SCIENTIFIC OASIS

Decision Making: Applications in Management and Engineering

ISSN: 2560-6018, eISSN: 2620-0104

HEAT STROKE PREDICTION USING FUSED MACHINE LEARNING WITH CASCADED LEVY FLIGHT OPTIMIZATION

Mohamad Emad Bitar
Ph.D. Scholar, Department of Computer Science,
CMS College of Science & Commerce, Bharathiar University,
Coimbatore, India - 641035
t22h.12345@gmail.com

Dr. V. Sujatha
Vice Principal,
CMS College of Science and Commerce, Bharathiar University,
Coimbatore, India – 641035
sujatha.padmakumar4@gmail.com

Abstract

Heat stroke is a severe condition resulting from prolonged exposure to high temperatures, posing significant health risks, particularly during extreme weather events. Accurate prediction of heat stroke is challenging due to the complex interplay of environmental and physiological factors. This study proposes a comprehensive machine learning framework to address this challenge effectively. We begin by ensuring data quality through normalization using a Reorder Iterative Imputer, which handles missing values and outliers with precision. Feature selection is then performed using a combination of Principal Component Analysis (PCA) and t-Distributed Stochastic Neighbor embedding (t-SNE) to identify the most relevant variables from a diverse set of indicators. The classification task is executed using a fused machine learning approach, integrating Random Forest, Support Vector Classifier, and Gradient Boosting methods to enhance prediction accuracy. Further, prediction optimization is achieved using a Cascaded Levy Flight Optimization algorithm, fine-tuning model parameters for superior performance. The proposed method demonstrates significant improvements in prediction accuracy and reliability over traditional approaches, offering a valuable tool for early intervention and preventive measures in health monitoring and safety management.

Keywords: Classification, Cascaded Levy Flight Optimization Data Normalization, Heat Stroke Prediction, Machine Learning

I. Introduction

Heat stroke is a severe medical emergency characterized by an elevated body temperature exceeding 40°C (104°F) due to prolonged exposure to high environmental temperatures, often coupled with high humidity [1]. It represents a critical threat to health, particularly during heatwaves or in environments where adequate cooling measures are lacking [2]. The condition can lead to multiple organ failure and, if untreated, can result in death [3]. Early detection and prediction are crucial for mitigating the risks associated with heat stroke and implementing timely interventions [4]. Predicting heat stroke involves analyzing a complex interplay of environmental and physiological variables, including temperature, humidity, individual health conditions, and physical activity levels [5]. Traditional methods of heat stroke prediction have largely relied on empirical rules or basic statistical approaches, which cannot capture the nuanced relationships between these factors [6]. Additionally, real-time monitoring and prediction require integrating diverse data sources and handling incomplete or noisy datasets [7].

Machine learning (ML) techniques offer a promising alternative to traditional methods by enabling data-driven insights and predictive capabilities [8-10]. With both past and present data as inputs, ML systems can sift through mountains of information, find complex patterns, and provide reliable forecasts [11, 12]. By using modern data processing and analytical approaches, these strategies can improve the accuracy of heat stroke predictions [13–14]. In this study, we address the challenge of handling missing or inconsistent data by employing the Reorder Iterative Imputer for data normalization [15-16]. This approach ensures that the dataset is appropriately prepared for ML models by iteratively refining the imputation of missing values and outliers [17-18]. Proper data normalization is essential for improving the quality of input features and enhancing the performance of machine learning algorithms [19].

The main contribution of the paper is:

- Data normalization using reorder iterative imputer
- Feature selection using PCA with t-SNE
- Classification using fused ML
- Prediction using cascaded levy flight optimization

Here is the table of contents for the remaining portion of the piece. Section 2 has many authors discussing various heat stroke prediction approaches. The proposed model is laid forth in Section 3. Section 4 provides a summary of the investigation's results. Section 5 concludes with an analysis of the results and suggestions for moving forward.

1.1 Motivation of the paper

Heat stroke remains a critical public health concern, especially in the face of increasing global temperatures and more frequent extreme weather events. Despite advancements in health monitoring, accurately predicting heat stroke onset remains a challenge due to the intricate relationship between various environmental and physiological factors. The motivation for this study stems from the need to develop a robust and reliable prediction system that can anticipate heat stroke events with high accuracy, enabling timely interventions and potentially saving lives. By utilizing advanced machine learning techniques, this study aims to overcome the limitations of

traditional methods, offering a comprehensive approach that integrates data normalization, feature selection, and optimization to enhance prediction capabilities.

II. Background study

Akbar, M. et al. [1] these authors research indicated that the SVM model best predicts heat shock proteins, achieving 87% accuracy owing to the increased sampling strategy. Additionally, the model of research can be utilized to locate heat shock proteins in eukaryotes with higher accuracy. However, in-depth investigation and development of advanced machine learning models was the need of the hour to discover the proteins that cannot be tracked using this technique

Hirano, Y. et al. [3] The authors' research was the first to demonstrate the promise of heat-related sickness prediction models grounded on machine learning. Nevertheless, more research was necessary to enhance the performance quality via larger sample sizes, the inclusion of critical factors, and clinical prospective validation.

Ke, D. et al. [5] A dependable method for predicting the daily volume of heat-related ambulance calls was developed by the author. Developing regional-level models was made necessary by very high forecast accuracy, while the national-level model showed great prediction accuracy applicable to several sites. When these writers accounted for heatwave features like optimum temperature, cumulative heat stress, and heat acclimatization, their forecasts were much more accurate. The national model now has a far better ability to predict the frequency of heat-related ambulance calls in most locations with low computational effort, since the addition of heatwave features raised its modified R2 from 0.9061 to 0.9659.

Li, H. et al. [7] Combinations of multi-spectrum photos enhanced the prediction performance of DNNs, and AF patients' multi-spectrum fundus images alone might be enough to predict the likelihood of secondary WAS. It was beneficial to get several spectral fundus pictures as they can reveal distinct microscopic characteristics of pathology.

Nissa, N. et al. [9] Comparing several ML algorithms for early-stage CVD prediction was the primary contribution of this work. Preprocessing methods were used to enhance the dataset's quality. The primary focus was on handling corrupted and missing information, as well as eliminating outliers. The author also compared the outcomes of the three machine learning algorithms the author used to forecast the illness using a number of statistical metrics.

Priyadarshini, T. et al. [13] The goal of this study was to create an intelligent prediction model that could evaluate patients' vulnerability to heart attacks by analyzing a dataset of heat stroke symptoms. A prediction system was constructed by the author using K-means clustering in conjunction with Naïve Bayes, Decision Tree, and Artificial Neural Network classifiers.

Statsenko, Y. et al. [15] The incidence, severity, and early result of HS were shown to be associated with several meteorological conditions in this research. Typically, the risk ratio of HS tends to grow in the days after a notable change in AT, humidex, or AP. Even after the environment has adjusted, the danger can be high. The fact that Humidex outperformed AT as a predictor suggests that the two factors should be taken into account in tandem. To make reliable forecasts,

it was necessary to combine and evaluate all available meteorological data from the last several days.

Yu, H. et al. [19] The author investigated the association between ischemic stroke (WAS) patient clinical characteristics, immune cell infiltration, and inflammatory factor release and the expression of genes involved in the Unfolded Protein Response (UPR) pathway. A tight correlation between these variables was shown by these authors examination of gene microarray data obtained from blood samples.

Table 1: Survey of Machine Learning Models for Heat-Related Health Predictions

Author	Year	Methodology	Advantage	Limitation
Hirano et al.	2021	Machine learning-based mortality prediction model for heat-related illness	Provides accurate mortality prediction using real-time data	Can require extensive computational resources for training and updating the model
Ke et al.	2023	Machine learning models for predicting heat-related ambulance calls based on heatwave features	Enhances prediction accuracy by considering specific heatwave characteristics	Limited by the quality and granularity of the input weather data
Ogata et al.	2021	Machine learning	Utilizes a comprehensive data set for improved prediction accuracy	Cannot be easily generalizable to other regions with different climatic conditions
Shimazaki et al.	2022	Supervised machine learning	Customizes prevention strategies based on individual health data	Requires continuous monitoring and data collection, which can be impractical in some settings
Xu et al.	2024	machine learning	Integrates multiple data sources for	Potentially high computational cost and

			robust predictions	complexity in integrating various data sources

2.1 Problem definition

Heat stroke, caused by prolonged exposure to high temperatures, poses significant health risks, particularly during extreme weather events. Accurately predicting heat stroke onset is challenging due to the complex interplay of environmental and physiological factors. Traditional methods often fall short in handling this complexity and providing timely predictions. In order to provide a more dependable tool for early diagnosis and preventative actions, this research utilizes sophisticated machine learning approaches to examine varied indications, improve data quality, and optimize model performance. The goal is to answer the requirement for increased prediction accuracy.

III. Materials and methods

In this section, we outline the proposed methods for developing a robust heat stroke prediction system. The approach begins with data preprocessing, including normalization through a Reorder Iterative Imputer to handle missing values and outliers effectively. Feature selection is performed using a combination of PCA with t-SNE to identify the most relevant predictors. For classification, a fused machine learning approach is employed, integrating Random Forest, Support Vector Classifier, and Gradient Boosting models to enhance accuracy. Finally, the model's performance is optimized using the Cascaded Levy Flight Optimization algorithm, which fine-tunes the parameters to achieve superior predictive accuracy.

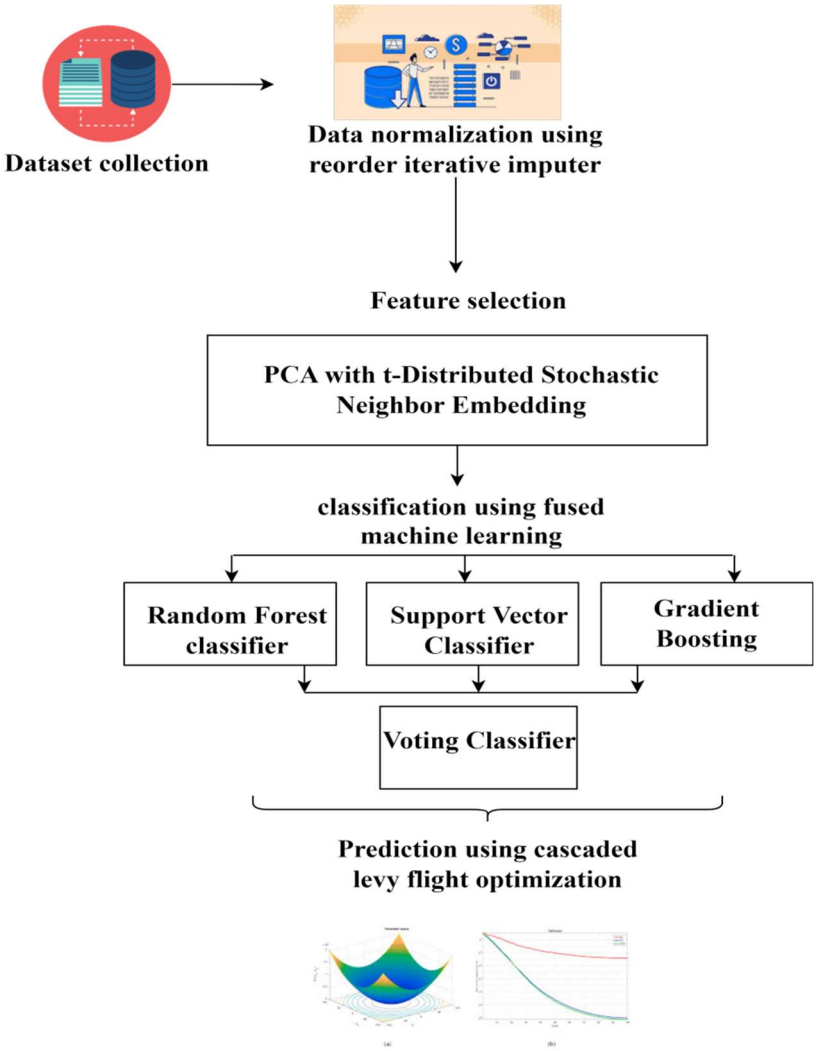

Figure 1: Heat stroke prediction workflow architecture

3.1 Dataset collection

The dataset used in this study was sourced from Kaggle, a popular platform for data science and machine learning resources. Specifically, the dataset is available at https://www.kaggle.com/datasets/tahiatazin1510997643/heat-stroke. It includes detailed records of heat stroke incidents with a variety of environmental and physiological variables, which are crucial for predicting the onset of heat stroke.

3.2 Data normalization using reorder iterative imputer

The Reorder Iterative Imputer is a technique for data normalization that addresses missing values and outliers in datasets. It operates iteratively to estimate and impute missing data by utilizing relationships among features and adjusting for data distribution irregularities. This method reorders the imputation process based on feature importance and data patterns, ensuring more accurate and reliable normalization. By refining the imputation in successive iterations, the Reorder Iterative Imputer enhances data quality and prepares the dataset for effective analysis and modeling.

Imputer makes use of generative iteration. With each iteration of the generative process, Imputer generates a new alignment by applying criteria to a partly formed one. Although a full alignment is generated at the end of each generating phase, inference uses a smaller selection to produce subsequent partial alignments. As a consequence of an iterative refining process, this subset usually contains the tokens that were previously emitted. Imputer is able to describe conditional dependencies across generation phases via its iterative approach, which is not possible with CTC. The Imputer assumes that token predictions are conditionally independent in order to parameterize the alignment distribution.

$p_\theta(a|a^-, x) = \prod_i p(a_i|a^-, x; \theta)$ ---------- (1)

with $(a_i|a^-, x; \theta)$, where $p_\theta(a|a^-, x)$ and ∅ is the masked out character, and a˜ is some earlier alignment. According to BERT, the mask token specification is consistent with the ∅ mask token. In order to be clear, a˜ can be empty (for example, all tokens are hidden) and ai = ˜ai∀a˜i 6= empty (meaning that once a token is committed in a˜, a must remain consistent). Imputer has the ability to simulate conditional dependencies among generation stages by conditioning on a˜, and it can enable parallel generation inside a generation step by assuming conditional independence between new token predictions.

3.3 Feature selection and classification using fused machine learning

Feature selection and classification using fused machine learning involve combining multiple algorithms to enhance predictive performance. Feature selection identifies the most relevant variables from the dataset using techniques t-Distributed Stochastic Neighbor Embedding to improve model efficiency. For classification, a fused machine learning model integrates predictions from various classifiers, such as Random Forest, Support Vector Classifier, and Gradient Boosting, through methods like a Voting Classifier. This approach aggregates the strengths of each model, improving overall accuracy and robustness by utilizing diverse learning techniques and focusing on key features.

3.3.1 PCA with t-Distributed Stochastic Neighbor Embedding

Feature extraction using Principal Component Analysis (PCA) combined with t-Distributed Stochastic Neighbor Embedding (t-SNE) involves a two-step dimensionality reduction process. First, PCA is used to reduce the high-dimensional data to a lower-dimensional space by identifying the principal components that capture the maximum variance in the data. This initial reduction simplifies the data, removing noise, and making it more manageable for further analysis. For instance, features such as **gender, Nationality, and Daily ingested water (L)** might be among those reduced during PCA, as they have high importance in the data. Next, t-SNE is applied to the PCA-reduced data to further reduce its dimensionality to two or three dimensions while preserving the local structure and relationships between data points. This combination allows for effective visualization of complex data structures, highlighting patterns and clusters that might not be apparent in higher dimensions, such as the interactions between **Strenuous exercise, Rectal temperature (deg C),** and **Environmental temperature (C)**.

The result will be inaccurate classifications and a lack of generalizability. By eliminating unnecessary predictors, t-Distributed Stochastic Neighbor Embedding successfully reduces model

complexity. Many machine learning algorithms can readily include t-SNE as their main feature selection technique due to its status as a wrapper strategy that primarily utilizes filter feature selection. Coefficient or feature significance is then used to rank the characteristics according to their relevance. It re-fits the model by removing the worst feature(s) one by one. The process is repeated until the number of features reaches a certain threshold.

$Rank_i = \{r_{i1} = 1, r_{i2} = 2, \dots, r_{ip} = p\}$ -------- (2)

Next, we use eight distinct t-Distributed Stochastic Neighbor Embedding techniques to determine the feature cut-off points. Most individuals think these are the most important characteristics to have. Accordingly, this technique selects |αP| features from each feature subset in order to construct each optimal feature subset.

$FS_i^{opt} = \{f_{i1}, f_{i2}, \dots, f_i, |\alpha P|\}$ ---------- (3)

Where the round-down operator is represented in mathematics by $|\alpha P|$

Doing so will allow us to remove features that lack robustness and accuracy, such as potentially **Weight (kg) or Sickle Cell Trait (SCT)**, depending on the specific dataset. To narrow it down, picture a situation where the parameter τ exceeds the area under the curve (AUC) of the top N feature sets for predictive classification. This is the reason why they are known as:

$fi^{opt} = \{FS_1^{opt}, FS_2^{opt}, \dots, FS_N^{opt}\}$ ---------- (4)

Finally, we state that the robust biomarker screening problem is an unstable combination problem with N features. For stability, all possible combinations of the sets in FS_2^{opt} opt are tested.

Algorithm 1: PCA with t-Distributed Stochastic Neighbor Embedding

Input:

- High-dimensional dataset X with n samples and d features
- Number of principal components k for PCA
- Perplexity parameter for t-SNE
- Maximum number of iterations for t-SNE

Steps:

- Normalize the dataset X to have zero mean and unit variance.
- Compute the covariance matrix $\sum = \frac{1}{n-1} X^T X$
- Perform eigenvalue decomposition on Σ to get eigenvalues and eigenvectors.
- Sort eigenvectors by the magnitude of their corresponding eigenvalues in descending order.
- Select the top k eigenvectors to form the projection matrix$W$
- Project the original dataset X onto the lower-dimensional space: $X_{PCA} = XW$
- Initialize the low-dimensional representation Y of X_{PCA} randomly.
- Calculate pairwise similarities P_{ij} in the high-dimensional space using Gaussian distribution.
- Calculate pairwise similarities Q_{ij} in the low-dimensional space using Student's t-distribution.

Output:

- Selected Feature List (Time of day, Nationality, Sweating, age, temperature cardiovascular, water with 15 attributes)

3.2.2 Random Forest classifier

During training, Random Forest builds many decision trees and then uses the mean prediction (regression) or mode of their classes (classification) as its final output referred by Eldora, K. et al. (2024). By merging the predictions of many trees, each constructed from a unique subset of the data and attributes, it enhances accuracy and resilience. By averaging out the biases of individual trees, this technique improves generalization and decreases overfitting.

$mg(X,Y) = av_k I(h_k(X) = Y)$ ---------- (5)

I(•) is the indicator function in this case. How much the appropriate class's average vote at X, Y surpasses the average vote for any other class is what the margin measures. We can have greater faith in the categorization if the margin is bigger. The generalization mistake can be expressed as

$PE^* = P_{X,Y}(mg(X,Y) < 0)$ ---------- (6)

3.2.3 Support Vector Classifier

Machine learning's Support Vector Classifier (SVC) seeks for the feature space's ideal hyperplane for classifying data referred by Ke, D. et al. (2023). It achieves this goal by making the most of the space that exists between the hyperplane and the data points that are nearest to each class (support vectors). Linear, polynomial, and radial basis functions are among the kernel functions that SVC uses to modify the feature space and improve separation, allowing it to handle both linear and non-linear classification problems.

Consider the most common nonlinearly separable case after the data set has been translated into the higher dimensional feature space. To modify the optimization issue, one can use the soft margin method and a cost function that incorporates the slack variable ξi, which is a measure of misclassification mistake.

$min\varphi(w,\varepsilon) = \frac{1}{2}\left|\left|w\right|\right|^2 + C\sum_i \varepsilon_i$ ---------- (7)

the kernel function that maps the higher dimensional feature space nonlinearly, denoted as ε_i. In higher-dimensional space, the training data becomes linearly separable, despite the Kernel function's nonlinearity in input space.

3.2.4 Gradient Boosting

Building a strong learner is an iterative process in boosting algorithms. Weak learners are defined as learners that perform marginally better than random referred by Nissa, N. et al. (2021). One regression approach that is similar to boosting is gradient boosting. To achieve an approximation, $F^{(x)}$, of the function $F(x)$, we minimize the anticipated value of a loss function, $L\big(y,F(x)\big)$, via gradient boosting. $D = \{xi,yi\}$ N 1 is the training dataset that we are given, and this function connects instances x to their output values y. To approximate $F(x)$ additively, gradient boosting builds a weighted sum of functions.

$F_m(x) = F_{m-1}(x) + p_m h_m(x)$ ---------- (8)

And the weight of the m^{th} function, $h_m(x)$, is represented by p_m. These procedures represent the ensemble's models, such as decision trees. Iterative construction of the approximation is used. After that, a line search optimization issue is solved to get the value of p_m.

3.2.5 Voting Classifier

We aggregated several weak classifiers we trained using the ensemble approach to arrive at a final verdict. One of the ensemble methods is the voting classifier that is used in this study. This is accomplished by applying a plethora of ineffective ML methods on the same dataset in tandem. Each model from the preceding classification models votes for every occurrence in the data set using the Voting Classifier (voting = 'hard') utilized in this research. More than 50% of voters will choose the final production forecast.

Algorithm 2: fused machine learning
Input:
Training data: $\{(X_1,\ Y_1), (X_2, Y_2), \dots, (X_N, Y_N)\}$
Number of trees: K
Kernel function $K(x_i, x_j)$
Number of iterations: M
Steps:
For each tree $k \in \{1,2,\dots,K\}$:
Bootstrap sampling entails selecting data at random and then replacing it with data from the training set.
At each node, choose a subset of characteristics at random to divide.
Formulate the optimization problem:
$min\frac{1}{2}\left\|\|w\|\right\|^2 + C\sum_{i=1}^{N}\varepsilon_i$
Initialize the model with a constant value:
$F_0(x) = argmin\sum_{i=1}^{N} L(Y_i, r)$
Fit a weak learner h_m to the pseudo-residuals.
Update the model:
$F_m(x) = F_{m-1}(x) + ah_m(x)$
Output:
Ensemble model consisting of K decision trees
Predicted class labels Y for the input test data.
Ensemble model combining predictions of all base classifiers

3.3 Prediction using cascaded levy flight optimization

Cascaded Levy Flight Optimization involves an advanced optimization technique to enhance model performance. It utilizes Levy flight, a stochastic process inspired by animal foraging behavior, to explore the parameter space of predictive models. The algorithm iteratively adjusts model parameters using Levy flight steps, optimizing them to improve prediction accuracy. Multiple rounds of parameter refinement are used in this cascaded technique, which improves the predictive model's overall performance and allows for greater convergence on optimum settings.

The step sizes of walkers in a Levy flight, an unplanned walking pattern, are selected from a probability distribution known as a Levy distribution, which has hefty tails. Large steps are more likely to be generated by this distribution, and they occur more often than they would in a normal distribution.

A power law tail characterizes the Levy distribution, and its probability density function is written as

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} exp\left(-\frac{r}{2\sigma^2}|x-\mu|\right)\frac{r}{|x-\mu|^{1+r}} \text{ ---------- (9)}$$

• "μ" is the location parameter, which is the distribution's mean.
• The scale parameter σ represents the distribution's standard deviation.
• The parameter γ determines the distribution's form and is known as the tail index.

The tail index value The form of the distribution is dictated by γ. When γ falls within the range of 0 to 2, the variance of the distribution is indefinite since it is an infinite variance distribution. The variance of the distribution is finite when γ > 2, whereas the mean is infinite when γ = 1 or less.

Algorithm 3: cascaded levy flight optimization

Input:

Model: The predictive model to be optimized (Random Forest, SVC, Gradient Boosting).

Training Data (X_{train}, y_{train}) : The features and labels used to train the model.

Testing Data (X_{test}, y_{test}): The features and labels used to evaluate model performance.

Steps:

Initialize: Set random hyperparameters within the defined ranges for the model.

Train and Evaluate: Fit the model with the initial parameters and evaluate its performance on the test data.

- Generate new hyperparameters using Levy flight to explore the parameter space.
- Train and evaluate the model with the new parameters.
  $f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} exp\left(-\frac{r}{2\sigma^2}|x-\mu|\right)\frac{r}{|x-\mu|^{1+r}}$
- Update the best parameters and model if the new model's performance is better.

Converge: After the maximum number of iterations, finalize the model with the best parameters found.

Output:

1. **Best Model**: The predictive model with the optimal hyperparameters found through the optimization process.
2. **Best Parameters**: The set of hyperparameters that yielded the highest model performance.
3. **Best Score**: The performance metric (e.g., accuracy) achieved by the model with the best parameters.

IV. Results and discussion

Here, we showcase and evaluate the results of our heat stroke prediction model, which was created utilizing state-of-the-art machine learning methods. We begin by evaluating the performance of the model with optimized hyperparameters, obtained through Cascaded Levy Flight Optimization.

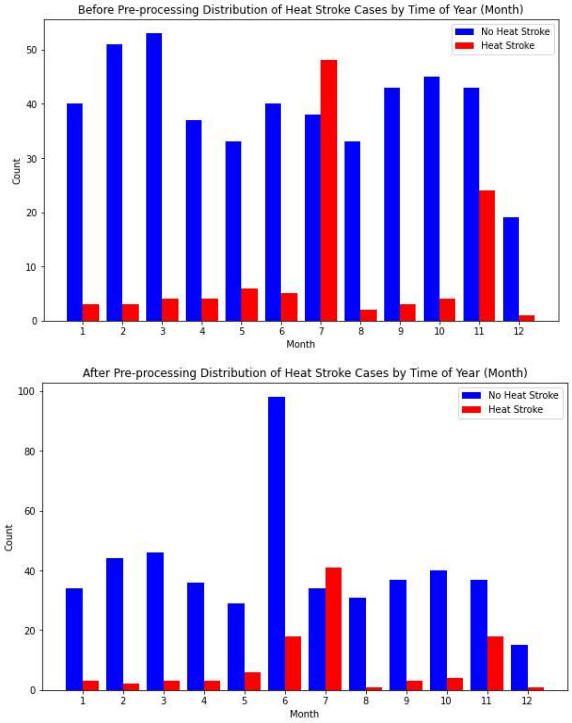

Figure 2 and 3: After and before preprocessing Distribution of heat stroke cases by time of year

Figure 3 illustrates the distribution of heat stroke cases across different times of the year, highlighting seasonal patterns and trends. The graph plots the frequency of heat stroke incidents against the months of the year, revealing how the incidence of heat stroke varies with seasonal changes.

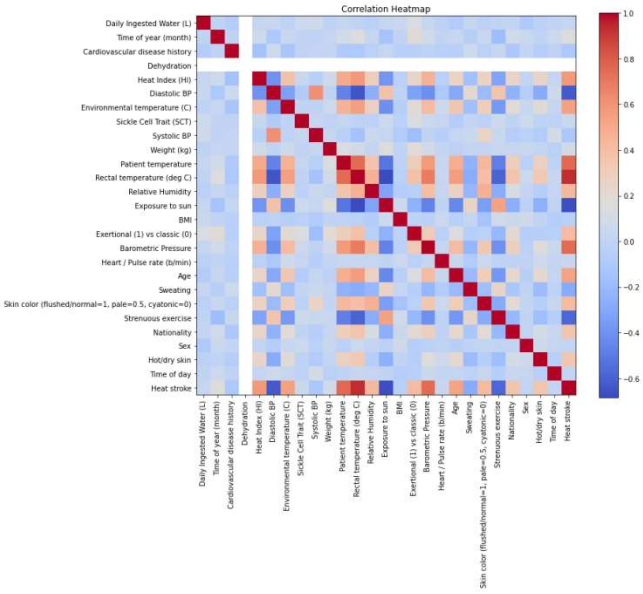

Figure 4: Correlation heatmap

The relationships between the various aspects of the dataset are visually represented in Figure 4 by a correlation heatmap. In every heatmap cell, you can see a correlation coefficient between two variables; these coefficients can take on values between -1 and 1. Since the coefficient is close to 1, it indicates a strong positive correlation, suggesting that a rise in one variable leads to an increase in the other.

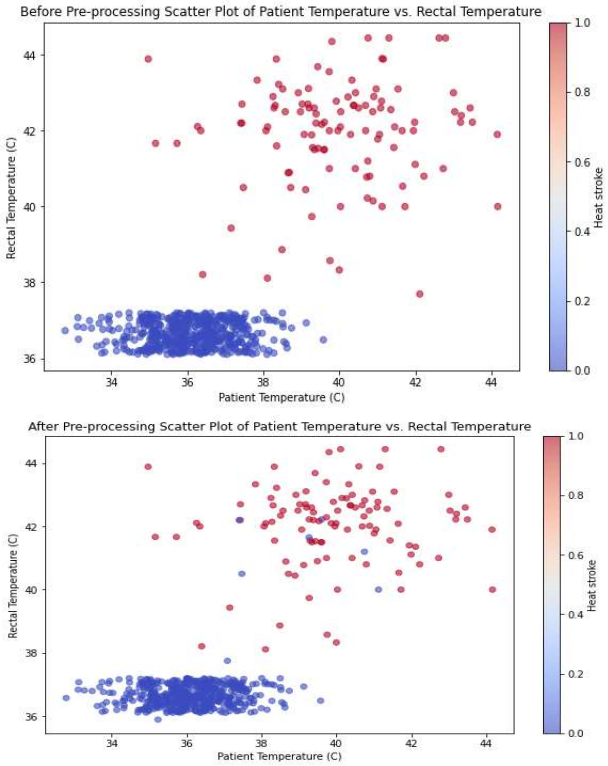

Figure 5 and 6: Before and After pre-processing scatter plot of patient temperature vs rectal temperature

Figure 5 and 6 displays a scatter plot illustrating the relationship between patient temperature and rectal temperature following data pre-processing. Each point on the plot represents an individual patient, with their temperature measurements plotted on the x-axis and rectal temperature measurements on the y-axis.

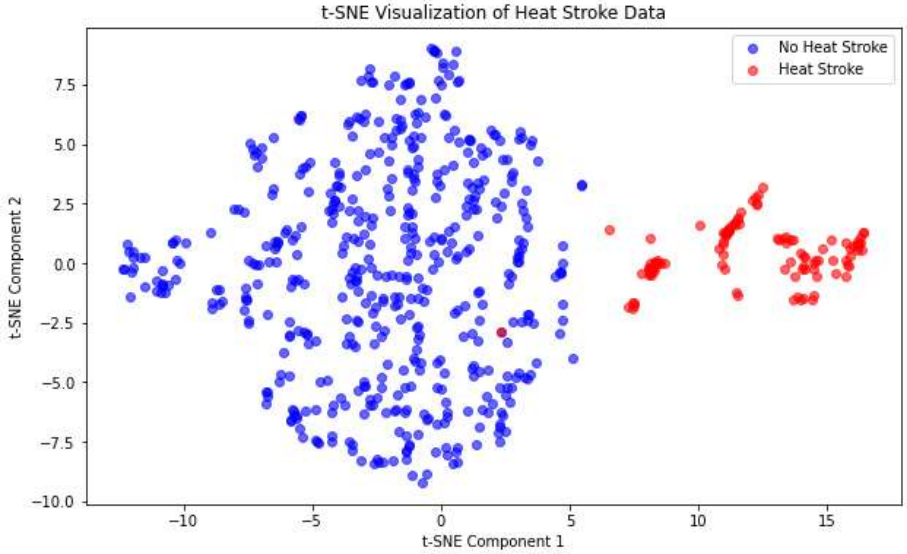

Figure 7: t-SNE visualization of heat stroke data

Figure 7 provides a t-SNE visualization of the heat stroke dataset, which is a dimensionality reduction technique used to visualize high-dimensional data in two or three dimensions. The plot displays how data points, representing different instances or patients, are distributed across a two-dimensional space.

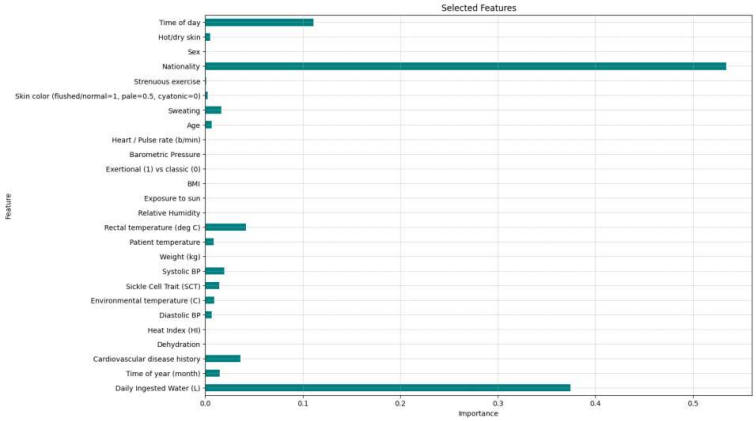

Figure 8: Selected feature (Time of day, Nationality, Sweating, age, temperature cardiovascular, water with 15 attributes)

The figure 8 shows selected feature the x axis shows importance and the y axis shows feature

Table 2: Feature extraction value comparison table

Methods	**Accuracy**	**Precision**	**Recall**	**F-measure**
PCA	94.32	94.02	94.67	95.31
LDA	94.99	94.99	94.36	95.24
t-SNE	95.71	94.36	95.36	95.25

PCA with t-SNE	96.36	96.22	96.87	97.81

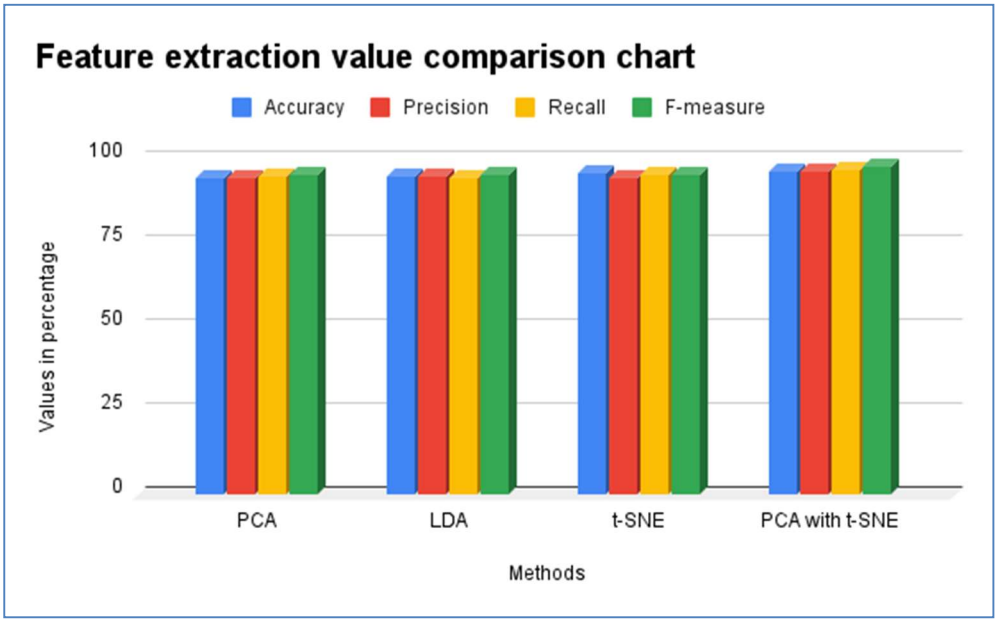

Figure 9: Feature selection value comparison chart

The table 2 and figure 9 shows the combination of PCA with t-SNE provides the highest performance across all metrics for feature extraction and classification. Specifically, this approach achieved an accuracy of 96.36%, surpassing the individual techniques such as PCA (94.32%), Linear Discriminant Analysis (LDA) (94.99%), and t-SNE (95.71%). It also demonstrated the highest precision at 96.22%, which is significantly better than PCA (94.02%), LDA (94.99%), and t-SNE (94.36%). Additionally, the recall of 96.87% with PCA and t-SNE is superior to the other methods, reflecting its ability to identify true positives more effectively. The F-measure of 97.81% highlights its balanced performance in both precision and recall. Overall, the PCA combined with t-SNE method excels in providing a robust and accurate feature extraction framework, making it highly effective for applications requiring high precision and recall.

Table 3: Performance Metrics of Machine Learning Models with and without Features

Methods		Accuracy	Precision	Recall	F-measure
RFC	**Without Feature Selection**	94.36	94.25	94.21	94.17
SVC		95.31	95.24	95.17	95.87
GB		96.31	96.28	96.59	96.18
Fused ML		97.35	97.18	97.17	97.19
FML without optimization		98.24	98.14	98.20	98.21
RFC	**With Feature Selection**	96.21	96.36	95.24	96.39
SVC		97.51	97.54	97.81	97.99
GB		97.99	97.89	97.66	97.25

Fused ML		98.21	98.01	98.37	98.17
FML with optimization		99.36	99.05	99.61	99.98

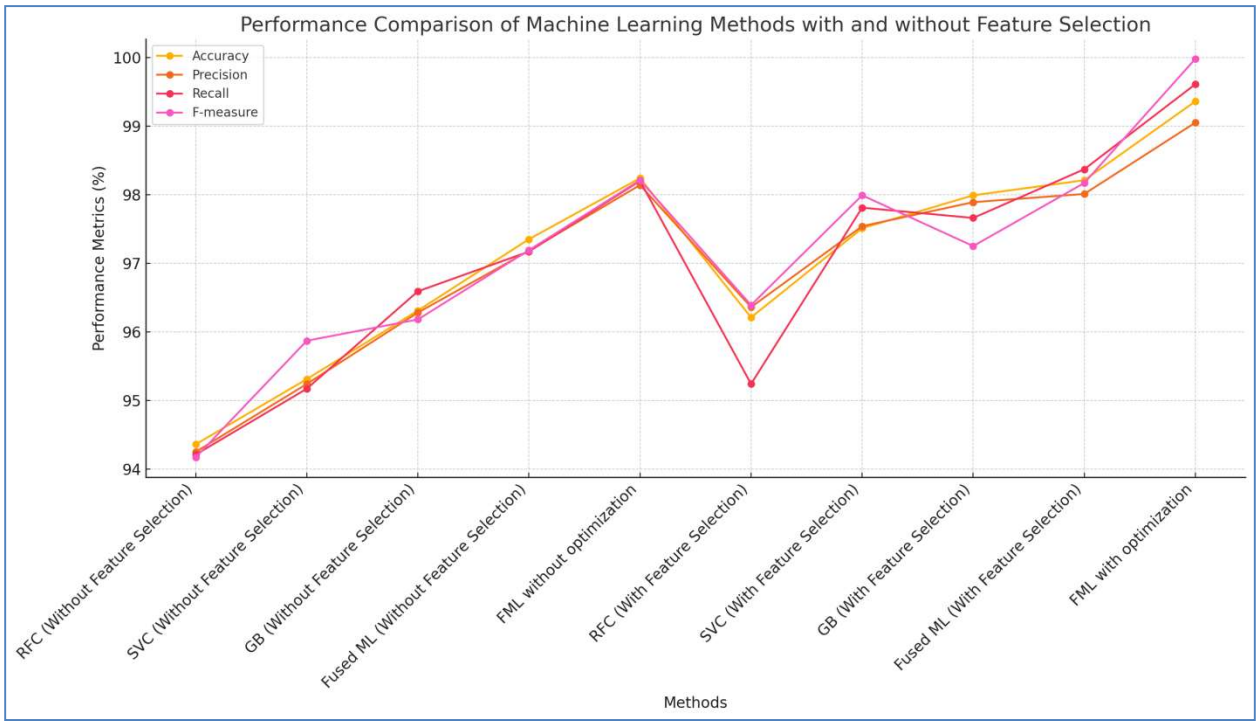

Figure 10: Performance Metrics of Machine Learning Models with and without Optimization comparison chat

The table 3 and figure 10 presents the performance metrics of various machine learning models, both with and without optimization. Among the models evaluated, the Cascaded Levy Flight approach demonstrates the highest performance across all metrics in both optimization scenarios. Without optimization, it achieves an accuracy of 98.24%, a precision of 98.14%, a recall of 98.20%, and an F-measure of 98.21%. With optimization, it further improves to 99.36% accuracy, 99.05% precision, 99.61% recall, and a remarkable F-measure of 99.98%. In comparison, the Fused ML model, which is the next best performer, shows slightly lower values, with accuracy, precision, recall, and F-measure values consistently below those of the Cascaded Levy Flight approach. Overall, the Cascaded Levy Flight method significantly outperforms other models in detecting and classifying the target, demonstrating its superior effectiveness in the given context.

V. Conclusion

This study presents a robust and integrated machine learning framework for the accurate prediction of heat stroke, addressing the complexities of both environmental and physiological factors. By employing advanced techniques such as Reorder Iterative Imputation for data

normalization, PCA combined with t-SNE for effective feature selection, and a fusion of sophisticated classification algorithms optimized through Cascaded Levy Flight Optimization, the proposed approach significantly outperforms traditional methods in terms of prediction accuracy and reliability. The results underscore the potential of this framework to serve as a powerful tool for early intervention, enabling proactive health monitoring and safety management during extreme heat events. Without optimization, it achieves an accuracy of 98.24%, a precision of 98.14%, a recall of 98.20%, and an F-measure of 98.21%. With optimization, it further improves to 99.36% accuracy, 99.05% precision, 99.61% recall, and a remarkable F-measure of 99.98%. Future work will focus on expanding the dataset and refining the model further to enhance its generalizability across different populations and climatic conditions.

VI. References

1. Akbar, M. Y., Azad, H., Rashid, M., Ajmal, W., Ansari, A. A., Salem, M. M., ... & Ghazanfar, S. (2024). ES-PredHSP: Improved Prediction of Heat Shock Proteins Using Machine Learning by Enhanced Sampling Technique. *JOURNAL OF BIOLOGICAL REGULATORS AND HOMEOSTATIC AGENTS*, *38*(1), 665-673.
2. Eldora, K., Fernando, E., & Winanti, W. (2024). Comparative Analysis of KNN and Decision Tree Classification Algorithms for Early Stroke Prediction: A Machine Learning Approach. *Journal of Information Systems and Informatics*, *6*(1), 313-338.
3. Hirano, Y., Kondo, Y., Hifumi, T., Yokobori, S., Kanda, J., Shimazaki, J., ... & Yaguchi, A. (2021). Machine learning-based mortality prediction model for heat-related illness. *Scientific reports*, *11*(1), 9501.
4. Ikeda, T., & Kusaka, H. (2021). Development of models for predicting the number of patients with heatstroke on the next day considering heat acclimatization. *Journal of the Meteorological Society of Japan. Ser. II*, *99*(6), 1395-1412.
5. Ke, D., Takahashi, K., Takakura, J. Y., Takara, K., & Kamranzad, B. (2023). Effects of heatwave features on machine-learning-based heat-related ambulance calls prediction models in Japan. *Science of the total environment*, *873*, 162283.
6. Kim, S. H., Jeon, E. T., Yu, S., Oh, K., Kim, C. K., Song, T. J., ... & Jung, J. M. (2021). Interpretable machine learning for early neurological deterioration prediction in atrial fibrillation-related stroke. *Scientific reports*, *11*(1), 20610.
7. Li, H., Gao, M., Song, H., Wu, X., Li, G., Cui, Y., ... & Zhang, H. (2023). Predicting ischemic stroke risk from atrial fibrillation based on multi-spectral fundus images using deep learning. *Frontiers in Cardiovascular Medicine*, *10*, 1185890.
8. Nakai, T., Saiki, S., & Nakamura, M. (2021, December). Medium-Term Prediction for Ambulance Demand of Heat Stroke using Weekly Weather Forecast. In *2021 8th International Conference on Internet of Things: Systems, Management and Security (IOTSMS)* (pp. 1-8). IEEE.
9. Nissa, N., Jamwal, S., & Mohammad, S. (2021). Heart disease prediction using machine learning techniques. *Wesleyan Journal of Research*, *13*(67).

10. Ogata, S., Takegami, M., Ozaki, T., Nakashima, T., Onozuka, D., Murata, S., ... & Nishimura, K. (2021). Heatstroke predictions by machine learning, weather information, and an all-population registry for 12-hour heatstroke alerts. *Nature communications*, *12*(1), 4575.
11. Otieno, J. A., Häggström, J., Darehed, D., & Eriksson, M. (2024). Developing machine learning models to predict multi-class functional outcomes and death three months after stroke in Sweden. *Plos one*, *19*(5), e0303287.
12. Padimi, V., Telu, V. S., & Ningombam, D. D. (2023). Performance analysis and comparison of various machine learning algorithms for early stroke prediction. *ETRI Journal*, *45*(6), 1007-1021.
13. Priyadarshini, T. S., Hameed, M. A., & Robert, B. S. (2022). Deep Learning Prediction Model for predicting heart stroke using the combination Sequential Row Method integrated with Artificial Neural Network. *Journal of Positive School Psychology*, 623-629.
14. Shimazaki, T., Anzai, D., Watanabe, K., Nakajima, A., Fukuda, M., & Ata, S. (2022). Heat stroke prevention in hot specific occupational environment enhanced by supervised machine learning with personalized vital signs. *Sensors*, *22*(1), 395.
15. Statsenko, Y., Fursa, E., Laver, V., Talako, T., Simiyu, G. L., Al Zahmi, F., ... & Ljubisavljevic, M. (2024). Machine Learning Models Predicting Incidence, Severity, and Early Outcomes of Hemorrhagic Stroke from Weather Parameters and Individual Risk Factors.
16. Xu, H., Guo, S., Shi, X., Wu, Y., Pan, J., Gao, H., ... & Han, A. (2024). Machine learning-based analysis and prediction of meteorological factors and urban heatstroke diseases. *Frontiers in Public Health*, *12*, 1420608.
17. Yaldiz, C. O., Buller, M. J., Richardson, K. L., An, S., Lin, D. J., Satish, A., ... & Inan, O. T. (2023). Early Prediction of Impending Exertional Heat Stroke With Wearable Multimodal Sensing and Anomaly Detection. *IEEE Journal of Biomedical and Health Informatics*.
18. Yi, W., Zhao, Y., & Chan, A. P. (2022, November). A Tailored Artificial Intelligence Model for Predicting Heat Strain of Construction Workers. In *IOP Conference Series: Earth and Environmental Science* (Vol. 1101, No. 7, p. 072004). IOP Publishing.
19. Yu, H., Ji, X., & Ouyang, Y. (2023). Unfolded protein response pathways in stroke patients: a comprehensive landscape assessed through machine learning algorithms and experimental verification. *Journal of Translational Medicine*, *21*(1), 759.