Nanotechnology Perceptions
ISSN 1660-6795
www.nano-ntp.com

Combination Of Fused Machine Learning And Cascaded Levy Flight Optimization In Heat Stroke Prediction

Mohamad Emad Bitar[1], Dr. V. Sujatha[2]

[1]*Ph.D. Scholar, Department of Computer Science, Bharathiar University, CMS College of Science & Commerce Coimbatore, India - 641035*
t22h.12345@gmail.com
[2]*Vice Principal, CMS College of Science and Commerce, Coimbatore, India – 641035*
sujatha.padmakumar4@gmail.com

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

A dangerous medical emergency known as heat stroke occurs when the core body temperature rises to more than 40 degrees Celsius (104 degrees Fahrenheit) as a result of being exposed to very hot and humid conditions for an extended period of time [1]. It poses a serious risk to public health, especially during heatwaves or in places without proper cooling infrastructure [2]. If left untreated, the illness may cause organ failure and ultimately death [3]. To reduce the risks of heat stroke and to intervene promptly, early identification and prediction are of the utmost importance [4]. A number of physiological and environmental factors, such as personal health, levels of physical activity, and humidity, interact intricately to determine the likelihood of heat stroke [5]. Because of their reliance on empirical principles or simplistic statistical techniques, traditional methods of heat stroke prediction have failed to capture the complex interactions between these components [6]. Furthermore, dealing with inadequate or noisy information and integrating multiple data sources are also necessary for real-time prediction and monitoring [7].

By facilitating data-driven insights and predictive capacities, machine learning (ML) approaches provide a potential substitute to conventional methodologies [8–10]. Machine learning (ML) algorithms can efficiently process massive amounts of data, uncover intricate patterns, and provide accurate predictions when fed recent and historical data [11, 12]. Heat stroke forecasts may be made more accurate with the use of these systems, which use contemporary data processing and analytical methodologies [13–14]. Using the Reorder Iterative Imputer to normalise the data, we tackle the problem of dealing with inconsistent or missing data in this work [15-16]. By repeatedly honing the imputation of missing values and outliers, this method makes sure the dataset is ready for ML models [17–18]. If you want better input features and better performance from your machine learning algorithms, you need to normalise your data properly [19].

Using an iterative imputer for data normalisation and PCA with t-SNE for feature selection are the primary contributions of the work.
Using fused ML for classification and cascaded levy flight optimisation for predictionFor what's left of the article, this is its table of contents. In Section 2, several writers explore different methods for predicting heat stroke. In Section 3, the model that has been suggested is outlined. The findings of the study are summarised in Section 4. Analysis of the findings and recommendations for further work make up Section 5.

1.1 Motivation of the paper

Despite efforts to mitigate the effects of climate change and other severe weather events, heat stroke continues to be a major problem in public health. The complex interplay between a wide range of environmental and physiological variables makes it difficult, even with improvements in health monitoring, to reliably anticipate when heat stroke will occur. This work is driven by the urgent need for a dependable prediction system that can accurately identify when heat stroke will occur, allowing for prompt actions that might save lives. This work employs state-of-the-art machine learning approaches to address the shortcomings of conventional methods. It provides a holistic strategy that improves prediction capabilities by integrating data normalisation, feature selection, and optimisation.

II. Background study

Akbar, M. et al. [1] According to its expanded sampling technique, the SVM model achieved 87% accuracy in predicting heat shock proteins, according to the study conducted by these scientists. Furthermore, heat shock proteins in eukaryotic organisms may be more precisely identified using the study model. To find the proteins that are untraceable using this method, however, urgent research and the creation of sophisticated machine learning models were required.

Hirano, Y. et al. [3] Prediction models for heat-related illnesses based on machine learning showed potential for the first time in the authors' study. Larger samples, crucial factor inclusion, and clinical prospective validation were all areas that might have used more investigation to improve performance quality.

Ke, D. et al. [5] The author came up with a reliable way to forecast the number of heat-related ambulance calls per day. Very high forecast accuracy necessitated the development of regional-level models, while the national-level model demonstrated exceptional prediction accuracy applicable to several locations. Writing about heatwave characteristics like optimal temperature, cumulative heat stress, and heat acclimatisation allowed these authors to make far more accurate predictions. With the inclusion of heatwave characteristics, the national model's adjusted R2 increased from 0.9061 to 0.9659, significantly improving its ability to forecast the frequency of heat-related ambulance calls in most places with little computing effort.

Li, H. et al. [7] it is possible that AF patients' multi-spectrum fundus photographs, on their own, may predict the chance of secondary WAS, and that DNNs had better prediction performance when combined with such photos. In order to uncover unique microscopic features of disease, it was helpful to get many spectral fundus photographs.

Nissa, N. et al. [9] This study mainly contributed by comparing several ML algorithms for early-stage CVD prediction. The quality of the dataset was improved by the use of preprocessing procedures. Dealing with faulty or missing data and removing outliers were the main concerns. Using a variety of statistical criteria, the author also compared the results of the three machine learning algorithms she employed to predict the disease.

Priyadarshini, T. et al. [13] The purpose of this research was to develop a smart prediction model that, using a dataset of heat stroke symptoms, could assess patients' susceptibility to heart attacks. The author used K-means clustering, Naïve Bayes, Decision Tree, and Artificial Neural Network classifiers to build a prediction system.

Statsenko, Y. et al. [15] The study found that certain weather conditions were linked to higher HS incidence, severity, and early results. In the days after a significant shift in AT, humidex, or AP, the risk ratio of HS often increases. There may still be a significant risk even after the ecosystem has stabilised. It seems that the two parameters should be considered together, as Humidex fared better than AT as a predictor. Combining and analysing all available meteorological data from the past few days was crucial for making accurate projections.

Yu, H. et al. [19] The author sought to determine if there was a correlation between the expression of genes related to the Unfolded Protein Response (UPR) pathway and ischaemic stroke (WAS) patient clinical features, immune cell infiltration, and inflammatory factor release. The authors demonstrated a strong relationship between these factors by analysing gene microarray data collected 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 robust predictions	Potentially high computational cost and complexity in integrating various data sources

2.1 Problem definition

Particularly during severe weather events, there is a considerable danger of heat stroke, which is caused by being exposed to high temperatures for an extended period of time. The complicated interaction of environmental and physiological variables makes accurate prediction of the onset of heat stroke problematic. When faced with such complexity and the need for rapid forecasts, traditional approaches often fail. This study employs advanced machine learning techniques to investigate many indications, enhance data quality, and optimise model performance; the goal is to provide a more reliable tool for early diagnosis and preventive measures. Improving the accuracy of predictions is the main objective.

III. Materials and methods

Here we lay out the steps that will be taken to create a reliable heat stroke prediction system. To properly manage missing values and outliers, the technique starts with data preparation, which includes normalisation by a Reorder Iterative Imputer. Combining principal component analysis (PCA) with t-SNE enables feature selection, which in turn identifies the most important predictors. In order to improve classification accuracy, a fused machine learning strategy is applied. This approach integrates Gradient Boosting, Random Forest, and Support Vector Classifier. Cascaded Levy Flight Optimisation, which optimises the model's parameters to produce greater predicted accuracy, is used to optimise the model's performance.

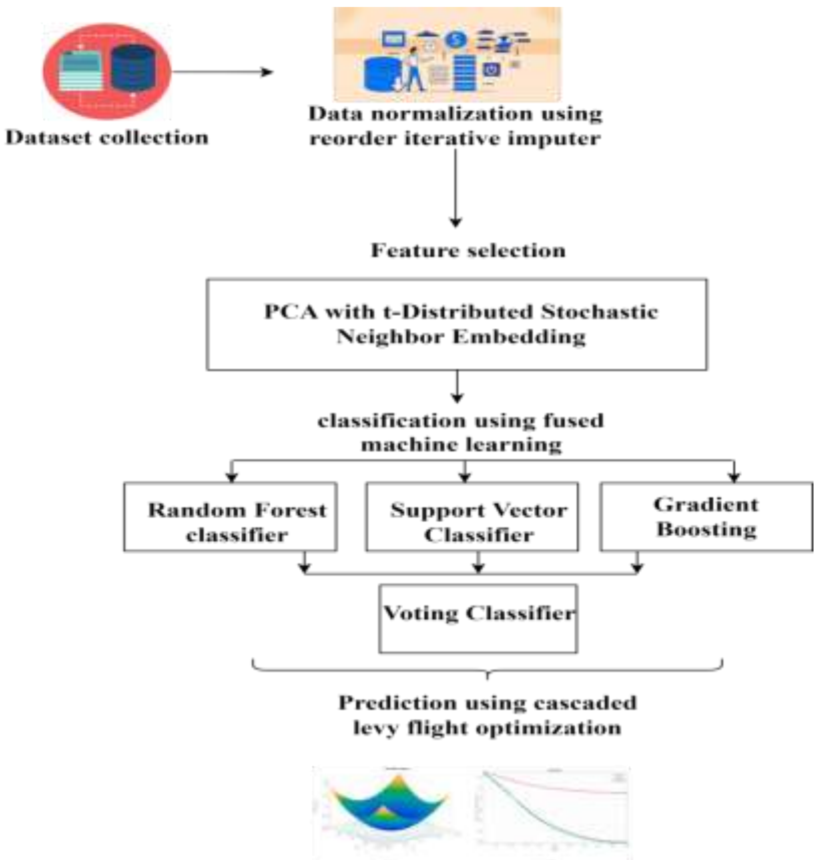

Figure 1: Heat stroke prediction workflow architecture

3.1 Dataset collection

Kaggle, a well-known website for data science and machine learning tools, provided the dataset used in this work. The dataset in question may be accessed at this URL: https://www.kaggle.com/datasets/tahiatazin1510997643/heat-stroke. Critical for forecasting the start of heat stroke, it contains comprehensive data of heat stroke episodes together with a range of physiological and environmental characteristics.

3.2 Data normalization using reorder iterative imputer

Data normalisation techniques such as the Reorder Iterative Imputer deal with issues including outliers and missing values. It uses feature-feature connections and distribution-distribution-adjustment repeatedly to estimate and impute missing data. Improved accuracy and reliability in normalisation are achieved by rearranging the imputation process according to the significance of features and data patterns. Data quality is improved and the dataset is prepared for effective analysis and modelling by iteratively refining the imputation using the Reorder Iterative Imputer.

Generative iteration is used by Imputer. Imputer applies criteria to a partially formed alignment and creates a new alignment with each iteration of the generating process. In each generating phase, a complete alignment is formed. However, in inference, successive partial alignments are produced by using a reduced selection. The previously issued tokens are often included in this subset as a result of an iterative refinement process. Unlike CTC, Imputer's iterative technique may specify conditional dependencies across generation stages. If the Imputer wants to parameterise the distribution of alignment, it will presume that token predictions are conditionally independent.

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

where $p_\theta(a|a^-, x)$and ∅ denote the masked out letter, and a˜ represents an earlier alignment. As per BERT's findings, the specification of the mask token aligns with the ∅ mask token. To clarify, after a token is committed in a˜, a must stay consistent; otherwise, a˜ might be empty (for instance, all tokens are concealed) and ai = ˜ai∀a˜i 6= empty. By conditioning on a˜, Imputer can model conditional dependencies between phases of generation, and by assuming conditional independence between predictions of new tokens, it can allow parallel generation inside a generation step.

3.3 Feature selection and classification using fused machine learning

In order to improve prediction performance, fused machine learning incorporates many algorithms into feature selection and classification. In order to increase the effectiveness of the model, feature selection uses t-Distributed Stochastic Neighbour Embedding methods to identify the most important variables from the dataset. To classify data, a fused machine learning model uses techniques like a Voting Classifier to combine the predictions of many classifiers, including Random Forest, Support Vector Classifier, and Gradient Boosting. Through the use of multiple learning approaches and an emphasis on critical characteristics, this strategy enhances overall accuracy and resilience by combining the capabilities of each model.

3.3.1 PCA with t-Distributed Stochastic Neighbor Embedding

There is a two-step dimensionality reduction procedure involved in feature extraction utilising PCA paired with t-Distributed Stochastic Neighbour Embedding (t-SNE). To begin, principle component analysis (PCA) is used to find the main components that account for the highest variation in the data, so reducing the high-dimensional data to a lower-dimensional space. The data is made more manageable for further analysis after this first reduction simplifies it by reducing noise. For example, due to their significant value in the data, characteristics like gender, nationality, and daily consumed water (L) may be lowered during principal component analysis (PCA). It is then possible to further decrease the data's dimensionality to two or three dimensions by applying t-SNE to the PCA-reduced data, all the while keeping the local structure and connections between data points intact. The interplay between Strenuous exercise, Rectal temperature (deg C), and Environmental temperature (C) is only one example of how this combination enables efficient visualisation of complicated data structures by drawing attention to clusters and patterns that may go unnoticed in higher dimensions.

Inaccurate categorisations and limited applicability are the end outcomes. Reduced model complexity is achieved by t-Distributed Stochastic Neighbour Embedding by the elimination of superfluous predictors. Since t-SNE is a wrapper approach that mostly uses filter feature selection, it may easily be used as the primary feature selection method in many ML algorithms. Next, the features are ranked based on their importance using coefficients or feature significance. The model is re-fitted by individually deleting the poorest feature or features. The procedure is iterated until the quantity of characteristics surpasses a certain limit.

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

Next, we employ eight different t-Distributed Stochastic Neighbour Embedding algorithms to identify the feature cut-offs. Most people believe these are the most essential attributes to have. This strategy picks |αP| features from each feature subset to create the best feature subset.

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

The round-down operator is represented mathematically as |αP|.This will enable us to exclude characteristics with poor robustness and accuracy, such as Weight (kg) or Sickle Cell Trait (SCT), depending on the dataset. Consider a scenario in which the parameter τ surpasses the AUC of the top N feature sets for predictive classification. This is why they are known as:

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

Finally, we define the robust biomarker screening issue as an unstable combination problem with N characteristics. To ensure stability, all permutations of the sets in FS_2^optopt are examined.

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 creates numerous decision trees and then utilises the mean prediction (regression) or mode of their classes (classification) as its final output, as cited by Eldora, K. et al. (2024). It improves accuracy and robustness by combining the predictions of several trees, each built from a different set of data and features. This strategy enhances generalisation while reducing overfitting by averaging out individual tree biases.

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

I(•) is the indicator function in this situation. The margin is defined as the amount by which the relevant class's average vote at X, Y exceeds the average vote for any other class. We may have more trust in the categorisation if the margin is larger. The generalisation error may be stated as:

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

3.2.3 Support Vector Classifier

The Support Vector Classifier (SVC) in machine learning looks for the optimum hyperplane in the feature space while classifying data, as alluded to by Ke, D. et al. (2023). It accomplishes this purpose by maximising the space between the hyperplane and the data points closest to each class (support vectors). Linear, polynomial, and radial basis functions are among the kernel functions used by SVC to change the feature space and increase separation, enabling it to handle both linear and nonlinear classification tasks.

Consider the most typical nonlinearly separable situation after translating the data set into the higher-dimensional feature space. To address the optimisation problem, consider using the soft margin technique and a cost function that includes the slack variable ξi, a measure of misclassification error.

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

A kernel function, designated as ε_i, maps the higher-dimensional feature space nonlinearly. The training data becomes linearly separable in higher dimensions, despite the Kernel function's nonlinearity in input space.

3.2.4 Gradient Boosting

Building a powerful learner is an iterative process for improving algorithms. Weak learners are those that do slightly better than random, according to Nissa, N. et al. (2021). Gradient boosting is a regression strategy similar to boosting. To approximate the function $F^{(x)}$, we use gradient boosting to minimise the expected value of a loss function $L\big(y, F(x)\big)$, The training dataset is supplied as$D = \{xi, yi\}$ N 1 , and this function relates occurrences x to their output values y. Gradient boosting creates a weighted sum of functions to approximate $F^{(x)}$

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

The weight of the m^{th}function, $h_m(x)$, is denoted by p_m. These techniques reflect the ensemble's models, like decision trees. An iterative approximation is applied. Following that, a line search optimisation problem is resolved to determine the value of p_m

3.2.5 Voting Classifier

To get our ultimate conclusion, we combined multiple weak classifiers trained using the ensemble technique. The voting classifier employed in this work is an example of an ensemble technique. This is performed by running a slew of useless ML techniques on the same dataset simultaneously. Each of the previous classification models votes for each occurrence in the data set using the Voting Classifier (voting = 'hard') used in this study. More than half of voters will choose the ultimate output prediction.

Algorithm 2: fused machine learning
Input: Training data: $\{(\mathrm{X}_1, \mathrm{Y}_1), (\mathrm{X}_2, \mathrm{Y}_2), \ldots, (\mathrm{X}_\mathrm{N}, \mathrm{Y}_\mathrm{N})\}$ Number of trees: K Kernel function $\mathrm{K}(\mathrm{x}_\mathrm{i}, \mathrm{x}_\mathrm{j})$ Number of iterations: M **Steps:** **For** each tree k ∈ {1,2, . . . , 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:

$\min \frac{1}{2}||w||^2 + C\sum_{i=1}^{N} \varepsilon_i$
Initialize the model with a constant value:
$F_0(x) = \operatorname{argmin} \sum_{i=1}^{N} L(Y_i, r)$
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 Optimisation uses complex optimisation techniques to improve model performance. It uses Levy fly, a stochastic process inspired by animal foraging behaviour, to investigate the parameter space of predictive models. The approach repeatedly optimises model parameters using Levy flight stages to increase prediction accuracy. This cascaded approach employs many rounds of parameter adjustment, improving the overall performance of the predictive model and allowing for higher convergence on optimal parameters.

The step sizes of walkers in a Levy flight, an unplanned walking pattern, are determined by a probability distribution known as the Levy distribution, which has large tails. This distribution generates larger steps more often than a regular distribution.

The Levy distribution has a power law tail and is represented by the probability density function

$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{r}{2\sigma^2}|x-\mu|\right)\frac{r}{|x-\mu|^{1+r}}$ -------------- (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 ($\mathbf{X_{train}}, \mathbf{y_{train}}$) : The features and labels used to train the model.
Testing Data ($\mathbf{X_{test}}, \mathbf{y_{test}}$): 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 present and analyse the outcomes of our heat stroke prediction model, which was developed using cutting-edge machine learning approaches. We begin by testing the model's performance using optimised hyperparameters produced via Cascaded Levy Flight Optimisation.

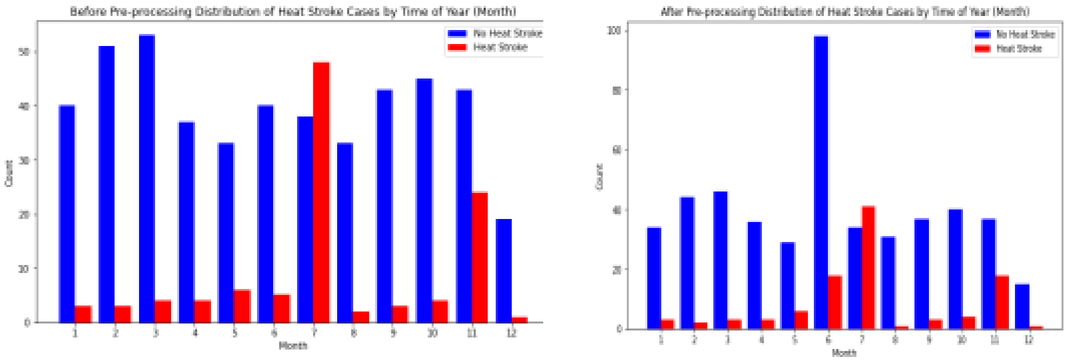

Figure 3 and 4: After and before preprocessing Distribution of heat stroke cases by time of year

Figure 3 depicts the distribution of heat stroke cases throughout the year, illustrating seasonal variations and changes. The graph displays the frequency of heat stroke episodes against the months of the year, demonstrating how the incidence of heat stroke fluctuates with the seasons.

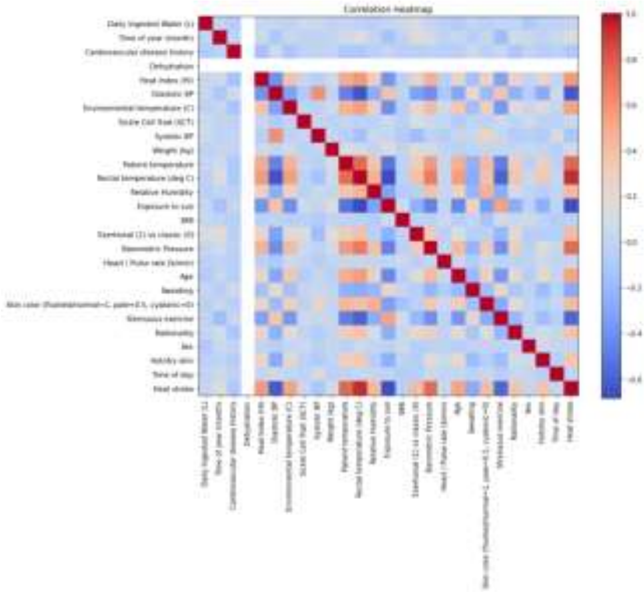

Figure 4: Correlation heatmap

Figure 4 depicts the relationships between the dataset's different features using a correlation heatmap. In each heatmap cell, you can observe a correlation coefficient between two variables, which might range from -1 to 1. Because the coefficient is near to one, it implies a significant positive correlation, implying that raising one variable causes an increase in the other.

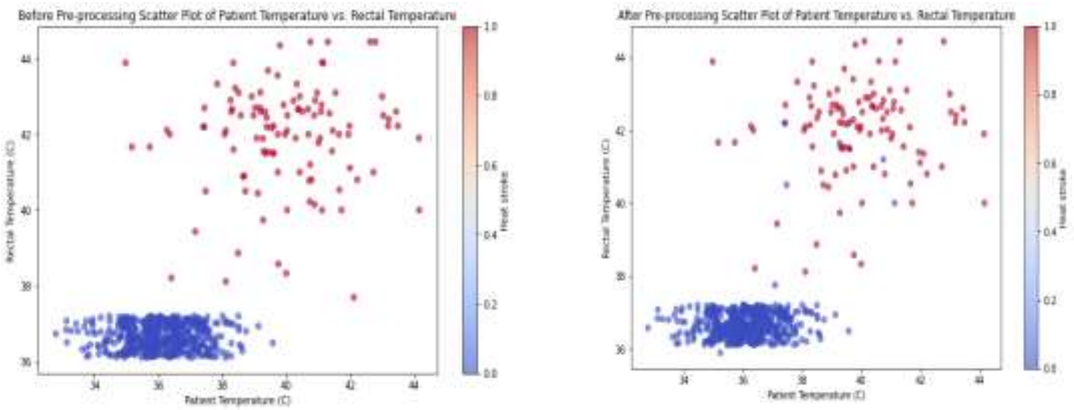

Figure 5 and 6: Before and After pre-processing scatter plot of patient temperature vs rectal temperature

Figures 5 and 6 show a scatter plot of the connection between patient temperature and rectal temperature after data preprocessing. Each point on the figure represents a unique subject, with temperature readings on the x-axis and rectal temperature measurements on the y-axis.

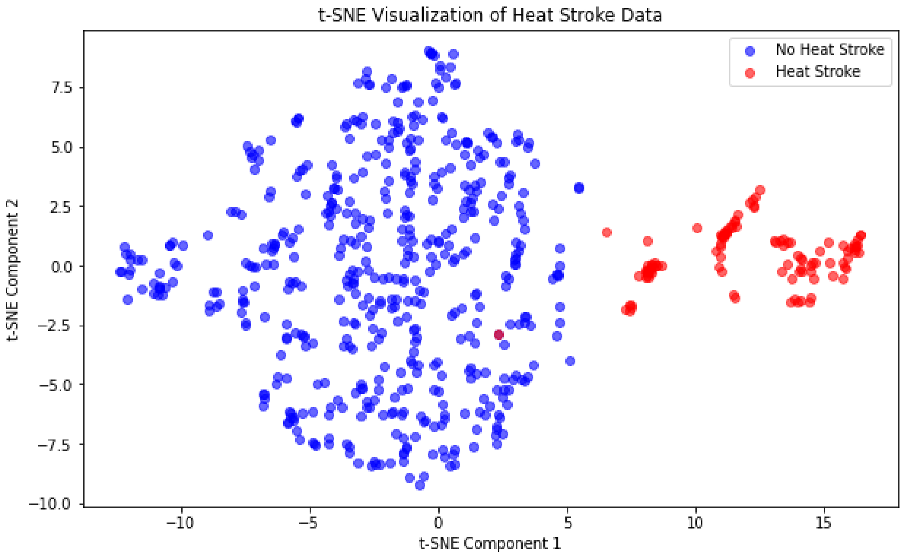

Figure 7: t-SNE visualization of heat stroke data

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.

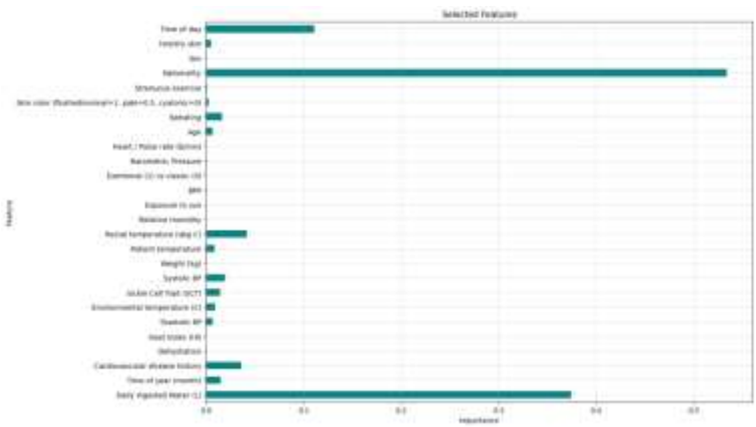

Figure 8: Selected feature (Time of day, Nationality, Sweating, age, temperature cardiovascular, water with 15 attributes)

Figure 8 illustrates the chosen feature. The x axis denotes significance, whereas the y axis displays 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

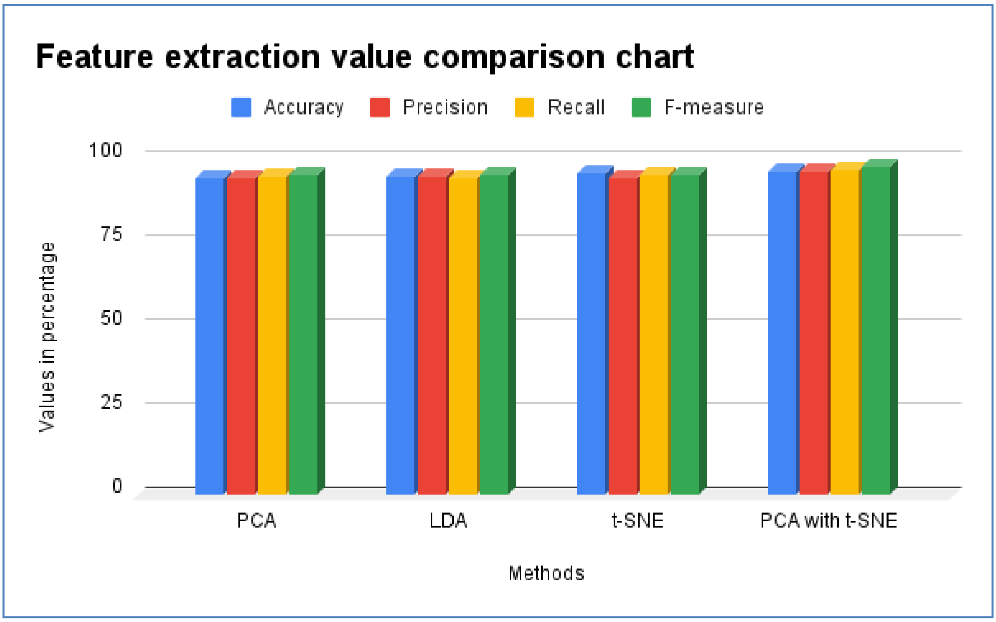

Figure 9: Feature selection value comparison chart

Table 2 and Figure 9 indicate that the combination of PCA and t-SNE achieves the best performance across all metrics for feature extraction and classification. Specifically, this strategy attained an accuracy of 96.36%, outperforming individual approaches such as PCA (94.32%), Linear Discriminant Analysis (LDA) (94.99%), and t-SNE (95.71%). It also had the best accuracy (96.22%), which outperformed PCA (94.02%), LDA (94.99%), and t-SNE (94.36%). Furthermore, PCA and t-SNE have a higher recall of 96.87% than the other approaches, indicating that they can better detect true positives. The F-measure of 97.81% demonstrates a balanced performance in both accuracy and recall. Overall, the PCA with t-SNE technique excels at providing a strong and precise feature extraction framework, making it ideal for applications that need high accuracy 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

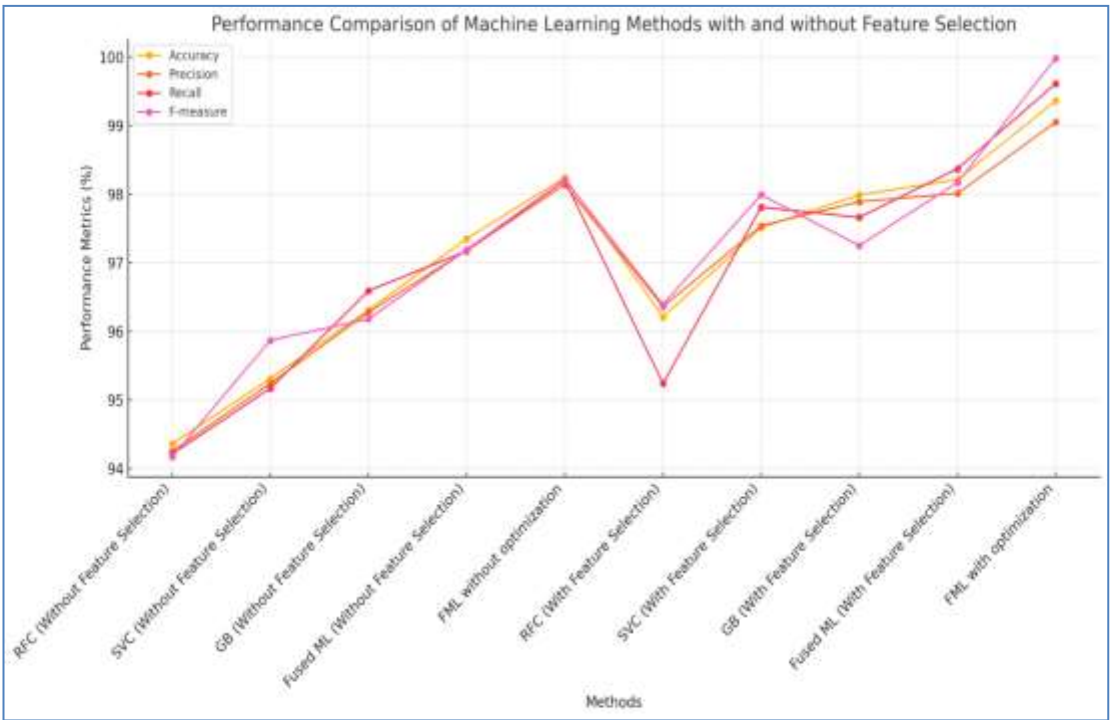

Figure 10: Performance Metrics of Machine Learning Models with and without Optimization comparison chat

Table 3 and Figure 10 show the performance characteristics of several machine learning models, with and without optimisation. Among the models tested, the Cascaded Levy Flight technique outperforms all metrics in all optimisation situations. Without optimisation, it has an accuracy of 98.24%, precision of 98.14%, recall of 98.20%, and F-measure of 98.21%. With optimisation, it achieves 99.36% accuracy, 99.05% precision, 99.61% recall, and an impressive F-measure of 99.98%. In contrast, the Fused ML model, the next best performer, has somewhat lower results, with accuracy, precision, recall, and F-measure values consistently lower than those of the Cascaded Levy Flight method. Overall, the Cascaded Levy Flight technique beats other models in identifying and categorising targets, proving its greater usefulness in the current situation.

V. Conclusion

This paper proposes a strong and integrated machine learning framework for accurately predicting heat stroke, taking into account the complexity of both environmental and physiological components. The proposed approach significantly outperforms traditional methods in terms of prediction accuracy and reliability by utilising advanced techniques such as Reorder Iterative Imputation for data normalisation, PCA combined with t-SNE for effective feature selection, and a fusion of sophisticated classification algorithms optimised via Cascaded Levy Flight Optimisation. The findings highlight the framework's potential as an effective early intervention tool, allowing for proactive health monitoring and safety management during severe heat events. Without optimisation, it has an accuracy of 98.24%, precision of 98.14%, recall of 98.20%, and F-measure of 98.21%. With optimisation, it achieves 99.36% accuracy, 99.05% precision, 99.61% recall, and an impressive F-measure of 99.98%. Future study will concentrate on increasing the dataset and improving the model to improve its generalisability across diverse demographics and climates.

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.