

COMP 551 PROJECT WRITE-UP

COMP 551, Applied Machine Learning McGill University

Anton Gladyr, Saleh Bakhit, Vasu Khanna

September 2019

Abstract

Summarize the project task and your most important findings. For example, include sentences like “In this project we investigated the performance of linear classification models on two benchmark datasets”, “We found that the logistic regression approach was achieved worse/better accuracy than LDA and was significantly faster/slower to train.” (Summarize the project task, the two datasets, and your most important findings. This should be similar to the abstract but more detailed. You should include background information and citations to relevant work (e.g., other papers analyzing these datasets).)

1 Introduction

In this miniproject you will implement two linear classification techniques—logistic regression and linear discriminant analysis (LDA)—and run these two algorithms on two distinct datasets. These two algorithms are discussed in Lectures 4 and 5, respectively. The goal is to gain experience implementing these algorithms from scratch and to get hands-on experience comparing their performance.

2 Datasets

Very briefly describe the and how you processed them. Describe the new features you come up with in detail. Highlight any possible ethical concerns that might arise when working these kinds of datasets. Note: You do not need to explicitly

verify that the data satisfies the i.i.d. assumption (or any of the other formal assumptions for linear regression).

3 Results

Describe the results of all the experiments mentioned in Task 3 (at a minimum) as well as any other interesting results you find. At a minimum you must report: 1. A discussion of how the logistic regression performance (e.g., convergence speed) depends on the learning rate. (Note: a figure would be an ideal way to report these results). 2. A comparison of the runtime and accuracy of LDA and logistic regression on both datasets. 3. Results demonstrating that the feature subset and/or new features you used improved performance.

4 Discussion and Conclusion

Summarize the key takeaways from the project and possibly directions for future investigation.

5 Statement of Contributions

State the breakdown of the workload across the team members.

References