**PREDICTING DIABETES IN HEALTHY POPULATON THROUGH ML**

**Abstract:**

Diabetes is a disease that has no permanent cure; hence early detection is required. Data mining, machine learning (ML) algorithms, and Neural Network (NN) methods are used in diabetes prediction in our research. We used the Pima Indian Diabetes (PID) dataset for our research, collected from the UCI Machine Learning Repository. The dataset contains information about 768 patients and their corresponding nine unique attributes. We used seven ML algorithms on the dataset to predict diabetes. We found that the model with Logistic Regression (LR) and Support Vector Machine (SVM) works well on diabetes prediction. We built the NN model with a different hidden layer with various epochs and observed the NN with two hidden layers

**Keywords:** Decision Trees, Random Forest Logistic Regression, ANN, Hybrid Model

**Introduction:**

The WHO (World Health Organization) reported that around 1.6 million people die due to diabetes every year. Diabetes is one kind of disease that occurs when the blood glucose/blood sugar level in the human body is very high. According to health experts, diabetes occurs when the human body’s gland called the pancreas cannot produce enough insulin (Type 1 diabetes), and the produced insulin cannot be used by the cell of the body (Type 2 diabetes) When we eat food, after the digestion process, glucose gets released. Insulin is a blood hormone that moves from blood to cells and instructs cells to consume blood glucose and transform it into energy. When the pancreas cannot produce enough insulin, the cells cannot absorb glucose, and the glucose remains in the blood. Hence the blood glucose/blood sugar increases in the blood at a very unacceptable level. Due to high blood sugar, some symptom arises in the human body, such as extreme hunger, intense thirst, and frequent urination. The usual range of glucose levels in the human body is 70 to 99 mg per deciliter. If the glucose level is more than 126 mg/dl, it indicates diabetes. A person is considered to have prediabetes if body glucose concentration is 100 to 125 mg/dl If

the human body’s blood sugar level becomes too high, the impending complications can be heart disease, kidney failure, stroke, and nerve damage. There is no permanent cure for diabetes. The most common long-term diabetes causes health problems, which are macrovascular and microvascular complications. The macrovascular complication is damage to the large blood vessels of the heart, brain, and legs. Microvascular complication damages the small blood vessels, causing problems in the kidneys, eyes, feet, and nerve. The efficient control of diabetes is possible if it can be detected early. Maintaining an effective fitness system and balanced eating habits can help to prevent diabetes. If a patient has prediabetes, losing bodyweight by getting physical activity can lower the risk of developing Type 2 diabetes. The Center for Disease Control and Prevention (CDC)-led National Diabetes Prevention Program, a lifestyle change program, can help to change a prediabetes patient’s lifestyle and prevent developing Type 2 diabetes. The healthcare industry collects an enormous amount of data include hospital records, medical records of patients, and results of medical examinations. For early disease diagnosis, the disease’s prediction is analyzed through a doctor’s experience and knowledge, but that can be inaccurate and susceptible. Hence the manual decisions can be alarming. The hidden pattern of data can be unnoticed, which can impact decision-making; therefore, patients become deprived of the appropriate treatment. Automated identification with better accuracy is essential for the early detection of diabetes.

Data mining and machine learning have been developing, reliable, and supporting tools in the medical domain in recent years. The data mining method is used to preprocess and select the relevant features from the healthcare data, and the machine learning method helps automate diabetes prediction. Data mining and machine learning algorithms can help identify the hidden pattern of data using the cutting-edge method; hence, a reliable accuracy decision is possible. Data Mining is a process where several techniques are involved, including machine learning, statistics, and database system to discover a pattern from the massive amount of dataset. According to Nvidia: Machine learning uses various algorithms to learn from the parsed data and make predictions.

**[1].G. Swapna, R. Vinayakumar, K.P. Soman, Soman KP diabetes detection using deep learning algorithms, ICT Express 4 (4) (2018) 243–246, http://dx.doi.org/10.1016/j.icte.2018.10.005, Elsevier B.V.**

Diabetes is a metabolic disease affecting a multitude of people worldwide. Its incidence rates are increasing alarmingly every year. If untreated, diabetes-related complications in many vital organs of the body may turn fatal. Early detection of diabetes is very important for timely treatment which can stop the disease progressing to such complications. RR-interval signals known as heart rate variability (HRV) signals (derived from electrocardiogram (ECG) signals) can be effectively used for the non-invasive detection of diabetes. This research paper presents a methodology for [classification](https://www.sciencedirect.com/topics/computer-science/classification) of diabetic and normal HRV signals using deep learning architectures. We employ long short-term memory (LSTM), [convolutional neural network](https://www.sciencedirect.com/topics/engineering/convolutional-neural-network) (CNN) and its combinations for extracting complex temporal dynamic features of the input HRV data. These features are passed into [support vector machine](https://www.sciencedirect.com/topics/engineering/support-vector-machine) (SVM) for classification. We have obtained the performance improvement of 0.03% and 0.06% in CNN and CNN-LSTM architecture respectively compared to our earlier work without using SVM. The classification system proposed can help the clinicians to diagnose diabetes using ECG signals with a very high accuracy of 95.7%

**Summary:** now a days humans can get unidentified diseases there is no medicine for like diseases as a discussion diabetes is also not having any treatment and medicine to cure but we are take some precautions for prevention

**[2].D. Sisodia, D.S. Sisodia, Prediction of diabetes using classification algorithms, Procedia Comput. Sci. 132 (2018) 1578–1585.**

Diabetes is considered as one of the deadliest and chronic diseases which causes an increase in blood sugar. Many complications occur if diabetes remains untreated and unidentified. The tedious identifying process results in visiting of a patient to a diagnostic center and consulting doctor. But the rise in machine learning approaches solves this critical problem. The motive of this study is to design a model which can prognosticate the likelihood of diabetes in patients with maximum accuracy. Therefore three machine learning classification algorithms namely Decision Tree, SVM and Naive Bayes are used in this experiment to detect diabetes at an early stage. Experiments are performed on Pima Indians Diabetes Database (PIDD) which is sourced from UCI machine learning repository. The performances of all the three algorithms are evaluated on various measures like Precision, Accuracy, F-Measure, and Recall. Accuracy is measured over correctly and incorrectly classified instances. Results obtained show Naive Bayes outperforms with the highest accuracy of 76.30% comparatively other algorithms. These results are verified using Receiver Operating Characteristic (ROC) curves in a proper and systematic manner.

**Summary:** for diabetes identification go the endocrinologists for earlier identification and fallow the precautions to cure there is no medicine for diabetes and fallow some diet and take healthy food

**[3].N.P. Tigga, S. Garg, Predicting type 2 Diabetes using Logistic Regression accepted to publish in: Lecture Notes of Electrical Engineering, Springer.**

Type 2 diabetes is possibly going to be the most significant plague in the history of mankind. If we bring all the diabetic people from around the world together, it can form the third-largest nation in the world. Lately, the predominance of diabetes has altogether expanded in India by 13%. While diet and change in the art of living are the foundation to control type 2 diabetes, most of the patients in the long run require medications to control glucose and related health problems. Therefore, early diagnosis and treatment are important to restrain from future complications. This study uses logistic regression, a popular machine learning classification algorithm to predict the risk of type 2 diabetes among individuals. The aim of this study is to improve prediction so that the logistic regression algorithm can be used on any dataset to give result with good accuracy. The Pima Indian Diabetes dataset is taken for analysis, and RStudio is used to process and visualize the result. Our model is showing pretty good prediction with an accuracy of 75.32%. This study will help the future researchers to develop new interventions to reduce the prevalence of diabetes.

**Summary:** basically in diabetes divided into two types type-1, type-2 both are same but type1 is more chances to get easily type2 is getting some health issues if anyone is suffering

**[4].Q. Zou, K. Qu, Y. Luo, D. Yin, Y. Ju, H. Tang, Predicting Diabetes Mellitus with Machine Learning Techniques, Vol. 9, Frontiers in genetics, 2018, p. 515**

Diabetes mellitus is a chronic disease characterized by hyperglycemia. It may cause many complications. According to the growing morbidity in recent years, in 2040, the world’s diabetic patients will reach 642 million, which means that one of the ten adults in the future is suffering from diabetes. There is no doubt that this alarming figure needs great attention. With the rapid development of machine learning, machine learning has been applied to many aspects of medical health. In this study, we used decision tree, random forest and neural network to predict diabetes mellitus. The dataset is the hospital physical examination data in Luzhou, China. It contains 14 attributes. In this study, five-fold cross validation was used to examine the models. In order to verity the universal applicability of the methods, we chose some methods that have the better performance to conduct independent test experiments. We randomly selected 68994 healthy people and diabetic patients’ data, respectively as training set. Due to the data unbalance, we randomly extracted 5 times data. And the result is the average of these five experiments. In this study, we used principal component analysis (PCA) and minimum redundancy maximum relevance (mRMR) to reduce the dimensionality. The results showed that prediction with random forest could reach the highest accuracy (ACC = 0.8084) when all the attributes were used.

**Summary:** more over the worldday to day life increasing diabetes peoples due to food and vitamins problem for this take healthy food and maintain body fit

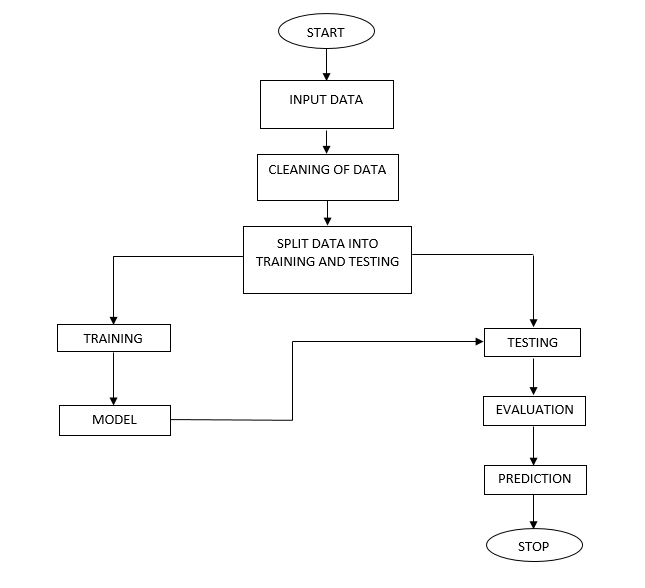
The increasing growth of machine learning, computer techniques divided into traditional methods and machine learning methods. This section describes the related works of Diabetes Prediction Using Different Machine Learning Classifiers and how machine learning methods are better than traditional methods. The existing method in this project have a certain flow is used for model development. Decision Trees, Random Forest, ANN, SVC are used algorithms in existing system. But it requires large memory and result is not accurate.

**Disadvantages:**

* Accuracy low
* Operating cost is high
* difficult to handle

**Proposed System:**

Proposed several machine learning models model to classify Diabetes or not, but none have adequately addressed this misdiagnosis problem. That can be used for this purpose are Stevens Multi a Diabetes Prediction Using Different Machine Learning Classifiers. Also, similar studies that have proposed models for evaluation of such tumours mostly do not consider the heterogeneity and the size of the data Therefore, we propose a machine learning-based approach which combines a new technique of pre-processing the data for features transformation, classifications hybrid algorithm, Logistic Regression. These both ML algorithms give the best accuracy techniques to eliminate the bias and the deviation of instability and performing classifier tests based.

**Fig. Block diagram of proposed method**

**Advantages:**

* Highest accuracy
* Reduces time complexity.

**Methodology:**

**1.RANDOM FOREST:**

**Ensemble methods** is a machine learning technique that combines several base models in order to produce one optimal predictive model. To better understand this definition lets take a step back into ultimate goal of machine learning and model building. This is going to make more sense as I dive into specific examples and why Ensemble methods are used.

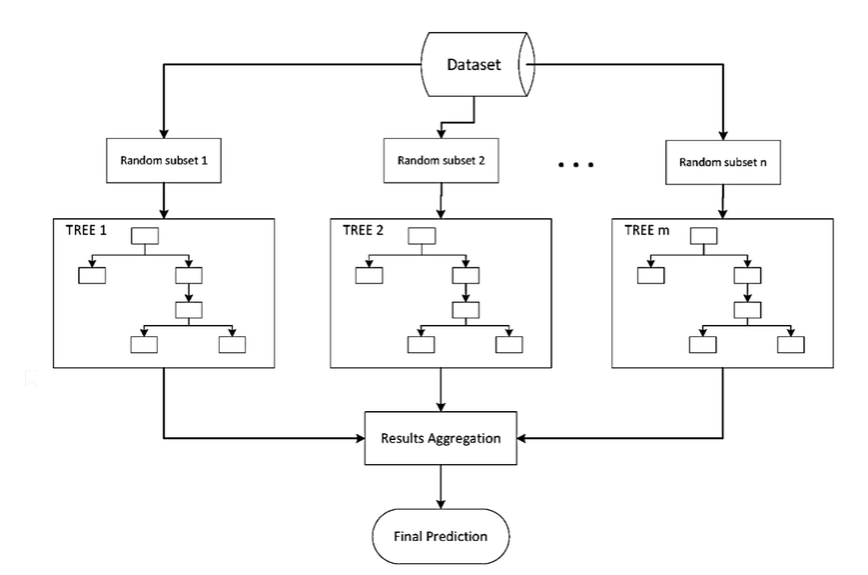
I will largely utilize Decision Trees to outline the definition and practicality of Ensemble Methods (however it is important to note that Ensemble Methods do not only pertain to Decision Trees).

A Decision Tree determines the predictive value based on series of questions and conditions. For instance, this simple Decision Tree determining on whether an individual should play outside or not. The tree takes several weather factors into account, and given each factor either makes a decision or asks another question. In this example, every time it is overcast, we will play outside. However, if it is raining, we must ask if it is windy or not? If windy, we will not play. But given no wind, tie those shoelaces tight because were going outside to play.

When making Decision Trees, there are several factors we must take into consideration: On what features do we make our decisions on? What is the threshold for classifying each question into a yes or no answer? In the first Decision Tree, what if we wanted to ask ourselves if we had friends to play with or not. If we have friends, we will play every time. If not, we might continue to ask ourselves questions about the weather. By adding an additional question, we hope to greater define the Yes and No classes.

This is where Ensemble Methods come in handy! Rather than just relying on one Decision Tree and hoping we made the right decision at each split, Ensemble Methods allow us to take a sample of Decision Trees into account, calculate which features to use or questions to ask at each split, and make a final predictor based on the aggregated results of the sampled Decision Trees.

BAGGing, or Bootstrap AGGregating. **BAGG**ing gets its name because it combines Bootstrapping and Aggregation to form one ensemble model. Given a sample of data, multiple bootstrapped subsamples are pulled. A Decision Tree is formed on each of the bootstrapped subsamples. After each subsample Decision Tree has been formed, an algorithm is used to aggregate over the Decision Trees to form the most efficient predictor. The image below will help explain:



**Random Forest** Models. Random Forest Models can be thought of as **Bagg**ing, with a slight tweak. When deciding where to split and how to make decisions, Bagged Decision Trees have the full disposal of features to choose from. Therefore, although the bootstrapped samples may be slightly different, the data is largely going to break off at the same features throughout each model. In contrary, Random Forest models decide where to split based on a random selection of features. Rather than splitting at similar features at each node throughout, Random Forest models implement a level of differentiation because each tree will split based on different features. This level of differentiation provides a greater ensemble to aggregate over, ergo producing a more accurate predictor. Refer to the image for a better understanding.

**2.DECISION TREE:**

A tree has many analogies in real life, and turns out that it has influenced a wide area of machine learning, covering both classification and regression. In decision analysis, a decision tree can be used to visually and explicitly represent decisions and decision making. As the name goes, it uses a tree-like model of decisions. Though a commonly used tool in data mining for deriving a strategy to reach a particular goal. A decision tree is drawn upside down with its root at the top. In the image on the left, the bold text in black represents a condition/internal node, based on which the tree splits into branches/ edges. The end of the branch that doesn’t split anymore is the decision/leaf, in this case, whether the passenger died or survived, represented as red and green text respectively.

Although, a real dataset will have a lot more features and this will just be a branch in a much bigger tree, but you can’t ignore the simplicity of this algorithm. The feature importance is clear and relations can be viewed easily. This methodology is more commonly known as learning decision tree from data and above tree is called Classification tree as the target is to classify passenger as survived or died. Regression trees are represented in the same manner, just they predict continuous values like price of a house. In general, Decision Tree algorithms are referred to as CART or Classification and Regression Trees.

So, what is actually going on in the background? Growing a tree involves deciding on which features to choose and what conditions to use for splitting, along with knowing when to stop. As a tree generally grows arbitrarily, you will need to trim it down for it to look beautiful. Let’s start with a common technique used for splitting

**3.LOGISTIC REGRESSION:**

Logistic Regression was used in the biological sciences in early twentieth century. It was then used in many social science applications. Logistic Regression is used when the dependent variable (target) is categorical.

For example,

To predict whether an email is spam (1) or (0)

Whether the tumor is malignant (1) or not (0)

Consider a scenario where we need to classify whether an email is spam or not. If we use linear regression for this problem, there is a need for setting up a threshold based on which classification can be done. Say if the actual class is malignant, predicted continuous value 0.4 and the threshold value is 0.5, the data point will be classified as not malignant which can lead to serious consequence in real time.

From this example, it can be inferred that linear regression is not suitable for classification problem. Linear regression is unbounded, and this brings logistic regression into picture. Their value strictly ranges from 0 to 1.

**Purpose and examples of logistic regression:**

Logistic regression is one of the most commonly used machine learning algorithms for binary classification problems, which are problems with two class values, including predictions such as “this or that,” “yes or no” and “A or B.”

The purpose of logistic regression is to estimate the probabilities of events, including determining a relationship between features and the probabilities of particular outcomes.

One example of this is predicting if a student will pass or fail an exam when the number of hours spent studying is provided as a feature and the variables for the response has two values: pass and fail.

Organizations can use insights from logistic regression outputs to enhance their business strategies so they can achieve their business goals, including reducing expenses or losses and increasing [ROI](https://searchcio.techtarget.com/definition/ROI) in marketing campaigns, for example.

An [e-commerce](https://searchcio.techtarget.com/definition/e-commerce) company that mails expensive promotional offers to customers would like to know whether a particular customer is likely to respond to the offers or not. For example, they’ll want to know whether that consumer will be a “responder” or a “non responder.” In marketing, this is called propensity to respond modeling.

Likewise, a credit card company develops a model to decide whether to issue a credit card to a customer or not will try to predict whether the customer is going to default or not on the credit card based on such characteristics as annual income, monthly credit card payments and number of defaults. In banking parlance, this is known as default propensity modeling.

**Uses of logistic regression:**

Logistic regression has become particularly popular in online advertising, enabling marketers to predict the likelihood of specific website users who will click on particular advertisements as a yes or no percentage.

* Logistic regression can also be used in:
* Healthcare to identify risk factors for diseases and plan preventive measures.
* Weather forecasting [apps](https://searchsoftwarequality.techtarget.com/definition/application) to predict snowfall and weather conditions.
* Voting apps to determine if voters will vote for a particular candidate.

Insurance to predict the chances that a policy holder will die before the term of the policy expires based on certain criteria, such as gender, age and physical examination.

Banking to predict the chances that a loan applicant will default on a loan or not, based on annual income, past defaults and past debts.

**Logistic regression vs. linear regression:**

The main difference between logistic regression and linear regression is that logistic regression provides a constant output, while linear regression provides a continuous output.

In logistic regression, the outcome, such as a dependent variable, only has a limited number of possible values. However, in linear regression, the outcome is continuous, which means that it can have any one of an infinite number of possible values.

Logistic regression is used when the response variable is categorical, such as yes/no, true/false and pass/fail. Linear regression is used when the response variable is continuous, such as number of hours, height and weight.

For example, given data on the time a student spent studying and that student’s exam scores, logistic regression and linear regression can predict different things.

With logistic regression predictions, only specific values or categories are allowed. Therefore, logistic regression can predict whether the student passed or failed. Since linear regression predictions are continuous, such as numbers in a range, it can predict the student’s test score on a scale of 0 -100.

**4.SUPPORT VECTOR MACHINES:**

The objective of the support vector machine algorithm is to find a hyper plane in an N-dimensional space (N — the number of features) that distinctly classifies the data points.

**Possible hyper planes :**

To separate the two classes of data points, there are many possible Hyper planes that could be chosen. Our objective is to find a plane that has the maximum margin, i.e. the maximum distance between data points of both classes. Maximizing the margin distance provides some reinforcement so that future data points can be classified with more confidence.

## Hyper planes and Support Vectors

**Hyper planes in 2D and 3D feature space**

Hyper planes are decision boundaries that help classify the data points. Data points falling on either side of the hyper plane can be attributed to different classes. Also, the dimension of the hyper plane depends upon the number of features. If the number of input features is 2, then the hyper plane is just a line. If the number of input features is 3, then the hyper plane becomes a two-dimensional plane. It becomes difficult to imagine when the number of features exceeds 3.

**Support Vectors**

Support vectors are data points that are closer to the hyper plane and influence the position and orientation of the hyper plane. Using these support vectors, we maximize the margin of the classifier. Deleting the support vectors will change the position of the hyper plane. These are the points that help us build our SVM.

**Large Margin Intuition**

In logistic regression, we take the output of the linear function and squash the value within the range of [0,1] using the sigmoid function. If the squashed value is greater than a threshold value (0.5) we assign it a label 1, else we assign it a label 0. In SVM, we take the output of the linear function and if that output is greater than 1, we identify it with one class and if the output is -1, we identify is with another class. Since the threshold values are changed to 1 and -1 in SVM, we obtain this reinforcement range of values ([-1, 1]) which acts as margin.

**Cost Function and Gradient Updates**

In the SVM algorithm, we are looking to maximize the margin between the data points and the hyper plane. The loss function that helps maximize the margin is hinge loss.

Hinge loss function (function on left can be represented as a function on the right)

The cost is 0 if the predicted value and the actual value are of the same sign. If they are not, we then calculate the loss value. We also add a regularization parameter the cost function. The objective of the regularization parameter is to balance the margin maximization and loss. After adding the regularization parameter, the cost functions looks as below.

**Loss function for SVM**

Now that we have the loss function, we take partial derivatives with respect to the weights to find the gradients. Using the gradients, we can update our weights.

Gradients

When there is no misclassification, i.e. our model correctly predicts the class of our data point, we only have to update the gradient from the regularization parameter.

Gradient Update — No misclassification

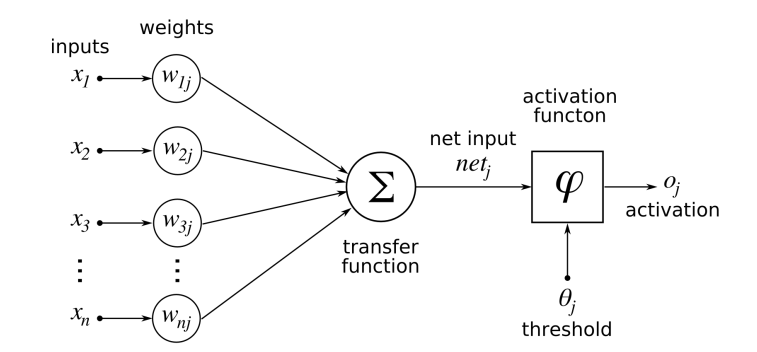
When there is a misclassification, i.e. our model make a mistake on the prediction of the class of our data point, we include the loss along with the regularization parameter to perform gradient update.

**5.ARTIFICIAL** **NEURAL NETWORKS:**

An artificial neural network (ANN) is the piece of a computing system designed to simulate the way the human brain analyzes and processes information. It is the foundation of artificial intelligence (AI) and solves problems that would prove impossible or difficult by human or statistical standards. ANNs have self-learning capabilities that enable them to produce better results as more data becomes available.

An ANN has hundreds or thousands of artificial neurons called processing units, which are interconnected by nodes. These processing units are made up of input and output units. The input units receive various forms and structures of information based on an internal weighting system, and the neural network attempts to learn about the information presented to produce one output report. Just like humans need rules and guidelines to come up with a result or output, ANNs also use a set of learning rules called backpropagation, an abbreviation for backward propagation of error, to perfect their output results.

An ANN initially goes through a training phase where it learns to recognize patterns in data, whether visually, aurally, or textually. During this supervised phase, the network compares its actual output produced with what it was meant to produce—the desired output. The difference between both outcomes is adjusted using backpropagation. This means that the network works backward, going from the output unit to the input units to adjust the weight of its connections between the units until the difference between the actual and desired outcome produces the lowest possible error.

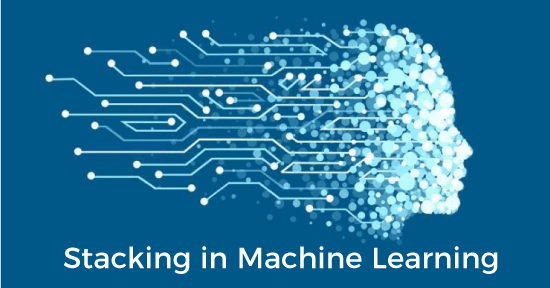


Whenever we increase the layers in our ANN then it is nothing but our Deep Neural Networks.

A **deep neural network** (DNN) is an artificial **neural network** (ANN) with multiple layers between the input and output layers. There are different types of **neural networks** but they always consist of the same components: neurons, synapses, weights, biases, and functions.

# **Stacking in Machine Learning**

There are many ways to ensemble models in machine learning, such as Bagging, Boosting, and stacking. Stacking is one of the most popular ensemble machine learning techniques used to predict multiple nodes to build a new model and improve model performance**.** Stacking enables us to train multiple models to solve similar problems, and based on their combined output, it builds a new model with improved performance.



In this topic, "**Stacking in Machine Learning**", we will discuss a few important concepts related to stacking, the general architecture of stacking, important key points to implement stacking, and how stacking differs from **bagging** and **boosting** in machine learning. Before starting this topic, first, understand the concepts of the ensemble in machine learning. So, let's start with the definition of ensemble learning in machine learning.

## **What is Ensemble learning in Machine Learning?**

Ensemble learning is one of the most powerful machine learning techniques that use the combined output of two or more models/weak learners and solve a particular computational intelligence problem. E.g., a Random Forest algorithm is an ensemble of various decision trees combined.

Ensemble learning is primarily used to improve the model performance, such as classification, prediction, function approximation, etc. In simple words, we can summarise the ensemble learning as follows:

**"**An ensembled model is a machine learning model that combines the predictions from two or more models.”

There are 3 most common ensemble learning methods in machine learning. These are as follows:

* Bagging
* Boosting
* Stacking

However, we will mainly discuss Stacking on this topic.

### **1. Bagging**

Bagging is a method of ensemble modeling, which is primarily used to solve supervised machine learning problems. It is generally completed in two steps as follows:

* **Bootstrapping:** It is a random sampling method that is used to derive samples from the data using the replacement procedure. In this method, first, random data samples are fed to the primary model, and then a base learning algorithm is run on the samples to complete the learning process.
* **Aggregation:** This is a step that involves the process of combining the output of all base models and, based on their output, predicting an aggregate result with greater accuracy and reduced variance.

**Example:** In the Random Forest method, predictions from multiple decision trees are ensembled parallelly. Further, in regression problems, we use an average of these predictions to get the final output, whereas, in classification problems, the model is selected as the predicted class.

### **2. Boosting**

Boosting is an ensemble method that enables each member to learn from the preceding member's mistakes and make better predictions for the future. Unlike the bagging method, in boosting, all base learners (weak) are arranged in a sequential format so that they can learn from the mistakes of their preceding learner. Hence, in this way, all weak learners get turned into strong learners and make a better predictive model with significantly improved performance.

We have a basic understanding of ensemble techniques in machine learning and their two common methods, i.e., bagging and boosting. Now, let's discuss a different paradigm of ensemble learning, i.e., Stacking.

### **3. Stacking**

Stacking is one of the popular ensemble modeling techniques in machine learning. Various weak learners are ensembled in a parallel manner in such a way that by combining them with Meta learners, we can predict better predictions for the future.

This ensemble technique works by applying input of combined multiple weak learners' predictions and Meta learners so that a better output prediction model can be achieved.

In stacking, an algorithm takes the outputs of sub-models as input and attempts to learn how to best combine the input predictions to make a better output prediction.

Stacking is also known as **a stacked generalization** and is an extended form of the Model Averaging Ensemble technique in which all sub-models equally participate as per their performance weights and build a new model with better predictions. This new model is stacked up on top of the others; this is the reason why it is named stacking.

**Conclusions:**

The project has concluded that, to know the Techniques for Predicting the person is diabetes person or they not having diabetes in her body for this we used 6 machine learning based methods such as Random Forest, Decision tree etc. techniques to get better accuracy. By using this algorithms we get the Match Results

**References:**

**REFERENCES:**   
[1] <https://www.who.int/health-topics/diabetes>.

[2]https://www.medicalnewstoday.com/articles/325018#how-is-the-pancre as-linked-with-diabetes.

[3] <https://www.webmd.com/diabetes/diabetes-causes>.

[4]<https://www.mayoclinic.org/diseases-conditions/prediabetes/diagnosis-treatment/drc-20355284>.

[5]https://www.niddk.nih.gov/healthinformation/diabetes/overview/sympto ms-causes.

[6]https://www.diabetes.co.uk/diabetes\_care/blood-sugar-level-ranges.html .

[7]https://www.healthgrades.com/right-care/diabetes/is-there-a-cure-for-di abetes.

[8]https://www.betterhealth.vic.gov.au/health/conditionsandtreatments/dia betes-long-term-effects.

[9] S.A. Kaveeshwar, J. Cornwall, The current state of diabetes mellitus in India, Australas. Med. J. 7 (1) (2014) 45.

[10] https://www.cdc.gov/diabetes/basics/prediabetes.html.

[11] C.L. Huang, M.C. Chen, C.J. Wang, Credit scoring with a data mining approach based on support vector machines, Expert Syst. Appl. 33 (4) (2007) 847–856, <http://dx.doi.org/10.1016/j.eswa.2006.07.007>.

[12] J. Chaki, S. Thillai Ganesh, S.K. Cidham, S. Ananda Theertan, Machine learning and artificial intelligence-based diabetes mellitus detection and self-management: A systematic review, J. King Saud Univ. - Comput. Inf. Sci. (2020).

[13] I. Contreras, J. Vehi, Artificial intelligence for diabetes management and decision support: Literature review, J. Med. Internet Res. 20 (5) (2018) e10775.

[14] G. Swapna, R. Vinayakumar, K.P. Soman, Soman KP diabetes detection using deep learning algorithms, ICT Express 4 (4) (2018) 243–246, http://dx.doi.org/10.1016/j.icte.2018.10.005, Elsevier B.V.

[15] M.W. Craven, J.W. Shavlik, Using neural networks for data mining, Future Gener. Comput. Syst. 13 (2–3) (1997) 211–229, http://dx.doi. org/10.1016/s0167-739x(97)00022-8.

[16] https://blogs.nvidia.com/blog/2016/07/29/whats-difference-artificialintelligence-machine-learning-deep-learning-a

17] T.M. Alam, et al., Informatics in medicine unlocked a model for early prediction of diabetes, Inform. Med. Unlocked 16 (2019) 100204.

[18] D. Sisodia, D.S. Sisodia, Prediction of diabetes using classification algorithms, Procedia Comput. Sci. 132 (2018) 1578–1585.

[19] N.P. Tigga, S. Garg, Predicting type 2 Diabetes using Logistic Regression accepted to publish in: Lecture Notes of Electrical Engineering, Springer.

[20] Salim Amour Diwani, Anael Sam, Diabetes forecasting using supervised learning techniques, Adv. Comput. Sci.: Int. J. [S.l.] (ISSN: 2322-5157) (2014) 10–18, Available at:

.