Statistics						
		Risk Score	Preparedness Level	Response Time	Economic Impact	Casualty Risk
N	Valid	30	30	30	30	30
	Missing	0	0	0	0	0
Median		3.69[a]	4.25[a]	4.29[a]	3.75[a]	4.14[a]
Mode		4	4[b]	5	4	4[b]
Percentiles	25	2.69	3.5	3.53	2.75	3.37
	50	3.69	4.25	4.29	3.75	4.14
	75	4.59	4.88	4.92	4.61	4.82

The statistical table provides a summary of key central tendency measures, including mean, mode, and percentages, for five variables: risk score, preparedness level, response time, economic impact, and accident risk. The dataset contains 30 valid observations for each variable, with no outliers. Median: The median represents the middle value of the dataset. Values range from 3.69 (risk score) to 4.29 (response time), indicating that most responses are skewed toward the higher end of the scale. Mode: The mode represents the most frequently occurring value. For risk score, preparedness level, economic impact, and accident risk, the mode is 4, while for response time, the most common response is 5. Many respondents rated response time as very high. Percentiles: The 25th percentile (Q1) represents the lowest quartile, below which 25% of the data fall. Risk score and economic impact have low Q1 values (2.69 and 2.75, respectively), indicating high variability in responses, while other variables have Q1 values above 3. The 50th percentile (Q2, or median) confirms

that the median values are generally higher across all variables. The 75th percentile (Q3) shows that 25% of responses are close to the maximum values, with readiness level, response time, and accident risk having Q3 values above 4.8, reflecting a trend toward higher ratings.

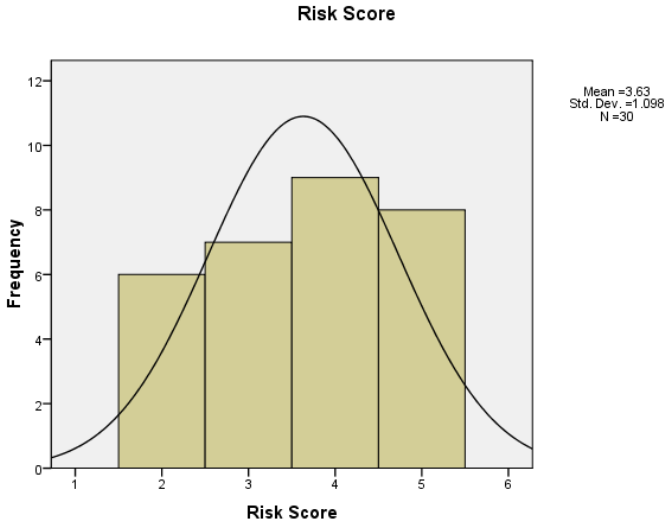

FIGURE 1: Risk Score

Figure 1 illustrates the distribution of risk scores across 30 observations. The x-axis represents the risk score values, while the y-axis shows the frequency of occurrence. The histogram bars indicate how often each score appears in the dataset, and the superimposed normal curve provides a visual representation of the distribution of the data. The mean risk score is 3.63 with a standard deviation of 1.098, indicating a moderate spread of responses around the mean. With a concentration of scores of 3 and 4, the distribution appears slightly skewed to the left. The presence of responses across the full range (approximately 2 to 5) indicates variation in perceived risk. The normal curve suggests that while the data is highly symmetrical, it may not be normally distributed. This visualization helps us understand the overall pattern of risk perception, with most scores clustering around the middle to upper range, suggesting a moderate to high risk assessment.

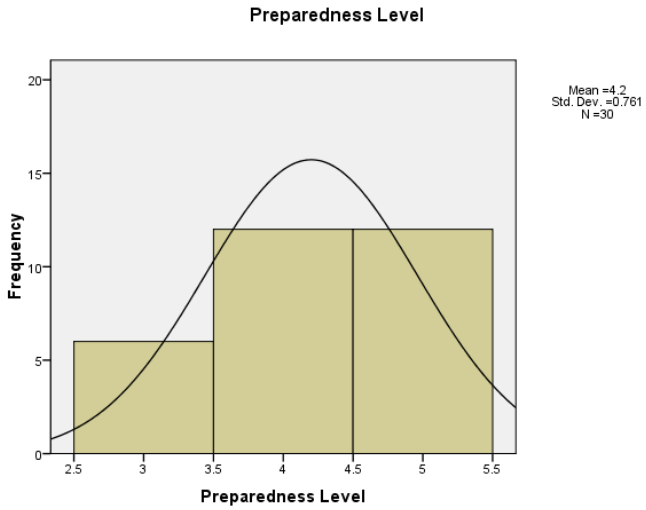

FIGURE 2: Preparedness Level

Figure 2: Distribution of Readiness Level Figure 2 shows a graph illustrating the distribution of readiness levels across 30 observations. The x-axis represents readiness level scores, while the y-axis represents the frequency of responses. The bars indicate how often each score appears in the dataset, with a normal curve overlay providing a visual indication of the distribution pattern. The mean readiness level is 4.2 with a standard deviation of 0.761, indicating that most responses are concentrated at the high end of the scale. The histogram suggests a slightly left-skewed distribution, with most responses falling between 3.5 and 5.5. The relatively low standard deviation indicates that there is little variation in readiness levels among respondents. This distribution suggests that most participants perceive a high level of readiness, with most scores

tending toward the upper range. The presence of a few low values indicates some variation, but overall, respondents generally appear to report a strong level of readjustment.

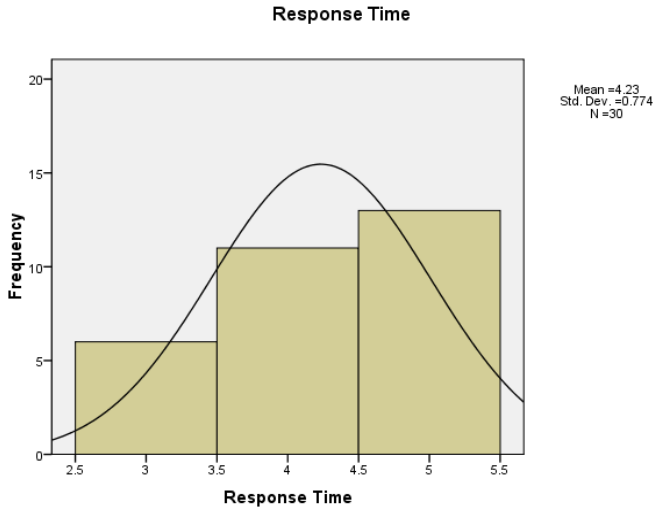

FIGURE 3: Response Time

Figure 3 illustrates the distribution of response times across 30 observations. The x-axis represents response time scores, while the y-axis represents the frequency of responses. The histogram bars show the frequency of different response times, with a normal curve overlay providing insight into the pattern of the distribution. The mean response time is 4.23, with a standard deviation of 0.774, suggesting that most responses are concentrated toward the high end of the scale. The distribution appears to be slightly skewed to the left, with most values between 3.5 and 5.5. The relatively small standard deviation indicates low variability, meaning that most respondents gave similar ratings of response time. Overall, the data suggests that respondents generally respond quickly and efficiently, with most ratings above 4.0. The shape of the distribution suggests that there are some low ratings, but they are not widespread, reinforcing the consistency of positive response time perceptions.

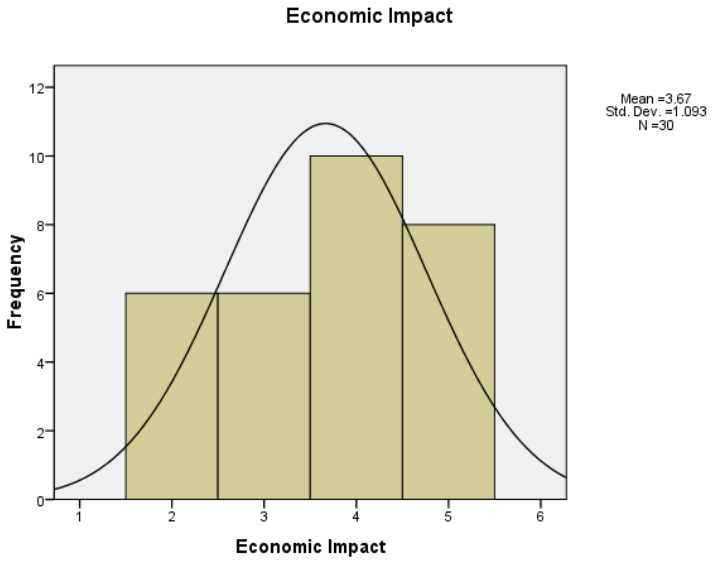

FIGURE 4: Economic Impact

Figure 4 presents a graph illustrating the distribution of economic impact across 30 observations. The x-axis represents economic impact scores, while the y-axis shows the frequency of responses. The histogram bars depict the occurrence of different economic impact ratings, with a normal curve overlay providing a visual reference for the distribution. The mean economic impact score is 3.67, with a standard deviation of 1.093, indicating moderate variation in responses. The distribution appears to be slightly skewed to the left, with most scores clustered between 2 and 5. Although many respondents rated economic impact as 4, the histogram suggests that there is a significant spread of responses across the

scale, indicating disagreement. Overall, the data suggest moderate economic impact, with responses showing a mix of opinions, although the majority lean toward a medium to high rating.

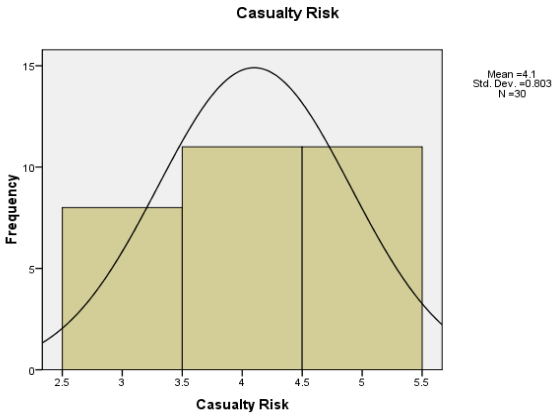

FIGURE 5: Casualty Risk

Figure 5 presents a histogram illustrating the distribution of Casualty Risk among 30 observations. The x-axis represents the Casualty Risk scores, while the y-axis displays the frequency of responses. The histogram bars indicate how frequently each score appears in the dataset, with an overlaid normal curve providing a reference for the distribution pattern. The mean Casualty Risk score is 4.1, with a standard deviation of 0.803, suggesting that most responses are concentrated around the higher end of the scale. The distribution appears slightly left-skewed, with the majority of responses falling between 3.5 and 5.5. The relatively low standard deviation suggests moderate variability, meaning that while responses show some dispersion, they are mostly within a close range.

TABLE 5. Correlations

		Risk Score	Preparedness Level	Response Time	Economic Impact	Casualty Risk
Pearson Correlation	Risk Score	1.000	.833	.835	.957	.903
	Preparedness Level	.833	1.000	.972	.870	.812
	Response Time	.835	.972	1.000	.869	.849
	Economic Impact	.957	.870	.869	1.000	.903
	Casualty Risk	.903	.812	.849	.903	1.000
Sig. (1-tailed)	Risk Score	.	.000	.000	.000	.000
	Preparedness Level	.000	.	.000	.000	.000
	Response Time	.000	.000	.	.000	.000
	Economic Impact	.000	.000	.000	.	.000
	Casualty Risk	.000	.000	.000	.000	.
N	Risk Score	30	30	30	30	30
	Preparedness Level	30	30	30	30	30
	Response Time	30	30	30	30	30
	Economic Impact	30	30	30	30	30
	Casualty Risk	30	30	30	30	30

The correlation table presents Pearson correlation coefficients between five key variables: Risk Score, Preparedness Level, Response Time, Economic Impact, and Casualty Risk, based on 30 observations. The Pearson correlation measures the strength and direction of the linear relationship between variables, where values close to +1 indicate a strong positive correlation, values near -1 suggest a strong negative correlation, and values around 0 imply no correlation. The results indicate strong positive correlations among all variables, with significance levels (p-values) of 0.000, meaning these relationships are statistically significant. Notable findings include: Risk Score and Economic Impact ($r = 0.957$, $p < 0.001$): This is the strongest correlation, suggesting that as risk perception increases, the perceived economic impact also rises. Risk Score and Casualty Risk ($r = 0.903$, $p < 0.001$): A high-risk score is closely associated with an increased perception of casualty risk. Preparedness Level and Response Time ($r = 0.972$, $p < 0.001$): This very strong correlation implies that

higher preparedness is linked to quicker response times. Casualty Risk and Economic Impact ($r = 0.903$, $p < 0.001$): The perception of casualty risk is strongly related to concerns about economic consequences.

TABLE 6. Regression analysis Model Summary

Model Summary										
					Change Statistics					
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	R Square Change	F Change	df1	df2	Sig. F Change	Durbin-Watson
1	.962[a]	.926	.914	.322	.926	77.973	4	25	.000	2.423
2	.976[a]	.953	.945	.178	.953	125.919	4	25	.000	1.982
3	.978[a]	.956	.949	.174	.956	136.846	4	25	.000	2.116
4	.968[a]	.937	.927	.295	.937	93.234	4	25	.000	2.855
5	.930[a]	.866	.844	.317	.866	40.309	4	25	.000	1.671

The Model Summary table provides key statistical indicators for evaluating the performance of multiple regression models predicting the dependent variable. The table includes R (correlation coefficient), R Square (coefficient of determination), Adjusted R Square, Standard Error of the Estimate, Change Statistics, and the Durbin-Watson statistic for different models. Key Findings: R and R Square Values: The highest R value (0.978) occurs in Model 3, indicating a very strong relationship between the independent variables and the dependent variable. The R Square values range from 0.866 to 0.956, meaning that the independent variables explain 86.6% to 95.6% of the variance in the dependent variable. Model 3 has the highest R Square (0.956), making it the best-fit model. Adjusted R Square values are slightly lower than R Square, adjusting for the number of predictors in the model. Standard Error of the Estimate (SEE): The lower SEE values indicate better model accuracy. Model 3 has the lowest SEE (0.174), meaning it provides the most precise predictions among all models. F-Change and Significance (p-value = 0.000 for all models): The F-statistic is significantly high across all models, confirming the overall significance of the models. The p-values (Sig. F Change) are 0.000, indicating that the predictors significantly contribute to the dependent variable. Durbin-Watson Statistic: The Durbin-Watson values range from 1.671 to 2.855, which are within the acceptable range (close to 2), suggesting no serious autocorrelation in residuals.

TABLE 7. ANOVA

ANOVA[b]						
Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	32.372	4	8.093	77.973	.000[a]
2	Regression	16.006	4	4.001	125.919	.000[a]
2	Regression	32.489	4	8.122	93.234	.000[a]
4	Regression	16.608	4	4.152	136.846	.000[a]
5	Regression	16.190	4	4.047	40.309	.000[a]

The ANOVA (Analysis of Variance) table provides statistical insights into the overall significance of regression models. It assesses whether the independent variables explain significantly the variance in the dependent variables. The main components of the table include the sum of squares, degrees of freedom (df), mean square, F-statistic, and significance (Sig.) values. Key findings: Sum of squares (regression and error): The regression sum of squares represents the variance explained by the independent variables. The higher the regression sum of squares, the more variance the model has in the dependent variable. Models with higher regression sums (e.g., 32.489 in Model 3) explain more variance than models with lower values (e.g., 16.006 in Model 2). Degrees of freedom (df): The degrees of freedom for the regression are 4, indicating that four independent variables are used in the models. Mean Square: The mean square is the regression sum of squares divided by the degrees of freedom. Model 3 has the highest mean square (8.122), indicating strong predictive power. F-Statistic (F-Value): F-Values measure the overall model fit, with higher values suggesting a stronger model. Model 4 has the highest F-Value (136.846), indicating the best explanatory power of all models. All models have very high F-Values, confirming the strong predictive power of the independent variables. Significance (p-Value or Sig.): The p-Values for all models are 0.000, which is highly significant ($p < 0.05$). This means that the independent variables contribute significantly to explaining the variance in the dependent variables.

TABLE 8. Factor Analysis Communalities

Communalities		
	Initial	Extraction
Risk Score	1.000	.895
Preparedness Level	1.000	.921
Response Time	1.000	.921
Economic Impact	1.000	.931

The Communalities table represents the proportion of variance in each variable that is accounted for by the extracted components in Principal Component Analysis (PCA). It helps determine how well each variable is explained by the principal components. Key Findings: Initial Communality: The initial value for all variables is 1.000, meaning that before extraction, 100% of the variance in each variable is considered. Extraction Communality: The extraction values indicate the proportion of variance retained by the principal components after extraction. Higher values (closer to 1) suggest that a variable is well represented by the extracted components, while lower values indicate weaker representation. Analysis of Variables: Risk Score (0.895): 89.5% of the variance in Risk Score is explained by the extracted components. This suggests that Risk Score is well captured by the principal components, with a small amount of variance remaining unexplained. Preparedness Level (0.921) and Response Time (0.921): Both variables have 92.1% of their variance accounted for, indicating a strong representation in the PCA model. Economic Impact (0.931): 93.1% of the variance is retained, making it the most strongly represented variable in the extracted components.

TABLE 9. Total Variance Explained

Total Variance Explained						
Component	Initial Eigenvalues			Extraction Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	3.669	91.716	91.716	3.669	91.716	91.716
2	.264	6.592	98.308			
3	.040	.990	99.298			
4	.028	.702	100.000			
Extraction Method: Principal Component Analysis.						

The Total Variance Explained table provides insights into the contribution of each principal component in explaining the variance in the dataset. This is essential in Principal Component Analysis (PCA) for determining how many components should be retained for meaningful data representation. Key Findings: Component 1 dominates the variance: The first principal component (PC1) has an Eigenvalue of 3.669 and explains 91.72% of the total variance in the dataset. This indicates that a single principal component captures almost all the variability, meaning the original variables are highly correlated. Component 2 contributes minimally: The second component has an Eigenvalue of 0.264 and accounts for only 6.59% of the variance, bringing the cumulative variance explained to 98.31%. While this component adds a small amount of additional information, it is relatively insignificant compared to PC1. Components 3 and 4 are negligible: Their Eigenvalues are below 1 (0.040 and 0.028, respectively), and they collectively contribute less than 1% of the total variance. These components do not add meaningful information and can be discarded in dimensionality reduction.

TABLE 10: Factor Analysis Communalities

Communalities		
	Initial	Extraction
Risk Score	1.000	.935
Response Time	1.000	.860
Economic Impact	1.000	.952
Casualty Risk	1.000	.913
Extraction Method: Principal Component Analysis.		

The communalities table represents how much of the variance in each variable is explained by the extracted components in the Principal Component Analysis (PCA). The initial communalities are all 1.000, indicating that each variable initially accounts for 100% of its variance before extraction. The extraction communalities represent the proportion of variance

retained in the extracted components. Risk Score (0.935): This high extraction value suggests that 93.5% of its variance is explained by the principal component, indicating a strong representation of this variable in the factor model. Response Time (0.860): With 86.0% of its variance explained, this variable is slightly less represented than the others but still holds a significant contribution. Economic Impact (0.952): The highest extracted communality, meaning 95.2% of its variance is explained, highlighting its strong association with the extracted factor. Casualty Risk (0.913): This variable also retains a high extraction value, with 91.3% of its variance explained, suggesting its importance in defining the principal component.

TABLE 11. Total Variance Explained

Total Variance Explained						
Component	Initial Eigenvalues			Extraction Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	3.660	91.496	91.496	3.660	91.496	91.496
2	.188	4.693	96.189			
3	.113	2.829	99.018			
4	.039	.982	100.000			
Extraction Method: Principal Component Analysis.						

The Total Variance Explained table provides insight into how much of the total variance in the dataset is accounted for by each principal component. The Initial Eigenvalues column represents the total variance each component initially explains, while the Extraction Sums of Squared Loadings column shows how much variance remains after extracting the principal component(s). Component 1 (Eigenvalue = 3.660, 91.496% of Variance): This component explains 91.50% of the total variance, indicating that most of the information in the dataset can be captured by a single factor. The high variance percentage suggests that the variables are strongly correlated and measure a common underlying construct. Component 2 (Eigenvalue = 0.188, 4.693% of Variance): The second component explains only 4.69% of the variance, which is a significant drop compared to the first component. Since its eigenvalue is below 1.0, it is not considered a meaningful factor in the Principal Component Analysis (PCA). Component 3 and Component 4 (Eigenvalues < 1.0): These components explain very little variance (2.83% and 0.98%, respectively) and do not contribute meaningfully to the overall data structure.

4. CONCLUSION

Statistical analysis provides insightful information about the connections and differences between important factors. such as risk score, preparedness level, response time, economic impact, and accident risk. Reliability analysis showed a high Cronbach's alpha and great internal consistency. values, indicating that the variables were well correlated and made a meaningful contribution to the dataset. Descriptive statistics showed that most variables had relatively high mean values, indicating a strong overall preparedness level and response capacity. Correlation analysis showed strong positive relationships between variables, especially between risk score and economic impact ($r = 0.957$), indicating that higher risk levels were closely associated with higher economic outcomes. Regression models revealed high R-squared values, demonstrating that the independent variables effectively predicted the dependent factors. ANOVA results confirmed that the models showed statistical significance ($p < 0.001$), enhancing the findings' dependability. Principal component analysis (PCA) suggests that a single principal component explains 91% of the variance, suggesting that the dataset can be effectively reduced while retaining essential information. This finding facilitates further analysis and decision-making by emphasizing the most influential factor. Overall, the study indicates a strong interaction between preparedness, risk, and impact factors, emphasizing the need for strategic risk management and economic planning. Future research could explore additional variables or external influences to further refine predictive models.

5. REFERENCE

[1]. Lotfi, Mojgan, Vahid Zamanzadeh, Leila Valizadeh, and Mohammad Khajehgoodari. "Assessment of nurse–patient communication and patient satisfaction from nursing care." Nursing open 6, no. 3 (2019): 1189-1196.

[2]. Yang, Hongwei. "The case for being automatic: introducing the automatic linear modeling (LINEAR) procedure in SPSS statistics." Multiple Linear Regression Viewpoints 39, no. 2 (2013): 27-37.

[3]. Jalolov, Tursunbek Sadriddinovich. "USE OF SPSS SOFTWARE IN PSYCHOLOGICAL DATA ANALYSIS." PSIXOLOGIYA VA SOTSIOLOGIYA ILMIY JURNALI 2, no. 7 (2024): 1-6.

[4]. Larson-Hall, Jenifer. A guide to doing statistics in second language research using SPSS and R. Routledge, 2015.

[5]. Kafle, Sarad Chandra. "Correlation and regression analysis using SPSS." The OCEM Journal of Management, Technology, and Social Sciences 1, no. 1 (2019): 125-132.

[6]. Rahman, Arifa, and Md Golam Muktadir. "SPSS: An imperative quantitative data analysis tool for social science research." International Journal of Research and Innovation in Social Science 5, no. 10 (2021): 300-302.
[7]. Preacher, Kristopher J., and Andrew F. Hayes. "SPSS and SAS procedures for estimating indirect effects in simple mediation models." Behavior research methods, instruments, & computers 36 (2004): 717-731.
[8]. Alferes, Valentim R., and David A. Kenny. "SPSS programs for the measurement of nonindependence in standard dyadic designs." Behavior Research Methods 41 (2009): 47-54.
[9]. Suresh, Devare. "Important of SPSS for social sciences research." Available at SSRN 2663283 (2015).
[10]. Espírito-Santo, Helena, and Fernanda Daniel. "Calcular e apresentar tamanhos do efeito em trabalhos científicos (2): Guia para reportar a força das relações." (2017).
[11]. Devi, B. Renuka, K. Rao, S. Setty, and M. Rao. "Disaster prediction system using IBM SPSS data mining tool." International Journal of Engineering Trends and Technology (IJETT) 4, no. 2 (2013): 3352-3357.
[12]. Bande, Swapnil, and Virendra V. Shete. "Smart flood disaster prediction system using IoT & neural networks." In 2017 international conference on smart technologies for smart nation (SmartTechCon), pp. 189-194. Ieee, 2017.
[13]. Wu, Che-I., Hsu-Yang Kung, Chi-Hua Chen, and Li-Chia Kuo. "An intelligent slope disaster prediction and monitoring system based on WSN and ANP." Expert Systems with Applications 41, no. 10 (2014): 4554-4562.
[14]. Kung, Hsu-Yang, Chi-Hua Chen, and Hao-Hsiang Ku. "Designing intelligent disaster prediction models and systems for debris-flow disasters in Taiwan." Expert Systems with Applications 39, no. 5 (2012): 5838-5856.
[15]. Yang, Lin, Chunrong Cao, Dehui Wu, Honghua Qiu, Minghui Lu, and Ling Liu. "Study on typhoon disaster loss and risk prediction and benefit assessment of disaster prevention and mitigation." Tropical Cyclone Research and Review 7, no. 4 (2018): 237-246.
[16]. Huang, Meng, Dian Zhang, Pan Li, YunHui Zhang, and RuoFei Zhang. "Study of earthquake disaster prediction system of Langfang city based on GIS." In IOP conference series: Materials science and engineering, vol. 224, no. 1, p. 012049. IOP Publishing, 2017.
[17]. Orozco, Michael M., and Jonathan M. Caballero. "Smart disaster prediction application using flood risk analytics towards sustainable climate action." In MATEC Web of Conferences, vol. 189, p. 10006. EDP Sciences, 2018.
[18]. Ayoub, N. A., A. A. Aziz, and W. A. Mustafa. "FloodIntel: advancing flood disaster forecasting through comprehensive intelligent system approach." Journal of Autonomous Intelligence 7, no. 1 (2023).
[19]. Boukerche, Azzedine, and Rodolfo WL Coutinho. "Smart disaster detection and response system for smart cities." In 2018 IEEE Symposium on Computers and Communications (ISCC), pp. 01102-01107. IEEE, 2018.
[20]. Kim, Dahee, Hee-sung Cha, and Shaohua Jiang. "The prediction of fire disaster using BIM-based visualization for expediting the management process." Sustainability 15, no. 4 (2023): 3719.
[21]. Kawasaki, Akiyuki, Akio Yamamoto, Petra Koudelova, Ralph Acierto, Toshihiro Nemoto, Masaru Kitsuregawa, and Toshio Koike. "Data integration and analysis system (DIAS) contributing to climate change analysis and disaster risk reduction." Data Science Journal 16 (2017): 41-41.
[22]. Kamati, Willbard, Valerianus Hashiyana, and James Mutuku. "Development of a near real-time early warning agricultural system for disaster prediction." Malaysian Journal of Computing (MJoC) 9, no. 1 (2024): 1706-1721.
[23]. Zhang, Xiang, Minghui Zhang, Xin Liu, Berhanu Keno Terfa, Won-Ho Nam, Xihui Gu, Xu Zhang et al. "Review on the progress and future prospects of geological disasters prediction in the era of artificial intelligence." Natural Hazards (2024): 1-41.
[24]. Javed, Rushba, Sidra Anwar, Khadija Bibi, M. Usman Ashraf, and Samia Siddique. "Prediction and monitoring agents using weblogs for improved disaster recovery in cloud." Int J Inf Technol Comput Sci 11, no. 4 (2019): 9-17.
[25]. Gogic, Asmir, Aljo Mujcic, Sandra Ibric, and Nermin Suljanovic. "Performance analysis of bluetooth low energy mesh routing algorithm in case of disaster prediction." Int. J. Comput. Electr. Autom. Control Inf. Eng 3 (2016): 1075-1081.
[26]. Singh, Jyoti Prakash, Yogesh K. Dwivedi, Nripendra P. Rana, Abhinav Kumar, and Kawaljeet Kaur Kapoor. "Event classification and location prediction from tweets during disasters." Annals of Operations Research 283, no. 1 (2019): 737-757.
[27]. Khalid, Mohamad Sukeri Bin, and Shazwani Binti Shafiai. "Flood disaster management in Malaysia: An evaluation of the effectiveness flood delivery system." International Journal of Social Science and Humanity 5, no. 4 (2015): 398.
[28]. Koshimura, Shunichi. "Establishing the advanced disaster reduction management system by fusion of real-time disaster simulation and big data assimilation." Journal of disaster research 11, no. 2 (2016): 164-174.
[29]. Tanaka, Kohji, Eisaku Yura, Masatoshi Sasaki, Takuya Shirahase, Akio Shimokawa, and Sho Kato. "Development of a Typhoon Search and Prediction System for Disaster Prevention and Mitigation Action Plans Based on Typhoon Course Analysis." Procedia Engineering 154 (2016): 1258-1266.
[30]. Cunha, João Carlos, Mário Zenha Rela, and João Gabriel Silva. "On the use of disaster prediction for failure-tolerance in feedback control systems." In Proceedings International Conference on Dependable Systems and Networks, pp. 123-132. IEEE, 2002.
[31]. Li, Haoran, Jun Sun, and Ke Xiong. "AI-Driven Optimization System for Large-Scale Kubernetes Clusters: Enhancing Cloud Infrastructure Availability, Security, and Disaster Recovery." Journal of Artificial Intelligence General science (JAIGS) ISSN: 3006-4023 2, no. 1 (2024): 281-306.

