AUTOMATED FAILURE DETECTION IN REFRIGERATORS USING MACHINE LEARNING ALGORITHMS

A Thesis

by

Selin Sarıal

Submitted to the
Graduate School of Sciences and Engineering
in Partial Fulfillment of the Requirements for
the Degree of

Master of Science

in the
Department of Data Science

Özyeğin University
May 2023

AUTOMATED FAILURE DETECTION IN REFRIGERATORS USING MACHINE LEARNING ALGORITHMS

Approved by:

Asst. Prof. Dr. İhsan Yanıkoğlu, Advisor
Dept. of Industrial Eng.
Özyeğin University

Asst. Prof. Dr. Erinç Albey
Dept. of Industrial Eng.
Özyeğin University

Asst. Prof. Dr. Tonguç Yavuz
Dept. of Industrial Eng.
Bilgi University

Date Approved: May 24, 2023

To Berkay Özkan...

ABSTRACT

The sustainable functioning of refrigerators is crucial in residential and commercial settings. These appliances are used continuously throughout the day, and any failure can lead to food spoilage, negatively impacting brand reputation. Therefore, having an efficient failure detection system that can identify and diagnose any problems instantly is essential. This paper proposes a novel machine learning pipeline that uses online sensor data from the refrigerators of anonymous customers and a feedback mechanism to inform customer service about the detected failure remotely. The performance of the system is evaluated through a real-life pilot project, and the results indicate that the proposed method achieves high accuracy in detecting various types of failure. Applying the proposed approach prevents food spoilage, reduces maintenance costs while increasing customer satisfaction, and enhances the reliability and safety of refrigerators.

ÖZETÇE

Buzdolaplarının sürdürülebilir çalışması, konut ve ticari alanlarda önemlidir. Bu cihazlar gün boyunca sürekli kullanılmakta ve herhangi bir arıza, gıda bozulmasına neden olarak marka itibarını olumsuz yönde etkileyebilmektedir. Bu nedenle, anında herhangi bir sorunu tespit edip teşhis edebilen verimli bir arıza tespit sistemine sahip olmak son derece önemlidir. Bu makale, anonim müşterilerin buzdolaplarından çevrimiçi sensör verilerini ve geri bildirim mekanizmasını kullanarak tespit edilen arızalar hakkında müşteri hizmetlerini uzaktan bilgilendiren bir makine öğrenimi akışı önermektedir. Sistem performansı, gerçek hayatta bir pilot proje aracılığıyla değerlendirildi ve sonuçlar, önerilen yöntemin çeşitli arızaları tespit etmede yüksek doğruluk sağladığını göstermektedir. Önerilen yaklaşımın uygulanması, gıda bozulmasını önler, bakım maliyetlerini azaltırken müşteri memnuniyetini artırır ve buzdolaplarının güvenilirliğini ve güvenliğini artırmaktadır.

ACKNOWLEDGEMENTS

I would like to express my sincere thanks to everyone who contributed to the completion of this thesis. Firstly, I am grateful to my thesis advisor, Assoc. Prof. İhsan Yanıkoğlu, for his invaluable guidance, patience, and endless support throughout my research. His insightful comments, constructive feedback, and enthusiastic encouragement were crucial in shaping the direction of this work, and I truly thank him for his mentorship.

I also want to thank my co-advisor from Arçelik, Dr. İsmet Arsan, for his technical expertise, valuable insights, and ongoing support throughout this project. His willingness to share his knowledge, expertise, and time was indispensable in helping me complete this work.

I am also grateful to my family and friends, who have always been by my side throughout my academic journey, for their support. Their love, encouragement, and belief in me have been a constant source of inspiration and motivation, and I could not have done this without them.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

Machine learning algorithms have shown promise in scrutinizing sensor data and identifying different types of failures in these systems. Supervised and unsupervised learning techniques have been used for predictive maintenance, but each has advantages and limitations. This paper proposes an ensemble machine learning approach to address these limitations and improve failure detection accuracy. The proposed approach does not require extensive experiments to collect labeled data and uses the sensory data of existing customers. The proposed failure detection pipeline is tested on a refrigerator by Arçelik.

1.1 Problem Description

Refrigerators are critical household appliances that preserve food quality and reduce waste. Ensuring that they operate continuously and promptly address any faults is of utmost importance. Detecting failures on time can prevent food spoilage and help maintain customer satisfaction. In addition, early intervention in refrigeration failures can result in replacing only the necessary components rather than replacing the entire product, which can be costly and wasteful. Timely detection and intervention of faults in products can increase their lifespan and provide fast and efficient solutions that can be accomplished in a single maintenance appointment. By adequately diagnosing and identifying necessary spare components and solution proposals, products can be repaired without multiple visits, improving customer satisfaction and reducing downtime.

An automatic alarm system can be implemented to improve the overall efficiency of the system significantly to achieve these goals. This system should be designed to be more intelligent and capable of accurately monitoring and detecting failures before they become significant problems. By analyzing data from various sensors and sources, such as temperature sensors, door sensors, and compressor sensors, the system can detect potential issues early and alert maintenance personnel to take necessary actions. It is essential to have access to relevant data sources that can provide information about the performance and operation of the refrigerator to implement a predictive maintenance system for refrigerators. This data can be collected through various means, such as IoT sensors or manual readings taken by maintenance personnel. The data can then be analyzed using machine learning algorithms to identify patterns and trends indicating potential failures.

Overall, a predictive maintenance system for refrigerators can help ensure that these essential household appliances operate efficiently and effectively, minimizing food waste and improving customer satisfaction. By implementing such a system, manufacturers can also improve their brand image by demonstrating a commitment to quality and reliability.

1.2 Previous Work

Over the past few years, the internet of things (IoT) has facilitated the collection of significant volumes of sensor data from home appliances, and machine learning (ML) algorithms have proven effective in scrutinizing these data and identifying different types of failures in such systems [1], [2]. ML techniques also gained popularity across various industries for their ability to analyze large volumes of data and identify anomalies that could indicate issues or failures [3]. As a result, these techniques have been applied in fields such as network systems [4], manufacturing [5], [6], [7], aerospace and aviation [8], [9], finance [10], natural resources and energy [11], transportation [12], and Industry 4.0 [13].

ML classification algorithms such as decision trees [14], [15], random forests [16], support vector machines [17], [18], [19], and extreme gradient boosting [20] have been utilized for failure detection by training labeled data. Additionally, deep learning methods such as convolutional neural networks (CNNs) have shown promise in recognizing complex patterns and correlations in data sets and achieving high accuracy in failure detection [21], [22], [23]. Unsupervised learning techniques have been proposed as an alternative to the aforementioned supervised methods for predictive maintenance, particularly in cases with minimal or no historical data available. These techniques, such as clustering and anomaly detection, can identify abnormal behavior in sensor data without being trained on specific failures [24]. A recent study examined various unsupervised learning algorithms applied to vibration data from an exhaust fan to test their accuracy, performance, and robustness [25].

Supervised learning algorithms have the potential for industrial applications as they can accurately detect failure in real time. However, detecting unknown failures not present in the training data is a challenge for these algorithms, which are only effective at identifying existing faults. Furthermore, collecting labeled data through physical experiments can be expensive, and processing the large data needed for training and inference can be computationally challenging. In contrast, unsupervised methods have an advantage over supervised methods because of their flexibility to detect both known and unknown failures. Nevertheless, these methods usually have higher positive rates than supervised algorithms, meaning they may have a higher rate of false positives. Ultimately, the choice between supervised and unsupervised methods depends on the specific needs and constraints of the application at hand. Additional challenges in using ML algorithms for failure detection include selecting appropriate features and interpreting the results. Some suggested remedies are creating synthetic data, feature selection, and feature engineering [26], [27] to tackle these issues; it is recommended to display and explain the outcomes, such as feature importance, decision rules, and decision boundaries to enhance the transparency and accountability of an ML algorithm [28].

This paper addresses the limitations of supervised and unsupervised learning algorithms for failure detection via an ensemble machine learning approach. The proposed method does not need extensive experiments to collect labeled data; instead, it uses the sensory data of existing customers. Utilizing a supervised ensemble machine learning approach to re-labeled failure data by the autoencoder yields highly accurate predictions on failure times and identifications. The proposed failure detection pipeline is tested on a particular type of refrigerator manufactured by Arçelik, a multinational home appliances company based in Turkey.

The rest of the paper is organized as follows. In Chapter II, the company description is presented. The proposed failure detection ML pipeline is given in Chapter III, and the experimental results are shown in Chapter IV. Finally, Chapter V provides the conclusion and future research directions.

CHAPTER II

COMPANY DESCRIPTION

Arçelik is a multinational home appliances company based in Turkey that has operated since 1955. The company manufactures and sells household appliances such as refrigerators, washing machines, dishwashers, ovens, and air conditioners. Arçelik is a leading brand in Turkey and has a strong presence in Europe, the Middle East, and Africa.

The corporation operates in over 145 countries and has 34 manufacturing facilities in 11 countries. Arçelik is devoted to sustainability and has adopted numerous steps to lessen its environmental effect, such as boosting the energy efficiency of its goods and reducing carbon emissions from its plants.

In addition to its commitment to sustainability, Arçelik strongly emphasizes innovation and technology. The company has made significant investments in research and development to generate cutting-edge goods that fulfill the changing needs of its clients. The products of Arçelik are known for their quality, reliability, and advanced features. Several honors and recognitions have been bestowed upon the firm for its unique goods and services.

Overall, Arçelik is a well-established and respected company dedicated to providing customers with high-quality, innovative, and sustainable home appliances worldwide.

CHAPTER III

METHODOLOGY

This chapter proposes the failure detection pipeline for predicting the time and identification of refrigerator failures. The related pipeline is illustrated by a flow chart in Fig. 1.

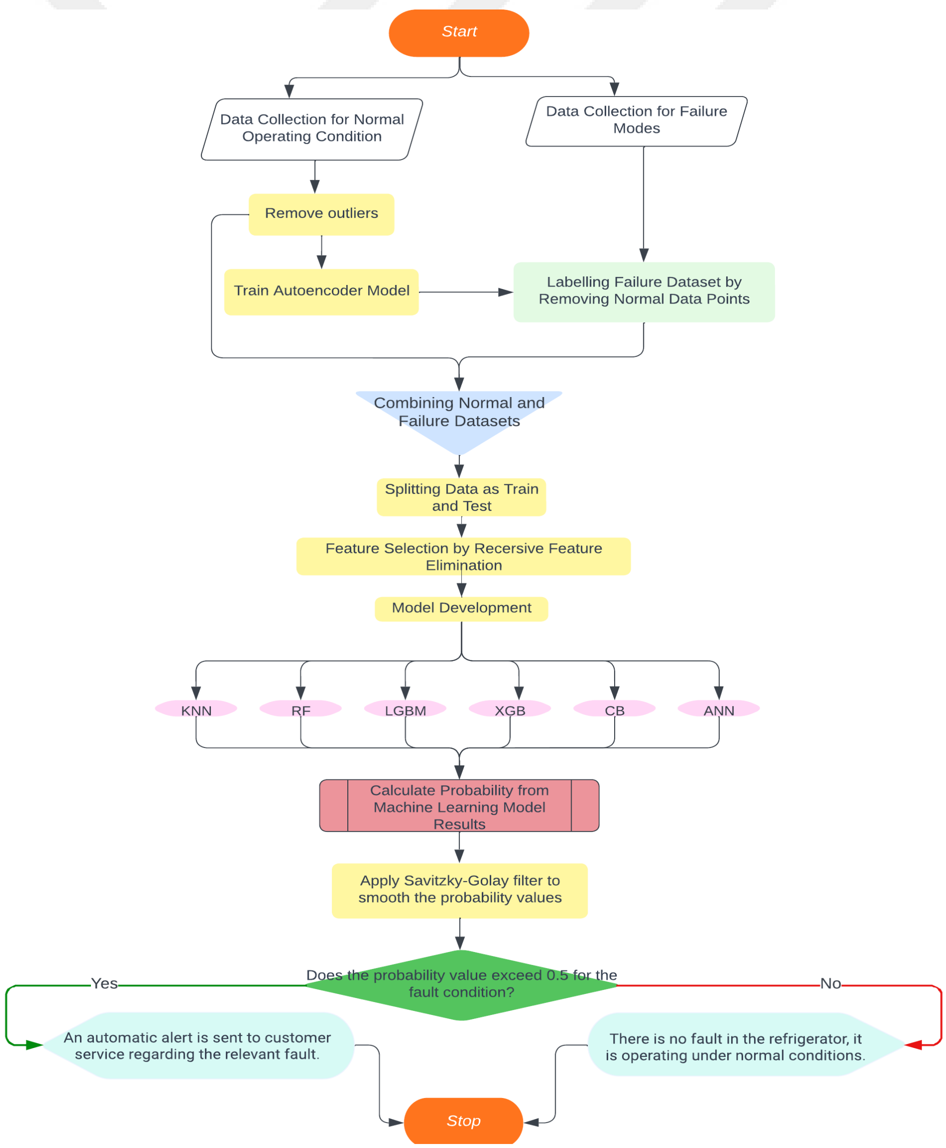

Figure 1: Failure detection pipeline

Initially, the proposed approach involves two phases: a training phase and a real-life application phase. The approach identifies four types of refrigeration failures: fan failure, compressor failure, gas leakage, and door gasket failure. In the training phase, sensor data from anonymous customers' refrigerators are used to train ML models, eliminating the need for expensive physical experimentation. Refrigerator failures diagnosed by field teams after customers' complaints are labeled according to expert failure reports, and the related data of such refrigerators are collected from the system. This data may contain normal operating conditions and failure modes since a customer ticket usually opens days after a refrigerator starts malfunctioning. Therefore, we adopt an autoencoder approach to extract failures from normal operating conditions. Later, the related data set is utilized in the proposed majority selection-based ensemble ML approach. The algorithms that are considered in ensemble ML include k-nearest neighbors (KNN), random forests (RF), light gradient boosting machine (LGBM), extreme gradient boosting (XGB), categorical boosting (CB), and artificial neural network (ANN). In real-life application phases, an automatic alarm system is developed using the ensemble ML approach using a majority selection-based probabilistic approach utilizing the Savitzky-Golay filter. The developed automatic alarm system is used daily for performance evaluation of all connected refrigerators and automatically diagnoses the fault remotely in case of any malfunction and takes the necessary actions.

In the remainder of this section, each step of the failure detection pipeline will be explained in detail. The organization of the subsections will be as follows. The data set used in the study is described in 3.1, while the data collection methodology is explained in 3.2. The data set undergoes a series of preprocessing steps to prepare it for ML model development, outlined in 3.3. The details of the ML model development and the analysis of the combined model results are presented in 3.4 and 3.5, respectively.

3.1 Dataset Description

This study collects sensor data from refrigerators every 6 minutes for failure detection, consisting of 91 features. However, some columns contain product-specific information, such as software versions, unsuitable for training ML models. Therefore, redundant columns are removed, resulting in a meaningful data set of 48 valid variables. These variables comprise necessary performance measures, such as fresh food (FF) and freezer (FRZ) air temperature values, compressor speed, and fan opening time. A sample example of the sensor readings for the FF and FRZ air temperature values is displayed in Fig. 2 to illustrate the data set.

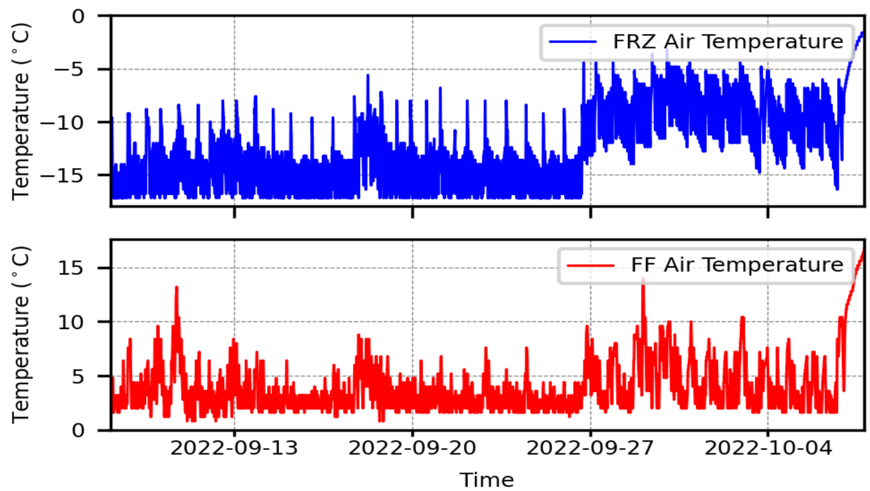

Figure 2: Example of a sensor data visualization

3.2 Data Collection

Collecting data from experiments for every type of failure can be costly and time-consuming, posing a significant challenge when replicating real-world usage patterns. A new approach is proposed to overcome these limitations, which involves using anonymous customer data, and service tickets generated by service technicians after their interventions. Data sets representing normal operating conditions and failure modes are collected and described below.

3.2.1 Normal Operating Conditions

During the data collection process aimed at representing the normal operating condition of refrigerators, records of refrigerators with failure reports are excluded from the list of all connected refrigerators. It is assumed that the remaining refrigerators can operate under normal conditions, and their data sets are collected to represent the normal operation of refrigerators. To eliminate the possibility of including data points with potential failure in the collected data that represent normal operating conditions, an outlier detection algorithm called the isolation forest model is developed using FF and FRZ air temperature values, which are considered two of the most fundamental features in refrigerators. After applying the isolation forest algorithm to the data set, 1817 data points are identified as outliers among the 18174 data points. The remaining 16357 data points are determined to be normal and are labeled as "F1".

3.2.2 Failure Modes

Before collecting data from faulty refrigerators, critical failures that affect the performance are identified by cooling experts. Table 1 provides a list of class names and their corresponding codes used in this paper to represent the normal operating condition and various failure modes.

Table 1: Class names and their codes

Class	*Code*
Normal Operating Condition	F1
Fan Failure	F2
Gas Leakage	F7
Compressor Failure	F14
FF Door Gasket Failure	F131
FRZ Door Gasket Failure	F133

The relevant field in the fault registration report is filtered and classified by type to represent the failures. However, it is possible to make mistakes while entering information into the field, such as recording a gas leakage in a refrigerator as a fan failure. Consequently, only refrigerators with material changes relevant to the malfunction are chosen to collect data representing the failure to ensure accuracy. All failure codes, except gas leakage, are verified based on the modified material code to implement this strategy. When collecting data on gas leakage failures, refrigerators with improved performance after the gas charging process are selected, as material changes are not typically carried out in such cases. For the representation of failure data, the start date of the data is selected as three days before the customer complaint, and the date of repair of the refrigerator is taken as the end date.

The reason for using a three-day window before the customer complaint is that the functioning of a refrigerator may have already been unsatisfactory at the time of the complaint, potentially resulting in data loss during the initial stages of the fault. On the other hand, extending the window to three days may misclassify normal operating conditions as faulty. An autoencoder architecture is employed to mitigate this risk, and its details will be explained in 3.3.3.

3.3 Data Preprocessing

3.3.1 Data Cleaning

Data cleaning is a crucial stage in the pipeline, aimed at improving the accuracy, completeness, and consistency of the raw data by detecting and correcting any inconsistencies, errors, or inaccuracies. All duplicate, missing, and undefined data points are identified and removed from the data set.

3.3.2 Data Scaling

Normalization is applied to the dataset using MinMaxScaler, which scales the data and converts it to a range between 0 and 1. This technique is frequently employed as a preprocessing step for numerous ML algorithms. It has been shown to improve performance during training while reducing the impact of outliers in the data [29].

3.3.3 Autoencoder

It is essential to eliminate data points that represent normal operation from data sets containing failure refrigerators data to enhance the effectiveness of failure mode classification. The autoencoder is a well-established deep-learning method that represents data in an encoded form. It comprises an encoder and decoder component that are jointly trained during training. The encoder compresses the data while the decoder decompresses it, ensuring that the output layer matches the input layer [30]. The autoencoder model trained on normal refrigerators data is applied to each type of failure data set to accurately label the data sets of faulty refrigerators. It involves calculating the reconstruction error (RE) between the input data $\mathbf{x}$ and the output of the autoencoder $\hat{\mathbf{x}}$ for each data point in the failure data set. RE represents the difference between the original input and its reconstructed output, given by the equation $||\mathbf{x} - \hat{\mathbf{x}}||^2$.

The autoencoder can identify failures in the data set by setting a threshold value for the RE. The threshold value is obtained by adding one standard deviation of the RE to the mean of the RE ($\mu_{\mathrm{RE}} + \sigma_{\mathrm{RE}}$).

Data points with RE above the threshold are labeled failures at the final stage. This approach ensures that the data sets of faulty refrigerators are accurately labeled, facilitating the development of robust machine-learning models for fault detection and diagnosis. Example results of the autoencoder model for fan failure and gas leakage are shown in Fig. 3 and Fig. 4.

In the final stage, the re-labeled faulty data is combined with normal data to create the data set for ML models. The number of data points for each class in the data set is shown in Table 2.

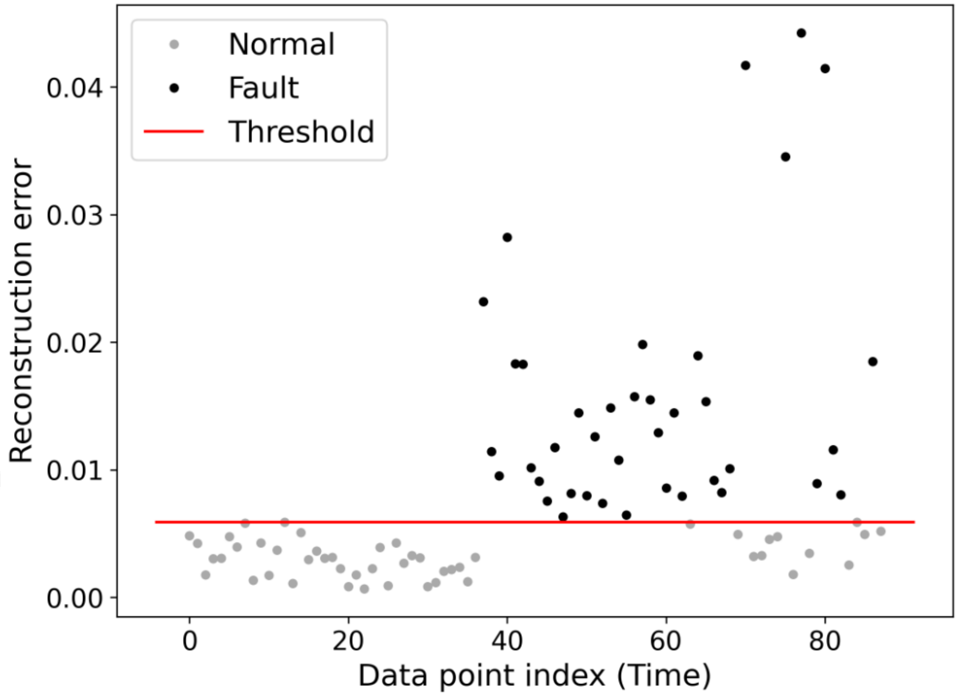

Figure 3: Fan failure labelling example

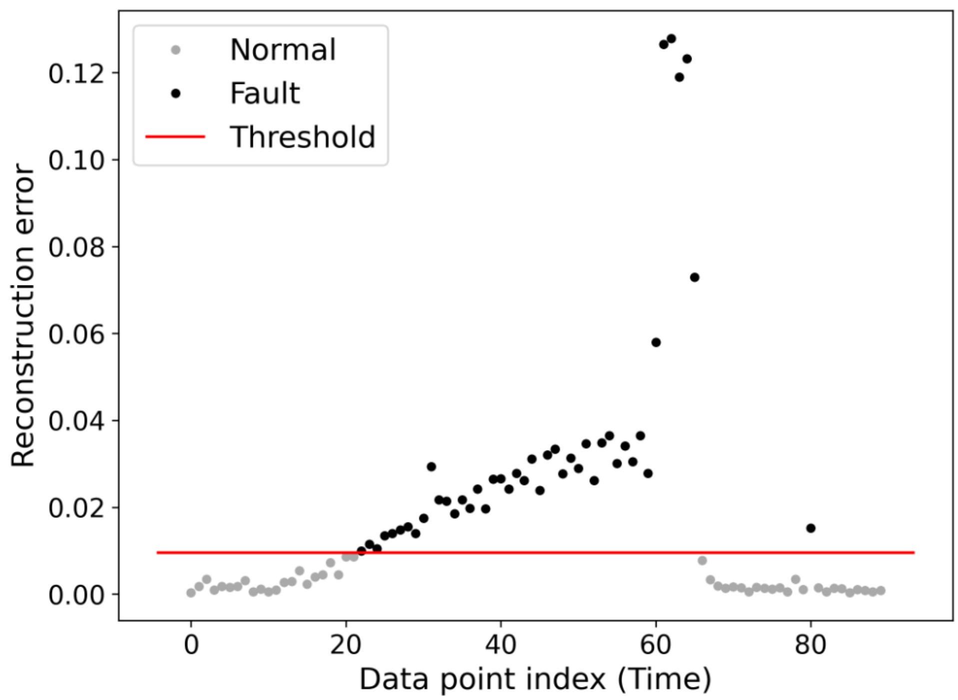

Figure 4: Gas leakage labelling example

Table 2: Class distribution table

FailureCode	*DataPoint*
F1	16357
F2	2629
F7	4155
F14	2306
F131	2077
F133	3031

3.3.4 Splitting

The data set is divided into an 80% training set and a 20% test set using the stratify parameter. It ensures that the target ratio is preserved in the train and test data sets as in the original data set.

3.3.5 Feature Selection

Identifying the independent variables with a high impact on the target variable is crucial in developing ML models. One of the heuristic methods for feature selection is the recursive feature elimination (RFE) method, which is employed to determine the independent variables to be used in the model. This method starts with the entire feature set and eliminates the worst-performing features step by step until the best feature subset is found. The RFE model is trained solely on the training data set, utilizing a 5-fold cross-validation method. During the application of this method, the LGBM classification model is used as the basis, and the accuracy performance evaluation criterion is employed. According to the results obtained from RFE, 11 features have been selected, and their importance is displayed in Table 3.

Table 3: Feature importance results

$FeatureName$	$Importance(\%)$
FF Air Temperature	14.03
FF Eva Temperature	13.95
FRZ Eva Temperature	13.76
FRZ Air Temperature	12.92
Compressor Speed	10.21
Solenoid Valve 2 Way	6.88
FF Fan Open Time	6.20
FRZ Fan Open Time	5.97
Solenoid Valve 1 Way	5.62
Compressor Open Time	5.25
FF Defrost Heater Time	5.21

3.4 Machine Learning Model Development

This study uses six machine learning models to predict six distinct classes, as shown in Table 2. These models include KNN, RF, LGBM, XGB, CB, and ANN, commonly used in practice classification tasks.

KNN is a non-parametric algorithm used for classification tasks that assign a class to an observation based on its k-nearest neighbors in a training dataset. It is a simple algorithm that is easy to understand and implement. However, its performance may degrade on high-dimensional datasets, requiring choosing an appropriate value for the k hyperparameter.

RF is an ensemble learning method that combines multiple decision trees for classification tasks. Each decision tree is constructed using a random subset of features and observations from the training dataset, making it less prone to overfitting. RF can handle large datasets and high-dimensional feature spaces and is generally robust to noisy and correlated data.

LGBM is a gradient-boosting framework used for classification tasks that employ weak learners, usually, decision trees, to build a strong learner. It can handle large datasets with high-dimensional features. Its fast and efficient implementation makes it suitable for real-time applications. LGBM uses a histogram-based approach for splitting continuous features, which can lead to improved performance compared to other boosting algorithms.

XGB is another gradient-boosting framework used for classification tasks, similar to LGBM in many ways. It uses a set of decision trees to build a strong learner. However, it employs a different optimization technique that can lead to improved performance on some datasets. XGB is also known for its speed and efficiency, and it can handle large datasets with high-dimensional features.

CB is a gradient-boosting framework for classification tasks that can handle categorical features. It uses a similar approach to LGBM and XGB. However, it incorporates a novel method for handling categorical features that can lead to improved performance. CB is also known for its speed and efficiency and can handle large datasets with high-dimensional features.

ANN is a class of machine learning models used for classification tasks inspired by the human brain's structure and function. It consists of layers of interconnected nodes (neurons) that perform simple computations on input data to generate output predictions. ANN can handle complex and non-linear relationships between input features and target variables. However, it may require large amounts of training data and careful tuning of hyperparameters to achieve good performance. ANN is a versatile model that can be used for various classification tasks, including image and speech recognition, natural language processing, and more.

These six machine learning models provide a comprehensive and diverse approach to classification tasks. It ensures that the results of the study are robust and reliable.

3.4.1 Hyperparameter Tuning

During model development, the hyperparameters for each model are selected using the grid search algorithm. Grid search is a frequently used approach for hyperparameter optimization that involves systematically training and evaluating the model across all feasible combinations of the predefined hyperparameter set [31]. The optimal set of hyperparameters is ultimately selected based on a thorough assessment of the performance outcomes associated with each combination. The performance of ML models and the examples of real-life application results are shown in Chapter IV.

3.5 *Ensembe ML Approach*

Utilizing a sole ML model to anticipate malfunctions is unsuitable for intricate systems such as refrigerators, where incorrect usage or model prognostications can lead to unnecessary failure notifications. To avert such risks and guarantee precise and dependable predictions, employing an ensemble of ML models to monitor the outcomes of multiple models instead of depending solely on a single model is recommended [32]. This approach aids in mitigating the perils linked to a solitary model and enhances the identification and detection of refrigerator malfunctions, resulting in more advanced failure prevention.

In Fig. 5, the ML mentioned above models were applied to refrigerator data containing compressor failure (F14). It was observed that FF Door Gasket Failure (F131) emerged in some data points by the models, especially in KNN and ANN, and disappeared instantaneously. It could be due to various factors, such as customer usage, such as placing hot items in refrigerators, or sometimes due to incorrect predictions made by the model.

Consequently, it was decided to evaluate the classification results by calculating the probabilities of each data point belonging to a specific class. For instance, when three of the results of the model indicate compressor failure, while the remaining

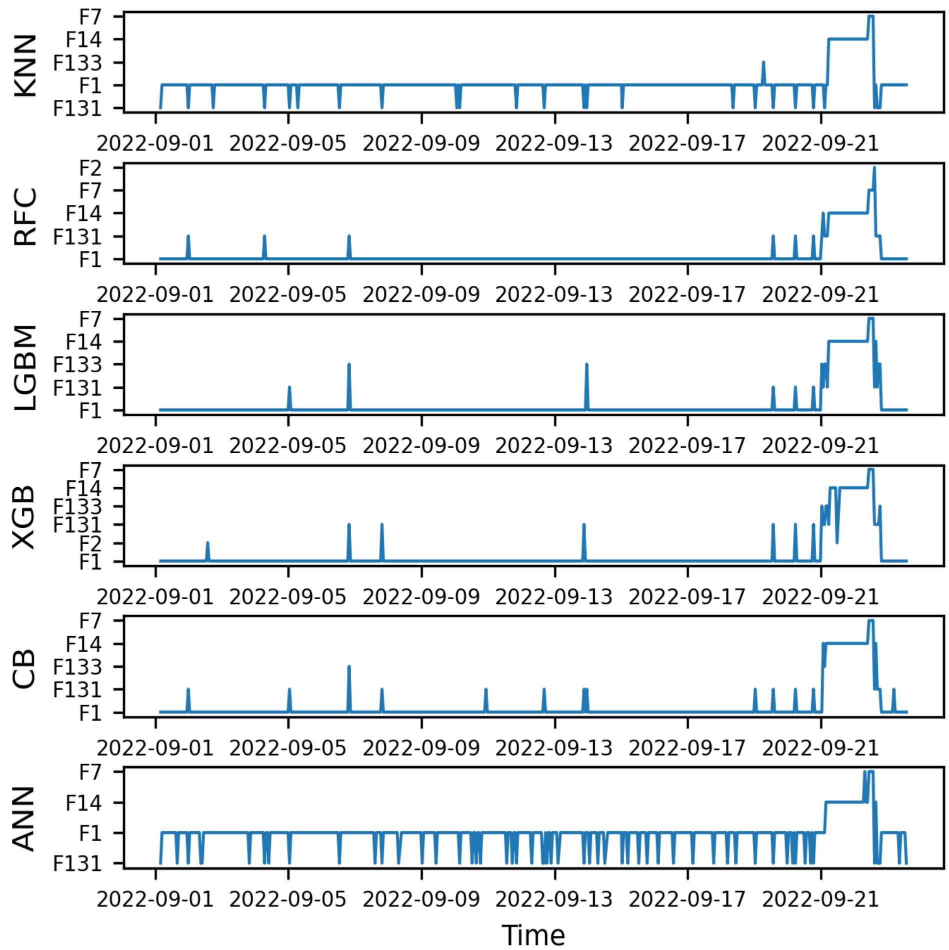

Figure 5: Performance evaluation of models for compressor failure prediction using real customer complaint data

results indicate normal operating conditions of a refrigerator, the probability of the compressor failure is calculated to be 50%. In contrast, the probability of normal operating conditions is also calculated to be 50%. According to Fig. 6, a probability calculation was presented for the same refrigerator mentioned earlier. However, failures can appear and disappear instantly. Similarly, in some cases, some failures can only be seen in specific models. Additionally, more than the probability graph is needed to provide sufficient explainability and traceability of the model results.

Therefore, to accurately monitor the numerical trends of the model results, it is necessary to evaluate the classification outcomes using filtering techniques on probability results.

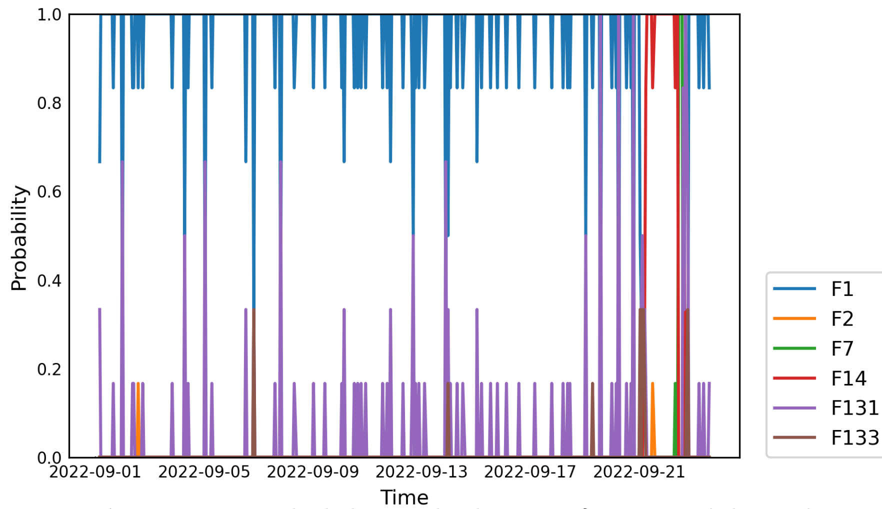

Figure 6: Probability calculation of ML model result

The probability outcomes of monthly refrigerator data are subjected to the one-dimensional Savitzky-Golay filter, a commonly used convolution filter in signal processing for noise reduction [33]. The application of this filter effectively smoothed the probability values, allowing for more accurate trend tracking. The use of this filter reduced the influence of noise and enhanced the precision of the results. Consequently, more dependable and interpretable outcomes were obtained, as demonstrated in Fig. 7.

According to the probability results in Fig. 7, some failures are observed with low probability. On the other hand, it is improbable for multiple failures to co-occur in a refrigerator. Consequently, it is deemed appropriate to exclude failures with probabilities not exceeding 33.33% from the display, to facilitate the identification of the dominant failure. This decision is based on the consideration that each model contributes approximately 16.67% to the overall probability, necessitating the presence of failure in at least two models to warrant display on the graph. Subsequently, an automatic alarm system is developed, activating an automatic alert to customer service

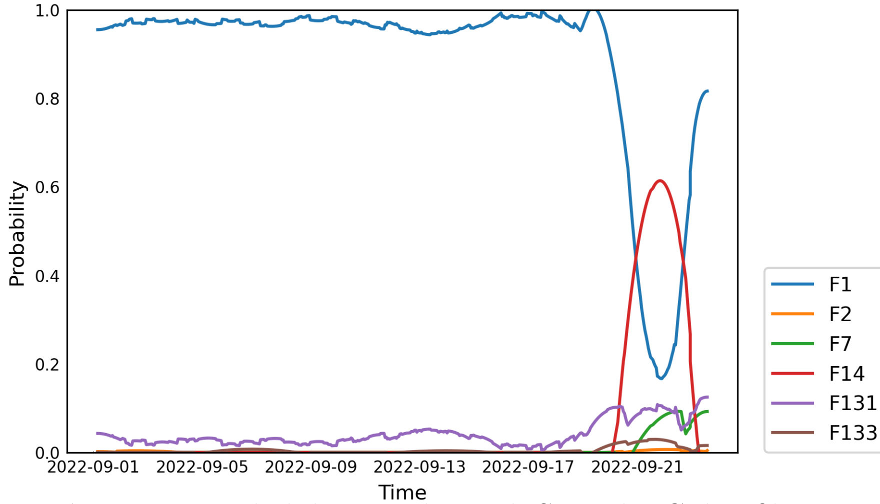

Figure 7: Probability results with Savitzky-Golay filtering

when the probability value for each fault condition exceeds 0.5. The probability of failures exceeding this threshold value is indicated by an 'X' sign in the reports. The primary objective of this mechanism is to prevent unnecessary alarms to customer services.

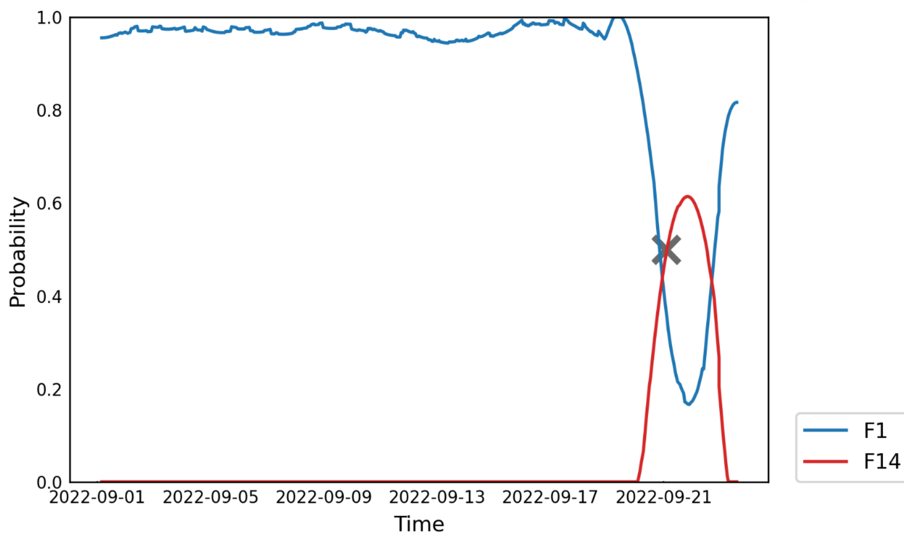

Figure 8: Illustrative example of an automatic alarm mechanism

An illustrative example of the proposed alarm mechanism is demonstrated in Fig. 8. On September 21, 2022, a report was made to customer service regarding a compressor failure (F14) in a specific refrigerator. Subsequently, on September 22, a visit was made to the customer with the necessary spare parts. The compressor failure in the refrigerator was diagnosed and repaired with the provided spare parts. As is evident from the graph, the probability of F14 decreased after the refrigerator was repaired, and it began to return to normal operating condition (F1). Additional failure detection examples will be presented in 4.2.

CHAPTER IV

EXPERIMENTAL RESULTS

The main work described in this chapter is implementing and evaluating automated diagnostics using machine learning models to increase product performance and customer satisfaction while reducing the costs of manual diagnostics and repair processes. The system uses an ensemble approach with six machine-learning models. Its performance is evaluated based on accuracy and a $\mathbb{F}1$ score to provide a more comprehensive assessment. The system, which can detect malfunctions such as fan failure and gas leakage, triggers automatic service calls to customer support, enabling faster failure detection and intervention. Implementing the automatic fault detection system is expected to increase customer satisfaction and product performance and reduce the costs of manual fault detection and repair processes.

4.1 Machine Learning Model Results

Evaluating the performance of an ML model is a crucial step in its development. However, relying solely on accuracy can be misleading, particularly in data sets with imbalanced class structures. In such cases, the $\mathbb{F}1$-score, the harmonic mean of precision and recall, is also evaluated to provide a more comprehensive assessment of the model performance.

The numerical results in Table 4 and Table 5 show that all the models that are used in the ensemble ML approach are successful in accuracy and $\mathbb{F}1$-score. Instead of selecting only one model with the best performance, it was determined that all six models would be combined during the decision-making process to leverage their collective success, which is called majority selection. Adopting this approach decreases the likelihood of making a prediction error significantly.

Table 4: Accuracy result of models

Model	*Accuracy(%)*
KNN	97.32
RF	99.33
LGBM	98.53
XGB	98.79
CB	97.01
ANN	95.57

Table 5: F1-score of class by models

Class	*KNN*	*RF*	*LGBM*	*XGB*	*CB*	*ANN*
F1	0.98	1.00	0.99	0.99	0.98	0.97
F2	0.99	1.00	1.00	1.00	1.00	0.97
F7	1.00	1.00	1.00	1.00	1.00	0.99
F14	0.99	1.00	0.99	0.99	0.98	0.99
F131	0.94	0.98	0.96	0.97	0.91	0.90
F133	0.92	0.98	0.95	0.96	0.90	0.84

4.2 Examples from Real Life Application

Fig. 9 showcases a real-life scenario in which the system detected a fan failure (F2) and initiated an automatic service call to customer support on March 22, 2023. Subsequently, the system ordered the required materials for the fan repair. The customer was informed about the fault, and a technician appointment was scheduled for March 29. During the customer visit, the technicians diagnosed the fan failure using the pre-ordered materials and successfully replaced the fan motor on the same day.

Fig. 10 presents another gas leakage (F7) detection example. The automatic fault detection system triggered customer service when it detected the gas leakage, and the customer was informed about the issue. At first, the customer declined the request for a technician visit and claimed no problem with the performance of the refrigerator. However, two days later, the frozen products in the freezer compartment had thawed,

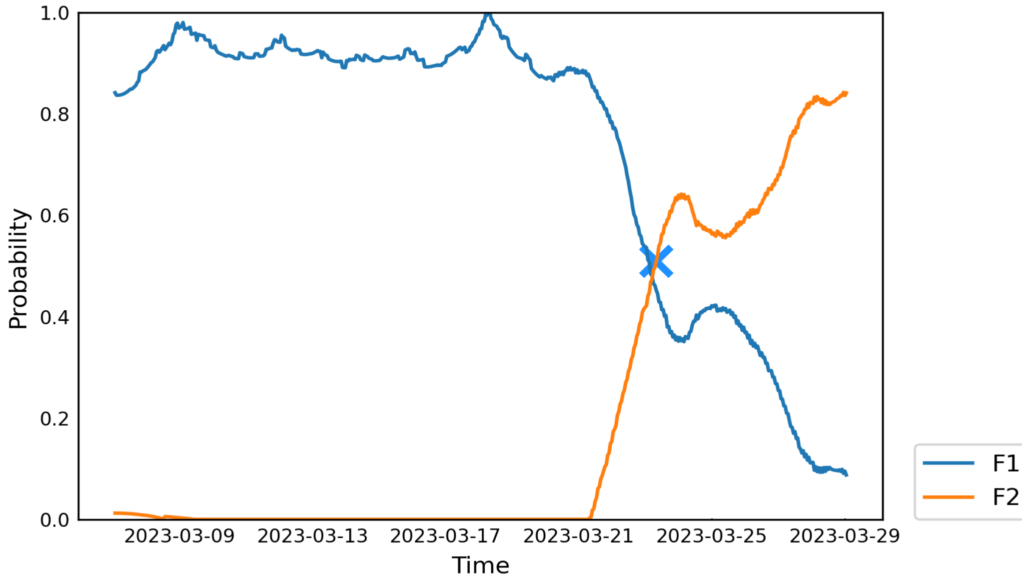

Figure 9: Example case of fan failure

and the customer contacted customer service to report the issue. The technicians were dispatched to inspect the refrigerator, and they detected the gas leakage. They resolved the problem by recharging the gas and restoring the performance of the refrigerator.

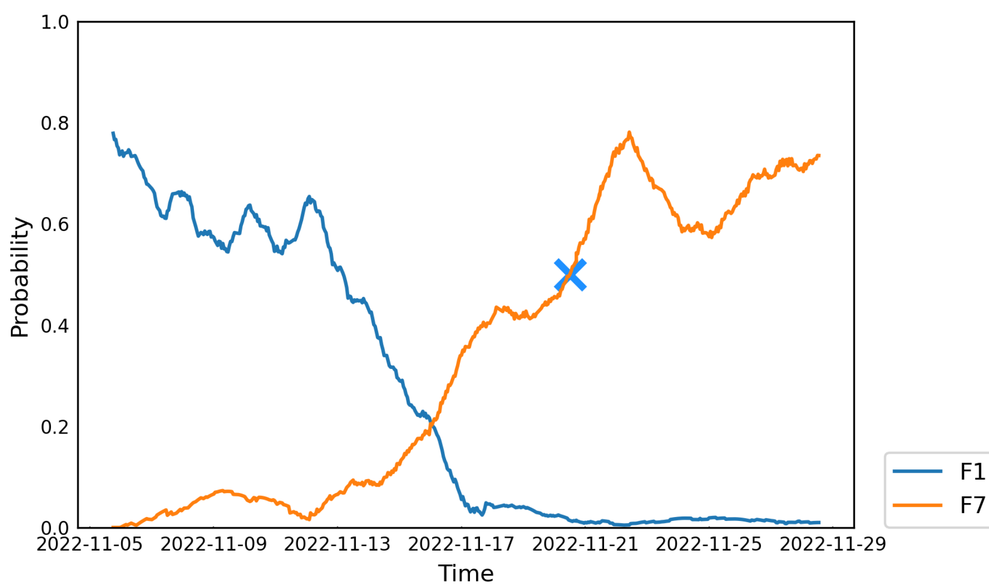

Figure 10: Example case of gas leakage

Implementing an automatic fault detection system alleviates the burden of customer service calls and technician visits, leading to heightened customer satisfaction and improved product performance. Additionally, it has been noted that this system bolsters customer service efficiency by decreasing the duration required for fault detection and intervention. The outcomes suggest that automatic fault detection system applications expedite the detection of faults and provision of customer support, resulting in enhanced product performance and increased customer satisfaction. Consequently, the automatic fault detection system implemented at Arçelik increases product performance and customer satisfaction and reduces the costs of manual fault detection and repair processes.

CHAPTER V

CONCLUSIONS

Integrating ML algorithms for refrigerator failure detection has been a significant breakthrough in appliance engineering. The study findings suggest that supervised learning algorithms, such as KNN, RF, LGBM, XGB, CB, and ANN, implemented through an ensemble ML approach, accurately detect anomalies in refrigerators, enhancing the safety and dependability of refrigerators while reducing food waste. Theresults of ML models demonstrated that the majority-based probabilistic selection approach that uses filtering was reliable and robust in real-life scenarios. Implementing ML techniques in refrigerators enabled manufacturers to design more intelligent and capable cooling systems that accurately monitor, detect, and identify malfunctions, positively impacting customers and the company.

The main contributions of this thesis to the existing literature are:

- The use of anonymous customer data for the ML models' datasets instead of conducting expensive experiments.
- The autoencoder algorithm labeled the faulty datasets.
- The joint decision of different ML models was used to achieve more accurate results by calculating the probability of the model results.
- The use of the Savitzky-Golay convolution filter to filter the probability values to track the trend of numerical values of faults.
- The development of an automatic alarm system to detect fault conditions using filtered probability results to prevent unnecessary fault alarms.

Overall, the labeling method used in this paper and the evaluation of ML model results differ from the existing literature. In future research, incorporating user feedback is a potential avenue for improving the reliability and accuracy of ML models. Soliciting user feedback on the accuracy of the model's predictions and incorporating user-reported issues and failures into the training data could improve model performance. Combining this feedback mechanism into the training process of models can ensure that the model is continuously updated and refined to reflect real-world use cases and user needs. This feedback-driven approach could also help identify new improvement and innovation areas, leading to a more reliable and effective model.

In summary, using ML techniques in refrigerators has enabled the development of more intelligent and capable cooling systems that accurately monitor, detect, and identify malfunctions. The findings of this study demonstrate that the majority-based probabilistic selection approach that uses filtering is reliable and robust in real-life scenarios. The critical contributions of this thesis to the literature include the use of anonymous customer data for the ML models' datasets, the autoencoder algorithm for labeling the faulty datasets, the joint decision of different ML models, the use of the Savitzky-Golay convolution filter, and the development of an automatic alarm system to detect fault conditions. Incorporating user feedback is potential avenue for improving the reliability and accuracy of ML models.

REFERENCES

[1] S. Munir and J. A. Stankovic, "Failuresense: Detecting sensor failure using electrical appliances in the home," in *IEEE 11th International Conference on Mobile Ad Hoc and Sensor Systems*, pp. 73–81, 2014.

[2] S. S. Hosseini, K. Agbossou, S. Kelouwani, A. Cardenas, and N. Henao, "A practical approach to residential appliances on-line anomaly detection: A case study of standard and smart refrigerators," *IEEE Access*, vol. 8, pp. 57905–57922, 2020.

[3] W. Yu, T. Dillon, F. Mostafa, W. Rahayu, and Y. Liu, "A global manufacturing big data ecosystem for fault detection in predictive maintenance," *IEEE Transactions on Industrial Informatics*, vol. 16, no. 1, pp. 183–192, 2020.

[4] J. B. Porch, C. Heng Foh, H. Farooq, and A. Imran, "Machine learning approach for automatic fault detection and diagnosis in cellular networks," in *2020 IEEE International Black Sea Conference on Communications and Networking (BlackSeaCom)*, pp. 1–5, 2020.

[5] H. Qiu, L. Yang, X. Hu, and J. Yao, "Equipment fault detection based on fusion algorithm lstm-cnn-gap," in *International Conference on Intelligent Computing, Automation and Applications (ICAA)*, pp. 42–46, 2021.

[6] B. Luo, H. Wang, H. Liu, B. Li, and F. Peng, "Early fault detection of machine tools based on deep learning and dynamic identification," *IEEE Transactions on Industrial Electronics*, vol. 66, no. 1, pp. 509–518, 2019.

[7] G. A. Susto, A. Schirru, S. Pampuri, S. McLoone, and A. Beghi, "Machine learning for predictive maintenance: A multiple classifier approach," *IEEE Transactions on Industrial Informatics*, vol. 11, no. 3, pp. 812–820, 2015.

[8] B. Bacon, A. Ostroff, and S. Joshi, "Reconfigurable ndi controller using inertial sensor failure detection & isolation," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 37, no. 4, pp. 1373–1383, 2001.

[9] S. Hussain, M. Mokhtar, and J. M. Howe, "Sensor failure detection, identification, and accommodation using fully connected cascade neural network," *IEEE Transactions on Industrial Electronics*, vol. 62, no. 3, pp. 1683–1692, 2015.

[10] N. Nguyen, T. Duong, T. Chau, V.-H. Nguyen, T. Trinh, D. Tran, and T. Ho, "A proposed model for card fraud detection based on catboost and deep neural network," *IEEE Access*, vol. 10, pp. 96852–96861, 2022.

[11] A. Stetco, F. Dinmohammadi, X. Zhao, V. Robu, D. Flynn, M. Barnes, J. Keane, and G. Nenadic, "Machine learning methods for wind turbine condition monitoring: A review," *Renewable Energy*, vol. 133, 10 2018.

[12] H. Chen and B. Jiang, "A review of fault detection and diagnosis for the traction system in high-speed trains," *IEEE Transactions on Intelligent Transportation Systems*, vol. 21, no. 2, pp. 450–465, 2020.

[13] M.-Q. Tran, M. Elsisi, K. Mahmoud, M.-K. Liu, M. Lehtonen, and M. M. F. Darwish, "Experimental setup for online fault diagnosis of induction machines via promising iot and machine learning: Towards industry 4.0 empowerment," *IEEE Access*, vol. PP, pp. 1–1, 08 2021.

[14] R. V and J. G, "Rolling average - decision tree - based fault detection of neutral point clamped inverters," *IEEE Journal of Emerging and Selected Topics in Industrial Electronics*, pp. 1–12, 2023.

[15] T. Yavuz, O. E. Kundakcioglu, and T. Ünlüyurt, "Exact and heuristic approaches to detect failures in failed k-out-of-n systems," *Computers & Operations Research*, vol. 112, p. 104752, 2019.

[16] S. Ma, M. Chen, J. Wu, Y. Wang, B. Jia, and Y. Jiang, "High-voltage circuit breaker fault diagnosis using a hybrid feature transformation approach based on random forest and stacked autoencoder," *IEEE Transactions on Industrial Electronics*, vol. 66, no. 12, pp. 9777–9788, 2019.

[17] Z. Yi and A. H. Etemadi, "Line-to-line fault detection for photovoltaic arrays based on multiresolution signal decomposition and two-stage support vector machine," *IEEE Transactions on Industrial Electronics*, vol. 64, no. 11, pp. 8546–8556, 2017.

[18] K. Jeong, S. B. Choi, and H. Choi, "Sensor fault detection and isolation using a support vector machine for vehicle suspension systems," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 4, pp. 3852–3863, 2020.

[19] S. H. Kim, C. Y. Kim, D. H. Seol, J. E. Choi, and S. J. Hong, "Machine learning-based process-level fault detection and part-level fault classification in semiconductor etch equipment," *IEEE Transactions on Semiconductor Manufacturing*, vol. 35, no. 2, pp. 174–185, 2022.

[20] X. Hu, H. Jia, Y. Zhang, and Y. Deng, "An open-circuit faults diagnosis method for mmc based on extreme gradient boosting," *IEEE Transactions on Industrial Electronics*, vol. 70, no. 6, pp. 6239–6249, 2023.

[21] C.-W. Hung, C.-H. Lee, C.-C. Kuo, and S.-X. Zeng, "Soc-based early failure detection system using deep learning for tool wear," *IEEE Access*, vol. 10, pp. 70491–70501, 01 2022.

[22] W. Silva and M. A. M. Capretz, "Assets predictive maintenance using convolutional neural networks," *20th IEEE/ACIS International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing (SNPD)*, pp. 59–66, 2019.

[23] L. Wen, X. Li, L. Gao, and Y. Zhang, "A new convolutional neural network-based data-driven fault diagnosis method," *IEEE Transactions on Industrial Electronics*, vol. 65, no. 7, pp. 5990–5998, 2018.

[24] G. Michau and O. Fink, "Unsupervised fault detection in varying operating conditions," in *IEEE International Conference on Prognostics and Health Management (ICPHM)*, pp. 1–10, 2019.

[25] N. Amruthnath and T. Gupta, "A research study on unsupervised machine learning algorithms for early fault detection in predictive maintenance," in *5th International Conference on Industrial Engineering and Applications (ICIEA)*, pp. 355–361, 2018.

[26] W. E. Neil H., "Using synthetic data to train an accurate real-world fault detection system," *The Proceedings of the Multiconference on "Computational Engineering in Systems Applications"*, vol. 1, pp. 483–488, 2006.

[27] K. Yan, "Feature selection and analysis on correlated gas sensor data with recursive feature elimination," *Sensors and Actuators B: Chemical*, vol. 212, 06 2015.

[28] T. Rahman, N. Ahmed, S. Monjur, F. M. Haque, and M. I. Hossain, "Interpreting machine and deep learning models for pdf malware detection using xai and shap framework," in *2nd International Conference for Innovation in Technology (INOCON)*, pp. 1–9, 2023.

[29] C. Zhang, J. Yella, Y. Huang, and S. Bom, "Learning from failures in large scale soft sensing," in *IEEE International Conference on Big Data (Big Data)*, pp. 6067–6070, 2021.

[30] G. Dong, G. Liao, H. Liu, and G. Kuang, "A review of the autoencoder and its variants: A comparative perspective from target recognition in synthetic-aperture radar images," *IEEE Geoscience and Remote Sensing Magazine*, vol. 6, no. 3, pp. 44–68, 2018.

[31] D. Belete and M. D H, "Grid search in hyperparameter optimization of machine learning models for prediction of hiv/aids test results," *International Journal of Computers and Applications*, vol. 44, pp. 1–12, 09 2021.

[32] O. Gokalp and E. Tasci, "Weighted voting based ensemble classification with hyper-parameter optimization," in *Innovations in Intelligent Systems and Applications Conference (ASYU)*, pp. 1–4, 2019.

[33] D. Acharya, A. Rani, S. Agarwal, and V. Singh, "Application of adaptive savitzky–golay filter for eeg signal processing," *Perspectives in Science*, vol. 8, pp. 677–679, 2016. Recent Trends in Engineering and Material Sciences.

VITA

Selin Sarıal received a B.Sc. degree in Industrial Engineering from Ozyegin University in 2021. She has been working as a Data Scientist in the R&D Department of Arçelik since 2021. Her research focuses on machine learning applications for failure detection and identification in refrigerators.