COPENHAGEN BUSINESS SCHOOL
HANDELSHØJSKOLEN

Job seeker profiling tool and its impact on society

Applied Economics and Finance

Tarek K. Ghanoum
Master's Thesis, Spring 2021

Student Number: 111781

Number of CBS standard pages: 72

Number of characters: 164,380

Supervisor: Monica Mundada

Date: 01.05.2021

Danish Agency for Labour Market and Recruitment

Abstract

Since 2015, the Danish Agency for Labour Market and Recruitment (STAR) has used a profile clarification tool on job seekers in order to help the unemployed return to the labor market. The tool assesses the citizen's risk of long-term unemployment. Based on the tool's results, the caseworker can, together with the citizen, organize a process that is adapted to the individual to get them back into work (this interaction takes place at the job centre, where the unemployed citizen receives unemployment benefits). This thesis outlines the motivation and logic behind the tool while also analyzing the social impact. The findings outline the difference in profiling across borders. While Denmark has chosen a transparent approach, other countries have decided not to disclose the features used in the analysis or the final classification, pointing at demotivation, stigmatization, and the fear of self-fulfilling prophecies. Another finding concerns the social impact, where I examine the excessive number of conversations that non-Western immigrants' descendants deal with based on origin, which suggests that more conversations are being held with these descendants than with citizens of Danish origin. Lastly, I present how STAR can remove potential discriminatory features by introducing a Random Forest model where origin does not have significant importance (removing origin from the data set decreases the precision from 70.13% to 70.10%). Based on the results, my recommendation to STAR is to undergo steps to substitute the current model with a Random Forest model.

Key words: Profiling tool, statistical discrimination, social impact, artificial intelligence, supervised machine learning

Acknowledgements

The thesis was written in collaboration with the Danish Agency for Labour Market and Recruitment. I would like to thank my supervisor Monica Mundada for the effort she put into supporting this study, my workforce for the data, and my wife for moral support.

Contents

CHAPTER 1

Introduction

1.1 Motivation

At the end of 2019 and during 2020, several complaints and a lawsuit were filed against Danish municipalities for discriminating against their citizens based on origin (*Retsinformation*, 2020). The municipalities use a profile clarification tool to help the unemployed return to the labor market. Based on several types of input data (generally referred to as *features* in the field of artificial intelligence), the profile clarification tool informs the citizen about their risk of long-term unemployment (henceforth LTU). Based on the tool's results, the caseworker can, together with the citizen, organize a process adapted to the individual to get them back into work (this interaction takes place at the job centre, where the unemployed citizen receives unemployment benefits). One of the most significant features used when calculating the risk of LTU is the citizen's origin. Based on fear of being discriminated against on the basis of origin, several people raised their voices on social media, advocating the removal of the origin variable from the profile clarification tool (Bostrup, 2021).

The criticism caught the Department of Human Rights' attention, which brought the Danish Agency for Labour Market and Recruitment to the Equal Treatment Board (Menneskeret, 2020). The agency, which is abbreviated as STAR (Styrelsen for Arbejdsmarked og Rekruttering), owns the profile clarification tool and is therefore seen as responsible for the discrimination.

Based on critique from human rights organizations, STAR is now considering removing the origin variable from its calculations while also using the situation as an opportunity to improve the profile clarification tool by exploring different machine learning tools and comparing it with similar tools used in foreign countries. As an employee at STAR, my attention was caught by the media coverage. Therefore, I volunteered to take on the responsibility of researching perceived statistical discrimination while also investigating the possibility of developing a new profile clarification tool.

1.2 Research Questions

The thesis seeks to contribute to the field of data science by showing how methods from machine learning can be used to categorize job seekers in terms of LTU while also expanding our knowledge of how statistical discrimination through the use of machine learning can have a social impact. Therefore, based on the wishes of STAR and my own curiosity, I chose to formulate the following research questions:

1. How can the Danish Agency for Labour Market and Recruitment remove potential discriminatory features while maintaining high precision?
2. What is the social impact of job seeker profiling tools?
3. How does the job seeker profiling tool compare to similar tools used in foreign countries?

1.3 Overview of the thesis

The thesis is comprised of four chapters, besides the introduction:

Chapter 2 In this chapter, I dive into the history behind the profile clarification tool, make an international comparison, and present related research and contributions to the subject of statistical discrimination. The chapter lays the foundations of the study and functions as a springboard into the subsequent chapters.

Chapter 3 The chapter consists of a theoretical and practical aspect. The theoretical aspect introduces the reader to supervised machine learning, while the practical part of the chapter takes the reader through the steps of data collection, preparation, modeling, and analysis.

Chapter 4 The results of the implementation are presented, and different models are compared and evaluated.

Chapter 5 The discussion, followed by the conclusion, presents a critical review of the methods used and the results. Afterward, I write about my contribution and the research limitations before concluding with recommendations for further research.

1.4 Delimitation

My quantitative analysis will be based on theory regarding supervised machine learning, where decision trees, random forest, and other tools will be inspected. Throughout the dissertation, parallels will sometimes be drawn with foreign countries in order to compare their features, tools, and solutions, but the dissertation will mainly deal with and focus on Denmark. This will be especially clear when I further investigate the use of the profile clarification tool in the job centres and the perceived discrimination among Danes.

The thesis does not seek to provide an exhaustive overview of machine learning methods, and for legal and ethical reasons, the thesis does not evaluate the performance of the presented machine learning tool among real users using online tests.

CHAPTER 2

Background

2.1 The Employment Reform

According to recent numbers from jobindsats.dk, which delivers data on the Danish labor market, the share in employment six months after unemployment is 49.2% (Appendix A, Figure A.2). The data, which is pre-corona, tells us that almost half of job seekers are workers who are only unemployed for a brief period before finding their way to a new job. Therefore, it is inefficient to allow these individuals to participate in costly activation programs. On the other hand, other job seekers remain unemployed for a long time because they either cannot or will not find a new job. It would seem sensible to focus active labor market policies on this group of unemployed individuals.

In June 2014, a new employment reform was created, which granted more freedom and flexibility to municipalities and unemployment insurance funds in their work with the unemployed. The new reform's focus was on individualizing the offers proposed by the municipalities and unemployment insurance funds. By organizing individually-tailored interventions for the unemployed, the reform has shifted away from the previous system, which offered the same activities for all unemployed individuals regardless of education, work experience, and connection to the labor market (STAR, 2014).

As part of the reform, a statistical profile clarification tool, called *profilafklaringsværktøjet* in Danish, was developed and implemented in ultimo 2015. The profile clarification tool, subsequently referred to as *the screening tool*, was developed and put forward by the Danish Agency for Labour Market and Recruitment, hereafter abbreviated to *STAR*. Based on administrative records and a voluntary questionnaire, a probability is calculated, stating the risk of LTU (the questionnaire was introduced in 2017). The results act as a basis for the initial meeting between the unemployed individual and the caseworker at the job centre.

According to STAR, the screening tool's purpose is to help the unemployed prepare themselves as best as possible for an active and job-oriented contact process and establish a common starting point for the process between the citizen and the caseworker in the subsequent job interviews. Based on the screening tool results and other relevant tools, the caseworker organizes an effort adapted to the unemployed individual's specific needs. The number of conversations with the job seeker, the choice of offers, or job support activities are all options that the caseworker can use to help the unemployed return to the labor market. However, according to the employment reform, a job seeker must receive a minimum of one conversation with a caseworker every month for the first six months (Appendix B).

2.2 Unemployment compensation process

There are several types of unemployment benefits, and the compensation process may differ depending on the type. The employment reform only applies to the insured unemployed, who are eligible to receive *'dagpenge'*. The entitlement conditions entail that a person must have been a member of an unemployment insurance fund (called *A-kasse* for at least one year - membership is free for college students. There are two types of dagpenge: full-time insurance, which requires that the unemployed has earned at least DKK 243,996 over the past three years, and part-time insurance, which requires earnings of DKK 162,660 over the same period (A-kasser, 2020).

Having met the conditions, the unemployed can now expect to receive an amount per month ranging between DKK 6,441 and 23,289 (before taxes) depending on the type of dagpenge and whether the individual is a provider or not. The amount can be received for a maximum period of two years, after which the individual will be required to do 1,924 hours of work before being eligible to receive dagpenge again (A-kasser, 2020).

My thesis' primary focus is on the insured unemployed; therefore, I will split the compensation process and highlight where the screening tool comes into play. Figure 2.1 provides an overview of the process and will be referred to in the sub-sections.

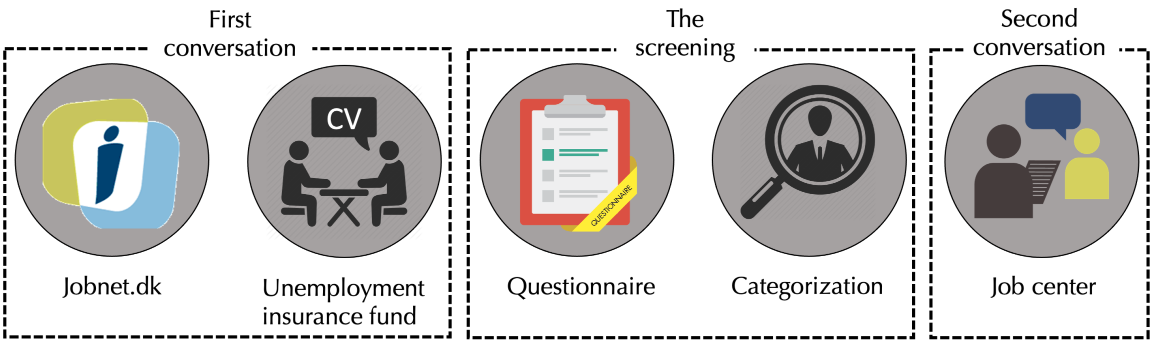


Figure 2.1: Unemployment compensation process

First conversation

The first step for the unemployed is to register as a job seeker at jobnet.dk, which is the public job center's website for all job seekers and employers in Denmark. A person can either register digitally or physically by showing up at the job center. The website gives the unemployed an overview of a person's CV, job search statistics, job search information, and a way to contact caseworkers.

A CV containing information about competencies, experiences, and other relevant matters must be filled out within the first two weeks after registration at jobnet.dk. The unemployment insurance fund must arrange an interview with the recipient of unemployment benefits no later than two weeks after his or her registration as a job seeker, where it is ensured that the CV information on jobnet.dk is adequate. In addition to ensuring that the CV is correctly filled in, the unemployment insurance fund will inform the job seeker about job search requirements every week and the overall process for the first six months (A-kasser, 2020).

The screening

After having registered as a job seeker a questionnaire becomes available to fill out at jobnet.dk. The questionnaire is made available for the first three months and is voluntary for the job seeker. According to STAR, the questionnaire's purpose is to uncover how the job seeker sees their job situation, to help the unemployed and the job consultant to prepare for the first interview, and to support the job consultant in assessing the unemployed's opportunities to get a job.

Given that the questionnaire is voluntary, the unemployment insurance fund can only recommend and remind the job seeker to fill out the form before the first conversation at the job center. The preparation form questions, displayed in table 2.1 below, were selected and formulated based on knowledge from research, STAR's own analyses, and the concrete experiences of municipalities and unemployment insurance funds (STAR, 2020).

No.	Questions	Reply options
1	What is your highest completed education qualification?	18 education categories
1a	In which field of study were you educated? (based on your highest completed education qualification)	• Humanities, religion or aesthetics • Non-vocational • Industry or craftsman • Agricultural-, fisheries- and food • Science • Pedagogical • Community-, office- and business • Health • Transport and communication engineering • Education on public safety • Other
1b	In which field of study were you educated? (based on your highest completed education)	• Non-vocational education (e.g. HF, STX) • Community education, office education and trade education (e.g. HHX) • Industrial and craft technical education (e.g. HTX)

No.	Questions	Reply options
2	How quickly do you think you will get a job (tick only one):	• I have a new job, but have not started yet • Within 1 month • Within 3 months • Within 6 months • It will take more than 6 months • I expect to go on maternity leave soon • I expect to retire soon • Do not know
3	Were you fired, or did you resign, from your previous job?	• I was fired • I resigned • Until now, I was self-employed • I did not have a job (e.g. if you have received sickness or maternity benefits, etc., or have interrupted your education) • I am a graduate • I have been temporarily sent home • Other
4	For how long have you known that you will be unemployed?	• I've known it for less than a week • I've known it for the past month • I've known it for the last three months • I've known it for the previous six months • I've known for over six months
5	How many jobs have you applied to since you knew you were unemployed until now?	• I have not applied for a job • I have applied for 1 job • I have applied for 2-5 jobs • I have applied for over 5 jobs
6	How do you assess your job opportunities in the areas where you are looking for work?	• My job opportunities are good • My job opportunities are fairly good • My job opportunities are not so good • My job opportunities are poor • Do not know
7	Do you know what job you would like to have?	• I know exactly what job I would like to have • I have an idea of what types of work I would like to do • I have a few vague ideas about what kinds of work I would like to do • I do not know what kinds of work I would like to do

No.	Questions	Reply options
8	How do you apply for / plan to look for work? (Feel free to put multiple ticks)	• Respond to advertisements in newspapers, magazines, the internet, etc. • Send unsolicited applications • Contact the employer directly or by telephone • Use social media (Facebook, LinkedIn, etc.) • Use my network (contacted friends, family, former study or work colleagues, etc.) • Temporary employment agency
9	Which of the following considerations have you taken to increase your job opportunities? (Feel free to put multiple ticks)	• I am considering changing industry • I am considering commuting long distances • I am considering moving • I am considering taking another education • I am considering going down in salary • I am not considering any of the above
10	Is there anything that makes it difficult for you to get a job?	• Yes • No • Don't know
10a	If "yes" in question 10: What makes it difficult for you to get a job? (Feel free to put multiple ticks)	• My physical health • My mental health • Personal or social matters • Financial conditions • Alcohol or drug abuse • Lack of reading, spelling or arithmetic skills • Lack of IT skills • Language barriers • Other, please specify

Table 2.1: Questionnaire (STAR, 2020)

In addition to the questionnaire, administrative records are used as part of the analysis. Figure 2.2 below illustrates the combination of the two sources of data. Information in the form of demographic characteristics, labor market history, and income is added to the analysis and displayed to both the unemployed and the caseworker as shown in Appendix A Figure A.1. Assuming that the unemployed filled out the questionnaire, a results section will be displayed at jobnet.dk with two possible outcomes. The first outcome is called *Dagpengemodtagere i risiko for langtidsledighed* which translates as 'Unemployment benefit recipients at risk of long-term unemployment' while the other possible outcome is *Øvrige dagpengemodtagere*, which translates as 'Other unemployment benefit recipients', meaning that the individual is not at risk of LTU.

Figure 2.2: The building blocks of statistical profiling models (OECD, 2018)

In addition to the results section a receipt is automatically generated, giving the unemployed an understanding of the categorization and ways to move forward. The receipt below is shown for people at risk of LTU (STAR, 2020):

> Thank you for completing the preparation form.
>
> Some unemployed people who have previously answered the same as you or who are similar to you in other ways (e.g., in terms of previous employment relationships) may have experienced a slightly longer path to a new job.
>
> To avoid this, it is important that you prepare thoroughly for your first interview with the job counselor. You can, for example, do this by bringing 3-5 specific positions that you would like to apply for within the next 14 days or that you have already applied for.
>
> By starting from some specific positions, you take an essential step towards a new job.

Second conversation

According to employment consultant Mikael Dehn Kristensen, who works at the job center Copenhagen, the second conversation occurs at the job center, where each discussion lasts 30 minutes with a 15-minute preparation time beforehand (Appendix B). The preparation time is spent looking through jobnet.dk to see what has been filled in and which parts are missing.

Based on statistical analysis, STAR was able to pinpoint the most significant input data. STAR found a clear correlation between the unemployed person's expectation of unemployment length and the unemployment period's actual length (see question 2 in Table 2.1). Unemployed people who think they will return to work quickly thus typically have a shorter period of unemployment than those who expect a long time to pass. When the unemployed expect it to take a long time before returning to work, this can indicate that this is a citizen with unique or special challenges. Regardless of the answer, it makes sense to address this question in the conversation.

Regarding the second conversation, Mikael made the following comment (Appendix B):

> It works well when the unemployed person has already reflected on the receipt text. However, it is our experience that it is rare for the citizen to have made up their thoughts about this before the [second] conversation. In general, our impression is that the citizens have not dealt in detail with the questionnaire when they answer and that they do not think of the questionnaire as a preparation tool for the interview. Therefore, it is important for us to ensure a good dialogue with them about this in the interview itself.

Another essential point made by Mikael is that some job seekers fill out the questionnaire with the information they think or hope Mikael wants to read. Another comment was given by consultant Jane Susan Ringberg, who works at job center Odense, stating that some people fill out the questionnaire without reading the actual questions. For example she mentioned that a person had indicated that he wanted to start a new education despite having recently finished one (Appendix B). Further problems and possible solutions will be elaborated on in another section.

Apart from the screening tool results, a caseworker has other tools at his or her display. I will specifically introduce two tools that have been made available by STAR and which were mentioned in the interviews.

Labour Market Balance

The Labour Market Balance, *arbejdsmarkedsbalancen* in Danish, is an online tool which shows the job opportunities for approximately 900 job titles in each of the eight regional labor market councils (usually referred to as RAR areas). For example, using the tool you can see if there are very good, good, or less good job opportunities for a social worker and health worker in RAR Zealand (which covers 17 municipalities). The Labour Market Balance is compiled based on various quantitative data describing the supply of and demand for labor in the RAR areas and can be accessed on www.arbejdsmarkedsbalancen.dk.

The Labour Market Balance helps to provide caseworkers with a reasonable basis for being able to target their effort towards the unemployed and show the unemployed the shortest possible path to a job. Assuming that a person has high or unreasonable expectations about their job opportunities, the Labour Market Balance can objectively assess the possibilities (STAR, n.d.-a).

Labour Market Barometer

While the Labour Market Balance gives insight into whether a specific job title has good or bad opportunities the Labour Market Barometer, *job barometer* in Danish, allows the caseworker to illustrate alternative employment paths and inform unemployed people about job opportunities locally and regionally (STAR, n.d.-b).

The tools' purposes are to structure the conversations with the unemployed at risk of LTU and ensure quality and a certain uniformity in the discussions. Based on overall assessments of the tools' results, a caseworker can propose an intensive interview process, combined with other offers such as a job search course or company internship.

According to Jørgen Brorsen, who works as a manager at job center Frederiksberg, if people are at risk of LTU, they will be called for several interviews (every 14 days, rather than once a month). They will also investigate the possibility that job seekers can change industry and possibly offer courses (CV courses, a coaching course, etc.). He especially mentioned the challenges of getting non-western women back into work because of language barriers. The information provided by Jørgen was received over a phone call and was therefore not transcribed.

While caseworkers at job center Odense frequently use the screening tool developed by STAR, others have chosen to use tools developed by unemployment insurance funds. A phone conversation with a job center manager in Aalborg revealed that they used a screening tool developed by the unemployment insurance fund HK. According to the manager, the job seekers in Aalborg are more homogeneous and are primary HK members, which attracts a specific group of workers. If they had a heterogeneous group of job seekers, they would have preferred the tool developed by STAR.

2.3 International comparison

In 1998 the Organisation for Economic Co-operation and Development (henceforth OECD) published a report entitled "Early identification of job seekers at risk of long-term unemployment." The report outlined the rise of unemployment across several OECD countries, stating that millions of people had become unemployed over the past years. Figure 2.3 shows the unemployment rate across OECD countries, where the '80s and '90s are especially noticeable. The cause can be traced back to a global recession caused by the international oil crises of 1973-74, 1978-80, and 1990-91. The crises were the result of oil-exporting countries, organized by the Organization of the Petroleum Exporting Countries (OPEC), choosing to reduce production and raise prices. The underlying causes were wars, crises, and unrest in the Middle East. The oil crises likely started based on regional or national problems, but they had economic consequences for several OECD countries in the globalized world. Other factors, such as a tightening of monetary policies and the weakness of lending institutions in the US, can also be mentioned (Farbøl et al., 2018).

The OECD report gives us insight into some of the earliest statistical screening tools developed in the wake of the global recessions. Australia was the first country to introduce a statistical profiling tool in the '90s called *The Job Seeker Classification Instrument (JSCI)*. The Netherlands presented *Work Profiler*, and the US rolled out the *Worker Profiling and Reemployment Services (WPRS)* tool. After the year 2000, several countries developed profiling models, which are either based on statistical models, caseworkers' professional assessments, or a combination of the two approaches. The national tools have primarily been developed based on inspiration from each other's and the predecessors in the field (OECD, 2018).

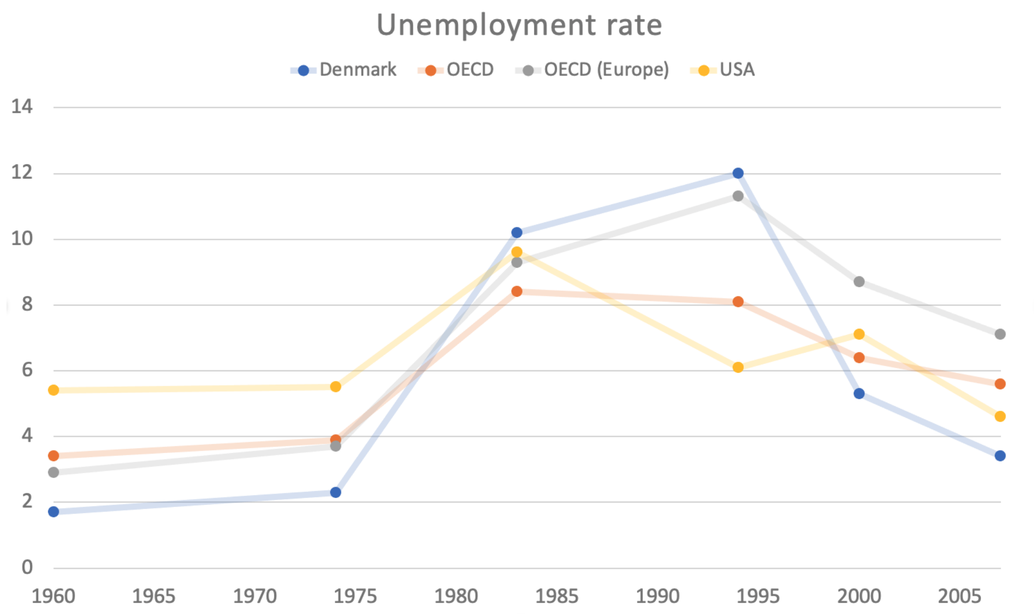


Figure 2.3: Numbers of unemployed people as a percentage of the labor force

Denmark was also struck by the global recessions and had the highest unemployment rate in the '90s compared to the OECD average (Figure 2.3). One method used to tackle unemployment was implementing a series of labor market reforms (1994, 1995, 1996, and 1998). To my surprise, I found out that Denmark was also inspired by countries abroad to develop a screening tool, which was made available in 2003/04 based on the characteristics and experiences of 1.2 million unemployed (VIVE, 2007). However, the reception of the screening tool at the caseworker level was not positive. An early evaluation in 2006, carried out by Rambøll for the Danish Labour Market Authority, showed that the employees did not find the tool credible and did not believe it provided new knowledge. That the use of the tool was obligatory in itself created resistance and was perceived as an inappropriate use of their resources. Based on the feedback, the screening tool was withdrawn in the spring of 2007 (VIVE, 2007).

Based on the same context as the countries mentioned above, other countries presented their screening tools after several years of research. France presented a screening tool in 2006, Finland in 2007, Germany introduced the *4-PM-model* in 2009, Sweden presented *Bedömningstöd* in 2012, the Netherlands introduced a tool in 2013, and Ireland rolled out the *Probability of Exit (PEX)* in 2013 (Eskelinen et al., 2015).

Variety of profiling tools

The different countries use various profiling tools, and I will mainly distinguish between three types of profiling: rule-based profiling, caseworker-based profiling, and statistical profiling.

- **Rule-based profiling** can be thought of as the traditional way for employment services to sort people into different groups. Based on age, education level, unemployment duration, etc., people will be handled differently. Looking at Flanders region, we see that people under the age of 25 will be reached within four months while job seekers aged 25 and above will be reached within twelve months (Desiere et al., 2019).

- **Caseworker-based profiling** relies solely on the judgment made by caseworkers. Although Germany introduced a statistical profiling model, it chose to abandon the model and rely on caseworkers' judgment instead. For instance, a caseworker will categorize job seekers as either 'easy' or 'hard-to-integrate' into the labor market after a one-hour interview. Other countries like Estonia, Greece, Luxembourg, Slovenia, and Switzerland also rely on this type of profiling. For example, Greece gives its caseworkers the freedom to categorize people as being at either high, medium, or low risk of LTU (Desiere et al., 2019).

- **Statistical profiling** means using a statistical model to predict LTU (Figure 2.2). The models are usually based on administrative and survey data, making it possible to individualize the profiling while being less time-consuming and, therefore, less expensive. This method is the main focus of the thesis and will be analyzed in greater detail in the Methods chapter.

The three types of profiling are not mutually exclusive, and in practice, we see that different countries have chosen to combine the approaches; this is illustrated below in Figure 2.4. Countries like Denmark, New Zealand, Sweden, France, and Finland use statistical profiling as a form of assistance while the final judgment lies in the hands of the caseworker. On the other hand, countries like Australia, Ireland, and the USA rely solely on statistical profiling. The caseworker has no or limited influence on the categorization of the job seeker. Meanwhile, countries like Austria combine all three profiling tools, where the initial categorization is rule-based depending on age, following which a statistical profiling is initiated, while the final judgment is in the hands of the caseworker, which can override results and fit people into categorizations they see as more appropriate.

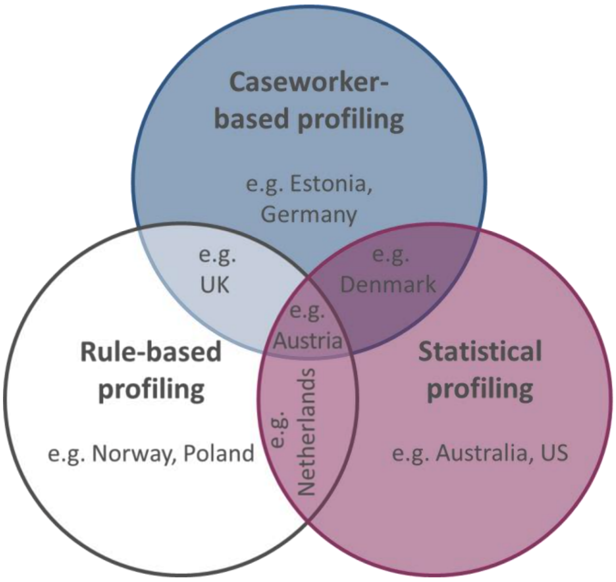


Figure 2.4: Types of profiling across the OECD (Desiere et al., 2019)

Model comparison

To fully assess the international comparison between the countries, it is necessary to dive into the statistical profiling models to understand where they differ or align. Table 2.2 provides an overview of 11 countries that use a screening tool in some way or another. The table gives a unique insight into the outcome variable, data collection method, type of data used, the statistical model, whether participation in profiling is compulsory for job seekers, whether caseworkers must use the results, and when the model is implemented (OECD, 2018).

The first column to the left gives an overview of the different definitions of LTU. The countries are divided into those which calculate the probability of LTU for 26 weeks (6.5 months) or for 52 weeks (12 months). The motivation for choosing one over the other might be connected to how fast the majority of job seekers get back into the workforce without major or any help from employment services. As mentioned earlier, almost half of the job seekers in Denmark find a job within six months of becoming unemployed (Appendix A, Figure A.2); in Sweden the number is 51% (Eskelinen et al., 2015).

The data collected to calculate the probability of LTU is derived from three sources: personal interview, administrative data, and questionnaire. The different sources give insight into demographic characteristics, family situation, health conditions, social conditions, education, work experience, unemployment history, motivation, residence, etc. The data is then passed on to a statistical model, which has traditionally been based on classic analytical techniques such as cross-tabulations or regression models. The majority of the countries below use Logistic Regression, while others use more sophisticated or advanced machine learning techniques. It might be worth noting that the original model used in Denmark, between 2003/04 and 2007, was based on Logistic Regression.

Terms like big data, machine learning, artificial intelligence, and deep learning have become hyped today, but the story can be traced several decades back. The recent breakthrough in artificial intelligence is due to the exponential improvement in computing power, the growing availability of data, the development in algorithms, and new data storage techniques. The advancement in the field has allowed countries to experiment and implement machine learning models such as Decision Tree, Random Forest, and Gradient Boosting (the different models and their pros and cons will be presented in the next chapter). As an example, Belgium has implemented a Random Forest model, and according to Desiere et al. (2019), the model is called *Next Steps* and is based on hundreds of variables. Unlike other countries that rarely update their model, the implementation of Next Steps is structured in a flexible way, and is regularly updated.

In terms of accuracy, Austria stands out, delivering an accuracy rate of 80-85%. The Austrian Public Employment Service (hereafter PES) introduced its first statistical profiling model in November 2018 and implemented the screening tool the following year. What makes the results unique is the use of only administrative data. In other words, they managed to outperform more advanced models by simply using existing data sources.

	Outcome (Probability of)	Data derived from	Statistical model	Accuracy	Compulsory/voluntary use by Job seeker	Caseworker	Implementation
Australia	LTU (12 months)	Personal interview; Online trial ongoing	Logistic Regression	-	Compulsory	Compulsory	Upon registration
Austria	Labour market integration probability	Administrative data	Logistic Regression	80%-85%	Compulsory	Compulsory	-
Belgium	LTU (>6 months)	Administrative data	Random forest	67% (AUC ~ 0.76)	Compulsory	Compulsory	-
Denmark	LTU (>26 weeks)	Online questionnaire; Administrative data	Decision Tree	>60%	Voluntary	Voluntary	Within 3 months
Ireland	Probability of exit to employment within 12 months	Questionnaire as part of benefit claim process; Administrative data	Probit regression	70%-86%	Compulsory	Compulsory	Upon registration
Italy	LTU (12 months)	Administrative data	Logistic Regression	-	Compulsory	Compulsory	-
Latvia	LTU (12 months)	Personal (individual) interview; Questionnaire; Administrative data	Factor analysis	-	Compulsory at PES & voluntary online	Compulsory	-
Netherlands	LTU (12 months)	Online questionnaire	Logistic Regression	70%	Voluntary	Compulsory	6-8 weeks after registration
New Zealand	Lifetime income support costs (LET), change in lifetime income support and staff costs from receiving a case management service (SEM)	Administrative data	Random forest (LET), Gradient boosting (SEM)	AUC: 0.63-0.83	Compulsory	Compulsory	-
Sweden	LTU (6 months)	Personal interview	Logistic Regression	-	Not aware	Voluntary	Upon meeting the caseworker
US	Exhausting the 26-week entitlement to UI benefits	Online questionnaire; Administrative data	Logistic Regression (/Depends upon state)	-	Compulsory	Compulsory	Upon registration

Table 2.2: International model comparison

The data input for Next Steps is described in Desiere et al. (2019) as the following:

> The model makes use of socio-economic variables (gender, age, nationality), information on job readiness (education, health limitations, care responsibilities), and opportunities (regional labour market development). A clear strength is the use of all available labour market history information, including detailed information on prior work experience (type and intensity), frequency and duration of unemployment, and participation in active labour market programmes.

An important lesson from this is that the type of model can have an impact on the performance, but the main focus should concern the type and quality of the data. On the other hand, it is unfortunate that my data regarding accuracy is limited to a handful of countries and that I was unable to gain insight into other evaluation metrics such as precision, recall or F1 score.

Regarding whether participation in profiling is compulsory for job seekers, and whether caseworkers must use the results, we see that Denmark clearly stands out by not requiring job seekers or caseworkers to use the screening tools. The job centers recommend using them and hope that the tools are seen as a form of support by the caseworker, but their use is not mandatory.

On the other hand, we see that job seekers in Sweden are not aware that they get profiled. The screening in Sweden occurs at registration in a conversation between the caseworker and the unemployed person. The caseworker reviews 12 questions that make up the profiling tool in Sweden. The screening tool (Bedömningstödet) indicates the result of the assessment immediately after the 12 answers have been entered by showing which of the four categories the unemployed person belongs to:

- Group 1: should have excellent opportunities to get into employment
- Group 2: should have good chances to get into employment
- Group 3: support should be considered to increase employment opportunities
- Group 4: needs support to increase employment opportunities.

Even if the assessment is conducted in a conversation situation, only the caseworker can see the result, i.e., in which category the system places the unemployed person. In addition to the total score and the category, the caseworker has access to the unemployed person's entire profile and can thus see how the total score was reached. The caseworker can use this knowledge to re-categorize the unemployed person.

Only the caseworker's final classification appears in the unemployed person's case information, and no information is given on whether it deviates from the categorization that the Bedömningstödet resulted in. Nor does it appear from the case documents whether the caseworker has implemented Bedömningstödet. In principle, this is a requirement, but it has no consequences if the caseworker does not use the screening tool, just as there is no control over whether early intervention is based on the assessment support or the caseworker's assessment. It is administratively possible to see if a particular unemployed person received early intervention. Nonetheless, it is impossible to see if this was based on profiling or the caseworker's categorization. The unemployed person's screening and screening results are not stored in the system; storage of this information requires the agreement of the unemployed individual (Eskelinen et al., 2015).

The unemployed person is neither informed that a screening is carried out nor told about the categorization. According to the developers of the model, this was a conscious choice because this information will not benefit the unemployed. The reason is that knowledge of one's risk category could lead to a self-fulfilling prophecy and could have a negative effect on the unemployed person in the event that the person in question belongs to the category with the highest risk of LTU. In this case, the risk category will be stigmatizing and will be perceived as discouraging. It is accordingly more relevant and constructive to use the results to talk to the unemployed person about what needs to be done in order for the person in question to get back into the workplace (Eskelinen et al., 2015).

The argument that the results of the screening tool might have a negative effect was also brought up in my interview with Mikael (Appendix B):

> **Mikael:** You just filled it out. You have not even had a meeting with the unemployment insurance fund (A-Kasse). You have just signed it. You read that you are welcome to fill out the questionnaire (Forberedelsesskema), as we call it. Strange term by the way, and then you get the result. You just graduated, you are so happy, you have faith in everything. Okay, there is corona. Let us pretend that we are not in the time of corona. You are so glad, and you are thinking, 'now I want to make money, damn I am good because I got 12.' You send applications, and you do not receive any reply. What is happening? Then you go through the unemployment benefit (dagpenge). Then you think, 'yes, a couple of months with unemployment benefits, and then it is on, yes.' However, the first thing you are told when you have filled out the questionnaire is that you are at risk of becoming long-term unemployed.
>
> **Tarek:** The motivation is completely down.
>
> **Mikael:** Yes. Therefore I am not an advocate for this. [...]

Neither Australia, Ireland, Sweden, or Germany openly uses the screening tools. This is reflected in the fact that the unemployed are not aware of the profiling tool's use and are not aware of the categorization. According to Eskelinen et al. (2015), the reason for not disclosing the profiling and categorization is that it will negatively affect the unemployed in cases where the screening places the unemployed in the categories furthest from the labor market. They state that it will be demotivating and stigmatizing in terms of efforts to reintegrate into the labor market and thus will become a self-fulfilling prophecy. The exact opposite situation is seen in Denmark, where transparency about the tool, the features used, and the categorization is displayed and told to the unemployed.

My interview with Jane also revealed a desire to open up the possibility of re-categorization throughout the unemployment process, in order to give caseworkers more control and influence over the end result, but also to use it as a motivational factor:

> **Tarek:** What is your general impression of the tool? What is your general opinion?
>
> **Jane:** I think it is good, but I wish that it was open for editing. People change their approach in the process. They can think 'These were my ideas before, now I am thinking it might take nine months, however I am on my way, because we have already started the process.' They might have started a basis course, it can be something else which has helped people, e.g. supplementary training which has an impact on how you perceive your job situation.
>
> **Jane:** [...] it would make sense for the tool to take account of the citizens' own perspective, because something happens with them when they have been unemployed for three months. Everyday hits them, they become discouraged, and they lose faith in their own abilities and professional qualifications. So there is a lot of work to do in terms of motivation. If you can keep track of their progress, then it is easier to keep them positive. 'Between the last meeting and now, this and this happened.' People will be able to visualize it. When they see their progress, they are amazed.

Feature comparison

In continuation of the international model comparison, I chose to dig into a handful of screening tools to reveal and understand the underlying features. Table 2.3 shows five countries and their individual choice of features (Eskelinen et al., 2015). Australia's tool is based on 19 features, Ireland's on 24, Sweden's on 13, Germany's on 15, and Denmark's on 6. I will also refer to the 12 features used by Austria, which were outlined in the previous section.

By diving into the different features, I noticed that countries could be distinguished based on whether they include information of a more subjective nature or not. In Australia and Ireland, information regarding self-assessed health is included, while the Swedish and Austrian tool is based solely on information that is also found in registers.

In Germany, the screening tool consists of a series of open-ended questions; the subjects are highlighted in the table. The German screening tool differs from the others in that the variables have a qualitative as well as a quantitative dimension. This is expressed by the fact that it is not only the registration of, for example, the unemployed person's housing problem, but to a greater extent, the caseworker's qualitative assessment of the scope of the housing problem that is important in the profile clarification.

An argument against including subjective variables in the profiling tool is that they depend on how the unemployed person understands the question and whether he/she is willing to give a sincere answer, given the implications a particular response may have. For example, the question might be about sensitive issues such as abuse, family relationships, or health, which can constitute significant barriers to getting a job.

Australia	Ireland	Sweden	Germany	Denmark
• Age and gender • Work experience • Education level • Transfer income history • Professional qualifications • English skills • Country of birth • Place of residence • Local employment opportunities • Distance to the labor market (city vs. country) • Population group (aboriginal) • Access to transport • Contactability by telephone • Disability/illness • Housing situation/permanent residence • Marital status/provider • Criminal record • Personal factors or characteristics	• Age and gender • Health (self-assessed) • Civil status • Whether a person has children • Spouse's income • Education level • Apprenticeship • Reading/arithmetic problems • English skills (self-assessed) • Unemployment history (time since last employment) • Temporary employment • Job mobility (willingness to move to get a new job) • Employment history (duration of employment) • Transfer income within the last 5 years • Transfer income for more than 12 months • Participation in local employment program within the last 5 years • Participation in local employment program within the last 12 months • Type of benefit (insured unemployed or non-insured unemployed) • The number of previous transfer income streams • Residence (urban or rural) • Transportation options • Proximity to public transport • Region	• Age and gender • Disability • Country of birth • Education level • Unemployment insurance (registered unemployment fund) • Length of previous unemployment • Latest unemployment • Registration month • Unemployment rate in the municipality • Sought work area • Experience in the sought work area • Education within sought work area	**Qualifications**: • School qualifications • Technical qualifications • Occupational qualifications • Language skills **Capacities** • Intellectual potential • Health potential • Social attitude and work behavior **Motivation** • Initiative/work attitude • Willingness to learn and study **General conditions** • Age • Personal situation • Geographical mobility • Housing situation • Family situation • Financial situation	• Age • Origin • Employment rate for the past 36 months • Aggregate wage for the past 12 months • Education (Q1a) • Expectations (Q2)

Table 2.3: International feature comparison

In Australia, a unique approach is used, which means that the unemployed person can avoid answering questions that may be perceived as sensitive to the person (e.g., health or criminal record). These aspects are elaborated at a later stage, with the possibility of supplementing information included in the unemployed person's profile with the possibility of re-categorization. Subjective variables are included in profiling in Denmark, Australia, Ireland, and Germany, and Sweden is in the process of expanding its Bedömningsstödet with subjective variables (Eskelinen et al., 2015).

The statistical profiling models behind the Australian JSCI, Austrian Next Steps, the Irish PEX, and the Swedish Bedömningsstöd focus on the unemployed person's barriers and weaknesses. This is because the models are designed to predict the risk of LTU, therefore identifying the most critical risk factors. Only in the German 4-PM-model are the unemployed person's strengths the first aspect that the caseworker has to deal with.

Taking a closer look at the features, I noticed that five of the countries, not including Denmark, focus on health data. Through administrative data, interviews, or self-assessments, the countries get insight into disabilities, illnesses, and other health-related disadvantages. The questionnaire in Denmark accounts for possible physical or mental health issues in question 10a (Table 2.1), but this aspect is not included in the final model. Another example is language skills, which Australia, Ireland, and Germany assess. According to Eskelinen et al. (2015), Sweden planned to add language and motivation in the future, which might have happened in the meantime. One reason why Denmark hasn't added language might be connected to the use of origin instead, which is correlated with language.

Subconclusion The primary intention behind introducing profile clarification tools in all countries has been to prevent long-term unemployment. While the motivation has been the same, we see that the development and implementation differ. Some of the areas in which Denmark clearly stands out in my analysis were whether participation in profiling is compulsory for job seekers and whether caseworkers must use the results. I showed that Denmark clearly stands out by not requiring job seekers or caseworkers to use the screening tools. Another example concerns transparency, where Denmark again chose to walk a separate path by displaying the categorization and the features used to the unemployed. Finally, we see that Denmark lacks some components, such as health and language skills, which are seen as essential features by other countries.

2.4 Social impact and related research

As stated in the introduction, several complaints and a lawsuit were filed against Danish municipalities in 2020 for discriminating against their citizens based on origin. The debate quickly became heated and caught the attention of the Department of Human Rights, which brought STAR before the Equal Treatment Board (*Retsinformation*, 2020). This section will examine the social impact of job seeker profiling tools by diving into the subjects of explainability, interpretability, legislation, fairness, bias, and discrimination in the field of artificial intelligence and machine learning.

I define artificial intelligence (henceforth AI) as a common term for the methods and technologies that enable computers to analyze, reason, and learn new knowledge. It is a support technology that can assist and expand the knowledge and experience we already have before making decisions. One of the founders of the discipline of AI, John McCarthy, described it in the following words: "Artificial intelligence is the science and engineering of making intelligent machines, especially intelligent computer programs." Having defined AI, we also need to mention machine learning (henceforth ML), which is a subfield of AI. According to Professor Andriy Burkov, ML is defined as the process of solving a practical problem by gathering a data set and algorithmically building a statistical model based on the data set [with the goal of prediction] (Burkov, 2019).

The recent breakthrough in AI is seen everywhere around us. It is AI that determines what we are presented with on the front page of Netflix. It can determine what we spend our Saturday night watching. Or what we most likely will listen to by selecting the playlist that Spotify suggests. It is the same with search results on Google, which determine what information we get, or the route on Google Maps, which outlines the direction. Based on a data-driven understanding of who we are as human beings and what others like us chose in the past, AI continuously tries to predict our next move.

Besides private companies like Netflix, Amazon, Google, and others, we can also see that the public sector is becoming aware of AI's potential. In addition to job centers' use of AI to calculate the risk of LTU, we also see that Sygehus Lillebælt, a hospital, uses AI to diagnose acute patients based on blood and urine samples and rank their disease severity. Another example is seen at Bispebjerg and Frederiksberg Hospital, which uses artificial intelligence to analyze X-rays of knees for osteoarthritis. This helps the almost 350,000 Danes who suffer from arthritis in the knees to get faster and better treatment in the health care system. In another context, Copenhagen Airport has used AI to optimize baggage delivery since 2016. AI is used to predict where the staff is most likely to empty planes and put suitcases on luggage belts, where the focus has been on optimizing time consumption. This has meant less crowding and a better experience for passengers (Digitaliseringsstyrelsen, 2020b).

The government allocated 200 million DKK in 2020, which will be used over the next few years to advance municipalities and regions' use of AI (Digitaliseringsstyrelsen, 2020a). Based on the rapid advancement of AI, it is vital to understand the societal issues related to AI use and to ML in particular.

Explainability and Interpretability

When using ML algorithms, one is clearly aware of the input (available data) and output (given outcome). Still, there is rarely an explanation for everything that has taken place in between. Therefore, it can be challenging to understand how and why an ML model has arrived at a particular output, and this complexity can make the process opaque. This is called the black-box problem, which has been exacerbated by the intensity of the current interest and development in deep learning (henceforth DL), a subfield of ML (Vellido, 2019).

As an example, I could mention Odense Municipality, which has used AI to find patterns that most effectively get unemployed people back into employment. More specifically, unsupervised machine learning was used, which I will introduce in the next chapter. The method showed that a change in caseworker during the unemployment compensation process positively affected the citizen. Why this was the case, however, was unclear. The municipality's practice was that citizens should not change caseworkers along the way. The caseworker and the citizen should instead build a healthy relationship by allowing the caseworker to follow the progress from start to finish. Therefore, the result did not harmonize with the municipality's practice, which had been built up over a more extended period based on previous experience. As there was no understandable explanation for the recommendation of changing caseworker, the municipality stopped using unsupervised machine learning instead of supervised machine learning. The example illustrates the need to understand the logic behind specific outputs or categorizations. Assuming that a citizen asks why the caseworker is being replaced in the middle of a process, it should be expected that a caseworker at least to some degree knows why it might be recommended to switch in the middle of a process (Petersen & Holm, 2020).

It is worth mentioning that some ML algorithms are considered white-box models, because you can easily understand the mechanism by which a model works. When working with white-box models, one can easily visualize which features had a say and which were disregarded in the analysis. You can easily understand and explain the logic behind certain outputs. As seen in Table 2.2, Denmark has chosen to implement a so-called decision tree, often referred to as a white-box model. STAR specifically mentions interpretability and explainability as a motivation for not picking more complex models. On the other hand, my interview with job counselor Marianne Krogh Fischer revealed that the caseworker is not aware of the logic or considerations behind the ML model's different categorizations. They have not received any description or presentation of the model, which leaves them with a feeling of blind trust in the ML model (Appendix B).

Why do users not always use a white-box model? By focusing on transparency, one would, at one point or another, face a trade-off between the performance of a model and its transparency (Barredo Arrieta et al., 2020).

Legislation

Technological developments have given people many great new opportunities: You can access qualified information from all over the world. One can share information with anyone at the moment when the incident occurs. You can track your health in real-time on several different parameters. You can get support for many practical functions in your home. One can easily access help to make a variety of decisions. Life has generally become more comfortable because there are a number of tedious routines that current generations have been exempted from performing thanks to technology.

However, if the technology is not designed with privacy in mind, there is a flip side to the coin that needs to be weighed against all the major new opportunities. While the discussion regarding AI and ML's potential in different fields has taken all the attention, another discussion regarding its legal boundaries and implications should not be ignored. According to the General Data Protection Regulation (henceforth GDPR), which was introduced in May 2018, there is an individual right to be given an explanation when decisions have been made by automated or AI algorithmic systems —this is referred to as *the right to explanation.* Article 13 of the directive states that (GDPR, 2018a):

> In addition to the information referred to in paragraph 1, the controller shall, at the time when personal data are obtained, provide the data subject with the following further information necessary to ensure fair and transparent processing: the existence of automated decision-making, including profiling, referred to in Article 22(1) and (4) and, at least in those cases, **meaningful information about the logic involved**, as well as the significance and the envisaged consequences of such processing for the data subject.

To rephrase the above paragraph, the Article states that any person using AI or ML systems must be able to interpret and explain how decisions were made to the humans affected by them. Looking at Recital 71, which provides additional information on Article 22, we find that people have the right to reject profiling (GDPR, 2018b):

> The data subject should have the **right not to be subject to a decision**, which may include a measure, evaluating personal aspects relating to him or her which is based solely on automated processing and which produces legal effects concerning him or her or similarly significantly affects him or her, such as automatic refusal of an online credit application or e-recruiting practices without any human intervention.
>
> Such processing includes 'profiling' that consists of any form of automated processing of personal data evaluating the personal aspects relating to a natural person, in particular to analyse or predict aspects concerning the data subject's performance at work, economic situation, health, personal preferences or interests, reliability or behaviour, location or movements, where it produces legal effects concerning him or her or similarly significantly affects him or her.

The Article relates to cases where a fully automated processing of personal data occurs, and thus not cases where only partially automated processing occurs. An example of a fully automatic decision is the decision to award SU (student benefit), which is decided digitally and without a caseworker's involvement. A partially automated processing will involve the presence of an element of human intervention. For example, suppose a caseworker uses AI as an aid tool to make a decision. In that case, the decision will, in principle, only be partially automated and thus fall outside the scope of Article 22 (Petersen & Holm, 2020). Can we regard the screening tool as solely partially automated? Unlike Sweden or other countries, caseworkers in Denmark cannot change the categorization or final output. Even though the screening tool is only seen as an aid, we cannot ignore that some caseworkers, especially inexperienced ones, might blindly trust the tool's output and act on it. Therefore it is essential to better equip job counselors to understand and explain the logic behind specific categorizations.

Related research The problem with black-box models and the requirements of the GDPR have led to attempts to develop new methods and techniques which attempt to break the trade-off between performance and transparency. The new methods and techniques are developed within the field of eXplainable AI (henceforth XAI), which may potentially solve the municipalities possible explanatory problems. The number of publications contributing to the field of XAI rose from approximately 50 in 2012 to over 350 in 2019 (Barredo Arrieta et al., 2020).

The techniques and methods developed within the field of XAI can be divided into two main categories: Transparent Models (white-box models) and Post-Hoc Explainability (black-box models), which is illustrated in Appendix A Figure A.3. Our primary focus is on the latter, labeled Post-Hoc, because we try to understand the model after it has been developed. We can further divide Post-Hoc Explainability into two further categorizations: Model-Agnostic and Model-Specific. Whereas the first type concerns techniques that are designed to explain ML models of any kind, Model-Specific techniques, as the name suggests, are only tailored or specifically designed to explain specific ML models (Barredo Arrieta et al., 2020).

An example of a Model-Agnostic technique is LIME, which stands for Local Interpretable Model-agnostic Explanations. The technique was presented back in 2016 and treats all models as black-box models. It can be applied to any classifier or regressor and will give you a visual representation of why your model behaves as it does (Ribeiro et al., 2016). I use LIME in my analysis and present some of these visual representations in the results chapter. Another example is SHAP, which stands for SHapley Additive exPlanations, and just like LIME, gives a visual representation of the model behavior, in addition to ranking the features based on their importance.

Techniques like LIME and SHAP are not meant to clarify a model's algorithm, but they give insight into how features impact the output. If a job seeker were to question the underlying logic of the assessment, a caseworker would be able to say that a specific age, gender, ethnicity had a negative/positive impact, etc., regardless of the model's algorithm.

Fairness and discrimination

AI is designed to imitate human traits without being affected by biases such as mood swings and biased assumptions; its calculations only relate to data. In other words, ML models never have a bad day and do not get tired from working overtime. As AI is not affected by such factors, it is reasonable to imagine that ML models cannot come up with discriminatory or unreasonable outputs. That is, the use of AI in itself should lead to compliance with the principle of fairness.

To illustrate how an ML model might discriminate, let us go back to 2013, when 20-year-old Dylan Fugett and 21-year-old Bernard Parker were arrested in Florida in the US. They had both previously committed crimes of the same nature and were arrested this time because of drug possession. Based on their criminal background, age, state, and other features, they were much alike. Using a risk assessment algorithm, an ML model calculated the probability of recidivism, or in other words, the likelihood of committing a future crime by giving a score between 1 and 10 (where ten is highly likely to re-offend). Despite the similarities, Bernard was scored as 10, and Dylan as 3. Some time later, Dylan was arrested for drug possession, while Bernard did not have any subsequent offenses. The only apparent difference was that Bernard is black and Dylan is white (Angwin et al., 2016).

The story of Bernard and Dylan is not unique, yet this has not limited the use of algorithmic risk assessment instruments; on the contrary, several states, including Oklahoma, Colorado, Washington, Louisiana, Kentucky, Arizona, Virginia, Delaware, and Wisconsin, use the algorithm to categorize criminals and hand over the results to judges during criminal sentencing. The bias against Afro-Americans is evident. While the introduction of new tools has tried to combat inherent biases, we still see problems through the correlation with other features. Removing race but keeping data on education and whether parents were sent to prison has been criticized as a stand-in for race (Jobes, 2018).

A similar example is the Danish labor market, which is very gender-segregated. Most men work in the private sector, while most women work in the public sector. Assuming that an ML-model is supposed to recommend specific job positions, it might end up reinforcing the division as it predicts a citizen's path into the labour market based on historical data. Another example is profiling, where statistical discrimination against job seekers puts migrants, people with disabilities, and the elderly at a disadvantage (Desiere & Struyven, 2020). Several countries include features that reveal the background of the unemployed. The features are labelled as the country of birth, nationality, or origin. While the aim of including these features is to improve performance, it frequently comes at the cost of discrimination. According to Desiere and Struyven (2020), discrimination in this context can be defined as:

> The proportion of job seekers who belong to a particular group and find a job ex-post, but are misclassified as high-risk job seekers, relative to this proportion among the dominant group.

Ideally, one would like to develop ML models which are accurate and not discriminatory, but the tension between accuracy and discrimination has forced some to choose between them. The trade-off is made even harder, given that discriminatory features depend on the cultural context. In 2018 a group of students decided to complain to their biology teacher at the University of Copenhagen because the teacher showed statistics which divided people based on gender. The teacher, who had 30 years of experience, had never received a similar complaint (Dandanell, 2019). The subject of discrimination was also raised in my interviews, where job counselor Marianne made the following comment regarding age:

> I am thinking about the issue of age. I have reached the age where I am interested in this. I am quite aware of the discrimination based on age that I think is part of the fact that we call people 'seniors' when they are 50. They still have 15 years in the labor market. I do not think we should continue with this. When we use this term, we place people in a box [...]. We need new terms, for instance, 'experienced' [instead of senior].

Others share her experience regarding discrimination based on age, and politicians are raising her concern (Olsen, 2021a). My interview with consultant Jane also revealed some strong opinions regarding origin:

> People already feel marginalized, they have pursued an education, and they have difficulties getting into the labor market. The reason could be because they have a different name. I know there are some fields where it is difficult to get a foothold. I think it is a bit sad. I have dealt with young Danish people, as I call them. They say, 'I speak Turkish, it is my mother tongue,' and Danish is ticked as 'experienced'. I ask them why they wrote this; they respond: 'Well, I am Turkish.' I tell them, 'Your origin is Turkish, you speak fluent Turkish, I also speak fluent English, but I am Danish. So think about it in terms of the parameters. You are Danish with Turkish origin.' If we keep throwing people into these categories, it will be tough to make this connection, as we have discussed in terms of work. For people who have a different ethnic origin, it is a bit sad. I also feel they become marginalized, 'because of my origin, and because of my name.' It is a pity. It makes me sad. [...]
>
> We try to get people into employment, and that includes people of a different ethnic origin. That is fine, but why should we keep reminding them that they are different? Why should people be reminded if they are from Jutland [a peninsula in Denmark]?

As noted earlier, we see that Denmark has chosen a separate path compared to other countries, by being transparent in regards to the categorization and the features used. This has raised concern among Danes who fear discrimination on the basis of origin, inspiring people to express their dissatisfaction on social media. I stumbled upon several examples but have only included the most viral instances. The tweet below by Telli Betül Karacan might be the earliest example where dissatisfaction was expressed online (the tweet is translated from Danish):

Figure 2.5: Example of dissatisfaction over the categorization (Bostrup, 2021)

Another example, found on Facebook, is shared below, where part of the post is translated (the full post can be found in Appendix A Figure A.4):

> Dear you, who was born in Denmark to parents from so-called "non-western" countries and are currently looking for a job. Jobnet, under the Danish Agency for Labour Market and Recruitment (STAR), has profiled YOU as a "descendant of non-Western descent." Thus, the effort they put

into you is different from the effort they put into Pia and Morten [typical Danish names]. In good old-fashioned terminology, this is called discrimination based on ethnicity [...].

The above Facebook post received several hundred reactions, 164 comments, and 140 shares. The evident dissatisfaction made people aware of the screening tool's features and might be the examples which caught the Department of Human Rights' attention. The Head of Equal Treatment, Maria Ventegodt, at the Department of Human Rights, had the following comment: "Emphasizing ethnic origin in the case processing of the individual citizen is discrimination unless it is a question of positive discrimination, and we do not consider this to be the case here" (*Retsinformation*, 2020).

I tried unsuccessfully to get a comment from Maria regarding her definition of positive discrimination, but unfortunately, without any luck. We are operating with a scenario where discrimination might be positive, depending on the services provided, or negative if the support is perceived as burdensome. A phone conversation with the job center manager in Frederiksberg municipality gave me the clear impression that more conversations with a caseworker were regarded as positive discrimination. However, if the decision is based on origin and not competencies, it might be considered as a burden by others.

To examine the case further, I decided to pull data from the job centers on all the unemployment processes between 2016 and 2019 to analyze the number of conversations per unemployment process. Based on a multiple linear regression analysis, I was able to get a baseline on the country's overall number of conversations and those in a specific selection of municipalities. The regression allowed me to compare the number of conversations divided by origin, gender, and education level, displayed in Figure 2.6 below.

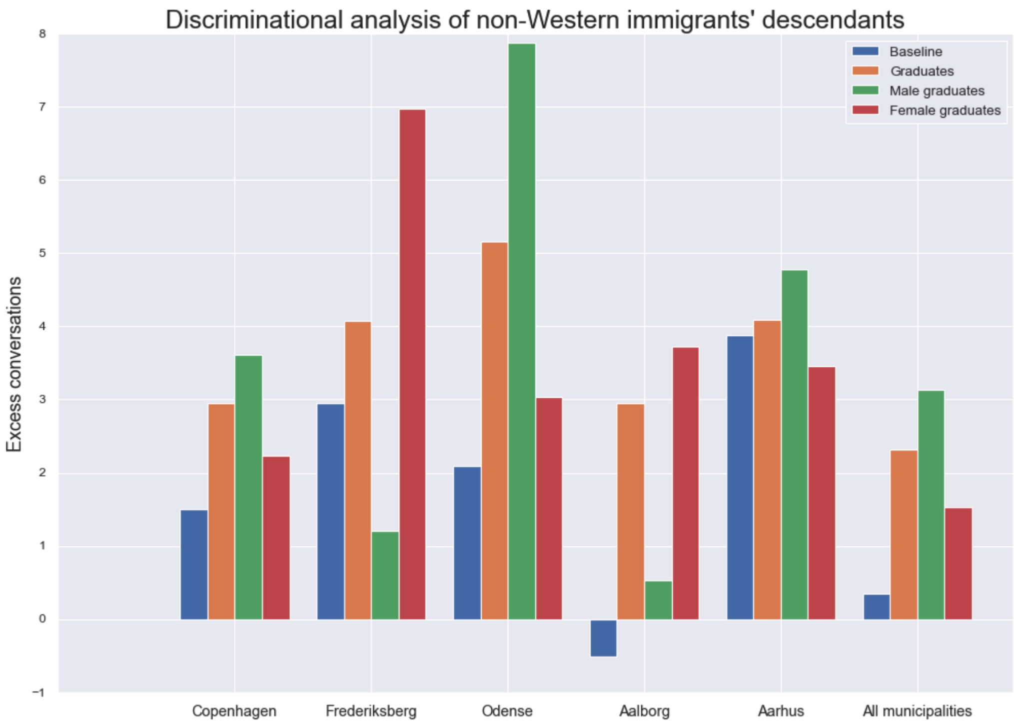


Figure 2.6: Excess number of conversations based on origin, gender and education

The length of the unemployment process can differ from person to person, depending on the type of benefit, social resources, language skills, geographic location, work experience, type of education, age, gender, grades, etc. To compare different groups, we would need to eliminate or minimize apparent differences. The above analysis starts with a baseline comparison where citizens with Danish origins are compared to non-Western immigrants' descendants. By using descendants as a comparison, I made sure that language wasn't a barrier, and neither was cultural understanding.

The blue bar in the figure shows the excess number of conversations descendants have had with caseworkers compared to citizens with Danish origins. The comparison is over-simplistic given the possible differences on the previously mentioned parameters. Nevertheless, we see that descendants received the same number of conversations as people of Danish origin in Aalborg municipality and in Denmark as a whole (presented as the *All municipalities* category in the figure). I say it is the same, given that I did not find a significant difference between the groups. On the other hand, we see a significant difference between the groups in Denmark's capital, Frederiksberg, Odense, and Aarhus. Both groups receive the same type of benefit, have the same language skills and geographical location.

Moving on to the orange bar, we now have a more comparable scenario. Both groups are graduates, have the same language skills, receive the same benefits, and have the same geographical location. We can see from the figure that descendants receive a considerably larger number of conversations with caseworkers across the country. Descendants who went through an unemployment process in Odense municipality had around five more conversations with a caseworker than citizens with a Danish background. By including gender, illustrated with the green and red bar, we make the groups even more comparable and see that male descendants were generally called into more conversations than female descendants. There are exceptions, as seen in Frederiksberg and Aalborg municipality.

It would have been desirable to dive even further into the comparison by looking at specific unemployment process length, age groups, and educational background. Although the data is available and forms part of my analysis in the next chapter, the number of samples was limited and would have given me unreliable results if I had divided it further. The data should be read with caution, given the possibility of omitted variables bias. A discussion regarding the low adjusted R-square can be found in the Discussion chapter. A proper introduction to the data will be given in the next chapter and a closer examination of my results can be found in Appendix C.

Even though my analysis suggests that more conversations are being held with non-Western immigrants' descendants than citizens with Danish origins, the big question is still whether this extra assistance is negative or positive. If a caseworker offers citizens with a specific origin more conversations, courses, or coaching sessions and helps the unemployed return to labor, should we categorize this as positive discrimination? In this case, we might have Danish-origin citizens who see the support offered to others as too generous and feel excluded and discriminated against by not getting the extra help needed. On the other hand, descendants might see the extra assistance as burdensome and unnecessary and unsuited to their needs.

Assuming that we want to adapt to the current environment continuously, it would mean excluding origin, gender, age, zip codes, etc. But excluding sensitive variables from our screening tool might still not reduce discrimination because of the correlation with other variables. These phenomena are referred to as discrimination-by-proxy (Vellido, 2019) or redlining (Kamiran & Calders, 2009). It would mean that even though a country like Australia removed information regarding the individual's origin, they might still get the same results since they have information on language skills, education, and postal code which is highly correlated with origin. However, removing all variables which correlate with each other will likely damage the accuracy of the algorithm, which leads us back to the trade-off between accuracy and discrimination. There is also a legal aspect, where the inclusion of certain sensitive information is prohibited. For example, Sweden has banned the use of gender in the context of profiling and the US has made it illegal to use variables like age, gender, and race (Desiere et al., 2019). But even these countries have not banned the use of zip codes, credit scores, language skills, education, and many other variables which serve as proxies for the omitted characteristics.

Related research The scientific research in this field can broadly be categorized into two groups: One group of researchers seeks to achieve fairness by modifying the underlying data. The other group aims to achieve fairness by modifying the classifiers. Examples of both groups are presented below.

One way of dealing with the situation is presented in a study by Pope and Sydnor (2009), which introduces a framework for implementing anti-discrimination policies in statistical profiling models. They offer a way to maintain efficiency in a model relative to eliminating the features completely. Instead of either allowing or banning sensitive attributes (such as origin, race, gender, age, etc.), they propose using the variable when training the model but adjusting the variables when performing actual predictions. They present their conceptual framework using Ordinary Least Squares (OLS), but the logic and method are also applicable to other models. The OLS model consists of two types of variables, one group being "socially acceptable predictors" (SAPs) and the other being "socially unacceptable predictors" (SUPs). We now have the following scenario:

$$y_i = \alpha + \beta * X_i^{SAP} + \theta * X_i^{SUP} + \epsilon_i \tag{2.1}$$

Here we have an output labeled y_i, which could be the number of weeks on unemployment benefits for person i. We have α, which is the baseline, and right next to it, we have the group of SAPs X_i^{SAP}, and their respective coefficients, β, and the group of SUPs X_i^{SUP}, and their individual coefficients, θ. Under normal conditions, we would remove all SUPs, which might lead to omitted variable bias (OVB). If a SUP correlates with the dependent variable, y_i, and at least one of the SAPs, removing it from the equation will impact the β coefficients, making them unreliable (Pope & Sydnor, 2009).

According to Pope and Sydnor (2009), we can avoid OVB by estimating the coefficient when SUPs are included and by using average values when individual predictions take place. As an example, we might have *age* as a SUP, which we use when building the model, but when performing the actual prediction, we use the average age of the population as a value. If the SUP is a categorical variable, we would then use dummy variables, but instead of multiplying with binary numbers, we would use the population proportion of minorities. The predictive power will decrease, but not to the same degree as if we had removed the variables before training the model. An empirical example from the study shows that the percentage of black job seekers in the high-risk group decreases from 22% to 16% when using this method. Given that most countries use logistic regression as a profiling tool, it would be easy to implement the framework in a real-world context.

In addition to the example above, it is worth mentioning a study by Kamishima et al. (2012), which proposes a so-called *prejudice remover regularizer*. By adding a regularizer to a logistic regression model, they manage to build a model which is less influenced by sensitive information. In other words, they show how to penalize prejudicial outcomes without losing substantial accuracy.

An example of modifying the underlying data is outlined in the study by Kamiran and Calders (2009), which introduces a new classification scheme for learning unbiased models on biased training data. They present a case where they start by "massaging" the data; in other words, they try to remove discrimination with the fewest possible changes to the training data. In the article, they use a Naive Bayesian classifier to calculate the class probability of all the samples; afterwards, they modify the target variable until discrimination is eliminated. They accordingly keep the changes to the data as minimal and unintrusive as possible. The last step is to train a model on the corrected data, referred to as *Classification with No Discrimination* (CND). Now the model should be ready to classify individuals without introducing discrimination.

Labor Market Discrimination Having addressed the issue concerning fairness and discrimination from an AI perspective, I find it natural to now approach the issue of labor market discrimination. Regardless of the number of conversations or coaching sessions offered by the job centre, if the employer discriminates based on age, gender, race, religion, or sexuality, etc., the job seeker may have a hard time getting back into employment.

It is desirable to build an AI tool that puts more weight on qualifications and less weight on apparent differences (such as gender, sexuality, religion, etc.). The same goes for the labor market, which has been proven to discriminate against certain minorities. A study by the University of Copenhagen shows that the chance of minority ethnic women wearing headscarves being called for an interview is significantly lower than for ethnic Danish women. The headscarf-wearing women must send 60 percent more job applications than an ethnic Danish woman to be called to the same number of interviews, despite having the same qualifications. A researcher examined the degree of discrimination in the Danish labor market by submitting 1,350 fictitious applications from ethnic Danish women, minority ethnic women, and minority ethnic women who wore headscarves. The results showed that ethnic Danish women were called to an interview in 30.6% of cases, while the numbers were 26% and 19.1% for minority ethnic women and minority ethnic women wearing headscarves, respectively. It is worth noting that even if a minority ethnic woman removed her headscarf, she would still have to work harder to compete with ethnic Danish women with the same qualifications (Marquardt & Carl, 2020). ‘’ Another study from the University of Copenhagen showed that applicants with Middle Eastern names had to submit 52% more job applications to be invited to a job interview. Two political science students sent 800 fictitious applications for 400 real positions, 400 with a Danish name and 400 with a Middle Eastern name. The applications were otherwise the same, but nonetheless, 33% of the applicants with Danish names were called for an interview compared to 22% of the applicants with Middle Eastern names (Lund & Pedersen, 2016).

Another issue is company dress codes, which have been discussed frequently in the media. Some companies have previously required women to wear high heels (Kay, 2019) and others require all men to shave their beards. An example is seen at Domino's Pizza, where a male Sikh applied for a managerial position but was denied employment after refusing to shave because of religious reasons (Renteln, 2008). By ignoring discrimination and corporate dress codes that might limit religious employees' right to wear symbols important for their identities, we end up treating symptoms and not the core problem.

One way of dealing with this from the job centers' side might be to recommend concealing certain CV information. Information regarding age, gender, home address, and civil status could easily be removed, and the same goes for pictures. The Danish Social-Liberal Party (Radikale Venstre) and the left-wing political party Independent Greens (Frie Grønne) are working to anonymize applications to hide information regarding age, gender, and origin. The hope is to shift the focus away from appearance and toward the applicant's qualifications. The process will start as an experiment in the public sector to see the response. In the meantime, firms in the public sector are now able to decide for themselves whether they require information regarding age and gender in a hiring process (Olsen, 2021b).

Subconclusion This section focused on the social impact of implementing AI and ML tools and presented related research along the way. The analysis aimed to clarify the trade-off between performance and transparency. When picking a more advanced ML model, a black-box problem might appear, which becomes even more critical to deal with when GDPR rules require adequate information on the logic involved. In this regard, I presented how eXplainable AI tries to solve this problem through the use of various techniques such as LIME and SHAP. Next, I introduced the trade-off between accuracy and discrimination, which is even harder to deal with, given the discrimination-by-proxy phenomenon, which referred to the possible indirect discrimination based on features correlating with discriminatory features. I distinguished between positive and negative discrimination and presented the excess number of conversations that non-Western immigrants' descendants deal with based on origin, which suggested that more conversations are being held with descendants than citizens of Danish origin. I also presented how researchers have tried to deal with discrimination by attempting to achieve fairness by modifying the algorithm or the underlying data.

Other subjects regarding the social impact of AI tools, such as economy, safety, robustness, accessibility, data management, or privacy, could also have been considered in my analysis. The next chapter will outline the process of developing an ML tool before presenting the findings in the Results chapter.

CHAPTER 3

Methodology

The chapter aims to outline the data mining process by laying the foundation of the theoretical approach. Inspired by an industry-accepted data analytics methodology, the subsequent sections will outline the process before presenting the results in a separate chapter. The research design is statistically founded with a large emphasis on quantitative measures, which will be used to answer my research question regarding the adaptation to potential discriminatory features while maintaining high precision. The theoretical techniques are implemented in python JupterLab, and the code is outlined in Appendix C. The reader should be aware that the results cannot be repeated as it requires access to the database used in the thesis.

3.1 Research strategy

To ensure a structured approach that makes it easy for the reader to understand my course, I chose to follow a data mining process called Cross-Industry Standard Process for Data Mining (abbreviated CRISP-DM). The methodology ensures a streamlined process that aims to help with the planning, organization, and implementation of a project. The process is outlined below in Figure 3.1, showing the six steps that I will go through in the chapter. An important distinction from other methodologies, such as SEMMA, is the iteration that is repeated as often as necessary to solve the business problem.

The first step, labeled business understanding, concerns an understanding of the problem at hand. Even though it might seem obvious, some problems might be opaque at the start but will become more apparent after having understood the data or even after a whole iteration through all six phases. Having a clear understanding of the problem, goal, expectations, success criteria, timeline, tools, techniques, etc., lays the foundation for the following steps. Having formed a clear and unambiguous understanding of the business problem, the next step is to understand the data by collecting the initial data, examining the data, performing an exploratory data analysis, and conducting a data quality assessment. At this stage, one might return to the first step and reaffirm, change or adapt the business understanding before moving on.

Having assessed the data, it would now be natural to prepare the data for analysis. This step will, in some cases, take up most of the project time depending on how raw the data is. Cleaning the data, engineering new features, scaling, constructing dummy variables, merging data from other sources, and formatting are all part of the data preparation step.

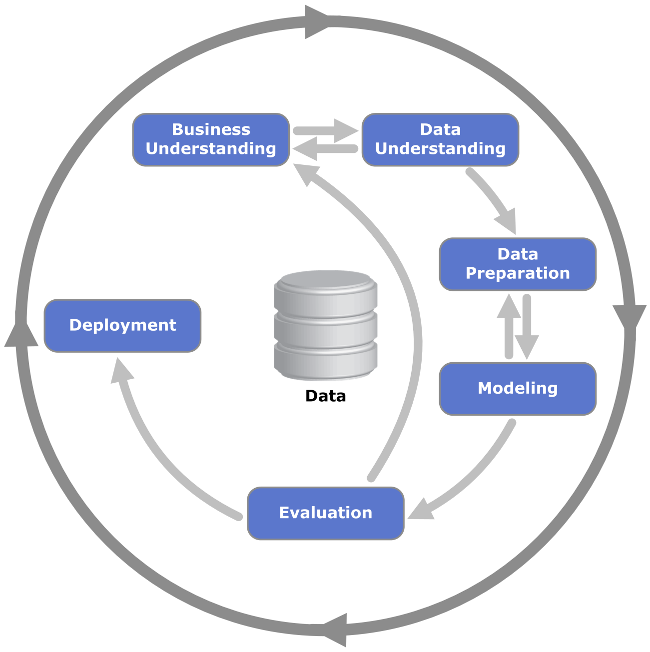


Figure 3.1: Cross-Industry Standard Process for Data Mining (Provost & Fawcett, 2013)

The different data mining techniques are applied in the modeling stage. According to the *no free lunch theorem* by David H. Wolpert (Raschka & Mirjalili, 2019), there is no single machine learning algorithm best suited for all situations. It is therefore recommended to try and compare a handful of models. Having built the models, one would need to evaluate them, which is connected to the business understanding. If the model does not meet the business criteria, it would require reiterating through the steps. Finally, the deployment phase is carried out, which could be as simple as publishing a report or making the model available for others on the net (Provost & Fawcett, 2013).

3.2 Business Understanding

In the wake of the complaints on social media, the articles written in the news, and the lawsuit filed against Danish municipalities for discriminating against their citizens based on origin, this project was brought to life. The Danish Agency for Labour Market and Recruitment is interested in evaluating the cost of removing potential discriminatory features from the present model. The cost in this regard should be interpreted as the cost of model precision. The management is willing to let me explore a range of ML models, which will be compared to the model implemented at the moment. Access to more data will be granted to compensate for the removal of other data (such as origin). The final outcome is not an actual model deployment but a report stating a recommendation for the management regarding future steps. The recommendations are expected to be delivered in May. Depending on the results, a trial period of two months is expected to occur in August and September, followed by model deployment in jobnet.dk in the fourth quarter of this year.

Limitations

I was granted access to JupyterLab (also referred to as an IDE), which is connected to a python database. My work in the IDE was limited to the conda package manager, which does not contain all ML models. As an example, I was not able to implement a Mixed Naive Bayes algorithm, but given the model assumptions, this is not seen as vital for the analysis.

3.3 Data Understanding

As explained in the second chapter's introduction, the screening tool was developed and implemented at the end of 2015. Parts of the questionnaire were made available from the very beginning, but the present form was introduced in 2017. In other words, the main part of my data had already been collected, although not preprocessed. The earliest unemployment process in my data started on January 4, 2016, and the last started on June 25, 2018. Unemployment processes ending in 2019 are included, while unfinished processes are labeled as the year '9999'. The data consist of 164,322 rows and 102 columns. An overview of the features is presented in Table 3.1, where I have split the features into eight categories. A more detailed overview of the features and their definition can be found in Appendix A Table A.1.

Count	Feature type	Label
1	Target variable	Target
4	General	ID, start, end, and duration [of unemployment process]
15	Personal and social characteristics	Gender, birthday, public_support_type, municipality, origin, age, parents_marital_status, provider, provider_count, marital_status, partner_occupational_status, age_first_child, negative_life_events, relocations_5, average_income_parents
2	Education	Education and graduate
34	Social benefit type	sd_all_12, sd_all_60, per_12, per_60, dp_12,...
9	Labor market affiliation	Unemployment_fund, unemployment_fund_duration, unemployment_fund_type, industry, employment_rate_12, employment_rate_36, aggregate_wage, aggregate_wage_missing_values and average_employment_period
30	Questionnaire	Q1, Q1a, Q1b, Q2, Q3, Q3a, Q3b,...
7	Others	Completion_of_questionnaire, load_date, run_date,...

Table 3.1: A categorical overview of my features

Exploratory Data Analysis

More than 70% of the data shown in the table above come from administrative data and the rest stems from the questionnaire. I will now perform an Exploratory Data Analysis (henceforth EDA) to unveil and explore possible patterns and relationships in the data. The analysis is not mutually exclusive but will act as a foundation for the following research. An EDA serves as an approach or technique to maximize insight into the data. By uncovering certain structures or patterns, we can detect anomalies and outliers or use it to perform a simple hypothesis test.

From Table 2.3 we can see that age and gender are the first mentioned features across countries. In my dataset, people range from 17 to 65 years of age, with a mean age of approximately 37. In Figure 3.2 I try to uncover the relationship between people's age and their unemployment duration (in weeks).

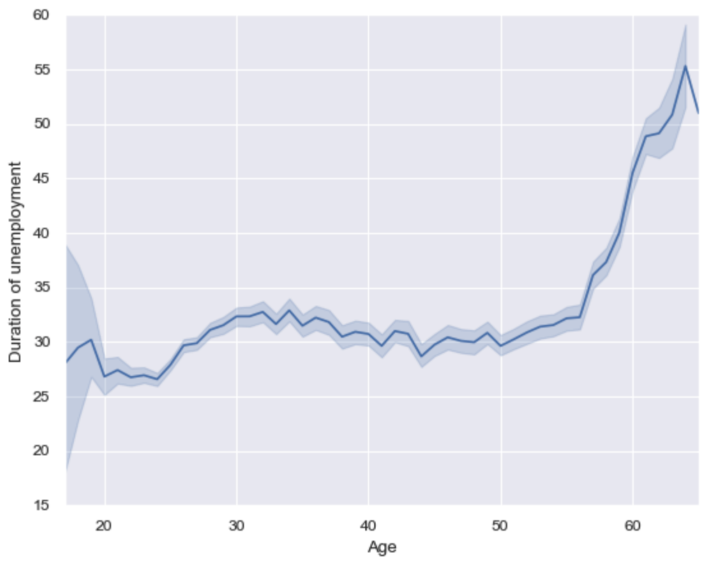


Figure 3.2: Age and duration (in weeks)

The first thing I notice is that people below 25 have the shortest unemployment process (henceforth duration), which might be connected to the employment reform introduced in chapter two. The reform required that people below 25 get sent to high school or university if they do not hold a degree. Between the ages of 25 and 30, we see a rise in the duration, which might be due to graduation and the struggle of landing one's first job. Between the age of 34 and 50, we see a fall that I can only assume is due to the rise in experience and network growth, making it easier to get back into the labour force. The last group, which is 50+, experiences a steep increase in duration.

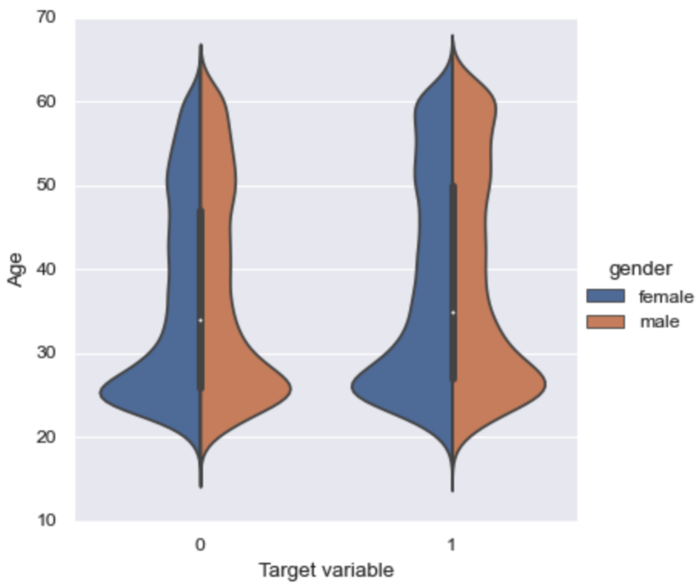


Figure 3.3: Age and gender

This development is in line with the statement made by job counselor Marianne, who said that people get labeled as 'seniors' despite their experience and knowledge in the field. Another problem might be that many of their skills have become outdated or been made obsolete. The duration tops at the age of 64, where people in the age category have an average unemployment duration of approximately 66 weeks or almost 1.5 years. Moving on to the next illustration (Figure 3.3), I next split people based on whether they are categorized as at risk of LTU (1) or not (0). The figure shows that most of the data samples consist of young individuals between 20 and 30 years of age. Comparing the two violin curves, we can see that people categorized as at risk of LTU are 'heavy' at the top, which is in line with our analysis of Figure 3.2. As an extra detail, I added information regarding gender, but the two genders look fairly equal to my surprise. I would have expected women to struggle more to get back into work, but the curves look very similar from this perspective.

An interesting feature in my data set concerns the existence of negative life events. The feature is defined as the death of parents, spouse, siblings, or children within the past five years. In Figure 3.4 people are categorized based on whether a negative life event has occurred recently, but also on age and whether they are at risk of LTU or not. Examining people with a negative life event, we see that people most at risk of LTU are elderly. On the other hand, we see that young individuals are primarily not affected by this, which does not come as a surprise. I do not consider this an important feature, but new data might make it more relevant given the current pandemic.

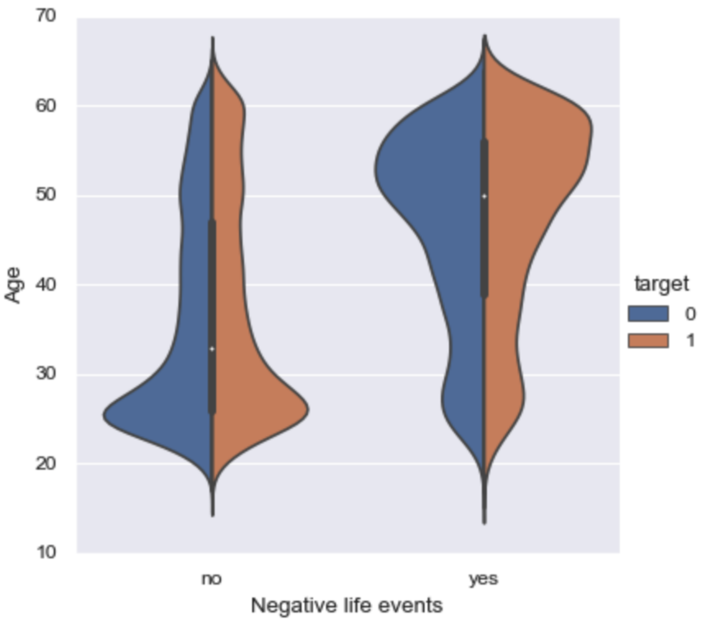


Figure 3.4: Age and life events

Given the discussion regarding discrimination, it was natural for me to analyze the data regarding origin. In Figure 3.5, people are split into five categories. Using box-plots, I tried to get an impression of the unemployment duration based on origin and gender. As a reminder, the boxes consist of three lines, with the upper line representing 75% of the data, the middle line representing 50% (the median value), and the bottom line 25%. The tails are referred to as upper-/lower whiskers, and the dots can be thought of as outliers or extreme values. If the box is centered between the whiskers, we say the data is normally distributed; otherwise the data is skewed. For example, 75% of the non-western male immigrants have a duration period below 40 weeks, which is the lowest among the groups in the table.

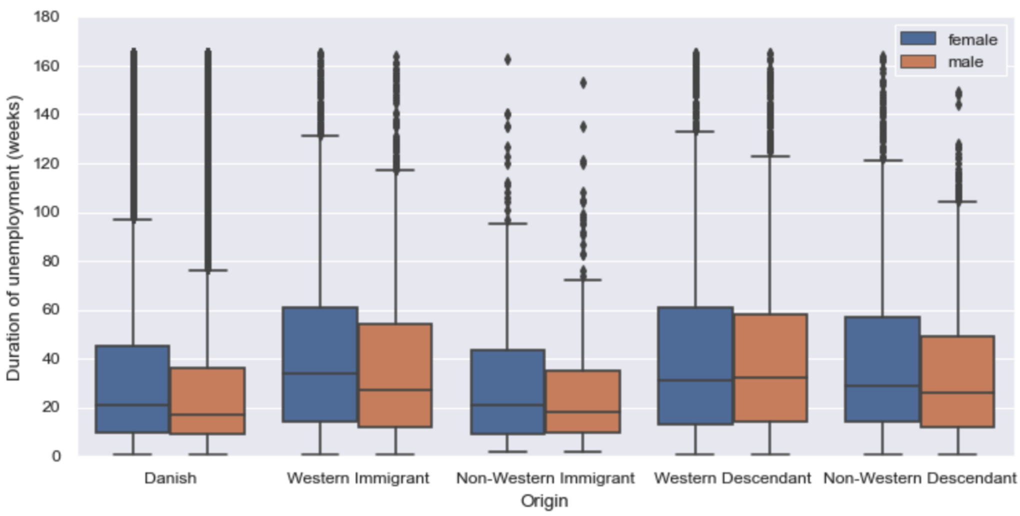


Figure 3.5: Duration and origin

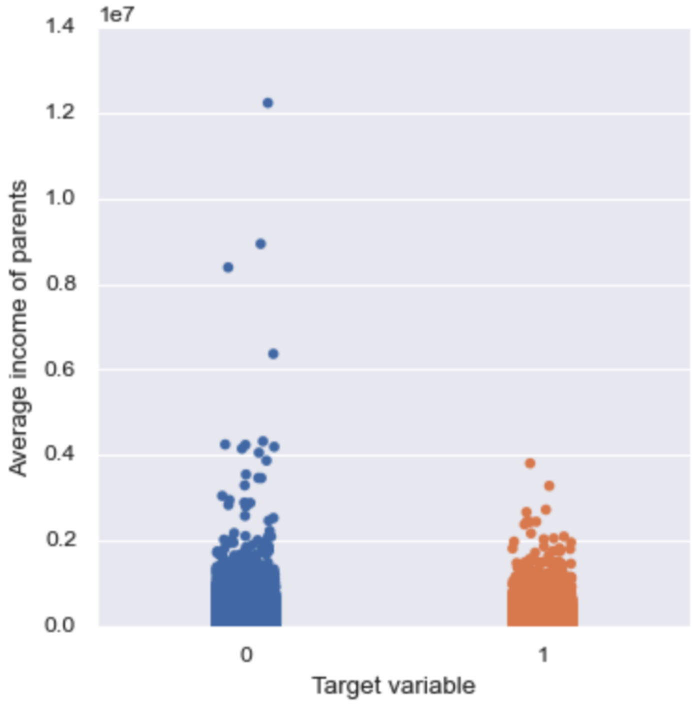


Figure 3.6: Average income of parents (in millions DKK)

On the other hand, their children, or the group of non-western descendants, experience a longer duration. This might be due to the difference in language and education level, enabling them to reach for more demanding jobs. The generation of non-western immigrants accepts low-paying jobs, working as cleaners, taxi drivers, chauffeurs, maids, and in housekeeping, etc. These positions are easy to fill and easy for the job centers to equip people for (e.g. with a driver's license). The next generation, which is born and raised in the country, aims for higher job positions, e.g. as doctors, lawyers, engineers, etc., positions that native-born citizens are also aiming for, which makes them more challenging to attain. They are generally fighting for positions where parents' network, experience, economy, etc., will be an advantage.

To clarify this point, I have added Figure 3.6, which illustrates the difference in parents' average income for people categorized as at risk of LTU or not. The average income of the parents of people not at risk of LTU is 135,694 DKK, while the number is 104,477 DKK for people at risk. The problem of labor market discrimination was raised in the previous chapter and might be worth considering when looking at the figure.

Unlike the impression gained from Figure 3.3, we now have another perspective on gender, revealing another story. Looking at the 75% line across origins, we see that women generally have a more extended duration period. This also goes for the median, except for western descendants, where men's median is higher.

Looking through the questionnaire, I found it interesting to look at the correlation between peoples' self-assessment of their duration and their actual duration period. Here I refer to the second question in Table 2.1, which is *How fast do you think you will get a job?*. Figure 3.7 illustrates the correlation between the answers and the duration period, and it is almost terrifying to see how close the correlation is. The figure shows that people expecting early retirement pay, a pension, or maternity leave spend the most time in the unemployment process, while citizens who expect to be back in work within 1, 3, or 6 months are almost always correct. The question seems to be highly relevant and is expected to be significant in my later analysis.

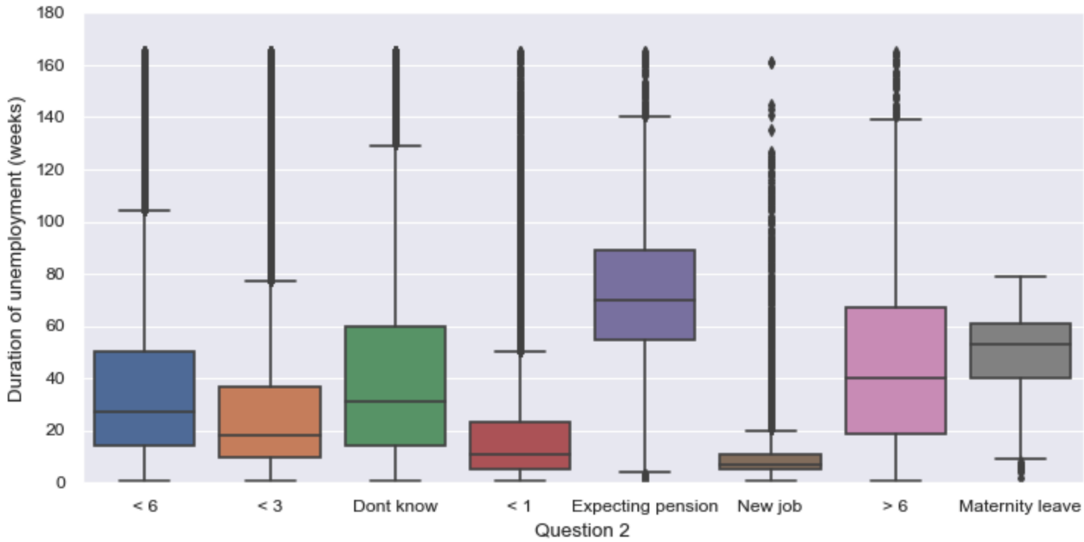


Figure 3.7: Q2 and duration (in weeks)

Other visual representations can be found in Appendix C. The EDA revealed some correlations and patterns which would have been hard to stumble upon when working with big data sets. Unfortunately, EDA still mainly consists of a continuous cycle of trial and error. A data scientist will have to experiment with different columns and visualization tools to gain insight. For this reason, it is mainly the start of our analysis.

3.4 Data Preparation

Data preparation, also known as data cleaning, is usually the most time-consuming step, depending on the data's state and size. It is vital to understand that data preparation is an art and that a single data set can be prepared in different ways. How the data is prepared has a direct influence on model performance.

Given that all the data is in Danish, the first step was to translate the data with appropriate labels and remove all duplicate rows. Afterward, I created a visual representation of the missing values in my DataFrame. Figure 3.8 shows all the features in my DataFrame, where yellow stripes represent missing values. In other words, if a column is entirely yellow, then it's empty.

Based on the visualization and other aspects, I had various choices to make. Some of the tools at my disposal would allow me to eradicate the column, fill it with values, make the missing values a separate category, import new data to outweigh the missing data, etc. Some tools made more sense given the size of the data set. For example, removing a thousand rows out of a million is like removing a drop of water from the ocean, while in other situations, it might damage the performance of a model.

In some scenarios, I replaced missing values with zeroes, while in others, it would make sense to use median values, mean values, or the most frequent value. As an example, we have 'aggregate wage', where the mean value of the entire column has replaced the missing values (this technique is referred to as *mean imputation*). This solution is seen as appropriate given the lack of outliers. When working with numerical features, the solution makes sense; if, on the other hand, I have categorical features, which contain text, I might create a new category, remove the rows with missing data, or use the most frequently occurring text.

Other scenarios might be simple mistakes, which do not appear in Figure 3.8 as a missing value. For example, we have the second question in the questionnaire, where a ninth category appeared despite only having eight possible responses. I found out that a single categorical answer was formulated in two different ways, so I found a way to group them.

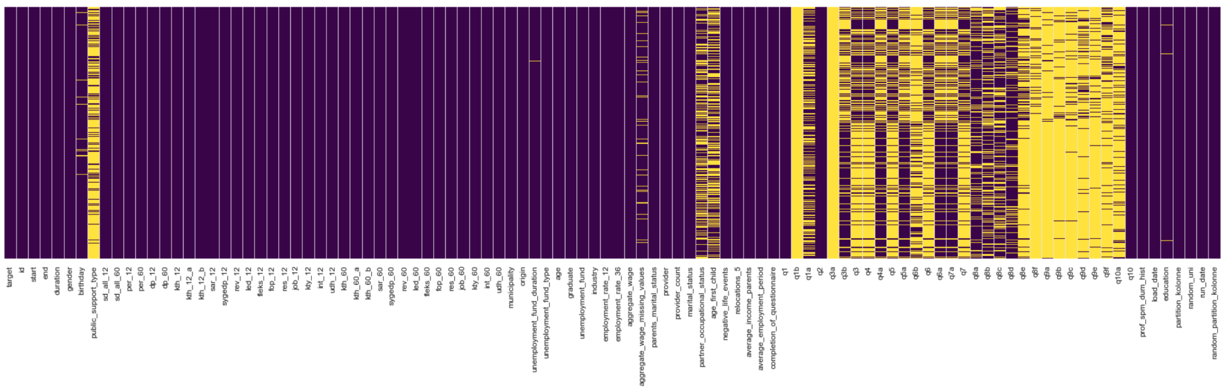


Figure 3.8: A visual illustration of missing values in my DataFrame

Problem of overfitting and multicollinearity When working with categorical data, we have to distinguish between ordinal and nominal data. Ordinal data is defined as data containing a hierarchy. An example is Q1, Q2, Q4, Q5, Q6, and Q7 from the questionnaire (Table 2.1). In this example I can rank the possible answers. On the other hand, nominal data is defined as data which do not have a natural ordering. An example is origin, gender, and education, where I cannot say that one is ranked above or below the others. To put them on the same level, I would have to create a separate column for each possible category. The columns will consist of binary data, i.e., ones and zeroes, referred to as dummy variables. Assuming that the citizen is female, then the female column will be 1, or 0 otherwise.

One problem with creating separate columns for each category is the problem of overfitting. For example, the municipality column contains 98 unique values, and given that it's a nominal data type, I would have to create 98 columns. Another example is education, where over 2,000 unique values are included. Going down this path, I ended up with over 3,000 columns. When training a model on the data, the algorithm might fit perfectly to the training set but perform poorly on the test set or other unseen data. In this situation, we say that the model has a high variance due to the significant change in estimate from one data set to another.
The opposite scenario is also possible, where underfitting occurs due to oversimplification. Here we say the model has a high bias, which is defined as the difference between our model's average prediction and the actual value we are trying to predict. One way of dealing with the risk of overfitting is to group the municipalities into eight regional labor market councils (RAR). The same goes for education, which is grouped into 19 categories

such as primary-/high school, bachelor's-/master's degree, Ph.D., etc.

A part of dealing with the risk of overfitting is dealing with intercorrelation or multicollinearity. When features are highly correlated among themselves, we say that a problem of multicollinearity exists. Not adjusting for or coping with multicollinearity will result in wrong parameter estimates and undermine the features' statistical significance. A typical example of multicollinearity is when all nominal categories of a feature are included in a regression model. By excluding a class, I can use it as a baseline to interpret the parameter estimates while avoiding multicollinearity. In Figure 2.6 I show the excess number of conversations. By excluding an origin and using it as baseline, I managed to get more accurate estimates.
Nonetheless, it's important to make a distinction between model interpretation and prediction. When the data preparation goal is interpretation, it's part of the model assumption that I do not have multicollinearity. However, I do not have to take multicollinearity into account when the goal is prediction. This distinction is outlined in the book *Applied Linear Statistical Models* by Kutner (2005):

> The fact that some or all predictor variables are correlated among themselves does not, in general, inhibit our ability to obtain a good fit nor does it tend to affect inferences about mean responses or predictions of new observations.

In other words, when preparing the data set for a machine learning context where the goal is prediction, I do not remove any columns.

Imputation and scaling Replacing missing data with numeric values is called imputation. As mentioned earlier, I used mean imputation to fill out the missing data in the aggregate wage column. Some missing values require more advanced imputation techniques to fill the gap. An example is my columns regarding questions 4 to 7, where almost half the data is missing. Removing the rows might have damaged my model performance, so I chose a more advanced approach called Multiple Imputation by Chained Equations (MICE). The logic behind the technique is quite simple. Using a feature with missing data as the target variable and all the others as independent variables MICE utilizes a linear regression model to estimate the missing values. The MICE algorithm is repeated several times to stabilize the results (Royston & White, 2011).

To use the MICE technique correctly, I would first need to split the data into a training and test set and afterward implement it on the two sets separately. This is illustrated in Figure 3.9. The reason for not using it on the whole set is due to fear of data leakage. Data leakage refers to a scenario where information from the training set is leaked to the test set, resulting in unrealistically high levels of model accuracy. Another example is the public_support_type feature in my data set, which tells me which support type a citizen is on after having finished their unemployment process on dagpenge (benefit type). The feature leaks certain information that an ML algorithm will quickly notice and therefore gives improper performance scores.

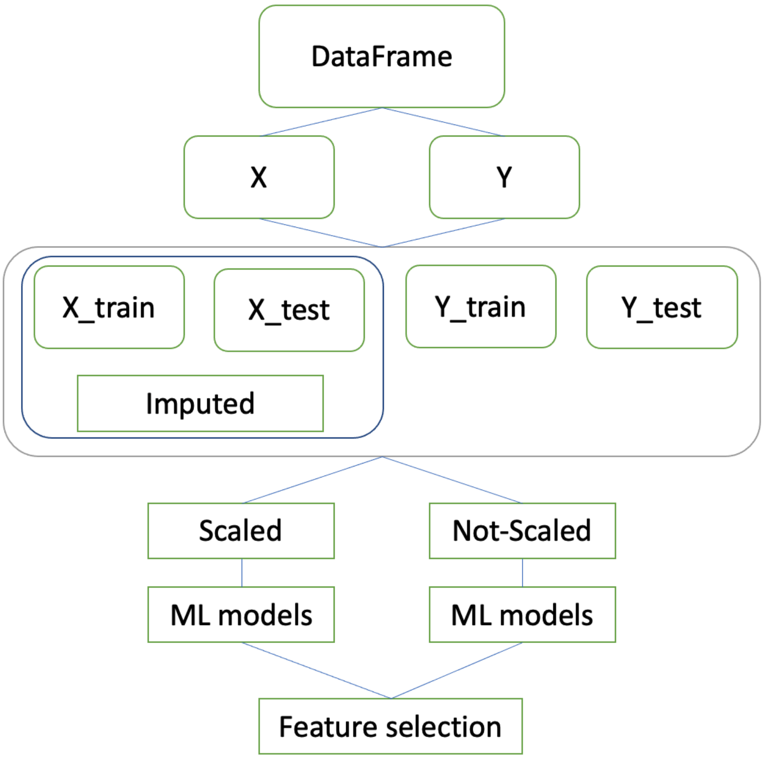


Figure 3.9: Process flow with imputation, scaling and feature selection

After dealing with missing data through row removal and imputation techniques, I was left with 159,087 rows (a reduction of 5,235). After creating new columns through merging and dummy variables, I had 133 columns (an increase of 31). Depending on the ML algorithm, I could now call it a day and move on to modeling. However, given that I chose to include ML models that are distance-based (which means that they categorize new data based on the distance to other data points), I needed to scale my data set. As illustrated in Figure 3.9, I chose to create two separate training and test sets, where one was scaled. I had different options for scaling: the most well-known are Min-Max Scaling and Z-score standardization. I chose to go with the first one, which transforms the data by scaling each feature to a certain range. The equation is shown below (Kutner, 2005):

$$X_{scaled} = \frac{X - X_{min}}{X_{max} - X_{min}} \tag{3.1}$$

The specific method was chosen to ensure that observations are on the same level as the dummy variables. The scaling was performed after the original data set had been split to avoid data leakage. A leakage could have happened by revealing the min and max values to the test set.

Having prepared the data, I was now ready for ML modeling. Several concepts and terms were mentioned in the previous chapter without a proper introduction. The following section will introduce the fundamental ML concepts and the underlying theory behind the chosen ML models.

3.5 Modeling

This section aims to introduce the reader to the different learning types and give an in-depth introduction to the ML algorithms. I previously defined ML as the process of solving a practical problem by gathering a data set and algorithmically building a statistical model based on the data set. The algorithm or the learning process depends on the type of data set. A data scientist or ML practitioner will generally apply two different types of learning approaches: supervised machine learning (SML) or unsupervised machine learning (UML).

Supervised machine learning When working on a supervised machine learning problem the goal is prediction. Based on a data set consisting of input variables x_i and output variables y_i the goal is to learn a mapping between the two. A data set is usually written as a set of labeled pairs $(x_1, y_1), \cdots, (x_N, y_N) = (x_i, y_i)_{i=1}^N$, which can also be expressed using matrices and vectors (notice that matrices are denoted with bold capital letters while vectors are denoted as bold characters):

$$\mathbf{X} = \begin{bmatrix} x_{11} & \cdots & x_{1D} \\ \vdots & \ddots & \vdots \\ x_{N1} & \cdots & x_{ND} \end{bmatrix}, \mathbf{y} = \begin{bmatrix} y_1 \\ \vdots \\ y_N \end{bmatrix}$$

The columns in the matrix represent the features, also called feature vectors. Other terms like independent variables, predictors, regressors, attributes, and explanatory variables are also used as references. The matrix consists of N rows and D columns (/dimension). The vector on the right-hand side is the target variable, also known as the dependent variable or the response variable. The target variable is not limited to a vector structure but is illustrated as a vector for simplicity. Relating it to my data, we see that the target variable has a vector structure, is a real number $y_i \in \mathbb{R}$ and belongs to a finite number of classes (at risk of LTU or not). Given that the target variable in my case is a category, we call it a classification problem. Had I chosen to work with the number of weeks a person is unemployed for, I would have had a regression problem.

Unsupervised machine learning The other possible scenario is having a data set without a target variable. In other words, we would be working with unlabeled data $(x_i)_{i=1}^N$. Through unsupervised machine learning models, one might try to discover hidden structures which can be visualized. As mentioned in the previous chapter, Odense Municipality used UML to find patterns that most effectively got unemployed citizens back into work. Other uses of UML include dimensionality reduction, density estimation, or a combination.

Besides SML and UML, we also have semi-supervised learning, reinforcement learning, and others. In my case, I only focused on SML, but other types of learning could have been used if time had allowed. I will now move on to introduce the ML algorithms used in my analysis with inspiration from Burkov (2019) and Raschka and Mirjalili (2019).

Logistic Regression

The Logistic Regression (LR) model is one of the most well-known algorithms of its kind. Despite the name, it is a classification learning algorithm. Its name is due to its similarity with the Linear Regression model, which is used for regression. The Linear Regression model consists of a linear combination of the features ($\mathbf{x}$) in a data set. One way of illustrating the linearity is through the following function:

$$f_{\mathbf{w},b}(\mathbf{x}) = \mathbf{w}\mathbf{x} + b \tag{3.2}$$

The vector $\mathbf{w}$ and the constant b are the parameters that will be estimated to optimize the model with the goal of making predictions. The values will be learned by the model when being trained on data. Values not learned by the model are called hyperparameters and are specified and tuned by the user of the model, although this is not the case for this specific model.

The above model can then be used for regression analysis to predict the target variable by assigning the result of the model to y, which can be expressed as follows: $y \leftarrow f_{\mathbf{w},b}(\mathbf{x})$. However, given that the results can range from minus infinity to plus infinity the prediction would not be appropriate for a classification problem where the target variable is either 0 or 1. One way around this is to wrap Equation (3.2) with another function that ensures a certain output. An example is the standard logistic function (also known as the sigmoid function):

$$f_{\mathbf{w},b}(\mathbf{x}) = \frac{1}{1 + e^{-(\mathbf{w}\mathbf{x}+b)}} \tag{3.3}$$

Here we have e, which is the natural logarithm, and we notice that it is raised to the power of $\mathbf{w}\mathbf{x} + b$, the equation from 3.2. The output can be interpreted as a percentage: if the output is 0.50 (50%) or above, we assign it to 1, or 0 otherwise. 0.50 is the standard threshold for all ML models. A model's threshold can be adjusted for specific situations, as I do later on in the results chapter.

To find the optimal values (labeled as $\mathbf{w}^*$ and b^*) for an ML model, we typically need to maximize or minimize an objective function. Other names, such as loss function and cost function, are also used. Think of the objective function as a way of expressing how Equation (3.3) fits the data or how badly our model is performing. When working with a Linear Regression model, one would typically optimize the model by minimizing the Mean Squared Error (MSE). When working with Logistic Regression, we optimize the model by finding the maximum likelihood. In other words, we try to fit a binomial distribution to the data by maximizing the following likelihood (L) function:

$$L_{\mathbf{w},b} = \prod_{i=1}^{N} f_{\mathbf{w},b}(\mathbf{x}_i)^{y_i}(1 - f_{\mathbf{w},b}(\mathbf{x}_i))^{(1-y_i)} \tag{3.4}$$

Translating Equation (3.4) into more simple terms, we see that the equation becomes $f_{\mathbf{w},b}(\mathbf{x}_i)$ when $y_i = 1$ and $(1 - f_{\mathbf{w},b}(\mathbf{x}_i))$ when $y_i = 0$. Here, we find that $f_{\mathbf{w},b}(\mathbf{x}_i)$ is the probability of becoming LTU ($p(X)$) and $(1 - f_{\mathbf{w},b}(\mathbf{x}_i))$ is the probability of not becoming LTU ($1 - p(X)$). For samples in our data which are categorized as at risk of LTU we try to estimate the parameters ($\mathbf{w}$ and b) such that $p(X)$ is as close to 1 as possible and vice versa.

The model manages to perform well when the data is linearly separable and can be adjusted for cases where the target variable has more than two possible outcomes (multinomial or multiclass classification). But, unfortunately, the model does not perform well with high dimensional data. Despite its limitations, Table 2.2 shows that six countries have chosen to use the model for job seeker profiling.

Support Vector Machine

Regardless of whether we talk about Linear Regression or Logistic Regression, the aim is to fit a model to the training data. With a Support Vector Machine (SVM), the aim is to find a model which separates the data. The separation can be as simple as a single numerical value (a scalar), a line, a plane, or a hyperplane depending on the data's dimensionality. Assuming that the data is not separable in a specific dimension, the algorithm will then try to move the data into a higher dimension to separate the observations. As an example, Figure 3.10 shows how two classes (blue and orange) are divided by the red line when working in a two-dimensional space.

The line used to separate the data is called the decision boundary and has two main functions, the first being a separation of the classes, and the other to stay as far away from the training instances as possible. The distance between the separated classes is called the margin, and as illustrated, we have three lines: a positive line, a negative line, and one in the middle. We want to maximize the distance between the positive and negative line, which is fulfilled by the red middle line. By maximizing the margin, we make sure the model is better at classifying unseen data.

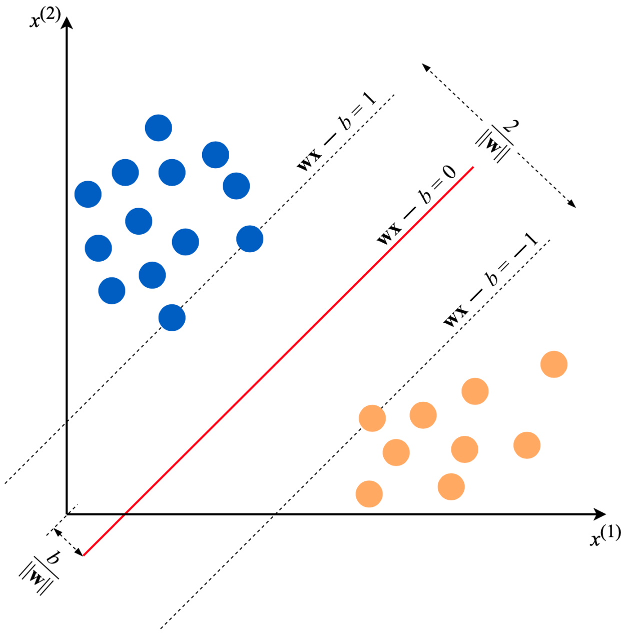


Figure 3.10: An example of an SVM model for two-dimensional feature vectors (Burkov, 2019)

When working with a classification problem, the SVM algorithm requires two distinct values for the target variable. My target variable was denoted {0, 1}, but given that the theoretical proof outlined in textbooks assumes a target variable denoted {-1, +1}, I assumed this was also the case for my data. In other words, I had $y = +1$ when citizens were at risk of LTU and $y = -1$ otherwise. To understand the assumption I will need to present the SVM function, which is illustrated below:

$$f_{\mathbf{w},b}(\mathbf{x}) = sign(\mathbf{wx} + b) \tag{3.5}$$

The SVM is a simple sign function (also known as signum function) that takes the values presented earlier and predicts the target value to be +1 if the input is positive, or -1 otherwise. In order to find the optimal values ($\mathbf{w}^*$ and b^*) I needed to solve an optimization problem that satisfied the following constraints:

$$\mathbf{wx}_{pos} + b \geq +1 \quad if \quad y_i = +1$$

$$\mathbf{wx}_{neg} + b \leq -1 \quad if \quad y_i = -1$$

The above constraints say that the sign function's input should be positive when the target value is positive and vice versa. They are the foundation for the positive and negative dotted parallel lines in Figure 3.10. To maximize the margin, we need to express the distance mathematically. We do this by subtracting the constraints from each other:

$$\Rightarrow \mathbf{w}(\mathbf{x_{pos}} - \mathbf{x_{neg}}) = 2$$

The difference is an expression of the distance between the positive and negative lines in Figure 3.10. We can normalize the above expression by dividing both sides by the length of vector $\mathbf{w}$. The length is given by the Euclidean norm:

$$\|\mathbf{w}\| = \sqrt{\sum_{j=1}^{D} \left(w^{(j)}\right)^2}$$

So, we arrive at the following equation:

$$\frac{\mathbf{w}(\mathbf{x_{pos}} - \mathbf{x_{neg}})}{\|\mathbf{w}\|} = \frac{2}{\|\mathbf{w}\|}$$

The distance is illustrated in Figure 3.10 and is a mathematical expression of the margin. We want to maximize the margin subject to the two constraints presented earlier. In this way, we hope the model manages to set all positive class examples on one side of the line and all negative on the other. The division might not always be possible for a straight line, so the SVM model incorporates kernels that can make a non-linear decision boundary. As hinted earlier, this happens by transforming the original space into a space of higher dimensions.

The model is chosen based on its ability to separate data in higher dimensions. To incorporate the model into my analysis, I had to take into account that it is a distance-based model. In other words, I had to scale my data as I have tried to illustrate on the left side of Figure 3.9. Another strength of the model is its ability to handle outliers and its speed when working with small data sets. Unfortunately, my data set was large, which makes such a process computationally expensive. Given its sensitivity to parameter and kernel choice, it would take a significant amount of time to process and tune.

Given the shortcomings, I decided to use *Stochastic Gradient Descent* to solve the optimization problem. Stochastic Gradient Descent is an optimization algorithm used to find a local or global minimum of a function. The optimization criterion of SVM is differentiable and has a convex form, which makes it suitable for Gradient Descent. The idea of the method is to give $\mathbf{w}$ and b random values and update the values by analyzing the entire data set repeatedly until a certain minimum is reached. Instead of analyzing the whole data set to update the parameters, I used a specific type of Gradient Descent called Stochastic Gradient Descent, which only explores a subset of the data to update the values. The method is also applicable to Logistic Regression and significantly reduces the time spent training the model.

Decision Tree

The third model I will introduce is the Decision Tree (DT), and as shown in Table 2.2, it is the model that is currently implemented in Denmark. The model has proven to be an asset based on its performance, interpretability, and explainability. An illustration of the current model can be found in Appendix A Figure A.5.

In general, a DT asks a question and classifies an observation based on the answer. You can think of it as an automated, rule-based profiling where job seekers are classified based on age, education level, unemployment duration, etc. This section will introduce a specific type of DT called ID3, which stands for Iterative Dichotomiser 3. As the name reveals, the algorithm iteratively (repeatedly) dichotomizes (divides) features into groups (Sakkaf, 2020). As illustrated in Figure 3.11, the model has a tree structure, where the top square is called a root node, and the final squares are called leaf nodes. There are usually squares in between, which are called internal nodes (a larger structure is presented in Figure A.5).

The ID3 algorithm starts by calculating the entropy of the whole data set. Think of entropy as the measure of disorder or randomness, with a value between [0, 1]. When it is 1, it means that we have a mixed data set where the target column has an equal number of values for both classes. On the other hand, it is 0 when the data set is homogeneous and only contains a single class of the target column. The entropy is given by the following equation:

$$H(S) = -f^S_{ID3} ln f^S_{ID3} - (1 - f^S_{ID3}) ln(1 - f^S_{ID3}) \tag{3.6}$$

where

$$f^S_{ID3} = Pr(y = 1|\mathbf{x}) = \frac{1}{|S|} \sum_{(\mathbf{x},y)\in S} y$$

Based on a set of labeled examples, denoted as S, the entropy is calculated as a difference in ratios. The ratio is presented by the f^S_{ID3}, where the mean average of the target values, in a set, is calculated. Notice that the ratio is a simple conditional probability which can be translated into the following question; *Given feature* $\boldsymbol{x}$*, what is the probability of becoming LTU?*. The entropy of my processed data set is calculated below:

$$f^S_{ID3} = Pr(y = 1|\mathbf{x}) = \frac{1}{159.087} * 67.585 = 0.4248$$

which gives me an entropy of:

$$H(S) = -0.4248 * ln(0.4248) - (1 - 0.4248) * ln(1 - 0.4248) = 0.6818$$

The algorithm will now try to split the data set by different features and thresholds to minimize the entropy. My analysis showed that Question 2 from Table 2.1 was chosen as the root node, which means that it reduced the entropy the most. The entropy of a split is shown in the equation below, which is a weighted sum of two entropies $H(S\_)$ and $H(S_+)$:

$$H(S\_, S_+) = \frac{|S\_|}{|S|} H(S\_) + \frac{|S_+|}{|S|} H(S_+) \tag{3.7}$$

The same procedure is illustrated below, where feature x^3 was chosen, and the threshold was set at 18.3. Although just an example, it shows that a conditional probability is calculated depending on whether the feature is above or below 18.3:

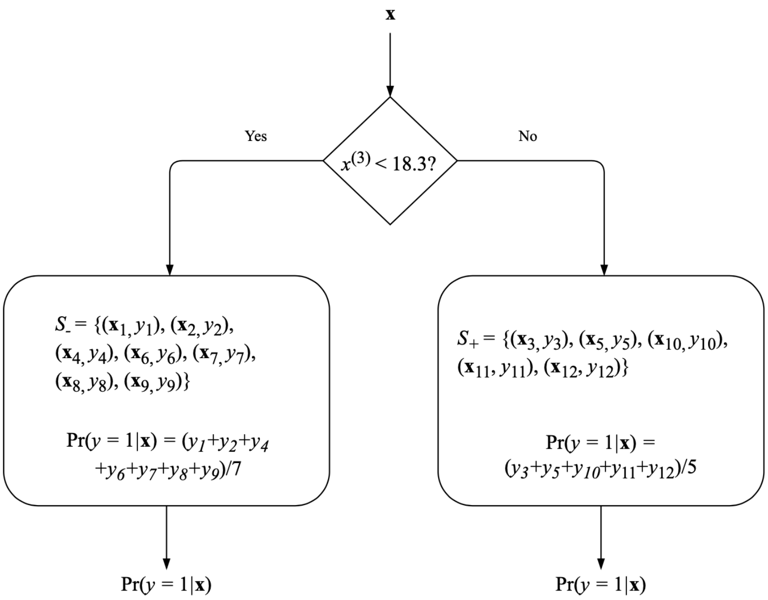


Figure 3.11: An example of a DT model (Burkov, 2019)

The algorithm will keep on splitting features until all are processed or until the entropy effect is gone. The entropy equation is an impurity function that often differs depending on the type of DT. Other algorithms evaluate a split on the Gini index or Information Gain. My analysis algorithm is called Classification And Regression Tree (CART), which produces either classification or regression depending on the features.

The DT model is simple to understand, visualize and most importantly, explain to others. With no assumption about the data and without any need to scale or remove outliers, the model can easily be implemented. However, the model is often prone to overfitting and is vulnerable to class imbalance. More advanced DT algorithms can handle overfitting by pruning, which means that leaf nodes replace branches with an insignificant impact. Regardless of the chosen algorithm, the main idea remains the same.

The DT's challenges and limitations can be overcome by simply combining a range of them in a so-called meta-model. In other words, instead of training a single model, we can combine multiple models and base a prediction on the overall results. The approach is known as ensemble learning, a learning paradigm that focuses on training many low-accuracy models and combining them to obtain an overall high accuracy. If the joined trees vary slightly from each other and perform somewhat better than a simple random guess, then it's expected that a combination will yield a high accuracy. There are two ensemble learning methods, boosting and bootstrap aggregation (also called bagging). I implemented both methods and will introduce them separately.

Random Forest

A well-known learning algorithm is the Random Forest (RF), which is based on the idea of bagging. The first step is to create B bootstrapped data sets S_b ($b = 1, ..., B$), which are copies of the original data set, but with minor variations. Each copy consists of randomly selected samples from the original data set and has the same length. The same observations can be picked more than once from the original data set, which means that some might be left out. In general, 1/3 of the original data will not end up in the bootstrapped data set. By using slightly different variations of the original data we reduce the variance of the final model.

Having created the copies, the algorithm will then train a unique DT model f_b on each bootstrapped data set S_b. An important detail is that the RF algorithm inspects a random subset of the features at each split in the learning process. By doing this, we avoid correlation between the trees and increase accuracy.

The prediction for unseen data is a simple average of B predictions if it's a regression problem. Alternatively, it's a majority vote in the case of classification. The method is known for its ability to handle a massive amount of data, and given that the final prediction is made by consulting a huge amount of trees, we do not have to worry about overfitting or outliers. On the other hand, we move away from white-box models and start entering the problem of black-box models. Although we can visualize the underlying trees, it wouldn't make sense to go through a hundred or a thousand trees to thoroughly understand a prediction.

Gradient Boosted Decision Tree

The other ensemble learning method is called Gradient Boosted Decision Tree (GBDT) or simply Gradient Boosted Machine (GBM) and is based on the idea of boosting. It's one of the most powerful machine learning tools where the algorithm iteratively creates a DT model that corrects the previous model's mistakes. The process differs depending on whether we have a classification or regression problem. Given that our target is binary, I will introduce GBDT for classification but draw some parallels along the way to the process of regression.

The initial move of the GBDT algorithm consists of making a single prediction for all observations. In other words, imagine that we modify the target label of all samples $i = 1, ..., N$ with the following:

$$\hat{y}_i = y_i - f(\mathbf{x}_i) \tag{3.8}$$

where

$$f(\mathbf{x}_i) = f_0 = ln(odds)$$

The f_0 is the initial prediction for every individual (had I had a regressional problem, I would have used the average as an initial prediction), and the difference between the actual value (y_i) and the predicted value ($f(\mathbf{x}_i)$) is called the residual (or error). Given that I have 67.585 samples where $y = 1$ and 91.502 where $y = 0$, the natural log is $ln(\frac{67.585}{91.502}) = -0.30$. In order to use it for classification we simply use the sigmoid function presented in the Logistic Regression section (Equation (3.2)):

$$Pr(y = 1|\mathbf{x}, f) = \frac{1}{1 + e^{-(ln(odds))}} = \frac{1}{1 + e^{-(-0.30))}} = 0.43 \tag{3.9}$$

Given a threshold of 0.50, the initial prediction is that all observations are not at risk of LTU ($y = 0$), but subtracting zero from the actual values wouldn't make sense. Instead, I deducted the prediction, which yielded the new target label:

$$\hat{y}_i = y_i - 0.43$$

The new data set with the modified target labels was then used to build a DT model, f_1. The model was trained in predicting the residual and was combined with the initial equation f_0 to make new predictions:

$$f(\mathbf{x}_i) = f_0 + \alpha f_1$$

The combination was accompanied by α, which is the learning rate used to scale the data (usually 0.1). Some in-between calculations were left out to simplify the overall process.

Let's take a step back and try to understand the intuition. I started with an initial prediction or guess of the target value (which was 0.43) and used it to calculate the residual, which is the difference between the actual value and the predicted value. Having calculated the residual, I then used it with the original data to train a DT model to correct or minimize the residual from the first prediction. I did this by simply combining the initial prediction (f_0) with the new model (f_1) and a learning rate (α). Afterward, I reevaluated the residual and kept on building and adding new DT models to correct the previous mistakes ($f_0 + \alpha f_1 + \alpha f_2 + ...$). I kept on adding trees until a predefined maximum was reached or until the effect of adding new trees was null.

Although powerful, the underlying mechanism can be hard to follow, and without proper hyperparameter tuning, overfitting can happen. It should now be clear that the idea of boosting is to correct the mistakes of former models while bagging builds models parallel to each other and gives each model a final say.

The models introduced so far can be thought of as shallow learning algorithms, which means that the parameters learned come from the features. I will now present a so-called deep learning algorithm, or deep neural network, which learns the parameters from the preceding layers.

Multi-Layer Perceptron

Moving on from the more traditional ML models, I will now dive into a subsection called Deep Learning (DL). The computational model I am interested in is called Artificial Neural Network (ANN) and is the building block of DL. The idea behind ANN can be traced back to 1943 when it was introduced by neurophysiologist Warren McCulloch and mathematician Walter Pitts. The motivation was to describe how neurons in the brain work. The model, like the brain, receives input through an input layer and processes the data in so-called hidden layers before delivering a prediction in the output layer. While the model does have some analogies with the human brain, it's not an actual depiction of how the brain functions. It would be more accurate to think of it as a simplified inspiration (Ghanoum, 2020).

I am interested in presenting a specific architecture of a neural network called Multi-Layer Perceptron (MLP). Other names such as vanilla neural network or feed-forward neural network (FFNN) cover the same theoretical foundation. Unlike the functions presented for LR or SVM, the MLP consists of nested functions:

$$y = f_{ANN}(\mathbf{x}) = f_3(\boldsymbol{f}_2(\boldsymbol{f}_1(\mathbf{x}))) \tag{3.10}$$

The number of nested functions is chosen by the analyst, and as presented in Equation (3.10), I decided to have three layers, two hidden and one output. The layers are visualized in Figure 3.12, where a feature vector consisting of two values is forwarded into the $f_{ANN}(\mathbf{x})$ function (illustrated as two green circles). Afterward, the input is received by the intermediate layers, where all the computation is done. We refer to them as hidden because they are unseen. Each layer consists of neurons; the figure shows two hidden layers with four neurons each (illustrated as blue and purple rectangles), although I activated 2x100 neurons in my actual analysis.

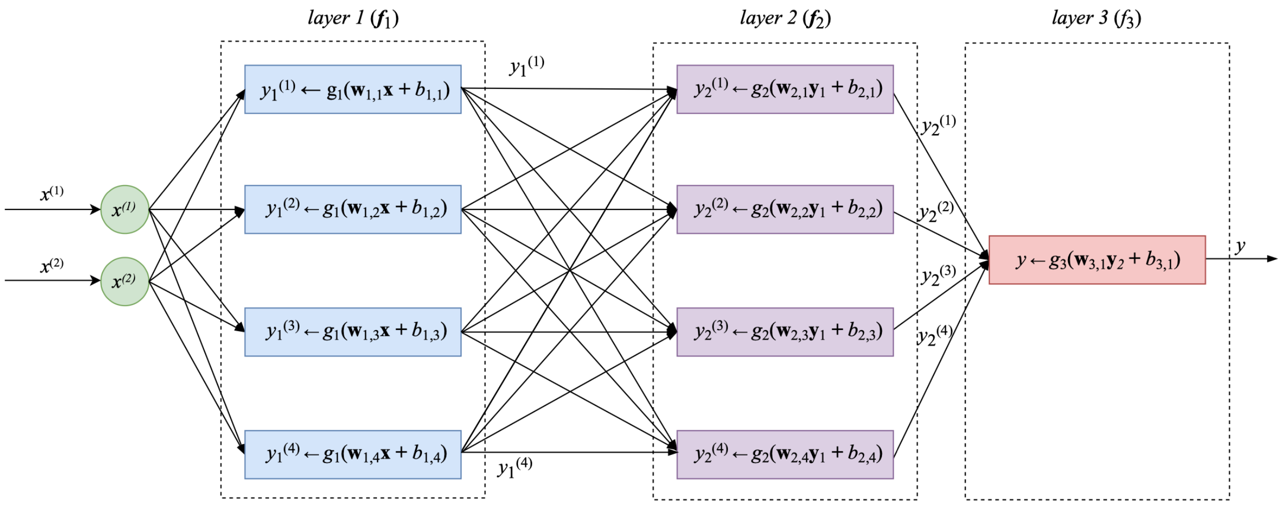


Figure 3.12: A multilayer perceptron with two-dimensional input, two layers with four neurons and one output layer with one neuron (Burkov, 2019)

The input values are fully connected to each neuron in the first hidden layer. After performing some computation in the first layers, the output is connected to every neuron in the second layer before being passed on to the last layer. The computation in the hidden layers is performed by vector functions where f_1 and f_2 are given by the following equation:

$$\boldsymbol{f}_l(\mathbf{x}) = g_l(\mathbf{W}_l\mathbf{x} + \mathbf{b}_l)$$

Here we have l, which is the layer index ($l = 1...\infty$), and g_l, an activation function that decides whether the neuron's output should be passed on or not. Inside the parentheses, we have what resembles a linear regression: a matrix $\mathbf{W}_l$ consisting of weights and a vector $\mathbf{b}_l$, which is referred to as the bias. While the analyst chooses the activation function upfront, the rest of the parameters are learned using gradient descent.

Focusing on the activation function, the purpose of the function is to make the MLP flexible. If we only added linear components, we would only produce linear models, but by adding non-linearity, we allow the model to approximate non-linear functions. There is a range of activation functions, and more than one can be added, but I chose to only implement an activation function called Rectified Linear Activation (ReLU):

$$ReLU(z) = \begin{cases} 0 & \text{if } z < 0 \\ z & \text{if } z >= 0 \end{cases}$$

Assuming that the linear regression inside the parentheses is negative ($\mathbf{W}_l\mathbf{x} + b_l < 0$), the ReLU activation function (g_l) will then output zero. On the other hand, if the regression is zero or above, the regression will be untouched by the function. The last layer has a separate activation function, depending on the type of problem. Given that I have a classification problem, the last activation function is logistic.

Every connection in Figure 3.12 is weighted, depending on if the input has a positive (**+w**) or negative (**-w**) impact, and the importance is reflected in the size of the weights. The initial weights are random numbers, but as the model trains, they get assigned more appropriate weights to minimize a cost function (this is done using an algorithm called backpropagation). I assumed that age would have a negative weight while a specific type of education would have a positive weight.

The ANN models can be used in all imaginable scenarios, although it comes at the cost of interpretability. The algorithm is computationally slow, which also makes it hard to tune and adjust. The MLP model used in my analysis relies solely on the computer's CPU, but other algorithms delivered by Keras can run on GPU, making the calculations faster. Unfortunately, I was limited to my workplace's packages, which is why I chose this specific model.

3.6 Evaluation

After choosing my models, the next step was to assess the models' performances. A British statistician named George Box once said: *All models are wrong, but some are useful* (Burkov, 2019). I did not expect to find or build the perfect model, but I hope that my results can give STAR a clear recommendation on changing the current model or keeping it. To assess the performance, I chose to split my data into a training and a test set (illustrated in Figure 3.9). Given the size of the data set, I only put 10% aside for training (15,908 rows) and stratified my data, which means that the proportion of values in the target variable will be the same in both the training and the test set. Afterward, I evaluated my models by exposing them to the training data. The evaluation process differs depending on the type of problem; I will present metrics such as accuracy, precision, recall, f1 score, AOC, and AUC-PR because I have a classification problem. I will use my RF model results as an example throughout this section and gather all models in the next chapter.

Regardless of the metric, it is wise to start by introducing the confusion matrix, a matrix which summarizes the performance of a model in a table, by comparing the predicted values with the true values. As an example, Figure 3.13 shows the Random Forest model's confusion matrix. The matrix consists of four squares because we have a binary classification problem. The squares report the counts of True Negative (TN), False Positive (FP), False Negative (FN), and True Positive (TP).

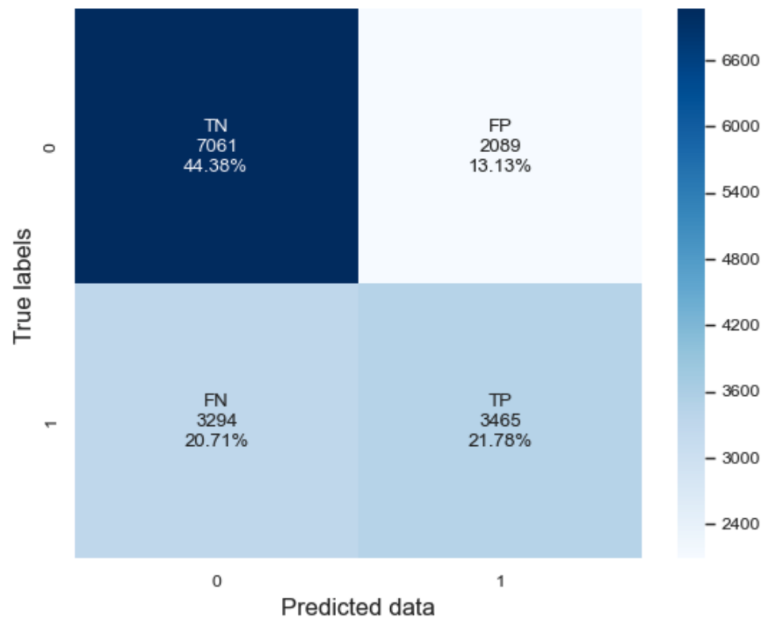


Figure 3.13: A confusion matrix for the Random Forest model

The perfect scenario has values only in the diagonal squares (TN and TP). Looking at the RF model's performance, we see that it misclassified 2,089 people as being at risk of LTU (also known as Type 1 error) and predicted that 3,294 people were not at risk, although they were (also known as Type 2 error). The confusion matrix lays the foundation for some of the most popular performance metrics, the first being the precision, given by the following formula:

$$Precision = \frac{TP}{TP + FP}$$

As we can see, precision is a simple ratio where we look at the people who are actually at risk of LTU out of all the people classified as being at risk. In this case, we see that RF has a precision of ($\frac{3.465}{3.465+2.089} =$) 62.39%. The next metric is the recall (also known as sensitivity):

$$Recall = \frac{TP}{TP + FN}$$

The equation looks very similar to the precision, I only replaced FP with FN, which shifts the focus to evaluate how many people the RF classified as at risk of LTU out of all the people actually at risk. The RF has a recall of ($\frac{3.465}{3.465+3.294}$ =) 51.26%. Looking at Table 2.2, we see that models, across borders, are compared on their accuracy and AUC score. The accuracy is relevant when both classes (0 and 1) are seen as equally important and is given by:

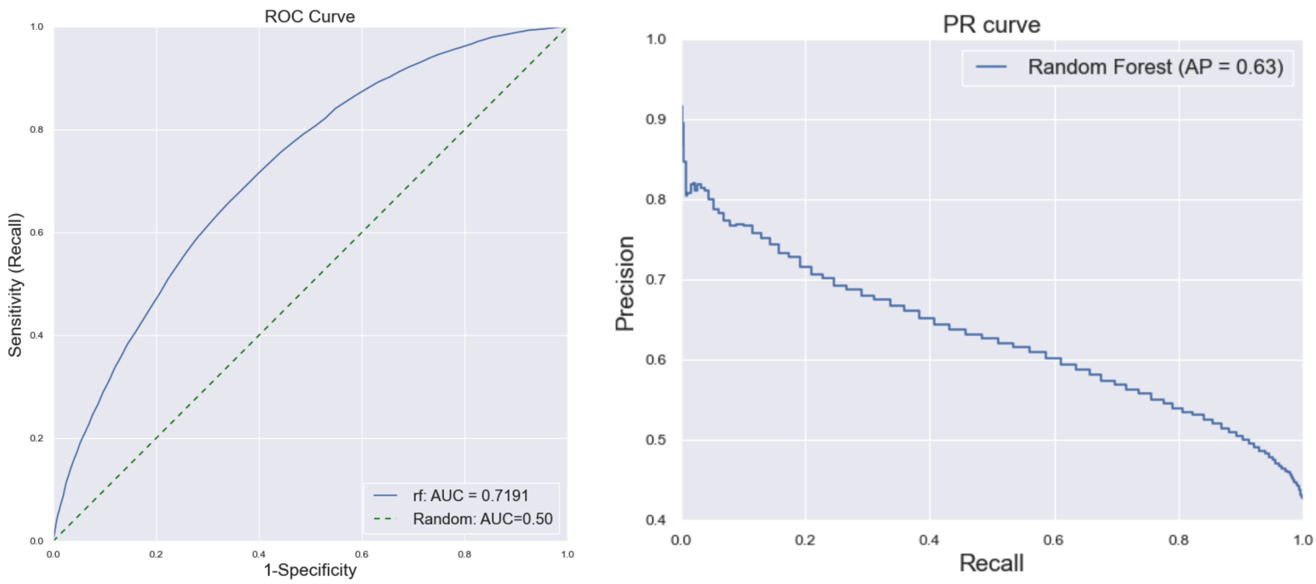
$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

In our case, the RF has an accuracy of ($\frac{3.465+7.061}{3.465+7.061+3.294+2.089}$ =) 66.16%, which is at the same level as Belgium (Table 2.2). Having a high precision often comes at the cost of a lower recall and vice versa. A way to balance the two is a combination where precision and recall are weighted equally. The metric is simply referred to as f1 score:

$$F1 = 2\frac{Precision * Recall}{Precision + Recall}$$

The RF has an f1 score of ($2 * \frac{0.6239*0.5126}{0.6239+0.5126}$ =) 56.28% (minor differences due to rounding).

It should be clear that the above metrics depend on the confusion matrix, and the matrix depends on the model threshold. As mentioned earlier, the baseline threshold is at 50, but what if we moved it? For each new threshold, a new confusion matrix appears, giving us new performance scores. I will now introduce two methods, illustrated in Figure 3.14, that measure performance across thresholds. The first one is the Receiver Operating Characteristic (ROC) curve, where sensitivity and 1-specificity are calculated at each threshold. The sensitivity is recall in disguise, and 1-specificity is simply the proportion of negative examples classified incorrectly ($\frac{FP}{FP+TN}$), also known as the false-positive rate.



(a) AUC

(b) AUC-PR

Figure 3.14: The ROC and PR curve for the Random Forest model

The ROC curve, Figure 3.14 (a), shows two lines. The green dotted line illustrates how a random classifier performs (think of it as classifying people by tossing a coin). The Area Under the Curve (AUC) is then measured and displayed. The more the curve goes to the northwest corner (upper left), the bigger AUC, the better. My RF model has an AUC score of 71.91%, which is lower than Belgium's score, but in the same range as Ireland and New Zealand. As a statistical instructor, I would usually tell my students that an AUC in the range of 70%-80% is 'okay', but given that it performs in the same range as other countries, I would interpret it as a good performance.

The second method, shown in Figure 3.14 (b), is simply a Precision-Recall (PR) curve, which clearly shows the tradeoff between the two. Assuming that I had a perfect model, I would see a curve that touches the top of the graph and the right side. We also have an AUC for the PR curve, which is called AUC-PR. In this case, the AUC-PR is 63%.

Having introduced a range of metrics, I will close the chapter by discussing the metrics' ranking. When hyperparameter tuning an ML model, it's worth considering which metric to optimize. The answer can differ depending on the use and the consequence of false positive and false negative. The procedure I used in my analysis was to start by comparing the f1 score, AUC and AUC-PR of my models; afterward, I picked the top two models and tuned their hyperparameters. The tuning is based on precision, which is valued higher than recall. This is due to the cost of a false positive compared to a false negative. By focusing on the precision, the hope is to lower the probability of falsely labeling people as at risk of LTU. The decision to value precision over recall was already made by STAR when implementing the original DT model. It was therefore necessary for me to do the same in order to have a basis for comparison.

Focusing on precision might be connected to the cost of investing in people at risk of LTU. It could also be a sign that the tool is seen as negative discrimination due to the possible burden on people categorized as at risk of LTU.

Having tuned my top two models, I then experimented with different thresholds, using the PR curve as inspiration. The final model was then trained on data with and without discriminatory features to assess the trade-off between precision and discrimination.

CHAPTER 4

Results

Having laid the foundation of the theoretical approach, this chapter presents my findings. Given the extensive range of features in my analysis, I will start by introducing the most significant features and compare them to the current model.

4.1 Feature importance

There is a range of feature selection methods, and I chose to use an intrinsic method. The method is naturally incorporated within the modeling process; I therefore decided not to dedicate space to it in my methodology chapter. In other words, all my models, besides MLP, could easily tell me which features they found most vital for the classification. These findings are presented in Table 4.1, where only the top ten features are included.

	RF	**DT**	**GBDT**	**LR**	**SVM**
1	Q2	Q2	Q2	Aggregate wage	Q2
2	Unemployment fund duration	Unemployment fund duration	Employment rate (36)	Kty (12)	Age
3	Employment rate (36)	Employment rate (36)	Q6	Q2	Q1a_1
4	Aggregate wage	Aggregate wage	Origin Danish	Kty (60)	Sygedp (12)
5	Age	Average employment period	Age	Age	Average employment period
6	Average employment period	Age	Dp (12)	Dp (60)	Kty
7	Employment rate (12)	Employment rate (12)	Q3_4	Sygedp (60)	Q6
8	Average income parents	Average income parents	Q10_2	Average employment period	Kth
9	Age first child	Age first child	Average income parents	Origin-Non-Western descendant	Q3_6
10	Relocations (5)	Relocations (5)	Q1a_1	Q6	Q10_1

Table 4.1: Top ten features importance

In accordance with the findings presented by STAR and also outlined in my Exploratory Data Analysis, the most significant feature is the second question from the questionnaire (Q2). The strong visual correlation between people's self-assessment of their duration and their actual duration period was also caught by the models. Only

LR found the aggregated wage income for the past 12 months to be more telling, followed by a specific benefit type (Kty = Kontantydelse). Another interesting feature is *Age*, which is included in all columns. The correlation between age and unemployment duration was also clearly presented in Figure 3.2.

Although not agreeing on the order, almost all the models agreed on the features. The only apparent surprise is the absence of gender; in sixteenth place in GBDT and fortieth place in RF. Likewise, origin only appeared in two models. GBDT mentioned Danish origin, and LR noted the origin of non-western descendants; in comparison, RF had origin as the thirty-seventh most crucial feature. This contrasts with STAR's current model, where origin is the second most important feature. The fact that my DT varies from the current model regarding feature importance is an indication of the difference in preprocessing methods.

Regarding educational background, I only found a trace of it within LR and GBDT. Both include Q1a_1, which asks whether a person has an education within the humanities, religion, or aesthetics. A more detailed view would have been interesting.

4.2 Model Performance

In this section, I will focus on the models' performance across metrics. Table 4.2 shows the six ML models presented earlier and how they performed. None of the models were tuned at this stage; all the default parameters were used. This doesn't apply to the MLP, which didn't have a default number of layers or neurons, so it can be regarded as tuned. The highest numbers are highlighted, and the table clearly shows that GBDT and MLP outperformed the rest. As mentioned earlier, my focus at this stage was on the f1 score, AUC, and AUC-PR. As mentioned earlier, the f1 score is a metric that balances recall and precision. With that in mind, we can see that MLP, followed by GBDT and RF, generalizes well when working on unseen data.

	RF	**DT**	**GBDT**	**LR**	**SVM**	**MLP**
Accuracy	66.32%	58.58%	**66.68%**	64.62%	64.26%	66.32%
Recall	51.00%	51.99%	51.15%	49.65%	49.15%	**57.24%**
Precision	62.75%	51.24%	**63.36%**	60.13%	59.63%	61.05%
F1	56.27%	51.61%	56.60%	54.39%	53.88%	**59.09%**

Table 4.2: Performance across metrics - Highest numbers are highlighted.

A visual representation of the performance can be seen in Figure 4.1. Regardless of how one measures the performance, one would conclude that MLP, GBDT, and RF outperformed the rest. This goes hand in hand with the presented trade-off between the performance of a model and its transparency.

Furthermore, in Figure 4.2 (a) I show the AUC score, which confirms my impression of the performance from the f1 score. Looking at the two graphs, it is evident that RF and MLP perform similarly. Noticeably, DT has a remarkably low AUC, almost in line with the green dotted line. The DT's overall performance is just barely better than tossing a coin in the air. From Figure 4.2 (b), we can see that sacrificing recall for precision wouldn't make a difference for the DT. However, reducing recall can significantly increase the precision for the rest of the models.

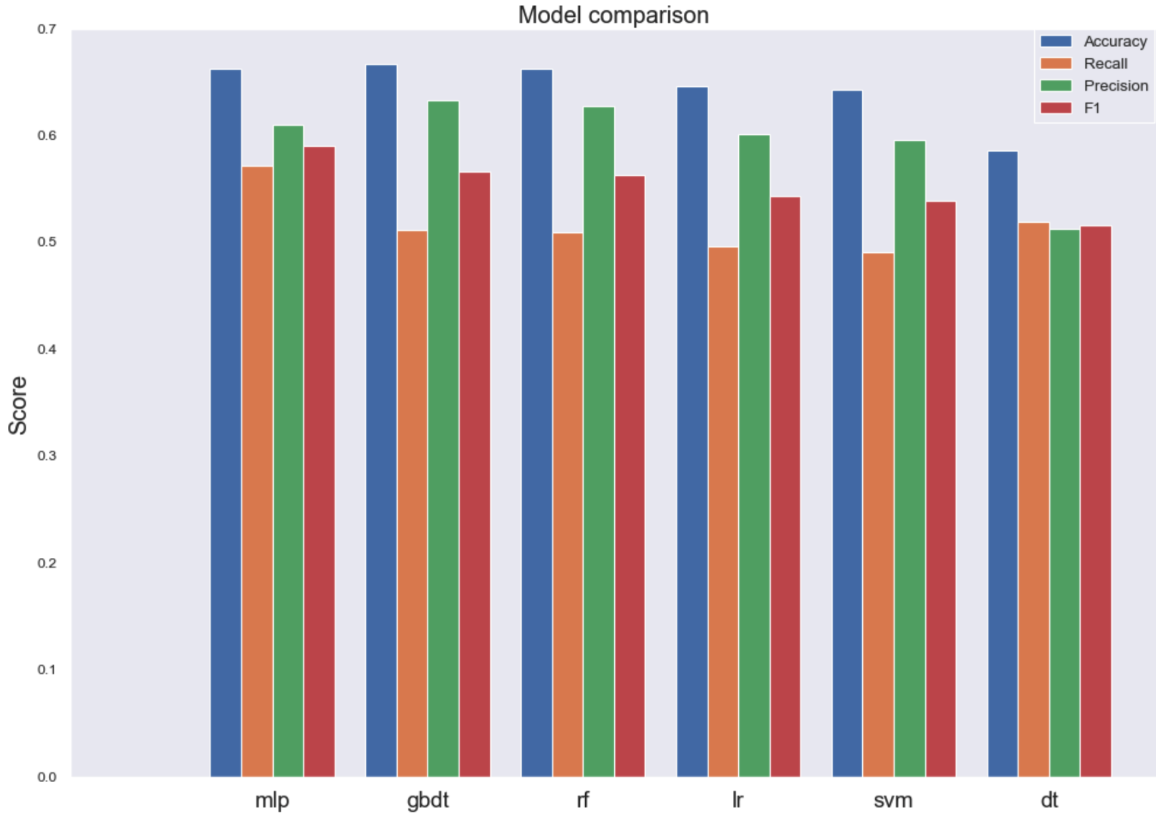
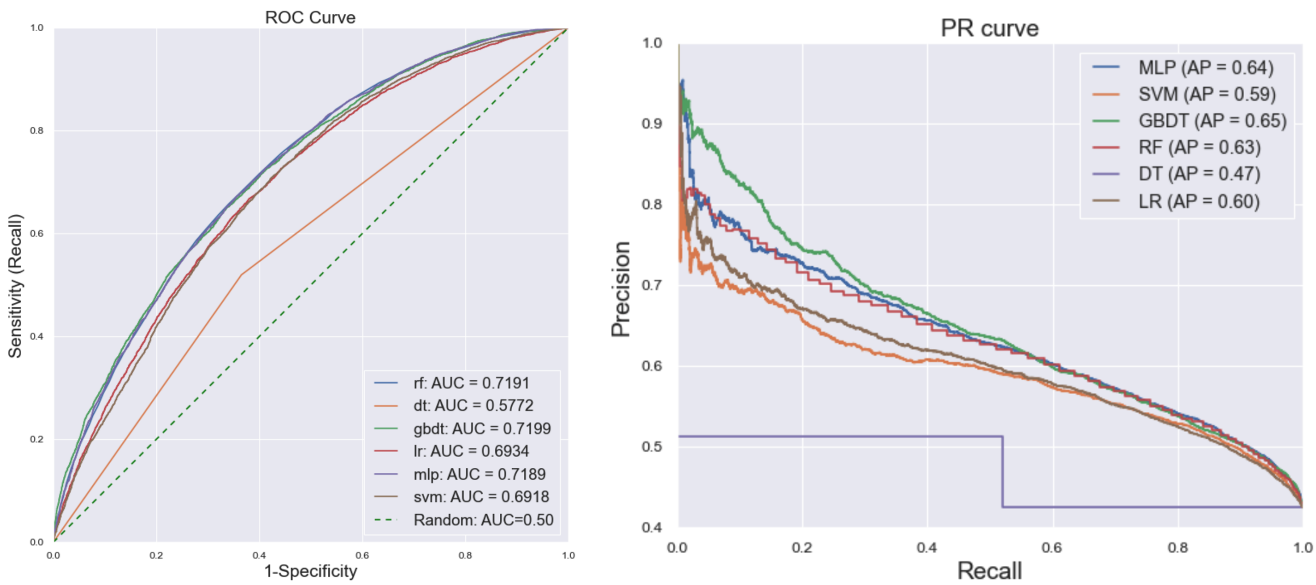


Figure 4.1: Performance across metrics



(a) AUC

(b) AUC-PR

Figure 4.2: The ROC and PR curves

After analyzing the initial performance, I took the top two performing models and tuned their hyperparameters. The two models I found most appropriate to carry on with were GBDT and RF. The decision was based on an overall assessment of the models' performance. Although MLP had a higher f1 score than RF, looking at the AUC and AUC-PR, they seemed identical. Given MLP's total lack of transparency, I found it more appropriate to move forward with the RF. The decision was supported by STAR, who expressed particular interest in RF.

I refer to the tuned models as GBDT- and RF-Tuned. Having tuned my models, I started experimenting with different thresholds to show the change in the trade-off. Moving from the base 50% threshold, I assigned a threshold of 60% to my tuned RF model and called it RF-Adj (*Adjusted*); I did the same for GBDT. The reason for picking 60% stems from the fact that the current model has a precision of 70%; I therefore tried to align my precision with the current model. The new and old results are displayed in Table 4.3.

Not surprisingly, the GBDT-Tuned had the most highlighted numbers. By sacrificing the recall to raise the precision, the f1 score naturally fell for the adjusted models. Although the precision was almost identical for the adjusted models, the underlying algorithms differed entirely. Combined with the fact that RF only ranked origin as the thirty-seventh most telling feature, I found it most optimal to take the RF model to the last stage. Should there be interest in moving on with GBDT, I would recommend going through my code in Appendix C, where RF and GBDT are analyzed.

	RF-Adj	**GBDT-Adj**	**GBDT-Tuned**	**RF-Tuned**	**GBDT**	**RF**
Accuracy	64.45%	65.92%	**66.89%**	66.58%	66.68%	66.32%
Recall	28.45%	35.24%	**53.56%**	52.29%	51.15%	51.00%
Precision	**70.13%**	69.51%	62.98%	62.82%	63.36%	62.75%
F1	40.48%	46.77%	**57.89%**	57.07%	56.60%	56.27%

Table 4.3: Performance across metrics after tuning the models and adjusting the threshold - Highest numbers are highlighted.

Moving on to the last stage with the RF-Adj model (referred to as *Baseline* in the table), I trained the model without the potential discriminatory features. To be more specific, I trained four RF models, one without origin (*RF-Origin*), one without age (*RF-Age*), one without gender (*RF-Gender*), and lastly, one without origin, age, or gender (*RF-All*). No effect on the RF-Adj model could be noticed when removing the origin or the gender (see Table 4.4). However, the trade-off between discrimination and precision became somewhat more apparent when I removed age or all the features simultaneously (a reduction of 2% and 2.2% respectively).

	Baseline	**RF-Origin**	**RF-Age**	**RF-Gender**	**RF-All**
Accuracy	**64.45%**	64.11%	63.77%	64.25%	63.71%
Recall	**28.45%**	27.09%	27.00%	27.79%	26.91%
Precision	**70.13%**	70.10%	68.74%	69.94%	68.56%
F1	**40.48%**	39.08%	38.77%	39.77%	38.65%

Table 4.4: Performance across metrics after adjusting for potential discriminatory features - Highest numbers are highlighted.

The current DT model has a precision of 70 and a recall of 35%. These numbers were only attained by RF and GBDT in my case, which indicates a difference in how the data were processed and might also be caused by a different hyperparameter tuning. Given that origin plays a major role in the current model, I would assume that removing the origin might severely damage the precision. However, removing origin from my model did not have a considerable impact (nor did removing age and gender). The difference in data processing might have allowed my models to find patterns in other features which work as a proxy for the origin.

4.3 Model Interpretability

I introduced eXplainable AI (XAI) in the background chapter and mentioned specific Model-Agnostic techniques that shed some light on black-box models. One of the mentioned techniques was LIME, which I used on the RF model to understand why it classifies specific individuals as at risk of LTU or not.

The first example, shown in Figure 4.3, is a case where the model classifies the individual as not being at risk of LTU with a probability of 91%. The specific combination of feature values that classified the person as not being at risk is displayed below. The figure shows the correlation between the feature value and the target. A red horizontal bar represents the strength of the correlation with the target value '0', while green bars represent the correlation with the target value '1'. In this case, the unemployed person answered Q2 with "*I have a new job, but have not started yet*". According to LIME, my RF model correlated the first three possible answers to Q2 with not being at risk of LTU. Although not displayed because of limited space, the data showed that her parents' average income was 454,311, which is relatively high. The overall self-assessment from the questionnaire combined with a young age is the reason for classifying the individual as not being at risk of LTU. It might be worth noticing that gender and origin did not play any role.

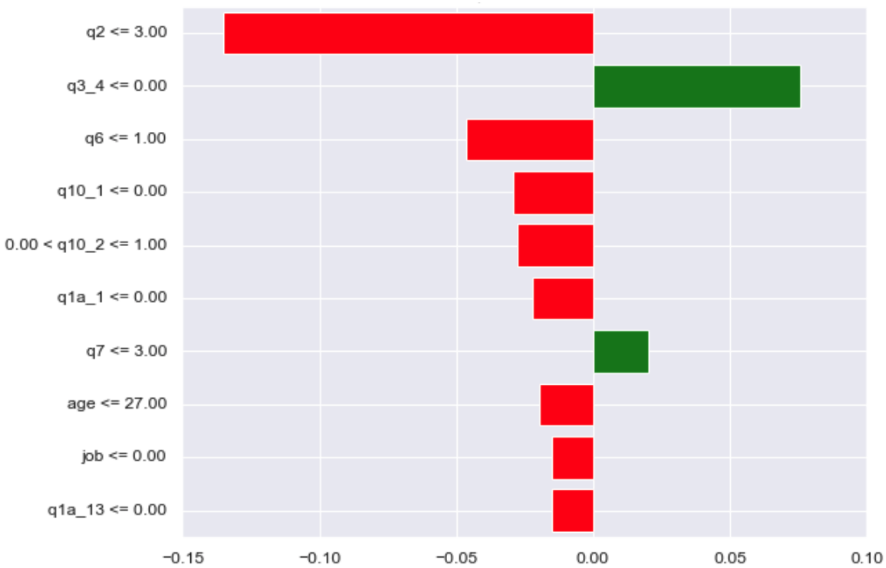


Figure 4.3: Person not at risk of LTU

Moving on to the next case, we now have an individual classified as being at risk of LTU with a probability of 70%. In this case, we have an elderly individual at 60, and the model correlates all aged above 48 as being at risk of LTU. The unemployed individual answered Q2 with the expectation of being unemployed for more than six months, and the model agreed with the self-assessment based on the feature values displayed below.

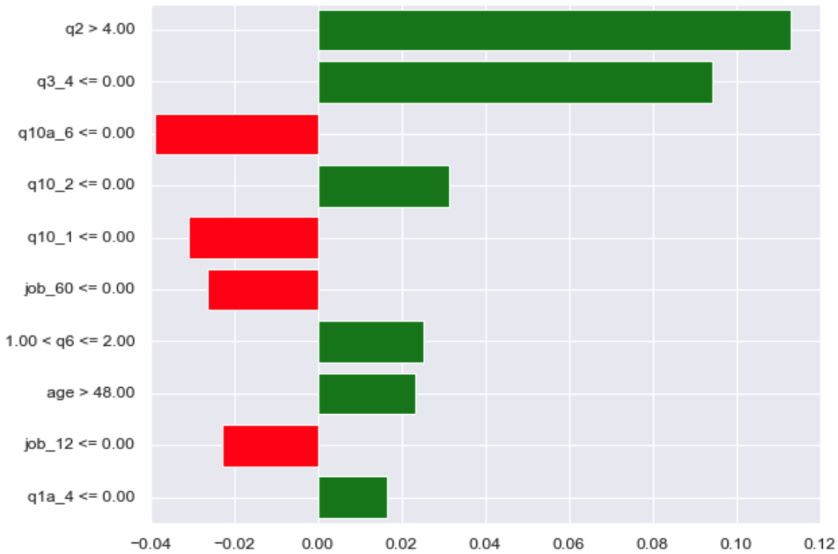


Figure 4.4: Person at risk of LTU

Understanding the logic behind the ML models is relevant from a GDPR aspect and should also be seen as a natural step to make caseworkers feel a sense of ownership of the tool. Enabling job counselors to explain the logic behind specific categorizations is essential and can help STAR further improve the model. Additionally, I have managed to build and present a model where removing specific potential discriminatory features doesn't damage the performance. It is therefore my recommendation to further improve this model in order to replace the current model with it. A discussion regarding how to improve the model will be presented in the next chapter.

CHAPTER 5

Discussion

The aim of this chapter is to present a critical review of the methods used and the results. Afterward, I write about my contribution and the research limitations of this study before concluding with advice on further research areas.

5.1 Evaluation of the models

In order to answer my second research question regarding the social impact of job seeker profiling tools, I presented the results of a multiple linear regression. The analysis showed the excess number of conversations job seekers had with job counselors based on origin, gender, and graduation. The data should be read with caution, given the low adjusted R-square and the possibility of omitted variables bias. The low adjusted R-square may be due to the existence of non-linear data. Another reason is the lack of data when dividing people based on origin. As an example, focusing on Frederiksberg Municipality, I only had 115 individuals of non-Western descent. Reducing this group further to graduates, the number fell to 64, and to only 31 when adding gender. The lack of data limited my ability to eliminate apparent differences between the groups.

Nonetheless, the adjusted R-square regards the correlation between the features and the target and not the model's regressional performance (Dunn, 2021). Focusing on the number of conversations and not the correlation, I concentrated instead on the confidence intervals, which show the range of the excess number of conversations with a 95% probability. Although there is room for improvement, I feel confident in my results, which suggest that more conversations are being held with non-Western immigrants' descendants than with citizens with Danish origins.

Moving on to the methodology chapter and the implementation of Logistic Regression, I noted that it wasn't one of the best performing models. This could simply be due to the harsh competition. Another explanation is not fulfilling the models' assumptions. The LR model requires the observations to be independent of each other. Not fulfilling this assumption might be the reason for not obtaining a higher performance.

Looking at SVM, I chose the model based on its ability to handle data in higher dimensions. Unfortunately, the model is slow when working with large datasets, which motivated me to utilize an SVM model that optimizes through Stochastic Gradient Descent. This reduced the computational time but also limited my choice of kernel to linear. In other words, the SVM presented only has a linear decision boundary. I experimented with another kernel (RBF), which did not improve the performance, and I would have tried several kernels if I had had more time.

Hyperparameter tuning my GBDT model took around 8-10 hours of runtime while tuning RF took more than 34 hours. Other tuning methods (such as RandomizedSearchCV) were tried out but at the cost of precision. Having more time would have allowed me to hyperparameter tune all models at the first stage. I nonetheless feel confident in my results and recommendations.

The performance of any model is constrained by the quality of the underlying data. The data preprocessing step was the most time-consuming step, and many decisions made along the way can severely impact the model performance. Data preparation is a work of art and can vary from person to person. While I might find it optimal to fill in missing values with the most frequent value, someone else might use the average or create a separate category. Unfortunately, I did not have insight into how the current ML model is tuned or how the data has been processed. On the other hand, this might have inspired me to find my own path instead of walking down the track already made by others.

5.2 Contributions

From a consulting perspective, the thesis' primary goal was to deliver my final product to STAR. The product, in this sense, is my recommendation on whether to proceed with the current model or change it. The main motivation for STAR to consider changing the model is the complaints from several people and the criticism by the Department of Human Rights concerning the use of origin in the analysis. Although I also analyzed other potential discriminatory features, the main focus was the use of origin. Since the current model relies heavily on origin, I was very pleased that I built a model where origin doesn't play any significant role while still keeping the precision.

While STAR hasn't experienced any complaints regarding the use of gender, I firmly believe that gender will be the next issue to deal with. I might add that if people knew the importance of age in the final categorization, they would probably stand in line to complain. Given that my model only loses a small percentage in the removal of both features, I therefore see it as a long-term solution. Therefore, my final recommendation is to take steps to move from the current model to an RF model. How to further improve it will be outlined in the last section.

5.3 Limitations

As hinted at in the methodology chapter, my analysis was limited by my work environment. The package manager and the IDE did not grant me access to all the well-known ML models. Some of the missed models are the Mixed Naive Bayes algorithm, XGBoost, and K-nearest neighbors. Another limitation was the omission of 2020 data; the worldwide pandemic pushed people into unemployment at a rapid rate and so while the data from 2020 would have been interesting to study, it could have skewed my models' performance since the pandemic was both an extreme and unique situation.

I did manage to get insight into other countries and their features, models, and performance, but it would have been desirable to gain further insight to analyze the performance across metrics. Showing how other countries perform and which features to consider is crucial to the improvement of the model in the long run. Information sharing across borders can save time and money in the further development of the model, especially if additional features will be removed due to perceived discrimination.

5.4 Future outlook

The social impact of job seeker profiling tools can be further investigated by gathering more data. Doing this would allow us to analyze the perceived discrimination with more accuracy. Having more data would allow us to compare people with the same social resources, language skills, geographic location, work experience, type of education, age, gender, grades, etc. Although more data is desirable, it might not always be the answer. If a linear model is not equipped for the challenge, I would recommend using other ML models, where the number of conversations is the target variable. Approaching it as a regression problem will allow us to implement models that can handle non-linearity and high dimensionality.

Concerning the further development of the RF tool, I would focus on quality over quantity. More of the same data might not make a significant difference, but acquiring data with regards to language skills, previous unemployment duration, health, criminal records, types of positions a job seeker is looking for, unemployment rates in the municipalities, and data regarding abuse will undoubtedly make a positive difference. There are some challenges with regards to gathering some of this information, and doing so might even be illegal. Still, seeing how other countries are benefiting from this, it is definitely worth investigating.

Another factor to (re-)consider is the question regarding transparency. Other countries have been hesitant in sharing the results of the profiling tool, saying that it is demotivating and stigmatizing while also fearing that the result might become a self-fulfilling prophecy. Being transparent has a cost, and if the cost out-weighs the benefits, it should be reevaluated. Having said that, it is worth noting that we might be setting an example for other countries by insisting on being transparent, and hopefully, it will pay off in the long run.

Based on my interview with Jane, I would recommend investigating the desire for and the possibility of giving the caseworker more control over the screening tool by allowing re-categorization. Adjusting the input data at the second conversation would enable the caseworker to improve the data quality and give them a sense of influence over the end result. Having categorized the job seeker, it would then be desirable to know how to move forward. My interview with Mie Hansen, consultant in the Employment and Social Services Administration in Odense Municipality, revealed that they are currently developing an AI tool that can deliver tailor-made recommendations on how to help a job seeker (Appendix B). Although still in its early stages, it is meant to give recommendations regarding whether a person should pursue additional education or apply for a company internship, or suggest other ways of improving qualifications. Combining this AI tool with the RF model presented in this study would give caseworkers a clear and straightforward way to help job seekers back into work.

CHAPTER 6

Conclusion

The thesis' goal was to contribute to the field of data science by showing how methods from machine learning can be used to categorize job seekers in terms of LTU while also considering the social impact. My research questions were answered from the bottom up, starting by introducing the unemployment process followed by an international comparison. Although all countries had the same intention of preventing long-term unemployment, I illustrated through tables and figures how they differ. While several countries have chosen not to be transparent regarding job seekers' classification, pointing at demotivation, stigmatization, and the fear of self-fulfilling prophecies as the reason for their lack of transparency, we see that Denmark has chosen a separate trajectory by being fully transparent. Besides being the only country to use a Decision Tree model, Denmark is also the only one to make answering the questionnaire and use of the results voluntary. Finally, there is a discussion regarding the features of the screening tools, where Denmark lacks some essential features such as health and language skills, which are seen as essential by other countries.

The preceding section was dedicated to the analysis of the social impact of job seeker profiling tools. I presented two types of trade-offs, the first being the trade-off between transparency and performance. The trade-off between the two becomes even more crucial when looking at the GDPR rules, which require adequate information about the logic involved if the process is automated. Another trade-off was between accuracy and discrimination, a struggle which was even harder to deal with given the discrimination-by-proxy phenomenon, which refers to the possible indirect discrimination from features correlating with discriminatory features. In this context, a figure was presented showing the excess number of conversations with job counselors that non-Western immigrants' descendants deal with based on origin, which suggests that more conversations are being held with descendants than with citizens of Danish origin. Besides differentiating between negative and positive discrimination, I presented how researchers try to deal with discrimination by either modifying the algorithm or the data.

The last question to be answered concerned the removal of potential discriminatory features while maintaining high precision. By presenting and implementing a range of different ML models, I showed how to build an ML model where origin is not a significant feature (a decrease in precision from 70.13% to 70.10%). The removal of other features (such as age and gender) was also displayed and the precise impact was presented in tabular form. Through the use of LIME, I also introduced the logic behind the Random Forest model. Based on the results, my recommendation to STAR is to undertake steps to switch from the current model to a Random Forest model.

Bibliography

A-kasser. (2020). Unemployment benefits. *A-kasser*. https://www.a-kasser.dk/benefits.html

Angwin, J., Larson, J., Mattu, S. & Kirchner, L. (2016). Machine bias. *Propublica*. https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing

Barredo Arrieta, A., Díaz-Rodríguez, N., Del Ser, J., Bennetot, A., Tabik, S., Barbado, A., Garcia, S., Gil-Lopez, S., Molina, D., Benjamins, R., Chatila, R. & Herrera, F. (2020). Explainable artificial intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. *Information Fusion*, vol. 58, 82–115.

Bostrup, J. (2021). Tellis oplevelse på jobcenteret fører nu til en optrapning i kampen mod diskriminerende algoritmer. https://politiken.dk/viden/Tech/art8140892/Tellis-oplevelse-p%C3%A5-jobcenteret-f%C3%B8rer-nu-til-en-optrapning-i-kampen-mod-diskriminerende-algoritmer

Burkov, A. (2019). *The hundred-page machine learning book*.

Dandanell, F. (2019). KU-lektor vil opgøre statistik i farver i stedet for køn, da det kan virke stødende at bruge begreberne mand og kvinde. *Berlingske*. https://www.berlingske.dk/samfund/ku-lektor-vil-opgoere-statistik-i-farver-i-stedet-for-koen-da-det-kan

Desiere, S., Langenbucher, K. & Struyven, L. (2019, February 20). *Statistical profiling in public employment services: An international comparison*. OECD. https://pdfs.semanticscholar.org/1081/e2f90fef96a02593171fb5f45152326dc17f.pdf

Desiere, S. & Struyven, L. (2020). Using artificial intelligence to classify jobseekers: The accuracy-equity trade-off. *Journal of Social Policy*, 1–19.

Digitaliseringsstyrelsen. (2020a). 13 nye signaturprojekter skal bruge kunstig intelligens i klimakampen og til bedre behandlinger på hospitalerne. *Digst*. https://digst.dk/nyheder/nyhedsarkiv/2020/september/13-nye-projekter-skal-bruge-kunstig-intelligens-i-klimakampen-og-til-bedre-behandling-paa-hospitalerne/

Digitaliseringsstyrelsen. (2020b). Kunstig intelligens. https://digst.dk/strategier/kunstig-intelligens/

Dunn, K. (2021). Avoid r-squared to judge regression model performance. *Towardsdatascience*. https://towardsdatascience.com/avoid-r-squared-to-judge-regression-model-performance-5c2bc53c8e2e

Eskelinen, L., Petersen, J. & Bolvig, I. (2015, December 21). *Profilering af ledige*. KORA. https://star.dk/media/1201/profilering-af-ledige-kora.pdf

Farbøl, R., Sørensen, A. & Olesen, T. (2018). Oliekriserne og deres betydning for dansk økonomi, 1973-1991. *Danmarkshistorien*. https://danmarkshistorien.dk/leksikon-og-kilder/vis/materiale/oliekriserne-og-deres-betydning-for-dansk-oekonomi-1973-1991/

GDPR. (2018a). Article 13 EU GDPR - information to be provided where personal data are collected from the data subject. *Privacy Regulation*. https://www.privacy-regulation.eu/en/article-13-information-to-be-provided-where-personal-data-are-collected-from-the-data-subject-GDPR.htm

GDPR. (2018b). Recital 71 EU GDPR. *Privacy Regulation*. https://www.privacy-regulation.eu/en/r71.htm

Ghanoum, T. (2020, November 30). *Analyzing amazon luxury & beauty product reviews* (Doctoral dissertation). CBS.

Jobes, R. (2018). The siren song of objectivity: Risk assessment tools and racial disparity. *Medium*. https://nacdl.medium.com/from-the-president-the-siren-song-of-objectivity-risk-assessment-tools-and-racial-disparity-fa5ccb0698a5

Jobindsats. (2020). Ledigheds- og jobindikatorer. shorturl.at/irtHN

Kamiran, F. & Calders, T. (2009). Classifying without discriminating. *2009 2nd International Conference on Computer, Control and Communication*, 1–6.

Kamishima, T., Akaho, S., Asoh, H. & Sakuma, J. (2012). Fairness-aware classifier with prejudice remover regularizer [Series Title: Lecture Notes in Computer Science]. In P. A. Flach, T. De Bie & N. Cristianini (Eds.). D. Hutchison, T. Kanade, J. Kittler, J. M. Kleinberg, F. Mattern, J. C. Mitchell, M. Naor, O. Nierstrasz, C. Pandu Rangan, B. Steffen, M. Sudan, D. Terzopoulos, D. Tygar, M. Y. Vardi & G. Weikum (**typeredactors**), *Machine learning and knowledge discovery in databases* (pp. 35–50). Springer Berlin Heidelberg.

Kay, K. (2019). Digging their heels in: Women wage war on footwear dress codes. *The Guardian*. https://www.theguardian.com/fashion/2019/jun/08/why-women-in-japan-are-fighting-high-heel-dress-codes-at-work-global-support

Kutner, M. H. (Ed.). (2005). *Applied linear statistical models* (5th ed). McGraw-Hill Irwin.

Lund, K. & Pedersen, P. (2016). Danske arbejdsgivere vælger 'mads' frem for 'muhammed'. *Politiken*. https://politiken.dk/debat/art5731006/Danske-arbejdsgivere-v%C3%A6lger-%E2%80%99Mads%E2%80%99-frem-for-%E2%80%99Muhammed%E2%80%99

Marquardt, F. & Carl, O. (2020). For første gang har forskning kortlagt, at kvinder med tørklæde vælges fra på arbejdsmarkedet – afnan var en af dem. *Berlingske*. https://www.berlingske.dk/samfund/for-foerste-gang-har-forskning-kortlagt-at-kvinder-med-toerklaede-vaelges?fbclid=IwAR0BnPWSCqFeMxWxogAQS7ftLAIiK4ZypeigJzXrCIxtUUuyLEOQfvhNUW4

Menneskeret. (2020). Institut for menneskerettigheder indbringer styrelsen for arbejdsmarked og rekruttering for ligebehandlingsnævnet. *Menneskeret*. https://menneskeret.dk/nyheder/institut-menneskerettigheder-indbringer-styrelsen-arbejdsmarked-rekruttering

OECD. (2018, December 1). *Profiling tools for early identification of jobseekers who need extra support*. OECD.

Olsen, M. (2021a). Nu er det nok: Anonyme jobansøgninger skal forhindre diskrimination. *Politiken*. https://politiken.dk/oekonomi/arbejdsmarked/art8137824/Anonyme-jobans%C3%B8gninger-skal-forhindre-diskrimination

Olsen, M. (2021b). et kæmpe skridt i den rigtige retning«: Oplysninger om alder og køn fjernet fra statens digitale ansættelsessystem. *Politiken*. https://politiken.dk/indland/art8145583/Oplysninger-om-alder-og-k%C3%B8n-fjernet-fra-statens-digitale-ans%C3%A6ttelsessystem

Petersen, J. & Holm, S.-K. (2020). *Retlige problemstillinger i forbindelse med anvendelse af kunstig intelligens i den offentlige forvaltning* (Doctoral dissertation). https://law.au.dk/fileadmin/Jura/dokumenter/forskning/rettid/Afh_2020/afh15-2020.pdf

Pope, D. G. & Sydnor, J. R. (2009). Implementing anti-discrimination policies in statistical profiling models, 28.

Provost, F. & Fawcett, T. (2013). *Data science for business*.

Raschka, S. & Mirjalili, V. (2019). *Python machine learning: Machine learning and deep learning with python, scikit-learn, and TensorFlow 2* [OCLC: 1155055478]. Retrieved March 25, 2021, from http://proquest.safaribooksonline.com/?fpi=9781789955750

Renteln, A. (2008). *Cultural defense*. Encyclopedia of Violence, Peace, & Conflict (Second Edition). https://www.sciencedirect.com/topics/economics-econometrics-and-finance/labour-market-discrimination

Retsinformation. https://www.retsinformation.dk/eli/accn/W20200993225

Ribeiro, M. T., Singh, S. & Guestrin, C. (2016). "why should i trust you?": Explaining the predictions of any classifier. *arXiv:1602.04938 [cs, stat]*. Retrieved March 6, 2021, from http://arxiv.org/abs/1602.04938

Royston, P. & White, I. (2011). Multiple imputation by chained equations (MICE): Implementation in *Stata*. *Journal of Statistical Software*, vol. 45no. 4.

Sakkaf, Y. (2020). Decision trees: ID3 algorithm explained. *Towardsdatascience*. https://towardsdatascience.com/decision-trees-for-classification-id3-algorithm-explained-89df76e72df1

STAR. (n.d.-a). Arbejdsmarkedsbalancen. *STAR*. https://star.dk/viden-og-tal/udvikling-paa-arbejdsmarkedet/arbejdsmarkedsbalancen/

STAR. (n.d.-b). Det lokale jobbarometer. *STAR*. https://star.dk/viden-og-tal/udvikling-paa-arbejdsmarkedet/det-lokale-jobbarometer/

STAR. (2014). The employment reform. *STAR*. https://www.star.dk/en/recent-labour-market-policy-reforms/the-employment-reform-2014/

STAR. (2020). Profilafklaring – det digitale forberedelsesskema på jobnet. *STAR*. https://star.dk/it/borger-it/profilafklaring-forberedelsesskema/

Vellido, A. (2019). Societal issues concerning the application of artificial intelligence in medicine. *Kidney Diseases*, vol. 5no. 1, 11–17.

VIVE. (2007). *Når politik bliver til virkelighed*. Copenhagen. https://www.vive.dk/media/pure/3866/334252

Appendices

APPENDIX A

Figures and Tables

A.1 Jobnet.dk

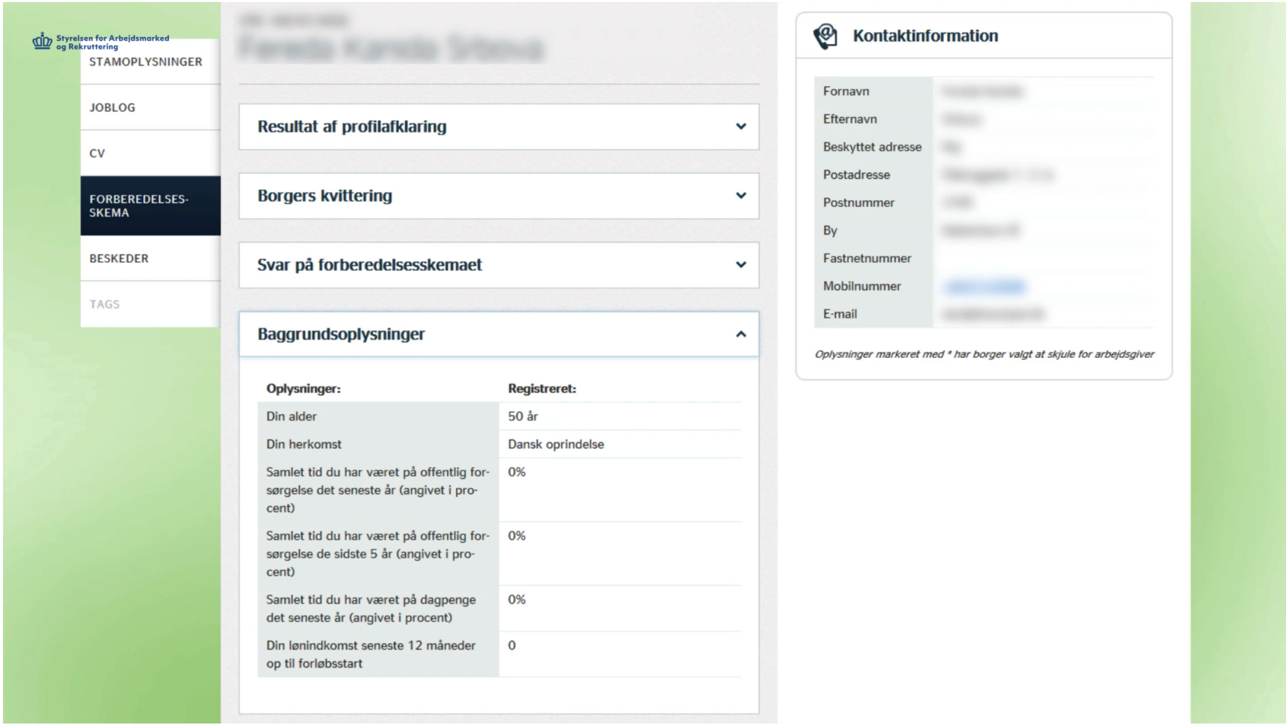


Figure A.1: Result section in jobnet.dk

A.2 jobindsats.dk

Figure A.2: Share in employment 3, 6, 9 and 12 months after unemployment (Jobindsats, 2020)

A.3 Explainable AI - XAI

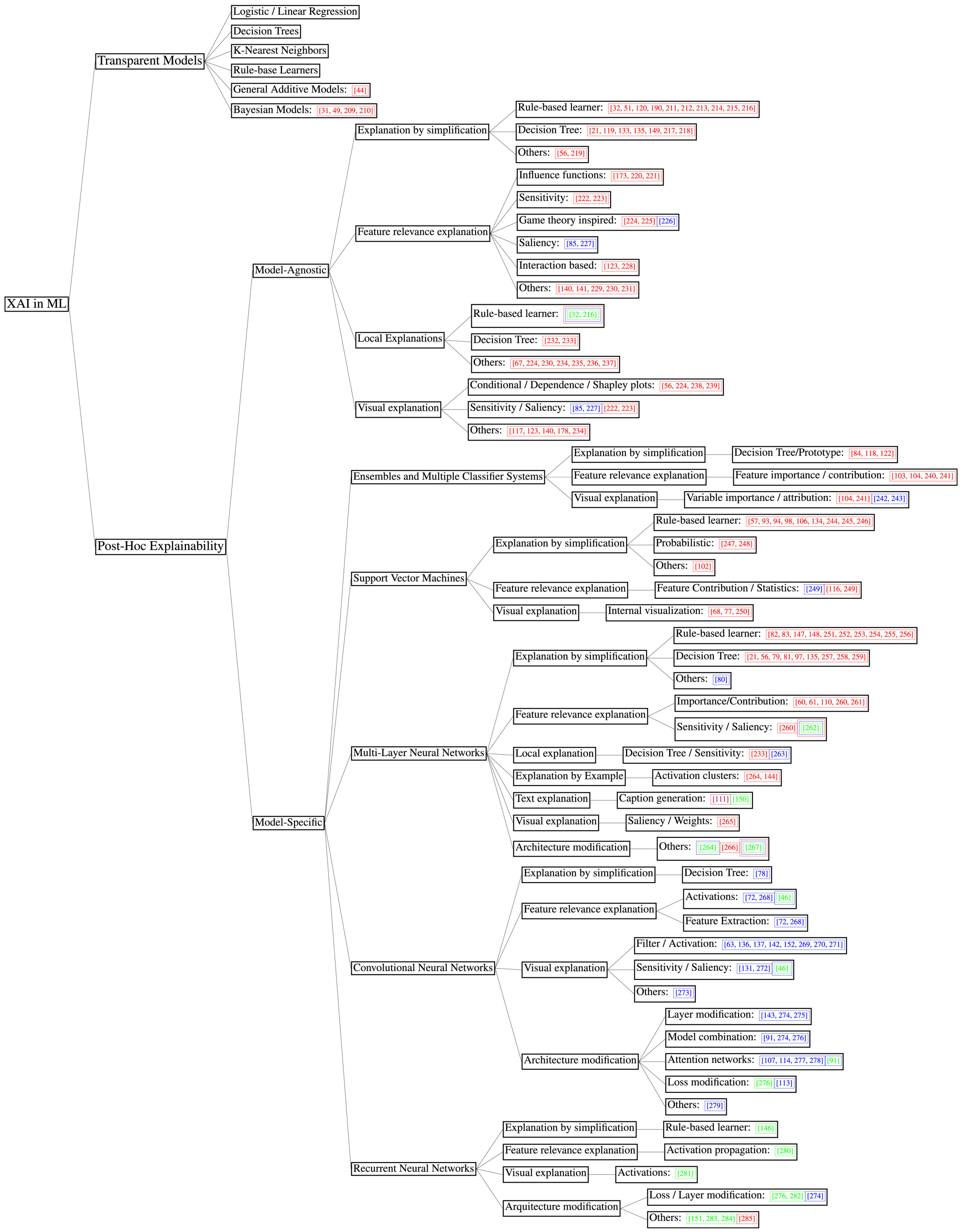


Figure A.3: Explainable AI (Barredo Arrieta et al., 2020)

A.4 Discrimination - Origin

[redacted]

5. december 2019 ·

OBS

Kære du, der er født i Danmark af forældre fra såkaldt "ikke-vestlige" lande og pt jobsøgende.

Jobnet (under Styrelsen for Arbejdsmarked og Rekruttering (STAR) har profileret DIG som "efterkommer af ikke-vestlig herkomst", og dermed er den indsats de retter mod dig anderledes, end den indsats de retter mod Pia og Morten.

Det er det, der i god gammeldagsforstand kaldes diskrimination på baggrund af etnicitet, men som Kammeradvokaten har sagt god for, fordi han åbenbart vurderede at "herkomst" og "etnicitet" er to vidt forskellige ting.

Det skulle så åbenbart også være "frivilligt", så hvis du er i denne situation, kontakt dit jobcenter og frabed dig at blive diskrimineret.

Ja, vi er i 2019 for de, der pludselig blev i tvivl.

Er det relevant for dig, så gå ind på jobnet --> Min Profil --> Forberedelsesskema --> Dine baggrundsoplysninger.

En anden hyggelig ting er, at er du kvinde med såkaldt "ikke-vestlig herkomst" så er du pt også en del af undersøgelsen 'Kortlægning af social kontrol som mulig barriere for etniske minoritetskvinders beskæftigelse', fordi du står registreret som så inde på jobnet.

Det vil blandt andet sige, at din sagsbehandler højst sandsynligt har fået eller vil få et spørgeskema tilsendt med spørgsmål om hvorvidt du ikke er i job fordi din mand eller andre familiemedlemmer ikke ønsker, at du skal arbejde med andre mænd (!).

Yes. Jeg har spørgeskemaet, hvis nogen er interesserede.

Undersøgelsen er besluttet og udmøntet gennem Beskæftigelses- og Integrationsudvalgets Integrationshandleplan
2019-22, under indsatsen til forebyggelse af social kontrol. (Se side 19 i udvalgets Integrationshandleplan 2019-22)

Til de der stadig er i tvivl, så er det her udtryk for strukturel racisme.

Det er også udtryk for, at kategoriseringen mellem "vestlig og ikke-vestlig herkomst" ikke har at gøre med om en person taler dansk og kan begive sig i det danske samfund, men om en racialisering af den synlige "Anden" som essentiel ikke-dansk!

Relevante artikler er linket i kommentarsporet.

[redacted] 355 andre 164 kommentarer 140 delinger

Figure A.4: Shows dissatisfaction over the use of origin in the danish screening tool

A.5 Decision Tree

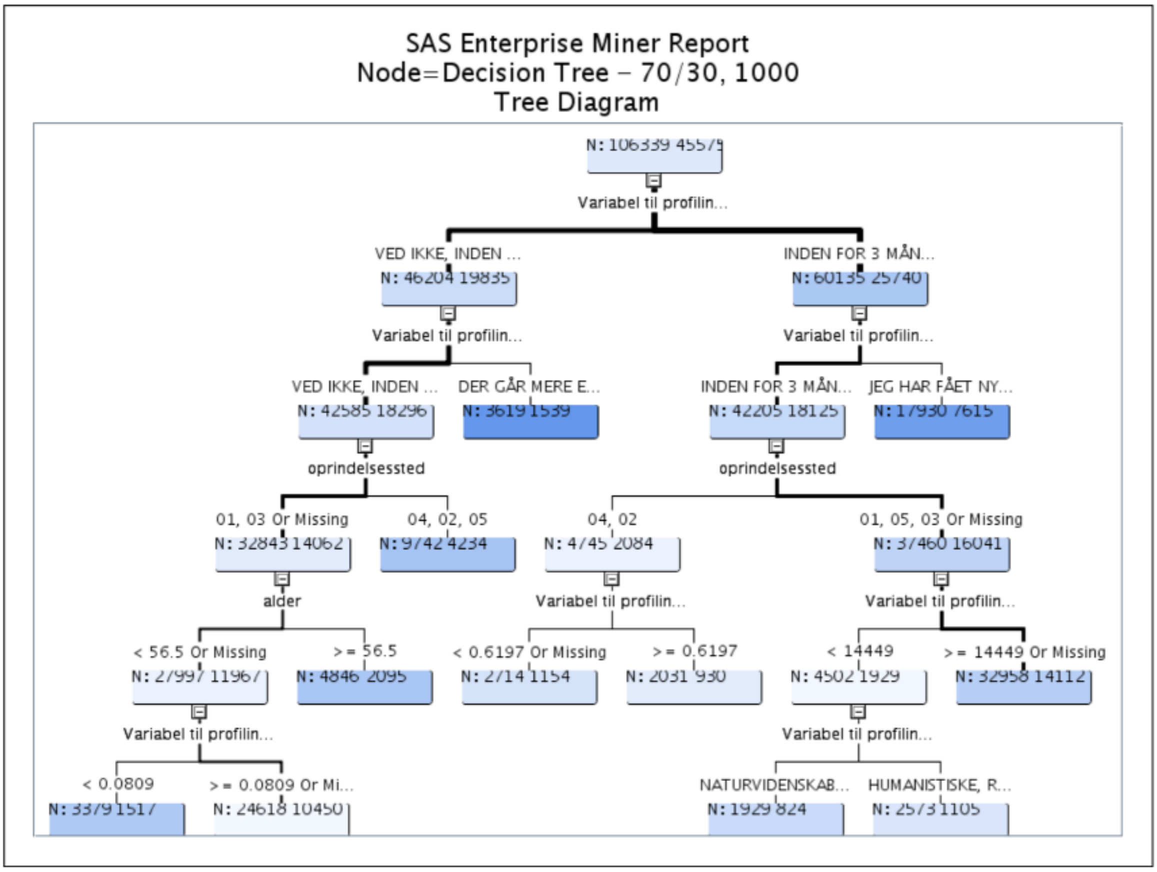


Figure A.5: The illustrations shows the current Decision Tree implemented by STAR

Feature	**Definition**
ID	Pseudo random numbers unique to every individual
Start	First day of receiving public benefit
End	Last day of receiving public benefit. Unfinished unemployment processes are labeled 31dec9999
Duration	Duration of the unemployment processes in weeks
Target	Target variable, 1 if the unemployment process is 26 weeks or above, else zero
Gender	Male or Female
Birthday	Day of birth
Punlic_support_type	Type of public support after the support period
sd_all_12, sd_all_60, per_12, per_60, dp_12,...	social benefit type
Municipality	The municipality that an individual lives in
Origin	Danish, western immigrant, non-western immigrant, western descendant, and non-western descendant
Unemployment_fund_duration	Number of days a job seeker has been member of an unemployment insurance fund
Unemployment_fund_duration	The unemployment fund type: Not insured, LO, FTF, AC, or others
Age	16-66
Graduate	Whether a person is a graduate or not
Unemployment_fund	Specific type of unemployment insurance fund
Industry	The last sector a person has worked within the past 12 months
Employment_rate_12	Employment rate for the past 12 months
Employment_rate_36	Employment rate for the past 36 months
Aggregate_wage	The aggregated wage for the past 12 months. Missing info is labeled zero
Aggregate_wage_missing_values	The aggregated wage for the past 12 months. Missing info is not labeled
Parents_marital_status	Parent marital status is categorized in six categories
Provider	A person is a provider if he/she has a person under the age of 18
Provider_count	Number of people to provide for
Marital_status	A persons marital status is divided into four categories
Partner_occupational_status	Whether partner is in employment, self-employed (self-supporting) or on a benefit
Age_first_child	Age when receiving first child
Negative_lifeevents	Whether a person has experienced death among relatives in the past five years
Relocations_5	Number of relocation for the past five years
Average_income_parents	Parents average income
Average_employment_period	The average employment period for work in the past twenty-four months

Table A.1: Overview of features and their definition - questionnaire not included

APPENDIX B

Interviews

B.1 Mikael Dehn Kristensen

Interview with Mikael Dehn Kristensen (MDK) – by Tarek Ghanoum (TG)

TG: Thank you very much again for taking your time. This is very new to me, some of the questions are basic, and from there I will try to take it to a higher level. Kindly begin by explaining your role in the Jobcenter and your experience?

MDK: I would like to do that. Briefly, I am a musician by training, when I was young. I have worked as a musician for 20 years. To make a long story short I end up in SAS3, which have the same function as the Jobcenter however they are a private company. I became an employment consultant. I received several training programs which made it possible for me to work here today. On a daily basis I work with FJS. FJS is an abbreviation for Fremskudt Jobservice (Advanced Jobservice). If you are going to talk to someone else from here you should know that all departments here have a three letter abbreviation. No one knows why, but it is like that. If you meet someone in the hall, they say 'I am from PMD', okay, yes, 'and what do you do?' It is like that. FJS stands for Fremskudt Jobservice (Advanced Jobservice), I am quite privileged. I love my job because I am fortunate enough to deal with highly resourceful citizens. I deal with young people typically holding a BA degree below 30 years. It is called Center for youth (Ungecentret). I prepared some things. Let me see what we found out. I just got the numbers. We have 3,000 employment benefit (dagpenge) recipients in my center; around 600, 700 to 800. The numbers fluctuate constantly. You anticipate that 20 to 25 percent, and this is not corona numbers, are in the risk of becoming long-term unemployed.

TG: Okay, yes.

MDK: But we have some methods we use when we build something up for the profile clarification tool and others. We have other tools to base it on.

TG: Okay, perfect. We will talk about that.

MDK: Yes, I thought so. On a daily basis, I teach, but I am also an employment consultant and have reset (nulstillende) interviews with unemployment benefit (dagpenge) recipients. We also deal with cash assistant recipients who hold a degree in our team; I meant in our citizen group. We just got the FUI group. It stands for

Fremskudte Unge Indstas (Advanced efforts for youth), which is for young people who are imposed to take an education. They are now part of the group. They were not part of our department before. We deal with more training than we have done before. I teach on a daily basis, and we have during these times of corona developed a training programme for our graduates. This is interesting for you to know about.

TG: Yes.

MDK: I am part of the idea generation group, where we try to get them through. I also deal with the club for graduates (Dimmitendklubben). The newly graduates have a degree, and they are very qualified, the only thing that they have not learned is to apply for jobs. You do not learn this during your education, for some odd reason. You skip that, you do not prepare them. They come here with a CV and an application which is hopeless. We walk them through a real job application process. I have experience in this, so I have my own way to do it. 'Tarek if you want to apply for a job here, it is fine if you send a CV. But if it is so bad, that I do understand what you have put on your CV, then you have not reached far.' It is all about communication, on written. What is the message you want to communicate in your CV, and how to keep it concise. This is what I deal with on a daily basis. We walk them through everything from job interviews, application materials, everything that will improve their chances.

TG: Let us talk about this. Let us say that I just got my degree, and I cannot find a job. How is the process? I reckon I have to go to Jobnet and register as unemployed? This is the first step?

MDK: It is the first step. You register as unemployed in Jobnet. If you are a member of an unemployment insurance fund (A-Kasse) then of course you will eventually receive unemployment benefit (dagpenge). The first interview you attend is at the unemployment insurance fund (A-Kasse). Then comes the easier part which is the Jobnet CV, and which looks more better now. You have to fill out, and which is oddly have to be approved by the unemployment insurance fund (A-Kasse), and not by the Jobcenter. I still cannot understand why, but I am not going to interfere in this. Then you have to have this mandatory, otherwise you will not receive your unemployment benefits (dagpenge). It has to be approved. It is filled out usually very quickly. Let us be honest we can tell. And we address the issue of the Jobnet CV, and make them aware that there are people who actually use it for recruiting. As you may know, there are companies which have access in order to recruit people. The more information you provide about yourself, the higher is the probability for you to get a job. They apply here. Eventually they are invited to the first interview which is an interview with the unemployment insurance fund (A-Kasse) participation, but it takes place here. We have the role of authority (myndighedsrolle), and the unemployment insurance fund (A-Kasse) does not have that. It has been decided to give the role of authority to the unemployment insurance fund (A-Kasse) as you know. They show up to the first interview here where we have to go through some things; availability, commitments, activation, mandatory activation, all of the judicial issues and a clarification of the citizen. I ask you Tarek you have a degree, let me ask you what is your degree?

TG: I study economics.

MDK: You study economics. Good that is a good idea especially in the time of corona. There are some who have to look at this from an economic point of view later.

TG: Exactly, exactly, yes.

MDK: I usually use myself as an example, I do not care what people make of it. I do it a lot because I always begin by saying ‘ I am a classical linguist, I write books, I do miscellaneous things, I am a person which deals with language, and if you ask me anything about finances, I will tell you I am glad that I am married and I have a wife.’ I will put it like that. Numbers and I, we do not work together. Nevertheless, it does not stop me from telling you what you have to do. First of all, you need some material in order to be taken from a possible employer. If you send a CV and your message is; ‘hi Mikael’, you work at The Danish Financial Supervisory Authority (Finanstilsynet). I have a degree in economics I would like to work here and make a lot of money.’ It is fine to state that you want to make a lot of money, you want to make money, I want to make money. I can see you live somewhere, where you have to pay your rent, and I do that as well. There is nothing wrong in that. Maybe you should talk about some other qualifications. We begin with that and make a clarification. What would you like? I can see you hold a BA degree would you like pursue an MA? If you want to do that then you have to go to another center. We have an interview tool and a process we have to follow.

TG: Then what do they talk about in the unemployment insurance fund (A-Kasse)? That is the first interview. Is it exclusively to ask them to fill out the Jobnet CV?

MDK: Yes. Exactly. Jobnet CV, and give them advice on what to include in your CV. More routine-like things, such as ‘remember to update your joblog by entering two jobs per week.’

TG: The go through more practical issues before having an interview with you?

MDK: Yes. Exactly. We take care of the more personal sparing. Our challenge is we only have half an hour with the citizen. Before that you have to analyze them half an hour, before you have another interview. This is what we do here. You ask the citizen, sometimes you get in to more depth if the material we look at - when I mention the word material. It is a bit nerdy when you are working in this field, with material I mean an application and a CV, which are what the interview is based on.

TG: When I have to attend my first meeting with you. Should I have a company in mind, and I should have prepared my CV and an application?

MDK: No. You should upload them to your joblog because the unemployment insurance fund (A-Kasse) will ask you to do that. You are not required to bring anything with you. You are not expected to have prepared anything. If you want to receive your unemployment benefits (dagpenge), you have to upload your material. We look at them. This is our responsibility. ‘Have you uploaded these? Have you updated your joblog?’ If not, we have to report it to the unemployment insurance fund (A-Kasse). If we find out that a person has not conducted what we call a realistic job search. If you forgot to do it last week and the week before, then I am very sorry...

TG: ... What is the time span from the first unemployment insurance fund (A-Kasse) interview until I have the first interview with you?

MDK: Two to three weeks.

TG: Two to three weeks.

MDK: Yes.

TG: Fair enough.

MDK: Sometimes a bit more, because it is a huge system which is in operation. But usually two to three weeks in average, no more than that. They come to us when they are recently unemployed or just graduated.

TG: Okay, fine, fine. I come from STAR, and we have made this questionnaire (Forberedelsesskema) which you can fill out. In our web site, we encourage everyone to tell the citizen to fill it out before they attend an interview with you. Is it something that you use?

MDK: I must admit that we do not use it that much. We do not do that because all the things that is in the conversation questionnaire (samtaleskema), the ten points, we address in the course of the first meeting. Our experience is that the citizen has not done it before they come here. The unemployment insurance fund (A-Kasse) does not remind them. I do not know where they should get this information. I think it is written in small letters in the bottom when you register as unemployed. 'Please fill out the questionnaire (Foreberedelsesskema).' Typically they have to fill out within three months? Please correct me if I am wrong.

TG: Exactly. They have three months to fill it out.

MDK: Yes they have three months, yes. Sometimes we look at it, but there is nothing in it. If you look up a citizen, then you would look at the last completed interviews, and the summaries which we send to My Plan (Min Plan). Everything you need to know is found there, as we have it from the horse's own mouth. This is how we do it.

TG: Some of the things we encourage people to do is to call the citizens and ask them beforehand, before the interview, 'we recommend that you fill it out.' You will get the answers eventually to the interview, as you mentioned. Thus there is no need to begin a process which is part of the process.

MDK: I can see the idea behind it. However, we do not have the resources to call the citizens. But the idea is good, it is not bad. It is an OK tool. The reason why we do not use it so much is because the result is revealed to the citizen. You are an economist?

TG: Yes.

MDK: You just filled it out. You have not even had a meeting with the unemployment insurance fund (A-Kasse). You have just signed it. You read that you are welcome to fill out the questionnaire (Forberedelsesskema), which we call it. Strange term by the way, and then you get the result. You just graduated, you are so happy, you have faith in everything. Okay, there is corona. Let us pretend that we are not in the time of corona. You are so glad, and you are thinking, 'now I want to make money, damn I am good because I got 12.' You send applications, and you do not receive any reply. What is happening? Then you go through the unemployment benefit (dagpenge). Then you think, 'yes, a couple of months with unemployment benefits, and then it is on, yes.' However, the first thing you are told when you have filled out the questionnaire is that you are at risk of becoming long-term unemployed.

TG: The motivation is completely down.

MDK: Yes. Therefore I am not an advocate for this. We deal with a lot of different people. We have a holistic approach without being naïve and stupid. We know when we have to push it. But we like the idea that we give empowerment to the citizen. We work a lot with a holistic approach, you are a whole human being; you are not your education. It is important that we can sense Tarek in this application so it is not only a list of 'I can do this, I can do this, I can do this.' I have many cases of this type. Every day I encounter this; professional qualifications, professional qualifications, professional qualifications. You have no sense of the human being behind this. They give you a computer generated CV. This is what we do.

TG: Okay. I attend a meeting with you. We have been through the questionnaire (Forberedelsesskemaet) orally. You have posed some questions, and I have talked about my ambitions, and concerning when I expect to be back on the labor market etc. Do you try to spot if people are in the risk of becoming long-term unemployed? Is it something you think about? And if you do, what do you use to spot it? Are there any specific factors you look for?

MDK: There are some industries, niche industries. Back then, if you were trained as multimedia designer then we knew that you will end up in the risk group. The other group of people, who belong to the risk group, which we spot, and from the first meeting we already talk Plan B and C with them. That is people from the creative industry; actors, musicians all of these. We have a substantial number of these. We talk with them about average jobs which are characterized by routines (rugbrødsjob), already from day one. We tell them 'you have taken an excellent education. You are competent. We can see that you just made a commercial and made a lot of money. You have received supplementary benefits (supplerende dagpenge). However, eventually time will come, and this is how it is in the creative industry, and you have to take an average job. So you need to have a Plan B and something that helps. Suddenly the two year, also three year, the two year unemployment benefit (dagpenge) will run out. And if you do not have any savings or (inaudible) anywhere, then it is cash benefit (kontanthjælp), and no one should take that road.

TG: You say that it is two years, but it is actually three years?

MDK: Yes it is a three year reference period, which means that you can re-earn your right to unemployment benefits (dagpenge) in order to get a three year reference period.

TG: How do I re-earn it?

MDK: By working. You accumulate extra employment benefit (dagpenge) hours trough the unemployment insurance fund (A-Kasse). So all in all you can have a three year reference period.

TG: Is it like a subsidized job (fleksjob)? Is it Jobparat (when the jobcenter assess a citizen to be ready for taking a job)? What is it?

MDK: No, everything we talk about here, it concerns people who are jobparat.

TG: Okay.

MDK: I do not deal with other than jobparate. If I use you as an example again?

TG: Yes.

MDK: You call, my wife teaches in the other room, she works for the Ministry of Foreign Affairs, and they need an economist to manage her team. I recommend you, networking; this is how you get a job. You begin you work there. You know that it is a temporary position, three and a half months. . .

TG: . . . I earn?

MDK: . . . you earn there, you are unregistered from Jobnet. When you return you will continue with the same unemployment benefit (dagpenge) period. You have to work a whole year before starting over. It is not all who are aware of this. It is therefore a good question you pose. You return to the same unemployment benefit (dagpenge) period. If you for instance had an eleven weeks' unemployment period, then the eleven weeks will proceed. It might be reset; however, it depends on the time span. I do not remember how you do the calculation. It is the unemployment insurance fund (A-Kasse) who is responsible for it.

TG: Okay. Yes, it makes sense, it makes sense.

MDK: You are an economist, thus you will be able to figure the division in earning. I spend a lot of time on this.

TG: I will just have a talk with you, and it will be settled.

MDK: I am hopeless in figuring that out. We have an earning module which explains how it works.

TG: Okay. You have some reason to be concerned when some people say 'I would like to be a musician, I would like to work in the movie industry', and you know it is difficult to find a job. Do you have any other tools than the profile clarification tool which you prefer to use? Or do you draw from your experience? 'We use the Labor Market Balance (Arbejdsmarkedsbalancen)'

MDK: Yes. We use The Labor Market Balance, and we know what to do. We have meetings where we can look at the risks in various jobs, we know the percentages. As we are dealing with highly resourced citizens, not all can work here as a consultant. You cannot get a job here if you are 21 years old and recently graduated. There is a certain weight for those who work here. I have an education in management. I have worked in the private sector for many years, employed 400 people. It is very important that people possess knowledge on how the labor market functions. We all know when we should be concerned. Our average age, we just found out, is 51 or 52. So you can figure out, we all have reached a high age. The knowledge we have acquired is nothing you can obtain just by taking an education. We have been out there and try it ourselves. 20 years ago, our team, The Employment Service (Arbejdsformidlingen), consisted of many social workers. Back then the idea was that they would be able to provide the help necessary. However, social workers have not necessarily been out in the real world, and do not know how things work. The people who work here in the Youth Centre (Ungecenter) have been hired according to their experience; your work experience is not only from working at a Jobcentre. Of course some people have worked here for 25 years but (inaudible). We can spot these things, because we have been there.

TG: Okay, yes. You mentioned the field or industry can be a factor. Is there anything else which may cause concern, in terms of long-term unemployment?

MDK: Yes. You can also spot it on the citizen's behavior. Depending on how much you have dealt with people. Again with age comes experience. If the joblog is not filled out one week after the other. If you do not fulfill the requirements which we and the unemployment insurance fund (A-Kasse) have put out, and society in general; this gives us something to worry about. We know what to expect before we sanction. The citizen might be careless. Remember if you receive unemployment benefits (dagpenge), you can still have 10 million Danish kroner in your account. It is legal. Because the unemployment benefit (dagpenge) is some kind of insurance, you know that.

TG: Yes.

MDK: An insurance which you can use. If someone breaks in to my home, and I am quite wealth; I will still make use of my insurance, if I got robbed. There is nothing wrong in receiving unemployment benefit (dagpenge). It is an insurance you have chosen to take. On the other hand, if you do not live up to the requirements in order to receive them, then you will be denied. It is a sign, how a person's behavior is which you are trained to spot. All in my centre can spot it. We are 29 employees in my centre just so you have an idea.

TG: Just to be sure. You are familiar with the profile clarification tool?

MDK: Yes.

TG: But have you used it? Have you read this reading? Have you thought 'I can use it'? Have you acted on the basis of it?

MDK: I have acted on the basis of it, and I have used it. It is not that we do not use it. But usually it goes down to the completed interview. Let us use you as an example. You have been to the first interview here. You send your applications, and you do what you are required to do. You go to the second interview. All youth below six, sorry, all young people who hold a degree in our centre have six interviews in six months according to the LAB legislation (Legislation on employment efforts). I can do the math; it is one interview per month. When they come to the second interview, if they are still unemployed, then you would be able to spot if some things have happened. Because you have to read the summary of the last interview before you invite a citizen to a meeting. If you wait outside, you have registered your arrival, of course physical attendance. I can see that you have registered your arrival. I book you in our system, and I read about you. You never attend a meeting without us being prepared. You know that, I know that, and I also know what your education is. I have also read the summary of the last interview. I have also checked your placement in terms of offers, according to the regulations on activation (aktivering). All of these things. I have made my preparation work on who you are. I have read what have been noted last time. The exact same information are also found in the profile clarification tool.

TG: Okay.

MDK: We have been through it. I might open it sometimes and see some things. Sometimes you can use it to spot if a citizen has a rather unrealistic approach to his or her goals. We spot that. You have to remember. I sound so didactic, but I think it is so interesting.

TG: Of course (inaudible). I also think it is interesting.

MDK: You just have to remember one thing. People differ. There are some who give a reply. Have you ever tried the DiSC test? When you are hired for a position then you have to fill out something that is similar to the profile clarification tool?

TG: Yes, yes.

MDK: It reveals something about you? I think most people have tried it. You have to remind yourself that some citizens have the tendency to give an answer which they think you want to hear. Thus it is not 100 percent accurate. Although they are not really cheating. 'Do you consider yourself as long-term unemployed?' 'No.' Even if everything indicates that you will become long-term unemployed. There are no jobs for you. My ex-wife studied philosophy. If you study philosophy, and when you take the Masters degree, you become cand. Phil. Then you know there is one thing you can do, and it is to teach others in philosophy. There is only one class, and it consists of only ten people. The professor position in Philosophy will be advertised once in a lifetime. There is a reason to end up as long-term unemployed, unless you expand your horizon a bit.

TG: I totally agree.

MDK: This is a good example on...

TG: ... I am attending a meeting with you. The tool states there is a risk that I can become long-term unemployed, what do you then do? What is your reaction? What is the process?

MDK: The process is to expand to Plan B and Plan C. You have taken a great education, and you have to figure out how you can use it. Dissect it a bit; try to take out the professional qualifications. That is, well to use an economist is not the best example, as you are not going to struggle.

TG: I do not know. Let us say I am a pedagogue or something else. It is okay.

MDK: Pedagogue is actually also easy. There are no problems there. Let us take a small thing; then I will tell you. My wife, or ex-wife, you have a degree in philosophy. That is quite interesting, and if you are lucky, there are only two and half positions for you in a whole year. What are your qualifications? It is a subject within humanities, so you have a methodology. When you take methodology which is a great part in humanities, you can have an overview of things. You have learned to put things in to systems in order to even go through this education. You have learned to work cross-disciplinary. There is a lot of that. We try to define all these professional qualifications, and then we try to figure out how you can use it in other places. My ex-wife, when she found out, well first you study this subject because you think it is very interesting, you think it is just what I want, and you are happy, then you realize 'I cannot find any jobs'. It is good to figure out that there are other opportunities. She ended up working as an editor in a publishing house. The ability to work in systems, have an overview, work with words, it suited well to this job, we figured that out at that time. Back then I was not an employment consultant. It is a good example actually.

TG: Okay.

MDK: We look at the professional qualifications, and try to figure out where you can use these qualifications elsewhere. Does it make any sense? TG: Perfect; that gives 100 percent sense. Some of the things that STAR recommend in their website as regards to the next step on the basis of the results from the profile clarification tool, is to combine it with more interviews, to combine it with perhaps company internship (virksomhespraktik), or to combine it with courses. Have you been in a situation, or some of your other colleagues, where you thought this person need some come to more interviews in order to give him a leg up? He needs more courses?

MDK: Let me put it this way; it is not a tool we use to combine with more interviews, it is not possible. They have six interviews in six months. You know that after 26 weeks you are "defined" as long-term unemployed.

TG: Yes.

MDK: However before that, they have been part of our club for graduates (Dimittendklub) according to the LAB legislation. We do it from an early phase, when they have their first meeting. They are assigned our CV workshop, because the CV is. Do you use Ballisager's recruitment analysis? To check some things? We use it a lot.

TG: Unfortunately it does not ring a bell.

MDK: It is okay. It is a recruitment analysis which is primarily aimed at the employment sector.

TG: Okay.

MDK: 4,000 companies around Denmark, big and small, have participated in a panel. 84 percent say that the CV is the catalyst for getting a job interview. The application, although you still have to enclose it, give if you can count 16 percent. People do not look at the application anymore as you did back in the old days. When I hire I check the application, I look at it and think nice. You do not do that anymore.

TG: Okay.

MDK: It is the CV that counts. That is why we have made a module for teaching CV. 'You know Tarek come here and sit with the other 25 people, you are all graduates'. You all introduce yourselves, and then I would go through the principles on signs, messages and values. I tell them how important they are. It is not the layout rather it is what you write and where. To write Tarek, this is just an example, with font size 750, and CV even bigger, that would not count. You are not going to be hired on this. It is a nice name, but you will not be hired on the basis of your name. You will be hired on your professional qualifications, and the ability to explain to me as brief as possible how good you are and of course how nice you are. As we have a holistic approach to this.

TG: If there are no possibility for further interviews.

MDK: Yes.

TG: The first thing is to take courses, CV courses, or writing courses. How about company internship (virksomhespraktik)? Is there any who make use of it?

MDK: We address this in the first interview. We have some kind of macros where it says 'write everything you experience.' You are informed that you have the opportunity to take a four weeks' company internship (virksomhedspraktik). You can easily ask as the applicant, sorry the company to fill out the link, (inaudible).dk. There are links to everything. We address it. We do this and this. If you wish to try company internship (virksomhedspraktik) then you have to know that it is four weeks and not two weeks, or eight weeks as it used to be. Today it is four weeks. There is also a possibility for a wage subsidy job (løntilskud). Actually it is not possible at the moment, but that is another story. It is possible to get a wage subsidy job (løntilskud) after 26 weeks. We do address this, I promise you. I tell them already when they are recently unemployed that after 26 weeks, God forbid, some of you will end as long-term unemployed or are still unemployed after the 26 weeks. But I have to tell you. I draw how the wage subsidy job (løntilskud) works. It is good to have it in the back of one's mind. I tell them that this is just a Plan C, it is important for them to know.

TG: Okay. Just to be sure, you say that there are six interviews during half a year? Is it because you are not allowed to have more? Or is it because of practical circumstances that it is not possible, because there are many other things to care of?

MDK: Because it is in the LAB (Legislation on employment efforts).

TG: Okay. Just because it is there?

MDK: It is a legislative framework, yes.

TG: Fine.

MDK: It is purely a matter of rules. The whole unemployment benefit (dagpenge) system is built on the LAB legislation. It is states that in order to obtain your right to unemployment benefits (dagpenge) you have to participate in six interviews at Jobcenter in six months, and after that. . .

TG: . . . yes, after that?

MDK: after that; if was the Prime Minister, and I am not unfortunately. If I was, I would change some things. It is political, and we can discuss politics many years ahead. This is interesting and I have never got the grasp of this, I refer to the LAB legislation. If you are still unemployed after 26 weeks, then we only invite the citizen for an interview every 13 weeks, and this makes no sense. It makes no sense that when they get in to trouble that we say 'off you go'. I am not in charge of this. The Jobcenter is not in charge of this. It is like that. I do not understand this, it makes no sense. You should demand more interviews.

TG: Yes, okay.

MDK: you are not in charge, and I am not in charge.

TG: I have another question, and it might be obvious. What is the most important thing for person to do in order to get a job? The answer is of course is to apply for a job. Is there anything else that can help in the process?

MDK: Yes, yes. Unsolicited contact, networking.

TG: Unsolicited, yes.

MDK: Unsolicited approach is how you will get a job. Two out of three, and this is also according to Ballisager's recruitment analyses, it is not just something I say, it is documented. Two out of three jobs are not advertised.

TG: Think about it.

MDK: Two out of three positions are not advertised. Precisely. When I teach in this, I have this nice graph where you can apply for jobs. It shows that over 72 percent of people get a job through an advertised position, and they think 'this is what we will do'. No. These jobs are advertised, they will be occupied. But what will happen when you apply for such a job. Of course you have to apply for advertised positions. It is also stated in the LAB legislation that you have to do that. Fine, but what happens? You are competing with 40, 50, 100, 200 people. When you call, or send an unsolicited application to me. I have a company. 'Hi I am Tarek, I think your company is interesting, do you have coffee get-together, I would like to get access here.' I teach them this. Do not do all this chit chat, say it as it is. 'This seems interesting, can I get inside.' That will do the trick. That is how you will get your job today. I have witnessed that. LinkedIn, LinkedIn, LinkedIn.

TG: Yes. Okay let me make sure that I have left out a question. Two seconds.

MDK: Tell me if I talk too much. Believe me I think it is very interesting, I teach in this.

TG: No, no, on the contrary. I am sitting on the other side of the table, and I have to work with this tool. I have not really got acquainted with it yet. This whole process is new to me, so I have to look up for every little detail.

MDK: Yes, yes. Of course.

TG: I have heard some rumors which say some of the case workers wait until the end of the day to report that they have had an interview, and some might forget it or perhaps ignore it. I do not know if it is just rumors or a myth. How does it usually work?

MDK: I would not say that it can get you fired. I am going to tell you how it works. If we take the example of physical attendance, we are in special times at the moment.

TG: Yes, it is.

MDK: You just became unemployed; you have your nice diploma in your hand. You hold a BA in economics and consider to study further, an MA, never mind. You enter, register your arrival at Jobcenter, you sit in the meeting room. By the way there is free coffee over there. The number 924 is on, you come to me. I say' hello, how are we to proceed on with this, because we have 15 minutes for this meeting, I have prepared myself. You have this and this education'. And I will go through everything we spoke about. When you exit the room, I have ten minutes to journalize, and to write all these things down.

TG: Okay, okay. So you have some time allocated to do this?

MDK: You will not be fired if you do not do that, but I do not know anyone who does not do it. You cannot remember the details which we have been through, after I have spoken to six other citizens. It is not possible. I do not think that anyone do reporting like that, I do not believe that. I have never, no. All these fine details, all these things between the lines that you need to write. They will slip from your mind if you only took notes.

TG: One of the reasons why I have been assigned this task, actually the main reason, is because STAR was actually exposed in the media last year. The tool itself reveals to people how they are categorized.

MDK: Exactly.

TG: One register variable says 'descendants of non-western immigrants' for instance.

MDK: I noticed that.

TG: There were some cases, and people say 'hey am I being discriminated? Will I be invited to more interviews because of this? Will I be treated differently?' I have been given the assignment to develop the tool without this factor.

MDK: Yes, yes. Of course.

TG: The question is; are we dealing with positive discrimination? Or has it all been negative discrimination? If I am invited to more interviews, and this has worried some people. But if this will help me to get a job, then I should be grateful?

MDK: I agree with you.

TG: The Danish Institute for Human Rights (Institut for Menneskerettigheder) says 'no no, this is not positive discrimination, it is negative discrimination.' Have you any thoughts on this? And it is okay if you do not have.

MDK: No, I have not. No, not that much, because I have not encountered it. I am so privileged, that when it comes to color of skin, head scarf, or what your origins are, then I am quite privileged that. Some of the other departments might have another opinion. I have not witnessed any kind of positive or negative discrimination. I have not witnessed any racism neither in the job search market. I have not witnessed that if you have a darker skin color, then it will take you more time to get a job. I am dealing with well-off and competent citizens who have a good education. We are dealing with strongly resourceful people. They just need to learn how to apply for a job. When I teach these classes, I sometimes get these questions. One dental clinic assistant would ask 'do you think I will have a problem getting a job because I wear a head scarf? Do you think I should take a photo without the scarf?' I say I think that it would be a stupid idea. 'First of all that would not be good for what you believe in, and have you thought of showing up the first day without a head scarf? No I have not. 'Good, be yourself.' It would not make any difference. A person with a head scarf. I have not witnessed any positive nor any negative discrimination. It does not exist in our center.

TG: I am pleased, I am pleased. I do not have any prejudice.

MDK: You might hear it from other departments, where they deal with people who are more pressured. Unfortunately.

TG: I hope to reach out to more. I have reached out to many Jobcenters. I have an interview tomorrow and we will see what they have to say about this matter. One last thing Mikael. I am responsible for developing this tool. There are many factors which I have to look at. I look at if these people had any negative incidents in their life, any kind of disability which can be the reason they are where they are today. I look at when they have their first child, how many children do they actually have. Their parents' income. There are extremely many variables. Nevertheless I have this tool which is honest to say not so strong. You can be impressed of the amount of variables you are looking at, nevertheless there is something missing in this tool. As I understand, the industry to which they are heading is important?

MDK: The industry is important. Yes.

TG: The industry is important. Is there anything else you can think of which could be important?

MDK: Yes, the one called the human factor. I do not know how you would have time for this? It is very important that you have in mind that we are dealing with a strongly resourceful team. Because there is a huge difference when compared to the cash assistance team (Kontanthjælpsteam) who only receive 3,000 kroner in allowance and have to do things that you do not want to know about. The colleagues in our team are competent. Early in the process, they will benefit of filling it out with a job consultant, an employment consultant. There must be something you can do that can be added. The human factor, as you have not had your experience yet. You attend a meeting here, and you have your certificate in your hand and have a BA in something. You fill some things online, and this and this. And remember you lose your motivation if you are told that you are in the risk of becoming long-term unemployed. That is not so cool to be told. But forget this. If you go through then perhaps 15 questions with me who actually have know about this and maybe be prepared by you to go through this together with the citizen, it might make more sense. We know some things, which they cannot know because they are 22, 23, 24, 25, 26 years old. I think the average age is 26 years. More cooperation instead of filling out alone as it is now.

TG: Yes.

MDK: Because I am not sure they are capable enough to use it properly.

TG: Okay. You have actually a good point. You say that you go through the questions in one way or the other. But you should press the buttons for the person?

MDK: Yes, we take in plenum.

TG: Take it in plenum, yes. He explains his thoughts, and you answer on his behalf? And maybe it would be a good idea that they do not see the result?

MDK: Exactly. That is important.

TG: I have analyzed how other countries use the tool. For example in Malmö they use the tool but they have chosen not to reveal the results.

MDK: Yes, yes.

TG: Or they do not show the factors which are used to calculate it. In Denmark we think that we should be transparent. We want all the cards on the table, and this might actually have negative consequences unfortunately.

MDK: Yes, we are trying to make our job easier, I respect that. But it does not work, if a citizen who is happy, loses motivation. And the first thing you tell him you can as well expect to be unemployed for a long time. I would not like this to be told. 'Nice to have you here Mikael, you will eventually learn in a couple years or maybe five.' 'Oh.'

TG: Especially you.

MDK: Exactly, and I can imagine that this might be important for a citizen. I always try, and this might sound naïve. Every day, you meet new people. I say to all the citizens before they leave the room already from the first day 'I hope never to see you again'. They first think 'Is it something personal?' And then they think 'Of course because I have to get a job, yes, exactly.' It is a good way of leaving the room.

TG: Yes.

MDK: And I do it every time. The ones who have been to a meeting with me, they ask me 'is it now that you will say this?' 'Yes it is.' You been here for a month, and we do not want that. It gives communication. All what you do is communication, I want the communication out. So we can assess, if you and I are to fill it out together. 'What are your thoughts?' You say 'As an economist I my prospects are good.' Yes but the last couple of years it has shown that you might expect three months of unemployment.

TG: Yes.

MDK: The term long-term unemployment should not be used. People do not like it.

TG: One of my tasks is to look at ethnicity or the potential discrimination concerning it.

MDK: Yes.

TG: Statistics are discriminating; it is, no matter what, you put things in boxes. Today it is not a problem that you categorize as male and female, of course it is beginning to become a problem. All companies are now using 'male', 'female', 'neuter', 'do not wish to state type of gender', there are a lot of options. I have to make a tool which possibly does not have gender, possibly does not have age, and possibly does not have ethnicity. Suddenly

I have five different tools which I need to present to my supervisor, worst case can I use this? Then let us use this. Can we use this? Then let us use this.

MDK: Yes.

TG: It would be a good idea if you fill it out together with them, in order to avoid that they feel discriminated because of this or that category?

MDK: Yes. It is something that you work out together, as I bring my own experience from the labor market. Because when they are asked 'How do you see yourself in the labor market?' They have no idea. They give an answer which they think you want to hear, or what we want to hear. It is not certain that you can use their answer. That is why I think (inaudible) tool. But you are going to be busy.

TG: It is my thesis. You know what, I hope so.

MDK: It is quite interesting.

TG: Mikael, thank you very much for your time, for walking me through this, it is much appreciated. It will make it easier when I am going to talk to the others. I do not need to hear the process one more time, and I can concentrate in what is important for me. Once again, thank you very much.

MDK: If there is anything, you are welcome to contact me.

TG: I appreciated it. If there is anything I can be of assistance here in STAR, please reach out to me.

MDK: I just had a thought. It is an interesting and huge assignment you are working with. Would it be of any benefit to you to observe our work here – the everyday routines? Well maybe when we have people here. It is not that interesting now where we only receive people by phone. Observe how we conduct our first, second and third interview with the citizen perhaps?

TG: That might be possible, that might be possible.

MDK: You could get an overview. There are several departments, and they differ significantly.

TG: Perhaps. Mikael. That might be a possibility. Perhaps when more facilities open again, and there is time for it, and without disturbing. It would be appreciated to see how the process works.

MDK: In our department, it would not be of any disturbance. However at FBI, Fremskudt Borger Indsats (Advanced efforts for citizens), the first interview takes place in the cell. And the second interview is with two police officers. That would be of disturbance. However, in our department we would say 'this is Tarek who is from the agency, he would like to go through some things, would it bother you if he is here?' 'No.' You are welcome. Please reach out to me.

TG: Yes. Thank you very much. It is rather appreciated. When this interview is completed, I think the video would be posted in the comment section. I believe you can send a link of the video. I do not exactly know how it works. It has been a while since I have worked with it.

MDK: I know how it works, I have done it twice. I can do it.

TG: If you could do that, I will transcribe it. It is nothing that you should be bothered with. Again thank you, and if there is anything please reach out.

MDK: And you too, if I can be of more assistance.

TG: Perfect, thanks.

B.2 Jane Susan Ringberg

Interview with Jane Susan Ringberg (JSR) – by Tarek Ghanoum (TG)

TG: Can you kindly introduce yourself so I can get an insight in to your experience and what you do in the Jobcenter.

JSR: I am a social worker by training, and I have been working with exactly this area in Fredericia right after my graduation. In the last five years I have been here in Odense. (inaudible). I have experience with it, yes.

TG: Sounds like you have the relevant experience. You have also worked with the profile clarification tool?

JSR: I spend a lot of time preparing for each citizen. Some of the citizens do not use it. My experience is that some of the HK members (union for salaried employees) who hold a low-level of education, not always use it. People who have a middle-level of education use it more often, and those with a high-level of education use it. There are few workmen but not many of them use it.

TG: When you say use it, you mean the questionnaire (Forberedelsesskema)?

JSR: . . . yes.

TG: . . . okay.

JSR: Yes it gives you an idea of what kind of profile the citizen has. There was once an elderly workman who used it, and who was not in the risk group . . .

TG: . . . okay.

JSR: . . . and then I had some competent people who answered incorrectly, that they ended up in the risk group. Then I had to ask them 'you state that you want to pursue a new education, are you aware of your answer? Is that an error?' 'It is because I have misunderstood the question.' I think that people think that this tool is not important; it is something that they need to get over with.

TG: Hypothetically, I am unemployed, I just graduated, I register as 'unemployed' in Jobnet, which is the procedure. Then I am invited to my first meeting at the unemployment insurance fund (A-Kasse). After I am told that I have to apply for two jobs per week, and I need to prepare a CV, then I am invited for an interview here. Meanwhile, I might have filled out the questionnaire (Forberedelsesskema), and now I am attending this meeting with you. Would you in the meantime have given me a call and asked me to kindly fill it out? Or what do you do?

JSR: I do not have the time for that. We are in a lot of time pressure. We have half an hour for each interview, and we have to complete six in one day. We have 15 minutes for preparation, and 15 minutes to make notes. If the system suddenly does not work, then we are already behind.

TG: Okay. Let us say that I have filled out the questionnaire (Forberedelsesskema), and I have to attend an interview with you. You can see that I am in the risk of becoming long-term unemployed. How do you carry on from there?

JSR: It depends on what you have answered that reached that result. 'I have social or personal challenges'; there are some people who struggle with that. I will pinpoint out that answer, I tell them: 'You have answered this questionnaire. I have to make a risk assessment, a realistic job-plan, and what way you are headed. You have provided us with this answer. Is it something that you would like you like to talk about? Do you think it is something that we can help you with? Or is it something that you are already working on?' We have to assess if it is something which is permanent, then it is something that should be reported to the unemployment insurance fund (A-Kasse) and it has to be registered as health impediments. It depends therefore on what they have answered. There is a trap, however; when they state that they want to pursue a new education. We ask them: 'Do you want to pursue a new education?' 'Oh, did I write this?' I do not think that people always really read the questions properly.

TG: Okay, okay. It is just an answer given quickly?

JSR: Yes, it is something that they want to get over with.

TG: Okay.

JSR: If they have not used the profile clarification tool, then I ask them during the first interview if they consider filling it out.

TG: Okay. Let us say that I have not filled it out, and I have my first meeting with you. Would you ask me to fill it out in the course of the meeting? Or perhaps after the meeting?

JSR: I would recommend you to do it, because it gives an idea of where you are headed. It is the process which the citizen has.

TG: How much impetus do you put in the categorization which is provided by the tool? How much does it provide? Hypothetically I am in a sector where the job options are not good. The tool places me in the risk of long-term unemployment. Would you invite me to more meetings? Or what is your experience?

JSR: My experience is that, if people have challenges in one of the parameters which place them in the risk of long-term unemployment; then I would usually suggest a sequence of meetings. If there is a need then we can have meetings more frequently. These are the tools which we have; we have these options. The biggest challenges, however, are usually when the citizen does not want to attend the meetings, if you can sense a distance. Then I would tell them 'We have to do this, you have the opportunity to use us, we are available for you, and we are your tool.'

TG: Yes. Okay. Have you other tools? Sorry yes?

JSR: What did you say?

TG: I asked if you have other tools, besides the profile clarification tool, which you use for an overall assessment.

JSR: Well I use the Labor Market Balance (Arbejdsmarkedsbalancen) a lot.

TG: Yes.

JSR: Especially when we are dealing with people who are not sure what they are doing. We simply guide them through the education guide (Uddannelsesguiden) and use the tool fully. I use what is known as the expert list (Ekspertlisten) if I am dealing with a workman. Or if a construction engineer says 'I do not know who I have to talk to, is it entrepreneurs or other construction companies?' Then it is a suitable tool for guidance: 'It is not the way to find a job, but it is a good way to find out what kind of company this is, how many employees they have, what is their website. This can guide you further; it can give you a better perspective in terms of unsolicited contact or to the place where you want to go'.

TG: Which factor could according to you lead to long-term unemployment?

JSR: If you do not have an overview of your time perspective, in terms of your expected unemployment period. It can be realistic or unrealistic. If you think that it would take you more than six months. 'Why do you think that it will take you more than six months?' You can face a recess period as there is a moment with Covid-19. However there might be jobs in other continents. I mean in for instance Jutland, for instance on Fyn it has been going well. On Zealand it might be more stagnating; then I wonder on the categorization of people on Zealand; are they less on their marks? They want things served on a silver platter? They are less proactive? That is what you learn in the media, and you use that to reflect on where the opportunities are found?

TG: What is your general impression of the tool? What is your general opinion?

JSR: I think it is good, but I wish that it was open for editing. People change their approach in the process. They can think 'These were my ideas before, now I am thinking it might take nine months, however I am on my way, because we have already started the process.' They might have started a basis course, it can be something else which have helped people, e.g. supplementary training which has an impact on how you perceive your job situation.

TG: The tool is used as an indicator?

JSR: Nevertheless, if the citizen is using it. I can use it. As the citizen can now. You can follow the process of the citizen.

TG: Okay, do you think? Yes?

JSR: You encounter a person who thinks 'I have no abilities I do not know what my professional qualifications are, I do not know anything'; being rather discouraged. These kinds of people end up with an RFL as we call it. Then you initiate a process, and after a month or two then they end up realizing that they have these opportunities. 'I will make use of this, now I am on this process', then you develop yourself, mindset wise. When you can see the progress, then you get a, I do not know what to call it...

TG: ... motivation? Or desire?

JSR: ... yes a better drive for moving on. Then I think it might be a good idea, it can be opened up in terms of job interviews, and lock it when you have booked a job interview. When you have been to a job interview you can reopen it until you have the next job interview. You can use it continuously. ‘Last time you wrote this, and now you are writing this.’ That becomes a parameter in the tool. ‘You are on the right track.’

TG: Have you ever encountered some people who provide answers they think you want to hear?

JSR: Yes. But not in the profile clarification tool, they cannot do that. They would give reasons which they think I want to hear. Then I would tell them: ‘You can answer what you want.’ But it is easy to spot…

TG: … are there some who would claim that they will have a job within a week?

JSR: … pleasers. ‘You should not please me. You are not here because of me. You are here because I am your tool’.

TG: Yes, yes.

JSR: I think it is a good idea to address this issue. The most important parameter in a job interview is to balance the expectations. If we do not follow the same track, then we would go separate ways and that will not end well. Then we would not develop.

TG: But you have not had any cases where some people have filled out the questionnaire (Forberedelsesskema) according to what they think you want to hear?

JSR: No, no. My impression is that they are rather honest. The questionnaire gives a here and now situation. If you had a bad day and you are thinking I have to do this and do this, I have to do this, I have to do this because I have to do it. I must do it because I have to.’ It will have an effect on the answers you provide. Therefore I wish it was an open tool, that it can be opened and locked. If you had a positive experience which makes you more positive, and it will give positive expectations in terms of finding work.

TG: It is not motivating if you are told that you are in the risk group? Is that what you had in minds?

JSR: Yes, when you address the issue ‘You have answered this, which concerns me, because I do not have this impression of you, but your answers is this and this, what are your thoughts? Can you explain the reason why?’ They would answer: ‘Well, I think it was a bit difficult and confusing when I became unemployed’; that is if you recently graduated for instance. ‘It was hard for me to figure out’. ‘Then let us talk about what your professional and personal qualifications are, in terms of the labor market’. ‘Oh I can do that’. ‘Yes you can’. Many people come back to me and say ‘Thank you, this gave me a push to look at it from a different angle.’ Based on that, it would be a good idea to be able to open and lock it, but …

TG: … Do you think that it would be better to fill out the questionnaire (Forberedelseskema) together, during the course of the first interview?

JSR: That would be difficult especially if it is done through a phone call, as is the case at the moment. Right?
TG: Yes.

JSR: It would be better, however it would take more time, because a 30 minutes' interview with all this information would be a challenge. We previously tried to hand out questionnaires which people had to fill out and return so we could add them to their file, however only 10 percent used it.

TG: Okay, okay.

JSR: If you made it mandatory, or told them 'This would be part of the job interview, next time you are here I would like you to fill it out.'

TG: Do you think it would help if it was mandatory to fill it out?

JSR: No. We have some people who are workmen, and some who are unskilled. They cannot see the idea behind it. Some of them cannot see the point in having a CV in Jobnet. 'I will eventually go and knock on carpenter Hansen's door, he knows me'.

TG: Yes, yes.

JSR: We are also dealing with groups of people which are difficult to work with. If there were more time, then you could have the conversation about this. 'If you had to fill this out, what are your thoughts?' Many of them are not capable of abstraction or reflection. It is usually this group who think 'I am about to get a job'. Like the elderly man I mentioned. He was not placed in the risk group, he thought he would find work soon.

TG: Yes, okay. One of your suggestions is that the tool could be unlocked so it can be edited. Have you had some other considerations?

JSR: When you have these job interviews, then maybe these job interviews should last for an hour. As you mentioned, in order to address, not the employment wheel (beskæftigelseshjul), a wheel with different parameters, to have more time to put it into perspective.

TG: Yes, okay.

JSR: When you go through the Labor Market Balance (Arbejdsmarkedsbalancen), and go through the different parameters which are available, this is rather helpful. However it would be a good idea to address how these tools are used via the unemployment insurance funds (A-Kasse).

TG: Yes, that is a good idea. Some of the things others have told me are, that an unemployed loose his or her motivation if they know that they belong to the risk group. Would it not be better not to reveal the result to the person who has filled out the questionnaire?

JSR: I think they need to be told the truth. It can be an error, it can be…

TG: … the wrong categorization…

JSR: … when a person informs you 'I am thinking of pursuing a new education.' There is something that is not right. 'You have stated this, why have you written this?' 'Oh did I write this?' As I told you, they are not always aware of their reply. I do not know if it is in the formulation, or maybe there should be some more blanks.

I think there are three different blanks: ‘I consider commuting’, ‘I consider pursuing a new education’, and there is one more. Maybe to sub-divide them.

TG: Okay, yes. You have a point there.

JSR: ... ‘are you professionally skilled?’ or are you, ‘what kind of education do you have?’. Perhaps divide them in that way, so it is easier to sort out. I know the questions but sort them out. Let us say we have a nurse, and she just completed her education, the probability for her to pursue a new education is not that big.

TG: No, of course.

JSR: No, nor a social worker. It is expected that they want to start working within their field. Nor a doctor who have been studying for maybe six years ...

TG: ... yes.

JSR: They want to continue with their field.

TG: Obviously.

JSR: It would be a good idea to separate these from others.

TG: Okay, yes. That makes good sense.

JSR: Yes. If you are dealing with unskilled people who consider pursuing an education, this will give you an idea of this person, does he/she want to or do not want to, and from there you can work with their motivation.

TG: Yes. It makes sense.

JSR: Yes.

TG: We have now discussed the tool in general, and you mentioned also other tools you use, for instance the Labor Market Balance (Arbejdsmarkedsbalancen)?

JSR: Yes.

TG: If we are to optimize the whole process. We have discussed this. If we divide the questionnaire (Forberedelsesskema) in to skilled/unskilled etc, from your experience is there something that you have wondered upon, something that should have been taken in to consideration? The tool itself has several parameters. I will for instance take in to account which kind of education a person has, what is the income of the parents. Factors such as gender, origin, and several other factors have been taken in to consideration. Are there any factors which according to you are very important to have in mind when a person becomes long-term unemployed?

JSR: No, I think there are factors which disqualify people.

TG: Yes.

JSR: And that is what a person’s origin is.

TG: Okay.

JSR: We discuss integration so much; then it is like a slap in the face.

TG: Okay, yes. Can you, yes. . .

JSR: . . . do you understand where I am going with this? You can have it in the questionnaire but it should not take up so much that it makes people suspect that it is a parameter which puts them in the risk zone. People already feel marginalized, they have pursued an education, and they have difficulties getting into the labor market. The reason could be because they have a different name. I know there are some fields where it is difficult to get a foothold. I think it is a bit sad. I have dealt with young Danish people, as I call them. They say, 'I speak Turkish, it is my mother tongue,' and Danish is ticked as 'experienced'. I ask them why they wrote this; they respond: 'Well, I am Turkish.' I tell them, 'Your origin is Turkish, you speak fluent Turkish, I also speak fluent English, but I am Danish. So think about it in terms of the parameters. You are Danish with Turkish origin.' If we keep throwing people into these categories, it will be tough to make this connection, as we have discussed in terms of work. For people who have a different ethnic origin, it is a bit sad. I also feel they become marginalized, 'because of my origin, and because of my name.' It is a pity. It makes me sad.

TG: Definitely, definitely. I think that people which have been categorized as 'descendants' or 'immigrants', might have the suspicion that they have been invited to more meetings because of this; in that way they felt discriminated. Have you had any such cases?

JSR: There are some Danish people who make a fuss, and there are people with a different ethnic background who make a fuss. I have dealt with both groups, and I have been scolded by both groups. I say things as they are.' I think you need to do that.' And I tell them 'You know what, this is not right, I need to report this to the unemployment insurance fund (A-Kasse). I need to report to the control group (Kontrolgruppen) because you are not doing what you are supposed to do. I cannot reach you, and I have called, and I have done this and that'. It does not matter if it is a Danish person or who it is. 'It does not help you.' People should not cheat the system. There are always some who do that. There are some workmen who do cash-in-hand jobs. If we know about it, then we report it to the control group, and we also notify the unemployment insurance fund (A-Kasse). We do not treat people differently.

TG: No, no. Of course not, of course not. It makes sense.

JSR: I do not believe that my colleagues are targeting a specific group, but we are attentive if there are any jobs available which match our professional groups. We call people and say 'We know that you are struggling with finding someone to take care of your children in the evening because your husband works in the evening and you work in daytime, and you are trying to make it work. I just received a job order, and I thought of you, are you interested? Then I will notify the company consultant (virksomhedskonsulent). Can you call her, I will text you with contact information on the company consultant (virksomhedskonsulent), so you can be informed about this job.' I will only reach out to people if I have a job order, or if there are some courses, or something that can benefit them. Otherwise I will not contact them. If people give you the impression that they want your help then they will get it.

TG: Yes, this makes sense. Perfect.

JSR: But we do not hunt people down. It might be the case in other places. I do not do that, neither my colleagues.

TG: I just have one last question. The results you get from the profile clarification tool, how much importance do you attach to the result? Do you take it with a pinch of salt? What are the most decisive in the results in terms of how to carry on from there?

JSR: To involve them in the job interview, and settle with them what is realistic and what is not realistic. If you help them balance their expectations. When you talk to them and you say: 'I can see that you have filled the questionnaire (Forberedelsesskema) and you have answered the questions. I have questions for you because it seems that you are in the risk of long-term unemployment.' That is a nice way to tell people, what is realistic and what is not realistic. You have opened up for a dialogue, and then people can explain what it is. 'I can see that you have some challenges with your mental health, you have social challenges, is it something that I need to report? Is it something that I need to deal with otherwise? Or what is it about?' You do not want to step on people. That is why I think that it should be possible to unlock the questionnaire to make it possible to edit peoples' prospects, as we are dealing with a here and now situation when you fill it out.

TG: You are absolutely right.

JSR: I think it becomes unambiguous and definitive and we use it as a basis. It can be rather toilsome. If you told your wife one day 'You are so stupid'. You are so mad because you had a really bad day at work. She would guard herself because she is thinking will he say something again the rest of the day. Life is not like that. It is just a here and now situation.

TG: Obviously, obviously.

JSR: Unlock it. 'Now we have discussed this and now we can see your perspective in a different way', it becomes an active tool rather than a dead one.

TG: Yes. It makes good sense.

JSR: Yes.

TG: You know what Jane, this is perfect. I just want to say thank you very much for taking your time to go through this with me.

JSR: Can you use this?

TG: Yes, yes, it is very beneficial. I will definitely work on this, especially your feedback.

JSR: I think it is a good tool, I use it every day, however as mentioned, in order to lead a proper conversation, I spend half an hour for preparation. I have a questionnaire, and in that questionnaire I have 'education', 'looking for a job as', 'CV', 'Job-plan' and 'questionnaire' (Forberedelsesskema).

TG: Yes, yes.

JSR: So it is on the top of my list. Then I have 'Jobnet CV', followed by 'Joblog'.

TG: It is quite useful what you have addressed here, and I will definitely use this for the development of the tool. I will use the feedback.

JSR: I think it is a quite excellent tool, and it is good that you have something you can base your work on. Otherwise you have to start from scratch. On the other hand it would be a good idea to introduce it via their unemployment insurance fund (A-Kasse), as they are also told to fill out their Jobnet CV, then you have to look at the...

TG: ... questionnaire (Forberedelsesskema).

JSR: ... questionnaire (Forberedelsesskema), your tool...

TG: ... Exactly.

JSR: ... This is something that will be used for continuous revision, as you have to revise your CV and Job-plan, it is not static. When you change the parameters, then we change that.

TG: Exactly, exactly. Yes it makes sense. That is true.

JSR: They are connected, they are connected. However the issue of people with a different ethnic background and non-western immigrants, I think we should stop it, otherwise we will never give access.

TG: Yes, it makes sense.

JSR: We try to get people into employment, and that includes people of a different ethnic origin. That is fine, but why should we keep reminding them that they are different? Why should people be reminded if they are from Jutland?

TG: Yes, yes, you are absolutely right, you are absolutely right. I hope in near future that it will be removed so the category is not visible, we are working on that.

JSR: We have it in facit, and we have it on momentum, it is stated there. Why should it also be included there?

TG: Yes, you are absolutely right, you are absolutely right. We have the intention of removing it.

JSR: I hope others will tell you the same thing. I think it is quite intimidating that I have this information. I am dealing with a person. I am dealing with a human being.

TG: Yes.

JSR: I am not dealing with a country.

TG: Of course, you are absolutely right, you are absolutely right. Thank very much for that. I follow you 100 percent and it makes sense. I have been given this assignment to develop this tool without this parameter, so we can manage without this categorization. It is something that we are working on. I am part of this team that is working on that.

JSR: This is a big request from me. If we can help doing something positive, it can help move many things as we need to see people as whole human beings, and where there are possible barriers. As I understand from you, there are some places where you use this part, well of course. When I meet Amina, or I meet Kassim, or I meet Trine who is worn out, 59 years old, my job is to try to help. It does not matter who they are.

TG: Yes, of course.

JSR: Even if it is Heinz from Germany, I have to help as well.

TG: You are absolutely right, you are absolutely right. I think we have discussed all the matters; the tool, your experience with it and how you use it. It has been very relevant and very beneficial.

JSR: Just remove them, so they do not stand together. People can misunderstand them, people who just received their degree state that they want to pursue a new education, something is not right.

TG: You are absolutely right, we look at the categorization of these questions, and how we can optimize them. We are working on them at the moment, definitely.

JSR: You should not have negative questions.

TG: Okay.

JSR: If you write, 'What will you NOT do?' Do not write these questions. Such questions can be misunderstood, people do not catch the 'Not'; people do not read the question properly.

TG: Yes. You are right, you are right.

JSR: Negating sentences are a no go.

TG: Okay.

JSR: Try to read it, you do not read the rest.

TG: Yes, you are absolutely right.

JSR: Even if it is reversed. There is a focus on the Not, and then they just answer it.

TG: Have your colleagues also experienced that? I.e. that people might have picked the wrong box?

JSR: Yes I think so. We laugh about it. When he or she has written something, and we think this does not make sense at all. This makes no sense.

TG: Okay.

JSR: We often experience that. Typically it is when people have misunderstood the question. It should be very brief and very clear, no parenthetical sentences. Very brief and clear.

TG: The categorizations are first provided when people have filled the questionnaire (Forberedelsesskema)? If you have not filled it out, you will not get the result?

JSR: When we have a job interview, and people have not filled it out, then we need to assess if they are in the risk group.

TG: Okay, so the tool ...

JSR: ... so we tell people, even if we do not have the questionnaire: 'Assess your way to obtain a job. We have to figure out if you are in the risk zone.' Right now we only use the statuary process; they have right and duty for 26 weeks of unemployment. Before corona, when people were in the risk group, e.g. graduates, above 60 - if there were any indications, and unskilled people who did not know their professional qualifications. Then we had one when they reached two months of unemployment, then we had one when they reached three months of unemployment, and one when they reached six months of unemployment, and one when they reached16 months of unemployment; plus something extra for helping and supporting them further. If we are dealing with graduates, we look at how they are helping themselves in writing an application. Are they invited to job interviews? The ones who do get job interviews are not in the risk zone as those who do not.

TG: Of course, of course, it makes sense.

JSR: We already work with the categorization. Therefore, it will make sense if it is open for getting the citizens' own perspective, because something happens with them when they have been unemployed for three months. Everyday hits them, they become discouraged, and they loose faith on their own abilities and professional qualifications. So there is a lot of work to do in terms of motivation. If you can follow the progress, then it is easier to keep them on the positive side. 'From the last time until now, this and this happened.' People will be able to visualize it. When they see it, they are amazed.

B.3 Marianne Krogh Fischer

Interview with Marianne Krogh Fischer (MF) – by Tarek Ghanoum (TG)

TG: Kindly explain what you work with, and your experience with people who are searching for work?

MF: Yes. I work as an employment consultant at BSF (Employment and Social Services Administration), and I work with Efforts for Academics (Akademikerindsatsen). My focus is on counseling academics. I have worked there for a year. Before that I have been working with other professional groups. My impression is that the questionnaire (forberedelsesskema) is not suitable for academics compared to other professional groups. However, we us it. But I must admit that I forget it sometimes.

TG: What makes you say that it is not suitable for academics?

MF: I do not think it provides much. I can use some of the blanks, for instance 'how long do you think it will take before you find a job?' TG: Yes.

MF: It gives you an idea of how they perceive their unemployment situation. As regards to the other blanks, I use their CV in Jobnet as the information provided there is more detailed. It does not provide much that that they have a higher level of education. I want to know what kind of education it is. There is a great difference if you have an education as a mechanical engineering or you work with communication. The opportunity for employment differs significantly.

TG: I agree, I agree. What do you think is the most important blank? Or the most important blanks in this tool?

MF: The most important question is their own prospects for finding a job. I think it is quite useful. And if they have applied for jobs, have they even started looking for jobs, it is rather important. We have many graduates. It reveals a lot. I think over half of the graduates have already found a job before they graduate. If you have not started applying for a job before you graduate, it says a lot about the person.

TG: Of course.

MF: There is a blank concerning mobility. How broad is the geographical range? And there is one on how you apply for jobs. Is it only advertised positions, unsolicited, do you use Linkedin? It is also important.

TG: Okay, these are good points. People have to go to the unemployment insurance fund (A-Kasse) before their first meeting with you. Meanwhile, it is up to them if they fill out the questionnaire (forberedelsesskema). Do you usually get in touch with them, or somehow remind them to fill it out before the first meeting?

MF: No.

TG: No, it does not happen.

MF: But I do not know if they do it, when they receive the welcome letter which is sent to all the unemployed when they register. I have no idea. However, I am notified when the questionnaire (forberedelsesskemaet) is filled out. I note 'The questionnaire (forberedelsskema) is completed'. When I have to prepare for the first interview with a given citizen, I will have a look at it.

TG: When I spoke with Jane yesterday, she told me that people do not fill it out properly. Some of the answers provided are bizarre, and some people are not really aware of their answer. Is that your experience?

MF: Yes. That is why I think it does not provide much. It is quite obvious who puts an effort in filling it out, and those who just want to get it over with.

TG: Okay. Yes. Jane suggested that it should be possible to be able to edit it together with the citizen. Do you think that it would be a good thing to do?

MF: I think we already have several things that we need to adjust. I think it is a questionnaire which should be available for our preparation to the first meeting, as we can use it to get an impression of a given citizen . . .

TG: . . . okay.

MF: . . . and from there we ask about the different elements. If we start adjusting the answers, then what is the point of it being there? We also have a job-plan and what should we do with it? I think it can be immense. There are already many things that we have to remember to do.

TG: I am thinking that the editing could help on the assessment of being in the risk of long-term unemployment?

MF: Yes.

TG: Do you think the tool helps to reach the final decision?

MF: Yes. If you do it thoroughly?

TG: Yes, if you could edit it a little together with the citizen. If a citizen tells you 'I expect to have a job within a week', and the Labor Market Balance (Arbejdsmarkedsbalancen) clearly shows that the job prospects are not good, perhaps one can adjust it a little with the citizen. 'Maybe you need to readjust your expectations a little bit?'

MF: Yes. That could be a possibility. I have not thought about this actually.

TG: That is okay. Another thing. Let us say that I just graduated, and I belong to the group who is in the risk of becoming long-term unemployed. What is the process which you will typically initiate?

MF: I would look at what the options are within your core qualifications. You usually fill out the questionnaire (forberedelsesskema) from what you imagine could be your dream job. The questionnaire would change significantly if you reached the conclusion that the dream job is not the first job, or the second job, but later in future. It is very important to address it to newly graduates.

TG: Okay, okay yes. There are some who would end up in the of risk of becoming long-term unemployed. Do you think that some people would be affected negatively, if they were told about that? They might take it personally? Or get annoyed or? Have you encountered that?

MF: I have not encountered it. I could imagine. I think it can be a bit tough to be told.

TG: Okay.

MF: It is definitely.

TG: Do you think that one should formulate it differently? Or perhaps not show the result? What do you think?

MF: No, but, but. I cannot remember what it says. But you tell them that they should broaden their search of jobs so it covers other areas where they can use their professional qualifications. On the other hand, it does not hurt to make an extra effort. At the moment, I am dealing with many graduates who have not even begun their job search. Some of them are quite honest about it 'I want 14 days off before I begin.' I am thinking 'You have not even thought of looking for a job while you were studying?' When you are told, then well.

TG: It is then a good opportunity to pull oneself together?

MF: Yes, yes.

TG: When you fill out the questionnaire (forberedelsesskema) you are informed how you have been categorized. The categorization can be according to age, according to gender, according to origin, there are many categories. Last year, the media revealed that people were very displeased with the origin category. . .

MF: . . . yes.

TG: . . . they were not pleased to be put in a specific box. For many who are born and raised in this country, nevertheless they are placed in the 'non-western origin' box. Have you encountered some who have been annoyed or wondered about it?

MF: No, I have not. But I have wondered myself. Our concern is the Danish language abilities.

TG: Yes.

MF: So maybe it should be changed to that instead. Origin. I think it is a little. Is there also a category for age?

TG: I do not know if it is visible to others, but the tool is also based on age. It is one of the factors which we take into account.

MF: Yes.

TG: There are several factors. What is visible to the citizen, and what I remember specifically is origin, as it was debated in the media.

MF: Yes. And it has negative connotations.

TG: Yes, unfortunately. Imagine that I am unemployed, and I belong to the risk group. Would you invite me to more interviews? What is your experience?

MF: No. We do not do that. But in the first six months of your unemployment, you have to attend an interview per month. It is quite intensive.

TG: Okay, okay. What happens after the six months typically?

MF: After that, they have an interview every three months.

TG: Three months?

MF: Depends on your assessment of the citizen, and we do that too. In these times of Covid-19, we are going to have more interviews. I am going to have eight to nine interviews within the first six months of unemployment. Everything has stopped. This is what we are told by the administration, what we are allowed to do. They are doing the calculation so we can complete all the interviews.

TG: Okay. If there is something you would like to optimize or change with the profile calculation tool, what would you suggest?

MF: The issue with origin, it is something that I have questioned. I think more prose-like, but you cannot use that form for searching. 'I have a high-level of education' it does not provide anything. You need a blank for field.

TG: Perhaps a category about where one is applying?

MF: Yes, I think it would be that. Yes, where you are searching. I do not know if it is possible, but within which field is your education, not just that it is a high-level of education. There is a great difference between a doctor and.

TG: Definitely, definitely. I agree.

MF: Yes. You should add a sub-category. If the Jobnet CV is not available, then it makes it easier to get a general picture.

TG: Yes, that makes sense, that makes sense. I am thinking, I just forgot the question.

MF: Yes. This questionnaire has been constructed so it actually fits all people?

TG: Yes, yes.

MF: I am thinking that it is a weak point?

TG: Some have actually suggested that one should be adjusted to fit academics, and one for unskilled people.

MF: Yes.

TG: Sorry skilled and unskilled people.

MF: Yes.

TG: That it should be done in that way. Because there are many unskilled workers who do not have the same approach to CV and job search as the skilled workers might have.

MF: Yes.

TG: That is something to consider. There is another thing I would like to know. How much faith do you have to the results which the tool provides?

MF: I do not have much faith. This is a reason why I agreed to give this interview. I just give it a quick glance. I do not get much from it without the other things.

TG: Okay.

MF: It is not a tool that I can use solely.

TG: You probably use the Labour Market Balance (Arbejdsmarkedsbalancen) as well? Employment Barometer (Jobbarometer)…

MF: … The Labour Market Balance (Arbejdsmarkedsbalance), Jobnet CV

TG: Yes.

MF: The personal CV. The tool cannot be used alone.

TG: You reach an overall assessment on the basis of several parameters?

MF: Yes, yes.

TG: Okay. Perfect. The last question. Do you know how the tool functions? What is behind the tool?

MF: No.

TG: No. Okay, okay.

MF: I can only guess…

TG: …yes, that is quite okay.

MF: … when I read the questions. Then I have some thoughts about how they answered what the options are.

TG: Yes.

MF: So, we need that.

TG: Okay, yes. It is okay. I have not encountered any who know how it actually works. Several methods have been chosen, because one had hoped that if someone would like to be more acquainted with it, then it would be easy for them. That is why I would like to know, if you had any courses, an introduction about it ….

MF: … no.

TG: … not there has not been any. Okay, okay. I think we have covered everything. Do you have anything?

MF: Are you going to revise it?

TG: Yes, exactly. One of the reasons why we are revising it, or reinvent it from scratch, it is among others to remove the ethnicity or rather the origin variable.

MF: Yes.

TG: We are trying to remove it, so nobody feels discriminated. We would also like to optimize it. We will try to add more variables, more data, and hopefully improve the questionnaire as well. There is a lot happening. I also know that Odense has been granted a couple of million Danish kroner to redevelop this tool.

MF: Yes.

TG: There is a lot happening, hopefully in near future it will be available to you.

MF: Yes. I am thinking about the issue of age. I have reached the age where I am interested in this. I am quite aware of the discrimination on age that I think is part of the fact that we call people 'seniors' when they are 50. They still have 15 years in the labor market. I do not think we should continue with this. When we use this term, we place people in a box. [...]

TG: . . . yes, categorizing.

MF: Yes, categorizing. Which is not necessarily positive. Add some other terms, for instance, 'experienced' [instead of senior].

TG: Yes, a good point. How about gender. Have you any experience with that?

MF: No, I have not. I am thinking it is . . .

TG: . . . quite normal. . .

MF: . . . it is not something that we should do today.

TG: I fully agree with you. Just to be sure; gender, age, origin, is there a fourth thing you have in mind?

MF: No, but if you have anything in mind, I might fully agree with you. What do you have in mind?

TG: No, it is actually these things that I am currently examining; the effect on the tool if I remove them. Think about it, it has a consequence. If I have an excellent tool which includes origin, age and gender. If I remove one of the variables, it will weaken the tool. The question is am I willing to pay the price for avoiding the risk of some people feeling discriminated? It comes to the same thing. I might take something out, but I might get something else.

MF: You could instead put some other parameters?

TG: That is what I am working on at the moment.

MF: Yes. There are other factors that may reveal other things. It says something about a person if you live in Korsløkkeparken or, yes.

TG: Yes, exactly. You are absolutely right. The residence address could be interesting to examine further. Usually, the data we have is on a municipality level. I have made sure that all of the data is in what is called (inaudible) level, so we are looking at it from a higher level. To my knowledge residence address has not been an issue, or potentially problematic. . .

MF: … but it is. It is indeed. Several years ago, I worked with young people who did not have an education… You had this list on young people, and you thought that those who have residence in Odense M, 'they are taking a sabbatical year', and there were others where you thought 'we need to get hold of them.'

TG: That is a good point. It is quite relevant; it is rather relevant. I have actually never heard anyone address this before, I am glad that you mention it. I will take it to my notes. Well Marianne, I think we have covered everything, I do not want to take more of your time …

MF: … I hope some of this has been useful?

TG: A lot, undoubtedly. Now I would like to ask you to stop the recording, then I will transcribe it and analyze it, in order to include all of your good points…

B.4 Mie Verning Maabjerg Hansen

Interview with Mie Verning Maabjerg Hansen (MH) - by Tarek Ghanoum (TG)

TG: If you could begin by introducing yourself briefly, I only have your name and title. So kindly explain what you do, and what your responsibilities are in terms of the Signature Project, and we will take it from there.

MH: Let us do that. My name is Mie Hansen, and I am chief consultant in the Employment and Social Services Administration (Beskæftigelses- og Socialforvaltningen) in Odense Municipality. Besides working as a project manager for the Signature Project, I work with what we call a Governance (inaudible) for the entire area of employment and businesses.

TG: Okay.

MH: This entails that I have the main responsibility for the Employment and Business Service; that they have the proper IT, and that it is exploited in the best possible way, and that the operations department is well supported.

TG: Yes.

MH: Everything from development projects, adjustments, solutions, supply management, negotiations, supplies, all the things which are dealt with in our Governance in the system landscape. Fortunately, I have four competent colleagues, also referred to as system contact persons. They know how our system works, they can (inaudible), and support our operations department. But I have the general overview.

TG: Okay.

MH: This also entails that I have proper knowledge about how we work within the field of Employment and Business. The Signature Project here is anchored in what we call Job 1, which deals with the unemployed who are assessed by the Jobcenter as ready to work (jobparat), and with focus on recipients of unemployment benefits (dagpenge).

TG: Yes, perfect.

MH: Yes.

TG: I got acquainted with the Signature Project through a small presentation, a brief introduction. I guess that your responsibility, and there are probably other elements, but part of it is to develop this tool? Some kind of screening tool? Can you kindly briefly explain for how long you have been working on this? Your considerations in the process for the initiation of developing it, as we already have one?

MH: Yes. As mentioned, I am the project manager, and have the main responsibility for the project, so I have insight in to all the elements of the project. We are not there where we have initiated the development . . .

TG: . . . Oh I see.

MH: . . . of a screening tool yet. I do not know where you got your information on the introduction to the project. However, we are in a phase of settling what we need, as it was important to have the operations department on board, as well as the strategic management. Our question is what do we want to obtain in terms of working with artificial intelligence? But also how it would benefit employees and citizens?

TG: Okay.

MH: As we want to work with algorithms and with this technology, we want to makes sure that we are able to develop a tool which can be of practical use and be value-creating for employees and users.

TG: Yes.

MH: We have spent this autumn, working together with a consultant firm, or rather an innovation firm, which is locally situated in Odense. From that process, based on different activities, that is discussions of strategies, in-depth interviews with citizens and employees, design workshops, and different types of activities, we are trying to figure out how we can be supported with artificial intelligence. We have put up a vision, and if the sky is the limit, this is how we would like to work with it. In this vision, there are several elements, which could be of interest to work further with. There is no screening tool as such. . .

TG: . . . Okay.

MH: However, the task of our Signature Project is to provide recommendations to types of efforts for the citizen. What kind of efforts will help a citizen in finding a job or pursuing an education in the fastest way possible? Company internship (virksomhedspraktik), (inaudible), improving qualifications; all kind of efforts which are stated in the LAB legislation (Legislation on employment efforts).

TG: Okay.

MH: That is the task of the project. However, in the work we have put, this autumn, we obtained various insights to other things. We are therefore examining the different components and insights further in order to assess where we think we can achieve the biggest value. What is realistic and can be achieved within the scope of the project, in terms of our data foundation, and the architecture of our solutions? We will reach a final decision in Week 9 ...

TG: . . . Okay

MH: . . . what kind of algorithm will be the one to work with in this development?

TG: If I understand you correctly, what you are doing is, that you are trying to find a tool or an algorithm which can help a citizen back in the labor market as quick as possible. Instead of developing an algorithm which reveals which citizen is in the risk of being long-term unemployed? Which is actually what is being used today? Do you agree on this?

MH: Yes.

TG: Okay. Why do you choose one tool rather than the other? They are more or less the same? Is it because you think that we are working with this, but we need the other as a supplementary tool, or is it going to replace the other?

MH: It is not going to replace the other. We also work with the profile clarification tool. The way we work with it in the organization; when we receive a reply from a given citizen we use it as a basis for a dialogue. We do not assign a high value to the results; however it is a tool for dialogue. How do you see yourself in terms of risk? And that is what is most decisive for the citizen; his or her own self-perception. So we use the profile clarification tool and we will continue using it as it is.

TG: Okay, okay.

MH: The focus of our project is to develop solutions which can give the unemployed choices of action. That is giving the unemployed an active role in their own case; provide them with some opportunities to figure out what he/she can do in the time between each interview. Some of the insights we obtained concerned, among others, what the unemployed do between the interviews in the Jobcenter, and what they do varies significantly. So we would like to figure out how we can help the unemployed to be more active. We have a group of unemployed people who can be more resourceful in finding a job than what they are today. The other part is, the employers here; i.e. our consultants, face challenges with keeping an overview of all the available offers and options which exist. The information is available, however it derives from different sources, or e-mails, or positive lists, or. The amount of information sources is rather vast.

TG: Yes.

MH: First of all to obtain an overview of the different options and then assess what could be relevant for the citizen I am dealing with. There is a vast amount of data that artificial intelligence can provide us with; something that a consultant would not be able to. That is why we decided to take that road.

TG: Okay. If we look at the profile clarification tool as it exists today. The tool gives you the result of a risk or the lack of risk. This is based on register data and a questionnaire as well. Thus, it is based on information we can acquire from Statistics Denmark (Danmarks Statistik) and information which is obtained from a person voluntarily or involuntarily. I could imagine that your tool is based on the same or would it be exclusively based on register data?

MH: We hope that the tool that we are going to develop will consist of several algorithms. Copenhagen Municipality had a project which aimed at answering the same questions which we would like to answer here. What kind of effort will it take for a citizen to find a job or pursuing an education? Their algorithm was based on especially the Jobnet CVs and other register data concerning the unemployed. What made them give up, that they had to stop and could not carry further on, are the categorization of the different efforts, the different existing offers.

TG: Okay.

MH: If we have la ook through DFDG, we can only see in a general level if it is improving qualifications or company internship (virksomhedspraktik). However the suggestions are not further specified to the unemployed. E.g. for the jobparate individuals an internship can be beneficial, we know what is beneficial for the different target groups. However we would like to go deeper and be able to offer a more specific type of internship, which field and with which purpose. This requires more information. First of all what kind of challenges is the citizen faced with? This can be obtained from register data and from questions posed to the citizen; this also includes what we already have from the profile clarification tool, the citizen's own self-conception. This can have an impact, if the citizen's own self-conception is that 'I have a difficulties finding a job', then an internship where some feeling of success could be a good start. However, if your self-conception is that 'I will get a job soon' then a course in writing a CV would be the push, and would be the solution for you. This is one part of it. However in order to find the matching offers, we need to categorize the different offers. We need to figure out how we get an overview of the content to each offer. Each municipality has its own methods. I expect that we will look in to what is in the systems (fagsystemer) in terms of a catalogue for offers, and what kind of descriptions have we provided to each. Moreover, the contracts we have settled with the suppliers in terms of what kind of service they offer, as well as the goals which the citizen can achieve through a given offer. What have we described to be the purpose which the citizen will achieve in taking a given offer? From these we can build up an algorithm which can categorize or classify some offers. This will be our first step…

TG: … sounds like a good idea. Yes.

MH: …so they can move on. Yes.

TG: Yes. If you move towards that direction, the first thing that comes to mind is that there are some people which have some challenges which are not so typical. There are some who might be discriminated for instance because of their ethnicity. We have cases where people omit a personal photo in their CV, because their ethnic identity would not be revealed, some omit their age, some omit their civil status, some omit their gender, some may go all the way and omit their middle name if they think it will help them. They may have more challenges because of these realities. Have you taken these issues in to consideration?

MH: We had a first hearing with the city council. The political angle from there is, as long as data is available to benefit the citizen, and as long as it is used in the algorithm to help the citizen closer, because of its high explanatory power; i.e. the data helps in specifying, then the data should be used. When we look further in to the different data, you mentioned for instance origin, should we add this or not? We have a committee which is interested in including this kind of information but we have not discussed specifics yet. That would be done in the next phase when the algorithm has been selected, and we have looked into the actual variables. From there we would take an ethical and judicial discussion about certain variables, what to include and not, and should that be taken up to a political level; or is it a decision that can be taken by the project.

TG: You might have heard that STAR has been in the media because they included origin…

MH: …Yes.

TG: … and when I call the different Jobcentres, among others Odense Jobcenter. I have also been in touch with Frederiksberg and Copenhagen and others, some of the response is 'we are dealing with positive discrimination, because it helps the citizen. So, if I discriminate a person because of his or her origin, in that a person is given more interviews with a case worker, it is considered positive discrimination because I am trying to help the citizen.' The Danish Institute for Human Rights (Institut for Menneskerettigheder) says 'no, that is negative discrimination.' When we look at the tool from a statistical point of view, undoubtedly origin is very important. It is actually the second most important variable in the tool. As you mentioned, you can look at this from different angles. You can look at it from a judicial angle, and you can look at it from a social angle, etc. Let us say that we agree to leave the origin question out. We do not want to end up like STAR. Let us pretend you had that in your mind, we remove it. My job is also to look at possible future discriminating factors, it is not enough to look at origin. We also have to look at gender e.g., male, female was all quite simple before. There are some people who say today 'I am binary; do not put me in a box of male or female.' Fine, I construct a tool, where I remove origin, and I remove gender. Then some would say 'how about age? there are a lot of people who would consider me senior etc., can you remove that?' I remove that as well. Then there is a case worker that point to the residence address factor. Actually, there was one from Odense who stated that address of residence can also be a discriminating factor. Suddenly we have a tool which loses much of its significance and explanatory power, because we want to be large. The reason why STAR was caught in that situation is because they actually want to be transparent. They revealed to the citizen which variables were actually used to categorize them. One of the solutions could be, and I am just sparing with you, and do not take it as a judgment on what direction you are taking, to conceal it? The more details you provide, the more for others to take presumptions. In Sweden they use the exact same, but they chose not to be transparent. 'We do exactly the same as they do it in Denmark, but we do not show it to the citizen.' The question is, now being in your shoe, I tell them that 'I am actually trying to help the citizen, this is positive discrimination. Based on this data you look very much like this person who has this origin, who has this age, who has this gender, who lives in this municipality, and you know what helped them? Company internship (virksomhedspraktik). So I am actually trying to help you.' When you provide this suggestion to the citizen, i.e. company internship is the way forward for you, maybe to hide how you have obtained that result?

MH: I think it is a very hard question . . .

TG: . . . It is, and there is no right answer, believe me. It is not to judge you or anything. I am just throwing this on the table.

MH: Yes. I think it goes down to how we use the algorithm. As we actually are looking for suggestions to efforts, it will give a result of consultant (inaudible). It would be something that the consultant would look in a system (fagsystem) or depends on how we choose to make it accessible for our consultants. In Odense we value freedom of method highly. You are able to have a dialogue with the citizen based on the results. The explanatory power, what are the exact variables, and how important they are, can be of less significance, because it is the dialogue and support in the decision which are the result of it. The consultants can include these in the dialogue with the citizen. Could this be the way forward because we can see that it has helped others? I think if you on the other hand expose the results directly to the citizen, and the citizen has to interpret the results on their own, as we do with the profile clarification tool. You can have a citizen who might get a result that states that they are in the risk of long-term unemployment. Therefore, I think that there is a need to explain the citizen the reasons behind the results. 'I cannot just call my consultant. I am not explained the results, and I have to interpret it myself and processes it in one way or the other.' However, what is right or wrong, I think there are as many answers as there are people you ask.

TG: Indeed. Some of the considerations that have been made in STAR, is to build a tool which one hoped to be so easy that a case worker also could understand the idea behind the tool. The Black Box model in machine learning has been discussed. (inaudible) however you have no idea how they reached the results. The profile clarification tool has been built as a decision tree. I do not know if you are aware of this? It is a decision tree, it is quite simple. E.g. either you are above 30 or below 30 and then you are placed in a specific category etc. In my interviews with some of the case workers, I ask them 'do you know how this tool works?' They have blind faith to the tool and no idea what is behind the tool. The question is now, if the case worker does not understand what the idea behind the tool is, and that is quite fair. It is not expected for the case worker to take a course in machine learning, however they might have been introduced to it. My impression is that they are not offered any courses concerning the tool, and they do not know what the idea behind the tool is, however they have faith to the tool. It has worked before, and still works today. The question is now, machine learning tools have been advanced with time. I can make tools that implement other tools; tools which are rather advanced, and those who have implemented the tools do not know what the thought behind the tool is when they have taken the decision to implement it. Have you had any considerations about which tools to use in the world of machine learning? Perhaps you had some thoughts? It has to be easy for everyone to understand it? Or was the thought 'let us go all the way where it takes us, let us try a number of tools?' Did you take some of these considerations?

MH: No, we have not discussed this directly; where to balance it. My impression is from the project owners and others involved, generally you have to have a rather high explanatory power to reach the necessary results. We as an organization are not mature enough to trust something which has a big Black Box. In our organization we focus rather on work, data and knowledge (inaudible). From a strategic point of view, we would like to take the results from an algorithm and convert them to strategic knowledge. This entails that there has to be some kind of explanatory power. What is actually decisive for the results which have been reached?

TG: Yes, yes.

MH: However, we have not discussed where we should balance specifically.

TG: Okay, okay. Fair enough. Let us go back to something you mentioned before. Copenhagen Municipality did something similar to what you actually are trying to implement. Can you clarify why they did not succeed? Or why they choose to stop it?

MH: It was because of the issue of offers. They could not get close enough to reach the answer of what offer would work for the given unemployed. They either reached results on a general level 'counseling/improving your qualifications' which actually comprises many things. However not specific enough for a proper recommendation. Or they could offer very specific recommendations; i.e. this supplier in this and this interval. They had these two outliers of data. . .

TG: . . . Okay.

MH: . . . that is why they could not carry on. They reached some of the way with the algorithm, i.e. to obtain some information on the citizen, and the variables which are decisive for the given citizen. However they lacked knowledge on the different offers. So if we could create knowledge about the offers. With these two factors, I think we can reach far.

TG: Okay. My last question. You have not personally worked as a case worker and dealt directly with people in concern. However, have you heard anything about the profile clarification tool as it exists today? Perhaps something positive, perhaps something negative? My impression is that you would like to proceed with using it? My impression is not that you want to stop using it soon, or that it is going to be replaced with your tool? Rather that it is going to be used as a supplementary tool? Have you formed any specific impression of this tool from what you hear from others?

MH: When I talk to the operations department, they state that the profile clarification tool is a rather good dialogue tool. When they get the results from the questions which the citizen has answered, then it is a good tool to use for a dialogue. The results which come out of the tool itself are less useful.

TG: They are less useful, okay. That makes good sense. It is a wake-up-call, however it does not tell you what step to take? And here you come in to the picture?

MH: Yes. As you mentioned, it is based on a decision tree, however based on a simple algorithm. Thus, the tool cannot provide us with any further information which the consultant do not already know. It does not provide new information.

TG: Okay, that is a good point. Mie do you have anything further to add? Or do you have any questions?

MH: I am curious, what are you going to use this for?

TG: Well my role is to redevelop the tool, the profile clarification tool. I am going to code them. We would like to see if we can optimize it. We are using the decision tree today. The result is so and so, what if we work with a more advanced tool which we know of today? Would we be able to obtain better results? Or would it be on the expense of something else which is lost on the way? We would like to experiment with different factors, potential discriminating factors. Can we remove them? What effect will it have on the results from the tool, the explanatory volume? So my function is on the one hand the coding part, and how that can be optimized, and on the other, reach out to case workers to understand how they are using it. There is no doubt that it is easy to dive in the coding part and develop something which you think is rather cool and relevant. However, it will not to be useful for others, and that is not good. I have also been in touch with municipalities which do not use the tool, and I am trying to understand the reason behind their choice. Some have developed their own. Some said that they did not know that it exists although it is several years old. So now they know. So it is interesting to get an impression of where it has been used, to what degree, and why it has been opted out etc. And as I understand some are trying to develop a new tool, however that was not the case with you. You are actually developing an extension to the tool. That is quite relevant. I would like to present that to my management group. They would also like to know how it is done in other countries. We are comparing our results with for instance Australia, Germany, Ireland, Switzerland etc. They have old tools which they have used since the 90's. We are looking in to what the differences are between us and them, performance wise. There are for instance some variables which we do not use. E.g. criminal record is used in other countries; that is not something we use today. We could consider adding criminal record. How about disabilities? We have a question which involves disabilities but we have not included it directly as a health issue. Through my thesis, these questions will be discussed; to look at other countries and see if we can learn from them. Simultaneously, of course reach out to you and others who would like to help. We could cooperate. We learned that Odense want to make a test phase, where a specific thing is made in Odense, that would be an idea to try to test together, perhaps a profile clarification tool which is adjusted. It is not my decision, just some ideas.

MH: Yes. That is quite interesting. We are in a dialogue with some of your colleagues. The offer issue is one component which we would like to work on. We are also very interested in when we get to the components which are citizen(inaudible). How do we present it to the citizen? We would like to use Jobnet which is already the platform which they use today. We have also looked in to other tools as JobMatchIt. As regards to Jobmatch and field shift and outliers who need something else from what they are educated in and worked with. Then it would be interesting, as you mentioned, to include some other variables. More personal sensitive, such as leisure time activities and other things. We have been in dialogue with your colleague Anne. When you obtain your information to the profile clarification tool, is there a possibility to obtain information simultaneously to other algorithms? As the citizen is already providing information to the given questions? You have the authority to obtain the information in the profile clarification tool. Can we use this authority? I think it is rather interesting. We are very curious, on how you work with artificial intelligence.

TG: Yes.

MH: Our ambition in Odense is not only to benefit Odense. We would like to share our results with as many as possible. That is why we have a trail in our Signature Project which concerns a common AE platform for all municipalities. We also have KL(inaudible) involved, which can ask when we have developed an algorithm, can we set up a platform which is shareable? As we work in the area of employment, we would like to know what is our common denominator. We have it in STAR...

TG: ...exactly...

MH: ... and Jobnet and the tools you have otherwise. So if you need any knowledge and cooperation, we are very interested in sharing.

TG: I am very pleased to hear that. Anne, as you mentioned earlier, she introduced me to your Signature Project. There are some other parties involved. It is something that we are working with at the moment, having discussions and several meetings here and there in Skype. In a few weeks hopefully we reach some results, and it might lead to future cooperation where we share knowledge with each other. Mie, thank you very much for your time. It is much appreciated to talk these things through with you. And if I can be of any help, then e-mail me, and we will take it from there.

MH: Thank you, have a nice day.

TG: I will, bye.

APPENDIX C

Computer Code