Individual Report

1. Problem Understanding

I was in charge of reading and breaking down Wang et al.'s paper, which used a NARX model for defect detection. After going through it, I realized their version of NARX doesn't actually use temperature as an exogenous input, it just slides a window over the signal and feeds the output back in, making it basically a standard RNN. I explained this to the group, and we decided to focus on LSTM models in future data modelling. This helped shift the team's direction early on and saved us time trying out a model that was not suited to our data.

2. Wave analysis

I focused on analyzing the guided wave data over the year to spot differences between pristine and defected signals. I split the real and imaginary parts and plotted them to compare visually. For the trace subtraction, I directly subtracted two consecutive rows to see how the wave changed over time. When there was no defect, the difference looked like a straight line, but over longer time gaps, small fluctuations appeared due to environmental noise. When a defect was introduced between two signals, I noticed a clear spike in part of the wave. This matched well with the defect labels we had and showed that the subtraction method works and is sensitive to damage, even with environmental changes. The full analysis is in "*pair3_analysis.ipynb"* file.

3. Temperature correlation test

I was also the one who tested the correlation between temperature and signal changes. I downloaded temperature data for all of 2012 (each month was a separate CSV file from the GitHub weather data) and matched it with the wave signals. After plotting and analyzing the results, I found a clear correlation and temperature shifts were affecting the signal patterns. This led the group to change direction and start comparing models trained with and without temperature as an input.

4. Model testing

I also tried to replicate the NARX method from Wang et al. for defect detection, but their approach was built entirely in MATLAB, which was out of our scope. Their model slides a window along the length of the wave (one signal) to predict the next segment, using batches of 20 signals. I tested a similar setup in TensorFlow, but the results were way off, mainly because TensorFlow treats each signal as a feature, not a timestep, so the input shape had to be fixed and did not work dynamically. Based on this, I suggested we change direction and treat each wave (each row) as a timestep instead, like in an ARIMA setup, so we could run LSTM over time using batches of rows. This helped restructure how we handled temporal modeling moving forward. I also tested two approaches using LSTM: one as a classifier with manually labeled data, and the other using a sliding window to reconstruct the signal. The classifier model completely failed as it predicted everything wrong, which showed that the wave shape didn't have enough clear features for LSTM to tell pristine and defect signals apart directly. The sliding window model, on the other hand, worked much better. It had low reconstruction error on pristine data and higher error on defected ones. Based on this, I shifted the whole group's modeling focus toward the sliding window method for better results.

5. Spectral Analysis and modelling

To get a better understanding of the signals, I transformed the wave data from the time domain to the frequency domain using Fast Fourier Transform. I manually labeled the data into defect and pristine and found that pristine signals in 2021 had much higher average amplitude than defected ones. I tested this by building a random forest model, which gave near-perfect results, which showed that the spectral features could clearly separate the two. But I noticed a problem: the overall signal magnitude in earlier years (2012–2015) was much higher, which could cause bias. So, I focused on a short, consistent window of data (the last 34 samples). When plotting the spectra, I saw that each defect type had a distinct pattern, even between different defect sizes. I trained a decision tree, and it achieved 100% accuracy. By checking the feature splits, I found the tree was using a specific frequency where the pristine data was totally different, making it a perfect threshold for classification. I also tested multi-class labeling for different defect sizes and still got high accuracy, with only slight confusion between similar defects. This showed we could define a clear threshold between healthy and damaged signals. However, due to limited data and the method only working in a short time range, this section was left out of the final report. Here is the full analysis of the spectral:

- **Methodology:**

For spectral analysis, the time-domain signals collected from sensor pairs were transformed into the frequency domain using the Fast Fourier Transform (FFT). This transformation highlights frequency characteristics that are often more indicative of structural defects compared to time-domain signals, as defects typically alter specific frequency components due to scattering and resonance effects. The last 34 FFT-transformed signals, consisting of 4 undamaged and 30 defected cases, were selected for classification. A Decision Tree Classifier was trained directly on these frequency-domain representations. The dataset was split into training and testing sets using stratified sampling to maintain the proportion of defect types across both sets.

Two classification tasks were performed: a binary classification distinguishing between undamaged and defected states, and a multiclass classification identifying specific defect types. The model's performance was evaluated using confusion matrices, and the decision process was visualized to interpret how frequency features contributed to defect detection.

- **Result and Discussion:**

The spectral analysis highlights how both the progression of defect severity and their position relative to Sensor Pair 3 influence the amplitude response in the frequency domain. As expected, the undamaged signals display consistent behavior, with a low mean amplitude of 0.30 and a maximum amplitude of 13.38. This reflects the stable propagation of guided waves through a pristine structure, where minimal scattering or reflection occurs.

Defect Type	**Mean Amplitude**	**Variance**	**Max Amplitude**
Undamaged	0.299397	1.829277	13.376265
Part depth 3.5mm	1.334966	27.562837	40.401338
3.5mm through hole	0.306413	1.895594	14.149841
7mm through hole	0.753560	10.292387	28.235318
Part depth 7mm (0.6m, 0.4m)	0.301079	1.846712	13.896287
Through hole 7mm (0.6m, 0.4m)	0.580971	6.213011	23.647424
6.5mm through hole in weld	0.336308	2.265248	16.532792

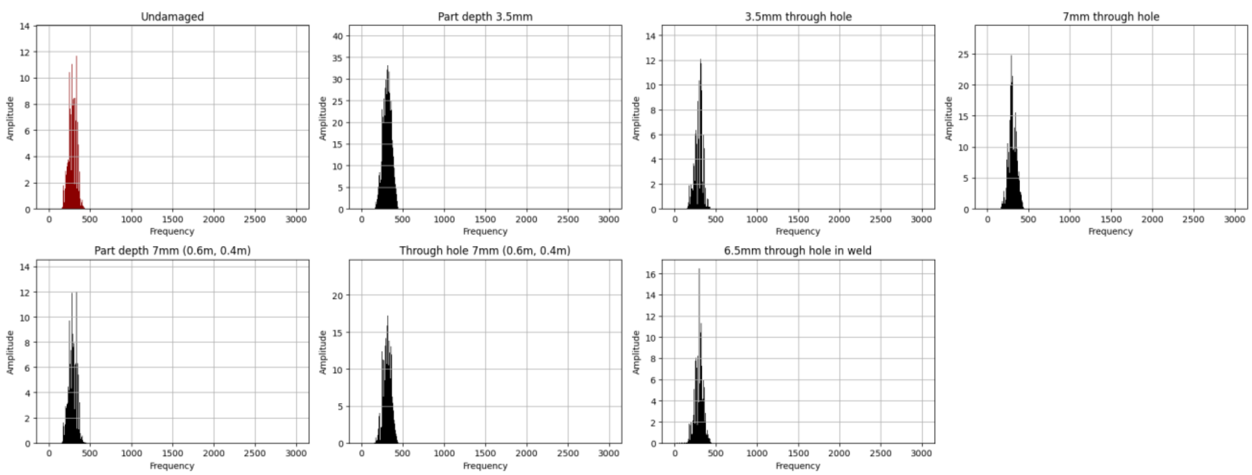

Figure 1 Frequency domain of different defect states

The first three defects in the dataset correspond to the same hole being progressively deepened and widened over time. These include the part depth 3.5mm, followed by a 3.5mm through hole, and finally a 7mm through hole. Although these defects are located near the center of the structure, they are not directly aligned with Sensor Pair 3. Interestingly, the part depth 3.5mm shows the highest amplitude response, with a mean of 1.33 and a peak reaching 40.40. This suggests that partial depth defects can cause irregular reflections and complex scattering patterns, leading to greater energy concentration at certain frequencies. As the defect becomes a complete through hole, the amplitude decreases but remains elevated. The 7mm through hole, due to its larger size, shows a stronger spectral response than the smaller 3.5mm through hole, with a mean amplitude of 0.75 and a maximum of 28.24.

In comparison, the defects located at 0.6m, 0.4m are positioned almost directly along the path between Sensor 1 and Sensor 4. The part depth 7mm defect at this location shows minimal spectral disturbance, with amplitude values similar to undamaged conditions. However, once this defect is drilled through, the amplitude increases noticeably, reaching a mean of 0.58 and a maximum of 23.65. Although this is a clear increase, it remains lower than the response observed for the centrally located 7mm through hole.

While undamaged signals maintain low and consistent amplitudes across frequencies, defects cause noticeable amplitude fluctuation. Overall, the presence of a defect disrupts the uniform energy distribution observed in pristine conditions, making spectral analysis an effective approach for distinguishing between healthy and damaged structures.

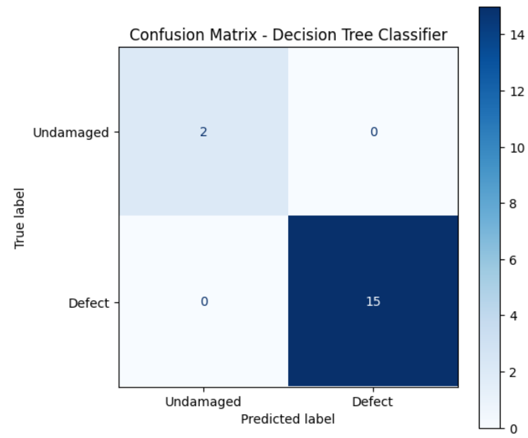

Figure 2 Confusion matrix of binary class classification Decision tree

After training the Decision Tree classifier on the frequency-domain data, the model's performance was evaluated using a confusion matrix. The confusion matrix shows perfect classification results, with all 17 samples correctly identified. Specifically, both undamaged signals (2 samples) and all defected signals (15 samples) were classified with 100% accuracy. This indicates that the spectral features extracted from the FFT provided a clear distinction between healthy and damaged structures, allowing the decision tree to separate the two classes without error.

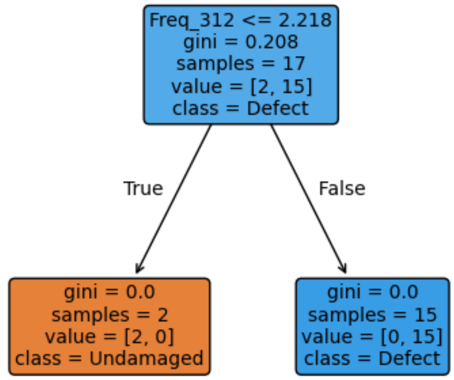

Figure 3 Tree split

To understand how the model achieved this result, the decision tree structure was examined. The tree revealed a very simple decision rule: it split the data based on the amplitude at a specific frequency point, labeled as Freq_312. The threshold was set at approximately 2.218, meaning that if the amplitude at this frequency was less than or equal to this value, the signal was classified as undamaged; otherwise, it was classified as defected. This demonstrates that a single frequency component carried enough discriminatory power to differentiate between the two classes effectively.

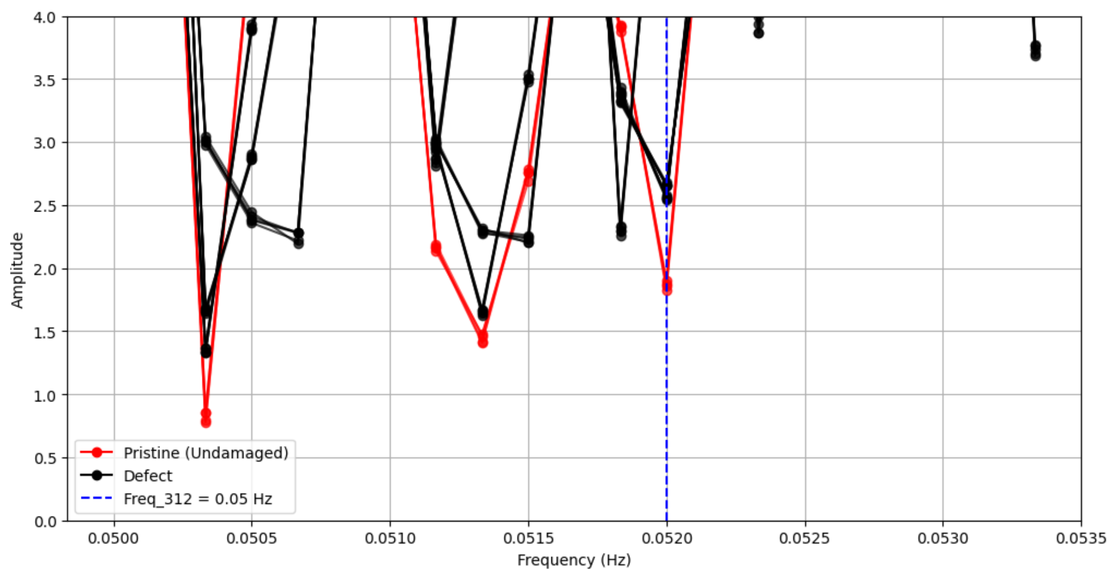

Figure 4 Zooming in the Frequency Domain

To validate this decision rule, a focused analysis was conducted by plotting the amplitudes around the critical frequency point. The visualization confirmed that undamaged signals consistently exhibited lower amplitudes at Freq_312, while defected signals showed significantly higher values. This clear separation supports the decision tree's choice of threshold and highlights how specific frequency responses are strongly indicative of structural defects.

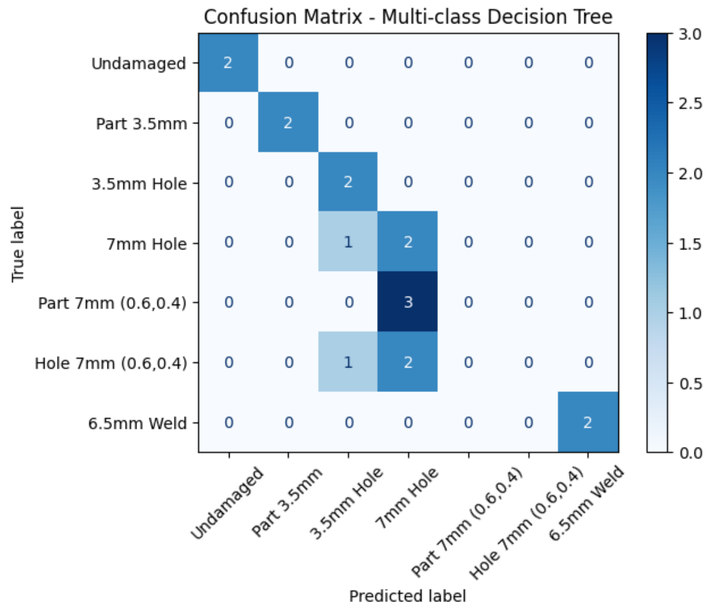

Figure 5 Confusion matrix of multiclass classification Decision tree

I extended the spectral analysis by building a multi-class Decision Tree classifier to not only detect defects but also distinguish between different defect types. The confusion matrix shows that the model performed really well, with most classes being classified correctly. It identified all undamaged and 6.5mm weld cases with 100% accuracy. It also did well on 3.5mm Hole**,** Part 3.5mm, and Part 7mm (0.6, 0.4). The only confusion happened between some 7mm Hole and Hole 7mm (0.6, 0.4) samples, which are very similar in defect type and size, so it's understandable the model had trouble telling them apart. Overall, this shows that the frequency features carry enough detail to not only detect damage but also differentiate between different structural defect types with high accuracy.