Algorithmic Bias

Alecia Vermeulen

T3 T4 | BRIEF

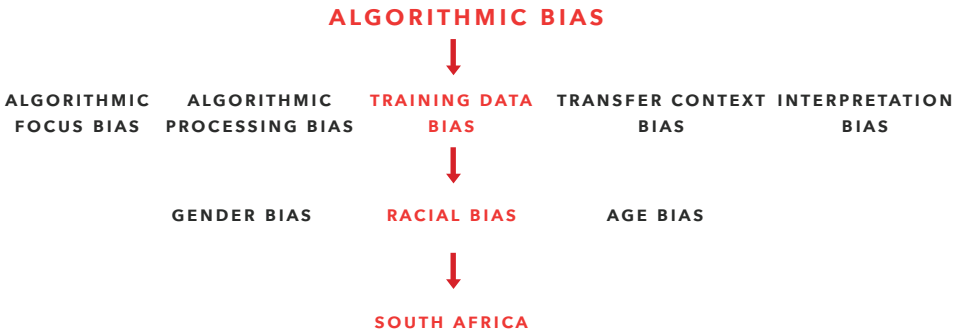
INTRODUCTION

PROJECT NAME:

Algorithmic bias in SA.

IDENTIFIED FIELD OF CONTEXT THAT IS REFERENTIAL OF RQ:

The practical component relates to the theoretical research component by also exploring the field of algorithmic bias. Specifically a training data bias will be explored. Moreover a training data bias which presents a racial bias will be explored. The practical component will explore the specified algorithmic bias in a South African context. Lastly the practical component will relate to the theoretical research component by using the findings of the theoretical research component to practically address the proposed problem.



ANALYSIS OF THE SELECTED FIELD OF CONTEXT:

There are five different types of algorithmic bias; algorithmic focus bias, algorithmic processing bias, training data bias, transfer context bias and interpretation bias. The practical component will focus on algorithmic models which exhibit a training data bias. A training data bias occurs when the input/training data used for an algorithmic model is incomplete or unrepresentative of a certain group. Moreover a training data bias could consequently present another bias such as a gender bias, race bias, age bias, etc depending on what group is incomplete and/or unrepresented in the input/training data. The practical component will solely be addressing a training data bias, which presents a racial bias.

WHY A TRAINING DATA BIAS WHICH PRESENTS A RACIAL BIAS?

Incomplete and unrepresentative data is a primary concern in South Africa where Apartheid has left a lasting effect on our data. The data therefore reflects human bias and historical injustice. Due to this the algorithmic bias, training data bias could easily manifest in a South African context . Due to the nature of why South Africa's data is unrepresentative and incomplete a subsequent bias such as a racial bias is of concern. Algorithmic bias is not something that can be circumvented overnight. There are many complexities involved in both detecting and eliminating biases from our technology. Therefore it is important to take small steps and start by evaluating and discussing one algorithmic bias at a time.

PROBLEM

THE PROBLEM:

How do I make the general public more aware of algorithmic bias, spesifically a training data bias, which presents a racial bias.

EXAMPLES OF A TRAINING DATA BIAS WHICH PRESENTS A RACIAL BIAS IN SA?

Facial recognition software

Prof Tshilidzi Marwala from the University of Johannesburg experienced complications using the application airbnb. The application required him to take a selfie, but then failed to detect his facial features several times, reinforcing that African faces are less represented in the input/training data when compared to European faces.

Voice recognition software

Prof Tshilidzi Marwala from the University of Johannesburg experienced complications using the voice-recognition software, "Google Assistant". The software could not pronounce his name due to unrepresentative data on African decendant languages and accents.

EXAMPLES OF A TRAINING DATA BIAS WHICH PRESENTS A RACIAL BIAS IN AMERICA?

The following examples are based on America because of the greater use of technology in that country, but this is not to say that similar issues and problems may not also arise in South Africa in the future.

'Predictive policing'

The algorithm is trained on historic crime data in order to predict where crimes might happen and who is most likely to commit the crime. Any past discrimination present in the collection of data will result in the software reflecting a racial bias. As a result, there is a risk that predictive policing software might not be objective, but might in fact perpetuate historic discrimination. Example: PredPol

Hiring algorithms

Algorithms used to screen job applicants are trained on data of whom succeeded in the past in the company. Therefore if there is not a lot of diversity in a company the algorithm would potentially reinforce discrimination by continuing to select individuals who reflect the companies current employees. Example: Amazon

SOLUTION

THE SOLUTION:

In order to solve the proposed problem I will be using my skills as a delevoper to create an interactive game using html, css and Javascript, which will allow the user to navigate and participate through a narrative in order to learn about algorithmic bias.

WHY AN INTERACTIVE GAME?:

I have chosen to create an interactive game becuase it is a dynamic environment that can allow users to learn, explore and respond to algorithmic bias through participation.

PROJECT DESCRIPTION:

A web-based interactive game where users can learn about algorithmic bias, spesifically a training data bias, which presents a racial bias will be created.

The game will be set in a South African context and present a scenario where a training data bias, presents a racial bias. The scenario of the game is to be confirmed during the research phase of the project.

The game will present research on algorithmic bias, spesifically a training data bias, which presents a racial bias. The necessary resources and information to be presented throughout the game will be confirmed during the research phase of the project.

The game will be interactive and allow users to participate in decision making processes. Necessary interactive components will be confirmed during the technical research phase of the project.

PHASES TO MAKE PRACTICAL PROJECT:

Week	To Do
3	Research phase
4	Research phase
5	Research phase

Week	To Do
6	Technical research phase
7	Technical research phase
8	Technical research phase

Week	To Do
9	Implementation phase
10	Implementation phase
11	Implementation phase
12	Implementation phase

Week	To Do
13	Debug & Test phase
14	Debug & Test phase

Week	To Do
15	Plan exhibition + practical document
16	Plan exhibition + practical document

DELIVERABLES

YOU WILL NEED TO HAND IN:

PRACTICAL COMPONENT

AN INTERACTIVE GAME CREATED WITH HTML, CSS AND JAVASCRIPT

PRACTICAL DOCUMENTATION

A DOCUMENT DETAILING THE PROCESS OF THE PRACTICAL COMPONENT

AIMS & OBJECTIVES

THE AIM:

The aim is to create an interactive game that will make the general public more aware of algorithmic bias, spesifically a training data bias, which presents a racial bias.

THE OBJECTIVES:

- Research interactive games and analyse how they are successful as educational tools.
- Research the technologies to be used in order to create an interactive web based game.
- Determine how algorihtmic bias, spesifically a training data bias, which presents a racial bias can be presented in an interactive game.

USERS

THE USERS:

Anyone in SA.

PHASE BREAKDOWN & TIMELINE

Week	Date	Have Ready	To Do
2	25 July, Thursday	Brief presentation	○ Fix based on feedback
3	2 August, Friday	Hand in final brief	○ Research phase
4	8 August, Thursday		○ Research phase ○ Work on Chapter 1
5	15 August, Thursday 16 August, Friday	Ch1 (With suggestions applied from proposal presentation)	○ Research phase
6	22 August, Thursday	CRIT session 1 – Show progress	○ Technical research phase ○ Work on Chapter 2
7	26 August, Thursday 29 August, Thursday	Ch2 (+ revisions for ch1 & ch2 completed)	○ Technical research phase
8	5 September, Thursday	CRIT session 2 – Show progress	○ Technical research phase

Week	Date	Have Ready	To Do
9	16 September, Monday	Ch3 draft (+ revisions for ch1 & ch2 completed)	○ Implementation phase
10	26 September, Thursday		○ Implementation phase
11	3 October, Thursday	CRIT session 3 – Show progress	○ Implementation phase ○ Work on Chapter 4
12	7 October, Monday 10 October, Thursday	Ch4 draft (+ revisions for all chapters) CRIT session 4 – Show progress	○ Implementation phase
13	17 October, Thursday		○ Work on Chapters ○ Debug & Test phase
14	21 October, Monday 24 October, Thursday	Final version (+ outline of Ch5) CRIT session 5 – Show progress + Hand in Ch5	○ Debug & Test phase ○ Finalise research paper
15	31 October, Thursday 1 November, Friday	Final paper submission	○ Plan exhibition ○ Practical document
16	7 November, Thursday	CRIT session 6 – Finalise practical & Exhibition	○ Plan exhibition ○ Practical document
	19 November, Tuesday	Exhibition & Practical document	

CHALLENGES & UNKNOWNS

THE CHALLENGE & UNKNOWNS:

A challenge of the project is to ensure that the problem of algorithmic bias is clearly communicated throughout the interactive game.

The storyline context of the game, which will frame the problem of algorthmic bias is unknown at this point. This unknown challenge will be solved during the research phase of the project. Possible storyline context options include; facial recognition, voice recognition, hiring/admission processes, recommendation procedures etc.

Lastly a big challenge for the practical project will be time management. Is is important to ensure that the schedule is followed in order to leave sufficient time to debug and test the final project before the exhibition.

RESOURCES & COSTS

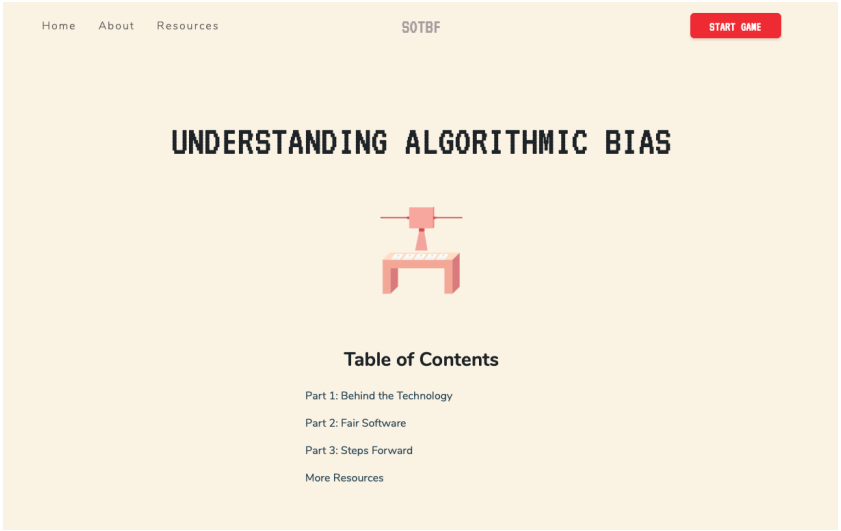
COSTS:

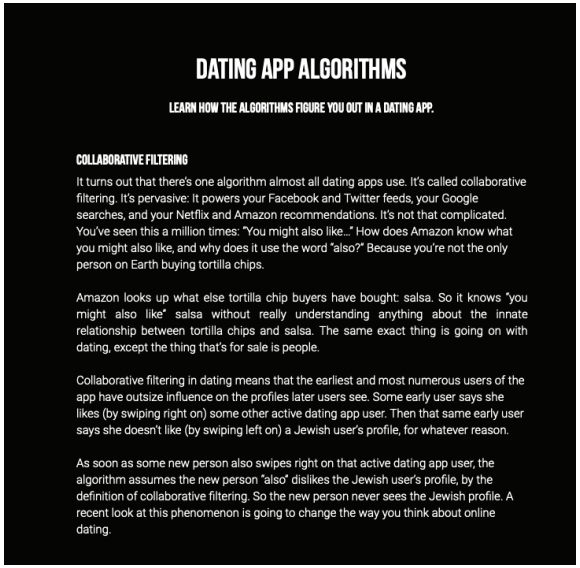
N/A

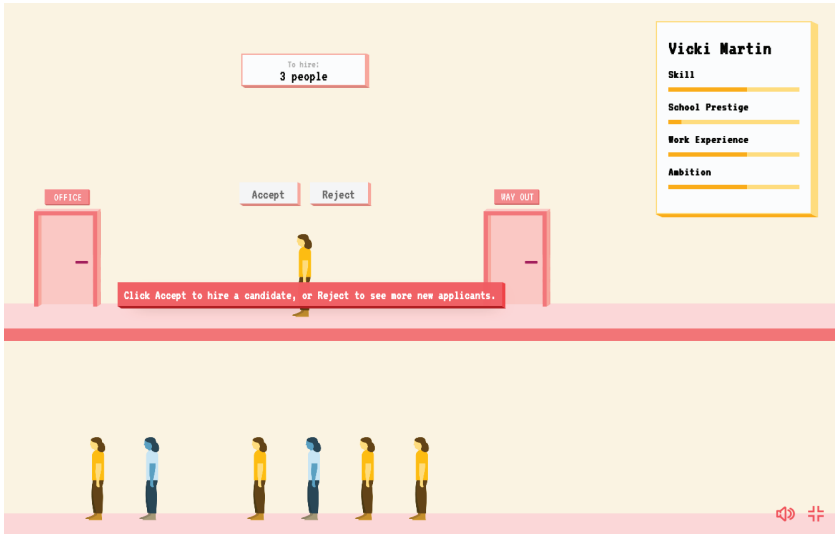
RESOURCES TO BE USED FOR PRACTICAL PROJECT:

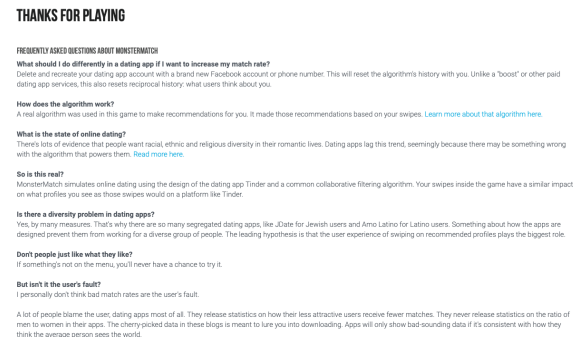
https://tutorialzine.com/2019/02/10-amazing-javascript-games
http://www.jsmadeeasy.com/javascripts/games/list_test.asp
https://www.survivalofthebestfit.com
https://github.com/survivalofthebestfit/survivalofthebestfit
https://monstermatch.hiddenswitch.com
https://implicit.harvard.edu/implicit/takeatest.html

INSPIRATION









Inspiration projects:

Survival of the Best Fit

MonsterMatch