



Universiteit  
Leiden  
The Netherlands

# The ProgRamifications on Society

*where COMPAS guides us*

Shabnam Raziuddin  
s2521105

Mentor: Prof.dr. S. Haring

Date: 28 August 2022

Major: Computer Science  
LIACS

# Acknowledgment

During this project a different route was chosen, by thinking of a subject myself and looking for a mentor who could support me with this chosen subject, rather than choosing from the available mentors or subjects, provided to us by LIACS.

By doing so, efforts had to be made to plan out this project period. Because both parties had to figure out what the appropriate step was at many points through this project.

First of all, I would like to thank Prof.dr. H.J. van den Herik, who has not only provided me with interesting sources, thanks to his expertise in both Computer Science and Law. But also guided me with bettering my way of writing and presenting this thesis.

Secondly, I would like to thank Prof.dr. S. Haring, who guided me and supported this project from the very first day. He has helped me divide this large project into smaller and manageable portions. He has always brought joy into this project. He has provided me with insights on writing and tools for tackling projects which I will continue to use throughout my academic career.

|  |           |
|--|-----------|
| <b>Acknowledgment</b>  | <b>1</b>  |
| <b>Introduction</b>  | <b>3</b>  |
| <b>The journey toward my Research Question</b>                             | <b>6</b>  |
| <b>Chapter 1: The COMPAS Software</b>                                      | <b>8</b>  |
| 1.1 Explanation of the COMPAS software                                     | 8         |
| 1.1.1 The input of the algorithm: The Questionnaire                        | 8         |
| 1.1.2 The output of the algorithm: The Scores                              | 9         |
| 1.1.3 The Trade Secret Law   | 9         |
| 1.2 Explanation of the Conducted Experiment by ProPublica                  | 10        |
| 1.2.1 The reasoning for conducting the experiment                          | 10        |
| 1.2.2 Access of Data   | 10        |
| 1.2.3 The Conducted Experiment   | 11        |
| 1.2.4 Results of the Conducted Experiment                                  | 12        |
| <b>Chapter 2: Society and Computer Science</b>                             | <b>16</b> |
| 2.1 Computer Science and the Judicial System                               | 16        |
| 2.2 Responsibility of Computer Scientists when it comes to AI              | 17        |
| 2.3 The effect Automation has on Society                                   | 17        |
| 2.4 The effect AI has on Society   | 18        |
| 2.4.1 How AI can effect the happiness within Society                       | 18        |
| 2.4.2 How AI can keep Society in the past                                  | 19        |
| <b>Chapter 3: COMPAS on Society</b>  | <b>21</b> |
| 3.1 The effect of longer incarceration on Society (RQ1)                    | 21        |
| 3.1.1 The purpose of incarceration and how it should protect Society       | 21        |
| 3.1.2 The effect of longer incarceration on Public Safety and Society      | 22        |
| 3.1.3 The effect longer incarceration has on the incarcerated individual   | 23        |
| 3.2 The Effect of incarceration of Race, Age, and Gender on Society        | 23        |
| 3.2.1 The effect of incarceration on the basis of Race on Society (RQ2a)   | 23        |
| 3.2.2 The effect of incarceration on the basis of Age on Society (RQ2b)    | 24        |
| 3.2.3 The effect of incarceration on the basis of Gender on Society (RQ2c) | 25        |
| 3.3 Other research on the COMPAS software                                  | 25        |
| <b>Conclusion</b>  | <b>27</b> |
| <b>Discussion</b>  | <b>29</b> |
| The interpretation of the results  | 29        |
| The implication of the results   | 29        |
| The limitation of the results  | 29        |
| The recommendation of the results  | 30        |
| <b>References</b>  | <b>31</b> |

# Introduction

Artificial Intelligence is a field known and used by many. From automated vehicles, that we use to travel to work, to the analysis of medical data, which are utilised to identify potential diseases, or for the writing of sports articles<sup>1</sup>. It can be found on the Pacific Ocean in an autonomous warship of the U.S. Navy<sup>2</sup>, while it can also be found in the palm of your hand, when you scroll through a social media feed, which makes use of an AI algorithm to create a feed which matches your social media behaviour. Despite the fact that it is used in plenty of different approaches, for numerous different purposes, all these implementations stem from one single idea.

Artificial intelligence is the science of mimicking human brain capability with an algorithm. This can be done in two specific subfields. First, we have Machine Learning (ML). The main concept of ML systems is that data is used to train and learn itself. Instead of coding a program with specific instructions yourself, an ML system can learn from data, recognise patterns and make predictions. A subset of ML is Deep Learning. Deep Learning uses ML methods by using a neural network which works similarly to human decision-making. But Deep Learning can be quite expensive since it requires quite a large dataset to train itself. This large data set is required, because there are a large amount of parameters that need to be comprehended by the Deep Learning algorithm, which can in return lead to a number of false positives, more on that later in the paper. A great example of this, is the example used by H. Reese. In order to recognise what a cat looks like, many images need to be analysed to find the patterns that differentiate a cat from a lion or a tiger or a fox<sup>3</sup>.

Now that we understand the basics of these Artificial Intelligence methods (i.e., finding patterns), we are able to explain why they are used for the analysis of data. This is because the analysis can be quite a time-consuming job. Automation can increase the speed of the methods. The time-consuming jobs for humans can be completed quickly by computers.

A topic which has not been mentioned in this introduction yet, is the use of AI in the judicial system. Judges have a very important task to decide which parties are wrong and right. And in some countries, by making these decisions, they are also putting case law into place, which determines the faith of parties in the following cases. An AI system taking care of this delicate and important decision could be quite handy. That is exactly what the company Northpointe had in mind, when designing and selling the COMPAS Algorithm to states all over the United States. The Correctional Offender Management Profiling for Alternative Sanctions algorithm, like any AI algorithm, takes a specific input, in this case, a questionnaire filled in by the defendants, and produces an output, which contains three scores, the Risk of recidivism, the Risk of violent recidivism and the Risk of failure to appear.

---

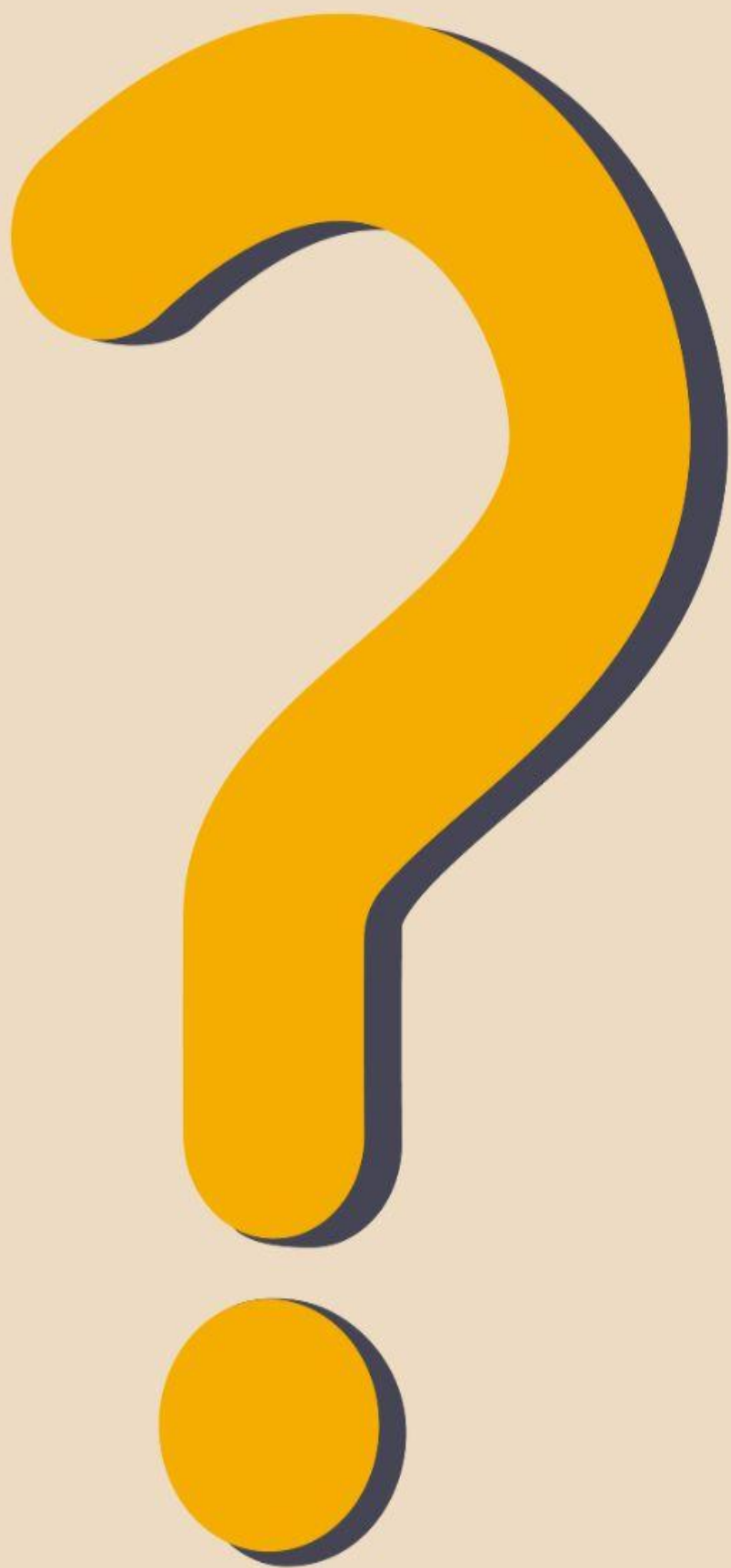
<sup>1</sup> Dysart J.(2019), 'Another AI Sportswriting Tool Hits The Market' *Robot Writers AI market*, <https://robotwritersai.com/2019/05/13/hello-world/>

<sup>2</sup> Pressman A.(2019), 'A.I., Captain': *The Robotic Navy Ship of the Future*, <https://fortune.com/longform/leidos-sea-hunter-ai-navy-ship/>

<sup>3</sup> Reese H.(2017), 'Understanding the differences between AI machine learning, and deep learning', TechRepublic p.2-4.

On these scores the judges base the shortening or extension of the sentence of the defendants. At this moment, one might believe that the exact workings of an algorithm that has quite a significant influence on the freedom of citizens, would also be known by the government. But this isn't the case. Because the producer of the COMPAS software is a private company, it is protected by the Trade Secret Law. This entails that the software, algorithm but also any data used are all protected from inspection by any third party. According to research done by the non-profit company ProPublica, there are quite some biases within the algorithm. Within this paper the effects of the COMPAS software on society will be investigated. Through this paper, we will try to figure out *The ProgRamifications on Society*.

In this paper, I will briefly explain the journey towards my research question, followed by the formulation of the research question at hand. Thereafter, in Chapter 1 we will analyse the COMPAS algorithm, as well as the experiment conducted by ProPublica. After that, in Chapter 2 there will be an explanation of the kind of effects the software has on society, according to the research that has been done. This will be supported by the effect computer science has on society followed by Chapter 3 about the effect the COMPAS software has on society. Then, a conclusion will be given where the research questions will be answered. Finally, a brief discussion on the research presented will end the thesis together with two recommendations.



## The journey toward my Research Question

Like many second-year Bachelor's Students, I too was looking for a way to utilise my third and final year of my Bachelor's. I could either opt for optional classes within my Computer Science course, or I could opt for a minor. As much as I was very interested in what these optional classes had to offer, I have always wanted to orient myself in another field, and this seemed like the perfect opportunity. So I chose the minor in Internet Law. In the following two semesters, I have seen the vast difference between both worlds. On the one hand in my Computer Science course, I learnt about the basics of programming in various languages, mostly C++. Next to programming, I've also learnt different types of algebra. And the inner workings of a computer, how it retrieves data, and how it executes the instructions we give. On the other hand, during this minor, I learnt about various other topics, from intricate guidelines set in place to protect our data to the principles of the internet, which still influence how humans interact with each other to this day. In addition, students were taught about the different ways that AI can manipulate the personal autonomy of humans.

All of these various topics left me wondering, what kind of influence programmers have on society. Better said, what kind of responsibility do programmers have towards society? Because this is quite a broad research question, I tried to focus on one of the causes for the responsibility programmers have towards society: the effect of creations of programmers write this very thesis. During my Law and AI course, I came across this very specific case about the COMPAS software in which the definition of what is fair was being questioned. It made me question not only what effect AI software has on the users, but also on the people around them and society as a whole which led me to use this case analysis for my thesis.

For this thesis, my Problem Statement is the following:

**What kind of effect could the COMPAS software have on Society when keeping the findings of ProPublica into account?**

This problem statement will be supported by the following research questions:

RQ1: What kind of effect does longer incarceration have on Society?

RQ2a: What kind of effect does incarceration on the basis of Race have on Society?

RQ2b: What kind of effect does incarceration on the basis of Age have on Society?

RQ2c: What kind of effect does incarceration on the basis of Gender have on Society?







# Chapter 1: The COMPAS Software

An automated system guiding judicial decisions could be of great use. In this chapter, an explanation of the workings of the algorithm, as well as the analysis of the experiment conducted by the company, ProPublica, will be given. We conclude this chapter with an explanation of the results of the experiment.

## 1.1 Explanation of the COMPAS software

The COMPAS software is a decision support tool. It is also used as a case management tool. *Case management*, in the legal context, is a term for all the methods which are utilised to process data which are elements within the case. An example of case management is the analysis of the data. However, primarily COMPAS is a decision support tool. A decision support tool or system, also known as DSS, is a system which supports companies with decision-making tasks.

The three basic elements of a DSS are, (1) the database, (2) the model and (3) the user interface. When speaking of models, one can think of the criteria for the decision.<sup>4</sup>

### 1.1.1 The input of the algorithm: The Questionnaire

As stated before, the input for the algorithm used by the COMPAS software is a questionnaire filled in by the defendant<sup>5</sup>. Questions asked in this questionnaire are of various different natures.

**The first series** of questions, contain questions about the current sentences. Afterwards, questions about the criminal history and non-compliance, during the parole or probation period of the defendants, are asked.

**The second series** of questions are related to the criminal defendants' family, peers, residency, education, job and leisure. Notable questions of the questionnaire are the following:

- If your parents were separated while you lived with them, at what age did this occur?
- How often do you have barely enough money to get by?
- How often did you move in the last 12 months?
- How often did you feel bored in the last 3-6 months?

---

<sup>4</sup> Power, D. J. (2002) 'Decision support systems: concepts and resources for managers' [Decision support systems: concepts and resources for managers](#), Westport, Conn., Quorum Books.

<sup>5</sup> Sample-COMPAS-Risk-Assessment-COMPAS-"CORE", contributed by Julia Angwin(ProPublica)

**The third series** of questions, contain several statements where the criminal defendant can either strongly disagree, disagree, not sure, agree, or strongly agree with. These statements are about social isolation, criminal personality, anger, and criminal attributes. Statements like the following are stated:

- I have never felt sad about things in my life.
- If people make me angry I lose my temper.
- I can be dangerous.
- A hungry person has the right to steal.

### 1.1.2 The output of the algorithm: The Scores

This algorithm compares the answers to the 137 questions and the attributes of the defendant with those of High Risk offenders. The COMPAS software then determines a number of scores, including the score for the Risk of Recidivism. Recidivism is the act of committing a crime again. These scores are given within the following three categories: Low Risk(1-4), Medium Risk (5-7) and High Risk(8-10).

The score can be found in the PSI report, which is given to the judge on the case of the defendant. Thus, the defendant's freedom is dependent on the COMPAS algorithm.

As stated previously, the defendant's faith is partially dependent on COMPAS. For that reason, one would deem it reasonable to know the workings of such an algorithm. Especially when it is used by the government, to determine such a sensitive matter as the freedom of people. But even the courts involved do not know the precise working of the algorithm. Which legally speaking is within Northpointe's rights.

### 1.1.3 The Trade Secret Law

Since the COMPAS software is developed by a private company, Trade Secret Law protects it. This implies, as mentioned before, that the algorithm itself, but also the software and used types of data are guarded against inspection from third parties<sup>6</sup>. These third parties include prosecutors, defendants themselves, but also judges who base the sentencing on the given scores.

But, the COMPAS algorithm being a trade secret wouldn't be as problematic if it handled all the defendants fairly. This, however, isn't quite the case when looking at the analysis of ProPublica.

---

<sup>6</sup> Lin, Thomas C.W. (2013). 'Executive Trade Secrets', Notre Dame Law Review. 87 (3): 911. SSRN 2047462.



Figure 3

## 1.2 Explanation of the Conducted Experiment by ProPublica

The explanation of the conducted experiment consists of four sections. In section 1.2.1 the reasoning for conducting the experiment is explained in order to get more background information on the intentions ProPublica had. In section 1.2.2 the methods used to access the data are explained in order to track back where how the data was collected and why certain decisions crucial to the experiment were made. In section 1.2.3 some more information about the conducted experiment is given. In the final section 1.2.4 the results of the experiment are discussed. In this section, we also figure out whether the COMPAS system has a bias within itself.

### 1.2.1 The reasoning for conducting the experiment

The reason why the nonprofit organisation ProPublica decided to conduct the experiment was that most of the studies done for the examination of recidivism risk methods, were done by the developers of the methods at hand themselves. According to their findings, there was no substantial set of studies which were done in the US which investigated the potential racial bias in these systems<sup>7</sup>.

### 1.2.2 Access of Data

The non-profit organisation ProPublica took it upon themselves to discover the accuracy of the algorithm as well as assess whether there were any biases within the algorithm towards a certain group.

In order to conduct this study, they needed to be able to gain access to records of a jurisdiction that was using COMPAS for pretrial release decisions. Broward County, in Florida, was a nearly perfect candidate. This is because, it is a large jurisdiction using

<sup>7</sup> Larson J., Mattue S., Kirchner L., Angwin J., (2016), 'How We Analyzed the COMPAS Recidivism Algorithm', ProPublica

COMPAS in the required scheme and it is within Florida, which has strong open-records laws<sup>8</sup>.

With a public records request, ProPublica gained access to two years of scores given by the COMPAS software. In total there were 18,610 scores, but these were reduced down to 11,757 because scores assessed at parole or probation were disregarded. Only focusing on scores which determine the release or detainment of a defendant before their trial.

As mentioned before the risk scores produced by the COMPAS algorithm could be put into three categories: Low, Medium and High.

Based on the COMPAS scores which were required, ProPublica made a profile of each person's criminal history. In order to match these profiles with the scores to the criminal records, first name, last name and date of birth were used. These criminal records were all available on the Broward County Clerk's Office Website.

A race classifier was used to determine the race. This same classifier was used by the Broward County Sheriff's Office. As stated in the report, this classifier identifies defendants as black, white, Hispanic, Asian and Native American.

### 1.2.3 The Conducted Experiment

In order to conduct the experiment, it was primarily important to determine the definition of recidivism. Within a study conducted by Northpoint themselves, they defined recidivism as "a finger-printable arrest involving a charge and a filing for any uniform crime reporting(UCR) code."<sup>9</sup> ProPublica interpreted that as a criminal offence that resulted in jail time, and also happened after the crime for which the person received the COMPAS score.

After this establishment, it was key to note down exactly which crime case was associated with the COMPAS score. When this couldn't be determined, these data points were dismissed in the experiment.

Traffic tickets and some acts of municipal ordinance violations were not seen as acts of recidivism. Municipal ordinances are rules, laws or regulations that are enacted in a town or city. Failing to appear or charges of crimes which were committed prior to the receiving of COMPAS scores, were also not considered to be acts of recidivism.

The recidivism was only seen as valid if it was within two years of receiving the scores. Since the Northpointe practitioner guide notes that the recidivism scores are supposed to predict a new misdemeanour or felony offence within two years of the COMPAS administration date.<sup>10</sup>

---

<sup>8</sup> Chapter 119, art.119.07 1(a) Public Records, The 2022 Florida Statutes

<sup>9</sup> Brennan T., Dieterich W., Ehret B.(2018), '*Evaluating the Predictive Validity of the Compas Risk and Needs Assessment System*'

<sup>10</sup> *Practitioner's Guide*,(2019), Northpointe Inc. d/b/a equivant

Next to that according to a study conducted by the U.S, Sentencing Commission, most acts of recidivism occur within the first two years of release<sup>11</sup>.

### 1.2.4 Results of the Conducted Experiment

ProPublica only analysed the Risk of Recidivism and Risk of Violence Recidivism scores.

For the risk of recidivism at first, they took a look at the distribution of scores for white and black defendants. The results can be seen in Figure 4. While the scores of black defendants were evenly distributed. The distribution for the white defendants were skewed to the left.

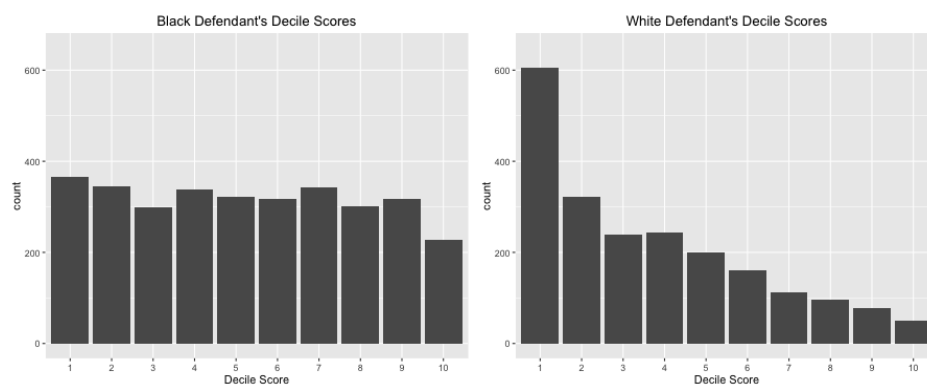


Figure 4

But in order to get a better understanding of the racial difference, ProPublica created a logistic regression model, which takes race, age, criminal history, future recidivism, charge degree, and gender into account. Logistic regression models use a type of statistical analysis which predicts a binary outcome. The result is either yes or no. It does this by analysing the relationships and correlations between the other variables<sup>12</sup>. In this case, the result was a yes or a no for a high COMPAS score and the variables were the ones stated above.

According to this logistic model, the factor which is most predictive of a higher risk score is age. Since defendants who were younger than 25 were 2.5 times more likely to get a higher score than middle-aged defendants. Keeping prior crimes, future criminality, race and gender into account. But as ProPublica was expecting, the race was also quite influential for a higher score. Even though black defendants had higher scores for recidivism, they were also 45% more likely to get a higher score than white defendants. But to their surprise, even though the female defendants had low-level crime overall. They were 19.4% more likely to get higher scores than men.

<sup>11</sup> Hunt K.S., Dumville R., (2005), 'Recidivism Among Federal Offenders: A Comprehensive Overview', United States Sentencing Commission

<sup>12</sup> Lawton G., Burns E., Rosencrane L., (2022), 'Logistic Regression',

For the analysis conducted by ProPublica, the presumed racial bias was analysed more than the other factors. In this analysis, the following was found.

As stated before, the COMPAS algorithm gave each pretrial defendant at least three scores including the Risk of Recidivism Risk of Violence and Risk of Failure to Appear. According to ProPublica's analysis, the first was predicted correctly 61% of the time, while the second was only correct 20% of the time. As one would hope, the recidivism rate for black and white defendants was approximately the same, 59% and 63%. When looking at the data like this, it may appear like there is no major bias within the algorithm, but when further analysing the cases, a different outcome is hidden behind these numbers. Though the wrong prediction percentage is similar, the way it predicted the score was on opposite sides of the spectrum. In order to explain this, we must take a look at the concept of false positives and false negatives.

*Binary classifications* are the classifying of elements of a set into either one of two groups<sup>13</sup>. The classification is done based on a classification rule. When evaluating the binary classifier, false positives and false negatives come into play. The assigned outcome of the classifier is either true or false for either one of the subgroups. Next to that, the actual outcome is either true or false. these two outcomes can be presented in a contingency table(Figure 5).

|                                 | Actual Outcome:<br>Negative | Actual Outcome:<br>Positive |
|---------------------------------|-----------------------------|-----------------------------|
| Predictive Outcome:<br>Negative | True Negative               | False Negative              |
| Predictive Outcome:<br>Positive | False Positive              | True Positive               |

Figure 5

Of these 4 outcomes for the analysis, we will focus on false positives and false negatives.

False positives are results that suggest that a certain condition exists when in actuality it does not. False positive errors are type I errors. In these cases, the test checks a single condition and wrongly indicates a positive decision.

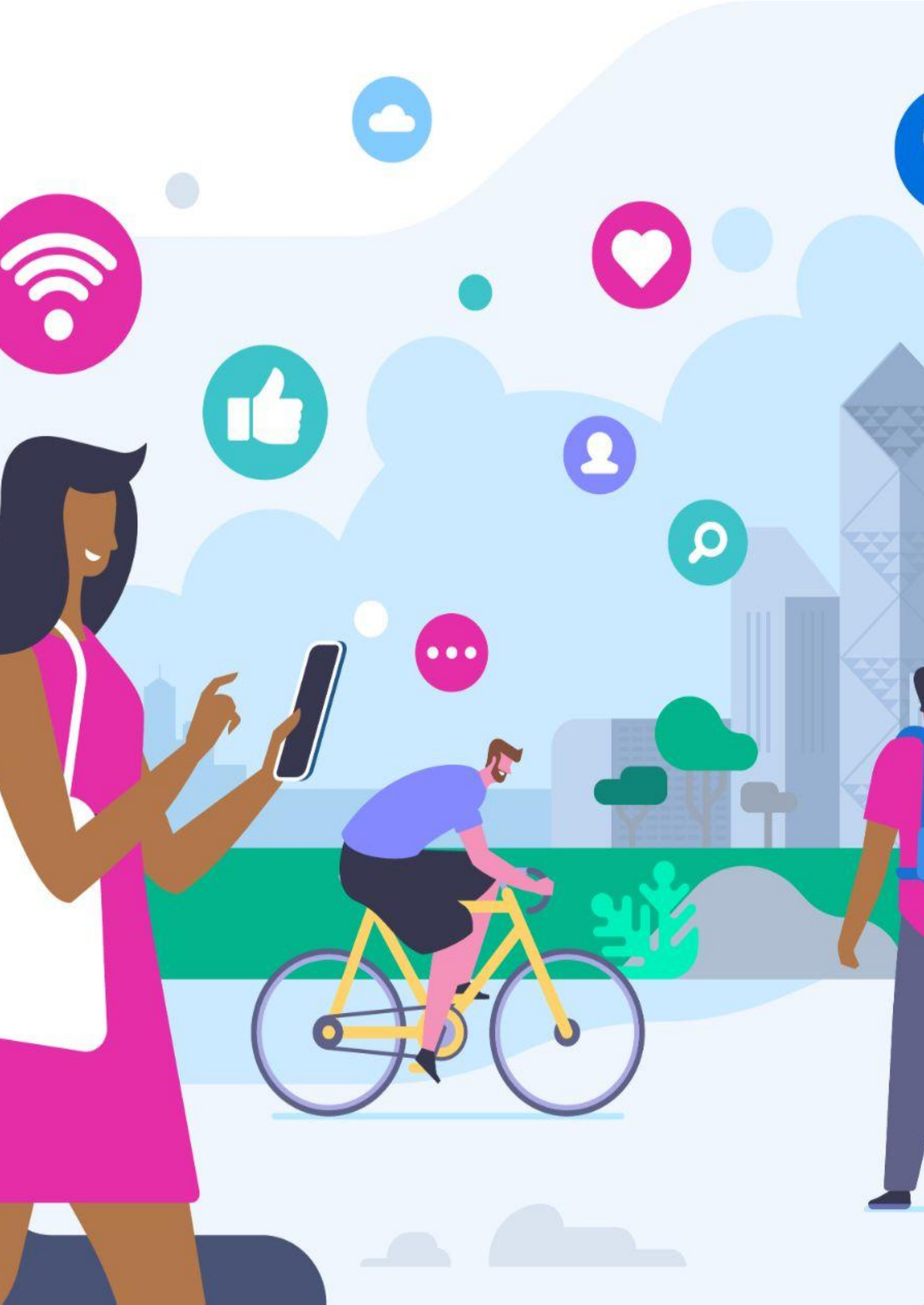
False negatives are results that incorrectly suggest that a certain condition does not exist. Likewise, false negative errors are type II errors. In those cases, the test checks a single condition and wrongly indicates a negative decision.

<sup>13</sup> Agresti A., (2018) 'Statistical Methods for Social Sciences', Pearson, Fifth Edition

As stated before, the false prediction rate of both black and white defendants was approximately the same. but the evaluation of these false predictions was not. According to ProPublica's analysis, the wrongfully predicted rates for black defendants were rated at a much higher risk than they actually were. Furthermore, the black defendants who did not recidivate within the two years were more or less twice as likely(45%) to be misclassified as High Risk compared to the white defendants (23%) in the situation.

Next to that, the wrongfully predicted rates for white defendants were rated at a lower risk than they actually were. Besides that, within the two years, white defendants who reoffended were wrongfully rated low risk almost twice as often(48%) as black re-offenders (28%).





## Chapter 2: Society and Computer Science

Firstly, I'll discuss the differences between the computer scientists and jurists. Thereafter, a brief part is devoted to the responsibility computer scientists have towards society. This is followed by the effect of automation on society. Lastly, I will discuss the effect Artificial intelligence has on society.

### 2.1 Computer Science and the Judicial System

When looking at a case which uses computer science with finding an outcome for a judicial task many thoughts may arise. Mostly, one is left to wonder, how this task is able to bridge both sides, since both fields take many years to become experts. Is it realistic to hand over judicial tasks to a computer system? Well, in order to hand over such a task to a computer system, one should notice that one is handing over the task to a computer scientist. But there are quite some differences between the two.

Firstly, a computer scientist thinks about processes, and, on the other hand, a jurist thinks about individual cases. Secondly, a computer scientist looks for similarities to automate the task at hand, while a jurist looks for differences since they are interested in the juridical aspect of a case. Thirdly, computer scientists create and test efficiency and performance while jurists don't see efficiency as a sign of success<sup>14</sup>.

These three differences are merely some of the differences between the two ways of working. The key problem is well described by J.C. Scholtes, according to him, the problem is the distance, the noise, the differences between both groups and the unknown. This problem could be solved by having a bridge, a so-called "LegalTech Bridge", which could help both parties connect. In order for this bridge to be a solution, it is important for the parties to cross the bridge. On one side, the jurists find the bridge new, unknown. They are sceptical of the unknown risks there could be. According to them, the current way of working is good as it is, so why should *they* take a new risk? On the other side, the computer scientists feel like their job is creating software and they should not have to meddle with legal matters. According to them, the jurists should cross the bridge to take care of the legal matters, while they can do what they understand and are meant to do, their niche: develop.

But this bridge needs to be crossed by both parties, only then can jurists be relevant in a digitalised world and can computer scientists make software which is accepted by society.

---

<sup>14</sup> Scholtes J.C. (2021), '*De LegalTech Brug*'. In: Custers B.H.M., Dechesne F. & Hof S. van der (Eds.) *Meesterlijk: liber amicorum ter gelegenheid van het emeritaat van prof.dr. Jaap van den Herik*. Meijers-reeks no. 376 Amsterdam: Ipskamp Publishing. 65-81.

## 2.2 Responsibility of Computer Scientists when it comes to AI

When thinking of software being accepted by society, one thing that quickly comes up are issues that question *morality*. These software programs are left to make moral decisions, which, of course, puts an important task in the hands of computer scientists.

Based on ethical principles and the Computer Science courses I've followed, Computer Scientists are responsible for the outcomes of the programs. According to them, they may keep a distance from the results, since they fed input data to the system and are only presenting what is correct according to the system. Yet they are the ones who select the data which gets used as input. They select what data gets used to train the software. They write the algorithm, which is the basis of the learning paradigm of the software. But, maybe most importantly, they are the ones which decide after x amount of tests, whether the software is accurate enough to deploy. Even when the software finds patterns independently, because of all the aspects around the algorithm, such as data selection, data supply, testing for efficiency etc. computer scientists are still very responsible for the discovered results and outputs.

## 2.3 The effect Automation has on Society

Society has become majorly dependent on automated processes. It influences the health sector, the educational sector, the government, journalism, and the environment and countless other parts of society.

When thinking of automation of society, many think of automated machines taking over the world. Or man becoming a slave to the machine. Yet this is nowhere near close because the automated systems simply are not advanced enough. In general there are three different types of AI:

- Narrow AI: These AI systems are able to do one single task very well.
- General AI: These AI systems demonstrate human intelligence.
- Super AI: These AI systems surpass human intelligence.

Currently, the used AI is at the Narrow AI level. So, thinking of the ideas shared in movies like the Terminator or Matrix is not something to worry about.

Something automation does have a significant effect on is employment. The automation of a certain task, is the replacement of human labour by an automated system. The long-term effect of automation should be unemployment. Yet the lost jobs are compensated by new positions, and automation is needed for the growing demand. These new positions could help improve the skill set of the employees. But this last point only holds as long as companies are growing fast enough to create new job opportunities for the employees.

## 2.4 The effect AI has on Society

In this section we discuss the effects that AI has on society. In section 2.4.1 we discuss how AI can effect the happiness within society by taking a look at how AI effects the workspace. In section 2.4.2 we investigate how AI can keep society in the past by unintentionally using human bias.

### 2.4.1 How AI can effect the happiness within Society

As mentioned before the use of AI will improve the productivity. According to a study done by Accenture<sup>15</sup>, the increase of productivity can be 40% while the profit for businesses will improve by an average of 38%. Next to that, AI robots can make workspace safer by performing these risky tasks. This can be jobs from exploring mars to mining work in the caves on earth.

Besides this AI can also improve the happiness of people, an individual's income has an influence on one's predicted general job and life satisfaction<sup>16</sup>. According to the conducted study by S. Moore an increase of an individual's income will only increase their utility if their position also increases. Therefore, getting a higher position because of the digitalisation of the businesses could lead to better utility and also increased happiness of employees.

The efficiency is not only beneficial to employees, but also to users. The service they utilise will work more efficiently and more accurately which in return will diminish frustration and increase time and happiness.

Even so, all these mentioned points will only be of effect in a positive manner, in the cases that the work conducted by the AI systems is done correctly. In the COMPAS case, the system does not give the correct results, for the tasks it is trying to do. Because of that, the problems it is trying to find a solution for, are left unanswered. When wrong answers or solutions are given by the AI systems, the users will still assume, because of the trust they have in such a system, that the solution and output is correct. The unanswered problem will be overshadowed by the incorrect solution. Not only is the problem unsolved, it also will stay unsolved, until someone finds a fault in the results of the system and tries to track down the problem. This will eventually lead up to either the data being at fault, the system or the interpretation of the system.

---

<sup>15</sup> Accenture (2017), 'How AI Boosts Industry Profits and Innovation', Accenture

<sup>16</sup> Boyce C.J., Brown G. D. A., Moore S. C., (2010) 'Money and Happiness: Rank of Income, Not Income, Affects Life Satisfaction'

## 2.4.2 How AI can keep Society in the past

As shown in the article of Science by Caliskan et al.(2017). If an ML system trains on human language, the machine takes over the human bias<sup>17</sup>. These biases and prejudices are fixed and anchored in our use of language. The algorithm was so accurate that neutral matter, like shops and birds, but also sensitive matter, like religion, or race were reproduced by the algorithm.

An AI system does not have the ability to utilise the desired societal changes, like women emancipation or equality of all races. This is because AI systems are trained on existing data, which makes the AI data inherently conservative. So existing power structure, patriarchy, slavery, communistic dictator but also discrimination is maintained<sup>18</sup>. When looking at the matter like this, the input data of an AI system will always be filled with bias of society, because humans living in this society are the ones feeding the system this particular data. So further development of society, bettering of society, and the future of society won't be taken into account when it comes to the characteristics of the data fed. Rather, current prejudice and biases which have been created in the past will still be underlying matter in the data. This could lead to society not being able to progress towards the desired goals, because the data which the AI systems learns from, will always be corrupted with the bias created by the past.

---

<sup>17</sup> Caliskan A. Bryson J., Narayan A., (2017), '*Semantics derived automatically from language corpora contain human-like biases*', Science

<sup>18</sup> De Bruyne J., Bouteica N., (2021) '*Artificiële intelligentie en maatschappij*', Gompel & Scavina. 1-66



## Chapter 3: COMPAS on Society

In order to answer the research question at hand, in this next chapter, a closer look will be taken into the effect of elongated incarceration on society and the effect of the main 3 factors which were most influential according to the ProPublica study: (1) race, (2) age and (3) gender, which were discussed in Chapter 1. By doing so we hope to get a better understanding of the effect the COMPAS software could have on society. Throughout these points, the connection of the effect on society will be explained by firstly understanding the effect on the individual itself.

In section 3.1 we discuss in what way longer incarceration can have an effect on society (RQ1). Within this section, we look at the reasons of incarceration, the effect of longer incarceration on society and the effect of longer incarceration on the incarcerated individual. In section 3.2 we discuss the effects of incarceration on the basis of race, age and gender on society (RQ2). Finally, in section 3.3 we discuss other research that was done on the COMPAS software.

### 3.1 The effect of longer incarceration on Society (RQ1)

Firstly, we start investigating the effect of longer incarceration on society, because we **assume** for this chapter that the COMPAS software led to longer incarceration of defendants.

This is done by first looking at the reasons for incarceration in section 3.1.1. Followed by an investigation of the effect of longer incarceration on the society in section 3.1.2. After which we discuss the effect of longer incarceration on the incarcerated individual in section 3.1.3.

#### 3.1.1 The purpose of incarceration and how it should protect Society

One may wonder what is the purpose of incarceration. Primarily there are four factors. Firstly we have **retribution**, which is the punishment an individual gets for committing a crime. Secondly, we have **rehabilitation**, which is the correcting of problematic behaviour. Thirdly we have the **safety**, which keeps the community safe, and fourthly, we have **deterrence**, which makes sure that individuals, but also others are scared off from breaking the law in the future.

Incarceration is a means to prevent inmates from committing crimes outside of prison. Next to the fact that it punishes the individual, it also protects society from harm. Around the 1970's the incarceration rate spiked under the name 'tough on crime'<sup>19</sup>. Yet crime did not reduce all too much, only a quarter of the reduction of crime is because of incarceration increasing. Next to that, the expansion of prisons in order to obtain this high incarceration rate leads to encouragement of incarceration of less serious offenders and it leads to lengthier prison terms to fill up the space as much as possible<sup>20</sup>. But lengthier sentences do

<sup>19</sup> Travis J., (2014), 'The Growth of Incarceration in the United States: Exploring Causes and Consequences', National Research Council.

<sup>20</sup> Western B., (2006), 'Punishment and Inequality in America'



not have a positive effect on society or public safety. This is because for certain crimes the individuals that are locked up can easily be replaced by others. A drug seller can easily be replaced in their community. Besides that, resources that could be used to better the community are spent on keeping certain individuals locked up longer than necessary.

### 3.1.2 The effect of longer incarceration on Public Safety and Society

Some individuals may be of the opinion that longer and harsher sentencing will make the world a safer place. Because longer sentencing gives prisoners time to think about their wrongdoings. Next to that, harsher punishment could be a motivator to not commit recidivism. But elongating the sentence only overcrowds prisons as stated before.

Longer Incarceration also leads to high costs which influence the taxpayers. According to a report by the New York University School of Law, it was estimated that \$200 billion could be saved within the last 10 years if 40% of the prison inhabitants were reduced<sup>21</sup>.

Other research shows that even with long sentencing, the reoffending remains high. According to the study done by C.L. Jonson for The University of Chicago Press, after three years in prison, 67% of the rearrested offenders were arrested for a new offence. And the time of rearrest is also not a long period<sup>22</sup>. According to the same study, 70% of the reconvicted were reconvicted within a year of their release.

Yet an important question remains: Does the length of a sentence effect Recidivism? According to a study conducted by Gottfredson et al. (1973) while on parole, offenders who had a longer sentence generally had a higher chance of committing recidivism, then offenders with a shorter sentence<sup>23</sup>.

**Incarcerating people for a longer time does not benefit the individual**, next to the fact it has a negative effect on the incarcerated it also has a negative effect on the community. As longer incarceration does not directly lead to more public safety. The costs to maintain the incarcerated also leads to more cost for taxpayers. And it does not necessarily help reduce the rate of recidivism, therefore incarcerating people for a longer time has a negative effect on society.

---

<sup>21</sup> Austin J., Eisen L., Cullen J., Frank J., (2016), *'How Many Americans Are Unnecessarily Incarcerated?'*, Brennan Center for Justice, New York University School of Law.

<sup>22</sup> Nagin D. S., Cullen F. T. Jonson C. L., (2009), *'Imprisonment and Reoffending'*, The University of Chicago Press Journal,

<sup>23</sup> Gottfredson, D.M., Neithercutt, M.G., Nuffield, J. and O'Leary, V. (1973). *'Four Thousand Lifetimes: A Study of Time Served and Parole Outcomes'* National Council on Crime and Delinquency, Davis, CA.

### 3.1.3 The effect longer incarceration has on the incarcerated individual

According to the study conducted by S. Wright, the primary problems with long sentencing is missing others who are outside of prison and the feeling of wasting and losing one's entire life<sup>24</sup>. These resulted in emotional and psychological fears related to the feelings described above.

## 3.2 The Effect of incarceration of Race, Age, and Gender on Society

According to a study conducted by the U.S. Department of Justice, Racism, Ageism and Sexism are prevalent in the Prison Community<sup>25</sup>. So the factors which were most influential according to the ProPublica study also happen to be the main causes of discrimination. These points will be used to further understand the effects the COMPAS software could have had on society.

In this next section, we are going to investigate the effect that incarceration on the basis of race has on society (RQ2a) in section 3.2.1. This section mainly focuses on the effect within certain communities. Afterwards, we will discuss the effect that incarceration on the basis of age has on society (RQ2b) in section 3.2.2. Finally, we discuss the effect that incarceration on the basis of gender has on society (RQ2c) in section 3.2.3.

### 3.2.1 The effect of incarceration on the basis of Race on Society (RQ2a)

Many would believe that removing criminals from the community is correct. Since people who do not respect the rules, that are set in place to protect the community, should not be able to benefit from the beauty of their community. But removing these individuals from the community could have a negative effect too<sup>26</sup>. This topic is thoroughly discussed by R. Crutchfield and G. Weeks in the Paper "*The Effect of Mass Incarceration on Communities of Color*" The concepts discussed are used for the following section.

If too many individuals are removed from their community and are put back there after their sentence, this could bring down the community as a whole as well. This is mainly because of the broken families that are created.

Another factor mentioned is the distrust within the community. This distrust is between the community and the law enforcement in the area, but also within the community itself. This distrust is created because of the fear of others snitching on you. But also because of the check-up rounds that are increased in the neighbourhood.

<sup>24</sup> Crewe B., Hulley S., Wright S., (2019) '*Experiencing long term imprisonment from young adulthood: identity, adaptation and penal legitimacy*', Institute of Criminology, University of Cambridge

<sup>25</sup> Goetting A., (1985) '*Racism, Sexism, and Ageism in the prison community*', Federal Probation Volume: 44 Issue: 3, U.S. Department of Justice

<sup>26</sup> Crutchfield R., Week G., (2015), '*The Effects of Mass Incarceration on Communities of Color*', Fall 2015, vol. XXII, NO. 1, Issues in Science and Technology

As one can expect it is often the poor coloured communities who have a large number of individuals who are incarcerated and later return to their community. This process is called 'coercive mobility' which could very well increase crime in the community. This could lead to crime rates going up in that area. As shown in the study conducted by D. Rose et al. there is "a positive relationship between the rate of releases one year and the community's crime rates the following year."<sup>27</sup>

The good thing about removing a criminal from these communities is that it stabilises a community again. Because living in fear and seeing your community only creates an opportunity for an unhealthy environment.

### 3.2.2 The effect of incarceration on the basis of Age on Society (RQ2b)

According to discoveries in criminology, an individual's criminal activity have a positive correlation<sup>28</sup>. According to the study conducted by A. Nellis, criminal behaviour starts in the mid-teen years and rises for a short time. For most crimes, criminal behaviour starts to decline around the early to mid-twenties.

Besides that, according to a study conducted by the Official Journal of The American Academy of Pediatrics, elongation of the incarceration period during adolescence and the early years of adulthood have a significant effect on the worsening of physical and mental health in later stages of adulthood<sup>29</sup>. As one can expect, having good mental and physical health is important to function well in society. It could lead to better integration into society. Which would help society flourish as a whole. So targeting younger defendants with longer incarceration will lead to a downfall within society.

---

<sup>27</sup>Clear T. R., Rose D. R., Waring E. & Scully K. (2003) 'Coercive mobility and crime: A preliminary examination of concentrated incarceration and social disorganization', Justice Quarterly, 20:1, 33-64, DOI: 10.1080/07418820300095451

<sup>28</sup> Nellis, A. (2017). 'Still Life: America's Increasing Use of Life and Long-Term Sentences'. The Sentencing Project, Washington, D.C.

<sup>29</sup>Dudovitz B. E., Nelsong R. B, Biely C.T., Li C., Chung N., (2017) 'How Does Incarcerating Young People Affect Their Adult Health Outcomes', Pediatrics.  
<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5260153/>

### 3.2.3 The effect of incarceration on the basis of Gender on Society (RQ2c)

When analysing gender only men and women will be regarded since those genders were regarded by ProPublica during their study. According to D.P. Mears et al.(2012) the gender of an individual has no effect on whether they do or do not commit recidivism<sup>30</sup>.

But the incarceration of women has a detrimental effect on themselves, their children and their community<sup>31</sup>. According to their conducted research, this is because of the importance of ties women have to their family and their community. Very often an effort is made to place inmates close to their relatives, but because of the lower amount of female accommodated prisons, this is not gender does not always get this possibility<sup>32</sup>. This could worsen the relationship the family members have with the individual. Especially the relationship between a mother and her children. According to R. Day et al. a mother's connectedness and involvement positively relate to the prosocial behaviours and hope of an adolescent<sup>33</sup>. These attributes could effect the community the child grows up in. Afterwards, it could effect the environment the adult becomes part of society in.

## 3.3 Other research on the COMPAS software

As discussed in the previous section, AI is only as good as the accuracy of the result it produces. It is true that the COMPAS software was very efficient for the judicial system. All the work conducted by an officer was taken over by the system at hand. But the efficiency is not of any use, since the results are not accurate.

The findings by ProPublica lead to a reaction from not only Northpointe, but also other researchers. The team led by C. Rudin from Duke University also invested their time into the COMPAS algorithm issue, according to their findings a simple algorithm with a small amount of basic set rules had the same prediction rate as the COMPAS software<sup>34</sup>.

Algorithms can have a positive effect on the justice system and society overall. But in order to have such a system, it should at least work accurately. After Chapter 3 it is clearly visible that an inaccurate biased system is not fair to the defendants, and society overall.

---

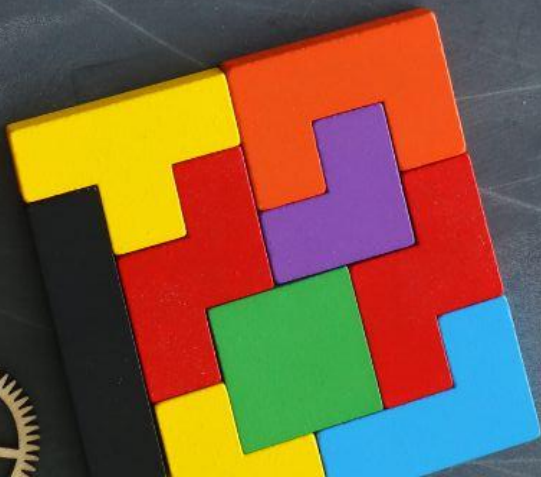
<sup>30</sup> Mears D.P, Cochran J. C., Bales W.D., (2012), 'Gender Differences in the Effects of Prison on Recidivism', *Journal of Criminal Justice* Volume: 40 Issue: 5 Dated: September/October 2012 Pages: 370-378

<sup>31</sup> Wolf A., (2006), 'Reducing the Incarceration of Women: Community-Based Alternative', U.S. Department of Justice, Office of Justice Programs,

<sup>32</sup> Ashraf S.(2022), 'What are Common Problems for Women in Prison?', *My Law Questions*

<sup>33</sup> Day, R. D., & Padilla-Walker, L. M. (2009). 'Mother and father connectedness and involvement during early adolescence'. *Journal of Family Psychology*, 23(6), 900–904.

<sup>34</sup> Angelino E., Larus-Stone N., Alabi D., Seltzer M., Rudin C., (2018) 'Learning Certifiably Optimal Rule Lists for Categorical Data', Duke University





## Conclusion

Throughout this paper, three points are emphasised during and were discussed extensively. They embody the effect of automation on society (in particular on race age and gender) and the abilities of AI at this moment. All the discussed points will help us build up to an answer for our Research Questions.

The COMPAS software is a good initiative and concept that could help the further automation of tasks. But like any automation, in order to be of any significant help, it should be accurate. The automation itself could have a positive effect on society, since it creates time for the previous employees who were responsible for the task, to do other more advanced tasks. Having a more advanced role with a higher position could also increase the work ethic, job satisfaction and happiness of the individual. Which leads to positive effects on society overall.

In the following section we will answer the supporting Research Questions:

RQ1: Longer incarceration, because of the algorithm, does not necessarily keep society safe. Since longer incarceration does not reduce recidivism. On top of that, it worsens the mental health and psychological integrity of the individual.

The three most influential factors of the algorithm according to ProPublica could also have a significant effect on society.

RQ2a: Keeping race into account while incarcerating could tear down a community, because of the instability within families and the community as a whole. The community gets filled with distrust within the community and towards law enforcement, which produces an unstable neighbourhood.

RQ2b: Keeping age into account while incarcerating could worsen the mental and physical health of an individual in their lives later, even though criminal tendencies and behaviour reduce by the mid-twenties.

RQ2c: Keeping gender into account while incarcerating could terribly influence the family and children of the incarcerated individual. Because a mother's influence is really important for the development of a child.

Overall the COMPAS software could have a bad influence on society when keeping ProPublica's findings into account. Even though it could automate a task, the automation is only as good as the accuracy of the results.

Algorithms can have a positive effect on the justice system and society overall. But in order to have such a system, it should at least work accurately. It is not fair to the defendants and society overall.





## Discussion

In this next part, the interpretation, implication, limitation and recommendation of this result will be discussed.

### The interpretation of the results

Firstly, the results that are conducted in this research are only as good as my research ability. Several papers and articles were read in order to answer the research question at hand, yet my personal bias and prior thoughts on the case at hand could have influenced the choice of research papers. Furthermore, the prior thoughts about this case also influenced the kind of answers I was looking for in these papers. It is notable that several points I had not thought of were mentioned in this paper, which were discovered throughout this research.

### The implication of the results

These results are important because when products or algorithms are created by private companies, profit and the client's wishes are kept in mind. Society's well-being should be given more importance. Especially when a product influences the freedom, mental and physical well-being of individuals, the flourishing and safety of their community, and society overall. It is important to thoroughly test the product and better its accuracy. The results are important because Society is not thought of enough when creating software products. And it is important to give it importance if the software products want support from society.

### The limitation of the results

As mentioned before, this may be less reliable since the findings are based on my own prior thoughts while going through this case. Next to that, the points which were further researched are based off of the results of the research conducted by ProPublica: race, age and gender. There could be a bias within their study, since the motivation mentioned by ProPublica for this specific study was finding racial bias within the assigning of the risk scores. Since they had not previously seen such a study before<sup>35</sup>, which examined racial bias within risk score methods. This could be a very clear example of confirmation bias. Which is a term used when seeking or interpreting findings in such a way, which affirm existing beliefs, the hypothesis or the expectations<sup>36</sup>.

---

<sup>35</sup> Larson J., Mattue S., Kirchner L., Angwin J., (2016), 'How We Analyzed the COMPAS Recidivism Algorithm', ProPublica

<sup>36</sup> Alloy, L. B. , & Abramson, L. Y. (1980). *The cognitive component of human helplessness and depression: A critical analysis*. In Garber, J. & Seligman, M. E. P. , Human helplessness: Theory and applications New York: Academic Press.

## The recommendation of the results

From the conducted research, the first recommendation would be how to diminish such problems within the software. Education of young Computer Science students could be an interesting point to start. Since I personally did not have many courses talking about the effect my work could have on society, and how I should keep myself and others within the community accountable. Therefore it would be an interesting point to research how to better educate the students.

A second recommendation would be to investigate the direct communities of the wrongly rated individuals, to get a better idea of the direct effect on these communities.

## References

In this section all used references for the images which were used can be found.

Figure 1: Problem and root cause analysis, research and leadership skill to find solution or answer for business problem concept, smart businessman analyst using magnifying glass to analyze question mark sign, Nuthawut

Figure 2: Artificial intelligence (AI), machine learning and modern computer technologies concepts. Business, Technology, Internet and network concept, Denis Putilov

Figure 3: Getty Images

<https://www.motherjones.com/crime-justice/2018/01/compas-software-racial-bias-inaccurate-predicting-recidivism/>

Figure 4: How We Analyzed the COMPAS Recidivism Algorithm (2016), J. Larson, S. Mattu, L. Kirchner and J. Angwin

Figure 6: Networking people concept, Ico Maker

Figure 7: Group side silhouette men and women of different culture, Sylisia

Figure 8: Education concept image. Creative idea and innovation. Crumpled paper as light bulb metaphor over blackboard, Tomertu

Figure 9: Bulles, Julien Eichinger