Calibration using AI

JUSTINE Jean   
 *Högskolan Väst*  
*EIGSI*Trollhättan Sweden  
jean.justine@aol.com

*Abstract*—The work we had to do was to find the inverse function to calibrate sensors. To do so, we will train neural network, to find the best models, that can predict efficiently the true value using the sensors output as training set, and the true value of the environment as target, then when the models are precise enough we will have our calibration algorithm. All of the code is available on GitHub here: https://github.com/Justine-IA/Deep\_learning-/tree/main/Project

# Data Analysis

Firstly, the project will be a Deep Learning project that consist of finding an algorithm that predict true value in function of the value that a sensor finds, for that we will use a dataset that measure pollution in a city the dataset is available here: <https://archive.ics.uci.edu/dataset/360/air+quality>.

## Requirement

The requirement will be to find the inverse function to find the real value, through the value detected by the sensors, because sensor data (y) is a function of true measurement (x).

We have to find the inverse function, which is: *f^{-1}(y) = x.*

This is an important task because having the real measurement can be crucial to take actions later on.

## Data

The dataset is composed of time feature, value that the sensor finds, like hourly average sensor response of CO, and also the real value (Ground Truth) like of CO, not the one the sensor detected. There is no missing value according to the website but there is some not correct value like negative number of pollution or -200 for sensors temperature so we removed them as well as for the GT data, there is a lot of missing value so we just replace them with the one closer in time to them, the time and date will be used for a Recurrent Neural Network later on because the sensor is detecting pollution in a city so it will be useful to train a model with the times feature, because of cars or anything else that influence pollution. We also have humidity and temperature in addition to all the other feature of pollution and else.

## Visualisation

Here we have some visualization for the data such as:

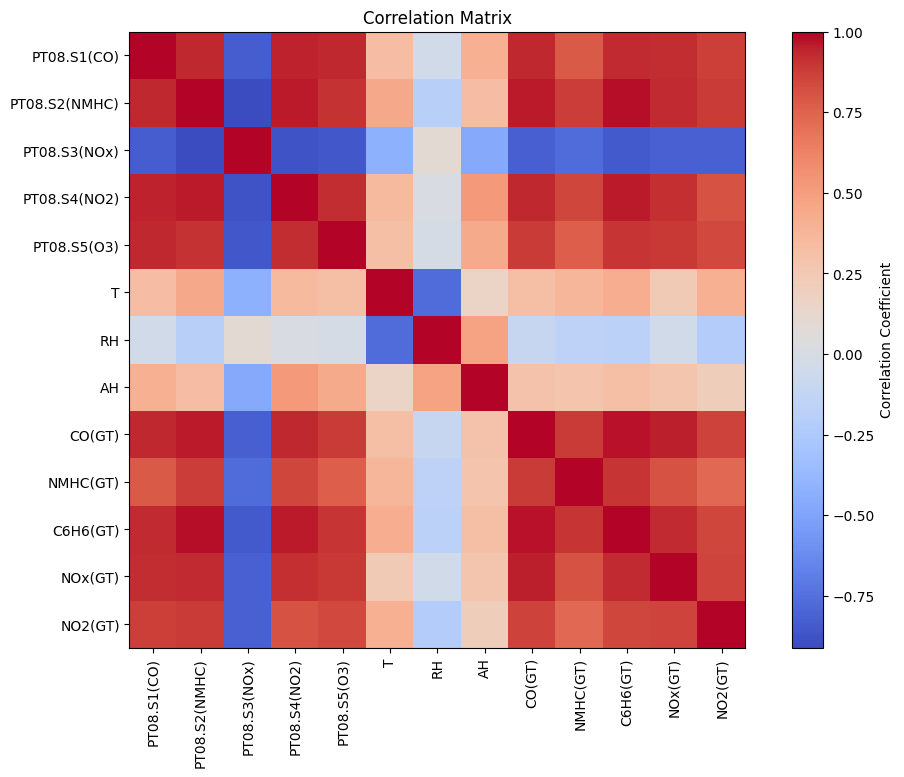


Fig. 1. Correlation Matrix

As shown in fig1, we can see the correlation between the polluting elements measured and its ground truth are very related to each other. We can also see that NOX and RH are not very related to anything but themselves.

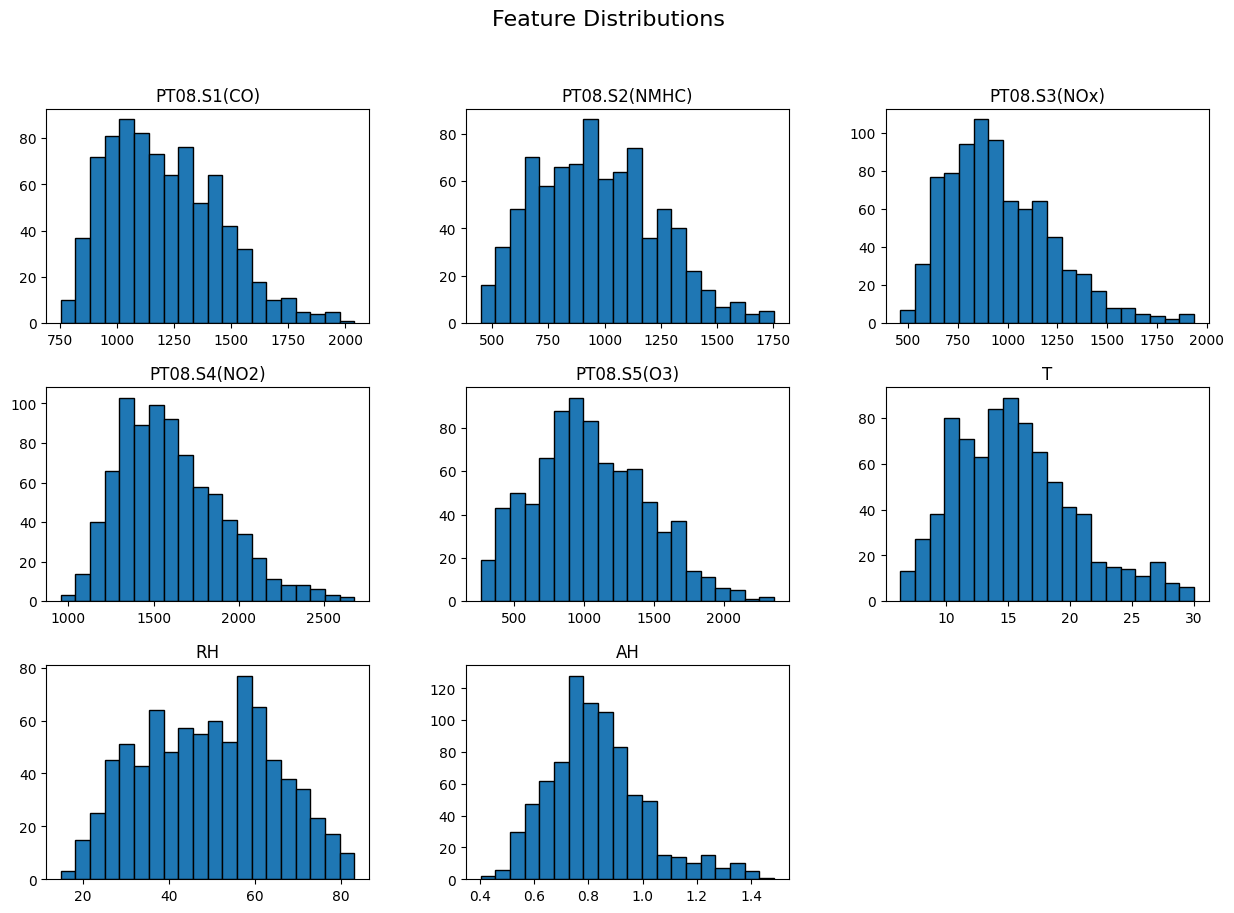


Fig. 2. Feature Distribution

We can see that in Fig 2, most of the data is correct and doesn’t need further data cleaning except for a scaling later on.

## Requirement

Some requirement that are in existing research. For temperature and light sensors some Neural Network (NN) were develop for [1].

Also ML based calibration can handle over time drift, and environment variability, improving long time accuracy [2].

List of requirement:

* Acquire Data
* Choosing Model
* Model Training and validation
* Figure of merits
* Documentation

# Algorithm Design

## Data Analysis

Firstly we fetch the data, then, as said in the first part, we then separate the feature form the target, y, the target, will take all of the GT value, x will take the rest, we will build a Recurrent NN so we keep time in a variable call X\_RNN, whereas X will not have time or date.

After that we need to scale the data, firstly we will need to transform the data into numerical value, like 2004 for year, 10 for month, 23 for day, 18 for hour instead of the usual 23/10/2004 and 18:00, we also have day of the week, after that we will add cyclical encoding for the time and day of the week, using cos and sin to have a cycle.

After all of that we need to do the real scaling we just import standards scaler and scale X and X\_RNN as well as y, we need to scale the data for having a better consistency, better metrics calculation, and it prevents certain data to be too dominant compare to others.

After we split the data between X train X test y train y test as usual but we also have X train RNN and y test RNN which contains the date and time.

# Prepare Your Paper Before Styling

Before you begin to format your paper, first write and save the content as a separate text file. Complete all content and organizational editing before formatting. Please note sections A-D below for more information on proofreading, spelling and grammar.

Keep your text and graphic files separate until after the text has been formatted and styled. Do not use hard tabs, and limit use of hard returns to only one return at the end of a paragraph. Do not add any kind of pagination anywhere in the paper. Do not number text heads-the template will do that for you.

## Abbreviations and Acronyms

Define

# Using the Template

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”.

##### 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

1. [1] Y. Zhang, L. O. H. Wijeratne, S. Talebi, and D. J. Lary, "Machine Learning for Light Sensor Calibration," Sensors, vol. 21, no. 18, pp. 6259, Sep. 2021. Available: <https://www.mdpi.com/1424-8220/21/18/6259>.
2. [2] Z. Wang, "Evaluating the Efficacy of Machine Learning in Calibrating Low-Cost Sensors," Applied and Computational Engineering, vol. 71, pp. 30–38, 2024. Available: <https://www.ewadirect.com/proceedings/ace/article/view/15136>.