INSTITUTO
UNIVERSITÁRIO
DE LISBOA

Machine Learning to predict turnover intention in social workers

Inês Isabel Santos Ferreira

Master's degree in Business Analytics,

Advisor:
Dra. Maria Gabriela Matias da Silva, Professora Auxiliar,
ISCTE

October, 2023

BUSINESS
SCHOOL

Departament of *Métodos Quantitativos para Gestão e Economia*

Machine Learning to predict turnover intention in social workers

Inês Isabel Santos Ferreira

Master's degree in Business Analytics

Advisor:
Dra. Maria Gabriela Matias da Silva, Professora Auxiliar,
ISCTE

October, 2023

Acknowledgments

I wish to express my gratitude to the people and organizations that have supported me as I've worked on my master's thesis.

My deepest gratitude goes out to Professor Maria Gabriela Silva, who worked as my thesis adviser, for her unwavering support, wise counsel, and unending encouragement.

I'd like to thank ISCTE for providing me with accessibility to the resources I required to complete this thesis.

In addition, I would also like to express my gratitude to workers of APAV, Liga dos Combatentes, O companheiro, Associação CRESCER, Aldeias SOS, and ILGA Portugal. Without their willingness to share their opinions and experiences, this research would not have been possible.

Throughout the challenging process of finishing this thesis, my friends and family's constant support, tolerance, and understanding have been my compass. I sincerely appreciate all the work my parents put towards helping my academic career and will always be grateful for it. Regarding my friends, I'm very grateful to have met some of the most wonderful individuals in the world, who have continuously provided encouragement to me during this journey.

Finally, I'd like to thank my coworkers for always being sympathetic and accepting of the fact that I was simultaneously working on my master's thesis and doing my professional duties.

The assistance, inspiration, and work of countless people allowed me to complete this thesis during a journey that was made possible by their support. I sincerely appreciate all the help I've had along the way.

Inês Ferreira

October 2023

Sumário

Este projeto dedicou-se a determinar os elementos que influenciam a decisão dos funcionários de organizações sociais a abandonarem o seu emprego. Para atingir esse objetivo, foi desenvolvido um questionário incorporando todos os antecedentes que, segundo a literatura, apresentavam capacidade para influenciar este fator. De forma a responder à pergunta de investigação, recorreu-se à metodologia CRISP-DM para organizar o desenvolvimento de uma solução analítica que se comprometesse a prever quais os funcionários com propensão para ter intenção de abandonar o emprego. Utilizando uma técnica C&R foi identificado um modelo com métricas satisfatórias (precisão=81%; sensibilidade=79%; especificidade=83%, precisão=85%; F-Score=0,81; e AUC=0,885). Através deste modelo, foram identificados dois perfis representativos dos grupos com maior probabilidade de equacionarem sair da organização em que trabalham.

Palavras-Chave: Cansaço; *Turnover Intention*; Trabalhores Sociais; Profissionais de apoio à Vítima; Perigos Ocupacionais; Análise Preditiva; *Machine Learning*

Códigos JEL C38; M12; Y40.

Abstract

The primary objective of this study was to ascertain the factors that exert an impact on the propensity of employees within social organizations to experience turnover intention. To do this, a questionnaire was developed considering all the antecedents that the literature indicated could have an impact on this factor. To answer the research question, the CRISP-DM methodology was used to organize the development of an analytical solution to predict which employees are prone to have intention of leaving their job. Using a C&R technique, a model with satisfactory metrics was identified (accuracy=81%; sensitivity=79%; specificity=83%, accuracy=85%; F-Score=0.81; and AUC=0.885). Through this model, two representative profiles of the groups most likely to demonstrate turnover intention were identified.

Key works: Fatigue; Turnover Intention; Social Workers; Victim Support Helpers; Occupational Hazards; Predictive Analytics; Machine Learning

JEL Codes C38; M12; Y40.

Contents

Figures index

Table index

Abbreviations and acronyms

RO – Role Overload

JC – Job Clarity

PS – Perceived Support

OJ – Organizational Justice

B – Burnout

CF – Compassion Fatigue

CS – Compassion Satisfaction

OV – Organizational values

OC – Organizational commitment

TI – Turnover Intention

AUC – Area Under the Curve

TP – True Positives

FP – False positives

TN – True Negatives

FN – False Negatives

ACC – Accuracy

SP – Specificity

PREC – Precision

SW – Social Worker

CRISP-DM – Cross Industry Standard Process for Data Mining

Et al. – et alia

APAV – Associação Portuguesa de Apoio à Vítima

C&R – Classification and Regression

CHAID – Chi-square automatic interaction detection

STSD – Secondary Traumatic Stress Disorder

CIG – Comissão para a Cidadania e Igualdade de Género

CPCJ – Comissões de Proteção de Crianças e Jovens

IPV – intimate partner violence

PTSD – Post-traumatic stress disorder

ILGA – International Lesbian, Gay, Bisexual, Trans and Intersex Association

1. Introduction

1.1. Context

The greatest significant resource for a company may be its employees. Therefore, concern for the welfare of the workforce should be a top priority in firms focused on establishing healthy and long-lasting organizational climates. This is crucial when the employees' primary responsibility at work is to care for others (Vîrgă, Baciu, Lazãr, & Lupșa, 2020).

Following a meeting with the technical advisor to the management and the general secretary of APAV, it was possible to identify a high prevalence of fatigue situations among employees and the lack of means to work on the problem which led to the loss of capacity of their most crucial resource: the employees. As is public knowledge, APAV's mission is to support victims of crime, their families, and friends by providing quality, free and confidential services, as well as contributing to the improvement of public, social, and private policies related to the victim status (APAV, 2022). Because a high turnover rate is directly linked to a decline in service quality, social worker retention is crucial in the social welfare sector (Cho & Song, 2017).

This study aims to serve as a potential mechanism for enhancing the welfare of employees and, subsequently, facilitating the delivery of improved societal services. In addition to the point mentioned above, it'll explore the organization self-recognized problem and deliver knowledge about its team and how to improve its internal culture and services available to employees. The use of a Machine Learning solution will enable the organization to make data-driven decisions and facilitate future recruitment processes (Grządzielewska, 2021). It is also in the interest of all those who benefit from this service, ensuring that they receive the best possible follow-up, and of the figures of power that delegate to the organization the responsibility of studying and supporting vulnerable communities.

Recognizing the recent pandemic as a new stressor that has particularly affected social service providers, the catalyst for investigating this subject (Ratzon, Moshe, Ratzon, & Adini, 2022) has gained significant relevance. Even when examining a specific organization, the study holds considerable importance in identifying feasible solutions that can be extrapolated to other organizations within the same sector. Furthermore, it is crucial to consider the significance of utilizing machine learning in domains that have yet to experience its advantages, as this can potentially provide new opportunities and advancements. After reviewing the literature, became apparent that predictive models of this sort are commonly associated with medical domains. Furthermore, as previously stated, these models are also commonly associated with the group that exhibits the symptoms under consideration (Maslach, Schaufeli, & Leiter, 2001). However, in the case of the analysis of the professional victim support community, the studies are mostly exploratory, not offering solutions that can act in an incisive way in combating the situation of emotional distress. Hence, it is imperative to commence the implementation of machine learning techniques in social domains, namely inside the realm in which APAV functions.

1.2. Research question

From an early stage of this investigation, it became apparent that this is a prevalent issue among social actors that aid the victim, thus signifying a significant matter in need of resolution (Caringi, et al., 2017). The exposure of social support service providers to people experiencing trauma often reveals, as would be expected, negative consequences for their mental health (Caridade, et al., 2022), which led to APAV asking for help identifying the factors that can predict the intention of turnover of their employees by studying how the different distress constructs harms them. (Baird & Jenkins, 2003). In the context of trauma-related social support positions, it is frequently observed that phenomena such as Secondary Traumatic Stress Disorder (STSD), compassion fatigue, and burnout manifest, resulting in discernible impacts on the efficacy of assistance rendered by these professionals, even when they possess adequate training (Caringi, et al., 2017). Hence, it is crucial to undertake a comprehensive analysis of the aforementioned factors and their correlation with the employee's decision to depart.

Based on the previously mentioned details, the research inquiry pertains to the **factors that contribute to social workers' turnover from their employment.** The objective of this study is to investigate the factors that contribute to distress and turnover intention among employees in non-profit organizations. The study aims to identify patterns of weariness that can serve as indicators of potential issues inside the company. Once the status of employees is comprehended, it becomes feasible to categorize them within the organization, with the objective of not only presenting these findings to the management but also identifying suitable strategies for each identified pattern.

1.3. Structure

The main objective of this study is to improve understanding of the factors that influence turnover intention among social professionals in Portugal.

The study commenced by conducting a comprehensive examination of existing literature (chapter 2) to gain a deeper understanding of the issue at hand, its implications, and the optimal utilization of machine learning techniques to offer insights and resolutions for the problem.

Having established a clear understanding of the situation, the subsequent task involved selecting the most suitable methodology to address the research inquiry, that can be found on chapter 3. The CRISP-DM methodology was chosen as the framework to guide the sequential phases involved in constructing the model. These steps encompassed gaining a comprehensive understanding of the business context and selecting the most suitable analytical approach. The following step involved identifying the data requirements for the data that needed to be gathered and selecting the most suitable instrument for this purpose. The next section of the study incorporates an analysis of the raw data acquired, encompassing the identification of all elements requiring treatment. Subsequently, a fresh analysis of the data is conducted following the process of cleaning and preparation. After the completion of data preparation and comprehension, the modeling process was initiated by introducing the employed techniques, namely CHAID, C&R, and QUEST, along with the metrics used for evaluating these strategies.

The results chapter encompasses the comprehensive findings pertaining to the quality of the model, as well as the outcomes derived from the decision tree.

In the fifth chapter, an analysis of the obtained results is presented, which is subsequently followed by the presentation of conclusions, limitations, and recommendations for further consideration.

1.4 Contributions

The objective of this study is to provide a viable mechanism for improving employee welfare and, consequently, enabling the provision of enhanced societal services. In addition to the aforementioned point, this study will delve into the self-identified issues faced by the organization, while also providing insights into its team composition and strategies for enhancing its internal culture and employee services. According to Grządzielewska (2021), the use of a Machine Learning solution offers the company the opportunity to make informed decisions based on data and streamline forthcoming recruitment procedures. It is also in the interest of all stakeholders who derive advantages from this service to ensure that they receive optimal follow-up. Additionally, it is in the interest of individuals in positions of authority who outsource the job of studying and helping vulnerable groups to the organization.

2. Literature review

New knowledge must be based on previously completed work. We gain an understanding of the scope and depth of the body of work already produced by evaluating pertinent literature and identifying research gaps or testing a particular theory/hypothesis by summarizing, analysing, and synthesizing a collection of related material, and we can assess the reliability and calibre of previously published work in comparison to a standard to identify gaps, contradictions, and discrepancies (Xiao & Watson, 2017).

The review encompasses a series of objectives aimed at establishing a theoretical foundation that support the upcoming practical actions. These objectives included:

1. Gain a clear comprehension of the theoretical underpinnings of the problem.
2. Situate the problem within the specific context of Portugal.
3. Establish the factors that serve as indicators of distress among professionals providing support to victims.
4. Ascertain the most suitable strategy for diagnosing the issue through effective data collection.
5. Examine the use of prediction models in the identification of wear profiles.

2.1. Social workers

To effectively support social workers, it is imperative to gain a comprehensive understanding of their everyday roles and responsibilities. The discipline of social work comprises a wide array of professionals who hold different roles and function within a variety of environments. According to Healy (2022), it is essential for social workers to possess a well-defined and flexible professional purpose that is based on theories, knowledge, and values. Social workers function in diverse situations, offering support to individuals facing a range of difficulties and experiencing different forms of marginalization. One of the primary duties assigned to social workers entails assuming the role of catalysts for instigating change across several domains, including direct practice, organizational structures, and broader societal contexts (Healy, 2022).

Individuals who opt for a profession in social work possess elevated aspirations to aid others and foster enhancements in the lives of their consumers. The professional autonomy of social workers, which allows them to apply their expertise and abilities in assisting individuals, has historically served as a gratifying, and affirming factor in their professional endeavours. Nevertheless, in the context of bureaucratic limitations, the need for cost-effective services, and increased responsibility towards external payers, the concept of autonomy becomes compromised, leading to adverse consequences such as reduced job satisfaction and the emergence of unfavourable behaviours like subpar work performance, absenteeism, tardiness, and high employee turnover (Kim & Stoner, 2008).

Social workers are assigned with the challenging duty of completing complex occupational duties. Various research studies have shed light on the difficulties faced in the domain of social work, which cover a variety of occupational demands including increasing administrative responsibilities, burdensome caseloads, and the complexities connected with managing challenging clientele. Furthermore, social workers frequently encounter challenges such as insufficient staffing and restricted availability of adequate supervision. Additionally, the widespread expansion of complex legislation and its corresponding regulations has led to a notable increase in conflicting and incongruous demands imposed on social workers (Kim & Stoner, 2008).

2.2. Violence in Portugal

Gaining insight into the challenges encountered by social workers is crucial for comprehending the extent of their influence. The term "victim" is employed within the context of APAV in a comprehensive manner. Therefore, it is crucial to understand the evolution of some crimes in recent years.

In Portugal, the prevalent manifestations of violence documented in 2022 entails domestic violence (77.4%), sexual offenses targeting children and adolescents (4.9%), and instances of intimidation (2.8%) (APAV, 2022). According to the CIG (2023), there has been a rise in the enrolment of individuals in programs designed for aggressors and an increase in the sentencing of domestic violence offenders to either pretrial detention or effective jail since 2018. The occurrence of domestic violence is primarily observed within intimate partnerships or filio-parental relationships, with an especially large effect on women (CIG, 2022).

In relation to the characteristics of the victims, APAV helped an average of 23 women per day in 2022, surpassing the numbers of children/adolescents (7), elderly (4), and adult men (4). Even though women constituted 77.7% of the individuals who sought aid as victims, there has been a noticeable increase (20.5%) in incidents of violence against men in recent years. Furthermore, there has been a rise in the number of individuals who identify as intersex, but they continue to represent a minority within the population affected by discrimination. Many individuals affected by the incident fell within the age range of 25 to 54 years. However, there has been a notable increase in the number of victims who were under the legal age, with the year 2022 being the highest recorded level of such cases. Although most victims had Portuguese nationality, individuals from Brazil constituted the largest proportion among the foreign demographic. Furthermore, there has been a notable increase in the number of victims from Germany and Moldova (APAV, 2022).

A notable distinction exists in the spheres of operation among social workers. In comparison to adult victims, children and adolescents infrequently seek direct aid from social workers (Hamama, 2012). However, according to Silva (2022), the Comissão de Proteção de Crianças e Jovens (CPCJ) documented a total of 43,075 instances of children in precarious situations in 2021. This represents a notable rise of 3,416 children compared to the previous year, indicating an increase of 8.6%. Additionally, the APAV aided a total of 1,959 children throughout the year 2021, surpassing the number of male victims, which stood at 1,842 (APAV, 2021).

Numerous studies have observed the unique expectations placed on child-care social workers in social service settings, which may cause them to feel greater stress, have a lower level of job satisfaction, have fewer resources available to them, and experience greater trouble coping. The ability to emotionally separate oneself from one's profession was lowest among childcare care providers. Young age, obliviousness, and vulnerabilities of children, as well as the need to collaborate with a complex array of other support systems, including parents, family, schools, health professionals, and co-workers, necessitate the upholding of especially high moral and professional standards when working with vulnerable groups. Also, due to the developmental requirements of young kids, children's social workers must summon all their creativity, knowledge, emotions, and empathy. (Hamama, 2012)

2.3. Burnout vs. Fatigue

Considering that burnout has been consistently mentioned by APAV as a noteworthy concern, it is imperative to grasp its definition to make an appropriate diagnosis and determine whether it is the sole negative emotion experienced by social workers.

The primary indicator of burnout is often characterized by emotional exhaustion. Consequently, there is a decline in the expression of passion, excitement, and confidence. Individuals experiencing hardship see a decline in their level of dedication towards their professional endeavours compared to previous periods. The second factor pertains to an absence of individual achievement. Human care professionals often tend to perceive themselves in a negative light, leading to feelings of inadequacy in their efforts to support others in need. The third component pertains to depersonalization, which involves a shift in employees' perception of their clients due to the development of hostile attitudes and negative emotions towards them. (Abdallah, 2009).

Burnout's signs and effects are not as obvious as some people might think. The effects of burnout can be felt by the individual, the customers, the organization, and society. On a personal level, burnout can be characterized by a variety of symptoms, including the desire to quit one's job, job dissatisfaction, absenteeism, decreased performance, less organizational commitment, depression, suspiciousness, isolation, anger, cynicism, anxiety, impatience, outbursts, exhaustion, insomnia, migraines, denial, blaming, displacing of feelings, declining clinical effectiveness, inefficiency, and fatiguability (Abdallah, 2009).

Additionally, burnout ultimately has an impact on companies and incurs expenditures due to greater levels of staff turnover, employee departures, the hiring and training of replacement employees, and overall decreased effectiveness and productivity. On a social and organizational support level, investing in opportunities for get-togethers, getaways, and group rituals as well as soliciting the assistance, consolation, wisdom, and humour of coworkers may be suggestions to consider. Organizations can increase resources, redistribute work, limit overtime, permit vacation time, and provide direct assistance when required (for example, through employee assistance programs). Since prevention can never be started too early, it would be wise to identify some suggested strategies since prevention seems to be the focus of much of the research (Abdallah, 2009).

Burnout has consistently been associated with occupations that foster a helping relationship with others, and new evidence has shown that this association is also true in the Social Work (Sánchez, Selle, Algarín, & J.L., 2022). Theoretical frameworks in the study of burnout propose that burnout plays a crucial role as a mediator in the association between chronic workplace stressors and a range of attitudinal consequences. One of the outcomes that has been scientifically substantiated as a significant consequence of burnout is turnover intention, as indicated by multiple research (Kim & Stoner, 2008).

Additionally, social workers experience other undesirable reactions other from burnout.

"Trauma" is used more widely to describe a reaction to a traumatic incident, which may involve negative effects on one's physical or mental health. Events that are considered traumatic (such as IPV, sexual assault, war, natural or man-made disasters) are extremely taxing and frequently involve threats of major emotional or bodily harm or actual actions that do so. Traumatic event reactions can range from slight interruptions to more severe and crippling reactions that can affect daily functioning (e.g., ability to work, relationships with others) (Tarshis & Baird, 2019).

People who have undergone trauma are commonly in contact with social workers and other mental health professionals (Levenson, 2017). These professionals run the risk of experiencing emotional disturbances due to their frequent intimate interaction with trauma sufferers. As a result, these frequent encounters may have consequences that result in indirect trauma experiences. Secondary trauma defined as the actions and feelings that follow learning about or experiencing a traumatic event, as well as the strain brought on by assisting others who have experienced trauma. Like PTSD, secondary trauma can manifest as flashbacks, nightmares, intrusive thoughts, detachment, and numbness and, if ignored or untreated, might eventually result in burnout or the choice to leave the industry, among other concerns (Tarshis & Baird, 2019).

It's crucial to distinguish between burnout and fatigue since, even though they both cause similar emotions, the latter is much more sudden and it may develop after just one encounter with a traumatic event, whereas burnout is a "process" that develops over time due to a prolonged exposure to high levels of stress (Conrad & Kellar-Guenther, 2006).

Considerable attention has been devoted to the elevated levels of stress that arise from participating in professional social work practice (Senreicha, Straussnerband, & Steen, 2020). Social workers often come across individuals who have experienced trauma, which may be described as exposed to an extraordinary event that poses a bodily or psychological danger to oneself or others, resulting in feelings of powerlessness and terror (Levenson, 2017).

Despite the existence of multiple studies that have demonstrated elevated levels of workplace stress among social workers compared to other occupations, the intrinsic rewards associated with the helping parts of this job give rise to a phenomenon known as compassion fulfilment (Senreicha, Straussnerband, & Steen, 2020).

2.4 Organizational commitment and job satisfaction

The correlation between job satisfaction and organizational commitment, as well as the inclination to remain employed within an organization, has been seen among social workers and other professionals in the field of human services. The concept of organizational commitment, as described in literature, refers to the psychological bond that an individual forms with an organization. The subject of organizational commitment has been extensively researched in numerous studies undertaken over the past few decades. It is crucial to investigate this topic in relation to both employees and employers, since it has significant implications for turnover intentions, attitudes towards organizational changes, and job satisfaction (Scales & Brown, 2020). In turn Job satisfaction is another significant indicator of turnover intention and is characterized by the extent to which employees perceive favourable attitudes and emotions towards their occupation. The correlation between job satisfaction and the attributes and requirements of one's occupation, as well as the extent to which the job meets the worker's needs, expectations, and values, has been established. Individuals who opt for a profession in social work possess elevated aspirations to aid others and promote enhancements in the lives of their clientele (Acker, 2018).

2.5. Voluntary turnover vs turnover as an outcome

Voluntary turnover refers to the deliberate decision made by an employee to either remain with or go from an organization. This phenomenon can have negative implications for organizations since it often involves the departure of highly skilled and talented workers. The departure of talented individuals, when done willingly, results in the loss of their knowledge, skills, and abilities, so impacting the efficiency and quality of work delivered by the company to the community (Scales & Brown, 2020).

The primary objective of APAV is twofold: firstly, to ascertain the underlying factors contributing to employee tiredness, and secondly, to implement measures aimed at mitigating turnover within the business.

Turnover intention refers to a purposeful and conscious decision to voluntarily terminate one's employment within a predetermined timeframe, frequently resulting in the individual leaving the organization (Kim & Stoner, 2008). There is a growing concern regarding the escalating rate of voluntary turnover among Social Workers (SWs). This trend is accompanied by a noteworthy decrease in employees' organizational dedication and appreciation for their work (Scales & Brown, 2020).

Several studies have been done on the factors that affect social workers' decisions to leave their professions and one of the main conclusions is that turnover- the act of leaving a job – is a process rather than an isolated event. Turnover is usually considered less impactful that the actual intention to leave. The desire of an individual to leave an organization is reported to be the best predictor of actual turnover, which led to most empirical studies to use the employees' intention to leave as the outcome variable instead of actual turnover (Wang, Jiang, Zhang, & Liu, 2021).

2.6. Turnover implications

Unsurprisingly, the organization is adversely impacted by employee turnover due to the detrimental effects it has on morale, continuity of care for both clients and individual workers, and the potential departure of workers from the profession (Kim & Stoner, 2008). The escalating expenses associated with employee turnover have imposed a significant load on human resource professionals. The significance of the role and the experience of happiness among social workers is crucial for the ongoing advancement of a robust workforce and the provision of high-quality services to their community. The commitment of social workers is crucial to both the organization and the profession (Scales & Brown, 2020).

The occurrence of voluntary turnover among social workers has a detrimental effect on both the agency and the people they are responsible for serving. The departure of a qualified licensed social worker has a significant influence on all levels of the organization. Primarily, the clients experience a significant impact as they are compelled to revisit their personal tragedy whenever a new social worker is assigned to their case. Furthermore, social workers (SWs) experience an impact, as the workload of the departing SW must be redistributed among the remaining staff members. Finally, the agency's service quality is adversely affected by an employee's heightened caseload (Scales & Brown, 2020).

The presence of high turnover rates poses a challenge in upholding the continuity of treatment, resulting in adverse effects on clients and their families. Youth who are engaged with the child welfare system have reported experiencing various consequences because of worker turnover. These consequences include a sense of mental and physical instability, as well as the loss of trusted relationships (Brown, Walters, & Jones, 2019).

2.7. Turnover prediction

Considering the employees’ intention to leave as the outcome, there are three categories that represent the antecedents of social workers’ TI (turnover intention) (Wang, Jiang, Zhang, & Liu, 2021) (Barak, Nissly, & Levin, 2001):

a. *Demographic factors, both personal and work-related:* <u>Age, education, job level, gender, and having children</u> are key indicators of turnover, according to several studies (Barak, Nissly, & Levin, 2001). Although the effect size is only moderate, age exhibits a greater magnitude compared to many individual difference predictor variables such as cognitive capacity, education, training, marital status, and others. Additionally, research has demonstrated that age can be of similar importance as commonly examined factors, such as perceptions of distributive justice, perceptions of role overload, and contentment with co-workers, when predicting turnover (Ng & Feldman, 2009). Tenure is another factor considered in this category, and it naturally shows an inverse association with turnover. Employees who have worked for an organization for a shorter period tend to depart more frequently than those who have worked there for a longer period, which is possibly due to them having a greater sense of commitment to the organization (Barak, Nissly, & Levin, 2001).

b. *Professional perceptions*: It has been shown that burnout is the main cause of low morale, which breeds the desire for turnover (Abdallah, 2009). On the other hand, the psychological and emotional support from family and friends outside the workplace acts as a buffer against work-related stress and subsequent TI. Other important factors are the professional commitment to the individuals who are being helped by the organization, which has a negative relationship with turnover (Barak, Nissly, & Levin, 2001) and the experience of conflict between the employees' professional values and those of the organization (Tham, 2007). As expected, job satisfaction is also a consistent predictor of turnover behavior (Acker, 2018), which acts not only on the TI but also on organizational commitment, another important antecedent of turnover (Barak, Nissly, & Levin, 2001).

c. *Organizational conditions:* Stress related factors, such as role overload and lack of clarity in the job descriptions are two important predicts of TI (Tham, 2007). Also, given the strong need for cooperation and team-based interactions to achieve good results in daily work, support from co-workers and supervisors (Maertz, Griffeth, Campbell & Allen, 2007), and satisfaction with the work of the others are essential to prevent employees from leaving. And finally, perceptions of positive procedural and distributive justice policies are truly important to prevent TI (Barak, Nissly, & Levin, 2001).

2.8. Machine learning for turnover prediction

The field of social work is characterized by its demanding nature, as practitioners frequently encounter emotionally charged and intricate situations. However, considering the current challenging cultural and political circumstances, social workers have an elevated risk of experiencing stress (Acker, 2018) and consequently to desire to leave their organization (Wang, Jiang, Zhang, & Liu, 2021).

In turn, machine learning is a computational approach that enables computers to gather information from extensive datasets and generate solutions. The categorization technique, a subfield of machine learning, is employed to discern unique patterns that may exhibit variation across diverse categories. The aforementioned patterns are afterwards employed for the purpose of assessing diverse metrics and assigning observations to their corresponding categories (Zhang, Zhang, Ren, & Jiang, 2023).

The utilization of machine learning techniques for the analysis of this subject is not extensively prevalent; yet there exist cases within the medical field that can be extended to the realm of social work. Scholarly investigations have focus on nurses who are actively involved in providing care to patients with life-threatening illnesses or those who are employed in the context of the Covid pandemic.

The basic aim of scientific psychology is to understand and anticipate human behavior. However, it is important to acknowledge that there has been a notable emphasis on the quest of explanation, while the issue of prediction or classification has received comparatively less attention over a prolonged period (Zhang, Zhang, Ren, & Jiang, 2023). In the field of psychology, classification algorithms have demonstrated significant efficacy in predicting the mental status of individuals. In a recent systematic review conducted by Zhang, Zhang, Ren, and Jiang (2023), the utilization of machine learning techniques in the field of mental health was examined and summarized. The review indicated that classification algorithms demonstrated satisfactory performance in predicting various mental health conditions, including depression, suicide, job-related stress, bipolar disorder, mood disorders, posttraumatic stress disorder, anxiety disorders, substance abuse, and schizophrenia.

Machine learning approaches have the potential to provide novel avenues for understanding patterns of human behaviour, detecting symptoms and risk factors related to mental health, making predictions about the evolution of diseases, and tailoring and optimizing therapeutic interventions (Thieme, Belgrave, & Doherty, 2020). It is proposed that directing more attention towards prediction, as opposed to explanation, has the potential to finally enhance our comprehension of human behaviour (Yarkoni & Westfall, 2017). To illustrate this a study conducted by Zhang, Zhang, Ren, and Jiang (2023) successfully elucidated the compassion fatigue profile exhibited by psychological hotline counsellors throughout the ongoing epidemic, and it placed significant emphasis on the utilization of categorization and prediction methods within the field of scientific psychology, as opposed to prioritizing explanatory approaches.

3. Methodology

The chosen approach for this study will be a modified version of the CRISP-DM framework, as it has been found to achieve the most favourable practical outcomes based on prior research. Instead of the five steps typically connected with it, the process will be presented in a more comprehensive manner. Therefore, it will be segmented into the subsequent stages (Logallo, 2019):

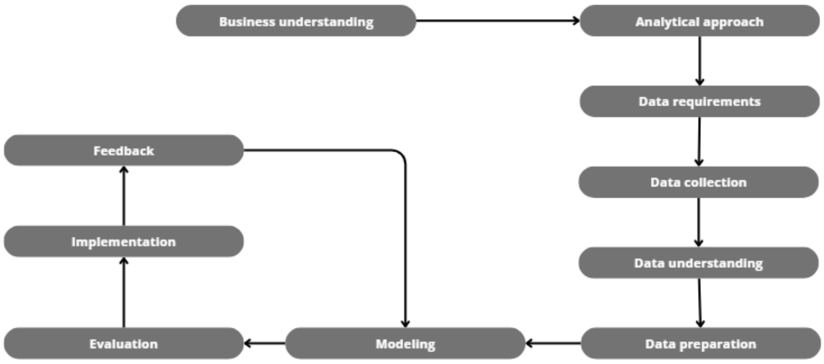

Figure 3.1. CRISP-DM (Logallo, 2019)

Understanding the business: Which facilitates the comprehension of APAV's present operations, its process management strategies, and the ongoing implementation of organizational support initiatives and employer branding measures. By following this procedure, it becomes feasible to determine the instances of business failures and establish specific objectives for the project. The comprehension of the business has been achieved by conducting research on the information available on the organization's social media platforms, as well as by engaging in direct communication with its representatives.

Analytical approach: When the objectives are accurately delineated, it becomes viable to identify the most appropriate procedures and gather data that align with these aims, such as employing diagnostic models that have been previously utilized in scholarly literature.

Data requirements: After gaining a comprehensive understanding of the business, delineating the objectives, and selecting the appropriate approach, the subsequent task involves determining the specific category of data that needs to be gathered. To effectively accomplish the objectives of the project, it is necessary to ascertain the most suitable content, formats, and fonts. In this instance, the selection of, for example, the burnout and fatigue evaluation model will require its customization to the organizational context and specific requirements.

Data collection: During this phase, data is collected by employing a tool, specifically Google Forms, to conduct a questionnaire and acquire the data intended for analysis.

Data understanding: Throughout this phase, a comprehensive evaluation will be undertaken to check the quality of the data that will be utilized to support the project, with the aim of identifying any deficiencies or inaccuracies that need correction.

Data preparation: Over this stage, it is imperative to process the data to prepare them for the analytical approach that has been selected beforehand. The data will undergo a process of cleaning and organization to enhance its quality and ensure its readiness for further analysis.

Modelling: The development of the model is focused on generating outcomes that facilitate the resolution of the initial challenge. The proposed model will employ a predictive algorithm with the primary aim of identifying wear profiles. This identification will be facilitated using decision trees, enabling the model to discern personal and professional qualities that exhibit a higher likelihood of experiencing the phenomena being investigated.

Model evaluation: After the process of developing the model is finished, it is crucial to undertake an analysis to determine its effectiveness and determine whether it demonstrates sufficient performance or requires more adjustments. For example, in a comparable investigation related to the anticipation of depression, a validation sample was employed, and the Area Under the Curve (AUC) was applied as a metric to evaluate the correctness of the established model. (Nishi, Michiyo, & Matoba, 2021).

Implementation: The implementation will enable the management council of APAV to get knowledge of the discovered profiles and the required mitigating measures for psychological tiredness experienced within a test group, to comprehend the effectiveness of its use in real settings.

Feedback: Ultimately, the culmination of the project entails the delivery of the generated solution to the intended audience, namely the end-users who initially specified their specific requirements and sought assistance in the creation of a suitable tool. To enhance the model and align it with the opinions of individuals involved, it is imperative to consider and incorporate criticisms. In this context, it is imperative to not only listen the suggestions of the management board, but also to comprehensively grasp the insights provided by the support teams, as they serve as the subject of examination.

3.1. Understanding the business: APAV

APAV's objective, widely acknowledged in the public domain, entails the provision of high-quality, cost-free, and confidential services to victims of crime, their families, and acquaintances. Additionally, APAV aims to contribute to the enhancement of public, social, and private policies pertaining to the victim status (APAV, 2022).

The organizations have numerous Victim Support Offices (GAV) that prioritize the fulfilment of the initial aspect of their purpose. This is achieved by establishing strong connections within their networks to provide victims with the most optimal response possible. The GAV is structured into teams, each led by a manager and supported by a group of professionals that has particular expertise in victim support. In addition to the corporate workforce, each team benefits from the support of volunteers to facilitate the implementation of their ideas (APAV, n.d.). According to APAV (n.d.), there exists a total of 54 GAV, which are accompanied by a mobile support unit and 15 specialized support networks. The organizational purpose is to enhance the efficiency of the association and guarantee that staff are in optimal condition to give the required support and effectively contribute to the well-being of the victims they assist (APAV, n.d.).

Refining the subject matter of APAV work proves challenging due to its encompassing of multiple types of crime, including but not limited to domestic violence, sexual offenses, and various others. Moreover, APAV extends its help to a wide range of victims, spanning from older individuals to children and adults, as well as those who have experienced prejudice based on factors such as sexuality, gender, race, or other. Nevertheless, it is evident that each circumstance is inherently challenging and intricate, necessitating a significant level of dedication from the professionals involved (APAV, n.d.).

Having a comprehensive understanding of the business operations and the large organization of the assistance center, together with familiarity with relevant literature, facilitates a deeper comprehension of the factors that have contributed to the issue of staff weariness. When APAV introduced the issue under assessment, it was presented as a recurring phenomenon that they had observed over a period but have not yet had the chance to handle adequately. The issue at hand pertained to a significant proportion of their workforce exhibiting indications of fatigue and burnout, which they hypothesized could be linked to challenges in both employee retention and recruitment.

3.2. Analytical Approach

The organization articulated a clear challenge about the identification of elements that contribute to the turnover of social workers from their employment. In this regard, as previously observed, classification models can serve as a valuable instrument for identifying, among the numerous factors that contribute to a heightened likelihood of departure, those that exhibit a greater probability within the given environment (Zhang, Zhang, Ren, & Jiang, 2023). Based on the previously listed data, the chosen analytical technique involved the utilization of a predictive model to determine the likelihood of an employee exhibiting turnover intention based on a set of specific constructs. To enhance comprehension of the interrelationships among various constructs, the decision tree algorithm was used to display the most optimal and visually appealing outcomes.

3.3. Data Requirements

Given the analytical methodology that was previously chosen, it became apparent that there was a need for comprehensive data that could contain not just all the factors previously identified as influencing turnover intention, but also the turnover intention itself. The primary data collection methods utilized in this area of research involve the application of questionnaires that incorporate pre-validated scales. After conducting a thorough review of relevant literature pertaining to the subjects under investigation, it was determined that the survey should be designed following the same methodology employed by previous researchers. The objective of the survey is to assess various domains concurrently, encompassing labour and demographic information, as well as the perceptions of social workers on the level of support they receive and their overall satisfaction (Hamama, 2012) (Sánchez, Selle, Algarín, & J.L., 2022). The primary source of data should be obtained from employees of APAV. However, due to the limited number of staff members, it is crucial to seek input from other organizations that share similar values and objectives.

The components of data included in this study were derived from the model described in chapter 2.9. This model was selected due to its alignment with a substantial body of existing literature and its demonstrated comprehensiveness in capturing the various constructs that have the potential to influence turnover intention. Based on that the collected should comprehend components of demographic value, professional perceptions, and Organizational conditions (Wang, Jiang, Zhang, & Liu, 2021) (Barak, Nissly, & Levin, 2001). Each of these components encompasses different constructs that propose hypotheses regarding their influence on the intention to leave an organization.

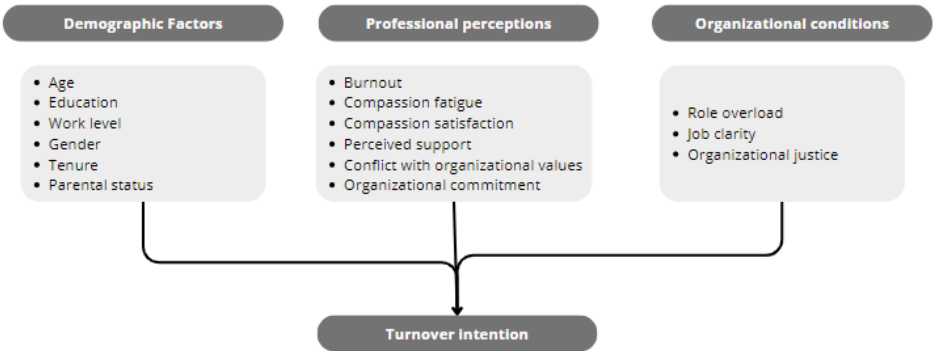

Figure 3.2. Research framework

Demographics: The demographic category should contain a range of individual attributes, including age, educational attainment, gender, parental status, and work-related variables such as job level and length of employment. These aforementioned factors have demonstrated their significant impact on employees' desire to depart from a firm, so qualifying as viable constructs for the purpose of this study (Wang, Jiang, Zhang, & Liu, 2021); (Barak, Nissly, & Levin, 2001). Demographic factors have frequently been found to be reliable indicators of employee turnover. Age, education, employment level, gender, and business tenure have been linked to turnover in several studies. It is generally accepted that younger, better educated, and less experienced workers are more likely to resign. Employees who differ from their coworkers in race, ethnicity, gender, or age are more likely to resign. Evidence suggests that higher-income individuals, and those with greater home social support networks have lower turnover rates. It also suggests that having children at home greatly influence turnover, especially for women (Barak, Nissly, & Levin, 2001).

Professional perceptions: This area encompasses constructs that merit examination, including burnout, perceived support, organizational commitment, conflict with organizational ideals, job satisfaction, and job commitment. Various categories that have been examined in past studies have demonstrated their influence on turnover intention (Scales & Brown, 2020); (Acker, 2018); (Wang, Jiang, Zhang, & Liu, 2021) ;(Barak, Nissly, & Levin, 2001). Burnout is a prevalent concern within the domains of mental health and social care, resulting in diminished morale and increased staff turnover. There is also evidence to suggest that receiving psychological and emotional support from those in one's personal network, such as family and friends, who are not affiliated with the workplace, may serve to alleviate job-related stress, and decrease employee turnover rates. In turn, the presence of organizational commitment has been found to have a detrimental effect on turnover rates. Individuals who experience a misalignment between their personal professional beliefs and the prevailing organizational culture are more inclined to voluntarily separate from the organization, whereas those who see a congruence between their principles and the organizational culture are more like to exhibit longer tenure;(Barak, Nissly, & Levin, 2001). Multiple research studies have also provided evidence suggesting that job satisfaction has a significant role in influencing both organizational commitment and the intention to leave, ultimately resulting in employee turnover (Acker, 2018). In another hand, research has proved that a committed employee demonstrates alignment with the organization's values and principles, exhibits a strong willingness to exert effort, and expresses a desire for long-term engagement. Decreased employee commitment is associated with heightened work unhappiness and an elevated probability of voluntary turnover within the organization (Scales & Brown, 2020); (Barak, Nissly, & Levin, 2001). The professional quality of life refers to the subjective perception of the quality of one's work as a helper. This encompasses another two key constructs: compassion fatigue and compassion satisfaction. Compassion fatigue refers to the emotional exhaustion and distress experienced by those in helping professions. On the other hand, compassion satisfaction is characterized by the pleasant emotions and fulfilment one derives from assisting others (Wells-English, Giese, & Price, 2019).

Organizational conditions: This group involves stress-related factors, including role overload, inadequate job descriptions, and an unsatisfactory sense of organizational justice (Wang, Jiang, Zhang, & Liu, 2021) ;(Barak, Nissly, & Levin, 2001). Job turnover has been linked in many studies to stress-related traits (Hamama, 2012) such as role overload and a lack of clarity in job description. There exists a negative relationship between turnover intentions and the way employees perceive positive procedural and distributive justice practices inside a business. Employees who hold the perception that an organization's compensation practices are equitable and impartial are more inclined to exhibit lower levels of turnover (Wang, Jiang, Zhang, & Liu, 2021).

Turnover intention: Turnover, which refers to the act of actually departing from a workplace, is commonly regarded as a gradual process rather than an abrupt event. The intention to quit, which refers to an individual's desire to leave an organization, is widely regarded as the most influential and accurate predictor of actual turnover within the turnover process (Wang, Jiang, Zhang, & Liu, 2021).

3.4. Data Collection

Data for the study was obtained through the utilization of a standardized survey questionnaire, utilizing a multidimensional approach as a methodological tool to examine the various constructs that contribute to turnover intention inside non-profit organizations operating in the domain of vulnerable populations. To develop the instrument for data collection, it was imperative to conduct an examination of research that employed and validated questionnaires deemed most appropriate for examining each of the predetermined constructs. Therefore, a set of questions and scales, that have been previously verified, were selected for the analysis of each construct. These items were then grouped into a questionnaire. Due to the comprehensive nature of the questionnaire, two assessment moments were established for the tool. During these moments, individuals affiliated with various institutions who worked in organizations that respected the lines of work of the professional area in focus, completed the questionnaire and provided their evaluations. Their contributions aimed to streamline the questionnaire by reducing the number of questions and refining the semantics to ensure clarity and objectivity for respondents, while also considering the specific objectives of each item. Initially, a prototype survey was administered to three individuals affiliated with a social organization. The respondents' feedback indicated that certain questions within the survey were perplexing and lacked clarity. Additionally, respondents were requested to provide suggestions for improvement after answering each set of questions, to address any potential uncertainty, they might've encountered. Following the implementation of the recommendations, a subsequent phase ensued wherein two additional employees undertook a similar task as the initial group, albeit with a more concise and tailored questionnaire.

A final set of questions on previously identified demographic categories, including gender (D1), age (D2), parental status (D3), marital status (D4), tenure (D5), and role (D6), were included in the questionnaire. Given the anticipated distribution of the survey among several organizations, an additional survey item was incorporated to gather information regarding the respondent's affiliation with their respective organization (D7). The questionnaire items were classified using a coding system that comprised the initials of the constructs and the order in which they appeared within each part. This coding system was implemented to streamline future analysis processes. (Appendix A)

The survey was conducted on August 3, 2023, employing Google forms as the chosen mode of dissemination. It was distributed among multiple organizations that satisfied the predetermined criteria of the study, by presenting a succinct summary of the research subject, outlining the project's aims, and including a hyperlink for accessing the questionnaire. The task of distributing the questionnaire to the staff members who were most aligned with the survey's aims rested with these organizations, as they possess a superior understanding of the challenges their employees face. The survey was available for participation until September 15th, at which point it was subsequently closed and transferred to file in .csv format to commence the analytical procedure.

3.5. Data Understanding – raw data

A comprehensive dataset consisting of 112 records was collected and analysed, encompassing a total of 68 questions and 11 pre-established constructs. Given that all the items had been previously codified, the same rationale was employed in the analysis of the survey responses.

Upon initial observation, it became apparent that the data exhibited indications of high quality. This was primarily attributed to the utilization of a questionnaire, which facilitated the assessment process by incorporating validated scales and mandating the completion of all questions. The process of data extraction was straightforward, and no instances of duplicate or blank records were identified during verification.

After analysing the open-ended questions provided in the demographic section, it became evident that the records exhibited a lack of consistency stemming from the absence of multiple response questions. This phenomenon occurs in survey items such as D6 and D7, which inquire about the respondent's role within their workplace and the specific organization in which they are employed. These items can be phrased in several ways, leading to several alternative expressions that should be classified to enhance comprehension.

At last, it matters most that the many constructions, which are assessed through a specific quantity of objects, can be determined with accuracy. Therefore, to assess the reliability of each construct, it was employed the Cronbach's Alpha coefficient and conducted an analysis to determine the effect of removing certain items from the construct. This analysis aimed to enhance the construct's ability to consistently measure the intended concept. When analysing the set of constructs that fall under the category of professional perception (Appendix B), it is evident that certain items must be removed from the analysis of each construct to ensure that the data being considered exhibits a Cronbach's Alpha coefficient greater than 0.7. Taking into consideration the aforementioned, it was determined that there is a requirement to remove B10, CF6, CS7, and PS6. However, the constructs pertaining to Conflict with organizational values and Organizational Commitment did not yield a suitable value in this metric and so cannot be included for examination in the subsequent stages.

In turn, the removal of JC3 and PS6 from the analysis leads to favourable outcomes for each of the three constructs within the context of organizational conditions (Appendix C).

The same analysis was conducted on Turnover Intention (TI) (Appendix D), which was likewise measured using a set of items. The study's findings indicate that the construct's reliability was significantly enhanced with the exclusion of item TI3.

Based on the above findings and the selected analytical methodology employed in this investigation, the further requirements for data treatment have been identified. The first step involves classifying the data collected from D6 and D7. Next, any items that have a significant negative impact on the reliability of each construct must be removed. Then, the mean score for each construct is to be calculated, and turnover intention must be converted into a binary form for use in the modeling process.

3.6. Data preparation

As previously mentioned, there was no requirement to delete any records during the data preparation phase, as most of the questions were closed-ended and mandatory.

Nevertheless, it was crucial to address the ambiguous queries and establish a uniform methodology for managing the data. The open response records pertaining to the role and organization have been classified to assure the data's uniformity and relevancy. Five key categories have been defined for the purpose of defining the job description. The spectrum of positions encompasses the subsequent classifications: management, psychosocial technician, psychologist, jurist, and individuals who have chosen not to disclose their line of work. Regarding the aspect of Organization (D7), a similar rationale was employed, resulting in the identification of six distinct organizations. Additionally, a separate category was established to encompass those who opted not to reveal their organizational affiliation. The aforementioned action gave rise to seven categories, namely Aldeias SOS, APAV, Crescer, ILGA Portugal, Liga dos combatentes, O Companheiro, and undisclosed.

Following the identification of the features that should be excluded and the determination of the item that should be deleted, the final step comprised the computation of the mean for each construct. The results of this technique yielded novel variables that corresponded to each of the constructs.

Table 3.1. .Descriptive analysis of the constructs

Construct	Min	Max	Mean	Std. Deviation
Role overload (RO)	1	5	2,87	1,06
Job Clarity (JC)	3	5	4,47	0,53
Perceived Support (PS)	2	5	3,92	0,74
Organizational Justice (OJ)	2	5	3,56	0,67
Burnout (B)	1	3	1,82	0,71
Compassion Fatigue (CF)	1	4	1,54	0,76
Compassion Satisfaction (CS)	2	4	3,47	0,63
Turnover Intention (TI)	1	5	2,43	1,22

To establish a binary target variable, two additional variables were generated based on Turnover Intention (TI). Given the absence of literature providing clear guidelines for interpreting this scale, the first variable categorized values above 4 as indicative of a propensity for turnover intention, while the second variable categorized values above 3 as more likely to indicate a propensity for turnover. The study introduces two distinct binary variables that can be employed as target variables. The decision to create two distinct target variables considered the limited sample size and the low average score of the turnover intention data. Tin_Bin1, where the value of 0 denotes a low inclination to have turnover intention (shown by mean scores below 4), while a value of 1 signifies a high inclination to have turnover intention (indicated by an average score equal to or above 4). The second variable, Tin_Bin2, assigns a value of 0 to indicate a low likelihood of having the intention to turnover, with the distinction that only values below 3 are included in this categorization. And a value of 1 is assigned to indicate a high likelihood of turnover intention, represented by values equal to or greater than 3.

Table 3.2. Target description

Target	Characterization
Tin_Bin1	1= records with an average equal to or above 4 0= records with an average below 4
Tin_Bin2	1= records with an average equal to or above 3 0= records with an average below 3

3.7. Data Understanding – After preparation

Given that all the items had been previously codified, the same rationale was applied during the analysis of the survey responses, resulting in the accurate identification of each item by discerning its type, label, measure, and role.

It is vital to comprehend the demographic attributes of the respondents who participated in the survey. A total of 81 individuals, accounting for 72.32% of the sample, were identified as women, while only 31 individuals, representing 27.68% of the sample, were identified as men. The largest age cohorts were individuals aged 35-44, accounting for 31.35% of the responses, followed by those aged 25-34 (28.57%) and 55-64 (22.32%). In relation to parental status, most of the participants, specifically 59.82%, do not possess any offspring. About 50.89% of the population is married, while 35.71% of individuals are classified as single. Approximately 29.46% of the population has been affiliated with the organization for a duration exceeding 10 years, while 24.11% have been employed by the organization for a period ranging from 3 to 5 years. In relation to job responsibilities, managers constituted the predominant category (51.79%). It is noteworthy that these individuals are responsible for coordinating both work teams and social initiatives. This implies that although they are primarily not directly engaged in the field, they maintain a significant level of engagement with the issues at hand. In relation to the respective affiliations of individuals, APAV emerged as the prevailing group, accounting for 55.36% of the total.

The next step was to analyse the correlation matrix (Appendix E), and based on the conducted analysis, it can be inferred that the constructs of Burnout (B) and Compassion Fatigue (CF) demonstrate a statistically significant positive correlation ($r = 0.540$), indicating a positive association between the two of them. Moreover, TIN_BIN1 demonstrates significant positive associations with "B," "CF," and "CS," indicating a propensity for these variables to exhibit positive correlation. Conversely, the variable "TIN_BIN2" demonstrates significant negative correlations with both "CS" and "B," suggesting a propensity for these variables to exhibit inverse behaviours. In turn, Role Overload (RO) does not demonstrate any statistically significant correlations with the remaining factors included in the matrix. This discovery suggests that there is no significant correlation between the component of interest, RO, and the other factors under investigation. The aforementioned observation also holds for the case of the Job Clarity. Based on the statistically significant correlations observed, it is evident that the constructs of "B," "CF," and "CS" possess substantial predictive capabilities for "TIN_BIN1" and "TIN_BIN2."

3.8. Modeling

Upon understanding the data, this step entails utilizing IBM SPSS Modeler to engage in the modeling process.

Two modeling approaches were developed to investigate the factors contributing to the turnover intention of social workers from their professional positions. The first viewpoint posits that turnover intention is deemed noteworthy only when it produces elevated outcomes, whereas the alternative perspective regards the less recurring contemplation over leaving the business as an interim phase that holds relevance and merits consideration.

Considering Tin_Bin1 as the target: To build the prediction model, a holdout approach was employed to partition the data for the purpose of model selection, resulting in the identification of the most advantageous outcomes. The holdout technique for selecting among competing structural models entails the partitioning of a sample Y into two distinct subsamples, namely Y_e and Y_{ho}. The models are subsequently estimated using Ye (training group), which represents the data obtained from the control group. Subsequently, these models are evaluated and assigned grades according to their prediction efficacy in relation to specific components of the holdout sample Y_{ho}, the testing part (Schorfheide & Wolpin, 2012). The prevailing approach reported in studies using prediction models typically involves partitioning 70% of the data for training purposes and setting aside the remaining 30% for testing. Nevertheless, because of the constrained sample size (n=112), a partition ratio of 80-20 was also utilized. The choice to abstain from employing a validation group was motivated by a scarcity of available data. To ascertain the pertinence of the study findings, it was crucial to acknowledge the previously documented substandard mean value of the variable under investigation. This was achieved by balancing the data, resulting in a shift from a distribution of 77.68% for class 0 (representing individuals with no turnover intention) and 22.32% for class 1 (representing those exhibiting turnover intention) to 65.41% and 34.59% respectively. In the setting of the 70-30 division, it was seen that CHAID, QUEST, and C&R exhibited superior performance. The findings derived from the implementation of the 80-20 division strategy revealed that three algorithms consistently demonstrated improved performance.

*Considering Tin_Bin2 as the target: T*he decision to choose Tin_Bin2 as the defined target variable is considered a more appropriate choice for addressing the research question, since it offers a more balanced and fairer dataset for analysis. The participants in the study were divided into two classes based on their turnover intention. Class 0, which consisted of individuals with no turnover intention, accounted for 50.89% of the participants. On the other hand, class 1 comprised individuals who exhibited turnover intention and made up 49.11% of the participants.

Value	Proportion	%	Count
0.000		50.89	57
1.000		49.11	55

Figure 3.3. Target distribution Tin_Bin2

In a manner akin to the selected methodology for the target Tin_Bin1, an examination was also undertaken on the optimal holdout technique. This involved not only utilizing a 70-30 partition, but also implementing an additional partitioning approach where 80% of the data was allocated for training purposes and the remaining 20% was set aside for testing. The purpose of this analysis was to assess the comparative effectiveness of these two techniques. In the setting of the 70-30 partition, it was seen that the CHAID, QUEST, and C&R algorithms displayed higher levels of performance. Conversely, while employing the 80-20 division technique, only CHAID and QUEST algorithms provided superior performance. All models were instantiated using the default settings.

3.9. Model evaluation

To enhance the understanding of the effectiveness of the constructed models, a range of metrics were employed during the assessment of the confusion matrix. The notion under consideration pertains to machine learning and encompasses data pertaining to both the actual and expected classifications performed by a classification system. The confusion matrix is a two-dimensional matrix that consists of two dimensions. One dimension is associated with the actual class of an object, while the other dimension is associated with the class predicted by the classifier (Denga, Liu, & Mahadevan, 2016). Based on the confusion matrix the evaluation of each model involved the use of various measures, including accuracy, sensitivity, specificity, precision, and F-Score.

Accuracy refers to the ratio of correct forecasts to the total number of predictions made (Denga, Liu, & Mahadevan, 2016). It is determined by adding the number of true positive (TP) and true negative (TN) predictions, and then dividing this sum by the total number of data sets (P + N). The highest level of accuracy is 1.0, while the lowest level is 0.00 (Vujović, 2021).

$$ACC = \frac{TP + TN}{P + N} \tag{3.9.1}$$

Sensibility (recall) – True positive rate: The term "recall" refers to a metric that quantifies the efficacy of a predictive model in identifying instances belonging to a specific class within a given dataset (Denga, Liu, & Mahadevan, 2016). It is computed by dividing the number of correctly predicted positive instances (TP) by the total count of positive instances (P) (Vujović, 2021).

$$Recall = \frac{TP}{P} \tag{3.9.2}$$

Specificity – True Negative Rate: The measure of specificity is determined by dividing the number of true negative predictions (TN) by the total number of negative instances (N) (Vujović, 2021).

$$SP = \frac{TN}{N} \tag{3.9.3}$$

Precision is a metric used to assess the accuracy of predictions made for a particular class (Denga, Liu, & Mahadevan, 2016). It is determined by dividing the number of correct positive predictions (TP) by the total number of positive predictions (TP + FP) (Vujović, 2021).

$$PREC = \frac{TP}{TP + FP} \tag{3.9.4}$$

F-Score is a metric that quantifies the accurateness of a given test. The calculation is derived from the metrics of precision and recall (Vujović, 2021).

$$FScore = 2 \times \frac{PREC \ \times Recall}{PREC + Recall} \tag{3.9.5}$$

Another metric used to evaluate the model is the *Area Under the Curve* (AUC), determined the quality of the ROC curve Score is a metric that evaluates the performance of a model in terms of its ability to rank predictions. This metric quantifies the likelihood that a positive instance, chosen at random, will be ranked higher than a randomly picked negative instance (Vujović, 2021).

In the process of model selection, it is important to evaluate the several metrics. However, among these metrics, F-Score and Sensitivity have significant importance and carry the largest weight in determining the optimal model for this research. The rationale over this is that the F-Score incorporates both accuracy and sensitivity, specifically in terms of achieving a balance between the model's capacity to accurately identify true positives (employees who exhibit an intention to leave the organization) and its ability to minimize false positives (employees incorrectly classified as individuals with an intention to leave the organization). The selection of sensitivity relies upon its emphasis on the class of individuals who are inclined to depart.

3.10. Implementation and feedback

The implementation has not yet been carried out; however, given that the model has been developed, it is planned to conduct a session with a test group from the organization to evaluate its performance in real-world scenarios. Based on the input that will be received, subsequent adjustments will be made, and the proposed solution will be submitted to the board. This will serve the purpose of not only promoting its utilization, but also obtaining feedback from individuals who are directly involved in the operational aspects daily. It is acknowledged that this proposed approach is not definitive and will require diligent monitoring to ensure its continued relevance.

4. Results

The research investigation focuses on the elements that contribute to the turnover of social workers from their employment. The primary aim of this research is to examine the various elements that are associated with the experience of distress and the intention to leave employment among individuals working in non-profit organizations.

4.1. Selection of the model

To address the research inquiry, many categorization methods were devised, and the aforementioned metrics were employed to evaluate the optimal approach.

Considering Tin_Bin1 as the target: Using Tin_Bin1 as the target of each model using a 70-30 partition demonstrated evidence of overfitting in both algorithms, which could not be resolved by employing bootstrapping techniques.

Table 4.1. Model evaluation Tin_Bin1 partition1

	Training					
	Accuracy	Recall	Specificity	Precision	F-Score	AuC
CHAID	100%	100%	100%	100%	1,00	1
C&R	98%	95%	100%	100%	0,98	0,973
QUEST	94%	95%	93%	90%	0,93	0,92
	Testing					
	Accuracy	Recall	Specificity	Precision	F-Score	AuC
CHAID	100%	100%	100%	100%	1,00	1
C&R	94%	67%	100%	100%	0,80	0,929
QUEST	94%	67%	100%	100%	0,80	0,923

The CHAID model has a flawless performance across all criteria, indicating a conspicuous manifestation of overfitting. Overfitting occurs when models excessively memorize the training data, including the inherent noise, rather than effectively capturing the underlying patterns and principles embedded within the data (Ying, 2019). The models exhibit a significantly high level of accuracy in their observed performance, even when subjected to minimum modification. Achieving a perfect accuracy rate of 100% in predictive modeling within real-world contexts is a rather rare phenomenon (Ying, 2019). Both the C&R and Quest algorithms produce consistent outcomes in terms of the metrics used, and they both reveal the same issue that was seen in the CHAID Model. Based on what was observed, it can be inferred that none of the proposed methods yielded a satisfying outcome, indicating a significant requirement for further data to achieve a truly accurate prediction.

Table 4.2. Model evaluation Tin_Bin1 partition2

	Training					
	Accuracy	Recall	Specificity	Precision	F-Score	AuC
CHAID	100%	100%	100%	100%	1,00	1
Quest	86%	87%	86%	77%	0,82	0,858
	Testing					
	Accuracy	Recall	Specificity	Precision	F-Score	AuC
CHAID	92%	60%	100%	100%	0,75	0,913
Quest	73%	17%	90%	33%	0,22	0,543

When implementing an 80-20 partition, it is seen that while CHAID continues to exhibit the same issue as previously mentioned, Quest does not demonstrate this problem. However, the obtained results are unsatisfactory as they reveal a significant disparity between the metrics obtained during the training phase and those observed during the testing phase. Moreover, the testing outcomes are notably subpar. This is evident in, for example, Recall, that exhibits a rate of 87% on the training set and 17% on the testing set. The aforementioned findings suggest that Tin_Bin1 may not be the most suitable subject for examination, as further data would be required to construct a model that can yield accurate insights. Based on the findings, it has been determined that incorporating this specific target would not significantly contribute to resolving the research question. Therefore, it has been decided to exclude it from the current study.

Considering Tin_Bin2 as the target: The modeling attempts conducted using a 70-30 partition exhibited indications of overfitting in both algorithms, which could not be rectified through the implementation of bootstrapping approaches. The presence of a significant discrepancy between the metrics obtained from the training and testing sets is evident upon observation. It is uncommon to observe such elevated results on the measures, particularly in samples of such limited size (Ying, 2019).

Table 4.3. Model evaluation Tin_Bin2 partition1

	Training					
	Accuracy	Recall	Specificity	Precision	F-Score	AuC
CHAID	94%	89%	100%	100%	0,94	0,991
C&R	91%	85%	100%	100%	0,92	0,915
	Testing					
	Accuracy	Recall	Specificity	Precision	F-Score	AuC
CHAID	81%	75%	92%	94%	0,83	0,91
C&R	81%	78%	86%	88%	0,82	0,832

However, when a partition of 80-20 is implemented, there exists an algorithm that demonstrates exceptional performance.

Table 4.4. Model evaluation Tin_Bin2 partition2

	Training					
	Accuracy	**Recall**	**Specificity**	**Precision**	**F-Score**	**AuC**
CHAID	84%	75%	100%	100%	0,86	0,943
C&R	**84%**	**79%**	**89%**	**90%**	**0,84**	**0,846**
Quest	81%	77%	87%	88%	0,82	0,846
	Testing					
	Accuracy	**Recall**	**Specificity**	**Precision**	**F-Score**	**AuC**
CHAID	73%	67%	88%	92%	0,77	0,888
C&R	**81%**	**79%**	**83%**	**85%**	**0,81**	**0,885**
Quest	69%	67%	73%	77%	0,71	0,725

Although the CHAID approach exhibits the issue of overfitting, the use of QUEST yields satisfactory results, although not as favourable as those obtained using C&R trees. This method demonstrates consistent and satisfactory outcomes in terms of assessment metrics, as there is minimal disparity between the metrics of the training and testing sets. In terms of accuracy, the model exhibits an 84% ratio of accurately predicted cases on the training set, with a subsequent decline of 3 percentage points observed on the testing set. In general, the statistic demonstrates a positive pattern.

When considering the metric of Recall, it is observed that both CHAID and Quest algorithms yield inferior results on the testing set, with a performance of 67%. This performance is notably lower when compared to the alternative method, C&R Tree. Both sets of data used on the model demonstrate a recall percentage of 79%, suggesting a highly satisfactory conclusion in terms of the predictive accuracy for positive situations (Denga, Liu, & Mahadevan, 2016), particularly individuals who express an inclination to resign from their current job. Regarding the concept of Specificity, it is observed that CHAID encompasses indications of overfitting. Additionally, while the model employing the QUEST algorithm demonstrates a satisfactory outcome in terms of this metric (87% on the training set and 73% on the testing set), it is not as favourable as the performance exhibited by the C&R model. For this model the training set exhibited a rate of 89%, while the testing set showed a rate of 83%. recall. This indicates that the model demonstrates superior predictive ability for negative classes compared to positive classes, which is not the most desirable outcome (Denga, Liu, & Mahadevan, 2016). However, considering that recall yields good results, it is not indicative of a disappointing result.

When examining precision, similar findings to the prior measures can be inferred, indicating that C&R remains the optimal selection with a precision rate of 90% on the training set and 85% on the testing set. This implies that the proposed method demonstrates a high level of accuracy in forecasting affirmative cases (Denga, Liu, & Mahadevan, 2016).

By evaluating the F-score, C&R demonstrates a lower performance on the training set (0.84) compared to CHAID (0.86), albeit with a minor disparity. However, while analysing the outcomes of the testing set, it is evident that the technique exhibits a higher value of 0.81, in comparison to the values of 0.77 from CHAID and 0.71 from Quest. In this evaluation, the C&R model demonstrates superior performance.

In the last metric under consideration, the area under the curve (AUC), CHAID demonstrates superior performance compared to the other methods in both the training set (0.943) and the testing set (0.888). However, the C&R model demonstrates similar and highly satisfactory results, with an accuracy of 0.846 on the training set and 0.885 on the testing set. The QUEST method demonstrates substantially lower values if compared to competing approaches.

In general, the C&R strategy is perceived as the superior method and was chosen as the primary tool to address the investigation's topic. Regarding the metrics of recall and f-score, which were previously identified as crucial to the objectives of this study, the obtained results were highly satisfying. Consequently, no discernible opposition was put out to challenge this judgment.

4.2. Employee profiles indicating high and low propensity to exhibit turnover intention

Based on the conducted analyses, it has been observed that the most suitable criterion for drawing valuable conclusions in this study is TIN_BIN2. This criterion identifies employees who express an intention to leave the company when their rating in the corresponding dimension reaches or exceeds 3. Additionally, the evaluation metrics of the various models indicated that the utilization of a C&R algorithm was a suitable approach for addressing the research topic that motivated this study.

Regarding the significance of predictors, none of the constructs appear to possess a particularly high value. The construct that exhibited the greatest significance in forecasting turnover intention was Perceived Support (0.35), which was followed by D2, a demographic variable representing the age of the employee (0.27), and D5, which pertains to tenure (0.16). Additionally, OJ demonstrated greater significance than the remaining constructs. However, its importance score is extremely low at 0.05. Regardless the extremely low importance of the predictors utilized in C&R, the remaining models exhibited comparable values with an equally low importance. So, C&R was utilized in the study with a consciousness of this limitation.

Based on the model's extracted output (Appendix F), the following conclusions can be drawn regarding employees who exhibit a greater propensity to express intention of turnover:

- With 95,24 percent confidence and twenty supporting records, employees who are either 18 to 24 years old or 35-54 years old, have been with the organization for less than a year, between three and five years, or have more than ten years of service, and scored below four in perceived support, have a greater propensity to express intention of turnover.
- With 75 percent confidence and nine supporting records, employees who are between the ages of 25 and 34 or 55 and 64 and obtained a score of 3 on the Perceived Organizational Justice Scale are more likely to indicate intention to leave.

In contrast, the following conclusions can be drawn regarding employees who demonstrate a minor propensity to express intention of turnover:

- With 100 percent confidence and three supporting records, employees who are either 18 to 24 or 35 to 54 years old and scored 5 in the perceived support scale demonstrate a minor propensity to express intention of turnover.
- With 66,67 percent confidence and two supporting records, employees who are either 18 to 24 or 35 to 54 years old, have been with the organization for one to two years or between six and eight years, and scored below four in perceived support, demonstrate a minor propensity to express intention of turnover.
- With 100 percent confidence and twenty supporting records, employees who are between the ages of 25 and 34 or 55 and 64 and obtained a score of 2 or between 4 and 5 on the Perceived Organizational Justice Scale, demonstrate a minor propensity to express intention of turnover.

5. Discussion

The findings indicate, as prior research has also established (Wang, Jiang, Zhang, & Liu, 2021) (Barak, Nissly, & Levin, 2001), that an extended tenure in the organization (more than ten years) is a significant predictor of intention to leave; however, it was unexpected to discover that employees with less than one year of service are also inclined to contemplate quitting their positions. This disclosure is likely attributable to the fact that this profile contains individuals aged 18 to 24.

An additional noteworthy discovery was that perceived support did, in fact, exert a substantial impact on turnover intention, with scores falling below 4 being sufficiently influential to prompt employees to contemplate quitting their jobs. Most of the research on employee turnover has concentrated on global attitudes toward the organization or work as antecedents, as opposed to employee relationships with managers and others (Maertz, Griffeth, Campbell & Allen, 2007). However, some scholars believe that the reduction of voluntary turnover is initiated by leadership at the highest level, as leaders play a pivotal role in shaping organizational commitment and fostering harmonious enthusiasm among personnel (Scales & Brown, 2020). As is evident from the profiles of individuals with a low propensity for turnover intention, perceived support scores are exceptionally high.

It is evident that social workers who possess the necessary abilities to effectively adapt and perform in their professional setting tend to experience favourable results, as evidenced by heightened levels of job satisfaction and a reduced likelihood of contemplating or enacting job resignation (Acker, 2018). The recognition of the significance of self-care within the realm of social work has been acknowledged; yet there exists a scarcity of empirical research that provides evidence for its advantageous influence on both personal well-being and professional job outcomes. The lack of a clear and practical definition of self-care is a challenge in identifying strategies that effectively promote the well-being of persons within their professional settings (Kim & Stoner, 2008).

6. Conclusion and recommendations

The research was limited by a specific constraint, namely the relatively small sample size that could be incorporated into the model. The aforementioned limitation had an adverse impact on the model's quality and may have restricted the extent to which conclusions could be derived from it. It is worth noting that most of the collaborating organizations have a limited workforce, which has posed significant challenges in obtaining a substantial number of timely responses from them. One further constraint related to the scarcity of literature dedicated to the utilization of these specific techniques within this domain, hence impeding a comprehensive understanding of optimal practices.

Once one has become acquainted with the findings of the research, it is crucial to focus subsequent endeavours on conducting a comprehensive examination of turnover intention and the way in which its various dimensions evolve in their influence over time. The overarching objective is to cultivate a resilient workforce that is equipped with all the necessary resources and support to confront any adverse emotions that may arise because of their occupation. This will facilitate the delivery of a solution that is more pragmatic for all participating organizations, particularly APAV.

Bibliography

Abdallah, T. (2009). PrevalenceC2:C51 and predictors of burnout among Palestinian social workers. International Social Work 52(2), 223–233.

Acker, G. M. (2018). Self–care practices among social workers: do they predict job satisfaction and turnover intention? SOCIAL WORK IN MENTAL HEALTH, 16(6), 713-727. doi:10.1080/15332985.2018.1494082

APAV. (2021). Estatísticas APAV Relatório Anual 2021. APAV.

APAV. (2022). Missão & Visão. Retrieved from APAV: https://apav.pt/apav_v3/index.php/pt/apav-1/visao-missao

APAV. (n.d.). GAV - O Que São? Retrieved from APAV: https://apav.pt/apav_v3/index.php/pt/gav/o-que-sao

APAV. (n.d.). Onde Estamos | Contactos. Retrieved from APAV: https://apav.pt/apav_v3/index.php/pt/gav/onde-estao

Baird, S., & Jenkins, S. R. (2003). Vicarious Traumatization, Secondary Traumatic Stress, and Burnout in Sexual Assault and Domestic Violence Agency Staff. Violence and Victims.

Barak, M. E., Nissly, J. A., & Levin, A. (2001, December). Antecedents to Retention and Turnover among Child Welfare, Social Work, and Other Human Service Employees: What Can We Learn from Past Research? A Review and Metanalysis. Social Service Review, 75(4), 625-661. doi:10.1086/323166

Brown, A. R., Walters, J. E., & Jones, A. E. (2019). Pathways to Retention: Job Satisfaction, Burnout, & Organizational Commitment among Social Workers. Journal Of Evidence-Informed Social Work, 16(6), 577-594. doi:10.1080/26408066.2019.1658006

Caridade, S., Oliveira, A., Saavedra, R., Ribeiro, R., Santos, M., Almeida, I., & Soeiro, C. (2022). Psychosocial risks factors among victim support workers during the COVID-19 pandemic: a study with the Copenhagen Psychosocial Questionnaire. BMC.

Caringi, J. C., Hardiman, E. R., Weldon, P., Fletcher, S., Devlin, M., & Stanick, C. (2017). Secondary Traumatic Stress and Licensed Clinical Social Workers. American Psychological Association.

Cho, Y. J., & Song, H. J. (2017). Determinants of Turnover Intention of Social Workers: Effects of Emotional Labor and Organizational Trust. Public Personnel Management, 46(1), 41–65. doi:10.1177/0091026017696

CIG. (2022). Igualdade de género em Portugal [pdf]: boletim estatístico / CIG- Comissão para a Cidadania e Igualdade de Género. Lisbon: CIG.

CIG. (2023). Portal da violência doméstica: Indicadores Estatísticos. Comissão para a cidadania e a igualdade de género.

Conrad, D., & Kellar-Guenther, Y. (2006). Compassion fatigue, burnout, and compassion satisfaction among Colorado child protection workers. Child Abuse & Neglect, 30, 1071-1080. doi: 10.1016/j.chiabu.2006.03.009

Cooper, C. L. (1998). Theories of Organizational Stress. OUP Oxford.

Denga, X., Liu, Q., & Mahadevan, S. (2016). An improved method to construct basic probability assignment based on the confusion matrix for classification problem. Information Sciences. doi: 10.1016/j.ins.2016.01.033

Ditzel, E. (2008). A study of perceived occupational stress, burnout and sense of community among New Zealand nurses. doi:10.13140/RG.2.2.19261.15849.

Doherty, A. S., Mallett, J., Leiter, M. P., & McFadden, P. (2021). Measuring Burnout in Social Work. European Journal of Psychological Assessment 37(1), 6–14.

Farmer, D. (2011). Increasing Job Satisfaction among Social Workers. International Journal of Humanities and Social Science, 1(18).

Fc, B., & Roodt, G. (2013). The validation of the turnover intention scale. SA Journal of Human Resource Management, 11(1). doi:10.4102/sajhrm.v11i1.507

Gómez-García, R., Alonso-Sangregorio, M., & Llamazares-Sánchez, M. L. (2020). Burnout in social workers and socio-demographic factors. Journal of Social Work, 20(4), 463-482. doi: https://doi.org/10.1177/1468017319837886

Gorenak, M., Edelheim, J. R., & Brumen, B. (2020). The influence of organizational values on job satisfaction of employees. 39(3), 1-15. doi:10.3233/HSM-190781

Grządzielewska, M. (2021). Using Machine Learning in Burnout Prediction: A Survey. Child and Adolescent Social Work Journal.

Hamama, L. (2012). Differences between Children's Social Workers and Adults' Social Workers on Sense of Burnout, Work Conditions and Organisational Social

Support. British Journal of Social Work, 42(7), 1333-1353. doi:10.1093/bjsw/bcr135

Healy, K. (2022). Social Work Theories in Context: Creating Frameworks for Practice (3 ed.). Bloomsbury Publishing.

Hossny, E. K., & Mohamed, F. (2020). Role clarity as a predictor of nurses' job satisfaction. Assiut Scientific Nursing Journal, 8(21), 65-73. doi:10.21608/asnj.2020.29533.1016

Huang, J., & Ling, C. X. (2005). Using AUC and Accuracy in Evaluating Learning Algorithms. Ieee Transactions on Knowledge And Data Engineering, 17(3), 299-310.

IBM. (2023). C&R Tree node. Retrieved from IBM: https://www.ibm.com/docs/en/cloud-paks/cp-data/4.0?topic=modeling-cr-tree-node

Kim, H., & Stoner, M. (2008). Burnout and Turnover Intention Among Social Workers: Effects of Role Stress, Job Autonomy and Social Support. Administration in Social Work, 5-25. doi:10.1080/03643100801922357

Kumar, S., K.P., T., & Thind, N. S. (2020). Contextualized Validation Of Social Support Scale. European Journal of Molecular & Clinical Medicine, 7(7), 3257-3266.

Kumar, S., K.P., T., & Thind, N. S. (2020). Contextualized Validation Of Social Support Scale. European Journal of Molecular & Clinical Medicine, 7(7), 3257-3266.

Levenson, J. (2017). Trauma-Informed Social Work Practice. Social Work, 62(2). doi: 10.1093/sw/swx001

Logallo, N. (2019, December 30). Data Science Methodology 101. Retrieved from Medium: https://towardsdatascience.com/data-science-methodology-101-ce9f0d660336

Maertz, C. P., Griffeth, R. W., Campbell, N. S., & Allen, D. G. (2007). The effects of perceived organizational support and perceived supervisor support on employee turnover. Journal of Organizational Behavior, 28, 1059-1075. doi:10.1002/job.472

Maslach, C., Schaufeli, W. B., & Leiter, M. P. (2001). Job Burnout. Annual Review of Psychology.

Milanović, M., & Stamenković, M. (2017). CHAID Decision Tree: Methodological Frame and Application. Economic Themes, 54(4). doi:10.1515/ethemes-2016-0029

Ng, T. W., & Feldman, D. C. (2009). Re-examining the relationship between age and voluntary turnover. Journal of Vocational Behavior, 74(3), 284-294. doi: doi.org/10.1016/j.jvb.2009.01.004

Nishi, M., M. Y., & Matoba, S. (2021). Prediction of well-being and insight into worklife integration among physicians using life integration among physicians using machine learning approach. PLoS ONE.

Omar, A., Salessib, S., Vaamonde, J. D., & Urteaga, F. (2018). Psychometric properties of Colquitt's Organizational Justice Scale in Argentine workers. Revista Peruana de Psicología, 24(1). doi: 10.24265/liberabit.2018.v24n1.05

Omar, A., Salessib, S., Vaamonde, J. D., & Urteaga, F. (2018). Psychometric properties of Colquitt's Organizational Justice Scale in Argentine workers. Revista Peruana de Psicología, 24(1). doi:10.24265/liberabit.2018.v24n1.05

Peraica, T., Petrović, &. Z., Barić, Ž., Galić, &. R., & Kozarić-Kovačić, D. (2021). Gender Differences among Domestic Violence Help-Seekers: Socio-Demographic Characteristics, Types and Duration of Violence, Perpetrators, and Interventions. Journal of Family Violence, 429–442.

Ratzon, A., M. F., Ratzon, N., & Adini, B. (2022, may). Resilience at Work, Burnout, Secondary Trauma, and Compassion Satisfaction of Social Workers Amidst the COVID-19 Pandemic. International Journal of Environmental Research and Public Health.

Richard T. Mowday, R. M., & Porter, L. W. (1979). The measurement of organizational commitment. Journal of Vocational Behavior, 14(2), 224-247. doi: doi.org/10.1016/0001-8791(79)90072-1

Sánchez, F. C., Selle, E., Algarín, E. B., & J.L., S. (2022). Burnout, apoyo social, ansiedad y satisfacción laboral en profesionales del Trabajo Social. Interdisciplinaria, 39(1), 179-194. doi:10.16888/INTERD.2022.39.1.11

Santos, J., Sousa, C., & Gonçalves, G. (2022). Job Satisfaction Scale of Warr, Cook and Wall (1979): The psychometric properties of the Portuguese version. BPA. doi:10.26387/bpa.2022.2

Scales, A. N., & Brown, H. Q. (2020). The Effects of Organizational Commitment and Harmonious Passion on Voluntary Turnover Among Social Workers: A Mixed Methods Study. Children and Youth Services Review. doi:10.1016/j.childyouth.2020.104782

Schorfheide, F., & Wolpin, K. I. (2012). On the Use of Holdout Samples for Model Selection. American Economic Review: Papers and Proceedings, 102(3), 477–481. doi:10.1257/aer.102.3.477

Senreicha, E., Straussnerband, S. L., & Steen, J. (2020). The Work Experiences of Social Workers: Factors Impacting Compassion Satisfaction and Workplace Stress. Journal Of Social Service Research, 46(1), 93-109. doi:10.1080/01488376.2018.1528491

Silva, R. (2022, June 22). Número de crianças e jovens em risco aumentou. Retrieved 2023, from Renascença: https://rr.sapo.pt/noticia/pais/2022/06/22/numero-de-criancas-e-jovens-em-risco-aumentou/289320/

Stamm, B. H. (2009). Professional Quality of Life: Compassion Satisfaction and Fatigue Version 5 (ProQOL).

Tarshis, S., & Baird, S. L. (2019). Addressing the Indirect Trauma of Social Work Students in Intimate Partner Violence (IPV) Field Placements: A Framework for Supervision. Clinical Social Work Journal, 90–102.

Tham, P. (2007). Why Are They Leaving? Factors Factors Affecting Intention to Leave among Social Workers in Child Welfare. British Journal of Social Work, 37(7), 1225-1246. doi:10.1093/bjsw/bcl054

Thiagarajan, P. C. (2006). A Confirmatory Factor Analysis of Reilly's Role Overload Scale. Educational and Psychological Measurement, 66(4), 657-666. doi: 10.1177/0013164405282452

Thieme, A., Belgrave, D., & Doherty, G. (2020). Machine learning in mental health: A systematic review of the HCI literature to support the development of effective and implementable ML systems. Acm Transactions on Computer-Human Interaction, 27(5), 1-53. doi: 10.1145/3398069

Vicente, C. S., Oliveira, R. A., & Maroco, J. (2013). Análise Fatorial Do Inventário De Burnout De Maslach (Mbi-Hss) Em Profissionais Portugueses. Psicologia, Saúde & Doenças 14(1), 152-167.

Vîrgă, D., Baciu, E.-L., Lazãr, T.-A., & Lupșa, D. (2020). Psychological Capital Protects Social Workers from Burnout and Secondary Traumatic Stress. Sustainability 2020, 12(6).

Vujović, Ž. Đ. (2021). Classification Model Evaluation Metrics. International Journal of Advanced Computer Science and Applications, 12(6). doi:10.14569/IJACSA.2021.0120670

Wang, Y., Jiang, N., Zhang, H., & Liu, Z. (2021, may). Organizational justice, burnout, and turnover intention of social workers in China. Journal of Social Work, 21(3), 456-475. doi:10.1177/1468017320911347

Wells-English, D., Giese, J., & Price, J. (2019). Compassion Fatigue and Satisfaction Influence on Turnover among oncology nurses at an urban cancer centre. CLINICAL JOURNAL OF ONCOLOGY NURSING, 23(5), 487-493. doi:10.1188/19.CJON.487-493

Xiao, Y., & Watson, M. (2017). Guidance on Conducting a Systematic Literature Review. *Journal of Planning Education and Research, 39*(1). doi:10.1177/0739456X17723971

Yang, F.-H., & Chang, C.-C. (2008). Emotional labour, job satisfaction and organizational commitment amongst clinical nurses: A questionnaire survey. International Journal of Nursing Studies, 45(6), 878-887. doi:10.1016/j.ijnurstu.2007.02.001

Yarkoni, T., & Westfall, J. (2017). Choosing Prediction Over Explanation in Psychology: Lessons from Machine Learning. Perspect Psychol Sci, 1100-1122. doi:10.1177/1745691617693393

Ying, X. (2019). An Overview of Overfitting and its Solutions. Journal of Physics: Conference Series. doi:10.1088/1742-6596/1168/2/022022

Zhang, L., Zhang, T., Ren, Z., & Jiang, G. (2023). Predicting compassion fatigue among psychological hotline counsellors using machine learning techniques. Current Psychology, 42, 4169–4180. Doi: 10.1007/s12144-021-01776-7

Appendix

Appendix A: Questionnaire structure

Construct	*Code*	*Item*	*Scale*	*Source*
Role overload (RO)	RO 1	I must do things that I do not really have the time and energy for.	Never (1) Always (5)	(Thiagarajan, 2006)
	RO 2	I need more hours in the day to do all the things that are expected of me.		
	RO 3	I do not ever seem to have any time for myself.		
Job Clarity (JC)	JC1	I know what my responsibilities are well.	Never (1) Always (5)	(Hossny & Mohamed, 2020)
	JC2	I know exactly what is expected of me.		
	JC3	There are clear explanations and goals for what is expected of me.		
Perceived support (PS)	PS1	How easy is it to talk with your immediate supervisor	Don't have such person (0) Very much (5)	(Kumar, K.P., & Thind, 2020)
	PS2	How easy is it to talk with your colleagues?		
	PS3	How easy is it to talk with your friends and family?		
	PS4	How much can your immediate supervisor be relied on when things get tough at work?		
	PS5	How much can your colleagues be relied on when things get tough at work?		
	PS6	How much can your friends and family be relied on when things get tough at work?		
Organizational justice (OJ)	OJ1	To what extent the evaluation of my performance provides a good assessment of the effort I have put into my work	Never (1) Always (5)	(Omar, Salessib, Vaamonde, & Urteaga, 2018)
	OJ2	To what extent the evaluation of my performance provides an appropriate assessment of the work I have completed		
	OJ3	To what extent I can appeal the assessments made by procedures used in my organization		
	OJ4	To what extent, the procedures used in my organization uphold ethical and moral standards		
	OJ5	To what extent, my supervisor treats me with respect		
	OJ6	To what extent, my supervisor refrains from improper remarks or comments		
	OJ7	To what extent, my supervisor explains procedures thoroughly.		
	OJ8	To what extent, my supervisor communicates details in timely manner		
	OJ9	To what extent, my supervisor tailors his/her/their communications to my specific needs		
Burnout (B)	B1	I feel emotionally drained from my work	Never (1) Always (5)	(Ditzel, 2008)
	B2	I feel fatigued when I get up in the morning and must face another day on the job.		
	B3	I feel burned out from my work.		
	B4	Working with people all day is really a strain for me.		
	B5	I feel I'm working too hard on my job.		
	B6	I worry that this job is hardening me emotionally.		
	B7	I feel I treat some recipients as if they were impersonal "objects".		
	B8	I don't really care what happens to some recipients		
	B9	I feel recipients blame me for some of their problems		
	B10	I can't easily create a relaxed atmosphere with my recipients		
	B11	I haven't accomplished many worthwhile things in this job.		
	B12	I feel frustrated by my job.		
	B13	I feel that I'm not positively influencing other people's lives through my work.		

Appendix A: Questionnaire structure

Construct	*Code*	*Item*	*Scale*	*Source*
Compassion fatigue (CF)	CF1	I have outburst of anger or irritability with little provocation	Never (1) Always (5)	(Stamm, 2009)
	CF2	I think that I need to "work through" a traumatic experience in my life		
	CF3	I have experienced intrusive thoughts of times with especially difficult people I helped.		
	CF4	I am losing sleep over a person I help's traumatic experiences.		
	CF5	I have a sense of hopelessness associated with working with those I help.		
	CF6	I have been in danger working with people I help.		
Compassion satisfaction (CS)	CS1	I have good peer support when I need to work through a highly stressful experience.	Never (1) Always (5)	(Stamm, 2009)
	CS2	Working with those I help brings me a great deal of satisfaction		
	CS3	I plan to be a helper for a long time.		
	CS4	I have joyful feelings about how I can help the victims I work with.		
	CS5	I feel like I have the tools and resources that I need to do my work as a helper.		
	CS6	I have thoughts that I am a "success" as a helper.		
	CS7	I depend on my co-workers to help me when I need it.		
	CS8	I trust my co-workers.		
Organizational values (OV)	OV1	Quality of work is important in our organization	Never (1) Always (5)	(Gorenak, Edelheim, & Brumen, 2020)
	OV2	Adaptation to different situations does not presents a problem for our organization		
	OV3	Immoral behaviour at work is not acceptable in our organization		
	OV4	In our organization we try to satisfy the needs of the people we help		
	OV5	At work in our organization, we behave responsibly towards others around us.		
Organizational commitment (OC)	OC1	I am willing to put in a great deal of effort beyond that normally expected to help the organization be successful	Never (1) Always (5)	(Richard T. Mowday & Porter, 1979)
	OC2	I am proud to tell others that I am part of this organization		
	OC3	I couldn't just work for a different organization, even if the type of work was similar		
	OC4	For me, this is the best of all possible organizations for which to work.		
Turnover intention (TI)	TI1	In the past 9 months, how often have you considered leaving your job?	Never (1) Always (5)	(Fc & Roodt, 2013)
	TI2	In the past 9 months, how often do you dream about getting another job that will better suit your personal needs?		
	TI3	In the past 9 months, to what extent do the benefits associated with your current job did not you from quitting your job?		
	TI4	In the past 9 months, how frequently do you scan the internet in search of alternative job opportunities?		

Appendix B: Reliability analysis (professional perceptions)

CONSTRUCT	CRONBACH'S ALPHA	ITEM	CRONBACH'S ALPHA IF ITEM DELETED
BURNOUT (B)	0,868	B1	0,846
		B2	0,843
		B3	0,850
		B4	0,850
		B5	0,854
		B6	0,851
		B7	0,856
		B8	0,865
		B9	0,850
		B10	**0,880**
		B11	0,877
		B12	0,855
		B13	0,872
COMPASSION FATIGUE (CF)	0,849	CF1	0,813
		CF2	0,826
		CF3	0,809
		CF4	0,834
		CF5	0,804
		CF6	**0,858**
COMPASSION SATISFACTION (CS)	0,662	CS1	0,566
		CS2	0,579
		CS3	0,541
		CS4	0,650
		CS5	0,597
		CS6	0,642
		CS7	**0,729**
		CS8	0,695
PERCEIVED SUPPORT (PS)	0,729	PS1	0,634
		PS2	0,707
		PS3	0,713
		PS4	0,662
		PS5	0,670
		PS6	**0,739**
ORGANIZATIONAL VALUES (OV)	**0,575**	OV1	0,501
		OV2	0,575
		OV3	0,586
		OV4	0,488
		OV5	0,462
ORGANIZATIONAL COMMITMENT (OC)	**0,169**	OC1	-0,024
		OC2	-0,269
		OC3	0,641
		OC4	-0,030

Appendix C: Reliability analysis (organizational conditions)

CONSTRUCT	CRONBACH'S ALPHA	ITEM	CRONBACH'S ALPHA IF ITEM DELETED
ROLE OVERLOAD (RO)	0,838	RO1	0,820
		RO2	0,738
		RO3	0,758
JOB CLARITY (JC)	0,671	JC1	0,643
		JC2	0,281
		JC3	**0,806**
ORGANIZATIONAL JUSTICE (OJ)	0,811	PS1	0,811
		PS2	0,802
		PS3	0,795
		PS4	0,788
		PS5	0,795
		PS6	**0,812**
		PS7	0,773
		PS8	0,774
		PS9	0,778

Appendix D: Reliability analysis (Turnover Intention)

CONSTRUCT	CRONBACH'S ALPHA	ITEM	CRONBACH'S ALPHA IF ITEM DELETED
TURNOVER INTENTION (TI)	0,637	TI1	0,350
		TI2	0,385
		TI3	**0,878**
		TI4	0,433

Appendix E: Correlation matrix

Correlations

		RO	JC	PS	OJ	B	CF	CS	TIN_BIN1	TIN_BIN2
RO	Pearson Correlation	1								
	Sig. (2-tailed)									
JC	Pearson Correlation	-0,125	1							
	Sig. (2-tailed)	0,189								
PS	Pearson Correlation	-0,083	,416**	1						
	Sig. (2-tailed)	0,384	0,000							
OJ	Pearson Correlation	-0,045	0,156	,459**	1					
	Sig. (2-tailed)	0,636	0,101	0,000						
B	Pearson Correlation	,540**	-0,130	-,233*	-,298**	1				
	Sig. (2-tailed)	0,000	0,170	0,013	0,001					
CF	Pearson Correlation	,426**	-,230*	-0,067	0,129	,512**	1			
	Sig. (2-tailed)	0,000	0,015	0,480	0,176	0,000				
CS	Pearson Correlation	-,188*	,265**	,433**	,368**	-,513**	-,328**	1		
	Sig. (2-tailed)	0,048	0,005	0,000	0,000	0,000	0,000			
TIN_BIN1	Pearson Correlation	,312**	-,315**	-,438**	-,195*	,407**	,472**	-,405**	1	
	Sig. (2-tailed)	0,001	0,001	0,000	0,039	0,000	0,000	0,000		
TIN_BIN2	Pearson Correlation	,277**	-,202*	-,355**	-,240*	,348**	,226*	-,286**	,546**	1
	Sig. (2-tailed)	0,003	0,033	0,000	0,011	0,000	0,017	0,002	0,000	

**. Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).

Appendix F: Decision Tree