Vani Rajasekar
Sathya Krishnamoorthi
Muzafer Saračević
Dzenis Pepic
Mahir Zajmovic
Haris Zogic

# ENSEMBLE MACHINE LEARNING METHODS TO PREDICT THE BALANCING OF AYURVEDIC CONSTITUENTS IN THE HUMAN BODY

**Abstract** *In this paper, we demonstrate the result of certain machine-learning methods like support vector machine (SVM), naive Bayes (NB), decision tree (DT), k-nearest neighbor (KNN), artificial neural network (ANN), and AdaBoost algorithms for various performance characteristics to predict human body constituencies. Ayurveda-dosha studies have been used for a long time, but the quantitative reliability measurement of these diagnostic methods still lags. The careful and appropriate analysis leads to an effective treatment to predict human body constituencies. From an observation of the results, it is shown that the AdaBoost algorithm with hyperparameter tuning provides enhanced accuracy and recall (0.97), precision and F-score (0.96), and lower RSME values (0.64). The experimental results reveal that the improved model (which is based on ensemble-learning methods) significantly outperforms traditional methods. According to the findings, advancements in the proposed algorithms could give machine learning a promising future.*

**Keywords** machine learning, artificial neural networks, diagnose, Ayurveda constituent, support vector machine

**Citation** Computer Science 23(1) 2022: 117–132

## 1. Introduction

Ayurvedic medicine is one of the world's oldest medical sciences – dating back over 5000 years. The term "Ayurveda" is a compound of two words: "Ayur" means existence, and "Veda" means the knowledge that results from the phrase "science of life." In today's high-tech civilization, each human being is trapped by disease at an early stage of life, so it is necessary to know the correct way to live healthily [9]. Each individual has unique features like physical, mental, and emotional characteristics as well as fingerprints (which differ from person to person).

The research of Prakriti aids doctors in guiding their patients toward a safe lifestyle. Prakriti personal study aids in getting to know an individual and his/her needs. Prakriti aids in the preservation of one's health as well as one's personal, family, and professional lives. Ayurveda-Tridoshas indicate three normal energies in the human body or beliefs that define how human body activities function, such as emotional levels, mental stability, etc. VAT, PITT, and KAPH are the three energies that are defined in Ayurveda. Each individual should have a unique combination of all three of these energies. A direct measurement of Ayurveda analytical procedures' and therapies' reliability is still lacking. An accurate analysis contributes to the success procedure. A questionnaire is used to collect a reliable data set of 22 different features that have been accepted by Ayurveda specialists. The existing system does not show the highest accuracy rate, precision, recall, or F-score rate. Ayurveda embraces all living things – human and non-human. Ayurveda discovers a correlation between the use of sensations and infectious agents. The inaccurate use of the senses causes a discord between nature and man, resulting in a mismatch between the two. The foundation principle of Ayurveda is the equality between nature and the human sense [15].

Before starting a treatment for a patient, a thorough examination about the condition of the patient is examined through the techniques that are available in Ayurveda and its advancement. To achieve predictive efficiency, ensemble methods implement various machine-learning algorithms. Ensemble techniques constitute a type of machine-learning methodology that integrates numerous base models in order to create a single best-fit predictive model. Ensemble learning is a technique for resolving difficult computational intelligence challenges. To achieve predictive efficiency, ensemble methods implement various machine-learning algorithms. Unlike machine-learning models, ensemble approaches are not limited to a single learning model [16]; they can being used to allocate trust in a model's determination, pick an optimal subset of features, train incrementally for optimal outcomes, and correct errors. The question of choosing or specifying a set of suitable hyperparameters for a classification algorithm and knowing it is known as hyperparameter optimization or refinement in machine learning. Those values can be estimated from the given data whose configuration variable is internal to the model are called the model parameters. While making predictions, these parameters are required by the model.

Ensemble learning helps to improve machine-learning results by combining several models and analyzing them effectively. This provides better performance and a better approach than a single model. Ensemble learning is used to enhance the classification outcome, prediction value, function approximation in a problem, etc. based on the model's performance (or reducing the likelihood of poor model data). Ensemble learning is used to improve the classification results. An optimized selection of parameters is used to improve the performance of the algorithm. Inter-rater heterogeneity has been addressed in several research studies that were focused on questionnaires, but a quantitative methodology was lacking. In the proposed method, a questionnaire was developed to collect important data. The data from the pilot study was verified for internal consistency before being used to create a trained method that used a machine-learning approach.

The major contribution of this chapter is that our findings were based on an existing data set, as raw data must go through a few measures before being used in machine-learning methods.

- Creating a data set based on a questionnaire collected from a healthy person. The questionnaire contained 484 specimens with 22 attributes and was structured in a compact format with multiple choices for each question.
- Ayurveda experts validated and checked the questionnaires that were collected.
- Ensemble machine-learning models were used to predict the Ayurveda constituents in the human body.
- The proposed approach was compared with various machine-learning models.
- The performance analysis showed the proposed system with high accuracy, precision, recall, and F-score rates.

The remainder of the paper is organized as follows. Section 2 describes related works, and Section 3 describes the proposed ensemble-learning technique. Section 4 describes the results and discussed the proposed method; in addition, the result is compared with state-of-the-art approaches. Section 5 describes our conclusions and future enhancement.

## 2. Related works

Verma et al. [17] looked at different dimensions of Ayurveda and studied the documentation on its history and practice. The authors addressed Ayurveda's growing popularity and importance in health care. They considered the factors that must be addressed for Ayurveda medication to advance and evolve. According to TodkariandLavekar et al. [5], Ayurveda is comprised of deciding an individual's constitution and disparities by using various modular strategies. The authors checked the performance of three separate questionnaires to determine a person's Prakriti standard. R. Chinthala, S. Kamble, and A. S. Baghel et al. proposed that Ayurveda possesses a unique identity by defining some distinct principles (in which Prakriti is one). An assessment of body constitution is the first and foremost factor in "Dashavidha-

parikshas," which consist of ten important aspects to be examined by a physician. Dashavidhaparikshas are very essential for assessing the strength of an individual as well as the severity of the involved doshas. Based on the descriptions that are available in the Ayurvedic literature, Prakriti is a conglomeration of anatomical, physiological, psychological, and sociological characters. Assessing the body constitution that is known as "Deha-Prakriti" is essential for every physician, as it helps when selecting drug or therapeutic procedures, determining drug dosages and modes of administration, and prescribing wholesome diets and lifestyles [4].

Almansour et al. [2] proposed that Ayurveda is a Sanskrit word that means "life science." Its expertise is not limited to medicine, cures, or therapies, and it is intended for laypeople, families, cultures, and physicians alike. Ayurveda and local health traditions have affected each other during their evolutionary developments. The effect of biomedicine on Ayurveda is contributing to its medicalization in contemporary times. This paper investigates how well the medicalization of Ayurveda influences Ayurveda knowledge, learning, and practice [10].

The question of selecting a set of suitable hyperparameters for a classification algorithm is known as hyperparameter optimization or refinement in machine learning. A hyperparameter is a reference for a parameter that is used to guide the learning experience [3]. On the other hand, other factors such as node scales are learned. To generalize distinct information patterns, the same machine-learning algorithm can demand different parameters, scales, or learning speeds. These parameters are known as hyperparameters, and they must be fine-tuned for a model to optimally overcome the machine-learning problem [7]. Dunlap et al. [6] measured Tridoshas (i.e., VATT, PITT, and KAPH) in living things on an ability model from a psychological standpoint. According to the writers, these doshas are made up of Pancha-Mahabhutas. At any given time, one or more doshas can be dominant. There is no situation in which one or more Pancha-Mahabhutas and Tridoshas are completely absent.

In another study, Woldaregay et al. [18] explored machine learning's capacity to address difficult operations in a dynamic world and with expertise attributed to its success. The jobs performed so far in the field of medicine and the use of machine learning in healthcare applications and rehabilitation has sparked interest in determining human Prakriti by using such specialist decision-making methods. To best derive valuable knowledge from Ayurveda, it is necessary to create a modern context in which such hypotheses can be tested, followed by the creation of methods that allow for systematic, quick, reliable, and scalable deployments [11, 12].

In their analysis, Khalil et al. [8] demonstrated that medical claims are full of inconsistencies. The author combed through a repository of patient requests on healthcare platforms that were rife with uncertainty; thus, the returned responses did not contain the requested information. To prevent this, they suggested a complete deep learning-based medical diagnostic framework to assist users in diagnosing diseases. Acharya et al. [1] discussed that bagging (bootstrap aggregation), boosting, and stacked generalization are some of the most frequently used ensemble-learning

algorithms (stacking). Bagging and boosting is a learning model for weak models. This model combines the expectations of many other sub-classifiers to generate testing techniques.

## 3. Proposed methodology

Tridosha refers to the three fundamental energies of VATT, PITT, and KAPH, which decide how our bodies work on a cognitive and psycho-social level. Ayurveda specialists were consulted to help compile a comprehensive data set. The data set is arbitrarily partitioned at a ratio of 7:3 into a training data set of 420 people and a validation set of 180 people using machine learning.

For classification analysis, the classifier is constructed by using various machine-learning algorithms such as support vector machine (SVM), k-nearest neighbor (KNN), naive Bayes (NB), artificial neural network (ANN), decision tree (DT), and gradient boosting algorithms. The flow of the proposed approach is shown in Figure 1.

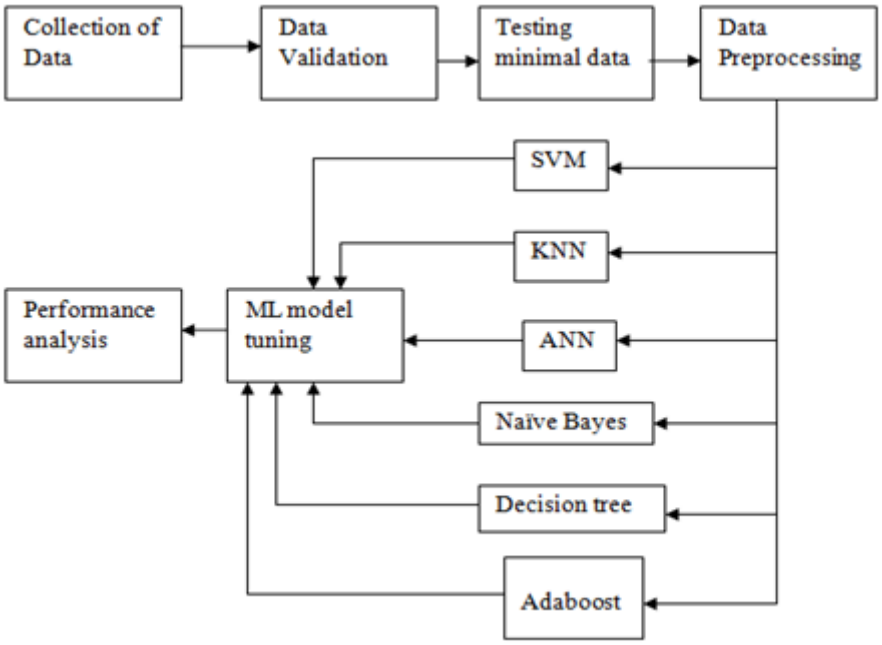


**Figure 1.** Proposed methodologies

### 3.1. Data collection, validation, and testing

Accurate and systematic data collection is a time-consuming undertaking for the training set to maintain accuracy. The methods that were used to collect the data

varied depending on the type of study; these could have been a review of documents, an observation, a conversation, a survey, or a combination of techniques. A questionnaire was designed by the authors to collect data on these traits from different people after determining the scheme and features of all of the distinct groups. Data collection via a questionnaire that was created by the authors is a cost-effective method of gathering information. It is simple to use and gives extensive coverage with minimal effort.

The data set contained 900 examples of 28 attributes that were collected from people between the ages of 20 and 70. An experimental tool that consisted of a series of questionnaires and other invitations were used to collect information from the respondents, and a questionnaire was used to sort the information. The authors undertook pilot research to assure data-gathering consistency. First, randomly chosen respondents from the overall population under investigation were used to test the questionnaire's significance. The authors continued with further data collecting by using the same approaches after attaining adequate accuracy from small-scale applications on this data set. Then, they continued to train and test the entire model based on the outcomes of the pilot research.

## 3.2. Data preprocessing

Only healthy participants were asked to participate, and the data was collected by using unbiased surveying procedures. The respondents for the data collection were chosen by using a simple random-sampling technique. Simple random sampling is based on selecting a small number of people from a larger group. A balanced collection of data points from various classes was seen to be the most beneficial for learning and accuracy. Even minor class imbalances might generate major training issues. The data was split into two sets (training and testing sets). Among the 900 samples, 700 samples were taken for the training; the remaining 200 samples were used for the testing.

## 3.3. Machine-learning models

### 3.3.1. Super Vector Machine (SVM)

SVM is one of the most widely used supervised-learning algorithms, and it can be used for solving both classification and regression challenges. It is mainly used throughout machine learning for classification tasks. The SVM computation's main objective is to build a distance measure that can divide n-dimensional space into categories so that new data points can be conveniently placed in the nominal ledger in the future. A hyperplane is used to describe this best judgment boundary.

SVM selects critical examples that aid in the construction of the hyperplane. SVMs are the names that are given to these extreme fringes, and the implementation is known as a support vector machine. The SVM approach is designed to find a hyperplane in an N-dimensional space that uniquely classifies the sets of data. The goal is to reduce the spacing between the hyperplane and the challenging places at

the decision boundaries as best as possible. The optimization function is specified in Equations (1) through (3):

$$W^T x + b \geq 1 \text{if} y = 1 \tag{1}$$

$$r = y \cdot \left[ \frac{(W^T x + b)}{|W|} \right] \tag{2}$$

$$w = \frac{2}{|W|} \tag{3}$$

where

$d$ – distance between hyperplanes,
$W$ – denoted as margin width,
$x$ – denoted as datapoint,
$y$ – denoted as class label to which it belongs.

### 3.3.2. KNN algorithm

KNN is one of the very few machine-learning techniques that can perform supervised learning. For binary classification, the K-NN algorithm is commonly used. Since K-NN makes no assumptions about the underlying data, it is referred to as a non--parametric methodology. KNN is also known as a lazy-learner method since it does not learn from the training sample right away; instead, it preserves the data set and conducts an operation on it when it comes time to classify it.

Initially deciding on the number of neighbors and the Euclidean distance need to be computed for the k-neighbors algorithm. The Euclidean distance is sorted as a rank-based metric, and the new data points are assigned to each category of the label based on a majority vote. The Euclidean distance measure is specified in Equation (4):

$$E = \sqrt{\sum_{i=1}^{k} (x_i - y_1)^2} \tag{4}$$

where

$E$ – denoted as Euclidean distance,
$(x_i - y_1)$ – denotes distance between two data points,
$i = 1, 2, 3, ..., k$ – denotes number of data points.

### 3.3.3. Naive Bayes algorithm

Naive Bayes is a supervised-learning model that solves classification and regression problems and is based on the Bayes theorem. It is a probability distribution, which means that it forecasts the likelihood of an object. The naive Bayes method treats problem instances as selected features that are grouped into separate groups by using the method. These features are not interdependent, which means that a value that is attributed to one does not affect the values of the others.

The naive Bayes classifier is denoted in Equation (5):

$$P(G|A) = \frac{P(A|G)XP(G)}{P(A)}, \tag{5}$$

where $A$ and $G$ denote the attributes and goals, respectively.

### 3.3.4. Artificial Neural Networks (ANN)

ANNs are algorithms for computing; they were designed to mimic the actions of biological processes that are made up of "neurons." An ANN is made up of a large number of interconnected processing units that work together to process data. The neural network can be used for more than just classification; it can also be used to predict continuous numerical characteristics. After the network has been trained, then it is put it to the test to see how well it generalizes.

ANNs provide input on whether or not the network was effective in classifying the data. The activation function that is used in the proposed approach is the exponential linear unit (ELU) function. The ELU method is used to accelerate neural network training just like a ReLU function. The ELU function's greatest benefit is that it can solve the degradation problem by using identities for positive values and enhancing a model's learning features. Negative ELU values bring the average unit activation closer to zero, reducing the computational complexity and increasing the learning speed. The ELU function is a good substitute for ReLU because it reduces bias changes during training by moving the mean activation to zero. The structure of an ANN is shown in Figure 2.

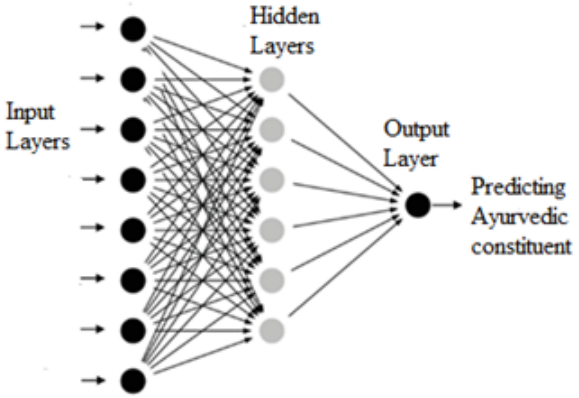


**Figure 2.** Proposed methodologies

A representation of ELU is given in Equation (6):

$$Elu(x) = (x, if x > 0 and exp(x), if x < 0,), \tag{6}$$

where $x$ is individual instance, and learning rate of proposed approach is 0.009.

The derivative of ELU is specified in Equation (7):

$$Elu'(x) = (1, if x > 0 and Elu(x), if x < 0) \tag{7}$$

### 3.3.5. Decision Tree algorithm

Decision tree algorithms are supervised-learning techniques that can be applied to regression and classification. Class labels denote the data set attributes, the branches represent the planning decisions, and each edge represents the outcome. This is a visual representation of all of the feasible solutions or determinations that are supported by certain circumstances. The root node of a decision tree is at the top, followed by lower child nodes, extensions, and internal nodes.

A tree's root binds the various groups. The groups are represented by the leaf nodes, the outcomes by the branches, and the processes by the inner leaves [13]. The routes from the root to the leaves are formed by classification principles. For each attribute entropy, an information gain will be calculated. The next node after the root is selected based on the attribute with a higher information gain.

### 3.3.6. AdaBoost algorithm

The AdaBoost algorithm (also known as adaptive boosting) is a boosting algorithm that is used in machine learning as an ensemble method. The weights are relocated to each instance, with higher weights being allocated to instances that have been wrongly categorized. In supervised learning, boosting is used to reduce bias and variance; this is based on the concept of the sequential growth of learners.

The AdaBoost algorithm learns from the errors of the misclassified data points, and the weights will be updated each time. The steps involved in AdaBoost are specified as follows:

Step 1: initialize weights of data points.

Step 2: train model using decision tree.

Step 3: identify weighted error rate, which is calculated as shown in Eq. (8).

Step 4: calculate weights of decision tree in ensemble learning (Eq. (9)); higher weighted error rate gives lower decision power on voting; lower weighted error rate gives higher decision power on voting.

Step 5: for wrongly classified points, weight values are updated.

Step 6: repeat step until several trees that were decided to train are completed, and finalize decision.

$$e = \frac{No.\,of\,wrong\,predictions}{Total\,no.\,of\,predictions} \tag{8}$$

$$weight = learning\,rate X log((1 - e)/e) \tag{9}$$

### 3.3.7. Hyperparameter tuning

The task of selecting a collection of ideal hyperparameter for a learning algorithm is known as hyperparameter optimization or tuning. A hyperparameter is a value for a parameter that is used to influence the learning process. On the other hand, other factors such as node values are also learned. The accuracy of the model can be adjusted by changing the number of base models or weak learners. More trees must be included to train the model so that the number of weak learners changes from a high bias to a low bias.

## 4. Results and discussions

The proposed method is implemented in net beans using the Java programming language and a system with an i5 processor and 8 GB RAM. Machine-learning models such as SVM, KNN, decision tree, ANN, naive Bayes, and Adaboost are implemented and tested under various conditions. It is critical to assess each algorithm's output, as no single approach is best for every challenge in machine learning. The following five parameters [14] are considered for predicting the effectiveness of the proposed research: TP (true positive), TN (true negative), FP (false pos.), and FN (false neg.):

1. Root mean square error(RMSE): this is used to measure the difference between an expected value and an observed value; it is given in Equation (10), where x is the expected value, and y is the observed value:

$$RSME = \sqrt{(x-y)^2} \tag{10}$$

2. Precision (positive predicted value): correct positive predictions over total positive predictions; this is also denoted as the number of relevant instances that are obtained over the total number of instances:

$$Precision = \frac{TP}{TP+FP} \tag{11}$$

3. Recall (true positive rate): correct positive predictions over actual positive values; this is also called sensitivity, where the best sensitivity value is 1, and the lowest sensitivity value is 0:

$$Recall = \frac{TP}{TP+FN} \tag{12}$$

4. F-Score: a measure of the test accuracy of a data set; this is used to evaluate binary classification systems, which classify examples into positive or negative:

$$F - Score = 2X\frac{PrecisionXRecall}{Precision + Recall} \tag{13}$$

5. Accuracy: a metric that is used by machine-learning models to assess which is the best at finding the correlations and patterns between the observed variables based on supervised information:

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \tag{14}$$

## 4.1. Performance evaluation of proposed methodology

All five of the above-mentioned performance metrics are tested against various proposed machine-learning models. Table 1 shows the results of RSME, Precision, Recall, F-Score, and Accuracy when compared for the various machine-learning algorithms.

**Table 1**
Performance analysis on various machine-learning algorithms

| Classification Models | RMSE | Precision | Recall | F-score | Accuracy |
|---|---|---|---|---|---|
| SVM | 3.4 | 0.83 | 0.89 | 0.77 | 0.81 |
| KNN | 0.96 | 0.79 | 0.82 | 0.84 | 0.79 |
| NAÏVE BAYES | 2.3 | 0.91 | 0.79 | 0.66 | 0.88 |
| ANN | 0.87 | 0.64 | 0.82 | 0.86 | 0.92 |
| DECISION TREE | 1.9 | 0.76 | 0.81 | 0.78 | 0.94 |
| ADABOOST | 0.77 | 0.93 | 0.94 | 0.94 | 0.94 |
| ADABOOST hyperparameter | 0.64 | 0.96 | 0.97 | 0.96 | 0.97 |

The following figures depict the overall results of all of the implemented machine-learning models. Figure 3 represents the RSME values for the various machine-learning algorithms. Figure 4 denotes a comparison of the precision among the various machine-learning algorithms.

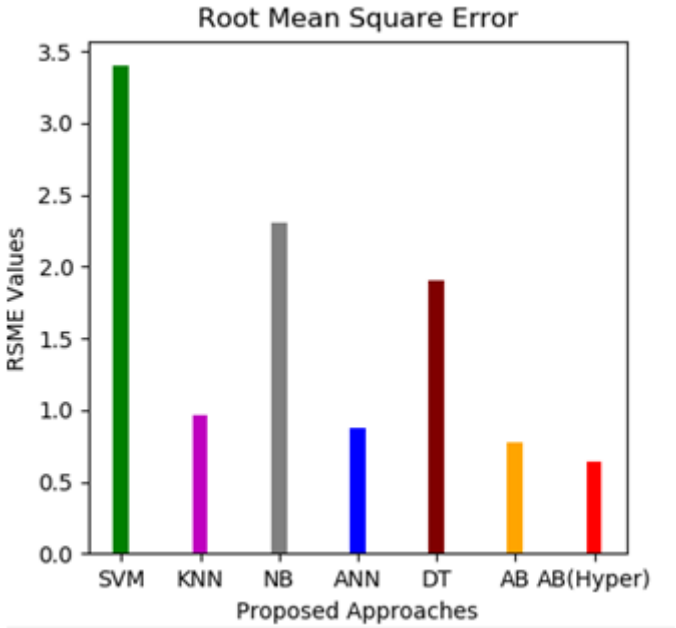


**Figure 3.** Values of RSME for various machine-learning algorithms

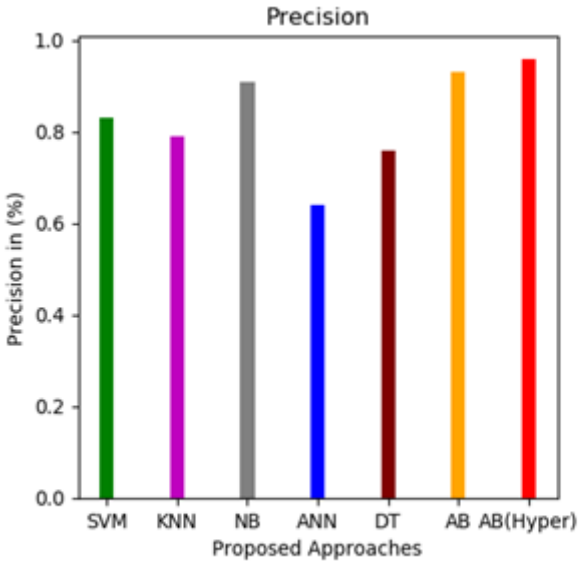


**Figure 4.** Comparison of precision among various machine-learning algorithms

The recall comparison is shown in Figure 5, and the F-score values in Figure 6.

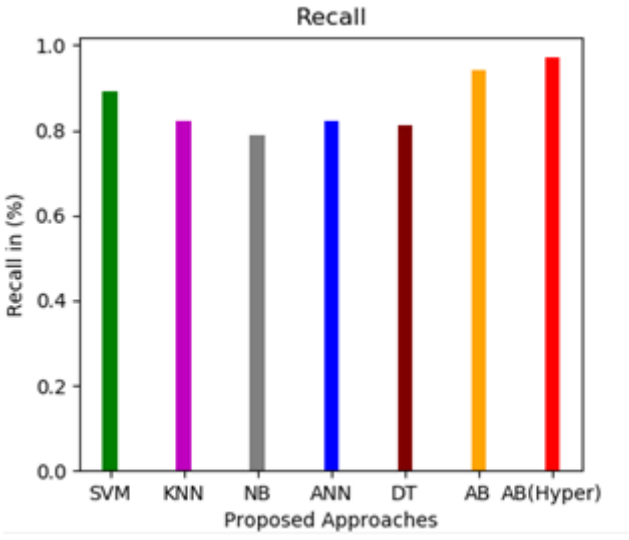


**Figure 5.** Recall comparison

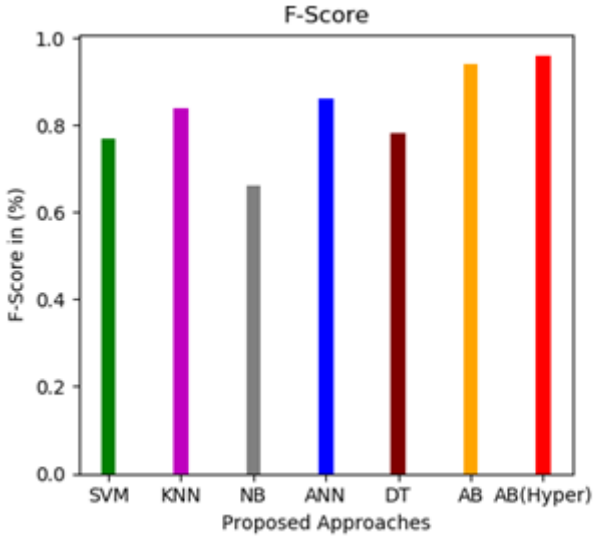


**Figure 6.** F-Score values

From the comparison, it is noted that the Adaboost algorithm with hyperparameter tuning produces a lower RSME value of 0.64, whereas the RSME value of the AdaBoost algorithm is 0.77. ANN shows a considerably lower RSME of 0.97. Among the compared algorithms, SVM shows a higher RSME value of 3.4.

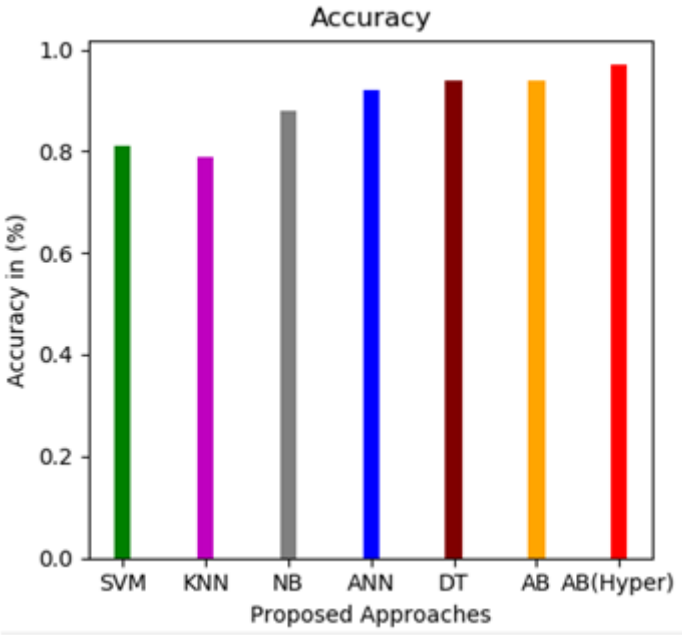


**Figure 7.** Accuracy

The AdaBoost algorithm with hyperparameter shows the highest precision value of 0.96 whereas the precision of the AdaBoost algorithm is 0.93. From the observed five algorithms, the naive Bayes shows a considerably higher precision of 0.91 (which is 0.05 lower than the highest obtained value). The AdaBoost algorithm with hyperparameter shows the highest recall value of 0.97, whereas the decision tree shows the poorest performance of recall of about 0.81 (see Figure 7).

## 5. Conclusion

The main objective of this proposed method is to predict Ayurveda-based constituents in the human body using different machine-learning methods. The data was collected for this approach after discussions with Ayurveda experts. A questionnaire was designed based on the information from the experts, and it was circulated to the end-users. The collected information was validated and implemented with machine-learning algorithms such as SVM, KNN, naive Bayes, ANN, decision tree, Adaboost, and Adaboost with hyperparameters. The result analysis demonstrated that the result of the ensemble-learning method of the AdaBoost algorithm with hyperparameter tuning provides a better result in terms of its better accuracy rate and recall of 0.97. Similarly, the same algorithm provides a better F-score and precision (0.96).

This proposed approach could be used by Ayurvedic medical professionals to help them recognize human constituents. In the future, this approach can be further fine-tuned with the help of transfer and reinforcement learning to identify Ayurvedic-doshas with improved accuracy.

## References

[1] Acharya U.R., Oh S.L., Hagiwara Y., Tan J.H., Adeli H.: Deep convolutional neural network for the automated detection and diagnosis of seizure using EEG signals, *Computers in Biology and Medicine*, vol. 100, pp. 270–278, 2018.

[2] Almansour N.A., Syed H.F., Khayat N.R., Altheeb R.K., Juri R.E., Alhiyafi J., Alrashed S., Olatunji S.O.: Neural network and support vector machine for the prediction of chronic kidney disease: A comparative study, *Computers in Biology and Medicine*, vol. 109, pp. 101–111, 2019.

[3] Andonie R.: Hyperparameter optimization in learning systems, *Journal of Membrane Computing*, vol. 1(4), pp. 279–291, 2019.

[4] Chinthala R., Kamble S., Baghel A., Bhagavathi N.: Ancient archives of Deha-Prakriti (human body constitutional traits) in Ayurvedic literature: A critical review, *International Journal of Research in Ayurveda and Pharmacy*, vol. 10(3), pp. 18–26, 2019.

[5] Dev V.A., Eden M.R.: Formation lithology classification using scalable gradient boosted decision trees, *Computers & Chemical Engineering*, vol. 128, pp. 392–404, 2019.

[6] Dunlap C., Hanes D., Elder C., Nygaard C., Zwickey H.: Reliability of self-reported constitutional questionnaires in Ayurveda diagnosis, *Journal of Ayurveda and Integrative Medicine*, vol. 8(4), pp. 257–262, 2017.

[7] Ibrahim Y., Kamel S., Rashad A., Nasrat L., Jurado F.: Performance Enhancement of Wind Farms Using Tuned SSSC Based on Artificial Neural Network, *International Journal of Interactive Multimedia and Artificial Intelligence*, vol. 5(7), pp. 118–124, 2019.

[8] Khalil R.A., Saeed N., Masood M., Fard Y.M., Alouini M.S., Al-Naffouri T.Y.: Deep Learning in the Industrial Internet of Things: Potentials, Challenges, and Emerging Applications, *IEEE Internet of Things Journal*, 2021.

[9] Madaan V., Goyal A.: Predicting Ayurveda-based constituent balancing in human body using machine learning methods, *IEEE Access*, vol. 8, pp. 65060–65070, 2020.

[10] Mathpati M.M., Albert S., Porter J.D.H.: Ayurveda and medicalisation today: The loss of important knowledge and practice in health?, *Journal of Ayurveda and Integrative Medicine*, vol. 11(1), pp. 89–94, 2020.

[11] Prasher B., Gibson G., Mukerji M.: Genomic insights into ayurvedic and western approaches to personalized medicine, *Journal of Genetics*, vol. 95(1), pp. 209–228, 2016.

[12] Prasher B., Varma B., Kumar A., Khuntia B.K., Pandey R., Narang A., Tiwari P., Kutum R., Guin D., Kukreti R., *et al.*: Ayurgenomics for stratified medicine: TRISUTRA consortium initiative across ethnically and geographically diverse Indian populations, *Journal of Ethnopharmacology*, vol. 197, pp. 274–293, 2017.

[13] Rajasekar V., Jayapaul P., Krishnamoorthi S., Saračević M.: Secure Remote User Authentication Scheme on Health Care, IoT and Cloud Applications: A Multilayer Systematic Survey, *Acta Polytechnica Hungarica*, vol. 18(3), pp. 87–106, 2021.

[14] Rajasekar V., Premalatha J., Sathya K.: Multi-factor signcryption scheme for secure authentication using hyper elliptic curve cryptography and bio-hash function, *Bulletin of the Polish Academy of Sciences: Technical Sciences*, vol. 68, pp. 923–935, 2020.

[15] Roopashree S., Anitha J.: Enrich Ayurveda knowledge using machine learning techniques, *Indian Journal of Traditional Knowledge (IJTK)*, vol. 19(4), pp. 813–820, 2020.

[16] Tiwari P., Kutum R., Sethi T., Shrivastava A., Girase B., Aggarwal S., Patil R., Agarwal D., Gautam P., Agrawal A., *et al.*: Recapitulation of Ayurveda constitution types by machine learning of phenotypic traits, *PloS one*, vol. 12(10), p. e0185380, 2017.

[17] Verma V., Agrawal S., Gehlot S.: Possible Measures to Assess Functional States of *Tridosha*: A Critical Review, *International Journal of Health Sciences and Research*, vol. 8(1), pp. 219–234, 2018.

[18] Woldaregay A.Z., Årsand E., Walderhaug S., Albers D., Mamykina L., Botsis T., Hartvigsen G.: Data-driven modeling and prediction of blood glucose dynamics: Machine learning applications in type 1 diabetes, *Artificial Intelligence in Medicine*, vol. 98, pp. 109–134, 2019.

## Affiliations

**Vani Rajasekar**
Kongu Engineering College, Department of CSE,India, vanikecit@gmail.com

**Sathya Krishnamoorthi**
Kongu Engineering College, Department of CT/UG, India, pearlhoods@gmail.com

**Muzafer Saračević**
University of Novi Pazar, Department of Computer Sciences, Serbia, muzafers@uninp.edu.rs, ORCID ID: http://orcid.org/0000-0003-2577-7927

**Dzenis Pepic**
University of Novi Pazar, Department of Computer Sciences, Serbia, dzenisinedz@gmail.com

**Mahir Zajmovic**
University VITEZ Travnik, Faculty of Information Technology, Bosnia and Herzegovina, mahir.zajmovic@unvi.edu.ba

**Haris Zogic**
University of Novi Pazar, Department of Economic Sciences, Serbia, haris.zogic@uninp.edu.rs

**Received**: 30.05.2021
**Revised**: 05.07.2021
**Accepted**: 19.09.2021