<center>**CO832 Data Mining and Knowledge Discovery**</center>

<center>**Assessment 1: Short Essay – Instructions and Marking Scheme, 2019**</center>

Introduction and Essay Topic

For this assessment you are required to write a short essay, with at most two pages of text (using a font size of at least 11pt), not counting the space taken by the bibliographical References section. If your essay has more than two pages of text, only the first two pages will be marked, the other pages will be ignored.

The topic for this short essay is:

<center>"Fairness-Aware Classification Algorithms"</center>

Most classification algorithms learn a model (a classifier) with the objective of maximizing the predictive performance of the model, regardless of how fair the model's predictions are. As a result, the application of such algorithms to real-world data (particularly data about people) often leads to predictions which have a good predictive accuracy but are unfair, in the sense of discriminating (being biased) against certain groups or types of people – characterized e.g. by values of attributes like gender or socio-economic attributes. In many cases this occurs because the data used to train the classifiers contains discrimination or biases, and a conventional classification algorithm has no incentive to remove or correct for such discriminations or biases – it simply tries to maximize predictive performance.

In the last few years, however, there has been a considerable amount of research on fairness-aware classification algorithms, which takes into account the trade-off between achieving a high predictive accuracy and a high degree of fairness (avoiding as much as possible predictions that are unfair or discriminate against some groups or types of people).

Your task is to find and read research articles about fairness-aware classification algorithms. Note that in the literature such algorithms can be described by different terms, like "fairness-aware" or "discrimination-aware" algorithms. The term "fairness-aware" is arguably a better term, because "discrimination-aware" can lead to confusion with the general use of the term "discrimination" as a synonym for "classification" – actually, the area of Statistics typically uses the term "discrimination" to refer to what in data mining and machine learning is called "classification". In any case, in your search for papers you can use both the terms "fairness-aware" and "discrimination-aware", as well as its variations, of course.

Note that this is NOT an essay about the legal or social meaning of "fairness". It is an essay about fairness-aware classification algorithms from a data mining / machine learning perspective. You may include some text in your short essay about the meaning of fairness in a kind of legal or social context, but this should be some relatively brief text. The focus and the majority of the text in your short essay must be about algorithmic aspects of fairness in the classification task of data mining / machine learning. That is, you should focus on how a (type of) classification algorithm has been modified to cope with discrimination or biases in the data. You can focus on a single (type of) classification algorithm – e.g., decision trees, which were discussed in some

detail in the lectures – or you can briefly consider multiple (types of) classification algorithms, within the constraint of the limited size of this short essay.

In order to write this short essay you will need to investigate the topic using whatever resources you can find, in particular the university library and the internet (e.g. using Google Scholar to search for relevant papers).

This is an individual assessment.

Assume that the reader is familiar with the type of classification algorithm(s) that you are discussing in your essay, i.e., how a conventional version of the algorithm works. So, do NOT waste space explaining how that conventional classification algorithm works. Focus on the topic of the essay, discussing how the algorithm was adapted to cope with discrimination / biases in the data. That is, you just need to explain the part of the algorithm that was designed to produce fair classifications, and how that differs from the conventional algorithm (which ignores discrimination / biases).

Practicalities

Please print out your essay and hand it in at the Student Administration Office. The deadline for submission is specified in the Student Data System.
This assessment is worth 10% of the total marks for this module.

Time Estimated to Complete the Assessment

The time that students take to write a short essay as required in this assessment varies significantly across students; but as a rough estimate, students can be expected to spend between 15 and 20 hours to do this assessment. This estimate refers to the total time, i.e., including the time to find and read relevant papers and the time to write the essay.

Notes on Plagiarism

Senate has agreed the following definition of plagiarism:

"*Plagiarism is the act of repeating the ideas or discoveries of another as one's own. To copy sentences, phrases or even striking expressions without acknowledgement in a manner that may deceive the reader as to the source is plagiarism; to paraphrase in a manner that may deceive the reader is likewise plagiarism. Where such copying or close paraphrase has occurred the mere mention of the source in a bibliography will not be deemed sufficient acknowledgement; in each such instance it must be referred specifically to its source. Verbatim quotations must be directly acknowledged either in inverted commas or by indenting.*"

The work you submit must be your own, except where its original author is clearly referenced. We reserve the right to run checks on all submitted work in an effort to identify possible plagiarism, and take disciplinary action against anyone found to have committed plagiarism. When you use other peoples' material, you must clearly indicate the source of the material.

Notes on essay writing

- An essay should have a structure. It should begin with an introduction which summarizes the essay so that the reader has some idea of what will be covered and how it will be structured. Next, it should have one or more sections discussing the main topic of the essay; you should mark out these sections explicitly by the use of section headings. You should end with some conclusions in a separate section, which summarizes the main points made and suggest questions which the reader might want to think about or outline directions for future reading.
- You may find it useful (once you have done your reading) to sit down with a few blank sheets of paper and "brainstorm" some ideas for the essay; the important ideas which have been stressed over and over in your reading will hopefully float to the top of your imagination during this process.
- Always acknowledge quotations (if you use any). An essay can be enlivened and an argument strengthened by using quotations from other people/books/papers, but you need to note carefully where these quotations have come from. The simplest way to do this is to make a numbered alphabetical list at the end of the essay of the books and papers used, then you can refer to any quotation you use by the number of the reference.
- You can quote references obtained from both the library and the Internet. However, you should consider that some references obtained from the Internet might be very "informal", reflecting the views of the author without a peer-review process – unless the reference was also published in a more formal format. By contrast, references in the form of papers published in the proceedings of conferences and academic journals (particularly the ones specialized in the research area of the essay) tend to be more reliable, because typically they have passed through a peer-review process before publication. When you get a reference from the Internet that has also been formally published in a journal or conference proceedings (or as a book / book chapter), quote the formal reference, instead of quoting the web page address. The next item explains how to quote a formal reference.
- Mention all the relevant details of each reference that you quote. This includes the name of the author(s), the title of the paper/book/book chapter, the title of the conference proceedings (in the case of a conference paper) or the title and the number of the volume/issue of the journal where the paper was published (in the case of journal papers), the year of publication, etc. See examples of references later in this document.
- If you can draw on multiple sources of information it makes the essay stronger. This demonstrates that you can read around a topic and bring different ideas together, rather than just recapitulating what another author has written.
- Illustrations can summarize important points, provide an overview of complex concepts, and demonstrate visual ideas. You can create your own illustrations by hand or using a computer drawing package, or you can use illustrations from elsewhere, again being certain to say where the illustration came from so that it does not look like you are passing off other work as your own. Content is more important than artistic merit for the purposes of this assessment.
- Using too many words to express a simple idea and distilling your writing to the point where it becomes hard to read are equally problematic. Try and use

enough words to express your ideas and make the sentences flow, but don't overwrite.

- Don't start writing too soon. Spend some time to read about the topic and make notes before you start writing the main text. It may be useful to block off a substantial block of time (e.g. a whole day in a weekend) to write the main draft of the essay, so that you can collect your thoughts and remain "in the experience" as you are working on it.
- After you have written a draft version don't revise it straight away; put it away for a couple of days and then revise it. One obvious ramification of this piece of advice is that you have to start writing more than a day before the deadline.
- Don't be afraid to cut material out as you revise your work. Don't be afraid to have a go at rewriting a problematic section from scratch rather than trying to revise it.

Reference

[1] Higham, Nicholas J., "Handbook of Writing for the Mathematical Sciences", Second edition, SIAM, August 1998

Examples of how to specify the bibliographical details of references for each category of reference

Conference paper
[2] B. Liu, W. Hsu and S. Chen. Using general impressions to analyze discovered classification rules. *Proc. 3rd Int. Conf on Knowledge Discovery and Data Mining (KDD-97)*, 31-36. AAAI Press, 1997.

Journal paper
[3] V. Dhar, D. Chou and F. Provost. Discovering interesting patterns for investment decision making with GLOWER – a genetic learner overlaid with entropy reduction. *Data Mining and Knowledge Discovery 4(4)*, 251-280. 2000.

Book
[4] J. Han and M. Kamber. *Data Mining: concepts and techniques.* Morgan Kaufmann, 2001

Book chapter
[5] U.M. Fayyad, G. Piatetsky-Shapiro and P. Smyth. From data mining to knowledge discovery: an overview. In: U.M. Fayyad et al (Eds.) *Advances in Knowledge Discovery and Data Mining*, 1-34. AAAI/MIT, 1996.

Marking Scheme

Your essay will be assessed based on three criteria, namely:

(A) Technical Quality – This involves the technical soundness and quality of the contents of the essay. In general, the more advanced (and correct) the technical concepts and arguments of the essay, the higher the mark. *This is the most important assessment criterion.*

(B) General Organization and Comprehensibility of the Text – The text should be divided into sections (introduction, section(s) with relevant title(s), conclusion). The text should consist of well-written sentences, which are understandable and meaningful. In addition, an essay should have a good logical sequence of ideas, say starting from simple concepts and gradually increasing the complexity of the concepts or methods described in the paper, or having any other logical order appropriate for the topic of the essay. A figure or table is usually helpful, facilitating the understanding of the text by the reader.

(C) Quality of the Bibliographical References – This involves quoting references that are relevant to the topic of the essay. Ideally the majority of the references should be prestigious, refereed references (e.g., academic journal papers, or papers published in proceedings of specialized conferences). Quoting only "informal" references such as web pages is not a good approach. All relevant bibliographic details of the references should be given (authors, title of the paper, name of conference/journal/book where it was published, number of the volume and issue in the case of journals, page numbers of the paper, publisher, year (or month and year for journals), etc.).

Your essay will be assigned a mark based on a *categorical marking scale* used by the University for marking essays and similar types of coursework. More precisely, marks will be allocated as follows:

Mark range: 100, 95, 85, 78, 75, 72
Marks within that range are allocated based on the extent to which the essay has the following characteristics:

(A) *Technical Quality*: text of excellent technical quality, with very advanced technical concepts, showing clear evidence of a deep bibliographical research.

(B) *General Organization and Comprehensibility of the Text*: an excellently written essay, very easy to understand and very well organised, with some good figure(s) or illustration(s).

(C) *Quality of the Bibliographical references*: an excellent set of references: many formal, prestigious references which are very relevant to the ideas and arguments discussed in the essay, and very few or no bibliographical details are missing

Mark range: 68, 65, 62
Marks within that range are allocated based on the extent to which the essay has the following characteristics:

(A) *Technical Quality*: text with a good amount of advanced technical detail, showing evidence of a good bibliographical research, and virtually all text is technically correct.

(B) *General Organization and Comprehensibility of the Text*: a well written essay, easy to understand and well organised, with some good figure(s) or illustration(s).

(C) *Quality of the Bibliographical references*: a good set of references: several formal, prestigious references which are very relevant to the ideas and arguments discussed in the essay, and few bibliographical details are missing.

<u>Mark range: 58, 55, 52</u>
Marks within that range are allocated based on the extent to which the essay has the following characteristics:

(A) *Technical Quality*: text with, overall, a reasonable amount of technical detail and mostly technical correct, although a few parts of the text have some technical errors.

(B) *General Organization and Comprehensibility of the Text*: text of reasonable writing quality overall, but some parts of the text are not easy to understand; a reasonably organised essay, with some reasonably good figure or illustration.

(C) *Quality of the Bibliographical references*: relatively few formal, prestigious references; citing mainly informal references; overall the references are relevant to the ideas and arguments discussed in the essay, and relatively few bibliographical details are missing.

Note: the below mark ranges are below the pass mark for this module (which is 50)

<u>Mark range: 48, 45, 42</u>
Marks within that range are allocated based on the extent to which the essay has the following characteristics:

(A) Technical Quality: text with little technical detail, leaving a lot to be desired; discussion based mainly on the contents of the lectures (i.e., no significant evidence of a bibliographical research that went beyond the contents of the lectures); presence of technical errors in several parts of the text.

(B) General Organization and Comprehensibility of the Text: the text is not well written, and not well organised; many sentences are difficult to understand, virtually no clear figure or illustration.

(C) Quality of the Bibliographical references: very few formal, prestigious references, the majority of references are informal web pages, several of them have little relevance to ideas and arguments discussed in the essay, many bibliographical details are missing.

<u>Mark range: 38, 35, 32, 20, 10, 0</u>

(A) Technical Quality: text with little technical detail, superficial discussions about the topic; or text with many technical errors.

(B) General Organization and Comprehensibility of the Text: text badly organized and badly written, with ideas mentioned in a more or less random order, without a coherent structure; many sentences that are difficult to understand.

(C) Quality of the Bibliographical references: very few references, almost all of them are informal web pages (almost no prestigious reference), almost all of them have little relevance to the ideas and arguments discussed in the essay, and most bibliographical details are missing.