■ RESEARCH ARTICLE

Estimating Types of Faults on Plastic Injection Molding Machines from Sensor Data for Predictive Maintenance

Gözde Aslantaş [a†], Tuna Alaygut [a], Merve Rumelli [a], Mustafa Özsaraç [a], Gözde Bakırlı [a], Derya Birant [b]

[a] VESTEL Electronics, Manisa, Turkey
[b] Department of Computer Engineering, Dokuz Eylul University, Izmir, Turkey
[†] gozde.aslantas@vestel.com.tr, corresponding author

RECEIVED JULY 27, 2022
ACCEPTED AUGUST 30, 2022

CITATION Aslantaş, G., Alaygut, T., Rumelli, M., Özsaraç, M., Bakirli, G., & Birant, D. (2023). Estimating types of faults on plastic injection molding machines from sensor data for predictive maintenance. *Artificial Intelligence Theory and Applications, 3(*1), 1-11.

Abstract

Fault type detection for the plastic injection molding machines is an important problem in order to take failure-specific actions to prevent any problem in production, hence providing continuity in procurement. In this study, we treat this problem as a multi-class classification task and proposed a novel machine learning model to achieve reliable and accurate results. We applied the Random Forest (RF) and Extreme Gradient Boosting (XGBoost) algorithms with and without SMOTE (Synthetic Minority Over-sampling Technique) to a real-world dataset for predictive maintenance. According to the results, XGBoost performed better than RF. With the combination of SMOTE method, the performances of both methods increased in terms of accuracy. XGBoost with SMOTE outperformed other techniques by achieving about 98% accuracy on average.

Keywords: machine learning; predictive maintenance; classification; plastic injection molding machines; manufacturing; sensors

1. Introduction

The development of classification techniques to predict electrical machine faults has become a significant area of research and interest in the manufacturing sector since the occurrence of a fault affects production processes and leads to high financial losses and reduced efficiency for the industries. Preventing failures is essential to avoid undesired effects such as vibration, overheating, voltage unbalance, costly machinery repair, reduced safety, and a stop of the production process.

This study aims to utilize machine learning methods to classify faults on plastic injection molding (PIM) machines using features extracted from sensor data. PIM machines are widely used in various industrial plants that especially produce medical devices, white goods, and household appliances. These machines require systematic, timely, and proper maintenance since different types of faults can occur, which influence the normal operation of the equipment. PIM process quality can deteriorate due to machine or tool wear, deviations in the condition of the input material, environmental change, operator

Artificial Intelligence Theory and Applications, ISSN: 2757-9778. ISBN : 978-605-69730-2-4 © 2023 University of Bakırçay

fatigue, and a change of material batch. A PIM machine part (i.e., piston, pump) can be damaged due to high speed, temperature, or vibration as a result of material defects, improper assembly, exceeding the initial design conditions, excessive force and overload during the production. For example, if the process-parameter settings are not properly adjusted by a process operator, unplanned PIM machine stops can occur. In this study, four fault types are detected: (i) hydraulic equipment (HE) malfunctions, (ii) heat application equipment (HAE) malfunctions, (iii) column and shear system hardware (CSSH) malfunctions, and (iv) plasticizing unit hardware (PUH) malfunctions.

The contributions of this study to the literature can be listed as follows. (i) It proposes a novel machine learning model that correctly estimates types of faults on plastic injection molding machines from sensor data. (ii) It compares different classification methods to determine which one is best suited to estimate failure types. (iii) Our study is useful for developing an effective model since it extracts useful features (i.e., entropy, kurtosis, skewness) from raw sensor data by aggregating samples over a time window.

The organization of the paper is as follows. In Section 2, we briefly review the related work in the literature. Section 3 explains our proposed approach and the details of the feature extraction process. Section 4 shows experimental results with a discussion. Finally, in Section 5, we present the conclusion and future work.

2. Related Work

Fault type classification has been studied in many different areas such as energy [1, 2], industry [3, 4], and manufacturing [5]. The literature involves some important and recent studies that use artificial intelligence methods to estimate types of faults in electrical machines and motors. More specifically, traditional machine learning techniques such as Bayesian networks (BN) [4], support vector machine (SVM) [6], k-nearest neighbors (KNN) [7], extreme learning machine (ELM) [3], logistic regression (LR) [3], decision trees (DT) [8], and neural network (NN) [1, 5] have been reported in many studies. Moreover, ensemble learning methods have also been utilized for fault type classification such as random forest (RF) [5], and extreme gradient boosting (XGBoost) [2]. Furthermore, deep learning (DL) [7] techniques have been successfully applied for predictive maintenance such as long short term memory (LSTM) [1] and convolutional neural networks (CNN) [6].

Table 1 shows a summary of the related work on fault detection and classification. Fault diagnosis has been performed for different types of machines, motors, or equipment such as power transformers [9], induction motors [10, 11], wind turbines [2], gearbox [3], steel plates [5], bearing [12], asynchronous machines [13], and rotating machinery [7]. Wang et al. [14] proposed a solution for defect diagnosis induction motors.

Chen et al. [3] demonstrated that the accuracy obtained by the optimized kernel-based extreme learning machine method was 93.97% on average for the detection and identification of faults in a gearbox. In fault detection for three types of wind turbine subsystems, Liu et al. [6] achieved 97.03% accuracy on average with the convolutional neural network. Morales et al. [8] applied different machine learning techniques for the automatic prediction of maintenance intervention types in roads and obtained a final accuracy of 93.4% with the decision tree method. Zhao et al. [9] used support vector machine to identify and classify the winding mechanical fault types and achieved an 83.3% accuracy rate on average.

Unlike the previous studies given in Table 1, our study aimed at the use of classification algorithms to estimate fault types for plastic injection molding machines. Furthermore, useful features (i.e., entropy, kurtosis, skewness) were extracted from raw sensor data by aggregating samples over a time window.

Table 1. Summary of related works

Reference	Year	Method	Description	#Faults	Equipment	Sector
Moradzadeh et al. [1]	2022	SVM, DT, KNN, CNN, LSTM	Identification of locations and types of faults	11	Transmission line	Energy
Leon-Medina et al. [2]	2022	XGBoost	Structural damage classification	4	Wind turbine	Energy
Chen et al. [3]	2021	LR, ELM	Multi-type and concurrent fault diagnosis in rotary machines	5	Gearbox	Industry
Bressan et al. [4]	2021	BN, SVM, KNN	Classification of types of faults on machines from acoustic signals	6	Induction motors	Industry
Trinh and Kwon [5]	2020	NN, KNN, SVM, RF	Classification of fault types and remaining useful life estimation	7	Steel plate	Manufacturing
Liu et al. [6]	2020	SVM, CNN	Fault detection in the process of power generation	5	Wind turbine	Energy
Liu et al. [7]	2018	KNN, NB, SVM, NN, DL	Fault diagnosis of mechanical equipment in modern industry	3	Rotating machinery	Industry
Morales et al. [8]	2018	DT, KNN, SVM, NN	Prediction of maintenance intervention types	5	Road	Transportation
Zhao et al. [9]	2017	SVM	Identification of winding mechanical fault types	3	Power transformer	Energy
Palacios et al. [10]	2015	NB, KNN, SVM, NN, DT	Fault identification in electrical machines	3	Induction motors	Industry
Aydin et al. [11]	2014	Fuzzy DT, NN, GA	Fault diagnosis in manufacturing equipment.	3	Induction motors	Industry
Ertunc et al. [12]	2013	NN	Detection and diagnosis of bearing faults	8	Bearing	Manufacturing
Barzegaran et al. [13]	2013	NN	Identification of winding failures	6	Asynchronous machines	Energy
Wang et al. [14]	2012	NB, KNN, SVM	Defect diagnosis of vital machine components	5	Induction motors	Industry

3. Material and Methods

3.1. Proposed Approach

This paper proposes a machine learning model that correctly estimates types of faults on plastic injection molding (PIM) machines from sensor data. The model is constructed by using classification algorithms. The aim of the study is to analyze sensor readings to predict PIM machine failures and their types before they occur. In this way, it is aimed to make appropriate scheduling of repairs and prevent unexpected failures of machines since machine malfunction affects production processes and leads to financial losses and reduced efficiency for the industries.

Fault detection in PIM machines is composed of a pipeline of various steps from the collection of raw data to the classification. In this process, the prediction of the remaining useful life (RUL) of the machines plays a vital role in unveiling the machines that are more likely to fail. RUL is the time interval between the current point and the point where a failure occurs or it needs maintenance action. For the time index (t) when the machine fails or reaches a maintenance action, the RUL is set as t, and after that RUL is linearly decreasing from t to zero at each time cycle. Therefore, RUL is calculated based on the historical maintenance and failure data of the machines. Many features of the machines affect this time interval such as clamping force, closing force, cycle time, holding pressures, the age of the machine, operational environment, the sequence of the active/idle periods, the quality of key equipment, and oil temperature. The values of these features can be obtained from sensors, process parameter records, alarm records, and tool exchange (planned or unplanned maintenance) records.

Figure 1 shows the pipeline of the proposed approach. In the first step, raw data is collected from PIM machines via sensors. In the data preprocessing step, missing values are handled either by removing or interpolating with the non-null values if the percentage of null values does not exceed a threshold. Outliers in the data are detected by z-score and discarded to obtain better results. After that, the feature extraction process is performed by using statistical methods. Later, features that contribute to the discrimination the most are selected from the data. Synthetic Minority Over-sampling Technique (SMOTE) [15] is applied to the data to re-balance it according to the fault types. Finally, in the last step, data is trained by machine learning methods. The constructed model is tested and the results are evaluated in terms of various measures such as accuracy, recall, precision, and F-score.

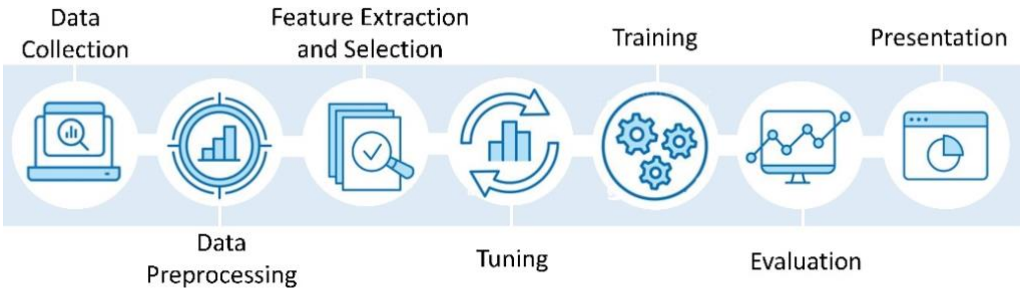

Figure 1. The pipeline of the proposed approach

3.2. Machine Learning Algorithms

In this study, two classification algorithms were used to construct machine learning models for predictive maintenance: random forest and extreme gradient boosting.

Random Forest (RF): It is an ensemble learning-based algorithm that is composed of multiple decision trees, which are called estimators. Decision trees are constructed based on a number of bootstrap samples drawn with replacements from the data. Each bootstrap is composed of a different subset of the data. Each decision tree votes on the classification of a sample, and the final estimation is obtained based on majority voting or averaging. This method is used for both regression and classification problems.

Extreme Gradient Boosting (XGBoost): It is also an ensemble learning-based method that uses gradient descent to optimize each decision tree. It finds the optimal parameter that minimizes the loss. In other words, it adds regularization terms into its loss function. The predictions of decision trees are combined with a voting mechanism. It has the ability to achieve a balance between computing speed and model performance.

3.3. SMOTE

In fault diagnosis applications, the collected data tends to be imbalanced since machines usually operate under healthy conditions prior to the occurrence of faults. When the class distribution of the data was imbalanced, the machine learning algorithms produce classifiers that may perform poorly on minority-class due to two reasons. First, since the majority class dominates a large proportion of the data, its samples are more likely to be selected for training the classifier. Second, the loss factor is calculated based on the ability of the classifier to recognize the majority class more than the minority class. In order to solve the data imbalance problem, oversampling or undersampling techniques are usually used.

SMOTE is a very popular and powerful method to expand minority sample data areas. For a sample *x* in a minority class, SMOTE searches its *k*-nearest neighbors (having the same class label) using a distance metric and randomly selects a sample *y* among them, and then creates a synthetic new sample by calculating linear interpolation between the samples *x* and *y*. The class label for the new sample is the minority class. Different synthetic samples are created based on different neighbor pairs.

In the implementation of SMOTE, the number of neighbors (*k*) is the key parameter for controlling the amount of oversampling of the minority class. For each sample belonging to the minority class in the original dataset, *k* new samples will be generated. In this study, we set *k* to 1, which doubles the number of minority samples. For example, for 100 original minority samples, SMOTE with *k*=1 produces a total of 200 minority samples (i.e. 100 original and 100 synthetics). In this study, the value for *k* was determined by GridSearchCV as an optimal value in the search space. In order to demonstrate the effect of the imbalanced class distribution, we applied RF and XGBoost algorithms with SMOTE and without SMOTE.

3.4. Feature Extraction

Raw sensor data usually do not carry sufficient information itself to describe a fault type since an observation contains one specific value at a particular time instant. For this reason, feature extraction is an important stage in the construction of a classification model and aims at the extraction of the useful information that characterizes each class. Table 2 shows the features extracted from the raw sensor data.

Feature extraction typically transforms the input data into a set of features to provide a compact but effective representation [16]. The determination of these features is a crucial issue as the choice of the classification method to be able to build a good classifier for predictive maintenance. In this study, raw sensor data was aggregated at 60-minute intervals and features were automatically extracted for each internal such as min, max, mean, skewness, kurtosis, and entropy.

Table 2. Features extracted from the raw sensor data

Feature	Description	Formula
Min	The minimum value over the segment	$\mathrm{MIN} = \min(X)$
Max	The maximum value over the segment	$\mathrm{MAX} = \max(X)$
Mean	The average value over the segment	$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$
Absolute Mean	The absolute average value over the segment	$\underline{x} = \frac{1}{n}\sum_{i=1}^{n} \lvert x_i \rvert$
Standard Deviation	The standard deviation of the values over the segment	$STD = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2}$
Median	The middle or central number in the sorted segment values	$MED = \begin{cases} X\left[\frac{n}{2}\right] & if\ n\ is\ even \\ \frac{X[\frac{n-1}{2}] + X[\frac{n+1}{2}]}{2} & if\ n\ is\ odd \end{cases}$
Peak-to-Peak Value	The difference value between the maximum and minimum values over the segment	$PP = \max(x) - \min(x)$
Root Mean Squared (RMS)	The quadratic mean of the discrete values over the segment	$RMS = \sqrt{\frac{1}{n}\sum_{i=1}^{n} x_i^2}$
Kurtosis	The measurement of the peakedness of values in the segment	$KV = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{x_i - \bar{x}}{\sigma}\right)^4$
Skewness	The measurement of the symmetry of the distribution of values in the segment	$SV = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{x_i - \bar{x}}{\sigma}\right)^3$
Crest Factor	The ratio of the maximum value to the root mean square value of samples in the segment	$CRF = \frac{\max(x)}{RMS}$
Clearance factor	The ratio of the maximum value to the squared mean of the sum of the square roots	$\mathrm{CLF} = \frac{\max(x)}{(\frac{1}{n}\sum_{i=1}^{n}\sqrt{\lvert x_i \rvert})^2}$
Shape Factor	RMS is divided by the absolute mean	$SF = \frac{RMS}{\frac{1}{n}\sum_{i=1}^{n}\lvert x_i \rvert}$
Impulse	Max values is divided by the absolute mean	$IMP = \frac{max(x)}{\frac{1}{n}\sum_{i=1}^{n}\lvert x_i \rvert}$
Entropy	The level of information of sensor data in the segment	$E = \sum_{i=1}^{n} P(x_i) log_2 P(x_i)$ $P(x_i) = \frac{x_i}{\sum_{j=1}^{n} x_j}$

4. Experimental Studies

In the experiments, we applied RF and XGBoost algorithms with SMOTE and without SMOTE versions to a real-world dataset. Hyperparameter values were selected according to the best performances of the algorithms in terms of accuracy metric by using GridSearchCV. For example, various alternative numbers of trees (50, 100, 150) were investigated for each algorithm separately. According to the experiments, the optimal values obtained for hyperparameters are given in Table 3. In addition to quality measure, four hyperparameters (*n_estimators*, *min_samples_leaf*, *max_features*, and *max_depth*) were tuned through grid search for the purpose of decreasing computational cost and saving the processing time of model training, similar to the previous studies [17, 18, 19]. Wang et al. [18] and Peng et al. [19] reported that the results were most affected by these hyperparameters.

Table 3. Hyperparameter values

Description	Parameter	Random Forest	XGBoost
Number of trees	n_estimators	150	100
Minimum number of examples to be a leaf	min_samples_leaf	5	5
Maximum depth of the tree	max_depth	12	12
Number of features to apply split	max_features	Log2	Auto
The quality measure	criterion	Gini	Mean Absolute Error (MAE)

In the evaluation phase, a 10-fold stratified cross-validation technique was used to avoid overfitting. In this technique, the original dataset is randomly split into 10 subsets with equal size. While nine subsets are used for training and the remaining subset is used for testing. The experimental results were evaluated in terms of four metrics: accuracy, precision, recall, and F-score. The performance metrics were calculated for each class, along with the macro average values of all classes.

4.1. Dataset Description

In this work, data was collected from the sensors of three plastic injection molding machines, namely HEP3204, HEP3207, and HEP3213, which are specialized in electronics of a Turkish home and professional appliances manufacturing factory. Data sizes collected from the injection machines are 251K, 205K, and 913K, and the collection time interval of each is between September 2019 - September 2021, June 2019 - September 2021, and May 2018 - September 2021, respectively. In this study, a separate predictive model was constructed for each plastic injection molding machine.

Sensor values collected from the injection machines are as follows: clamping force peak value, clamping force set value, closing force Skx value, cycle time ZU-sets value, hold pressures between 1st and 10th steps, hydraulic holding pressure peak value, material cushion smallest value, injection time, and oil temperature. A sample part of the raw data is given in Table 4. Here, the last column (fault type) is the output to be estimated while the others are the inputs.

Table 4. A sample part of the raw data

Machine	Material Cushion Smallest Value	Oil Temp.	Cycle Time	Zone1 Temp.	Zone2 Temp.	Injection Time	Closing Force	RUL	Fault Type
HEP3213	64.5	299.7	189.2	49.9	290.1	0.80	0	1	HE
HEP3213	64.6	299.8	189.2	49.9	289.9	0.80	0	1	HE
HEP3213	64.5	299.6	187.4	50.0	290.0	0.80	0	1	HE
HEP3213	20.6	280.1	46.6	20.0	284.9	1.10	0	1	HE
HEP3213	57.8	309.9	174.4	49.9	290.1	0.62	0	2	HE
HEP3213	22.4	274.9	61.4	18.9	284.9	0.60	2	2	HE
HEP3213	64.6	140.2	344.4	58.0	255.0	0.86	2	2	HE
HEP3213	20.5	280.1	47.8	20.0	284.9	1.14	0	3	HE
HEP3213	24.4	140.2	78.2	58.0	255.0	0.86	2	3	HE
HEP3213	32.5	284.9	67.8	23.0	280.0	0.64	0	1	HAE
HEP3213	20.6	284.9	48.2	18.0	270.0	0.62	0	1	HAE
HEP3213	20.6	284.9	45.8	18.0	269.9	0.62	0	1	HAE
HEP3213	21.6	280.1	49.6	21.9	244.2	0.62	0	2	HAE
HEP3213	20.6	268.1	25.2	23.0	270.0	0.62	0	2	HAE
HEP3213	22.3	285.0	46.0	20.0	269.9	0.62	0	2	HAE
HEP3213	21.6	279.6	49.0	22.0	245.7	0.62	0	3	HAE
HEP3213	21.8	265.6	25.8	21.9	275.1	0.66	0	3	HAE
HEP3213	65.3	289.7	156.6	49.9	279.9	0.92	0	3	CSSH
HEP3213	65.4	289.7	154.6	50.0	280.2	0.92	0.	3	CSSH

A large number of missing values in a data column or data row could mislead the classification problem since they increase uncertainty and unreliable conclusions. On the other hand, the removal of rows or columns having a small amount of missing data leads to a loss of important data and can cause bias in the results. Thus, determining a strategy for handling missing values has an important place in terms of making the most out of the data while having no information loss. According to previous studies like [17], the replacement of missing values with suitable ones is an important step in building an effective classifier since it avoids data loss, provides a better understanding of patterns hidden in data, and usually improves classification accuracy. Motivated by the studies in the literature [17], we successfully handled the missing values problem by using data imputation for numerical features. It is necessary to find a trade-off between the benefit of filling missing data and classification accuracy. Based on our experiences, we determined this trade-off as 40% to obtain satisfactory accuracy in the classification of the fault types. Therefore, in this study, the attributes of the dataset that have null values for less than 40% were filled with the interpolated values of the not-missing values of the same sensor. Furthermore, the rows that include sensor values with high missing rates (>40%) were eliminated.

In the dataset, there are 15 distinct failure messages in total. These messages are grouped into 4 fault types according to the supplied domain knowledge. Errors are listed in Table 5 and associated with the corresponding fault types.

Table 5. The types of faults in the dataset

Failure Message	**Fault Type**
Filter Error Mold Opening and Closing Error Hydraulic Safety Error of Mold Closing Mold Stroke Error Vise Adjustment Error Vise Piston Failure Vice Stroke Error Pump Failure High Oil Temperature Error	Hydraulic Equipment (HE)
Group Cylinder Temperature Error	Heat Application Equipment (HAE)
Colon Failure Shear System Error Vise Opening/Closing Error	Column & Shear System Hardware (CSSH)
Machine Failure to Inject	Plasticizing Unit Hardware (PUH)
No Error	No Error

The data was enhanced with the remaining useful life (RUL) feature. RUL was calculated by subtracting the collection date from the maintenance/failure date. After that, the data was annotated based on the condition where RUL is smaller than or equal to 3 days. The residual failure types along with their count for the three machines are listed in Table 6.

Table 6. Fault types with counts

Machine\Fault Type	**Hydraulic Equipment**	**Heat Application Equipment**	**Column & Shear System Hardware**	**No Error**
HEP3204	203	40	185	2166
HEP3207	757	338	104	1725
HEP3213	939	343	157	8425

As seen in Table 6, the class distribution in the dataset is imbalanced. The majority of samples are labeled as "No Error" for all machines. Samples that need maintenance due to hydraulic equipment type of failure are in the first rank amongst the most seen fault types. Such a class imbalance problem could mislead the classification results by showing a tendency to bias toward the class that has more data. For this reason, we used the SMOTE technique to avoid imbalanced data problems.

4.2. Experimental Results

Table 7 shows the average accuracy, precision, and recall metric values obtained by the methods: Random Forest (RF), XGBoost, RF with SMOTE, and XGBoost with SMOTE, respectively. According to the results, it is possible to say that the algorithms had no difficulty in predicting fault types successfully. The XGBoost + SMOTE algorithm achieved the best performance with an accuracy of 98% on average. In other words, the comparison in this table depicts that XGBoost + SMOTE outperformed other methods in terms of all metrics.

Table 7. Average performance results of each method

Method	Accuracy (%)	Precision	Recall
Random Forest	94	0.936	0.840
XGBoost	95	0.930	0.870
Random Forest + SMOTE	97	0.976	0.970
XGBoost + SMOTE	**98**	**0.980**	**0.980**

Figure 2 shows average macro F-score values obtained by alternative methods. While RF and RF + SMOTE achieved 0.88 and 0.97 according to the performance metric, XGBoost and XGBoost + SMOTE obtained values of 0.90 and 0.98, respectively. With the use of SMOTE technique, the performances of both methods increased. According to the results, the XGBoost + SMOTE algorithm is seen to have better performance than others on average.

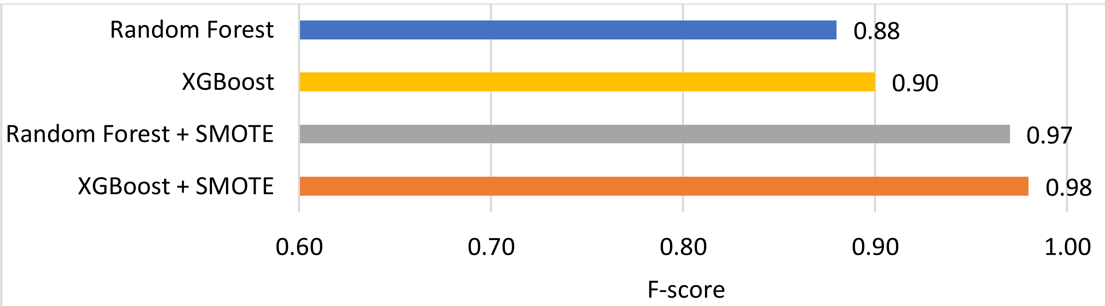

Figure 2. Average macro F-score values

5. Conclusion and Future Work

In this study, we performed an estimation of fault types of plastic injection molding machines by considering the problem as a multiclass classification for predictive maintenance. Raw sensor data was collected from three machines and then preprocessed and analyzed by using machine learning methods. In the data preprocessing step, missing values were handled since the high presence of missing values leads to uncertainty in the results, and therefore, they cause difficulty in extracting meaningful information. While filling a large amount of missing data causes unreliable conclusions, the removal of rows and columns having a small amount of missing data leads to a loss of important data. A widespread strategy is that a dataset is considered reliable if the rate of missing values is below a threshold, which was determined as 40%

in this study. Based on our experiences, the missing values were filled with the suitable interpolation of the observed values in that attribute per product type. As mentioned in Section 3.4. and listed in Table 2, useful features (i.e., entropy, kurtosis, skewness) were extracted from raw sensor data by aggregating samples over a time window. Furthermore, remaining useful life (RUL) values were calculated from data and then used to annotate data for classification. In this study, we grouped 15 distinct failure messages into 4 fault types. We applied Random Forest (RF) and Extreme Gradient Boosting (XGBoost) algorithms with and without SMOTE technique. According to the results, XGBoost performed better than RF. With the use of SMOTE technique, the performances of both methods increased since the dataset is imbalanced. XGBoost with SMOTE achieved the highest accuracy (98%) on average whereas XGBoost without SMOTE predicted with 95% accuracy.

In future work, we aim to expand this study by including other plastic injection molding machines in the manufacturing factory. An application will be implemented to show the classification results to the managers via a user interface for giving feedback about the status of the machines. In this way, the output of the model will be taken into consideration by a manager for decision-making.

References

[1] Moradzadeh, A., Teimourzadeh, H., Mohammadi-Ivatloo, B., & Pourhossein, K. (2022). Hybrid CNN-LSTM approaches for identification of type and locations of transmission line faults. *Electrical Power and Energy Systems*, *135*, 1-13.

[2] Leon-Medina, J.X., Anaya, M., Pares, N., Tibaduiza, D. A., & Pozo, F. (2021). Structural damage classification in a jacket-typewind-turbine foundation using principal component analysis and extreme gradient boosting. *Sensors*, *21*(8), 1-29.

[3] Chen, Q., Wei, H., Rashid, M., & Cai, Z. (2021). Kernel extreme learning machine-based hierarchical machine learning for multi-type and concurrent fault diagnosis. *Measurement*, *184*, 1-12.

[4] Bressan, G. A., de Azevedo, B. C. F., dos Santos, H. L., Endo, W., Agulhari, C. M., Goedtel, A., & Scalassara, P. R. (2021). Bayesian approach to infer types of faults on electrical machines from acoustic signal. *Applied Mathematics & Information Sciences*, *15*(3), 353-364.

[5] Trinh, H-C., & Kwon, Y-K. (2020). A data-independent genetic algorithm framework for fault-type classification and remaining useful life prediction. *Applied Sciences*, *10*(1), 1-20.

[6] Liu, Z., Xiao, C., Zhang, T., & Zhang, X. (2020). Research on fault detection for three types of wind turbine subsystems using machine learning. *Energies*, *13*(2), 1-21.

[7] Liu, R., Yang, B., Zio, E., & Chen, X. (2018). Artificial intelligence for fault diagnosis of rotating machinery: A review. *Mechanical Systems and Signal Processing*, *108*, 33-47.

[8] Morales, F. J., Reyes, A., Caceres, N., Romero, L., & Benitez, F. G. (2018). Automatic prediction of maintenance intervention types in roads using machine learning and historical records. *Transportation Research Record*, *2672*(44), 43-54.

[9] Zhao, Z., Tang, C., Zhou, Q., Xu, L., Gui, Y., & Yao, C. (2017). Identification of power transformer winding mechanical fault types based on online IFRA by support vector machine. *Energies*, *10*(12), 1-16.

[10] Palacios, R. H. C., da Silva, I. N., Goedtel, A., & Godoy, W. F. (2015). A comprehensive evaluation of intelligent classifiers for fault identification in three-phase induction motors. *Electric Power Systems Research*, *127*, 249-258.

[11] Aydin, I., Karakose, M., & Akin, E. (2014). An approach for automated fault diagnosis based on a fuzzy decision tree and boundary analysis of a reconstructed phase space. *ISA Transactions*, *53*, 220-229.

[12] Ertunc, H. M., Ocak, H., & Aliustaoglu, C. (2013). ANN- and ANFIS-based multi-staged decision algorithm for the detection and diagnosis of bearing faults. *Neural Computing & Applications*, *22*(1), 435-446.

[13] Barzegaran, M., Mazloomzadeh, A., & Mohammed, O. A. (2013). Fault diagnosis of the asynchronous machines through magnetic signature analysis using finite-element method and neural networks. *IEEE Transactions on Energy Conversion*, *28*(4), 1064-1071.

[14] Wang, J., Liua, S., Gaoa, R. X., Yanb, R. (2012). Current envelope analysis for defect identification and diagnosis in induction motors. *Journal of Manufacturing Systems*, *31*, 380–387.

[15] Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). Smote: Synthetic minority over-sampling technique. *Journal of Artificial Intelligence Research, 16*(1), 321–357.

[16] Zebari, R.R., Abdulazeez, A. M., Zeebaree, D. Q., Zebari, D. A., Saeed, J. N. (2020). A comprehensive review of dimensionality reduction techniques for feature selection and feature extraction. *Journal of Applied Science and Technology Trends*, *1*(2), 56-70.

[17] Charoen-Ung, P., & Mittrapiyanuruk, P. (2019). Sugarcane yield grade prediction using random forest with forward feature selection and hyper-parameter tuning. In: Unger, H., Sodsee, S., & Meesad, P. (eds) Recent Advances in Information and Communication Technology 2018. *Advances in Intelligent Systems and Computing*, *769*, 33-42.

[18] Wang, R. Gao, W., & Peng, W. (2021). Spatial downscaling method for air temperature through the correlation between land use/land cover and microclimate: A case study of the Greater Tokyo Area, Japan. *Urban Climate*, *40*, 1-16.

[19] Peng, W., Yuan, X., Gao, W., Wang, R., & Chen, W. (2021). Assessment of urban cooling effect based on downscaled land surface temperature: A case study for Fukuoka, Japan. *Urban Climate*, *36*, 1-18.

[20] Baitharu, T.R., & Pani, S.K. (2013). Effect of missing values on data classification. *Journal of Emerging Trends in Engineering and Applied Sciences*, *4*(2), 311–316.

Acknowledgement

This study has been supported by the project numbered 9190028 carried out within the scope of the TUBITAK.