

Project name

Project owner _

Purpose

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

If AI is to enhance human decision making, be it at the supreme court or on the high-way, the data used in these contexts must be as objective and free of bias as possible. In this project, we will train a neural network on the COMPAS data-set...

In AI, it is super importent that data is objective, which is very hard. If we want AI

to help human proffesions, it needs to be less biased than human, otherwise we

investigate bias in data, and implement adjustments in the algorithm, in order to try

might as well use only humans. Therefor, we are going to train an algorithm on the data,

Scope

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

Project contains: - Description of data

- Statistic analysis of data

- Implementing classification algorithm, and training it on data

- Bias adjustments of algorithm

- Examine data, by for instance investige the influence the different features has

on the classification, and see what happens if we manipulates the features - Ethic discussion of the subjects mentioned in "Purpose"

Success Criteria

What do we need to achieve in order for the project to be successful?

How can the Success Criteria be measured?

First succes criteria: To train an agloithm on our data, and investigate whether or not the algorthms has bias, and how the bias is traced back to the data.

Second succes criteria: That our discussion contains multiple opinions, and discuss from multiple views. And in addition, contains our discussion with the philosopher.

Milestones

We find it important to discuss how we can use algorithms ethical in sociaty. Discussing the ethical aspects, could give jectus and it 26thd on take person le think more thou march investige data.

When will we start the project and whitestead longitusting leverything called in and designation of june What are the key milestones and when will they occur?

and remove bias.

How can the milestones be measure Reflect upon the topic: what happens if we replace judges with AI judges, or at least happens to the meaning of bias in data (reading articles)

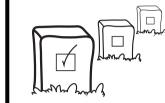
1. of march: Find information for discussion and understanding of the concept bias in data

Beginning of week 11: Implement and train algorithm

Well on our way on 6. th of april: Find bias using statistics

6. th of april: Find and read articles for discussion

Well on our way on 6th of april: Implement adjustments to remove bias



Actions

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



Find articles: Search on google, Findit, google scolar

Decide which statistics to compute Write python code in order to:

- Import data - Write a neural network

- Statistics analysis: Confusion matrix etc.

Talk to Sune Hannibal Holm



result?

- A book - A website - An event

Repport Presentation Pitch Obtain knowledge

Team

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

> Sunniva Olsrud Punsvik: Invastigator Rasmus Stokholm Bryld: Investigator Matilde Maria de Place: Investigator



Stakeholders

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

Aasa Feragen Sune Hannibal Holm Melanie Ganz Morten Mørup



Users

Who will benefit from the outcome of the project?

Society Data analysts Aasa Feragen Sune Hannibal Holm Melanie Ganz



Resources

What resources do we need in the project?

- Physical (office, building, server) DTU's supercomputer - Financial (money) Sune Hanniba Holn - Human (time, knowledge)

Aasa Feragen (superviser) Morten Mørup Overleaf 13 weeks + 3 weeks



Constraints

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

Time constraint: We only have limited time, and we have other projects. Knowledge of some of the needed libraries in python Programming experience



Risks

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

Subjective discussion What bias in data means, could be different from person to person Some might mean, that the bias we remove, is not bias.

