K-NEAREST NEIGHBOURS LABORATORY

Gabriele Francesco Berruti

University of Genoa
Via Roma 37/6
Genova, Italy
gabriele.berruti@unige.edu

Ibrahim Altufayli

University of Genoa
Via Roma 37/6
Genova, Italy
ibrahim.altufayli@unige.edu

ABSTRACT

This lab is about the implementation and analysis of the KNN algorithm for classification problems

This is the template file for the proceedings of the 2024 International Conference on Auditory Display, which will be held June 24 to 28, 2024 at Rensselaer Polytechnic Institute, Troy, NY, USA. This template has been generated from the ICAD 2024 template and aims at producing conference proceedings in electronic form. Please use either LaTeX or Microsoft Word formats when preparing your submission. All questions concerning ICAD 2024 submissions should be addressed to the paper chairs (papers@icad2024.icad.org). These guidelines and templates are adapted from previous editions of ICAD conferences, so an experienced author who has published something in these proceedings will find them familiar. The templates are available in electronic form at the website: https://icad2024.icad.org/
The length of the abstract should be approximately 150 words.

1. INTRODUCTION

Classification may be defined as the process of predicting class or category from observed values or given data points. The categorized output can have the form such as “Black” or “White” or “spam” or “no spam”.
- Summarize relevant works in the field (even in scientific report it is expected you possess a grasp on the state of the art) . the state-of-art to implement similar operations is the use of libraries such as sklearn which allow you to evaluate neural networks of classifiers using nearest neighbor also carrying out predictions and calculating the error .
- Explain your choices (algorithms and experiments) .
- Summarize the structure of the rest of the report, section by section .
The aim of the laboratory was to train a KNN classifier which is based first of all on a generated (training) set composed of 2D points randomly sampled on a plane and to which a binary label is assigned according to their position with reference to a linear separator .
This is the procedure by which our dataset is created. Another operation that is applied on our dataset consists of applying a noise to each of the samples to allow perturbing the labels and making the task more complicated .
The first section of the report allows to use the KNN algorithm to estimate the classification function. The implementation sees the algorithm train on a training set and make predictions on the test set. In this section, metrics are also defined that evaluate the goodness of the classifier.
The second section essentially evaluates the classifier in the case in which the creation of the dataset occurs without noise. And it consists of 4 points:
generate the training set (without noise) and display the separation curve which derives from the predictions made by the trained model on the training set .
The second section essentially evaluates the classifier in the case in which the creation of the training and test dataset occurs without noise. And it consists of several points:
the first one generate the training set (without noise) and display
the separation curve which derives from the predictions made on the model
trained on the training set and evaluated on the same.
The next point trains the model on the training set and evaluates it on the test set, thus calculating the prediction error as done previously.
The last point allows you to compare the results obtained through the previously described implementation with that of the state-of-art algorithm (scikit-learn) .

2. ALGORITHMS AND METHODS

All manuscripts must be formatted for white A4 paper (8.27 in x 11.02 in). Please do **not** use US letter-size papers. All printed material, including text, illustrations, and charts, must be kept within a print area of 6.27 inches (169 mm) wide by 9.2 inches (233 mm) high. Do not write or print anything outside the print area. The top margin must be 1 inch (25 mm), except for the title page, and the left margin must be 0.75 inch (19 mm). All *text* must be in a two-column format. Columns are to be 3.29 inches (83.5 mm) wide, with a 0.31 inch (8 mm) space between them. Text must be fully justified.

3. EXPERIMENTAL ANALYSIS

The paper title (on the first page) should begin 0.98 inches (25 mm) from the top edge of the page, centered, completely capitalized, and in Times New Roman 14-point, boldface type. The authors’ name(s) and affiliation(s) appear below the title. Papers

This work is licensed under Creative Commons Attribution – Non Commercial 4.0 International License. The full terms of the License are available at http://creativecommons.org/licenses/by-nc/4.0/

with multiple authors and affiliations may require a greater amount of space for this information.

4. TYPE-STYLE AND FONTS

To achieve the best rendering in the proceedings, we strongly encourage you to use Times New Roman font. In addition, this will give the proceedings a more uniform look. Use a font size that is no smaller than nine point type throughout the paper, including figure captions.

In nine point type font, capital letters are 2 mm high. **If you use the smallest point size, there should be no more than 3.2 lines/cm (8 lines/inch) vertically.** This is a minimum spacing; 2.75 lines/cm (7 lines/inch) will make the paper much more readable. Larger type sizes require correspondingly larger vertical spacing. Please do not double-space your paper. True-Type 1 fonts are preferred.

The first paragraph in each section should not be indented, but all the following paragraphs within the section should be indented as these paragraphs demonstrate.

5. CONCLUSIONS

Major headings, as for this section, should appear in all capital letters, bold face, centered in the column, with one blank line before, and one blank line after. Use a period (".") after the heading number, not a colon.

5.1. Subheadings

Subheadings should appear in sentence case and in boldface. They should start at the left margin on a separate line.

5.1.1. Sub-subheadings

Sub-subheadings, as for this paragraph, are discouraged. However, if you must use them, they should appear in they should appear in sentence case and start at the left margin on a separate line, with paragraph text beginning on the following line. They should be in italics.

6. PAGE NUMBERING, HEADER, AND FOOTER

Please do **not** paginate your manuscript. Page numbers, session numbers, and conference identification will be inserted when the manuscript is included in the proceedings. In addition, please do **not** change or remove the header and footer.

7. ILLUSTRATIONS, GRAPHS, AND PHOTOGRAPHS

Illustrations must appear within the designated margins. They may span the two columns. If possible, position illustrations at the top of columns rather than in the middle or at the bottom. Caption and number every illustration. All halftone illustrations must be clear black and white prints. Colors may be used, but they should be selected so as to be readable when printed on a black and white-only printer.

Since there are many ways, often incompatible, of including images (e.g., with experimental results) in a LaTeX document, an example of how to do this is presented in Fig. **??**.

8. EQUATIONS

Equations should be placed on separate lines and consecutively numbered with equation numbers in parentheses flush with the right margin, as illustrated in (1), which gives the homogeneous acoustic wave equation in Cartesian coordinates [1],

$$\nabla^2 p(x,y,z,t) - \frac{1}{c^2}\frac{\partial^2 p(x,y,z,t)}{\partial t^2} = 0, \qquad (1)$$

where $p(x,y,z,t)$ is an infinitesimal variation of acoustic pressure from its equilibrium value at position (x,y,z) and time t, and where c denotes the speed of sound.

Symbols in your equation should be defined before the equation appears or immediately following it. Use (1), not Eq. (1) or equation (1), except at the beginning of a sentence: "Equation (1) is ..."

9. FOOTNOTES

To help your readers, avoid using footnotes and instead include necessary peripheral observations in the text (within parentheses, if you prefer, as in this sentence). Footnotes that do appear should be placed at the bottom of the column on the page on which they are referenced. Use Times New Roman 9-point type, single-spaced.

10. REFERENCES

List and number all bibliographical references at the end of the paper. The references should be numbered in order of appearance in the document. When referring to them in the text, place the corresponding reference number in square brackets as shown at the end of this sentence [2], [3]. For LaTeX users, the use of the BibTeX style file IEEEtran.bst is recommended, as included in this template.

11. ACKNOWLEDGMENT

The preferred spelling of the word acknowledgment in America is without an "e" after the "g". Try to 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.

12. REFERENCES

[1] E. Williams, *Fourier Acoustics: Sound Radiation and Nearfield Acoustic Holography*. London, UK: Academic Press, 1999.

[2] A. Bee, C. Player, and X. Lastname, "A Correct Citation," in *Proc. of the 1st Int. Conf. (IC)*, London, UK, 2001, pp. 1119–1134.

[3] M. Smith, "A Good Journal Article," *J. Acoustical Soc. Am.*, vol. 110, no. 3, pp. 1598–1608, March 2001.