Some super concise and informative title
Data Analysis Project for
Machine Learning: Basic Principles

November 27, 2017

Abstract

Precise summary of the whole report, previews the contents and results. Must be a single paragraph between 100 and 200 words.

1 Introduction

Background, problem statement, motivation, many references, description of contents. Introduces the reader to the topic and the broad context within which your research/project fits

- *What do you hope to learn from the project?*
- *What question is being addressed?*
- *Why is this task important? (motivation)*

Keep it short (half to 1 page).

2 Data analysis

Briefly describe data (class distribution, dimensionality) and how will it affect classification. Visualize the data. Don't focus too much on the meaning of the features, unless you want to.

- *Include histograms showing class distribution.*

3 Methods and experiments

Explain your whole approach (you can include a block diagram showing the steps in your process).

- *What methods/algorithms, why were the methods chosen.*
- *What evaluation methodology (cross CV, etc.).*

4 Results

Summarize the results of the experiments without discussing their implications.

- *Include both performance measures (accuracy and LogLoss).*
- *How does it perform on kaggle compared to the train data.*
- *Include a confusion matrix.*

5 Discussion/Conclusions

Interpret and explain your results

- *Discuss the relevance of the performance measures (accuracy and LogLoss) for imbalanced multiclass datasets.*
- *How the results relate to the literature.*
- *Suggestions for future research/improvement.*
- *Did the study answer your questions?*

6 References

List of all the references cited in the document

7 Appendices

Additional information that is not essential to explain your findings, but supports your work. For example, source code, additional images, mathematical derivations, etc. If you include source code, don't include the whole code, focus only on the most important parts, for example, a function implementing a specific algorithm