Cavitation Detection in pumps using Machine learning

Markus, Valtiner
Supervisor

Yu, Kinoshita
Student ID:01623806

Thomas, Gärtner
Co-Supervisor

TU Wien

e1623806@student.tuwien.ac.at

1 Abstract

This project was carried out in cooperation with Senzoro GmbH.

Pumps maintained in industrial systems may get into a state of cavitation, causing vibration and shock waves that can damage the pump components.

This project aims to develop a machine learning model that can classify cavitation in pumps using an ultrasound sensor that measures acoustic signals induced by the bursting bubbles.

The model will be trained on data that was collected through several experiments designed to replicate the state of cavitation. The performance of the ML model will be evaluated using standard metrics such as accuracy, precision, and especially recall.

The findings in this report suggest that a ML model trained on data obtained through experiments could indeed detect unseen cavitation data originating from industrial pumps. Although, it needs to be mentioned that the test set size was rather small and it remains to be seen if the model will perform well on a larger, more diverse set of data.

Contents

2 Data acquisition

Data was collected by conducting different experiments using an ultrasound sensor provided by Senzoro GmbH. The goal of the experiments was to obtain signals from collapsing bubbles. The following experiments were designed with the help of Professor Markus Valtiner and Matteo Olgiati.

2.1 Experiment 1: Oil pump cavitation

A rotary vane pump with oil as a lubricant is used in this experiment. By introducing small amounts of water into the pump we attempt to create cavitation. Out of the experiments, this setup is by far the closest to cavitation in a real world industrial pump.

Figure 1: rotary vane pump

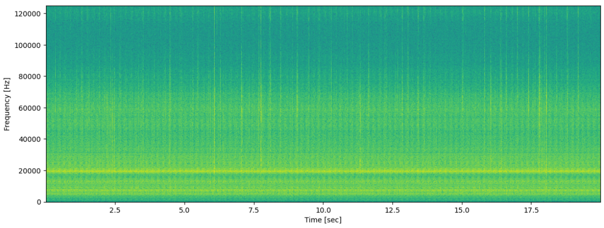

Figure 2: rotary vane pump assumed cavitation

2.2 Experiment 2: Degas saturated solution by pulling vacuum

In this method, a desiccator with hydrophobic walls is filled with the saturated solution (Argon (Ar) or Carbon dioxide (CO2)) The water will be degassed by pulling a vacuum. This forces bubbles in the desiccator. The ultrasound sensor is placed in various positions to obtain data in various settings. The bubbling on the CO2 induced water is much more reactive and thus deliver stronger ultrasound signals while the argon induced water deliver very weak signals.

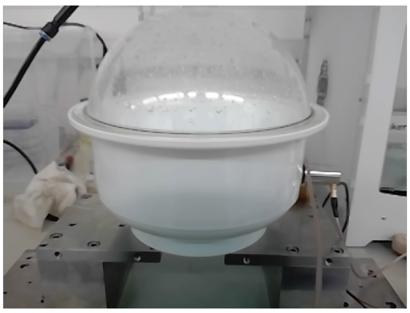

Figure 3: Degas saturated solution by pulling vacuum

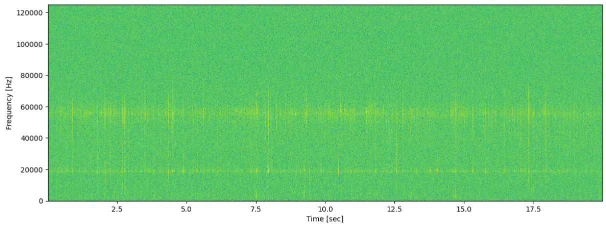

Figure 4: vacuum degassing bubbling

2.3 Experiment 3: Teflon beaker

An argon emitting pipe was placed into a Teflon beaker filled with water, to create bubbling on the surface. Measurements are done using different pressure strength of the argon pipe.

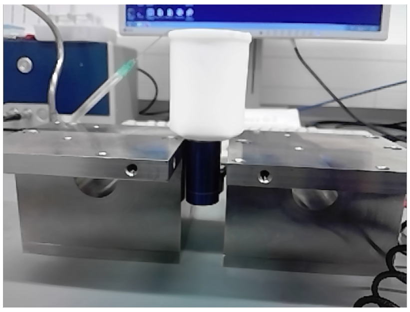

Figure 5: Teflon beaker

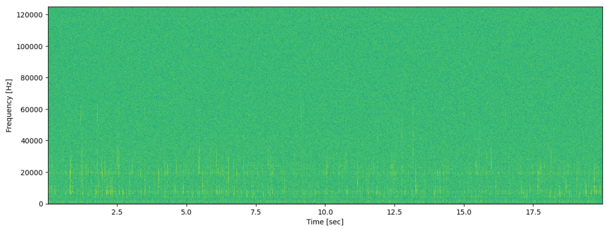

Figure 6: Teflon beaker bubbling

2.4 Experiment 4: Electrochemical water splitting

In this method, we decompose water into oxygen and hydrogen by electrolysis using electricity. For this, a stainless steel immersed in a water-based liquid (3% HNO_3) will be charged with electricity. Bubbles are formed when producing gaseous hydrogen.

	Samples in seconds
Experiment 1	620
Experiment 2	1280
Experiment 3	720
Experiment 4	1800

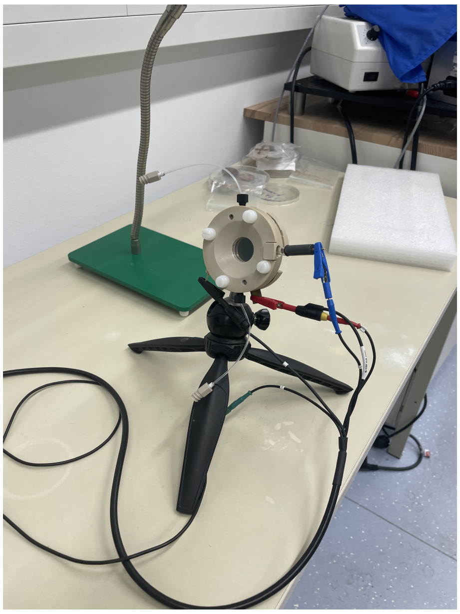

Figure 7: Electrochemical water splitting

Three out of four experiments do not replicate cavitation, instead they measure the signals of collapsing bubbles. One minor goal of this research is also to see if collapsing bubbles can be associated with cavitation signals.

3 Methodology

3.1 Model pipeline

First, the collected data is split into a training and testing set with a ratio of 80:20. The data is divided into equal sized smaller samples and Fourier transformed into the frequency domain. Next, the features are extracted from the transformed data and the corresponding class labels ("cavitation" or "no cavitation") are added. The model will be trained using the train set, which will be optimized on recall using 5-fold cross-validation and the grid search algorithm if necessary. Recall (True Positive Rate) is chosen as metric because it is better to predict falsely cavitation than not to avoid potential cavitation damage on the pumps. Different window sizes of the signals will also be evaluated.

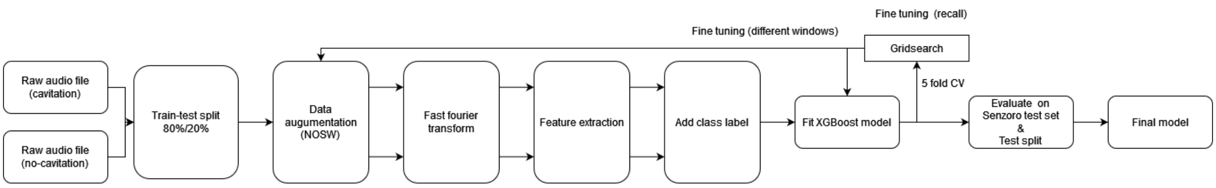

Figure 8: XGBoost Model pipeline

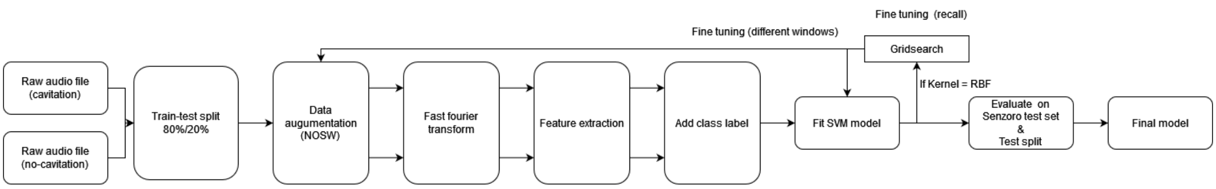

Figure 9: SVM Model pipeline

3.2 Data Augmentation

The data is split into equal sized smaller samples using a method based on Non - Overlapping Sliding Window (NOSW). Most of the original data samples are 20 seconds long. This is split into smaller lower second samples. When doing so, one should be careful on deciding the second value the data should be divided into as features can be lost when a too low value is chosen.

3.3 Data provided by Senzoro GmbH

Two audio sample of possible cavitation and three samples of non-cavitation data, measured on an industrial pump was provided by Senzoro for testing purposes.

The data collected in the experiment are also split into train and test set for evaluation purposes.

3.4 Feature Extraction

To deal with non-essential information statistical feature extractions is introduced. [1] This consists of Central Trend Statistics[1], such as mean, median, low quartile and upper quartile, the Dispersion Degree Statistics[1] such as minimum, maximum, inter quartile range standard deviation, root mean square and square root amplitude, and finally the distribution shape statistics [1] such as kurtosis, skewness, shape factor, clearance factor and crest factor.

The features are calculated for each Fourier transformed sample.

4 Model

Models are trained with three different training sets. Training set one consists of experiment 1 (rotary vane pump) data. The second set consists of experiment 1 and experiment 2 data (desgas in vacuum). Lastly, the third set consists of experiment 1, experiment 2 and experiment 3 (Teflon beaker). Data from the water splitting experiment (4) is not used for the cavitation models as the measured signals are vastly different in comparison to the signals of the other experiments. The signals mostly consists of background noise and very weak bubbling signals.

4.1 XGBoost

EXtreme Gradient Boosting (XGBoost) is selected as the first approach. XGBoost is a supervised decision tree ensemble learning algorithm. It creates decision trees in sequential form and builds a strong classifier from the multiple "weak" classifiers (decisions trees). Using a gradient descent optimization the loss is minimized by adding weak models. The loss function used in XGBoost is the 2nd order Taylor's expansion.

	Experiment 1 (rotary vane pump)	**Experiment 2 (Vacuum degas)**	**Experiment 3 (Teflon beaker)**	**Experiment 4 (Water splitting)**
model 1	✓	x	x	x
model 2	✓	✓	x	x
model 3	✓	✓	✓	x
model 4	x	x	x	✓

4.2 Support Vector Machine

For the second model, the Support Vector Machine (SVM) is introduced. SVMs separate the class by constructing linear decision boundaries based on hyperplanes. The main difference to a Perceptron is that the SVMs transform the feature space into a higher-dimensional space where the data is linearly separable. Transforming the features into higher dimensions would be computationally costly (see figure 10). This is avoided with the help of Kernels. Kernels can represent the data through a set of pairwise similarity comparisons between the original observations.

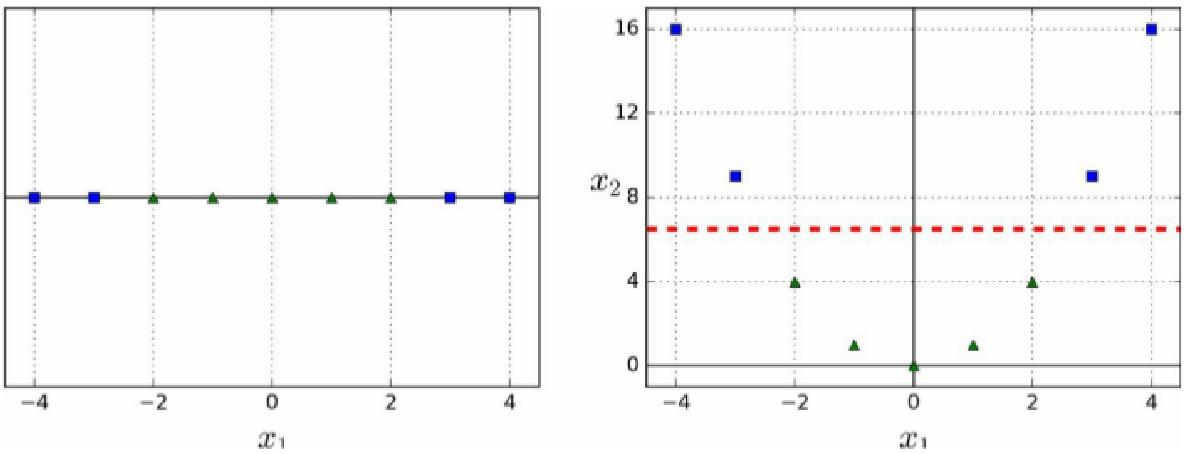

Figure 10: Adding features to make a dataset linearly seperable [2]

Models with polynomial, sigmoid and radial basis function (rbf) kernel are trained. The polynomial kernel is used to map features into a different space, it is one of the kernel functions used when the data is non-linear separable.

The rbf is a similarity function that calculates the Euclidean distance between two points. This kernel has two hyperparameters sigma and gamma, which is tuned using grid search.

The sigmoid kernel is widely used in neural networks. SVM with a sigmoid kernel is the equivalent of a 2-layer perceptron. [3]

5 Results

The models are evaluated on several unseen test sets. They consist of different experiment combination and data from industrial pumps provided by senzoro GmbH.

5.1 Metrics

Accuracy

$$Accuracy = \frac{truepositves + truenegatives}{total} \tag{1}$$

Actual and predicted classes are compared with each other and each match counts as correctly predicted. Accuracy can be very misleading when the data set is imbalanced.

Precision

$$Precision = \frac{truepositves}{truepositves + falspositives} \tag{2}$$

Precision depicts the ratio of how many of the positive classes are correctly classified.

Recall

$$Recall = \frac{truepositves}{truepositives + falsenegatives} \tag{3}$$

Recall is similar to precision but the focus lies on false negatives. This metric is important if it is better to predict a certain class than not (e.g. it is better to predict that a person is infected with a virus than not).

F1

$$F1 = 2x\frac{precisionxrecall}{precision + recall} \tag{4}$$

F1 measure combines precision and recall metrics.

5.2 XGBoost Evaluation

5.2.1 Evaluation of XGBoost Model on test dataset split:

Table 1 depicts the evaluation results of the XGBoost model trained on 1 second window sizes, evaluated on own test dataset split including data from the rotary vane pump, degas vacuum and Teflon beaker experiment.

Models	**Accuracy**	**Precision**	**Recall**	**F1**
1	0.259	1	0.166	0.2857
1 + 2	1	1	1	1
1 + 2 +3	1	1	1	1

Table 1: Evaluation of XGBoost Model on test dataset split

The model which was trained using only the rotary vane pump does not perform well on the test set. It is able to classify unseen cavitation data originating from the experiment using the rotary vane pump. Any other data is classified as non-cavitation. Interestingly, when looking into the model which was trained using the vacuum degassing and rotary vane pump experiment, the model is able to classify every data point correctly, including data from the Teflon beaker experiment which should be completely new to the classification model. Similar behaviour can be observed on other window sizes with the metric varying only by a negligible amount.

This behaviour does not change when evaluating the model trained with experiment 1, 2 and 3 on the test set only containing data from rotary vane pump (1) and degassing vacuum (2) experiments.

5.2.2 Evaluation on Senzoro GmbH test data:

Models	**Accuracy**	**Precision**	**Recall**	**F1**
1	0.4	0.4	1	0.57
1 + 2	0.38	0.387	0.95	0.55
1 + 2 +3	0.38	0.387	0.95	0.55

Table 2: Evaluation of XGBoost Model on Senzoro GmbH test data

At first sight the results look quite good considering the models has never seen an industrial grade pump but the Precision being 0.4 implies that all data has been classified as 1 since the test data is imbalanced with two cavitation files and 3 non-cavitation files. Further, when training the model with the additional experiment 2 and 3 data it returns similar results, classifying all non-cavitation data as cavitation and miss-classifying few cavitation data as non-cavitation.

All of the above does not improve when applying different window sizes between 0.25 and 5 seconds.

5.2.3 Feature importance

Feature importance in percentage for the best performing model (rotary vane pump & vacuum degas) on the Senzoro test set can be found in table 3.

Inter quartile range and shape factor are the two most important features with over 20%. Standard deviation, upper quartile and maximum are also important with over 10%. The remaining features havae significant lower percentage values ranging from 3.7 % to 0.

Features	**Importance**
IQR	0.245788
Shape factor	0.232523
Standard deviation	0.175498
Upper quartile	0.148632
Maximum	0.131642
Lower quartile	0.037502
Minimum	0.011761
Skewness	0.009944
Kurtosis	0.004148
Crest factor	0.001539
Clearance factor	0.001024
Mean	0.000000
Median	0.000000
Root mean square	0.000000
Square root amplitude	0.000000

Table 3: Feature importance

5.3 SVM Evaluation

5.3.1 Evaluation of SVM Model on test dataset split:

Table 4 depicts the evaluation results of the SVM model trained on 1 second window sizes, evaluated on the experiment test data split including data from the rotary vane pump, degas vacuum and Teflon beaker experiment.

Similar to the XGBoost results, the models only trained on the rotary vane pump does not perform well on the test set. It classifies the majority as "no-cavitation". Also similar to the XGBoost results models trained on vacuum degassing, teflon beaker and rotary vane pump experiment performs good on the test set. This behaviour can be observed on all models with different kernels except for the RBF kernel model. All RBF kernel models with different training data result in equal performance on the test set.

Models	Accuracy	Precision	Recall	F1
Poly (1)	0.212	0.936	0.122	0.217
Poly (1+2)	0.888	0.888	1.0	0.941
Poly (1+2+3)	0.888	0.888	1.0	0.941
RBF (1)	0.888	0.888	1.0	0.941
RBF (1+2)	0.888	0.888	1.0	0.941
RBF (1+2+3)	0.888	0.888	1.0	0.941
Sigmoid (1)	0.209	0.784	0.152	0.254
Sigmoid (1+2)	0.861	0.885	0.968	0.925
Sigmoid (1+2+3)	0.851	0.884	0.958	0.92

Table 4: Evaluation of SVM Model on test data

5.3.2 Evaluation on Senzoro GmbH test data:

The Support Vector Machine models performs significantly better than the XGBoost model on the Senzoro test data set. The model with the polynomial kernel achieves an accuracy of 0.8 and is also able to correctly classify all of the "cavitation" data, having a recall of 1.0 and precision of 0.66. On the other hand, the RBF model performs quite poorly with an accuracy of 0.4, which is caused by predicting only "cavitation" label. Thus, the precision is also at 0.4 and recall at 1. Lastly, the model with the sigmoid kernel is able to predict all data correctly, regardless of class, even though the model has never seen cavitation data originating from industrial pumps before. It is able to achieve an accuracy of 1.0 , recall of 1.0 and precision of 1.0 . The model delivers similar results when the model is trained on different window sizes between 2 and 0.25 seconds and evaluated in the corresponding window sizes.

At 0.25 seconds the accuracy only drops to 0.92. Although, it needs to be mentioned that one should be careful with reducing the window sizes too much since too low window sizes can lead to the loss of information.

Models	Accuracy	Precision	Recall	F1
Poly (1)	0.38	0.387	0.95	0.55
Poly (1+2)	0.8	0.66	1.0	0.8
Poly (1+2+3)	0.8	0.66	1.0	0.8
RBF (1)	0.4	0.4	1.0	0.57
RBF (1+2)	0.4	0.4	1.0	0.57
RBF (1+2+3)	0.4	0.4	1.0	0.57
Sigmoid (1)	1.0	1.0	1.0	1.0
Sigmoid (1+2)	1.0	1.0	1.0	1.0
Sigmoid (1+2+3)	1.0	1.0	1.0	1.0

Table 5: Evaluation of SVM Model on Senzoro GmbH test data

5.4 Implication & Limitation

5.4.1 XGBoost models

From the evaluation we can tell that models trained with vacuum degassing and Teflon beaker can seemingly reliably detect each others data. We can especially see that when it was able to classify the Teflon beaker experiment data even tough the model was never trained with the Teflon beaker experiment data.

Further, no XGBoost model was able to classify the industrial pump data well. It may have been able to correctly classify the cavitation labels but it was not able to classify non-cavitation correctly. Several reason could explain these behaviors. First, data gathered in the experiments completely lacked the noise from industrial pumps and machines except experiment 1 to some extend. Experiment 2 and 3 had a very calm environment, except for the bubbling itself. Also, bubbles bursting in the experiments may have not been as strong as in industrial pumps due to the lack of pressure. Even in the rotary vane experiment which is most similar to industrial pumps, it was noticeable with the human ear that the intensity of the bubbles bursting were not similar as to cavitation examples that can be found online. Similarly, shock waves generated in industrial pumps when cavitation is present may be vastly different than those created in the experiments.

5.4.2 SVM models

In comparison to the XGBoost model, classifies test data from the same experiment slightly less well. But the SVM models are able to classify cavitation data from industrial pumps surprisingly good. The model with the polynomial kernel yields a very promising result, and the models with the sigmoid kernel are able to almost perfectly predict cavitation regardless of window size, as long as the window size does not reduce the features too much. It is also surprising that the sigmoid model trained with only the rotary vane pump data is able to predict cavitation almost flawlessly.

However, the major drawback of these models is that it is difficult to extract the feature importance, as the feature space is transformed into different dimensions through the use of kernels, making it challenging to identify specific cavitation features in the ultrasound data.

5.4.3 Limited test data

Although the SVM model performs well, the Senzoro test data set is relatively small, and therefore it cannot be definitively stated that the model can accurately predict cavitation in pumps. To evaluate a model that can reliably detect cavitation in any environment, a larger, sufficient amount of testing data should be collected and used for verification and evaluation.

5.4.4 Limited experiments & improvements

To create a model that can reliably classify cavitation in any pump, incorporating data of both cavitation and non-cavitation from a variety of different pumps can further improve the model's accuracy. This approach can help avoid overfitting the model to a specific pump, as noise levels can vary among pumps. As seen in our experiment, the non-cavitation data from the rotary vane pump appeared distinct from the data from the industrial pumps provided by Senzoro.

Additionally, while the vacuum degassing and Teflon beaker experiments generated signals of bursting bubbles, they did not recreate actual cavitation, which may have contributed to poor performance by some models. The rotary vane experiment was the only experiment that somewhat replicated an industrial pump.

5.5 Bubble formation on electrochemical water splitting

As a secondary task, a machine learning model was trained to detect bubble formation (bursts) in an electrochemical water splitting process. The available 30-minute measurement data contained only 12 seconds of signals indicating bubble bursts (as seen in Figure 6).

The XGBoost model and SVM models with different kernels were trained on 80% of the available data and evaluated on three different test datasets to assess their ability to correctly identify data without bubbling signals. To evaluate the model's performance with such a small amount of data, three different test split methods were used: a "balanced" dataset with 50% bubbling and 50% non-bubbling data (total of 6 samples), an "imbalanced" dataset with 25% bubbling and 75% non-bubbling data (total of 12 samples), and an "extremely imbalanced" dataset with 10% bubbling and 90% non-bubbling data (total of 37 samples).

SVM:
Even though the SVM models perform well on the cavitation task, they don't perform well when predicting bubbles in the electrochemical experiments. The polynomial kernel model predicts every entry as "non-bubbling." The RBF and sigmoid kernels perform better on the balanced data set, but their metrics get much worse when evaluating the imbalanced and extreme imbalanced test sets.

XG Boost:
The XGBoost model performs better than the SVM model. When evaluated on the balanced data set, only one entry was misclassified. The imbalanced testing set also resulted in one misclassified value, and in the extreme imbalanced data set, five values were misclassified. It can be said that the XGBoost model's classification of this particular data was quite consistent, and with more data, it should get even better.

Future research
The next step in this research should be to gather more data from the electrochemical water splitting experiment, and to conduct the experiment with different types of metal. This will increase the amount of data available for training and evaluation, and ensure that the model can detect bubble bursts in a consistent manner across different electrochemical setups.

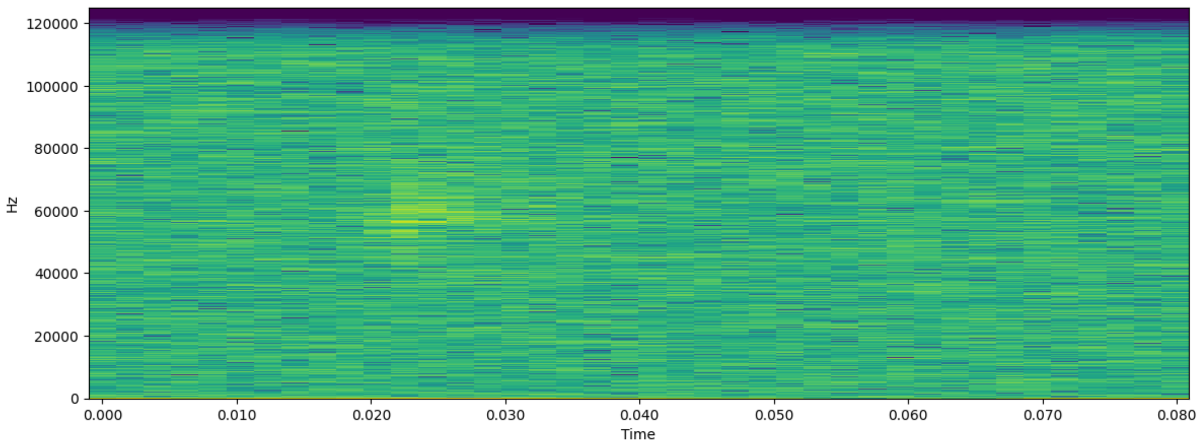

Figure 11: Bubble bursting in experiment 4

	Accuracy	Precision	Recall	F1
model	0.66	0.66	0.66	0.66
imbalanced	0.75	1	0.857	0.85
superimbalanced	0.864	0.375	1	0.545

Table 6: Experiment 4 XGBoost

Models	Accuracy	Precision	Recall	F1
Poly (balanced)	0.5	0	0	0
Poly (imbalanced)	0.58	0	0	0
Poly (superimbalanced)	0.83	0	0	0
RBF (balanced)	0.66	0.66	0.66	0.66
RBF (imbalanced)	0.58	0.33	0.66	0.44
RBF (superimbalanced)	0.56	0.11	0.66	0.2
Sigmoid (balanced)	0.83	1	0.66	0.8
Sigmoid (imbalanced)	0.33	0.2	0.66	0.3
Sigmoid (superimbalanced)	0.37	0.08	0.66	0.14

Table 7: Experiment 4 SVM

5.6 Comparison to state-of-the-art

To the best of our knowledge, there have been no publicly available efforts to train a model that can effectively classify cavitation in any type of pump, regardless of the conditions. The study "An acoustic signal cavitation detection framework based on XGBoost with adaptive selection feature engineering"[1] deals with a similar problem which trains a XGBoost model using experimental data from a test rack and evaluate it on valve acoustic signal test set provided by SAMSON AG [1]. The final model was able to achieve an accuracy of 0.9356. Although, the methodology of this paper may not be the same, it does suggest that classifying cavitation can work. It would be interesting to see, how the SVM model would perform on the same test set used in the study [1].

6 Conclusion & future research

In this project, XGBoost and SVM models were trained to predict the presence or absence of cavitation using ultrasound data. The XGBoost model performs rather poorly when predicting cavitation on unseen data from industrial pumps, while the Support Vector Machine models with the sigmoid kernel is able to predict all data points correctly. However, this performance is based on a small amount of testing data and the lack of real pump data makes it difficult to determine the model's true ability to predict cavitation.

The success of the SVM model suggests that training a general model for cavitation prediction using data from experiments simulating cavitation-like behavior is a worthwhile direction for further research.

Future research should collect more cavitation data from multiple types of industrial pumps to better train and evaluate the model. Actual cavitation signals or a simulated pump environment would provide better data for improving the model's performance. Even a small amount of actual cavitation data would improve the evaluation process quality immensely. In order to get better measurement results the design of the pumps should be taken into consideration before placing the sensor by either doing more research on the specific pump or consult a specialist for pumps.

Lastly, the SVM model with the sigmoid kernel is equivalent to a two-layer neural network, indicating that larger neural networks may have the potential to further improve prediction performance, and possibly performing even better than SVM and XGBoost models.

References

[1] Y. Sha, J. Faber, S. Gou, B. Liu, W. Li, S. Shramm, H. Stoecker, T. Steckenreiter, D. Vnucec, N. Wetzstein, W. Andreas, and Z. Kai, "An acoustic signal cavitation detection framework based on xgboost with adaptive selection feature engineering," 2022. [Online]. Available: https://arxiv.org/pdf/2202.13226.pdf

[2] A. Geron, *Hands-on machine learning with Scikit-Learn and TensorFlow : concepts, tools, and techniques to build intelligent systems.* Sebastopol, CA: O'Reilly Media, 2017.

[3] A. Navia-Vázquez, "Support vector perceptrons," *Neurocomputing*, vol. 70, no. 4-6, pp. 1089–1095, 2007.