FUDMA Journal of Sciences (FJS)
ISSN online: 2616-1370
ISSN print: 2645 - 2944
Vol. 10 No. 9, May, 2026, pp 103 – 109
DOI: https://doi.org/10.33003/fjs-2026-1009-5315

Review of Classification Methods For Machine Learning Modeling

***[1]Agu C. Sunday and [2]Augustina M. Ukeje-Okorie**

[1]Department of Computer Science, Benson Idahosa University, Benin, Edo State, Nigeria.
[2]Department of Computer Science, Federal College of Agriculture, Ishiagu, Ebonyi State, Nigeria.

*Corresponding authors' email: sagu@biu.edu.ng
ORCID: https://orcid.org/0000-0003-0635-1939

Received: 16 May 2026
Accepted: 01 June 2026
Published: 16 June 2026

ABSTRACT
Logistic Regression (LR), Support Vector Machine (SVM), Decision Tree (DT) methods and their corresponding ensembles, strengths and weaknesses were surveyed, considering machine learning accuracy matrices, probability estimation as well as other features, and determining the frequency of their applications in different sectors. Descriptive statistics was used to analyze the survey which revealed that SVM yielded better predictive performance in different application areas. The study pointed out that no algorithm out rightly outperforms the other as the performance depends on the dataset, and recommended further work on the ensembles of DT and LR to leverage their fundamental advantages. Following the report on the strength and weaknesses, a simple workflow was suggested. Finally, the application areas of the algorithms were reviewed which pointed out that social media is not primarily an application area but a web application for generating datasets for predictive analytics.

Keywords: Decision tree, Ensembles, Logistic regression, Machine Learning, Predictive Analysis, Support Vector Machine

INTRODUCTION
The increasing spreading of machine learning models has become a key area of concern for all types of organizations as they are assisting to influence all activities from cybersecurity, defense, marketing, customers care services, to recruitment, chatbots, and more (Ellen, 2019). Data mining, a method of discovering patterns and anomalies from massive datasets could be applied to identify multiple groups and clusters of data and forward them for further analysis through machine learning for predictive analytics (Lakshay & Prakita, 2017). The benefits which have aroused the interest of many businesses in deciding how to implement ML gave rise to a wide range of techniques for ML algorithms such as Linear Regression, Logistic Regression, DT, SVM, Naive Bayes, K-Nearest Neighbors (KNN), Random Forest, Dimensionality Reduction algorithms, Gradient Boosting with its various ensembles GBM, XGBoost, LightGBM, and CatBoost (Sunil, 2017). DT, LR, and SVM are the most basic and widely used supervised algorithms in machine learning big data analytics. The algorithms build from externally supplied inputs to develop specific new model can be applied for predicting future outcome around a specific life problem (Lakshay & Prakita, 2017). Hence, the objective of this research is to examine DT, LR and SVM algorithms and their ensembles, review their strengths and weaknesses and application areas.
There are controversies over one predictive algorithm been better than another. All algorithms are theoretically equivalent when their performance is averaged across all possible variables. In reality, solving for all possible variables is not applicable (Claudia, 2017). Machine learning task was introduced as the most practical way to group algorithms instead of their learning mechanisms. As such, no one algorithm works best for every problem as it all depends on the problem and factors under consideration (Elite Data Science, 2019). Nevertheless, some algorithms provide a comparative advantage over others when considering interpretability, robustness, predictive performance to accuracy, ranking, and probability estimation, and so on. One basic difference between linear regression and other ML models is that linear regression can only support linear solutions. Moreover, no model outperforms others since efficiency is based on the type of training data distribution which invariably determines the appropriate model (Danny, 2018). However, some learning algorithms offer relative benefits over others since different algorithms make different assumptions about data and have different rates of convergence (Jack, 2019). The algorithm which minimizes some cost function of interest, such as cross-validation, makes assumptions that are consistent with the data and has sufficiently converged to its error rate.
The algorithm to choose for a particular problem depends on a simplified 2-D data which has to be extrapolated for an understanding of a higher dimensional data (Lalit, 2015). Hence, the distinction between logistic regression, decision tree, and support vector machine (SVM) is anchored on the shape of the produced class decision boundary graph. Logistic regression is most appropriate where the class decision boundary is linearly separated.

Examining LR, DT and SVM Algorithms and their Ensembles
According to Claudia (2017), there is no substantial evidence in formal scientific terms about the relative expected performance of one algorithm over the other, and their ensembles as most justifications in this regard are 'theoretically' centred on bias-variance trade-off. Thus, this study reviewed some specific works in different life issues where each of the DT, LR, and SVM algorithms and their corresponding ensembles have comparatively yielded better results and have proved more suitable than others and also to their respective traditional statistical approaches.

i. Landslide susceptibility assessment: considering the geological, morphological, and environmental attributes, a specific area was selected to compare the DT, LR, and SVM algorithms. SVM outperformed others, including the common Analytical Hierarchy Process (AHP) approach in all evaluation metrics (κ index, area under ROC curve and false-positive rate in stable ground class), creating a landslide susceptibility map of the relevant area (Milos et al, 2011).
ii. Enterprise credit risk assessment: many statistical and intelligent methods have been used to assess enterprise credit risk without consistent conclusions on which methods are better. A hybrid ensemble approach based on Support Vector Machine demonstrated the effectiveness and feasibility of RSB-SVM which revealed an alternative method for enterprise credit risk assessment (Gang & Ma, 2012).
iii. Customer reviews for customers' purchase decisions: SVM is revealed as a frontier over social networks, forums and blogs for categorizing customer's review into polarity levels such as positive, negative, and neutral which has emerged as a critical factor for the customers' purchase decisions (Cagatay & Nangir, 2017).
iv. Predicting Breast Cancer Recurrence: DT (C4.5), SVM, and Artificial Neural Network ANN) were used to discover hidden patterns and relationships for breast cancer recurrence in patients who were followed-up for two years. The result indicated that the SVM classification model predicts breast cancer recurrence with the least error rate and highest accuracy, achieved using 10-fold cross-validation for measuring the unbiased prediction accuracy of each model (Ahmad et el, 2013).
v. Prediction of lung cancer patient survival: SVM and Logistic regression were applied to Surveillance, Epidemiology, and End Results (SEER) program dataset which indicated result value from low to moderate survival times, with SVM generating a distinctive output and informing patient care decisions (Lynch, et al., 2017).
vi. Customer churn prediction: DT and LR are considered the most popular algorithms for customer churn prediction. The challenges that they both face have necessitated a new hybrid ensemble, the logit leaf model (LLM) (Arno, Kristof, & Koen, 2018), (Nadeem, Umar, & Shahzad, 2018). The area under the receiver operating characteristics curve (AUC) and top decile lift (TDL) were used to measure the predictive performance for which LLM scores significantly better.
vii. Comparison of algorithms on diabetes dataset: Seven machine learning algorithms were used for classification and comparison on diabetes dataset out of which SVM had the most precision and accuracy considering the time, kappa statistic and Mean Absolute Error (MAE) (Osisanwo, et al., 2017).
viii. Semi-Supervised Decision Tree (SSDT): A typical decision tree splits nodes, based only on the target variable which causes drawback. As a result, a hybridsemi-supervised decision tree ensemble (SSDT) was modeled to splits internal nodes by utilizing both labels and the structural characteristics of six different datasets from real-world problems for subspace partitioning. The result proved that subspace partitioning using SSDT provided better training sets to build classifiers (Kyoungok, 2016).
ix. Predicting performance of postgraduate students: several learning algorithms were applied to predict the performance of MSc students in a class. Faced with a small size of the datasets and the imbalance in the distribution of class values problems, some techniques such as resampling and feature selection were applied for the first time in a performance prediction application. In the end, Naïve Bayes and 1-NN achieved the best prediction results (Koutina & Kermanidis, 2011).
x. Flood susceptibility analysis and its verification: to overcome the problem inherent in using statistical and machine learning methods individually, a novel ensemble method integrated support vector machine (SVM) and frequency ratio (FR) to produce spatial modeling in flood susceptibility assessment. DT was also integrated with FR for comparative analysis. The results proved the efficiency of the SVM+FR as rapid, accurate, and reasonable in flood susceptibility assessment (Tehrany, Pradhan, & Jebur, 2015).
xi. Heart disease prediction: a hybrid rule-based model combined with two or more of the DT, LR, and SVM algorithms are proposed to be more efficient against using a single algorithm. Results indicated that the rule-based model increased the accuracy of predicting the early onset of cardiovascular diseases based on accuracy, sensitivity, and specificity (Mythili, Dev, Nikita, & Abhiram, 2013).

Table 1: Usage of DT, LR, and SVM Algorithms

Description	**Algorithm**
Landslide susceptibility assessment	SVM
Enterprise credit risk assessment	RSB-SVM
Customer reviews for customers' purchase decisions	SVM
Predicting breast cancer recurrence	SVM
Prediction of lung cancer patient survival	SVM
Customer churn prediction	LLM
Prediction of diabetes dataset	SVM
Predicting credit information of clients	SSDT
Predicting image signs of diabetic retinopathy	SSDT
Predicting age of abalone from physical measurement	SSDT
Predicting bank direct telemarketing campaign	SSDT
Predicting molecular musk	SSDT
Predicting forest cover type from cartographic variables	SSDT
Predicting performance of postgraduate students	1-NN
Flood susceptibility analysis and its verification	SVM+FR

Method

Table 1.0 showed 15 application areas of DT, LR, and SVM algorithms and their ensembles that were randomly reviewed. Each of the algorithms and its corresponding improvement was merged as one algorithm. As a result, three classes of the algorithms (DT, LR, and SVM) were used for the analysis. To avoid bias, the 6 SSDT algorithms which were the ensemble algorithms to DT were the results gotten from one research experiment for different application areas, and hence it was counted as 'one' for DT. Again, the LLM algorithm was an ensemble over both the DT and LR algorithms, as a result, LLM is counted 'once' each for both DT and LR algorithms respectively. Figure 1.0 depicted the Frequency Distribution for the usage of DT, LR, and SVM.

Remark 1

The Frequency Distribution for Usage of DT, LR, and SVM showed in Figure 1.0 revealed that SVM yielded better predictive performance in different application areas. It also revealed that the SSDT ensemble to DT outperformed SVM.

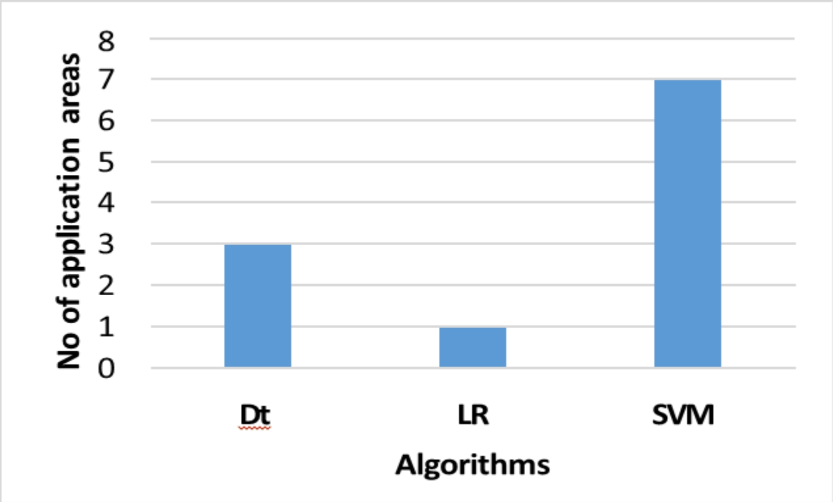

Figure 1: Frequency Distribution for Usage of DT, LR, and SVM

Strengths and Weaknesses

Recent researches in the field of Machine Learning Predictive Analytics (Scikit-learn, 2019), (Claudia, 2017), (Elite Data Science, 2019), (Brownlee, 2019), (Lars, 2019), (Dhiraj, 2019), (queirozf, 2019), (Lalit, 2015), (Danny, 2018), have shown the puissance and weaknesses of DT, LR, and SVM resulting from different factors.

Decision Tree

Decision trees construct classification or regression models like a tree structure. It splits a dataset into smaller subsets and incrementally develops an associated decision tree. The resultant effect is a tree with decision nodes and leaf nodes.

Strengths of Decision Tree

i. Simple to understand and to interpret as trees can be visualized.
ii. Great accuracy when predictive performance is based on classification problems.
iii. DT can learn non-linear relationships by deriving hyper-rectangles in input space to solve the problem of the non-linear relationships.
iv. The decision tree handles collinearity better than LR.
v. Ensembles are scalable and perform very well in practice (Elite Data Science, 2019).
vi. Less effort for data preparation during pre-processing. Hence, it does not require normalization of any scaling of data, dummy variables do need to be created, and blank values do not have to be removed.
vii. DT is preferable if the signal to noise ratio is high, which is obtained when the dataset of the problem is inherently predictable, such as customers' responses to direct telemarketing campaign (Claudia, 2017).
viii. Capable of handling both numerical and categorical data.
ix. Performs better on a large dataset (Claudia, 2017).
x. Able to handle multi-output problems and are robust to outliers.
xi. Possible to validate a model using statistical tests. That makes it possible to account for the reliability of the model.
xii. Performs well even if its assumptions are partly violated by the true model from which the data were generated.

Weaknesses of Decision Tree

i. Less accuracy when predictive performance is based on ranking and probability estimation problems.
ii. Overfitting, the chances of creating over-complex trees that do not generalize the dataset. This can be handled by using(1) pruning (2) setting the minimum number of samples required at a leaf node or (3) setting the maximum depth of the tree.
iii. Instability where small variations in the data might result in a completely different structure of the decision tree. This can be alleviated by using decision trees within an ensemble.
iv. Trees generally have a harder time coming up with calibrated probabilities as they mostly return biased probabilities that would have given some confidence in the prediction.
v. DT provides NP-complete which is the problem of learning an optimal decision tree. As a result, heuristic algorithms such as the greedy algorithm where locally optimal decisions are made at each node are applied to DT learning algorithms. With a greedy algorithm, returning a globally optimal decision tree is not assured. This can be leveraged by using ensemble learner, where the features and samples are randomly sampled with replacement.

vi. Some concepts are hard to learn because decision trees do not express them easily, such as XOR, parity, or multiplexer problems.
vii. Decision tree learners create biased trees if some classes dominate. It is therefore recommended to balance the dataset before fitting with the decision tree.
viii. Information gain in a decision tree with categorical variables gives a biased response for attributes with a greater number of categories.

Logistic Regression

Logistic Regression is a classification algorithm, when given a set of independent variables, is used to predict a binary outcome. It is an enhancement to linear regression when the outcome variable is categorical, where a log of odds is used as the dependent variable to predict the probability of occurrence of an event by fitting data to a logit function.

Strengths of Logistic Regression

i. Outputs have a nice probabilistic interpretation.
ii. It can be regularized to avoid overfitting.
iii. It can be easily updated with new data using stochastic gradient descent (Elite Data Science, 2019).
iv. Good if the signal to noise ratio is low, which is obtained when the dataset of the problem is inherently unpredictable such as the stock market.
v. They make assumptions that simplify the learning process with a set of parameters of fixed size.
vi. Simpler, the method is easier to understand and interpret results.
vii. Speed, as a parametric model it is very fast to learn from data.
viii. Less data, it does not require as much training data and can satisfactorily execute even if the fit to the data is not perfect.
ix. The multi-collinearity issue can be handled with L2 regularization to an extent
x. There is a widespread industry comfort for logistic regression solutions

Weaknesses of Logistic Regression

i. Under performs when there are non-linear decision boundaries. (Elite Data Science, 2019).
ii. Not suitable when the signal to noise ratio is high, which is obtained when the dataset of the problem is inherently predictable such as customers' responses to a direct telemarketing campaign (Claudia, 2017).
iii. Could only handle categorical data through conversion using integer encoding one-hot encoding (Brownlee, 2019).
iv. Drifts in handling multi-output problems.
v. LR is constrained. By choosing a functional form, it is highly constrained to the specified form.
vi. Limited Complexity, as it is more suitable for simpler problems.
vii. Poor Fit, in practice it is unlikely to match the underlying mapping function.
viii. Collinearity and outliers interfere with the accuracy of the LR model.
ix. A proper selection of features is required.

SVM

A support vector machine (SVM) uses a classification algorithm for two-group classification problems. After receiving a set of labeled training data for either of two categories, new examples are categorized (Smriti, 2019), (Stecanella, 2017).

Strengths of SVM

i. Effective in high dimensional spaces.
ii. Still effective in cases where several dimensions are greater than the number of samples.
iii. Uses a subset of training points (called support vectors) as it does not rely on entire data in the decision function and hence, it is also memory efficient.
iv. SVM handles outliers better as it derives maximum margin solution. It handles outliers using soft margin constant (C) hyperparameter that decides the level of penalty over the outliers.
v. SVM uses kernel trick to solve complex solutions.
vi. Versatile: different Kernel functions can be specified for the decision function. Common kernels are provided, but it is also possible to specify custom kernels.
vii. SVM's can model non-linear decision boundaries, and there are many kernels to choose from.
viii. They are also fairly robust against overfitting, especially in high-dimensional space.
ix. SVM uses a convex optimization function, due to which global minima are always achievable.
x. Hinge loss provides higher accuracy. Hinge loss in SVM outperforms log loss in LR, that is, the loss metric of the deviation between expected and predicted output.

Weaknesses of SVM

i. SVMs do not directly provide probability estimates, these are calculated using an expensive five-fold cross-validation.
ii. SVM's are memory intensive
iii. Trickier to tune due to the importance of picking the right kernel
iv. Do not scale well on larger datasets.
v. Longer training time for larger datasets
vi. Hinge loss leads to sparsity
vii. Hyperparameters and kernels are to be carefully tuned for sufficient accuracy.

Simple Workflow

Having reviewed the strengths and challenges, a simple workflow to decide which algorithm to use is suggested thus:

i. Always start with logistic regression, then to use the performance as a baseline.
ii. Determine whether the decision tree provides a significant improvement. Even if the resultant model is not utilized, the result of the DT could remove noisy variables.
iii. Then apply SVM if there are a large number of features and the number of observations is not a limitation for available resources and time.

Remark 2

The workflow implies that SVM should be given the highest priority amongst the three algorithms for performing predictive analysis to achieve the most reliable results.

Application Areas

Decision tree, logistic regression and support vector machine algorithms are today being applied in almost different sectors but not limited to the ones outlined below:

Healthcare Management

(1) Application of evidence-based medicine for making clinical decisions (Jong-Myon, 2014). (2) Predicting the

possibilities of a disease in an area for a speedy response (Banumathi & Aloysius, 2017). (3) Discovering hidden patterns and relationships for predicting breast cancer recurrence in patients (Lakshay & Prakita, 2017). (4) Predicting lung cancer patient survival (Rajesh, 2018). (5) Heart disease prediction (Ahmad, Eshlaghy, Poorebrahimi, Ebrahimi, & Razavi, 2013). (6) Prioritization of emergency room patient treatment (Lynch, et al., 2017), (Mythili, Dev, Nikita, & Abhiram, 2013).

Education

(1) Predicting slow learners among students. (2) Predicting the vulnerability of students to deviant ideologies that might lead to terrorism. (3) Increasing retention capacity of students though not yet specific nor tested by real-world data (Parneet, Manpreet, & Gurpreet, 2015). (4) Predicting the final performance of students in a specific course in a semester and (5) Predicting dropping out of students (Fatimetou, 2017).

Business

(1) Customer segmentation for predicting likely buyers of a product during marketing campaign to enable targeting of limited advertisement budget (Roberto, 2020). (2) Risk assessment for predicting profitable business operations such as customer's eligibility to loan facilities (Lakshay & Prakita, 2017). (3) Customer churn for predicting customers' relationship with the organization before leaving for another organization to prevent such occurrence (Rajesh, 2018). (4) Predicting request to cover future purchasing (Ahmad, Jafar, & Aljoumaa, 2019). (5) Predicting purchases for better sales decisions. (6) Predicting stock market indications. (7) Enterprise Resource Planning (ERP) systems prediction for decision making in sales, levels of production, and distribution plans (Arno, Kristof, & Koen, 2018). (8) Forecasting event logs and processes to predict the workflow time of an organization to meet the Service Level Agreements (SLAs) signed with the customers (Nadeem, Umar, & Shahzad, 2018). (9) Already deployed by LinkedIn and Yahoo, real-time data warehousing predictive modelling to synchronically order all the campaigns in the organization and simplify the comparison between different offers and operations across the various product to create centralized campaign analytics for delivering faster decision making, (Fatimetou, 2017).

Social Media

The data that is used for predictive analytics originates from either the existing dataset or social media. Hence, social media is not primarily an application area but a web application that enables users to create and share content or to participate in social networking. Through users' participation on social media, a huge amount of data is generated which can be used for predictions. Recent research has shown how data emanating from the widespread adoption and use of social media channels such as Facebook, Twitter, Google Trends, Google News, Forum, Youtube, Blogs, Tumblr, IMDB, Flixster, Yahoo Movies, HSX, and RottenTomatoes.com, can be used to predict outcomes such as Hollywood movie revenues, Apple iPhone sales, seasonal moods, Consumer spending, Political election outcome, socio-economic events, epidemic outbreaks, and behavioural marketing where consumers are targeted based on the websites they search for a commodity (Niels, Lisbeth, & Ravi, 2016).

Governance

Predicting the frequency of future attacks to minimize the loss of human life from the drone. Predict the chances of success and failure of a project during and after its development (Fatimetou, 2017). The police department could predict the crime scenes and hotspots through the analysis of data patterns. Government can predict fraud detection, weather, child abuse and child protection (Lakshay & Prakita, 2017). Humanitarian emergencies related to natural disasters such as earthquake, migrant and political crises that strike some part of the world (Anwaar, Junaid, Raihan, Arjuna, & Jon, 2016),

Heavy Industry

Remotely sensed satellite image classification (Joshi, Richa, & Aniruddha, 2013). Predicting power consumption and predicting the times of vehicle arrival in the short and long term (Fatimetou, 2017). Project Risk Management for cost and benefit analysis of a project (Lakshay & Prakita, 2017).

CONCLUSION

From the study, the appropriate ML algorithm to be utilized is a function of the type of the dataset and the application area, as no algorithm out rightly outperforms the others considering some trade-offs. However, a simple workflow recommended starting with LR, DT, or its ensemble and then SVM. Finally, SVM been an enhanced algorithm to the weaknesses inherent in both DT and LR, more ensembles of DT and LR are needed if grants would be awarded to bear the running cost, to still leverage their fundamental advantages, are recommended for future work.

RECOMMENDATION

It is therefore recommended that future research should be done to determine the reasons why SVM tends to yield better predictive performance and then device better DT ensembles that will further enhance predictive performance for the benefits of their simplicity of interpretation and visualization.

REFERENCES

Ahmad, A., Jafar, A., & Aljoumaa, K. (2019, March 20). Customer churn prediction in telecom using machine learning in the big data platform. Journal of Big Data volume, 6(26). Retrieved February 21, 2020, from https://link.springer.com/article/10.1186/s40537-019-0191-6#auth-1

Ahmad, L. G., Eshlaghy, A. T., Poorebrahimi, A., Ebrahimi, M., & Razavi, A. R. (2013). Using Three Machine Learning Techniques for Predicting Breast Cancer. Health & Medical Informatics, 4(2). Retrieved February 4, 2020, from https://s3.amazonaws.com/academia.edu.documents

Anwaar, A., Junaid, Q., Raihan, u. R., Arjuna, S. A., & Jon, C. (2016). Big data for development: applications. Big Data Analytics (Cross Mark), 1(2). Retrieved 20 February 2020, from https://bdataanalytics.biomedcentral.com/track/pdf/10.1186/s41044-016-0002-4

Arno, D. C., Kristof, C., & Koen, D. B. (2018, September 1). A new hybrid classification algorithm for customer churn prediction. European Journal of Operational Research, 269(2), 760-772. Retrieved February 4, 2020, from https://www.sciencedirect.com/science/article/abs/pii/S0377221718301243

Banumathi, S., & Aloysius, A. (2017). Predictive Analytics Concepts In Big Data- A Survey. International Journal of Advanced Research in Computer Science, 8(8), 27-30.

Retrieved February 25, 2020, from http://ijarcs.info/index.php/Ijarcs/article/view/4628

Brownlee, J. (2019, May 20). Why One-Hot Encode Data in Machine Learning? Retrieved from Machine Learning Mastery: https://machinelearningmastery.com/why-one-hot-encode-data-in-machine-learning/

Cagatay, c., & Nangir, M. (2017, January). A sentiment classification model based on multiple classifiers. Applied Soft Computing, 50, 135-141. Retrieved February 4, 2020, from https://www.sciencedirect.com/science/article/abs/pii/S1568494616305919

Claudia, P. (2017, June 19). What Are The Advantages Of Logistic Regression Over Decision Trees? Quora. Retrieved January 10, 2020, from https://www.forbes.com/sites/quora/2017/06/19/what-are-the-advantages-of-logistic-regression-over-decision-trees/#198b0f922c35

Danny, V. (2018, December 6). Comparative Study on Classic Machine learning Algorithms. Retrieved from Towards Data Science: https://towardsdatascience.com/comparative-study-on-classic-machine-learning-algorithms-24f9ff6ab222

Dhiraj, K. (2019, May 26). Top 5 advantages and disadvantages of the Decision Tree Algorithm.Retrieved from Medium: https://medium.com/@dhiraj8899/top-5-advantages-and-disadvantages-of-decision-tree-algorithm-428ebd199d9a

Elite Data Science. (2019, February 10). Modern Machine Learning Algorithms: Strengths and Weaknesses. Retrieved from https://elitedatascience.com/machine-learning-algorithms

Ellen, T. (2019, February 20). What are the benefits of white-box models in machine learning? Retrieved from the Silicon Republic: https://www.siliconrepublic.com/enterprise/white-box-machine-learning

Fatimetou, Z. M. (2017). THE APPLICATION OF PREDICTIVE. International Journal of Scientific and Research Publication, 7(5), 549-566. Retrieved from http://www.ijsrp.org/research-paper-0517/ijsrp-p656 4.pdf

Gang, W., & Ma, J. (2012, April 5). A hybrid ensemble approach for enterprise credit risk assessment based. Expert Systems with Applications, 39(5), 5325-5331. Retrieved from https://www.sciencedirect.com/science/article/pii/S0957417411015338

Jack, R. (2019). What are the advantages of logistic regression over decision trees? Are there any cases where it's better to use logistic regression instead of decision trees? Quora. Retrieved January 15, 2020, from https://www.quora.com/What-are-the-es-of-logistic-regression-over-decision-trees-Are-there-any-cases-where-its-better-to-use-logistic-regression-instead-of-decision-trees

Jong-Myon, B. (2014). The clinical decision analysis using a decision tree. Epidemiology and Health, 36, 1-7. Retrieved February 20, 2020, from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4251295/

Joshi, K. P., Richa, S., & Aniruddha, G. (2013). Decision tree approach for classification of remotely. J. Earth Syst. Sci. (Indian Academy of Sciences), 122(5), 1237–1247. Retrieved February 20, 2020, from https://www.ias.ac.in/article/fulltext/jess/122/05/1237-1247

Koutina, M., & Kermanidis, K. L. (2011). Predicting Postgraduate Students' Performance Using Machine Learning Technique. IFIP Advances in Information and Communication Technology, 364, 159-168. Retrieved February 5, 2020, from https://link.springer.com/chapter/10.1007/978-3-642-23960-1_20#citeas

Kyoungok, K. (2016). A hybrid classification algorithm by subspace partitioning through a semi-supervised decision tree. Pattern Recognition, 60, 157-163. Retrieved February 4, 2020, from https://www.sciencedirect.com/science/article/abs/pii/S0031320316300620

Lakshay, S., & Prakita, T. (2017). Predictive Modelling Analytics through Data Mining. International Research Journal of Engineering and Technology (IRJET), 4(9), 5-11. Retrieved January 10, 2020, from https://pdfs.semanticscholar.org/adf7/cf1e5b7702cbf1dff223862171ee59cb65d8.pdf

Lalit, S. (2015, October 5). Logistic Regression Vs Decision Trees Vs SVM: Part I. Edvancer. Retrieved January 13, 2020, from https://www.edvancer.in/logistic-regression-vs-decision-trees-vs-svm-part1/

Lalit, S. (2015, December 6). Logistic Regression vs Decision Trees vs SVM: Part II. Retrieved from Edvancer Eduventures: https://www.edvancer.in/logistic-regression-vs-decision-trees-vs-svm-part2/

Lars, H. (2019, March 14). Black-box vs. white-box models. Retrieved from Towards Data Science: https://towardsdatascience.com/machine-learning-interpretability-techniques-662c723454f3

Lynch, C. M., Abdollahi, B., Fuqua, J. D., de Carlo, A. R., Bartholomai, J. A., Balgemann, R. N., & Frieboes, H. B. (2017). Prediction of lung cancer patient survival via supervised machine learning. International Journal of Medical Informatics, 1-8. Retrieved February 4, 2020, from https://www.sciencedirect.com/science/article/abs/pii/S1386505617302368

Milos, M., Milos, K., Branislav, B., & Vít, V. (2011). Landslide susceptibility assessment using the SVM machine learning algorithm. Engineering Geology, 225-234. Retrieved February 4, 2020, from https://www.sciencedirect.com/science/article/abs/pii/S0013795211002195

Mythili, T., Dev, M., Nikita, P., & Abhiram, N. (2013). A Heart Disease Prediction Model using SVM-Decision Trees-Logistic Regression (SDL). International Journal of Computer Applications in Technology, 68(16), 11-15. Retrieved from https://www.researchgate.net/publication/273261237_A_Heart_Disease_Prediction_Model_using_SVM-Decision_Trees-Logistic_Regression_SDL

Nadeem, A. N., Umar, S., & Shahzad, S. M. (2018). A Review on Customer Churn Prediction Data Mining Modeling Techniques. Indian Journal of Science & Technology, 11(27). Retrieved February 21, 2020, from http://www.indjst.org/index.php/indjst/article/view/121478

Niels, B. L., Lisbeth, L. C., & Ravi, V. (2016). Predictive Analytics with Social Media Data. In A. Quan-Haase (Ed.), The SAGE Handbook of Social Media Research Methods (pp. 328-341). Retrieved 2 15,2020, from https://www.inet.ox.ac.uk/files/SLOAN_QUAN-HAASE-Chp20_2pp.pdf

Osisanwo, F., Akinsola, J., Awodele, O., Hinmikaiye, J. O., Olakanmi, O., & J., A. (2017, June). Supervised Machine Learning Algorithms: Classification and Comparison. International Journal of Computer Trends and Technology (IJCTT), 48(3), 128-138.Retrieved Feb 4, 2020, from https://www.researchgate.net/publication/318338750_Supervised_Machine_Learning_Algorithms_Classification_and_Comparison

Parneet, K., Manpreet, S., & Gurpreet, S. J. (2015). Classification and prediction based data mining algorithms to. Procedia Computer Science.57, pp. 500-508. Elsevier (Science Direct). Retrieved February 21, 2020, from https://www.sciencedirect.com/science/article/pii/S1877050915019018

Queirozf. (2019, January 24). Evaluation Metrics for Ranking problems: Introduction and Examples. Retrieved from queirozf.com: http://queirozf.com/entries/evaluation-metrics-for-ranking-problems-introduction-and-examples

Rajesh, B. S. (2018, October 26). Introduction to Decision Trees: Retrieved from Greyatom: https://medium.com/greyatom/decision-trees-a-simple-way-to-visualize-a-decision-dc506a403aeb

Roberto, L. A. (2020). Six Applications of Predictive Analytics in Business Intelligence. Retrieved February 15, 2020, from Neural Designer: https://www.neuraldesigner.com/blog/6_Applications_of_predictive_analytics_in_business_intelligence

Scikit Learn. (2019). Probability calibration. Retrievedfrom Scikit Learn: https://scikit-learn.org/stable/modules/calibration.html

Scikit-learn. (2019, February 11). Decision Trees. Retrieved from https://scikit-learn.org/stable/modules/tree.html

Smriti, D. (2019, May 28). Top 5 Predictive Models and Their Applications. Retrieved from Grazitti Interactive: https://www.grazitti.com/blog/top-5-predictive-models-and-their-applications/

Stecanella, B. (2017, July 22). An introduction to Support Vector Machines (SVM). Retrieved February 29, 2020, from Monkey Learn: https://monkeylearn.com/blog/introduction-to-support-vector-machines

Sunil, R. (2017, September 9). Commonly used Machine Learning Algorithms (with Python and R Codes). Retrieved from Analytics Vidhya: https://www.analyticsvidhya.com/blog/2017/09/common-machine-learning-algorithms/

Tehrany, M. S., Pradhan, B., & Jebur, M. N. (2015). Flood susceptibility analysis and its verification using a novel ensemble support vector machine and frequency ratio method. Stochastic Environmental Research and Risk Assessment, 29(4), 1149–1165.

Retrieved February 4, 2020, from https://link.springer.com/article/10.1007/s00477-015-1021-9#cit