Project report

Vilmos Szirmai

December 2024

1 Introduction

The topic of my project is related to the automotive industry and is done in a fellowship program at Bosch. In automotive manufacturing, it is critically important to detect both design flaws and faulty components. One approach is to listen to the noise: undesirable noise in a car component can be used to infer various manufacturing or design defects. The aim of the project is to develop machine learning models to localize the source of noise from measurements of vibrations in order to locate the errant parts. To get signals for modeling purposes, we excited parts with an automatic hammer and recorded time signals at several points. The noise is measured with a laser vibrometer, resulting in a signal of velocities as a function of time.

In this report, several approaches are presented to the above problem, discussing modeling choices, metrics of error, model performances, and also different scopes of generalizations.

2 Approach

The approach is to find out where the measurement was made. This is because it is equivalent to localizing the excitation. We went around this using two methods:

- A regression approach to estimate distance from the excitation point
- Discretizing by dividing the component into parts

The first method aims to (relatively) precisely determine the location of the measurement point. In this case, the error-measuring metric is the Euclidean distance. The second approach may be somewhat more practical. However, the problem here arises from merely considering how accurately the model is able to predict the label of the range. In this scenario, we might also count as an error when the model misclassifies a point that was located on the boundary of the range.

3 Formal description of the task

Since our task in both approaches is to predict where the measurements have been taken, we need to take a coordinate system. The origin of this coordinate system should be placed at one of the corners of the plate or part, and in addition a positive x- and y-direction should be taken. This is an absolute system, after taking them together with the measurement and excitation points, then retaining them for each measurement. Both the excitation and measurement points are stored with coordinates, plus they have a point index assigned to them by the measurement software.

In case of coordinate estimation of a measurement point, it is given what we are referencing and what will provide the labels for our supervised learning task. But when we discretize, we need some kind of system by which to group the points.
To do this, we have several built-in algorithms in Python, such as K-means or Gaussian Mixture, both of which divide the points into a fixed number of clusters based on the planar coordinates. Now let us move on to possible generalizations.

4 Generalizations

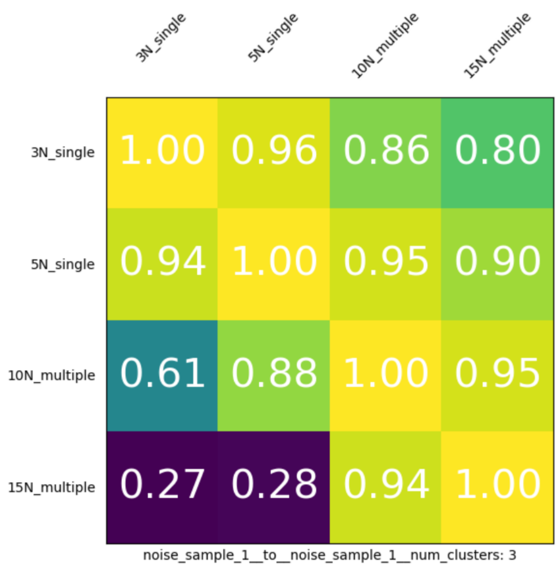
- **Generalization over different forces (measured in Newton)**
  We investigated whether the models can generalize between different forces. The result was that they can if they have to predict from smaller forces to bigger ones. We got much different accuracy scores when we used nearest neighbor interpolation and simple feedforward neural networks (this will be shown on the next page).

- **Generalization over different excitation shapes**
  Models are not too sensitive to this, so this generalization seems to be solved.

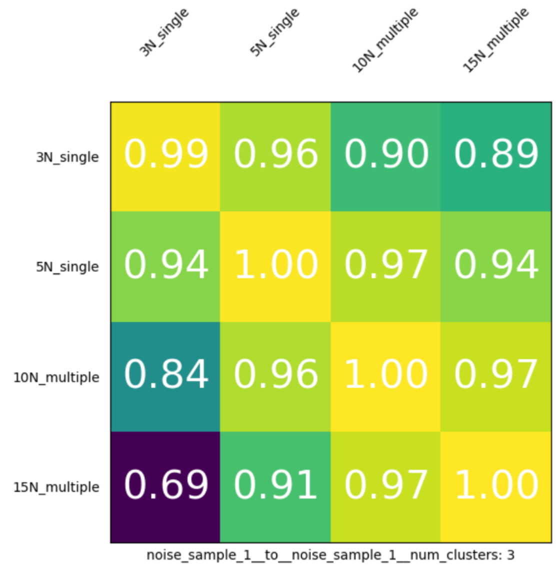
- **Generalization over distance prediction from different excitation points**
  A harder instance, models only achieve from 1 cm to 7 cm mean error.

- **Generalization over distance estimation on different automotive parts**
  It looks almost impossible to solve. Models had from 18 cm to 77 cm mean error.

- **Generalization by finding an appropriate embedding**
  It is thought to be a valid approach to use self-supervision to find a proper embedding. It is by exploiting the inner structure of time signals to predict both discrete clusters and distances.



(a) Interpolation accuracy scores



(b) Neural network accuracy scores

Figure 1: Accuracy matrices with different generalizations, interpolation and neural network

In Figure 1 it can be seen that the neural network outperforms interpolation in most cases, and the matrices also reveal that the models generalize better from lower forces to higher forces than vice versa.

5 Elements of the measurement setup

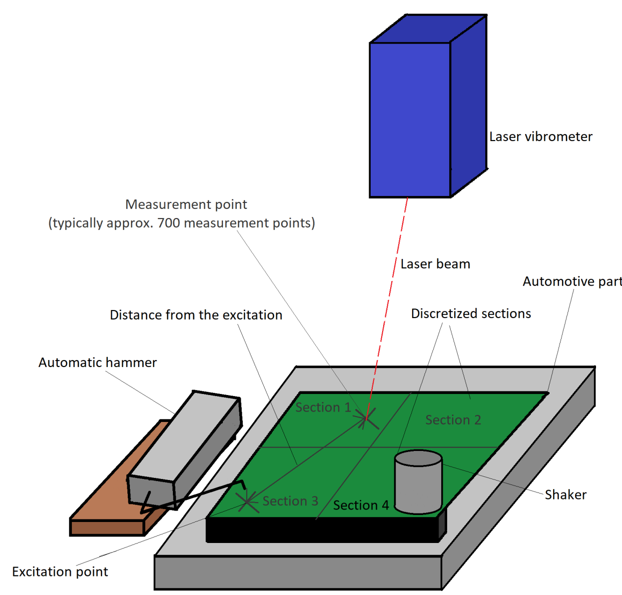


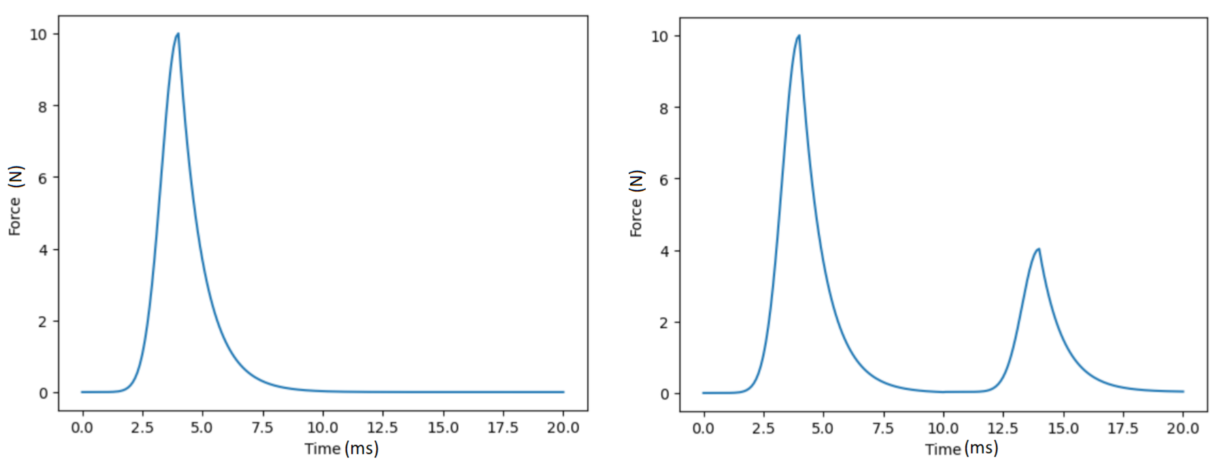
Figure 2: Schematic figure of the measurement setup

Figure 2 depicts the measurement setup. As the hammer strikes the component, a wave is generated. The laser vibrometer calculates from the temporal changes in the laser beam how fast the component vibrates at a given point at each moment in time. These temporal signals are later utilized to solve both classification to predict the label of the part of the measurement point, and regression tasks to estimate distance.

6 Categorization

There are four categorization aspects for the data.

- **Excitation shape**: There are various excitation shapes, including single, double, triple and multiple, among others.
- **Force applied**
- **Direction of the excitation**: This is a very simple one, because it has only two versions: +Z and -Z depending on the direction in which the hammer strikes in the coordinate system.
- **Location of the excitation**: This aspect serves the purpose of logically associating the estimated distances with the points of excitation.



(a) Single hit

(b) Double hit

Figure 3: Schematic figures of single and double hit

In Figure 3 we can see the schematic figures of single and double hit. The single and double hit phenomena (or alternatively, single and double impact) refer to scenarios where, in the case of a single hit, the body is struck by just one excitation, whereas in the case of a double hit, it is struck twice in a short period. This occurs, for example, in our measurement setup when the component rebounds and strikes the hammer again.

7 The prediction

7.1 Distance estimate by interpolation

The current main estimation method, interpolation, employs a nearest neighbor approach. This method assumes that closely located measurement points will yield similar time signals. It utilizes Euclidean distances between points in a matrix to assign values to new points. Initially, there was a notable improvement in accuracy, with average error dropping from 5 cm to 1 cm. Though some challenges have arisen, this progress indicates a promising new direction.

7.2 Triangulation

This method utilizes three excitation points, each estimating the distance to a given measurement point. The predicted measurement point is determined by the common intersection of the circles drawn from these excitation points. While theoretically sound, numerical errors and their accumulation can lead to deviations. To address this, a straightforward approach was adopted: if circles should intersect but do not due to distance estimate errors, the midpoint between the intersections of the segments connecting their centers with the circles will be the predicted point. This method ensures robustness by providing estimates even in cases where direct intersection is not feasible, initially discarded to avoid ambiguity.

7.3 Working with triggered signals

At this point, it should also be mentioned that we abandoned the idea of working with the whole time series, as we realized that the resting phase at the beginning contains a lot of information about the distance from the excitation point, but this is due to the fact that the measurements were always made in a predefined time window. In reality, the excitation will hit the sensor unexpectedly, so we have switched to working with triggered signals, which are generated by cutting off the portion of the signal at the beginning that does not reach a given level in absolute value (usually set to 0.005).

7.4 Discretization

This approach involves decomposing the component into clusters based on x-y coordinates. Future adaptations aim to discretize potential error sources into clusters. The program code resembles previous distance estimation methods but operates on cluster labels instead. A strong correlation between the number of clusters (varied between 3 and 8) and model accuracy was observed.

8 Models and their performance

8.1 Triangulation

Method	Min error	Max error	Mean error	Std	Number of estimations	Under 0.5 cm
Interpolation	0.0	0.38	0.003	0.03	174	172
Pol. regression	$4.58 \cdot 10^{-17}$	0.0016	0.0002	0.0004	542	542

Table 1: Performance of triangulation models

The minimum error, maximum error, mean error, and standard deviation (std) are all measured in meters. It is important to note that the second set of measurements is significantly superior to the first. This improvement stems from its application to automotive parts, which are smaller in scale compared to the metal sheet. As a result, all the data exist within a smaller range, contributing to a more precise assessment.

8.2 Cluster prediction

As it was mentioned before, a correlation was observed between the number of clusters and the accuracy score. However, it is evident that this model struggles to generalize across distinct excitation points and various automotive parts. This limitation arises from the diverse signal shapes generated by different excitation points and the considerably varied discretizations present in different automotive components.

9 Feature engineering

In this section, we will delve into the concept of feature engineering, a critical step in the development of effective machine learning models. Feature engineering involves the extraction, selection, and transformation of key attributes from raw data to create meaningful inputs for machine learning algorithms. Using relevant features, we can enhance the predictive power and efficiency of our models.
We will specifically compare feature-based models with those that rely on full- or partial signal shapes. While signal-based approaches directly utilize raw or minimally processed data, feature-based models leverage carefully crafted characteristics that summarize the underlying patterns in the data. This distinction is particularly important when dealing with complex or high-dimensional datasets, as feature engineering often helps reduce dimensionality, improve interpretability, and reduce computational requirements.
Furthermore, we will examine the effectiveness of feature-based models in various applications. These models often demonstrate superior performance, especially when the chosen features capture domain-specific knowledge or highlight critical trends within the data. By emphasizing feature-based approaches, we can achieve a balance between model accuracy and computational efficiency, making them a preferred choice in many scenarios.
Through this discussion, we aim to highlight the advantages and challenges of feature-based modeling and how it contributes to the broader landscape of data-driven problem-solving.

9.1 Features of the signal

Time signals can be described by a wide range of numerical properties, each capturing different aspects of their behavior. Among these features, we selected those that show the strongest correlation with distance, as they are likely to provide the most meaningful insights for our analysis. Let us take a closer look at these selected features.

- **Trigger position:** This refers to the point in time when the signal exceeds a certain threshold in absolute value (when its absolute velocity becomes greater). This threshold value was empirically selected and is set to 0.005 m/s. The correlation of this feature with distance is remarkably strong; however, it does not exhibit significant variance across signals, typically ranging between 3000 and 3070.
- **Trigger velocity:** This refers to the absolute velocity at the trigger position.
- **Decay position:** This feature is designed to describe where the signal ends. In other words, it identifies the position—or, if you prefer, the point in time—beyond which nothing relevant happens. Defining this feature accurately proved challenging, as a specific calculation rule often failed to produce reliable results under different measurement configurations.
- **Decay velocity:** This refers to the absolute velocity at the decay position.
- **Maximal velocity:** This feature apparently refers to the maximal velocity in absolute value. It shows a negative correlation with distance, as expected, although there are exceptions to this as well.
- **Delta:** This is the difference of decay- and trigger position. This feature is intended to describe the duration of the signal.
- **Total displacement:** Since we have a velocity vs. time signals, total displacement can be obtained as an integral. This feature represents the value of the integral taken over the section between the trigger- and the decay position.

9.2 Envelope curve

The determination of the decay position often faced challenges and could not be generalized across multiple measurement configurations. For example, it could not be defined in the same way as the trigger position, such as when the signal's absolute value drops below a certain threshold. Therefore, we adopted a new method: we fitted an envelope to the signal and examined when its slope falls below a specific threshold. At that point, we considered it the limit up to which the signal should be analyzed.

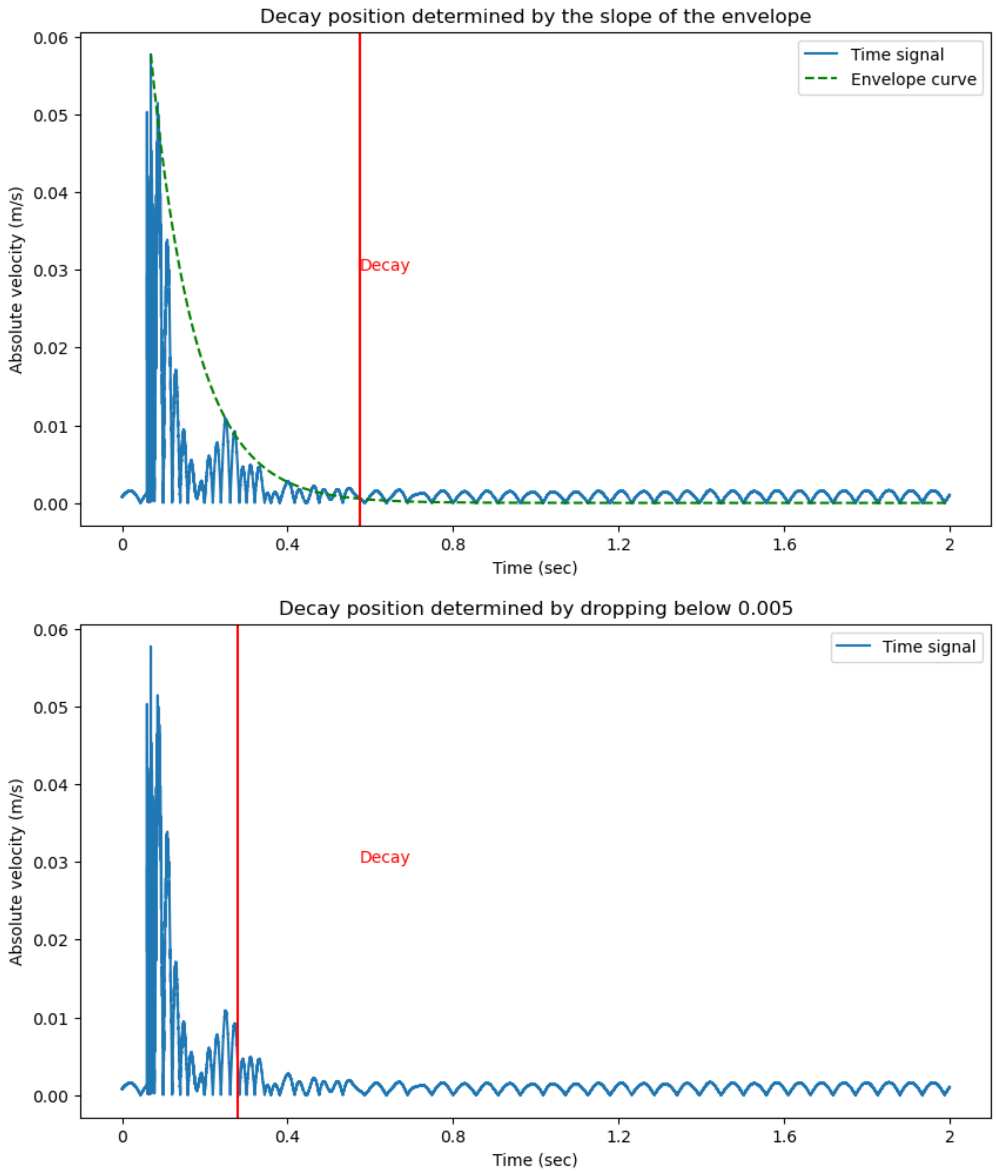
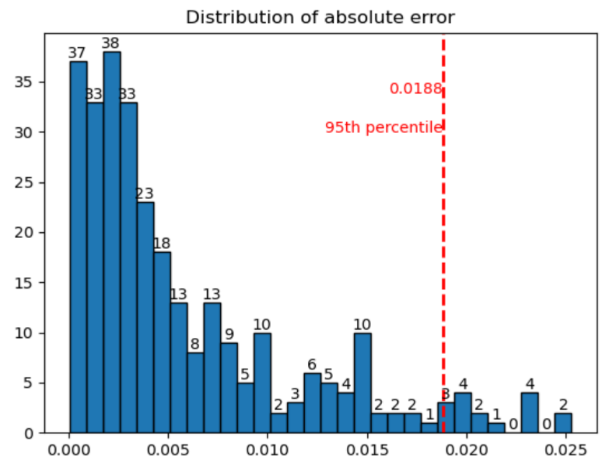


Figure 4: Comparison of decay positions

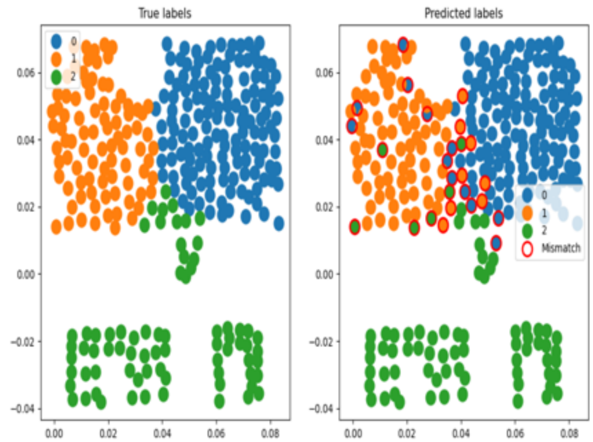
In Figure 4 we can see the difference between the descriptiveness of decay positions defined in different ways. It is can be seen that with the first method, where the decay position is determined by the slope of the envelope, the section of the signal to be analyzed can be described much more effectively.

10 Decision trees

Decision trees are among the simplest machine learning models, as they rely solely on less-than, greater-than, or equal-to comparisons. They can be used for both classification and regression tasks, which is exactly what we did. The extracted features were fed into a decision tree classifier and regressor, and we evaluated how accurately they performed.



(a) Distribution of absolute error



(b) Classification accuracy (accuracy score: 0.913)

Figure 5: Decision tree regressor and classifier performances

It can be observed that in the case of regression, the error roughly follows an exponential distribution. We also conducted further experiments to test the models, specifically by varying the number of features used.

- **Three features:** We first experimented with three features: the trigger position, the decay position, and the delta.
- **Four features:** Next, we experimented with four features: trigger velocity, maximal velocity, decay velocity, and delta.
- **All features:** Finally, we also tested the case where we used all the features along with the total displacement.

The evaluations show that the model based on the fewest features generally performed the best. However, this result could be misleading, as the model with three features used the trigger position, which is not a valid piece of information. In reality, we will not know when a measurement begins, as the sensor will continuously or record signals at specific intervals.

11 Convolutional neural networks

We also examined how distance estimation works with convolutional neural networks. To do this, we identified the trigger position of the signals and then took 10,000 samples from each signal starting from that point. These samples were used as input for a convolutional neural network, which consisted of three convolutional layers and one linear layer.

11.1 Hyperparameter tuning

Here, it was possible to adjust the kernel sizes of the different convolutional layers, the output sizes, the size of the output of the linear layer, the learning rate, the number of epochs and the search mode (grid search and random search). The evaluations show that the highest accuracy is achieved when the size of the convolutional layers increases, the learning rate is small (e.g. 0.01), and the number of epochs is not too high (e.g. 10). Mean absolute error was about 0.01 m.

12 Conclusion and next steps

In this project, we explored the challenge of localizing the source of noise in automotive components using machine learning methods. Through both regression and classification approaches, we demonstrated the potential of models to predict measurement points with varying degrees of accuracy. While methods such as triangulation and interpolation proved effective in estimating distances under certain conditions, challenges arose when attempting to generalize across different excitation forces, shapes, and components.

Feature engineering played a crucial role in enhancing model performance, as carefully selected features provided meaningful insights into the underlying patterns of the time signals. Despite the progress, limitations in generalization, particularly for diverse automotive parts, highlight the need for further research. Future work could focus on self-supervised learning techniques and advanced embeddings to improve model robustness and adaptability. Overall, the results of this study offer valuable contributions to the application of machine learning in the automotive industry, paving the way for more accurate and efficient defect detection systems.

In the near future, we will also conduct 3D measurements, using a shaker for excitation instead of the automatic hammer. These data can later be used to build multi-output models capable of identifying not only the type of component but also the type of rattling or, potentially, multiple rattling scenarios.