ORİJİNAL ARAŞTIRMA ORIGINAL RESEARCH

DOI: 10.5336/biostatic.2022-89894

Bayesian Additive Regression Trees for Predicting Colon Cancer: Methodological Study (Validity Study)

Kolon Kanserini Öngörmede Bayesian Eklemeli Regresyon Ağaçları: Metodolojik Çalışma (Geçerlik Çalışması

Oyebayo Ridwan OLANIRAN[a], Saidat Fehintola OLANIRAN[b], Jumoke POPOOLA[a], Ifeyinwa Vivian OMEKAM[a]

[a]Department of Statistics, Faculty of Physical Sciences, University of Ilorin, Ilorin, PMB 1515, Kwara State, Nigeria
[b]Department of Statistics and Mathematical Sciences, Faculty of Pure and Applied Science, Kwara State University, PMB 1530 Malete, Kwara State, Nigeria

ABSTRACT Objective: The occurrence of colon cancer starts in the inner wall of the large intestine. The survival of colon cancer patients strongly relies on early detection. Diagnosing colon cancer using clinical approaches often takes longer, especially in most developing countries with limited facilities. The recent use of microarray technology has presented a new approach for the oncologist to diagnose cancer cells using non-clinical machine learning methods. In this paper, the aim is to predict the status of colon cancer tissues using the Bayesian Additive Regression Trees (BART) and 2 other machine learning methods. **Material and Methods:** The development and comparative analysis of BART alongside 2 other competing methods (Random Forest: RF and Gradient Boosting Machine: GBM) were implemented. The dataset used for the analysis is the microarray colon cancer data which consists of 2,000 gene expression measurements for 62 tissue samples. **Results:** The methods are compared based on overall metrics (accuracy, balance accuracy, detection rate, F-measure and AUC) and class-specific metrics (sensitivity, specificity, positive predictive value and negative predictive value). The overall metrics results showed that the best method is RF. The class-specific metrics results showed that BART is better than RF. **Conclusion:** On average, BART is more sensitive in detecting the presence of colon cancer cells, while RF is more accurate and specific in detecting the presence or absence of colon cancer cells.

Keywords: Colon cancer; Bayesian trees; random forest; gradient boosting

ÖZET Amaç: Kolon kanseri kalın bağırsağın iç duvarında başlar. Kolon kanseri hastalarının sağ kalımı kuvvetle erken tanıya dayanır. Kolon kanserine klinik yaklaşımlarla tanı koyulması özellikle sınırlı kaynakları olan gelişmekte olan ülkelerde sıklıkla uzun zaman alır. Son zamanlarda mikrodizilim teknolojisinin kullanımı onkologlara klinik olmayan makine öğrenme yöntemleri kullanılarak kanser hücrelerini tanımaları için yeni bir yaklaşım sunmaktadır. Bu yazının amacı Bayesian Eklemeli Regresyon Ağaçları [Bayesian Additive Regression Trees (BART) ve diğer 2 makine öğrenme yöntemi kullanılarak kolon kanseri dokularının durumunun öngörülmesidir. **Gereç ve Yöntemler**: Diğer 2 hesaplama yöntemi olan Rastgele Orman (Random Forest: RF) ve Gradyan Artırma Makinesi (Gradient Boosting Machine: GBM) yanı sıra BART'ın geliştirilmesi ve karşılaştırmalı analizi uygulandı. Analiz için kullanılan veri seti, 62 doku örneği için 2.000 gen ekspresyon ölçümünden oluşan mikrodizi kolon kanseri verisidir. **Bulgular**: Yöntemler, genel ölçülere (doğruluk, terazi denge doğruluğu, saptama oranı, F-ölçüm ve AUC) ve sınıfa özgü ölçülere (duyarlılık, özgüllük, pozitif tahmin değeri ve negatif tahmin değeri) dayalı olarak karşılaştırıldı. Genel ölçüm sonuçları, en iyi yöntemin RF olduğunu göstermiştir. Sınıfa özel ölçü sonuçları, BART'ın RF'den daha iyi olduğunu göstermiştir. **Sonuç**: Ortalama olarak, BART kolon kanseri hücrelerinin varlığını tespit etmede daha duyarlıyken, RF kolon kanseri hücrelerinin varlığını veya yokluğunu tespit etmede daha doğru ve özgüldür.

Anahtar kelimeler: Kolon kanseri; Bayesian ağaçları; rastgele orman; gradyan artırma

Recent research in biostatistics and bioinformatics focuses on diagnosing diseases using non-clinical approaches that involve machine learning methods. Several algorithmic procedures have been applied to solve various experimental problems that involve simulation and modelling of deoxyribonucleic acid (DNA) and ribonucleic acid (RNA) proteins.[1-6] The DNA and RNA are essential biological measurements used to monitor the abnormal cell growth in genetic sequencing, which serves as the bedrocks in non-clinical diagnoses.

Correspondence: Oyebayo Ridwan OLANIRAN
Department of Statistics, Faculty of Physical Sciences, University of Ilorin, Ilorin, PMB 1515, Kwara State, Nigeria
E-mail: rid4stat@yahoo.com

Peer review under responsibility of Turkiye Klinikleri Journal of Biostatistics.

Received: 22 Mar 2022 ***Received in revised form:*** 24 May 2022 ***Accepted:*** 25 May 2022 ***Available online:*** 08 Jun 2022

Most machine learning algorithms aim to efficiently identify essential biomarkers that are useful for classifying disease groups. This is achieved by focusing on optimising the loss functions. The earlier developed methods have been designed to work on low dimensional data (*n*>*p*, where *n:* number of biological samples, *p:* number of genes). Gene expression datasets such as the colon cancer dataset do not conform to this criteria because the number of the observed tissue samples is far less than the number of observed genes.[7] This scenario is termed high-dimensionality or "low *n* or large *p*" syndrome. Baseline classification methods such as logistic regression and linear discriminant analysis often break down in this situation. The foremost solution to this problem is to perform stage-wise analysis by ensuring the number of selected genes is less than the number of tissue samples. This form of analysis is suboptimal because it ignores the interaction between genes.[5,6]

The decision trees algorithm proffers an interim solution to this issue. However, its performance is not stable based on findings from several empirical studies.[5,8] This led to the development of ensemble methods that combine homogenous or heterogeneous machine learning methods to build a new algorithm.[9,10]

Bagging was one of the earliest development of the ensemble of regression trees model.[11] Bagging combines multiple bootstrapped trees to improve the classification performance of a single tree.[12,13] Boosting improves the performance of weak learners by iteratively updating the model at different stages.[14] On the other hand, Random Forest (RF) updates the performance of a single tree using subsets of the original variables used in building a tree.[11] Breiman showed that RF is better than the single tree method and Bagging since a random subset of trees would result in forests with uncorrelated trees.[11] While RF performances are incredible in most machine learning problems, it still faces acceptance issues among biostatisticians due to its black-box nature. RF has no known probabilistic framework or model but only a step by step algorithmic problem-solving approach.[14]

Moreover, Bayesian approaches are the new emerging probabilistic approach that is realistic and provides lower error than other classical maximum likelihood-based approaches.[4-6,15-19] Therefore, in this paper, we aim to compare the predictive performance of the Bayesian additive regression trees Bayesian Additive Regression Trees (BART) with RF and gradient boosting machine (GBM) using micro-array colon cancer data.[14]

MATERIAL AND METHODS

To illustrate the applicability and comparability of BART, the microarray colon dataset was employed.[7] The dataset consists of an experiment that consists of 62 biological samples, of which 40 revealed malignant presence while the rest 22 samples show the absence of any tumorous cells. The dataset was preprocessed using the standard log_{10} transformation which removes outliers. The following subsections briefly describe the non-clinical diagnostic procedures employed in this study.

BART: Chipman et al. introduced the BART as a sum of tree models using a Bayesian formulation.[14] BART models the tissue sample outcomes as the target (*Y*) and the genes (*X*). This led to the introduction of the sum of trees model given below:

$$Y = \sum_{i=1}^{m} g(X; T_i, M_i) + \varepsilon, \ \ \varepsilon \sim N(0, \sigma^2) \qquad (1)$$

where T_i is the decision tree $i = 1, 2, . . . , m$ total trees, and M_i is the associated terminal node parameters for each tree. The model fitting is achieved via a combination of the back-fitting algorithm and Gibbs-sampler. This procedure simultaneously generates the posterior samples of terminal node parameters and their standard deviations. Chipman et al. gave the posterior distribution $p(T, M, \sigma | X, Y)$ for the BART model as[14]

$$p(T, M, \sigma | X, Y) \propto L(T, M, \sigma | X, Y) \times p(T, M, \sigma) \qquad (2)$$

where $L(T, M, \sigma | X, Y)$ is the joint likelihood function of the parameters, $p(T, M, \sigma)$ is the prior joint probability of the parameters.[14] Based on the principle of independence and symmetry, (2) becomes;

$$p(T, M, \sigma | X, Y) \propto L(T, M, \sigma | X, Y) \times \prod_i p(M_i | T_i)\, p(T_i) p(\sigma) \qquad (3)$$

where $p(M_i | T_i)$ is the conditional prior distribution of each terminal node M_i which is distributed $N(\mu_{M_i}, \sigma^2_{M_i})$, $p(T_i)$ is the prior distribution of each tree T_i which according to Chipman et al. is $p(T_i) = \alpha(1 + d)^{-\beta}, \alpha \in (0,1), \beta \in [0, \infty)$, and is the prior distribution for the model variance which is chi-squared v degrees of freedom distributed χ^2_v.[14] Equations (1) and (2) are valid for the regression problem. On the other hand, the classification or diagnostic model is based on probit regression modelling, defined as

$$p(Y = 1 | X) = \Phi(\sum_{i=1}^{m} g(X; T_i, M_i)) \qquad (4)$$

where Φ is the cumulative distribution function of standardised normal distribution. The final diagnostic prediction is obtained by computing the trees' average.

RF

Breiman (2001) introduced RF as a sum of trees model, which is an update over the earlier developed Bagging, a solution provided for the instability in the terminal nodes estimates of decision trees. RF iteratively select random subsets of predictors to build trees that make the forest. For classification forest applicable to disease diagnostic, the recommended threshold for the number of random subsets to be selected is (*p/3*).

The algorithm below shows the step by step approach used by RF to diagnose colon cancer in the study.

RF algorithm

Step 1: Resampling the original sample size (*n=62*) *B* number of times to generate bootstrap samples.

Step 2: For each bootstrapped sample generated, a classification tree $\hat{f}_b(y = \{0,1\} | \boldsymbol{x_k})$ is fitted to a maximal depth, where $\boldsymbol{x_k}$ corresponds to the gene subset used to build each tree.

Step 3: Obtain the final estimate using majority votes averaging procedures which imply each tree in the forest will vote for a predictive class (malignant or normal) to which the most representative class win the vote.

Stochastic gradient boosting: Freidman presented the gradient boosting procedures as an ensemble approach that focuses on improving the predictive performance of weak or base models.[20,21] The gradient boosting machine was designed to minimise the loss function $\hat{k} = \varphi(y, \hat{g})$ in general function estimation problems. The Freidman gradient boosting algorithm is summarised below.

Gradient boosting machine algorithm

```
1. Let $\hat{g}(x) = 0$ and $k_d = y_d$ for all $d$ in the training dataset.
For $m$ in $1, \ldots, M$ do
2. Fit a tree $\hat{g}_h(x)$ with $q$ splits ($q + 1$ terminal nodes) to the training data $(X, k)$.
3. Update $\hat{g}_m$ by adding in a regularised form of a new tree:
        $\hat{g}(x) \leftarrow \hat{g}(x) + \lambda\hat{g}_m(x)$
4. Update the residuals,
        $k_d \leftarrow k_d - \lambda\hat{g}_m(x)$
```

5. Print the final model

$$\hat{g}(x) = \sum_{m=1}^{M} \lambda \hat{g}_h(x)$$

Performance Metrics: Given a cancer diagnostic confusion matrix as in Table 1,

TABLE 1: Confusion matrix.

True class	Predicted class Normal	 Tumour	Total
Normal	*a*	*b*	*a*+*b*
Tumour	*c*	*d*	*c*+*d*
Total	*a*+*c*	*b*+*d*	*n*

where a: represents the actual number of normal tissues predicted as normal tissues, b: is the number of normal tissues predicted as tumour tissues, c: is the number of tumour tissue predicted as normal tissue and d: is the number of tumour tissues predicted as tumour tissue. Also, (a+c) represents the tissue samples that were predicted as normal, (b+d): is the tissue samples predicted as tumourous. Similarly, (a+b) is the actual number of normal samples, and (c+d) is the actual number of tumour samples. Therefore, the following metrics according to were computed:[22]

Accuracy $= \frac{a+d}{n}$

Sensitivity $= \frac{d}{c+d}$

Specificity $= \frac{a}{a+b}$

Balance Accuracy $= \frac{Sensitivity+Specificity}{2}$

Positive Predictive Value $= \frac{d}{a+c}$

Negative Predictive Value $= \frac{a}{b+d}$

Prevalence $= \frac{c+d}{n}$

Detection Rate $= \frac{d}{n}$

F-Measure $= 2\left(\frac{Positive\ Predictive\ Value \times Sensitivity}{Positive\ Predictive\ Value + Sensitivity}\right)$

The last classification metric employed for the diagnostic procedures is the area under the receiver operating characteristics curve (AUC).[23] Hanley & McNeil presented the estimate using a statistic similar to the non-parametric Mann-Whitney U.[24]

$$AUC = \frac{U}{n_0 n_1}$$

where n_0, n_1 represent the number of normal and malignant tissue samples, respectively.

Comparison: Different performance measures were used based on the different formulations of loss functions used in the development of each algorithm. Another reason for using different measures is that the data is unbalanced; there are more tumourous cells (65%) than normal cells (35%). This is expected to influence the likelihood of predicting tumorous cells compared to normal cells. Thus, the Freidman test and the Nemenyi test were employed to compare each algorithm's ranks across the metrics to harmonise the metric results.[9,25]

RESULTS

Table 2 presents the performances of the algorithms using the performance metrics defined earlier. The results represent the average of a 10-folds cross-validation of the original dataset. The R statistical package "Caret" was used for the cross-validation, package "bartMachine" for BART, package "randomForest" for RF and "gbmboost" for GBM. Furthermore, the results are based on holdout samples and not on the training test. Table 3 presents each algorithm's corresponding ranks based on performance metrics. The Nemenyi test is presented in Table 4. This test compares the pair of ranks for the methods.

TABLE 2: Performance measures for BART, RF, GBM based on the average of 10 folds cross-validation.

Performance metrics	BART	RF	GBM
Sensitivity	0.95	0.90	0.85
Specificity	0.50	0.70	0.68
Positive predictive value	0.81	0.90	0.88
Negative predictive value	0.91	0.87	0.78
F-measure	0.87	0.90	0.87
Detection rate	0.62	0.59	0.55
Accuracy	0.80	0.84	0.79
Balance accuracy	0.73	0.80	0.76
AUC	0.92	0.93	0.86

BART: Bayesian Additive Regression Tree; RF: Random Forest; GBM: Gradient boosting machine; AUC: Area under the receiver.

TABLE 3: Ranks of the three non-clinical diagnostic algorithms. First: 1, second: 2, and third: 3.

Performance metrics	BART	RF	GBM
Sensitivity	1	2	3
Specificity	3	1	2
Positive predictive value	3	1	2
Negative predictive value	1	2	3
F-measure	2.5	1	2.5
Detection rate	1	2	3
Accuracy	2	1	3
Balance accuracy	3	1	2
AUC	2	1	3
Average	2.05	1.44	2.61

BART: Bayesian Additive Regression Tree; RF: Random Forest; GBM: Gradient boosting machine; AUC: Area under the receiver.

Table 3 ranks the method in decreasing order of magnitude. The method with the highest metrics receives a value of 1, the next receives a value of 2, and the least receives 3. If there is a tie in the metrics, the rank is average, and the method takes the same average rank. This ranking arrangement implies that the method with the least average ranking is the best. Thus, in Table 3, the best method in terms of least average ranking is RF. The comparison is further explored using the Freidman and Nemenyi tests.

Table 4 presents three pairwise comparison test p-values. Pair 1 is GBM vs BART, Pair 2 is RF vs BART and Pair 3 is RF vs GBM.

TABLE 4: Nemenyi posthoc p-values.

	BART	GBM
GBM	0.466	-
RF	0.276	0.018

BART: Bayesian Additive Regression Tree; RF: Random Forest;
GBM: Gradient boosting machine; AUC: Area under the receiver.

DISCUSSION

The performance metrics can be subdivided into 2 groups. The first group consists of overall metrics, while the other group consists of class-specific metrics. For the overall (accuracy, balance accuracy, detection rate, F-measure and AUC), the best method is RF on most except the detection rate. BART competes with RF and takes the 2nd position, with GBM in this category least. BART and RF compete in this category with leading positions in 2 metrics for the class-specific metrics (sensitivity, specificity, positive predictive value and negative predictive value). While BART is the most sensitive in detecting the presence of colon cancer cells among tumorous cells, RF is the most specific in detecting the absence of colon cancer among normal cells. Again, GBM has the least performance in this group. Overall, the average of the ranks in Table 3 shows that the best algorithm based on all metrics used is RF, closely followed by BART. The competing results were tested for the difference using the Friedman test, and it shows a significant difference at the 5% level (*Q(2)=7.6*, *p=0.0224*). The significance was further explored using the Nemenyi pairwise posthoc comparison, which is presented in Table 4.[25] The result in Table 4 showed a significant difference between RF and GBM at the 5% level. This implies that the rejection of the Freidman test null hypothesis is due to the difference between RF and GBM. Also, there is no difference between the pair of (GBM vs BART) and (RF vs BART).

The findings of this study are related to what was obtained in, where it was found that RF achieved 100% accuracy for predicting breast cancer.[26] However, reported the inadequacy of relying on accuracy for comparing algorithms as the metrics strongly depend on the ratio of the 2 classes in the dataset.[27] They also recommended using statistical tests as done here to compare the significance of the difference observed among the tests. Furthermore, the sensitivity of BART was found to be similar to the findings in; it was reported that Bayesian methods are new effective procedures for assessing the sensitivity of diseases such as cancer.[28] This is due to the inclusion of relevant prior information during model prediction formulations.

CONCLUSION

This paper presented the BART, RF and GBM machine learning algorithms for predicting 62 colon cancer cells based on 2,000 gene expression profiles. The results revealed that BART is the most sensitive method for detecting the presence of colon cancer cells among tumourous cells. However, RF is more accurate and specific for detecting both presence and absence of colon cancer cells in the overall sample. The best classifier using all the metrics is RF, followed by BART and GBM.

Source of Finance

During this study, no financial or spiritual support was received neither from any pharmaceutical company that has a direct connection with the research subject, nor from a company that provides or produces medical instruments and materials which may negatively affect the evaluation process of this study.

Conflict of Interest

No conflicts of interest between the authors and/or family members of the scientific and medical committee members or members of the potential conflicts of interest, counseling, expertise, working conditions, share holding and similar situations in any firm.

Authorship Contributions

***Idea/Concept:** Oyebayo Ridwan Olaniran; **Design:** Oyebayo Ridwan Olaniran; **Control/Supervision:** Oyebayo Ridwan Olaniran; **Data Collection and/or Processing:** Oyebayo Ridwan Olaniran, Saidat Fehintola Olaniran; **Analysis and/or Interpretation:** Oyebayo Ridwan Olaniran, Ifeyinwa Vivian Omekam; **Literature Review:** Jumoke Popoola; **Writing the Article:** Saidat Fehintola Olaniran, Ifeyinwa Vivian Omekam; **Critical Review:** Oyebayo Ridwan Olaniran.*

REFERENCES

1. Sim AY, Minary P, Levitt M. Modeling nucleic acids. Curr Opin Struct Biol. 2012;22(3):273-8. [Crossref] [PubMed] [PMC]
2. Banjoko AW, Yahya WB, Garba MK, Olaniran OR, Dauda KA, Olorede KO. Efficient support vector machine classification of diffuse large B-cell lymphoma and follicular lymphoma MRNA tissue samples. Annals Computer Science Series. 2015;13(2):69-79. [Link]
3. Olaniran OR, Abdullah MAA. Gene selection for colon cancer classification using bayesian model averaging of linear and quadratic discriminants. Journal of Science and Technology. 2017;9(3):140-4. [Link]
4. Olaniran OR, Abdullah MAA. BayesRandomForest: An R implementation of Bayesian Random Forest for Regression Analysis of High-dimensional Data. Romanian Statistical Review. 2018;66(1):95-102. [Crossref]
5. Olaniran OR. Abdullah MAA. Bayesian variable selection for multiclass classification using Bootstrap Prior Technique. Austrian Journal of Statistics. 2019;48(2):63-72. [Crossref]
6. Olaniran OR, Abdullah MAA. Bayesian analysis of extended cox model with time-varying covariates using bootstrap prior. Journal of Modern Applied Statistical Methods. 2020;18(2):7-17. [Crossref]
7. Alon U, Barkai N, Notterman DA, Gish K, Ybarra S, Mack D, et al. Broad patterns of gene expression revealed by clustering analysis of tumor and normal colon tissues probed by oligonucleotide arrays. Proc Natl Acad Sci U S A. 1999;96(12):6745-50. [Crossref] [PubMed] [PMC]
8. Lin G, Shen C, Shi Q, Van den Hengel A, Suter D. Fast supervised hashing with decision trees for high-dimensional data. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Columbus, OH, USA; 2014. p.1963-970. [Crossref]
9. Zhou ZH. Ensemble Methods: Foundations and Algorithms. 1st ed. London: Chapman & Hall/CRC; 2012. [Crossref]
10. Yang P, Hwa Yang Y, Zhou B, Zomaya Y. A review of ensemble methods in bioinformatics. Current Bioinformatics. 2010;5(4):296-308. [Crossref]
11. Breiman L. Random forests. Machine Learning. 2001;45:5-32. [Crossref]
12. Breiman L. Stacked regressions. Machine Learning. 1996a;24:41-64. [Crossref]
13. Breiman L. Bagging predictors. Machine Learning. 1996b;26:123-40. [Crossref]
14. Chipman HA, George EI, McCulloch RE. BART: Bayesian Additive Regression Trees. Annals Applied Statistics. 2010;4(1):266-98. [Crossref]
15. Linero AR. Bayesian regression trees for high-dimensional prediction and variable selection. Journal of the American Statistical Association. 2018;113(522):626-36. [Crossref]
16. Hernández B, Raftery AE, Pennington SR, Parnell AC. Bayesian additive regression trees using bayesian model averaging. Stat Comput. 2018;28(4):869-90. [Crossref] [PubMed] [PMC]
17. Yahya WB, Olaniran OR, Ige SO. On Bayesian Conjugate Normal Linear Regression and Ordinary Least Square Regression Methods: A Monte Carlo Study. Ilorin Journal of Science. 2014;1(1):216-27. [Crossref]
18. Olaniran OR, Olaniran SF, Yahya WB, Banjoko AW, Garba MK, Amusa LB, et al. Improved Bayesian Feature Selection and Classification Methods Using Bootstrap Prior Techniques. Anale SerialInformatică. 2016;14(2):46-52. [Link]
19. Olaniran OR, Yahya WB. Bayesian hypothesis testing of two normal samples using bootstrap prior technique. Journal of Modern Applied Statistical Methods. 2017;16(2):618-38. [Crossref]
20. Friedman JH. Greedy function approximation: a gradient boosting machine. The Annals of Statistics. 2001;29(5):1189-232. [Crossref]
21. Friedman JH. Greedy function approximation: a gradient boosting machine. Ann Statist. 2002;29(5):1189-232. [Crossref]
22. Powers DMW. Evaluation: From precision. Recall and F-measure to ROC informedness, markedness & correlation. Journal of Machine Learning Technologies. 2011;2(1):37-63. [Crossref]
23. Fawcett,T. An introduction to ROC analysis. Pattern Recognition Letters. 2006;27(8):861-74. [Crossref]
24. Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. Radiology. 1982;143(1):29-36. [Crossref] [PubMed]
25. Demsar J. Statistical comparisons of classifiers over multiple data sets. Journal of Machine Learning Research. 2006;7:1-30. [Link]
26. Octaviani TL, Rustam DZ. Random forest for breast cancer prediction. In AIP Conference Proceedings. 2019;2168(1):020050. [Crossref]
27. Statnikov A, Wang L, Aliferis CF. A comprehensive comparison of random forests and support vector machines for microarray-based cancer classification. BMC Bioinformatics. 2008;9(1):1-10. [Crossref] [PubMed] [PMC]
28. Jackins V, Vimal S, Kaliappan M, Lee MY. AI-based smart prediction of clinical disease using random forest classifier and Naive Bayes. The Journal of Supercomputing. 2021;77(5):5198-219. [Crossref]