

ITI105: Final Report Guidelines

Due date: The final report is due on Polymall on Wednesday, 21st Aug 2019 at 11:59 pm.

Abstract

The abstract is optional, depending on your available space. It should consist of 1 paragraph consisting of the motivation for the project and a high-level explanation of the methodology you used/results obtained.

Introduction

What is the problem you are solving? What is your formulation of a ML problem? Clearly state what the input and output is. Be very explicit: e.g. the input to our algorithm is an image, amplitude, patient age, rainfall measurements, grayscale video, etc., we then use a SVM, linear regression, etc. to output a predicted age, stock price, cancer type, music genre, etc.

Related work

You should find existing works in your area of choice and discuss their approaches and the strengths and weaknesses.

Dataset and Features

Describe your dataset: how many training/validation/test samples do you have? Is there any preprocessing you did? What about normalization or data augmentation? What is the resolution of your images? How is your time-series data discretized? Include a citation on where you obtained your dataset from. Depending on available space, show some examples from your dataset. You should also talk about the features you used. If you extracted features using Fourier transforms, word2vec, PCA, etc. make sure to talk about it. Try to include examples of your data in the report (e.g. include an image, show a waveform, etc.).

Methods

Describe your learning algorithms. Make sure to include relevant mathematical notation. For example, you can briefly include the SVM optimization objective/formula or say what the softmax function is. It is okay to use formulas from the lecture notes. For each algorithm, give a short description of how it works. Again, we are looking for your understanding of how these machine learning algorithms work. Additionally, if you are using a niche or cutting-edge algorithm (e.g. those not covered in the class), you may want to explain your algorithm. You need to explain clearly the rationale of choosing the algorithms for the chosen project.

Experiments/Results/Discussion

You should also give details about what (hyper)parameters you chose (e.g. why did you use X learning rate for gradient descent, what was your mini-batch size and why) and how you chose them. Did you do cross-validation, if so, how many folds? Before you list your results, make sure to list and explain what your primary metrics are: accuracy, precision, recall, AUC, etc. Provide equations for the metrics if necessary. For results, you want to have a mixture of tables and plots. If you are solving a classification problem, you should include a confusion matrix or AUC/AUPRC curves. Include performance metrics such as precision, recall, and accuracy. For regression problems, state the average

error. You should have both quantitative and qualitative results. To reiterate, you must have both quantitative and qualitative results! This includes unsupervised learning (talk with your project mentor on how to quantify unsupervised methods). Include visualizations of results, heatmaps, examples of where your algorithm failed, and a discussion of why certain algorithms failed or succeeded. In addition, explain whether you think you have overfit to your training set and what, if anything, you did to mitigate that. Make sure to discuss the figures/tables in your main text throughout this section. Your plots should include legends, axis labels, and have font sizes that are legible when printed.

Conclusion/Future Work

Summarize your report and reiterate key points. Which algorithms were the highest-performing? Why do you think that some algorithms worked better than others? For future work, if you had more time, more team members, or more computational resources, what would you explore?

Contributions

This section should describe what each team member worked on and contributed to the project.

References/Bibliography

This section should include citations for: (1) Any papers mentioned in the related work section. (2) Papers describing algorithms that you used which were not covered in class. (3) Code or libraries you downloaded and used. Acceptable format includes MLA, APA, IEEE. If you do not use one of these formats, each reference entry must include the following (preferably in this order): author(s), title, conference/journal, publisher, year.

Notes about Formatting

Feel free to adjust the specific sections according to your needs (e.g. combine introduction and related work or separate the experiments from the discussion. Please use single column layouts. The paper size is standard A4. Your font size must be greater than or equal to 10pt. When you submit your final report, it must be in **PDF format**.