ISSN No:-2456-2165

# Credit Card Score Prediction Using Machine Learning

Simon Williams
M. Sc IT (Artificial Intelligence) Department of Information Technology University of Mumbai, India.

**Abstract:- This study used a bank loan database to check the applicability of the borrower classification model and examined machine learning techniques. I developed auxiliary vector machine models, decision trees, and random forests, and compared their prediction accuracy with benchmarks based on logistic regression models. They analyzed the performance indicators based on the overall ranking. My results show that the performance of Random Forest is better than other models. In addition, the performance of the support vector machine model is poor when using linear and non-linear kernels. My results show that banks have the opportunity to create value. Improve standard predictive models by researching machine learning techniques.**

***Keywords:*** *Machine Learning, Artificial Intelligence, Supervised learning, Classification, Regression, Tensorflow.*

## I. INTRODUCTION

**Credit Scoring**

The term "credit rating" is used to describe how the customer's default in financial obligations is assessed (Hand & Henley, 1997). The aim is to divide customers into two groups: good and bad. Your financial obligations are thought to be reimbursable. Members of unscrupulous groups are believed to be unable to fulfil their financial obligations.

A credit card is the simplest form of a series of features used to assign a credit rating to customers that shows their risks. The rating can be compared to the loan decision threshold. Because credibility is basically a problem of discrimination (good or bad). The scoring is traditionally divided into two main types according to the tasks and data used (Bijak and Thomas, 2012). Firstly, the level of application used in the loan application process to evaluate the candidate's qualifications. Requesting. The data used for this task are based on the repayment rate of the customer and, at a later date, the good-/bad situation. Banks must precisely predict the possibility of customer defaults in different periods in order to be profitable (1 month, 3 months, 6 months, etc.). High default risk customers can be banned to allow the bank to take adequate measures to protect or control itself against loss.

**MACHINE LEARNING**

The area of research includes a wide variety of fields including computer science, mathematics and information theory, cognitive psychological studies and philosophy (AI), McCarthy et al1955. (Cook & Holder, 2001). Intelligent biological systems are often regarded as prerogative to solutions. Due to the multidisciplinary characteristics of artificial intelligence systems, they reflect multiple technologies and development strategies (Mira, 2008): (ii) formal introduction of these models, (iii) implementation of programming and hardware strategies for those models.

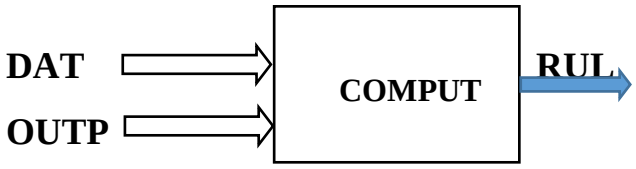
AI-based system is essential for adapting to changes in the environment. The AI discipline for automatic adaptation programmes and data based learning is machine learning. Or experience before (Mitchell, 1997). To this end, an algorithm can be defined defining the sequence of instructions converting input to output (Alpaydin, 2004). Machine learning uses algorithms to differentiate important patterns from irrelevant data patterns. For example machine learning applications involve accurate medical diagnosis (e.g. breast cancer), reliable environmental disaster monitoring (e.g. forest fires) and sensory manufacturing process monitoring (e.g. mechanics).

Machine learning is mainly about supervised learning, and the aim is to learn how input and output are displayed. The data input is data describing a large number of single objects, usually called examples or instances. The result is the supervisor's result or result. Classification is a supervised learning form in which different classes are separated by a mapping (or discriminating function). There are several classes in the results. The results are called class names in machine learning. A classifier is called the discriminating function. The training set is called a series of cases with known class names. The model is defined by many parameters during classification that are optimised for mapping instances to new invisible names of the instance.

One-class classification (OCC) is an appreciation-based method that relies on a single example category to determine the normal or expected behavior of the target category. This is a semi-managed form of classification because the training data contains examples that are only labeled for the target category. This is in contrast to the standard-driven classification method, which uses a discriminative-based method to distinguish examples of different categories. Various real-world problems, such as machine failure detection (Sarmiento et al., 2005), fraud

detection (Juszczak et al., 2008), and authentication (Hempstalk, 2009).

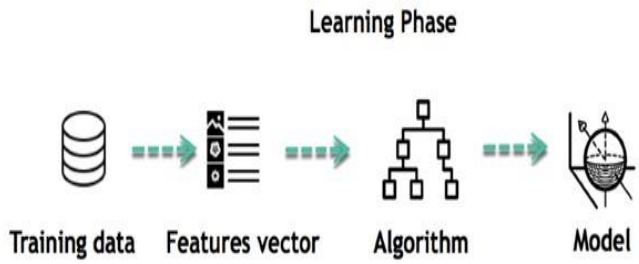
Machine learning is a system that can learn from examples through self-improvement without the need for explicit programming by the programmer. This advancement is based on the idea that machines can learn from data in unique ways (for example, through examples) to obtain accurate results. Combine data with statistical tools to predict results. The company used this discovery to draw practical conclusions. Machine learning is closely related to data mining and Bayesian predictive modelling. The machine receives data as input and uses algorithms to formulate a response. It is recommended to use typical machine learning tasks. For users with a Netflix account, any movie or TV show recommendations are based on the user's historical data. Technology companies are using unsupervised learning to enhance the user experience through personalized guidance. Machine learning is also used for various tasks, such as fraud detection, preventive maintenance, asset portfolio optimization, task automation, etc. Traditional programming is very different from machine learning. In traditional programming, programmers encode all rules after consulting with experts in the industry for which software is being developed. Each rule is based on a reason; after the logic is declared, the machine executes the output. As the system becomes more complex, you need to write more rules. During operation, this can quickly become unstable.




Machine

**How does Machine learning work?**

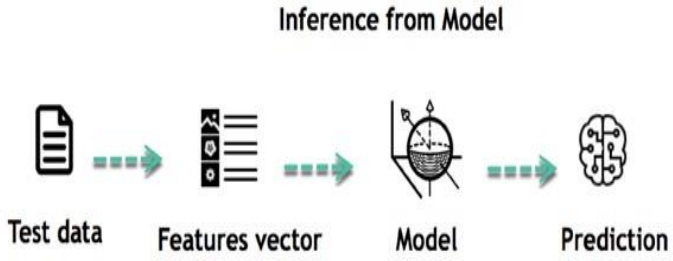
The brain in which all learning takes place is machine learning. Like humans, machines learn. People learn from their own experience. The more we know, the easier it is to predict. For example, when we face an unknown situation. The success rate is lower than the known situation. The way of machine learning is the same. In order to make an accurate prediction, the machine will consider an example. If we give the machine a similar example, it can calculate the result. However, just like humans, it is difficult for machines to make predictions when you provide examples that have never been seen before. Learning and reasoning is the main objective of machine learning. The machine first learns by recognising patterns. Data are attributed to this discovery. As a data scientist, selecting carefully the data you wish to supply to the computer is the most important thing. List of problem-solving attributes. This is known as a vector of function. Feature vectors can be considered a subset of data for solving the problem. In order to simplify reality and translate it into a model, the machine uses complex algorithms. The training phase is therefore used to describe and consolidate the data into the model.




For example, a machine tries to understand the relationship between a person's salary and the possibility of going to a high-end restaurant, and it turns out that the machine finds a positive correlation between a person's salary and going to a high-end restaurant: this is a model.

Inferring

After building the model, you can see how it performs on previously invisible data. Convert the new data into feature vectors, refine the model and make predictions. training. No need to update rules or retrain the model. You can use previously trained models to draw conclusions about new data.

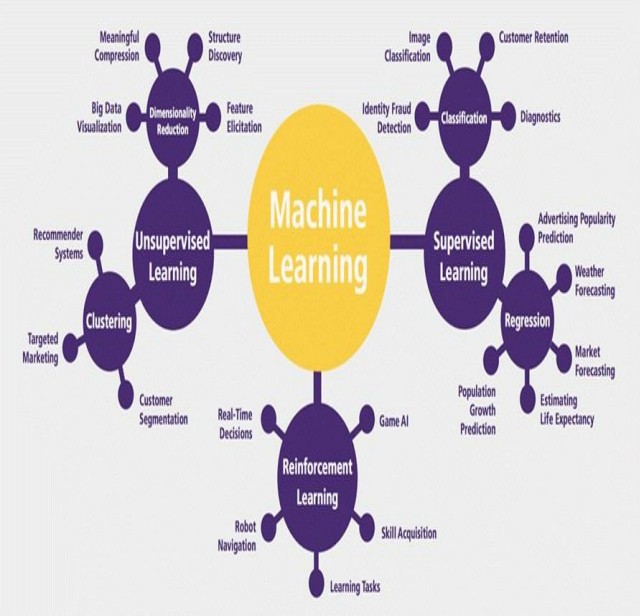



The life of a machine learning programme, summed up as follows

- Define the problem,
- Data gathering,
- Data display,
- Algorithm training,
- Algorithm testing,
- Feedback collection,
- Algorithm optimization.
- Cycle 4-7 to a satisfactory outcome.
- Use the predictive model.

Once the algorithm is successful, the learned knowledge is applied to a new data set.

Machine learning Algorithms and where they are used?




Machine learning may be separated into two types of learning objectives: supervised and unsupervised. There are other different algorithms available.

**Supervised Learning.**

The algorithm finds the association between a certain input and a specified output using feedback and training data. Professionals, for example, can deduct marketing costs. Forecasting sales is done with the use of weather predictions. You can utilise supervised learning when the outcome is known. New data can be predicted by the programme.

There are two types of supervised learning:

Ø Classification problem

Ø Regression problem

**Classification**

Assume you want to anticipate a customer's gender for a salesperson. They'll take your client database's height, weight, employment, salary, shopping cart, and other information. Each customer's gender is known to you. It has to be either male or female. The classifier's goal is to assess whether a person is more likely to be male or female (i.e. tags) based on the information provided (i.e the features you collect).

When the model learns to distinguish between men and women, it may make predictions based on fresh information.

For instance, suppose you've recently gotten fresh information from an unknown customer and you're trying to figure out if it's a male or a woman. If the classifier predicts a male-to-female ratio of 70%, the computer can conclude that this consumer is 70% male and 30% female.

There might be two or more classes in a label. In the case above, however, there are just two classes. When the classifier wants to forecast an item, though, it has hundreds of options (such as glass, table, shoes, etc.). A class is an object).

**Regression**

The issue is regression when the output is a continuous value. Financial analysts, for example, may be required to forecast the value of stocks depending on a variety of factors (such as stocks, past stock returns, and macroeconomic indexes). Errors are possible.

| **Algorithm Name** | **Description** | **Type** |
|---|---|---|
| **Linear regression** | Enables each feature to be correlated to the result to predict future values. | Regression |
| **Logistic regression** | Linear regression extension used for classification tasks. The output variable 3 is binary instead of continuous (e.g. just black and white) (e.g., an infinite list of potential colors) | Classification |
| **Decision tree** | Classification or regression model that separates data-feature values into branches at decision nodes (e.g., if a feature is a colour, each potential colour becomes a new branch) until a final decision output is determined. | Regression Classification |
| **Naive Bayes** | Classification or regression model that separates data-feature values into branches at decision nodes (e.g., if a feature is a colour, each potential colour becomes a new branch) until a final decision output is determined. | Regression Classification |
| **Support vector** | For classification, the Support Vector Machine, or SVM, is commonly employed. The SVM method finds a hyperplane that divides the classes in the most efficient way. It | Regression (not very common) |

| | | |
|---|---|---|
| **machine** | works best when combined with a nonlinear solution. | Classification |
| **Random forest** | The technique is based on a decision tree, which greatly improves accuracy. Random forest builds numerous basic decision trees and choose one label to return using the'majority vote' approach. The final forecast for the classification job will be the one with the most votes, while the final prediction for the regression job will be the average forecast of all the trees. | Regression Classification |
| **AdaBoost** | A classification or regression strategy that assesses a variety of models depending on their accuracy in predicting the outcome before making a choice. | Regression Classification |
| **Gradient-boosting trees** | Gradient-boosting trees are a cutting-edge classification and regression method. It focuses on the preceding trees' mistakes and attempts to fix them. | Regression Classification |

**Unsupervised learning**

The algorithm checks input in an uncontrolled learning without explicit variables of output (for example, checking customer demographics to identify patterns). You can use this to classify the data or to identify patterns for your algorithm and to classify the information.

| Algorithm | Description | Type |
|---|---|---|
| **K-means clustering** | Puts data into groups (k), each of which contains data with a comparable set of attributes (as determined by the model, not in advance by humans) | Clustering |
| **Gaussian mixture model** | A k-means clustering extension that allows for more flexibility in the size and structure of group clusters. | Clustering |
| **Hierarchical clustering** | Forms a categorization system by splitting groups along a hierarchical tree.<br>Can be utilised for customers that have a Cluster loyalty card. | Clustering |
| **Recommender system** | This helps to identify relevant data to make recommendations. | Clustering |
| **PCA/T-SNE** | Usually used to reduce the data's dimensionality. The techniques decrease the amount of features to 3 or 4 greatest variance vectors. | Dimension Reduction |

## II. APPLICATION OF MACHINE LEARNING

**Augmentation**:

- Machine learning can help people solve daily personal or business problems without having to fully control the results. Machine learning is utilised in a variety of applications, including virtual assistants, data analysis, and software solutions. Reduce human error.

**Automation**:

- Machine learning can work completely autonomously in each field without human intervention, such as B. Robots that perform important steps in production facilities.

**Finance Industry**

Ø Machine learning is becoming more and more popular in the financial industry. Banks mainly use machine learning to find data patterns and prevent fraud.

**Government organization**

Ø The government uses machine learning to manage public safety and government services. Taking China as an example, he has a big face. The government uses artificial intelligence to prevent troublemakers from appearin.

**Healthcare industry**

Ø Healthcare is one of the first industries to use machine learning together with image recognition.

**Marketing**

Ø Because of its enormous data access, artificial intelligence is increasingly employed in marketing. Researchers employed powerful mathematical methods such as Bayesian analysis to estimate consumer benefits before the era of big data. The marketing department is relying on AI to optimise customer interactions and marketing operations as data develops.

**Examples of the application of machine learning in the supply chain.**

Machine learning produces great visual pattern recognition results and opens up a slew of new physical inspection and maintenance applications across the supply chain. Unsupervised learning allows you to quickly find comparable patterns in different data sets. During transportation, the machine can perform quality inspections on the damage and wear of the entire logistics centre. For example, the IBM Watson platform can detect damage to shipping containers.

Watson integrates vision and system data for real-time tracking, reporting and recommendations. In the past year, reserve managers have relied heavily on the main methods of estimating and forecasting reserves. By combining big data and machine learning, more advanced prediction methods are introduced (20% to 30% higher than traditional prediction tools). Indicates that it has increased by 2-3% due to possible reduction in storage costs.

Example of machine learning in Google Car.

For example, everyone knows Google's car. The car's roof is full of lasers, which tell the car's position relative to the environment. There is a radar in the front, which can inform the car of the speed and movement of all surrounding cars. With all this data, not only can you figure out how to drive a car, but you can also find and predict how a potential driver will drive a car. The machine processes nearly 1 GB of data per second.

**Deep learning**

Deep learning is a sort of computer software that simulates the brain's network of neurons. Because deep neural networks are used, this is a subset of machine learning known as deep learning. The data is checked at several stages by the machine. The number of layers in the model represents the depth of the model. Deep learning is a relatively recent term in the artificial intelligence field. The learning step of deep learning is carried out via a neural network.

**Reinforcement Learning**

Reinforcement learning is a subset of machine learning that trains the system by receiving virtual "rewards" or "punishments" (essentially trial and error). Google DeepMind used reinforcement learning to defeat the human champion in the game of Go. It also improves gameplay by providing smarter robots in video games.

One of the most famous algorithms:

- Q-learning
- Deep Q network
- State-Action-Reward-State-Action (SARSA)
- Deep Deterministic Policy Gradient (DDPG)

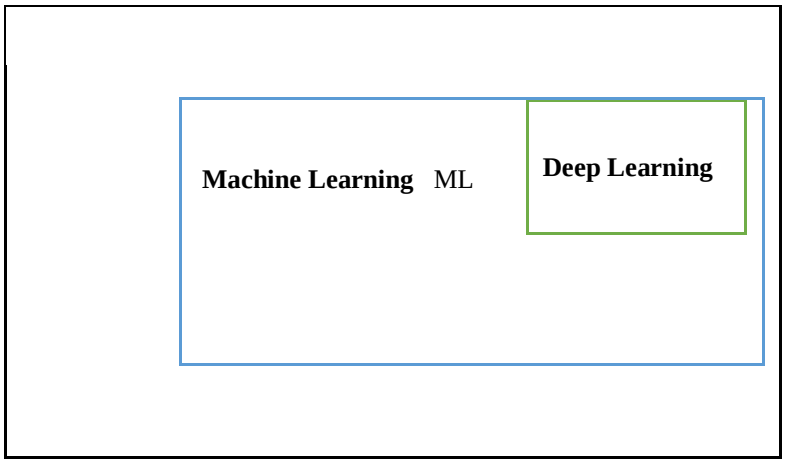
**Applications/ Examples of deep learning applications**

**AI in Finance:** The fintech industry is already using artificial intelligence to save time, reduce costs and increase value. Deep learning changes loans by improving creditworthiness. Use artificial intelligence to better assess risks to assess the types and skills of job applicants. Underwrite is a financial technology company that provides artificial intelligence solutions to lenders. His method is fundamentally superior to traditional methods.

**AI in HR:** Underwear Armour is a sportswear company that is modernizing and modernizing its recruitment experience through artificial innovation. In fact, Under Armour reduced its retail store order time by 35%. The popularity of Under Armour increased in 2012. An average of 30,000 resumes per month. Reading all these apps and starting the selection and interview process took too long. The long process of recruiting and matching employees affected Under Armours' ability to fully staff its retail stores. At the time, Under Armour had all the necessary HR technologies, such as B. Transactional solutions for procurement, implementation, tracking and onboarding, but these tools were not useful enough. Under Armor, choose HireVue (an artificial intelligence provider for human resource issues) for on-demand and real-time interviews. The result is bluffing. They were able to reduce filling time by 35%, thereby hiring more qualified employees.

**AI in Marketing:** Artificial intelligence is a valuable tool for managing customer service and personalization issues. By using artificial intelligence technology to improve voice recognition and call routing in call center management, it can bring customers a smoother experience. For example, deep learning audio analysis. Enable the system to measure the mood tone of customers. If the client does not respond well to the AI chatbot, the system can route the conversation to a real human operator, who will be responsible for solving the problem. In addition to the previous three examples, artificial intelligence is also widely used in other industries.

**Artificial Intelligence**

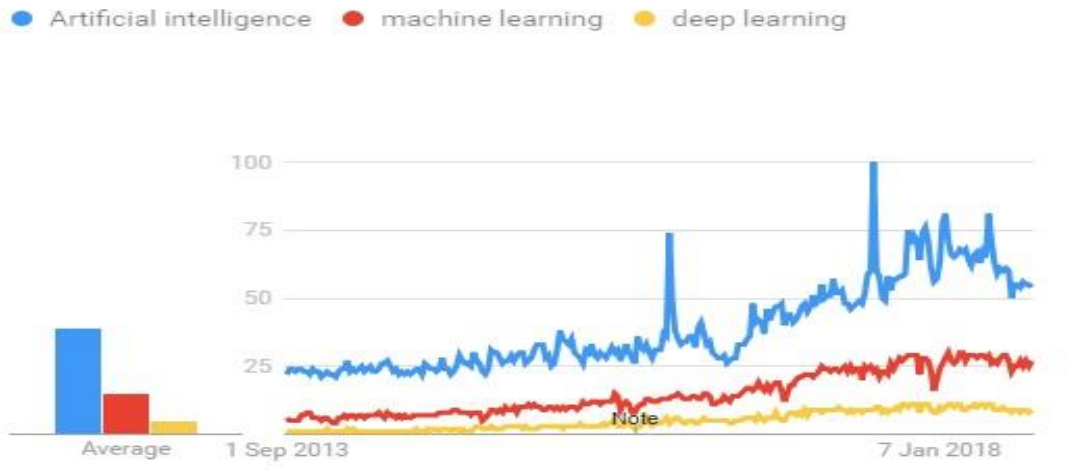


ISSN No:-2456-2165

Difference between Machine Learning and Deep Learning.

| | **Machine Learning** | **Deep Learning** |
|---|---|---|
| **Data Dependencies** | On a small/medium dataset, excellent results were achieved. | Excellent performance with large amount of data. |
| **Hardware dependencies** | Work on a low-cost computer. | DL requires a powerful processor, preferably one with a GPU, because it conducts a lot of matrix multiplication. |
| **Feature engineering** | The characteristics that represent the data must be understood. | There's no need to figure out which feature best describes the data. |
| **Execution time** | From few minutes to hours | Several weeks are possible. A large number of weights must be computed using a neural network. |
| **Interpretability** | Some algorithms are simple to understand (logistic, decision tree), while others are nearly hard to comprehend (SVM, XGBoost) | Difficult to impossible |

Compared with deep learning, machine learning requires less data to train algorithms. Deep learning requires a large number of diverse data sets to define the infrastructure. In addition, machine learning also provides faster model training. Exercise from a few days to a week. The advantage of deep learning over machine learning is that it is very accurate. It is not necessary to know which characteristics the data best represents. Neural networks have learned to highlight key characteristics. With machine learning, you need to choose the features to include in the model.




TensorFlow

Google's TensorFlow is the world's famous deep learning library. The products used to improve search engines, translations, headlines or recommendations in all products by Google machine learning. To give a particular example: We work faster and faster for Google users. Use search enhancement AI. Google will give you suggestions for the next word when the user inserts a keyword in the search bar. Google hopes that its extensive data sets are used by machines and users can enhance their user experience.

Different groups use machine learning:

- Researchers
- Data scientists
- Programmers.

Everyone can work together using the same instruments to improve its efficiency. Google doesn't just have information. You have the world's largest computer, so TensorFlow is largely built. TensorFlow is a Google Brain team library designed to speed up machine training and the exploration of the deep neural networks. It can be used with

ISSN No:-2456-2165

several CPUs or GPUs, including mobile systems. It has multiple conteners, such as Python, C++ or Java, which use different languages.

**TensorFlow Architecture**

Tensorflow architecture works in three parts:

- Preprocessing the data
- Build the model
- Train and estimate the model

Tensorflow gets its name from the fact that it takes input in the form of a multidimensional array (also called a tensor). You can use this information to create a flowchart (also known as a chart) to run. Then proceed through this versatile system to the other end and exit as an exit. Tensors enter, iterate over the list of operations, and then exit on the other side, which is why it's named TensorFlow.

It can be trained on several machines, then executed on an alternative machine, after the trained model has been created.

Both GPUs and CPUs can be utilised to train and run the model. GPUs were created with video games in mind. Stanford researchers discovered in late 2010 that GPUs are also very good at matrix operations and algebra, making them incredibly quick for these types of tasks. A lot of matrix multiplication is used in deep learning. Because TensorFlow is written in C++, it is extremely quick at computing matrix multiplication. TensorFlow, despite being written in C++, can be accessed and controlled using various languages, most notably Python.

Finally, the TensorBoard is an important feature of TensorFlow. The TensorBoard lets you monitor what TensorFlow does graphically and visually.

## III. MAIN CONCEPTS

This section provides a formal expression of the elements involved in supervised learning. In a typical supervised learning environment, a training set $S$ of example $x \in X$ and related output values $y \in Y$ are provided. $X$ is the set of all possible examples in the input space, where $X = \{x_1,. , x_i,. , x_n\}$. Usually each example of $x$ describe with a vector of features or attribute values. Generally, in machine learning text, a function can be regarded as one of two data types:

(i) number: the feature value is a real number; or

(ii) categorical: the feature value is a member of a predetermined finite set.

The statistical text is different in the expansion of data types, including:

(i) Nominal value: the characteristic value is a member of an unordered set, such as {tenant, owner, other};

(ii) ordinal: the characteristic value is ordered The members of the set S. {high, medium, low};

(iii) Interval: the characteristic value is measured in a fixed and equal unit, and is a member of an ordered set, for example The temperature is in degrees Fahrenheit;

(iv) Ratio: The characteristic value has the attribute of the interval data type, but has an absolute zero point (that is, no negative value). For instance, Income to expenditure. It is a collection of all possible outcomes. The value training set S in the output space consists of n tuples (or instances).

$$S = \{(x_1, y_1),. , (x_i,y_i),. , (x_n, y_n)\} \quad (1)$$

It is important to note that assuming that the examples of $x$ in $S$ are independent and the same as (i.) of $X$, $X$ is an unknown but fixed common probability distribution function $P(x, y)$.

**THE LEARNING PROBLEM**

Using the training set S, the goal of supervised learning is to approximate a function $h : X \rightarrow Y$, and the example $x_i$ can be mapped to its output value $y_i$. A learning algorithm commonly called an inductor is used to perform the mapping function. The sensors of a particular training set are called classifiers (Rokach, 2010). The space $H$ of classifiers or functions is called *the space of classifiers or the space of hypotheses*. According to the result of $Y$, the type of learning problem is usually defined as: (i) Regression learning, where $Y = R$; (ii) Learning classification, where $Y = C$, making $C$ represent a set of categories, where $C = \{c_1,.,c_n\}$. The focus of this article is the latter, that is, classification research. It is worth noting that the general task of learning how to classify multiple categories can be decomposed into many binary classification problems (Xu & Chan, 2003). In the binary classification problem, the two categories can be represented as 0 and +1 respectively. For example, the type of borrower in the credit rating (i.e. Error) can be expressed as $Y = \{0, +1\}$.

**MINIMIZING RISK**

In order to select the best classifier for the hypothesis space, the loss function is used as a quantitative measure of the corresponding relationship between the predicted $h(x)$ and the expected result y. The best function h is the minimum expected error (risk),

$$R(h) = L(h(x),y) = \int L(h(x),y) \, dP(x,y) \quad (2)$$

Where L represents the correspondingly selected loss function. For binary classification, the loss function is usually 0/1 loss. If $y = h(x)$, then $L(h(x), y)$ is 0, otherwise it is 1. $(X, y)$ is unknown, and the risk cannot be minimized directly, but the solution is close to the expected minimum error of the available training set S. and is discriminatory. Classification basis (Cunningham et al.) The generation-based method checks common $P(x,y)$ or $P(y|x)P(x)$ probability models, and uses Bayes' theorem to obtain the required posterior probability.

$$P(y|x) = \frac{P(x|y)P(y)}{\sum_j P(x|j)P(j)} \quad (3)$$

ISSN No:-2456-2165

Thus the sum in the denominator is taken over to all categories. Discriminant method research assigns the input attribute *x* directly to the class name *y*, that is, the back of *P (y | x)*.

## IV. SUMMARY

Banks use credit ratings to classify and assess the risk of loss due to actual or perceived changes in customer capabilities and willingness to pay financial liabilities.

Initially, the credibility of private clients was taken as a subjective measure.

The normal operation and improvement of the rating system is an obvious subject of concern to banks, customers and regulators, such as the latest developments in the world economy.

A scorecard is a digital scale used to assign levels to customer characteristics to obtain a value that represents the risk that the customer is considering you have not fulfilled your financial obligations to other customers.

The scorecard development process is divided into three main stages: (i) creating a data set; (ii) modeling; (iii) documentation. This introduces various methods and methods used in the creation and modeling of data sets. The development of a scorecard is a detailed process and many aspects need to be considered. Demographic trends and economic events can cause various situations. Which standards and accepted dashboard design methods are not suitable.

One of these problems arises when the sample contains few default values, which makes it difficult to construct a reliable scorecard. After all, due to the sensitivity of privacy laws and companies, obtaining actual credit information is a challenge for many scholars. With the help of artificial data, scientists can overcome these limitations and create special conditions for studying specific problems.

## REFERENCES

[1]. Li, Xiao-Lin, and Yu Zhong. "An overview of personal credit scoring: techniques and future work." (2012).
[2]. West, Jarrod, and Maumita Bhattacharya. "Some Experimental Issues in Financial Fraud Mining." Procedia Computer Science 80 (2016): 1734-1744.
[3]. Yap, Bee Wah, Seng Huat Ong, and Nor Huselina Mohamed Husain. "Using data mining to improve assessment of creditworthiness via credit scoring models." Expert Systems with Applications 38.10 (2011): 13274-13283.
[4]. Jayagopal, B. "Applying Data Mining Techniques to Credit
[5]. Hand, David J., and William E. Henley. "Statistical classification methods in consumer credit scoring: a review." Journal of the Royal Statistical Society: Series A (Statistics in Society) 160.3 (1997): 523-541.
[6]. Sabeti, Behnam, et al. "Credit Risk Rating Using State Machines and Machine Learning." *International Journal of Trade, Economics and Finance*, vol. 11, no. 6, 2020, pp. 163–168.
[7]. Wallis, Mark, et al. "Credit Rating Forecasting Using Machine Learning Techniques." Advances in Data Mining and Database Management Managerial Perspectives on Intelligent Big Data Analytics, 2019, pp. 180–198.
[8]. Tsai, Chih-Fong, and Ming-Lun Chen. "Credit Rating by Hybrid Machine Learning Techniques." *Applied Soft Computing*, vol. 10, no. 2, 2010, pp. 374–380.
[9]. Weston, Liz Pulliam. *Your Credit Score: How to Improve the 3-Digit Number That Shapes Your Financial Future*. FT Press, 2016.
[10]. *Understanding Your Credit Report and Credit Score*. Financial Consumer Agency of Canada, 2012.