International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

A Review on Ensemble Learning Methods: Machine Learning Approach

Dipanshu Mishra[1*†], Shrikant Mani Tripathi[1], Akash Chaurasia[2], Pawan Kumar Chaurasia[2†]

[1*]Department of Computer Application and Sciences, School of Management Sciences, Lucknow, 19th Kilometer Stone, Sultanpur Road, Gosainganj, Lucknow, 226501, Uttar Pradesh, India.

[1]Department of Mathematics, Sanjay Ghodawat University, Kolhapur- Sangli Highway, Kolhapur, 416118, Maharashtra, India.

[2]Department of Information Technology, Babasaheb Bhimrao Ambedkar University, Lucknow, Vidya Vihar, Raebareli Road, Lucknow, 226025, Uttar Pradesh, India.

E-mail(s): dipanshum.28@gmail.com; Contributing authors: mani.shrikant201@gmail.com; chaurasiaa788@gmail.com; pkc.gkp@gmail.com;

DOI : https://doi.org/10.55248/gengpi.6.0225.0971

ABSTRACT

Unbalanced datasets make it difficult for machine learning to forecast the right class, but the ensemble approach offers a state-of-the-art workaround. Instead of basing predictions on a single model, the ensemble technique combines the pre- dictions of several models to forecast the proper class. This paper's goal is to examine the conventional ensemble approaches, such as bagging, boosting, and stacking generalization, and how they might be used to address the present prob- lems posed by unbalanced datasets. To address these approaches' shortcomings, such as limited diversity in bagging, and overfitting in boosting, Researchers also discuss and compare other modifications of these classic techniques. From a variety of recently released works, Researchers highlight several new areas for ensemble classifier research. The study reviews several earlier theoretical research to offer a fuller understanding of the ensembles themselves.

Keywords: Bagging, Boosting, Stacking, Ensemble learning, Machine learning

Fig. 1 Google Trends database shows 'ensemble' searches over 5 years.

1. Introduction

A machine learning method is ensemble learning, in which many classifiers are employed to forecast a single output. It is also called a multiple-classifier system. To outperform other machine learning algorithms, ensemble learning combines mul- tiple algorithms for machine learning to provide weakly predictive results based on attributes extrapolated from data and results using a variety of voting processes. Ensemble learning involves combining the prediction on multiple individual models to make a final prediction or decision. Ensemble classifiers are more accurate than individual classifiers many in cases. Ensemble learning's concept may be matched to actual life circumstances. Multiple experts are considered when making a crucial decision rather than relying just on one opinion. In many instances, the ensemble has proven to be more precise than every one of the classifiers, however, integrating the model is never successful. A classifier with improved accuracy that differs as much as feasible makes up the optimal ensemble. The overall error will decrease if each classifier produces distinct errors.

Recognizing that models based on machine learning have constraints and can be wrong is the fundamental principle of ensemble learning [1]. To enhance classification performance, ensemble learning makes use of the advantages of many base models. ML algorithms have some drawbacks, such being the fact that the models they pro- duce have high variation, high bias, and poor accuracy. [2],[3] However, several study studies have revealed that ensemble models typically obtain more accurate results than single ML models [4]. The variation and bias mistakes caused by single ML models can be limited using ensemble approaches; for instance, bagging decreases variance without raising bias while boosting decreases bias [5],[6].

Overall, ensemble classifiers outperform individual ensemble learners and are more reliable. This review will give a brief description of the ensemble model and a sur- vey of current ensemble learning techniques. Researchers also discuss deep neural network ensemble learning and training ensemble models using various ensemble tech- niques. The paper's remaining sections are organized as follows: Before discussing the various ways of developing an ensemble methodology and forecasting outcomes by merging various classifiers, the Researcher first covers the background, guiding princi- ples, and benefits of ensemble learning. The different ensemble approaches employed by researchers are then described in the comparison between different ensemble tech- niques and the simplification and reduction of the ensemble. The final portion covers the conclusion and recommendations for further research. In the previous five years as shown in Fig. 1, the search term "ensemble" has shown various degrees of interest among internet users, according to data taken out of the ”Google Trends” database. In this usage, the word ”ensemble” probably refers to a collection of objects, compo- nents, or performers who operate well together. Over the given time range, it appears that the search trend has fluctuated, with periods of rising curiosity followed by rel- ative drops. This pattern could imply that the changes in relevance or focus have affected the term ensemble’s popularity. One may hypothesize that the term’s pop- ularity is related to elements like seasonal fluctuations, cultural events, or business developments that cause individuals to look up the idea of ensembles with additional research and context.

2. Ensemble Learning

The concept that it is the wisdom of the crowd serves as one of the fundamental prin- ciples behind ensemble learning. Ensemble learning tries to provide judgments that are frequently superior to those produced by a single model by mixing knowledge from other models. In ensemble learning, independence, decentralization, variety of view- points, and aggregation are requirements for reaching the wisdom of crowds, when fitting the original data with parallel regression algorithms and the additional data with a separate linear regression model. [7] Tukey initially suggested the application of ensemble learning methods in 1977. Over time, researchers created many methods for fusing various classifiers for training the model. Schapire developed the Adaboost method, which combines several weak learners to produce a powerful learner, in 1990. Using ensemble learning techniques, researchers have recently been able to find solu- tions to a variety of real-world issues in industries including petrochemical products bio informatics healthcare, satellite imagery, higher education, including software bug detection. Ensemble learning is an efficient method for approaching complicated issues because of its adaptability and efficiency.

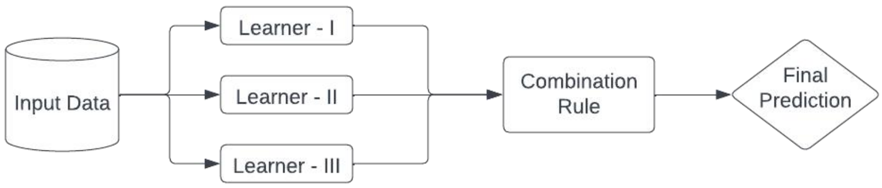

Fig. 2 Parallel Ensemble Model Block Diagram.

Fig. 3 Sequential Ensemble Model Block Diagram.

Ensemble learning is the act of integrating several different models based on machine learning to get results that are superior to those produced by the algorithms when used independently. Instead of relying just on one machine learning model, a combination of rules is used to integrate the predictions from many models, yielding a single forecast that is more accurate than those from different models of machine learning. In general, there are two categories for the ensemble learning approach, 1. A parallel approach and 2. Sequential techniques. The parallel technique in Fig. 2 trains various base learners individually and then uses a combiner to combine their predic- tions. Bagging, a parallel ensemble technique that extends random forest algorithms, is highly popular [8]. Sequential ensemble models in Fig. 3. cannot independently fit the basic model. They get regular training so that every model learns how to correct the error made by the preceding model. The boosting algorithm is one of the com- mon sequential algorithm subtypes [9]. The precision of the base learner has a major factor in the effectiveness of ensemble learning [10]. Any machine learning algorithm is only regarded as correct if it has an effective generalization strategy to previously unobserved examples. Recognizing the limitations and potential for mistakes of the usual machine learning model is the core concept behind ensemble learning. There- fore, by combining the capabilities of many base models, ensemble learning is utilized to increase both the precision and the efficiency of classification issues.

3. Ensemble Model

The ensemble model consists of several separate classifiers that aggregate their outputs using a variety of ensemble methods, including weight averaging, voting, and proba- bilities. Each classifier first trains on a specific dataset. The work from the ensemble learning project is displayed in Fig. 2,3. The precision attained by ensemble learning has been theoretically demonstrated by Dietterich TG [11] and Galar M et al [12]. Researchers describe the many training methods for the ensemble approach to obtain the required result. Researcher’s have to stick to two concepts, primarily diversity and classifier predictor performance, to create a successful ensemble model. Bi[13] noted that the varied classifier would not always increase accuracy, but it does help reduce bias and variance.

3.1 Selection of the Ensemble

A technique called ensemble selection is used to create a classifier with ensembles from a group of base models. Because choosing a good portion of the base model improves performance compared to using all of the models to build the ensemble classifier, ensemble selection is a crucial subject in ensemble learning [14]. The related performance would change since the underlying model has been created using multiple machine learning methods or varied subsets of training data. Some could perform better than others, while others might perform poorly. Selecting only a portion of models having good performance, as opposed to combining good and terrible models, will be advantageous and improve the performance of the ensemble [15]. Caruana et al.[15] developed the initial forward model selection method to choose the initial model subset with the highest performance. Below is a list of the basic steps.

1. Start with a blank outfit.
2. Using a validation set, choose the base classifiers from the library that.
3. Continue performing step II until every model in the collection has been looked at.
4. Returned the model's component with the best validation set performance.

Researchers have found that there are both benefits and limitations to the model selection strategy. While it is generally a fast and productive method, there are many times when it may overfit and result in poor ensemble performance. In response to this issue, there have been several ensemble selection strategies proposed, such as the bagging ensemble selection developed by Sun and Pfahringer [16]. A static or dynamic strategy should be used, with static techniques picking a particular set of base models throughout training and then implementing it across all unseen instances. This is another crucial decision. These static selection approaches fall under two further categories: ordering-based and optimization-based methods. Only the model with the highest score will be picked as the ideal subset according to ordering-based algorithms, which rank the base models according to certain criteria. The validation error is one criterion for ranking the basic model [14]. A technique to arrange the base model for ensemble learning has been developed by Guo et al [17], which uses an assessment measure that considers margins and diversity. The selection process is formulated as an optimization issue via optimization-based approaches, which may be resolved by mathematical programming. The dynamic model dynamically chooses cert

4. Frequently Used Ensemble Methods

Machine learning relies heavily on ensemble approaches, which combine predictions from many models to increase accuracy and robustness. Bagging produces different information subsets to use for individual machine learning models with CatBoost, XGBoost, as well as LightGBM, while Random Forest builds a collection of decision trees with reliable predictions. By correcting prior model flaws, gradient-boosting and AdaBoost incrementally improves predictions. Vote Classifiers/Regressors and Stack- ing integrate the results of several models to make a single forecast while making use of each model's advantages. These methods, which balance bias and variance, improve performance on a variety of tasks, confirming the importance of ensemble approaches in obtaining increased generalization and predictive ability. Recently, the adoption of ensemble learning models composed of many machine learning components has increased. Bagging, boosting, and stacking are the models. Bagging is employed to decrease the variation error in decision trees, boosting is used to decrease bias errors, and stacking is utilized to increase the predictive accuracy of the model.

4.1 Bagging

Bagging is a commonly used ensemble learning technique. It is a method that is used for randomly selecting data from a single dataset; it is sometimes referred to as boot- strap aggregating [18]. Each learner receives a unique random sample of training data from the system. Bagging updates the classifier by replacing n samples from the train- ing data with new ones, making it ideal for models with high levels of complexity like the Decision Tree algorithm. Run it through a few iterations before finally joining every learned classifier. The Random forests have a very similar architecture that makes use of bootstrapping. One of the earliest, most understandable, as well as simplest ensemble-based algorithms, bagging—bootstrap aggregating—has a very high perfor- mance [18]. Each ensemble used in the bagging process has developed using a distinct set of training data, and the predictions were then combined using regular averaging or voting. Each dataset is created by randomly selecting N items from each example of the whole N data set and replacing them with samples from each dataset. [19] The term ”bagging” originated from the term ”bootstrap,” which is used to describe each sample bootstrap building up. Therefore, a bootstrap examines N items. The chance of any given data point with value p = (1-1/N)N is not selected when replacement is distributed at random. Therefore, it is projected that a single bootstrap will pro- vide a high N including around 63.2% of the starting collection, whereas 36.8% of the population that was originally collected is not selected.

A decision tree approach referred to as Random Forest [20] uses two techniques to broaden the range of ensemble-bagging and randomly chosen predictor variables. Randomly chosen statistics derived from p variables are used for variables, and a sample with replacement has been added for datasets. In other words, Random Forest provides diversity in variables as well as data.

The best models for bagging, like many ensemble techniques, are unstable mod- els, which show different generalization behaviors with tiny changes to the training data. Decision trees and artificial neural networks are two examples of what is fur- ther referred to as high-variance models. As a result, bagging frequently fails with really straightforward models. In practice, selecting samples at random from the set of potential models to compose the ensemble sampling produces nearly similar (low variety) forecasts for very simple models. According to the ”Springer Link” database, Fig. 4 shows the development of publications that have used the Bagging technique

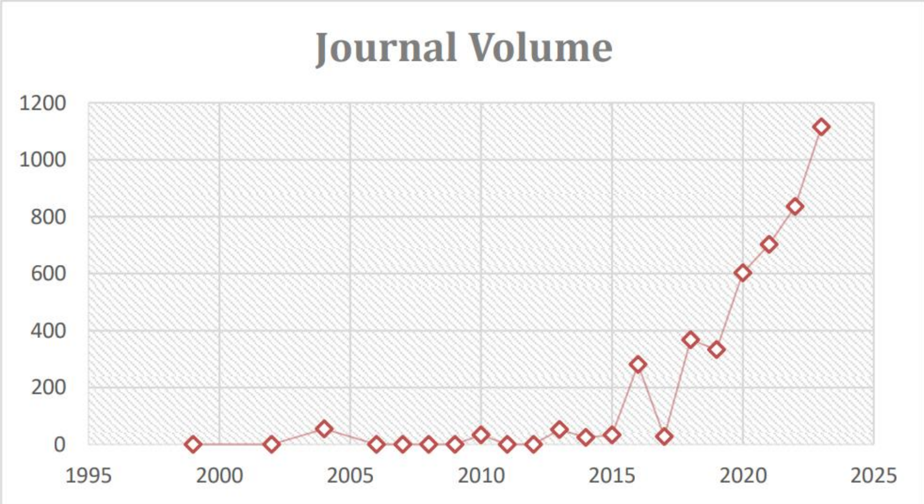

Fig. 4 “Springer Link” database shows publications using the “Bagging” approach through time.

throughout time. The machine learning ensemble strategy known as bagging, or boot- strap aggregating, includes training several models using built portions of the original data set and merging their predictions to enhance overall performance and lower vari- ance. The Bagging technique is becoming more popular among researchers, as seen by the chart in Fig. 5 The trendline shows a steady rise in publications over time, under- scoring the importance of this method as a useful tool for predictive modeling and data analysis. As scholars continue to investigate and utilize the Bagging technique, this rise of papers underlines its ongoing importance.

4.2 Boosting

In a 1990 publication, Schapire [21]initially introduced the boosting method. Increas- ing the goal of the value within a particular point that the prior model did not fit is a technique known as ”boosting.” AdaBoost uses this technique to change the weight of data selection. By using certain measures, the boosting approach makes sure that it stands out significantly from the other members. An additive model (ensemble) is often fitted using a gradient boosting machine (GBM) [22] in a forward stage-wise fashion. A weak learner is added at each stage to make up for the weaknesses of the weak learners already there. Gradients are used in GBM to identify ”shortcomings”. Both gradients and high-weight data points indicate areas have our model needs improvement. Gra- dient boosting generates several approximations. Boosting attempts to increase model accuracy by progressively introducing weak learners, but since it targets data that earlier models failed to successfully predict, it is vulnerable to data that is noisy, out- liers, and overfitting. Popular techniques to minimize variance are crossvalidations as well as bagging (bootstrapped aggregated ensemble). However, boosting is commonly used for reducing bias [23]. Improved algorithms for enhancing machine learning were introduced in the following: XGBoost[24], LightGBM [25], and CatBoost [26]. Based

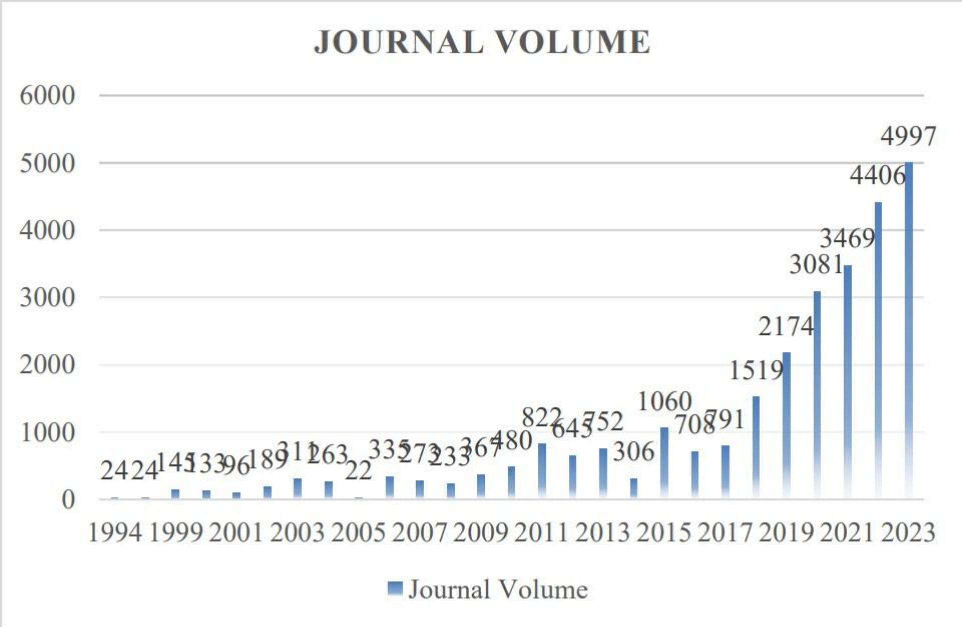

Fig. 5 "Springer Link" database shows publications use the "Boosting" approach through time.

on GBM, XGBoost is a method for processing enormous volumes of data fast. Tech- niques like depth-first tree pruning, caches comprehending, out-of-core computation, regularization to prevent overfitting, integrated cross-validation capabilities, etc. are introduced by XGBoost. LightGBM [25] uses methods including gradient-based one- side samples and special feature bundling. Using conventional GBM, determining the knowledge gain of each potential split point needs screening every piece of information instances for each feature. It takes a while since it has to scan each feature and each data instance. A gradient-based one-side sample is introduced in LightGBM, focusing on objects with large gradients while ignoring those with tiny gradients. The use of exclusive feature bundling, which minimizes the number of features by grouping par- ticular attributes into a single variable, is a different approach. CatBoost [26] adds components to the traditional boosting approach, such as organized target data as well as ordered boosting. If targeted statistics are used in place of the classified feature, the organized target statistics technique is utilized to address the target data leakage issue by using an instance's target value to calculate its feature value. By employing the ordered boosting technique, which necessitates distinct training samples for each stage, the prediction shifting problem is eliminated.

The long-term pattern of publications using the "Boosting" technique, as recorded within the "Springer Link" database, is shown graphically in Fig. 5 Boosting is a potent ensemble learning method that focuses on successively training several weak learners, with each learner aiming to rectify the mistakes of the one before it, thereby improving predicting accuracy. The information displayed in this graph demonstrates a notable increase trend in the total number of publications using the Boosting method over time. This development highlights Boosting's continued importance and broad use in the artificial intelligence training and data science fields. Boosting algorithms are still being explored and improved by researchers, which contributes to their continued significance in enhancing the functionality of different machine-learning algorithms and tackling.

4.3 Stacking

An ensemble learning framework called "stacked generalization" (also known as "Stacking") trains a different machine learning algorithm to integrate several ensem- ble members' predictions. Wolpert [27] has initially introduced in 1992 to lower generalization errors in machine learning problems. To choose whether to use the pre- dictions from different models, another ML model would then be used in the stacking procedure. When several ML models have specialized skills that are individually distinct for a given task, stacking is successful [28].

Creating models primarily involves using several level-0 models, a meta-learning a strategy that tells another model to add predictions using the base models is often referred to just as "base models." A level-1 model of the meta-model is used [29]. The core idea underlying stacking is to train level-0 base learners with the training dataset before exposing them to out-of-sample or unobserved data. The input, as well as the output pairs of a fresh dataset, are then utilized for training the meta-learner by com- bining their anticipated targeted labeling using the real labels on the hidden data [30].

Using the outcomes of additional algorithms for machine learning, meta-learning algorithms are taught to create predictions that are more accurate than those gen- erated by the other basic classifiers [31]. The meta-learner, which learns the model that generates the final prediction, is a crucial component of the stacking system. The metaclassifier discovers the ideal way to integrate the predictions of basic learners[32]. The stacking technique is effective because it combines the benefits of many effec- tive classifiers to provide classifications that are superior to those produced by the ensemble's models.

By employing many base algorithms as well as the same data, stacking also produces models that are unique and tackle the forecasting issue in a new way. In contrast to bagging, which primarily employs decision tree algorithms that are trained on a subset of the input data, the stacking models use various approaches and were developed on the same datasets.

The number of academic papers using the popular ensemble learning technique called ”Stacking” during a certain period is likely depicted in Fig. 7 as a temporal trend. This diagram will show how the academic community's interest in and use of stacking techniques has changed over time. One would anticipate changes and increases in the number of articles over time, indicating the ongoing investigation, study, and use of stacking techniques in various domains. Figure 8 analysis may shed light on Stacking is of ongoing importance and has a growing role in improving the performance

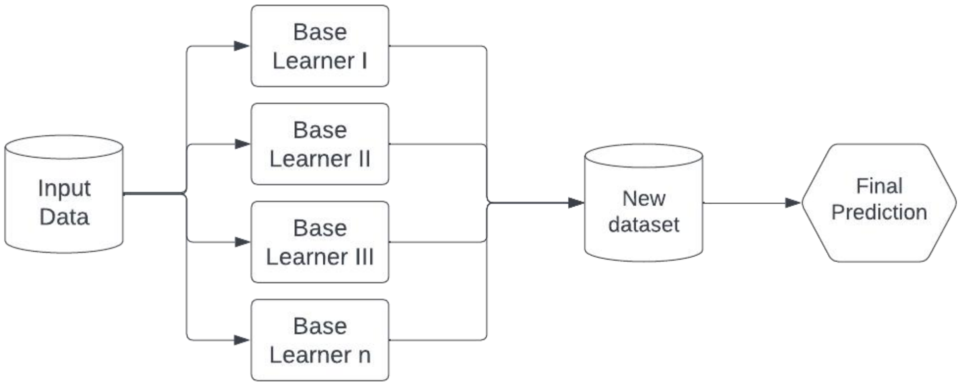

Fig. 6 Block diagram of Stacking approach in Ensemble Methods.

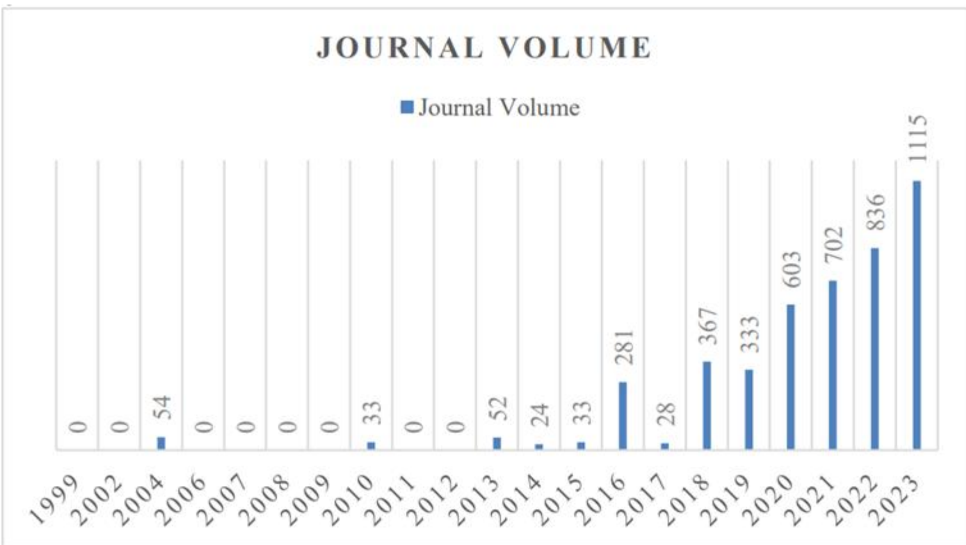

Fig. 7 Springer Link database shows publications using the "Stacking" approach through time. of machine learning models and solving real-world challenges in the fields of data science and machine learning.

5. Comparison Between Different Ensemble Methods

Over the past 20 years, a large number of academics have examined and assessed how various ensemble techniques perform over a wide range of categories.

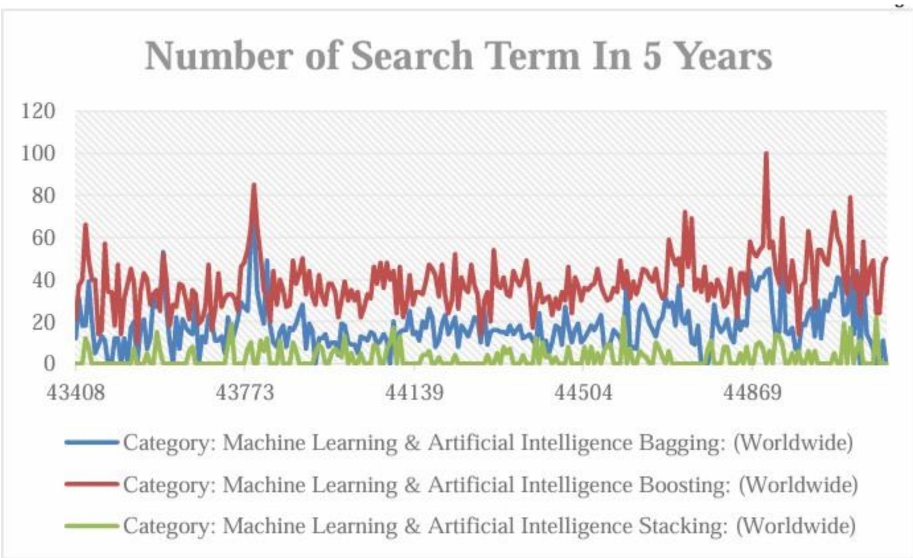

Fig. 8 Google Trends database shows the searches in 5 years in the context of "Machine Learning & Artificial Intelligence"

Dietteric [10] shows that Bagging, which employs a decision tree as its base learner, is less accurate than the sequential method of increasing predictions. The sole restriction for Boosting is that there shouldn't be much noise in the data. Banfield et al. [33] compared bagging against an algorithm that employed seven randomizations and studied how boosting could overfit noisy data. The most efficient randomization technique Bagging outperformed Random Forests and Boosting in only 7 out of 57 datasets. Fernandez et al.'s [34] comparison of 179 algorithms from 17 families over 121 datasets likewise concluded that sequential algorithms beat parallel techniques. There are a few considerations to ponder while selecting the best ensemble strategy for a task. Here are some of them:

Dataset Type: Identify the different types of datasets, such as unbalanced datasets, multiclass datasets, noisy data, and high dimensionality data. This makes it easier to select the machine learning method that is best for the dataset.

Time in Execution: According to Bifet et al. [35], real-time analysis depends heav- ily on execution. Therefore, before choosing a machine learning method for the task, Researcher's must assess the relationship between prediction computation time and accuracy.

Platform Compatibility: Certain platforms offer their ML algorithm, which poses certain difficulties in adapting it to a specific application.

textbfUsefulness: The model's tuning parameters should be transparent in terms of how they affect behaviors. This improves a model's suitability for use with many kinds of issues.

By examining data as shown in Fig. 8 pulled from the Google Trends database, Researcher's mapped out a thorough five-year journey of patterns of searches within the field of 'Machine Learning & Artificial Intelligence.'" Our graphic depiction high- lights the fluctuation in levels of interest, providing insightful information about the changing level of curiosity surrounding these cutting-edge topics. The graph shows the shifting terrain of technological research by highlighting peak times as well as more general fluctuations in popular curiosity.

Table 1 Comparison of Different Ensemble Methods

Ensemble Method	**Bagging**	**Boosting**	**Stacking**
Aim	Decrease Variability	Decrease Bias	Increase Accuracy
Training Base Model	Parallel/Independent	Sequential	Meta Learner Model
Base Learner	Homogeneous (Weak Learner)	Homogeneous (Weak Learner)	Heterogeneous (Strong Learner)
Aggregation	Averaging/Max Voting	Weighted Averaging	Weighted Averaging
Disadvantages	Prone to Underfitting	Prone to Overfitting	Needs much time and computing power

6. Conclusion

In situations when there is an imbalance between multiple forms of data, such as when receiving medical care, purchasing a vehicle or home, etc., To create very effective machine learning algorithms that work with any type of data, an ensemble is employed. The concept of ensembles is based on how individuals make decisions in real-life situ- ations, where the majority seeks out many viewpoints before choosing their course of action. With the use of the "Springer Link" Database, we examined commonly used ensemble models including bagging, boosting, and stacking in this review paper. In recent years, boosting has been the most popular. Comparing the three approaches mentioned above, bagging reduces variance, boosting reduces bias, and stacking mostly improves accuracy. The accuracy of forecasts and a straightforward, all-encompassing approach should be preserved in newly developed ensemble models. Researchers may concentrate on the final two points we made, which include employing neural sys- tems more frequently to create ensembles with multiple generalizations that are more successful.

References

[1] Mienye, I.D., Sun, Y.: A survey of ensemble learning: Concepts, algorithms, appli- cations, and prospects. IEEE Access **10**, 99129–99149 (2022) https://doi.org/10. 1109/ACCESS.2022.3207287

[2] Mishra, S., Sun, Y., Sharma, A.: Improving the accuracy of ensemble machine learning classification models using a novel bit-fusion algorithm for healthcare ai systems. Front. Public Health **10**, 858282 (2022) https://doi.org/10.3389/fpubh.2022.858282

[3] Sun, Y., Li, Z., Li, X., Zhang, J.: Classifier selection and ensemble model for multiclass imbalance learning in education grants prediction. Applied Artifi- cial Intelligence **35**(4), 290–303 (2021) https://doi.org/10.1080/08839514.2021.1877481

[4] Langford, J., Smith, A.D., Brown, T.J.: Ensemble learning. In: Sammut, C., Webb, G.I. (eds.) Encyclopedia of Machine Learning, pp. 312–320. Springer,Boston, MA (2011). https://doi.org/10.1007/978-0-387-30164-8 252

[5] Doroudi, S.: The bias-variance tradeoff: How data science can inform educational debates. AERA Open **6**(4), 233285842097720 (2020) https://doi.org/10.1177/2332858420977208

[6] Alelyani, S.: Stable bagging feature selection on medical data. Journal of Big Data **8**(1), 11 (2021) https://doi.org/10.1186/s40537-020-00385-8

[7] Kumar, S., Kaur, P., Gosain, A.: A comprehensive survey on ensemble meth- ods. In: 2022 IEEE 7th International Conference for Convergence in Technol- ogy (I2CT), pp. 1–7. IEEE, Mumbai, India (2022). https://doi.org/10.1109/ I2CT54291.2022.9825269

[8] Liu, H., Gegov, A., Cocea, M.: Ensemble learning approaches. In: Rule-Based Systems for Big Data. Studies in Big Data, vol. 13, pp. 63–73. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-23696-4 6

[9] Ravichandran, T., Suresh, S., Kumar, R.: Ensemble-based machine learning approach for improved leak detection in water mains. Journal of Hydroinformatics **23**(2), 307–323 (2021) https://doi.org/10.2166/hydro.2021.093

[10] Flach, P.A., Bie, T.D., Cristianini, N.: Machine Learning and Knowledge Discov- ery in Databases: European Conference, ECML PKDD 2012. Lecture Notes in Artificial Intelligence, vol. 7523–7524. Springer, Berlin; New York (2012)

[11] Dietterich, T.G.: An experimental comparison of three methods for construct- ing ensembles of decision trees: Bagging, boosting, and randomization. Machine Learning **40**(2), 139–157 (2000) https://doi.org/10.1023/A:1007607513941

[12] Zhou, Z.-H.: Ensemble learning. Encyclopedia of Biometrics, 270–273 (2009) https://doi.org/10.1007/978-0-387-73003-6 153

[13] Chhabra, A., Malhotra, R.V.: Ensemble learning: A review of deep learning tech- niques for malware detection. Journal of Information Security Applications **55**, 102683 (2020) https://doi.org/10.1016/j.jisa.2020.102683

[14] Chollet, F.: Deep Learning with Python. Manning Publications, ??? (2018)

[15] Polikar, R.: Ensemble based systems in decision making. IEEE Circuits and Sys- tems Magazine **6**(3), 21–45 (2006) https://doi.org/10.1109/MCAS.2006.1688199

[16] Rokach, L.: Ensemble-based classifiers. Artificial Intelligence Review **33**(1), 1–39 (2010) https://doi.org/10.1007/s10462-009-9124-7

[17] Freund, Y., Schapire, R.E.: A decision-theoretic generalization of online learning and an application to boosting. Journal of Computer and System Sciences **55**(1), 119–139 (1997) https://doi.org/10.1006/jcss.1997.1504

[18] Ho, T.K.: The random subspace method for constructing decision forests. IEEE Transactions on Pattern Analysis and Machine Intelligence **20**(8), 832–844 (1998) https://doi.org/10.1109/34.709601

[19] Wolpert, D.H.: Stacked generalization. Neural Networks **5**(2), 241–259 (1992) https://doi.org/10.1016/S0893-6080(05)80023-1

[20] Krizhevsky, A., Sutskever, I., Hinton, G.E.: Imagenet classification with deep convolutional neural networks. Communications of the ACM **60**(6), 84–90 (2017) https://doi.org/10.1145/3065386

[21] Shannon, C.E.: A mathematical theory of communication. Bell System Tech- nical Journal **27**(3), 379–423 (1948) https://doi.org/10.1002/j.1538-7305.1948. tb01338.x

[22] Kittler, J., Hatef, M., Matas, J.: On combining classifiers. IEEE Transactions on Pattern Analysis and Machine Intelligence **20**(3), 226–239 (1998) https://doi. org/10.1109/34.667881

[23] Freund, Y., Schapire, R.E., Lee, W.S.: Boosting the margin: A new explanation for the effectiveness of voting methods. Annals of Statistics **26**(5), 1651–1686 (1998) https://doi.org/10.1214/aos/1024691352

[24] Breiman, L.: Bagging predictors. Machine Learning **24**(2), 123–140 (1996) https://doi.org/10.1007/BF00058655

[25] Breiman, L.: Random forests. Machine Learning **45**(1), 5–32 (2001) https://doi. org/10.1023/A:1010933404324

[26] Chen, T., Guestrin, C.: Xgboost: A scalable tree boosting system. In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 785–794 (2016). https://doi.org/10.1145/2939672.2939785

[27] Prokhorenkova, L., Gusev, A.V., Gulin, A.: Catboost: Unbiased boosting with categorical features. arXiv preprint (2017) https://doi.org/10.48550/arXiv.1706. 09516

[28] Witten, I.H., Hall, E.F.: Data Mining: Practical Machine Learning Tools anTechniques, 3rd edn. Morgan Kaufmann, ??? (2011)

[29] O., S., L., R.: Ensemble learning: A survey. WIREs Data Mining and Knowledge Discovery **8**(4) (2018) https://doi.org/10.1002/widm.1249

[30] al., M.L.: A stacking ensemble learning framework for genomic prediction. Fron- tiers in Genetics **12**, 600040 (2021) https://doi.org/10.3389/fgene.2021.600040

[31] Hospedales, T.M., Antoniou, P.M., A.J., S.: Meta-learning in neural networks: A survey. IEEE Transactions on Pattern Analysis and Machine Intelligence (2021) https://doi.org/10.1109/TPAMI.2021.3079209

[32] R., V., Y., D.: A perspective view and survey of meta-learning. Artificial Intelli- gence Review **18**(2), 77–95 (2002) https://doi.org/10.1023/A:1019956318069

[33] Banfield, R.E., Hall, K.W.B.L.O., W.P., K.: A comparison of decision tree ensem- ble creation techniques. IEEE Transactions on Pattern Analysis and Machine Intelligence **29**(1), 173–180 (2007) https://doi.org/10.1109/TPAMI.2007.250609

[34] Bifet, A., Holmes, B.P.R.K.G., R., G.: New ensemble methods for evolving data streams. In: Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 139–148 (2009). https://doi.org/10. 1145/1557019.1557041

[35] Fern´andez-Delgado, M., Cernadas, S.B.E., D., A.: Do we need hundreds of clas- sifiers to solve real-world classification problems? Journal of Machine Learning Research **15**, 3133–3181 (2014)