Discrimination of reflected sound signals

Master of Engineering in Information Technology  
Machine Learning  
line 3: Given Name(s) Surname line 4: matriculation number  
line 5: student email address (…@stud.fra…)

*Abstract*—Machine learning (ML) in the acoustics and signal processing domain has experienced rapid advancements and developments with persuasive outcomes over a course of years. The statistical techniques of Machine Learning offer detection of data patterns, which helps in the identification of the convoluted relationship between features and further discrimination, or classification based on these features. One of the ML techniques, called Binary classification is usually used to discriminate between two class labels. This paper provides an ML-based solution for the discrimination of reflected sound signals, which are reflected from two different objects. Firstly, data pre-processing is performed on the reflected time signals to render the dataset. Secondly, Quadratic Time-Frequency representation (QTFR) of the reflected sound signal is generated and features extraction is performed on it. Afterward, four different Machine Learning classification models; namely, K-Nearest Neighbors, Random Forest, Logistic Regression, and Decision Trees are utilized for data training and prediction for the realization of binary classifier or discriminator. Finally, an assessment of classification results based on accuracy and various other measures are presented and discussed.

Keywords—Machine Learning, Quadratic Time-Frequency Representation, Binary Classification, Discriminator, K-Nearest Neighbors, Random Forest, Logistic Regression, Decision Trees

# Introduction

The recent inventions of smart and intelligent systems and devices have increased the utilization of Machine Learning and computational intelligence-based algorithms. A wide range of research on the popular topic of Machine Learning (ML) algorithms and techniques are available, which have brought considerable progress and ease to routine activities in this digital era. Machines are being trained to perform the task that humans do. The innovative algorithms and techniques of Machine learning (ML) have empowered advancements in automatic-data processing and pattern recognition across many sciences and engineering fields. ML offers intelligence-based solutions to complex engineering challenges, in the same manner as the processing of the human brain. Additionally, ML deals with a variety of diverse big datasets such as image, video, audio, time-series signals, 1D signals, text, etc, which are vastly produced and stored by intelligent systems [1]. ML techniques in the domain of sound signals and acoustics have gained much attention for its persuasive solutions towards crucial tasks such as identification and validation of different sound signals. These ML algorithm detects data patterns by extracting useful attributes and features from the given dataset. Data labeling is performed afterward by utilizing these patterns. ML-based statistical methods enable the system or machine to learn and predict based on pattern recognition [2].

ML classification technique works through estimating the mapping pattern that logs the training dataset to the target class or label [3]. One of the ML classification techniques which acquire only two class labels is referred to as the Binary classification technique. The input dataset samples are classified into two states by computing specific classification measurements. There are 2 disjoint classes available for binary classification [4]. Some of the popular ML algorithms that can be used to realize binary classifier are Logistic Regression, k-Nearest Neighbors (k-NN), Decision Trees, Support Vector Machine (SVM), Naive Bayes, etc.

The inspiration behind this study is to deliver a classification-based solution to discriminate between the reflected sound signals coming from different objects. The considered use-case scenario is the acoustic signals that incident on the surface of some objects and reflects back. The time signal is formed after recording those reflected sound signals. The resultant time signal is the convolution of the incident sound waveform along with the reflecting object’s surface properties. By analyzing the properties of the reflected time signal, the knowledge about the reflecting object can be achieved, on the basis of which the discrimination between reflected time signals is possible.

The ML-based solution for the aforementioned challenge is the main goal of this study. The aim is to realize a Binary Classifier or discriminator by applying different ML algorithms on the labeled dataset. The employed ML algorithms include Logistic Regression, k-Nearest Neighbors (k-NN), Decision Trees, and Random Forest (RF). This Binary classifier or discriminator is used to classify or distinguish the reflected sound signals, which belong to two different objects, named Object#1 and Object#2. Before implementing the binary classification algorithm, the Quadratic time-frequency representation is created for given sound signals, based on which the associated features are extracted to perform classification.

This paper presents an ML-powered solution for the discrimination of reflected sound signals by using a binary classification technique. The framework of this paper is as follows: Section II provides a brief description of the utilized techniques for the realization of the Binary classification or Discriminator model. Section III illustrates the workflow and step-by-step approach and methodology to achieve the discriminator of the reflected signal. Section IV explains the python-based code implementation for the binary classification model. In Section V, the result analysis and assessment for the implemented binary classification model based on some evaluation metrics will be discussed. Finally, the paper is concluded in Section IV.

# Techniques used in Binary Classification Model Implementation

Before explaining the main implementation approach, a brief overview of the techniques utilized in achieving the Binary Classification model is presented in this section.

## Quadractic Time-Frequency Representation

First, confirm that you have the correct template for your paper size. This template has been tailored for output on the A4 paper size.

## Machine Learning Classification Models

The template is used to format your paper and style the text. All margins, column widths, line spaces, and text fonts are prescribed; please do not alter them. You may note peculiarities. For example, the head margin in this template measures proportionately more than is customary. This measurement and others are deliberate, using specifications that anticipate your paper as one part of the entire proceedings, and not as an independent document. Please do not revise any of the current designations.

# Methodology

As stated earlier, the main inspiration of this study is to provide an ML-based solution for the discrimination of reflected sound signals occurring from two different objects. In this section, a complete workflow of the implemented approach for the realization of the Binary Classification model is presented. The step-by-step description of the applied model is provided. Furthermore, for the performance assessment of the trained classification models in testing phase, some evaluation metrics are used, which will be explained later in this section. To provide a convenient and simple usage of the proposed model, a graphical user interface (GUI) for the Binary Classification model is created. The implementation steps for the creation of GUI is also discussed in this section.

## The workflow of Proposed Binary Classification Model

The block diagram of the proposed experimental setup to implement the Binary classification model is illustrated in fig x.

### Data Processing: The training dataset is based on the comma-separated values (csv) files, which contain a time signal per row. These csv files, containing different time signals are allocated for each object i.e., Object#1 and Object#2. The data processing is performed on these csv files, as depicted in figure x.

#### CSV files merging: All the csv files, containing time signals for Object#1, are merged into one csv file. In the same manner, the data of csv files for Object#2 are also merged into a single csv file.

#### Conversion into excel file: The merged csv files for Object#1 and Object#2 are then converted into excel files (.xlsx format). These excel files for both objects serve as the input and are fed into the proposed experimental setup for further processing.

### Quadratic Time-Frequency Representation (QTFR): In this step, the system starts reading the given time signal one by one from the input excel files, and creates quadratic time-frequency representation for each time signal. As discussed earlier, the QTFR method is used to calculate the energy of a signal as a function of time and frequency. The resultant signal is termed as the quadratic time-frequency representation of the signal. Since the QTFR signal depicts the signal’s energy density in the time-frequency domain, this can also be referred to as Spectrogram. There are different quadratic time-frequency signal analysis methods, out of which the Short-time Fourier transform (STFT) spectrogram method is chosen to implement in this experiment. Each one-dimensional input time signal is mapped into the two-dimensional time-frequency signal, which creates the spectrogram. These spectrograms are utilized in the feature extraction phase.

### Feature Extraction: The feature extraction phase is the most vital element of any classification model and it serves as the first building block for designing and training of any ML-based classification model. The STFT spectrograms are subjected to the feature extraction phase. Two distinct features are extracted from the STFT spectrograms for training and testing of the model, which are as follows.

* Maximum Frequency (the maximum frequency of the spectrogram of one time signal)
* Maximum Spectrum Sum (the sum of all the elements in the spectrogram of one time signal)

### Data Frame Formation: In the feature extraction phase, the feature extraction process is applied on each time-frequency spectrogram, coming from both objects. To organize these extracted features, the data frame formation method is used. In this step, all the “Maximum Frequency” & “” features extracted from the QTFR spectrograms of Object#1 are organized in a tabular format, as shown in Table I. The target value “0” is assigned to these extracted features of Object#1.

1. Data Frame Formation of Object#1

| Spectrogram Array | Max Frequency | Max Spectrum Sum | Target |
| --- | --- | --- | --- |
| 0 | Max Frequency0 | Max Spectrum Sum0 | 0 |
| 1 | Max Frequency1 | Max Spectrum Sum1 | 0 |
| … | … | … | 0 |

In the similar manner, the extracted features from the QTFR spectrograms of Object#2 are subjected to Data formation and target value “1” is assigned to them, as shown in Table II.

1. Data Frame Formation of Object#2

| Spectrogram Array | Max Frequency | Max Spectrum Sum | Target |
| --- | --- | --- | --- |
| 324 | Max Frequency | Max Spectrum Sum | 1 |
| 325 | Max Frequency | Max Spectrum Sum | 1 |
| … | … | … | 1 |

### Splitting Dataset for Testing and Training phase: At this step, the dataset is split up in “testing dataset” and “training dataset”. The training dataset is used to train the classification model. The testing dataset is utilized to evaluate the final performance of the model. This testing dataset helps the model to learn for future predictions.

### Binary Classification Model: In the previous step, the extracted feature based training and testing datasets are obtained. In classification phase, two operations are performed parallely on classification models, namely “Model Training” and “Model Testing”.

In the “Model Training” operation, the training dataset is utilized to train different classifiers for the realization of binary classification. The applied classification models are as follows.

* Logistic Regression
* k-Nearest Neighbors (k-NN),
* Decision Trees
* Random Forest (RF)

Each classifier train itself with the given training data according to its algorithm as described in previous section. Based on the distinct features and class labels of two objects, the classifiers learn the discrimination logic. As the binary classification is the main goal in this study, therefore the classifiers learn from the extracted features and associated labels that the input signal belongs to Object#1 or Object#2.

In the “Model Testing” operation, the testing data is fed into the classifiers to test whether the classifier’s prediction is accurate or not. Based on accurate predictions, the efficiency of the classification model is evaluated. If the prediction accuracy of the classification model is greater than 85%, then the best model will be saved for future usage. If the evaluated accuracy is less than 85%, then it is assumed that the classification model needs to learn more, so this triggers model training operation again. The wrong predictions are used to further update and train the classification model in order to achieve better efficiency.

## Performance Evaluation Metrics

The predicted results are evaluated using Confusion Matrix, which depicts the comprehensive performance of the Binary Classification model. In this 2-D matrix, rows and columns represents the Classes and the diagonal depicts the accurate classification. The evaluation metrics highly depends on the values of True positive **tp**, True negative **tn**, False positive **fp** & False negative **fn**. The **tp** are the number of positive predicted tests, which are originally positive. Similarly, the **tn** are the number of negative predicted tests, which are originally negative. The **fp** are the number of positive predicted tests, which are originally negative and the **fn** are the number of negative predicted tests that are originally positive. The following measures are used to in the main evaluation metrics.

* False Discovery Rate (FDR)
* Negative Predictive Value (NPV)
* True Positive rate (TPR)
* True Negative Rate (TNR)

The evaluation metrics used in this project are described below.

### Accuracy: Classification accuracy is defined as the percentage of the tests which are correctly predicted from total conducted tests. Accuracy can be evaluated by using eq.

### Precision: This evaluation metric is defined as the ratio of number of tests which actually have labelled class to the overall tests classified as labelled class. Precision can be evaluated by using eq.

### Recall: This evaluation metric is defined as the ratio of number of tests which are classified as labelled class to overall tests which truly lie in labelled class. It is calculated by using eq.

### F1 Score: This evaluation metric is defined as the ratio of number of tests which are classified as labelled class to overall tests which truly lie in labelled class. It is calculated by using eq.

### ROC: A receiver operating characteristic curve (ROC) curve, is a graphical plot that depicts the prediction performance of a binary classification model, when its “discrimination threshold” is changed. The ROC curve plots a parametric graph of True Positive Rate of a threshold parameter versus False Positive Rate of a threshold paramenter, while the threshold serves as the variable parameter.

## Graphical User Interface (GUI)

* The subscript for the permeability of vacuum **0, and other common scientific constants, is zero with subscript formatting, not a lowercase letter “o”.
* In American English, commas, semicolons, periods, question, and exclamation marks are located within quotation marks only when a complete thought or name is cited, such as a title or full quotation. When quotation marks are used, instead of a bold or italic typeface, to highlight a word or phrase, punctuation should appear outside of the quotation marks. A parenthetical phrase or statement at the end of a sentence is punctuated outside of the closing parenthesis (like this). (A parenthetical sentence is punctuated within the parentheses.)
* A graph within a graph is an “inset”, not an “insert”. The word alternatively is preferred to the word “alternately” (unless you really mean something that alternates).
* Do not use the word “essentially” to mean “approximately” or “effectively”.
* In your paper title, if the words “that uses” can accurately replace the word “using”, capitalize the “u”; if not, keep using lower-cased.
* Be aware of the different meanings of the homophones “affect” and “effect”, “complement” and “compliment”, “discreet” and “discrete”, “principal” and “principle”.
* Do not confuse “imply” and “infer”.
* The prefix “non” is not a word; it should be joined to the word it modifies, usually without a hyphen.
* There is no period after the “et” in the Latin abbreviation “et al.”.
* The abbreviation “i.e.” means “that is”, and the abbreviation “e.g.” means “for example”.

An excellent style manual for science writers is [7].

# Code Description

After the text edit has been completed, the paper is ready for the template. Duplicate the template file by using the Save As command, and use the naming convention prescribed by your conference for the name of your paper. In this newly created file, highlight all of the contents and import your prepared text file. You are now ready to style your paper; use the scroll down window on the left of the MS Word Formatting toolbar.

## Authors and Affiliations

**The template is designed for, but not limited to, six authors.** A minimum of one author is required for all conference articles. Author names should be listed starting from left to right and then moving down to the next line. This is the author sequence that will be used in future citations and by indexing services. Names should not be listed in columns nor group by affiliation. Please keep your affiliations as succinct as possible (for example, do not differentiate among departments of the same organization).

### For papers with more than six authors: Add author names horizontally, moving to a third row if needed for more than 8 authors.

### For papers with less than six authors: To change the default, adjust the template as follows.

#### Selection: Highlight all author and affiliation lines.

#### Change number of columns: Select the Columns icon from the MS Word Standard toolbar and then select the correct number of columns from the selection palette.

#### Deletion: Delete the author and affiliation lines for the extra authors.

## Identify the Headings

Headings, or heads, are organizational devices that guide the reader through your paper. There are two types: component heads and text heads.

Component heads identify the different components of your paper and are not topically subordinate to each other. Examples include Acknowledgments and References and, for these, the correct style to use is “Heading 5”. Use “figure caption” for your Figure captions, and “table head” for your table title. Run-in heads, such as “Abstract”, will require you to apply a style (in this case, italic) in addition to the style provided by the drop down menu to differentiate the head from the text.

Text heads organize the topics on a relational, hierarchical basis. For example, the paper title is the primary text head because all subsequent material relates and elaborates on this one topic. If there are two or more sub-topics, the next level head (uppercase Roman numerals) should be used and, conversely, if there are not at least two sub-topics, then no subheads should be introduced. Styles named “Heading 1”, “Heading 2”, “Heading 3”, and “Heading 4” are prescribed.

## Figures and Tables

#### Positioning Figures and Tables: Place figures and tables at the top and bottom of columns. Avoid placing them in the middle of columns. Large figures and tables may span across both columns. Figure captions should be below the figures; table heads should appear above the tables. Insert figures and tables after they are cited in the text. Use the abbreviation “Fig. 1”, even at the beginning of a sentence.

1. Table Type Styles

| Table Head | Table Column Head | | |
| --- | --- | --- | --- |
| Table column subhead | Subhead | Subhead |
| copy | More table copya |  |  |

1. Sample of a Table footnote. (*Table footnote*)
2. Example of a figure caption. (*figure caption*)

Figure Labels: Use 8 point Times New Roman for Figure labels. Use words rather than symbols or abbreviations when writing Figure axis labels to avoid confusing the reader. As an example, write the quantity “Magnetization”, or “Magnetization, M”, not just “M”. If including units in the label, present them within parentheses. Do not label axes only with units. In the example, write “Magnetization (A/m)” or “Magnetization {A[m(1)]}”, not just “A/m”. Do not label axes with a ratio of quantities and units. For example, write “Temperature (K)”, not “Temperature/K”.

# Results

# Conclusion

##### Acknowledgment *(Heading 5)*

The preferred spelling of the word “acknowledgment” in America is without an “e” after the “g”. Avoid the stilted expression “one of us (R. B. G.) thanks ...”. Instead, try “R. B. G. thanks...”. Put sponsor acknowledgments in the unnumbered footnote on the first page.

##### References

The template will number citations consecutively within brackets [1]. The sentence punctuation follows the bracket [2]. Refer simply to the reference number, as in [3]—do not use “Ref. [3]” or “reference [3]” except at the beginning of a sentence: “Reference [3] was the first ...”

Number footnotes separately in superscripts. Place the actual footnote at the bottom of the column in which it was cited. Do not put footnotes in the abstract or reference list. Use letters for table footnotes.

Unless there are six authors or more give all authors’ names; do not use “et al.”. Papers that have not been published, even if they have been submitted for publication, should be cited as “unpublished” [4]. Papers that have been accepted for publication should be cited as “in press” [5]. Capitalize only the first word in a paper title, except for proper nouns and element symbols.

For papers published in translation journals, please give the English citation first, followed by the original foreign-language citation [6].

1. G. Papakostas, K. Diamantaras and F. Palmieri, “Emerging Trends in Machine Learning for Signal Processing”, Computational Intelligence and Neuroscience, vol 2017, pp. 1-2, November 2017, Article ID 6521367, DOI: [10.1155/2017/6521367](https://www.researchgate.net/deref/http%3A%2F%2Fdx.doi.org%2F10.1155%2F2017%2F6521367?_sg%5B0%5D=XWGBoldQewsJBnvG7YxetYkUFg7uR1V_B2pA9ndw1fmCp5l9hf_N_L2AHjlGfLtWdy2andutwgqn1ROjRCw9Vl3FKQ.iq11NVTyIfEmKHmDS6lxbaTc2yhelQsJoEkQ_XFhShVW0TLLugDF7CPeRBu3IGgHjj2L01xBrET0e0a1g2SvAQ)
2. M. J. Bianco, P. Gerstoft, J. Traer, E. Ozanich, M. A. Roch, S. Gannot, and C. A. Deledalle, "Machine learning in acoustics: Theory and applications", The Journal of the Acoustical Society of America 146 , pp. 3590-3628, 2019, DOI: 0.1121/1.5133944
3. A. C. de Carvalho and A. A. Freitas, "A tutorial on multi-label classification techniques," in Foundations of Computational Intelligence, vol. 5, , ed: Springer, pp. 177-195, 2009
4. M. Er, R. Venkatesan, & N. Wang, “An Online Universal Classifier for Binary, Multi-class and Multi-label Classification”, September 2016.
5. J. Clerk Maxwell, A Treatise on Electricity and Magnetism, 3rd ed., vol. 2. Oxford: Clarendon, 1892, pp.68–73.
6. I. S. Jacobs and C. P. Bean, “Fine particles, thin films and exchange anisotropy,” in Magnetism, vol. III, G. T. Rado and H. Suhl, Eds. New York: Academic, 1963, pp. 271–350.
7. K. Elissa, “Title of paper if known,” unpublished.
8. R. Nicole, “Title of paper with only first word capitalized,” J. Name Stand. Abbrev., in press.
9. Y. Yorozu, M. Hirano, K. Oka, and Y. Tagawa, “Electron spectroscopy studies on magneto-optical media and plastic substrate interface,” IEEE Transl. J. Magn. Japan, vol. 2, pp. 740–741, August 1987 [Digests 9th Annual Conf. Magnetics Japan, p. 301, 1982].
10. M. Young, The Technical Writer’s Handbook. Mill Valley, CA: University Science, 1989.

**This template contains guidance text for composing and formatting seminar paper. Please ensure that all template text is removed from your paper prior to submission to the examination office. Failure to remove template text from your paper may result in your paper being devaluated.**

We suggest that you use a text box to insert a graphic (which is ideally a 300 dpi TIFF or EPS file, with all fonts embedded) because, in an MSW document, this method is somewhat more stable than directly inserting a picture.

To have non-visible rules on your frame, use the MSWord “Format” pull-down menu, select Text Box > Colors and Lines to choose No Fill and No Line.