

**Faculty of Natural and
Mathematical Sciences**
Department of Informatics

The Strand
Strand Campus
London WC2R 2LS
Telephone 020 7848 2145
Fax 020 7848 2851



7CCSMDPJ

Individual Project Submission

2016/17

Name: Inzamamul Haque
Student Number: 1664749
Degree Programme: MSc Data Science
Project Title: ** TITLE TOO LONG ** .
Supervisor: Dr. Nishanth Sastry, Dr. Amos Folarin, Dr. Davide Danovi
Word Count: 0

RELEASE OF PRODUCT

Following the submission of your project, the Department would like to make it publicly available via the library electronic resources. You will retain copyright of the project.

☐ **I agree** to the release of my project

☐ **I do not** agree to the release of my project

Signature:

Date: August 10, 2017

Abstract

This thesis looks at utilising Google TensorFlow's machine learning library to try and classify stem cells. Images were recorded hourly over 24 hours of cell wells containing stem cells from anonymous donors that had been exposed to various concentrations of protein. The raw images of the cell wells were processed and masked in order to segment the individual stem cells from the raw images. Stem cells that were segmented from the first raw image captured in the 24 hour period were labelled as 'normal', i.e. cells which had not yet been affected by the protein concentrations.

These individual stem cells were then trained upon by a convolutional neural network, in order to be able to predict and classify whether a new unseen stem cells was 'normal' or 'abnormal', or in other words, whether or not a stem cell had been affected by the protein concentrations. Using the count of 'normal' and 'abnormal' stem cells in a specific hour of the experiment, we are able to identify interesting metrics to assign to the activity levels of the cells exposed to varying levels of protein.

Contents

List of Figures

List of Tables

Acknowledgements

I would like to thank my supervisors Dr. Davide Danovi, Dr. Amos Folarin and of course Dr. Nishanth Nastry, for their incredible support and patience that was conveyed throughout this project process. I would also like to thank my close friends and family for supporting me throughout this project and my life thus far, and I hope for their continued warm support going forward.

1 Introduction

The Introduction is the first content section of your report. You should describe the general area (e.g., application domain) in which your project research is conducted, the motivation for conducting the research and the overall aims of the research. Be sure to outline your research questions and give a brief summary of the conclusions drawn, though the conclusions will be detailed later in the report. With the Introduction, you want to interest your reader and tell them why they should care about your research and why they should read the rest of the report. The report will be read (marked) by examiners with a technical Computer Science background, but not necessarily any knowledge of your domain, so make sure that you provide enough information for a naive reader.

2 Background

The Background section of your report should provide the reader with enough technical background so that they understand the area in which your research is conducted. This should be the kind of information that you might find in a textbook that teaches someone about the area. The next section of the report ("Related Work") is where you describe new research in your area, so think of this Background section as where you provide enough information so that the reader will be able to understand the important details contained in the Related Work section.

Example of glossary entry is SVM. CNN. Example of bibliography entry is given by Johnstone [?]. Further information can be found at: [?].

3 Related Work

The Related Work section of your report should provide a review of recent literature in the area of your research. This is distinguished from the Background section because it is typically newer and more experimental. If there are standard terms or techniques mentioned in the literature, then you can define what these are in the Background and use the Related Work section to explain how researchers have used the standard techniques as benchmarks or fundamental methodologies for their research. For example, if you review an article that describes using k-means clustering for finding appropriate groups of patients with similar sets of symptoms, then you could describe what k-means clustering is in your Background section and describe how researchers used that technique on patient data in your Related Work section. When you review literature, be sure to explain how the articles you cite are relevant to your project. Be critical—outline pros and cons of the work you are reviewing. Be clear to explain how the work you review is different from your own work. Note that you may find it easier to compare and contrast others' techniques with yours later in the report, after you have explained your own work. That is fine—just be sure to forward reference in the Related Work where you will compare to your own work (and backward reference in the later sections back to the Related Work). This can include information that you had in your Project Proposal report that was due in April, but should typically be substantially expanded from what you had in your proposal.

Examples of articles we might cite are [?] and [?].

4 Approach

The Approach section of your report should describe what you did. You should discuss your research questions in detail here, explaining for each question how you addressed each question (i.e., what techniques you used) and how you evaluated the success (or failure) of your investigation. This should include a description of the data set(s) that you used for your research (e.g., what you included in your Data Acquisition report that was due in March).

5 Results

The Results section of your report basically contains the answers to your research questions. This section should present the results of your evaluation, provided quantitatively, qualitatively and/or visually, as appropriate, followed by an analysis of the results. Discuss with your project supervisor(s) and/or domain advisor(s) how best to present your results. The main point is to make sure that it is clear to the reader what the answers to your research questions are and how you arrived at these answers.

6 Conclusion

The Conclusion is the last section of your report (other than Appendixes). In this section, you can revisit the research questions and summarise your answers. Clearly explain how your investigation and your answers are a contribution—why your work is worthy of a passing mark. Also in the Conclusion section, it is good to have subsections that highlight (a) Future Work, in case you were going to keep working on the same line of research or you wanted to recommend follow-up investigation for another student to pursue next year; and (b) Lessons Learned, where you can explain how you might do things differently if you started over, because you’ve learned valuable things along the way (these could be technical, but they could also be personal, such as organising your time better or listening to the project coordinator who told you to BACK UP your work frequently).

References

- [1] I. M. Johnstone, *Gaussian estimation: Sequence and multiresolution models*. 2011.
- [2] F. Inc., “Phage lambda: description & restriction map,” November 2008.
- [3] J. Doe, *The Title*. PhD thesis, University of Mars, 2011.
- [4] I. Johnstone and B. Silverman, “Ebayesthresh: R programs for empirical bayes thresholding,” *Journal of Statistical Software*, vol. 12, no. 8, pp. 1–38, 2005.

A Appendix: Review of Data Mining Techniques

A.1 Classification

A.2 Clustering