Economic Sciences
https://economic-sciences.com
ES(2026) | Vol. 22, No. MIC'26 | Pg: 214-222 | ISSN:1505-4683

Comparison of Random Forest, Neural Network and Naïve Baye's in Classification Problem– an Illustration using Orange

Dr. Deepak Dagar[1], Mr. Praveen Kumar Singh[2], Dr. Ajay Sharma[3], Dr. Gaurav Aggarwal[4], Dr. Manoj Verma[5]

[1]Associate Professor, Dept of Business Admin,
Maharaja Agrasen Institute of Management studies, Delhi (INDIA), deepakdagar.faculty@maims.ac.in
[2]Dept of Commerce, Maharaja Agrasen Institute of Management studies, Delhi (INDIA),
praveenkumarsingh.faculty@maims.ac.in,
[3] Associate Professor, GNIOT Institute of Professional Studies, Greater Noida (INDIA), ajay0202@gmail.com
[4]Head, Department of Economics, Maharaja Agrasen Institute of Management studies, Delhi (INDIA),
hodeconomics@maims.ac.in
[5]Head, Department of BBA, Maharaja Agrasen Institute of Management studies, Delhi (INDIA),
hodbusinessadmin@maims.ac.in

Abstract

In this study, we use the orange data mining platform to examine the performance of popular classification algorithms: Random Forest, Neural Network, and Naïve Bayes. Assessing and comparing these models' precision, interpretability, and computational effectiveness on a standard classification task is the aim. The tests were conducted using a classical iris dataset comprising of 150 instances of iris setosa, iris virginica and iris versicolor. Models were trained and evaluated using 10-fold cross-validation using Orange's visual programming environment, and performance metrics such as recall, accuracy, precision, F1 score, and AUC were noted. The findings show that Random Forest continuously produced the best accuracy and robustness, Neural Networks performed competitively but needed to be carefully adjusted to prevent overfitting, and Naïve Bayes provided ease of use and quick training times but had somewhat worse predictive accuracy. Orange uses drag and drop techniques and assist in the implementation of model which otherwise required coding and expertise in programming language. The insights are useful in identifying the appropriate method and choose the best algorithms.

Keywords: *Random Forest, Neural Network, Naïve Baye's, Illustration using Orange*

I. Introduction

The objective of classification, a subset of supervised machine learning, is to use patterns discovered in past data to predict a label or category for incoming data.

Examples of classification include email filtering (spam or not spam), medical diagnostic, approval of loan application and image recognition. The Classification works in 2 phases – Training and prediction phase. During the training phase, the model is fed with the labelled data ie. Features and it understand and learn based on patterns and rules to separate different classes.

In the prediction phase when a new input is fed to the model, the model predicts based on the features it learns from training phase and predict the correct class label.

Problem are classified in classified as binary classification (Yes/No), multiclass classification (more than 2 classes ie. Dog, cat, bird etc) and multi-label classification (eg. Movie can be action, comedy, horror etc). Logistic regression, decision trees, random forest (RF), Naïve Bayes, support vector machines (SVM), K-Nearest Neighbors (KNN), and neural networks (NN) are examples of popular classification techniques. In this paper, we will use Random Forest (RF), Neural Network (NN) and Naïve Bayes algorithm for solving the classification problem in the iris dataset and will compare the result-based recall, accuracy, precision, F1 score, and AUC.

Economic Sciences
https://economic-sciences.com
ES(2026) | Vol. 22, No. MIC'26 | Pg: 214-222 | ISSN:1505-4683

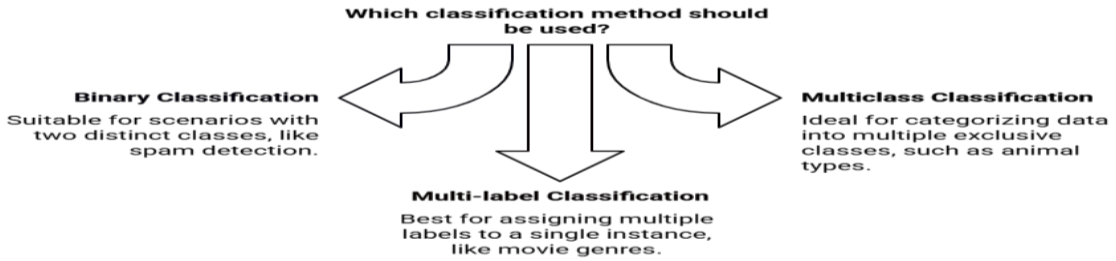


Fig 1 Types of classification

II. Literature Review

The usefulness of machine learning algorithms for predictive analytics and decision assistance has grown due to the quick expansion of data across several areas. The classification involves Naïve Bayes (NB), Random Forest (RF), and Artificial Neural Networks (ANN). These algorithms have been used in many fields, such as sentiment analysis, cybersecurity, educational data mining, and healthcare prediction. Model complexity, feature selection, and dataset characteristics all affect how well they function. The advantages, disadvantages, and relative performance of these algorithms are highlighted in recent research that has been published in peer-reviewed publications.

II(a) Random Forest related articles in Recent Research

Recent studies indicate that Random Forest performs particularly well with structured datasets containing multiple variables. For example, research on exoplanet detection using Kepler time-series data evaluated several machine learning classifiers and reported that Random Forest achieved the highest accuracy of 94.2%, outperforming logistic regression, decision trees, and Naïve Bayes classifiers. The algorithm also demonstrated strong precision and recall values, indicating balanced performance in classification tasks. [1]

In a similar vein, research on educational data mining has demonstrated that Random Forest is capable of accurately categorizing student performance. Random Forest outperformed Naïve Bayes in handling high-dimensional academic datasets, as seen by its around 99.38% accuracy in predicting student grades. [2]

Random Forest has demonstrated excellent predictive performance in healthcare analytics as well. For example, Random Forest outperformed neural network models and other classifiers with an accuracy of 98.58% in a study on machine learning algorithms for stroke prediction.

Random Forest's efficacy has also been shown in cybersecurity applications like intrusion detection systems, where ensemble learning techniques were able to identify hostile activity with high accuracy and dependability.

II(b) Artificial Neural Network related articles in recent Research

This article focusses on neural network and its importance while dealing with big datasets. For instance, the deep learning models (convolution neural networks 'CNN') had given tremendous result and outperformed conventional machine learning algorithms on health misinformation detection with metrics above 98%[6]

A study conducted on heart disease detection, gives a significant accuracy of 82.3% using neural network, over other machine learning models when feature selection techniques were implemented.[3]

II(c) Naïve Bayes related articles in recent Research

Naïve Bayes has demonstrated remarkable efficacy in a variety of classification problems, especially in text mining and sentiment analysis, despite its simplicity. For example, a study comparing Naïve Bayes and Random Forest for predicting student admission found that Naïve Bayes outperformed numerous other models in the dataset with an accuracy of 99.79%.[7]

Economic Sciences
https://economic-sciences.com
ES(2026) | Vol. 22, No. MIC'26 | Pg: 214-222 | ISSN:1505-4683

Naïve Bayes achieved an accuracy of 78.19% in sentiment analysis of social media debates related to mental health stigma, demonstrating its effectiveness in natural language processing tasks requiring large textual datasets.[4]

Naïve Bayes is frequently used in sentiment analysis and text classification applications due to its ease of computation and ability to handle enormous text corpora.

Table 1: Classification in supervised learning

Type	Advantages	Limitations
Random Forest	• Efficiently manages high-dimensional data. • Resistant to ensemble averaging-induced overfitting. • Offers interpretability feature significance scores. • Performs well when surrogate splits are used to fill in missing data.	● Harder to understand than individual decision trees. • Predictive time is slower than with simpler models, such as logistic regression. • It can be difficult to adjust hyperparameters (such as tree depth and number of trees).
Neural Networks	• Excels in handling complicated data, including time series, text, and graphics. • Reduces the need for manual engineering by automatically extracting features. • Cutting-edge results in numerous fields (e.g., computer vision, NLP).	• Large datasets are needed for training. computationally costly (requires GPUs and TPUs). • Because of its black-box nature, interpretability is challenging.
Naive Bayes	• Quick prediction and training (linear complexity). • Performs well when working with high-dimensional data, like text.	• Strong assumption of independence, which is frequently unreal. • Inadequate performance when dealing with complicated data, such pictures.

However, when dealing with huge datasets and intricate feature connections, more sophisticated models like neural networks and ensemble techniques can attain more accuracy.[6]

Comparative research generally shows that no algorithm consistently performs better than others on every dataset. A number of variables, including dataset size, feature relationships,

preprocessing methods, and model optimization tactics, have a significant impact on each algorithm's success.

III. Classification in Supervised Learning

The objective of classification, a supervised learning method in machine learning (ML), is to forecast distinct class labels from input features. Numerous categorization algorithms have been created, each with advantages and disadvantages. This paper examines the uses, benefits, and drawbacks of three well-known classification techniques: Random Forest (RF), Neural Networks (NN), and Naive Bayes (NB).

Table 2: Description of dataset

Name of Attributes	Type of Attribute	Description of Attribute
'Sepal-Length'	Numeric (continuous)	Length of sepal(cm)
'Sepal-width'	Numeric (continuous)	width of sepal(cm)
'petal-Length'	Numeric (continuous)	Length of petal(cm)
'petal-width'	Numeric (continuous)	width of sepal(cm)
Iris (Target Variable)	Discrete(categorical)	Species of flower

Economic Sciences
https://economic-sciences.com
ES(2026) | Vol. 22, No. MIC'26 | Pg: 214-222 | ISSN:1505-4683

Table 3: Classification of Dataset

Target Variable (Class)	**Features (Input Variables)**	**Target (Output Variable)**
iris has 3 categories: *Iris-setosa* *Iris-versicolor* *Iris-virginica* Each class has **50 samples** → perfectly balanced dataset	✓ sepal length ✓ sepal width ✓ petal length ✓ petal width	iris (species)

Table 4: Sample of iris dataset

Sepal Length(cm)	Sepal Width(cm)	Petal Length(cm)	Petal Width(cm)	Iris Species
5.0	3.5	1.4	0.2	Iris-setosa
4.9	3.0	1.4	0.2	Iris-setosa
4.7	3.2	1.3	0.2	Iris-setosa
6.4	3.2	4.5	1.5	Iris-versicolor
6.9	3.1	4.9	1.5	Iris-versicolor
5.5	2.3	4.0	1.3	Iris-versicolor
7.1	3.0	5.9	2.1	Iris-virginica
6.3	2.9	5.6	1.8	Iris-virginica

4. Orange for Classification

Orange is a drag-and-drop visual programming tool for machine learning and data mining that is open-source. In Orange, widgets are grouped according to the tasks they carry out. To create models, test them, and illustrate the results, the various categories include data, transform, model, evaluate, unsupervised, image analytics, text mining, etc.

4. Description of Data-set

Below is a description of the attributes, their types along with their description

Table 2 is a description of the attributes, their types along with their description.

Table 3 is description of classification of Dataset.

Table 4 is an illustration of a sample of iris dataset.

5. Work-Flow

A file widget is used to retrieve the data from the dataset. There are 150 examples of iris setosa, iris virginica, and iris versicolor in the iris dataset. Iris is the target variable (categorical), while the input variables (feature-numerical) are sepal length, sepal width, petal length, and petal width. The three model widgets—RN, NN, and Naïve Baye's—are fed the data after it has been imported into the system. These are used to learn the model based on the following characteristics:

Table 5: Model's and their Attributes

Model	Attributes
Random Forest (RF)	No of Trees(10) – no of attributes considered at each split (5) – Replicable training Growth Control – Limit depth of individual tress (3) – Do Not Split Subset more than (5)
Neural Network (NN)	Neuron in hidden layer (100) Activation (Relu) Solver (Adam) Max. no of iterations (200)
Naïve Bayes	Default attributes

Economic Sciences
https://economic-sciences.com
ES(2026) | Vol. 22, No. MIC'26 | Pg: 214-222 | ISSN:1505-4683

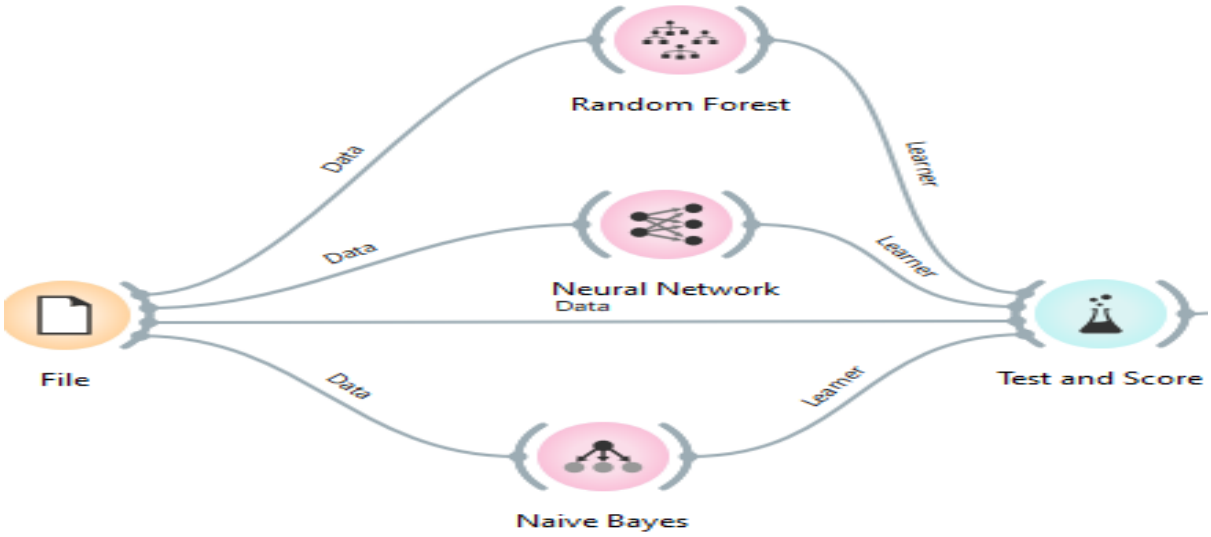


Figure 2: Widget (orange) depicting the flow of data in the model

5. Result

The Test and Score Widget is used to acquire the results. The results are obtained using the Test and Score Widget. During the cross-validation procedure, the data is divided into a fixed number of folds (10). The approach is tested one-fold at a time, with the model being induced from other folds and examples from the held-out fold being categorized. Every fold pass through this once again.

5.1 Result - Using Cross Validation

The following values are provided in the cross validation:

No of Folds: 10 Sampling: Stratified
Repeat Train/Test: set of 10 Training Test Size, which is around 66%
OUTCOME:

Table 5: Result using Cross validation

Model	AUC	CA	F1	Prec	Recall	MCC
Random Forest	0.991	0.947	0.947	0.947	0.947	0.92
Neural Network	0.993	0.947	0.947	0.948	0.947	0.92
Naive Bayes	0.981	0.893	0.893	0.894	0.893	0.84

- AUC: The area beneath the receiver-operating curve is known as the area under ROC.
- CA: The percentage of correctly classified examples is known as classification accuracy.
- F-1 is a weighted harmonic mean of recall and precision
- The percentage of genuine positives among cases categorized as positive, such as the percentage of Iris virginica accurately identified as Iris virginica, is known as precision.
- The percentage of genuine positives among all positive instances in the data, such as the number of sick people among all those with a diagnosis of illness, is known as recall
- Mathew's correlation coefficient, or MCC, accounts for both true and spurious positives and negatives.

5.2 Result: Using Area under ROC

After the data is randomly separated into the training and testing sets in the proper ratio (e.g., 70:30) using random sampling, the procedure is repeated a predefined number of times.

Economic Sciences
https://economic-sciences.com
ES(2026) | Vol. 22, No. MIC'26 | Pg: 214-222 | ISSN:1505-4683

You can select the Target class for classification at the bottom of the widget. When the Target class is set to (Average over classes), methods return weighted average scores for all classes.

Table 6: Comparison of Models based on AUC

	Random Forest	Neural Network	Naïve Bayes
Radom Forest		0.500	0.924
Neural Network	0.500		0.857
Naïve Bayes	0.076	0.143	

A graph that shows how well a model can differentiate between several things is called the Receiver Operating Characteristic (ROC) curve. It displays the frequency with which the model effectively avoids identifying negative items as positive and the frequency with which it accurately identifies positive items. In essence, it shows how effectively the model works on binary classification issues. AUC (Area Under the Curve) is a single metric that describes the overall performance of a binary classification model based on the area under the ROC curve.

Table 7: Sensitivity, FNR and Specificity in comparison

Sensitivity / True Positive Rate / Recall	$Sensitivity = \frac{TP}{TP + TN}$	The percentage of the positive class that was correctly classified is known as sensitivity. percentage of false-negative results
False Negative Rate (FNR)	$FNR = \frac{FN}{TP + FN}$	The percentage of the positive class that the classifier incorrectly classified is shown by FNR.
Specificity / True Negative Rate	$Specificity = \frac{TN}{TN + FP}$	Specificity shows the proportion of the negative class that was correctly classified.

A Receiver Operating Characteristic Curve, or ROC Curve, is a graphical representation that illustrates the trade-off between True Positive Rate and False Positive Rate at various classification levels.

A higher TPR and a lower FNR are desired since we want to correctly classify the positive class.

5.3 Result: Using Confusion Matrix (Predicted vs Actual: Proportion of Predicted)

Based on result, using confusion, the result are obtained (actual vs predicted) based on the three methods: - random forest, Neural network and Naïve Bayes.

iris-setosa dataset are correctly classified in Random Forest and Naïve Baye's. 94% Iris-versicolor are classified correctly using neural network and random forest, whereas 92% iris-verginica are correctly classified in neural network.

Table 8: Result comparison using confusion matrix (Random Forest)

A C T U A L		Iris-Setosa	Iris-Versicolor	Iris-Virginica	∑
	Iris-Setosa	100%	0.0%	0.0%	50
	Iris-Versicolor	0.0%	94.0%	6.0%	50
	Iris-Virginica	0.0%	8.0%	92.0%	50
∑		49	52	49	150

Economic Sciences
https://economic-sciences.com
ES(2026) | Vol. 22, No. MIC'26 | Pg: 214-222 | ISSN:1505-4683

Table 9: Result using Confusion matrix (Neural Network)

		Iris-Setosa	Iris-Versicolor	Iris-Virginica	∑
A C T U A L	Iris-Setosa	98.0%	2.0%	0.0%	50
	Iris-Versicolor	0.0%	94.0%	6.0%	50
	Iris-Virginica	0.0%	8.0%	92.0%	50
∑		49	52	49	150

Table 10: Result using Confusion matrix (Naïve Baye's)

		Iris-Setosa	Iris-Versicolor	Iris-Virginica	∑
A C T U A L	Iris-Setosa	100.0%	0.0%	0.0%	50
	Iris-Versicolor	0.0%	82.0%	18.0%	50
	Iris-Virginica	0.0%	14.0%	86.0%	50
∑		49	52	49	150

5.4 Result: Using Scatter chart

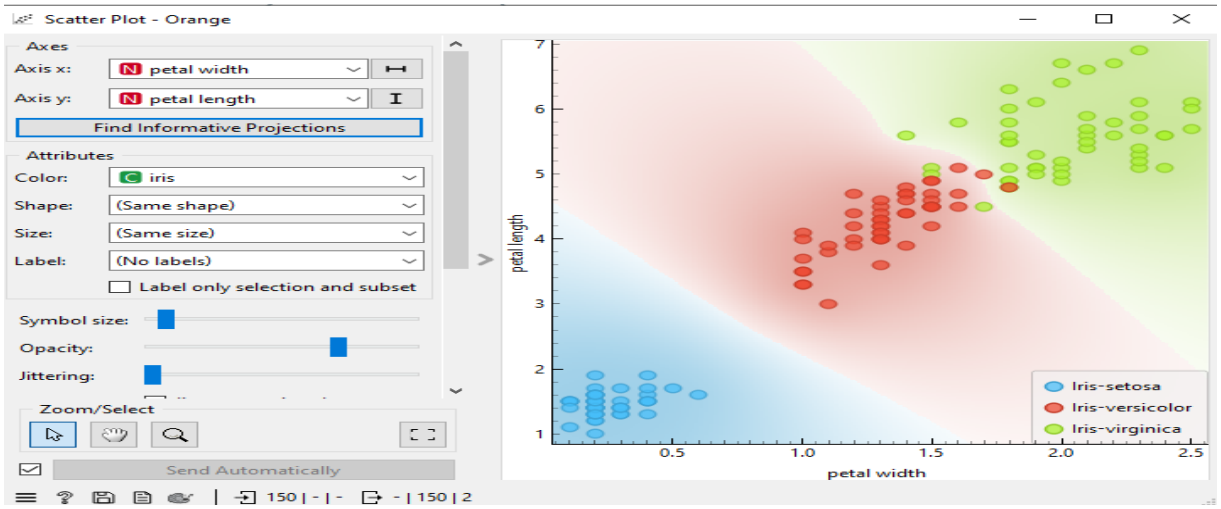
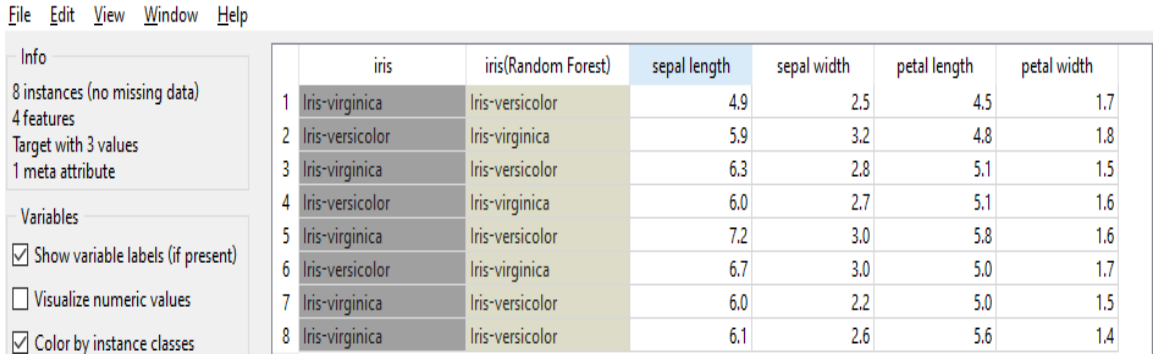


Fig 3: Result using Scatter chart

A 2-dimensional scatter plot is seen using the Scatter Plot widget. The information is shown as a set of points, each of which has a value for the x-axis attribute that indicates its location on the horizontal axis and a value for the y-axis attribute that indicates its location on the vertical axis. The left side of the widget allows you to change the graph's many characteristics, including colour, point size and shape, axis names, maximum point size, and jittering. The scatter plot of the Iris dataset with the class attribute's colouring matching is displayed in the screenshot below.

Economic Sciences
https://economic-sciences.com
ES(2026) | Vol. 22, No. MIC'26 | Pg: 214-222 | ISSN:1505-4683



	iris	iris(Random Forest)	sepal length	sepal width	petal length	petal width
1	Iris-virginica	Iris-versicolor	4.9	2.5	4.5	1.7
2	Iris-versicolor	Iris-virginica	5.9	3.2	4.8	1.8
3	Iris-virginica	Iris-versicolor	6.3	2.8	5.1	1.5
4	Iris-versicolor	Iris-virginica	6.0	2.7	5.1	1.6
5	Iris-virginica	Iris-versicolor	7.2	3.0	5.8	1.6
6	Iris-versicolor	Iris-virginica	6.7	3.0	5.0	1.7
7	Iris-virginica	Iris-versicolor	6.0	2.2	5.0	1.5
8	Iris-virginica	Iris-versicolor	6.1	2.6	5.6	1.4

Fig 4 Data Table – for misclassified (Random Forest)

For finding the incorrectly predicted values in RF (for illustration)
Open Confusion matrix -> Select Misclassified -> Open Data-Table Widget

5.5 Data Table

One or more datasets are fed into the Data Table widget, which displays them as a spreadsheet. It is possible to order data instances according on attribute values. Additionally, the widget allows for manual data instance selection. The dataset's name, which is often the input data file.

Only 8 out of 150 instances are misclassified, and their values of input attributes are described. Similarly we can choose using NN and Naïve Bayes method, and their misclassified instances.

6. Conclusion

Criteria	Result
Classification Accuracy (CA) and Precision	NN better than RF and Naïve Baye's
No of Mis-classifier Labels	RF – 8; NN – 8; Naïve Bayes – 16 Both RF and NN are equally good as compared to Naïve Bayes method.

For structured data, Random Forest is a robust and flexible classifier. Neural networks are dominant in difficult fields like NLP and vision, but they need a lot of data. For text and tiny datasets, Naive Bayes is quick and effective, but it has trouble with dependencies. Through data gathering and processing, data mining aims to extract meaningful information from data. With the use of statistical calculations and mathematical procedures, data mining software can be used to collect and extract data. By facilitating the extraction of insights and the development of prediction models, Orange aims to optimize the value of your data. No matter their level of experience, it enables all members of your business to take part in data-driven decision-making. Orange's visual style makes it simple to explore and work with datasets, preprocess data, display outcomes, and assess models.

7. References:

1. Hesar, F. F., & Foing, B. (2024). Evaluating classification algorithms: Exoplanet detection using Kepler time series data. arXiv preprint arXiv:2402.15874.
2. D. Kurniasari, R. N. Hidayah, N. Notiragayu, W. Warsono, And R. K. Nisa, "Classification Models For Academic Performance: A Comparative Study of Naïve Bayes and Random Forest Algorithms In Analyzing University Of Lampung Student Grades", J. Tek. Inform. (Jutif), Vol. 5, No. 5, Pp. 1267–1276, Oct. 2024.

3. Ahmad, B., Chen, J., & Chen, H. (2025). Feature selection strategies for optimized heart disease diagnosis using ML and DL models. ArXiv, abs/2503.16577.
4. P. Elisa and A. R. Isnain, "COMPARISON OF RANDOM FOREST, SUPPORT VECTOR MACHINE AND NAIVE BAYES ALGORITHMS TO ANALYZE SENTIMENT TOWARDS MENTAL HEALTH STIGMA", J. Tek. Inform. (JUTIF), vol. 5, no. 1, pp. 321–329, Feb. 2024.
5. Nurhachita, N., & Negara, E. (2021). A comparison between deep learning, naïve bayes and random forest for the application of data mining on the admission of new students. IAES International Journal of Artificial Intelligence (IJ-AI), 10(2), 324-331. doi:http://doi.org/10.11591/ijai.v10.i2.pp324-331.
6. Sikosana, M., Ajao, O., & Maudsley-Barton, S. (2024, September). A comparative study of hybrid models in health misinformation text classification. In Proceedings of the 4th International Workshop on Open Challenges in Online Social Networks (pp. 18-25).
7. "Evaluating Machine Learning Approaches: A Comparative Study of Random Forest and Neural Networks in Grade Classification", ijodas, vol. 6, no. 1, pp. 73–80, Mar. 2025, doi: 10.56705/ijodas.v6i1.240.
8. A Novel Performance Measure for Machine Learning Classifications: Mingxing Gong (28 Feb 2021-International Journal of Managing Information Technology)
9. Shi, T., & Horvath, S. (2006). Unsupervised Learning With Random Forest Predictors. Journal of Computational and Graphical Statistics, 15(1), 118–138. https://doi.org/10.1198/106186006X94072
10. Yogesh Mali, & Tejal Upadhyay. (2023). Fraud Detection in Online Content Mining Relies on the Random Forest Algorithm. SciWaveBulletin, 1(3), 13-20.
11. Tin-Chih Toly Chen, Hsin-Chieh Wu, Min-Chi Chiu,A deep neural network with modified random forest incremental interpretation approach for diagnosing diabetes in smart healthcare, Applied Soft Computing, Volume 152,2024,111183, ISSN 1568-4946,https://doi.org/10.1016/j.asoc.2023.111183.
12. Jackins, V., Vimal, S., Kaliappan, M. et al. AI-based smart prediction of clinical disease using random forest classifier and Naive Bayes. J Supercomput 77, 5198–5219 (2021). https://doi.org/10.1007/s11227-020-03481-x
13. Nurhachita, N., & Negara, E. S. (2021). A comparison between deep learning, naïve Bayes and random forest for the application of data mining on the admission of new students. International Journal of Artificial Intelligence, 10 (2), 324Pause similarity match. -341. Resume similarity match. https://doi.org/10.11591/ijai.v10.i2.pp324-331
14. A comparative study of classification techniques for intrusion Detection. (2013, August 1). IEEE Conference Publication | IEEE Xplore. https://ieeexplore.ieee.org/abstract/document/6724320