MACHINE LEARNING-BASED RISK PREDICTION AND SAFETY MANAGEMENT FOR OUTDOOR SPORTS ACTIVITIES

YAN LU

Abstract. Participant safety is becoming increasingly important as outdoor sports activities gain popularity. A machine learning-based strategy for risk assessment and safety control in outdoor sports activities is presented in this paper. Our framework uses predictive modelling, sophisticated algorithms, and historical data analysis to identify potential dangers and improve safety procedures. It also considers participant profiles and environmental conditions. Comprehensive testing and validation are used to examine the model’s efficacy, showing that it can offer risk evaluations in real-time and support preventive safety measures. Our approach entails placing sensor-based Internet of Things (IoT) devices at building sites to gather extremely detailed temporal and geographic weather, building, and labour data. This data is then cooperatively used on the edge nodes to train Deep Neural Network (DNN) models in a cross-silos way. The present study makes a valuable contribution to sports safety by offering a clever approach that integrates technology and outdoor leisure to ensure participants have a safe and pleasurable experience. The experiment’s outcomes show how well the suggested strategy works to increase the adoption of construction safety management systems and lower the likelihood of future mishaps and fatalities. As a result, the system has improved speed and responsiveness, an important feature for time-sensitive applications like safety prediction.

Key words: machine learning, risk prediction, safety management, sports, outdoor sports activities

1. Introduction. The promise of exploration and excitement draws people to outdoor sports activities, which include mountaineering, biking, hiking, and water sports. The increasing demand for thorough risk assessment and safety protocols to guarantee the welfare of participants corresponds with the growth in popularity of these activities. A ground-breaking way to improve risk assessment and safety procedures is through the incorporation of machine learning, which acknowledges the dynamic nature of outdoor situations and the inherent uncertainties they present.

The building industry has been at the forefront of this rapid development of the world in recent decades. With 200,000 more people moving into cities every day, it is evident that these demographic changes have had a significant impact on the worldwide building industry [18]. Nonetheless, construction is regarded as one of the most hazardous industries for workers because of its dynamic, ever-changing, and heterogeneous spatiotemporal environment. Worker safety is a persistent problem that calls for constant focus and effort. Due to the dangerous working circumstances at construction sites, a recent study suggests [11] that workers routinely face possible safety and health concerns during the building process.

Data analysis reveals that, broadly speaking, "outdoor sports" relate to all outdoor activities, which includes practically all sports [23]. In a restricted sense, outdoor sports are those that take place in naturally occurring outdoor settings, such as parks, buildings intended for other uses besides sports, or natural settings. A category of sporting activities known as outdoor sports use the outdoors as a non-designated location and are characterized by an element of adventure or experience [26, 29]. Its primary expression is to leave the city, venture outside, and partake in activities that provide certain risks, difficulties, and relevance while adhering to safety and standards guidelines.

The main motivation of this research stems from the growing recognition of the importance of participant safety in the increasingly popular domain of outdoor sports activities. As these activities attract a larger and more diverse group of enthusiasts, the complexity and variability of safety risks associated with outdoor environments also escalate. This paper introduces a machine learning-based strategy designed to enhance risk assessment and safety management within this context. Leveraging the power of predictive modeling, advanced
Machine Learning-based Risk Prediction and Safety Management for Outdoor Sports Activities

Our approach aims to proactively identify potential hazards and refine safety protocols tailored to the unique demands of outdoor sports. Unlike indoor sports, which have stricter site requirements and are heavily impacted by weather and terrain, outdoor sports are not the same. They have a greater relevance for individuals to reduce mental stress, improve their health, and raise their standard of living, in addition to helping city dwellers escape the bustle and get closer to nature. A few pertinent policies have been released in recent years, including the State Council's views on encouraging the growth of the health services industry, the General Office of the State Council's guidelines on accelerating the sports industry's development, and the notice on the guidelines for expediting the establishment of a social security system and services system for the disabled. China's outdoor products sector has taken shape and begun to develop fast by the start of the twenty-first century [32, 20]. Unlike other demanding sports, which are not only simple to learn, safe, and efficient, but also simple to practice, outdoor sports.

The main contribution of the proposed method is given below:

1. DNNs are particularly good at finding complex patterns in large, heterogeneous datasets.
2. When it comes to outdoor sports, where dangers can take many different forms depending on a range of factors like weather, topography, and participant behavior, DNNs help by quickly identifying intricate patterns that lead to more precise risk evaluations.
3. Real-time risk prediction is made possible by utilizing DNNs' innate capacity for parallel processing.
4. Instantaneous risk evaluations that dynamically adjust to changing conditions during outdoor activities are provided by these networks, which are capable of quickly analysing continuously evolving environmental data.

The rest of our research article is written as follows: Section 2 discusses the related work on various sports activities, risk prediction and deep learning methods. Section 3 shows the algorithm process and general working methodology of proposed work. Section 4 evaluates the implementation and results of the proposed method. Section 5 concludes the work and discusses the result evaluation.

2. Related Works. Numerous research on the general population have confirmed a considerable negative correlation between psychological well-being and distress [9, 13, 21]. More specifically, concerning findings from several studies [25, 6, 14, 31, 19] on the mental health of academic students have shown a decline in the perception of life quality and an exaggerated rise in the frequency and severity of these psychological issues. Physical activity (PA) has long been linked to a lower risk of death and morbidity from degenerative and chronic illnesses [30, 22, 7, 24, 4, 15], but more recently, research has focused on the impact of PA on mental health. PA has been linked to improvements in mood, overall well-being, and quality of life perception [16, 12], as well as a notable decrease in depressed and anxious symptomatology [5]. Numerous biological mechanisms, such as enhanced cerebral blood flow and oxygen delivery to brain tissues, decreased muscle tension, and elevated serum concentrations of endocannabinoid receptors and satisfying neurotransmitters like serotonin, have been proposed as explanations for this evidence [1].

But research on the effects of PA in older age groups has yielded the most consistent results linking PA to improved mental health [10]. Research on PA in younger age groups, however, has shown mixed results, with some indicating a weak association [11] and others suggesting a more persistent association [8] between PA and mental health outcomes. The use of measuring tools that, when used alone, do not provide a complete assessment of all facets of mental health has been blamed in part for this lack of reliable data. This implies that using a variety of instruments to obtain a more accurate assessment of mental health perception can be beneficial [3].

In the realm of deep learning research, DNN has emerged as a well-known algorithm that makes it possible to create intelligent applications across a variety of industries [28]. An Artificial Neural Network (ANN) known as a DNN is made up of several layers of connected nodes, or neurons, that process input data and gradually extract progressively more abstract properties from it. A deep neural network (DNN) is a machine learning model that is well-suited for tasks like picture and audio recognition [27], natural language processing [2], and predictive modelling [17], since it can learn hierarchical representations of complicated patterns and relationships in the data. Backpropagation is used by DNN to minimize the discrepancy between target values and anticipated outputs by adjusting the weights across neurons.
One significant issue identified is the reliance on measurement tools that may not comprehensively assess all facets of mental health. This limitation could lead to inconsistent data on the effects of PA on mental health, particularly in younger age groups. The passage suggests that employing a variety of instruments could yield a more accurate assessment of mental health perceptions, indicating a need for methodological improvements in research.

3. Proposed Methodology. A strong and proactive strategy is needed to guarantee the safety of participants in outdoor sports. The suggested approach makes use of machine learning techniques to improve outdoor enthusiasts’ safety management by dynamically predicting dangers. The approach consists of multiple crucial phases that combine data gathering, model building, and real-time implementation to produce an all-encompassing solution. Initially, the data is collected and then the collected data is pre-processed. Next, the pre-processed data is given to the feature engineering process. Finally, risk prediction is carried out using deep neural networks (DNN). In figure 3.1 shows the architecture of the proposed method.

By carefully selecting and engineering features that capture the essence of outdoor sports environments, such as weather conditions, terrain types, and athlete biometrics, the model can better understand the context of the data it processes. This helps in accurately interpreting variations in the input data. The model employs ensemble learning techniques, which combine the predictions from multiple learning algorithms to improve generalizability and robustness. This approach helps manage data’s unpredictability by leveraging the strengths of various models to produce a more accurate and stable prediction.

3.1. Data collection and pre-processing. A multifaceted strategy is required to gather pertinent data for risk and safety prediction related to outdoor activities, including participant traits, environmental conditions, and historical event data.

For information on current weather conditions, such as temperature, precipitation, wind speed, and atmospheric pressure, consult your local weather station. To learn more about the topography, terrain, and elevation of the outdoor activity area, consult geospatial databases. Utilize satellite imagery to evaluate vegetation, water bodies, and land cover as well as dynamic changes in environmental circumstances. Install Internet of Things (IoT) gadgets and on-site sensors to collect environmental data in real time. Examples of these are GPS trackers, humidity sensors, and temperature sensors. Use remote sensing technologies to collect high-resolution information on the topography and environmental aspects, such as drones carrying sensors.

3.1.1. Data Pre-processing. Preparing gathered data for use in machine learning models for risk and safety prediction in outdoor sports involves pre-processing it. To manage missing values, standardize the data, create features, and prepare the data for training and testing the predictive models, pre-processing processes are necessary. Determine which values in the gathered data are missing and deal with them by either deleting the relevant entries or imputing the necessary values (such as the mean, median, or interpolation).
3.2. Feature Engineering. Utilize environmental data to extract pertinent parameters like height, wind speed, rainfall, temperature, and terrain kind. Transform time-related data (date, time of day, etc.) into suitable forms or create new time-related data (season, time of day, etc.) that could affect the weather outside. Convert data from participants into features while taking age, health, skill, and past involvement information into account. Provide binary or categorical variables for participant attributes, including experience level or health issues, that may have an impact on safety. Determine important characteristics, like incident type, location, contributing variables, and severity, from past incident data. To identify potential patterns, engineer temporal characteristics (e.g., time of day, day of week, season) associated to incident incidence.

3.2.1. Normalization. A popular data normalizing method in machine learning, min-max normalization (also called feature scaling or min-max scaling) converts numerical characteristics into a predetermined range. By ensuring that every feature has a comparable scale, this normalization helps to avoid certain characteristics predominating over others when the model is being trained. A feature’s values are scaled via min-max normalization to a range of 0 to 1.

$$p_i = \frac{(q_i - \min(q))}{(\max(q) - \min(q))}$$  

3.3. Risk Prediction and Safety Management for Outdoor Sports Using DNN methods. Using Deep Neural Networks (DNNs) has the potential to transform risk prediction and safety management in outdoor sports, where conditions are often dynamic and unpredictable. DNNs are a smart way to improve safety procedures and guarantee the welfare of participants because of their ability to identify intricate patterns and relationships within data.

3.3.1. Risk Prediction Using DNN. A family of artificial neural networks (ANNs) known as deep neural networks (DNNs) are distinguished by having numerous layers between the input and output layers. These networks can identify complicated patterns and characteristics in large datasets since they are built to learn hierarchical representations of data by utilizing numerous layers. Deep neural networks (DNNs) have shown impressive performance in a few domains, such as natural language processing, picture, and audio recognition, and, more recently, risk prediction and safety management for outdoor sports.

The first layer that gets the data as input in its raw form. Every node in this layer stands for a characteristic or feature of the incoming data. Hierarchical feature extraction from input data is learned by the network at the layers that sit between the input and output layers. Multiple hidden layers are characteristic of deep networks, which allow them to catch intricate patterns, parameters related to the connections made by nodes in various tiers. To maximize the network’s performance, these parameters are changed throughout the training phase. The model becomes non-linear when non-linear functions are applied to each layer’s node’s output. Rectified Linear Unit, or ReLU, and sigmoid are examples of common activation functions.

The last layer that generates output for the network. Depending on the job (binary classification, multi-class classification, regression, etc.), this layer has a different number of nodes. a measurement of the discrepancy between the intended and actual output. The objective of the training process is to reduce this loss function. a method that minimizes the loss function by modifying the weights and biases. Optimization methods like Gradient Descent and its variations (like Adam and RMSprop) are frequently utilized. In figure 3.2 shows the structure of DNN.

3.3.2. Safety Management using LSTM. Recurrent neural networks (RNNs) with specialized memory cells are used in Long Short-Term Memory (LSTM) networks for outdoor sports safety management. This allows RNNs to capture temporal connections in data. Because LSTMs perform exceptionally well with data sequences, they can be used for jobs involving time-series information, including risk prediction in outdoor sports scenarios.

Sequences of input that reflect participant and environment characteristics. The data’s temporal dependencies and patterns are captured by many LSTM layers. Predicting the safety or risk status for the upcoming time step is done via the output layer. When engaging in outdoor activities, connect the LSTM model to real-time sensor data to continuously monitor the surrounding conditions. Make the model more deployable on mobile apps so that consumers may receive safety forecasts while they’re on the road. Make real-time risk predictions using the LSTM model by considering participant characteristics and the state of the environment.
Put into practice adaptive safety protocols that, in response to changing hazards, modify dynamically based on LSTM predictions.

\[
I_{n}^{t} = f(W_{e_{i_{x}}x^{t}} + We_{i_{h}}h^{t-1} + We_{i_{c}}C^{t-1} + b_{i}) \tag{3.2}
\]

\[
F_{O}^{t} = f(We_{f_{o_{x}}}x^{t} + We_{f_{o_{h}}}h^{t-1} + We_{f_{o_{c}}}C^{t-1} + b_{f_{o}}) \tag{3.3}
\]

\[
C_{E}^{t} = F_{O}^{t} . C_{E}^{t-1} + I_{n}^{t}(We_{c_{e_{x}}}x^{t} + We_{c_{e_{h}}}h^{t-1} + We_{c_{e_{c}}}C^{t-1} + b_{c_{e}}) \tag{3.4}
\]

\[
O_{P}^{t} = f(We_{o_{p_{x}}}x^{t} + We_{o_{p_{h}}}h^{t-1} + We_{o_{p_{c}}}C^{t-1} + b_{o_{p}}) \tag{3.5}
\]

\[
h^{t} = O_{P}^{t} \cdot g(C_{E}^{t}) \tag{3.6}
\]

Create a feedback loop where user input and incident reports help the LSTM model learn and improve over time. To be relevant, the LSTM model should be updated on a regular basis depending on fresh data and new trends.

4. Result Analysis. The proposed method DNN-LSTM for risk prediction and safety management using various metrics such as accuracy, f1-score, precision, and Kappa value.

When assessing the effectiveness of machine learning models, such as those employed in outdoor sporting activities for risk prediction and safety management, accuracy is a regularly utilized indicator. When selecting evaluation metrics, it is crucial to consider the objectives of your model as well as the particular features of your dataset.

Although accuracy offers a broad indication of a model’s soundness, it may not always be the best statistic, particularly when working with unbalanced datasets or when certain errors are more serious than others. Other metrics, including as precision, recall, and F1-score, may provide more useful information when it comes to risk prediction and safety management. In figure 4.1 shows the evaluation of accuracy.

Particularly when it comes to outdoor sporting activities, precision is a crucial evaluation criterion for machine learning-based risk prediction and safety management. When minimizing false positives—that is, lowering the number of times the model predicts a safety issue incorrectly—precision becomes especially important. The ratio of true positive predictions to all positive predictions (true positives plus false positives) is known as precision. In figure 4.2 shows the evaluation of precision.

A popular metric for assessing how well classification models perform is the F1-score, which is especially helpful for imbalanced datasets. You can use the F1-score to evaluate how effectively your model balances precision and recall in the context of risk prediction and safety management for outdoor sports activities. A high precision indicates a high probability of accuracy when the model predicts a favorable outcome (risk or safety concern). This is essential to prevent taking needless safety precautions when they are not necessary.

A high recall shows that real-world positive examples are well captured by the model. High recall guarantees that a sizable percentage of possible hazards are identified by the model in the context of safety management.
Finding a balance between recall and precision is aided by the F1-score. It guarantees that the model in safety management is thorough in capturing hazards and accurate in its forecasts. In figure 4.3 shows the evaluation of F1-score.

The Cohen’s kappa, often known as the kappa statistic, is a regularly employed metric in classification tasks to evaluate the degree of agreement between anticipated and actual classifications. It accounts for chance agreement in inter-rater agreements. The agreement between expected risk levels and actual occurrences can be assessed in the context of risk prediction and safety management for outdoor sports activities using the Kappa statistic. In figure 4.4 shows the evaluation of Kappa Value.

5. Conclusion. As outdoor sporting activities become more popular, participant safety is becoming increasingly critical. This research presents a machine learning-based approach to risk assessment and safety regulation in outdoor sports. Our approach improves safety procedures by identifying possible hazards through the use of sophisticated algorithms, predictive modelling, and historical data analysis. It also takes the surroundings and participant profiles into account. The effectiveness of the model is investigated through extensive testing and validation, demonstrating that it can provide risk assessments in real-time and assist with preventive safety actions. Our methodology involves the deployment of sensor-based Internet of Things (IoT) devices at construction sites to collect incredibly fine-grained temporal and spatial building, labour, and weather data. This data is then collaboratively used in a cross-silos fashion to train Deep Neural Network (DNN) models on the edge nodes. The current study adds much to the field of sports safety by providing a novel strategy that
combines outdoor recreation and technology to guarantee participants’ enjoyment and safety. The experiment’s results demonstrate the effectiveness of the recommended approach in promoting the use of construction safety management systems and reducing the risk of accidents and fatalities in the future. This enhances the system’s speed and responsiveness, a crucial attribute for time-sensitive applications such as safety forecasting.

REFERENCES


Machine Learning-based Risk Prediction and Safety Management for Outdoor Sports Activities


Edited by: Rajanikanth Aluvalu

Special issue on: Evolutionary Computing for AI-Driven Security and Privacy

Accepted: Feb 9, 2024