Assignment 1: Understanding the Literature

31005 Advanced Data Analytics
32513 Advanced Data Analytics Algorithms
Spring 2018

TASK

The goal of this assignment is to develop your skill in reading the data analytics literature. Being able to read and understand journal and conference papers and to critique them is an essential skill in your education in data mining, machine learning and artificial intelligence. Also, you can take this skill to other areas of computing.

You will write a report on the paper you choose. Your report will explain in your own words what the paper is about and will critique it. Your report should be in one of the two formats: Format 1: around 10 pages in 11 or 12 point single line spaced text with *NO table of contents or title page* and the following structure:

Title

Introduction

Content Describe in your own words what the paper is about. Some of the things you can consider here are: What area of machine learning or data mining is the paper concerned with? Does the title of the paper correspond to the content? What data mining or data analytics methods are considered in the paper?

Innovation How innovative is the research work in the paper? Here you should describe the "novelty" of the paper. What does the paper contribute: a new method? an algorithm? a methodology? a comparison between methods? or something else?

Technical quality How would you rate the technical quality of the work in the paper? Here you should consider the quality of the work done. For example, a paper comparing classifiers on the basis of their accuracy on a training set would be of poor quality because generally the accuracy measured on a training set is higher than the actual test error. Other indicators of poor quality might be that the results in the paper could not be replicated by someone reading the paper (because they were not described clearly enough); a comparison was made between two things but the two things couldn't fairly be compared; or conclusions were drawn from too few experiments.

Application and X-factor Do you think the application domain is appropriate for the propose technique? What other application domains could the research work be applied? Also in this section, give a couple of suggestions for further developments of the research work. Do you think the work described in the paper could spark a good discussion in class? What did you find interesting about the work?

Presentation How would you rate the quality of the presentation of the paper? Things you could consider here include: how easy it was to follow the argument in the paper; presentation style; depth of the argument; and clarity of the presentation.

References To answer the questions above in your report you will need to read other books or papers. List any references to other papers and books you read here. Referencing should be in the Harvard format. See the information at the UTS Library http://www.lib.uts.edu.au/help/referencing.

Format 2: The contents are similar to those in Format 1, but organised in individual cells of a Jupyter notebook. An example will be shown on class in Week 1.

DUE DATE

Due date 11:59pm 31 Aug 2018

How to submit Please submit using the UTS Online in the Assignments folder. If you choose Format 2, you should print out the jupyter notebook as a PDF to submit, as well as provide a link or attachment of your ipynb file. Extensions may be granted for assignments after consultation with the Subject Coordinator before the 11:59pm 24 Aug 2018.

A late penalty of up to 50% may be applied unless prior arrangements have been made with the subject coordinator. If your performance in an assessment item or items has been affected by extenuating or special circumstances beyond your control you may apply for Special Consideration. Information on how to apply can be found at http://www.uts.edu.au/current-students/managing-your-course/classes-and-assessment/special-circumstances/special.

ASSESSMENT

Group work This assignment is individual work.

Return I will endeavour to return marked assignments within three weeks.

Contribution to final mark This assignment contributes: 30% towards your final mark.

Objectives This assignment supports objectives 1 and 2 in the subject outline and Graduate Attributes B6 and E1.

Academic Standards Please see the subject outline for details on the ethical standards we expect from you.

Hours An average student should expect to spend around 20 hours to get a 50P result on this assignment.

Marking Scheme Your report will be marked based on how well you address each of the issues on the previous page as well as the quality of your report. Note that the teaching team will take the technical difficulty of the paper you choose into consideration when performing assessment. That is, we will try to maintain a consistent criteria among students regarding to the effort, skill and quality in reading, understanding the literature and presenting your own critics in reports, but *regardless* the paper being read (as long as the paper is from our list or you have discussed the choice with the teaching team).

Addressing “Content”	30%
Addressing “Innovation”	15%
Addressing “Technical Quality”	15%
Addressing “Application and X-Factor”	15%
Addressing “Presentation”	5%
Layout and presentation of your report	20%
	100%

CONTENT	100	full understanding of the work
	75	sufficient understanding of the main content (workflow or theory), but with some details of technical components missed
	50	roughly understand the content, with key components missed
	25	get the topic, with significant mis-understanding
	0	off-topic or missing this part entirely
INNOVATION	100	complete understanding the motivation (background), and the innovative point of the paper, with critical / retrospective (if the paper is of some time old) review
	75	understand the innovation compared to background, critical comments may be slightly imprecise
	50	get the innovation, and provide own comments
	25	mostly re-wording authors claim
	0	missed the innovation
TECHNICAL QUALITY	100	complete understanding the technical details, and identify advantage/disadvantage
	75	good understanding of technical details, with mostly relevant comment
	50	follow technical material, provide some comments but miss critical / challenging part
	25	rough description the technique, critical comments are bare or missing
	0	missed technical stuff
APPLICATION & X-FACTOR	100	relevant comments on the proposed application domain of the paper; proposing new applications or research ideas, which are based on careful argument and background review;
	75	relevant comments on the proposed application domain of the paper; reasonable new applications or research ideas;
	50	relevant comments on the proposed application domain of the paper
	25	identify the combination application and technique, comments missing or off-topic
	0	failed to link application with the technique
PRESENTATION	100	accurate comments on the representation and easy-to-confuse points, and on how to improve
	50	reasonable comments on the structure of the paper
	0	missing the point
PRESENTATION OF REPORT	100	Well structured; Accurate writing; Arguments are well supported either by theoretical discussion, quotation from the literature or relevant experiment results; Format / quality of submitted file (including inserted graphics) is high; NO Cover Page / Title Page or Table of Contents;
	75	Well structured; Mostly accurate writing; Arguments are supported but may contain flaws; Minor format / quality issues; NO Cover Page / Title Page or Table of Contents;

	50	Non-standard structure (harming understanding or violating standard for no good reason, such as including a cover/title page or ToC); Writing clearly needs improvement by more careful proofreading; Arguments are supported but may contain flaws; Format / quality issues.
	25	(ANY OF THE FOLLOWING ISSUE) Unclear structure; Writing contains errors that could be removed by *automatic* tools (such as Microsoft Word);
	0	Text part is very difficult to understand, non-English or contains automatically translated stuff; (If you have 0 in this section, unless you can successfully appeal on this section, we don't accept appeal on results of other sections)