Spectrum of Engineering Sciences
ISSN (e) 3007-3138 (p) 3007-312X Volume 3, Issue 6, 2025

AUTOMATED DETECTION OF CELIAC DISEASE USING MACHINE LEARNING ALGORITHMS

Fatima Bukhari[1], Hafiza Iqra Younas[2], Aiza Shabir[3], Muzaffar Hameed[4], Meherwar Fatima[5], Sehrish Raza[6], Khadija Kanwal[*7]

[1,2]*National University of Modern Languages, Multan Campus, Multan, 60000, Pakistan*
[3,5,6,*7]*Institute of Computer Science and Information Technology, The Women University, Multan, 6000, Pakistan*
[4]*Department of Computer Science, Bahauddin Zakariya University, Multan, 6000, Pakistan*

[*7]khadijakanwal.6022@wum.edu.pk

DOI: https://doi.org/10.5281/zenodo.15749919

Keywords
Celiac Disease, Machine Learning, Naïve Bayes, Random Forest, Logistic Regression

Article History
Received on 19 May 2025
Acceptsed on 19 June 2025
Published on 26 June 2025

Corresponding Author: *
Khadija Kanwal

Abstract
The belief that weather influences chronic pain is widespread among patients, yet empirical evidence remains limited due to challenges in collecting large-scale, high-resolution symptom data across varying meteorological conditions. Leveraging the ubiquity of smartphones, the Cloudy with a Chance of Pain study systematically recorded daily pain levels and environmental data from 2,206 individuals over a 15-month period. Statistical analysis revealed modest but statistically significant associations between pain intensity and meteorological variables—particularly relative humidity, barometric pressure, and wind speed—even after controlling for confounders such as mood and physical activity. This research demonstrates the feasibility and scientific value of citizen-science approaches in generating large, ecologically valid datasets. The findings contribute to a deeper understanding of weather-related pain patterns and offer a foundation for predictive tools to support personalized pain management.

INTRODUCTION

Celiac disease primarily affects the small intestine, but it can also impact other systems such as the skeletal structure. Currently, the most reliable method for diagnosing celiac disease is the histological examination of duodenal tissue collected through upper gastrointestinal endoscopy [3]. However, this method is invasive and time-consuming. To enhance the efficiency and accuracy of diagnosis, there is growing interest in using machine learning (ML) algorithms for automated detection. One promising avenue involves analyzing capsule endoscopy images using techniques like Naïve Bayes (NB). Min et al. (2019) provide an overview of machine learning applications in gastrointestinal endoscopy, highlighting its potential to automate diagnostic processes based on visual data [1]. Their review discusses various ML-driven diagnostic systems tested in contexts such as identifying anatomical features in esophagogastroduodenoscopy images, detecting colorectal polyps, recognizing hookworm infections and celiac disease, and analyzing small intestine motility in capsule endoscopy images.

These studies collectively demonstrate how artificial intelligence (AI) can support endoscopy specialists by automating the detection and classification of abnormalities, ultimately improving diagnostic accuracy. A noteworthy contribution comes from

Stoleru et al. (2022), who proposed an ML-based approach for diagnosing celiac disease using video capsule endoscopy images [2]. Their method employs kernel-based techniques to detect specific visual markers of celiac disease, including mucosal atrophy with exposed submucosal vascular patterns, fissures resembling dry land, loss of duodenal folds (partial or complete), and a "submerged" appearance at the Kerckring folds characterized by reduced villi. Their system achieved 94.1% accuracy and a 94% F1 score, showing performance comparable to more complex algorithms. This research supports the feasibility of diagnosing celiac disease with high accuracy through automated image processing using relatively simple ML methods.

Another line of investigation leverages deep learning to analyze histological images from duodenal biopsies. Wei et al. (2019) developed a deep learning model based on residual convolutional neural networks to identify celiac disease in biopsy images [3]. Their model demonstrated 95.3% accuracy in detecting celiac disease, 91.0% in identifying normal tissue, and 89.2% in classifying nonspecific duodenitis. The area under the ROC curve exceeded 0.95 for all categories, indicating excellent diagnostic performance. Gadermayr et al. (2018) reviewed the latest developments in computer-aided diagnosis for celiac disease based on upper endoscopy, emphasizing the promise and challenges of deploying such systems in clinical environments [4].

Further advancement came from Wang et al. (2020), who introduced a recalibration module to enhance ML performance in capsule endoscopy-based celiac disease diagnosis [5]. Their method achieved 95.94% overall accuracy, with 97.20% sensitivity and 95.63% specificity—markedly better than traditional approaches. Overall, current evidence strongly supports the use of machine learning—especially deep learning—in automating the detection of celiac disease. These technologies can accurately interpret endoscopic and histological images, potentially reducing the reliance on invasive procedures and improving both speed and precision in diagnosis.

Literature Review:

Celiac disease is an autoimmune disorder that primarily targets the small intestine, though its effects can extend to other systems. Currently, diagnosis relies on invasive procedures such as endoscopy and histological analysis of duodenal biopsies. While effective, these methods are time-consuming, uncomfortable, and not universally suitable for all patients. Recent advancements in machine learning (ML) have opened new avenues for automating the diagnosis process, offering the potential for faster and less invasive alternatives.

One study proposed an innovative AI-based approach using convolutional filters to analyze capsule endoscopy images. By evaluating the texture of the intestinal wall, the system achieved 94.1% accuracy in detecting celiac disease—demonstrating the feasibility of rapid, automated diagnosis without relying on complex ML architectures [1]. Another comprehensive review outlined various computer-aided diagnostic techniques used in clinical trials, including traditional feature extraction methods and deep learning models such as JRip, Random Forest, ZeroR, and Naïve Bayes [2]. The review highlighted how AI can significantly enhance diagnostic accuracy for celiac disease.

In a separate study, researchers developed an AI-based image analysis platform using biopsy images to classify celiac disease and other small intestinal disorders. This system achieved 98% classification accuracy, further supporting the clinical viability of automated diagnosis [4]. Several other investigations focused specifically on video capsule endoscopy. One deep learning method achieved perfect sensitivity and specificity (100%) when distinguishing between patients with celiac disease and healthy controls [6]. Another study applied fractal analysis and scale-invariant texture descriptors to classify duodenal endoscopy images, achieving state-of-the-art performance in several scenarios [7-9].

Collectively, these studies underscore the effectiveness of machine learning—especially deep learning—in detecting and diagnosing celiac disease. These techniques have demonstrated high accuracy across a variety of medical imaging modalities, including biopsy slides and capsule endoscopy footage. By automatically identifying and classifying pathological features, AI shows promise in enabling more precise, timely diagnoses [10].

A 2022 Scientific Reports article introduced a machine learning-based method for detecting celiac disease from capsule endoscopy images using

enhanced filtering techniques. This method achieved 94.1% accuracy and a 94% F1 score, confirming its robustness [11]. A 2020 study in Computer Methods and Programs in Biomedicine employed deep residual learning and feature extraction using ResNet50 and Inception-v3 models, integrated with a novel recalibration module. The model achieved 95.94% accuracy, 97.20% sensitivity, and 95.63% specificity in diagnosing celiac disease [12].

An earlier 2018 paper in Computers in Biology and Medicine reviewed recent progress in upper endoscopy-based computer-aided diagnosis for celiac disease, discussing the strengths, limitations, and future directions of this emerging field [13]. A 2021 publication in Translational Gastroenterology and Hepatology focused on software advancements in capsule endoscopy, emphasizing the transformative potential of high-accuracy ML algorithms for aiding real-time diagnostic decisions [14].

Another study, published in Computer Methods and Programs in Biomedicine in 2021, developed an ML framework to automate biopsy image analysis for identifying villous atrophy based on modified Marsh scores. Using Steerable Pyramid Transform to extract subband features, the model reached 88.89% accuracy in two-class classification of villous abnormalities [15]. Lastly, a 2019 study in Journal of Pathology Informatics used a deep residual convolutional neural network to detect celiac disease in duodenal biopsy slides, achieving 95.3% accuracy [16,17].

Finally, a study published in Current Medical Imaging assessed the severity of celiac disease using video capsule endoscopy by comparing machine learning algorithms to the evaluation methods of experienced human readers. The results showed a strong agreement between the predictions made by machine learning models and the assessments of human experts across the small intestine [18].

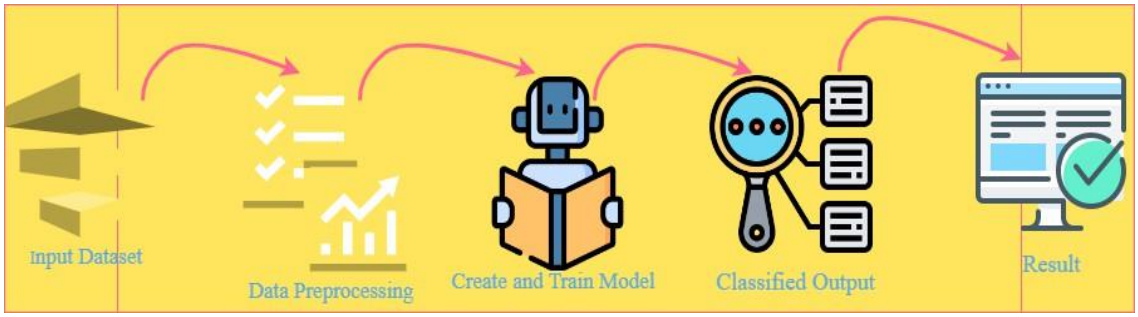


Figure 1 Proposed Methodology

Methodology:

Figure 1 illustrates various research efforts and enabling methodologies currently employed for binary classification tasks. Accurately diagnosing cancer through histopathological images remains a significant challenge for pathologists, particularly when time and precision are critical. Our proposed study seeks to overcome the limitations of prior work by introducing a dependable and robust method for classifying invasive ductal carcinoma (IDC). The approach leverages an ensemble of classification models—including Random Forest, JRip, Naïve Bayes, and ZeroR—to improve diagnostic accuracy. In particular, the Random Forest model benefits from transfer learning by adapting a pre-trained architecture to the specific task.

To enhance performance and reliability, our system integrates multiple classifiers, and the experimental setup was implemented using the Weka platform. While celiac disease classification serves as a test case, the method is broadly applicable to similar binary classification challenges in histopathology. The proposed model is designed to streamline diagnostic workflows, reduce human error, and provide a scalable automated solution. A publicly available dataset was utilized for validation, and the detailed steps of the proposed technique are outlined in Figure 1. The remainder of this section provides a step-by-step breakdown of our methodological framework.

Dataset:

The dataset utilized in this study was provided by the Biotechnology Department at Wageningen University & Research. Its primary aim is to facilitate the prediction of celiac disease based on a range of clinically relevant diagnostic features. As depicted in Figure 2, the dataset comprises 2,206 records and 15 variables, including both demographic and medical attributes. Key features include patient age, gender, and diabetes type, among others, while the target variable indicates the presence or absence of celiac disease. The dataset is well-suited for predictive modeling and risk stratification, offering a rich foundation for evaluating the performance of machine learning classifiers in the context of celiac disease diagnosis.

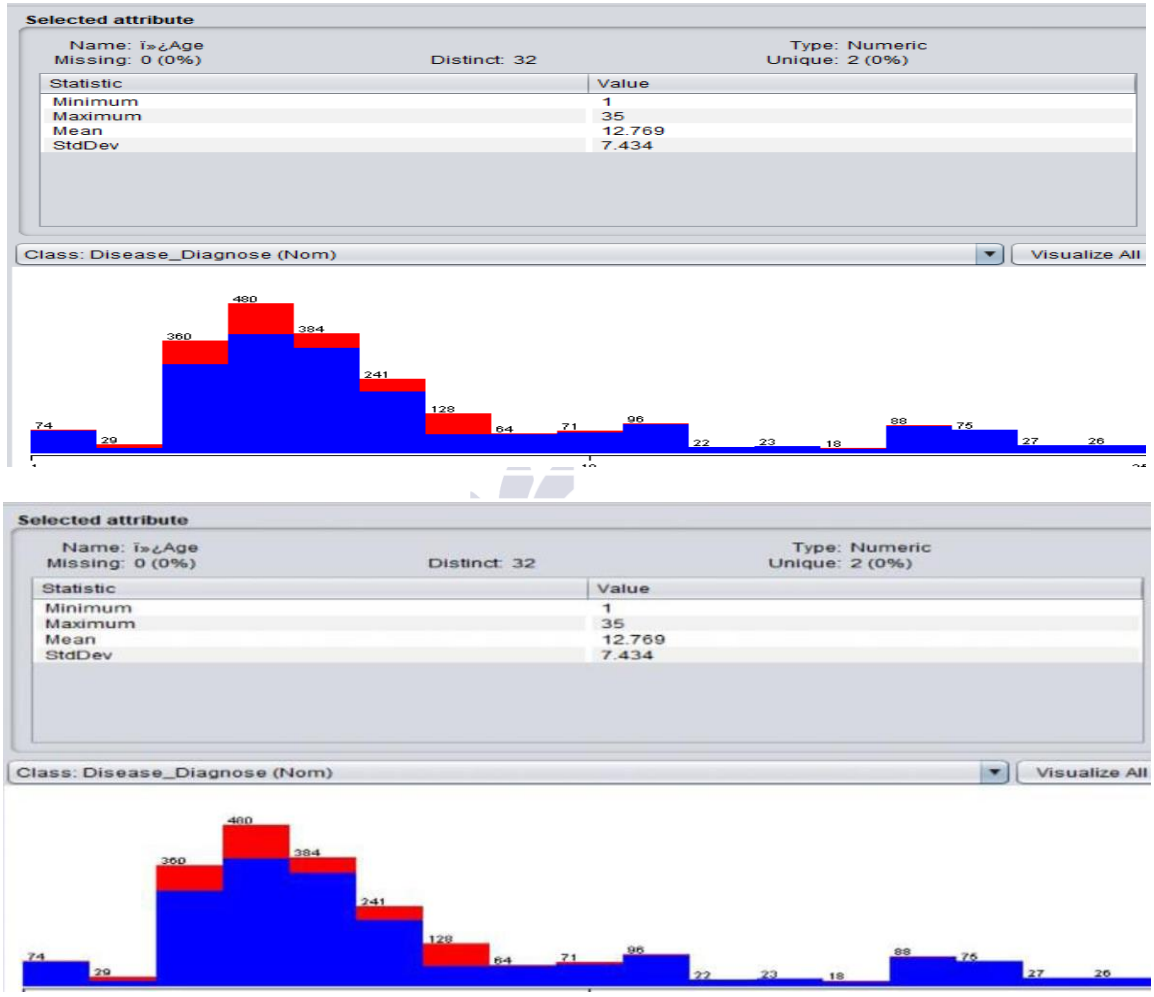
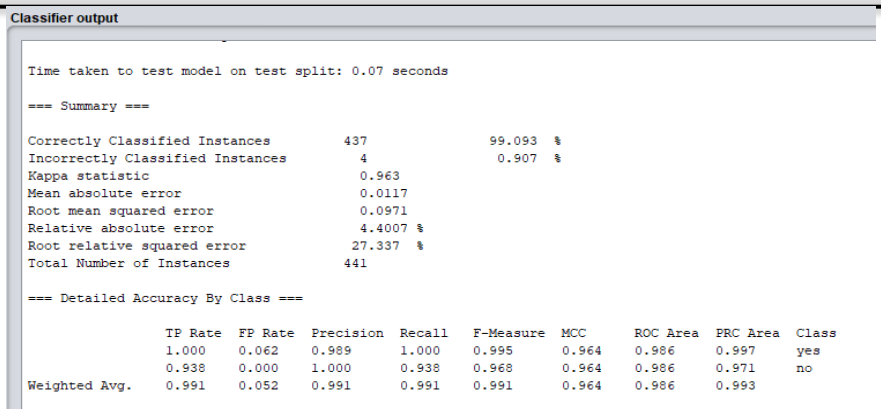


Figure 2 Data Set Statistics

Machine Learning Models:

The proposed method leverages multiple machine learning models [9,10] to detect celiac disease with high accuracy. Specifically, the approach incorporates Naïve Bayes (NB), JRip, Random Forest (RF), and ZeroR classifiers. To ensure optimal performance, each model is fine-tuned through careful parameter optimization, enhancing the reliability and predictive accuracy of the system.



Classifier output

```
Time taken to test model on test split: 0.07 seconds

=== Summary ===

Correctly Classified Instances         437               99.093  %
Incorrectly Classified Instances         4                0.907  %
Kappa statistic                          0.963
Mean absolute error                      0.0117
Root mean squared error                  0.0971
Relative absolute error                  4.4007 %
Root relative squared error             27.337  %
Total Number of Instances              441

=== Detailed Accuracy By Class ===

               TP Rate   FP Rate   Precision   Recall  F-Measure  MCC      ROC Area  PRC Area  Class
               1.000    0.062    0.989      1.000    0.995      0.964    0.986     0.997     yes
               0.938    0.000    1.000      0.938    0.968      0.964    0.986     0.971     no
Weighted Avg.  0.991    0.052    0.991      0.991    0.991      0.964    0.986     0.993
```

Figure 3 Naive Bayes

Naïve Bayes (NB):

The Naïve Bayes (NB) classifier is a probabilistic model based on Bayes' theorem, designed to predict class membership probabilities for a given input. It selects the class with the highest posterior probability as the predicted outcome. Rather than outputting deterministic labels, NB provides a probability distribution across possible classes, making it especially useful for uncertainty-aware classification tasks.

The term "naïve" arises from the simplifying assumption that all input features are conditionally independent given the class label. In other words, it assumes that the presence or absence of a particular feature does not influence the presence or absence of any other feature. Although this assumption rarely holds in real-world scenarios, the classifier often performs surprisingly well, especially in high-dimensional spaces.

Naïve Bayes is widely used in various domains due to its simplicity, efficiency, and interpretability. It is often employed as a baseline model in classification tasks to establish a reference performance level before applying more complex algorithms, as illustrated in Figure 3. Its lightweight computational requirements also make it well-suited for large-scale applications and deployment in distributed or parallel processing environments.

Despite these advantages, Naïve Bayes has certain limitations. Its core assumption of feature independence can lead to suboptimal performance in datasets where features are highly correlated. Additionally, for complex classification problems, it may not match the accuracy of more advanced models, such as deep learning approaches. Nonetheless, Naïve Bayes remains a valuable tool in many machine learning pipelines—particularly where speed, scalability, and simplicity are critical, or in text classification tasks where it often excels (Figures 4 and 5).

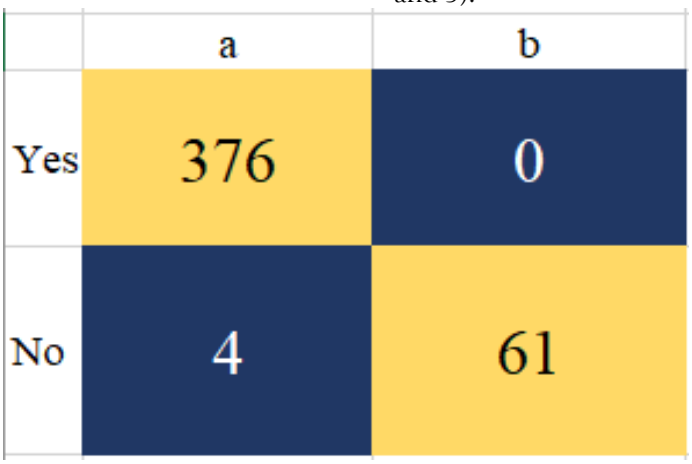


Figure 4 Confusion Matrix

Classifier output

```
Time taken to test model on test split: 0.07 seconds

=== Summary ===

Correctly Classified Instances         438               99.3197 %
Incorrectly Classified Instances         3                0.6803 %
Kappa statistic                          0.9724
Mean absolute error                      0.0093
Root mean squared error                  0.0842
Relative absolute error                  3.4851 %
Root relative squared error             23.7011 %
Total Number of Instances              441

=== Detailed Accuracy By Class ===

               TP Rate  FP Rate  Precision  Recall  F-Measure  MCC      ROC Area  PRC Area  Class
               1.000    0.046    0.992      1.000   0.996      0.973    0.975     0.992     yes
               0.954    0.000    1.000      0.954   0.976      0.973    0.975     0.961     no
Weighted Avg.  0.993    0.039    0.993      0.993   0.993      0.973    0.975     0.987
```

Figure 5 Random Forest

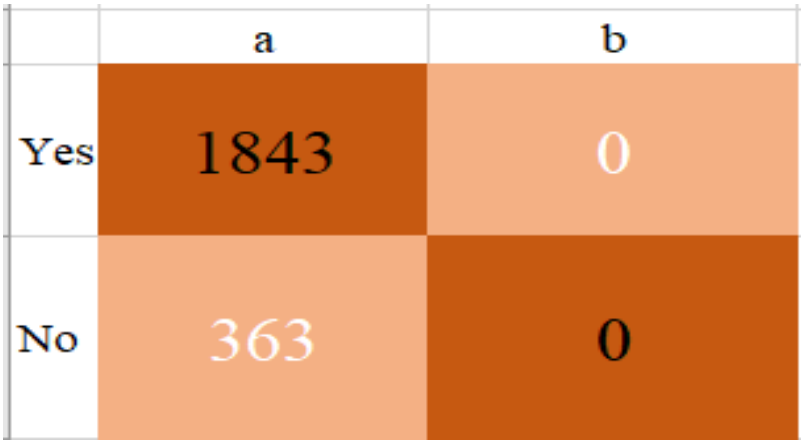
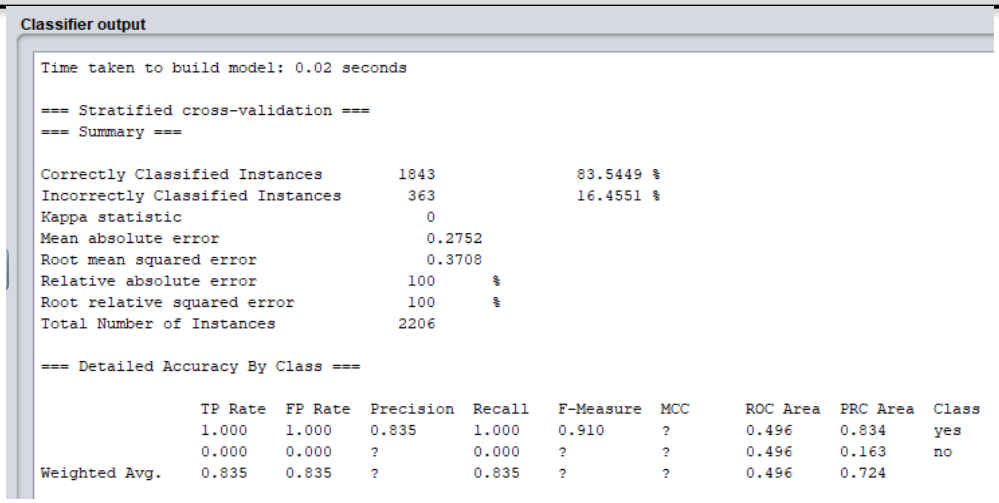


Figure 6 Confusion Matrix

Random Forest (RF):

The Random Forest classifier is a widely used machine learning technique for both classification and regression tasks, as illustrated in Figure 6. It belongs to the family of ensemble learning methods and works by aggregating the predictions of multiple decision trees to improve overall performance. By combining these individual learners, Random Forest significantly reduces the risk of overfitting while enhancing predictive accuracy. Owing to its robustness and versatility, Random Forest has been successfully applied across diverse domains, including natural language processing, healthcare diagnostics, and financial modeling.

ZeroR:



Classifier output

```
Time taken to build model: 0.02 seconds

=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances        1843               83.5449 %
Incorrectly Classified Instances       363               16.4551 %
Kappa statistic                          0
Mean absolute error                      0.2752
Root mean squared error                  0.3708
Relative absolute error                100      %
Root relative squared error            100      %
Total Number of Instances             2206

=== Detailed Accuracy By Class ===

                 TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
                 1.000    1.000    0.835      1.000    0.910      ?        0.496     0.834     yes
                 0.000    0.000    ?          0.000    ?          ?        0.496     0.163     no
Weighted Avg.    0.835    0.835    ?          0.835    ?          ?        0.496     0.724
```

Figure 7 ZeroR

ZeroR, also known as the "zero rule" classifier, is one of the simplest and most fundamental models in machine learning. It operates by predicting the majority class in the dataset, without considering any input features. As such, it does not attempt to uncover patterns or relationships within the data. ZeroR is typically used as a baseline model to benchmark the performance of more sophisticated classification algorithms. Its role is essential in evaluating whether a proposed method provides meaningful predictive power beyond random or majority-class guessing, as illustrated in Figure 7.

ZeroR earns its name by making predictions without applying any rules or utilizing feature data—it relies solely on the class distribution in the training set. Specifically, it identifies the most frequently occurring class in the training data and assigns this class label to every instance in the test data, regardless of its attributes. In doing so, ZeroR completely ignores all feature information and treats every test instance as belonging to the majority class. The resulting confusion matrix for this classifier is shown in Figure 8.

While ZeroR is extremely simple and easy to implement, it is not suitable for real-world applications where meaningful features are present. Instead, its primary role is to serve as a performance benchmark for evaluating more sophisticated classification models. If an advanced machine learning algorithm[23] fails to outperform ZeroR in a given classification task, this could signal underlying issues such as poor dataset quality, inadequate feature engineering, or inappropriate evaluation metrics.

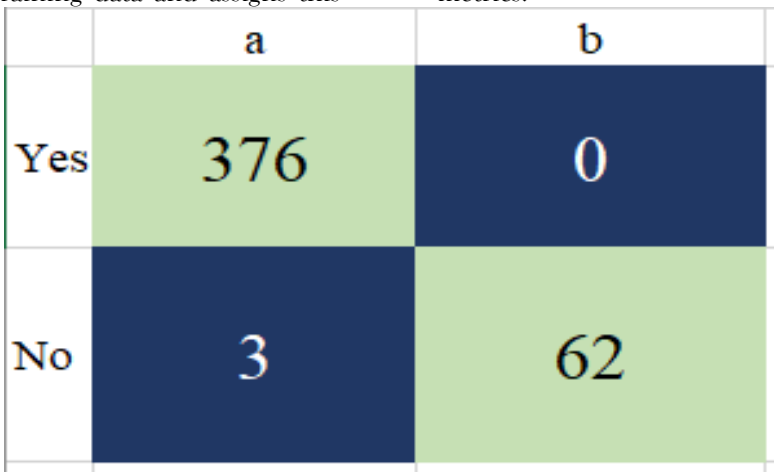


Figure 8 Confusion Matrix

Classifier output

```
Time taken to build model: 0.38 seconds

=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances        2198               99.6374 %
Incorrectly Classified Instances         8                0.3626 %
Kappa statistic                          0.9867
Mean absolute error                      0.0074
Root mean squared error                  0.0604
Relative absolute error                  2.6825 %
Root relative squared error             16.2976 %
Total Number of Instances             2206

=== Detailed Accuracy By Class ===

               TP Rate   FP Rate   Precision   Recall  F-Measure  MCC      ROC Area  PRC Area  Class
               1.000    0.022    0.996      1.000    0.998      0.987    0.981     0.992     yes
               0.978    0.000    1.000      0.978    0.989      0.987    0.981     0.982     no
Weighted Avg.  0.996    0.018    0.996      0.996    0.996      0.987    0.981     0.990
```

Figure 9 JRip

JRip:

JRip, short for "JRip Rules," is a machine learning algorithm used for classification tasks, as illustrated in Figure 9. It is a rule-based classifier that constructs a set of logical rules to predict outcomes based on input data. Known for its clarity and interpretability, JRip is particularly valued in applications where transparency and explainability are important. The algorithm is part of the Weka machine learning suite and is an implementation of the well-established RIPPER (Repeated Incremental Pruning to Produce Error Reduction) algorithm.

JRip works by generating a series of rules from the training dataset. Each rule consists of a class label—indicating the predicted outcome—and a set of conditions based on input feature values. To enhance model generalizability and reduce overfitting, JRip applies a pruning process to simplify the rule set, as shown in Figure 10. This step ensures that only the most effective and generalizable rules are retained. The resulting rules are evaluated based on their predictive accuracy and conciseness, and are ranked accordingly. During classification, JRip sequentially applies the rules to a new instance until a matching rule is found, at which point the associated class label is assigned.

JRip offers several advantages, including ease of use, strong interpretability, and compatibility with both categorical and numerical data. It is particularly well-suited for scenarios where human-readable decision rules are needed to better understand the underlying classification process. This makes JRip an excellent choice for domains such as healthcare, finance, and any field where model transparency is critical.

	a	b
Yes	1843	0
No	8	355

Figure 10 Confusion Matrix

BayesNet:

Bayes network structure of Bayesian classifier which is a probabilistic graphical model that represents a set of variables and their conditional dependencies in fig 12. This can be seen in Weka classifier graph visualizer and confusion matrix as in fig 11.

	a	b
Yes	1843	1
No	12	351

Figure 11 Confusion Matrix

```
Time taken to build model: 0.42 seconds

=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances        2192               99.3654 %
Incorrectly Classified Instances        14                0.6346 %
Kappa statistic                          0.9768
Mean absolute error                      0.0089
Root mean squared error                  0.0795
Relative absolute error                  3.2272 %
Root relative squared error             21.4316 %
Total Number of Instances             2206

=== Detailed Accuracy By Class ===

               TP Rate   FP Rate   Precision   Recall  F-Measure  MCC      ROC Area  PRC Area  Class
                 0.997    0.025     0.995     0.997     0.996      0.977    0.994     0.998     yes
                 0.975    0.003     0.986     0.975     0.981      0.977    0.994     0.983     no
Weighted Avg.    0.994    0.021     0.994     0.994     0.994      0.977    0.994     0.995
```

Figure 12 BayesNet

Logistic Regression:

Logistic regression is a statistical model that is used to predict the probability of a binary outcome based on one or more predictor variables fig 13. The goal of logistic regression is to find the best fitting model that can predict the probability of the outcome variable being 1 given the predictor variables. In logistic regression, the dependent variable is a binary variable that takes the value of 0 or 1. The independent variables can be continuous, categorical, or a combination of both fig 14. The logistic regression model uses a logistic function to model the relationship between the independent variables and the probability of the dependent variable being used [14].

	a	b
Yes	1838	5
No	9	354

Figure 14 Confusion Matrix

Random Tree:

Random Tree is a supervised learning algorithm that belongs to the family of ensemble methods, as depicted in Figure 15. Based on the principles of bagging (bootstrap aggregating), it constructs a large ensemble of decision trees by training each one on a randomly sampled subset of the data. While traditional decision trees determine the best split at each node using the full feature set, Random Trees introduce an element of randomness by selecting the optimal split from a randomly chosen subset of features at each node. This approach promotes model diversity and reduces the risk of overfitting.

The concept of Random Trees was pioneered by Adele Cutler and Leo Breiman, and it forms the foundation of the widely used Random Forest

algorithm. Random Trees can be applied to both classification and regression problems. In classification tasks, the input feature vector is passed through each tree in the ensemble, and the final class label is determined by majority voting. In regression tasks, the prediction is computed as the average output of all trees in the ensemble.

Random Trees effectively merge ideas from Random Forests and model trees—an advanced type of decision tree where each leaf node contains a linear regression model tailored to the local data subspace. The strength of Random Trees lies in their dual-randomization strategy: first, by using bootstrapped subsets of training data for each tree, and second, by selecting a random subset of features to determine the best split at each node. These two mechanisms enhance the diversity among individual trees, significantly improving performance over single-tree models. This technique has shown strong applicability in classification problems, offering robustness and scalability across datasets with high dimensionality. The integration of Random Forests and model tree principles has led to the development of random model trees, which represent a hybrid approach combining structural simplicity with predictive power. A confusion matrix illustrating the performance of the Random Tree classifier is presented in Figure 16.

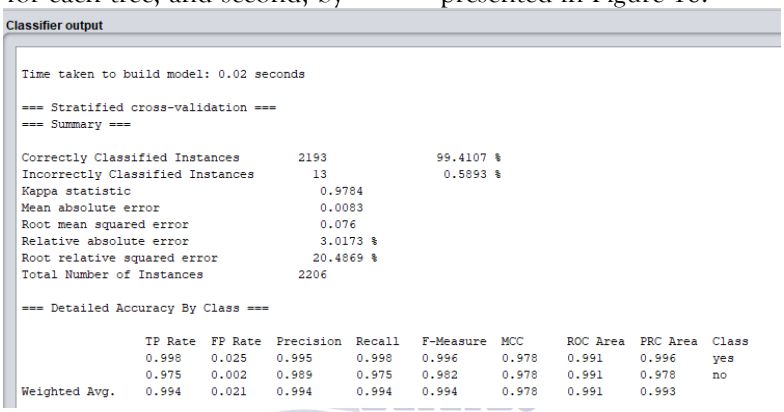


Figure 15 Random Forest

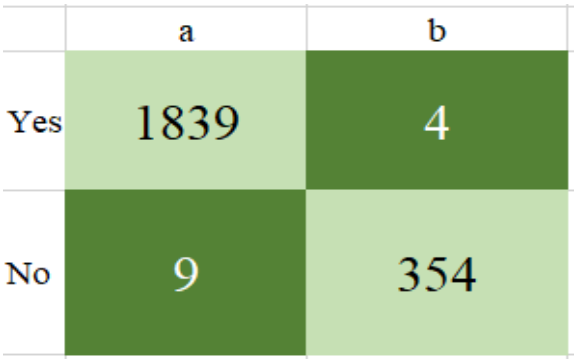


Figure 16 Confusion Matrix

Table 1: Classification of Classifiers

Classifier	Time Complexity	Precision	Recall
Naïve Bayes	0.07 sec	0.991	0.991
Random Forest	0.07 sec	0.993	0.993
ZeroR	0.02 sec	0.982	0.835
JRip	0.38 sec	0.996	0.996
BayesNet	0.17 sec	0.994	0.994
Logistic	0.42 sec	0.994	0.994
RandomTree	0.02 sec	0.994	0.994

Figure 17 Precision

Conclusion:

The automated detection of celiac disease using machine learning algorithms holds significant potential as a non-invasive, efficient, and accurate diagnostic approach. Numerous studies have reported high performance—demonstrating strong accuracy, sensitivity, specificity, and agreement with expert assessments—based on various types of medical imaging, as illustrated in Figures 17–20. These advancements offer promising avenues to enhance both the diagnostic rate and clinical efficiency in identifying celiac disease. Despite these encouraging outcomes, further validation on larger, more diverse datasets is essential. Additionally, continued research is required to optimize algorithmic performance and ensure seamless integration into real-world clinical settings. With rigorous validation and refinement, machine learning-based diagnostic tools could play a transformative role in the early and reliable detection of celiac disease.

REFERENCES:

Min JK, Kwak MS, Cha JM. Overview of Deep Learning in Gastrointestinal Endoscopy. Gut Liver. 2019 Jan;13(4):388- 393.

Stoleru CA, Dulf EH, Ciobanu L. Automated detection of celiac disease using Machine Learning Algorithms. Sci Rep. 2022 Mar 8;12(1):4071.

Wei JW, Jackson CR, Ren B, Suriawinata AA, Hassanpour S. Automated Detection of Celiac Disease on Duodenal Biopsy Slides: A Deep Learning Approach. J Pathol Inform. 2019 Mar 8;10:7. Gadermayr M,Wimmer G,Kogler H,Vécsei A,

Merhof D,Uhl A.Comput Methods Programs Biomed.Automated classification of celiac disease during upper endoscopy: Status quo and quo vadis.Comput Biol Med.November (2018)102:221-226

Wang X,Qian H,Ciaccio EJ,Lewis SK,Bhagat G, Green PH,Xu S,Huang L,Gao R,Liu Y.Celiac disease diagnosis from videocapsule endoscopy images with residual learning and deep feature extraction.Comput Methods Programs Biomed.April (2020)187:105236

Stoleru CA, Dulf EH, Ciobanu L. Automated detection of celiac disease using Machine Learning Algorithms. Sci Rep. 2022;12(1):4071.

Molder A, Balaban DV, Jinga M, Molder CC. Current Evidence on Computer-Aided Diagnosis of Celiac Disease: Systematic Review. Front Pharmacol. 2020;11:341.

Hu, J. L., Li, Y., & Chew, J. C. (2025). Industry 5.0 and Human-Centered Energy System: A Comprehensive Review with Socio-Economic Viewpoints. Energies, 18(9), 2345.

Janjua, J. I., Zulfiqar, S., Khan, T. A., & Ramay, S. A. (2023, November). Activation Function Conundrums in the Modern Machine Learning Paradigm. In 2023 International Conference on Computer and Applications (ICCA) (pp. 1-8). IEEE.

Alqarafi, A., Batool, H., Abbas, T., Janjua, J. I., Ramay, S. A., & Ahmed, M. (2024, December). Estimating Uncertainty in Deep Learning Methods and Applications. In 2024 International Conference on Computer and Applications (ICCA) (pp. 1-6). IEEE.

Phillips F, Beg S.Transl Gastroenterol Hepatol.Video capsule endoscopy: pushing the boundaries with software technology.Transl Gastroenterol Hepatol.Jan 5;6:17.

Syed S,Ehsan L ,Shrivastava A,Sengupta S,Khan M,Kowsari K,Guleria S,Sali R,Kant K,Kang SJ,Sadiq K,Iqbal NT.Cheng L,Moskaluk CA.Kelly P.Amadi BC.Asad Ali S.Moore SR.Brown DE.Artificial Intelligence-based Analytics for Diagnosis of Small Bowel Enteropathies and Black Box Feature Detection.J Pediatr Gastroenterol Nutr.2021 Jun 1;72(6):833- 841.

Wei JW,Wei JW,Jackson CR,Ren B,Suriawinata AA,Hassanpour S.Automated Detection of Celiac Disease on Duodenal Biopsy Slides: A Deep Learning Approach.J Pathol Inform.2019 Mar 8;10:7.

Wang X,Qian H,Ciaccio EJ,Lewis SK,Bhagat G,Green PH,Xu S,Huang L,Gao R,Liu Y.Celiac disease diagnosis from videocapsule endoscopy images with residual learning and deep feature extraction.Comput Methods Programs Biomed.2020 Apr;187:105236.

Koh JEW,De Michele S,Sudarshan VK,Jahmunah V,Ciaccio EJ,Ooi CP,Gururajan R,Gururajan R,Oh SL,Lewis SK,Green PH,Bhagat G,Acharya UR.Automated interpretation of biopsy images for the detection of celiac disease using a machine learning approach.Comput Methods Programs Biomed.2021 May;203:106010.

TenórioJM,Hummel AD,Cohrs FM,Sdepanian VL,Pisa IT,Fátima Marin H.Artificial intelligence techniques applied to the development of a decision-support system for diagnosing celiac disease.Int J Med Inform.2011 Nov;80(11):793-802.

Zhou T,Han G,Li BN,Lin Z,Ciaccio EJ,Green PH,Qin J.Quantitative analysis of patients with celiac disease by video capsule endoscopy: A deep learning method.Med Image Anal.2017 Jun 1;17(4):458-74.

Hegenbart S,Uhl A,Vécsei A.Systematic assessment of performance prediction techniques in medical image classification: a case study on celiac disease.Inf Process Med Imaging.2011;22:498-509.

Hegenbart S,Uhl A,Vécsei A,Wimmer G.Scale invariant texture descriptors for classifying celiac disease.Med Image Comput Comput Assist Interv.2014;17(Pt 2):454-61.

Kwitt R,Hegenbart S,Rasiwasia N,Vécsei A,Uhl A.Do we need annotation experts?A case study in celiac disease classification.Med Image Comput Comput Assist Interv.2014;17(Pt 2):454-61.

Kalmegh, S. (2015). Analysis of weka data mining algorithm reptree, simple cart and randomtree for classification of indian news. International Journal of Innovative Science, Engineering & Technology, 2(2), 438-446.

Desai, A., & Sunil, R. (2012). Analysis of machine learning algorithms using WEKA. International Journal of Computer Applications, 975, 888

Janjua, J. I., Irfan, M., Abbas, T., Ihsan, A., & Ali, B. (2024, December). Enhancing Contextual Understanding in Chatbots and NLP. In 2024 International Conference on TVET Excellence & Development (ICTeD) (pp. 244-249). IEEE.