

AQI Prediction Report

(COMP3125 Individual Project)

*Note: Do not used sub-title

Mohamed Mohamed
dept. name of organization

Abstract—This electronic document is a “live” template and already defines the components of your paper [title, text, heads, etc.] in its style sheet. ***CRITICAL: Do Not Use Symbols, Special Characters, Footnotes, or Math in Paper Title or Abstract.** (provide a short abstract)

Keywords—example1, example2, example3, example 4, example 5 (provide 3-5 keywords)

I. INTRODUCTION (HEADING 1)

Air pollution remains one of the most pressing environmental challenges of our time, with significant implications for human health, ecosystems, and climate. The quality of air we breathe is influenced by various factors, including industrial activities, vehicle emissions, and weather conditions. As urban populations continue to grow, so does the demand for robust monitoring systems that can predict and mitigate the risks associated with poor air quality. This project focuses on the analysis and prediction of Air Quality Index (AQI) levels using historical environmental data and machine learning techniques.

Accurate AQI prediction not only empowers local governments and environmental agencies to make data-driven decisions but also enhances public awareness and safety. To achieve this, we explore the relationships between AQI levels and environmental variables such as PM2.5, ozone (O₃), carbon monoxide (CO), nitrogen dioxide (NO₂), temperature, humidity, and wind speed. By identifying which pollutants contribute the most to air quality deterioration and how these patterns vary seasonally and geographically, we can better understand the dynamics of air pollution.

Previous studies have demonstrated the potential of machine learning models, such as regression and classification algorithms, in predicting air quality outcomes. For example, linear regression has been used effectively to model AQI as a continuous variable, while classification models such as decision trees, support vector machines (SVM), and neural networks have shown promise in categorizing AQI levels into health-related brackets. Evaluation metrics such as root mean square error (RMSE), mean absolute error (MAE), accuracy, and F1-score are commonly applied to assess the performance of these models [1][2][3].

This project aims to answer several interrelated questions: Which pollutants are the strongest indicators of poor air quality? How accurately can we predict AQI levels using historical and environmental data? What influence do seasonal changes and geographic differences have on pollution patterns? And finally, how can this information inform actions cities can take to reduce harmful exposure to polluted air? Through rigorous data analysis and modeling, this work intends to contribute to the growing body of research on predictive air quality systems and offer practical insights for future environmental policy and public health planning.

Datasets

The datasets used in this project are sourced from publicly available, reputable platforms including Kaggle and Data.gov. These platforms are widely used in the data science community and are known for hosting high-quality, credible datasets submitted by organizations, researchers, and verified contributors.

□ **Air Quality Data Set (Kaggle)**
This dataset was published on Kaggle and consists of hourly-averaged responses from a chemical sensor array in an Italian city. The original data was collected between March 2004 and February 2005 using metal oxide sensors designed to detect air pollutants such as CO, benzene, and NOx. The dataset was created by researchers to support air quality analysis using sensor technology.

□ **U.S. Pollution Data (2000–2022)**
This dataset was generated by compiling air quality monitoring data from the U.S. Environmental Protection Agency (EPA). It includes pollutant readings (e.g., O₃, CO, NO₂, and SO₂) across the United States from 2000 to 2022. The data was collected from various EPA monitoring stations and submitted for public access and research purposes.

□ **Air Quality System (AQS) Monitoring Data (Data.gov)**
This dataset is hosted on Data.gov and maintained by the EPA. It contains geospatial metadata and sensor readings from monitoring stations across the U.S., capturing pollutants such as PM2.5, NO₂, O₃, and meteorological variables like wind speed and barometric pressure. The AQS is a reliable government data source used for environmental and public health research.

Example: XXXX

A. Character of the datasets

Describe the dataset's format and size. Additionally, provide an overview of the dataset's characteristics, including its features, size, structure, and any relevant attributes that are important for your analysis. Describe the dataset's format and size, as well as its key features, including the parameters, columns, rows, and character attributes along with their respective units. Using a table to present this information is recommended for clarity. Explain whether you cleaned the data or converted any units, specifying the formulas or rules applied. If multiple datasets were combined, describe how they were merged. Additionally, mention if you created any new categories for analysis, detailing what they are and how they were generated. Providing this background ensures transparency and helps readers understand the reliability and relevance of your data.

Example: XXXX

II. METHODOLOGY

In this part, you should give an introduction of the methods/model. First, what's the method/model. What's the assumption of this method/model. What's the

advantage/disadvantage of this method/model. Why did you choose it. What Python module or function do you apply to apply this method/model. Any optional input/extra work did you adjust to make the results better. If you have multiple methods, feel free to use subsection A., B. to separate them.

Example: 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.

A. Method A

Example: The equations are an exception to the prescribed specifications of this template. You will need to determine whether or not your equation should be typed using either the Times New Roman or the Symbol font (please no other font). To create multileveled equations, it may be necessary to treat the equation as a graphic and insert it into the text after your paper is styled.

$$a + b = \gamma$$

(1)

Note that the equation is centered using a center tab stop. Be sure that the symbols in your equation have been defined before or immediately following the equation. Use “(1)”, not “Eq. (1)” or “equation (1)”, except at the beginning of a sentence: “Equation (1) is . . .”

B. Method B

- Bulletin 1
- Bulletin 2.
- Bulletin 3

C. Method C

Example: The equations are an exception to the prescribed specifications of this template. You will need to determine whether or not your equation should be typed using either the Times New Roman or the Symbol font (please no other font). To create multileveled equations, it may be necessary to treat the equation as a graphic and insert it into the text after your paper is styled.

$$a + b = \gamma$$

(1)

Note that the equation is centered using a center tab stop. Be sure that the symbols in your equation have been defined before or immediately following the equation. Use “(1)”, not “Eq. (1)” or “equation (1)”, except at the beginning of a sentence: “Equation (1) is . . .”

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

III. RESULTS

In this section, present your findings using an appropriate method, such as equations, numerical summaries, or visualizations like charts and graphs. Clearly explain all results and provide guidance on how to interpret them. If any unexpected results arise, discuss possible reasons or contributing factors. To improve clarity and organization, consider using subsections (e.g., A, B) to separate different aspects of your results.

Example: 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.

A. Result A

Example: XXX

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

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

- a) Selection: Highlight all author and affiliation lines.
- b) 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.
- c) Deletion: Delete the author and affiliation lines for the extra authors.

B. Results B

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

C. Results C

a) 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.

TABLE I. TABLE TYPE STYLES

Table Head	Table Column Head		
	Table column subhead	Subhead	Subhead
copy	More table copy ^a		

^a Sample of a Table footnote. (Table footnote)

Fig. 1. 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”.

IV. DISCUSSION

Every method/project has its shortage or weakness. Please discuss the unsatisfied results in your project. And discuss the feasible suggestions of future work to revise/improve your result.

Example: xxx

V. CONCLUSION

In this part, you should summarize your project. What important results did you find for your topic and what's the effect of this result on the real-world?

Example: xxx

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

Use the IEEE format for the citation. 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 ...” 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.

- [1] G. Eason, B. Noble, and I. N. Sneddon, “On certain integrals of Lipschitz-Hankel type involving products of Bessel functions,” *Phil. Trans. Roy. Soc. London*, vol. A247, pp. 529–551, April 1955. (*references*)
- [2] J. Clerk Maxwell, *A Treatise on Electricity and Magnetism*, 3rd ed., vol. 2. Oxford: Clarendon, 1892, pp.68–73.
- [3] 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.
- [4] K. Elissa, “Title of paper if known,” unpublished.
- [5] R. Nicole, “Title of paper with only first word capitalized,” *J. Name Stand. Abbrev.*, in press.
- [6] 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].
- [7] M. Young, *The Technical Writer’s Handbook*. Mill Valley, CA: University Science, 1989.

IEEE conference templates contain guidance text for composing and formatting conference papers. Please ensure that all template text is removed from your conference paper prior to submission to the conference. Failure to remove template text from your paper may result in your paper not being published.

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.