



Comparative Analysis of Deep Learning and Machine Learning Techniques for Obesity Classification

Prince Chauhan^{1*} and Shilpee Srivastava²

^{1,2} Department of Mathematics, Chandigarh University, Mohali, Punjab - 140413
*princechauhan324u@gmail.com, ²shilpee2007@rediffmail.com

Abstract. The number of obese individuals has risen to alarming levels throughout the world and today obesity can be considered an illness that causes numerous diseases such as diabetes, cancer and heart disease. Effective Obesity classification can help in early detection of the disease and better planning of preventive measures for lifestyle diseases. The study aims to compare Machine Learning (ML) techniques: Logistic Regression, Decision Tree, Random Forest, XGBoost and Support Vector Machine (SVM) with Deep Learning models: Recurrent neural network (RNN) and Neural network for obesity classification using anthropometric data and lifestyle. The dataset which is used here has features of age, gender, weight, height, body mass index (BMI) and physical activity level. Pre-processed the data (scaling, missing value treatment), to deal with them before model training. We determine the performance of each model through accuracy, precision, recall and F1- score with respect to three obesity categories: normal weight, overweight, and obese. Although models like Decision Tree, Random Forest and XGBoost reached an accuracy of 1.0000 which indicates they are overfitting to the data at hand hence their poor generalization capability. Deep learning models, especially the RNN (which captures long-term non-linear dependencies in time), generally outperformed traditional machine-learning techniques thanks to their ability to make more general conclusions without overfitting. The current research unveils the promise in those techniques, with future work required to mitigate overfitting and mixed hybrid models for additional reliability during practical healthcare deployments.

Keywords: Machine learning, Obesity classification, Decision Tree, Deep learning, Logistic Regression, Support Vector Machine (SVM), Random Forest, XGBoost, Recurrent Neural Network (RNN), Neural Network, Overfitting.

1 Introduction

Overweight has become a global problem that has a critical impact on rises in chronic illnesses such as diabetes, cardiac diseases, and certain cancers [1], [2]. This study showed that the incidence of obesity increases with age and that there is a strong indication of the need to identify and categorize individuals early for health care purposes. It is critical to be able to predict obesity on the individual level because the timely prevention and subsequent individualized approach to therapy are among the

key priorities in solving the problem of improving the population's health, as well as in reducing the growth rate of the health care system's costs [3].

The present study aims to use the ML and DL strategies to differentiate obesity based on the various anthropometric and peoples' lifestyle data. Several models such as Random Forest, Decision Trees, Support Vector Machine (SVM), Logistic Regression and the advanced model XGBoost and a Deep Learning model using Artificial Neural Networks shall be applied. The use of these models is meant to measure how well they can assign persons to groups of non-obese, obese and overweight [3], [4].

The data used in this study include the effective factors like age, height, gender, weight, BMI, and physical activity that play part in obesity. Missing values based on data preprocessing techniques will also undergo certain treatments and features to eliminate training the data using ill formatted and treated data [5], [6]. The performance of each model will be assessed legally and metrically via accuracy, precision, recall and F1-score which would give us a clear perception on how each model performs in its predictions.

Therefore, with the help of literature analysis and the comparison between traditional ML models and DL methods, this work will attempt to reveal the advantages and shortcomings of each approach to the chosen problem of obesity classification. The findings from this analysis are expected to add to existing literature on healthcare analytics and enhance future approaches to combating obesity.

2 Literature Review

The machine learning (ML) applications in the given field of healthcare has moved up a notch in disease prediction and therapeutics especially in the context of obesity classification. Over the years, several algorithms in association with neural and fuzzy networks have been worked on, and the effectiveness of each has been demonstrated. K-Nearest Neighbor (KNN) has also witnessed a broad acceptance because it is capable of managing large datasets. Unlike the Support Vector Machine (SVM) that has been popular because of its high ability to manage high dimensionality data. Alzubaidi et al. [3] also used SVM for predicting obesity related health risks where SVM yielded accuracy more than 90%. This showcases how SVM is capable of capturing data relationships hence can be used to classify obesity patients. Supplied its high tolerance to noise and its capacity for scaling up for massive datasets, Random Forest (RF) has also been widely addressed in obesity prediction studies. RF has been described by Rodríguez et al. [4] as superior in terms of features selection and model interpretability and capable of detecting the variable related to obesity. From their study they concluded that RF does not only yield high accuracy but also acts to explain the predictors of obesity. The other basic algorithm often used is the K-Nearest Neighbors (KNN) has been found to work well in many real-life problems. Another work done by the same author, Soni, Patidar and Khandelwal employed and got good results from KNN for obesity prediction [5]. However, the performance of KNN is very sensitive to the choices of distance metrics and the value of k, thus

making them sensitive to the optimization tuning. Logistic Regression (LR), even though it is an older statistical tool, remains the standard for comparison when it comes to analyzing predictions. In a recent study, Nguyen et al. [6] narrated the various applications of LR in aspects of healthcare, and the results displayed high reliability and interpretability. Their research therefore emphasized LR's importance in obesity prediction most especially where the decision-making processes were transparent. New developments in Deep Learning (DL) have taken predictive modelling into other higher levels. Chen et al. [7] described the ability of the neural networks to model nonlinearity where the complexity of the datasets increases. Due to this they are flexible which is important when dealing with healthcare issues such as classification of obesity. In a comparative study done by Ahmed et al. [8] where several algorithms for ML, such as SVM, RF, as well as KNN, were compared and have been used for the prediction of obesity. The research established that while SVM merely delivered the highest common accuracy, RF was consistent in almost all the datasets. As such, it was recommended that model selection was problem dependent in the context of healthcare studies. The prevalent literature also stresses on the integration of basic ML and advanced DL approaches to improve the predictive capability. Stacking and boosting, two other techniques related to ensemble methods, have been used to enhance the performance with the help of effectiveness in various algorithms [9]. Taken together, these investigations highlight important findings in using machine learning methodologies for obesity categorization and support the continued investigation towards achieving better model accuracy and construct clarity for clinical application.

3 Methodology

3.1 Dataset

The data collected for this study include health and lifestyle characteristics from people concerning obesity categorization. It includes several features:

- 1) Age: The number of years he/she has achieved.
- 2) Gender: The sex of the individual; Male/Female.
- 3) Height: The height is the anthropometric parameters of the heights of the individual in centimeters.
- 4) Weight: It represents the weights in kilograms.
- 5) BMI: BMI stands for Body Mass Index. It is the weight in kilograms divided by square of height measured in meters.
- 6) Physical Activity Level: A categorical measure indicating the member's participation level in relation to the group's activities was also used in the studying the amount of physical exercise (from negligible to active).
- 7) Obesity Category: The target class which is going to decide the baseline of obesity category (for instance Normal, Overweight, Obese etc.).

3.2 Preprocessing

Some preprocessing that was done before training the models include:

- 1) Handling Missing Values: Due to the research, some samples have missing values in the Obesity Category variable; thus, samples with missing labels were deleted. In other features, there were no missing values among them and therefore no further action of data imputation was needed [10].
- 2) Data Scaling: Age, Height, Weight, BMI and Physical Activity Level are the continuous variables and after collecting the data, Standard Scaler has been applied on these variables to make them into standardized data. The transformation also scales all the features such that they have a mean of zero and standard deviation of one required for required algorithms like SVM and Neural Networks [11].
- 3) Encoding Categorical Data: The characteristic Gender was one hot encoded so that it could be put into a numerical form suitable for modeling. This method results in creating binary values for each of the categories; it would be easier for the model to use this kind of data [12].

3.3 Models Used

The present research study involves comparative evaluation using seven different Deep Learning (DL) and Machine Learning (ML) models, which are briefly described in the following with reference to their specific features and methodological underpinnings.

- 1) Decision Tree: This model uses a technique that divides datasets according to the values of certain variables. Because of its tree-like structure, the flow of decision making becomes easily understandable from the diagram [8], [18].
- 2) Logistic Regression: Although mainly acknowledged as a statistical method for binary pattern classification, logistic regression is, nonetheless, transferable to multiclass classification paradigms. This model works based on probability estimation with the help of a logistic function that helps to classify numerous outputs [17], [18], [19].
- 3) Random Forest: On the classification front, Random Forest is an ensemble learning method. It builds many decision trees in the training phase and a final classification result is achieved by making use of a voting system. This approach on the same premise works well to improve the rate of prediction as well as reducing the rate of overfitting thus improving the generality of the model [17], [19].
- 4) Support Vector Machine (SVM): It is known for its strengths in classification analysis because it is capable of establishing the right hyperplane that best sets apart data points within a high dimensional space. In linearly non-separable data, this model shows high effectiveness through the use of kernel functions [3], [17].
- 5) XGBoost: As an enhanced gradient boosting work, XGBoost is considered to be characterized with high performance and speed. This is particularly so in classification type of issues especially when dealing with big data scenarios [13][14].

- 6) **Neural Network (Deep Learning):** Like human brains, neural networks are designed to analyze patterns in datasets, and they perform these duties well. In the present study, feedforward neural network with multiple hidden layers was used for the classification of obesity [7].
- 7) **Recurrent Neural Network (RNN):** This architecture is useful for understanding patterns in a sequential data structure due to its deep learning design. As RNNs are designed to make use of sequential information, they are also beneficial for analysis of time series data, or data sets containing consecutive elements wherein several elements depend upon prior elements. In this study, a RNN was trained to handle the sequences of input features relevant to the classification of obesity [15], [16].

Based on the current research, all these models were trained and tested using the same dataset. Using the above metrics of precision, accuracy, F1-score and recall the efficiency of the above said models in classifying the obesity categories was analyzed. These models were deployed using libraries that are widely available, which includes Scikit-learn and TensorFlow/Keras; this guaranteed the method was standardized and could easily be replicated in the study.

3.4 Training and Testing of Model

The methodology of training and testing of the model is very critical in assessing performance as it decides the generalization ability to reach a solution in unencountered data as yet. Below are systematic processes undertaken along with the division of the dataset and effectiveness of model's evaluation:

- 1) **Data Splitting:** The data has been split into 80:20 as train and test. We trained the models on 80 percent of installations and tested them with the rest 20 percent. This division allows the models to be tested with data that they have not yet seen.
- 2) **Model Training:** Logistic Regression, Decision Tree, Random Forest, XGBoost, SVM, RNN and Neural Network models were trained on the training set. With that in mind, the models learned exactly what were such patterns and relationships of data with input features. These updates of the model parameters are what training is all about to minimize some loss function by different optimization techniques like gradient descent for neural net and many more optimizations in machine learning models.
- 3) **Model Testing:** After the training phase, all models were followed by evaluating them using the testing set. Moving further, depending upon the characteristics of the test data, predictions were made which were then compared with actual labels of the target variable – Obesity Category.
- 4) **Performance Metrics:** These are the metrics used to measure the performance of the models:

Precision: Known as actual positive divided by all predicted positive in a classification, precision helps to determine how accurate positive prediction is or actually is.

- Accuracy: It is defined as the no. of correctly classified instances divided by the total no. of test instances.
- Recall (Sensitivity): This is a true positive divided by the actual classes that are positive of the total amount of positive classes, which depicts the expertise of the model on recognizing positive cases.
- F1-Score: In general, the F1-proportion is the mean of specific and generative attendance, which balance is more effective when working with plates with few false positives and false negatives.

These performance metrics were systematically computed for each of the models used hence permitting a good comparison of their efficiency in the prediction of obesity categories. The conclusions and findings of this evaluation are therefore very crucial to identifying the relative merits and drawbacks of each of the models and more importantly to determine the most suitable model of use in classification of obesity.

4 Results

The section of the results then established the measures of performance used in the evaluation of the different various models used in obesity classification. This comprises the display of accuracy as well as the detailed classification report and the comparison of the given models.

4.1 Model Performance

The accuracy scores for each model are summarized in the following table 1:

Table 1. Model Performance Comparison

Model	Accuracy
RNN Model	0.9700
Logistic Regression	0.9884
Decision Tree	1.0000
Random Forest	1.0000
XGBoost	1.0000
SVM	0.9651
Neural Network	0.9942

The following figure (i.e. Fig.1) illustrates the accuracy of deep learning and machine learning models used in the obesity classification.

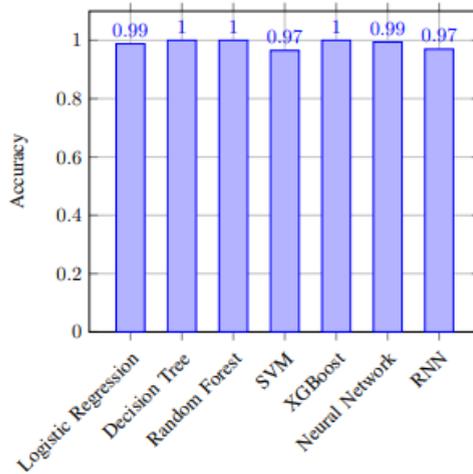


Fig. 1. Comparison of Model Accuracy

4.2 Classification Reports

The classification reports for each model provide insights into precision, F1-score and recall which are crucial for evaluating model performance in detail. The performance of the model's has been given in table 2, 3, 4, 5 and 6

Table 2. Classification Report (Decision Tree)

Class	Precision	Recall	F1-Score	Support
0.0	1.00	1.00	1.00	75
1.0	1.00	1.00	1.00	59
2.0	1.00	1.00	1.00	38
Accuracy	1.00(172)			
Macro avg	1.00	1.00	1.00	172
Weighted avg	1.00	1.00	1.00	172

Table 3. Classification Report (Logistic Regression)

Class	Precision	Recall	F1-Score	Support
0.0	1.00	0.99	0.99	75
1.0	0.97	1.00	0.98	59
2.0	1.00	0.97	0.99	38
Accuracy	0.99(172)			
Macro avg	0.99	0.99	0.99	172
Weighted avg	0.99	0.99	0.99	172

Table 4. Classification Report (Random Forest)

Class	Precision	Recall	F1-Score	Support
0.0	1.00	1.00	1.00	75
1.0	1.00	1.00	1.00	59
2.0	1.00	1.00	1.00	38
Accuracy	1.00(172)			
Macro avg	1.00	1.00	1.00	172
Weighted avg	1.00	1.00	1.00	172

Table 5. Classification Report (SVM)

Class	Precision	Recall	F1-Score	Support
0.0	1.00	0.99	0.99	75
1.0	0.93	0.97	0.95	59
2.0	0.97	0.92	0.95	38
Accuracy	0.97(172)			

Macro avg	0.96	0.96	0.96	172
Weighted avg	0.97	0.97	0.97	172

Table 6. Classification Report (XGBoost)

Class	Precision	Recall	F1-Score	Support
0.0	1.00	1.00	1.00	75
1.0	1.00	1.00	1.00	59
2.0	1.00	1.00	1.00	38
Accuracy	1.00(172)			
Macro avg	1.00	1.00	1.00	172
Weighted avg	1.00	1.00	1.00	172

4.3 Comparative Analysis

From the findings it is clear that there are considerable disparities in the efficiency of the models being compared. In detail, the Decision Tree, Random Forest, and XGBoost models received the highest accuracy of 1.0000. However, this observation gives rise to questions regarding model overfit, since such models can barely generalize sufficiently enough to other unseen data and even more so to a more intricate set of data.

On the other hand, Support Vector Machine (SVM) and Logistic Regression (LR) models returned slightly lower accuracy of 0.9651 and 0.9884 respectively. Nevertheless, these models showed less variance in terms of precision and recall in all the classes and implied more creativity in dealing with the classes.

For the RNN and Neural Network (Deep Learning) models, accuracy levels of 0.9700 and 0.9942 confirmed the ability of the models to capture complexities within the parameters of the data. However, the RNN is differentiated by its ability to handle sequential data, which plays a crucial role in obesity prediction owing to its capability in tracking measured lifestyle changes over time.

Therefore, while the Decision Tree and Random Forest as well as XGBoost models live up to high accuracy, the problem of overfitting is an important concern that should be minded when applying it. However, it is crucial to perform additional experiments on independent validation sets to assess their real potentials. It may be expected that the decision to integrate both approaches belonging to the ML and DL will lead to superior results for obesity classification in realistic scenarios.

5 Discussion

5.1 Model Comparison

The present section provides the reader with the strengths and weaknesses of different types of models of Deep Learning (DL) and Machine Learning (ML) used for obesity classification. Cohort study or comparative analysis is useful in establishing the applicability of various approaches in dealing with various forms of data and features interrelationships.

1. Logistic Regression:

- a) Strengths: The proposed Logistic Regression model was slightly less accurate than the tree models, yet had a desirable and stable accuracy of 0.9884. This model is simple, fast, and easy to interpret, which is convenient for binary and multiclass classification problems, and the prediction function is no more complex than linear or approximately linear.
- b) Weaknesses: The prime limitations of Logistic Regression are it may be hard to effectively train on highly nonlinear relationships between features and labels. What is more, it demonstrates a bit lower results concerning feature interactions compared with such elaborated models as XGBoost and Random-Forest.

2. Decision Tree:

- c) Strengths: From the result, the Decision-Tree model obtained a 1.0000 accuracy score for both sets of datasets used in the findings – training and testing. One feature that makes Decision Trees easily understandable is the fact that they come with easily understandable decision paths. In addition, they can map both nominal and interval data effortlessly.
- d) Weaknesses: Achievement of the perfect accuracy deserves worry owing to its tendency to lead to overfitting. Decision Trees are known to grow excessive complex decision trees that while they can fit the training data points perfectly, they hardly generalize well on other unrelated data sets. This shows that, although the accuracy of the model is high when working with the current dataset, the model may not perform well in a fluctuating and noisy environment.

3. Random Forest:

- e) Strengths: Also like the Decision Tree model, the Random Forest achieved a score of prediction accuracy of 1.0000. Random forest reduces the possibility of over-fitting because unlike the single decision tree, random forests employ an ensemble of decision trees to make the final prediction. Further, it outperforms other methods in terms of handling high-dimensional data, and is less sensitive to outliers.
- f) Weaknesses: Despite the rather general nature of the results due to the ensemble of models, the case with perfect accuracy again leads to overfitting. However, Random Forest is less prone to overfitting than an individual Decision Tree, the enhancements in this study suggest the likelihood of overfitting to the particular

dataset analyzed. More assessment using the unseen data is required in order to support this argument.

4. Support Vector Machine (SVM):

- g) Strengths: This model attained accuracy of 0.9651. They are widely acknowledged as strong classifiers that can provide efficient performance within a small to medium scale of the arrays and as suitable tools for dealing with nonlinearity with the help of the kernel tricks. This characteristic makes the networks especially advantageous especially when the data are not separable by a linear function.
- h) Weaknesses: Although the accuracy of the Support Vector Machine (SVM) model is reasonable the result showed that there was slightly lower accuracy compared to other models. SVMs tend to work very intensively with large datasets, mainly because of the computational complexity of the kernel functions Used, and so many tuning parameters often need to be optimized (like the choice of kernel function, the regularized values etc.).

5. XGBoost:

- i) Strengths: The final accuracy obtained by the XGBoost was 1.000 indicating its efficiency of the gradient boosting algorithm known for its scoring ability in structured data classification. At the same time, this model menacingly copes with cases of missing data and can potentially reveal intricate interconnections between the features during boosting.
- j) Weaknesses: Like the Random Forest and Decision Tree algorithms, reporting of the best possible accuracy indicates overtraining. As we already know, the XGBoost model is complex; thus, it can quickly shift to training data to improve performance, highly likely to be over-fitted, especially when working with fewer datasets. Hence, k-fold cross validation and further tuning of the model is a must so as not to be over-fitted to this particular database.

6. Neural Network:

- k) Strengths: This model gave 0.9942 accuracy which placed it among the other best models. Popular in 'big data' analytics, mainly because they are effective at handling very large data volumes with thousands of features, which can easily reveal highly nonlinear relationships in data through multiple layers of abstraction.
- l) Weaknesses: Neural networks have a heavy computational load which are even more time consuming when training than other models of machine learning. Moreover, he explained that such models are less understandable than, for example, Decision Trees or Logistic Regression, which hinders interpretation of the prediction making processes. Moreover, most of the deep learning models require careful tuning of its hyperparameters and large data sets to avoid overfitting of the algorithm.

7. Recurrent Neural Network (RNN):

- m) Strengths: The RNN model had an accuracy of 0.9700. RNNs are made to capture serial data, hence RNN is perfect for Time Series Analysis or Data with sequence nature in its core. RNNs may be able to capture changing physical activity, dietary habits or lifestyle as patterns in time will be relevant for obesity classification.
- n) Weaknesses: Like any other deep learning model, RNNs require a large amount of computational power to train and have some quirks such as vanishing gradients which need careful tuning. Since RNNs are known to capture the time dependency relatively more efficiently, their accuracy in this study was slightly less than that of the Neural Network (DL), and a dataset-specific formulation suggests non-sequential models could be used with an advantage for any given set.

5.2 Summary of Model Strengths and Weaknesses

A range of predictive models are analyzed in order to determine the various strengths and downsides linked with regard to their performance. However, in our case Random Forest, Decision-Tree and XG-Boost models achieved 1.0 i.e. perfect accuracy but with a highest possibility of overfitting to the dataset which will make predictions less generalizable. On the other hand, Support Vector Machines (SVM) and Logistic Regression yielded a more balanced performance. While these other models achieved slightly lower accuracies, they had the most promise for generalization to unseen data.

Neural Networks (DL) and the especially” promising” Recurrent Neural Networks (RNN), employed a range of techniques suitable for capturing non-linear patterns in data which are much more complex than what is evident with simpler machine learning approaches. On the other hand, these models require significant computational resources and training time takes long, so they may not be widely applicable.

5.3 Insights

In analyzing model capabilities, we find that Random Forests and XGBoost, along with Neural Networks perform well in detecting feature connections making them ideal for datasets with linear relationships. On the other hand models, like Logistic Regression are easier to interpret and less likely to overfit but may struggle with nonlinear data patterns.

Moreover, deep learning algorithms require capacity but can identify intricate patterns essential for analyzing extensive datasets or sequential information. The comparison highlights the balance between clarity of results versus accuracy and speed in computations when choosing a model, for practical uses like predicting obesity trends in the real world.

6 Conclusion

This paper classifies obesity with a combination of varied machine learning and deep learning models, which are: Logistic-Regression, Decision- Tree, Support Vector Machine (SVM), Random-Forest, XG-Boost, Neural Network (Deep Learning), and Recurrent Neural Network (RNN)-and, using precision, accuracy, F1-score and recall metrics, does comparative analysis of the above models to explain the effectiveness of each method.

The analysis yielded several significant findings:

1) High Accuracy but Overfitting: The tree-based models, such as Random-Forest, Decision-Tree and XG-Boost, achieved an exemplary accuracy rate of 1.0000. However, these results suggest that there is a tendency towards overfitting, whereby these models may have learned the noise and particular patterns in the training data, thereby reducing their capacity to generalize well to other data sources.

2) Balanced Performance: Though Logistic Regression and SVM had slightly lower levels of accuracy, these latter models showed much better generalization abilities. Their relative interpretability as well as the lower chance of overfitting make them more suitable to datasets characterized by simpler or linear relationships.

3) Deep Learning Potential: The models of Neural Network (Deep learning) and Recurrent Neural Network (RNN) were capable enough to discover good patterns within the data set with accuracy rates as high as 0.9942 and 0.9700, respectively. These models have some advantageous features to work on non-linear relationships; however, they require more computations and careful adjustments of parameters.

In summation, comparative analysis of the models has unveiled their own merits and demerits respecting the obesity classification.

References

1. World Health Organization: Obesity and overweight. World Health Organization (2021). Available: <https://www.who.int/news-room/fact-sheets/detail/obesity-and-overweight>.
2. Bray, G.A., Bouchard, C.: Handbook of Obesity: Clinical Applications, 3rd ed., Taylor & Francis (2008).
3. Alzubaidi, K., et al.: Obesity Prediction Using Support Vector Machine, 2020.
4. Rodriguez, J. et al.: Feature Importance in Random Forest for Obesity Prediction, Journal of Medical Systems, 2021.
5. Soni, P., et al.: K-Nearest Neighbors for Obesity Prediction. International Journal of Health Informatics, 2019.
6. Nguyen, T., et al.: Logistic Regression Applications in Predictive Modeling, International Journal of Data Science and Analytics, 2020.
7. Chen, M., et al.: Deep Learning for Healthcare: A Review. Journal of Biomedical Informatics, 2021.
8. Ahmed, A., et al.: Comparative Analysis of Machine Learning Algorithms for Obesity Prediction, Healthcare Analytics Journal, 2022.
9. Zhang, H., et al.: Ensemble Learning Techniques for Disease Prediction, International Journal of Machine Learning and Applications, 2019.

10. Han, J., Kamber, M., and Pei, J.: *Data Mining: Concepts and Techniques*, 3rd ed. Elsevier, 2011.
11. Goodfellow, I., Bengio, Y., Courville, A.: *Deep Learning*, MIT Press, 2016.
12. Cortes, C., Vapnik, V.: Support-vector networks, *Machine Learning*, vol. 20, no. 3, pp. 273–297, 1995.
13. Breiman, L.: Random forests, *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001.
14. Chen, T., Guestrin, C.: XGBoost: A Scalable Tree Boosting System, in *Proc. 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 785–794, 2016.
15. LeCun, Y., Bengio, Y., Hinton, G.: Deep learning, *Nature*, vol. 521, no. 7553, pp. 436–444, 2015.
16. Graves, A.: Supervised Sequence Labelling with Recurrent Neural Networks. *Studies in Computational Intelligence*, vol. 385, 2012.
17. N. Yadav and S. Srivastava: Validation of Sophisticated Machine Learning Model for Early Detection of Acute Kidney Injury, *2024 International Conference on Advances in Computing Research on Science Engineering and Technology (ACROSET)*, pp. 1–6, Indore, India (2024).
18. Manisha, M., Srivastava, S. and John, D. P.: Heart Disease Predication Using Various Machine Learning Algorithms. *2024 International Conference on Healthcare Innovations, Software and Engineering Technologies (HISSET)*, pp. 422–428, Karad, India (2024).
19. Sriramoju, S.P. and Srivastava, S.: Enhanced Glioma Disease Prediction Through Ensemble-Optimized Machine Learning. In *2024 International Conference on Communication, Computer Sciences and Engineering (IC3SE)* (pp. 460–465). IEEE, 2024.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits any noncommercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

