Available online at www.eccomasproceedia.org
Eccomas Proceedia UNCECOMP (2021) 159-167

ECCOMAS
Proceedia

UNCECOMP 2021
4th ECCOMAS Thematic Conference on
Uncertainty Quantification in Computational Sciences and Engineering
M. Papadrakakis, V. Papadopoulos, G. Stefanou (eds.)
Streamed from Athens, Greece, 28 -30 June 2021

SENSOR FAULT LABEL IDENTIFICATION FOR ROBUST STRUCTURAL HEALTH MONITORING

Andreea-Maria Oncescu[1], Alice Cicirello[2]

[1]Department of Engineering Science, University of Oxford
Parks Road, Oxford OX1 3PJ, UK
e-mail: andreea-maria.oncescu@sjc.ox.ac.uk

[2] Department of Engineering Structures, Delft University of Technology
Stevinweg 1, Delft 2628 CD, Netherlands
e-mail: a.cicirello@tudelft.nl

Abstract. *Health Monitoring strategies rely on tracking the health status of critical engineering structures (Structural Health Monitoring) and of people (monitoring of medical conditions) to detect anomalies in the measurements and make inferences on the health condition for supporting decisions on preventive actions to be implemented to restore normal conditions. In these applications, the health monitoring devices are subjected daily to various events that can damage internal electrical components and sensors. As a result, the quality of the data collected can be compromised and therefore lead to a wrong health assessment. Therefore, robust health monitoring strategies need to be capable of automatically detecting sensors failures. Having the sensors' data is often not enough to gain insights into a monitoring system failure since the data variation can be related to changes in operating and environmental conditions. Alternatively, a supervised machine learning approach can be used. However, this requires an engineer to label the data in real-time, which rarely happens. Nonetheless, the common practice when a system fails is to write failure reports from which information about the failure can be extracted. Manually extracting comprehensive labels from the failure reports can be time-consuming. A strategy for automatically extracting failure labels from a set of failure reports written to describe failures of different types of sensors of a monitoring device is presented. This strategy consists in transforming the reports in their word vector form, processing each failure report to reduce the list of important words and identifying clusters of reports. The feasibility of the proposed approach is shown through its application to the failure reports compiled to describe seven types of failure of a low-cost wearable device based on an Arduino programmable board. Comparisons between manually extracted labels, and labels extracted with the proposed strategy when considering semi-supervised and unsupervised clustering strategies are presented. It is shown that the proposed strategy is capable of identify the failure label of a cluster of reports with a good accuracy. Therefore, enabling the development of a self-supervised classification algorithm for sensor fault identification for robust Structural Health Monitoring.*

Keywords: Monitoring device failure, Failure reports, Natural Language Processing, Wearable device failure, Automatic failure label extraction.

ISSN:2623-3339
Peer-review under responsibility of the organizing committee of UNCECOMP 2021.
doi: 10.7712/120221.8029.19011

1 INTRODUCTION

In engineering and healthcare applications, effective monitoring strategies are being developed tracking the health status of critical engineering structures (Structural Health Monitoring, SHM) [1, 2] and of people [3, 4] to make inferences on the health condition and support decisions, such as preventive actions to restore normal conditions. Therefore, the measurements obtained with the monitoring system must be informative, reliable and accurate. However, a monitoring device can fail during operating conditions because of poorly manufactured sensors and/or electronics, problems with cable harnesses, ageing effects, improper handling, electromagnetic interference, and environmental factors [5]. Unnoticed failures of the monitoring device undermine the quality of the measurements and consequently compromise inferences and decisions making. In SHM applications, a faulty monitoring device could lead to a wrong assessment of the remaining useful life of a structure [5]. This is one of the key bottlenecks undermining the reliable deployment of SHM technologies. A faulty wearable health monitoring devices can cause fatal conditions to be missed, over-treatment and it might produce health anxiety or fatigue [3, 6].

Failures of the monitoring device may not be detected during inspections [5]. The implementation of an additional monitoring system can be costly and prone to the same problems. Several investigations have been carried out for automatically detecting a faulty monitoring device for chemical process monitoring [7], in aircraft control applications [8, 9], in wearable health monitoring devices [10] and in SHM applications [1, 2, 5, 11, 12]. Broadly speaking, the approaches for sensor validation [5, 13] can be grouped into model-based approaches, knowledge-based approaches and data-driven approaches. Currently the monitoring device health status cannot be reliably identified and/or distinguished from structural failures and/or operating and environmental conditions by using only measurements [14, 15, 16, 17, 18, 19]. Alternatively, a supervised machine learning approach can be used where discriminative features in the measurements are paired with failure labels. However, this would require an engineer to label the data in real-time, which rarely happens and might be inaccurate [1, 2, 20], and to asses the discriminative features. Nonetheless, the common practice when a system fails is to write failure reports [21, 22] from which information about the failure of the device can be extracted. Manually extracting comprehensive labels from the failure reports can be time-consuming. Therefore, this work focusses on automatically extracting failure labels from a set of failure reports written to describe failures of different types of sensors of a monitoring device. This strategy consists in transforming the reports in their word vector form, processing each failure report to reduce the list of important words and identifying clusters of reports for each failure type. The feasibility of the proposed approach is shown through its application to the failure reports compiled to describe the sensor failures of a low-cost wearable device based on an Arduino programmable board. The chosen application displays similar challenges encountered in SHM applications, such as: (i) the sensors employed record various quantities at different rates; (ii) the measurements are influenced by operational and environmental conditions; (iii) similar failure types can occur for the same sensor; (iv) only a limited dataset of recorded failures is available; and (v) the number of elements in the training dataset for each failure type is imbalanced. Comparisons between manually extracted labels, and automatic extraction based on semi-supervised and unsupervised clustering strategies are presented. Finally, the implications of using these labels to train a self-supervised classification algorithm for sensor fault identification are discussed.

2 BRIEF DESCRIPTION OF THE MONITORING DEVICE AND FAILURES CONSIDERED

A low-cost wearable device that includes typical sensors used in wearable applications is chosen for investigating several failure types while keeping the costs low. This monitoring device is composed of a programmable Printed Circuit Board (Adafruit Metro Mini 328), a temperature sensor (digital Dallas Temperature Sensor), a humidity sensor (digital Grove - Temperature & Humidity Sensor Pro), an accelerometer (analog Triple Axis Accelerometer BMA220(Tiny)) and a Galvanic Skin Response (GSR) sensor. Seven types of failure are manually induced for a total of 117 failure instances. Specifically, three failure types are considered for the GSR sensor and two for the accelerometer, a failure type for the temperature sensor and another for the humidity sensor. Moreover, different number of failure instances are considered for each failure type.

Wearable devices, and in general small electronic devices, experience predominantly failures related to the solder joints and to the the sensor connectors [23]. These failures can be caused by improper soldering, ageing, improper handling of the wearable device or cracks in the solder at the connection point caused by a bent Printed Circuit Board (PCB). Within the current setup these failures can be easily reproduced by disconnecting wires at the interface with the PCB. Depending on the sensor and which pin was disconnected, the effects on the recorded signal varied. Moreover, another common failure type is related to burnt resistors. This failure type is induced by adding a resistor to the analog and power pins of the GSR sensor. The induced failures are summarised in Table 1.

Failure Type	Effects on measurements	Occurrences
(GSR, analog, pin)	jumps to 521	24
(GSR, ground, pin)	jumps above 1000	24
(GSR, burnt, resistor)	signal distorted	16
(accelerometer, ground, pin)	jumps to higher values	11
(accelerometer, power, pin)	jumps to lower values or zeros	11
(humidity, power, pin)	jumps to different values or -300%	18
(temperature, ground, pin)	jumps to different values or -127 ° C	13

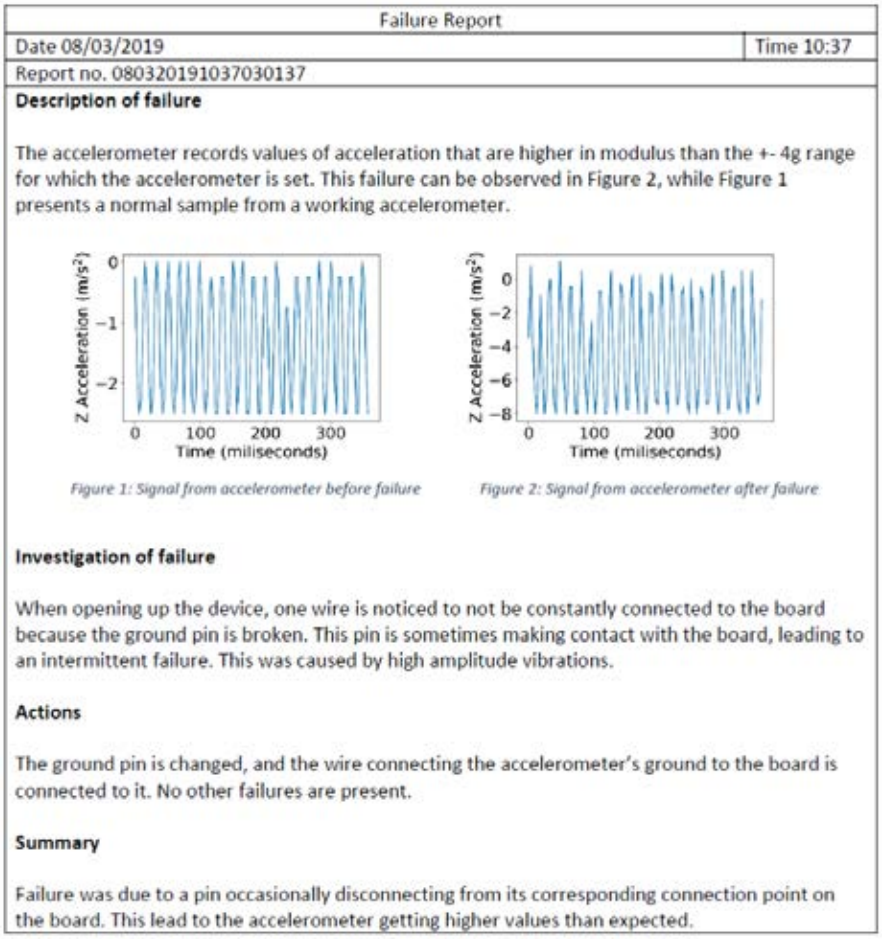
Table 1: Induced failures and effects on recorded data

Data was recorded during controlled and operating conditions, and a failure report was written each time a failure occurred, for a total of 117 failure instances. Data and reports are stored within a Structured Query Language database for easy retrieval of information.

3 FAILURE INVESTIGATIONS AND FAILURE REPORTS

Failure investigations of a structure/system are carried out by an expert to identify the root-cause of failure and suggest remedial actions [21, 22]. Each failure investigation includes the measurements collected in operation, an analysis of the patterns observed in the measurements before and after the failure occurred, the lab experiments and steps required to identify the root-cause of the failure, and a failure report. Currently, the information collected during failure investigations is used for quality assessment, to support decisions about design changes and schedule maintenance [22]. The information collected during these investigations can be also used to improve SHM technologies.

Failure reports are documents with a standard outline [21, 22] and with sections written as free text. The first section focuses on the description of the failure effects observed during operating conditions, and it includes images and plots, and a brief description of the patterns observed in the measurements. Other sections focus on describing the steps taken to identify the root-cause of failure and to reproduce it; the remedial actions implemented; and how to manage similar failures in the future. One example of a failure report for the low-cost monitoring device under investigation is provided in Figure 1.


Failure Report

Date 08/03/2019 | Time 10:37

Report no. 080320191037030137

Description of failure

The accelerometer records values of acceleration that are higher in modulus than the +- 4g range for which the accelerometer is set. This failure can be observed in Figure 2, while Figure 1 presents a normal sample from a working accelerometer.

Figure 1: Signal from accelerometer before failure

Figure 2: Signal from accelerometer after failure

Investigation of failure

When opening up the device, one wire is noticed to not be constantly connected to the board because the ground pin is broken. This pin is sometimes making contact with the board, leading to an intermittent failure. This was caused by high amplitude vibrations.

Actions

The ground pin is changed, and the wire connecting the accelerometer's ground to the board is connected to it. No other failures are present.

Summary

Failure was due to a pin occasionally disconnecting from its corresponding connection point on the board. This lead to the accelerometer getting higher values than expected.

Figure 1: Example of a failure report.

Manual extraction of the information from reports can be time consuming and costly. Therefore a strategy for automatically grouping and extracting the failure labels from failure reports is presented.

4 APPROACH FOR REPORTS CLUSTERING AND LABEL EXTRACTION

A strategy is proposed for automatically grouping the reports according to the failure type described and extract the failure labels by pre-processing the failure reports and applying Natural Language Processing (NLP) techniques. Each document is represented as a vector in a multi-dimensional space, the so-called document embedding [24]. Initially text is extracted from the failure reports (word documents) by using the *docxpy* python package. The number

of representative words of each failure report is reduced to the most relevant ones by applying pre-processing techniques [24] such as: tokenization; reducing list of tokens; part of speech tagging; and lemmatization. Then, each failure report is represented as a vector in a word-space model. In particular, the Term Frequency-Inverse Document Frequency (TF-IDF) [24] is used in combination with Bag of Words (BoW) [24] to refine the list of words. BoW considers the raw frequency of that specific word within the report [24] and therefore selects the most frequent words. However, some words, such as *figure, failure or sensor*, are not helpful for distinguishing the group of reports. These non-informative words are then eliminated by using the TF-IDF approach [24] which considers how many times each word appears in one document and also how many times that word appears in all the documents of the corpus. Words that appear in all the documents being processed will be given a zero TF-IDF score.

These weights can be calculated by first finding the term frequency (tf) [24]:

$$\text{tf(word)} = \frac{\text{Number of times the word appears in document}}{\text{Total number of words in document}} \tag{1}$$

Next, the inverse document frequency term is needed (idf) [24]:

$$\text{idf(word)} = \log\left(\frac{\text{Total number of documents in corpus}}{\text{Number of documents containing the required word}}\right) \tag{2}$$

Finally, the TF-IDF score (which takes values in the interval [0,1]), is calculated as [24]:

$$\text{TF-IDF} = \text{tf(word)} \times \text{idf(word)} \tag{3}$$

The words with TF-IDF scores above a certain threshold are then used to represent each document as a vector. Once this vector representation is obtained, groups of reports belonging to different failure types can be then obtained by applying semi-supervised and unsupervised K-means clustering [25]. While in the unsupervised clustering the initial cluster centres are randomly allocated, in the semi-supervised clustering the initial cluster centres are assigned by selecting one report for each failure type. Once the K-means algorithm has been run to determine each cluster centre and the reports belonging to that cluster, the label of each cluster is manually extracted by selecting a single report within that cluster that is close to the identified cluster centre.

5 APPLICATION OF THE PROPOSED APPROACH

A TF-IDF score threshold of 0.0019 was set. The K-means implementation from the sklearn package [26] was used where the cluster number was set to 7. A brute-force algorithm was implemented to quantify the performance of K-means clustering. This performance was assessed in terms of "accuracy", that is the ratio of correct failure type predictions to total predictions made. Since multiple classes are considered and each class has an unequal number of observations, the confusion matrix is also considered. These matrices display the count values of the correct and incorrect failure labels predictions, and they are defined such that rows display the expected class, while the columns represent the predicted class obtained with the clustering algorithm. The goal is to maximise the count values obtained on the main diagonal since they correspond to the total number of failures for that class.

For the unsupervised K-means clustering with 100 starting points, a maximum accuracy of 83.7% was observed, and a lowest of 70.1%. The clustering with the lowest accuracy is shown in Table 2.

Labels	L1	L2	L3	L4	L5	L6	L7
L1= (GSR, analog, pin)	12	0	0	0	12	0	0
L2= (GSR, ground, pin)	12	12	0	8	0	0	0
L3= (GSR, burnt, resistor)	0	0	16	0	0	0	0
L4= (accelerometer, ground, pin)	0	0	0	11	0	0	0
L5 = (accelerometer, power,pin)	0	0	0	11	0	0	0
L6= (humidity, power, pin)	0	0	0	0	0	18	0
L7= (temperature, ground, pin)	0	0	0	0	0	0	13

Table 2: Unsupervised clustering, confusion matrix with accuracy of 70.1%.

It can be observed that the failure types (GSR, analog, pin), (GSR, ground, pin), (accelerometer, ground, pin), and (accelerometer, power, pin) can be miss-clustered due to the similarity of the failure reports and to the reduced number of reports to learn from. In Table 3 it is shown that even when an accuracy of 83.7% is obtained, the failure type (accelerometer, power, pin) can still be entirely miss-clustered.

Labels	L1	L2	L3	L4	L5	L6	L7
L1= (GSR, analog, pin)	24	0	0	0	0	0	0
L2= (GSR, ground, pin)	0	24	0	0	0	0	0
L3= (GSR, burnt, resistor)	0	0	16	0	0	0	0
L4= (accelerometer, ground, pin)	0	0	0	11	0	0	0
L5 = (accelerometer, power,pin)	0	0	0	11	0	0	0
L6= (humidity, power, pin)	0	0	0	0	6	12	0
L7= (temperature, ground, pin)	0	2	0	0	0	0	11

Table 3: Unsupervised clustering, confusion matrix with accuracy of 83.7%.

These results could be improved by considering pre-knowledge on the labels and/or relationships between words at the TF-IDF stage, before running the clustering algorithms, or by increasing the number of available documents.

For example, when the initial cluster centre was set manually by assigning one failure report to each failure type, the accuracy was improved as shown in Table 4 and the (accelerometer, power, pin) was correctly clustered.

Labels	L1	L2	L3	L4	L5	L6	L7
L1= (GSR, analog, pin)	20	4	0	0	0	0	0
L2= (GSR, ground, pin)	0	24	0	0	0	0	0
L3= (GSR, burnt, resistor)	0	3	13	0	0	0	0
L4= (accelerometer, ground, pin)	0	0	0	2	9	0	0
L5 = (accelerometer, power,pin)	0	0	0	0	11	0	0
L6= (humidity, power, pin)	0	0	0	0	0	18	0
L7= (temperature, ground, pin)	0	2	0	0	0	0	11

Table 4: Semi-supervised clustering, confusion matrix obtained when initial cluster centres are assigned by specifying one report belonging to each cluster.

Therefore, when the labels are extracted automatically, some reports may be incorrectly labelled. As a result, this would reduce the performance of a self-supervised classification algorithm. Nonetheless, the overall performance of the proposed approach can still reach a certain target accuracy while at the same time reducing the setting up time by making the labelling process fully automatic or semi-automatic for the user.

6 CONCLUSIONS

An approach for extracting information obtained during failure investigations is proposed with the aim to facilitate the implementation of a self-supervised machine learning strategy that enables to detect if a monitoring device failed, and if so, to classify which sensor failed and the type of sensor failure.

Within the proposed approach, the process of extracting labels from failure reports, and therefore assigning labels to the corresponding measurements, is sped up by pre-processing the failure reports and applying Natural Language Processing (NLP) techniques to create a vector representation of each failure report in the word-space. The failure report vector representations are clustered together using K-means clustering, and a failure label is assigned to each cluster.

The applicability of the proposed approach was shown by analysing the reports collected when performing failure investigations of a low-cost health monitoring device. This application displays similar challenges encountered in SHM applications, such as: (i) the sensors employed record various quantities at different rates; (ii) the measurements are influenced by operational and environmental conditions; (iii) similar failure types can occur for the same sensor; (iv) only a limited dataset of recorded failures is available; and (v) the number of elements in the training dataset for each failure type is imbalanced. Seven types of failures were manually induced, and measurements with different sensors were recorded during operating and testing conditions. Failure reports for each failure investigated were written, and paired with the recorded data. A small dataset of 117 failures was produced. This limited dataset was characterised by four different faulty sensors, two of which displayed multiple failure types and an imbalanced number failure were considered for for each failure type.

It was shown that the proposed label extraction procedure when using unsupervised clustering can miss-cluster entirely one of the failure types even if yielding an overall high accuracy. As a result, this would reduce the performance of the self-supervised classification algorithms. Nonetheless, the overall performance of the proposed approach can still reach a certain target accuracy while at the same time reducing the setting up time by making the labelling process fully automatic or semi-automatic for the user. It was concluded that when dealing with small failure datasets, with unbalanced classes and similar failure types that the semi-supervised clustering procedure should be preferred.

Indeed, depending on the complexity of the failure reports, the extraction of the failure type labels using NLP strategies can lead to wrong labels assignment, with the risk of not including a particular failure type in the training dataset. Moreover, a failure type can potentially be wrongly identified in the failure report itself, and in fact it might not be supported by the features observed in the data. In turn, this will affect the capability of the proposed approach to detect and isolate the correct failure type for new, unseen data. This is of particular importance for SHM applications. The assessment of the quality of the features-label pairs for improving the training of the classification algorithm is the subject of current research investigations.

REFERENCES

[1] C.R. Farrar, K. Worden, *Structural Health Monitoring: a Machine Learning Perspective*. Wiley, 2013.

[2] R.J. Barthorpe, K. Worden, Emerging Trends in Optimal Structural Health Monitoring System Design: From Sensor Placement to System Evaluation. *Journal of Sensor and Actuator Networks*, **9(3)**, 1–31, 2003.

[3] S. Patel, H. Park, P. Bonato, L. Chan and M. Rodgers, A review of wearable sensors and systems with application in rehabilitation. *J Neuroeng Rehabil.*, **9**, 1–21, 2012.

[4] D. Dias, J. Paulo Silva Cunha, Wearable Health Devices - Vital Sign Monitoring, Systems and Technologies. *Sensors*, **8**, 1–28, 2018.

[5] T.H. Yi, H.B. Huang, H.N. Li, Development of sensor validation methodologies for structural health monitoring: A comprehensive review. *Measurement*, **109**, 200–214, 2017.

[6] Academy of Medical Royal Colleges, *Artificial Intelligence in Healthcare*. Academy of Medical Royal Colleges, 2019.

[7] R. Dunia, J.S. Qin, E.F. Thomas, T.J McAvoy, Identification of faulty sensors using principal component analysis. *AIChE Journal*, **42**, 2797–2812, 1996.

[8] L. Van Eykeren, Q.P. Chu, Sensor fault detection and isolation for aircraft control systems by kinematic relations. *Control Engineering Practice*, **31**, 200–210, 2014.

[9] B.M. de Silva, J. Callaham, J. Jonker, N. Goebel, J. Klemisch, D. McDonald, N. Hicks, J. Nathan Kutz, S.L. Brunton, A.Y. Aravkin, Physics-informed machine learning for sensor fault detection with flight test data. *arXiv*, 2020.

[10] H. Zhang and J. Liu and N. Kato, Threshold Tuning-Based Wearable Sensor Fault Detection for Reliable Medical Monitoring Using Bayesian Network Model. *IEEE Systems Journal*, **12**, 1886–1896, 2018.

[11] M.I. Friswell and D.J. Inman, Sensor Validation for Smart Structures. *Journal of Intelligent Material Systems and Structures*, **10**, 973–982, 1999.

[12] G. Kerschen, P. De Boe, J. Golinval, K. Worden, Sensor validation using principal component analysis. *Smart Materials and Structures*, **14**, 36–42, 2004.

[13] Da. Li, Y. Wang, J. Wang, C. Wang, Y. Duan, Recent advances in sensor fault diagnosis: A review. *Sensors and Actuators A: Physical*, **309**, 2020.

[14] H. Sohn, Effects of environmental and operational variability on structural health monitoring. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, **365**, 539–560, 2007.

[15] J. Kullaa, Eliminating Environmental or Operational Influences in Structural Health Monitoring using the Missing Data Analysis. *Journal of Intelligent Material Systems and Structures*, **20**, 1381–1390, 2009.

[16] J. Kullaa, Distinguishing between sensor fault, structural damage, and environmental or operational effects in structural health monitoring. *Mechanical Systems and Signal Processing*, **25**, 2976–2989, 2011.

[17] J. Kullaa, Robust damage detection in the time domain using Bayesian virtual sensing with noise reduction and environmental effect elimination capabilities. *Journal of Sound and Vibration*, **473**, 2020.

[18] E.J. Cross, K. Worden, Q. Chen, Cointegration: a novel approach for the removal of environmental trends in structural health monitoring data. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, **467**, 2712–2732, 2011.

[19] L.D. Avendano Valencia, E.N. Chatzi, D. Tcherniak, Gaussian process models for mitigation of operational variability in the structural health monitoring of wind turbines. *Mechanical Systems and Signal Processing*, **142**, 2020.

[20] L.A. Bull, T.J. Rogers, C. Wickramarachchi, E.J. Cross, K. Worden, N. Dervilis, Probabilistic active learning: An online framework for structural health monitoring. *Mechanical Systems and Signal Processing*, **134**, 2019.

[21] M.L. Perry, *Electronic Failure Analysis Handbook: Techniques and Applications for Electronic and Electrical Packages, Components, and Assemblies*. McGRAW-HILL, 1999.

[22] J.S. Otegui, *Failure Analysis: Fundamentals and Applications in Mechanical Components*. Springer, 2014.

[23] N. Jiang, L. Zhang, Z.Q. Liu, L. Sun, W.M. Long, P. He, M.Y. Xiong, M. Zhao, Reliability issues of lead-free solder joints in electronic devices. *Science and technology of advanced materials*, **20(1)**, 876–901, 2019.

[24] D. Jurafsky, J.H. Martin, *Speech and language processing - an introduction to natural language processing, computational linguistics, and speech recognition*. Prentice Hall series in artificial intelligence, 2000.

[25] M. Bishop, *Pattern Recognition and Machine Learning (Information Science and Statistics*. Springer-Verlag, 2006.

[26] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel ,M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, E. Duchesnay, Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*, **12**, 2825–2830, 2011.