ATLANTIS PRESS
Advances in Intelligent Systems Research, volume 175
2nd International Conference on Industry 4.0 and Artificial Intelligence (ICIAI 2021)

Smart Feature Selection for Fault Detection in the MEMS Sensor Production Process Using Machine Learning Methods

Itilekha Podder[1,*] Tamas Fischl[2,] Udo Bub[1]

[1] *Eotvos Lorand University, Budapest, Hungary*
[2] *Robert Bosch Ltd. Budapest, Hungary*
[*] *Corresponding author. Email: itilekha19@inf.elte.hu*

ABSTRACT
Micro-electromechanical systems (MEMS) manufacturing is a highly complex process consisting of several hundred steps. The real-time data captured during those process control steps results in a huge data base. Analysis of that enormous amount of data in real-time with high sample rate during production for eventual fault detection and prediction is very challenging. The parameters are highly nonlinear and complex in nature. This makes it difficult for the traditional methods to find this hidden pattern. Advances in Machine Learning (ML) paves the path to investigate the vast dataset and find the hidden complex pattern for early failure prediction and root cause analysis. In the paper, we focus on exploring the applicability of ML methods for the prediction of the affected MEMS inertial sensors using different ML methods. We use statistical analysis to investigate the results to learn about the root cause effect. Finally, we investigate the optimal set of sub-parameters needed for the chosen methods to achieve maximum performance without over-fitting and redundancy.

Keywords: *MEMS Inertial Sensor, BIG Data Analysis, Feature Engineering, Feature Selection, Machine Learning, Recursive Feature Algorithm, Principal Component Analysis.*

1. INTRODUCTION

Micro-electromechanical systems (MEMS) are silicon based electro-mechanical systems which are produced using micro-fabrication technologies (such as lithographic process, etching, polishing, milling). This allows the flexibility of creating complex mechanical and electrical systems. A MEMS wafer contains multiple chips which are then combined with application specific integrated circuit (ASIC) to create a lot. The final MEMS wafers go through front-end testing for instant feedback on any wafer level processing issues. This is to avoid the packaging of the defected wafers. Inertial MEMS sensors which are used for automotive applications must be tested over temperature. It is because the thermal expansion effects influence the performance of the sensors. Thus, to have accurate sensors, factors like thermal stability and uniformity must be tested [1].

With the evolution of the MEMS industry, there is an ongoing demand for more accurate sensor production. Several thousands of technological and product specific parameters such as time-sequential, static, temperature, pressure data etc. are generated during the manufacturing process due to numerous production steps and long supply chain processes. Due to the overwhelming amount of data, timely fault detection during the production process using the traditional method is difficult. Thanks to recent technological advancement, it is now possible to efficiently process large amount of streaming data with Machine Learning (ML) methods. The difficulty lies in the proper selection of the parameters and algorithms to do so.

2. PROBLEM DESCRIPTION

In the paper, we investigated the temperature drift of an internal monitoring parameter (P) of the affected inertial sensors. It is because the temperature drift of the parameter is monitored over lifetime for safety reasons; therefore, understanding how to reduce the drift in the production would be beneficial. Internal Monitoring Signal (P) can be interpreted as a mixture of modulated signal.

Copyright © 2022 The Authors. Published by Atlantis Press International B.V.
This is an open access article distributed under the CC BY-NC 4.0 license -http://creativecommons.org/licenses/by-nc/4.0/.

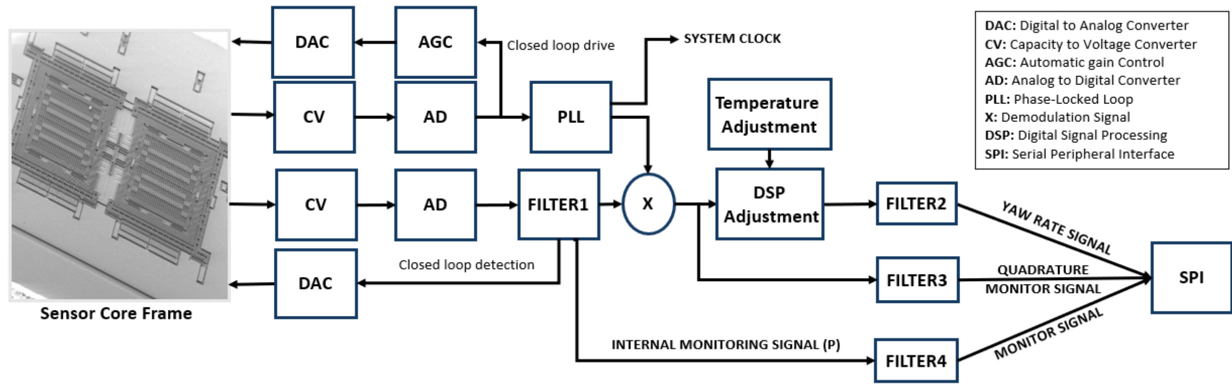

Figure 1 Block diagram of a yaw rate sensor with control loops, adjustments, and filtering.

It contains all physical effects acting on the detection mass like Quadrature and Coriolis forces. MEMS sensors operate on a specific drive frequency (around 15kHz). The frequency region above this drive frequency is damped at "FILTER1" and it generates the internal monitoring signal P as shown in Figure 1. Typically signals, which are two times or three times of the drive frequency, are difficult to distinguish from the drive frequency. This high frequency effect can be detected as the Direct Current (DC) signal effect by the "MONITOR SIGNAL" coming from "FILTER4".

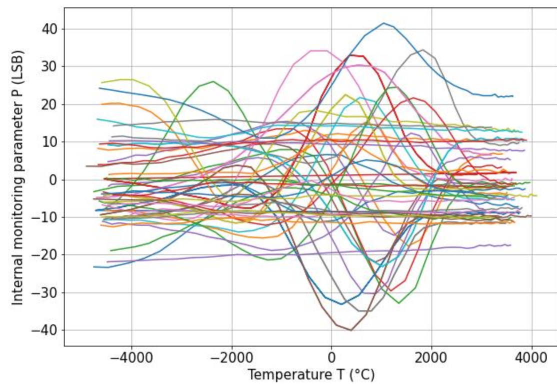

Figure 2 Temperature characteristic of internal monitoring parameter P for some sensors.

Figure 2 shows the DC value behaviour over dynamic environmental temperature T. These high signals from those effects should be avoided as they overload the signal chain and distort the rate signal.

Drift effect is a measure of this nonlinear behaviour of P which occurs due to high frequency components.

$$DFT = \max\{P(T)\} - \min\{P(T\} \quad (1)$$

Here P is the temperature T dependent monitoring parameter and DFT is the drift threshold. The drift here is calculated as the difference between the maximum and minimum value of the internal monitoring parameter over temperature as shown in Equation (1). If the drift value is higher than the predefined threshold (25 LSB is considered here) then those sensors are categorized as affected sensors.

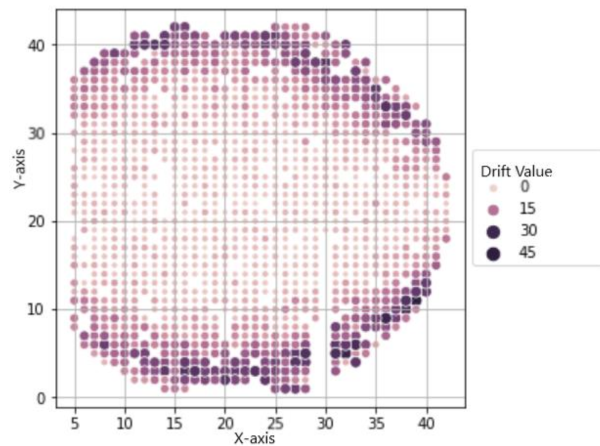

Figure 3 Spatial distribution of drift on wafer level.

Figure 3 shows the spatial distribution of the drift behaviour of the sensors for a single wafer. The darker value in the heat map represents higher drift value, which in this case can be seen around the edges.

3. RELATED WORKS

The sensor manufacturing industry has captured the attention of the Data Science researcher long back due to the high process complexity and constant thrive for process and product improvement. In the review paper Stanisavljevic et al. [2] have shown that application of ML is highly beneficial for optimizing manufacturing process. Wang et al. [3] have provided a comparative study of different ML applications in the domain of additive manufacturing which shows the significance of incorporating ML into different production processes.

Munirathinam et al. [4] have proposed predictive models for equipment fault detection in manufacturing using classical ML algorithms and neural network. Wang et al. [5] have shown that using PCA model for manufacturing stage recognition of machine tools has improved the performance significantly.

The novelty in this paper is that using different ML algorithms along with statistical analysis, we were able to find the most important factors contributing to drift anomalies of the wafers. We showed that using a combination of algorithms, it was possible to pinpoint those features which saved a lot of time for further analysis. This approach is very useful for any further complex, nonlinear data analysis. It further led us to early prediction of the affected sensors using the best performing model.

4. RESULTS

4.1. Experimental Setup

Our experimental dataset consists of information related to yaw rate sensors, inertial sensors, acceleration sensors, process measurements, temperature measurements, infrastructure measurements etc. The temperature drift effect is visible on sensor level as shown in the Figure 2. Our aim is to understand which parameters are the most important in determining the high drift effect.

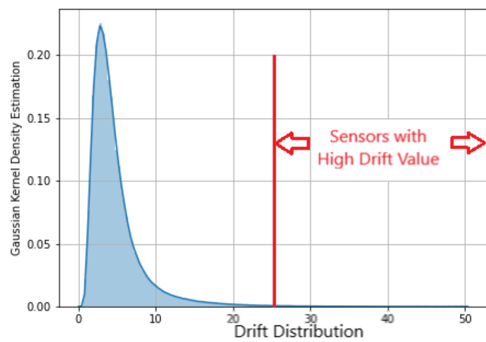

Figure 4 Classification of affected sensors based on temperature drift.

Approximately 1.5 million of data points, consisting of 212 attributes have been used for the analysis. The target variable is classified into two categories given a threshold over the temperature drift value: we call the part affected or failed if it has a high drift value and passed conversely. This classification is demonstrated by the red line over the curve on Figure 4.

For the data preparation, the data with missing values have been removed from the dataset according to the business clarification. Univariate and bi-variate analysis to understand the behaviour of every input variable or features and their dependencies on the target variable. Using Subject Matter Expert (SME) knowledge and Correlation Analysis, we have reduced the number of input features to 42. Z-score method has been used for the detection of outliers as it is the simplest and widely used statistical method. To balance the dataset Synthetic minority Oversampling Technique (SMOTE) has been used to increase the number of minority classes [6].

4.2. Results and Analysis

Recursive Feature Elimination (RFE) [7] has been used for feature ranking and DecisionTree (DT) has been used as the estimator. We have also used Principal Component Analysis (PCA) for feature reduction. For the baseline model, the input features have been chosen randomly and incrementally. Four ML algorithms, 1) XGBoost Classification, 2) Decision Tree, 3) Support Vector Machine (SVM) and 4) Gaussian Naive Bayes (GNB) have been used. These four algorithms have been trained and tested using the previously mentioned three feature selection techniques.

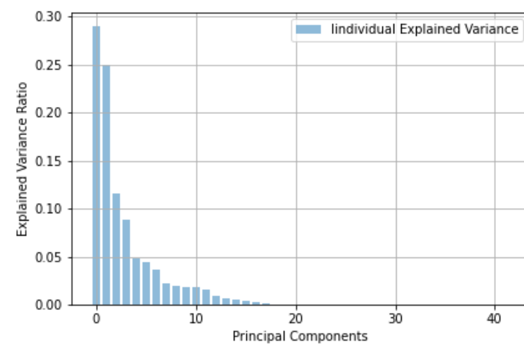

Figure 5 Variance Ratio of the principal components for the input features.

Figure 5 shows the variance of every principal component and 15 components explains 90% variance of the data.

Figure 6 shows the performance comparison of all the four algorithms with varying number of input features using "*Accuracy*" as the performance measure. As expected, when the features have been chosen arbitrarily (which is shown with the red line in the graphs), the performance of the models are at the worst and unstable. When more sophisticated techniques, such as PCA (shown with orange lines) and RFE (shown with green lines) have been used, the performance of the models have been improved drastically.

From careful observation, it can be noticed that not all the features contribute to the accuracy improvement of the models. Rather the accuracy of the models does not improve beyond 15 input features. It means that adding more features beyond this point will be useless. This is where we can use this result as a suggestion for selecting the least amount of the most important features. The XGBoost Classification model was further tuned. In combination with the RFE method, the model achieved 95% accuracy.

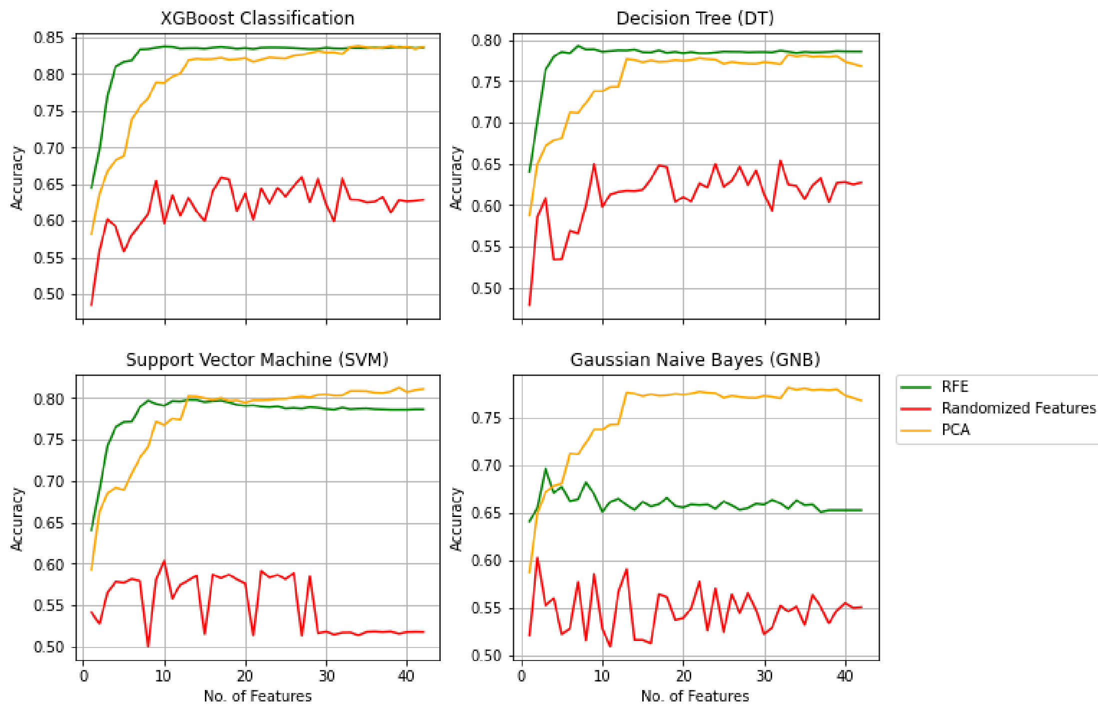

Figure 6 Performance Comparison of different ML Algorithms for Smart Feature Selection using RFE, PCA and Randomized Features.

Further scrutinizing the top two features from the model feature importance, the subject matter expertise was able to see the strong impact of these the parameters on the drift behaviour which helped narrowing down the root cause analysis.

This proposed approach will not only save time and effort, but also least amount of data will be required to analyse and collect. It further helps to reduce digital waste.

5. DISCUSSION AND FUTURE WORK

In this work, we have tried to tackle the hard problem of predicting failures in MEMS sensor production. After introducing the topic and mentioning some of the traditional methods applied for this problem, we have utilised our own approach for greatly reducing the parameters needed to achieve our goals. From an initial set of 212, with the help of subject matter analysis and pre-processing, we have narrowed down to the 42 most important factors.

From here on, we used unsupervised techniques to rank these factors according to their respective importance and by an incremental approach, we have narrowed it further down to a handful. As further work, we are interested in seeing how much further we can reduce the number of parameters by more advanced learning techniques, like neural networks.

AUTHORS' CONTRIBUTIONS

Itilekha Podder and Tamas Fischl conceived the original idea. This was also discussed with Udo Bub. Eventually all authors conceived and designed the study. The experiments in the section 4 were performed by Itilekha Podder. Tamas Fischl contributed in result interpretation as the subject matter expert. All authors contributed to the manuscript version. All authors read and approved the final manuscript.

REFERENCES

[1] V.K. Lindroos, M. Tilli, A. Lehto, T. Motooka: Handbook of Silicon Based MEMS Materials and Technologies (Third Edition), pp. 845--850 Elsevier. (2020). DOI{10.1016/B978-0-12-817786-0.00043-8}.

[2] Stanisavljevic, Darko & Spitzer, Michael. (2016). A Review of Related Work on Machine Learning in Semiconductor Manufacturing and Assembly Lines.

[3] C. Wang, X.P. Tan, S.B. Tor, C.S. Lim. Machine learning in additive manufacturing: State-of-the-art and perspectives, Additive Manufacturing, Volume 36, 101538, ISSN 2214-8604. (2020). DOI: https://doi.org/10.1016/j.addma.2020.101538.

[4] S. Munirathinam and B. Ramadoss. Predictive models for equipment fault detection in the semiconductor manufacturing process. IACSIT

International Journal of Engineering and Technology, 8, August 2016.

[5] Jingshu Wang, Tianbao Cheng, Bin Xing, and Xiaolin Hu. 2020. Manufacturing State Recognition of Machine Center Based on Revised WPT and PCA. In Proceedings of the 4th International Conference on Computer Science and Application Engineering (CSAE 2020). Association for Computing Machinery, New York, NY, USA, Article 104, 1–5. DOI: https://doi.org/10.1145/3424978.3425086

[6] Arjun Puri, Manoj Kumar Gupta, Knowledge discovery from noisy imbalanced and incomplete binary class data, Expert Systems with Applications, Volume 181, 2021, 115179, ISSN 0957-4174, DOI: https://doi.org/10.1016/j.eswa.2021.115179.

[7] Xiaojian Ding, Fan Yang, Sheng Jin, Jie Cao, An Efficient Alpha Seeding Method for Optimized Extreme Learning Machine-based Feature Selection Algorithm, Computers in Biology and Medicine, 2021, 104505, ISSN 0010-4825, DOI: https://doi.org/10.1016/j.compbiomed.2021.104505
.