Conceptual Design on Computer Sentencing Simulation Based on SVM

Da Ke

University of Tampere
Faculty of Natural Sciences
Degree Programme in Computer Sciences
Software Development
M. Sc. thesis
Supervisor: Zheying Zhang
May 2018

University of Tampere
Faculty of Natural Sciences
Degree Programme in Computer Sciences
Software Development
Da Ke: A Conceptual Design on Computer Sentencing Simulation Based on SVM
M.Sc. thesis, 54 pages, 3 reference pages, 6 index and appendix pages
May 2018

Abstract

The criminal law in China is a relatively uncertain statutory punishment law, and the judge exercise the equitable discretion within the extent for discretionary action of sentencing. However, influenced by many objective and subjective factors, the punishment imparity exists inevitably. To farthest implement the justice goal that criminal law pursues and get the largest benefit from criminal penalty, the Support Vector Machine (SVM), one of the machine learning method that newly emerged in the artificial intelligence theory, is adopted for the application of measurement method research of penalty in this thesis, and the SVM measurement model of penalty (SVM sentencing model) is presented, which attempted to decrease the imparity in the measurement of penalty through the improvement of sentencing method. Based on the SVM sentencing model as the core measurement method of penalty, the machine learning based sentencing expert system's general frame is described. Finally, the theft crime is taken as an example, the realization procedures and details of expert system are illustrated.
Key words: sentencing, sentencing circumstances sentencing method, machine learning, support vector machines.

Ackowledgement

I would first like to thank my thesis supervisor Dr. Zheying Zhang of the Faculty of Natural Sciences at University of Tampere. The door to Prof. Zhang's office was always open whenever I ran into a trouble spot or had a question about my research or writing. She consistently allowed this paper to be my own work, but steered me in the right direction whenever she thought I needed it. Her patience and professional guidance inspired my perseverance on the research and this thesis.

I would also like to thank the teachers and experts who were involved in my learning process in University of Tampere and Kirsi-Marja Tuominen. Without their passionate participation and input in helping me during the long education journey in Software Development and the study life in Finland, the thesis could not have been successfully presented.

I would like to acknowledge Professor Martti Juhola of the Faculty of Natural Sciences at University of Tampere as the reviewer of this thesis as well, and I am gratefully indebted to him for his very valuable comments on this thesis.

Finally, I must express my very profound gratitude to my parents and to all my friends for providing me with unfailing support and continuous encouragement throughout my years of study and through the process of researching and writing this thesis. This accomplishment would not have been possible without them. Thank you.

Da Ke
29/5/2018

Contents

1. Introduction

Justice is the basic value pursuit and basic code of conduct of human society with eternal significance. In some sense, the pursuit of justice is the process of human society development from backward to advanced and from unreasonable to reasonable. The realization of justice is one of human ambitions.

Initially, the meaning of justice in Ancient Greek philosophy is to conduct lawfully. Plato [Lamb 1925] thinks that justice should be a moral code of human virtue, represents as each taking its place and each taking its share.

The most primitive and simple form of justice is the natural pursuit of reciprocity. In the field of criminal law, this reciprocity manifests as the balance between crime and punishment, which is to suit punishment of crime and to punish in keeping with crime.

Nevertheless, the justice in legislation is general and popular, which applies to everyone. Individual justice can only be revealed in judicial. It is the kind of justice according to some individuals and individual cases under the guidance of general justice. Individual justice is important because general justice owns the limitation of legal norms, and it can hardly be applied to all circumstances in a natural and perfect way. The limitation can only be remedied through judicial actions. Even a highly reasonable law still has the primness from stability, generality and abstraction of legal norms. The judiciary is obliged to maintain the consistency with legislation and it will generate.

In order to better achieve fairness and justice and to pursue accurate sentencing, this research concerns the use of the power of machine learning and the SVM method in sentencing estimation.

The main part of the thesis consists of three chapters. In chapter 2, the concept and characteristics of sentencing is briefly summarized firstly, which pointed out that as a criminal sentencing system, the justice that sentencing pursues can be realized only through correct measurement of punishments. Then the requirements that implement accurate sentencing is described, and the current existed penalty imparity status is analyzed. The analyses indicate that it is emergent and significant to update the sentencing method and develop sentencing application technique. It can guide the judge to realize the accurate sentencing, consequently decrease the imparity of sentencing and implement the balance of sentencing. In addition, this chapter summarizes the current research status of sentencing methods and points out that it is feasible to apply newly emerged machine learning theory to the development of sentencing expert system.

In Chapter 3, the machine learning and support vector machine theory is briefly introduced firstly, and then the feasibility of applying SVM to the development of sentencing method is analysed.

In Chapter 4, the model building procedure of the SVM sentencing model is presented. During the model building process, firstly the expert evaluated samples that are relative correct and can represent the system characteristics are collected. In order to obtain the

relative corrected sentencing samples, the advantages of sentencing of Common Law system are referred to optimize the current sentencing scheme that China adopts, and the correlated sentencing scheme theory are analysed. After the samples are obtained, the sentencing circumstances are extracted and quantified to get the quantity representation of the act, which are then fed as the input to support vector machines for training to get the sentencing model. When a new criminal case comes, the act of which are extracted and quantified firstly, then they are sent to the SVM sentencing model to obtain the referred sentence.

Chapter 5 takes the above built SVM sentencing model as the core inferential machine, the sentencing expert system's general framework is described.

In Chapter 6, the theft crime is taken as an example to illustrate the realization procedures and details of expert system with a focus on the concrete implementation details of SVM sentencing model.

In Chapter 7, the existing problems and further research directions of the research are discussed.

In summary, the machine learning theory is adopted for the development of a sentencing assistant, and the SVM based sentencing expert system realized the crossover between the subjects of criminal law and computer science. However, essentially, the thesis is researched and written from the viewpoint of Chinese criminal law, which put the emphasis on the building of SVM sentencing model and the application of machine learning on sentencing.

2. Sentencing and Sentencing Methods

The thesis involves the computer science field, legal field and statistics field. So, it is a cross-disciplinary research and there is a great need to illustrate and explain some background concepts in legal fields about sentencing well. The following sections will give a brief but necessary explanation of sentencing and sentencing methods.

2.1 Concept and characteristics of sentencing

The concept of sentencing is expressed in various Chinese and foreign legal works, but the general contents are similar. Japanese scholars believe that the so-called sentencing refers to the type and amount of penalties that should be announced for specific decisions. "The process of selecting a particular sentence is called the measurement of the punishment. Specifically, it means the process of deciding the announced penalty." [Kahan & Nussbaum 1996] For announcing specific penalties, the court first selects the types of penalties that should be applied, decides whether or not to apply any legally-reduced cause of exemption, and whether it can be mitigated accordingly. Then, the penalties that should be announced are specifically determined within the scope of the penalties and the sanctions will be made. In addition, discretionary exemption of the penalty, whether or not to allow probation, is also determined according to discretion. The amount of punishment relies on the discretion of the judge. The specific circumstances of the crime are varied and it is difficult to regulate by the general provisions of the law. Therefore, the proper and appropriate penalty must not be imposed on the judge's individual judgment. So, the specific and appropriate penalty has to rely on the judge's individual judgment. However, even if it is discretionary, it does not allow the judge to act arbitrarily. The judge must work hard to determine the reasonable penalty [Kahan & Nussbaum 1996].

German scholars believe that sentencing is a determination of the legal consequences of crime [Jescheck 2004]. It includes the choice of system (such as imprisonment penalty, fines, etc.), the determination of sentencing standards (such as the duration of the freedom sentence), and if necessary, a verdict on the delivery of punishments or the probation of security measures. In specific circumstances, most of the laws give the court a wide range of space for sentencing. Only in the case of murder and genocide crimes, mandatory lifelong imprisonment penalties are stipulated. In general, the specific criminal law regulations only stipulate a penalty range, i.e. where the penalties that shall be imposed in the penalty range. The law does not make specific provisions, but only sets forth some general principles and rules of use that apply to specific circumstances. Therefore, people have concluded that the amount of penalty is the issue of the judge's discretion, and at the same time it reflects the “personal ability” of the presiding judge. Today, people agree that under specific circumstances, the choice and determination of sanctions is a legally binding decision [Jescheck 2004].

The concept of sentencing in Chinese academia is not consistent in the text. The general view is that there is a broad and strict sense of sentencing. The strict sentencing refers to the people's court's trial of specific criminals' discretion and the determination of specific penalties. Sentencing in a broad sense refers to the entire process in which the people's court decides to give criminals specific punishments or exemptions from punishment. In addition to the narrow sentencing, the broad sentencing also includes discretionary punishment and probation discretion. Specifically, the sentencing is a special activity that the People's Court decides whether or not to impose criminal punishment and what kind of punishment is imposed on criminals according to the offender's facts of the crime, the nature of the crime, the circumstances of the crime, the degree of harm to the society, and other circumstances. The sentencing includes the following steps: the disciplinary division of punishment, which is to determine whether the offender is sentenced to criminal punishment or not after the conviction; the choice of punishment to determine the type of penalty that should be applied based on the facts and circumstances of the crime; the determination of the degree of punishment, which is to determinate the punishment according to the penalty range in the corresponding law; and the measurement of penalties, which includes all kinds of matters of heaviness, lightness, mitigation and exemption are applied in accordance with the law, and a final declaration of punishment when penalties are imposed.

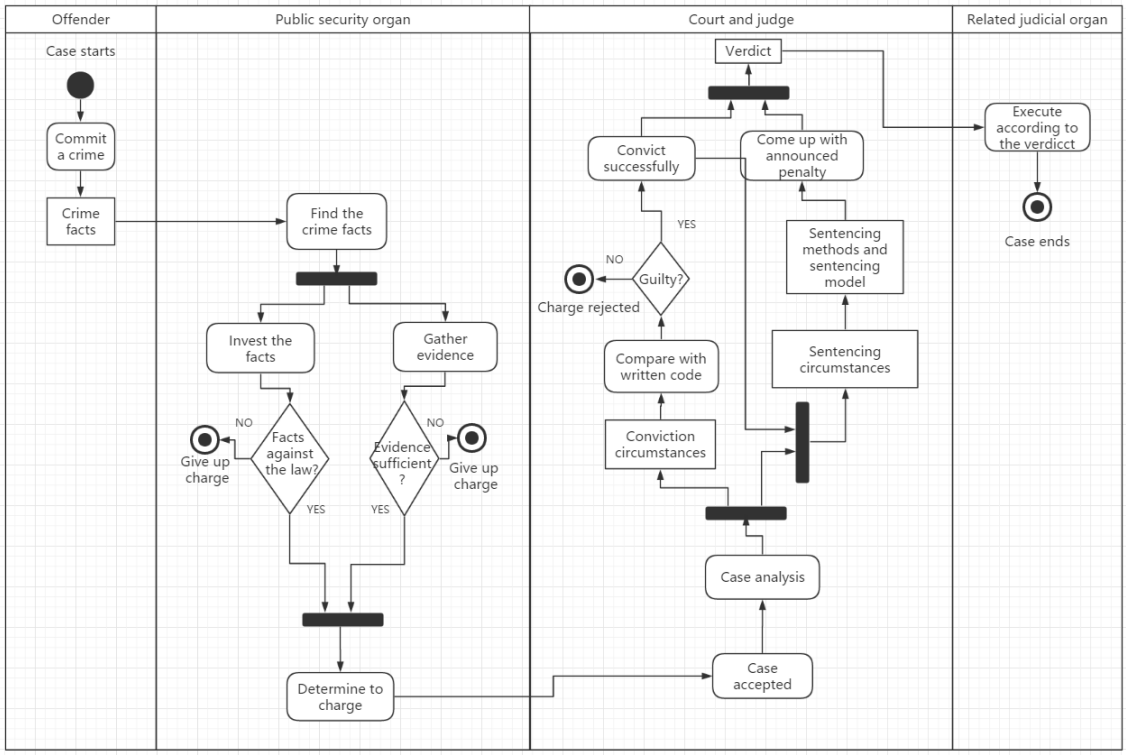


Figure 2.1 The process from a case starts to its ends including sentencing phase in China

As Figure 2.1 shows, the whole process from a case starts to its ends is a very complicated process. Four subjects: offender, public security organ, court and judge and related judicial organ are related. A case starts when the offender commits a crime. Then the crime facts are somehow found by public security organs either by themselves or

reported by others. The public security organs on one side invest the facts and on the other hand gather evidences. Either the facts are not against the law or the evidences are not sufficient, the charge will be given up. If both conditions fit in the case, the public security organs determine to charge and the case is moved to the court. The court accepts the case and analyzes it. Both conviction circumstances and sentencing circumstances are extracted. The conviction circumstances are compared to the written code and if they to determine whether the offender is guilty of some kind of crime. If the offender is thought to be not guilty, then the charge is rejected. If the offender is somehow thought to be guilty, then the court and the judges analyze the sentencing circumstances using certain kinds of sentencing methods and sentencing model to come up with a suitable announced penalty. The convict of the offender of some certain crime and the announced penalty consist of a verdict. After the verdict takes effect, the related judicial organs are going to execute exactly according to the verdict.

From the concept of sentencing, it is easy to find out that the sentencing in the Chinese legal system has the following characteristics:

The clarity of the subject of sentencing is the first characteristic. The power of sentencing is an important part of the judicial power of the country. As one of the important links in trial activities, sentencing must be conducted by the People's Court. As the judicial organ, the people's court is the only judicial authority that has the power to act on behalf of the state to exercise the power of sentencing. No other agency, group or individual has the right to measure.

The specificity of the objects of sentencing is the second characteristic. As the direct target of sentencing, the actual bearer of the specific penalty is the perpetrator of the criminal act, i.e. the offender. In other words, the objects in each sentencing process are specific. Only those who have committed crimes are the objects of sentencing.

The diversity of sentencing forms is the third characteristic. From the carrier form, sentencing can be either expressed as a form of criminal judgment or a form of criminal adjudication; From the substantive content, sentencing can be not only expressed as a life sentence, but it can also be expressed as an imprisonment penalty. It can even be expressed as a property penalty or a qualification penalty.

The certainty of the nature of sentencing is the last characteristic. Sentencing is the decision of people's court to determine the offender and determine the penalty according to the facts of the crime, the nature of the crime, the circumstances and the degree of harm to the society, and with reference to the criminal’s personal circumstances, according to the relevant provisions of the criminal law. Therefore, the nature of sentencing is a criminal justice activity.

2.2 The pursuit and the present situation of sentencing

The present situation and the pursuit of sentencing is the main reason why this research is necessary. The following content will give a detailed, vivid and professional introduction to the pursuit and the present situation of sentencing.

2.2.1 The pursuit of sentencing: accurate sentencing

Sentencing is to ensure that the legal relationship between crime and punishment provided in the criminal law becomes a real crime-related relationship, so that the legislature's penalties for a class of crimes in the legislation become an important part of punishment for criminal acts in specific cases in social reality. Only with correct sentencing, legal punishment can truly become a realistic, enforceable sanction measure. Sentencing is also a prerequisite for execution. Whether or not the sentence is correct is decisive for execution. When the sentence is accurate, execution will not only have the correct direction, but also be relatively smooth to obtain good results. A wrong sentencing not only makes the execution deviate from the correct direction but increases the resistance to execution so as to have adverse consequences. If sentencing is improper, the more stringent the execution of the sentencing penalty is, the more unfair the consequences to the society may be.

Accurate sentencing is an important means to achieve the task of criminal law in China. If the sentencing is not accurate, it will not only fail to fulfill the task of the criminal law, but also hinder the smooth realization of the task of the criminal law. Besides, the correct measurement of the penalty is an important guarantee for the realization of the purpose of punishment. One of the effects of punishment is to achieve individual prevention and general prevention through punishing and educating criminals. Whether this prevention goal can be achieved depends to a large extent on the accuracy of sentencing. For criminals, by accurate sentences, they will receive punishment that they deserve as well as education reform, so that they will no longer commit crimes. At the same time, by penalizing criminals, it gives potential offenders in society vigilance education so that they no longer embark on the criminal road. The realization of the purpose of punishment cannot be achieved merely by applying the penalty but must be based on accurate sentencing. If an innocent person was sentenced, the legitimate interests of citizens would be infringed; If it were a misdemeanor sentence, it would not allow the criminal to plead guilty to sin, but also would increase the resistance, and then they might take the risk and continue to commit crimes; If a felony got punished a minor sentence or if a criminal gets no sentence, it would make the criminals feel lucky and even commit crimes again without fear. At last, correct sentencing is an important guarantee for improving the quality of case handling. The importance of sentencing is not only no less than conviction, but also to some extent more important. The ultimate goal of criminal trials by judges is to impose criminal punishments on criminal elements, and whether or not the penalty is effective depends on whether the penalty is correct

accurate and reasonable. Inadequate and unreasonable sentencing will not only seriously undermine the image of judicial justice, but it will also lead to a waste of national resources.

2.2.2 The requirements of accurate sentencing

To make the best use of the penalty, it is imperative to implement the principle of impartiality in sentencing activities and achieve accurate sentencing. Just as Francis Bacon once said:" An unfair trial results worse than ten crimes. Because crime is ignoring the law - it is like polluting the water, but unfair trials ruin the law - it is like polluting the water source." [Su *et al*., Sentencing and Computers: A Fair and Rational Application of Sentencing, 量刑与电脑:量刑公正合理应用论, 1989] Accurate sentencing require that the sentencing must be unified, balanced, coordinated and fair. First of all, for crimes with the same nature and circumstances, the same range of penalties should be chosen and the appropriate statutory penalties should be imposed without great disparity. Second, if the circumstances are the same for the same type of case, the severity of the sentence should be roughly the same.

Finally, the sentencing of justice requires that no matter who, as long as the crime is committed, it must be sentenced in accordance with the law, sentencing in equal measure, and opposing the privilege in addition to the law.

When cultivating, you can't just care about sowing and not care about harvest. Similarly, the judge can't just ignore the social effects of sentencing. There are two kinds of social effects of sentencing: one is a benign social effect, that is, a positive effect. This is through accurate sentencing, so that criminals get punished and reformed, and become law-abiding citizens that no longer commit crimes. At the same time, it also deters potential criminals in society from committing crimes. The other is a non-benign social effect, that is, a negative effect, which is completely opposite to the above effect.

To make the sentencing produce a benign social effect, then first of all the sentence must be lawful and timely. Late justice is unjust. Secondly, the sentence must be properly and correctly. Accurate sentencing also shows the fairness of sentencing, and the social effects received are generally benign.

2.2.3 Sentencing deviation

Incorrect and unreasonable penalties result in an imbalance of sentencing, that is called sentencing deviation. This refers to the phenomenon that, in the same temporal and spatial conditions where crimes with the same nature and the circumstances are equivalent to each other, there is a great difference in the penalty in the sentence results from the judicial organs when the same law is applied. [Zhang Z. 1999]

Sentencing deviation is a common problem in the world. As long as judges have discretionary power, deviation from sentencing is inevitable.

The reason why sentencing issues has attracted the attention of all countries is because after the issue of conviction has been resolved, the sentencing issue becomes

particularly prominent. Judging from the judicial practice, the rate of changing guilty judgement is extremely low, and most defendants are more concerned with their sentence (prison term). The prison term often carries the individual subjective color of the judge, and there is a certain degree of flexibility within the legal margin. Some scholars have conducted investigations on the crime of rape. For the same case, the minimum sentence for judges is 3 years, and the maximum is 8 years, there is a difference of 5 years [Ke 1989].

The author once assisted Higher People's Court to conduct a sentencing survey, and deeply felt the imbalances among different courts and different judges. For example, when the other circumstances are approximately the same, the penalty for theft is directly proportional to the amount of theft. That is to say, when the penalty is similar, the amount of theft should be roughly the same. In the sentencing procedure of six theft cases, it is possible to extract some of the facts as the sentencing circumstances, i.e. theft amount, theft frequency, confession, whether the offender is a recidivist, whether the offender has an accomplice, the amount that the offender gives up ill-gotten gains actively or passively, other circumstances and the announced penalties. We extracted the mentioned facts and listed them in Table 2.1. We can see that: the six theft cases are ordinary thefts and the crimes are accomplished and the criminals are recidivists without confession or turning themselves in.

Amount	Frequency	Confession	Recidivist	Accomplice	Give up ill-gotten gains actively	Give up ill-gotten gains passively	Other circumsta-nces	Penalties
16808	2		y	y				5
28455	multiple		y					5
45800	6		y					6.5
46002	2		y	y		30400	good attitude	6
52050	multiple		y	main				6
415000	15		y				good attitude	5

Table 2.1 Sentencing circumstances extracted from 6 theft cases

Then we can try to figure out the relationship between the number of years of imprisonment and the amount of theft in these six cases in Figure 2.1:

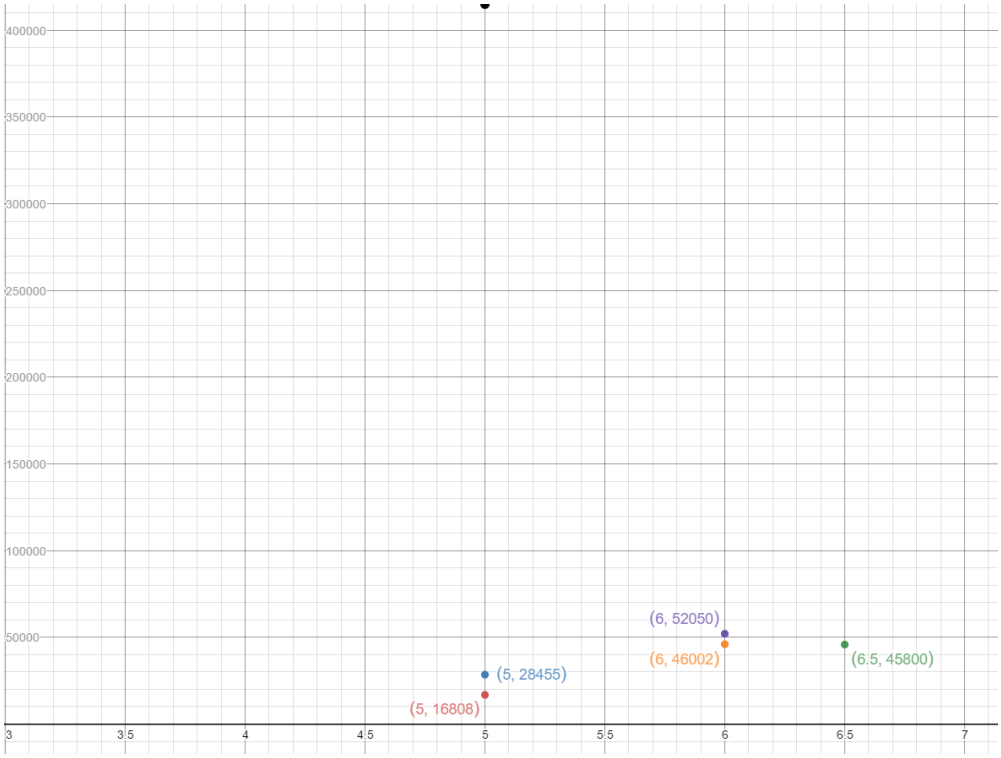


Figure 2.2 Sentencing circumstances extracted from 6 theft cases

In Figure 2.2, we can see that there is a point (5,415000) with different peaks. This is a very obvious deviation. In judicial practice, judges' use of discretionary powers within the scope permitted by law is undoubtedly legal, but not necessarily reasonable. The deviation of sentencing caused by this unreasonable sentencing penalty makes the value goal pursued by the law impossible to achieve. Besides, for the general public, who are usually not familiar with and are not proficient in law, it is very difficult for them to judge the fairness from the results of an isolated case, but they will judge whether the referee is fair or not by comparing the results of the same or similar cases. Can we insist on the equality of all people before the law in the judgment of the case? This is the most sensitive and most concerned issues during the public judgment on the justice of the judiciary, are also the ones that most strongly reflect the injustice of justice [Chiongson *et al*., 2012].

Sentencing is the activity of judges in applying the law. Therefore, the best way to eliminate deviations from sentencing is to start with the law and the judges. The first is to improve the sentencing provisions in criminal legislation and limit the freedom of judges. The second is to improve the quality of judges [Ma, Improper use of penalties and their countermeasures, 刑罚适用失当及其对策, 2002]. However, the law is limited and endless. Legislation cannot exhaust every sentencing scenario and stipulate and the overall improvement of the quality of judges is not a task that can be accomplished overnight. Hence, at present, we can only provide methodological help for judges to accurately measure sentences through the update of sentencing methods and the development of sentencing techniques, thereby reducing deviations from sentencing and achieving a balance of sentencing.

2.3 Development and evolution of sentencing methods

The method of sentencing refers to the sum of the steps, procedures, and means by which judges arbitrate criminal decisions according to law. All procedures and means for properly determining and determining penalties fall within the category of sentencing methods. With the increasing attention paid to the problem of sentencing deviation and the increasing development of science and technology, the method of sentencing is constantly developing and evolving.

2.3.1 Traditional methods

There used to be some traditional methods that is used during the process of sentencing by judges. Two of them are illustrated in the following content. They are comprehensive assessment of sentencing methods and benchmarking sentencing methods.

The sentencing methods of comprehensive assessment are very wildly used in China. It is a traditional sentencing method [Fan 1994]. The judge judges the offender based on his own understanding of the law and past experience in handling cases. Generally, the procedure is as following. The judge first heard the case and mastered the case. Then, on the basis of conviction, within the scope of legal punishment, and with reference to the past experience of judicial practice, the judge roughly estimated the penalty that should be imposed on the current case. After that, the judge considered the cases of mitigation, heaviness, lightness, and exemption from punishment. And finally, a comprehensive assessment of the penalty that the perpetrator should perform is announced. The advantage of this sentencing method is its simplicity and flexibility. It is used and familiarized by the actual staff of the judicial department, and it can also give full play to the subjective initiative of judges. However, due to the fact that China’s criminal law does not stipulate the limits of lightening and other statutory circumstances, there is no specific requirement for the application of discretionary circumstances. Judges have greater discretion, often with the influence of their own political quality, professional quality and psychological quality, it will produce blindness, contingency, and subjective arbitrariness when it comes to sentencing. Together with other subjective and objective factors, it tends to appear to be less biased, and distorted. Therefore, such a sentencing method lacks objectivity, standardization and scientificity, and it will result in unequal disparities in sentencing, and in contravention of the principle of appropriate adaptation of crimes, it cannot achieve the goal of justice pursued by criminal law.

The benchmark sentencing method, also known as the basic criminal penalty method, is to first determine the basic penalty within the scope of the corresponding legal penalty, find out the benchmark for the penalty, and then consider whether the case has any effect or not, and clearly divide the severity and in the final stage, the basic penalties that have already been determined are made to fluctuate, and the sentence to which the crime is due is determined [Ma, General Theory of Penalty, 刑罚通论, 1995]. The scholars who proposed this method believe that although China's criminal law stipulates that we should

not only explicitly refer to punishment, it must not be explicitly sentenced. The heavy or light penalties for criminals should be established on a certain amount of standard, which is the basic penalty. “The so-called basic punishment is to temporarily ignore the various circumstances of the strict punishment, and only in accordance with the degree of social harmfulness of the crime itself, the sentence is imposed within a certain range of punishment.” “The basic penalty is a reference point that emphasizes lightness, and if it is uncertain, the basic penalty cannot be punished by widening and strict punishment, because it has no basis; basic punishments are not allowed, high or low, and it will also lead to lenient punishment.” [He 1995] This sentencing method obviously has the following two problems: First, the issue of how to establish basic penalties is the benchmark for sentencing. There are quite a lot of differences among the theoretical circles. The main points are as follows [Zhou 1999]: 1. The midline theory, that is, the reference point should be fixed at one-half of the legal penalty range, from above the midline, from below the midline; 2. The theory of sub-grid, that is, a certain number of divisions within the statutory penalty range, adding several benchmarks to deal with complex situations such as heavier and lighter; 3. Situational theory, that is, determining the benchmarks based on the severity of the security situation. The benchmark is floating with the security situation. 4. The main factor theory, the assertion that the determination of the reference point for the use of legal punishment should be based on the factors that play a major role in the size of social harm and demonstrate by examples of investigation statistics. Therefore, those who hold this view emphasize discussing issues through empirical analysis; 5, Focus theory, that the statutory reference point is a major factor in the size of the behavior of social harm, this factor is the focus of the abstract sin. The legal punishment corresponding to the abstract sin's focusing point is the benchmark of sentencing [Zheng 1998]. Therefore, since there is no recognized method for how to establish a benchmark, it is obviously not possible to use the benchmark to commensurate with the sentencing. Secondly, even if a unified benchmark for sentencing is established, how to deal with the severity of punishment on a benchmark basis in a specific case is determined according to the judge's discretion. This sentencing method can only reduce the sentencing deviation to a certain extent, but it cannot fundamentally avoid the occurrence of sentencing bias.

2.3.2 Mathematical methods

Due to various shortcomings of the traditional method of comprehensive assessment of sentencing, mathematical methods are introduced into sentencing.

As Max said, “Any science can only become a true science when it is fully used.” [Su *et al*., Study on the Method of Sentencing Methods, 量刑方法研究专论, 1991] With its wide applicability, high degree of abstraction, and strict logic, mathematical methods make the objective and unity of sentencing possible. The currently known mathematics

penalties like mathematical models, analytic hierarchy process, weighted average test method and penalty points method are mainly introduced in the following.

Mathematical models decompose and quantify crimes and penalties separately. They specify the "crime punishment scales" and "crime punishment scales" and identifies the corresponding points in the "crime punishment scale" according to the scores obtained in the "crime punishment scale." The value is then converted into the corresponding penalty.

The specific method of analytic hierarchy process is improved based on the mathematical model sentencing method. The difference is that designers have used the "multi-layered weighted analysis and decision method" that has emerged in recent years to quantify the social harm of crime. Its quantitative value is more accurate and effective, and it is deduced with a certain mathematical formula to make it reliable in science. Based on the logical reasoning and precision calculations, it is more accurate than the mathematical model of the sentencing method.

The weighted average test method consists of weighted average evaluation and fuzzy comprehensive evaluation. They are used to classify crime scenarios into several levels according to the circumstances of punishment, and then to classify the corresponding number of grade sentences. In accordance with the principle of appropriate punishment for crimes, then with the level of the specific crime scene and check and sentenced to the appropriate sentence. [Yu 1993]

The method of calculating the penalty for penalty points is proposed later in Wuhan [Cai & Xu 1996]. This sentencing method can be summarized as: 1. The statutory penalization of space, on the basis of conviction, regards the legal punishment corresponding to a crime as a space whose length is a number of scales (one scale corresponds to the latter one); 2. The circumstances are divided into degree points, each of which examines each circumstance of severity, then scores, and calculates the total points of the circumstance in the case; 3, from the heavy circumstance points and counterbalance points from the light circumstance to find the total points, if negative, it means that the need for heavy punishment; for the rule is a leniency punishment; 4, from the total score for the best moderate declaration of punishment, if the point is negative, the starting point of the point is the lower limit of the spatial legal limit, if it is positive, it is the upper limit. If one point of the activity indicated by the points is within the legal penalty space, the best moderation is the penalty corresponding to the middle point of the remaining space [Ma, General Theory of Penalty, 刑罚通论, 1995]. The output of this method is the result of the non-consecutive announcement and is related to the precision of the integration of points. For example, for theft, the law provides that the upper limit is 15 years and the lower limit is 6 months. If 100 scales are defined, each scale corresponds to 1.74 months. The output of this method will be proportional to 1.74 months. That is, the points on the penalty space are not in one-to-one correspondence

with the output values of the model method. Therefore, the maximum accuracy of sentencing cannot be achieved.

2.3.3 Expert system

With regard to artificial intelligence, there is currently no clear definition. Professor Nilsson of the Artificial Intelligence Research Center at Stanford University believes that artificial intelligence is a science about knowledge—how to express knowledge and how to acquire knowledge and use knowledge. “Artificial intelligence is the study of how to make computers to do the smart work that only people can do in the past.” [Yan 1995] “Artificial intelligence is a branch of computer science that involves the research, design, and application of intelligent machines. Its immediate goal is to study the use of machines to imitate and implement certain intellectual functions of the human brain and develop related theories and techniques.” [Cai & Xu 1996]

In a broad sense, it is generally accepted that the use of computers to simulate human intelligence behavior falls within the category of artificial intelligence. Artificial intelligence has been widely used in knowledge engineering, expert systems, decision support systems, pattern recognition, natural language understanding, and intelligent robots. Expert system (ES) is one of the most mature applications. The so-called expert system is actually a (or a group of) computer programs capable of solving the difficulties in the field at the level of human experts in a specific field. It has a lot of expert knowledge and experience in a certain area and can use the knowledge of human experts and problem-solving methods to solve problems in this field [Yan 1995]. In other words, the expert system is a program system with a large amount of specialized knowledge and experience. Artificial intelligence technology is used to reason and judge according to the knowledge and experience provided by one or more human experts in a field to simulate the decision process of human experts to solve complex problems that require expert decisions.

The first practical application of the expert system in law was the legal adjudication system (LDS) developed in 1981 [Naik & Lokhanday 2012]. Researchers explored to use it as a practical tool for the application of laws to detect certain aspects of the American civil law system, using models such as strict liability, relative negligence, and damage compensation to calculate the value of compensation for liability cases and demonstrated how to simulate the law experts’ opinions. There came then a lot of all kinds of expert system in law field, including in Chinese law field, such as Judgement System by Technological Intelligent Criminal Law Engineering (JUSTICE) [Steinwart & Christmann 2008].

In general, the sentencing expert system is mainly composed of several components and they are knowledge base, database, inference engine and other parts (which includes knowledge acquisition part, human-machine interface, explanation part and so on). [Su

et al., Sentencing and Computers: A Fair and Rational Application of Sentencing, 量刑与电脑:量刑公正合理应用论, 1989]

The knowledge base is the memory of domain knowledge. It stores expert experience, specialized knowledge and common-sense knowledge, including three parts: legal library, empirical library, and case library. Legal laws, regulations, legislative interpretations, and judicial interpretations related to legal deposits and sentencing are stored in legal library, which is the core of the expert system. The experience library is mainly stored by expert judges, how to correctly apply the experience of legal sentencing, as well as the correct understanding of the law and the theoretical summary of the trial experience. The case library mainly stores typical cases that have been verified by the Supreme People's Court, those have been proved to be accurate in conviction, and those cases reasonable judged by experts. The knowledge base can be modified and supplemented by the knowledge engineer based on the abolition, modification, establishment of the law, the further accumulation of experience, and the increase in the number of cases. Knowledge is the main factor that determines the performance of an expert system. The knowledge base must have good usability, correctness, and perfection.

The database is used to store the initial data in the field and all kinds of information obtained during the reasoning process. The contents stored in the database are some facts that the expert system currently processes, such as the quantitative data of the circumstances of the penalty in the new case.

The inference engine is used to control and coordinate the expert's entire expert system. Based on the current input data, ie the information in the database, knowledge in the knowledge base is used to provide decision-making information according to certain inference strategies. In other words, the criminal facts are combined with all the laws and regulations related to sentencing, such as quantitative sentencing scenarios, discretionary quantitative sentencing scenarios, and professional knowledge and experience of expert judges in the specific use of the sentencing circumstances. The result of combination shall be several "If <condition>, then <form> (if ... then statement) form of the expression of the rules. These rules must be complete and compatible. That is, this set of rules embodies the relationship between all the available evidence and the logical conclusion that can be obtained from the information. When the facts provided by the judges were put into the system, under the control of a certain strategy, the network searched for relevant knowledge from the knowledge base, conducted reasoning judgments and obtained results.

The knowledge acquisition part transforms and processes the knowledge about sentencing into the internal representation of the computer, thus providing means for modifying inappropriate knowledge in the knowledge base, deleting unnecessary knowledge in the knowledge base, and expanding new knowledge in the knowledge base.

The Human-Machine Interface takes the role of communicating with users. It receives sentencing information and translates it into an acceptable internal form of the system, and outputs a penalty result. It can also provide the user with the useful knowledge that the inference engine outputs from the knowledge base.

The explanation part gives the necessary explanation to the inference part, i.e. the sentencing output, so as to provide the convenience for the user to understand the reasoning process and to learn and maintain the system.

The sentencing expert system summarized the experiences of the vast number of judges in handling cases and comprehensively analyzed the basic factors and specific factors related to sentencing in the facts of the case. Based on these factors, the expert knowledge stored in the system is used to make inferences and judgments, and the sentencing conclusions of the expert group on a particular case are obtained, which helps the judge to overcome the interference of non-legal factors outside the court and improve the fairness of sentencing.

However, with the development of computer science and technology, especially artificial intelligence in these years, new artificial intelligence theories and application technologies are emerging, such as machine learning and support vector machine theory. Therefore, it is possible and fantastic to try to apply these newly emerged artificial intelligence theories to computer-assisted sentencing to improve the accuracy and efficiency of computer-assisted sentencing.

3. Machine learning and SVM

In this chapter, the machine learning and support vector machine theory is briefly introduced firstly, and then 4 different frequent used algorithms are introduced, after that the feasibility that apply SVM to the development of sentencing method is analyzed.

3.1 Brief introduction of machine learning and data mining algorithms

Learning is the main symbol of human intelligence and the basic means to gain wisdom. It is an important intelligent behavior that humanity has. According to the AI master H. Simon, learning is the ability of the system to enhance or improve its ability to perform its work in repeated work, to make the system perform better or more efficiently than it did the next time it performs the same or similar tasks [Jian 2004].

3.1.1 Machine learning

Ever since computers were invented, people wanted to know if they could learn. Present computer systems and artificial intelligence systems do not have any learning ability. At most, they have only a very limited ability to learn, and thus cannot meet the new requirements of technology and production. To this end, people have conducted various studies on machine learning with the goal of simulating the basic mechanism of human intelligence and developing more "smart" computer systems. Machine learning is another important research field of artificial intelligence application following the expert system, and it is also one of the core research topics of artificial intelligence and neural computing. Scientists at NASA's JPL Laboratory wrote in "Science" (September 2001): "Machine learning is increasingly supporting the entire process of scientific research.... In a few years, stable and rapid development will be achieved." The purpose of machine learning research is to hope that computers have the ability to acquire knowledge from the real world like human beings. At the same time, they will establish learning computing theory, construct various learning systems, and apply them to various fields. For example, let the computer learn from medical records and obtain the most effective method to treat new diseases; the residential management computer system analyzes the electricity consumption patterns of households to reduce energy consumption; The personal software assistant system tracks the user's interests and selects the online news that is of most interest to them. In 1959, Samuel of the United States designed a chess program [Russell & Norvig 2016]. This program has the ability to learn, and it can improve its chess skills in continuous playing. Four years later, this program defeated the designer himself. After another three years, this procedure defeated the United States' undefeated champion that has been unbeaten for eight years. This program shows people the power of machine learning and put forward many thought-provoking social and philosophical issues. Currently machine learning has been widely used in many fields, such as training computer-controlled vehicles to make it run properly on various types of roads. For example, the ALVINN system [Cuingnet *et al*., 2011] has used its learned

strategy to sprint between the other vehicles on the freeway and traveled 90 miles at 70 mph.

What is Machine Learning? So far, there is no unified definition. In general, machine learning is a discipline that studies how to use machines to simulate human learning activities. The more rigorous formulation is that machine learning is a study of machines that acquire new knowledge and new skills and identify existing knowledge. The "machine" mentioned here refers to a computer. In the traditional sense, machine learning evaluates the dependence of a given system's input and output based on a given training sample, enabling it to make as accurate an estimate of the unknown output as possible. It can be described as: Let W be a problem space and $(x,y) \in W$ be called a sample or object, where x is an n-dimensional vector and y is a value in a category field. Due to the limitation of observation ability, we can only obtain a true subset of W, denoted as $Q \in W$ as the sample set. Thus, an optimal model M is established based on Q, and it is expected that the prediction accuracy of this model for all samples in W is greater than a given constant. This process is called training of the model. After training, it is used to evaluate new samples. In general, machine learning uses numerical modeling methods that are summarized by Wiener as the "black box" principle. That is, the test of the problem space established by the model is only consistent with its input and output, and the model itself does not explain the actual world observed by the problem space [Wang & Shi 2003]. In this way, the modeling process can be described as follows: for a subset of a given problem space, understand it as a function y=f(x), the modeling task is to obtain f so that all the samples in the sample set satisfy a given objective function, and the non-samples in the problem space satisfy a certain accuracy rate.

3.1.2 Decision tree

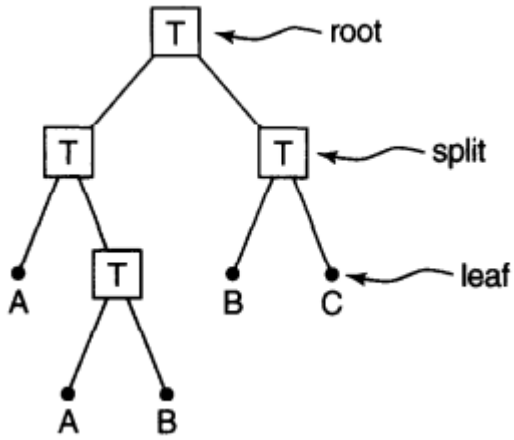


Figure 3.1. A decision tree classifier [Friedl & Brodley 1997]

In Figure 3.1 the decision tree classifier, each box is a node at which tests (T) are applied to recursively split the data into successively smaller groups. The labels (A, B, C) at each leaf node refer to the class label assigned to each observation.

“A decision tree is defined as a classification procedure that recursively partitions a data set into smaller subdivisions on the basis of a set of tests defined at each branch (or node) in the tree ” [Pal & Mather 2001] (Figure 3.1). The tree consists of a root node which is formed from all data, a set of internal nodes and a set of end nodes which is leaves in tree. Only one parent node or more descendant nodes belong to each node in a decision tree. In a decision tree framework, “a data set is classified by sequentially subdividing it according to the decision framework defined by the tree and a class label is assigned to each observation according to the leaf node into which the observation falls.” [Friedl & Brodley 1997]

The so-called decision tree, as its name implies, is a tree, a tree built on the basis of strategic choices. In machine learning, decision tree is a predictive model. It represents a mapping relationship between object attributes and object values. Each node in the tree represents an object. And each forked path (branch) represents a possible attribute value. Each leaf node corresponds to the value of the object represented by the path from the root node to the leaf node. Decision tree has only a single output. If multiple outputs are needed, independent decision trees shall be created to handle different outputs. The machine learning technology that generates decision trees from data is called decision tree learning, generally speaking, this technology can be called decision tree algorithm. To put it plainly, this is a predictive tree algorithm that relies on classification and training. Based on known predictions, it classifies the future.

In other words, the simple strategy of a decision tree is like the screening of a person’s resume during the company’s recruitment interview. If one’s condition is quite good, for example, a Ph.D. graduate from an elite university, then just call him over for an interview. If one graduate from a not famous university, but with rich experience in actual project, then should be also considered to be called and interviewed. That is, the so-called decision making accordingly to specific situation. However, each unknown option can be categorized into existing classification categories.

One example is from the book <Machine Learning> written by Tom M.Michell [Mitchell 1999]. The purpose of the researcher is to find out in what situation will people prefer to play golf through the weather forecast. He learned that the reason that people decide whether to play or not depends on the weather situation. As we can see in Figure 3.2, the weather can be fine, clouds or rain; the temperature is expressed in Fahrenheit; Relative humidity is expressed as a percentage; if it is windy on the day. In this way, we can construct a decision tree as follows.

As Figure 3.2 shows, the numbers in the nodes of the tree are scores or values that determines the decisions in individual leaves for playing or not playing and the greater value in a node gives its result.

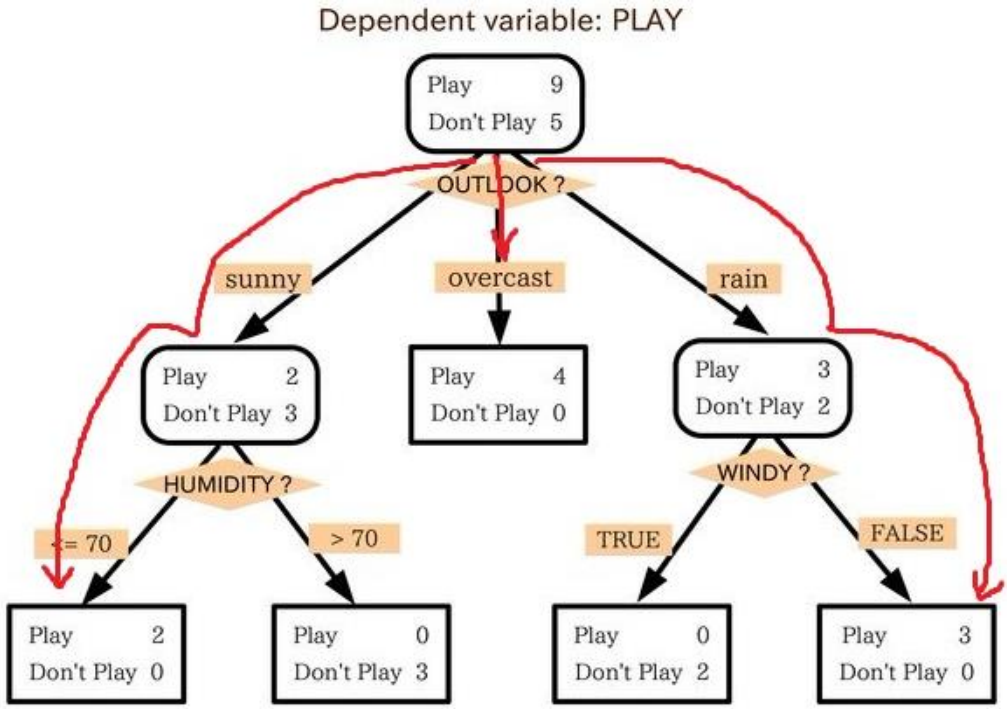


Figure 3.2 Decision tree of people playing golf or not based on weather forecast

The above decision tree corresponds to the following expression:

(Outlook=Sunny ^Humidity<=70) V (Outlook = Overcast) V (Outlook=Rain ^ Wind=Weak).

Decision tree algorithm has several advantages over traditional supervised classification procedures. In particular, decision trees are strictly nonparametric and do not require assumptions regarding the distributions of the input data. In addition, they handle nonlinear relations between features and classes, allow for missing values, and are capable of handling both numeric and categorical inputs in a natural fashion [Fayyad & Irani 1992]. Finally, decision trees have significant intuitive appeal because the classification structure is explicit and therefore easily interpretable. First of all, decision tree algorithm is pretty easy to interpret and explain, people are usually able to understand the meaning expressed by the decision tree after interpretation; Secondly, for decision tree algorithm, data preparation is often simple or unnecessary. Other algorithms often require that the data be generalized first, such as removing redundant or blank attributes; Thirdly, decision tree algorithm can handle both data and conventional attributes. Other algorithms often require single data attributes; Fourthly, decision tree algorithm is a white box model. Given an observed model, it is easy to derive the corresponding logical expression based on the resulting decision tree; Fifthly, it is easy to evaluate the model through static tests; Sixthly, in a relatively short period of time it can produce feasible and well-performing results from large data sources; Lastly, decision trees scale well into large databases, and their size is independent of the size of the database.

Despite all the advantages of decision tree algorithm, it has several disadvantages. Firstly, for data with inconsistent sample sizes, the information gains in the decision tree are biased toward those with more values [Fayyad & Irani 1992]; Secondly, decision tree encounters difficulties when processing missing data; Thirdly, there is an overfitting issue; Lastly, correlations between attributes in the dataset are tending to be ignored.

3.1.3 Naive Bayes algorithm

In machine learning, naive Bayes classifiers are a family of simple "probabilistic classifiers" based on applying Bayes' theorem with strong (naive) independence assumptions between the features. Naive Bayes has been studied extensively since the 1950s. It was introduced under a different name into the text retrieval community in the early 1960s [Russell & Norvig 2016] and remains a popular (baseline) method for text categorization, the problem of judging documents as belonging to one category or the other (such as spam or legitimate, sports or politics, etc.) with word frequencies as the features. With appropriate pre-processing, it is competitive in this domain with more advanced methods including support vector machines [Rennie *et al*., 2003]. It also finds application in automatic medical diagnosis [Rish 2001]. In the statistics and computer science literature, naive Bayes models are known under a variety of names, including simple Bayes and independence Bayes. All these names reference the use of Bayes' theorem in the classifier's decision rule, but naive Bayes is not (necessarily) a Bayesian method [Hand & Yu 2001].

Bayes classifier is based on Bayes' theorem. Naive Bayes classifiers assume that the effect of an attribute value on a given class is independent of the values of the other attributes. This assumption is called class conditional independence. It is made to simplify the computation involved and, in this sense, is considered "naive" [Murphy 2006].

If let $X = \{x_1, x_2, \ldots, x_n\}$ be a sample, whose components represent values made on a set of n attributes. In Bayesian terms, X is considered "evidence". Let H be some hypothesis, such as that the data X belongs to a specific class C. For classification problems, our goal is to determine P (H|X), the probability that the hypothesis H holds given the "evidence", (i.e. the observed data sample X). In other words, we are looking for the probability that sample X belongs to class C, given that we know the attribute description of X. [Murphy 2006] P (H|X) is the posteriori probability of H conditioned on X. In contrast, P (H) is the a priori probability of H. Similarly, P (X|H) is the posteriori probability of X conditioned on H. P(X) is the a priori probability of X.

According to Bayes' theorem, the probability that we want to compute P (H|X) can be expressed in terms of probabilities P (H), P (X|H), and P (X) as Formula 3.1 shows:

$$P(H|\mathbf{X}) = \frac{P(\mathbf{X}|H)\ P(H)}{P(\mathbf{X})}$$

Formula 3.1 Bayes' theorem

And these probabilities may be estimated from the given data.

This is the basic method of the Naive Bayes classifier: on the basis of statistical data, according to certain characteristics, the probability of each category is calculated to achieve classification.

Naive Bayes algorithm has multiple advantages. Naive Bayes model originates from the classical mathematical theory, has a solid mathematical foundation, and stable classification efficiency. Besides, it requires few parameters to estimate, is less sensitive to missing data, and is relatively simple. If the Naive Bayes conditional independence assumption actually holds, a Naive Bayes classifier will converge quicker than discriminative models like logistic regression, so less training data is needed. Theoretically speaking, the Naive Bayes model has the smallest error rate compared to other classification methods. But it's not always the case. This is because the Naive Bayes model assumes that the attributes are independent of each other. This assumption is often not true in practical applications. This has brought some influence on the classification accuracy of the Naive Bayes model. When the number of attributes is large or the correlation between attributes is large, the classification efficiency of the Naive Bayes model is less than that of the decision tree model. Otherwise, the Naive Bayes model has the best performance when the attribute correlation is small. Meanwhile, the priori probability needs to be known and classification decision has a certain error rate.

3.1.4 KNN algorithm

The K-nearest neighbors algorithm (KNN) is a non-parametric method used for classification and regression. KNN can be defined as lazy learning or instance-based learning, which means that not only the function is only approximated locally but all computation is deferred until classification as well [Tan 2006]. The KNN algorithm is one of the simplest algorithms among all in machine learning fields. Either for classification or regression, a useful technique can be to assign weight to the contributions of the neighbors, in order to ensure that the nearer neighbors are able to contribute more compared to the more distant ones. The neighbors are taken from a set of objects. The class for KNN classification and the object property value for KNN regression for these objects are known. Although no explicit training step is required, this can still be regarded as the training set for the KNN algorithm [Tan 2006].

The KNN algorithm is to find the closest K records from the training set and the new data, and then determine the new data category according to their main classification. The algorithm involves three main factors: training set, distance or similar measure, size of K. The main idea of KNN algorithm is like a Chinese old saying: “Jin zhu zhe chi, jin mo zhe hei.” Which means “lies down with dogs must rise up with fleas.” It is an algorithm that infers your category according to your neighbors.

There are three main procedures:

1. Distance calculation: Given the test object, calculate the distance between it and each object in the training set;

2. Neighbor defining: Delineate the nearest K training objects as the nearest neighbors to the test object;

3. Classification: Making classification of test objects based on the main categories of the k nearest neighbors. [Zhang M. L. 2007]

In the process of applying KNN algorithm, as it implies, two definitions are of most importance, “Distance” and “K”.

What is the proper distance measure? The closer the distance means that the more likely these two objects belong to one category. Usually, Euclidean distance is used as the distance measurement.

Whether the value of K is appropriate relates closely to the accuracy of the result of KNN algorithm. An example will be illustrated.

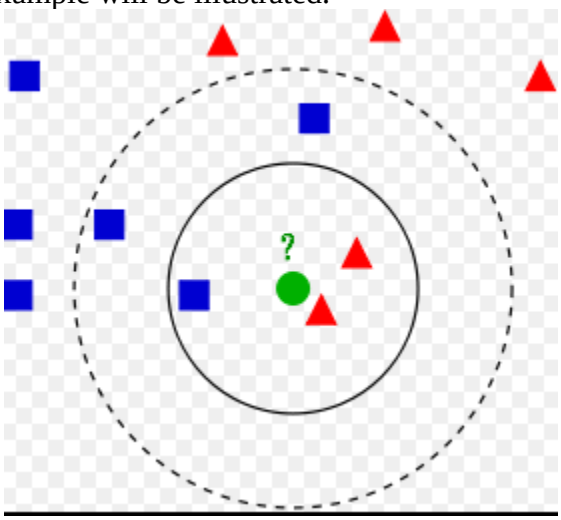


Figure 3.3 Sample points layout for KNN explanation [Mani 2003]

In Figure 3.3, the green circular is the test object that is waiting for a classification. There are two categories in the system: blue square and red triangle. When K is set as 3, actually it is 3 nearest neighbors are to be found around the test object (the green circular), and thus the neighbor circle is the solid line circle. Among the 3 nearest neighbors, 2 of them are red triangles and 1 is a blue square. Thus, 2 > 1, the green circular is classified as more likely to be a red triangle. When K is set as 5, actually it is 5 nearest neighbors are to be found around the test object (the green circular), and thus the neighbor circle is the dotted line circle. Among the 5 nearest neighbors, 3 of them are blue squares and 2 are red triangles. Thus, 3 > 2, the green circular is classified as more likely to be a blue square. It reveals that the value of K has a great influence of the classification result. In this sense, the core in KNN algorithm is to acquire the most suitable K value to achieve an accurate classification.

The KNN algorithm is simple and easy to understand and implemented. Because the KNN algorithm mainly depends on the surrounding limited samples, instead of determining the category by means of classifying the class, the KNN method is more

suitable than other methods for the sample sets with more cross or overlap of class fields. The KNN algorithm is more suitable for the automatic classification of class domains with large sample sizes, while the class domains with smaller sample sizes are more prone to misclassification using this algorithm. The KNN algorithm also shares many disadvantages. It is a lazy learning algorithm which means it lack the process of machine learning. Meanwhile, the output is not that interpretable and the amount of calculation is very large since distances from the test object to every single sample objects need to be calculated. The main disadvantage of this algorithm in classification is that when the sample is unbalanced, for example if a sample has a large sample size, while other samples have a small sample size, it is possible that when a new sample is entered, the samples of the large capacity class in the K neighbors of the sample are in the majority. The algorithm only calculates "nearest" neighbor samples. If the number of samples in a certain class is large, then either such samples are not close to the target sample or such samples are close to the target sample. Both of the two situations will lead to a result that a new test sample is likely to own more neighbors of the certain class than any other classes even if the test sample is much nearer to other classes.

3.2 Support Vector Machine Theory

Machine learning studies look for patterns from observational data and use these rules to predict future or unobservable data. The statistical learning theory is a machine learning rule that specializes in the study of finite sample conditions in practical applications and has developed the supportive vector machine (SVM). [Chen *et al*., 2004]

3.2.1 Brief introduction of SVM

The core idea is that learning machines are adapted to a limited number of training samples and are mainly used in classification and regression problems. The support vector in support vector machines is obtained by solving a convex quadratic optimization problem, which can ensure that the solution found is globally optimal. The so-called optimization refers to the calculation of a specified error function, and the resulting functional relationship fits the “best” (smallest cumulative error) of the sample dataset, thereby minimizing the “total deviation” of all sample points from the hyperplane. In the specific implementation process, the support vector machine transforms the problem of finding the optimal regression hyperplane into a quadratic programming problem and obtains the final regression function of the SVM by solving the optimization problem.

SVM is a type of machine learning method proposed by Vapnik *et al*. [Wikipedia, Support vector machine, From Wikipedia, the free encyclopedia, 2018]. Due to its excellent learning performance, this algorithm has become a research hotspot in the machine learning community. And SVM has been successfully applied in many areas, such as face detection [Osuna *et al*., 1997], handwriting digital recognition [Shanthi & Duraiswamy 2010], text automatic classification [Joachims 1998].

SVM is a statistically based learning method. It is the perfect embodiment of the principle of minimization of structural risks [LeCun *et al*., 1998].

3.2.2 Using SVM to deal with linear problems

Imagine this, one put a lot of balls of two different colors with some regularity on table as Figure 3.4 shows. Then he is supposed to try to separate the balls according to their color using only one stick making the separation of the stick still applicable after more balls are put in. The man tried as Figure 3.5 shows. Then more balls are put in on the table and seemed on ball just laid on the wrong side as Figure 3.6 shows.

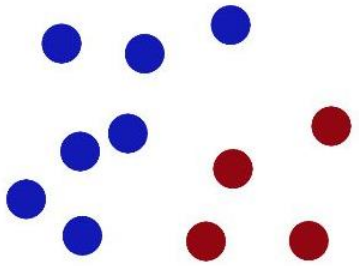


Figure 3.4 Balls layout [Andrew 2000]

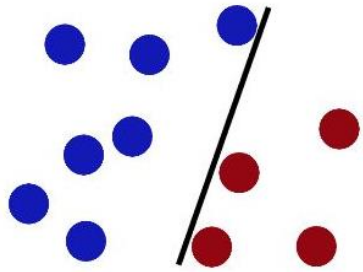


Figure 3.5 Division of balls using a stick [Andrew 2000]

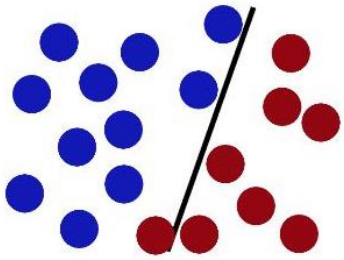


Figure 3.6 Division goes wrong when more balls put in [Andrew 2000]

SVM is the algorithm trying to put the stick in the optimal position so that there is as much separation space as possible on both sides of the stick. In this case, when the optimal position is found, even the devil put more balls onto the table as Figure 3.7 shows, the stick still separates the ball with different colors well as Figure 3.8 shows.

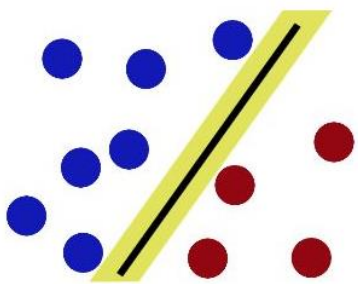


Figure 3.7 The optimum division of balls [Andrew 2000]

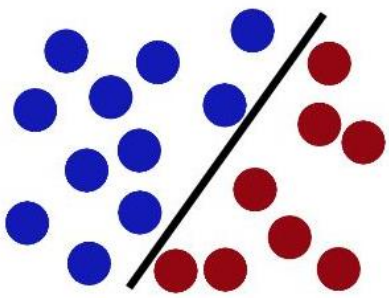


Figure 3.8 Working as more balls put in [Andrew 2000]

Map the case into SVM algorithm, the balls are equivalent to data, the stick is equivalent to classifier or hyperplane. Therefore, the main problem in SVM is trying to find the “stick” which is equivalent to classifier or hyperplane with the training data.

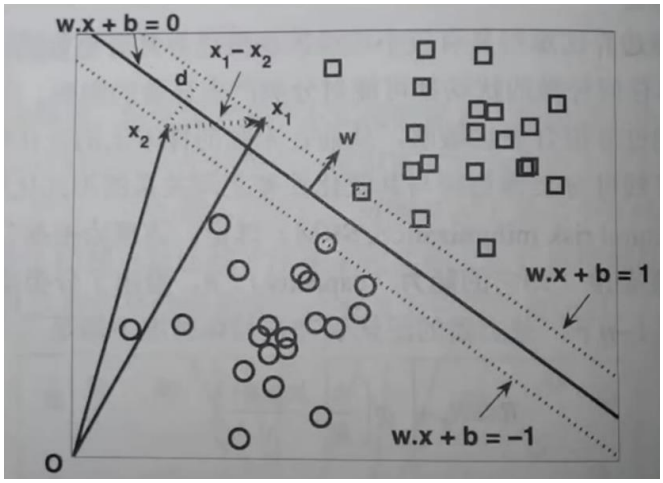


Figure 3.9 SVM geometric presentation [Steinwart & Christmann 2008]

As Figure 3.9 shows in the case, the purpose is trying to separate the circulars and squares. Suppose the solid line L1 is the hyperplane demanded. Slowly pan it down until it meets the circular the very first time and the bottom dotted line L2 can be got. In a similar way, the top dotted line L3 can be got. The circulars and squares on L2 and L3 are called support vectors. L2 and L3 are called supporting plane. L1 is called the hyperplane. The space between L2 and L3 is called isolation zone. The vertical distance between L2 and L3 is called margin. When the isolation zone is taken the maximum, the hyperplane is optimal. It's like the case in which we try to tell if a human is male or female, seldom mistakes are made in this kind of separation because the difference between category "male" and category "female" is huge. The difference in this case is equivalent to isolation zone. Making the distance between two supporting planes the maximum is the core idea of SVM. It is called Maximum Marginal and it is one of the most important theoretical foundations.

If L1's mathematical expression is defined as $w \cdot x + b = 0$ and the mathematical expression of L2 and L3 are defined as $w \cdot x + b = -1$ and $w \cdot x + b = 1$. The margin is defined as d. A vertical vector of L1 is defined as w. A support vector on L2 and L3 are defined as X1 and X2.

Then $\gamma_{1=}\ w^T \cdot X_2 + b = 1$; $\gamma_{2=}\ w^T \cdot X_1 + b = -1$. $\gamma_1 - \gamma_2 \Rightarrow w^T \cdot (X_2 - X_1) = 2$;

$\| w \| \cdot d = 2$

$\therefore d = 2 / \| w \|$

The maximum margin is equivalent to solving the following formula:

$$\max\left(d = \frac{2}{\|\mathbf{w}\|}\right) \Rightarrow \max\frac{1}{\|w\|} \Rightarrow \min\frac{1}{2}\|w\|^2$$

Then the solution will be:

$$\min\frac{1}{2}\|w\|^2 \quad s.t., y_i(w^T x_i + b) \geq 1, i = 1, \cdots, n$$

This is a convex optimization problem that can be solved with the help of the Lagrangian multiplier method using the theory of Karush-Kuhn-Tucker (KKT) conditions, which will not be discussed further in this thesis.

3.2.3 Using SVM to deal with non-linear problems

Back to the case mentioned in the above. What if the placement of ball is as Figure 3.10 shows?

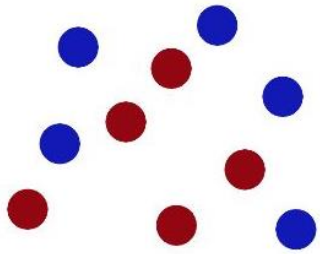


Figure 3.10 New layout of balls of different colours [Andrew 2000]

It seems there won't be a stick in the world that is able to separate the balls of different colours well. It is called a non-linear classification case. In a similar way, the former case that balls can be separated by a stick is called linear classification case. We can try to slap and flip the table throwing the balls into the air and grab a sheet of paper and slip it between the balls as Figure 3.11 shows. Looking at the balls from where the man is standing, the balls will seem split by some curvy line as Figure 3.12 shows.

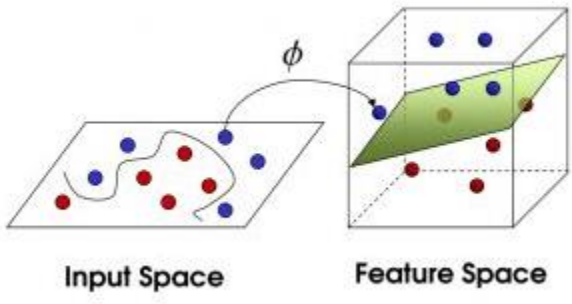


Figure 3.11 The process of mapping the balls into a new hyperplane [Andrew 2000]

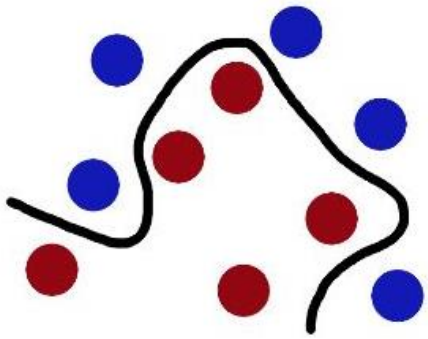


Figure 3.12 The actual division line to separate the balls of different colours [Andrew 2000]

The solution is SVM is to transform the original linear space into a higher-dimensional space. In this high-dimensional linear space, we divide it by a hyperplane can be found to finish the division.

To illustrate it well, the following case is an non-linear classification case as Figure 3.13 shows, the points of different colours cannot be separated with a line and this is called an non-linear problem.

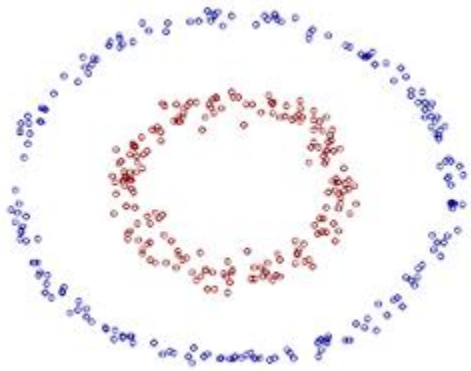


Figure 3.13 A classification problem that linear method cannot solve [Andrew 2000]

The solution is to map two different types of points similar to ellipses to a high dimensional space and the mapping function is:

$Z_1 = X_1^2$, $Z_2 = X_2^2$, $Z_3 = X_2$

Use this function to map the points in the plane above to a 3D space (Z_1, Z_2, Z_3) and after the mapped coordinates are rotated, a linearly separable set of points can be obtained as Figure 3.14 and Figure 3.15 shows. Figure 3.18 is the initial state of coordinates: the points are not possible to be separated in linear way. Figure 3.19 shows the result of the coordinates after the rotation of the coordinates which indicates the process of mapping the initial coordinates and points to a plane and in this way, the points become linear divisible.

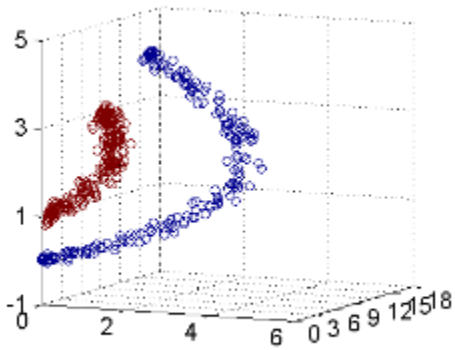


Figure 3.14 Original Coordinates with sample points of two categories [Andrew 2000]

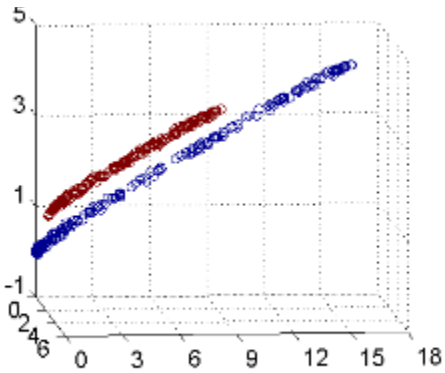


Figure 3.15 Rotated Mapped coordinates with sample points of two categories [Andrew 2000]

The idea behind this solution of SVM to solve non-linear classification problem is easy to understand. It is a truth with many examples in real life. Philosophically speaking, there are no two identical objects in the world, they can be finally divided by adding dimensions. For example, two different books might be the same by dimension of content and book cover colour. But if the dimension of author is added, they might already be different. In case this does not work, more dimensions such as pages, owners, purchase place and so on can be added to finally make the two different books separable.

The method of feature mapping does help solve non-linear classification problems but because a lower dimension problem is converted to a higher dimension problem, the computational complexity and algorithm complexity increase exponentially. SVM provides kernel trick which significantly reduces algorithm complexity. The following case will illustrate what a kernel trick is.

$X = (x_1, x_2, x_3)$ $Y = (y_1, y_2, y_3)$

$F(X) = (x_1x_1, x_1x_2, x_1x_3, x_2x_1, x_2x_2, x_2x_3, x_3x_1, x_3x_2, x_3x_3)$

$K(X, Y) = (<X,Y>)^2$

$X = (1, 2, 3)$ $Y = (4, 5, 6)$

$F(X) = (1, 2, 3, 2, 4, 6, 3, 6, 9)$

$F(Y) = (16, 20, 24, 20, 25, 36, 24, 30, 36)$

$<(F(X), F(Y)> = 16 + 40 + 72 + 40 + 100 + 180 + 72 + 180 + 324 = 1024$

$K(X, Y) = (4 + 10 + 18)^2 = 32^2 = 1024$

As shown above, K(X, Y) ends the same result as <F(X), F(Y)> while it is of much smaller computational complexity and algorithm complexity. And this K(X,Y) is actually a kernel trick in SVM.

3.2.4 Implementation steps, advantages and applications of SVM

The SVM algorithm is based on a strong foundation of statistics and mathematics, so its classification accuracy is unmatched by other similar algorithms. The SVM algorithm implementation steps can be summarized as: (1) Obtain a learning sample; (2) Select a kernel function that performs nonlinear transformation and a penalty factor that punishes the wrong division; (3) Form a quadratic optimization problem; (4) Obtain a support

vector and related parameter values using an optimization algorithm to obtain the above a regression model. [Chen *et al*., 2004]

SVM' application is able to significantly reduce the need for labeled training instances in both the standard inductive and transudative settings so that they are helpful in text and hypertext categorization. Classification of images can also be performed using SVMs. [Gaonkar & Davatzikos 2013] According to experimental results, SVM achieves so much higher search accuracy compared with traditional query refinement schemes after around 3 to 4 rounds of relevance feedback [Decoste & Schölkopf 2002]
. It is same as to image segmentation systems. Some of them are using a modified SVM algorithm that uses the privileged approach which was suggested by Vapnik [Barghout 2015]. Using SVM can help with hand-written characters recognition [Decoste & Schölkopf 2002]. The SVM algorithm has been widely applied in many fields, especially biological fields as well as other sciences. For example, the classification of proteins with up to 90% of the compounds is dealt with in a brilliant accuracy using SVM algorithms. Permutation tests based on SVM weights have been suggested as a mechanism for interpretation of SVM models. [Gaonkar & Davatzikos 2013] [Cuingnet *et al*., 2011] Support vector machine weights have also been used to interpret SVM models in the past [Statnikov *et al*., 2006]. Posthoc interpretation of support vector machine models in order to identify features used by the model to make predictions is a relatively new area of research with special significance in the biological sciences.

The SVM algorithm is widely used and owns an excellent reputation because of its great advantages. The SVM algorithm works great on the classification of clear boundaries because of its solid statistical and mathematical foundation. Besides, the SVM algorithm performs excellent when dealing with non-linear problems and high dimension problems due to its kernel trick. The SVM algorithm has a relatively small sample dependence. It only cares about the support vectors, so when samples are not that rich but filtered, the SVM algorithm can always work well. But of course, most algorithms work better with a bigger sample set with higher quality.

3.3 The feasibility of using SVM algorithm in sentencing

Modern scientific prediction methods have developed rapidly. In recent years, as the development of artificial intelligence technology has matured, machine learning has achieved great success in pattern recognition and complex system control. Machine learning related techniques and algorithms have been well applied in various fields such as biology, medicine, economics, and education. Most of the machine learning and datamining related algorithms like decision tree algorithm, KNN algorithm and Naive Bayes algorithm applies pretty well in the above fields but they all requires to build an accurate, professional and scientific mathematical model to achieve success. The legal professionalism of legal-related issues is extremely high. The lexical sentences in statutory codes are standardized and stylized descriptions of the facts, circumstances, and

sentencing of cases after repeated researching, consideration and discussion by legal experts. Besides, it is a complex procedure to determine what a case really is and how the accurate sentencing should be and needs the accurate elicitation of circumstances, comparison circumstances with the written code, judge through plaintiff and defendant's statement and debate on the factual evidence of the case and using the right of discretion. Thus, not many applications using machine learning methods are realized in legal fields. Even though the application of the algorithms such as KNN algorithm, Naïve Bayes algorithm and so on could overcome the gap between mathematical modeling and feature extraction in the legal field, they still have many flaws. They are either training-only without a learning process, or the algorithm structure needs to be specified in advance or found during training through a heuristic algorithm. The adjustment and determination of the weight coefficient of special certificates is bad unachievable in most algorithms. The training process easily falls into a local minimum point and will have a problem of over-fitting.

The series of excellent machine learning algorithms performed poorly mainly because the theoretical basis is traditional statistics. This is an asymptotic theory when training sample volume tends to infinity. But in real problems, the sample cannot be infinite, sometimes even very limited. The theoretical basis of SVM algorithm is statistical learning theory which provides a unified framework for solving finite sample problems. It incorporates many methods and solve many problems that were difficult to solve. Only those support vector samples matter and if the samples are well-organized and carefully selected i.e. most of the samples are support vectors or are very approximate to support vectors, then there won't be too many samples needed to get the prediction. In legal fields, there have been a lot of cases and verdicts. But first of all, the information of the cases and verdicts is sometimes very sensitive and not easy to achieve. Secondly, the sentencing tendency is changing with the development of society and economy, so there are limited well-judged cases and verdicts in a certain relatively stable period. This forms the very first reason why SVM is feasible for sentencing estimation.

Meanwhile, in a real case, there are a lot of different circumstances such as conviction circumstances, sentencing circumstances, sentencing circumstances for liberal punishment, sentencing circumstances for strict punishment, extenuating sentencing circumstances, severe sentencing circumstances, whether the defendant gives up ill-gotten gains *et al*. that determines the result of sentencing. The circumstances are converted as features in this study. These features are corresponding dimensions of samples and most case samples are high-dimension data. The SVM algorithm has an excellent performance in solving high-dimensional issues. Because of the existence of kernel tricks, the SVM algorithm solves high-dimensional issues with a relatively lower algorithm complexity and takes less time. This forms the second reason why SVM is feasible for sentencing estimation.

The circumstances elicited from cases may have different degree of influence in the sentencing process. For example, there are three sentencing circumstances among all in theft crimes, theft amount, theft times and guilty plea. Obviously, the first two circumstances have a relatively much greater influence on the sentencing results than the circumstance of guilty plea. That means, different features of case samples in sentencing estimation needs to own their own weights. To assign a weight to a feature, the SVM algorithm works better than other algorithms such Naive Bayes algorithm, KNN algorithm and so on. This forms the third reason why SVM is feasible for sentencing estimation.

4. SVM sentencing model, sample acquisition and circumstance extraction

A favorable environment has been provided for SVM to be used in sentencing simulation. Based on SVM theory, SVM model is developed. Sample acquisition becomes the next problem since different countries own different sentencing mode and China's sentencing mode is quite unique. Under the unique sentencing mode of China, sentencing circumstances are important to be considered in a case. The characteristic and classification of sentencing circumstances needs to be understood so that circumstances extraction can be dealt with well.

4.1 The application of machine learning in sentencing: SVM model

In the process of hearing a criminal case, the first is to determine the crime committed by the offender and to weigh his/her criminal liability. That is, to determine the conviction and the corresponding legal punishment for the offense and under the premise of determining the crime correctly, based on the sentencing circumstances of the case, a certain sentencing method is used to correctly evaluate the offender's criminal responsibility. That is, for any criminal case, what is the statutory punishment of the defendant's crime is first confirmed, and within the scope of the legal punishment, all possible circumstances of sentencing are investigated. Finally, according to these circumstances and based on a certain sentencing method, the sentence was derived. The activity diagram is given in Figure 4.1. With a new coming case, two types of circumstances can be extracted from it, one is conviction circumstances and the other is sentencing circumstances. Conviction circumstances are used to determine if the offender is guilty or not. If not, charge shall be rejected, if yes, it comes to the determination of charge. On the other hand, the extracted sentencing circumstances are input to sentencing model to get an announced penalty. The announced penalty together with the determination of charge constitute the verdict or trial.

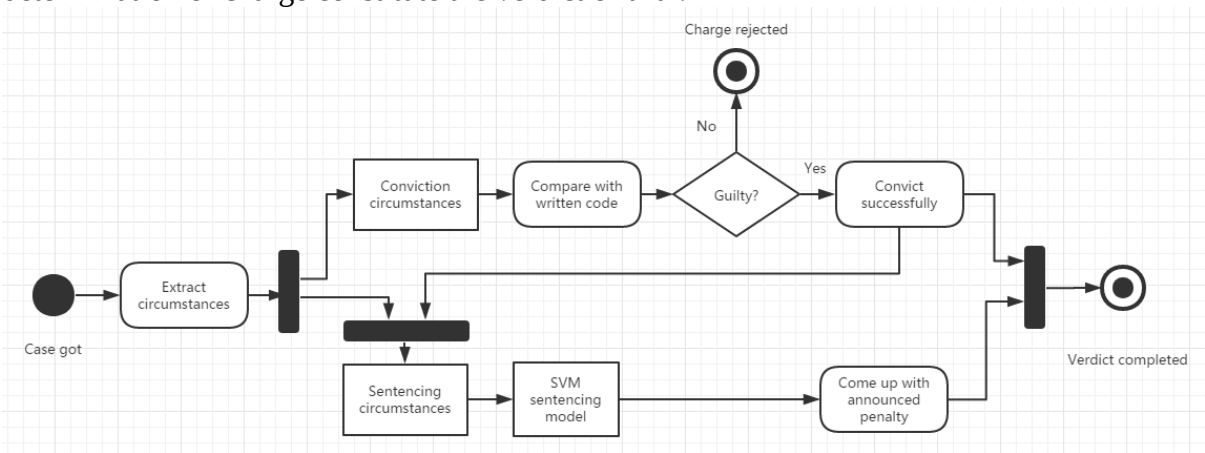


Figure 4.1 The schematic diagram of the process of hearing a case

To truly achieve accurate sentencing, it is necessary to adopt more accurate sentencing methods to reduce sentencing bias under the premise of correct sentencing.

The preceding part of the thesis summarizes various methods of sentencing. The mathematic sentencing method introduces mathematics modeling into sentencing. It provides a way to seek reasonable and fair sentencing. However, the sentencing circumstances in real life are often characterized by uncertainty, cross-correlation, and ambiguity. Therefore, a simple linear mathematical model is difficult to adapt to the comprehensive evaluation of the circumstances.

Support vector machine is a machine learning method proposed for this purpose. Therefore, the introduction of machine learning theory in the field of sentencing is feasible and meets the needs of academic research and the development of sentencing theory. For this purpose, SVM theory is used as a method of measurement in sentencing and it is called SVM sentencing model.

As Figure 4.2 shows: We first collect a number of cases that have been evaluated by experts. These cases must be relatively accurate and able to represent the characteristics of the entire system, so as to ensure that the announced penalty given for the new case SVM model are relatively accurate. Then we extract its circumstances and quantifies the circumstances into numerical representations for use in support vector machine training, until the training result falls within the actual required error range, the sentencing model described above is obtained. When there is a new case, the sentencing circumstance of the new case is extracted and imported in the model. The result of the comprehensive evaluation of the circumstance can be finally mapped out and the sentence can be declared so the announced penalty comes out. So, in essence, the SVM sentencing method is an automated learning process. It can effectively learn useful patterns from previous cases and acquire knowledge for the purpose of penalizing new cases. And once the legislative or judicial interpretation is changed, new cases (samples) can be organized to train the support vector machine so that the model can be adjusted adaptively with changes in the law.

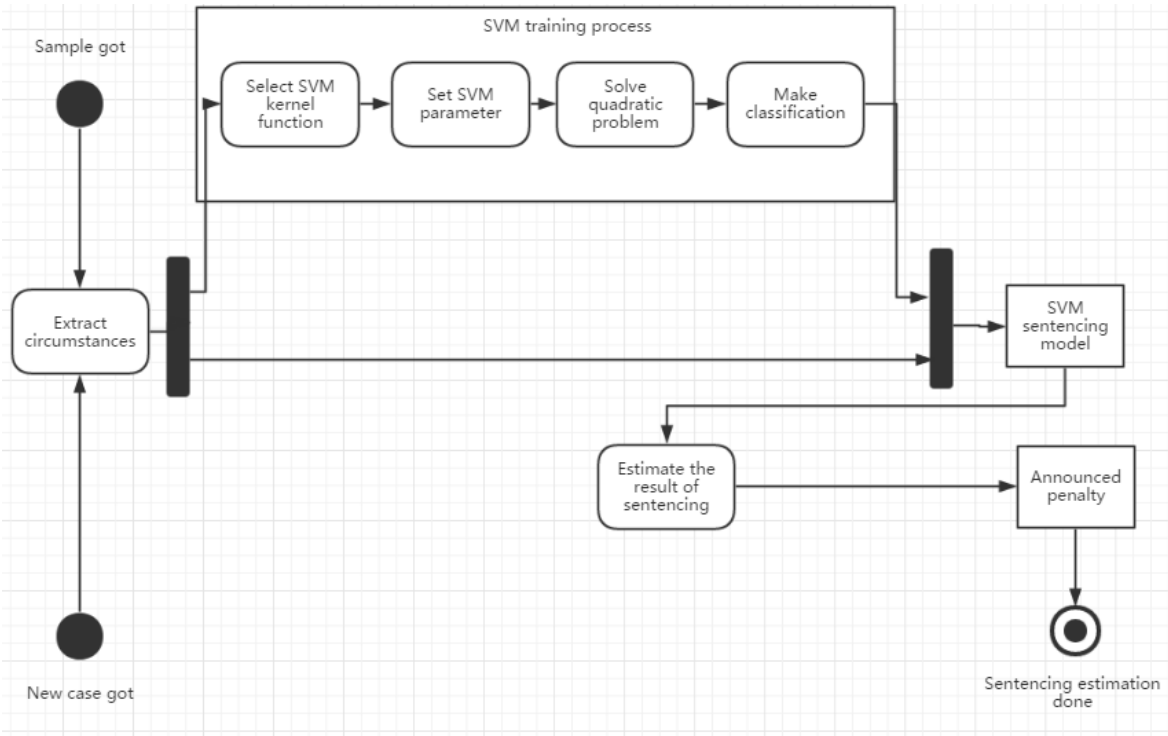


Figure 4.2 The process of how SVM sentencing model works

4.2 Sample acquisition

The most important issue for constructing SVM sentencing models is to collect sample features that is representative enough. The type and the range of sentencing of the sample case should have great coverage and good balance so as to ensure the accuracy and reliability of the model output. How should these cases be obtained? The current sentencing model in China is basically similar to the sentencing model of the Continental Law system. The court is based on facts and the law is the criterion. Conviction and sentencing are strictly based on criminal laws. The jurisprudence is not recognized as one of the sources of the law, and the jurisprudence has no legal effect. Therefore, in order to obtain accurate cases needed to construct the SVM model, it is necessary to draw on the strengths of the sentencing mode of the Common Law system and supplement it, and optimize the current sentencing mode in China.

4.2.1 Continental Law sentencing mode

Written code is the main source of the law of Continental Law countries. They believe that written code is rigorous in terms of wording, generality, wide application, and convenient reference. Sentence by law is the best way and method to maintain legal consistency with the judicial system and to achieve fair and reasonable sentencing. The sentencing mode based on written code is philosophically speaking, the inference procedure from general to individual.

From the above, it can be seen that the Continental Law sentencing mode can also be called deductive reasoning sentencing mode. The advantages of this mode are: 1. It can effectively avoid the loss of control of the legal operation mechanism; 2. It is conducive to the judicial unification and the stability and authority of the law; 3. It is convenient to

quote written code directly. The disadvantages are: 1. The provisions of written code are general, abstract and not specific. When it stipulates that the legal punishment is absolutely determined, it is too rigid and restricts the activism of the judge. When its provision is a relatively definite legal punishment, the amplitude of penalties is so flexible that there is a great difference in sentencing between different judges, resulting in unfair sentencing; 2. The written code is relatively stable, and the legislative process is complex. When socio-economic and political changes develop rapidly, there is a lack of rapid and sensitive resilience.

4.2.2 Common Law sentencing mode

In Common Law countries, the jurisprudence is the main source and form of the law, and the case law is applied. According to the jurisprudence system, the legal rules in a judgment apply not only to the case, but also as a precedent that applies to future cases under the jurisdiction of a court or lower court. As long as the basic facts of the case are the same or similar, it should be handled by the rules set by the precedents. This sentencing mode is philosophically speaking, the inference procedure from individual to individual reasoning. The sentencing operation uses an analogy method. That is:

A certain case of binding has several attributes A, B, C, ... and imposes a penalty X.

After being identified, case M being heard also has several attributes A, B, C... or similar to these attributes.

Therefore, a penalty similar to X should be imposed on case M.

From the above we can see that the sentencing mode in Common Law system can also be referred to as the analogous penalty mode. The advantages of this mode are: 1. The judiciary and the legislation are integrated, and the legislation can be readily adjusted to guide the trial in time; 2. The cases are more specific and it is easy for the judges to imitate; 3. Based on the jurisprudence, it is possible to ensure equal punishments for similar cases and to avoid greater inconsistencies. The disadvantages are: 1. The cases are too complex and inconvenient to use or quote; 2. Judges are different from each other in citing jurisprudence, often resulting in deviation and judicial inconsistency.

4.2.3 The sentencing mode in China and sample acquisition

The sentencing modes of the Continental Law system and the Common Law system cut both ways. At present, the sentencing modes of the two major legal systems show a tendency of mutual integration. Therefore, given the shortcomings of the Continental Law sentencing mode, we can take advantage of the merits of the sentencing mode of the Common Law system and under the trend of the complementary integration of the two large-scale criminal modes, we should optimize our current sentencing mode. We still insist that written code is the main basis for sentencing. On this basis, judges who have experience in trials and criminal experts organized by the Supreme People's Court and various higher people's courts shall recognize, collect, and organize judgments that are representative and well handled. These expertly-recognized cases can be used as samples

for the SVM model we have made. With the development and changes in economic and social life, these sample cases can also be supplemented, updated, and replaced at any time.

4.3 Circumstances extraction in SVM sentencing model

From the point of view of mathematics, after determining the crime, the circumstances of the penalty (variables of the function) become the only factor that influent the penalty and the only factor that determines the best moderation (results of function operations). In the SVM sentencing model, circumstance extraction is an extremely critical link in the application of the sentencing circumstance.

4.3.1 The conception and characteristics of sentencing circumstances

Circumstance refers to the existence and change of things. In criminal jurisprudence, circumstances can be divided into conviction circumstances and sentencing circumstances according to the effect of the circumstances on the criminal activities in each stage. The sentencing circumstance is the situations which the court takes as a basis for deciding the severity of the execution or exemption from punishment. It also refers to the extent of the social harm that affected by the behavior and the degree of the anti-social attributes of the crime actor, based on which to decide whether or not to execute the sentence and the severity of the punishment.

From this conception, the sentencing circumstance has the following characteristics: Firstly, the sentencing circumstances are the actual factual situation concerning crimes and offenders. The sentencing circumstances only have an impact on whether to impose a penalty, what kind of penalty is imposed, and whether the penalty is executed immediately. If a factual situation is a condition for the establishment of a crime, then it cannot become a sentencing. Secondly, the sentencing circumstance is a factual situation that reflects the social harmfulness of the behavior and the personal danger of the crime actor. The social harmfulness of the behavior and the personal danger of the crime actor are the two bases for sentencing. The facts existing in the case, whether it influences the degree of social harmfulness of the behavior or the personal danger of the crime actor, are sentencing circumstances. Thirdly, the sentencing circumstances are objective and exist as criminal acts or have the meaning of criminal law with the implementation of criminal acts. These objectively existing circumstances not only refer to tangible facts that are visible, such as criminal means, criminal consequences, etc. It also includes intangible facts about the subjective factors of the offender, such as the motivation of the crime, the attitude of confession and repentance, and so on. Some of these facts occur when crimes are committed. Some exists before the criminal act, however, only after criminal behavior has been committed, that provides criminal law meanings to them. For example, the status of a national staff member is used as a sentencing circumstance only after the perpetrator commits a crime. Lastly, the sentencing circumstances are factual

situations that are expressly provided for in criminal law or recognized by judicial practice. This is the two forms of confirmation of sentencing circumstances.

4.3.2 The effect of sentencing circumstances

First of all, the sentencing circumstance is the basis for determining the declaration of punishment within the scope of the legal penalty. China's criminal law stipulates a relatively definite statutory punishment for the vast majority of crimes and stipulates several kinds of penalties and rates of punishment for statutory punishments for specific crimes. The statutory punishment has resolved the overall difference problem between a crime and the others in the application of penalty, but it does not solve the problem of how to apply the penalty for crimes of the same nature but in different situations. The verdict of the specific offender can only be finally decided according to the sentencing circumstances of a specific crime.

Meanwhile, the sentencing circumstance is the basis for changing the legal punishment. In general, once the legal punishment is established, it will have unrestricted restrictions on the judge. However, legislators in the determination of the statutory penalties are only concerned with the general situation of a particular crime and cannot reflect all the details of the crime. Therefore, in order to enable the sentencing to take into account the special circumstances that may arise in a specific case, legislators inevitably have to specify certain special factors that can exceed the legal penalty in determining the legal punishment of the general situation. These special factors are the exceptions to the sentencing circumstances. Therefore, in a few cases, the sentencing circumstances have the function of changing the legal punishment and declaring the punishment.

4.3.3 The classification of sentencing circumstances

According to different standards, the sentencing circumstances can be classified differently from various perspectives.

According to whether the criminal law provides express provisions, the sentencing circumstances can be divided into legal circumstances and discretionary circumstances. The legal circumstance is the circumstance that the criminal law expressly stipulates in the measurement of sentencing. The legal circumstances can be divided into must-circumstances and could-circumstances based on the absolute nature of their functions. The must-circumstance is the circumstance that should result lenient or severe punishment based on the provisions of the law. It usually has a necessary influence to the sentencing result. The could-circumstance is the circumstance that can result lenient punishment based on the provisions of the law. It might have an influence to the sentencing result. The discretionary circumstance is a circumstance where the criminal law is not expressly prescribed, and the judge shall consider the facts and the law at the time of measurement. Many circumstances lack clear and specific provisions in the legal provisions. This requires judges to sum up the experience of sentencing and to exercise in the discretion of sentencing.

According to the degree of social harmfulness and personal dangers marked by the circumstances and the severity of the execution, the sentencing circumstances can be divided into sentencing circumstances for liberal punishment and sentencing circumstances for strict punishment. Sentencing circumstances for liberal punishment refers to a circumstance that reflects a light degree of social harmfulness and personal danger which is beneficial to the perpetrator that tending to result in a liberal punishment for the perpetrator. Sentencing circumstances for strict punishment refers to a circumstance that reflects a heavy degree of social harmfulness and personal danger which is not beneficial to the perpetrator that tending to result in a strict punishment for the perpetrator.

According to the time of appearance of the circumstance and the relationship with the crime, the sentencing circumstances can be divided into the pre-crime circumstances, the in-crime circumstances and the post-crime circumstances. The pre-crime circumstance is a circumstance that already exists before the commission of the criminal act and affects the sentencing, such as the usual performance and whether the perpetrator is a recidivist. In general, the pre-crime circumstance only has an impact on the personal danger status of the actor and does not affect the social harmfulness status of the behavior. The in-crime circumstance is a circumstance that occurs during the commission of criminal act and affect the sentencing, such as the criminal motive, the criminal means, and the criminal consequences. The in-crime circumstance not only has an impact on the personal danger status of the actor but affects the social harmfulness status of the behavior as well. The post-crime circumstance refers to the circumstance that affects the sentencing after the execution of the crime. It is mainly the attitude of the offender to the crime that has already been completed, such as surrender, meritorious service, guilty plea and active giving up ill-gotten gains. The post-crime circumstance mainly affects the personal danger status of the actor. For example, when a person surrenders himself after committing a crime, the act of surrendering himself does not have any effect on the social harmfulness of the perpetrator's prior criminal behavior, but merely indicates that the possibility that the perpetrator re-enforces the crime and the personal danger of the perpetrator has been reduced.

4.3.4 Application of sentencing circumstances and circumstances extraction in SVM sentencing model

The process of sentencing is to a large extent the process of application of various sentencing circumstances. In criminal trials, there is a large number of phenomena with multiple sentencing circumstances in one case, and the specific situation will be more complicated. Therefore, summarizing a scientific application method of sentencing circumstances, in order to do orderly judgement and appropriate sentencing, is a thorny issue in trial practice. In the current practice of criminal trials in China, when a single case has multiple sentencing circumstances, there are several applicable methods.

The offset method applies when there are both sentencing circumstances for liberal punishment and for strict punishment in one case, the two circumstances with different effects offset. The drawback of this method is that it requires high levels of cancellation for the circumstances. The nature of liberality and strictness should correspond, i.e. the effect on sentencing must match.

The prior circumstance method works when there are multiple sentencing circumstances in a case. For example, the judge makes a choice based on his own legal values. In the sentencing, he only considers one of the superior sentencing circumstances, and ignores other circumstances. Obviously, this method is one-sided.

The similar term merging applies for several circumstances leading the same direction of sentencing in a case, they are not considered separately, but are added together as another sentencing circumstance to be considered. This method will encounter technical difficulties in how to add the circumstances leading the same direction of sentencing.

The above methods have their own advantages, but all share the same disadvantage, that is, when several sentencing circumstances are coexisted, scientific and reasonable sentencing is hard to be achieved. This, on one hand, highlights the complexity of this issue and on the other hand, it also creates difficulties for the judicial practice, leading to differences in results of the substantive handling of criminal cases on this issue.

However, no matter which method to apply, how to apply a variety of sentencing circumstances and the scope of each penalty sentiment is strictly based on the discretion of the individual judge on the circumstances of the case. Therefore, the status quo of non-standard sentencing has not yet been improved. As Figure 4.3 indicates, in the SVM sentencing model, the sentencing circumstances of each sample case are extracted and quantified into a numerical form. Numerical values are used to replace the role of circumstances in specific cases.

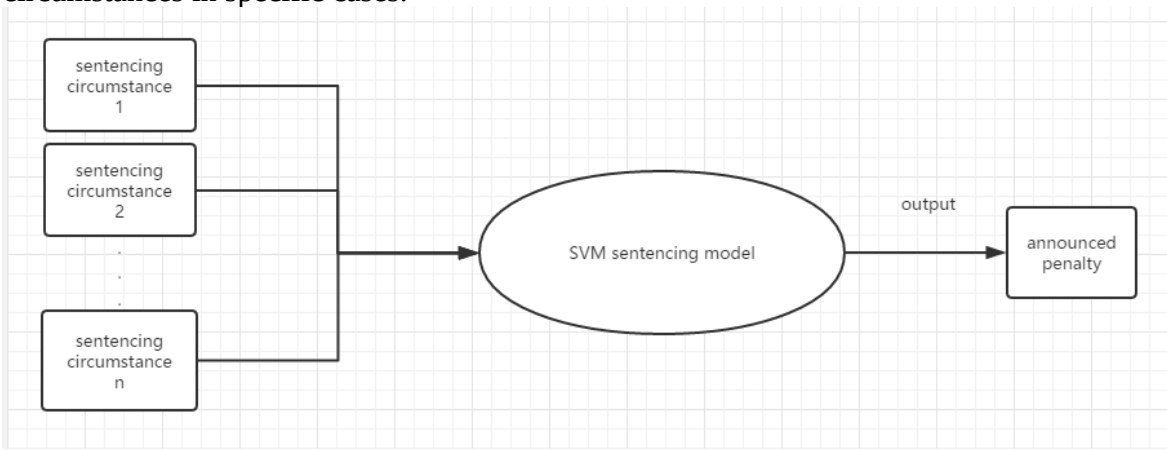


Figure 4.3 The input and output of SVM sentencing model

In this way, there is no need to determine the order of application of each sentencing circumstance. Instead, it just sends the quantified value as an input weigh variable to the SVM sentencing model to get an output declaration sentence. Thus, the differences

caused by judges' perception of the sentencing circumstances can be avoided to a certain extent.

5. Overall framework of sentencing expert system based on SVM

The basic principles of the SVM sentencing model has been discussed. This section is going to present the overall framework of a sentencing expert system using this model as an inference engine for expert systems. It is mainly composed of three parts: an interface part, a data processing part and a database storage part.

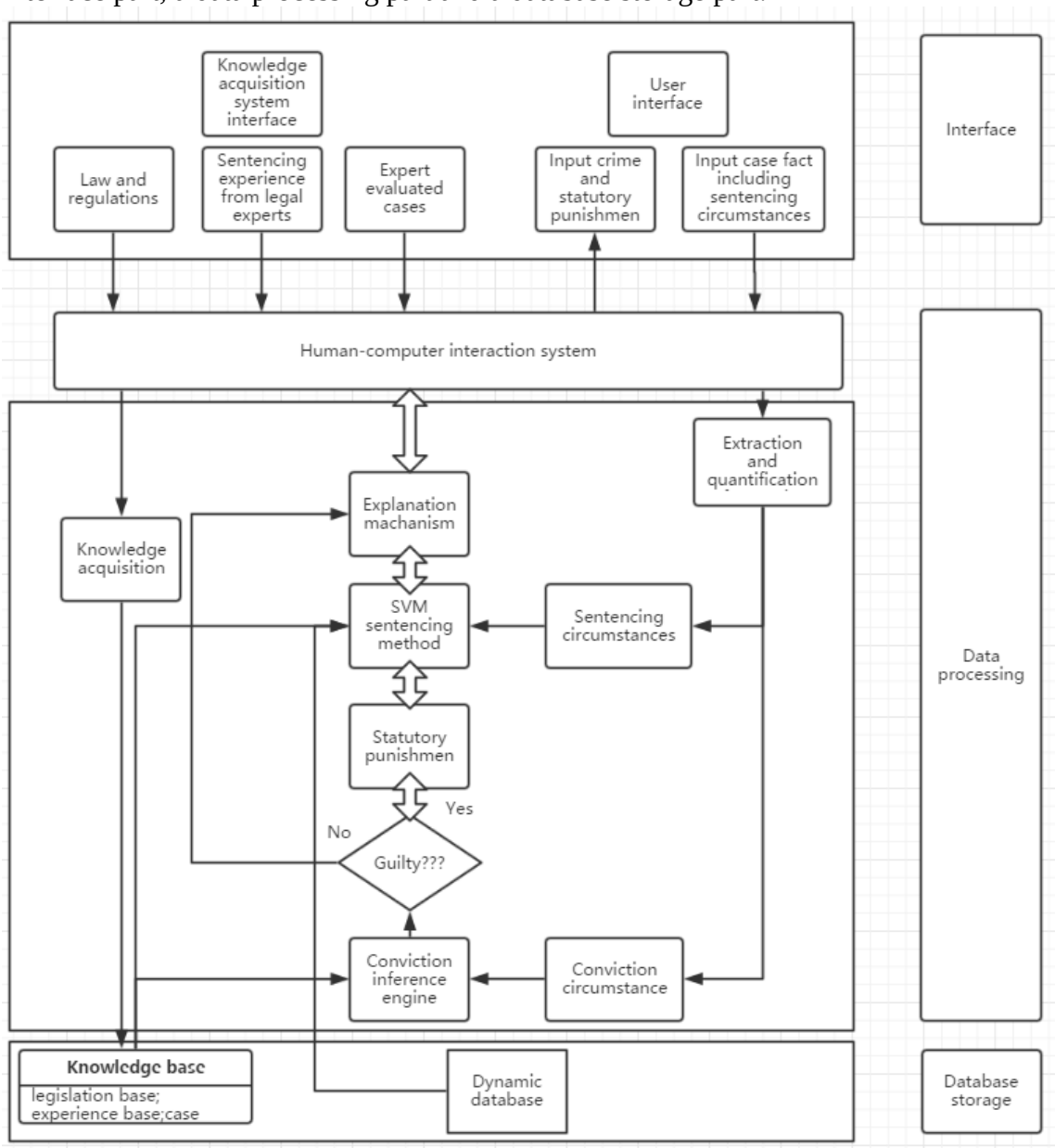


Figure 5.1 Overall framework of sentencing expert system based on SVM sentencing model

As Figure 5.1 shows, this is the framework of the sentencing expert system based on SVM sentencing model.

Expert system engineers collaborate with legal experts to transfer knowledge about the conviction and sentencing laws and regulations, expert judges' sentencing experience, and cases where the expert evaluation is accurate after the sentence is sent to the knowledge base of the expert system through the knowledge acquisition system. The knowledge includes conviction knowledge and sentencing knowledges. The conviction

inference engine judges the offender's criminal constitution based on the convictions and the conviction knowledge in the knowledge base, i.e., the determination of charges and the corresponding legal punishment.

For example, according to the provisions about theft crimes of Article 264 of the Chinese Criminal Law: Theft of public or private property, if the amount is large or multiple thefts, the offender shall be sentenced imprisonment, criminal detention or control of less than three years and shall be concurrently imposed or a single fine; If the amount is huge or there are other serious circumstances, the offender shall be sentenced to fixed-term imprisonment of not less than three years but not more than ten years and shall be concurrently punished with a fine; If the amount is particularly huge or there are other particularly serious circumstances, the offender shall be sentenced to fixed-term imprisonment of not less than 10 years or life imprisonment, and be punished with fines or confiscation of property; In any of the following cases, life imprisonment or the death penalty is imposed and the property is confiscated:

(1) Theft of financial institutions with particularly large amount;

(2) Theft of precious cultural relics with a serious circumstance. [Law 2017]

For a case of theft, the conviction inference engine will judge whether it can constitute a crime and the determination of charges based on the input facts of the crimes entered by the user, the criminal law provisions in the knowledge base, judicial interpretations, and jurisprudence, that is, the process of qualitative judgment of the completion of the crime. Considering that most of the difficult cases in China's judicial practice have problems with the determination of charges, in this system, the conviction inference engine is mainly determined by judges to do the part of determination of charges according to the conviction circumstances by human judgment.

After the completion of the conviction, the SVM sentencing model will pass the sentencing circumstances and the relevant provisions and judicial interpretations on the statutory punishment of the crime, and a reference announced penalty will be given through calculation. If the results obtained are reasonable, then the trial results and sentencing scenarios of this case can be stored as a sample in the case base in the knowledge base. In this system, the knowledge base and the SVM model are separate. Therefore, when the legislation changes, such as when the starting point of sentencing for theft is changed, only the knowledge in the corresponding knowledge base needs to be changed and a new sample (a case for trial under the new law) can be retrained by the SVM sentencing model.

6. Case analysis of the sentencing expert system based on SVM

The Chinese criminal law provides a total of more than 420 criminal charges, and some crimes are also equipped with several sets of a statutory punishment. This thesis selects cases of theft that should be sentenced to fixed-term imprisonment as examples. In the example case analysis, we mainly consider the specific operation of the SVM sentencing model, which assumes that the case has been determined to constitute theft.

6.1 Conceptual SVM sentencing model for theft crimes

The basic steps for establishing computer-assisted sentencing for theft are as the following. First, conduct an in-depth investigation on the status of sentencing for theft and the sentencing experience of the sentencing experts, and collect a number of representative judgments of thefts that have been recognized by experts. Then, from these judgments, the elements of sentencing are extracted and quantified to support the training of the SVM sentencing model. After the completion of the training, the SVM sentencing model serves as a comprehensive evaluation method of the sentencing circumstances. For new cases, the similar circumstances are extracted and quantified, and they are substituted into the trained sentencing model above to obtain the announced penalty of the case. There is open source code of SVM algorithms on many books and websites as well packages in programming languages like Python, such as scikit-learn package which almost includes most machine learning algorithms to be used for classification, regression, clustering, dimensionality reduction, model selection and preprocessing. Scikit-learning package is simple and efficient tools for data mining and data analysis and it is accessible to everyone and reusable in various contexts. Besides, it is open source and commercially usable with a BSD license. Since the package helps solve the realization of the SVM algorithm, the specific steps and implementation are not discussed here.

The author has collected 343 verdicts of theft penalty, and seventeen sentencing circumstances are extracted and quantified as Figure 6.1 shows. The definition of the extracted features, the feature extraction process and quantification process are presented as follows.

	Sentencing Circumstanc	Value	Description
1	Age	0	Adult
		1	Juveniles
2	Criminal capacity	0	person with full assumption of criminal responsibilities
		1	person with limited assumption of criminal responsibilities
3	Deaf-blind	0	Not deaf-blind
		1	Deaf-blind
4	Ordinary theft times	0	No ordinary theft
		Integer	Actual times
		*100	Several times
5	Ordinary theft amount	0	No ordinary theft
		Real number	Actual amount
6	Burglary times	0	No burglary
		Integer	Actual times
		*100	Several times
7	Burglary amount	0	No ordinary theft
		Real number	Actual amount
8	Pickpocket times	0	No pickpocketing
		Integer	Actual times
		*100	Several times
9	Pickpocket amount	0	No pickpocketing
		Real number	Actual amount
10	Turning in	0	Not turning him/herself in
		1	Turning him/herself in
11	Recidivism	0	Not recidivism
		1	Recidivism
12	Crime pattern	0	Accomplished offense
		1	Attempted offense
		2	Preparing offense
		3	Discontinued offense
		4	Both accomplished and attempted offense
13	Joint offense	0	Single criminal
		1	General co-criminal
		2	Principal criminal
		3	Accessory
14	Given up ill-gotten gains type	0	Does not give up ill-gotten gains
		1	Give up ill-gotten gains actively
		2	Give up ill-gotten gains passively
15	Given up percentage	Percentage	Return amount/theft amount*100%
16	Meritoriousness	0	Not meritorious
		1	Meritorious
17	Guilty plea	0	General
		1	Good
		2	Bad

Table 6.1 Sentencing circumstances extraction of theft crime

Burglary (also called *breaking and entering* and sometimes *housebreaking*) [Wikipedia, Burglary, From Wikipedia, the free encyclopedia, 2018] is an unlawful entry into a building or other location for the purposes of committing an offence. Pickpocketing is a form of theft crime that involves the stealing of money or other valuables from the person of a victim without them noticing the theft at the time. Other thefts are usually referred to ordinary theft. Ordinary theft times means the number of times that the offender has committed ordinary theft. Ordinary theft amount refers to the actual amount of the crime or the amount involved in the case. It is same with burglary and pickpocket. All the theft times circumstances are extracted, quantified and then classified into 3 types: 0 means no theft, 100 means several times and actual times of theft will be quantified as the corresponding integer. The theft amount circumstances are extracted, quantified and

then classified into 2 types: 0 means no theft and actual amount will be quantified as the corresponding real number. According to Article 264 of the "Criminal Law of China", theft refers to the act of stealing public or private property for the purpose of illegal possession, or theft of several times, burglary, stealing carrying a weapon, or pickpocket stealing public and private property [Law 2017]. The amount of personal theft of public and private property is "large", starting from $1,000 to $3,000, the offender shall be sentenced criminal detention, control or fixed-term imprisonment within three years, with a concurrent or single fine. The amount of personal theft of public and private property is "huge", starting from $30,000 to $100,000, the offender shall be sentenced fixed-term imprisonment more than three years less than ten years, with a concurrent fine. The amount of personal theft of public and private property is "particularly huge", starting from $300,000 to $500,000, the offender shall be sentenced fixed-term imprisonment not less than ten years or life imprisonment, with a concurrent fine or confiscation of property. [Law 2017] It is easy to be found that ordinary theft times, ordinary theft amount, burglary times, burglary amount, pickpocket times and pickpocket amount are very important sentencing circumstances that will have a great influence on how the sentencing goes. Thus, they are extracted as the feature sentencing circumstances in theft crime.

A joint offense is a person who actively participates in the commission of a crime, even if they take no part in the actual criminal offense. All joint offense circumstances are extracted, quantified and then classified into 4 types: 0 means single criminal, 1 means general co-criminal which indicates that all criminals in the joint offense case share the similar criminal influence, 2 means principle criminal which indicates the criminals in the joint offense case with a larger criminal influence and 3 means accessary which indicates the criminals in the joint offense case with a smaller criminal influence. In joint offense theft crimes, each accomplice is responsible for joint criminal acts based on joint criminal intentions and are responsible for the consequences of harm caused by joint thefts. [Law 2017] Obviously, different announced penalty will be given to joint criminals as different roles in a joint offense case. So, the situation of joint offense is extracted as the feature sentencing circumstance in theft crime.

Age means the actual age of the criminals when they are committing crimes. Deaf-blind means criminals with a disability in listening, seeing or speaking. Criminal capacity means the physical capacity of a person to commit a crime. Age circumstances are extracted, quantified and then classified into 2 types: 0 means adult (>=18) and 1 means juveniles (<18). Deaf-blind circumstances are extracted, quantified and then classified into 2 types: 0 means the offender is not deaf or blind and 1 means the disability lay on the offender. Criminal capacity circumstances are extracted, quantified and then classified into 2 types: 0 means the offender is with full assumption of criminal responsibilities and 1 means the offender is with limited assumption of criminal

responsibilities. Article 17 of the “Criminal Law of China” stipulates: “A person who has reached the age of 16 years commits a crime and shall be held criminally liable. A person who has reached the age of 14 years and under the age of 16 shall be guilty of intentional homicide, intentional injury causing serious injury or death, rape, robbery, and trafficking drugs, setting fires, explosions, and the release of dangerous substances. Person who has reached the age of 14 but under the age of 18 that commits crimes shall be given a lighter or reduced punishment.” [Law 2017] Article 19 of the “Criminal Law of China” stipulates “ If a deaf and mute person or a blind person commits a crime, he or she may be given a lighter, reduced or exempted punishment.” [Law 2017] Criminal capacity refers to the ability of an actor to recognize and control his own behavior in the sense of criminal law necessarily to constitute a crime and bear criminal responsibility. A person who does not possess criminal capacity cannot be held criminally liable even if he or she commit acts that harm the society [Law 2017]. As the above articles indicate, age, deaf-blind and criminal capacity are very related to the sentencing result that they shall be extracted as the feature sentencing circumstances in theft crime.

Turning in means that the offender informs the relevant authorities of the crime and accepts the referee before the crime is detected. After the criminals are arrested, they may report and expose other people's criminal acts, including criminals in joint crime cases who expose other crimes other than the joint offender's joint crimes, and are verified to be true. They provide important clues to detect other cases and are verified to be true; and prevent others from committing criminal activities. To assist the judicial authorities in arresting other criminal suspects (including co-offenders); having other prominence for the country and society. The above situations should be deemed to be considered as mertoriousness. Guilty plea means the attitude towards what the criminals have done after they got caught. Turning in circumstances are extracted, quantified and then classified into 2 types: 0 means no turning in circumstances and 1 means the existence of turning in circumstances. Meritoriousness circumstances are extracted, quantified and then classified into 2 types: 0 means no meritoriousness circumstances and 1 means the existence of meritoriousness. Guilty plea circumstances are extracted, quantified and then classified into 3 types: 0 means general attitude, 1 means good attitude and 2 means bad attitude. For criminals who turn themselves in, punishment can be reduced or mitigated. For those whose crimes are relatively minor, they may be exempted from punishment. (Law, 2017) According to the provisions of Article 68, paragraph 1 of the Criminal Law of China, criminals help to reveal other people's crimes (verified to be true) or provide important clues so that other cases are solved, the behavior is called meritoriousness. Those who has achieved meritoriousness can be reduced or mitigated in sentencing. [Law 2017] Criminal with a good guilty plea can be given a lighter penalty [Law 2017]. As the above articles indicate, turning in, meritoriousness and guilty plea shall be extracted as the feature sentencing circumstances in theft crime.

The recidivism refers to a criminal who has been sentenced to a certain penalty within a statutory period after he has been punished by certain penalties, after the execution of the penalty is completed or after the excuse. Recidivism circumstances are extracted, quantified and then classified into 2 types: 0 means the offender is not a recidivism and 1 means the offender is. Recidivism must be severely punished. [Law 2017] For recidivism, no probation is applicable [Law 2017]. Recidivist does not apply to be paroled [Law 2017]. As the above articles indicate, recidivism shall be extracted as the feature sentencing circumstances in theft crime.

The criminal pattern refers to various criminal forms that deliberately commit crimes have stopped due to subjective and objective reasons at each stage of the process of their occurrence, development and completion [Law 2017]. The criminal pattern includes the completion form of crimes and the unfinished form of crimes. The completion form of crimes is also referred to accomplished offense (whose quantification code or value is 0). The unfinished form of crimes includes attempted offense (whose quantification code or value is 1), preparing offense (whose quantification code or value is 2) and discontinued offense (whose quantification code or value is 3). As the above articles indicate, crime pattern shall be extracted as the feature sentencing circumstances in the theft crime.

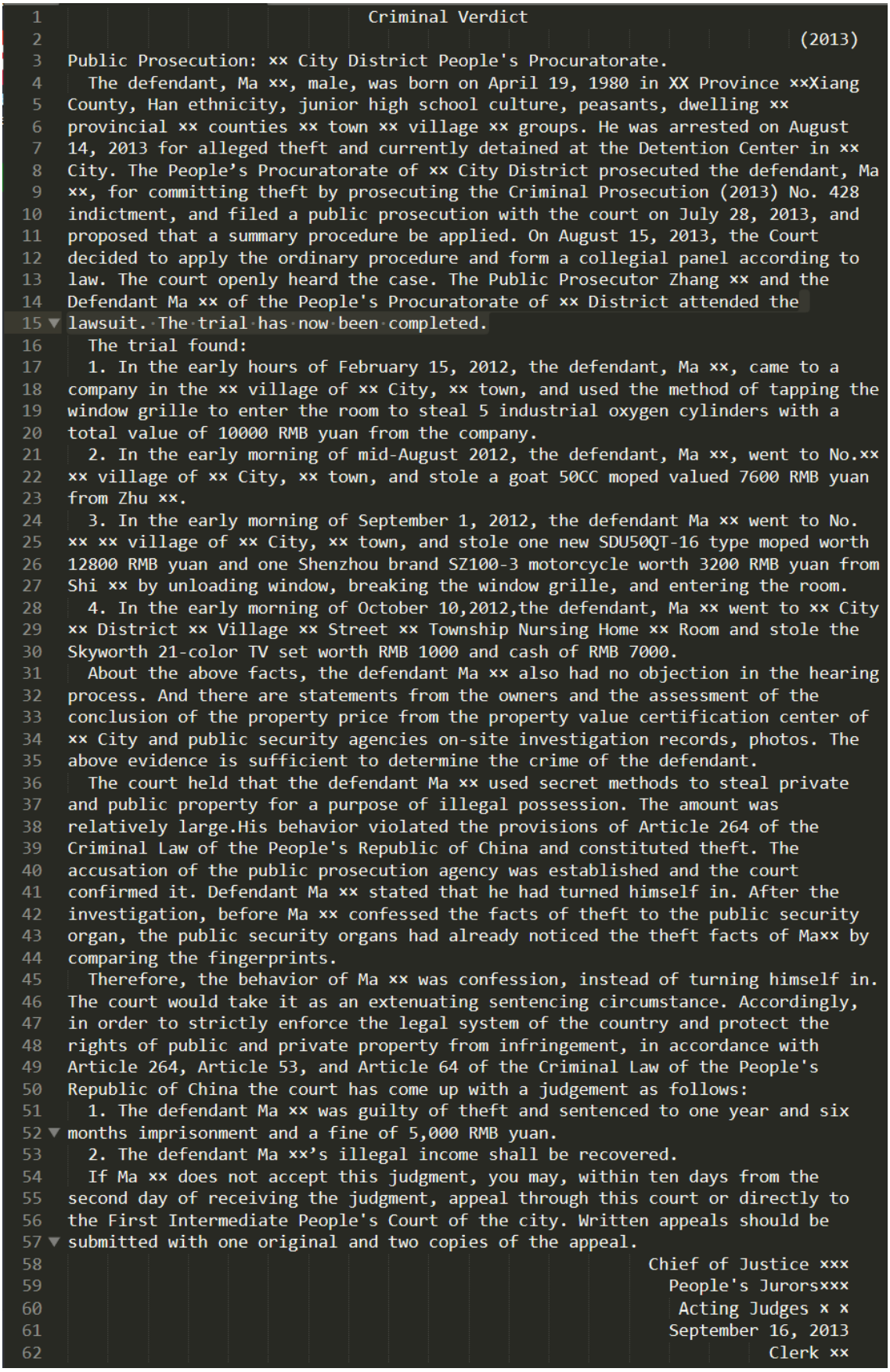
There is no written code which indicates if the offender gives up the ill-gotten gains positively or actively with a what percentage can or must be given a lighter punishment. But usually in daily sentencing activities, if the offender gives up the ill-gotten gains especially actively, he or she is possibly to get a lighter punishment. Besides, the higher the percentage of the ill-gotten gains are returned, the higher the possibility is for the offender to receive a lighter punishment and the lighter the punishment is likely to be given to the offender. In this sense, given up ill-gotten gains type and amount percentage shall be extracted as the feature sentencing circumstances in the theft crime. Given up ill-gotten gains type circumstances are extracted, quantified and then classified into 3 types: 0 means the offender does not give up ill-gotten gains, 1 means the offender gives up ill-gotten gains actively and 2 means the offender gives up ill-gotten gains passively. Given up percentage circumstances are extracted, quantified to the percentage that return amounts weigh in the whole theft amount.

All 17 feature sentencing circumstances of crime theft extracted are listed in the Table 6.1 above.

6.2 Case analysis of the conceptual SVM sentencing model

To illustrate the process of extracting and quantifying sentencing, two verdicts are quoted as examples. Both two verdicts and the following 343 cases of theft crime mentioned in Chapter 6.2, Appendix A and Appendix B are collected from the 2rd Criminal Court of Zhejiang Higher People's Court.

Verdict One:



Criminal Verdict

(2013)

Public Prosecution: ×× City District People's Procuratorate.

The defendant, Ma ××, male, was born on April 19, 1980 in XX Province ××Xiang County, Han ethnicity, junior high school culture, peasants, dwelling ×× provincial ×× counties ×× town ×× village ×× groups. He was arrested on August 14, 2013 for alleged theft and currently detained at the Detention Center in ×× City. The People's Procuratorate of ×× City District prosecuted the defendant, Ma ××, for committing theft by prosecuting the Criminal Prosecution (2013) No. 428 indictment, and filed a public prosecution with the court on July 28, 2013, and proposed that a summary procedure be applied. On August 15, 2013, the Court decided to apply the ordinary procedure and form a collegial panel according to law. The court openly heard the case. The Public Prosecutor Zhang ×× and the Defendant Ma ×× of the People's Procuratorate of ×× District attended the lawsuit. The trial has now been completed.

The trial found:

1. In the early hours of February 15, 2012, the defendant, Ma ××, came to a company in the ×× village of ×× City, ×× town, and used the method of tapping the window grille to enter the room to steal 5 industrial oxygen cylinders with a total value of 10000 RMB yuan from the company.

2. In the early morning of mid-August 2012, the defendant, Ma ××, went to No.×× ×× village of ×× City, ×× town, and stole a goat 50CC moped valued 7600 RMB yuan from Zhu ××.

3. In the early morning of September 1, 2012, the defendant Ma ×× went to No. ×× ×× village of ×× City, ×× town, and stole one new SDU50QT-16 type moped worth 12800 RMB yuan and one Shenzhou brand SZ100-3 motorcycle worth 3200 RMB yuan from Shi ×× by unloading window, breaking the window grille, and entering the room.

4. In the early morning of October 10,2012,the defendant, Ma ×× went to ×× City ×× District ×× Village ×× Street ×× Township Nursing Home ×× Room and stole the Skyworth 21-color TV set worth RMB 1000 and cash of RMB 7000.

About the above facts, the defendant Ma ×× also had no objection in the hearing process. And there are statements from the owners and the assessment of the conclusion of the property price from the property value certification center of ×× City and public security agencies on-site investigation records, photos. The above evidence is sufficient to determine the crime of the defendant.

The court held that the defendant Ma ×× used secret methods to steal private and public property for a purpose of illegal possession. The amount was relatively large.His behavior violated the provisions of Article 264 of the Criminal Law of the People's Republic of China and constituted theft. The accusation of the public prosecution agency was established and the court confirmed it. Defendant Ma ×× stated that he had turned himself in. After the investigation, before Ma ×× confessed the facts of theft to the public security organ, the public security organs had already noticed the theft facts of Ma×× by comparing the fingerprints.

Therefore, the behavior of Ma ×× was confession, instead of turning himself in. The court would take it as an extenuating sentencing circumstance. Accordingly, in order to strictly enforce the legal system of the country and protect the rights of public and private property from infringement, in accordance with Article 264, Article 53, and Article 64 of the Criminal Law of the People's Republic of China the court has come up with a judgement as follows:

1. The defendant Ma ×× was guilty of theft and sentenced to one year and six months imprisonment and a fine of 5,000 RMB yuan.

2. The defendant Ma ××'s illegal income shall be recovered.

If Ma ×× does not accept this judgment, you may, within ten days from the second day of receiving the judgment, appeal through this court or directly to the First Intermediate People's Court of the city. Written appeals should be submitted with one original and two copies of the appeal.

Chief of Justice ×××
People's Jurors×××
Acting Judges × ×
September 16, 2013
Clerk ××

Figure 6.1 The content of Verdict One in Sublime Text 3

Extracted and quantified sentencing circumstances are given in Table 6.2 and the announced penalty is given in Table 6.3. For example, to extract and quantify the sentencing circumstance of Age mentioned in Table 6.2, it is easy to be found that the offender was born on April 19 1980 according to Line 4 in Figure 6.1. The verdict is announced on September 16 2013 according to Line 61 and all the crimes were committed in 2012 according to Line 17-30. Thus, the offender was 32 when he committed the crime and 33 when he was sentenced. Obviously, the offender was adult and the corresponding quantification code of adult in Age is "0". It is the same way with the rest extraction and quantification.

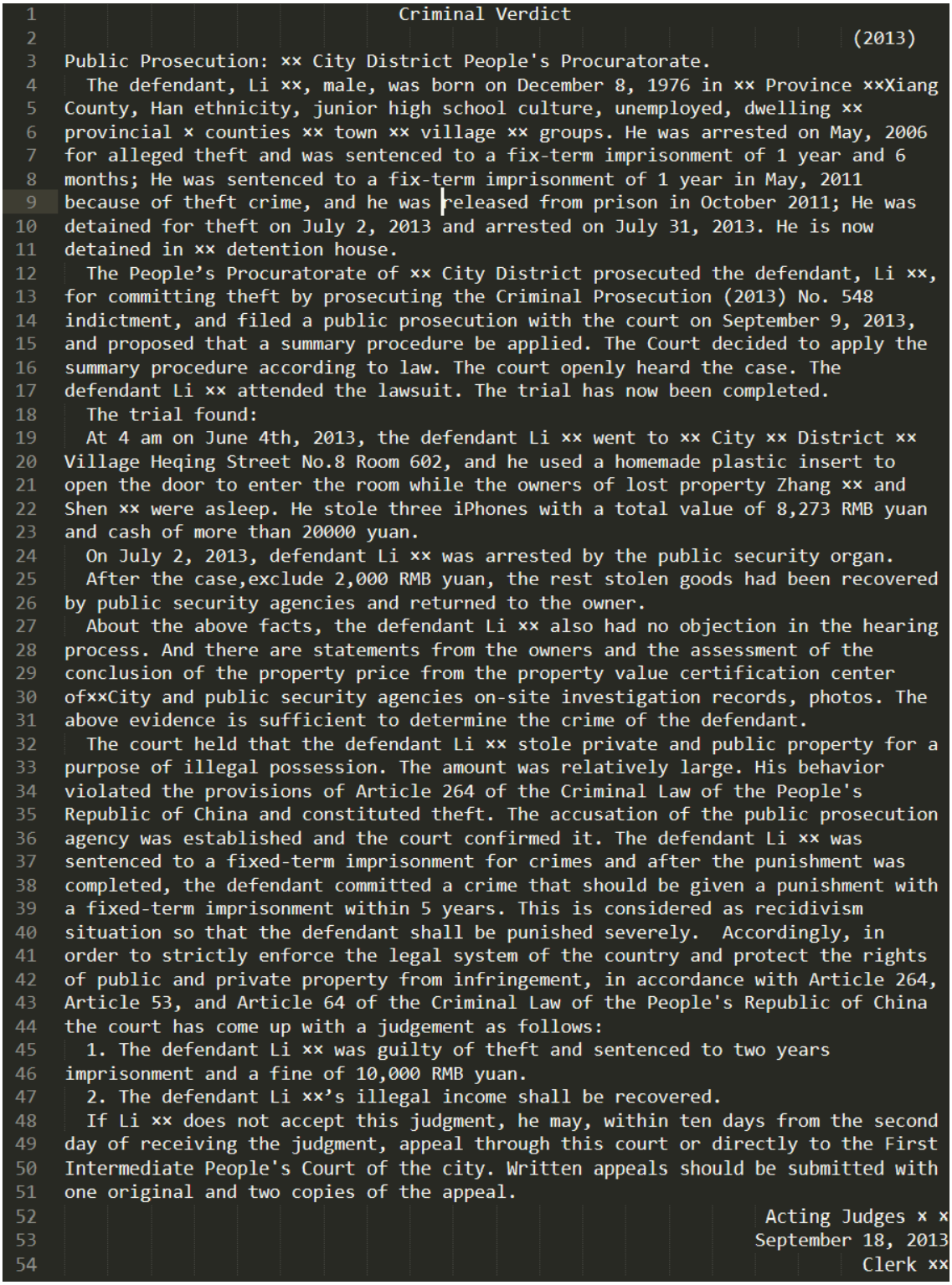
	Sentencing Circumstances	Quantification Code	Source Line
1	Age	0	4, 17-30, 61
2	Criminal Capacity	0	4,5
3	Deaf-blind	0	
4	Ordinary theft times	2	24-30
5	Ordinary theft amount	18000	24-30
6	Burglary times	2	17-23
7	Burglary amount	23600	17-23
8	Pickpocket times	0	
9	Pickpocket amount	0	
10	Turning in	0	
11	Recidivism	1	17-30
12	Crime pattern	0	
13	Joint offense	0	
14	Given up ill-gotten gains type	0	
15	Given up amount percentage	0	
16	Meritoriousness	0	
17	Guilty plea	1	45

Table 6.2 Sentencing circumstances extraction and quantification from Verdict One

Feature Name	Quantification Code	Source Line
Sentencing Result (month)	18	51-52

Table 6.3 Sentencing result from Verdict One

Verdict Two:



Criminal Verdict

(2013)

Public Prosecution: ×× City District People's Procuratorate.

The defendant, Li ××, male, was born on December 8, 1976 in ×× Province ××Xiang County, Han ethnicity, junior high school culture, unemployed, dwelling ×× provincial × counties ×× town ×× village ×× groups. He was arrested on May, 2006 for alleged theft and was sentenced to a fix-term imprisonment of 1 year and 6 months; He was sentenced to a fix-term imprisonment of 1 year in May, 2011 because of theft crime, and he was released from prison in October 2011; He was detained for theft on July 2, 2013 and arrested on July 31, 2013. He is now detained in ×× detention house.

The People's Procuratorate of ×× City District prosecuted the defendant, Li ××, for committing theft by prosecuting the Criminal Prosecution (2013) No. 548 indictment, and filed a public prosecution with the court on September 9, 2013, and proposed that a summary procedure be applied. The Court decided to apply the summary procedure according to law. The court openly heard the case. The defendant Li ×× attended the lawsuit. The trial has now been completed.

The trial found:

At 4 am on June 4th, 2013, the defendant Li ×× went to ×× City ×× District ×× Village Heqing Street No.8 Room 602, and he used a homemade plastic insert to open the door to enter the room while the owners of lost property Zhang ×× and Shen ×× were asleep. He stole three iPhones with a total value of 8,273 RMB yuan and cash of more than 20000 yuan.

On July 2, 2013, defendant Li ×× was arrested by the public security organ.

After the case,exclude 2,000 RMB yuan, the rest stolen goods had been recovered by public security agencies and returned to the owner.

About the above facts, the defendant Li ×× also had no objection in the hearing process. And there are statements from the owners and the assessment of the conclusion of the property price from the property value certification center of××City and public security agencies on-site investigation records, photos. The above evidence is sufficient to determine the crime of the defendant.

The court held that the defendant Li ×× stole private and public property for a purpose of illegal possession. The amount was relatively large. His behavior violated the provisions of Article 264 of the Criminal Law of the People's Republic of China and constituted theft. The accusation of the public prosecution agency was established and the court confirmed it. The defendant Li ×× was sentenced to a fixed-term imprisonment for crimes and after the punishment was completed, the defendant committed a crime that should be given a punishment with a fixed-term imprisonment within 5 years. This is considered as recidivism situation so that the defendant shall be punished severely. Accordingly, in order to strictly enforce the legal system of the country and protect the rights of public and private property from infringement, in accordance with Article 264, Article 53, and Article 64 of the Criminal Law of the People's Republic of China the court has come up with a judgement as follows:

1. The defendant Li ×× was guilty of theft and sentenced to two years imprisonment and a fine of 10,000 RMB yuan.

2. The defendant Li ××'s illegal income shall be recovered.

If Li ×× does not accept this judgment, he may, within ten days from the second day of receiving the judgment, appeal through this court or directly to the First Intermediate People's Court of the city. Written appeals should be submitted with one original and two copies of the appeal.

Acting Judges × ×

September 18, 2013

Clerk ××

Figure 6.2 The content of Verdict Two in Sublime Text 3

Extracted and quantified sentencing circumstances are given in Table 6.4 and the announced penalty is given in Table 6.5. The process of extraction and quantification is as same as while dealing with Verdict One that is mentioned earlier.

	Sentencing Circumstances	Quantification Code	Source Line
1	Age	0	4-11, 19-26, 53
2	Criminal Capacity	0	4,5
3	Deaf-blind	0	
4	Ordinary theft times	0	
5	Ordinary theft amount	0	
6	Burglary times	1	19-23
7	Burglary amount	28273	19-23
8	Pickpocket times	0	
9	Pickpocket amount	0	
10	Turning in	0	
11	Recidivism	1	6-11
12	Crime pattern	0	
13	Joint offense	0	
14	Given up ill-gotten gains type	2	25-26
15	Given up amount percentage	0.93	25-26
16	Meritoriousness	0	
17	Guilty plea	0	

Table 6.4 Sentencing circumstances extraction and quantification from Verdict One

Feature Name	Quantification Code	Source Line
Sentencing result (month)	24	45-46

Table 6.5 Sentencing result from Verdict One

After all data has been extracted, 343 samples are obtained (Appendix A). Then put the samples into the SVM sentencing model for training, until the sum of the error value predicted by the SVM sentencing model and the actual sentencing value is very small, stop training. The model obtained in this way serves as a comprehensive assessment model for sentencing circumstances. Specific training steps are not discussed in this paper. There are many training algorithms in the field of machine learning. When there are new cases that need to be judged, the SVM sentencing model can give a reference sentencing result based on the value of the extracted circumstances. As a preliminary experiment, it is assumed that we have judged from the conviction and inference engine of the expert system that the crime constitutes theft and that the statutory punishment is fixed-term imprisonment. If it is a different crime or legal punishment, it will be judged by other SVM sentencing models (the training samples used to train the SVM are different, and the model gained is different).

7. Conclusion

The thesis will be of benefit to the development of sentencing methods and an important link connecting law and computer science. We live in the information era and under the Internet+ background, software is such a subject that should be linking with one after another practical science in social and daily life fields. Using computer science and technology to help promote social progress from all angles and perspectives, isn't this the mission that the era has given to us?

There has never been any similar SVM-based sentencing expert systems or conceptual models for such expert systems in Chinese criminal law field. But actually, the sentencing deviation is so common and the society really misses the true justice. Law itself is a complex field, not to mention that sentencing is not just about case, for example, evidences must be argued to determine to be acceptable or not, which has already made SVM-based sentencing expert systems not easy to develop. From this perspective, the idea of using machine learning methods or SVM algorithms to assist sentencing is like guiding lights in the Pacific Ocean, which points a direction where developing artificial intelligence knowledge can be combined tight to legal issues, to the ship of law and judge who seems have lost direction to march to future land. On the other hand, unlike English or other languages, Chinese language is a way complex language with extraordinary situations. For example, one Chinese character can have several meanings in different contexts, one meaning can be expressed in several different Chinese words and characters etc.. From this perspective, the natural language processing is not easy to be applied in Chinese criminal law field. The thesis developed a conceptual model for theft crimes based on SVM to help analyze a case of theft crime accurately and extract and quantify the necessary sentencing circumstances from the case to be input in the SVM sentencing model and sentencing expert system, to get an estimation of optimal sentencing range for lawyers, judges and society for reference, so that the announced penalty won’t be too far away from the punishment that is proper for the offender. The conceptual sentencing model is an outstanding model that makes sentencing expert systems possible to deal with cases of theft crimes. Of course, more conceptual sentencing model for other crimes shall be developed in the future. But just for the record, there is way too many differences between different crimes, so the sentencing circumstances and their weighs differ greatly. This won’t be an easy task, but still, it is a significant task and will be an outstanding work.

Compared with previous sentencing expert systems, the expert system presented in the thesis is based on advanced SVM sentencing models and has the ability to self-learn and implement complex comprehensive evaluations. However, there are still many issues that need to be further studied. This not only includes legal issues but also issues that have not yet been completely resolved in the study of machine learning and support vector machines.

First of all, from the perspective of computer-assisted sentencing, it is to rule out human interference, but such exclusion cannot be complete.

Second, the issue of the extraction and quantification of the SVM sentencing model circumstance, as a tentative experiment, only a dozen sentencing scenarios have been drawn. However, actual cases are much more complicated than this. Many factors that affect sentencing cannot be identified only from the analysis of the verdict and are far from being summarized by these ten scenarios.

At the same time, the SVM model is a black box model. It learns from previous successful cases (samples) and obtains relatively accurate sentencing models (embodied in model parameter adjustments). It is then used in the new case's comprehensive evaluation of sentencing scenarios. The model itself has no explanation for the actual world observed in the problem space.

Finally, the SVM model is to train the model by learning previous cases. Therefore, whether the case is good or not and whether the case can represent the characteristics of the entire system will affect the accuracy of SVM sentencing. In addition, the system must be constructed and maintained by knowledge engineers. The deviations in the process that knowledge engineers gain knowledge understanding and expression from expert communication also affect the accuracy of the system. This requires legal experts and knowledge engineers to work closely and work together.

As social life and legal relations become more and more complex, legal practice requires new thinking tools. Otherwise, whether lawyers, prosecutors or judges will not be able to bear the burden of increasing legal documents and law cases. Although the legal reasoning is very complicated, it has relatively stable objects (cases), relatively clear preconditions (legal rules, legal facts) and strict procedural rules, and a definitive conclusion of the judgment must be drawn. It provides extremely favourable conditions for artificial intelligence simulation and machine learning applications. Moreover, long-term accumulation of legal knowledge and complete archives provide rich and accurate information for the simulation of acquisition, expression, and application of legal knowledge. All these provide impetus and favourable conditions for the development of machine learning-based sentencing expert systems. As computer-assisted sentencing provides relatively uniform reasoning standards and evaluation criteria for judicial trials, it can assist judges in making consistent judgments, thus achieving the justice goals pursued by criminal law. The machine learning sentencing expert system referred to in this paper is a conceptual design on the basis of predecessors, trying to do the best to exclude the influence of human factors, making the sentencing balanced, offering a kind of original justice. However, there are still many problems in the study of sentencing expert systems that remain unresolved. There is still a long way to go before it is widely accepted and practically used by the whole society. This requires legal researchers and experts in the field of artificial intelligence to work together. It can be said that the

sentencing expert system is the ideal target for the pursuit of sentencing research. In the existing research, people seem to witness the dawn of a fusion of people (judge) and computers (computer-assisted sentencing). In this difficult but fascinating field, the efforts to create a sentencing expert system that is fully human-intelligent may be like chasing the rainbow on the horizon. Although it will never be possible to catch up, people will find numerous precious treasures in the process of chasing the rainbow.

References

Andrew, A. M. (2000). *An Introduction to Support Vector Machines and Other Kernel-based Learning Methods.* Cambridge: Cambridge University Press.

Barghout, L. (2015). Spatial-Taxon Information Granules as Used in Iterative Fuzzy-Decision-Making for Image Segmentation. *Granular Computing and Decision-Making, Springer*. Cham.

Cai, Z., & Xu, G. (1996). *Artificial Intelligence and Its Applications, 人工智能及其应用.* Beijing: Tsinghua University Press.

Chen, Y., Yu, X., Gao, x., & Feng, h. (2004). A New Approach to Dealing with Nonlinear Classification and Regression Problems: An Introduction to Support Vector Machine Methods, 处理非线性分类和回归问题的一种新方法: 支持向量机方法简介. *Journal of Applied Meteorology*(3), 345-354.

Chiongson, R. A., Desai, D., Marchiori, T., & Woolcock, M. (2012). Role of law and justice in achieving gender equality. *World Development Report*, 7778105-1299699968583.

Cuingnet, R., Rosso, C., Chupin, M., Lehéricy, S., Dormont, D., Benali, H., & Colliot, O. (2011). Spatial regularization of SVM for the detection of diffusion alterations associated with stroke outcome. *Medical image analysis, 15*(5), 729-737.

Decoste, D., & Schölkopf, B. (2002). Training invariant support vector machines. *Machine learning, 46*(1-3), 161-190.

Fan, F. (1994). General Theory of Penalty. *China University of Political Science and Law Press, 刑罚通论*, 481.

Fayyad, U. M., & Irani, K. B. (1992). On the handling of continuous-valued attributes in decision tree generation. *Machine learning 8*(1), 87-102.

Friedl, M. A., & Brodley, C. E. (1997). Decision tree classification of land cover from remotely sensed data. *Remote sensing of environment 61*(3), 399-409.

Gaonkar, B., & Davatzikos, C. (2013). Analytic estimation of statistical significance maps for support vector machine based multi-variate image analysis and classification. *Neuroimage 78*, 270-283.

Hand, D. J., & Yu, K. (2001). Idiot's Bayes—not so stupid after all? *International statistical review 69*(3), 385-398.

He, B. (1995). *Criminal law textbooks, 刑法教科书.* Beijing: China Legal Publishing House.

Jescheck, H.-H. (2004). The general principles of international criminal law set out in Nuremberg, as mirrored in the ICC Statute. *J. Int'l Crim. Just. 2*, 38.

Jian, H. L. (2004). Machine Learning Research and Application in Artificial Intelligence, 人工智能中的机器学习研究及其应用. *Journal of Jiangxi Science & Technology Normal University*, 84-86.

Joachims, T. (1998). Text categorization with support vector machines: Learning with many relevant features. *European conference on machine learning*. Berlin, Heidelberg.

Kahan, D. M., & Nussbaum, M. C. (1996). Two conceptions of emotion in criminal law. *Columbia Law Review 96*(2), 269-374.

Ke, G. (1989). Analysis of the Relationship between the Circumstances of Rape and Sentencing, 强奸罪情节与两性关系的调查分析. *Politics and Law*(6), 16-19.

Lamb, W. R. (1925). *Plato in twelve volumes.* Cambridge, MA, and London.

Law, C. e. (2017). *Criminal law of the People's Republic of China.* http://www.66law.cn/tiaoli/9.aspx: Retrieved on June 21.

LeCun, Y., Bottou, L., Bengio, Y., & Haffner, P. (1998). Gradient-based learning applied to document recognition. *Proceedings of the IEEE, 86*(11), 2278-2324.

Ma, K. (1995). *General Theory of Penalty, 刑罚通论*. Wuhan: Wuhan University Press.

Ma, K. (2002). Improper use of penalties and their countermeasures, 刑罚适用失当及其对策. *People's Justice*(10), 42-43.

Mani, I. Z. (2003). kNN approach to unbalanced data distributions: a case study involving information extraction. *Proceedings of workshop on learning from imbalanced datasets* .

Mitchell, T. M. (1999). Machine learning and data mining. *Communications of the ACM 42*(11), 30-36.

Murphy, K. P. (2006). *Naive bayes classifiers.* University of British Columbia Press.

Naik, V. M., & Lokhanday, S. (2012). Building a Legal Expert System for Legal Reasoning in Specific Domain-A Survey. *International Journal of Computer Science & Information Technology, 4*(5), 175.

Osuna, E., Freund, R., & Girosit, F. (1997). Training support vector machines: an application to face detection. *Computer vision and pattern recognition, 1997. Proceedings*, 130-136.

Pal, M., & Mather, P. M. (2001, 11). Decision tree based classification of remotely sensed data. *Proceedings of 22nd Asian Conference on Remote Sensing*, p. p9.

Rennie, J., Shih, L., Teevan, J., & Karger, D. (2003). *Tackling the poor assumptions of Naive Bayes classifiers (PDF).* ICML.

Rish, I. (2001). An empirical study of the naive Bayes classifier. *IJCAI 2001 workshop on empirical methods in artificial intelligence, 3*(22), 41-46.

Russell, S. J., & Norvig, P. (2016). *Artificial intelligence: a modern approach.* Malaysia: Pearson Education Limited.

Shanthi, N., & Duraiswamy, K. (2010). A novel SVM-based handwritten Tamil character recognition system. *Pattern Analysis and Applications, 13*(2), 173-180.

Statnikov, A., Hardin, D., & Aliferis, C. (2006). Using SVM weight-based methods to identify causally relevant and non-causally relevant variables. *sign, 1*(4).

Steinwart, I., & Christmann, A. (2008). *Support vector machines*. Springer Science & Business Media.

Su, H., Zhang, G., & Shi, J. (1989). *Sentencing and Computers: A Fair and Rational Application of Sentencing, 量刑与电脑:量刑公正合理应用论*. Shanghai: Baijia Publishing House.

Su, H., Zhang, G., Shi, J., & Hu, J. (1991). *Study on the Method of Sentencing Methods, 量刑方法研究专论*. Shanghai: Fudan University Press.

Tan, S. (2006). An effective refinement strategy for KNN text classifier. *Expert Systems with Applications*, 290-298.

Wang, Y., & Shi, C. (2003). Machine Learning Research, 机器学习研究. *Journal of Guangxi Normal University*(2), 1-15.

Wikipedia. (2018, May 3). *Burglary, From Wikipedia, the free encyclopedia*. (Wikipedia) Retrieved from https://en.wikipedia.org/wiki/Burglary

Wikipedia. (2018, May 23). *Support vector machine, From Wikipedia, the free encyclopedia*. Retrieved from https://en.wikipedia.org/wiki/Support_vector_machine

Yan, C. (1995). Summary of Expert System, 专家系统综述. *Shanxi Electric Power Technology*(6), 61-64.

Yu, W. (1993). *The General Theory of Sentencing, 量刑通论*. Wuhan: Wuhan University Press.

Zhang, M. L. (2007, 7). ML-KNN: A lazy learning approach to multi-label learning. *Pattern recognition*, pp. 2038-2048.

Zhang, Z. (1999). *A Comparative Study of Chinese and Western Legal Culture, 中西法律文化比较研究*. Nanjing: Nanjing University Press.

Zheng, W. (1998). *Study on Misdemeanor Misconduct, 重罪轻罪研究*. Beijing: China University of Political Science and Law Press.

Zhou, G. (1999). Sentencing Benchmark Study, 量刑基准研究. *Chinese Law*, 127-138.

Appendix A: Case data of theft crimes

Actual meanings for the acronyms in the following long table are as follows.

CN = Case number

A = Age

CC = Criminal capacity

DB = Deaf-blind

OTT = Ordinary theft times

OTA = Ordinary theft amount

BT = Burglary times

BA = Burglary amount

PT = Pickpocket times

PA = Pickpocket amount

TI = Turning in

R = Recidivism

CP = Crime pattern

JO = Joint offense

GIT = Give up ill-gotten gain type

GIAP = Give up ill-gotten gain percentage

M = Meritoriousness

GP = Guilty plea

AP (Month) = Announced penalty

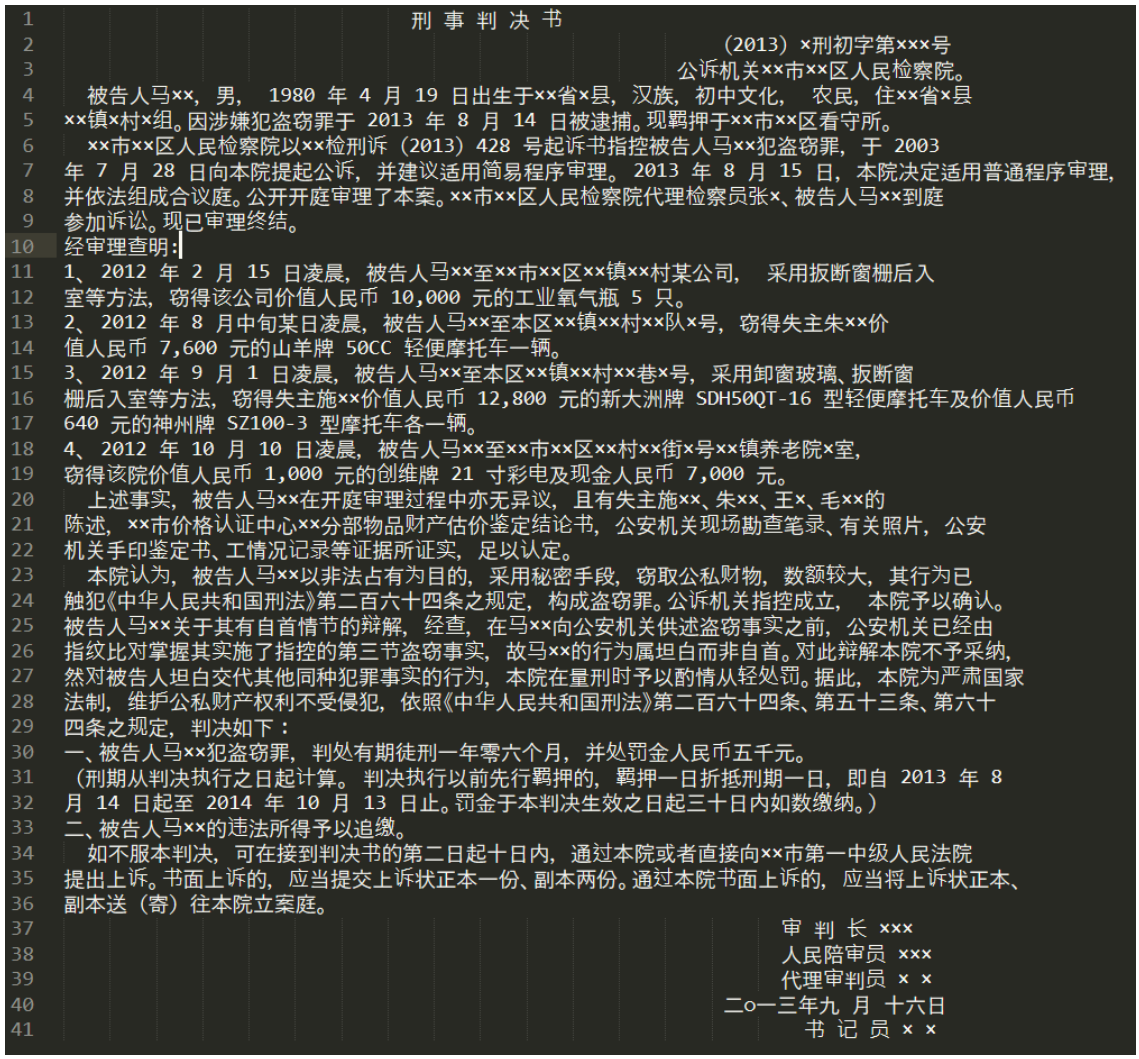
The following long table includes the extraction and quantification of 343 case of theft using the conceptual SVM sentencing model for theft crimes mentioned in Chapter 6 Section 1. The extraction process and meaning of each values can be found in Chapter 6 Section 1 and 2.

CN	A	CC	DB	OTT	OTA	BT	BA	PT	PA	TI	R	CP	JO	GIT	GIAP	M	GP	AP
1	0	0	0	0	0	0	0	1	1125	0	0	0	0	2	1	0	1	7
2	0	0	0	0	0	0	0	1	1350	0	0	0	0	0	0	0	1	7
3	0	0	0	0	0	0	0	2	2272	1	0	0	0	0	0	0	0	6
4	0	0	0	1	2375	0	0	0	0	1	0	0	1	0	0	0	0	6
5	0	0	0	0	0	0	0	2	3345	1	0	0	0	0	0	0	1	7
6	0	0	0	6	3910	0	0	0	0	1	0	4	0	0	0	0	1	8
7	0	0	0	0	0	1	7550	0	0	0	0	0	0	1	1	0	0	7
8	0	0	0	0	0	0	0	1	1620	0	0	0	0	0	0	0	1	9
9	0	0	0	4	2750	0	0	0	0	1	0	0	0	0	0	0	1	6
10	0	0	0	1	3268	0	0	0	0	0	0	0	0	0	0	0	1	6
11	0	0	0	100	4403	0	0	0	0	0	0	0	0	0	0	0	1	8
12	0	0	0	0	0	0	0	1	2300	0	0	0	0	0	0	0	1	8
13	0	0	0	1	5110	0	0	0	0	0	0	1	0	0	0	0	0	9
14	0	0	0	0	0	2	1440	0	0	1	1	0	0	0	0	0	0	6
15	0	0	0	0	0	4	2430	0	0	0	0	0	0	0	0	0	1	6
16	0	0	0	0	0	0	0	1	1360	0	0	0	0	2	1	0	1	6
17	0	0	0	2	3910	0	0	0	0	1	0	0	0	0	0	0	0	6
18	0	0	0	2	3080	0	0	0	0	1	1	0	0	0	0	0	0	8
19	0	0	0	1	3000	0	0	0	0	1	1	0	0	0	0	0	0	9
20	0	0	0	1	5270	0	0	0	0	0	0	0	0	2	1	0	0	9
21	0	0	0	2	10650	0	0	0	0	0	4	1	0	0	0	0	0	8
22	0	0	0	0	0	2	1390	0	0	1	1	0	1	0	0	0	0	7
23	1	0	0	11	14605	0	0	0	0	1	0	0	0	0	0	0	0	10
24	1	0	1	1	2546	0	0	0	0	0	1	0	0	0	0	0	0	6
25	0	0	0	1	8546	0	0	0	0	0	0	0	1	2	1	0	1	6
26	0	0	0	1	7930	0	0	0	0	0	0	0	0	0	0	0	0	8
27	0	0	0	0	0	1	6000	0	0	0	0	0	0	0	0	0	0	8
28	0	0	0	0	0	0	0	1	1850	0	0	0	0	0	0	0	0	6
29	0	0	0	1	3076	0	0	0	0	0	0	0	1	2	1	0	1	6
30	0	0	0	1	2800	1	1045	0	0	0	1	1	0	0	0	0	0	10
31	0	0	0	1	2399	0	0	0	0	0	1	0	0	0	0	0	0	6
32	0	0	0	1	5905	0	0	0	0	0	1	0	0	0	0	0	0	6
33	0	0	1	0	0	0	0	1	1100	0	0	0	0	0	0	0	1	6
34	0	0	0	1	2118	0	0	0	0	0	0	1	0	0	0	0	1	8
35	0	0	0	2	6660	0	0	0	0	1	0	0	1	1	0	0	0	10
36	0	0	0	1	3700	0	0	0	0	1	0	0	1	1	0	0	0	6
37	0	0	0	2	4030	0	0	0	0	1	0	0	0	0	0	0	0	8
38	0	0	0	2	3568	0	0	0	0	1	0	0	0	0	0	0	0	9
39	0	0	0	2	4432	0	0	0	0	1	0	0	0	0	0	0	0	6
40	0	0	0	1	2490	0	0	0	0	0	1	0	0	2	0.9	0	0	6
41	0	0	0	1	5000	0	0	0	0	0	0	0	0	2	0.6	0	0	6
42	0	0	0	1	8000	0	0	0	0	0	0	0	0	1	0.6	0	1	10
43	0	0	0	0	0	3	3712	0	0	1	0	0	1	0	0	0	0	8
44	0	0	0	3	5311	0	0	0	0	1	0	0	1	0	0	0	0	10
45	0	0	0	0	0	4	4100	0	0	1	0	0	0	0	0	0	0	6
46	0	0	0	0	0	0	0	1	1040	0	1	0	0	0	0	0	0	8
47	0	0	0	0	0	6	1850	0	0	1	1	4	0	0	0	0	0	10
48	0	0	0	0	0	0	0	3	1990	1	1	4	0	0	0	0	0	9
49	0	0	0	0	0	0	0	1	1000	0	0	0	0	0	0	0	0	6
50	0	0	0	1	3063	0	0	0	0	0	1	0	0	2	1	0	0	6
51	0	0	0	0	0	0	0	4	5586	1	0	0	1	2	0.5	0	0	10
52	0	0	0	1	5100	0	0	0	0	0	0	0	0	2	1	0	0	7
53	0	0	0	7	3860	0	0	0	0	1	0	0	0	2	1	0	0	7
54	0	0	0	0	0	3	3900	0	0	1	1	0	0	0	0	0	0	6
55	0	0	0	1	2685	0	0	0	0	0	1	0	1	0	0	1	0	6
56	0	0	0	0	0	4	5210	0	0	1	1	0	0	0	0	0	0	8
57	0	0	0	1	2042	0	0	0	0	0	1	0	0	0	0	0	1	6

58	0	0	0	0	0	0	0	1	1256	1	0	0	0	0	0	0	0	6
59	0	0	0	0	0	0	0	1	1100	0	0	0	0	2	1	0	0	8
60	0	0	0	1	2016	0	0	0	0	0	0	0	1	0	0	0	0	7
61	0	0	0	1	2100	0	0	0	0	0	1	0	0	0	0	0	1	10
62	0	0	0	0	0	2	4845	0	0	0	0	0	1	0	0	0	1	9
63	0	0	0	1	3200	0	0	0	0	1	0	0	0	0	0	0	0	6
64	0	0	0	2	3220	0	0	0	0	0	0	0	1	0	0	0	1	10
65	0	0	0	5	3570	0	0	0	0	0	0	0	0	2	0.2	0	1	6
66	0	0	0	1	2600	0	0	0	0	0	0	0	0	2	1	0	1	6
67	0	0	0	1	3900	0	0	0	0	0	0	0	0	2	0.4	0	1	6
68	0	0	0	0	0	4	1210	0	0	1	0	4	0	0	0	0	0	8
69	0	0	0	3	1920	1	1700	0	0	1	1	0	0	0	0	0	0	10
70	0	0	0	0	0	0	0	1	1900	0	0	0	0	0	0	0	1	8
71	0	0	0	0	0	0	0	1	1125	0	0	0	0	0	0	0	1	6
72	0	0	0	0	0	0	0	1	1400	0	0	0	0	0	0	0	1	6
73	0	0	0	0	0	0	0	1	1140	0	0	0	0	0	0	0	1	6
74	0	0	0	0	0	0	0	1	1400	0	0	0	0	0	0	0	0	9
75	0	0	0	2	1399	0	0	0	0	0	0	0	0	2	1	0	0	8
76	0	0	0	1	4900	0	0	0	0	0	0	0	1	0	0	0	0	6
77	0	0	0	1	4900	0	0	0	0	0	0	0	1	0	0	0	0	6
78	0	0	0	1	11009	0	0	0	0	0	0	0	0	0	0	0	0	10
79	0	0	0	1	12063	0	0	0	0	1	0	0	0	1	0.4	0	0	12
80	0	0	0	0	0	0	0	1	2430	0	0	0	0	0	0	0	2	12
81	0	0	0	5	2577	0	0	0	0	0	1	0	0	0	0	0	0	12
82	0	0	0	0	0	3	2140	0	0	1	1	0	0	0	0	0	0	12
83	0	0	0	3	2500	0	0	0	0	0	1	0	1	0	0	0	1	12
84	0	0	0	1	7000	0	0	0	0	1	0	0	0	0	0	0	2	12
85	0	0	0	0	0	1	4400	0	0	1	1	0	0	0	0	0	0	12
86	0	0	0	0	0	2	1306	0	0	0	1	0	0	0	0	0	0	12
87	0	0	0	1	15200	0	0	0	0	0	0	1	0	0	0	0	1	12
88	0	0	0	2	5200	0	0	0	0	1	1	0	0	0	0	0	0	12
89	0	0	0	1	10184	0	0	0	0	1	0	0	1	0	0	0	0	12
90	0	0	0	2	3360	0	0	0	0	1	1	0	0	0	0	0	0	12
91	0	0	0	0	0	9	12148	0	0	1	1	4	0	0	0	1	0	12
92	0	0	0	0	0	2	2176	0	0	0	0	0	2	0	0	0	1	12
93	0	0	0	0	0	3	6449	0	0	1	0	0	0	0	0	0	0	12
94	0	0	0	2	7095	0	0	0	0	1	0	0	0	0	0	0	0	12
95	0	0	0	1	5130	0	0	0	0	0	0	0	0	0	0	0	0	12
96	0	0	0	6	100000	0	0	0	0	0	0	0	0	0	0	0	1	12
97	0	0	0	0	0	0	0	5	1300	0	1	0	1	0	0	0	0	12
98	0	1	0	2	2727.6	8	2068	0	0	1	1	0	0	0	0	0	0	12
99	0	0	0	0	0	0	0	1	1370	0	1	0	0	0	0	0	1	12
100	0	0	0	0	0	0	0	2	3607	0	0	0	1	0	0	0	1	12
101	0	0	0	1	1315.2	0	0	0	0	0	1	0	1	0	0	0	1	12
102	0	0	0	0	0	0	0	1	1440	0	1	0	0	0	0	0	0	12
103	0	0	0	0	0	0	0	1	1330	0	1	0	0	0	0	0	0	12
104	0	0	0	0	0	0	0	4	7070	0	0	0	0	0	0	0	1	12
105	0	0	0	0	0	0	0	1	1710	0	1	0	0	0	0	0	0	15
106	1	0	0	0	0	9	37502	0	0	0	0	0	1	0	0	0	0	15
107	0	0	0	0	0	0	0	4	2400	0	1	0	0	0	0	0	2	15
108	0	0	0	0	0	0	0	4	2160	1	0	0	0	0	0	0	0	15
109	0	0	0	0	0	3	4840	0	0	0	0	0	1	0	0	0	0	15
110	0	0	0	0	0	3	4840	0	0	1	0	0	1	0	0	0	0	12
111	0	0	0	1	9700	0	0	0	0	0	0	0	0	0	0	0	0	15
112	0	0	0	0	0	4	4358.2	0	0	1	0	0	0	0	0	0	0	15
113	0	0	0	1	7000	0	0	0	0	0	0	0	0	0	0	0	0	15
114	0	0	0	0	0	5	7650	0	0	1	0	0	1	0	0	0	0	15
115	0	0	0	1	9170	0	0	0	0	0	0	0	0	0	0	0	0	15

290	0	0	0	1	95500	0	0	0	0	0	0	0	0	2	1	0	0	108
291	0	0	0	1	116136	0	0	0	0	1	0	0	1	1	1	0	0	86
292	0	0	0	3	235505	0	0	0	0	1	0	0	0	0	0	0	0	108
293	0	0	0	0	0	23	54939	0	0	1	1	0	0	0	0	0	0	108
294	0	0	0	4	88215	0	0	0	0	1	0	0	1	0	0	0	0	84
295	0	0	0	14	54400	0	0	0	0	0	1	0	1	0	0	0	0	84
296	0	0	0	0	0	8	48000	0	0	1	0	0	0	0	0	0	0	90
297	0	0	0	0	0	15	76000	0	0	1	1	0	0	0	0	0	0	104
298	0	0	0	0	0	7	36500	0	0	0	1	4	1	1	0.6	0	1	90
299	0	0	0	0	0	7	24500	0	0	0	0	0	1	0	0	0	1	96
300	0	0	0	0	0	20	48270	0	0	0	0	0	0	2	0.6	0	1	96
301	0	0	0	15	97200	0	0	0	0	0	0	0	1	0	0	0	0	96
302	0	0	0	0	0	2	121400	0	0	0	1	0	1	0	0	0	0	96
303	0	0	0	3	146909	0	0	0	0	0	0	0	1	0	0	0	0	132
304	0	0	0	2	109691	0	0	0	0	0	0	0	0	1	0	0	1	120
305	0	0	0	0	0	23	114000	0	0	1	1	0	0	2	0.3	0	0	138
306	0	0	0	6	118000	1	44500	0	0	1	0	0	0	2	0.5	0	0	132
307	0	0	0	3	157705	0	0	0	0	0	0	0	1	2	1	0	1	120
308	0	0	0	6	220740	0	0	0	0	0	1	4	2	0	0	0	1	132
309	0	0	0	6	220740	0	0	0	0	0	0	0	2	0	0	0	0	144
310	0	0	0	33	370000	0	0	0	0	0	0	4	0	0	0	0	0	138
311	0	0	0	0	0	11	217605	0	0	0	0	0	1	2	0.3	0	1	126
312	0	0	0	2	218211	0	0	0	0	1	0	0	1	0	0	0	0	120
313	0	0	0	58	350000	0	0	0	0	0	0	0	2	0	0	0	0	144
314	0	0	0	0	0	77	96000	0	0	0	0	4	0	0	0	0	2	132
315	0	0	0	0	0	46	73000	0	0	0	1	0	2	0	0	0	0	132
316	1	0	0	1	229400	0	0	0	0	0	0	0	0	1	1	0	1	120
317	0	0	0	3	127000	0	0	0	0	0	0	4	2	0	0	0	1	120
318	0	0	0	2	102000	0	0	0	0	1	0	4	3	0	0	0	0	96
319	0	0	0	10	189000	0	0	0	0	0	0	0	2	1	0.5	1	0	132
320	0	0	0	12	180000	0	0	0	0	0	0	4	2	0	0	0	0	132
321	0	0	0	16	340000	0	0	0	0	1	0	4	2	0	0	1	0	132
322	0	0	0	11	288000	0	0	0	0	1	1	4	2	0	0	0	0	132
323	0	0	0	1	345000	0	0	0	0	0	0	0	0	2	1	0	0	132
324	0	0	0	0	0	40	110000	0	0	0	0	4	2	0	0	0	1	132
325	0	0	0	0	0	1	102300	0	0	0	1	0	0	0	0	0	0	138
326	0	0	0	2	12200	0	0	0	0	0	1	0	0	0	0	0	0	138
327	0	0	0	1	116136	0	0	0	0	0	0	0	1	2	1	0	0	120
328	0	0	0	0	0	57	116624	0	0	0	0	0	1	0	0	1	1	144
329	0	0	0	5	399082	0	0	0	0	0	0	0	1	2	0.8	0	1	174
330	0	0	0	4	393802	0	0	0	0	1	0	0	1	2	1	0	1	162
331	0	0	0	2	321255	0	0	0	0	0	0	0	1	2	0.5	0	1	144
332	0	0	0	5	238000	0	0	0	0	0	0	0	0	0	0	0	1	156
333	0	0	0	0	0	22	195058	0	0	0	0	0	0	2	0	0	1	144
334	0	0	0	27	815476	0	0	0	0	0	1	0	1	2	0.3	0	1	168
335	0	0	0	13	791373	0	0	0	0	0	0	0	1	2	0.2	0	1	144
336	0	0	0	0	0	74	160898	0	0	0	0	0	1	0	0	0	1	168
337	0	0	0	0	0	28	123769	0	0	1	0	0	1	0	0	0	1	168
338	0	0	0	16	72750	0	0	0	0	0	1	0	0	0	0	0	1	144
339	0	0	0	7	18160	21	105923	0	0	0	1	0	0	2	0.5	0	1	156
340	0	0	0	0	0	67	130000	0	0	1	0	0	0	0	0	1	0	156
341	0	0	0	2	171290	0	0	0	0	0	0	0	0	0	0	0	1	144
342	0	0	0	0	0	29	87000	0	0	0	1	0	0	2	0.2	0	1	156
343	0	0	0	0	0	27	90674	0	0	0	0	0	1	2	0	0	1	150

Appendix B: The Chinese version of two verdicts mentioned in Chapter 6.2



刑 事 判 决 书

（2013）×刑初字第×××号

公诉机关××市××区人民检察院。

被告人马××，男，1980 年 4 月 19 日出生于××省×县，汉族，初中文化，农民，住××省×县××镇×村×组。因涉嫌犯盗窃罪于 2013 年 8 月 14 日被逮捕。现羁押于××市××区看守所。

××市××区人民检察院以××检刑诉（2013）428 号起诉书指控被告人马××犯盗窃罪，于 2003 年 7 月 28 日向本院提起公诉，并建议适用简易程序审理。2013 年 8 月 15 日，本院决定适用普通程序审理，并依法组成合议庭。公开开庭审理了本案。××市××区人民检察院代理检察员张×、被告人马××到庭参加诉讼。现已审理终结。

经审理查明：

1、2012 年 2 月 15 日凌晨，被告人马××至××市××区××镇××村某公司，采用扳断窗栅后入室等方法，窃得该公司价值人民币 10,000 元的工业氧气瓶 5 只。

2、2012 年 8 月中旬某日凌晨，被告人马××至本区××镇××村××队×号，窃得失主朱××价值人民币 7,600 元的山羊牌 50CC 轻便摩托车一辆。

3、2012 年 9 月 1 日凌晨，被告人马××至本区××镇××村××巷×号，采用卸窗玻璃、扳断窗栅后入室等方法，窃得失主施××价值人民币 12,800 元的新大洲牌 SDH50QT-16 型轻便摩托车及价值人民币 640 元的神州牌 SZ100-3 型摩托车各一辆。

4、2012 年 10 月 10 日凌晨，被告人马××至××市××区××村××街×号××镇养老院×室，窃得该院价值人民币 1,000 元的创维牌 21 寸彩电及现金人民币 7,000 元。

上述事实，被告人马××在开庭审理过程中亦无异议，且有失主施××、朱××、王×、毛××的陈述，××市价格认证中心××分部物品财产估价鉴定结论书，公安机关现场勘查笔录、有关照片，公安机关手印鉴定书、工情况记录等证据所证实，足以认定。

本院认为，被告人马××以非法占有为目的，采用秘密手段，窃取公私财物，数额较大，其行为已触犯《中华人民共和国刑法》第二百六十四条之规定，构成盗窃罪。公诉机关指控成立，本院予以确认。被告人马××关于其有自首情节的辩解，经查，在马××向公安机关供述盗窃事实之前，公安机关已经由指纹比对掌握其实施了指控的第三节盗窃事实，故马××的行为属坦白而非自首。对此辩解本院不予采纳，然对被告人坦白交代其他同种犯罪事实的行为，本院在量刑时予以酌情从轻处罚。据此，本院为严肃国家法制，维护公私财产权利不受侵犯，依照《中华人民共和国刑法》第二百六十四条、第五十三条、第六十四条之规定，判决如下：

一、被告人马××犯盗窃罪，判处有期徒刑一年零六个月，并处罚金人民币五千元。

（刑期从判决执行之日起计算。判决执行以前先行羁押的，羁押一日折抵刑期一日，即自 2013 年 8 月 14 日起至 2014 年 10 月 13 日止。罚金于本判决生效之日起三十日内如数缴纳。）

二、被告人马××的违法所得予以追缴。

如不服本判决，可在接到判决书的第二日起十日内，通过本院或者直接向××市第一中级人民法院提出上诉。书面上诉的，应当提交上诉状正本一份、副本两份。通过本院书面上诉的，应当将上诉状正本、副本送（寄）往本院立案庭。

审 判 长 ×××

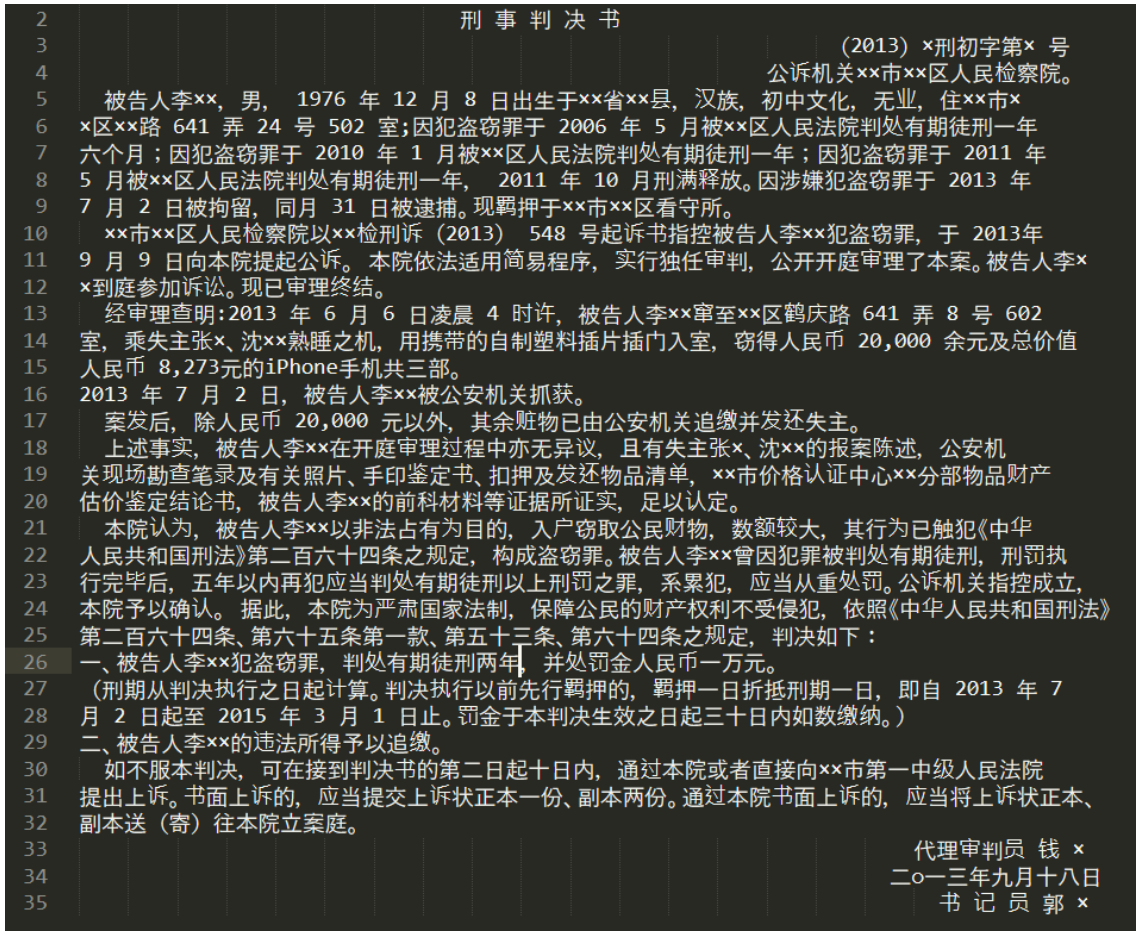
人民陪审员 ×××

代理审判员 × ×

二o一三年九 月 十六日

书 记 员 × ×

Figure B1 The Chinese version of verdict one mentioned in Chapter 6.2



刑 事 判 决 书

(2013) ×刑初字第× 号

公诉机关××市××区人民检察院。

被告人李××，男， 1976 年 12 月 8 日出生于××省××县，汉族，初中文化，无业，住××市××区××路 641 弄 24 号 502 室；因犯盗窃罪于 2006 年 5 月被××区人民法院判处有期徒刑一年六个月；因犯盗窃罪于 2010 年 1 月被××区人民法院判处有期徒刑一年；因犯盗窃罪于 2011 年 5 月被××区人民法院判处有期徒刑一年， 2011 年 10 月刑满释放。因涉嫌犯盗窃罪于 2013 年 7 月 2 日被拘留，同月 31 日被逮捕。现羁押于××市××区看守所。

××市××区人民检察院以××检刑诉（2013） 548 号起诉书指控被告人李××犯盗窃罪，于 2013年 9 月 9 日向本院提起公诉。本院依法适用简易程序，实行独任审判，公开开庭审理了本案。被告人李××到庭参加诉讼。现已审理终结。

经审理查明:2013 年 6 月 6 日凌晨 4 时许，被告人李××窜至××区鹤庆路 641 弄 8 号 602 室，乘失主张×、沈××熟睡之机，用携带的自制塑料插片插门入室，窃得人民币 20,000 余元及总价值人民币 8,273元的iPhone手机共三部。

2013 年 7 月 2 日，被告人李××被公安机关抓获。

案发后，除人民币 20,000 元以外，其余赃物已由公安机关追缴并发还失主。

上述事实，被告人李××在开庭审理过程中亦无异议，且有失主张×、沈××的报案陈述，公安机关现场勘查笔录及有关照片、手印鉴定书、扣押及发还物品清单，××市价格认证中心××分部物品财产估价鉴定结论书，被告人李××的前科材料等证据所证实，足以认定。

本院认为，被告人李××以非法占有为目的，入户窃取公民财物，数额较大，其行为已触犯《中华人民共和国刑法》第二百六十四条之规定，构成盗窃罪。被告人李××曾因犯罪被判处有期徒刑，刑罚执行完毕后，五年以内再犯应当判处有期徒刑以上刑罚之罪，系累犯，应当从重处罚。公诉机关指控成立，本院予以确认。据此，本院为严肃国家法制，保障公民的财产权利不受侵犯，依照《中华人民共和国刑法》第二百六十四条、第六十五条第一款、第五十三条、第六十四条之规定，判决如下：

一、被告人李××犯盗窃罪，判处有期徒刑两年，并处罚金人民币一万元。

（刑期从判决执行之日起计算。判决执行以前先行羁押的，羁押一日折抵刑期一日，即自 2013 年 7 月 2 日起至 2015 年 3 月 1 日止。罚金于本判决生效之日起三十日内如数缴纳。）

二、被告人李××的违法所得予以追缴。

如不服本判决，可在接到判决书的第二日起十日内，通过本院或者直接向××市第一中级人民法院提出上诉。书面上诉的，应当提交上诉状正本一份、副本两份。通过本院书面上诉的，应当将上诉状正本、副本送（寄）往本院立案庭。

代理审判员 钱 ×

二o一三年九月十八日

书 记 员 郭 ×

Figure B2 The Chinese version of verdict two mentioned in Chapter 6.2