project canvas

Project name: Fairness in Classification

Project owner: Anders Henriksen, Oskar Eiler Wiese Christensen og Dagh Mikael

Purpose

What is the intent of this project?
Why are we doing this project?

The intent of this project is to learn about fairness in classification as well as how to write a good project. We are doing this project because we think this is an important topic for the future of the Machine Learning and AI field. As ML and AI implementations becomes a more common part of our everyday; the world calls for fair algorithms that ensure the human rights and ethics that we all deserve.

Scope

What does this project contain?
What does this project not contain?

This project contains a classifier, that will be trained with the COMPASS data. The classifier will classify the decile score from low, medium or high. Two different datasets will be passed through the classifier to find existing biases. A bias correction algorithm will then be implemented to exterminate the bias. A discussion of biasees and ethics will then be provided to ensure fair classification algorithms in the future.

Success Criteria

What do we need to achieve in order for the project to be successful?
How can the Success Criteria be measured?

In order for the project to be succesful all of the members in the group become experts within the field of fairness in classification. Furthermore, existing biases in the COMPASS data is confirmed by a classifier implemented by us and this classifier will then be corrected in order to erase the bias in the classifier The measure of succes can be measured by the knowledge we have after the project is written and done.

Milestones

When will we start the project and when is the final deadline ?
What are the key milestones and when will they occur?
How can the milestones be measured?

The project starts 12/02/2020 and ends in week 3 which can be seen in the "Kursusplan".

Key milestones:

Milestone 1: 11/03 Discorvered the bias in the data used for the COMPAS algorithm. Understood what it is and how to prove it. Found out how it makes the algorithm biased and what it means that the data is biased.

Milestone 2: 18/03 Midway - having an introduction, methods and data sections written

Milestone 3: 12/04 Implementing a classifier on the COMPASS dataset. Using correction algorithm for biases and check that it works. Implementing different classifiers and having a candidate for the final classifier that is going to be used in the project.

Milestone 4: 25/03 written feedback as well as constructive feedback (+ questions)

Milestone 5 09/05: Obtain a good understanding of project applications as well as safe AI and ethics.

Actions

Which activities need to be executed in order to reach a certain milestone?

Milestone 1:

Actions:

- Playing with data in python to get a visual understanding
- Making plots of different variables in the data
- Using the ground truth from 2-year after studies from ProPublica to prove bias.
- Writing down what the bias is, how to check it and what it means that the data is biased - furthermore, what happens to algorithms trained on biased data.

Then, the milestone is reached.

Milestone 2:

Actions:

- Searching web for State of the Art with newly learnt researching skills.
- Writing "State of the Art" section
- Having group meeting to agree upon the methods that are to be used in the project
- Writing methods section
- From milestone 1, at this point we have a good understanding of the data and should therefore be able to write the data section.

Then, the milestone is reached.

Milestone 3:

Actions:

- Using python to implement a classifier (PyTorch, Numpy, Sklearn, etc.)
- Training classifiers in GoogleColab
- Comparing the candidate classifiers.
- Use the correction algorithm on Classifier
- Get results i.e. check if the bias still exist in the classifier

Then, the milestone is reached.

Milestone 4:

Actions:

- Getting feedback from inside
- Using the feedback to improve introduction, methods and data section

Then, the milestone is reached.

Milestone 5:

Actions:

- Using supervision to get an understanding of bias and ethics in general
- Having a metting with a Philosopher will widen our knowledge within the field of ethics and AI
- Research other studies on Fairness as well as safe AI and other important areas
- Finnish the writing process

Then, the milestone is reached.

Outcome

What is the end result?

- A book
- A website
- An event

A project is the end result as well as a poster and a powerpoint presentation.

Team

Who are the team members?
What are their roles in the project?

The team members of this project are the following people: Anders Henriksen, Dagh Mikael and Oskar Wiese. They all have equal power and control of the project and all of these three people are main authors. Their roles are to code and document existing biasees and show how to correct these.

Stakeholders

Who has an interest in the success of the project?
In what way are they involved in the project?

The stakeholders of this project are the following people: Anders Henriksen, Dagh Mikael and Oskar Wiese which are all writers and manegers of the projekct.

Aase Feragen, Melanie Ganz og Sune Hannibal Holm which are the supervisors of the project.

Users

Who will benefit from the outcome of the project?

We are the ones who stands to gain from this project. We will learn how to write a good project, cite, in general learn more abou AI and ethics, confirm existing bias, how to mathematically correct them and more. Therefore, we as a group will benefit most from the project and hopefully learn a lot during the process.

Resources

What resources do we need in the project?
- Physical (office, building, server)
- Financial (money)
- Human (time, knowledge)

Physical: We need DTU buildings in order to hold meetings and work.

Financial: We do not need any money

Human: We need both time and our collabrative knowledge in order to create a succesful project.

Constraints

What are the known limitations of the project?
- Physical (office, building, server)
- Financial (money)
- Human (time, knowledge, politics)

Our physical constraints are our avaible apartments, DTU buildings and public places where we can meet as a group and work. Financially we have no constraints since the project does not have any costs. We do have some human contraints to an extent since one of the group members are trapped in China due to Corona Virus.

We have agreed on the fact that, Anders and Oskar will send assignments for Dagh until he has returned from China. We will also have Skype sessions if needed. We have constructed a group chat on Facebook which we will use as our main communication source. We have also created a Git repository where all documents and code will be available.

Risks

Which risks may occur during the project?
How do we treat these risks?

One of the risks that may occur during the project is the fact that we do not find the existing biases in the COMPASS dataset. The way to treat theese risk is to seek help from our supervisors and consult our code. Another risk is that we do not work hard enough during the 13 week period which is avoided by having weekly meeting, giving homework in the group.