PROGRESS OUTCOME 8

Can algorithms be biased?

Context

Erana has been working on a project in her digital technologies class. Through it she has developed a strong interest in programming and algorithms.

Insight 1: A digital concept in practice

This year I began researching the concept of the social impacts of algorithms after reading articles about how algorithms affect human behaviour. Headlines included "Algorithmic Discrimination", "Algorithms for Design" and "Algorithmic Personality Detection". Just reading the headlines was disturbing, so I wanted to find out more. I decided to investigate policing, as it would be scary if computers started policing … or would it?

Insight 2: Heat-map algorithms

I did some research into how 'heat-map' algorithms work. In policing, a heat-map algorithm tells the police where to look for criminals, based on the locations in which people have been arrested in the past. The police claim it helps allocate resources more effectively. But I discovered that data is only useful for predictive processing if you are working with a large sample or big data.

The data that police collect is like a tally: they count all the arrests they've made and display the most common locations. While most algorithms might appear to be just programs, some are designed to find trends. This is how an algorithm that identifies dominant trends could be discriminatory.

To test this, I conducted some experiments on a small scale. I looked at Google™ predictive searching and how it works. I then wrote a simple program that searched for common words to show how this might work on a large scale.

Insight 3: The key problem

After I tested my simple program to show how data can be gathered and analysed, I looked at how this could be improved. I examined the concept of "sample size disparity" and discovered that if you want to make a prediction, you need to be careful about how your algorithm treats outlying data.

For example, if a disproportionate number of people in a specific ethnic group behave in a particular way, an algorithm may connect all people from that group with that behaviour. This is like saying everyone from the "rough side of town" is a criminal.

To test this, I built a data set of different ethnicities and included their street addresses and their favourite fruit. I put in a disproportionate number of English people liking bananas, and when I created a simple query, my program predicted that English people ate bananas.

Insight 4: Applying algorithmic predictions in society

I wanted to use the results from my research and experiments to find out the wider applications. As a society, we want to reduce crime and better resource police to deal with crime. It would be great if we could predict where crimes are likely to occur, but I found that algorithms are not unbiased lines of code – an algorithm designed to pick up statistical patterns in data will report back what it is exposed to.

This means that if data includes a social bias against a minority, an algorithm is likely to incorporate this bias. This can lead to unfair or bad decisions by police or the government for members of the minority.

Insight 5: Critical evaluation of the concept and its future use

I was amazed at what I found out from my project. Nobody can agree on how an algorithm can be 'fair'. I thought algorithms were immune from bias, but they may not be because they try to make meaning of data and will predict outcomes as they process it.

There are many problems associated with this and no easy solutions, particularly because it's difficult to know how some algorithms are used. Many companies keep their algorithms secret in case they are accused of being biased or to prevent their algorithm being stolen. I think that algorithms need to be carefully developed to consider the range of social conditions that may influence their predictions.

We all want less crime and to be able to prevent crime, but we don't want to unfairly target specific groups. As computer science progresses, I think we will continue to use algorithms in fighting crime. If we keep checking the results and keep the public informed, the algorithms will become more refined.

MINISTRY OF EDUCATION
TE TĀHUHU O TE MĀTAURANGA