PROCEEDINGS OF
INTERNATIONAL CONFERENCE ON INTELLIGENT SYSTEMS AND NEW APPLICATIONS
https://proceedings.icisna.org/
2nd International Conference on Intelligent Systems and New Applications (ICISNA'24), Liverpool, April 26-28, 2024.

DETECTION OF MACHINE FAILURES WITH MACHINE LEARNING METHODS

Ilkay CINAR[1], Yavuz Selim TASPINAR[2]

[1] *Department of Computer Engineering, Selcuk University, 42250 Selcuklu, Konya, Türkiye*
ilkay.cinar@selcuk.edu.tr, ORCID: 0000-0003-0611-3316

[2] *Doganhisar Vocational School, Selcuk University, 42930 Doganhisar, Konya, Türkiye*
ytaspinar@selcuk.edu.tr, ORCID: 0000-0002-7278-4241

***Abstract*— This study aims to evaluate the effectiveness of classification algorithms such as Naive Bayes (NB), k-Nearest Neighbours (kNN) and Artificial Neural Networks (ANN) for machine fault detection and investigates the importance of feature selection. The dataset is analysed using cross-validation and the performance of the algorithms is evaluated in terms of AUC, accuracy, F1 Score, Precision and Sensitivity. Naive Bayes and ANN models have the highest AUC values and achieved 99.9% accuracy. kNN model has a lower AUC value than the others (76.0%), but has an accuracy of 97.2%. The feature selection analysis revealed that certain features such as HDF, OSF and PWF contribute significantly to the classification performance. These features have an important role to improve the effectiveness of classification algorithms in detecting faults. These results emphasize the effectiveness of algorithms and the importance of features in machine fault detection and contribute to the development of more reliable and efficient fault detection systems in industrial systems.**

***Keywords*— Artificial intelligence, Detection, Feature selection, Machine learning, Machine failure**

I. INTRODUCTION

Nowadays, the reliability and efficiency of industrial machinery is a major concern in many industries. Especially in production processes, reducing downtime or downtime is critical to reducing costs and increasing productivity. Unexpected situations caused by machine failures can lead to loss of production, occupational safety risks and increased costs. Therefore, the early detection and prevention of machine failures is of vital importance for the sustainability and competitive advantage of industrial enterprises. In this context, data-driven approaches for predicting and preventing machine failures play an important role in the development of preventive maintenance strategies [1].

The aim of this study is to create and analyse a machine learning model fed with fault data as a step towards the development of fault prediction and preventive maintenance strategies. In this context, motor machine failure conditions are determined by examining Torque [Nm], Tool Wear [min], Rotational Speed [rpm], Power Failure (PWF), Tool Wear Failure (TWF), Heat Dissipation Failure (HDF), Overstrain Failure (OSF) along with engine failure conditions. In addition, the most effective features on the above-mentioned features that cause machine failures are determined by various feature selection models and their analysis is performed.

This study consists of Related Works section where previous researches are reviewed, Materials and Methods section where detailed explanation of data, methods and analysis techniques used in the study are included, Experimental Results section where the main findings of the study are presented and finally Discussion and Conclusion section where the findings are analysed and the study is summarized.

II. RELATED WORKS

This section contains a review of previous research on the topic. Summary information about how similar studies deal with similar problems, what methods they use, and what findings they reach is given below.

Abu-Samah and colleagues present a Bayesian-based methodology for learning and associating failures with potential fault occurrences. The data set used includes factors such as product process waiting time, equipment capacity (Cm), overall equipment efficiency (OEE), customer satisfactory product delivery time (OEE Time), total working time (empty and full time), productive time. The highest accuracy obtained in the study was achieved by the Bayesian Network (BN) with a classification accuracy of 97.2% [2].

Janssen et al. proposed a feature learning model for condition monitoring of rotating machines using tremor analysis. Using convolutional neural networks (CNN), fault detection features are learned from the data. In this study, a total of 40 test runs were conducted, five bearings were tested for each of them. They carried out each test using accelerometers to capture vibration data in the x and y planes during the last 10 minutes of the one-hour running time. They have achieved 93.61% classification accuracy with the CNN algorithm [3].

In their study, Goswami and Roy used three different classification algorithms: Decision Tree, K-Nearest Neighbour, Support Vector Machine for the classification of 11 different faults in the synthetic dataset for 100-kilometer-long power transmission lines. According to the results they obtained, they achieved the best performance with 91.6% test accuracy with the SVM algorithm [4].

Chen and colleagues propose an error diagnosis approach that integrates CNN and an Extreme-Learning Machine (ELM). They have collected 2100 pieces of data including a gearbox dataset and an engine bearing dataset. The method integrated with CNN on the data achieved 99.83% classification accuracy [5].

Khalil et al. have proposed a method for early failure prediction in circuits. The method is based on detecting faults using Fast Fourier Transform (FFT), Principal Component Analysis (PCA) and CNN. The data set includes voltage, current, temperature, noise and delay values. The data were used for training using Tensorflow. Using FFT, PCA, CNN, it provides a fault prediction accuracy of 98.93% and 98.91% for comparator and amplifier circuits, respectively [6].

Orrù and colleagues used machine learning algorithms for early failure prediction of a centrifugal pump in the oil and gas industry. This study is based on real-life historical data obtained from the process and equipment sensors installed on the machine. Support Vector Machine (SVM) and Multilayer Sensor (MLP) are used as machine learning algorithms. MLP achieved 98.2% and SVM achieved 98.1% classification accuracy [7].

When the studies found in the literature are evaluated, various researchers have used various machine learning methods to predict the failures of machines in different industrial areas. It is seen that among these methods there are algorithms such as Bayesian Network, Convolutional Neural Networks, Decision Tree, K-Nearest Neighbour, Support Vector Machine and Multilayer Sensor. It was not possible to find any studies performed using the data set used in this study.

III. Materials and Methods

This section contains a detailed description of the data set used in the study, the research method and the analysis techniques used.

A. *Dataset*

The data set used in this study is called "Machine Failure Prediction Cleaned Dataset" [8]. This data set has been specially created for the purpose of predicting machine failures. It contains the final version of a different dataset containing the failure states of various vehicles after clearing outliers and feature selection. The data set consists of 9515 rows belonging to 7 features in total. The property information in the data is as follows: Torque [Nm], Tool Wear [min], Rotational Speed [rpm], Power Failure (PWF), Tool Wear Failure (TWF), Heat Dissipation Failure (HDF), Overstrain Failure (OSF). The data set consists of 2 classes: Durable and Failure. A detailed description of the features in the data set is presented in Table 1.

TABLE I

Detailed description of the features in the data set.

Features	Description
Rotational speed [rpm]	Calculated from a power of 2860 W, overlaid with a normally distributed noise.
Torque [Nm]	Torque values are normally distributed around 40 Nm with a Ïƒ = 10 Nm and no negative values.
Tool wear [min]	The quality variants H/M/L contribute 5/3/2 minutes of tool wear, respectively, to the utilized tool during the process.
TWF	The tool will be replaced of fail at a randomly selected tool wear time between 200 ~ 240 mins.
HDF	Process failure occurs due to inadequate heat dissipation when the temperature difference between the air and the process falls below 8.6 K, and the rotational speed is lower than 1380 rpm.
PWF	The process fails if the power required for the process, calculated as the product of torque and rotational speed (in rad/s), falls below 3500 W or exceeds 9000 W.
OSF	The process fails due to overstrain if the product of tool wear and torque exceeds 11,000 min Nm for the L product variant (12,000 for M, 13,000 for H).
Machine failure	Whether the machine has failed in this particular data point for any of the following failure modes are true.

B. *Naive Bayes (NB)*

The Naive Bayesian classifier is a classification algorithm based on the Bayes theorem. The class uses a strong assumption of independence when making predictions. It is usually effective with high-dimensional inputs in supervised learning processes. In short, the Naive Bayesian classifier assigns objects to certain classes using a simple probabilistic model [9, 10].

C. *K-Nearest Neighbour (kNN)*

The kNN algorithm is a classification method. This method is used to classify a new sample based on training data. The steps of the algorithm are as follows:

Step 1. A new example is given.

Step 2. With each example in the training data, it is compared with the attributes of the new example and the distance is calculated according to these attributes.

Step 3. The nearest neighbouring samples are determined.

Step 4. The new sample is classified using the classification of the nearest neighbouring samples.

The advantages of the kNN algorithm include its simplicity, ease of implementation and the fact that no parameter tuning is required during the training phase [11].

D. *Artificial Neural Network (ANN)*

ANN is an artificial intelligence model inspired by the neural networks of the human brain and is used to solve nonlinear and complex problems. ANNs can be of various types with different architectures consisting of an input layer, one or more hidden layers and an output layer [12]. ANN can also be used in signal processing, processing audio, video and other signals.

E. *Confusion Matrix*

It is a widely used tool to evaluate the performance of the model in classification problems. Confusion matrix allows the model's predictions to be compared with the actual classes and allows this comparison to be analysed with different metrics [13]. The representation of the confusion matrix is shown in Fig. 1.

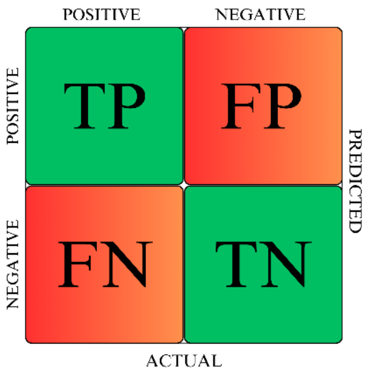


Fig. 1. Confusion matrix

This matrix divides the model's true and false predictions into four categories: True Positives (TP), False Positives (FP), True Negatives (TN), and False Negatives (FN). TP represents the number of instances where the model predicts true positives, FP represents false positive predictions, TN represents true negative predictions, and FN represents false negative predictions. These four categories form the basis for evaluating the classification performance of the model with various metrics [14].

F. *Cross Validation*

Cross validation is a method used to assess the generalizability of a machine learning model. This method divides the data set into different parts and evaluates how the model performs on each part. Basically, the dataset is divided into a certain number of parts and each part is used as a test set while the other parts are used as a training set. This process is repeated until every part is used as a test set. As a result, the performance of the model is measured for each part and these performance measures are combined to obtain an overall performance measure [14, 15]. In this study, 10-fold cross validation is used.

G. *Performance Metrics*

Performance metrics are measurements used to evaluate the performance of a machine learning model or a classifier or regression model. These metrics are used to understand how well the model works and evaluate the accuracy, precision, sensitivity, specificity and other performance characteristics of the model [16, 17]. The formula for performance metrics is shown in Table 2 [18].

TABLE II
PERFORMANCE METRICS FORMULA

Measure	Formula	
Accuracy	(TP+TN)/(TP+TN+FP+FN)	(1)
Precision	TP/(TP+FP)	(2)
Recall	TP/(TP+FN)	(3)
F1-Score	2*(Precision*Recall)/(Precision+Recall)	(4)

Accuracy: The ratio of correctly predicted samples to the total number of samples. It is generally a good performance measure for balanced classes. However, it can be misleading in case of class imbalance [19, 20].

Precision: The ratio of positively predicted instances to actually positive instances [19, 20].

Recall: The ratio of true positives (TP) to all positive samples [19, 20].

F1 Score: The harmonic mean of precision and sensitivity. It is preferred over accuracy in cases with unbalanced classes [19, 20].

Receiver Operating Characteristic Curve (ROC): It is a graphical method used to evaluate the performance of the classification model. This curve shows the relationship between recall and specificity. An ideal classifier is expected to be closer to the upper left corner of the ROC curve [21].

Area Under the Curve (AUC): It refers to the area under the ROC curve. AUC is a measure used to summarize classifier performance. The AUC value usually ranges between 0 and 1. The better a classifier is, the higher the AUC value will be. The ROC curve and AUC are widely used to evaluate model performance, especially in imbalanced classification problems [21].

H. *Feature Selection Methods*

Feature Selection Methods are used in the feature selection process to help identify the most appropriate subset of features from the datasets, and are also used to rank the importance of features in the dataset and select the most informative ones [22]. Methods such as Gain Ratio, Gini, Chi-Square are widely used metrics to evaluate the classification performance of features. The metrics used in this study are Gain Ratio, Gini, Chi-Square and Fast Correlation-Based Feature.

1) Gain Ratio: Gain Ratio is a measure of information gain used in data mining and is an important criterion for feature selection in algorithms such as decision trees. Gain Ratio measures how much information a feature can gain in a given classification task. Higher Gain Ratio values mean better feature selection [22, 23]. Gain ratio formula is given in Equation 5.

$$Gain\ Ratio = \frac{Information\ Gain}{Split\ Information} \quad (5)$$

In equation 5, information Gain measures the information gain when using a particular feature. Split Information is a measure of how data is split with a given feature.

2) Gini: Gini is a measure of the homogeneity or irregularity of a data set. It is widely used in fields such as statistics and economics. In particular, it is an important metric for assessing the quality of splitting nodes in classification algorithms such as decision trees. Gini is the sum of the probabilities that two randomly selected items in a given dataset belong to different classes. A small Gini indicates that the dataset is more homogeneous or more organized, while a large Gini indicates that the dataset is more heterogeneous or more complex [24]. If there are K different classes in a dataset, the Gini index is calculated as in Equation 6:

$$Gini = 1 - \sum_{i=1}^{K} p_i^2 \quad (6)$$

In Equation 6, pi is the probability that instances in the dataset belong to class *i.*

Algorithms such as decision trees use the Gini index as a criterion for splitting a node. The smaller the Gini index, the more homogeneous or pure the split. Therefore, splits with a smaller Gini index are considered better splits.

3) Chi-Square (X^2): Chi-square is a test and measure used in statistical analysis to quantify the relationship between variables or to determine how observed data differ from expected values. The chi-square test is often used in the analysis of categorical data and can be used to assess the significance of the relationship between two categorical variables. The chi-square test measures the difference between observed frequencies and expected frequencies. If there is a significant difference between the observed and expected frequencies, it can be considered that there is a relationship or a change between these variables [22, 25].

4) Fast Correlation-Based Feature (FCBF): FCBF is an algorithm for feature selection. Feature selection is used to improve the performance of a machine learning model and filter out features that contain redundant or unnecessary information. The FCBF algorithm evaluates the relationship between features and selects the most important features based on these relationships [26].

İ. *Experimental Setup*

In this section, the parameters of the machine learning algorithms used in the study are discussed. Experiments were conducted using different parameters for ANN and kNN. The most successful results obtained from the algorithms as a result of classification were achieved using the parameters given in Table 3.

TABLE III
TRAINING OPTIONS

	Parameters	**Values**
kNN	Number of neighbours	5
	Metric	Euclidean
	Weight	Uniform
ANN	Neurons in hidden layers	100
	Activation	ReLu
	Solver	Adam
	Regularization	0.0001
	Maximal number of iterations	200

IV. EXPERIMENTAL RESULTS

Within the scope of the study, NB, kNN and ANN algorithms were used to detect machine faults and classification processes were performed. The results obtained are thoroughly analysed using various evaluation tools such as confusion matrix, ROC curve and performance metrics. This comprehensive analysis aims to provide a detailed evaluation and understanding of the classification performance of the models.

The dataset cross-validation method was used to analyse the effectiveness of the classification algorithms. In the cross-validation method, the dataset is divided into 10 parts and each part of the dataset is run 10 times as a test set in each iteration. The results were then averaged to obtain the average performance metrics of the models. The performance metrics obtained from the algorithms are given in Table 4.

TABLE IV
PERFORMANCE METRICS FOR ALL ALGORITHMS

	AUC	**Accuracy**	**F1 Score**	**Precision**	**Recall**
ANN	0.980	0.999	0.999	0.999	0.999
NB	0.989	0.999	0.999	0.999	0.999
kNN	0.760	0.972	0.963	0.963	0.972

Table 4 shows that the highest AUC value is obtained from the Naive Bayes model (0.989), followed by the ANN model (0.980). kNN model has a lower AUC value (0.760). Looking at the Accuracy results, it can be said that the ANN and NB algorithms achieved almost 100% success. kNN algorithm achieved 97.2%, which can be considered as a high classification success. This shows that all models in general correctly classified most of the examples in the dataset. In terms of F1 Score, it is possible to say that the results obtained provide a balanced classification performance. The confusion matrix results of the algorithms are shown in Fig. 2 and the ROC curves are shown in Fig. 3.

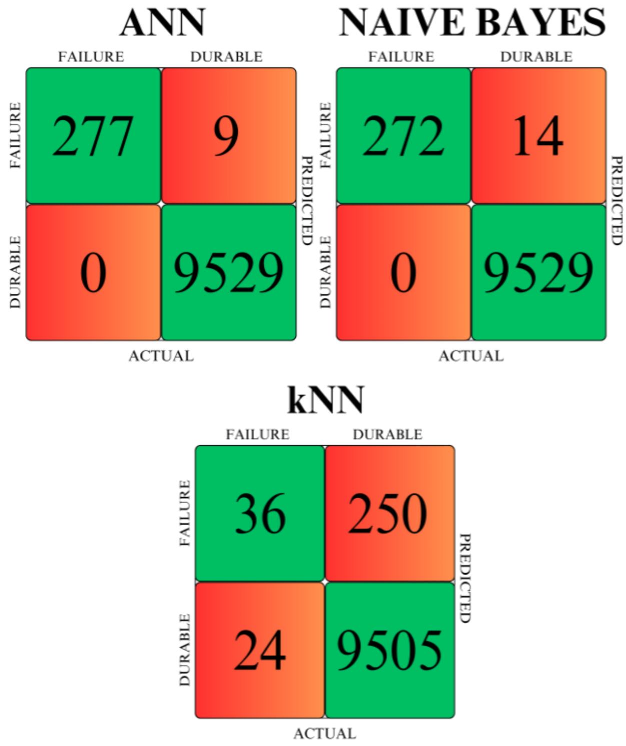
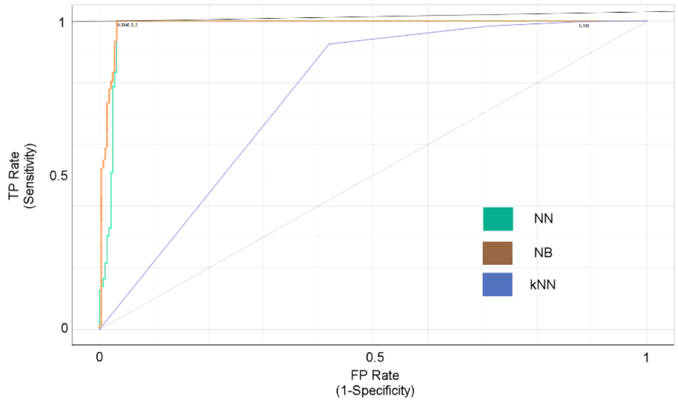
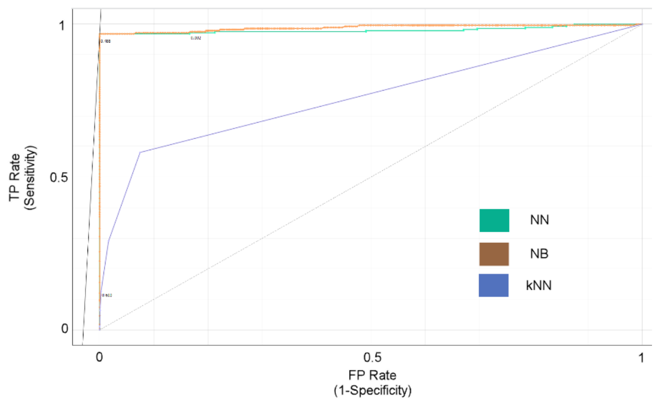


Fig. 2. Confusion matrix for all models



(a)



(b)

Fig. 3. ROC Curve (a) Target: Failure, (b) Target: Durable

When Figure 2 is analysed, it is seen that the TN and TP values of the algorithms are high and the FN and FP values are low. Confusion matrix results show that the most successful algorithm is ANN. The results obtained show that the algorithms perform well in general and correctly distinguish between Durable and Failure states.

Figure 3 shows that the highest AUC value is in the ANN algorithm followed by NB. The lower AUC value of the kNN algorithm compared to the other algorithms indicates that the area under the ROC curve is smaller and the performance of the model is weaker than the others.

Table 5 shows the scores obtained by four different feature selection methods (Gain Ratio, Gini, Chi-Square and FCBF) to show the effect of the features in the dataset on the Durable and Failure classes. The scores obtained for each feature are used to assess how informative that feature is for the feature importance or its relationship with the target variable. The Gain Ratio allows for a fair comparison between features by normalizing the information gain by the uncertainty reduction in a feature. Higher Gain Ratio values mean better feature selection. A low Gini value indicates that the data points in the node have a high probability of belonging to a class, i.e. homogeneous. The Chi-square test is used to assess the relationship between a target variable and each feature. In the case of a classification problem, the Chi-squared test measures how related each feature is to the target variable. A FCBF value of 0 may indicate that the selected features have a low correlation with the classes or are highly correlated with each other. This means that these features do not contribute to classification or modelling performance.

TABLE V

IMPORTANCE OF THE FEATURES IN THE DATASET

Features	**Gain Ratio**	**Gini**	**Chi-Square**	**FCBF**
HDF	0.692	0.022	3831.591	0.823
OSF	0.659	0.018	3131.909	0.621
PWF	0.564	0.008	1466	0.253
TWF	0.564	0.008	1466	0.253
Rotational Speed (rpm)	0.016	0.003	269.462	0.03
Torque (Nm)	0.014	0.003	251.019	0
Tool Wear (min)	0.005	0.001	74.372	0

Table 5 shows that the HDF feature scores higher than all feature selection methods. The HDF feature has a stronger relationship with the target variable than the other features and can be said to contribute significantly to the classification performance. The Gain Ratio, Gini and Chi-Square values of the OSF feature are quite high, indicating that the OSF feature has a strong relationship with the target variable and can improve classification performance. The Gain Ratio and Gini values of the PWF feature are quite high, indicating that the PWF feature has a strong relationship with the target variable and can improve classification performance. The Chi-Square value is also quite high, which supports that the PWF feature has a significant relationship with the target variable. The FCBF value is 0.253, which means that the feature has a low correlation with the classes but is still informative. Similar to PWF, the TWF feature has high Gain Ratio, Gini and Chi-Square values, indicating that the TWF feature also has a strong relationship with the target variable and can improve classification performance. As for the Rotational Speed feature, the Chi-Square value is low and the FCBF value is 0.030, indicating that this feature has a weak relationship with the target variable. However, the Gain Ratio and Gini values are slightly higher than the other features, suggesting that they still have some contribution to the classification. Finally, Torque and Tool Wear features score low for all feature selection methods. Their FCBF values are 0, indicating that these features have no relationship with the classes, have a weak relationship with the target variable and are not very informative to improve classification performance.

V. Discussion and Conclusion

This study aims to evaluate the effectiveness of various classification algorithms for machine fault detection and examine the impact of feature selection on classification performance. The results obtained provide important findings and valuable contributions to identify future research directions.

Analyzing the performance of classification algorithms such as NB, kNN and ANN, it is observed that NB and ANN models achieve high AUC and accuracy values. This shows that the models have a strong potential in detecting machine faults. However, even though the lower AUC value of the kNN algorithm compared to the other two algorithms indicates that it performs relatively poorly compared to the others, it makes a significant contribution.

Feature selection analysis and feature importance ratings revealed that certain features such as HDF, OSF and PWF make significant contributions to classification performance. These features are identified as critical components to be considered in the process of detecting machine faults. However, other features such as Torque and Tool Wear were found to contribute very little to the classification performance.

Limitations of the study and future research directions are also discussed. Issues such as the size of the dataset, the selection of features and the impact of different hyperparameters of the classification algorithms offer potential research areas for future work. Furthermore, issues such as how these algorithms perform in real-world applications and what challenges they may face on an industrial scale need to be further investigated.

In conclusion, this study highlights the effectiveness of classification algorithms for machine fault detection and the importance of feature selection. The findings provide a strong foundation for industrial applications and a framework for future research.

Acknowledgment

I would like to thank Selcuk University Scientific Research Coordinator for their support.

References

[1] Salawu, E.Y., et al. Impact of Maintenance on Machine Reliability: A Review. in E3S Web of Conferences. 2023. EDP Sciences. DOI: 10.1051/e3sconf/202343001226.

[2] Abu-Samah, A., et al., Failure prediction methodology for improved proactive maintenance using Bayesian approach. IFAC-PapersOnLine, 2015. 48(21): p. 844-851. DOI: 10.1016/j.ifacol.2015.09.632.

[3] Janssens, O., et al., Convolutional neural network based fault detection for rotating machinery. Journal of Sound and Vibration, 2016. 377: p. 331-345. DOI: 10.1016/j.jsv.2016.05.027.

[4] Goswami, T. and U.B. Roy. Predictive model for classification of power system faults using machine learning. in TENCON 2019-2019 IEEE Region 10 Conference (TENCON). 2019. IEEE. DOI: 10.1109/TENCON.2019.8929264.

[5] Chen, Z., K. Gryllias, and W. Li, Mechanical fault diagnosis using convolutional neural networks and extreme learning machine. Mechanical systems and signal processing, 2019. 133: p. 106272. DOI: 10.1016/j.ymssp.2019.106272.

[6] Khalil, K., et al., Machine learning-based approach for hardware faults prediction. IEEE Transactions on Circuits and Systems I: Regular Papers, 2020. 67(11): p. 3880-3892. DOI: 10.1109/TCSI.2020.3010743.

[7] Orrù, P.F., et al., Machine learning approach using MLP and SVM algorithms for the fault prediction of a centrifugal pump in the oil and gas industry. Sustainability, 2020. 12(11): p. 4776. DOI: 10.3390/su12114776.

[8] Singh, I. Machine Failure Prediction Cleaned. 2024 [cited 2024 20.04.2024]; Available from: https://www.kaggle.com/datasets/ikjotsingh221/machine-failure-cleaned.

[9] Yang, F.-J. An implementation of naive bayes classifier. in 2018 International conference on computational science and computational intelligence (CSCI). 2018. IEEE.DOI: 10.1109/CSCI46756.2018.00065.

[10] Erdem, K., et al., A Detailed Analysis of Detecting Heart Diseases Using Artificial Intelligence Methods. Intelligent Methods In Engineering Sciences, 2023. 2(4): p. 115-124. DOI: 10.58190/imiens.2023.71.

[11] Cunningham, P. and S.J. Delany, K-nearest neighbour classifiers-a tutorial. ACM computing surveys (CSUR), 2021. 54(6): p. 1-25. DOI: 10.1145/3459665.

[12] Ebrahimi, E., K. Mollazade, and S. Babaei, Toward an automatic wheat purity measuring device: A machine vision-based neural networks-assisted imperialist competitive algorithm approach. Measurement, 2014. 55: p. 196-205. DOI: 10.1016/j.measurement.2014.05.003.

[13] Görtler, J., et al. Neo: Generalizing confusion matrix visualization to hierarchical and multi-output labels. in Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems. 2022. DOI: 10.1145/3491102.3501823.

[14] Kaya, I. and I. Cinar, Evaluation of Machine Learning and Deep Learning Approaches for Automatic Detection of Eye Diseases. Intelligent Methods In Engineering Sciences, 2024. 3(1): p. 37-45. DOI: 10.58190/imiens.2024.85.

[15] Kursun, R., et al. Flower recognition system with optimized features for deep features. in 2022 11th Mediterranean Conference on Embedded Computing (MECO). 2022. IEEE. DOI: 10.1109/MECO55406.2022.9797103.

[16] Padilla, R., S.L. Netto, and E.A. Da Silva. A survey on performance metrics for object-detection algorithms. in 2020 international conference on systems, signals and image processing (IWSSIP). 2020. IEEE. DOI: 10.1109/IWSSIP48289.2020.9145130.

[17] Taspinar, Y.S., et al., Gender Determination Using Voice Data. International Journal of Applied Mathematics Electronics and Computers, 2020. 8(4): p. 232-235. DOI: 10.18100/ijamec.809476.

[18] Gencturk, B., et al., Detection of hazelnut varieties and development of mobile application with CNN data fusion feature reduction-based models. European Food Research and Technology, 2024. 250(1): p. 97-110. DOI: 10.1007/s00217-023-04369-9.

[19] Saritas, M.M., et al., Railway Track Fault Detection with ResNet Deep Learning Models. In 2023 International Conference on Intelligent Systems and New Applications (ICISNA'23)

[20] Cinar, I. Detection of Chicken Diseases from Fecal Images with the Pre-Trained Places365-GoogLeNet Model. in 2023 IEEE 12th International Conference on Intelligent Data Acquisition and Advanced Computing

Systems: Technology and Applications (IDAACS). 2023. IEEE. DOI: 10.1109/IDAACS58523.2023.10348829.

[21] Taspinar, Y.S. and I. Cinar. Detection of Broken Eggs with Deep CNNs and Performance Analysis of Models. in 2023 IEEE 12th International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS). 2023. IEEE. DOI: 10.1109/IDAACS58523.2023.10348756.

[22] Cinar, I. and M. Koklu, Determination of effective and specific physical features of rice varieties by computer vision in exterior quality inspection. Selcuk Journal of Agriculture and Food Sciences, 2021. 35(3): p. 229-243. DOI: 10.15316/SJAFS.2021.252.

[23] Karegowda, A.G., A. Manjunath, and M. Jayaram, Comparative study of attribute selection using gain ratio and correlation based feature selection. International Journal of Information Technology and Knowledge Management, 2010. 2(2): p. 271-277.

[24] Gastwirth, J.L., The estimation of the Lorenz curve and Gini index. The review of economics and statistics, 1972: p. 306-316. DOI: 10.2307/1937992.

[25] Kiliç, S., Chi-square test. Psychiatry and Behavioral Sciences, 2016. 6(3): p. 180. DOI: 10.5455/jmood.20160803110534.

[26] Gharavian, D., et al., Speech emotion recognition using FCBF feature selection method and GA-optimized fuzzy ARTMAP neural network. Neural Computing and Applications, 2012. 21: p. 2115-2126. DOI: 10.1007/s00521-011-0643-1.