

Survey on Classification Algorithms in Machine Learning

Prachi Vilas Kothawade
 Department of Computer Science
 School of Engineering and Applied
 Sciences, University of Mumbai
 Kalyan, India

Dr. Uttara Gogate
 Department of Computer Science
 School of Engineering and Applied
 Sciences, University of Mumbai
 Kalyan, India

Dr. Satishkumar Varma
 Department of Computer Science
 School of Engineering and Applied
 Sciences, University of Mumbai
 Kalyan, India

Abstract— Supervised learning problems can be classified into regression problems and classification problems. A problem is said to be a regression problem when outputs are continuous whereas a classification problem is when outputs are categorical. There is a wide range of machine learning applications where classification techniques are used. So, this study attempts to compare different types of classification algorithms. We will discuss some of the major types of classification methods, including Bayesian networks (BN), Naïve Bayes (NB), Stochastic gradient descent (SGD), Decision tree (DT), k-nearest neighbor (KNN) classifiers, Neural network, Random forest (RF), Logistic Regression (LR), Support vector machines (SVMs) and their advantages, disadvantages, and applications. The purpose of this study is the comprehensive review of various classification methods in machine learning. This work is useful for both Academia and newcomers in the field of machine learning to further strengthen the basics of classification methods.

Keywords—Machine Learning, Classification Algorithms, Machine Learning Applications

I. INTRODUCTION

Machine learning, by definition [1], is a field of computer science that evolved from the study of pattern recognition and computational learning theory in artificial intelligence. This is the training and construction of algorithms that can be learned and predicted from datasets. These steps work by building a model from an example input to make data-driven predictions or selections, rather than following the instructions of a solid static program [2]. With hundreds of statistics-based algorithms built into machine learning, choosing the right algorithm or combination of algorithms for your job has always been a challenge for anyone working in this area. However, it is important to understand three comprehensive categories of machine learning before considering a particular algorithm. These three categories are

1. Supervised
2. Unsupervised
3. Reinforcement.

In supervised learning, make predictions using a known dataset, also known as a training dataset. The training dataset contains input data and response values. From there, the algorithm attempts to build a model that can predict the response value of the new dataset. The program is "trained" in a predefined set of "training examples" to promote the ability to

reach accurate conclusions when given new data. Supervised techniques can be further divided into two main categories. Classification and regression. In regression, output variables take continuous values, but in classification, output variables take class labels [3]. Classification is used to classify a record. It is used for questions which can have only a limited number of answers (fix no of output (1/0)). Regression algorithms are used to calculate numeric values.

Unsupervised learning is an independent learning process. For unsupervised learning, not all variables and data patterns are categorized. Instead, the machine should use an unsupervised learning algorithm to reveal hidden patterns and create labels. This approach is based on absolute error measurement because there is no supervisor. This is useful if you need to learn how to group (cluster) a set of elements according to their similarity (or distance scale). Unsupervised Learning can be further divided into Clustering and Anomaly detection algorithms. Anomaly detection is a technique for identifying rare events or observations that can be suspicious because they are statistically different from other observations. Clustering is use to find structure behind problem. Clustering is similar to a classification algorithm but uses a different type of dataset. Classification uses labeled datasets, while clustering uses unlabelled datasets.

Reinforcement learning is based on feedback provided by the environment, even in the absence of a real supervisor. However, in this case, the information is more qualitative and does not help the agent determine accurate measurements of the error. In reinforcement learning, this feedback is usually called reward (negative feedback is sometimes defined as a penalty) and helps to understand whether a particular action taken in a given condition is positive.

Classification of input data is a very important process in machine learning, such as whether the email is genuine or spam, and whether the transaction is fraudulent, etc. hence, we will go through the most commonly used classification algorithms in Machine Learning.

II. CLASSIFICATION ALGORITHMS

Classification is the process of classifying a particular dataset into classes that can be run on both structured and unstructured data. The process starts by predicting the class of a particular data point. Class is usually referred to as target, label, or category.

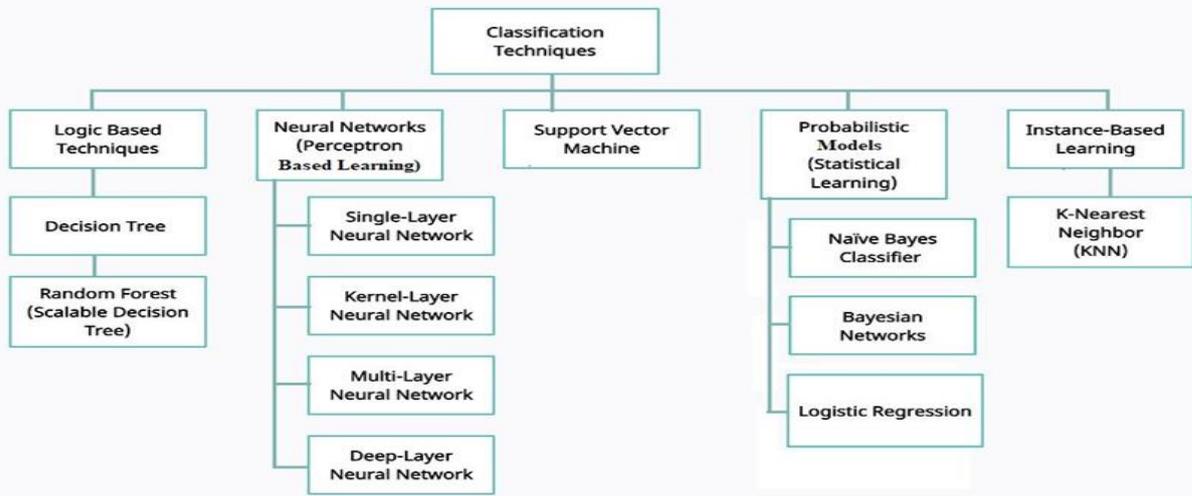


Fig. 1. Classification algorithms

A. Probabilistic models for classification

Statistics-based algorithms generalize problems with help of distributive statistics and look into the distribution structure to continue the predicting task. Here, Naïve Bayes has been explained as a popular example of statistics-based algorithm. This is a collection of classification techniques based on Bayesian theorem of probability. [4]

Graphical model [5] is a probabilistic model in which a graph shows a conditional independent structure between random variables. Graphical models provide an easy way to visualize the structure of stochastic models and can be used to design and motivate new models.

1) Naïve Bayes Classifier

The naive Bayes classifier generates probabilities for all cases. Then predict the highest probability of outcome. The naive assumption is that the features are independent and that the classification of all pairs of features is independent of each other. [4] The naive Bayes classifier is a simple stochastic classifier based on applying Bayes' theorem with a strong (naive) independence assumption.

Advantages	Disadvantages
A small size of training data is sufficient here to estimate the required parameters Classifier is very fast Can be used to solve diagnostic problems efficiently [4]	Known to be a bad estimator [4]

2) Logistic Regression

This is a classification function built using classes and using a single multinomial logistic regression model with a single estimator. Logistic regression usually indicates where boundaries between classes exist, and certain approaches also indicate that the probability of a class depends on the distance from the boundary. This means that the larger the dataset, the faster it will move towards the extremes (0 and 1). These statements about the probability of making logistic

regression more than just a classifier. Logistic regression uses the concept of predictive modeling as regression. Therefore, it is called logistic regression and is used to classify samples. Therefore, it corresponds to the classification algorithm

Advantages	Disadvantages
highly interpretable easy to use easy to normalize, & is regularized, very efficient [6]	cannot solve nonlinear problems relies heavily on the proper display of data vulnerability to overfitting [6]

3) Stochastic Gradient Descent

After explaining the basics of logistic regression, it is useful to introduce an SGD classifier class that implements a very well-known algorithm that can be applied to several different loss functions. The idea behind stochastic gradient descent is to repeat weight updates based on the gradient of the loss function. Stochastic gradient descent is a very popular algorithm used in many machine learning algorithms, and most importantly, it forms the groundwork of neural networks.

Advantages	Disadvantages
Computation is fast For larger datasets, the parameters are updated more frequently, resulting in faster convergence	noisy steps computationally expensive

4) Bayesian Networks

Bayesian Network (BN) refers graphical model for probability associations between a set of variables [7]

Hidden Markov Model

Here, we introduce the hidden Markov model, which is a special case of the Bayesian network. In the hidden Markov model, the hidden states are linked as a chain and dominated by the Markov process. The observable values are then generated independently given the hidden states that form the sequence. Hidden Markov models are widely used in speech recognition, gesture recognition, and part-of-speech tagging, and hidden classes are interdependent. Therefore, the class label of an observation depends not only on the observation, but also on the adjacent states.[5]

Markov Random Fields (MRF)

Next, we look at another major class of graphical models that are written in undirected graphs and specify both factorization and a set of conditional independence relationships. The

Markov Random Field (MRF), also known as an undirected graphical model, has a set of nodes, each node corresponding to a variable or group of variables, and a set of links, each connecting a pair of nodes. [5]

Conditional Random Fields (CRF)

So far, we have explained the Markov network representation as a joint distribution. One notable variant of MRF is the Conditional Random Field (CRF). In this case, each variable may be conditioned on a set of global observations. More formally, the CRF is an undirected graph, and its nodes can be divided into two disjoint sets of observation variable X and output variable Y . These can be parameterized as a set of factors in the same way as a regular Markov network. The main advantage of CRF over HMM is its conditional nature, which relaxes the independence assumptions that HMM requires to ensure manageable inference. [5]

Advantages	Disadvantages
Flexible applicability: can be used to solve both regression and classification problems Handling missing data Smoothness characteristics: Minor changes in the Bayesian network model do not affect the behaviour of the system	noise lack of information awareness of changing attributes for class variables

B. Logic based learning

Logic-based algorithms address the problem of step-by-step data streaming by letting the logic work at each step. Here decision tree has been overviewed as a classical example of logic-based algorithm.

1) Decision Tree

Decision tree algorithms are most commonly used algorithms in classification. Decision tree provides an easily understandable modeling technique and it also simplifies the classification process [3]. The decision tree is transparent mechanism it facilitates users to follow a tree structure easily in order to see how the decision is made. The core objective of decision tree is to produce a model that calculates the value of a required variable based on numerous input variables [3]. Usually, all decision tree algorithms are constructed in two phases [3]:

- (i) tree growth; in which training set based on local optimal criteria is splitting recursively until most of the record belonging to the partition having same class label
- (ii) tree pruning; in which size of tree is reduced making it easier to understand.

ID3, C4.5, CART are examples of widely used decision tree algorithms.

In decision tree learning, ID3 (Iterative Dichotomiser 3) is an algorithm invented by Ross Quinlan that is used to generate decision trees from datasets [8]. A tree is built using the decision tree technique to model the classification process. When the tree is built, it applies to each tuple in the database, and as a result, that tuple is categorized. [8] It is one of the most widely used algorithms in the fields of data mining and machine learning because of its effectiveness and simplicity [9]. The ID3 algorithm is based on information gain

C4.5 is a popular algorithm for decision tree formation. This is an extension of the ID3 algorithm that minimizes the

shortcomings caused by ID3 [3]. The decision tree generated by the C4.5 algorithm can be used for classification. For this reason, C4.5 is also known as a statistical classifier [8]

CART stands for Classification and Regression Trees (Breiman et al., 1984). This is characterized by the fact that it builds a binary tree. That is, each internal node has exactly two output edges. Splits are selected using towing criteria and the obtained tree is pruned by cost and complexity pruning. When provided, CART can take into account the cost of misclassifying tree leads. Users can also provide prior probability distributions. [8]

Advantages	Disadvantages
Ease of interpretation [10] can easily handle heterogeneous data simple fast require less data pre-processing, can process both categorical and numeric data. [4]	risk of overfitting can lead to complex tree structures that are not well generalized. Noise Impact Accuracy is often not state-of-the-art [10]

2) Random Forest

As the name implies, "Random forests are classifiers that contain a large number of decision trees in different subsets of a particular dataset and take an average to improve the predictive accuracy of that dataset." Instead of relying on one decision tree, it gets the predictions from each tree based on the majority of the predictions and predicts the final output.

It is based on the concept of ensemble learning. This is the process of combining multiple classifiers to solve complex problems and improve model performance. Many trees can outperform the performance of individual trees by reducing the errors that typically occur when considering a single tree. If one tree doesn't work, the other tree may work. This is an additional benefit, and this ensemble formed is known as Random Forest. [11]

Advantages	Disadvantages
Can handle large, high-dimensional datasets. Improve model accuracy and prevent overfitting robust and versatile algorithm. Can be used to handle missing values in specified data [11]	High calculation cost. Difficult to interpret Creating forecasts is often very time consuming [11]

C. Support Vector Machine

Definition: This represents a dataset item or record with "n" features plotted as points in n-dimensional space separated into classes by the widest possible margin known as a hyperplane. The data items are then mapped to the same n-dimensional space, and predictions of the categories to which they belong are obtained based on the aspects of the hyperplane to which they belong. [4]

SVMs are considered one of the most prominent and useful techniques for solving problems related to data classification and learning and prediction. Given a series of training examples, each marked as belonging to one of two categories,

the SVM training algorithm builds a model that predicts whether the new example will fall into one of the categories. [3]

There are many types of SVMs, including linear, polynomial, and sigmoid.

Advantages	Disadvantages
memory efficient [12] can handle a variety of classification problems, including high-dimensional problems that are not linearly separable [3]	not suitable for large datasets [12] does not work very well if the dataset is noisy [12] training time is so long [4]

D. Artificial Neural Network

Artificial neural networks, also known as neural networks, are a common machine learning technique for processing data through layers of analysis. A typical neural network can be divided into an input layer, a hidden layer, and an output layer. The data is first received by the input layer, where a wide range of features are detected. The hidden layer then analyses and processes the data. Based on previous calculations, the data is streamlined by going through each hidden layer. The final result is displayed as an output layer.

Neural networks are applied when you do not know the exact nature of the input-output relationship. An important feature of neural networks is that they learn the relationship between inputs and output during training.

1) Single-Layer Neural Network

It has one input layer and one output layer. The simplicity of these networks makes them suitable for introducing the internal behavior of neural networks. Some standard single-layer networks widely used for classification problems are perceptron, Adaline networks, and linear vector quantization (LVQ).[5]

2) Kernel-Layer Neural Network

We used the term kernel-based neural network for classes of neural networks that use the kernel within the architecture. A well-known example of this type is the radial basis function network. Radial basis function networks (RBFNs) implement a two-layer architecture featuring a hidden-layer kernel unit and an output-layer linear unit. [5] RBFNs can achieve powerful classification functions with a relatively small number of units. As a result, RBFNs have been successfully applied to various data classification tasks such as handwritten digit recognition, image recognition, speech recognition, process failure detection, and various pattern recognition tasks

3) Multi-Layer Neural Network

A Multilayer Perceptron (MLP) consists of an input layer and an output layer, with one or more hidden layers of non-linearly activated nodes or sigmoid nodes. This is determined by the weight vector and the network weights need to be adjusted. This is a classifier that finds network weights by solving quadratic programming problems using linear constraints, rather than solving non-convex, unconstrained minimization problems as in standard neural network training.

4) Deep-Layer Neural Network

Deep learning techniques aim to learn and leverage the capabilities of neural networks to generate internal feature representations of the data in hidden layers. The output layer can be thought of as a spatial classification machine that spans the output of the last hidden layer. More hidden layers and more units within those layers generate a stronger network with the ability to generate a more meaningful internal representation of the data.

Advantages	Disadvantages

Store information throughout the network work with incomplete knowledge Fault Tolerance Parallel processing function [13]	Hardware Dependencies Network unexplained behavior Determining the proper network structure [13]
--	--

E. Instance-Based Learning

"Instance-based learning" that delays the generalization process until the classification task is performed. This algorithm is tagged as "lazy" and is widely known as the lazy learning algorithm. The KNN or K-nearest neighbor algorithm is a popular case of instance-based learning. KNN is one of the most understandable and very simple classification algorithms. Despite its simplicity, it can produce very competitive results. It can handle classification and regression type prediction problems [4]

1) K-Nearest Neighbors

A supervised learning technique used to classify new data points based on their relationship to nearby data points. Methods in another class can be associated with statistical methods. These are called nearest neighbor methods or, in some cases, memory-based methods. The k-nearest neighbor classifier considers the k-nearest neighbor points and uses the majority of the signs to classify the samples.

Advantages	Disadvantages
Powerful for noisy training data and works well with large training data [4] Simplicity, transparency, easy to understand and implement [3]	for each new instance, all distances from the K-nearest neighbors have to be calculated multiple times, which increases the calculation time. To reduce the error rate, you need to determine the value of K correctly. [4] Computational complexity memory limitations poor run-time performance for large training sets irrelevant attributes can cause problems [3].

III. CLASSIFICATION ALGORITHM’S APPLICATIONS

TABLE I. APPLICATIONS OF CLASSIFICATION ALGORITHMS

Classification Algorithms	Applications
Naïve Bayes Classifier	recommendation system [14] text classification [15] spam filtering [16]
Logistic Regression	medicine [17] social sciences [18]
Bayesian Networks	Traffic accident detection [19] Signature verification [20]
Decision Tree	Air traffic control systems [21] financial analysis [22]
ID3	Predict student performance [23] Fraud detection application [24] Tolerance-related knowledge acquisition [25]
C4.5	Debtor's decision to apply for a loan

	[26] Predicting software defects [27] Thrombosis Collagen disease [28] Question pool selection [29]
CART	intrusion detection [30]
Random Forest	Classify fraudulent and non-cheating.[31] Examine stock market trends.[32] Remote sensing.[33] Medicine [34]
Support Vector Machines	face detection, [35] image classification, [36] handwriting recognition, [37]
Artificial Neural Networks	Typed character recognition [38] Image processing [39] Aircraft navigation [40] Speech synthesis [41]
K-Nearest Neighbours	recommended systems [42] search applications [43]

IV. MACHINE LEARNING APPLICATIONS

Three of the most important applications that utilize machine learning are image processing, speech recognition, and text interpretation.

1. Image Recognition

The most important requirement for a machine when it comes to image processing is - similar to human vision and thinking - to be able to interpret the images made available to it and to recognize various objects on these. This process is also called labelling and this is one of the most widely applicable areas of artificial intelligence

2. Speech Recognition

Voice/speech recognition allows machines to interpret human speech and voices from other sources. This enables us to handle our smart devices more quickly and more comfortably, and to give our devices the ability to recognize sounds.

3. Text interpretation

The process of text interpretation basically means information retrieval from large amounts of text-based data. Natural language processing (NLP) is essential for this procedure, which enables the machine to understand and process everyday language manifestations. Without this, computers might be able to understand the meaning of each word, but NLP helps machines to interpret words in context

TABLE II. MACHINE LEARNING APPLICATIONS

Application Domain	Applications	Used Classification methods
Image Processing	Image processing and machine learning in the morphological analysis of blood cells [44]	Neural Networks, Decision Trees, and Support Vector Machines
	Horizon Detection Using Machine Learning Techniques [45]	SVM, J48, and naive Bayes classifiers
	Image Processing and Machine Learning for Automated Fruit Grading System [46]	Support vector machines, K-Nearest neighbour Artificial Neural Network, C4.5
	Image Processing-	Support Vector

	Based Detection of Pipe Corrosion Using Texture Analysis and Metaheuristic-Optimized Machine Learning Approach [47]	Machines
Speech Recognition	Automatic bird species recognition based on birds vocalization [48]	hidden Markov model, K-Nearest neighbour
Text Interpretation	Multi-class Document Classification using Support Vector Machine (SVM) Based on Improved Naïve Bayes Vectorization Technique [49]	SVM, Naïve Bayes
	Sentiment Analysis on Twitter Data [50]	SVM, Naïve Bayes
	Machine Learning Methods For Spam E-Mail Classification [51]	Bayesian classification, k-NN, ANNs, SVMs

V. CONCLUSION

In this study various popular classification techniques of machine learning have been discussed with their basic information, advantages, disadvantages & applications. The discussed classification techniques can be implemented on different type of data set i.e., health, financial, educational etc. It is difficult to find out which technique is superior to other because each technique has its own merits, demerits and applications. The selection of classification technique depends on user problem domain. So, we can conclude special features of different classification techniques from above study. Naive Bayes classifier has been widely used in many classification problems, especially when the dimensionality of the features is high. SVM has capability to deal with wide variety of classification problems includes high dimensional and not linearly separable problems. KNN is used in search applications where "similar" items are searched by the user. Neural networks are used when the exact nature of the relationship between inputs and output is not known. A key feature of neural networks is that they learn the relationship between inputs and output through training. Logistic regression is used extensively in numerous disciplines, including the medical and social science fields. Decision trees can deal with both categorical and numerical data, it may have an overfitting issue. Random forest enhances the accuracy of the model and prevents the overfitting issue. It is capable of handling large datasets with high dimensionality. The most interesting feature of Bayesian Networks, compared to decision trees or neural networks, is most certainly the possibility of taking into account prior information about a given problem, in terms of structural relationships among its features. Stochastic gradient descent is a very common algorithm used in various machine learning algorithms, and most importantly, it forms the basis of neural networks. Naïve Bayes Classifier, Support Vector Machine, and Logistic Regression algorithms are good to use for small datasets. Stochastic gradient descent and K-Nearest Neighbours algorithms are effectively working for large datasets. From

Table II, Support Vector Machine and decision tree classifiers are commonly used algorithms in the image processing domain. In-text interpretation Support Vector Machine and Naïve Bayes have commonly used classification algorithms. So according to Table I and Table II Support Vector Machine has wide range of applications but it is especially good for image processing and naïve bayes is good for text interpretation.

In this study, we have compared classification methods theoretically. We can apply them on different types of data sets having different types of input types like images, text and voice and properties and can attain a best result by knowing that which algorithm will give the best result on a specific type of data set. We can compare them practically to check the efficiency of each algorithm and decide which algorithm is best for which application.

REFERENCES

- [1] William L. Hosch Machine learning artificial intelligence
- [2] Ron Kohavi; Foster Provost (1998). "Glossary of terms". Machine Learning 30: 271–274
- [3] Aized Amin Soofi* and Arshad Awan "Classification Techniques in Machine Learning: Applications and Issues"
- [4] Pratap Chandra Sen, Mahimarnab Hajra and Mitadru Ghosh "Supervised Classification Algorithms in Machine Learning: A Survey and Review"
- [5] Data Classification Algorithms and Applications Edited by Charu C. Aggarwal IBM T. J. Watson Research Center Yorktown Heights, New York, USA
- [6] The Logistic Regression Algorithm by Niklas Donges <https://machinelearning-blog.com/2018/04/23/logistic-regression-101/>
- [7] Phyu TN. Survey of classification techniques in data mining, in Proceedings of the International MultiConference of Engineers and Computer Scientists 2009; pp. 18-20.
- [8] Use of ID3 Decision Tree Algorithm for Placement Prediction Hitarthi Bhatt#1, Shraddha Mehta#2, Lynette R. D'mello#3 #Dept. of Computer Engineering, D.J. Sanghvi College of Engineering, Mumbai University Mumbai, India
- [9] Quinlan JR. Induction of decision trees. Machine learning 1986; 1: 81-106. <https://doi.org/10.1007/BF00116251>
- [10] Machine Learning Algorithms Machine Learning Algorithm's for Classification Rob Schapire Princeton University
- [11] A Guide To Random Forests : Consolodating Decision Trees by Vihar Kurama
- [12] Top 4 advantages and disadvantages of Support Vector Machine or SVM by DhirajK
- [13] Artificial Neural Networks Advantages and Disadvantages Maad M. Mijwel Computer science, college of science, University of Baghdad Baghdad, Iraq maadalnaimiy@yahoo.com January 2018
- [14] Rahul S. Gaikwad Sandeep S. Udmale Vijay K. Sambhe E-commerce Recommendation System Using Improved Probabilistic Model © Springer Nature Singapore Pte Ltd. 2018, https://doi.org/10.1007/978-981-10-3920-1_28
- [15] Johnson Kolluri, Shaik Razia Text classification using Naïve Bayes classifier © 2020 Elsevier Ltd. All rights reserved https://doi.org/10.1016/j.matpr.2020.10.058_2
- [16] Rohit Kumar Solanki; Karun Verma; Ravinder Kumar Spam Filtering Using Hybrid Local-Global Naive Bayes Classifier 2015 International Conference on Advances in Computing, Communications and Informatics (ICACCI) 978-1-4799-8792-4/15/\$31.00 c 2015 IEEE
- [17] S. C. Bagley, H. White, and B. A. Golomb. Logistic regression in the medical literature: Standards for use and reporting, with particular attention to one medical domain. Journal of Clinical Epidemiology, 54(10):979–985, 2001.
- [18] J. Antonakis and O. Dalgas. Predicting elections: Child's play! Science, 323(5918):1183–1183, 2009.
- [19] J. Hwang, Y. Lee and S. Cho, "Structure evolution of dynamic Bayesian network for traffic accident detection," 2011 IEEE Congress of Evolutionary Computation (CEC), New Orleans, LA, USA, 2011, pp. 1655-1671, doi: 10.1109/CEC.2011.5949815.
- [20] Xuhong Xiao, Graham Leedham, Signature verification using a modified Bayesian network, Pattern Recognition, Volume 35, Issue 5, 2002, Pages 983-995, ISSN 0031-3203, [https://doi.org/10.1016/S0031-3203\(01\)00088-7](https://doi.org/10.1016/S0031-3203(01)00088-7)
- [21] Malakis, S., Psaros, P., Kontogiannis, T. et al. Classification of air traffic control scenarios using decision trees: insights from a field study in terminal approach radar environment. Cogn Tech Work 22, 159–179 (2020). <https://doi.org/10.1007/s10111-019-00562-7>
- [22] Adrian Gepp, Kuldeep Kumar, Predicting Financial Distress: A Comparison of Survival Analysis and Decision Tree Techniques, Procedia Computer Science, Volume 54, 2015, Pages 396-404, ISSN 1877-0509, <https://doi.org/10.1016/j.procs.2015.06.046>
- [23] Ramanathan L., Saksham Dhandra, Suresh Kumar D Predicting Students' Performance using Modified ID3 Algorithm Ramanathan L et.al / International Journal of Engineering and Technology (IJET) ISSN : 0975-4024 Vol 5 No 3 Jun-Jul 2013
- [24] K. Zou, W. Sun, H. Yu and F. Liu, "ID3 Decision Tree in Fraud Detection Application," 2012 International Conference on Computer Science and Electronics Engineering, 2012, pp. 399-402, doi: 10.1109/ICCSEE.2012.241.
- [25] Xinyu Shao, Guojun Zhang, Peigen Li, Yubao Chen, Application of ID3 algorithm in knowledge acquisition for tolerance design, Journal of Materials Processing Technology, Volume 117, Issues 1–2, 2001, Pages 66-74, ISSN 0924-0136, [https://doi.org/10.1016/S0924-0136\(01\)01016-0](https://doi.org/10.1016/S0924-0136(01)01016-0).
- [26] R. K. Amin, Indwiarti and Y. Sibaroni, "Implementation of decision tree using C4.5 algorithm in decision making of loan application by debtor (Case study: Bank pasar of Yogyakarta Special Region)," 2015 3rd International Conference on Information and Communication Technology (ICoICT), 2015, pp. 75-80, doi: 10.1109/ICoICT.2015.7231400.
- [27] L. D. Panjer, "Predicting Eclipse Bug Lifetimes," Fourth International Workshop on Mining Software Repositories (MSR'07:ICSE Workshops 2007), 2007, pp. 29-29, doi: 10.1109/MSR.2007.25.
- [28] S. A. Soliman, S. Abbas and A. M. Salem, "Classification of thrombosis collagen diseases based on C4.5 algorithm," 2015 IEEE Seventh International Conference on Intelligent Computing and Information Systems (ICICIS), 2015, pp. 131-136, doi: 10.1109/IntelCIS.2015.7397209.
- [29] A. M. Seet and I. A. Zualkernan, "An Adaptive Method for Selecting Question Pools Using C4.5," 2010 10th IEEE International Conference on Advanced Learning Technologies, 2010, pp. 86-88, doi: 10.1109/ICALT.2010.31.
- [30] Miao Li Application of CART decision tree combined with PCA algorithm in intrusion detection 2017 8th IEEE International Conference on Software Engineering and Service Science (ICSESS)
- [31] M. S. Kumar, V. Soundarya, S. Kavitha, E. S. Keerthika and E. Aswini, "Credit Card Fraud Detection Using Random Forest Algorithm," 2019 3rd International Conference on Computing and Communications Technologies (ICCCT), 2019, pp. 149-153, doi: 10.1109/ICCCT2.2019.8824930.
- [32] Khaidem, Luckyson, Snehanthu Saha, and Sudeepa Roy Dey. "Predicting the direction of stock market prices using random forest." *arXiv preprint arXiv:1605.00003* (2016).
- [33] Mariana Belgiu, Lucian Drăguț, Random forest in remote sensing: A review of applications and future directions, ISPRS Journal of Photogrammetry and Remote Sensing, Volume 114, 2016, Pages 24-31, ISSN 0924-2716, <https://doi.org/10.1016/j.isprsjprs.2016.01.011>.
- [34] Guy Cafri, Luo Li, Elizabeth W. Paxton & Juanjuan Fan (2018) Predicting risk for adverse health events using random forest, Journal of Applied Statistics, 45:12, 2279-2294, DOI: [10.1080/02664763.2017.1414166](https://doi.org/10.1080/02664763.2017.1414166)
- [35] Hyungkeun Jee, Kyunghee Lee and Sungbum Pan, "Eye and face detection using SVM," Proceedings of the 2004 Intelligent Sensors, Sensor Networks and Information Processing Conference, 2004., 2004, pp. 577-580, doi: 10.1109/ISSNIP.2004.1417525.
- [36] S. Agrawal, N. K. Verma, P. Tamrakar and P. Sircar, "Content Based Color Image Classification using SVM," 2011 Eighth International Conference on Information Technology: New Generations, 2011, pp. 1090-1094, doi: 10.1109/ITNG.2011.202.
- [37] Hamid, Norhidayu Abdul, and Nilam Nur Amir Sjarif. "Handwritten recognition using SVM, KNN and neural network." arXiv preprint arXiv:1702.00723 (2017).
- [38] Ahmad M. Sarhan, and Omar I. Al Helala Arabic Character Recognition using Artificial Neural Networks and Statistical Analysis World Academy of Science, Engineering and Technology 27 2007
- [39] Pandelea, Alexandrina-Elena, Mihai Budescu, and Gabriela Covatariu. "Image processing using artificial neural networks." Buletinul Institutului Politehnic din Iasi. Sectia Constructii, Arhitectura 61.4 (2015): 9.
- [40] H. Baomar and P. J. Bentley, "Autonomous navigation and landing of large jets using Artificial Neural Networks and learning by imitation," 2017 IEEE Symposium Series on Computational Intelligence (SSCI), 2017, pp. 1-10, doi: 10.1109/SSCI.2017.8280916.
- [41] E. V. Raghavendra, P. Vijayaditya and K. Prahallad, "Speech synthesis using artificial neural networks," 2010 National Conference On Communications (NCC), 2010, pp. 1-5, doi: 10.1109/NCC.2010.5430190.
- [42] G. Li and J. Zhang, "Music personalized recommendation system based on improved KNN algorithm," 2018 IEEE 3rd Advanced Information Technology, Electronic and Automation Control Conference (IAEAC), 2018, pp. 777-781, doi: 10.1109/IAEAC.2018.8577483.

- [43] Hassanat ABA (2018) Two-point-based binary search trees for accelerating big data classification using KNN. PLoS ONE 13(11): e0207772. <https://doi.org/10.1371/journal.pone.0207772>
- [44] J. Rodellar S. Alférez A. Acevedo A. Molina A. Merino Image processing and machine learning in the morphological analysis of blood cells <https://doi.org/10.1111/ijlh.12818>
- [45] Sergiy Fefilatyeu; Volha Smarodzinava; Lawrence O. Hall; Dmitry B. Goldgof Horizon Detection Using Machine Learning Techniques 2006 5th International Conference on Machine Learning and Applications (ICMLA'06)
- [46] Rashmi Pandey Sapan Naik Roma Marfatia Image Processing and Machine Learning for Automated Fruit Grading System: A Technical Review International Journal of Computer Applications (0975 – 8887) Volume 81 – No 16, November 2013
- [47] ND Hoang, VD Tran Image Processing-Based Detection of Pipe Corrosion Using Texture Analysis and Metaheuristic-Optimized Machine Learning Approach Volume 2019 |Article ID 8097213 | <https://doi.org/10.1155/2019/8097213>
- [48] Jiri Stastny, Michal Munk & Lubos Juranek Automatic bird species recognition based on birds vocalization EURASIP Journal on Audio, Speech, and Music Processing volume 2018, Article number: 19 (2018)
- [49] Hajah T. Sueno¹ , Bobby D. Gerardo² , Ruji P. Medina³ Multi-class Document Classification using Support Vector Machine (SVM) Based on Improved Naïve Bayes Vectorization Technique Volume 9, No.3, May - June 2020 International Journal of Advanced Trends in Computer Science and Engineering Available Online at <http://www.warse.org/IJATCSE/static/pdf/file/ijatcse216932020.pdf> <https://doi.org/10.30534/ijatcse/2020/21693202>
- [50] Varsha Sahayak Vijaya Shete Apashabi Pathan Sentiment Analysis on Twitter Data International Journal of Innovative Research in Advanced Engineering (IJIRAE) ISSN: 2349-2163 Issue 1, Volume 2 (January 2015)
- [51] W. Awad, S. Elseuofi MACHINE LEARNING METHODS FOR SPAM E-MAIL CLASSIFICATION International Journal of Computer Science and Information Technology DOI:10.5121/IJCSIT.2011.31