# Predicting the threat of death in Stalking cases through Artificial Neural Network

by

Simone Icardi

A thesis submitted to the
School of Graduate and Postdoctoral Studies in partial
fulfillment of the requirements for the degree of

**Master of Arts in Criminology**

Faculty of Social Science and Humanities

University of Ontario Institute of Technology (Ontario Tech University)

Oshawa, Ontario, Canada

October 2020

**THESIS EXAMINATION INFORMATION**

Submitted by **Simone Icardi**

**Master of Art in Criminology**

| |
|---|
| Thesis title: Predicting threat of death in Stalking cases through Artificial Neural Network |

An oral defense of this thesis took place October, 19 2020 in front of the following examining committee:

**Examining Committee:**

| | |
|---|---|
| Chair of Examining Committee | Dr. Leigh Larking |
| Research Supervisor | Dr. Steven Downing |
| Examining Committee Member | Dr. Karla Dhungana-Sainjiu |
| Examining Committee Member | Dr. Christopher O' Connor |
| Thesis Examiner | Dr. James Walsh, Assistant Professor |

The above committee determined that the thesis is acceptable in form and content and that a satisfactory knowledge of the field covered by the thesis was demonstrated by the candidate during an oral examination. A signed copy of the Certificate of Approval is available from the School of Graduate and Postdoctoral Studies.

## ABSTRACT

Stalking is a complex phenomenon that needs to be explained through several frameworks of research. During the years, scientists from the psychological, criminological, psychological, and legal fields made important steps to understand the nature of Stalking to find how to stop this type of behavior. The new concept that this research introduces is the possibility to predict if there will be the threat of death against a victim by a stalker through Artificial Intelligence and, more specifically, using an Artificial Neural Network (ANN). The thesis analyses variables impact on the learning process and the accuracy of this new system about predicting the threat of death. Publicly available and collected into the database *Violence and Threats of Violence Against Women and Men in the United States* built by Tjaden and Thoennes, the secondary data have been analyzed through the ANN. The result showed that variables impact fully support the theories, and the prediction of the threat of death against the victim is possible in this simulated environment with real data.



**Keywords:** stalking; threats; variables; machine learning; Artificial Neural Network

## AUTHOR'S DECLARATION

I hereby declare that this thesis consists of original work of which I have authored. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I authorize the University of Ontario Institute of Technology (Ontario Tech University) to lend this thesis to other institutions or individuals for the purpose of scholarly research. I further authorize University of Ontario Institute of Technology (Ontario Tech University) to reproduce this thesis by photocopying or by other means, in total or in part, at the request of other institutions or individuals for the purpose of scholarly research. I understand that my thesis will be made electronically available to the public.

_____ Simone Icardi

## STATEMENT OF CONTRIBUTIONS

I hereby certify that I am the sole author of this thesis and that no part of this thesis has been published or submitted for publication. I have used standard referencing practices to acknowledge ideas, research techniques, or other materials that belong to others. Furthermore, I hereby certify that I am the sole source of the creative works and/or inventive knowledge described in this thesis.

## ACKNOWLEDGEMENTS

To my supervisor, Dr. Steven Downing. It has been the first professor that recognize my idea as valid enough to support. He gave me the opportunity to develop an idea that it could be crazy or unrealizable most of the time. He provided tips, suggestions, and he spent big time helping me swim through problems and lack of time. I hope to keep work with you in our next step of research.

I want to say thank you to my committee members, Dr. Karla Dhugana-Sainju for your instant interest in my project and her help during the thesis writing. Dr. Christopher O'Connor has supported all my ideas about the cross-collaboration with Computer Science that could develop into a really interesting project for the near future at Ontario Tech University.

A special thank you to my classmates Lawrence, Riley, Nicole, and Allison. As an international student in his 30s, the impact of a new reality abroad to my Country was tough. They considered me a good classmate and a good speaker even though I wasn't good, and I killed the English language. It will be said to don't see you in class anymore.

In the end, thanks to my wife Angela for following me everywhere and trust me in all that I do even though the risk of unsuccess is high every time. A little thank you to my son Ryan for helping me to stay up during the night.

I thanks to my father and my mother for supporting me every time. It is not possible to write what it does mean for me your presence.

**TABLE OF CONTENTS**

## LIST OF TABLES

## CHAPTER 3

## LIST OF FIGURES

## CHAPTER 3

## LIST OF ABBREVIATIONS AND SYMBOLS

| | |
|---|---|
| AI | Artificial Intelligence |
| ANN | Artificial Neural Network |
| ML | Machine Learning |
| CBT | Control Balance Theory |
| U.S. | United States of America |
| RECON | Relationship Context-Based |
| PTSD | Post-traumatic stress disorder |
| NIJ | National Institute of Justice |
| NCIP | National Center for Injury Prevention and Control |
| CDC | Center of Disease Control and Prevention |
| NVAW | National Violence Against Women |
| NA | Not Available |

# Chapter 1: Introduction

There are two types of stalking: intimate stalking and non-intimate stalking. For a woman, the risk of being stalked converges with the category *intimate*. In most stalking cases, women are pursued by an ex-partner (Budd et al., 2000; Tjaden & Toennes, 2000). Cupach and Spitzberg (2007) found that the former partner is the most common category of the stalker observable. The category of *intimate stalking* includes cases where the victim is harassed by the former partner (Mullen et al., 2009). Victims of intimate stalking, where the former partner is the harasser, report violence more times than other types of stalking victims. (Purcell et al. 2000). Mohandie reported in 2006 that stalkers with a past relationship with the victim could become more violent and are more inclined to use weapons, abuse alcohol, and abuse drugs than non-intimate stalkers (Boehnlein et al., 2020).

Smith et al. (2017) report that 16% of women and 5% of men around the world have been victims of stalking at least one time in their life. Although women are more likely to be affected by it, stalking can affect anyone. While women more commonly experience stalking by men, men could be stalked in equal measure by women or men (Smith et al., 2017).

This phenomenon causes several significant negative consequences for victims. Stalking is associated with crimes such as homicide, assault, property damage, and related economic issues caused by forced lifestyle modification such as rescheduling work shifts or switching jobs (Logan, Walker, 2010; Logan & Cole., 2007). In addition to financial problems such as losing workdays, stalking victims report suffering physical health issues, physical violence, psychological disorders such as PTSD, depression,

stress, and mood, and panic disorders (Baum et al., 2009; DuMonthier, 2017; Geistman, 2013).

One of the most critical problems related to stalking is police reporting. Although stalking can cause several problems to a victim in all aspects of their lives, few cases are reported to the police (Boehnlein et al., 2020). Numerous studies affirm that unreported stalking cases account for 50% to 80% of all stalking cases (Baum et al., 2009; Fisher et al., 2002; Tjaden & Thoennes, 1998). There is a correlation between the perceived danger of the stalking case and the likelihood of it being reported to law enforcement. More dangerous situations are more likely to be registered (Campbell & Moore, 2011).

On the other hand, studies have discovered that women victims to a former partner tend not to report the case to the police (Reyns & Englebrecht, 2010). A stalking victim might avoid contacting law enforcement because of fear of possible retaliation from the stalker. Furthermore, most of the time, a victim does not believe that police intervention would help to solve the situation (Baum et al. 2009; Fisher et al. 2002; Tjaden and Thoennes 1998). Logan et al. (2006) report that 98% of women stalking victims do not believe police intervention will produce a positive outcome and 15% of women stalking victims believe that the related economic cost to having justice has been an obstacle in receiving support. For example, in order to obtain justice for this situation, a victim must hire an advocate, and she needs to meet several times with law enforcement, which could cost money and job time.

Another problem with stalking is the perception of it and decision-making process behind the analysis of a case from law enforcement and special dedicated services for victims. Law enforcement or victim services do not often perceive stalking cases as

dangerous to the victim (Botuck et al. 2009; Jordan et al. 2003; Logan & Cole 2007; Logan & Walker 2017). Moreover, law enforcement and victim support centers tend to present former partner stalkers as less dangerous than strangers (Sheridan et al. 2003; Scott & Sheridan 2011; Scott et al. 2010; Weller et al. 2013). Additionally, studies conducted by Tjaden and Thoennes (2000) in their database *Violence and Threats of Violence Against Women and Men in the United States*, Baum (2009), Jordan (2003), and Brady and Nobles (2017) highlighted that law enforcement response and justice processes are inefficient when managing stalking cases, which puts victims of intimate partner stalking at greater risk.

There is a considerable quantity of data about stalking. Databases, archives, and information collected through interviews and past research can be useful for better understanding of the phenomenon, its evolution, and its critical points.

The major problem is finding the best way to manage databases with more than 10,000 cases and more than 100 labeled variables. The difficulty is to analyze and understand each variable impact on the phenomenon and their weight about each variable and implications. In order to overcome this problem, social scientists and criminologists have started to use Artificial Intelligence and Machine Learning (ML). Machine Learning is the field of AI study that uses particular algorithms for making machines (such as computers, robots, etc.) learn from the information they process. One of the most powerful ML systems for deep learning is the Artificial Neural Network (ANN). Currently, it is easy to find tools for developing a simple ANN or ML system. The purpose of this thesis is to see if it could be possible to apply Artificial Intelligence for studying a complex phenomenon such as stalking from a database that has not been

designed directly for the task and lay the foundation for a more complex analysis in the future, with a more sophisticated algorithm and database. Moreover, the idea is to lay the foundation for developing new prediction models that could improve the decision-making process on this type of crime. Most of the time, it could be very tough to decide what is the best intervention for managing a stalking case. McFarlane et al. (2002) reported that intimate partners had murdered 80% of stalking victims. It seems that there is a problem during the decisional process about the risk evaluation of a case. At the end of the thesis, there will be a future perspective and ideas for a possible new, more complex system.

### 1.1 Complexity of stalking

Stalking is one of the most challenging issues in studies of criminology, psychology, and violence by law enforcement and victim service professionals (National Victim Assistance Academy, 2000). Percentages (McFarlane et al., 1999) estimates that stalking became more common than its onset in the past. Its consequences to victims are more inward than we can superficially imagine. The demographic characteristics of stalkers and victims are different than that of the general population. The complexity of stalking behavior and its motivations make it problematic to collocate in a single frame of the study.

Consequently, it is difficult to comprehend all the dynamics that comprise stalking, making it a difficult problem to solve. Professionals in criminology, psychology, and victimology started to develop intervention strategies based on their studies and experience with stalkers and stalking behavior (National Victim Assistance Academy, 2000).

In order to complete the experiment with the Artificial Neural Network, it is necessary to compose a reliable and robust battery of theories from the sociological, criminological, and psychological fields. It is impossible to categorize stalking behavior under a single research frame because several dynamics play an essential role in this phenomenon. In order to overcome this complexity and develop a system able to analyze and evaluate every aspect of stalking deriving from various fields of research, the variables selected for training the ANN are drawn from different theories.

From the psychological field, the classifications made by Mullen et al., the RECON system (Mohandie et al., 2006), and other classifications, as explained below, provide an essential list of observed behaviors that compose the obsessive part of stalking. Although classifications give a complete picture of stalking behavior, there is nothing mentioned in these models about the power and the level of control inside relationships between actors. Control Balance Theory by Tittle (1995) and its revision by Piquero and Hickman (2003) provide an essential explanation about the balance of control, power, and strength inside a relationship, and what control-balance environment could lead a person to become a stalker. There is evidence that violence is often correlated with stalking cases. Physical and psychological violence are recognized as one of the signatures of stalking cases. Classifications and control balance theory are not sufficient to provide a full overview of the phenomenon. The presence of criminal behaviors in terms of physical violence and psychological pressure needs to be analyzed differently. Agnew's (1992) strain theory focuses critically on circumstances that can lead to violence within relationships, including stalking. Strain theory examines aspects such as social context, situations, and events that could negatively develop the relationship.

All these aspects about stalking are represented by variables that I selected for this project from the *Violence and Threats of Violence Against Women and Men in the United States* study, and are fully supported by the literature review in order to avoid possible bias in judgment. The bias problem, variables, and theories are discussed further in this thesis.

## Chapter 2: Literature review

In the late 1970s, a pattern of stalking was identified among sexual harassment cases (Cupach & Spitzberg, 2004). At the beginning of the decade, stalking cases were without a proper definition, with cases being reported as rapes, obsessive behavior, and sexual harassment. After the 1990s, this behavior pattern has been defined as “stalking” (Lowney & Best, 1995).

Stalking was a phenomenon new to studies by Criminologists and Psychologists past century. Nevertheless, there are some notable cases which happened between 1980 and 1990. In literature, one of the most famous cases is about the Lennon homicide committed by Mark David Chapman. Chapman shot famous Beatles musician John Lennon on December 8, 1980, pushed by feelings of envy and believed he would gain popularity or notoriety. Chapman himself has confirmed this motivation in 2000. Another famous case is about John Hinckley, who attempted to assassinate American President Ronald Reagan by shooting him on March 30, 1981, with the purpose of gaining attention from actor Jodie Foster, whom he was stalking (Miller, 2012). In 1989, an obsessed fan killed television celebrity, Rebecca Schaeffer. He had purposely written numerous letters, followed her everywhere, and gave her several gifts for over two years. Because of the lack of a framework about this obsessive behavior, Schaeffer did not obtain any legal protection against her stalker. Stalking, as a crime, was not classified until the end of the 1990s (Boon & Sheridan, 2002).

## 2.1 Defining Stalking

Stalking is a complex phenomenon, and the traditional definition includes several typologies. Although there is a basic pattern of behavior observable in various types of stalkers, it does not mean that all the perpetrators become stalkers for the same reasons.

The definition of stalking originates from the University of Melbourne Center for Youth Mental Health, thanks to the research work of psychologists Mullen, Purcell, and Pathé and Mackenzie, who laid the foundations in 1999 for the recognition of stalking as a behavioral pathology. Psychologists tentatively outlined the frequencies and repetitions of the harassing behavior in relation to the victim's response (Mullen et al. 2009).

The first framework elaborated to understand stalking introduced three generic variables: an actor, a repeated series of behavior, and the victim selected by the actor. (Mullen et al., 2001). The actor identifies a person and begins developing an intense, aggressive attraction toward the object of desire (Mullen et al. 1999, 2001). The repeated series of behavior includes surveillance, stalking, and attempts at communication with the object of desire. In relation to these behaviors, there will be different types of responses from the target. After the response of the victim, there will be a couple-specific relational and communicative dynamic. These dynamics generate what is called "the ritualized obsessive behavior" (Mullen et al., 2001). The person identified by the actor perceives these behaviors as unwelcome and intrusive and associates them with a sense of fear and threat. At the same time, the stalker feels pleasure in thinking and seeing the other as afraid or distressed (Mullen et al., 2001).

Post-intimate stalking is a continuous and intrusive series of violations. The stalker is committed to the objective of having a relationship with an unwilling partner.

Undesired behaviors will become a crime when their persistence reasonably generates fear or serious apprehension in the victim (Finch, 2002).

Considering the similarity between the predatory behavior where the hunter analyzes and checks the environment in which he finds himself, chases it repeatedly until the final act, before catching prey, the term *stalking* has been used to refer to the actions of a prowler of a poacher (Oxford English Dictionary).

Meloy and Gothard (1995) used the term "obsessional follower" to describe particular cases. The individuals showed a series of behavior by pointing out the repeated nature of an obsessive attitude toward a person that could be a former partner, a friend, an acquaintance, or an unknown person.

Mullen (1999) redefined stalking as the set of behaviors that take the form of a series of repeated and continuous attempts, over time, aimed at finding communication and contact with a person who does not match the same interest. The modus operandi includes repetitive telephone calls, letters, SMS, e-mails, and messages written on the walls observable in public places frequented by the victim. Therefore, the final aim of the stalker is to establish contact and monitor the victim at the same time.

The Melbourne group (Mullen et al., 2002) subsequently delves into the meaning of the term stalking using the expression "harassing harassment syndrome," an expression also used by Curci in 2003 in his studies on stalking in Italy. Curci divides the stalker's behavioral pattern into three categories:

- Unwanted contacts
- Unwanted communication

- Associated behavior

Unwanted contacts encompass any way the stalker physically approaches the victim. It consists of the traditional following, approaching the victim's house, surveillance, and going to the public places where the victim is present. Unsolicited communication includes all attempts by the stalker to make direct or indirect contact with the victim (Mullen et al., 2002; Curci, 2003).

Unsolicited communication can escalate during the development of stalking. It starts with simple contacts to get information on the status or whereabouts of the victim. These attentions soon turn into real threats to the victim's life and may also affect the victim’s colleagues, family, and friends (Mullen et al., 2002; Curci, 2003).

Associated behaviors include actions likely to affect the life of the victim directly, such as delivering objects of various kinds to the person's address at late at night to instill fear. Additional behaviors may include the cancellation of the victim’s electricity and water supply contracts, destruction of private mail, revocation of credit cards, and closure of the victim's bank accounts (Mullen et al., 2002; Curci, 2003).

In recent years, with the development of studies of the phenomenon, stalking has no longer been defined only as a series of actions performed by the stalker towards the victim. It is also analyzed from the victim’s point of view. Mullen (2009) states that stalking occurs when the stalking-like behavior of an individual, of either sex, creates enormous discomfort in the victim or prevents them from freely deciding how to manage their daily life. In effect, stalking compromises the victim’s ability to control their own life.

The five profiles categorization elaborated by Mullen is still the best framework for classifying stalkers. The practical function of identifying these profiles is to categorize the criminal and develop a specific plan of intervention to allow operators to intervene in time, advise, and decide more effectively on how to deal with reported cases.

The *intimacy seeker* is the first profile. This profile includes stalkers driven by the need for a romantic relationship (Mullen et al., 2009). The individual exploits developed dynamics in strictly professional situations such as the relationship between patient and psychotherapist. The stalker looks for a victim who responds to his partner model. This happens after a superficial observation of the victim's characteristics (Mullen et al., 2009). The stalker thinks that the designated person needs help. Afterward, the stalker concludes that he will be able to solve the victim's problems (Mullen et al., 2009). In this imagined situation, success occurs when the victim falls in love with the stalker. The established relationship, even it does not exist in real life, pushes the stalker to distort their interpretation of any behavior by the victim that relates to him (i.e., a normal help request) (Mullen et al., 2009).

The stalker’s cognitive distortion exacerbates the situation. The stalker is convinced that a relationship has already started between the two individuals (Linehan, 1993). There are two types of courtship: veiled and direct. The evolution of the scenario depends on the subsequent action of the victim (Mullen et al., 2009). If the victim denies the relationship, the stalker denies the refusal, and subsequently, reinterprets the situation. The process leads the stalker to develop a wrongful belief, for example, that the victim's emotional unavailability is due to the awkward moment experienced, and there is a need time to overcome these problems. According to the stalker's mind, the victim will regain

lucidity, recognizing the closeness and help received by the stalker, and agree to establish a love relationship (Mullen et al., 2009). As the scenario progresses, the stalker's realization of the victim's refusal is seen as a forceful attack on the stalker's self-esteem. The stalker's psychological defenses reject reality, replacing it with an imagined scenario acceptable for his fantasy that could restore the lost self-esteem (Mullen, Pathé, Purcell, MacKenzie, 1999, 2001, 2006 2009). This category of behavior is included in the psychopathological form of this behavior defined as "erotomaniac delusion" (Leong, 1994). The stalker's mind eroticizes affection, and he perceives in the responses of the victim a sexual desire that does not exist (Leong, 1994).

The *incompetent suitor* is the profile of the classic stalker met among work colleagues, close friends, people who hang out in everyday life, schoolmates, or university peers (Mullen et al., 2009). The behavior of this stalker begins with an attempted courtship that can be considered limited enough to be considered harmless. It does not trigger any alarm to the victim. Over time, the stalker's attentions become more insistent and induce a feeling of discomfort to the victim with consequent social distancing (Mullen et al., 2009). The stalker's inadequate relational capacity makes him indifferent to rejection and less insistent in perpetuating the stalking behavior, which often ceases after a short time (Mullen et al., 2009).

If there are no significant psychopathologies, the stalking scenario ends in a short time, from a matter of days to a number of weeks. The typology is not considered dangerous to the victim's life (Mullen, Pathé, Purcell, MacKenzie, 1999, 2001, 2009).

*The predator* category is among the most violent recorded in literature. The stalker is an admirer who may have seen the victim at any time in everyday life. The condition

that there may have been some connection between the two people is not mandatory. It is not mandatory that the victim knew the stalker before the behavior begins (Mullen et al., 2009).

The stalker's desire for the victim develops into a sexual plan. In such cases, the stalker will try to bring the victims as close as possible to themselves. The most frequent dynamics are surveillance, pursuit, and induction of fear. The peculiarity of this category is the technique used by the stalker. He plans how to approach the victim as well as a hunter organizes a hunting trip. This type of stalker feels excitement by instilling fear in the victim, and the stalker often manifests psychological disorders or psychopathologies such as fetishism and pedophilia (Mullen et al., 2009).

This type of behavior can afflict individuals of different ages, including puberty, and can be a sign of serious psychological problems (Mullen, Pathé, Purcell, MacKenzie, 1999, 2001 2009).

The *resentful* is a type of stalker who generally takes the form of an ex-partner eager for revenge against the person who has decided to end the previous relationship, apparently without a justifiable reason (Mullen et al., 2009).

A resentful stalker's goal is often to destroy the victim's public image through a series of harmful actions that affect the victim's intimate and personal relationships. Dissemination media such as social networks, chats, and public places (both virtual and non-virtual) are used to damage the victim's reputation by, for example, publishing images of intimate moments on the web, divulging private conversations through various

types of chats and forums, or revealing the exact name and surname of the person and attaching a recognizable photo (Mullen et al., 2009).

Resentful stalker's actions are not limited to virtual places. The stalker might print and post photos and personal conversations gathered during the couple's life to disclose them in the areas frequented by the victim, such as around neighborhood, on shop shutters, or in workplaces (Mullen et al. 2000).

An escalation of aggression is observable in the behavior of the resentful stalker. The actions soon become physical. Property such as the victim's car, apartment, and personal belongings may be damaged by the stalker (Mullen, 2000).

This type of stalker is the most dangerous because he is unable to analyze reality rationally. He does not try to establish a real dialogue with the ex-partner, but rather he insists on this stalking behavior. His behavior is based solely and exclusively on the rancorous feelings brought up due to the end of the relationship and perceived abandonment by the victim (Meloy, 2006).

The *rejected* is a stalker who reveals himself as such in reaction to a possible refusal by the person with whom he is in contact. This type of stalking can occur in both ex-partners and suitors. In most cases, it is an ex-partner who tries to re-establish a badly ended relationship or tries to take revenge for an abandonment immediately (Mullen et al., 2009). Two simultaneous stimuli cause the instability of the stalker: the desire for revenge on the ex-partner and the desire to re-establish the previous relationship. Generally, he is not intimidated by the refusals or by the constant denial by the victim (Mullen et al., 2009). The established relationship between the two individuals feeds the

stalker's self-esteem and satisfies his desire for closeness. The stalker's thought towards the relationship remains positive as the stalker evaluates two possible scenarios: the end of the relationship and the stalking relationship. Most of the time, the final decision leads the stalker to prefer the existence of a bond rather than the total dissolution of the relationship (Mullen, 2000).

### 2.1.1 Other classifications

Another pillar for studying stalking from the aggression perspective is represented by the RECON method (Mohandie, Meloy, McGowan, William, 2006; Racine, 2014).

RECON is the acronym composed of RE: relationship and CON: context-based (Meloy, 2006), Racine 2014). This classification is used to address the stalking problem starting from the analysis of the previous relationship between stalker and victim (Racine, 2014). *Type I* stalkers had a previous relationship with the victim. By contrast, *Type II* stalkers had no prior relationship with the victim (Racine, 2014). The classification is based on a study of 1005 stalking cases recorded in North America. The data was gathered from the various security agencies and police departments. The analysis gave four groups of stalkers: Type I i) intimate, ii) acquaintance. Type II: i) public figures and ii) private strangers (Mohandie, Meloy, McGowan, William, 2006; Racine, 2014). The assignment of categories has a multifactorial basis. The considered factors are demographic, clinical, harassment, threats, and types of violence between groups (Mohandie, Meloy, McGowan, William, 2006; Racine, 2014).

*Intimate stalkers* are the most violent group. The stalkers belonging to this group often have a criminal record and abuse drugs and alcohol. After the first approach to the target, their behavior quickly degenerates. The stalking behavior, including the intensity

and frequency of contacts with the victim, increases over time. The escalation quickly turns into threats and insults (Mohandie, Meloy, McGowan, William, 2006; Racine, Billick, 2014). Of these individuals, 50% physically assault the victim, and one in three demonstrate suicidal tendencies. Intimate stalkers are more likely to be repeat offenders and offend faster than other groups. One in three use weapons to threaten their victim. The sample used for the analysis confirms that if there was a previous sexual relationship between stalker and victim, the risk of suffering sexual violence is 50% (Mohandie, Meloy, McGowan, William, 2006; Racine & Billick, 2014). The aggression is not due to previous severe psychopathologies but seems to be generated by some personality disorders or an insecure attachment to the mother (Miller et al., 2012). The growth of obsessive behavior among intimate stalkers appears during the first weeks after the break of the relationship (Mohandie, Meloy, McGowan, William, 2006).

Into the *acquaintance stalkers* category, 21% are female stalkers (Meloy & Boyd, 2003). The remaining cases show similarities with the behaviors of the first group. One-third of acquaintance stalkers assault the victim, cause damage to property, and engage in extreme stalking behavior. The duration of the behavior is approximately two years. If the relationship with the victim was intense, the risk of violence decreases. The biggest desire of these stalkers is to have or maintain a relationship with the victim. Emergency management protocol indicates that law enforcement officers must forbid the stalker from contact with the victim and pursue a path of mental health rehabilitation. For female stalkers, the most frequent diagnosis is *borderline personality disorder* (Meloy & Boyd, 2003).

The *public figure stalker* group includes 271 of the 1005 cases analyzed by Meloy (2006). Many of them come from the work of Dietz et al. (1991) regarding stalking against celebrities. 27% of the members of this group are female, and 30% of the victims are male. Although there is a more significant presence of female stalkers, the gender disparity in stalking continues to lean towards the activity being predominantly male.

Public figure stalkers are generally older, are less likely to have criminal records than the other three groups and are more likely to suffer from psychotic disorders. The majority show a history of severe mental disorders. The probability of an escalation of violence is low (Dietz, 1991; Meloy, 2006; Racine, 2014). Although the data collected by Dietz are quite outdated, they are parallel to Meloy's 2006 studies, and it appears that only 14% gave rise to violent behavior previously threatened. Violence towards celebrities occurs at an extremely low rate of 2% compared to the 74% violence rate of intimate stalkers. For this type of stalker, the most significant deterrent is probably due to the high quality of security measures provided for celebrities. People with mental disorders are unable to approach their victims of public relevance easily. Meloy (2003) noted that when the victim is a public figure, the violent stalker whose behavior fits in the predatory category uses a firearm. Risk mitigation must be managed by mental health professionals who must diagnose and perform targeted therapy to limit the likelihood the stalker’s behavior will escalate to violence (Meloy, 2003).

*Private foreign stalker* group records 10% of the cases analyzed by Meloy. The stalking starts in a private context, even though the stalker does not directly relate to the victim. Therefore, the characteristics of this group are similar to the characteristics of both the intimate stalker and the public figure stalker. Many stalkers in this category are

affected by mental illness, and 12% show a propensity to commit suicide. However, this class of stalker has a lower probability of abusing alcohol and drugs and a lower level of criminal intent than the intimate stalker; only a third of them go through the process of stalking and performing violence upon their victims (Meloy, 2006; Racine 2014). Of the 103 cases analyzed, 1/7 is violent, and the recidivism rate is moderate (25%). The risk management protocol provides a specific psychiatric treatment. The prevention of contact between harasser and victim is indicated when the potential for violence and the severity of the psychiatric disorder might be lethal (Meloy 2006; Racine 2014).

The RECON typology shows an excellent analytical ability regarding the likelihood a manner in which how various types of stalking lead to aggression over time. This study also confirmed the validity of the Zona (1998) and Mullen (2009) classifications as regards the behavioral developments of the stalkers and their observable characteristics (Meloy, 2006, Racine, 2014).

### 2.1.2 Victimology and social impact of stalking

Stalking is a crime that passes through several phases of harassment before exploding into violence. For most victims, the damage caused by stalking is not physical, but psychosocial. A stalking victim's sensation of powerlessness and fear could cause significant trauma that could persist for a long time (McEwan, Mullen, & MacKenzie, 2008). Unlike celebrity stalkers or acquaintance stalkers, ex-intimate stalkers maintain the pattern longer, causing more psychosocial and sometimes even physical trauma to the victim (Budd & Mattinson, 2000). In a stalking scenario, both perpetrator and victim change, both psychologically and mentally.

Stalking is socially dangerous because it denies freedom to the victim, who often adapts their lifestyle to avoid the stalker (Mullen et al. 2009). A stalking victim changes their lifestyle first, usually changing when they leave the house, their path to work, and limiting the people with which they spend time during the day (Mullen et al., 2009). Next, the stalking victim limits their social activities, avoiding preferred public places such as the gym, bars, parties, and other activities unrelated to working life (Mullen et al., 2009).

In more extreme cases, regular visits by the stalker could incite a victim to change jobs in order to find one that can be better adapted to their new routine (Mullen et al., 2009). As the victim's lifestyle is affected so greatly, additional medical costs could become a new problem. Stalking victims are often forced to go to specialized centers that provide methods of psychological therapy useful for overcoming the trauma of stalking, albeit at an extra cost that can be relatively high (Mullen et al., 2009). A person who has broken off a relationship is ready to make a new life with another person. The violence perpetrated by a stalker extends to the victim's relatives, friends, and new partners. Therefore, victims often prefer to prematurely end new relationships to avoid endangering the safety of their new loved one (Mullen et al., 2009). Additional costs come to the victim's tangible assets, most often to their cars. In these cases, various types of damage cause the victim to spend money on either repairs or even a new vehicle unfamiliar to the stalker (Mullen et al., 2009). Frequent attendance in court or police also affects the victim. The victim is often forced to go to the police stations to file a complaint. These kinds of events increase the victims' level of anxiety and mental distress (Mullen et al., 2009).

If the seriousness of the situation increases, and law enforcement intervention is not sufficient to maintain the victim safe from the threat of the stalker, some stalking victims take more radical measures. Sometimes, the victim will make changes to their appearance to lower the possibility of being recognized by their stalker. Many victims change their haircuts, the color of the skin, use tools to change their apparent height or change their type of clothing (Mullen et al. 2009).

If the situation becomes unbearable, the victim often decides for a drastic modification of life. A victim might move their residence to another area of the city, abandon the city for another one, or even migrate abroad (Mullen et al., 2009).

It is undeniable that this kind of coercive change causes damage to the mental balance of the stalking victim (Mullen et al., 2009). The stalking victim suffers significant psychological damage due to the situation. The most common symptoms found in the victims are high states of anxiety, fear, phobias, guilt, shame, and post-traumatic stress (De Fazio & Galeazzi, 2004).

Post-traumatic stress disorder is caused by the brain's response to threatening stimuli, which implement a series of biological processes. Observable symptoms of PTSD include avoidant behavior, increased alertness, and mental activation, the activation of the sympathetic division of the autonomic nervous system, and the adrenal glands' release of cortisol into the blood (Bear, Connors, Paradiso, 1999).

Women are most at risk of becoming a victim of stalking and are especially at risk of physical violence from former intimate partners (Budd & Mattinson, 2000). The level of violence registered on ex-intimate partners ranges from assault to homicide. The

percentage of women killed is 70%, and 85% is the percentage of attempted murders (McFarlane, 1999). It seems that their partners would stalk victims in the year before their murders (McFarlane, Campbell & Watson, 2002). In addition to that, there is a significant correlation between stalking, women murders, and attempted homicide (Cupach & Spitzberg, 2004). Studies reported that the average duration of post intimate stalking activity is 2.2 years (Tjaden & Thoennes, 1998). If stalking behavior lasts for an extended period, it is more likely to end in severe, even lethal, violence (Cupach & Spitzberg, 2004). Reliable data on predictors of stalking duration, including psychological pathology, are lacking, due in part to the nascent state of research on post-intimate stalking. Research focusing on the etiology of post intimate stalking, including the characteristics of intimate partner stalkers' personality, can provide an essential contribution to the management of an intractable crime (McCutcheon, Aruguete, Scott, Parker, & Calicchia, 2006).

### 2.1.3 Evolution of Stalking in Internet Era

The Internet rapidly increased its spread in this first part of the century, and people can access to the world wide web everywhere and everytime. Although the access to Internet on large scale has brought incredible benefit to society, it also open to new criminal opportunities (Pittaro, 2007). A new criminal figure raised into web in some cases as an extention of traditional stalker and in other cases as new stalker and its called cyber-stalker (Pittaro, 2007).

Cyber-stalker pursues the obsessive behavior through a wide use of technology (Petherick, 2007). The behavior is similar to traditional stalkers but the differences is the physical presence. If a traditional stalker must be physically close to the victim in order to

act harassments, a cyber-stalker does not need this risky part of the process. It is possible to harass a person through the use of several internet features.

As definition of cyber-stalking, Bocij and McFarlane (2003) provided this one: “cyber stalking as behaviors in which an individual, group of individuals or organization, uses information and communications technology to harass another individual, a group of people or organization. This action leads to emotional suffering to the victims. Cyber stalking is also defined as exploitation of telecommunication technologies such as pagers, cell phones, emails and the Internet, to bully, threaten, harass, and intimidate a victim” (p.139).

### 2.1.3 The impact of social network in Stalking phenomenon

Web 2.0 is the technology where thw wide social networking system is born. Social networks provide more social, collaborative and interactive experience of Internet surf (Haron & Yusof, 2010). After the evolution of the Internet in Web 2.0, this tool changed from a simple information organizer to a fundamental part of people's lives. Social Networking market introduced several platform for conneting people all around world: Facebook, Instagram, Twitter, YouTube, blogs and interactive websites (Haron & Yusof, 2010). This list of services was the first wave of innovation brought by Internet. After that, it has been the raise of a second type of interactive communication. WhatsApp, Telegram, Facebook Messenger, Viber, Signal, Skype, and other thousand communication systems appeared in people's lives for remaining and being permanently part of society.

Social networks enables the emergence of social networking medium which provides easy content creation and sharing among the users (Haron & Yusof, 2010;

p.237). Although one of the most important characteristics of social network is the transparency of contents and the free use for everybody, the dark side of this technology is represented by the uncontrolled exposure of user information (personal data, photos, videos, life details, live events, etc.) to cyber world and cyber-criminals.

In literature, there is evidence that explain how cyber-stalking starts to damage a person. Love obsession, a relationship broken, jealousy, desires of revenge, desires to contact or a simple friendship that becomes something more serieous in terms of feelings can lead a person to cyber-stalk. Motivations that lead people to cyber-stalk a person are the same for the traditional stalking with the auxilium of technology.

People that become victim of cyber-stalking or experience cyber-harassment in addition to traditional stalking adopt three different strategies. Three principal coping strategies are ignoring the cyber-stalker, confronting or sharing problems.

Haron and Yusof (2010) found that most victims prefer to share their problem to close friends or parents. What it seems to be missed is the report to law enforcement. This behavior directly generates the grey zone of stalking cases. This zone includes all those cases not reported but existent, creating gaps with the available data and the real number of the phenomenon. Ignoring or confronting the stalker is very rare (Haron & Yusof, 2010)

## 2.2 Criminological theories and perspectives

Stalking is not a simple and schematized phenomenon, and it is not possible to understand and analyze all the dynamics by applying only one framework of the study. The first part of the literature review examined all the scientific frameworks developed in the discipline of psychology. Pioneers such as Mullen and Zona traced the path for

analyzing this deviant behavior and put a frame around a phenomenon still new and hard to predict.

The second part of the literature review is about another essential cluster of stalking: social circumstances and characteristics observable in the stalking phenomenon. It is necessary to review a variety of extant criminological theories to attain a broad understanding of stalking and its dynamics. This section will introduce several theories that explain how stalking develops and the causes that lead a person to become a stalker. The self-control theory by Gottfredson and Hirschi (1990) provides an explanation about how a person becomes a victim by having a low level of self-control. Social learning theory by Albert Bandura (1971) explains how some behaviors can be learned socially, which could lead to becoming either a victim or a stalker, as well as the level of risk of acting in a deviant manner against someone. Control Balance theory proposed by Charles Tittle in 1995 (Piquero & Hickman, 2003) works similarly to the general theory of crimes but follows a different path of variables for framing the victimization process. The last theory reviewed is the strain theory (Durkheim, 1964; Agnew, 1992) that tries to connect the psychological world with the sociological perspective for providing a complete framework for studying stalking and the processes of both victimization and harassment.

### 2.2.1 Control Balance Theory: review

The dangerousness of stalking as a social problem increased because of the level of violence and physical assault increases. Tjaden (1997) found that 80% of stalking victims experienced physical attacks by a partner who later stalked them after the end of the relationship. At the end of 1997, studies reported that while there were 1.4 million cases of stalking, and the majority of offenders were male (Tjaden, 1997, U.S. Department of

Justice, 1998). The growing number of stalking cases over time alerted scientists and law enforcement to this phenomenon and its fast spread among society. After the classification and typology of stalking, one new goal of criminologists was to understand the motivation and the dynamics that push a person in a relationship to act violently and then stalk the victim. Meloy's research (2000) supports the hypothesis that stalkers have a history of failed heterosexual relationships. In addition to that, it has been clear that most stalking victims were sexually intimate with their stalkers (Meloy, Rivers, Siegel, Gothard, Naimark, Nicolini, 2000). In an older study, Meloy asserted that there was no correlation between domestic violence and the possible post-stalking behavior (Meloy, 1996). Newer studies have, however, found that stalking and domestic violence are indeed linked.

Criminologists, psychologists, and psychiatrists collaborated to classify stalkers by their personal traits and their relationship with their victims. The goal was to understand how to predict the probability of deviant behavior and the process of victimization. Stalking someone, surveilling them, and making repeated attempts at unwanted social and physical contact can be defined as abnormal. Tittle's control balance theory (1995) produced a new framework for understanding crime and deviance. The CBT analyzes, as the first step, the relationship between perceived control and the level of control to which one believes they are subjected (Tittle, 1995). The theory shows how the control dynamic can be imagined, like a ratio on a spectrum that measures the total control power exerted by two individuals. Starting to the negative side of the spectrum (e.g., the stalker has a negative level of control on the victim) to control balance, which represents the balanced level of power into exerted into a relationship (both individuals involved apply the same

level of power) to total control exerted over the victim (extreme level of power applied on the victim)(Higgins et al. 2005). According to Tittle's revised theory (2004), people with a deficit or surplus of control risk being more likely to commit crimes or deviant behaviors with the purpose of obtaining, hold or increase their control or power over other people (Nobles, 2013).

According to Piquero and Hickman (2003), people with a low level of control also have a low level of self- confidence and low self-protection ability. These skills are necessary in order to be safe from victimization. On the other hand, people who exhibit a high level of control gain a sense of invincibility. This sense of invincibility can be considered as a warning of a future perpetrator (Piquero & Hickman, 2003). Following this analysis, stalkers may be attracted to people who appear weak, submissive, and vulnerable; in other words: people with a deficit of control. An interesting observation made by Nobles and Fox (2013) is that individuals who show a high level of control may inadvertently attract harassers whose goal is to overcome the victim's level of control in order to improve their own personal strength. The application of CBT provides a partial explanation of how the level of control in a relationship affects the victimization dynamic. What it does not discern is the gender difference in stalking cases, nor is there a numerical analysis that can explain the level of risk that people with a high level of control could incur. Testing the CBT with a significant number of people with a high level of control with the goal of finding a correlation with the risk of becoming a stalking victim could improve the strength of the theory as it applies to stalking.

CBT has been empirically tested for finding correlations in gender differences. According to the U.S. Department of Justice (1998), most stalking cases involve a female

victim harassed by a man. The CBT applied by Nobles, Fox, and Fisher (2016) found a correlation between gender and stalking. Overall, the criteria of control balance dynamics did not find any association with stalking victimization. Their findings have been focused on stalking cases. The obtained results reveal the validity of the theory in cases where the victim is a female, and the offender is a male. According to the U.S. Department of Justice, the number of stalked men was over 400,000, and the CBT is not able to explain men's victimization in these cases.

### 2.2.2 Strain Theory

Strain theory is one of the most authoritative theories for explaining violence and criminal behavior. At the end of the 1990s, Durkheim introduced the term *anomie* for describing the reaction of people to the imposed control and morality by the surrounding society (Macionis & Garber,  2018)

Strain theory starts from Durkheim's idea of anomie (Langton & Piquero, 2007). Durkheim was a pioneer in criminological studies because he moved the focus of the observation from the classic psychological theory of crime to a possible social explanation of the problem. Durkheim (1964) described the anomie concept as an inappropriate level of social control perceived by people as inadequate. With this statement, he described the environment in which deviance and crimes flood society (Curran & Renzetti, 2001).

Durkheim (1964) theorized that a concatenation of social situations or empirical facts describes societal tendencies and individual qualities (Smith, 2008). He found a correlation between suicide rates and social variables from several statistics. As a result of this research, the highlight was that suicide could be egoistic or anomic. He also stated

that suicide happens when there is weak social integration and failed moral regulation (Durkheim, 1964; Smith, 2008). The following conclusion was that between religious people such as Jews or Catholics and protestants, intellectuals, and single people, the anomic suicide was more spread (Sminth, 2008). Individuals and people without connection with other persons showed a higher tendency for suicide compared to those full of connections to the community (Smith, 2008)

Robert Merton (1938) extrapolated a new application of the anomie from Durkheim's theory in order to explain the relationship between opportunities and the level of equality in resource distribution. According to Merton, the notion of anomie is the outcome of the unbalanced distribution of opportunities in the social structure because of the lack of equality in social supports (Bernburg, 2002). Durkheim states that anomie is a more complex concept. The absence of social authority is the cause of anomia because people continue to feel a sensation of bottomless insatiability. (Bernburg, 2002). Anomie is also generated when predetermined social objectives are impossible to reach. In order to achieve an unattainable goal, society condemns its people to a state of continuous unhappiness (Bernburg, 2002).

Robert Agnew (1992) revisited the Strain theory by approaching the problem of strains as an observable socio-psychological dynamic rather than a traditional social structure variable. In essence, from this new perspective, it is possible to state that strains are the cause of deviant criminal behavior. Adding a degree of separation between strains and criminal actions, and removing low socio-economical situation as strain generator, the new approach by Agnew changes the basis of the theory (Langton & Piquero, 2007). Agnew integrates into his new approach most of the strains that have a high correlation

with criminal actions (Agnew, 1992). The strains that show a high level of perceived injustice correlated with low social control are those which generate criminal behavior. (Agnew, 2001). General strain theory states that anomie causes several strains. The consequence is the production of negative emotions, including aggression and anger. Each person shows a different level of intensity of negative feelings, and this process is the cause of negative action as a coping mechanism. Only one coping process, conveyed by negative feelings, can be converted into criminal deviance (Agnew, 1992).

General strain theory is organized into two classifications, which are objective and subjective. Objective strains include those situations and circumstances that are typically experienced as unpleasant and unacceptable, such as repudiation, loss of jobs or friendships, and the conclusion of an intimate relationship (Agnew, 2006). Subjective strains include events and situations that induce a measurable level of negativity directly into the subject (Agnew, 2006). The breakdown of a relationship with a partner might be categorized as an objective strain. This type of strain would have a limited-time effect if its resolution usually occurs. However, the outcome might be different if the subject starts to experience distress and negative perception after the break of the relationship. According to strain theory, if this event generates anger, hate, resentment, or other negative feelings, the strain will be classified as subjective. Applying the strain theory to an analysis of the situation, the subjective strain could show a strong level of correlation with domestic violence and stalking (Froggio & Agnew, 2007).

Agnew (1992) affirms that negative cases where the risk of strains is present may lead to deviant behavior. Interpersonal violence, theft, and property destruction are the primary consequence of strains that could become criminal actions. (Agnew, 2002).

Furthermore, Agnew (2001) observed that frustration and anger are commonly-recurrent emotional responses to most life events. Childhood abuse, dashed expectations, injustice, unfairness, failed personal relationships, lay-offs, and other negative stimuli beyond a person's power to control could lead to criminal behavior.

The series of factors such as delinquent peers, sociopathic traits, and non-normative beliefs found by Agnew (1992) have been later statistically correlated with domestic violence and other criminal deviance (Langton & Piquero, 2007). Criminal actions can assume the shape of revenge against the source of strain. For the perpetrator, these actions and can be useful for releasing and overcoming negative emotions produced by the “strain” situation. Strains of greater magnitude could be resolved quickly through criminal actions such as stalking and domestic violence (Agnew, 2001, 2006).

### 2.3 Integrated theoretical model

The input variables that have been selected for this research are related to the theories reviewed. Part of the variables is supported by the literature review, and other variables have been selected because important for defining the profile of the stalking actors. The list of questions is reported in section Data and Variable and Appendix B.

The variable A8 about the stalking confirmation case clarifies that the victim did not experience family violence or sexual assault but stalking, and B2 about age is useful for defining the profile of the victim. The system will analyze all the variables, and it will calculate their weight on process learning.

Variables B4, B6, B7, D6, D7, D8, D11, D12, N2, N14 are variables that define the profile of stalker and victim in terms of social status and position. Strain theory gives

important information about the impact of social status, and the feeling of injustice correlates to criminal actions. The first strain theory includes a low socio-economical situation as a cause of possible deviant behavior. Furthermore, Agnew's theory of strain explains that the feeling of injustice correlated to low social control can lead to possible criminal actions. This theory suggests that the variables selected can provide a fascinating profile of the social status of the stalker that experiences an injustice during the relationship with the victim.

Variables E2, E3, E4, E5, E6, E7, E8, L39, N8, N13, and L40 are variables that design the relationship between stalker and victim and measure the different level of power internal to the relationship. Control Balance theory by Tittle led me to select this series of variables because they provide essential information about perceived control by the victim and what are actions lead stalker to control the victim showing jealousy, possession, provoking arguments, limiting the freedom. CBT explains that all these typical behavior observable in a stalking case can be generated by the loss of control from the victim. The stalker starts to be obsessed with victim distancing. In order to obtain back the control over the victim, the stalker would act, starting to destroy the victim's personality (i.e., question E7). According to the theory, all these patterns of action help design an important series of information for training the ANN about a final threat of death. The threat of death could be considered another way for taking back into the relationship the victim through the fear of die.

Variables E9, E10, E11, E12, E13, E14, H1A, L2, L16 are variables that give a frame to a stalker behavior. These series of questions provide every single possible action that stalker acts during the stalking period against the victim. Mullen and Al. and the

RECON system are the two classifications that support the presence of these variables because they accurately. These variables are a list of specific behaviors showed by an individual during stalking harassment. It includes sending numerous SMS, messages, and email, giving an unexpected and unwanted gift, and all those actions that define a traditional stalking pattern.

Variables D13, L8, L9, L12, L13, L14, L17, L43, L49, L52, N7, N10, N15 are variables that provide information about the perception of the victim on the situation. For example, the question D13 inquires about the abuse of alcohol that could lead to obsessive behavior, or question L52 that asks the victim is the harassment is over. It is important for building a complete model of AI to have also this information because they help to create a complete picture of the situation from the classic pattern behaviors listed before and the current situation at the moment of the report. For example, it was knowing a possible motivation that led a person to become a stalker, or if the victim left the apartment for finding peace. In literature, PSTD is recurrent in case of stalking, and it does mean that the harassment has been important. All this information connects other theories, and they could be essential for reaching the correct model for predicting the output.

# Chapter 3: Methodology

The purpose of this section is to outline the research design for predicting stalker behavior. Because the nature of stalking leads to observed violence in every instance, there will be no distinctions among the stalker categories (Mullen et al., 2009). An innovative methodological system is discussed for developing the analytical software responsible for measuring, analyzing, and predicting the output variables. Definitions of the dataset, variables, algorithm, input and output selection and statistical relevance of the analysis will also be discussed.

The dataset has been built with data collected between 1994 and 1996 under the supervision of the National Institute of Justice (NIJ), the National Center for Injury Prevention and Control (NCIP), the Center of Disease Control and Prevention (CDC) and sponsored by the National Violence Against Women (NVAW) survey. This dataset has been chosen because the most complete in terms of stalking case collection. An Artificial Intelligence (AI) system runs the analysis of the dataset and its variables. The Machine Learning (ML) field from AI science and its deep learning technology are applied to the study in order to perform AI training and for making predictions based on this experimental research.

The research is entirely led by the research question and not vice-versa (Shavelson & Towne, 2002). The goal of the study is to create a system able to predict some specific variables extrapolated from the secondary data. In this case, the application of machine learning and, more specifically, deep learning is appropriate for designing the research. The artificial neural network and its function, development, training, and application will be discussed below.

Then dataset selected as the source for developing the system has been used in several studies about stalking and domestic violence, including seminal studies in this field of research.

### 3.1 Conceptual Framework

Stalking studies across years show a specific series of behaviors that make the frame of this phenomenon. Mullen (2009) made a list of specific behavioral traits observable when a stalker starts to harass his victim. The RECON system and Mullen's categorization show that stalkers can be framed in four categories (Mohandie et al., 2006; Mullen et al., 2009), and all those categories show the same behavioral traits. Consequently, stalker behavior has a linear evolution during the process of harassment activity. The criminological theories provide an important framework for the influence and the effect of social status and about the level of power between stalker and victim.

### 3.2 Research Design

The research is structured to follow three important steps. The first step has to find a dataset that includes the variables selected for defining stalking and domestic violence. For this research, I selected the Violence and Threats of Violence Against Women and Men in the United States, 1994-1996. The selection falls on this database because, among the available datasets, it is the most complete in terms of questions related to the stalking literature.

The second step was to select the development environment for building the ANN. For this research, I decided to use RStudio software, including several powerful libraries and tools for programming ANN.

The thirst step was to program the ANN using the code and the rules of artificial intelligence. For this research, the ANN is a supervised neural network.

### 3.3 Artificial Neural Networks: introduction

One of the most ambitious goals created by science is to visualize knowledge through informatics. The human brain contains acquired knowledge, and the idea of informatics is to reproduce brain function using a computer. In order to create an intelligent machine, scientists tried to copy the brain's architecture (Géron. 2019). Initially, developers used to call the components of ANN with neuroscientific nomenclature, such as “substrate of neurons” for defining the layer of units that compose the first part of the ANN. Currently, words such as "neurons" have been substituted with informatics terminology such as "units" for identifying them (Géron. 2019).

ANNs are the most significant product of Machine Learning and Deep Learning because of their power to make predictions, their scalability, and their ability to handle large and sophisticated Machine Learning tasks. There are numerous examples of famous machine learning applications that people use every day, such as Google Images, Apple's Siri Assistant, YouTube’s Recommended Videos feature, and Google’s Deep Mind.

#### 3.3.1 Artificial Neural Networks

Artificial Intelligence can be considered one of the most advanced innovations reached by informatics and neuroscientific fields in this era.  Nonetheless, computer scientists have discussed ANNs since 1943. The first architecture of an ANN was presented by McCulloch and Pitts. It was a simplified computational model of neurons that simulated their biological functions (Géron. 2019). Between 1960 and 1970, scientists believed they were close to creating truly intelligent machines. The XOR

problem puts the ANNs in a dark era. The XOR or “exclusive OR” is a classic problem in the ANN field. This problem appears when scientists try to predict an output from two binary inputs. If the two inputs are not equal, the XOR function should give a *true* value. If the two inputs are equal, the result should be *false.* The situation changed in 1980 when the AI field gained interest and attention again, acquiring new architectures and algorithms. ANNs have recently seen renewed interest in several areas of research, such as economy, medicine, and social science.

Today, scientists and researchers have a vast quantity of data for training ANNs, and computing power has drastically increased in the last decade. Computers can make calculations faster than ever, and they can support heavy training with an enormous amount of data. In addition to this fundamental progress in technology, algorithms for training ANNs have improved in recent years, making this method of analysis and knowledge representation more efficient and accurate.

### 3.3.2 Biological and artificial neurons

In order to have a general idea about artificial neural networks, it is important to visualize where the concept of reproducing and simulate the brain circuits comes from. Biological neurons are cells present in every animal cortex. They are composed of a *cell body* containing the nucleus and other complex components and short extensions called *dendrites.* Also, there is a long extension called an *axon*. The axon splits off in several branches called *telodendria* that end at *synaptic terminals*. A synapse is connected to another neuron through its dendrites. Neurons receive short electrics signals which contain information. Neurons transmit information after collecting a certain quantity of impulses, at which point the receiver transmits its signal to another neuron.

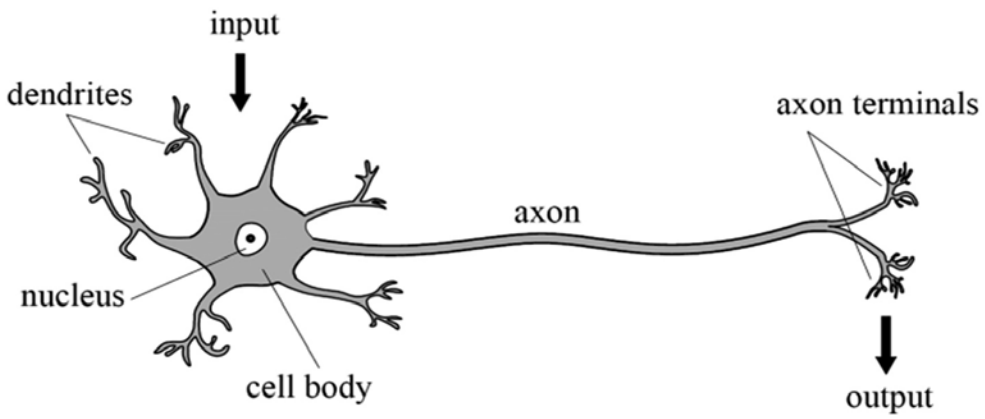


*Figure 1: Biological neuron*

The brain contains a vast network of neurons, in which each neuron is connected to thousands of others. Because the computational power of a healthy brain means it can potentially perform numerous complex and combined calculations, the brain has inspired the idea of artificial neural networks.

The ANN model is an adaptive system modeled following the functioning processes of the human brain (Grossi & Buscema, 2007). The model can alter its internal network in relation to a function objective. ANN systems are most often used to solve nonlinear mathematical problems (Grossi & Buscema, 2007).

Nodes, also known as process elements (PE), are the base for building an ANN. Each PE has its input and output, where it receives communication from other nodes and sends information to other nodes (Grossi & Buscema, 2007). Each node that composes the network has a function (*f*) that transforms its input into an output (Figure 2).

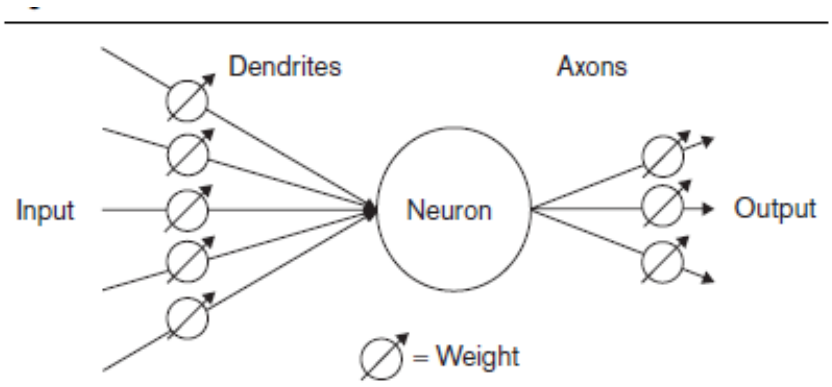


*Figure 2: Process Element (PE) containing input and output*

Every connection has a parameter which determines the inhibition or the activation. The strength of the connection determines the PE condition. Positive values result in excitatory connections. Negative values show inhibitory connections (Grossi & Buscema, 2007). The connections between the nodes have the property to modify their structure during the training process. This ability allows an ANN to begin the learning process. The process which determines the change in an ANN is called the “Law of Learning.” (Grossi & Buscema, 2007). The entire process is directly dependant on time. For allowing an ANN to modify its connections, it is necessary repeating the learning process with data. Data are essential core for teaching to an ANN how to change its structure in order to become more efficient and accurate. The learning process represents one of the critical mechanisms considered an adaptive processing system in the ANN field (Grossi & Buscema, 2007). This process called "learning" is a method that ANN uses to adapt the connections to the data structure.

The process of learning started by ANN on the environment generates a series of information collected into a log and printable at the end of the procedure, which is the

key for understanding the relationships between variables that compose environment Grossi & Buscema, 2007).

In other words, neural networks are adaptive, and they are able to find the fuzzy rules that connect various sets of data. At the beginning of the learning process, ANNs receive data following a specific set of rules. If they acquire new information at a later time, they are able to adjust their set of rules to integrate the new data with the old (Grossi & Buscema, 2007). This process occurs without external instruction.

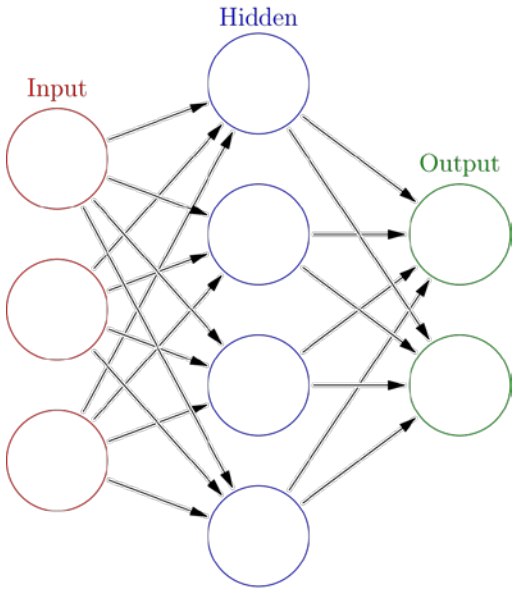


*Figure 3 Multiplayer Neural Network, Norwig, 2005*

An ANN continuously refines its understanding of input data to improve its knowledge of the problem under analysis. This process constantly updates and generates a dynamic bank. The system extracts from this bank a new and improved knowledge (Grossi & Buscema, 2007).

The ANN independently manages this process from the first categorization to a new and potentially more complex categorization. It uses new data to learn about the new possible category (Grossi & Buscema, 2007).

After the training with suitable data, the ANN finds the hidden rules which underlie a problem or phenomenon. A trained ANN can correctly generalize information it has never seen them before so that it can then predict and recognize similar information in the future (Grossi & Buscema, 2007).

### 3.3.3 Application of ANNs

The most typical problem that an ANN can deal with can be expressed as follows: given *N* variables (for which data collection is easy) and *M* variables (for which data collection is difficult), assess whether it is possible to predict the values of *M*-type variables on the basis of the *N*-type variables. When the *M* variables occur later in time than the *N* variables, the problem is described as a prediction problem. When the *M* variables another factor, the problem is described as a recognition problem. To correctly apply an ANN to this type of question, we need to run a validation protocol. We must start with a good sample of cases, in each of which the *N* variables (known) and the *M* variables (to be discovered) are both known and reliable.

The sample of complete data is needed to K train the ANN, and K assess its predictive performance

### 3.3.4 Description of the standard validation protocol

The procedure for compiling and validating an ANN has several stages.

In the first stage, the scientist splits the original database into two random sub-datasets. The first dataset, containing a majority of the data, will be used for training the

ANN. The second dataset, smaller than the training set, will be used for testing the ANN's prediction abilities.

The second stage is to choose the nature of the ANN (i.e., classification, categorization, decision trees, etc.), which will be trained using the training set. In this phase, the ANN learns to associate the input variables with the output variables.

In the final stage of the training process, the ANN generates a weight matrix table that contains the variable weights and all the parameters used for training the system.

The weight matrix, generated after the learning cycle shows the evaluation expressed by the ANN for each case, based on the previously carried out training. This process runs for each input vector and the output vector that it is unknown to the ANN.

The ANN is, in this way, evaluated only on the generalization ability that it has acquired during the training phase.

A new ANN is built with identical architecture to the previous one, and the procedure is repeated from the first stage. The final step is critical to find the correct architecture validated by the standard protocol; for this research, the procedure generated 100 ANNs with a different number of hidden neurons and different values of threshold.

### 3.3.5 Machine learning types: supervised and unsupervised

The machine learning system is known to have two major categories of learning: supervised and unsupervised.

Supervised learning includes those systems that the scientist feeds the algorithm with data that includes all the solutions. All the data have specific *labels.* A traditional type of supervised ML is used for classification. Another example of supervised learning

is the creations of systems used for predicting a numerical target, such as the increase in the cost of a car.

Unsupervised learning, on the other hand, does not need any labels in the solution dataset. In this category, the ML system receives unlabeled data. This learning process is most used for the clusterization problem or for finding a possible hidden pattern underneath the data. Clusterization is used, for example, for finding similar categories of data such as the preferences or typology of visitors to a specific store.

Because of the structure of the dataset that I used for this research, the best solution is to develop an ANN-based on supervised learning. I will give to the system all the information about the variable in terms of labels and typology of data.

### 3.3.6 Differences between statistical learning and Machine Learning

Statistics and Machine learning work for solving many of the same problems in data analysis. Both fields collaborate, and they have much to offer to a scientist. Although, for a junior user that starts with ML, it seems that there is no difference between these two fields, the first distinction recognized in data science is that ML is more practical than statistics and it is used for developing a model.

Both methods are data dependant, but if Statistics is formalized in the form of relationship between variables, Machine learning learns from data without the need for programming. Statistical learning works with small databases and a small number of attributes rather than ML can learn from billions of records and not necessarily labeled under a specific variable name.

Statistical learning must observe some strict rules for being efficient. Statistic research must respect assumptions as homoscedasticity, no multicollinearity or normality. ML ignores these rules working on predictions.

Math is essential in statistics because it is based on a coefficient estimator and requires a high level of reading data ability. ML has the ability, through algorithms, to detect patterns behind datasets analyzing the iterations between variables without the scientist effort.

Because this research is directed to predict an output, the choice to experiment with a model based on the machine learning algorithm seemed more appropriate than a traditional application of Statistics, especially because the heterogeneity of variables and their relationship could not follow the strict rules required in Statistics. Also, a secondary analysis will be on the weight of variables on prediction power and accuracy. This information will be extracted from the data bank generated during the learning process. It a prerogative of ML and ANN that it is not available in traditional statistics.

### 3.4 Research questions

This research aims to answer a specific question: is it possible, through the support of ANN, to predict the threat of death toward the stalked victim or someone close? The main purpose of this research is to experiment if it is possible to predict from a supervised dataset if a stalker will menace the death of the victim. The

In addition to this research question, is it possible to measure the weight of specific variables for predicting a potential threat of death in stalking cases? Does this question examine the extent to which variables such as race, financial situation, and alcohol abuse

are determinants for leading a stalker to threat of death the victim or someone close to? Stalking classifications and criminological theories for explaining the stalking phenomenon do not mention anything about racial background influence to the threat of death the victim. The research question for this variable has been formulated for confirming that stalking and its characteristics are independent of racial background. Financial situation and alcohol abuse are two variables reported in the control balance theory and strain theory as a determinant for the developing of threat toward the victim. Through the analysis of the ANN, I want to understand the real weight of these two variables in this specific model for learning and predicting the threat of death toward victims.

This research will try to answer these research questions Artificial Intelligence as innovative method of analysis.

### 3.5 Database and variables

The database built by Tjaden and Thoennes (1996) provides a massive amount of information about violence and threats against women and men. The database includes 14 sections that collect information about respondents' geographical information, age, and family situation. Sections are structured for treating aspects of violence such as fear of violence, accommodation behavior, relationship status (present and past), the current partner characteristics in terms of behavior, social and financial situation, relationship with power, level of control inside the relationship related to emotional abuse, rape, physical assault, the threat of death and stalking victimization and, at the end of the database, another detailed report about physical assault, rape, threat, stalking and violence in the victim’s current domestic relationship.

The database collected information 16,000 people in the U.S. who decided to respond. Complete interviews have been obtained by 8000 women and 8005 men residing in the U.S. The age of respondents started from 18 years old. The sample has been drawn dialing random phone numbers across the United States. Non-working and non-residential numbers have been deleted from the dataset.

According to the literature review and the principal study frameworks from the social science and psychological field, I selected 118 variables that collected the essential data for building the sub-dataset. Variables include all the aspects that it is possible to find in victim reports of stalking, incorporating possible episodes of harassment, physical and psychological violence, obsessive behavior, violence in the family, etc.

As has been previously explained, for training and predicting with an ANN, it is necessary to predict the output. According to the research question, the output that this ANN will try to predict is configurated into question L15, section: Stalking Victimization (Codebook, L-6). Data included in the variable L15 has been collected from the question: *Did [PERPETRATOR] ever threaten to harm or kill you or someone close to you when he/she was following or harassing you?*

The explanation of the output choice will be detailed late in the paper.

Following the questions from the codebook (Tjaden & Thoennes, 1996):

*Input*

- A8: Have you ever been stalked by anyone?
- B2: Age?
- B4: Respondent employment status;

- B6: Respondent level of education;
- B7: Respondent racial background;
- D6: Partner/spouse age;
- D7: Partner employment status;
- D8: Partner level of education;
- D9: Partner racial background;
- D11: Partner annual income;
- D12: Partner health status;
- D13: Partner frequency of alcohol abuse;
- E2: Has a hard time seeing things from your point of view?
- E3: Is jealous or possessive?
- E4: Tries to provoke arguments?
- E5: Tries to limit your contact with family or friends?
- E6: Insists on knowing who you are with at all times?
- E7: Calls you names or puts you down in front of others?
- E8: Makes you feel inadequate?
- E9: Is frightened of you?
- E10: Shouts or swears at you?
- E11: Frightens you?
- E12: Prevents you from knowing about or having access to the family income even when you ask?
- E13: Prevents you from working outside the home?
- E14: Insists on changing residences even when you don't need or want to?

- H1A: Mullen's matrix traits of traditional stalking
- L2: Traditional obsessive behavior (detailed)
- L8: Did these incidents happen while you were still involved with [perpetrator] or after your relationship with him/her ended?
- L9: Did you ever follow or harass [perpetrator] on more than one occasion?
- L12: Why do you think he/she started doing these things?
- L13: Did [PERPETRATOR] ever get someone else to help him/her follow or harass you?
- L14: Who was this person who helped him/her? He/she was following or harassing you?
- L16: Did [perpetrator] ever approach you (make himself/herself visible to you) when he/she was following or harassing you?
- L17: Did you ever believe you or someone close to you would be seriously harmed or killed when [perpetrator] was following or harassing you?
- L39: Did you get a restraining order against him/her as a result of these incidents?
- L40: To your knowledge, did he/she ever violate this restraining order?
- L43: Were criminal charges ever filed against him/her for following or harassing you?
- L49: How satisfied were you with the way you were treated during the court process?
- L52: Has he/she stopped following or harassing you?
- N2: Did you ever leave your current husband/partner because he/she was violent towards you

- N3: How many different times did you leave?
- N5: Where did you stay?
- N7: After you returned, did your husband/partner's violence towards you...
- N8: Did your current husband/partner ever leave you because he/she was violent towards you?
- N10 Did you have any children living with you when your husband/partner was violent towards you?
- N13 Did your husband/partner ever receive counseling for his/her violent behavior?
- N14: Was the counseling voluntary or mandatory (court-ordered)?
- N15: Do you think your current husband/partner's violent behavior towards you has stopped?

  Output
- L15: Did [perpetrator] ever threaten to harm or kill you or someone close to you when he/she was harassing you?

**3.6 The output variable. The importance of the threat of death in stalking cases.**

Two important factors have driven the choice of the output variable. The first factor is that this dataset has not been collected data from possible murder of victim or somebody close during the stalking period.

The second factor is that there is important evidence about the threat of death role in stalking cases that could lead to a possible homicide.

McFarlane et al. (1999) found that 76% of femicide and 85% of attempted femicide victims had experienced stalking. They reported that every year more than 5000 women

suffer life-threatening violence. In their study, they report that stalker left threatening messages on the phone 22% in femicide and 12% on attempted femicide. Stalker threats to harm kids 13% of the time in femicide and 11% of time in attempted femicide. Also, the harasser menaced the victim's family in 24% of the time in femicide and 31% of the time in attempted femicide. As of last, victims have been threatened with a firearm weapon 39% of the time in femicide and 40% of the time in attempted femicide cases.

The threatening behavior seems to play an important role in the development of stalking that could end with an attempted homicide or femicide.

The question L15 "Did [perpetrator] ever threaten to harm or kill you or someone close to you when he/she was harassing you?" includes threats received by the victim; I decided to experiment this model to predict a possible threat.

Because the idea of this project is to build a model that could be helpful for people that decide how to manage e stalking report before an attempted homicide or a homicide, the output about threats against victim and family or friends can be an important step for better understanding the evolution of the reported case. It could help to evaluate the dangerousness of the situation.

### 3.7 Sampling Design

From the original dataset, I extrapolated the columns that contain the information that I needed to utilize for training the ANN. The new dataset includes 16,000 cases with 119 columns. In order to make the process of programming easier and quicker, I changed the name of columns from the original code (i.e., A8) to a more comfortable acronym of variables (i.e., V1).

Following the rules of the learning process in Artificial Intelligence, I checked whether the column *OUTPUT* was complete. Unfortunately, many data were missing. For overcoming this problem, according to Rubin and Roderick (2002), when a dataset shows missing data, it is possible to deal with it through manipulation of the dataset explained in the next section *dataset manipulation*. Thus, I selected those cases where the outputs were known. The new dataset contains the same number of columns but with only 1,920 cases. ANN does not have an established number of cases for working. It does mean that 16000 or 1920 can work properly for reaching the final result. The learning process depends on the quality of data rather than the quantity. It could work better a smaller dataset with complete data rather than a bigger database with errors of recording or missing labels.

### 3.8 Dataset Manipulation

The major problem exhibited in the dataset was missing data. An ANN needs to be fed with data, avoiding incomplete cases. The original dataset presented numerous *NA* data that could create learning issues during the training of the ANN. For overcoming this problem, according to Rubin et al. (2002), the best way is to change *NA*, when data works in binary (YES/NO), with value *0*. Any weights connecting with the *NA* data will not influence the training, because $0 \times wij = 0$.

The second problem was that not all the 16,000 cases in the original dataset contained the output. It means that respondents did not answer the question L15, or data was not reported correctly. According to the properties of ANNs, a clear output must be present to facilitate the learning process. To fixed this problem, I discarded all the cases without a clear output. The number of cases available for training was 1,921.

Other manipulations were about the input section. The format of data collected in the original database was not compatible for training the ANN. The incompatibility was due to the fact that the data reported in each case was ordinary. It does mean that the number inside each variable was simply a position number. In order to make the data usable for this experiment, I changed them from ordinary numbers to a classic binary input. For example, the variable that reports the stalker behavior has several numerated sub-variables, I changed their value from the list number to 1 (if the behavior was reported as active) or 0 if the behavior was not active.

### 3.9 Development of the ANN

The development of the ANN called *iv4* was made in the RStudio environment. RStudio is a powerful tool for programming artificial intelligence because it contains important libraries useful for doing this kind of operation.

The first step was to upload all the necessary libraries for starting the development of the ANN. The libraries used for this experiment were *lattice, ggplot2, neuralnet,* and *caret*.

*Neuralnet* library is the package that allows scientists to develop supervised ANNs in this environment. Because the dataset contains labels, I used an algorithm for running supervised learning experiments.

For the second step, I uploaded the dataset into the developmental environment. After that, I selected the columns that contain the necessary variables, as described above.

The third step was to separate the OUTPUT column from the remainder of the dataset in order to prepare the correct output vector for the training.

After that, according to the methodology for programming and training ANNs, I was ready to split the database into two parts. Using the *function* for creating partitions, I split the dataset into two sides; specifically, 75% of the cases (variable input+output) was destined for the ANN training. The remaining 25% was set aside to be used to test the efficacy of the ANN.

The next step is to prepare the environment to find the best architecture for the ANN based on the dataset that I was using for this research. I generated a parameter for creating combinations in terms of hidden neurons and thresholds. Because this ANN is a multi-layer perceptron network, I needed to find the best number of hidden neurons and the best parameter for the activation threshold in order to reduce, at minimum, the number of errors generated during the learning process.

Now, I have all the objects ready for starting the research of the best architecture for the dataset. Creating a classical *for cycle*, I started the *neuralnet* function for finding the best structure for my research. Following a branch of the code:

```
for (i in 1:nrow(combinations)){
    intersectV4<-neuralnet(OUTPUT ~ v1+v2+v3+v4+v5+v6+v7+v8+v9+[…],
    data=t_set, hidden=combinations[i,2], lifesign="minimal", linear.output=FALSE,
    rep=5, threshold=combinations[i,1])
    iv4<-intersectV4$result.matrix[1,]
    error<-rbind(error,iv4)
    }
err_med<-apply(error,1,mean); err_med<-as.data.frame(err_med)
var_med<-apply(error,1,var); var_med<-as.data.frame(var_med)
arc_report<-cbind(combinations,err_med,var_med)
```

This process generates a table containing all the information about the possible combinations among the parameters imposed by the developer. For this research, because the number of variables is 118, I decided to test the architecture with 125, 140, 150, and 160 hidden neurons with a threshold set up between 0.1 and 0.9. Thresholds will be tested with an increase of 0.1 for every cycle of the test. For obtaining the final report, the ANN needs to cycle every solution that I proposed, and it must reach *convergence*. *Convergence* is the moment when the ANN stops generating learning errors.

The table shows that the best architecture for my ANN is 140 hidden neurons with a threshold of 0.2.

| **tshould** | **hidden_n** | **err_med** | **var_med** |
|---|---|---|---|
| 0.2 | 140 | 0.95316 | 0.163367 |
| 0.1 | 125 | 1.343193 | 1.310534 |
| 0.2 | 160 | 1.522434 | 0.59757 |
| 0.3 | 125 | 1.589476 | 1.386225 |
| 0.4 | 140 | 2.081355 | 0.775284 |
| 0.8 | 140 | 2.182033 | 1.746068 |
| 0.3 | 150 | 35.61544 | 6068.735 |
| 0.1 | 140 | 35.76308 | 6058.156 |
| 0.1 | 160 | 36.01764 | 6035.307 |
| 0.3 | 160 | 36.20385 | 6020.547 |
| 0.3 | 140 | 36.23722 | 6016.333 |

| | | | |
|---|---|---|---|
| 0.1 | 150 | 36.27454 | 5971.832 |
| 0.4 | 125 | 36.39036 | 5957.221 |
| 0.8 | 125 | 36.69526 | 5957.252 |
| 0.5 | 140 | 36.76167 | 5971.858 |
| 0.4 | 160 | 70.50717 | 9053.808 |
| 0.7 | 150 | 70.57264 | 9006.167 |
| 0.2 | 125 | 70.60001 | 9034.032 |
| 0.5 | 150 | 70.90849 | 8941.624 |
| 0.5 | 125 | 70.98192 | 8920.959 |
| 0.4 | 150 | 71.03008 | 9007.36 |
| 0.5 | 160 | 71.24284 | 8960.24 |
| 0.7 | 125 | 71.35052 | 8897.617 |
| 0.9 | 140 | 72.50771 | 8715.159 |
| 0.6 | 160 | 105.6153 | 9022.461 |
| 0.7 | 140 | 105.7978 | 8928.415 |
| 0.8 | 160 | 105.9147 | 8942.157 |
| 0.8 | 150 | 105.9169 | 8939.747 |
| 0.7 | 160 | 105.9746 | 8928.354 |
| 0.9 | 150 | 106.0565 | 8871.544 |
| 0.6 | 125 | 106.3203 | 8787.553 |
| 0.6 | 150 | 139.8619 | 5992.359 |
| 0.9 | 125 | 140.6798 | 5746.738 |

| | | | |
|---|---|---|---|
| 0.9 | 160 | 140.7254 | 5845.949 |
| 0.2 | 150 | 144.7016 | 55850.95 |
| 0.6 | 140 | 145.664 | 55562.08 |

*Table 1: Table generated by RStudio, Med Err, and architecture*

This setup generates a median error of 0.95316 and a median variance of 0.163367.

The obtained parameters will be inserted into the *neuralnet* function for building the ANN.

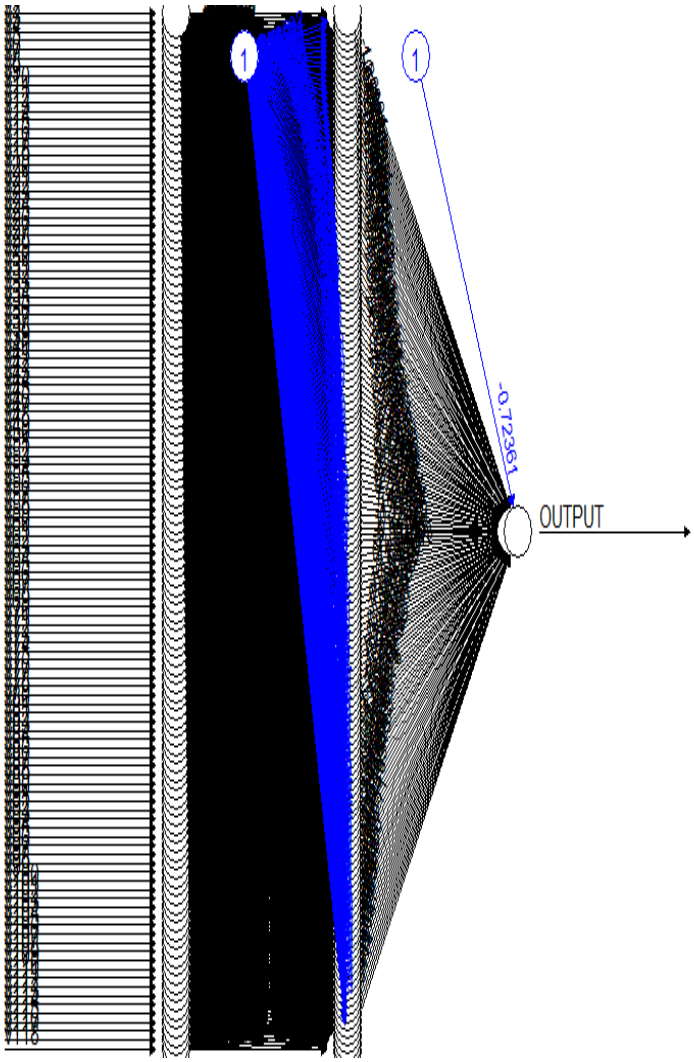

*Figure 4 iv4 best architecture, RStudio*

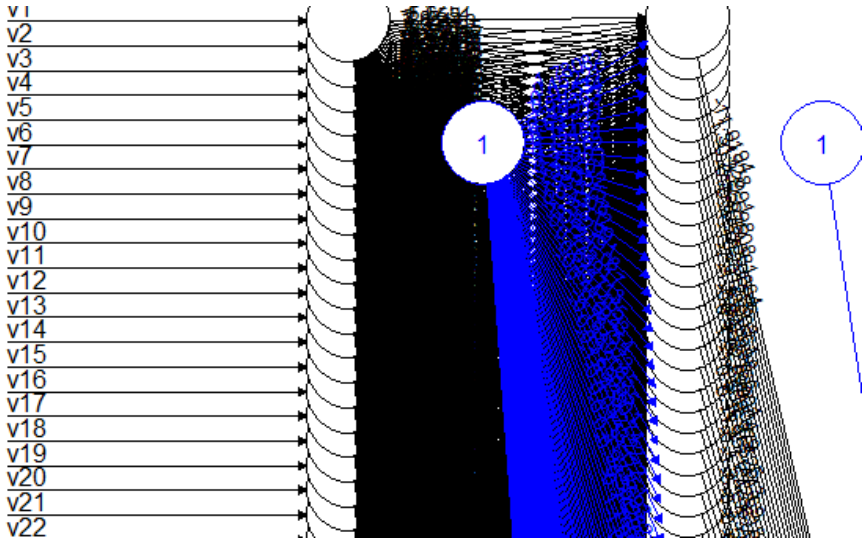


*Figure 5 Zoom on Input neurons, Rstudio*

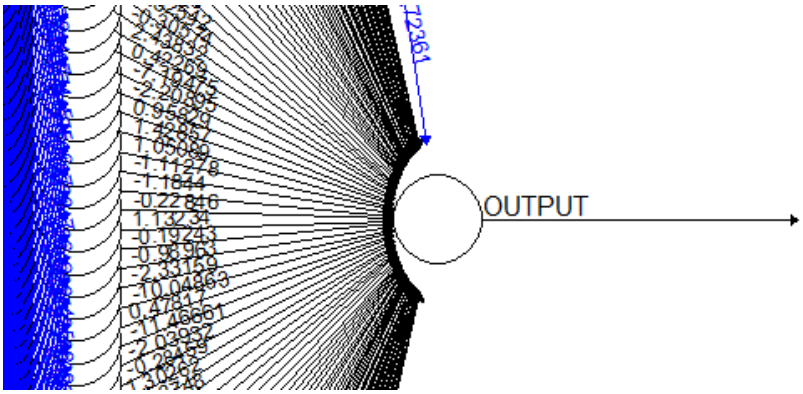


*Figure 6 Zoom on hidden neurons, Rstudio*

Upon completion of this process, the model was ready to be tested in order to verify its prediction power.

## Chapter 4: Data discussion

I previously split the dataset into two in order to have a comparison dataset. The dataset that I needed to use for testing the ANN needed to be complete, containing all input and output data. Unlike during the training phase, for the purposes of verification, the ANN will not see the real output, but it will generate a predicted output based on learned information obtained during the training process. The ANN makes its predictions about the possible output.

The predicted output is compared to the real output in order to measure the accuracy of the ANN.

| **Id case** | **Real Data** | **Predicted Data** |
|---|---|---|
| **4** | 0 | 1 |
| **6** | 0 | 1 |
| **8** | 1 | 0 |
| **12** | 0 | 1 |
| **13** | 0 | 1 |
| **20** | 1 | 0 |
| **25** | 0 | 1 |
| **27** | 1 | 0 |
| **29** | 0 | 1 |
| **31** | 1 | 0 |
| **39** | 1 | 0 |

| | | |
|---|---|---|
| **41** | 1 | 0 |
| **42** | 1 | 0 |
| **46** | 1 | 0 |
| | | |

*Table 2: Part 1 Comparison Table generated in RStudio*

---

| | | |
|---|---|---|
| | | |
| **674** | **0** | **0** |
| **676** | **1** | **1** |
| **687** | **1** | **1** |
| **691** | **0** | **0** |
| **706** | **0** | **0** |
| **708** | **1** | **1** |
| **710** | **0** | **0** |
| **715** | **0** | **0** |
| **716** | **1** | **1** |
| **717** | **0** | **0** |
| **719** | **0** | **0** |
| **727** | **0** | **0** |
| **735** | **0** | **0** |
| **758** | **0** | **0** |
| | | |

*Table 3: part 2 Comparison Table generated in RStudio*

The column *real data* reports the real output from the original database. In this research, 0 means that no threat of death has been reported, while 1 is the value which represents the presence of a threat of death during the harassment.

It is possible to observe from *Figure 8* and *Figure 9* how the ANN has committed mistakes in its prediction, missing the correct prediction.

Reading the whole table, we can see how the ANN correctly predicted the output for some other cases. The ANN has been tested for 480 cases. These cases compose the subset dedicated to testing the ANN. The *iv4* correctly predicted 350 cases on 480, with a total accuracy of 73%.

The subsequent analysis is about the weight of variables that I used for composing the dataset. The ANN has its internal data bank, where it calculates the influence of each variable on the learning process. The following table shows the positive or negative impact on the prediction power.

| **x.names** | **rel.imp** | **v72** | 0.132275 | **v10** | 0.039676 | **v85** | 0.001629 |
|---|---|---|---|---|---|---|---|
| **v101** | 1 | **v17** | 0.130013 | **v70** | 0.038911 | **v1** | 0.001577 |
| **v51** | 0.992355 | **v90** | 0.116117 | **v116** | 0.03168 | **v26** | 0.001277 |
| **v111** | 0.739635 | **v20** | 0.112448 | **v23** | 0.030481 | **v14** | 0.000453 |
| **v94** | 0.634172 | **v67** | 0.107873 | **v82** | 0.030326 | **v87** | 0 |
| **v99** | 0.632445 | **v91** | 0.0928 | **v29** | 0.025936 | **v105** | -0.00186 |
| **v98** | 0.558063 | **v40** | 0.086432 | **v28** | 0.025855 | **v6** | -0.00217 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| **v73** | 0.553011 | **v52** | 0.084173 | **v95** | 0.023892 | **v75** | -0.00269 |
| **v58** | 0.552251 | **v100** | 0.073449 | **v22** | 0.015692 | **v11** | -0.00373 |
| **v74** | 0.479842 | **v112** | 0.066842 | **v5** | 0.013593 | **v46** | -0.00389 |
| **v103** | 0.441549 | **v30** | 0.06618 | **v113** | 0.012762 | **v69** | -0.00492 |
| **v42** | 0.423248 | **v33** | 0.065064 | **v31** | 0.010741 | **v108** | -0.00648 |
| **v54** | 0.369165 | **v32** | 0.055041 | **v27** | 0.009349 | **v9** | -0.00657 |
| **v92** | 0.279709 | **v104** | 0.052573 | **v2** | 0.007994 | **v77** | -0.00672 |
| **v96** | 0.260454 | **v7** | 0.049329 | **v35** | 0.007264 | **v4** | -0.00872 |
| **v61** | 0.249698 | **v12** | 0.049017 | **v106** | 0.007066 | **v44** | -0.00998 |
| **v84** | 0.210658 | **v8** | 0.048536 | **v80** | 0.006199 | **v68** | -0.01125 |
| **v41** | 0.194916 | **v102** | 0.044006 | **v37** | 0.005902 | **v47** | -0.01133 |
| **v48** | 0.180766 | **v38** | 0.041699 | **v34** | 0.004277 | **v13** | -0.01162 |
| **v45** | 0.140296 | **v79** | 0.04085 | **v53** | 0.00427 | **v65** | -0.01175 |

*Table 4: part 1 Weight Table plot in RStudio*

| | | | |
|---|---|---|---|
| **v81** | -0.01492 | **v15** | -0.19395 |
| **v76** | -0.01507 | **v109** | -0.22092 |
| **v107** | -0.01607 | **v56** | -0.26683 |
| **v19** | -0.01884 | **v78** | -0.26948 |
| **v118** | -0.0236 | **v25** | -0.27766 |
| **v3** | -0.02709 | **v97** | -0.32001 |
| **v18** | -0.03424 | **v88** | -0.32981 |
| **v36** | -0.03704 | **v55** | -0.33043 |
| **v66** | -0.04254 | **v43** | -0.34444 |

| | | | |
|---|---|---|---|
| **v86** | -0.04346 | **v83** | -0.36353 |
| **v24** | -0.06019 | **v59** | -0.42343 |
| **v39** | -0.0617 | **v110** | -0.49513 |
| **v62** | -0.07303 | **v57** | -0.52149 |
| **v71** | -0.07724 | **v93** | -0.53891 |
| **v114** | -0.0827 | **v89** | -0.81425 |
| **v115** | -0.10264 | **v64** | -0.85958 |
| **v21** | -0.1056 | **v63** | -0.87031 |
| **v16** | -0.10608 | **v117** | -0.89671 |
| **v50** | -0.15987 | **v60** | -0.98016 |
| **v49** | -0.19113 | | |

*Table 5: part 2 Weight Table plot in RStudio*

The weight table (Figure 10 and 11) shows which variables positively or negatively impact the prediction. The series of numbers reported in the table is the same that is observable in *figure 7*. The range is between 1 and -1. This information helps to understand the variables that contributed to training the ANN and represent the strength of the connection between units. The variables close to 1 and -1 have a strong impact on the learning process in positive and negative. It means that weight brings down the importance of the input value. Weights near zero means changing this input will not change the output. Negative weights mean increasing this input will decrease the output. A positive weight means that increasing the input will increase the output. A weight decides how much influence the input will have on the output. Because all the variables that I used are binary, the weight must be interpreted in terms of presence. Positive inputs will move the output toward the 1 that represents the threat of death, rather than inputs

with negative weight move the output toward the 0 that represents the absence of the threat of death.

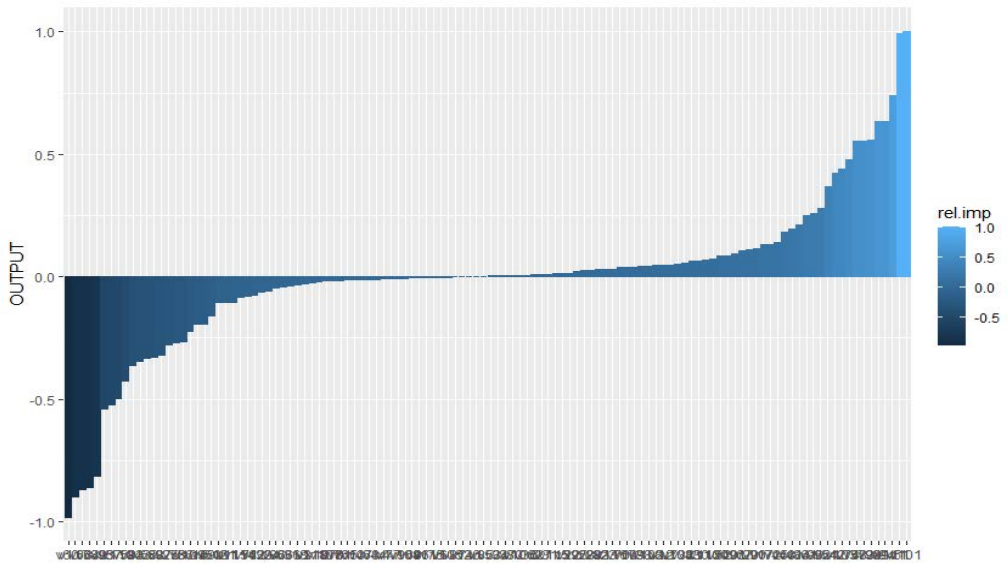


*Figure 7 Weight impact on the learning process, RStudio*

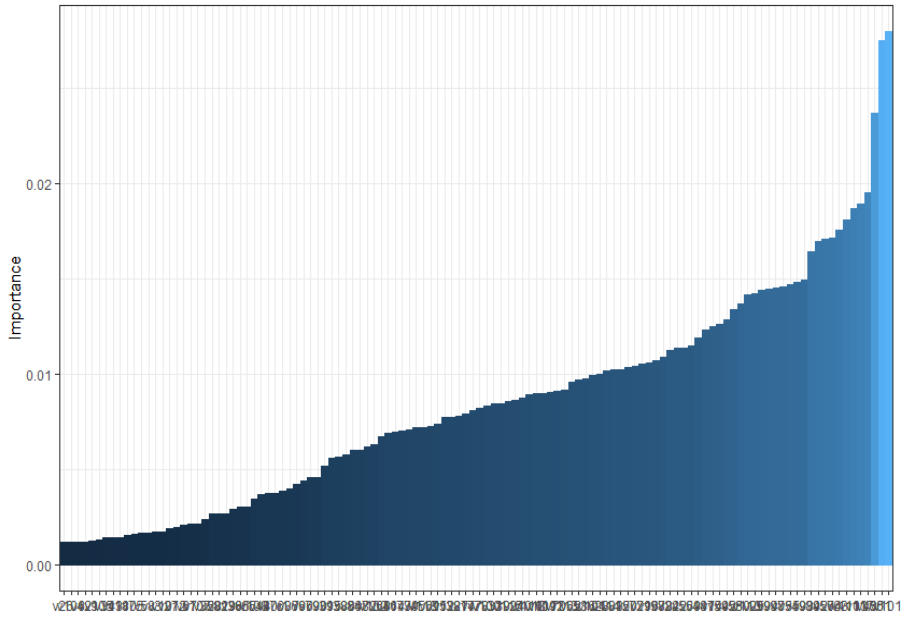


*Figure 8 Another representation of the impact of the weight on ANN learning*

One of the research questions asks if the stalker's racial background could be an important parameter for predicting a future threat of death. For the *iv4* system, this specific variable related to the whole process weights at -0.00657 (Figure 11).

Being close to the 0, according to the property of ANN and weight interpretation, the value of this variable is irrelevant in terms of learning to reach the correct output. Based on the assessment made by *iv4*, the racial background of the stalker is not correlated to a threat of death.

Another variable that I wanted to focus on is the employment status of the offender. Most of the time, violence can be observed in a context where the economic situation is problematic. Based on the *iv4* analysis, the variable *v7* (generated by the question D7) about the victim partner's employment, the weight is only 0.049329. Compared to the racial background variable, the impact is slightly more significant, but it is very close to the 0, and it does mean that it does not have an effect on the output prediction. The third variable worth paying attention to is alcohol abuse. In literature, the alcohol problem is associated with violence. *Iv4* system shows that the variable *v12* (question number D13) about the partner alcohol frequency, the weight is only 0.049017, very close to 0. As with the previous variables, alcohol frequency is not correlated to a possible homicide.

This information highlights that traditional variables studied in Criminology do not affect predicting threat of death in stalking cases.

### 4.1 Extra variable analysis

*Iv4* did not terminate its analysis to a simple prediction. The data bank matrix table generated during the training process by *iv4* shows each variable's impact on learning process (Figure 13). The first five variables that have an impact, positively or negatively, are v101 (L40 Question), v51, v60, v43 (L2 Question), and v117 (N14).

Question L40 is about the violation of a restraining order. The question asks the respondent if the stalker has ever violated a restriction order imposed by a judge (v101, L40 question). This variable has the heaviest impact on *iv4* process learning. According to the properties of weight in ANN, the presence of this behavioral trait moves the output toward the 1 (threat of death toward the victim).

The second group of variables (v51, v60, v43, L2 question) falls under the same question. The question is Mullen's framework of patterns in stalking behavior. It is not a surprise that this series of variables is crucial for the learning process because it is the main framework used in stalking studies. For the ANN, categorization, including these behavioral traits, are decisive for moving the output toward the 0 (absence of the threat of death) or 1 (threat of death).

The last variable is a little bit unexpected. The question is about court-ordered counselors. It seems that the intervention of a court positively influences the output. There is no evidence about a possible relationship between the threat of death and counselor order by a judge. This factor needs further analysis.

### 4.2 Further analysis and a new ANN model: iv5

The weight table is also useful for making a further selection of variables to improve the research. The purpose improves prediction accuracy can be achieved by

developing a new ANN using the information obtained from the previous training. With the weight analysis, it is possible to write a new, lighter dataset, ensuring it is more focused on the problem of stalking. For doing this operation, according to the weight table, I selected the first 50 variables that had an important impact during the learning process.

Applying the same code with new parameters and dataset, I obtained an ANN composed of 50 hidden neurons, with a threshold of 0.1. Further training and testing led to a 7% increase in predictive accuracy, passing from 73% to 80%.

The last extra analysis was to use the *iv5* to predict the remaining part of the dataset. As I mentioned above, a large number of cases did not have a clear output. *Iv5* can be used for predicting those missed outputs.

After the process of building the complete dataset with all the output for starting the training process, 14070 cases remained unused because of missed output. If they were not usable for training the network in the previous part of the development of the system, they could be used for finding the missed output.

Using the command *predict* in *neuralnet* library, the ANN, through its structure and after the learning process, starts to study cases. At the end, it will make a prediction of the possible threat of death against the victim (all the input data required are given to the ANN). I asked *iv5* to predict the other case outputs. At the end of the prediction process, I discovered that among the other 14079 cases, there is 80% of the possibility that 53 cases ended with threats of death against victims.

### 4.3 Finding discussion

Before discussing the findings, my first assumption is that this has been an experiment with a new methodology for obtaining new knowledge on the stalking phenomenon. I will not repeat that every finding is related to this simulated environment, with all the limitations highlighted in the next section.

Theories and literature support all the variables that compose the experiment with this intelligent system. Technically, findings should confirm or not confirm views. This methodology that I classify as new in this field of studies does not follow the traditional rules. The analysis made through the ANN *iv4* and *iv5* provides interesting addition about stalking for all the support theories that I summarized in the literature review section. Stalking is a complex phenomenon that includes several research fields, such as Psychology, Criminology, and Sociology. For supporting this research, I started from the psychological perspective introducing several classifications. The Melbourne team and the RECON system provide their framework about every type of stalking without gender differentiation. Also, minor classifications follow this type of structure for understanding stalkers. It is common to associate stalking with the idea of men against women. This association happens because there is evidence in the literature that the most famous stalker type is the former partner or the intimate partner. As we know, the recognized types of stalkers are five, and they could be equally dangerous for the victim. In order to avoid a gender bias into the ANN learning process, the number of cases is equally distributed between genders because no evidence in the literature defines stalking as a gendered phenomenon.

The first claim that I have to do is that in positive or negative, all variables except one have an impact on defining the risk for a victim to receive a threat of death. Starting from the variable that does not influence this system for predicting the risk of death threats, figure 11 shows that v87 has no impact on the learning process. Variable v87 is a column of the question that highlights if a perpetrator has ever approached the victim during the harassment. Mullen (2009) and Meloy (2006) include the victim's approach in their frameworks for classifying stalker behavior. Approaching the victim during the harassment can also be explained by the Control Balance Theory (Tittle, 1995). This type of pressure imposed by the stalker on the victim is a sort of control that unbalances the relationship between victim and harasser. The first look to this result suggests that approaching the victim does not play a role in harassment and consequent dangerous behavior.

I went deep on this specific result, and I found that variable v87 is one of the five columns that collect data for the question L16 (v85, v86, v88, and v89). Reading the table of weight in figure 11, we can see where other variables are allocated. The v85 is close to variable 87, with a small positive impact on the death threat's risk. The other variables have a negative effect on the output. It does mean that this behavior limits the risk of the death threat. Including this variable seems very important because it confirms that classifications include an essential trait for defining the stalker type. In addition to that, it seems critical to understand how the control on the victim (Tittle, 1995) managed by approaching during the harassment could re-establish the balance that the stalker felt lost and avoid possible verbal violence. The negative impact suggests that approaching the

victim lows the risk of the death threat. The limitation of the dataset section can explain the variable that obtained a null impact on the phenomenon.

Clarified the variable with a null value, we can pass to analyze all the variables that lead the stalker to threaten the victim's death. Figure 11 shows how the last seven variables have a substantial impact on this pattern. All these variables belong to Mullen's classification framework, and they explain the behavioral traits observable in a stalker. The most common behavior is sending numerous SMS, following the victim everywhere, and all the obsessive actions that a perpetrator uses to have against the victim. It could sound incredible that these variables suggest that stalkers will not threat of death the victim. It would be possible that acting in this way the obsessional behavior will not lead the harasser to menace the victim. The CBT can explain again why these traits could avoid the risk of the threat of death. Every time a stalker exercises power against the victim, the relationship's control turns back, and the anger can be kept under control. Taking a quick look at the other side of the table (Figure 11), we can observe that the first eight variables that strongly impact the positive path are related to interference between the victim and the stalker by something. The presence of external factors as a restraining order released by a court, a victim's satisfaction about a restrictive order obtained against the stalker could lead to a threat of violence against a close person or people (v94) and the end of harassment. In this case, the end of harassment because of an external interference would not be the best scenario. The strain theory and the CBT agree with the idea of control and power in the hand of the harasser play a fundamental role at the beginning of a possible violent scenario. We can hypothesize that when the stalker loses control of the victim, the anger grows because obsessions can not find a relief valve. The

next step could be the threat or harm against people close to the victim. In the end, it can finish with the threat of death against the victim and a possible homicide.

In addition to the innovative idea to predict the next phase of the stalker behavior, this research suggests that we can not understand stalking starting from only one perspective, but we must need to put together several theories for having a complete panoramic of it. Also, the presence of the traditional traits of behavior (Mullen et al., 2009; Meloy et al., 2006) is not necessarily a negative path. The presence of these traits may mean that the stalker is not dangerous yet. On the other hand, when the stalk stops because of external interferences, the decision-makers might pay attention to the next stalker action because it can hide a powerful and violent reaction because of the total interruption of the relationship with the victim. As criminological theories suggest, control and power are very important for a perpetrator who wants to keep active the victim's relationship. If the essential precautions are not adopted in time, interrupting this connection can be very dangerous for the victim.

### 4.4 Limitations of this study

Because the dataset is secondary data, and I did not create and collected data that I used for training the ANN. I did not make the variables. Although the dataset is vast with an important number of cases, most of the variables are missed. This characteristic was the more significant limitation of this dataset. For this experiment, the primary purpose is to predict if there would be a possible threat of death for the victim or close people.

The database represents the major limitation even though I choose it because it was the best available option for performing this research. It was not complete enough in terms of variables. The ANN reached a good percentage of accurate predictions, but not

enough to guarantee many cases outcomes correctly predicted. Although 73% of prediction accuracy is decent and 80% is a good level of accuracy, in order to use this system for improving the decision making of people who manage stalking cases, the risk to have a not enough accurate prediction power could lead the decision-maker to make a mistake of evaluation. The purpose of ML is to try to reach 100% accuracy.

The best way to improve this system would be to re-design a dataset that includes variables and measures easier for the learning process. Also, the number of final outputs used for training the ANN was limited. The *NA* data were excessive, and this problem limited the analysis of the ANN's data weight. A significant data manipulation (changing NA to 0) was required to fit the dataset into the ANN's learning parameters. The database codebook explains that some variables are hidden because sensible data must present a special request for accessing all data. Even though results exceeded my expectations, especially under the weight analysis, it is not yet possible to use the *iv4* and *iv5* for analysis of real cases. Further improvements are necessary to validate the tools.

In addition to the technical part of this study's limitation, two other important limits must be discussed. The obsolete dataset represents the first. The dataset has been assembled at the beginning of 1990. It does mean that the important aspect of cyber-stalking has not been included in the questionnaire because it was not already studied. This limitation is essential because, in these years, stalkers started to use the technology widely for harassing victims. Social networks provide an important way to remain in contact with a person against the willingness to be contacted. Fake social profiles, different phone numbers associated with instan messaging platforms, chats, and emails are determinants for stalkers that want to persecute and damage the victim. This aspect

must be included in further research to not be anchored to old classifications and theories that do not fulfill the stalking phenomenon complexity.

The second essential limit of this research is not available data in terms of victim reports. It is known that the stalking phenomenon hides a big problem behind the data collections. Victims use to do not report the status of harassed by somebody to law enforcement. The motivation behind these problems has been the object of studies for years. Some factors lead a person to do not denounce the harasser. The idea that the situation could degenerate in negative. Distrust of law enforcement and the idea that nothing will be done for fixing the situation. Shame to be unable to react to the newly established situation with the stalker. These three factors generate a grey zone composed of all those active cases that scientists could only estimate without having an accurate report.

### 4.5 Question about bias during the learning process

One of the questions that could be asked about an intelligent system that is able to learn and provide prediction is if there is a risk of bias. The answer to this question is yes. The risk of bias is possible because the system learns from data. What is interesting is to find where the bias could be generated. The ANN is “perfect”. It does mean that after the learning process and the development of the architecture, the ANN will provide predictions based on the dataset, variable iterations, and algorithm selected for the learning process. The scientist provides this information but the ANN does not judge data. Also, the ANN does not work in terms of traditional math, where $a+b=c$, and it is what scientists expect. The ANN works on the iteration between all the selected variables and how much every variable impacts the learning process.

On the other hand, the bias can be generated while building a questionnaire for collecting data and during the build of the dataset.

When we talk about supervised learning, as it has been explained in the previous section, scientists decide what type of data is interesting to give to the ANN, and the ANN will make a process of analysis and learning from those data. If a scientist decides to cut off a variable because it has been judged useless without scientific literature support, it could mistakenly lead the research.

The problem of bias is central in criminological field. Machines can help to do not discriminate data during and after the analysis, but not before it. When a scientist has to decide the variables that will be included or excluded in the training dataset. The risk is to commit two biases into the process of variable choice. Omitted-variable bias happens when results are skewed if the model does not control all relevant factors (Jung et al., 2018). On the other hand, included-variable bias can skew results if there is no control of irrelevant factor. In statistics, it is possible to run a simple test called *risk-adjusted regression* to evaluate though three steps the decision making accuracy in variable analysis and selection of factors.

In this research, I did not run this type of test because I based my variable selection on the theoretical framework. The selection of relevant variables has been made by applying criminological and psychological theories that provide the best frame for stalking phenomenon. Technically, there was no exclusion of variables but a selection of the best data that compose stalking and, after that, the ANN analyzed the weight of variables in process learning without my direct intervention or manipulation. The algorithm did not suffer any human modification.

## Chapter 5: Conclusion

This experiment delivered two valuable contributions to the field. The technical contribution is about the possibility of implementing artificial neural networks in studies of criminological phenomena. Stalking is relatively easy to categorize even though the several disciplines and theories are needed for having a complete panoramic about stalking and its dynamics. Stalking shows a series of specific behaviors, and the relevant social contexts are described by frameworks in the literature. Most of the framing studies agreed with the same structure of stalking even though the particular name of categories slightly changes. This research shows that it is possible to merge two fields of research as disparate as computer science and criminology for developing predictive models of criminal phenomena and understanding hidden relationships behind variables. Even using a dataset that was not built specifically for algorithmic analysis, *iv5* can predict the risk of life threats in stalker scenarios with an accuracy of 80%. Additionally, *iv4* provided interesting information about what it would be crucial to learn from 118 variables. The analysis delivered confirmed the validity of the categorization of stalkers developed by Mullen and through RECON, showing that while racial background, alcohol abuse, and economic situation are often key for understanding crimes, they are not necessarily central to stalking cases.

This thesis is a starting point finding new applications for Artificial Intelligence in Criminology, and its limitations point towards a need to develop a more accurate system, trained by a dataset built with all the necessary properties and variables required for this specific task.

The second contribution is about how to evaluate the stalking, its context, and its evolution starting from different points of its development. In literature, there are numerous contributions for categorizing stalking, behavioral traits that highlight this phenomenon, social contexts and criminological perspectives help to understand when what and why physical violence starts against a victim, psychological pressure, and related pathologies (i.e., PSTD observed in stalking victims) and sociological dynamics that lead to becoming a stalker. This research puts all these variables in an informatics model that works differently than traditional statistics models for extrapolating new information. The model from a specific dataset, supported by literature, tells us that even though several aspects and actions compose stalking, not all of these components lead the development of the harassment to a threatening scenario. Some variables that could include actions that apparently could be judged as worse than others have a negative effect on the future treats of death. As McFarlane and al. (1999) observed, the threat of death is one of the most dangerous variables in stalking because it could generate a strong PSTD to a victim and, also, being the last step before a homicide. Moreover, variables that could be apparently harmless have a strong impact on the future threat of death.

The bottom line remains to elaborate a system that would be able to help law enforcement, scientists, and organizations that help victims of stalking to improve their decision-making about the dangerousness of stalking cases. Stalking cases follow the same structure, but they do not show the same variables every time. An informatics system, as the ANN *iv5,* should support investigators by helping them understand and categorize reported cases in order to prevent situations that degenerate in homicide.

In this way, this thesis has completed its task, and it opens new scenarios around the possibility of predicting serious cases.

### 5.1 Future ideas

The next step must be to build a new dataset from a new dedicated questionnaire built for collecting data with value more usable for the ANN learning process. Variables must strictly respect the theoretical frameworks found in the literature and ANN properties. Through the internet release of a new survey, it will be quicker to provide questionnaires and collect data.

After having built a new and more robust database, one interesting option to pursue might be a change of the algorithm used for training the ANN. There are several algorithms that could fit better for this type of analysis. For example, the *random forest* algorithm and the *decision tree* algorithm should be studied deeply and implemented in future related studies. The *decision tree* algorithm would allow us to build different categorization of cases that could present different traits of behavior or social situation. The dynamicity of this algorithm would allow us to generate a new classification closer to real cases rather than embedded data under fixed labels and let empty spots where data are absent.

## References


Agnew, R. (1992). Foundation for a general strain theory of crime and delinquency.

Agnew, R. (2001). Building on the foundation of GST: specifying the types of strain

Agnew, R. (2006). Pressured into crime: An overview of general strain theory. Los Angeles: Roxbury.

Bandura, A. (1971). "Social Learning Theory." General Learning Corporation. 24 October 2013.

Baum, K., Catalano, S., Rand, M., & Rose, C. (2009). Stalking victimization in the United States (NCJ 224527). Bureau of Justice

Bear, M. F., Connors, B. W., & Paradiso, M. A. (1999). *Neuroscience: Exploring the brain*. Philadelphia, PA: Lippincott Williams & Wilkins.

Bernburg, J.G. (2002). Anomie, Social Change, and Crime: A Theoretical Examination of Institutional-Anomie Theory. British Journal of Criminology. 42:729–42.

Bocij, P., & McFarlane, L. (2003). Seven fallacies about cyber stalking. Prison Service Journal, 149(1), 37-42.

Boehnlein, T., Kretschmar, J., Regoeczi, W. et al. (2020). Responding to Stalking Victims: Perceptions, Barriers, and Directions for Future Research. J Fam Viol 35, 755–768 https://doi.org/10.1007/s10896-020-00147-3

Brady, P. Q., Nobles, M. R., & Bouffard, L. A. (2017). Are college students really at a higher risk for stalking? Exploring the generalizability of student samples in victimization research. Journal of Criminal Justice, 52, 12–21.

Budd, T., & Mattinson, J. (2000). The extent and nature of stalking: Findings from the 1998 British Crime Survey. The Home Office: London

Campbell, J., & Moore, R. (2011). Self-perceptions of stalking victimization and impacts on victim reporting. Police Practice and Research, 12, 506–517, *Criminology, 30*, 47-87.

Cupach, W. R., & Spitzberg, B. H. (2004). *The dark side of relational pursuit: From attraction to obsession and stalking*. Mahwah, New Jersey: Lawrence Erlbaum

Curci, P., Galeazzi, G.M., Secchi, C. (2003). *La sindrome delle molestie assillanti*. Boringhieri, Torino.

Curran, Daniel J. & Renzetti, Claire M. (2001). Theories of crime. Boston : Allyn and Bacon

De Fazio, L., & Galeazzi, G. M. (2004). Women victims of stalking and helping professions: Recognition and intervention in the Italian context. In G. Mesko, M. Pagon, & B. Dobovsek (Eds.), Policing in Central and Eastern Europe: Dilemmas of contemporary criminal justice (pp. 606–619). Slovenia: Faculty of Criminal Justice, University of Maribor.

Dietz, P.E., Matthews, D.B., Martell D.A., Stewart T.M., Hrouda, D.R., Warren J., (1991). Threatening and otherwise inappropriate letters to members of the United States Congress, *J Forensic Sci.* Sep; 36(5):1445-68.

DuMonthier, A., Dusenbery, M, & Gonzalez, S. (2017). Intersections of stalking and economic security. Report # B365, Institute for Women's Policy Research.

Durkheim, E. (1964). The division of labor in society. New York, Free Press of Glencoe.

Finch, E. (2002). Stalking the perfect stalking law: An evaluation of the efficacy of The Protection from Harassment Act 1997. Criminal Law Review, 14, 703-718.

Fisher, B. S., Cullen, F. T., Turner, M. G. (2002). Being pursued: Stalking victimization in a national study of college women. Criminology & Public Policy, 1, 257-308.

Fox, K. A., Nobles, M., Fisher, S.B. (2016) A Multi-Theoretical Framework to Assess Gendered Stalking Victimization: The Utility of Self-Control, Social Learning, and Control Balance Theories, *Justice Quarterly*, 33:2, 319-347, doi:10.1080/07418825.2014.902985

Froggio, G., & Agnew, R. (2007). The relationship between crime and ―objective‖ versus ―subjective‖ strains. *Journal of Criminal Justice, 35,* 81-87.

Geistman, J., Smith, B., Lambert, E., & Cluse-Tolar, T. (2013). What to do about stalking: a preliminary study of how stalking victims responded to stalking and their perceptions of the effectiveness of these actions. Criminal Justice Studies, 26, 43–66.

Gerber, John J. Macionis, Linda M. (2010). Sociology (7th Canadian ed.). Toronto: Pearson Canada. p. 97. ISBN 978-0-13-700161-3

Géron, A. (2019). Hands-On Machine Learning with Scikit-Learn and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems.

Gottfredson, M. R., & Hirschi, T. (1990). A general theory of crime. Stanford, CA: Stanford University Press.

Grossi E, Buscema M: Introduction to artificial neural networks. Eur J Gastroenterol Hepatol 2007, 19:1046-1054.

Haron, H., & Yusof, F.B. (2010). Cyber stalking: The social impact of social networking technology. 2010 International Conference on Education and Management Technology, 237-241.

Higgins, E., Lauterbach, C., Tewksbury, R. (2005). Control Balance Theory and Violence: an examination of contingencies. Sociological Focus, Vol. 38, No. 4 (November 2005), pp. 241-260

Jordan, C., Logan, T., Walker, R., & Nigoff, A. (2003). Stalking victimization: the efficacy of criminal justice response. Journal of Interpersonal Violence, 18, 148–165.

Jung, J., Corbett-Davies, S., Shroff, R., & Goel, S. (2018). Omitted and Included Variable Bias in Tests for Disparate Impact. arXiv: Applications.

Langton, L., & Piquero, N. L. (2007). Can general strain theory explain white-collar crime? A preliminary investigation of the relationship between strain and select white-collar offenses. *Journal of Criminal Justice, 35*, 1-15.

Leong, G.B. (1994). Syndrome (Erotomania) in the criminal justice system. Another look at this recurring problem, journal of Forensic Sciences, n.39

Linehan, M. M. (1993). Cognitive behavioral treatment of borderline personality disorder. New York, NY: Guilford Press.

Logan, T., & Cole, J. (2007). The impact of partner stalking on mental health and protective order outcomes over time. Violence and Victims, 22, 546–562.

Logan, T., & Walker, R. (2010). Toward a deeper understanding of the harms caused by partner stalking. Violence and Victims, 25, 440– 455.

Logan, T., Walker, R., Jordan, C., Leukefeld, C. (2006). Women and victimization: Contributing factors, interventions, and implications. Washington DC: American Psychological Association.

Lowney, K. S., & Best, J. (1995). Stalking strangers and lovers: Changing media typifications of a new crime problem. In J. Best (Ed.), Images of issues: Typifying contemporary social problems (, 2nd ed., pp. 33–57). New York, NY: Aldine de Gruyter.

Macionis, J. J., & Gerber, L. M. (2018). *Sociology*.

McCutcheon L.E., Aruguete M., Scott Jr. V.B., Parker J.S., Calicchia J. (2006). The development and validation of an indirect measure of celebrity stalking. North American Journal of Psychology, 8(3), 503-516.

McEwan, T.E., Mullen, P.E., & MacKenzie, R. (2008). A study of the predictors of persistence in stalking situations. Law and Human Behavior, [Online only] doi 10.1007/s10979-008-9141-0

McFarlane, J. M., Campbell, J. C., Wilt, S., Sachs, C. J., Ulrich, Y., & Xu, X. (1999). Stalking and intimate partner femicide. Homicide Studies, 3(4), 300-316. https://doi.org/10.1177/1088767999003004003

McFarlane, J., Campbell, J. C., & Watson, K. (2002). Intimate partner stalking and femicide: urgent implications for women's safety. *Behavioral sciences & the law*, *20*(1-2), 51–68. https://doi.org/10.1002/bsl.477

Meloy, J. R. (1992). Violent attachments. Northvale, NJ: Jason Aronson, Inc.

Meloy, J. R., & Boyd, C. (2003). Female Stalkers and Their Victims. *Journal of the American Academy of Psychiatry and the Law, 31*(2), 211–219.

Meloy, J. R., Gothard S. (1995). Demographic and clinical comparison of obsessional followers and offenders with mental disorders, *American Journal of Psychiatry*, n.152

Meloy, J.R. (1996). Stalking (obsessional following): a review of some preliminary studies. *Aggress. Viol. Behavior*, 1:147–62.

Meloy, J.R., Mohandie, K., McGowan, M.G., Williams, J. (2006). The RECON Typology of Stalking*: Reliability and Validity Based Upon a Large Sample of North American Stalkers*, J Forensic Sci, January 2006, Vol. 51, No. 1 doi:10.1111/j.1556-4029.2005.00030.x Retreived from: www.blackwell-synergy.com

Meloy, J.R., Rivers, L., Siegel, L., Gothard, S., Naimark, D., Nicolini, R. (2000). A replication study of obsessional followers and offenders with mental disorders. *J Forensic Sci*; 45 (1):147–152.

Meloy, J.R.., Rivers L., Siegel L., Gothard S., Naymark D., Nicolini R.M. (2000). A replication study of obsessional followers and offenders with mental disorders, *Journal of forensic science*, n.45

Merton, R.K.(1938). Social Structure and Anomie. American Sociological Review. 1938;3:672–82

Miller, L. (2012). Stalking: Pattern, motives, and intervention strategies, Aggression and violent behavior, 17, 495-506

Mohandie, K., Meloy, J. R., McGowan, M. G., & Williams, J. (2006). The RECON typology of stalking: reliability and validity based upon a large sample of north American stalkers. Journal of Forensic Sciences, 51, 147–155.

Mullen, P. E. (2000). Dangerousness, risk and the prediction of probability. In M. G. Gelder, J. J. Lopex-Ibor, N. C. Andreasen, (Eds.). New Oxford textbook of psychiatry (pp. 272-281). Oxford, UK: Oxford University Press.

Mullen, P. E. (2003). Multiple classifications of stalkers and stalking behavior available to clinicians. Psychiatric Annals, 33, 651-656.

Mullen, P. E., & Pathe, M. (2002). Stalking. In M. Tonry's (Ed.) Crime and justice: A review of research (pp. 273 – 318). Chicago, IL: University of Chicago Press.

Mullen, P. E., Mackenzie, R., Ogloff, J. R., Pathe, M., McEwan, T., & Purcell, R. (2006). Assessing and managing the risks in the stalking situation. Journal of the American Academy of Psychiatry and the Law, 34, 439-450.

Mullen, P.E., Pathé M., Purcell,R. (2009). Stalkers and their victims, Cambridge University Press

Mullen, P.E., Pathe, M., & Purcell, R. (2001). Stalking: New constructions of human behaviour. Australian and New Zealand Journal of Psychiatry. 35, 9-16. 135

Mullen, P.E., Pathé, M., Purcell, R. (2000). Stalkers and their Victims. *Cambridge University Press*, Cambridge

Mullen, Paul E.; Pathé, Michele; Purcell, Rosemary; Stuart, Geoffrey W. (1999). A Study of Stalkers. *American Journal of Psychiatry*. Philadelphia, Pennsylvania: American Psychiatric Association (156): 1244–1249.

Nobles, M., Fox, K. A. (2013). Assessing stalking behaviors in a control balance theory framework. *Criminal Justice and Behavior*, 40, 737-762.

Petherick, W. (2007). Cyber stalking. Retrieved May 9, 2007, from CrimeLibrary.com Web site: http://www.crimelibrary.com/criminology/cyberstalking/

Piquero, A. R., Hickman, M. (2003). Extending Tittle's control balance theory to account for victimization. *Criminal Justice and Behavior*, 30, 282-301.

Pittaro, M. (2007). Cyber stalking: An Analysis of Online Harassment and Intimidation. *International Journal of Cyber Criminology, 1*, 180-197.

Racine C., Billick S. (2014). *Classification systems for stalking behavior*. J Forensic Sci. 2014 Jan; 59(1):250-4. doi: 10.1111/1556-4029.12262

Reyns, B. W., & Englebrecht, C. M. (2010). The stalking victim's decision to contact the police: a test of Gottfredson and Gottfredson's theory of criminal justice decision making. Journal of Criminal Justice, 38, 998–1005.

Rubin, Donald B.; Little, Roderick J. A. (2002). Statistical analysis with missing data (2nd ed.). New York: Wiley

Scott, A. J., Lloyd, R., Gavin, J. (2010). The influence of prior relationship on perceptions of stalking in the United Kingdom and Australia. Criminal Justice and Behavior, 37, 1185-1194. doi:10.1177/0093854810378812

Scott, A. J., Sheridan, L. (2011). "Reasonable" perceptions of stalking: The influence of conduct severity and the perpetrator-target relationship. Psychology, Crime & Law, 17, 331-343. doi:10.1080/10683160903203961

Shavelson, R. J., and Towne, L. (Eds.). (2002). *Scientific research in education*. Washington, DC: National Research Council, National Academy Press.

Sheridan LP, Blaauw E, Davies GM. Stalking: Knowns and Unknowns. *Trauma, Violence, & Abuse*. 2003;4(2):148-162. doi:10.1177/1524838002250766

Sheridan, L. P., & Boon, J. (2002). Stalker typologies: Implications for law enforcement. In J. Boon & L. Sheridan (Eds.). Stalking and psychosexual obsession: Psychological perspectives for prevention, policing and treatment (pp. 63-82). West Sussex, UK: Wiley.

Smith, P. (2008). Durkheim and Criminology: Reconstructing the Legacy. Australian & New Zealand Journal Of Criminology (Australian Academic Press), 41(3), 333-344. doi:10.1375/acri.41.3.333

Smith, S. G., Chen, J., Basile, K. C., Gilbert, L. K.,Merrick, M. T., Patel, N., Walling, M., & Jain, A. (2017). The National Intimate Partner and sexual violence survey (NISVS): 2010–2012 state report. Atlanta: National Center for Injury Prevention and Control, Centers for Disease Control and Prevention.

Spitzberg, B., Cupach, W.R. (2007). The state of the art of stalking: Taking stock of the emerging literature. *Aggression and Violent Behavior*. London, England: Elsevier. 12 (1): 6486. doi: 10.1016/j.avb.2006.05.001.

Tittle, C.R. (1995). Control balance: Toward a general theory of deviance. *Boulder*, CO: Westview.

Tjaden, P. (1997) The crime of stalking, how is big the problem? U.S. Dept. of Justice, Office of Justice Programs, National Institute of Justice, [1997]

Tjaden, P., & Thoennes, N. (1998). *Stalking in America: Findings from the National Violence Against Women Survey.* Washington, D.C.: National Institute of Justice and Centers for Disease Control and Prevention. [Electronic version]. http://www.ncjrs.gov/pdffiles/169592.pdf

Tjaden, P., & Thoennes, N. (2000). The role of stalking in domestic violence crime reports generated by the Colorado Springs Police Department. Violence and Victims, 15, 427–441.

U.S. Department of Justice (1998). Stalking in America: findings from the National Violence Against Women Survey. National Institute of Justice Research in

Weller, M., Hope, L., Sheridan, L. (2013). Police and public perceptions of stalking: The role of prior victim-offender relationship. Journal of Interpersonal Violence, 28, 320-339. doi:10.1177/0886260512454718

Zona, M.A., Palarea, R. E., & Lane, J. C. (1998). Psychiatric diagnosis and the offender–victim typology of stalking. In J. R. Meloy (Ed.), *The psychology of stalking: Clinical and forensic perspectives* (pp. 69–84). San Diego, CA: Academic Press.