

This is the first of two assignments for CSM6420, and comprises 40% of the total marks for the module. It will be assessed according to the Department's assessment criteria for essays (see online Appendix AC of the Student Handbook) and also, subject to the performance of the oral presentation. In particular, marks will take account of understanding of the problem, completion of the task, and quality of both written and oral presentations. Other marks will cover knowledge of the literature, justification of the answers given, and amount of work involved.

Please submit your report through Blackboard before 1pm on Thursday, 26th March 2019.

Oral Presentations will be held in a seminar session on 20th March 2020.

The Exercise

You are tasked with producing a presentation and a report on the assessment of one of the following six papers. The choice of paper will be assigned in class.

- Giusti A, Guzzi J, Ciresan D, He FL, R, Juan P, Fontana F, Faessler M, Forster C, Schmidhuber J, Di CG, Scaramuzza D and Gambardella L, A Machine Learning Approach to Visual Perception of Forest Trails for Mobile Robots. IEEE Robotics and Automation Letters, 1:2, 2016.

The authors developed an image classifier using deep neural networks for perceiving forest trails for mobile robots, and demonstrated some preliminary results on the control of a quadrotor micro aerial vehicle.

- Eng CL., Tong JC, Tan TW, Predicting host tropism of influenza A virus proteins using random forest. BMC Med Genomics 7(S1), 2014.

This article describes the development of a computational system that uses random forests to predict the viral host for influenza given their protein sequences.

- Della Vedova ML, Tacchini E, Moret S, Ballarin G, DiPierro M and de Alfaro L, Automatic Online Fake News Detection Combining Content and Social Signals, 22nd Conference of Open Innovations Association (FRUCT), Jyväskylä, 272-279, 2018.

This paper reports a machine learning method for fake news detection in social networks by combining news content and social context features.

- Guo W, Fukatsu, T and Ninomiya S. Automated characterization of flowering dynamics in rice using field-acquired time-series RGB images, BMC Plant Methods, 11:7, 2015.

This paper presents the development of a high throughput phenotyping system for automatically extracting features for flowering dynamics in rice, using time-series image data.

- Covington P, Adams J, and Sargin E. Deep Neural Networks for YouTube Recommendations. In Proceedings of the 10th ACM Conference on Recommender Systems (RecSys '16). ACM, New York, NY, USA, 191-198, 2016.

This paper presents Google's large scale YouTube recommendation system based on deep learning.

- Schrittwieser J, Antonoglou J, Hubert T, Simonyan K, Sifre L, Schmitt S, Guez A, Lockhart E, Hassabis D, Graepel T, Lillicrap T, Silver D, Mastering Atari, Go, Chess and Shogi by Planning with a Learned Model, arXiv:1911.08265, 2019

The authors present the MuZero algorithm, which by combining tree-based search with a reinforcement learning model, achieved comparable performance of AlphaZero on various Atari games even without the knowledge of the game rules.

Your individual report will be assessed as 75% of this first assignment. Your report should address the following questions (amongst any other issues which you might like to consider):

1. Describe the main problems being solved by this paper and state the main solutions. (10%)
2. Identify the main solution component(s) relevant to machine learning in the paper and describe the methods involved. (20%)
3. What are the main issues to be considered (technical or ethical, whenever appropriate) for the application of machine learning in this scenario? (10%)
4. Select and propose an alternative machine learning approach to one of the identified components in Question 2. The alternative approach could involve using either a different model type, or a variation of the original one with modification in one or several components of the approach (e.g. input/output representation, feature extraction/selection, model architecture, optimisation methods, distance metrics, learning algorithm). Justify your proposal. (20%)
5. Discuss the limitations and propose directions for the future development of the system. (5%)

Give reasons for your answers. To complete the assignment, you may need to conduct your own research into the area, in order to understand the methods used in system and to propose alternative solutions.

The quality of report will be marked as 10%, assessing the structure and presentation, including writing style, citation and formatting issues. As a guideline, 3000 words (excluding bibliography) might be appropriate for this report. You should structure it as if it was produced for a reader who wishes to learn about the basic concepts of machine learning and their application to the development of intelligent systems. Note that a report which is not in PDF format will be penalised.

Each student is required to join a group of three or four members to make a group oral presentation to the rest of the class on your analysis. This group presentation will be assessed as 25% of this first assignment. The presentation should introduce the paper and the main methods and issues you want to discuss, and should also describe an alternative machine learning approach agreed within the group (what you chose and why you chose it). Each presentation will be 12 minutes. There will then be 3 minutes for questions and discussion following each presentation. The remaining members of the class will be the audience for the presentation and all members of the class are required to attend.

All members of the audience are strongly encouraged to provide feedback to the presenters by participating in discussions. All members of the audience will also be involved in peer-marking the presentation. Final marks for group presentation will be the average of peer and examiner marks. Individuals absent in the presentation will receive no mark for presentation. The marksheet for peer marking is available on Blackboard.

Plagiarism: One of the dangers of this assignment is the temptation to use paragraphs from web documents or papers that you have read. Please resist this temptation and do not do it. Otherwise, you will be heavily penalised. The report should be completely in your own words. If it is appropriate and absolutely necessary to include sentences and materials from elsewhere, then they should be clearly indicated as quotes, and references should be cited.

Please do not show your report to any other students.

Chuan Lu (cul) 26th February 2020