outerbottom

FREIE UNIVERSITÄT BERLIN

MASTER THESIS

Predicting German's Attitudes towards Refugees

Author:
Rahaf Gharz Addien

Supervisor:
Prof. Dr. David Richter

Examiner:
Prof. Dr. Katinka Wolter

February 14, 2023

Declaration of Authorship

I, Rahaf Gharz Addien, declare that this thesis titled, "Predicting German's Attitudes towards Refugees" and the work presented in it are my own. I confirm that:

- This work was done wholly or mainly while in candidature for a research degree at this University.
- Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated.
- Where I have consulted the published work of others, this is always clearly attributed.
- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work.
- I have acknowledged all main sources of help.
- Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.

Signed:

Date:

FREIE UNIVERSITÄT BERLIN

Abstract

Master of Data Science

Predicting German's Attitudes towards Refugees

by Rahaf Gharz Addien

This study detects broad relationships between attitudes toward refugees of 6795 participants living in Germany and various variables focusing on the influence of the big five personality traits on predicting these attitudes. The results revealed a dominant influence of the perceived cultural threats on shaping attitudes toward refugees. The big five personality traits influenced attitudes toward refugees, and their impact differed based on attitudes strength, whereby the big five personality traits had a notable effect on predicting hostile attitudes.
The study relies on statistical methods of correlation analysis and ordinary least squares regression analysis to examine linear relationships. To explore further possible non-linear relations, the efficient XGBoost model was trained and fed into the tree explainer of Shap to determine feature importance for predicting attitudes based on attitudes strength in three classes hostile, moderate, and positive.

Acknowledgements

As I write the final words, I am grateful for the great opportunity offered by the Free University of Berlin to engage in the interdisciplinary path of data science in the social sciences. It enabled me to combine my mathematics background, knowledge, and interest in applying machine learning in social sciences and studying its ethical aspects.

I would like to express my sincere gratitude to my supervisor Prof. Dr. David Richter, for his efforts in monitoring this work and providing courses on fundamentals and principles dealing with data in Psychology. To Prof. Dr. Katinka Wolter for her recommendations and guidance to data science students, especially in the difficult time of the lockdown.

This work is inspired by the world-weariness and people who experienced different hardships and traumas. Nevertheless, they still open their hearts to alleviate other people's suffering and struggle to make the world a better place for us humans regardless of color, ethnicity, or gender.

I am thankful to my partner and friends for their ongoing support, patience, and interesting discussions. I am also thanking my parents and friends in Syria and hope that achieving this master's degree would give hope to many people where wars kill dreams and potential.

Contents

List of Figures

List of Tables

Dedicated to my parents, Amr, and Helena

the breezes that kept my fire on ...

to Kefah always
my home and my peaceful universe ...

Chapter 1

Introduction

1.1 Overview

A growing body of literature focuses recently on studying attitudes toward refugees coinciding with the global refugee crisis. "Attitudes toward refugees" is a hot important social and political topic whose consequences are reflected in several life domains beginning in the economy and political life and do not end with the co-existence of various social groups in the hosted countries. As attitudes result from complex cognitive and affective processes that affect behavior mutually, understanding factors influencing attitudes toward refugees is essential for research on attitudes in Psychology, Sociology, and political sciences. Furthermore, policies dealing with this issue affect the well-being and security of both the citizens and the newcomers. Therefore, the findings of such studies are worthy of being considered by decision-makers.

Using a large sample from the German Socio-economic Panel (SOEP), this study aims to explore German [1] attitudes toward refugees and detect the impact of different variables on shaping these attitudes. In addition to the well-being and worries about economic and security, the study focuses mainly on the big five personality traits as predictors of attitudes.

On the other hand, the study employed advanced machine learning techniques of Extreme Gradient Boosting (XGBoost) and SHapley Additive exPlanations (Shap)

[1] The study utilizes the term "German attitudes" because the respondents live in Germany and 86% of them have German citizenship.

to provide information on variables contributing to predicting attitudes categorized into three classes presenting negative, moderate, and positive attitudes. The results of Shap were then discussed in light of those resulting from the usually used statistical models.

1.2 Background

This section explains the basic concepts used in this study and draws the headlines of the broached problem. First, it highlights the importance of research on attitudes toward refugees in light of the refugee crisis before presenting the concept of attitude, its definitions, and components. Next, it briefly introduces the integrated threat theory (ITT) and the big five personality traits to clarify the link between attitudes and the examined predictor variables.

1.2.1 The refugee crisis

Every minute of every day during 2015, 24 people worldwide were forced to leave their homes due to civil war, generalized violence, conflict, persecution, or human rights violations. By the end of 2015, overall (63.9 million) people had been forcibly displaced. The majority (11.7 million) of the protection seekers were Syrian, and more than (2 million) were Afghans, Colombians, Congolese, Iraqis, Nigerians, Somalis, Sudanese, South Sudanese, and Yemenis. Turkey hosted the largest number of refugees (2.5 million), mostly from Syria and Iraq (UNHCR, 2016).

A convention refugee is defined as "a person who is outside his or her country of nationality or habitual residence; has a well-founded fear of being persecuted because of his or her race, religion, nationality, membership of a particular social group or political opinion; and is unable or unwilling to avail him- or herself of the protection of that country, or to return there, for fear of persecution" (Lomo, 2012).

This unexpected influx of refugees resulted in many challenges encountering the refugees themselves, the host countries' citizens, and the policymakers, to be described as the refugee crisis. The newcomers usually face difficulties integrating into the new societies and mostly the new cultures. They have to learn a new

language to engage in the job market, which is a challenge, especially for refugees who experienced harsh conditions because of the war and the travel risks to arrive at their destination.

On the other side, the attitudes of the host countries' citizens toward refugees vary between solidary and hostile, as refugees have often been perceived as a threat arousing many concerns about the economy and employment opportunities.
In addition, although hosting refugees from different countries plays an essential role in enriching the host countries' cultural diversity, it evokes at the same time severe worries about the national identity, society's homogeneity, and, most important, security, given the increasing Islamophobia (Foroutan, 2013). These perceived threats lead to discrimination toward the newcomers, which causes adverse social and economic effects on the host country and the refugees.

Between welcoming refugees and controlling the borders strictly, the host countries proposed several policies dealing with the refugee crisis over time, which differed across countries and affected the public opinions on the issue of refugees. Germany was one of the most preferred destinations by refugees, notably after German Chancellor Angela Merkel announced a welcome policy in August of 2015. According to the Federal Ministry of the Interior (BMI), Germany harbored (890 000) refugees in 2015, and the number exceeded a million by the end of 2016. Based on the migration theory as reported by Brücker et al., 2016, refugees favored Germany as a destination due to the pull factors of respect for human rights, the German educational system, and feeling welcome.

1.2.2 Attitudes - Definition and Formation

Studying attitudes in the social sciences began in the 1920s and witnessed an increasing interest in social psychology and sociology in the second half of the 20th century. Attitude research attempts to define the term "attitude" compatibly and precisely in different contexts to explain attitudes formation, measurement, changes, and their relationship to the behavior (Olson and Kendrick, 2008).

Many definitions of attitudes were proposed and developed cross time. According to Merriam Webster's Online Dictionary, the sociological definition of an attitude is "a mental position with regard to a fact or state or a feeling or emotion toward a fact or state". Whereas social psychologists use the term **attitude** for people's favorable or unfavorable evaluations of issues, ideas, actions, people (individual/ collective), or objects, "attitude is a psychological tendency that is expressed by evaluating a particular entity with some degree of favor or disfavor" (Eagly and Chaiken, 1993).
One of the differences between the sociological and psychological approaches is that the sociological approach prioritizes the behavior in studying attitudes aiming at eliminating prejudice and discrimination by changing the behavior. While the psychological approach is concerned with changing attitudes considering behavior is one of many complex components that interact to form an attitude (Chaiklin, 2011).

Research on attitudes formation commonly follows a tripartite model approach which explains attitudes as a multidimensional concept that encapsulates three components: cognition, affect, and behavior (Katz and Stotland, 1959). The cognitive component refers to the mental processes, perception, thoughts, and beliefs. Additionally, it involves knowledge and information about the attitude object. The affective component reflects feelings and emotions towards the attitude object. Finally, the behavioral part indicates actions or intention expressions (the tendency to act). For example, a person's attitude toward refugees can be inferred by determining the beliefs regarding refugees (that could be knowledge based on reports, statistics, or analysis), the feeling toward refugees (such as empathy or fear), and the engagement in practice (like participating in demonstrations for/against refugees, donating, etc.). The attitudes could be explicit or implicit, and many hidden factors could significantly impact their formation.

In the light of the tripartite model, researchers introduced that "An attitude represents an evaluative integration of cognitions and affects experienced concerning an object. Attitudes are the evaluative judgments that integrate and summarize these cognitive/ affective reactions. These evaluative abstractions vary in strength,

which in turn has implications for persistence, resistance, and attitude-behavior consistency" (Crano and Prislin, 2006). Nevertheless, researchers recently are also interested in studying the interaction between attitudes and beliefs, affect, and behavior rather than dealing with them as a part of the attitudes (Albarracín et al., 2005; Olson and Kendrick, 2008).

Attitudes strength is one of the most important elements for assessing and comparing attitudes. Researchers from different perspectives described the properties of a strong attitude. Attributes that identify attitude strength are: accessibility, certainty, affective-cognitive consistency, crystallization, direct experience, importance, stability, generalized attitude strength, intensity, latitude of rejection, and vested interest (Raden, 1985), durability and impact (Krosnick and Petty, 1995), extremity, importance, accessibility, and ambivalence (Bassili, 2008), extremity, certainty, and personal experience (Baron, Byrne, and Branscombe, 2006). In the context of the topic "attitudes towards refugees", strong attitudes are more related to:

- Extremity: degree of passivity or negativity towards the attribute object in terms of an attitude scale.
- Certainty: knowledge about the attitude object.
- Importance: the extent to which the attitude object has priority.
- Personal or direct experience.

1.2.3 Integrated threat theory

The integrated threat theory was proposed first by Stephan and Stephan, 2000, who explained prejudice and negative attitudes of the in-group toward out-groups by four types of factors: realistic threats (physical well-being and economic concerns), symbolic threats (cultural concerns due to differences in cultures and values between the last two groups), negative stereotype (negative stereotypes that the in-group members form about the out-group), and intergroup anxiety (in-group's anxiety of interacting with the out-group) (Stephan and Renfro, 2002).
Many studies examined predictors of negative attitudes toward migrants based on ITT and provided evidence of the solid relationship, especially between negative

attitudes and the perceived threats (Stephan et al., 2005; Talay and De Coninck, 2020).

1.2.4 The big five personality traits

While research on the influence of specific personality traits on prejudice and political attitudes gained growing interest from researchers, few studies tested associations between attitudes toward refugees and personality traits. This section presents the big five personality traits and describes their related characteristics.
Relying on the lexical hypothesis proposed by Francis Galton in the 1800s that "every natural language contains all the personality descriptions that are relevant and important to the speakers of that language", Gordon Allport and Henry Odbert extracted 4500 terms describing personality traits (Vinney, 2018). The big five personality traits model was a result of the contribution of many Psychologists, who depended on factor analysis to confirm that personality traits can be reduced to five, which are Conscientiousness, Agreeableness, Neuroticism, Openness to Experience, and Extraversion (A, 2020).
In other words, the big five model is a descriptive model conceptualizing traits as a spectrum between two extreme ends. Briefly, as reported in (A, 2020), (Vinney, 2018), and (John, Srivastava, et al., 1999), individuals who score high on:

- Conscientiousness (vs. Lack of Direction) have good impulse control and tend to be deliberate, controlled, organized, disciplined, thoughtful, careful, and detail-oriented.
- Agreeableness (vs. Antagonism) have a positive and altruistic orientation toward others, reflecting traits of trusting, forgiving, and sympathetic to others.
- Neuroticism (vs. Emotional Stability) tend to have negative emotions, including anxiety and tension, hostility and irritability, depression, self-consciousness and shyness, and lack of self-confidence.
- Openness (Closedness to Experience) are perceived as creative and artistic, where Openness "describes the breadth, depth, originality, and complexity of an individual's mental and experiential life" (John, Srivastava, et al., 1999).

- Extraversion (vs. Introversion) gain their energy from interacting and being with others, making them usually active, sociable, assertive, and emotionally positive.

1.3 Statement of the problem

Attitudes formation is a complex process that has to be studied carefully based on the specific context of the case study to be covered. Moreover, analyzing the origins of attitudes toward political issues and refugees particularly requires considering the interaction of many social, individual, and even historical factors. Therefore, the findings differ between countries and sometimes between regions of the same country.

This study attempts to introduce a better understanding of German attitudes toward refugees and examine the findings of a few studies on the impact of the big five personality traits on attitudes toward refugees. Moreover, the study tries to enrich insights on the topic utilizing advanced machine learning algorithms of XGBoot and Shap.

1.3.1 Data

The study used a large sample of 6795 respondents from SOEP to detect different relations between German public attitudes toward refugees and a wide range of variables. The extracted dataset is a cross-sectional data collected mainly in 2016 and consists of variables in five categories covering socio-demographics, life satisfaction, concerns, perceived threats, and personality traits. Table 1.1 presents the variables of each category.
SOEP measured attitudes toward refugees by five items:

- Is it generally good or bad for the German economy that refugees are coming here?
- Will refugees erode or enrich cultural life in Germany?
- Will Germany become a better or worse place to live because of the refugees?

- Does a large influx of refugees mean more risks or more opportunities in the short term?
- Does a large influx of refugees mean more risks or more opportunities in the long term?

Will Germany become a better or worse place to live because of the refugees? is the dependent variable of the regression analysis and the target for machine learning models, whereas the variables of "Is it generally good or bad for the German economy that refugees are coming here? " and "Will refugees erode or enrich cultural life in Germany?" represent the perceived economic and cultural threats, respectively De Coninck, d'Haenens, and Joris, 2019.

1.3.2 Research methods

Different statistical and machine learning methods were applied to surface relationships between attitudes and possible predictors.

After exploring the data statistically, correlation analysis between the dependent variable of attitude and the independent variables was conducted, exposing the Pearson correlation coefficients and their significance levels. Next, OLS linear multiple regression was utilized for estimating the coefficients of the linear regression equations describing the associations between the dependent variable and the groups of the independent variables.

Finally, the well-known XGBoost classifier was applied. Nevertheless, the goal of using XGBoost is not the classification task per se but discovering the contribution of each feature to the resulted prediction using the inventive SHAP.
XGBoost is an efficient boosting-based algorithm that combines a set of weak learners to derive a strong model. It creates decision trees sequentially and adjusts weights frequently, achieving high performance, where the misclassified outcomes get higher weights.
Shap is "a game-theoretic approach to explain the output of any machine learning model. It connects optimal credit allocation with local explanations using the Shapley values from game theory and their related extensions" (Lundberg and Lee, 2017). In other words, Shap offers high explainability levels of machine learning models by computing features' importance for each outcome and each class.

Variable Category	Variable Description
Identification	Person ID
	Current wave household ID
	Survey year
Socio-demographic	Age
	Gender
	Federal state
	Migration background
	Amount of education or training (in years)
	Employment Status
	Monthly salary gross amount
Well-being	Current life Satisfaction
	Satisfaction with personal income
Concerns	Worried about economic development
	Worried about finances
	Worried about own retirement pension
	Worried about cohesion in society
	Worried about peace
	Worried about crime in Germany
	Worried about immigration to Germany
	Worried about hostility to foreigners
Personality traits	
conscientiousness	Thorough worker
	Carry out tasks efficiently
	Tend to be lazy
Extraversion	Am communicative
	Am sociable
	Reserved
Agreeableness	Able to forgive
	Friendly with others
	Am sometimes too coarse with others
Openness	Am original
	Value artistic experiences
	Have lively imagination
Neuroticism	Worry a lot
	Somewhat nervous
	Deal well with stress
Perceived threats	
Economic threat	Is it generally good or bad for the German economy that refugees are coming here?
Cultural threat	Will refugees erode or enrich cultural life in Germany?

TABLE 1.1: Overview of the used independent variables.

1.3.3 Research questions

- Which variables are more likely to predict German attitudes toward refugees?
- Which effect do the big five personality traits have on attitudes toward refugees?
- How informative can the Shap explainer be for evaluating feature importance in social studies?

1.4 Objectives and hypotheses

Hypotheses to be tested are driven by findings of other studies on attitudes toward refugees and factors influencing them.

Regarding the socio-demographic variables, educational levels and political orientation have been exposed as the most prominent predictors of attitudes toward refugees, whereas findings on age, gender, and income vary between studies (Crawley, 2009; Dempster and Hargrave, 2017; Cowling, Anderson, and Ferguson, 2019). Remarkably, differences in prejudicial tendencies between East and West Germany were observed (Gehrsitz and Ungerer, 2017, Sola, 2018). Unfortunately, political orientation is excluded from the analysis because the data does not offer enough informative items in this regard.

H_1 : Holding less negative attitudes is more likely to be related to being young, female, having a high educational level, having a migration background, and living in West Germany.

On the other hand, preserved threats and concerns reported high associations with attitudes. As individuals scoring high levels of worry about security and economic development tend to feel threatened by the newcomers and hold negative attitudes as a result (Crawley, 2009; Dempster and Hargrave, 2017). Well-being has been rarely tested relating to attitudes; therefore, no clear hypothesis in this regard can be made.

H_2 : Holding negative attitudes is related to high levels of worry about the economics and security in Germany.

H_3 : Holding negative attitudes is related to perceiving refugees as a cultural and economic threat.

Moreover, scholars on personality traits and prejudicial attitude (Sibley and Duckitt, 2008), particularly, toward refugee (Talay and De Coninck, 2020; Carlson et al., 2019; Pruysers, 2020) found traits of Openness, Agreeableness, and Conscientiousness to be related to attitudes.

H_4 : Holding less negative attitudes is positively related to reporting high scores in Openness and Agreeableness and negatively to scoring high in Consciousness.

1.5 Structure of the study

The study consists of five chapters. After introducing the research problem, its importance, and the key terms, the second chapter is a literature review on attitudes toward migrations and refugees, considering a wide variety of factors that have been studied as predictors of attitudes. The third chapter presents the data in detail, the measures, data prepossessing steps, and the exploratory data analysis. The algorithms of the research methods, the analysis, and the results are in chapter four, ending with the discussion, limitations, and conclusion in the last chapter.

Chapter 2

Literature Review

This chapter reviews studies basically about attitudes toward refugees and migrants, focusing on factors that could predict these attitudes, especially after the refugee crisis. It is essential to point out that studies use different questions in their surveys to measure attitudes toward refugees. In addition, many studies mentioned that respondents usually do not differentiate between migrants, asylum-seekers, and refugees. Therefore, I tried to highlight the most relevant insights to the research questions.
In light of the complexity of factors influencing attitudes toward refugees, I chose studies from Germany and other countries to enable observing the role of the contextual effect.

Public attitudes toward refugees

In the recent decade, many countries worldwide have encountered the refugee crisis as a critical challenge that evokes several public concerns about national identity, culture, economy, employment opportunities, terrorism, and integration of the newcomers. That affected general policies, especially in Europe, and opened controversial debates on how to deal ideally with this crisis. Furthermore, it gives researchers plenty of data to study public attitudes toward refugees. However, the research on attitudes toward migrants has a long history, mainly in countries with high diversity. The following two studies offer an assistant framework for understanding the factors shaping these attitudes.

Motivated by the increasingly British hostile attitudes toward immigrants generally and asylum-seekers in particular, Crawley, 2009 drew a guide for researchers based on the statistics and findings of several studies. She proposed that factors influencing attitudes toward migrants are strongly interconnected and multifaceted, namely demographic, economic, political, social, cultural, psychological, and even geographical factors. In addition, she highlighted the challenge of deducing any causal relationships among these associations.
Generally, studies exposed that positive attitudes are more likely to be related to women, young people, highly educated people, and those who have ties with the newcomers (people with migration backgrounds or who belong to ethnic minorities). Nevertheless, these results are not generalizable since they vary based on the context.

On the other side, individuals' attitudes are not only affected by their characteristics and personal experiences. Rather social networks have an essential role in creating social norms that impact the cognitive processes of attitudes formation.
Moreover, researchers should consider factors such as the received information credibility on the topic and how policy-makers and media introduce and relate the topic to the public threats. She disclosed why the newcomers left their countries as an influential factor where people in the UK had less negative attitudes toward economic reasons immigrants, especially the high-skilled, compared to the asylum-seekers and refugees who had been considered illegal in the country.

Cooperating with experts on the topic and based on 160 selected studies from a wide range of cross-disciplinary literature, Dempster and Hargrave, 2017 attempted to understand the complexity of public attitudes toward refugees by engaging the real-world concerns, emotions, and values. Aiming at enriching the insights, they recommend considering factors beyond the demographic. Their approach emphasizes the importance of rooting attitudes toward refugees in the national and regional contexts.
The study covered studies from many countries (the UK, some European countries, the USA, Canada, Australia, Turkey, etc.). Overall, findings on the sociodemographic variables of age, gender, and education are consistent with Crawley,

2009. Additionally, positive attitudes toward refugees and migrants are less common among politically liberal.

Notably, European attitudes toward refugees are ambivalent compared to the severely hostile attitudes in Turkey, with remarkable differences based on the origin countries of refugees, their religious or ethnic groups, and their educational level. Attitudes in the same country varied regionally according to the history of each region with migrants, its ethnic diversity, and the hosted number of refugees. Whereby people reported less negative attitudes toward refugees in urban areas in the UK compared to other regions (Crawley, Drinkwater, and Kauser, 2013), in West Germany compared to the east of the country (Gehrsitz and Ungerer, 2017), and in Turkish cities with (10%) and above refugee population (Keleş et al., 2016). The real-world concerns about security, culture, and economics affect attitudes toward refugees strongly. People who have economic concerns or suspect integrating the refugees successfully into society are more likely to hold negative attitudes. Further, there is plenty of evidence that people relate hosting refugees from specific groups to terrorism.

2.1 Socio-demographic and ideological factors

This section focuses on the findings of the correlation analysis between attitudes and different socio-demographic and ideological variables conducted on global data. It also broaches examples of some studies that concluded contradictory results on associations between attitudes and socio-demographic factors in other contexts.

Cowling, Anderson, and Ferguson, 2019 extended the meta-analysis by Anderson and Ferguson, 2018 who reviewed 32 studies on the correlations between demographic and ideological factors and negative attitudes toward asylum seekers in Australia to conduct a review of 70 global studies.
The results revealed that negative attitudes have a small significant negative correlation with the level of education and small significant positive correlations with being male, politically conservative, religious, and nationally identified. Contrary

to the previous studies, they found no relationship between negative attitudes and age, but consistently with the previous studies strongly, negative attitudes were strongly correlated with perceiving refugees as a threat.
The influence of social justice beliefs was also important as people who believe that the resources are equally distributed in fair societies had more likely positive attitudes toward asylum seekers. Furthermore, the study surfaced the essential impact of refugee misrepresentation in media and policies that aim to reduce the number of refugees and control the border strictly.

Elsewhere, Bansak, Hainmueller, and Hangartner, 2016 conducted a conjoint experiment requesting (18000) Europeans from fifteen European countries for assessing (180000) profiles of asylum-seekers. Humanitarian concerns invoked the voters to accept asylum claims of people who fled war or experienced severe vulnerabilities compared to economic migrants. Though, anti-Muslim sentiment affected how the voters evaluated the requests, reflecting security concerns.
The voters also considered employability as asylum-seekers who have high education and language skills were preferred over the others, reflecting public concerns about the whole country's economy. However, the analysis discovered no differences between the fifteen countries nor any role of voter demographic variables of gender, age, education, political orientation, or income.

Noteworthy, findings of studies on attitudes in Australia, the USA, and Europe seem to be more compatible than results from studies on attitudes in Turkey, where the extreme negativity of Turkish attitudes prevents comparing the influence of the previous factors. Keleş et al., 2016 tested the relationship between youth´s attitudes toward Syrian refugees in Turkey and the demographic variables using a sample of 638 young adults (age range: 18-35, M = 25.6, SD= 4.6). The analysis showed no significant relationships between negative attitudes and gender or income, while postgraduate participants reported significantly less negative attitudes toward refugees than high school or undergraduate participants.
In this regard, Yelpaze and Güler, 2018 examined the influence of the education type on the attitudes of 340 students from seven different faculties and found that students of Theology had positive attitudes compared to students from the other

faculties. Alici, 2021 attempted to examine the impact of cultural sensitivity on attitudes toward the Syrian refugees in a sample of nursing students (N = 412). Although the mean age was 20.48 years (SD = 1.65), with most females (80%), most students held unfavorable attitudes.

The contextual factors mentioned before can explain some variations in the results to some extent, but also, it might be due to differences in the chosen samples and surveys.

2.2 Individuals' characteristics

Many studies focused on different demographic and ideological factors, concerns, and perceived threats. Otherwise, few studies handled other societal concerns or went beyond that to broach the well-being variables.

Albada, Hansen, and Otten, 2021 presented a well-designed three-dimensional framework for understanding polarization of attitudes toward refugees in the Netherlands. These dimensions are the individual and social self (income, education, political orientation, and relative deprivation), perceptions and experience regarding refugees and migrants (perceived cultural distance, frequent contact), and the societal context (societal discontent).

In line with the previous studies, the researchers concluded that most Dutch (63%) have moderate attitudes toward refugees, whereby negative attitudes had significant associations with less educational levels and political right-orientation. Over and above, negative attitudes were more likely common among people who perceived more cultural distance and experienced more relative deprivation and societal discontent. Relative deprivation refers to a personal judgment or experience when people think that they or their groups receive fewer benefits than others. Whereas social discontent or social unease is "a latent concern among citizens that fundamental aspects of society are unmanageably deteriorating" (Steenvoorden,

2015). In other words, social discontent reflects public concerns regardless of personal issues.

Renner et al., 2018 tested life satisfaction and optimism, the perceived economic threat, the perceived cultural threat, and modern racism as predictors of negative attitudes toward asylum seekers in two samples (Austrians and Germans: N = 349, and Slovak: N = 307). The analysis discovered that life satisfaction predicted the perceived threats and modern racism negligibly while Optimism did not. Whereas the perceived cultural threat was the only significant predictor of negative attitudes toward asylum seekers in both samples, the negativity in the German sample was lower and modern racism was a better predictor than in the Slovakian one.

2.3 Personality traits

The interest in personality traits as predictors of political attitudes began recently, especially in political sciences and sociology. Related to this study are two research directions. First, since the literature on attitudes toward refugees shows political orientation as a significant predictor, it will be informative to study the relationship between attitudes, political orientation, and personality traits. Second is the influence of personality traits on prejudicial attitudes.
However, studies on prejudice are primarily interested in attitudes toward immigrants, not toward refugees. The findings differ between the two cases as a refugee is forced to displace, and reasons for fleeing or migration were found to be contributed significantly to shaping attitudes.

In their meta-analysis of 71 studies on personality traits and prejudice, Sibley and Duckitt, 2008 documented that conservatism is associated with low Openness and high Conscientiousness. At the same time, low Openness and low Agreeableness could predict prejudice. The results vary between studies based on the personality traits' chosen model; for instance, studies that used the BFI (Big Five Inventory) reported lower associations than studies that used other models.

Findings regarding personality traits and attitudes toward migrations were sometimes inconsistent with the meta-analysis results. For example, Gallego and Pardos-Prado, 2014 disclosed that positive attitudes in a Dutch sample are associated with Agreeableness but not with Openness, while negative attitudes are moderately associated with Neuroticism followed by Conscientiousness. They conclude that the socio-demographic variables are more effective predictors of attitudes toward migrants.

More specifically, I am aware of three studies concerning attitudes toward refugees and personality traits; nevertheless, none of them studied a German sample.
Talay and De Coninck, 2020 focused on the relationship between personality traits and European attitudes toward refugees using data from Belgium, Sweden, France, and the Netherlands in 2017. The results revealed a high significant but weak positive correlation with Openness and Agreeableness ($r = .21$, $r = .17$, respectively), a significant negative correlation with Neuroticism ($r = -.05$), and no significant correlation with Extraversion and Conscientiousness. Consequently, attitudes toward refugees were found positively and strongly associated with Openness and Agreeableness, and negatively with Extraversion and Conscientiousness. While no evidence of the influence of Neuroticism on attitudes toward refugees since it was associated with positive attitudes but correlated negatively with them.
The findings of Talay and De Coninck, 2020 came to support the findings of Carlson et al., 2019 on American attitudes toward the Syrian refugees.

Pruysers, 2020 examined the associations between prejudicial attitudes toward refugees and different personality traits, including the big five and the dark triad (narcissism, Machiavellianism, and psychopathy) in a Canadian sample.
In this study, negative attitudes toward refugees were positively associated with the traits of Conscientiousness, Machiavellianism, narcissism, and psychopathy; negatively with Openness and honesty-humility; whereas no relations between attitudes and Agreeableness and Neuroticism were exposed.

On the other hand, many studies are interested in the influence of personality traits on being liberal or conservative, often regardless of the contextual factors or

variations across issue domains that could remarkably affect the results.
For instance, Gerber et al., 2010 utilizing a large national sample in the USA revealed positive associations between (Openness and socially liberal attitudes), (Neuroticism and economically liberal attitudes), and (Conscientiousness and socially conservative attitudes). Moreover, the relationship between Extraversion and Agreeableness and political attitudes were either weak or insignificant.

Other studies argued against relating personality to political attitudes, given that no strong associations could be proven. Hatemi and Verhulst, 2015 examined the causality between personality traits and political attitudes by observing the development of personality traits and political attitudes over ten years to reject the causality assumption between them. Nevertheless, they uncovered that personality traits and political attitudes develop parallel, manifesting the same latent genetic influence. Interestingly, though Openness to experiences was higher in the adolescents´ sample, it was not a meaningful predictor of liberal attitudes.

2.4 German attitudes toward refugees

Plenty of studies on German attitudes toward refugees broached different aspects of the refugee crisis and its consequences but rarely considered individuals' characteristics. Overall, German attitudes are influenced fundamentally by reasons for fleeing and changes in the welcome policy of 2015. Despite the high diversity of the German society, negative attitudes toward migrants even before 2015 are strongly related to public concerns about economic, national identity, and especially Islamophobia after the terrorist attacks of September 2001 (Foroutan, 2013).

According to the SOEP survey of German attitudes in 2016, most respondents (81%) held positive attitudes toward admitting refugees of war and political reasons with preferences based on those refugees' religion and origin country. Notwithstanding, the majority (75%) associated the recent influx of refugees with more risks than opportunities (Gerhards, Hans, and Schupp, 2016). These statistics come in line with many studies.

Meidert and Rapp, 2019 compared German attitudes toward different groups of refugees using data from ALLBUS [1] collected in 2016. Considering motivations for displacement, the respondents reported more negative attitudes toward economic refugees than war refugees. At the same time, almost half agreed that migration should be allowed but restricted. The study disclosed significant associations between negative attitudes and national pride, less education, political right-orientation, unemployment, and high income. While females and people, who have contact with refugees, tended to show more positive attitudes toward war refugees as well as political refugees.

Plener et al., 2017 used different survey items to assess German attitudes (N = 2524) toward the Unaccompanied Refugee Minors (URM). Only (22.8%) of the participants accepted that Germany could host more URM. In line with the above, favorable attitudes are more likely to be related to being young, graduating from high school, having a higher monthly income, and do not have German citizenship. However, these associations differ across questions, e.g., gender effect was apparent when asking if URM should send back to their country home. Also, strong associations between negative attitudes and Islamophobia and supporting right wings were observed in all given questions.

Liebe et al., 2018 explored citizens' acceptance of refugee and migrant homes in their vicinity (N = 418) in November 2015 and 2016. About (80%) of the respondents had negative attitudes with clear preferences for non-Muslims. Furthermore, the attitudes of the approval group were less stable over time, which could be explained by the public debate on the topic and the changes in the government's policy dealing with it. Among many variables, the study asserted only the relationship between positive attitudes and high education and having contact with refugees.

Sola, 2018 estimated the short-term causal effect of the refugee crisis in Germany on public concerns about migration using a large sample of SOEP data. Alessandro uncovered that concerns about migration after the refugee crisis (the

[1] Allgemeine Bevolkerungsumfrage der Sozialwissenschaften

second half of 2015) increased by around (22%) compared to before, more saliently in East Germany. In addition to an inverted U-shaped relationship between age and concerns about migration, the analysis exposed that concerns are negatively correlated with higher educational levels, being single, having a direct migration background, and positively with unemployment and disability. Interestingly, this increase in concerns is significantly correlated with the rising popularity of the German right-wing party Alternative für Deutschland (AFD), but no evidence explains causality.

Chapter 3

Data and descriptive statistics

This chapter presents detailed information on the used dataset, the measures, and the descriptive statistics. It explains the required data preprocessing steps for the correlation analysis and the final models and reports the correlations between the independent variable of attitudes toward refugees and groups of the dependent variables.

3.1 Database and measures

The study utilizes data from the German Socio-economic Panel (SOEP). SOEP is an official representative panel survey that began in 1984 with an original sample size of almost 6,000 households and above 12,000 individual respondents. As a multidisciplinary longitudinal survey, SOEP data enables researchers to study a wide variety of scientific topics (Wagner, Frick, and Schupp, 2007).

The used data is a cross-sectional data from 2016 on attitudes, well-being, concerns, and data from 2017 on the big five personality traits.

3.1.1 Attitudes and the perceived threats

At the beginning of 2016, SOEP conducted a monthly survey on German attitudes toward refugees and concerns about immigration, interviewing more than 20000 respondents. SOEP survey measured attitudes toward refugees also in 2018, presenting the following questions:

- Is it generally good or bad for the German economy that refugees are coming here?

- Will refugees erode or enrich cultural life in Germany?
- Will Germany become a better or worse place to live because of the refugees?
- Does a large influx of refugees mean more risks or more opportunities in the short term?
- Does a large influx of refugees mean more risks or more opportunities in the long term?

The dependent variable /the target concerning attitudes toward refugees is **Will Germany become a better or worse place to live because of the refugees?** while the variables "Is it generally good or bad for the German economy that refugees are coming here?" and "Will refugees erode or enrich cultural life in Germany?" show attitudes in light of the perceived economic and cultural threats, respectively. Respondents could choose from 1 to 11 on the Likert scale (1: a worse place to live, 11: a better place to live), (1: bad for the economy, 11: good for the economy), and (1: erode cultural life, 11: enrich cultural life) (Berlin et al., 2020).

3.1.2 The big five personality traits

SOEP measures the big five personality traits irregularly since 2015 using the short version BFI-S which contains 15 questions (three questions for each trait) that can be completed within two minutes (Gerlitz and Schupp, 2005).

However, this study focuses on how attitudes could be affected by the personality traits that have been measured in 2017. Therefore, data from 2016 was chosen. Respondents selected answers between 1 and 7 on how each statement fits them (1: does not describe me at all, 7: describes me perfectly). The items of each trait are: Conscientiousness (thorough worker, carry out tasks efficiently, tend to be lazy), Extraversion (I am communicative, I am sociable, reserved), Agreeableness (I am sometimes too coarse with others, able to forgive, friendly with others), Openness (I am original, value artistic experiences, have lively imagination), and Neuroticism (worry a lot, somewhat nervous, deal well with stress).

3.1.3 Well-being and concerns

SOEP surveys offer annual data on concerns and life satisfaction, reflecting well-being levels and worries about personal and national issues (Goebel et al., 2019).

The will-being variables present how satisfied the respondents are with their current life and personal income measured on the 11 points Likert scale (0: completely dissatisfied, 10: completely satisfied).

On the other hand, the variables on concerns survey how worried the participants are about the following issues: economic development, finances, own retirement pension, cohesion in society, peace, crime in Germany, immigration to Germany, and hostility to foreigners. Levels of worry are assessed using a three-point scale (1: very concerned, 2: somewhat concerned, 3: not concerned at all).

3.2 Participants

The study sample consists of (N = 6795) participants living in Germany and aging between 18 and 82 years (45.3 ± 11). Among them, 3441 (51 %) were females.
On average, participants had approximately 12.7 years of education, and a monthly income of 2884.43 €, table 3.1.
The sample is imbalanced regarding federal states (West vs. East), employment status, and migration background, where: 5506 (81%) respondents were from West Germany and (19%) from East Germany.
In addition, 4269 (63%) worked full time, 1896 (28%) worked part-time, 378 (almost 6%) were not employed, and 252 (about 4%) had other employment statuses. Regarding having a migration background, 4775 (70%) had no migration background, 1585 (23%) had a direct migration background, and 435 (7%) had an indirect migration background.

Variable label	Mean	Std	Min	Max	Median
Age	45.3	11.0	18.0	82.0	46.0
Amount of education	12.7	2.8	7.0	18.0	12.0
Income	2884.43	2083.0	500.0	43000.0	2458.0
Satisfaction: Current life	7.6	1.5	0.0	10.0	8.0
Satisfaction: personal income	6.9	2.1	0.0	10.0	7.0
Worried: economic development	2.1	0.6	1.0	3.0	2.0
Worried: finances	2.3	0.7	1.0	3.0	2.0
Worried: own retirement pension	2.1	0.7	1.0	3.0	2.0
Worried: cohesion in society	1.8	0.6	1.0	3.0	2.0
Worried: peace	1.5	0.6	1.0	3.0	2.0
Worried: crime in Germany	1.7	0.7	1.0	3.0	2.0
Worried: immigration to Germany	1.8	0.7	1.0	3.0	2.0
Worried: hostility to foreigners	1.6	0.7	1.0	3.0	2.0
Thorough worker	6.2	1.0	1.0	7.0	6.0
Tend to be lazy	2.5	1.6	1.0	7.0	2.0
Carry out tasks efficiently	5.8	1.0	1.0	7.0	6.0
Am communicative	5.6	1.3	1.0	7.0	6.0
Am sociable	5.2	1.4	1.0	7.0	5.0
Reserved	4.0	1.6	1.0	7.0	4.0
Sometimes too coarse with others	3.2	1.7	1.0	7.0	3.0
Able to forgive	5.5	1.3	1.0	7.0	6.0
Friendly with others	5.8	1.0	1.0	7.0	6.0
Am original	4.8	1.3	1.0	7.0	5.0
Value artistic experiences	4.3	1.8	1.0	7.0	4.0
Have lively imagination	4.9	1.5	1.0	7.0	5.0
Worry a lot	4.3	1.7	1.0	7.0	4.0
Somewhat nervous	3.5	1.7	1.0	7.0	3.0
Deal well with stress	4.6	1.5	1.0	7.0	5.0
Economic concerns	2.2	0.5	1.0	3.0	2.0
Security concerns	1.7	0.5	1.0	3.0	1.6
Conscientiousness	5.8	0.9	1.7	7.0	6.0
Extraversion	4.9	1.1	1.0	7.0	5.0
Agreeableness	5.4	1.0	1.7	7.0	5.3
Openness	4.7	1.1	1.0	7.0	4.7
Neuroticism	3.7	1.2	1.0	7.0	3.7
Economic threat	5.7	2.7	1.0	11.0	6.0
Cultural threat	5.8	2.8	1.0	11.0	6.0
Attitude: better place to live	5.2	2.5	1.0	11.0	6.0

TABLE 3.1: Descriptive statistics of all variables of the present study.

3.3 Data preparation

Many data preprocessing steps were required to produce the final dataset to be used for the exploratory analysis, the regression analysis, and the classification. It included combining data from different databases, data cleaning, factor analysis, and computing the composite scales if needed.

3.3.1 Data collecting

SOEP-Core provides data from a comprehensive variety of surveys on the individual level in different databases. Therefore, extracting the data of the above variables requires, first, selecting respondents who participated in the SOEP survey on German attitudes toward refugees in 2016—second, filtering the items of interest by individuals' Ids and the year 2016 from several offered databases and merging them. Finally, adding the items from 2017 on personality traits corresponding to the previous individuals to constitute a dataset of 40 columns and (19006) observations.

3.3.2 Data cleaning

Invalid answers of SOEP data have negative values, indicating different cases of missingness (-1: No answer, do not know, -2: Does not apply, -3: Implausible value, -4: Inadmissible multiple responses, -5: Not included in this version of the questionnaire)(Goebel et al., 2019). The negative values were replaced by "NaN" and removed.
Additionally, respondents who reported a monthly income of fewer than 500 € were excluded from the analysis to result in the final sample of (6795) participants.

3.3.3 Data re-coding

SOEP presents data on the employment status in six categories full-time, regular part-time, not employed, vocational training, marginal, irregular part-time employment, and sheltered workshop. To ease the result's interpretations and as the study is interested in the first three categories that cover most of the respondents,

the last four categories were considered one group of "Other" employment status. Moreover, the big five personality traits items of "Tend to be lazy", "Reserved", "Am sometimes too coarse with others", and "Deal well with stress" were recorded in the direction of the traits Conscientiousness, Extraversion, Agreeableness, and Neuroticism, respectively.

3.3.4 Exploratory factor analysis

The table 1.1 shows eight items measuring worries about different life domains. Exploratory factor analysis was applied to these items. Looking at the eigenvalues and according to the Kaiser criterion, eigenvalues larger than 1 imply that the factor explains more than one item. Two factors can be extracted to represent the eight items. It is also what the screen plot 3.1 suggests. The first factor represents

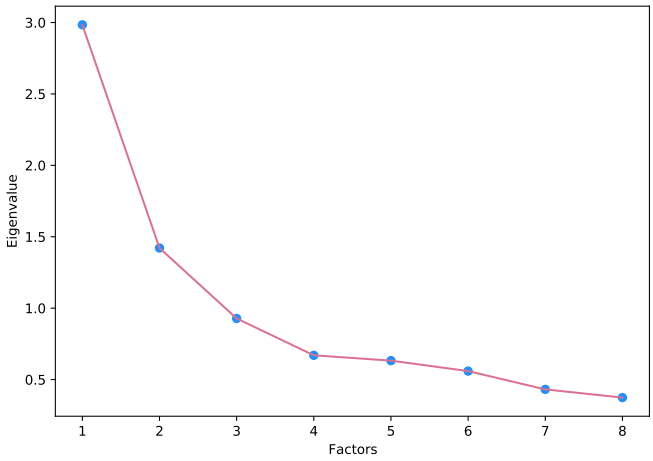

FIGURE 3.1: Scree plot of the factor analysis

the items "Worried about economic development", "Worried about finances", and "Worried about own retirement pension". While the second factor loaded on the items "Worried about cohesion in society", "Worried about peace", "Worried about crime in Germany", "Worried about immigration to Germany", "Worried about hostility to foreigners". Cronbach's α takes values of (.7) and (.73), respectively.

3.3.5 The composite scale

After applying the exploratory factor analysis to the items on worries, the composite scale for each factor was constructed by computing the mean value of the items loaded on this factor. In the rest of the study, "economic concerns" and "security concerns" will refer to these two factors. Similarly, the composite scale of each trait of the big five was computed based on the mean of the three related items.

3.4 Exploratory data analysis

Consistently with studies on attitudes toward refugees in Europe and German, the attitudes in the given sample were moderate. Katwala and Somerville, 2016 suggested evaluating attitudes based on the strength of these attitudes in three levels. Accordingly, (37%) of the respondents hold hostile attitudes rejecting that refugees make Germany a better place to live (four or less on the 1-11 scale), (46%) reported moderate attitudes (between five and seven), and only (17%) had welcoming attitudes (eight or higher).
Furthermore, figure 3.2 shows that respondents from West Germany, compared to East Germany, tended to hold more welcoming attitudes, whereas there were no apparent differences between males‘ and females‘ attitudes. Figure 3.3 illustrates the distributions of attitudes based on the employment status and having a migration background.

3.4.1 Descriptive Statistics

Table 3.1 presents descriptive statistics of variables of interest for this study, including the mean, the standard deviation, the minimum, the maximum, and the medium values.

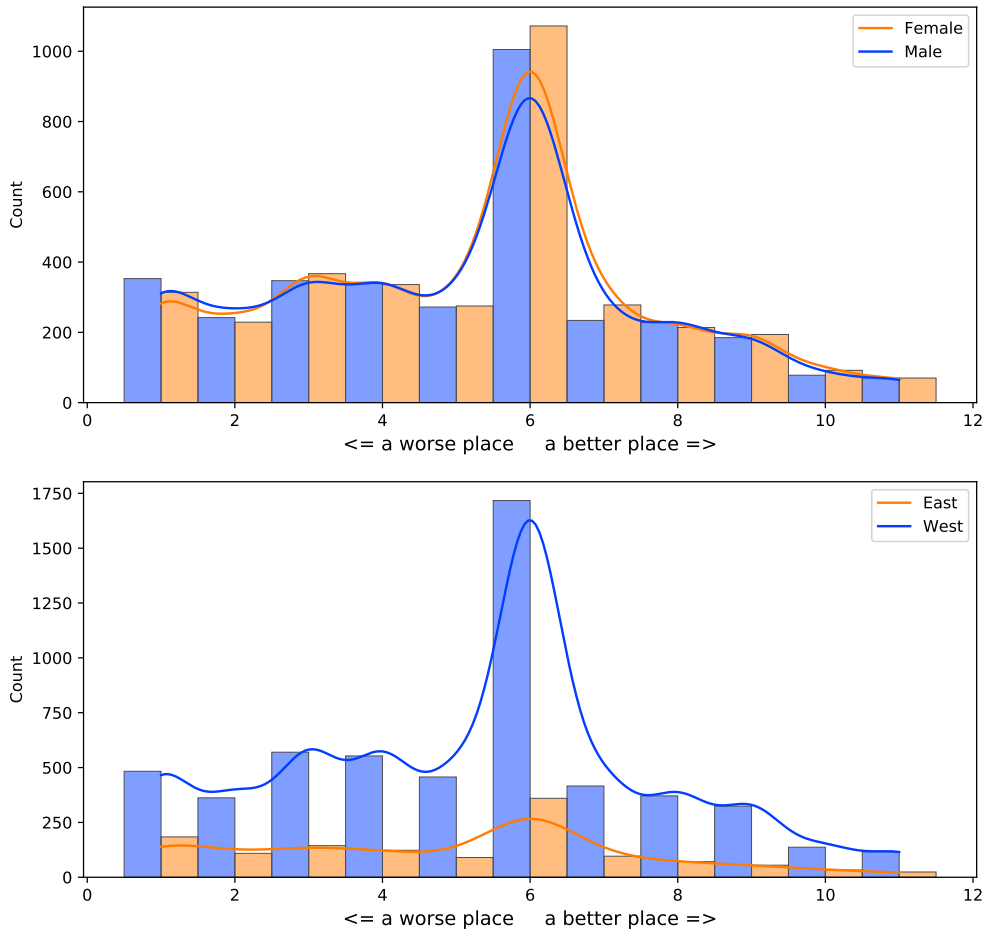

FIGURE 3.2: Distribution of German Attitudes toward Refugees based on sex and federal states

3.4.2 Correlation Analysis

Correlation analysis between the dependent variable of attitudes "Will Germany become a better or worse place to live because of the refugees?" and groups of the independent variables 1.1 were conducted. All of the following reported Pearson's correlation coefficients were statistically significant.

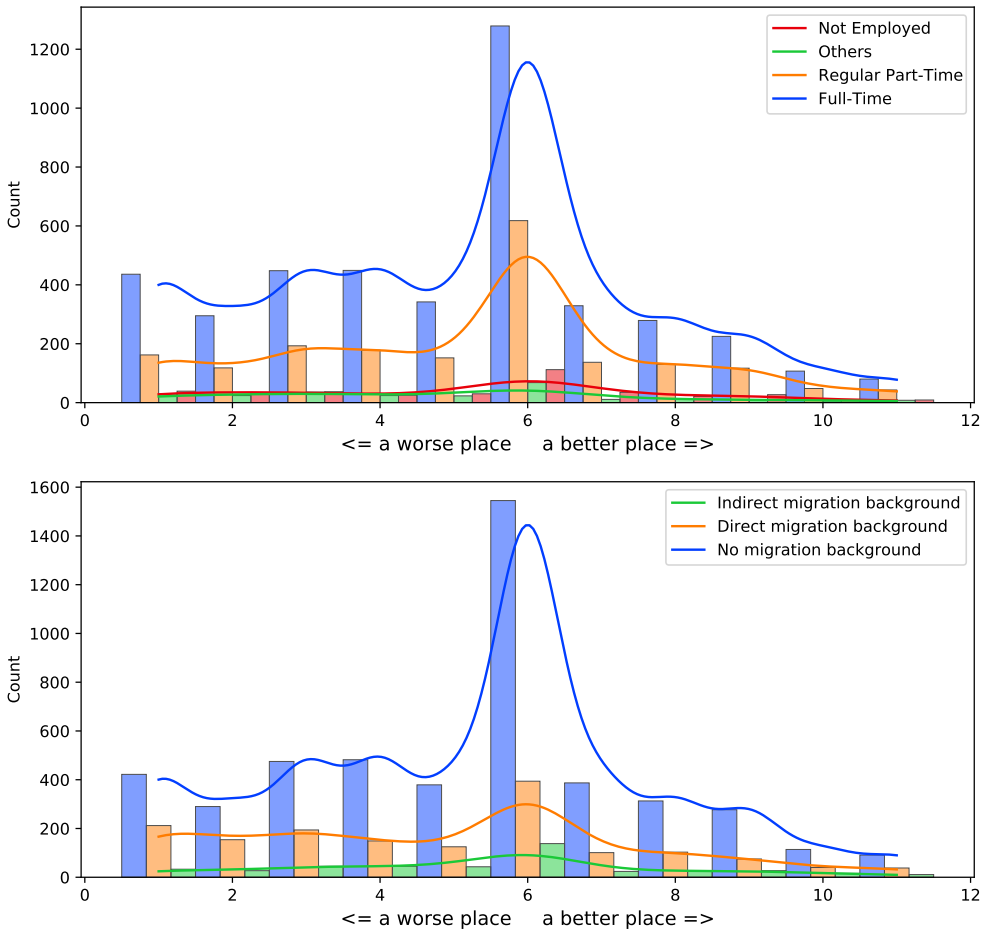

FIGURE 3.3: Distribution of German Attitudes toward Refugees based on respondents' employment status and migration background

Attitudes vs. the socio-demographic variables

Counteracting to the reviewed studies, weak positive correlations between attitudes and the socio-demographic variables 3.4 age ($r = .04$), the amount of the educational years ($r = .29$) and the income ($r = .15$) were observed, although it shows the educational years as a possible predictor of attitudes.

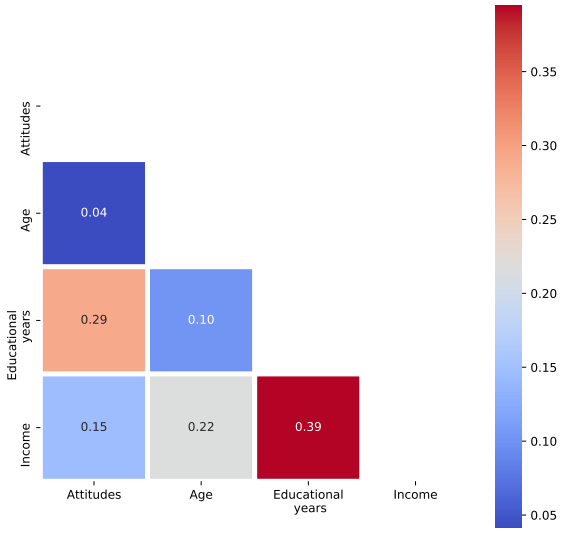

FIGURE 3.4: Correlation heatmap. Attitudes vs. socio-demographics

Attitudes vs. well-being variables

The current life satisfaction and satisfaction with personal income found also to be weakly correlated to attitudes ($r = .13$, $r = .12$), respectively, figure 3.5.

Attitudes vs. worries

Among the worries variables, only the concern about hostility to foreigners correlated negatively to attitudes ($r = -.11$). Attitudes were strongly correlated to the concern about immigration ($r = .53$) and moderately to the concern about crime in Germany ($r = .36$).
While the correlation's coefficients between the related economic concerns (economic development, finances, and retirement pension) ranged between ($r = .16$) and ($r = .21$), figure 3.6.

As respondents chose answers about worries items on a three-point scale (0: Very concerned, 1: Somewhat concerned, 3: Not concerned), the correlation indicates that respondents with less concern, especially about immigration and crime

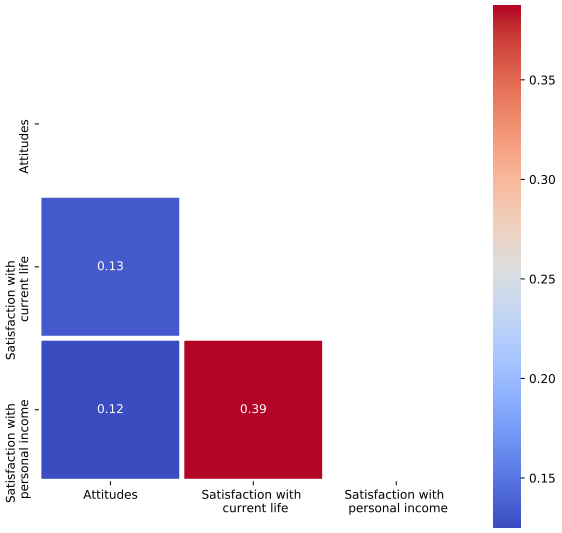

FIGURE 3.5: Correlation heatmap. Attitudes vs. well-being variables

in Germany, are expected to hold less negative attitudes.

Attitudes vs. personality traits

The big five personality traits were found weakly correlated with attitudes. Supporting to the meta-analysis (Sibley and Duckitt, 2008), attitudes were positively correlated to Agreeableness ($r = .08$) and Openness to experiences ($r.08$), and negatively to Consciousness ($r = -.09$). In addition, negative correlations between attitudes and Neuroticism ($r = -.12$) and Extraversion ($r = -.03$) were observed, figure 3.7.

Attitudes vs. the perceived threats

The correlation analysis also revealed strong positive correlations between attitudes and perceived threats. Respondents who thought that hosting refugees is good for the economy and enriches the cultural life agreed that Germany becomes

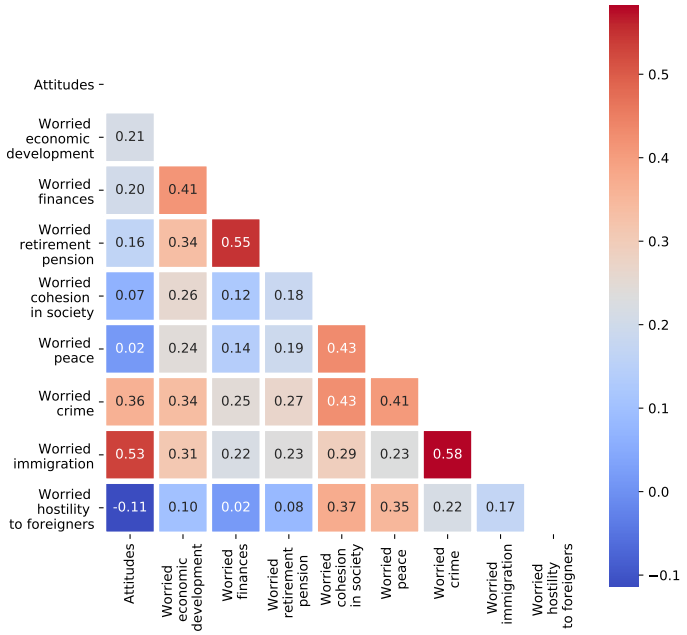

FIGURE 3.6: Correlation heatmap - Attitudes vs. concerns

a better place because of refugees. Pearson's correlation coefficients are ($r = .70$) for the economic threat and ($r = .77$) for the cultural threats (table 3.8).

The correlation between the independent variables

Figure 3.9 provides Pearson's correlation coefficients between the independent variables. The following pairs of variables were strongly correlated: the cultural threat and the economic threat ($r = .69$), the cultural threat and the worry about immigration to Germany ($r = .53$), the worry about immigration to Germany and the worry about crime in Germany ($r = .58$), and the worry about own retirement pension and the worry about finance ($r = .55$).

Moreover, exposed were moderate correlations between: worry about finance and (satisfaction with personal income: $r = .42$, worry about economic development: $r = .41$);

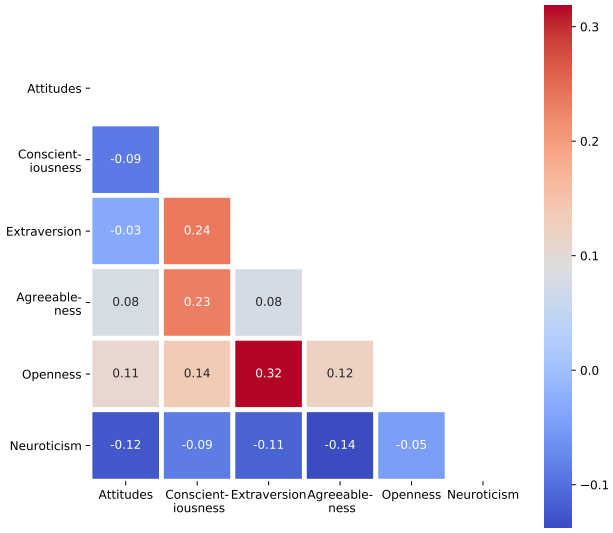

FIGURE 3.7: Correlation heatmap - Attitudes vs. personality traits

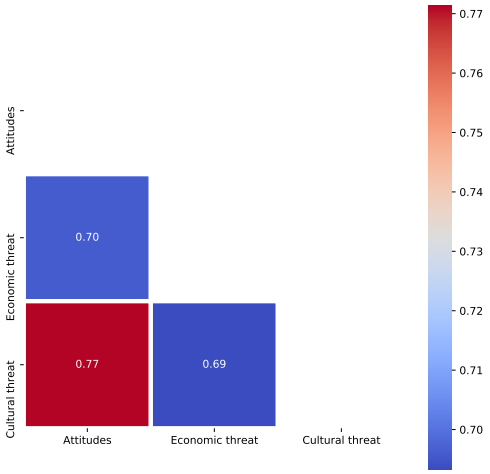

FIGURE 3.8: Correlation heatmap. Attitudes vs. the perceived threat

amount of educational years and income ($r = .39$);
satisfaction with personal income and (income: $r = .34$, current life satisfaction:

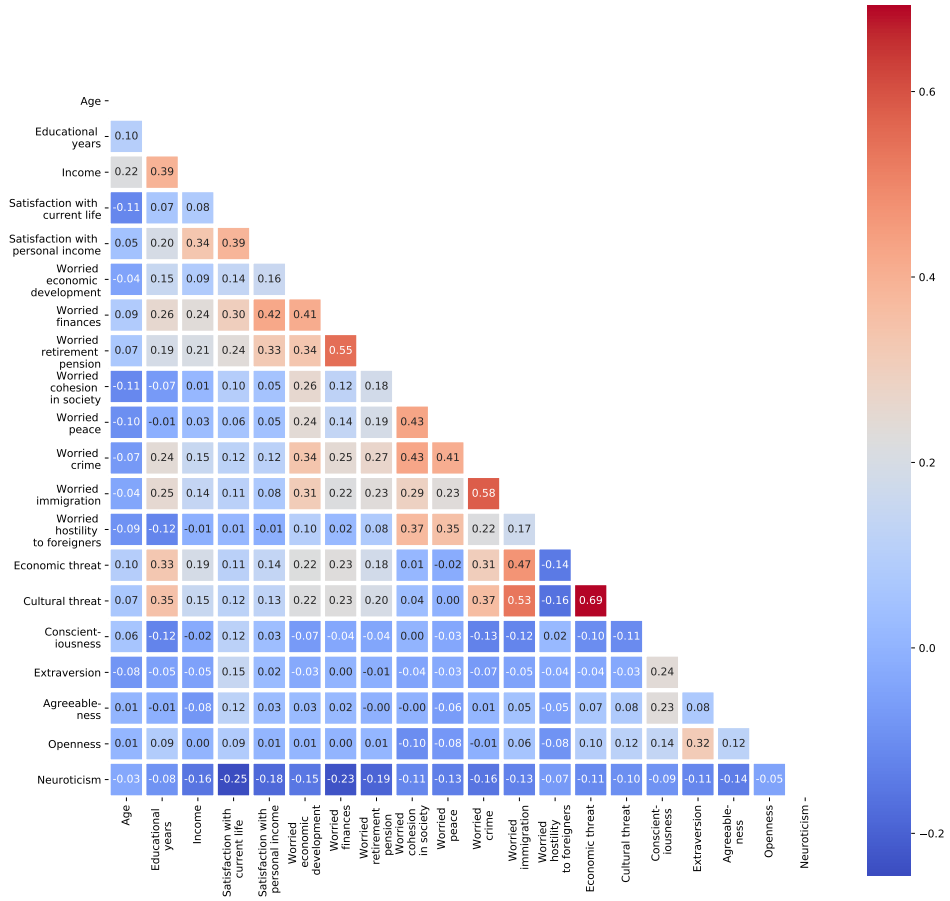

FIGURE 3.9: Correlation heatmap. The independent variables

$r = .39$);
worry about own retirement pension and (satisfaction with personal income: $r = .33$, worried about economic development: $r = .34$); worry about peace and worry about cohesion in society ($r = .43$);
worry about crime in Germany and (worry about economic development: $r = .34$, worry about cohesion in society: $r = .43$, worry about peace: $r = .41$);
worry about immigration to Germany and worry about economic development

($r = .31$);
worried about hostility to foreigners and (worry about peace: $r = .35$, worry about cohesion in society: $r = .37$);
economic threat and (amount of educational years: $r = .33$, worry about crime in Germany: $r = .31$, worry about immigration to Germany: $r = .47$);
cultural threat and (amount of educational years: $r = .35$, worry about crime in Germany: $r = .37$, worry about immigration to Germany: $r = .45$);
Extraversion and Openness ($r = .32$).

Chapter 4

Methods and Results

The reviewed studies focus mainly on correlates and associations between attitudes and possible predictors using ordinary least squares regression analysis. However, predicting attitudes by machine learning classifiers is not the goal given the complexity of the attitude's concept and the wide variety of the interacted factors influencing attitudes where some of these factors can not even be measured. The chapter clarifies the used algorithms and presents the results and their interpretation. In the study, ordinary least squares regression analysis was first conducted to evaluate findings in light of the related studies. Next, the extreme gradient boosting model was trained and explained by Shap.

4.1 Methods

This section provides simplified and sufficient explanations of the analysis methods approaches OLS, XGBoost, and Shap.

4.1.1 Ordinary Least Squares

OLS regression is one of the most fundamental, commonly used techniques, especially in social sciences. Its simple modeling idea is finding the best fitting line for a set of data points $\{X_1, X_2, ..., X_n\}$. A simple OLS regression determines the contribution of an explanatory variable X to predicting the response variable Y 4.1:

$$\hat{Y} = \alpha + \beta X + \epsilon \tag{4.1}$$

Where $\hat{Y}$ is the predicted value, $alpha$ is the intercept (the value of $\hat{Y}$ when X=0), β is the slope that indicates the change in $\hat{Y}$ associating with a unit change in X, and ϵ is the error term. Whereas OLS regression with multiple explanatory variables can be expressed by 4.2:

$$\hat{Y} = \alpha + \beta_1 X_1 + \beta_2 X_2 + ... + \beta_n X_n + \epsilon \tag{4.2}$$

The residual sum of squares (RSS) $RSS = \Sigma_{i=1}^{n}(y_i - \hat{y}_i)^2$ provides a simple measure of the model's goodness of fit, where models of poor fit result in large RSS referring that the data deviates markedly from the model overall. Therefore, optimizing the OLS regression model requires minimizing the sum of the squared residuals between the actual and predicted values. Another important measure of goodness of fit is R-square which can be computed by comparing RSS to the total sum of squares 4.3:

$$R^2 = \frac{\Sigma_{i=1}^{n}(y_i - \hat{y}_i)^2}{\Sigma_{i=1}^{n}(y_i - \bar{y})^2} \tag{4.3}$$

where $\bar{y}$ is the mean value of the variable Y.

To measure the effect of adding new explanatory variables to a model, the significance of the deviance difference between the two models can be tested by F-statistic 4.4:

$$F = \frac{RSS_p - RSS_{p+q}}{(df_p - df_{p+q})(RSS_{p+q}/df_{p+q})} \tag{4.4}$$

where RSS_p is the residual sum of squares of the first model of p variables ($p < q$) , RSS_{p+q} is the residual sum of squares of the extended model, df_p and df_{p+q} are the degrees freedom of the previous two models respectively.
F-statistic and the degrees of freedom enable determining the corresponding p-value. If the p-value is less than α, the extended model fits the data statistically better than the parsimonious Hutcheson, 2011.

4.1.2 Extreme Gradient Boosting

XGBoost is the implementation of the gradient boosted decision trees algorithm proposed by researchers at the University of Washington. It is an efficient supervised machine learning technique used for regression and classification problems.

A decision tree depends on a binary recursive partitioning logic. It splits the dataset into subsets based on a value test attribute at the nodes' level producing two branches from each node. This process repeats until each data point gets a class label at the final leaves' level based on the given stopping criterion.
A random forest consists of multiple decision trees to make the final decisions based on the small trees' voting in the case of the classification task and based on the aggregation (mean of the small trees' outputs) in regression problems. In comparison, gradient boosting 4.1 creates a series of these weak classifiers where each classifier corrects the prediction error of the previous classifier. The weights

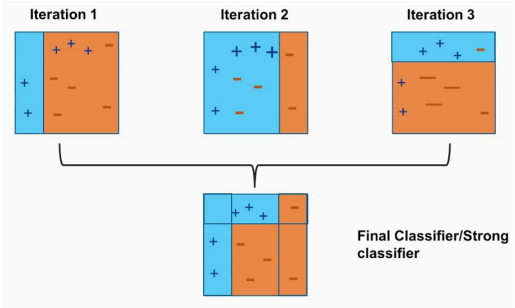

FIGURE 4.1: The boosting process in classification (Banerjee, 2019)

in the gradient boosting play an essential role as each classifier uses the predictions of the successor classifier giving higher weights to the misclassified data points to reduce the error and minimize the loss function (Chen and Guestrin, 2016; Documentation, 2021). For a dataset D of n observations and m features $D = (x_i, y_i)(|D| = n, x_i \in \mathbb{R}^m, y_i \in \mathbb{R})$, the final prediction of a tree ensemble model is given by 4.5 :

$$\hat{y}_i = \Phi(x_i) = \Sigma_{t=1}^{T} f_t(x_i), f_t \in \mathbb{F} \tag{4.5}$$

Where $F = f_1, f_2, ..., f_T$ is a set of decision trees. The equation 4.6 expresses the objective function of the previous model to be optimized over the parameters:

$$\mathbb{L}(\Phi) = \Sigma_i l(\hat{y}_i, y_i) + \Sigma_t \Phi(f_t) \tag{4.6}$$

Where l is the loss function that measures the difference between y and $\hat{y}_i$, and $\Phi(f) = \gamma T + \frac{1}{2}\lambda \| w \|^2$ is the regularization term that smooths the final weights w

to avoid overfitting. To minimize the prediction error in 4.6, the model trains in an additive manner with respect to the predictions of the previous iteration 4.7 (Chen and Guestrin, 2016):

$$\mathbb{L}^{<t>} = \Sigma_{i=1}^{n}(y_i, \hat{y_i}^(t-1) + f_t(x_i)) + \Phi(f_t) \tag{4.7}$$

4.1.3 SHapley Additive exPlanations

Shap was proposed first in 2017 by Lundberg and Lee based on Shapley values from game theory computing each player's contribution to the result of a game. The model outcome can be seen as a game in machine learning, and each feature is a player. Using marginal contributions, Shap enables explaining complex models such as gradient boosting (tree explainer), support vector machine (kernel explainer), and deep neural network (deep explainer) efficiently.

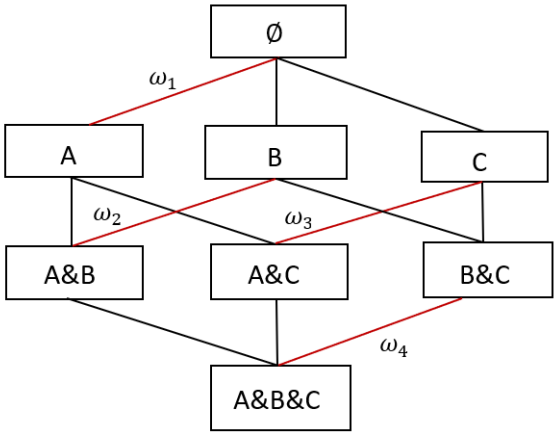

FIGURE 4.2: Illustrative example presenting marginal contributions of feature A - in case of three features

For a dataset of m features, Shap retrains the predictive model f using all possible features' subsets $S \subseteq M$, where M ($|M| = 2^m$) represents the power set of all features, to get the predicted values $\hat{y_{ij}}, j = 1, 2, ..., 2^m$ corresponding to each subset S for each data point x_i. For example, if we have three features A, B, C in our dataset, hence $M = \{\Phi, A, B, C, A\&B, A\&C, B\&C, A\&B\&C\}, |M| = 2^3 = 8$. The predicted value for Φ (no features are given) is the average of the target.

Based on the marginal contributions, determining the importance of feature k requires comparing the predictions of training the model f using all subsets $S \cup \{k\}$ ($f_{S\cup\{k\}}$: indicates that the feature k is among the used features for training the model) and the predictions of training the model with features' subset excluding k ($f_S : S \subseteq M \setminus \{k\}$). Shap values then the average weighted of all possible differences 4.8, (Lundberg and Lee, 2017):

$$Shap_k = \Sigma_{S\subseteq M\setminus\{k\}} \frac{|S|!(M-|S|-1)}{M!}[f_x(S\cup\{k\}) - f_x(S)] \tag{4.8}$$

Figure 4.2 shows a simple example of the possible feature combinations and the marginal contributions in the case of three features. Shap value of feature A for a data point x_i can be given by 4.9:

$$Shap_A = \omega_1 MC_{A,\{A\}}(x_i) + \omega_2 MC_{A,\{A,B\}}(x_i) + \omega_3 MC_{A,\{A,C\}}(x_i) + \omega_4 MC_{A,\{A,B,C\}}(x_i) \tag{4.9}$$

Where MC refers to the marginal contributions, ω_i are the weights, and $MC_{A,\{A\}}(x_i) = predicted_{\{A\}}(x_i) - predicted_{\Phi}(x_i)$, (Mazzanti, 2020).
As a local interpreter, Shap computes features' importance for each observation, whereas the global features' importance is the average of Shap values of all observations. In this study, applying the Shap tree explainer to the predictive model of the XGBoost classifier returns a matrix of the shape (p, n, m), where p is the number of classes, n is the number of observations, and m is the number of features.

4.2 Analysis and Results

This section presents the main analysis steps, including the OLS nested models and the XGBoost model for the classification task. In addition, it highlights the results of the regression analysis and the explanations of the made predictions based on the Shap values.

4.2.1 Ordinary Least Squares

For the regression analysis, the generated models utilized categorical variables of the federal state (west, east), having a migration background (no, direct, indirect

migration background), employment status (full-time, part-time, not employed, other), and sex (female, male). Whereby the models handled the other variables as continuous variables and which have been standardized before entering the regression models.
The base model included only the socio-demographic variables. Further four models were built by adding groups of the variables on well-being, concerns, big five personality traits, and perceived threats, respectively, to the base model. Comparing the models' well fit depends on the Anova test from statsmodels in python that compares the performance of two models based on the significance levels of the F test.

The base model- the socio-demographic variables

Results of the base model (table 4.1) indicated that the socio-demographic variables could explain almost 10% of the variation in the dependent variable of attitudes (adjusted $R^2 = 0.10$, $F = 70.63$, $p < .001$).

The intercept of 5.20 ($t = 96.50$, $p < .001$) refers to the value of the dependent variable corresponding to being male, from West Germany, having no migration ground, having full-time job, and of average age, income, and educational years.

Being from East Germany compared to West Germany decreased the dependent variable significantly by 0.47 units ($t = -6.29$, $p < .001$). Compared to having no migration ground, having a direct migration ground, and an indirect migration ground associated to an increase of 0.09 and 0.13 units in the dependent variable ($t = 1.19$, $p = 0.23$) and ($t = 1.10$, $p = 0.27$), respectively. The categories of the employment status of having a part-time job, having no job and having other types of work compared to having a full-time job increased the dependent variable by 0.23, 1.15, and 0.11 units.

Moreover, the model expected that being a female had almost no effect on attitude as it increased the dependent variable by only 0.01 units ($t = 0.08$, $p = 0.93$). The same for age, the results revealed an association between a change of one standard deviation in the age variable and a change of only 0.03 units in the dependent variable ($t = 0.85$, $p = 0.40$). The educational years variable was the most important predictor of attitudes among the socio-demographic variables. A one

standard deviation change in the number of educational years was associated with a statistically significant change of 0.69 units in the dependent variable ($t = 20.60$, $p < .001$). Followed by the income with a significant positive coefficient of 0.11 ($t = 2.96$, $p < .01$).

	model 1	**model 2**	**model 3**	**model 4**	**model 5**
Intercept	5.20^{*}	5.21^{*}	5.15^{*}	5.16^{*}	5.26^{*}
East Germany	-0.47^{*}	-0.47^{*}	-0.41^{*}	-0.42^{*}	-0.11^{***}
Direct migration background	0.09	0.03	-0.04	-0.07	0.05
Indirect migration background	0.13	0.13	0.16	0.13	-0.04
Part-time job	0.23^{**}	0.23^{**}	0.20^{**}	0.20^{**}	-0.01
Other employment status	0.15	0.16	0.8	-0.02	-0.37^{*}
Not employed	0.11	0.15	0.10	0.05	-0.05
Female	0.01	-0.01	0.14^{***}	0.16^{***}	-0.03
Age	0.03	0.06^{***}	0.12^{*}	0.11^{*}	-0.05^{*}
Educational years	0.69^{*}	0.67^{*}	0.57^{*}	0.51^{*}	0.00
Income	0.11^{**}	0.06	0.04	0.06	-0.02
Current life satisfaction		0.26^{*}	0.19^{*}	0.17^{*}	0.06^{**}
Satisfaction with personal income		0.06			
Economic concerns			0.19^{*}	0.17^{*}	-0.05^{***}
Security concerns			0.53^{*}	0.53^{*}	0.18^{*}
Extraversion				-0.10^{*}	-0.04
Neuroticism				-0.10^{*}	-0.06^{**}
Openness				0.26^{*}	0.06^{**}
Agreeableness				0.18^{*}	0.02
Conscientiousness				-0.20^{*}	-0.01
Perceived economic threats					0.76^{*}
Perceived cultural threats					1.34^{*}

TABLE 4.1: Coefficients of the nested OLS regression models.
$x^{*} : p < .001$, $x^{**} : p < .01$, and $x^{***} : p < .05$

Model of the we11-being variables

The second model tested the effect of adding the items of current life satisfaction and satisfaction with personal income to the base model, which increased the explained variance in the dependent variable slightly by 1% (adjusted $R^2 = 0.11$, $F = 67.92$, $p < .001$).

Current life satisfaction showed a statistically significant effect on the dependent variable, where a one-standard-deviation change was associated with a change of 0.26 units in the dependent variable ($t = 8.37$, $p < .001$). At the same time, a change of one standard deviation in satisfaction with personal income was expected to increase the dependent variable by only 0.06 units ($t = 1.89$, $p = 0.06$) (table 4.1).

Comparing this model to the base model, the results of the Anova test ($F = 50.59$, $p < 001$) suggested rejecting the nu11 hypothesis (the two models are equally helpful for predicting the dependent variable). Thus, the larger model should be preferred, and current life satisfaction was added to the further models.

Model of the worries variables

Extending the previous model by adding the composite scores of the variables on worries (economic concerns and security concerns) could explain 16% of the variance of attitudes (adjusted $R^2 = 0.16$, $F = 103.7$, $p < .001$).
A one standard deviation change in related economic concerns and related security concerns were significantly associated with changes of 0.19 and 0.53 units in the dependent variable ($t = 6.09$, $p < .001$; $t = 17.66$, $p < .001$), respectively (table 4.1).
The Anova test of ($F = 239.76$, $p < .001$) implies preferring the extended model.

Model of the big five personality traits

The fourth model included additionally the composite scales of the big five personality traits that resulted in an increase of 3% in the explained variance (adjusted $R^2 = 0.19$, $F = 86.43$, $p < .001$).

As presented in table 4.1, personality traits were statistically significant for predicting attitudes. A change of one standard deviation in Openness to Experience and Agreeableness was expected to increase the dependent variable by 0.26 and 0.18 units ($t = 9.04$, $p < .001$; $t = 6.39$, $p < .001$), respectively.
In contrast, changes of one standard deviation in Conscientiousness, Neuroticism, and Extraversion were negatively associated with attitudes, decreasing the dependent variable by 0.20, 0.10, and 0.10 units, respectively.

Compared to the previous model, the result of the ANOVA test ($F = 34.88$, $p < .001$) confirmed the well fit of the personality traits extended model.

Model of the perceived threats variables

The last model used all the independent variables, including the perceived cultural and economic threats (table 4.1), and was able to explain 65% of the variance in attitudes (adjusted $R^2 = 0.65$, $F = 103.7$, $p < .001$).

A one standard deviation change in the perceived economic threats increased the dependent variable significantly by 0.76 units ($t = 30.44$, $p < .001$). At the same time, a change of one standard deviation in the perceived cultural threat was associated with increasing attitudes by 1.34 units ($t = 52.29$, $p < .001$). Based on the Anova test ($F = 4575.43$, $p < .001$), the final model fits the data better than the previous model.

The final model results exposed a statistically significant negative association between the attitude variable and being from East Germany, being old, and having an "Other" employment status (vocational training, irregular part-time job, sheltered workshop).
Furthermore, having fewer worries related to security had a highly significant positive association with attitudes.
In this model, only Neuroticism and Openness were significantly associated with attitudes. A change of one standard deviation in Neuroticism was expected to decrease the dependent variable by 0.06 units, and a change of one standard deviation in Openness was expected to increase the dependent variable by 0.06 units.

4.3 The classification - XGBoost

For the classification task, the target variable (attitudes) was recorded based on the strength of attitudes, as explained in the third chapter, in three classes hostile (four or less on the 1-11 scale), moderate attitudes (between five and seven), and welcoming (eight or higher).

The XGBoost model was trained using 80% of the data (training data) to predict the attitude class of the remaining 20% (test data).

The complex algorithm of XGBoost requires setting a large number of the parameters Parameters, 2021 for its interacted parts. In addition to the regularization parameters, the booster parameters differentiate between trees and linear models. The learning and command line parameters set the number of trees, their depths, and the used samples.
Generally, applying parameters' tuning approaches such as the Bias-Variance trade-off enables determining the best model by balancing the predictive power and complexity.
However, this study utilized an XGBoost model built with the default set of parameters. Only the parameter "object" was set to "multi:softprob" for the multi-class classification task.

	precision	recall	f1-score	support
0	0.79	0.76	0.77	522
1	0.66	0.72	0.69	617
2	0.58	0.49	0.53	220
accuracy			0.70	1359
macro avg	0.68	0.66	0.66	1359
weighted avg	0.70	0.70	0.70	1359

FIGURE 4.3: XGBoost - Classification report

After training the model and predicting the attitude class of 1359 respondents (20% of the data), the predicted values were compared to the actual value to assess the model performance. The results were evaluated in terms of accuracy and f1 scores because we have imbalanced classes.
The classification report 4.3 shows that the model achieved a good accuracy of 70% in predicting attitudes and macro average f1 scores of 66%. It discloses that the model was more efficient in predicting the moderate and negative classes (f1 scores of 77% and 69%, respectively) compared to its performance in predicting the positive class (53%).

4.4 Shap

Shap provides data scientists with a wide variety of visualization charts explaining feature importance for predicting the class of a single data point. Additionally, it enables summarizing the contribution of each feature to a specific class by aggregating the Shap values corresponding to the given sample of data points.

After fitting the XGBoost model, the model was inserted into the TreeExplainer of Shap that returns a three-dimensional matrix representing the number of classes, the number of observations (data points), and the number of features.

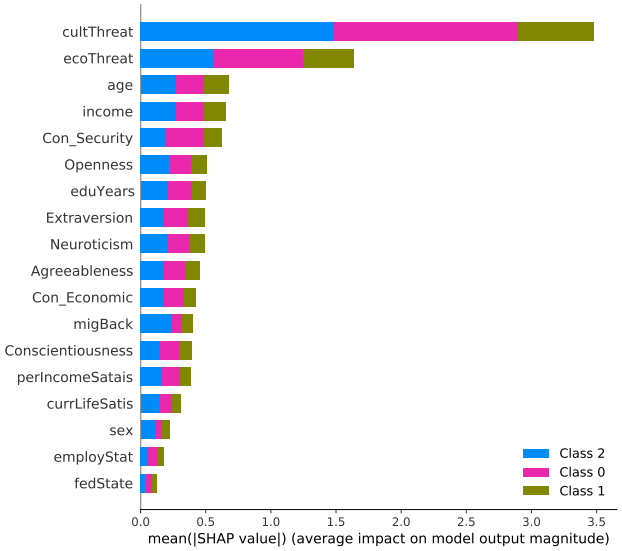

FIGURE 4.4: Shap values - Class 0: negative, Class 1: moderate, Class 2: positive. Con_Security and Con_economic refer to the related security and related economic concerns.

In figure 4.4, the features are listed on the y-axis in the rank order of importance based on the sum of Shap values for each class illustrated on the x-axis, where higher values indicate more contribution in predicting the target.
Overall, the perceived cultural threat had the most salient impact on predicting attitudes remarkably. Its Shap value (3.6) was almost double the Shap value of the second most important feature of the perceived economic threat (1.7).

On the other hand, the Shap values of the other features did not exceed (0.7). The variables of age, the related security concerns, and income had a high contribution to predicting attitudes, followed by the personality trait Openness and the amount of educational years. The features of sex, employment status, and federal state had the most negligible influence.
The colors in figure 4.4 illustrate feature importance for each class. Notably, all features had a less important role in predicting moderate attitudes (class 1) compared to negative (class 0) and positive (class 2) attitudes.

Aiming to determine feature importance for predicting each class, the mean of Shap value per class was calculated for the whole dataset. In the figures 4.5, 4.6, and 4.7, the bar plots on the left side include the features ordered base on contribution in predicting the corresponding class.
As shown in 4.4, the perceived threats had the most prominent role in predicting the three classes of negative, moderate, and positive attitudes; lowest contribution was to the moderate class. While the order of the remaining features differed between the classes.

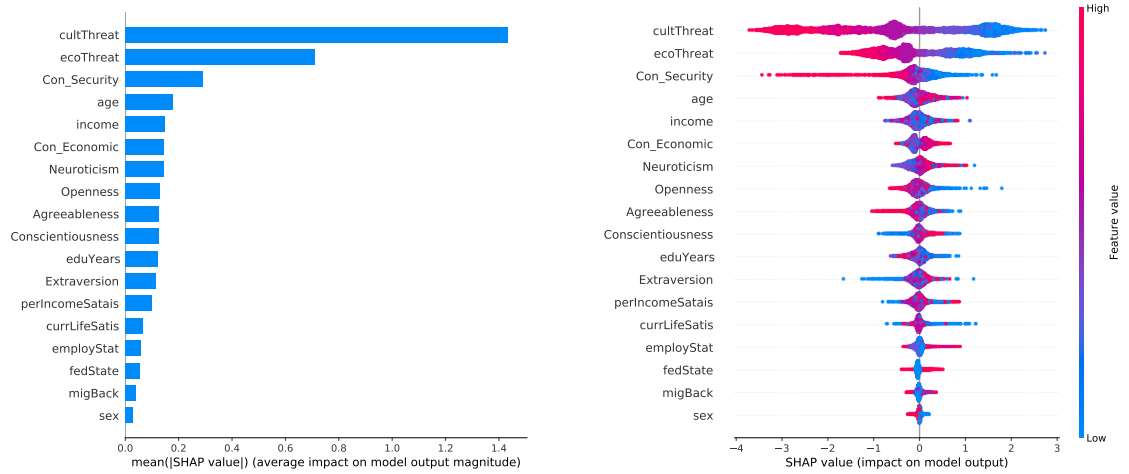

FIGURE 4.5: Feature importance for predicting negative attitudes

Furthermore, the related security concerns came directly after the perceived threats in the feature importance order of the hostile class, followed by the sociodemographic variables (4.5). The Shap values of the related security concerns were 0.30 for the negative class vs. almost 0.10 and 0.15 for the moderate and positive

classes, respectively. Among the socio-demographic variables, income and age had high Shap values for all classes. In addition, having a migration background contributed considerably only to the positive class and had a Shap value close to values of age and income.

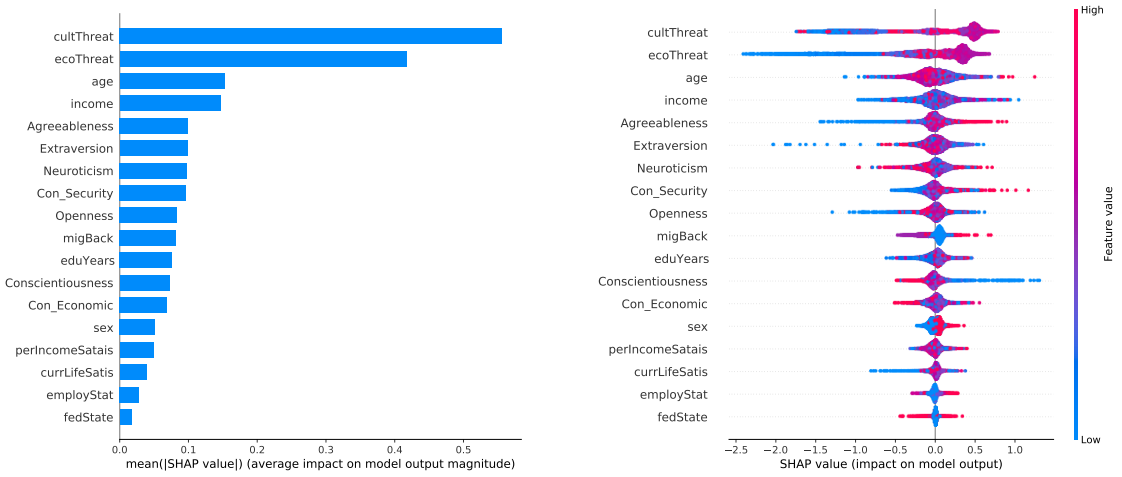

FIGURE 4.6: Feature importance for predicting moderate attitudes

The big five personality traits followed the previous features in the order, and their Shap values were not much divergent. However, the order of the features varied between the classes to be (Neuroticism, Openness, Agreeableness, Conscientiousness, Extraversion) for the negative class and (Agreeableness, Extraversion, Neuroticism, Openness, Conscientiousness) for the moderate class. For the positive class 4.7, Openness and Agreeablenes were as important as the demographic variables of age, income, and migration background, followed by Extraversion, Neuroticism, and Conscientiousness.
Current life satisfaction and satisfaction with the personal income had small Shap values, indicating low contribution to predicting the three classes.

Moreover, Shap provides information about relationships between feature values and the classes, whether they are associated with an increase or decrease in each attribute. In the scatter density plots on the right side of figures 4.5, 4.6, and 4.7), the y-axis presents the features ordered based on their importance, and the

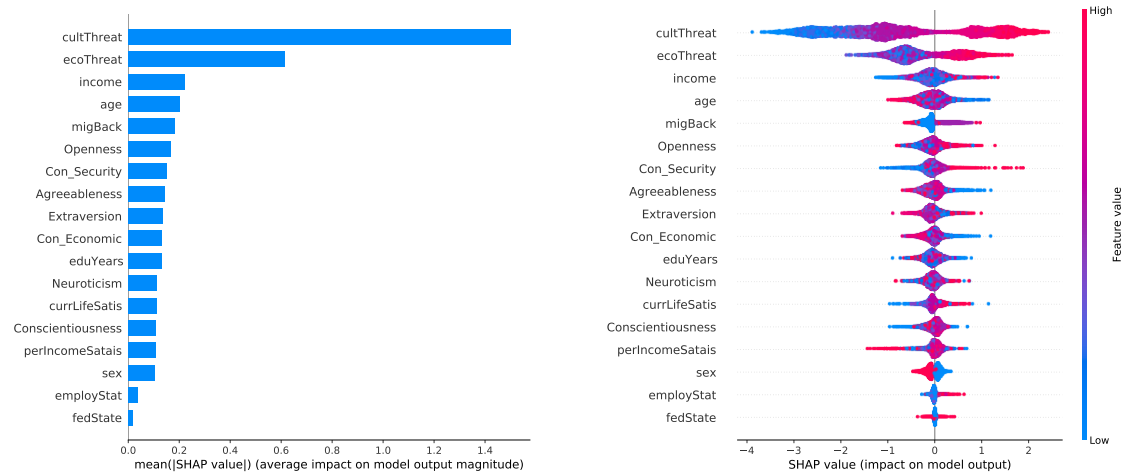

FIGURE 4.7: Feature importance for predicting positive attitudes

x-axis illustrates the average of Shap values. The color indicates the feature value (red: high values, blue: low values).

This study attempts to add findings on respondents of each class and how high or low values of specific features could impact attitudes by looking at the distribution of the colored data point and their density.

The scatter charts of the figures 4.5 and 4.7 expose that high perceived threat values pulled the data points considerably to the positive class. It reveals the remarkable impact of the belief that hosting refugee effects enrich the cultural life and improve the economy in Germany on holding positive attitudes. In addition, low scores of related security concerns indicate being very worried, being old, reporting low income and low educational levels, and being less satisfied with current life had a notable contribution to the negative class.

In the plot of the moderate class (figure 4.6), the overlapped colors of the clustered data points indicate that respondents holding moderate attitudes were more likely to choose answers in the middle of the given range scale. Not only answering items on the perceived threat but also the other questions. Although scoring low in Agreeableness and Openness had a clear negative impact on holding moderate attitudes. Otherwise, scoring low in Conscientiousness contributed positively to holding moderate attitudes.

The influence of the big five personality traits can be best observed in the negative class because the colors are less overlapped. Figure 4.5 shows that holding negative attitudes was influenced by scoring low in Openness and Agreeableness and high in the other traits. The degree of importance is illustrated in the bar plots as explained above.

It is worthy to point out that interpreting the results of such plots in the case study context is a critical task given the interaction between data points density, feature values, and Shap values, especially for large samples.

Chapter 5

Discussion

This chapter summarizes the analysis results evaluating the hypotheses and the findings based on the reviewed literature. Furthermore, it assesses the insight gained by machine learning and discusses the role of the most influential factors in predicting German attitudes toward refugees. Next, observations on the relationships between respondents' attitudes and their behavior toward refugees are reported and discussed. The chapter also highlights different technical and ethical considerations concerning applying advanced machine learning in social sciences.

5.1 The influence of the tested variables on attitudes toward refugees

In line with the meta-analysis by Cowling, Anderson, and Ferguson, 2019, this study found strong correlations between attitudes and the perceived threats, a small positive correlation between attitudes and the amount of educational years, and a very weak correlation with age.
In contrast to Talay and De Coninck, 2020 the only study that reported correlations between attitudes and the big five personality traits, attitudes toward refugees were found to be weakly correlated with Conscientiousness and Extraversion. In addition, the Pearson's coefficients of the other traits in this study were lower than in Talay and De Coninck, 2020 Openness ($r = .21$ vs. $.11$), Agreeableness ($r = .17$ vs. $.08$), and Neuroticism ($r = -.05$ vs. $-.12$).

On the other side, OLS regression analysis presents information on the role of the independent variables in explaining the dependent variable. While OLS tests the linear relationship between the dependent variables and possible predictors, the goal of applying complex machine learning algorithms is to explore further associations between the target (the dependent variable) and the features (the independent variables). The study fulfills this purpose by utilizing the ensemble model of XGBoost and Shap explainer. The trained XGBoost model was fed into the Shap explainer to determine feature importance based on the contribution of each feature to predicting a specific class by comparing the results of the trained models without a given feature to those trained with it. The higher the accuracy and f1 scores are, the more information on the features' effect.

It is essential to mention that the study does not compare the results of Shap to those of OLS. Instead, it attempts to gain a better understanding of various possible relationships between the chosen variables.
In addition to the regression models presented in the previous chapter, the study tested OLS models for each class of attitudes to explore the associations between variables for negative, moderate, and positive attitudes. The models were built in the same manner by extending the baseline model of the socio-demographic variables. The independent variables in the final models explained only 28%, 14%, and 18% of the variance in attitudes for the negative, moderate, and positive attitudes, respectively. Furthermore, none of the independent variables had a statistically significant association with attitudes except the perceived threats variables, the related security concerns for moderate attitudes, and having a direct migration background for positive, figures A.1, A.2, and A.3. Thus, Shap was informative in explaining the effect of feature values on each class.

5.1.1 The perceived threats and concerns

The OLS regression analysis revealed a considerable effect of the perceived threats on explaining attitudes toward refugees, especially the perceived cultural threat, indicating that respondents holding positive attitudes tended to agree with the statements "refugees will enrich the cultural life in Germany", and "it is generally good for the German economy that refugees are coming here" strongly.

Consistently with the related studies, the concerns also significantly affected attitudes in all of the tested models, especially the related security and immigration concerns. The results exposed that the less the respondents were worried about related economic (economic development, finances, and retirement pension) and related security and immigration (cohesion in society, peace, crime in Germany, immigration to Germany, and hostility to foreigners) issues, the more respondents tended to have positive attitudes.

The Shap tree explainer affirmed the perceived threats' prominent contribution to predicting classes of attitudes, and the perceived cultural threat was the most important feature for predicting all classes. The significant effect of the perceived threats on attitudes was observed across the studies. However, the survey and data collecting process could also explain the strong correlations and associations between the perceived threats and attitudes. While data on different variables were collected from several surveys, items on attitudes, including the perceived threats, were presented in the same survey and explicitly mentioned refugees.

Hence, the results confirmed the second and third hypotheses on the influence of concerns and perceived threats.

H_2 : Holding negative attitudes is related to high levels of worry about the economics and security in Germany.

H_3 : Holding negative attitudes is related to perceiving refugees as a cultural and economic threat.

Perceiving refugees as a cultural threat was linked to the concerns about related security and immigration issues, especially the concern about immigration to Germany that was found strongly correlated to the perceived cultural threat ($r = .53$) versus ($r = .47$) for the perceived economic threat. The regression analysis of the model of the socio-demographics and concerns as categorical variables A.4, showed that being not worried at all about immigration to Germany was associated with an increase of 3.05 units in attitudes, whereas being not worried at all about crime in Germany was expected to increase attitudes by 0.89 units.

In the light of the integrated threat theory, the perceived cultural distance justifies these findings. As most of the refugees in 2015 were from Syria and Iraq, having different religions, languages, and social norms evoked concerns about German society's national identity and values. Moreover, the explored relations between attitudes and the concerns about security and immigration are influenced by the promoted negative stereotypes, fear of terrorism, and Islamophobia.
On the other hand, perceiving refugees as an economic threat is linked to how the people evaluate the general situation of the country's economy (Vinney, 2018), which generally explains the moderate attitudes of Europeans toward refugees. The economic threat is not only related to concerns about the skills of the newcomers and their ability to learn German to integrate into society and the job market. The economic threat and concerns are additionally about the costs of the required needs for public and health services for hosting refugees.

5.1.2 The socio-demographic variables

The socio-demographic variables presented the second important group for predicting attitudes in the used sample after the perceived threats.

In the regression models, being from East Germany was the only statistically significant predictor in all models, indicating that respondents from East Germany were more likely to hold negative attitudes. However, the effect of these variables decreased after controlling for the perceived threats.
Having a part-time job and the educational years had significant positive associations with attitudes before controlling for the perceived threats. However, the impact of the other variables was either not significant or fluctuated across the tested models.

Based on the Shap values, the variables of age and income influenced mainly predicting hostile attitudes. The effect of age was notable, indicating that older people were more likely to hold less positive attitudes. In contrast, the overlapped colors representing high and low-income values prevented concluding the effect of income. The histograms of the age and income variables (figure 5.1) justify these results. While age had an almost normal distribution in the sample, the income values were skewed to the left with a mean value of 2884.43 € and a large standard deviation of 2083. Thus, the income of most of the respondents was considered

low.
Furthermore, the contribution of the amount of educational years was less than expected in light of the reviewed studies. Nevertheless, the Shap charts revealed that higher educational levels were related to holding more positive attitudes. Because the sample is imbalanced regarding the federal state (East vs. West), the contribution of being from East Germany was not clear due to this imbalance.

The results supported the first hypothesis that positive attitudes are more common among younger people, those who live in West Germany and have high educational levels. Whereas the effects of gender, income, employment status, and migration background were observed to be negligible.
H_1 : Holding less negative attitudes is more likely to be related to being young, female, having a high educational level, having a migration background, and living in West Germany.

The educational levels predict attitudes toward refugees and immigrants, especially in rich countries, where well-educated people are less likely to be worried about job opportunities, and the competition with the newcomers in the job market Mayda, 2006. Thus, high education eliminates worries about own economy and perceiving refugees as an economic threat. Further, having high education could improve different life aspects, including how people deal with public issues.

On the other side, the differences in prejudice between East and West Germany can be ascribed to many interacted social and historical factors. First, the Federal Republic of Germany (West Germany) has followed a welcoming policy toward migrants since the second world war. In contrast, the former German Democratic Republic (East Germany) had a restrictive policy in this regard until the reunion in 1990, in which migrants were allowed to stay in the country only for five years (Hellmann et al., 2019).

Moreover, Germany hosted almost a million migrants from Turkey in the 1970s (Vollmer and Karakayali, 2018), and most of them have settled in the Western German states. According to the integrated threat theory, contact with out-group members plays an essential role in accepting the refugees as newcomers and reducing

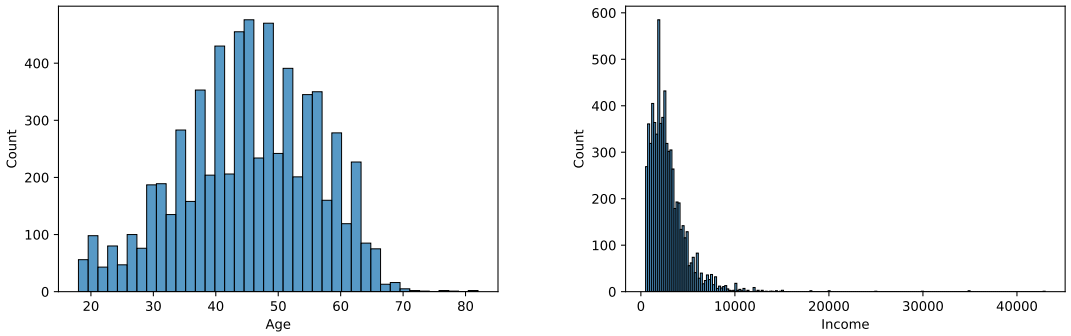

FIGURE 5.1: Histograms of the variables age and income

intergroup anxiety and the intensity of the evoked concerns and threats. In addition, many studies related the hostile attitudes toward refugees in East Germany to the rising popularity of the right-wing parties and the increasing concerns about migration (Hellmann et al., 2019, and Sola, 2018). In addition, differences in the social structure between East and West Germany, such as the deprivation (measured by educational levels and unemployment rate) and the apathy (measured via voter abstention rate), also contribute to perceiving this influx of refugees as severe cultural and economic threats (Jäckle and König, 2017).

5.1.3 Well-being and the big five personality traits

Concerning the well-being variables, the current life satisfaction positively affected attitudes toward refugees in the regression models, indicating that the higher the satisfaction with life, the more respondents tended to have positive attitudes.
Based on Shap plots, satisfaction with personal income was generally more important than the current life satisfaction. However, the effect of its values (low vs. high) was unclear, while high scores in current life satisfaction had a notable impact on holding positive attitudes.

According to the regression analysis, the big five personality traits affected attitudes significantly. Openness had the most noticeable positive effect, followed

by Agreeableness, while Conscientiousness had the highest negative impact, followed by Extraversion and Neuroticism. These results are aligned with the findings of Sibley and Duckitt, 2008. Nevertheless, only the effects of Openness and Neuroticism were statistically significant after adding the perceived threats.

Consistently, Openness was the most important trait for predicting attitudes toward refugees based on the Shap values, and the influence of the traits differed between attitudes classes.
Holding hostile attitudes was associated with high scores in Neuroticism and low scores in Openness and Agreeableness. While moderate attitudes were affected by high scores in Agreeableness and low scores in Conscientiousness, welcoming attitudes were related to high scores in Openness.

Overall, the results supported the hypothesis on the big five personality traits:
H_4 : Holding less negative attitudes is positively related to reporting high scores in Openness and Agreeableness and negatively to scoring high in Consciousness.

Researchers on the topic introduced explanations of the observed effect of the big five personality traits on prejudice. Crawford and Brandt, 2019 found that Agreeableness is the most related trait to prejudice as it is the only trait that determines people's orientations toward others. Sibley and Duckitt, 2008 argued on the role of Agreeableness in prejudice that people low in Agreeableness perceive the social world as a competition, and they are attempting to accomplish their goals and prove their power even when it conflicts with the public interest or the desire and interests of other people. In contrast, people high in Agreeableness tend to be more tender toward other people and less sensitive to competitive situations. Accordingly, people scoring high in Agreeableness are more likely to sympathize with refugees and less likely to feel threatened due to hosting refugees in their country.

Furthermore, Flynn, 2005 discussed that the relationship between prejudice and Openness is not independent of the influence of Openness on other social and political beliefs, as Openness was found to be associated with religious fundamentalism and political right orientation. Moreover, people high in Openness are more

likely to reject negative stereotypes because they tend to be cultured, curious, and broad-minded. In contrast, people scoring low in Openness tend to avoid new experiences or things and have strong opinions. This also explains that people low in Openness tend to perceive refugees as cultural and economic threats and generally have more concerns.

Talay and De Coninck, 2020 and Carlson et al., 2019 justified the influence of Conscientiousness on prejudice by its facets of orderliness, efficiency, and dutifulness. In the context of attitudes toward refugees, people low in Conscientiousness are less likely to rely on the negative media representation of refugees.

The findings on the influence of Neuroticism on prejudice and attitudes differ across the studies. However, the negative effect of Neuroticism on attitudes toward refugees is explainable by the nature of the trait. Otherwise, this study found Extraversion to affect attitudes negatively. This is aligned with (Talay and De Coninck, 2020, and Carlson et al., 2019), but there is no justification for these results in light of the warm and emotional facets of Extraversion. Further insights in this regard could be concluded utilizing the facets of the big five personality traits rather than the factors as suggested by Ekehammar and Akrami, 2007.

5.2 Attitudes and behavior

SOEP survey also offers data on behavior related to refugees asking the respondents whether they have worked directly with refugees, helped refugees by donating money or goods, and went to demonstrations or collecting signatures for initiatives to help refugees since last year (2015); and whether they would do that in the future.

Among the 6795 participants, 6596 answered this group of items, and only 1% of respondents who reported supportive behavior toward refugees in 2015 said they would not do the same in the future.

As expected, the behavior reflected notably the attitudes toward refugees. The more positive the attitudes were, the more the people tended to work with refugees or donate to them.

Attitudes	value	Working	Donating	Activities
negative	Yes	4%	18%	2%
	No	96%	82%	98%
moderate	Yes	8%	38%	4%
	No	92%	62%	96%
positive	Yes	16%	56%	15%
	No	84%	44%	85%

TABLE 5.1: The distribution of respondents who participated in related refugee activities over attitudes classes.

Table 5.1 shows the portion of respondents who were a part of concerning refugees activities, distributed based on the studied attitudes' classes negative, moderate, and positive. 56% of respondents who held welcoming attitudes donated money or goods to refugees compared to 18% and 38% of respondents who reported negative and moderate attitudes.

Moreover, donating to refugees was the most common supportive behavior even by people holding hostile attitudes. It could be explained by the humanitarian motivations, especially since, in 2015, thousand of protection seekers on their dangerous roads to Europe attracted the attention of the world and media. On the other hand, the agreement that hosting refugees makes Germany a worse place to live could be influenced objectively by the social and economic consequences of hosting this influx of refugees and does not reflect necessarily negative sentiment toward refugees. In other words, holding negative attitudes, in this case, is a result of a cognitive process, and the respondents have reported their objective opinions, not their emotions toward refugees. Therefore, feeling empathy for refugees and donating to them is not contradictory, in this case, to the belief that hosting refugees makes Germany a worse place to live.

5.3 Machine learning in social sciences

Machine learning, making predictions, and the competition for achieving high performance are nowadays essential tasks in almost every domain. Although the recent decade has witnessed a growing interest in applying machine learning algorithms in social science, including Psychology, Sociology, and political science, such applications are still a critical challenge concerning many technical and ethical issues.

One can argue whether making predictions is needed at all in social sciences. Supervised machine learning methods train the models using labeled data to learn hidden patterns and then predict the output of new unseen data (test data). The performance of these models can easily and understandably be evaluated using many performance measures such as the accuracy and area under the curve (AUC). Moreover, the cross-validation techniques enable a good level of results verification compared to the classical statistical methods Orrù et al., 2020. Whereas unsupervised machine learning is essential for discovering underlying patterns and structures in unlabelled data.
Hence, the predictive models offer researchers the opportunity to predict the outputs of new data, deal with big data [1], and make personalized predictions (individual's level) which might open new research directions in many branches such as clinical Psychology.

Molina and Garip, 2019 contended that applying machine learning in Sociology goes beyond the confirmatory research that attempts to get insights from data to confirm the theories. Instead, it establishes an exploratory direction that could address new assumptions and research questions inspired by what the data tells and may contribute to developing the standard statistical methods. In the same context, Yarkoni and Westfall, 2017 discussed that making predictions in Psychology will support and enrich the gained explanations and insights, and the efficiency of machine learning algorithms can be employed in diving deeper into studying

[1] Machine learning models are more efficient than statistical methods in dealing with high dimensionality data (datasets with a large number of variables).

causality.

On the other side, high-performance machine learning models are rather complex. They require skilfully feature engineering[2], parameters tuning, and choosing the model of the best performance using ensemble learning or combining many learners Orrù et al., 2020. This complexity results in a lack of interpretability. Therefore, the simple statistical methods that offer accurate parsimonious models are still preferred Hindman, 2015. Furthermore, the complex models could produce overfitting, which is hard to avoid even using cross-validation techniques.

It is worth mentioning that the data per se in social sciences is also a matter as data quality affects the performance of any model. All data preprocessing steps, including data cleaning and outliers detection, should be conducted compatibly with the theories of the domain study.

Another critical point concerning utilizing machine learning in social sciences is generalizability. The models can achieve high accuracy for a given sample, but testing the validity of the results for new data or other samples should also be based on theories and consider possible contextual differences.
As in many applications of machine learning, biased and discriminant patterns can be due to the data itself, and the predictive models are prone to reinforce such patterns. This study presents a good instance in this regard. First, the factors influencing attitudes toward refugees differ remarkably between samples, even samples from the same country. Second, conclusions such as relating negative attitudes to East Germany or males should be regarded in their context because generalizing them regardless of the other factors could root for new patterns of prejudice against males or people living in East German.

In the light of these considerations, interdisciplinary efforts are necessary to accomplish the maximum gain of employing machine learning in social science, where both quantitative and qualitative analysis is required. In other words, utilizing machine learning does not underestimate the role of the domain experts, and

[2]Including feature selection, combining variables, etc.

it can not be an alternative; instead, machine learning can be applied to develop theories based on real-world data.

In addition, most of the studies in social sciences are influenced by a wide variety of several interacted and even underlying hidden factors. This makes machine learning vulnerable to being misused. Therefore, the ethical aspects of using machine learning in social sciences and making predictions, exceptionally those personalized, should be seriously studied and carefully established Dwyer, Falkai, and Koutsouleris, 2018 via cooperation between academia, the governments, and the industry.

5.4 Limitations and conclusion

Attitudes toward refugees result from complex affective and cognitive processes sensitive to many variables. Understanding prejudicial attitudes helps decision-makers take actions that could alleviate public concerns evoked by hosting refugees, reduce prejudice and discrimination, and increase societal harmony. This also facilitates refugees' integration and lightens their challenges to be productive in their new societies.

This study analyzed the effect of demographics, well-being, concerns, the big five personality traits, and perceived threats on shaping German attitudes toward refugees. The findings exposed a fundamental role of the perceived cultural threats in predicting attitudes toward refugees. They highlighted the influence of the big five personality traits on attitudes, especially Openness and Agreeableness. However, the big five personality traits have a modest effect on predicting attitudes toward refugees compared to the impact of the perceived threats, concerns, and socio-demographic variables.

Moreover, studying attitudes should consider historical and political factors concerning policies dealing with immigration and refugees. It also should deem demographics on the number of refugees and immigrants and the diversity of the studied population. Other factors to be observed are the impact of social media,

networks, and the credibility of the received news and information on refugees that could promote negative stereotypes.

However, there are plenty of possible predictors of attitudes worth to be studied. Including ideological variables on religiosity and political orientation, and values of trust and fairness provides a better understanding of attitudes. Furthermore, measuring attitudes toward refugees annually by an official side like SOEP enables observing the change of attitudes over time and controlling for the influence of the personal experience with refugees, as refugees of 2015 are today a part of the German society. The annual measuring of the big five personality traits supplies insights into possible causal relationships between attitudes and personality traits and the effect of age on the development of both attitudes and personality.

On the technical side, the study proposed utilizing Shap, which ensures one of the most important ethical aspects concerning artificial intelligence, the high explainability of complex machine learning algorithms. Shap offers a solution for the challenge facing data scientists of the trade-off between interpretability and accuracy when complex but high predictive models are applied. However, considering the complexity of the concept of attitudes, the study is mainly interested in insights on feature importance rather than the accuracy of the predictions[3].

Finally, it seems valuable to study the influence of facets of the big five personality traits on attitudes, which could explain the observed associations, especially regarding Extraversion. In addition, employing Shap as an explainable and transparent tool in social studies, applying hyperparameters tuning to find the best predictive XGBoost model, and examining the robustness of the Shap results compared to the current are interesting questions for further work.

[3]The code is available on GitHub https://github.com/Rahaf66/MasterArbeit

Appendix A

Further results

A.1 OLS - negative class

A.2 OLS - moderate class

A.3 OLS - positive class

A.4 Concerns as categorical variables

	coef	P_value
Intercept	2.79	<.001
East Germany	-0.08	0.06
Direct migration background	-0.04	0.29
Indirect migration background	-0.04	0.59
Part-time job	-0.03	0.57
Other employment status	-0.12	0.16
Not employed	-0.02	0.78
Female	0.02	0.66
Age	0.01	0.59
Educational years	0.04	0.15
Income	-0.04	0.17
Current life satisfaction	0.04	0.04
Satisfaction with personal income	0.02	0.20
Economic concerns	<.001	0.95
Security concerns	<.001	0.85
Extraversion	<.001	0.86
Neuroticism	<.001	0.85
Openness	<.001	0.83
Agreeableness	0.04	0.03
Conscientiousness	-0.03	0.08
Perceived economic threats	0.24	<.001
Perceived cultural threats	0.46	<.001

TABLE A.1: OLS Regression Results - Negative attitudes (four or less on the 1-11 scale)

	coef	**P_value**
Intercept	5.88	<.001
East Germany	0.01	0.61
Direct migration background	0.02	0.55
Indirect migration background	-0.05	0.23
Part-time job	-0.04	0.10
Other employment status	-0.11	0.07
Not employed	<.001	0.10
Female	0.02	0.30
Age	0.01	0.57
Educational years	0.01	0.20
Income	-0.01	0.28
Security concerns	0.04	<.001
Openness	0.01	0.254
Perceived economic threats	0.14	<.001
Perceived cultural threats	0.20	<.001

TABLE A.2: OLS Regression Results - Moderate attitudes (between five and seven on the 1-11 scale)

	coef	**P_value**
Intercept	8.25	<.001
East Germany	<.001	0.99
Direct migration background	0.21	<.01
Indirect migration background	0.07	0.51
Part-time job	-0.07	0.35
Other employment status	0.12	0.48
Not employed	-0.01	0.94
Female	-0.03	0.60
Age	-0.03	0.30
Educational years	-0.04	0.19
Income	-0.06	0.05
Perceived economic threats	0.20	<.001
Perceived cultural threats	0.53	<.001

TABLE A.3: OLS Regression Results - Positive attitudes(eight or higher on the 1-11 scale)

	coef	P_value
Intercept	3.78	<.001
East Germany	-0.26	<.001
Direct migration background	-0.09	0.16
Indirect migration background	0.11	0.27
Part-time job	0.14	0.04
Other employment status	-0.03	0.81
Not employed	-0.02	0.83
Female	0.02	0.67
Worried about economic development[some what]	0.47	<.001
Worried about economic development[not at all]	0.44	<.001
Worried about economic development[some what]	0.11	0.17
Worried about economic development[not at all]	0.17	0.07
Worried about own retirement pension[some what]	-0.03	0.64
Worried about own retirement pension[not at all]	-0.06	0.43
Worried about cohesion in society[some what]	-0.09	0.13
Worried about cohesion in society[not at all]	-0.37	<.001
Worried about peace[some what]	-0.37	<.001
Worried about peace[not at all]	-0.51	<.001
Worried about crime in Germany[some what]	0.53	<.001
Worried about crime in Germany[not at all]	0.89	<.001
Worried about immigration to Germany[some what]	1.86	<.001
Worried about immigration to Germany[not at all]	3.05	<.001
Worried about hostility to foreigners[some what]	-0.71	<.001
Worried about hostility to foreigners[not at all]	-1.15	<.001
Age	0.09	<.001
Educational years	0.22	<.001
Income	0.02	0.45
Current life satisfaction	0.16	<.001

TABLE A.4: OLS Regression Results of concerns as categorical variables

Bibliography

A, Lim (2020). http://www.simplypsychology.org/big-five-personality.html.

Albada, Katja, Nina Hansen, and Sabine Otten (2021). "Polarization in attitudes towards refugees and migrants in the Netherlands". In: *European Journal of Social Psychology*.

Albarracín, Dolores et al. (2005). "Attitudes: Introduction and scope". In: *The handbook of attitudes* 2005, pp. 3–19.

Alici, Nilgun Kuru (2021). "Cultural sensitivity and attitudes towards refugees of Turkish nursing students: A cross sectional descriptive study". In: *International Journal of Intercultural Relations* 80, pp. 1–6.

Anderson, Joel and Rose Ferguson (2018). "Demographic and ideological correlates of negative attitudes towards asylum seekers: A meta-analytic review". In: *Australian journal of psychology* 70.1, pp. 18–29.

Banerjee, Prashant (2019). *XGBoost Classification with Python and Scikit-Learn*. https://github.com/pb111/XGBoost-Classification-Project.

Bansak, Kirk, Jens Hainmueller, and Dominik Hangartner (2016). "How economic, humanitarian, and religious concerns shape European attitudes toward asylum seekers". In: *Science* 354.6309, pp. 217–222.

Baron, Robert A, Donn Byrne, and Nyla R Branscombe (2006). "Social psychology, 11/E". In: *Aufl, Boston*.

Bassili, John N (2008). "Attitude strength". In: *Attitudes and attitude change* 237, p. 260.

Berlin, DIW et al. (2020). *SOEP-Core-2018: Deceased individual (with reference to variables)*. Tech. rep. SOEP Survey Papers.

Brücker, Herbert et al. (2016). "Forced migration, arrival in Germany, and first steps toward integration". In: *DIW Economic Bulletin* 6.48, pp. 541–556.

Carlson, Marianne Millen et al. (2019). "We welcome refugees? Understanding the relationship between religious orientation, religious commitment, personality,

and prejudicial attitudes toward Syrian refugees". In: *The International Journal for the Psychology of Religion* 29.2, pp. 94–107.

Chaiklin, Harris (2011). "Attitudes, behavior, and social practice". In: *J. Soc. & Soc. Welfare* 38, p. 31.

Chen, Tianqi and Carlos Guestrin (2016). "Xgboost: A scalable tree boosting system". In: *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*, pp. 785–794.

Cowling, Misha Mei, Joel R Anderson, and Rose Ferguson (2019). "Prejudice-relevant correlates of attitudes towards refugees: A meta-analysis". In: *Journal of Refugee Studies* 32.3, pp. 502–524.

Crano, William D and Radmila Prislin (2006). "Attitudes and persuasion". In: *Annu. Rev. Psychol.* 57, pp. 345–374.

Crawford, Jarret T and Mark J Brandt (2019). "Who is prejudiced, and toward whom? The big five traits and generalized prejudice". In: *Personality and Social Psychology Bulletin* 45.10, pp. 1455–1467.

Crawley, Heaven (2009). "Understanding and changing public attitudes: A review of existing evidence from public information and communication campaigns". In.

Crawley, Heaven, Stephen Drinkwater, and Rukhsana Kauser (2013). "Regional variations in attitudes towards refugees: evidence from Great Britain". In.

De Coninck, David, Leen d'Haenens, and Willem Joris (2019). "Investigating intergroup attitudes in Europe: Cross-national data on news media, attitudes towards newcomers, and socio-psychological indicators". In: *Data in brief* 26, p. 104535.

Dempster, Helen and Karen Hargrave (2017). "Understanding public attitudes towards refugees and migrants". In: *London: Overseas Development Institute & Chatham House*.

Documentation, XGBoost (2021). https://xgboost.readthedocs.io/en/stable/.

Dwyer, Dominic B, Peter Falkai, and Nikolaos Koutsouleris (2018). "Machine learning approaches for clinical psychology and psychiatry". In: *Annual review of clinical psychology* 14, pp. 91–118.

Eagly, Alice H and Shelly Chaiken (1993). *The psychology of attitudes*. Harcourt brace Jovanovich college publishers.

Ekehammar, Bo and Nazar Akrami (2007). "Personality and prejudice: From Big Five personality factors to facets". In: *Journal of personality* 75.5, pp. 899–926.

Flynn, Francis J (2005). "Having an open mind: the impact of openness to experience on interracial attitudes and impression formation." In: *Journal of personality and social psychology* 88.5, p. 816.

Foroutan, Naika (2013). "Identity and (Muslim) integration in Germany". In: *Washington, DC: Migration Policy Institute* 4.

Gallego, Aina and Sergi Pardos-Prado (2014). "The big five personality traits and attitudes towards immigrants". In: *Journal of Ethnic and Migration Studies* 40.1, pp. 79–99.

Gehrsitz, Markus and Martin Ungerer (2017). "Jobs, crime, and votes: A short-run evaluation of the refugee crisis in Germany". In.

Gerber, Alan S et al. (2010). "Personality and political attitudes: Relationships across issue domains and political contexts". In: *American Political Science Review* 104.1, pp. 111–133.

Gerhards, Jürgen, Silke Hans, and Jürgen Schupp (2016). "German public opinion on admitting refugees". In: *DIW Economic Bulletin* 6.21, pp. 243–249.

Gerlitz, JY and J Schupp (2005). "To survey the Big Five-based personality traits in the SOEP". In: *DIW Research Notes* 4, p. 2005.

Goebel, Jan et al. (2019). "The German socio-economic panel (SOEP)". In: *Jahrbücher für Nationalökonomie und Statistik* 239.2, pp. 345–360.

Hatemi, Peter K and Brad Verhulst (2015). "Political attitudes develop independently of personality traits". In: *PloS one* 10.3, e0118106.

Hellmann, Jens H et al. (2019). "Support for refugee integration in west and east germany". In: *Social Psychology*.

Hindman, Matthew (2015). "Building better models: Prediction, replication, and machine learning in the social sciences". In: *The ANNALS of the American Academy of Political and Social Science* 659.1, pp. 48–62.

Hutcheson, Graeme D (2011). "Ordinary least-squares regression". In: *L. Moutinho and GD Hutcheson, The SAGE dictionary of quantitative management research*, pp. 224–228.

Jäckle, Sebastian and Pascal D König (2017). "The dark side of the German 'welcome culture': investigating the causes behind attacks on refugees in 2015". In: *West European Politics* 40.2, pp. 223–251.

John, Oliver P, Sanjay Srivastava, et al. (1999). "The Big Five trait taxonomy: History, measurement, and theoretical perspectives". In: *Handbook of personality: Theory and research* 2.1999, pp. 102–138.

Katwala, Sunder and Will Somerville (2016). "Engaging the anxious middle on immigration reform: evidence from the UK debate". In: *Washington, DC: Migration Policy Institute*.

Katz, Daniel and Ezra Stotland (1959). "A preliminary statement to a theory of attitude structure and change". In: *Psychology: A study of a science* 3.423-475.

Keleş, Seray Çağla et al. (2016). "Attitudes of Turkish youth toward Syrian refugees in respect to youths' gender, income, education, and city: A Scale Development Study". In: *Turkish migration*, pp. 173–181.

Krosnick, Jon A and Richard E Petty (1995). "Attitude strength: An overview". In: *Attitude strength: Antecedents and consequences* 1, pp. 1–24.

Liebe, Ulf et al. (2018). "From welcome culture to welcome limits? Uncovering preference changes over time for sheltering refugees in Germany". In: *PloS one* 13.8, e0199923.

Lomo, Zachary A (2012). *The 1951 Convention Relating to the Status of Refugees and its 1967 Protocol: A Commentary. Edited by Andreas Zimmermann.*

Lundberg, Scott M and Su-In Lee (2017). "A Unified Approach to Interpreting Model Predictions". In: *Advances in Neural Information Processing Systems 30*. Ed. by I. Guyon et al. Curran Associates, Inc., pp. 4765–4774. URL: http://papers.nips.cc/paper/7062-a-unified-approach-to-interpreting-model-predictions.pdf.

Mayda, Anna Maria (2006). "Who is against immigration? A cross-country investigation of individual attitudes toward immigrants". In: *The review of Economics and Statistics* 88.3, pp. 510–530.

Mazzanti, Samuele (2020). *SHAP Values Explained Exactly How You Wished Someone Explained to You*. https://towardsdatascience.com/shap-explained-the-way-i-wish-someone-explained-it-to-me-ab81cc69ef30.

Meidert, Nadine and Carolin Rapp (2019). "Public attitudes towards refugees in Germany: What drives attitudes towards refugees in comparison with immigrant workers from European Union countries?" In: *Journal of Refugee Studies* 32.Special_Issue_1, pp. i209–i218.

Molina, Mario and Filiz Garip (2019). "Machine learning for sociology". In: *Annual Review of Sociology* 45, pp. 27–45.

Olson, Michael A and Richard V Kendrick (2008). "Origins of attitudes". In: *Attitudes and attitude change* 111, p. 130.

Orrù, Graziella et al. (2020). "Machine learning in psychometrics and psychological research". In: *Frontiers in psychology* 10, p. 2970.

Parameters, XGBoost (2021). *XGBoost Parameters*. https://xgboost.readthedocs.io/en/latest/parameter.html.

Plener, Paul L et al. (2017). "Unaccompanied refugee minors in Germany: attitudes of the general population towards a vulnerable group". In: *European child & adolescent psychiatry* 26.6, pp. 733–742.

Pruysers, Scott (2020). "Personality and attitudes towards refugees: evidence from Canada". In: *Journal of Elections, Public Opinion and Parties*, pp. 1–21.

Raden, David (1985). "Strength-related attitude dimensions". In: *Social Psychology Quarterly*, pp. 312–330.

Renner, W et al. (2018). "Threat perception and modern racism as possible predictors of attitudes towards asylum seekers: Comparative findings from Austria, Germany, and Slovakia". In: *International Journal of Business and Social Research* 7.12, pp. 10–22.

Sibley, Chris G and John Duckitt (2008). "Personality and prejudice: A meta-analysis and theoretical review". In: *Personality and Social Psychology Review* 12.3, pp. 248–279.

Sola, Alessandro (2018). "The 2015 refugee crisis in Germany: Concerns about immigration and populism". In.

Steenvoorden, Eefje H (2015). "A general discontent disentangled: A conceptual and empirical framework for societal unease". In: *Social Indicators Research* 124.1, pp. 85–110.

Stephan, Walter G and C Lausanne Renfro (2002). *The role of threat in intergroup relations. In D. M. Mackie & E. R. Smith (Eds.), From prejudice to intergroup emotions: Differentiated reactions to social groups (pp. 191–207)*. Psychology Press.

Stephan, Walter G and Cookie W Stephan (2000). *An integrated threat theory of prejudice. In "Reducing Prejudice and Discrimination"(S. Oskamp, ed.)*

Stephan, Walter G et al. (2005). "The effects of feeling threatened on attitudes toward immigrants". In: *International Journal of Intercultural Relations* 29.1, pp. 1–19.

Talay, Louis and David De Coninck (2020). "Exploring the link between personality traits and European attitudes towards refugees". In: *International Journal of Intercultural Relations* 77, pp. 13–24.

UNHCR (2016). *Forced displacement in 2015. Retrieved September 14, 2016 from*. http://www.unhcr.org/576408cd7.pdf.

Vinney, Cynthia (2018). "Understanding the big five personality traits". In: *ThoughtCo, www. thoughtco. com/big-five-personality-traits-4176097*.

Vollmer, Bastian and Serhat Karakayali (2018). "The volatility of the discourse on refugees in Germany". In: *Journal of immigrant & refugee studies* 16.1-2, pp. 118–139.

Wagner, Gert G, Joachim R Frick, and Jürgen Schupp (2007). *The German socio-economic panel study (SOEP): Scope, evolution and enhancements*. Tech. rep. SOEP-papers on Multidisciplinary Panel Data Research.

Yarkoni, Tal and Jacob Westfall (2017). "Choosing prediction over explanation in psychology: Lessons from machine learning". In: *Perspectives on Psychological Science* 12.6, pp. 1100–1122.

Yelpaze, İsmail and Deniz Güler (2018). "The relationship between attitudes towards asylum seekers and compassion levels of university students". In: *International Journal of Assessment Tools in Education* 5.3, pp. 524–543.