Project Title (fill in)
Data Set(s) (specify which): Indoor Wireless Localization, Hand Postures

Author 1, email   Author 2, email

May 1, 2020

Please organize your report along the lines of this template; you may use any word processing software you like, as long as you submit your report as the required pdf file described below.

Your report must be typewritten and submitted as a pdf document, in machine readable form (no scans or screen shots).

The report should not exceed 15 pages. Please note that this is **not** the recommended length **nor is it a hard limit**. You should favor readability (proper font and image sizes) over possible page issues. However, be aware that part of your grade depends on the written quality of your report, and this includes your capabilities of summarizing your work and your results (both clarity and conciseness).

You must properly cite your references. If you are following ideas from papers or from online discussion forums (such as Kaggle), be sure to clearly distinguish between what is someone else's work and what is yours and remember to cite the source. Wherever you include material that came from elsewhere, include a citation at that location (e.g., [1], or [author's last name(s)]). At the end of the report, include the reference. We suggest the IEEE format (easily found in Word or LaTeX) like this [1] or this [2]. Failure to properly acknowledge sources of what you report amounts to plagiarism and can result in substantial penalties.

For your code, if you use code from elsewhere, you must cite the source in the pdf file of your code. This can be done informally, by stating in a comment where it came from (e.g., "thanks to Kaggle kernels for the following code for function norm() [www.kaggle.com/.../...]"). Here also, failure to do this could result in plagiarism penalties.

Your code must be submitted in two different formats: a pdf file, all code in the one file, also required to be machine readable (no scans or screenshots), **and all the original code files in a zip file**, so we can run your code.

1 Abstract

A brief, informative description of your project. Include the problem, dataset(s) you used, approach (naming the pattern recognition methods you used and how you compared them), and key results (be sure to include at least your best result and the method with which you obtained it, for each dataset).

The abstract should be considered a "stand alone" section – it should be understandable on its own and include only information that is described (and supported) elsewhere in the report. The report should also be able to "stand alone" without the abstract.

Tip: many people find it works better to write the abstract last, even though it will be read first.

2 Introduction

2.1 Problem Assessment and Goals

A brief description of the data set(s) you chose, and the goals you want to achieve.

2.2 Literature Review (Optional)

If you have investigated previous or existing approaches to your problem, briefly describe them here.

3 Approach and Implementation

Report your approach and implementation details in the following subsections. You should mention which libraries and functions you used but avoid including code in your report. Your description of what your system does should be readable and understandable to a reader that isn't familiar with the functions and libraries you used but is familiar with the algorithms and techniques that were covered in EE 559. (For example, stating "we standardized all real-valued features", and also stating the functions used in your code for this, is fine; stating only the functions used in your code is not fine.)

If your project has more than one dataset:

- We suggest that you follow the outline below, describing your work on all datasets in each subsection. If you used a similar approach for different datasets, you can describe it only once for one dataset and then take a paragraph to highlight the differences to the other dataset. However, if you think it would make more sense for your project to follow this sequence of subsections separately for each dataset (e.g., 3.1-3.5 for Wireless Localization, then 3.6-3.10 for Hand Postures; or create subsubsections such as 3.1.1: Preprocessing Wireless Locations, 3.1.2 Preprocessing Hand Postures), you may do that instead.

- It is helpful to provide a table describing which classifiers and preprocessing methods were used on which dataset.

3.1 Preprocessing

Describe in detail any pre-processing techniques you used. This could include for example, data normalization, re-casting different variable types, and dealing with missing data. Also justify or explain your choice of methods.

If you used different pre-processing for different pattern recognition methods, or if you tested the same pattern recognition method with different pre-processed inputs, state so. A table or plot can be useful in these cases.

While it's unlikely that compensation for unbalanced data is pertinent to your project, if you did do any such compensation, describe the method(s) you chose, and why you chose it (or them), and show any outcomes.

3.2 Feature engineering (if applicable)

If you developed any new features, describe them here. If you developed a set of them but then refined the set afterwards, describe your methods, state any intermediate results that helped to choose among the features, and give the final set that was chosen.

If you did not develop any new features, state "Not applicable" for this section.

If your refinement of features you developed was combined with your feature dimensionality adjustment, you may combine both Sections 3.2 and 3.3 into one section, entitled "Feature engineering and dimensionality adjustment".

3.3   Feature dimensionality adjustment

Explain your approach and implementation. If you optimized over some parameter(s), or compared different methods, show and describe your work and outcomes.

3.4   Dataset Usage

Describe the procedure you followed in the use of your dataset. The description must include all stages of your work, from feature preprocessing, feature engineering, and dimensionality adjustment, to training and model selection including training and validation steps. Feel free to create subsections for each part if it seems appropriate to do so.

You should clearly state how many data points were used for training, validation set(s), and testing. For cross validation, also specify the number of folds and number of runs.

Describe how validation sets were used, and where in the process cross validation was used. If cross validation was used multiple times, specify how they were arranged (sequential loops or nested loops), and where decisions were made based on validation-set results.

Also describe where in the process the test set was used, and how many times the test set was used.

3.5   Training and Classification

Describe how you trained your model(s), the classifiers you used, and the parameters you chose.

For each classifier or learning algorithm you used, include the following. (you can optionally use a separate subsection (3.5.1, 3.5.2, etc.) for each classifier/learning algorithm you used.)

- Describe your model and your algorithm. It generally isn't necessary to repeat equations given in EE 559 just to describe the model and algorithm; however, you must give enough information to clearly define which model and algorithm (and which version of the model and algorithm) you are using. Also, if you want to refer to any equations (e.g., for your interpretation or analysis), you must include those equations in your report. If applicable, also explain what you did to adapt the method to your case.

- State the parameters of the model and how they were chosen. If a parameter is chosen by heuristics, state so. If a parameter is chosen by some model selection, optimization, or validation process, state so and describe the method.

- If you compared the degrees of freedom with the number of constraints or data points you have, describe that here.

- If you have sets of results to show for this pattern recognition method, include them here. (For a comparison of results from different pattern recognition methods, use the next subsection.)

4   Analysis: Comparison of Results, Interpretation

Present performance comparison of your different models and methods here. Include a comparison with the given baseline system(s), and with any results you found in the literature or internet. For each result, be sure to clearly state whether it is from training, cross-validation, or test set. Use table(s) and/or plots. **Do not paste print screen images**. If you used more than one dataset, you can either have different subsections for different datasets or you can directly compare how each classifier performed on each dataset.

Include your analysis and interpretation of these results. Can you explain what you observe? If not, any conjectures? Did you observe anything particularly unexpected? Note that your analysis shows your understanding of what you have done, and is an important part of your project score.

5 Contributions of each team member

If a team project, state here what the contribution of each team member was (i.e., who did what).

6 Summary and conclusions

Briefly summarize key findings, and optionally state what would be interesting or useful to do as follow-on work. Optionally, summarize some of the key things you learned while doing the project.

References

[1] *Bibliography management with bibtex*, available at `https://www.overleaf.com/learn/latex/bibliography_management_with_bibtex`

[2] Vitor Cerqueira, et al., *Combining Boosted Trees with Metafeature*, in Advances in Intelligent Data Analysis XV: 15th International Symposium, Stockholm, 2016.