Crime Prediction Using Data Mining and Machine Learning

Shaobing Wu[1], Changmei Wang[2(✉)], Haoshun Cao[1], and Xueming Jia[1]

[1] Institute of Information Security, Yunnan Police College, Kunming 650223, China
[2] Solar Energy Institute, Yunnan Normal University, Kunming 650092, China
823804919@qq.com

Abstract. In order to predict the crime in YD county, data mining and machine learning are used in this paper. The aim of the study is to show the pattern and rate of crime in YD county based on the data collected and to show the relationships that exist among the various crime types and crime Variable. Analyzing this data set can provide insight on crime activities within YD county. By introducing formula and methods of Bayesian network, random tree and neural network in machine learning and big data, to analyze the crime rules from the collected data. According to the statistics released by the YD county From 2012-09-01 to 2015-07-21, The crime of smuggling, selling, transporting and manufacturing drugs, Theft, Intentional injury, Illegal business crime, Illegal possession of drugs, Rape, Crime of fraud, Gang fighting, manslaughter, Robbery made the top ten list of crime types with high number of crimes. The crime rate of drugs was the highest, reaching 46.86%, farmers are the majority, accounting for 97.07%, people under the age of 35 are the subject of crime. Males accounted for 90.17% of crimes committed, while females accounted for 9.83%. For ethnic groups, the top five were han, yi, wa, dai and lang, accounting for 68.43%, 23.43%, 1.88%, 1.67% and 1.25% respectively. By adopting random forest, Bayesian networks, and neural network methods, we obtained the decision rules for criminal variables. By comparison, the classification effect of Random Trees is better than that of Neural Networks and Bayesian Networks. Through the data collection of the three algorithms, the validity and accuracy of the random tree algorithm in predicting crime data are observed. The performance of the Bayesian network algorithm is relatively poor, probably due to the existence of certain random factors in various crimes and related features (the correlation between the three algorithms is low).

Keywords: Crime prediction · Data mining · Machine learning

1 Introduction

For almost everyone, machine learning (ML) is still a very mysterious field that sounds complicated and difficult to explain to a person without any skills [1]. However, this is very important today and will continue in the next few years.

Q. Liu et al. (Eds.): CENet 2018, AISC 905, pp. 360–375, 2020.
https://doi.org/10.1007/978-3-030-14680-1_40

ML is a fairly multidisciplinary field that deals primarily with programming and mathematics (mainly involving probability and density functions). In addition, because it is new and quite complex, it requires good research skills.

For crime detection issues, the game organizer provided a huge database of crime training in San Francisco. The database is tagged, that is, it contains the correct category for each entry (e.g. theft, assault, bribery, etc.), so it is a supervised learning problem. With this in mind, the algorithms used to solve this problem are: Random tree, neural network, Bayesian network.

The US Federal Bureau of Investigation (FBI) defines violent crime as a crime involving violence or threats. The United States Federal Bureau of Investigation (FBI) Unified Crime Report (UCR) program defines each type of criminal behavior as: (i) Murder-intentional (non-faulty) murder. UCR does not include deaths caused by accidents, suicides, negligence, proper homicides, and attempts to murder or assault murder (all of which are classified as serious attacks), in this crime classification [2]. (ii) Forced rape-rape is a sexual assault that violates the will of women. While attempting or attacking rape by threat or force is considered a crime under this category, statutory rape (without the use of force) and other sexual offences are excluded from [3]. (iii) Robbery-threatening or violent by force or force and/or placing the victim in fear, gaining or attempting to obtain anything of value from the care, custody or control of one or more persons. Crimes that aggravate the crime of personal assault and theft are crimes of robbery. Unfortunately, these types of crimes seem to have become commonplace in society. Law enforcement officials have turned to data mining and machine learning to help fight crime prevention and enforcement.

Miquel Vaquero Barnadas [13] proposed machine learning applied to crime prediction. In this paper, he plans to use different algorithms (such as K-Nearest neighbour, Parzen windows and Neural Networks) to solve the real data classification problem (the San Francisco crime classification).

Gaetano Bruno Ronsivalle [14] presented Neural and Bayesian Networks to Fight Crime: The NBNC Meta-Model of Risk Analysis. In his paper, he used this tool with the specific goal of providing an effective model for Italian bank security managers to "describe" variables and define "robbing" phenomena; "interpret" calculations (i) "exogenous", (ii) "Endogenous and (iii) methods of global risk index for each branch; through simulation modules to "predict" composite risk and changes in different branch security systems.

Jeffrey T. Ward, James V. Ray, Kathleen A. Fox [15] developed a MIMIC model for Exploring differences in self-control across sex, race, age, education, and language, and draw a conclusion that apart from race, testing group differences in self-control with an observed scale score is largely unbiased. Testing group differences in elements using observed subscores is frequently biased and generally unsupported.

In this research, we developed the Random Trees, Neural Networks, and Bayesian Networks algorithms using the same finite set of features, on the communities and crime un normalized dataset to conduct a comparative study between the violent crime patterns from this particular dataset and actual crime statistical data for the state of YD County. The crime statistics used from collected data. Some of the statistical data that was provided by YD County people's procuratorate such as the population of YD

County, population distribution by age, number of violent crimes committed, and the rate of those crimes are also features that have been incorporated into the test data to conduct analysis.

The rest of the paper is organized as follows: Sect. 2 gives an overview of data mining and machine learning. Section 3 provides information about the Crime Classification in YD County. Section 4 presents the results from each of the algorithms and Sect. 5 concludes with the findings and discussion of the paper results.

2 Data Mining and Machine Learning Algorithms

2.1 Data Mining

Data mining is part of the interdisciplinary field of knowledge discovery in databases [9]. Data mining consist of collecting raw data and, (through the processes of inference and analysis); creating information that can be used to make accurate predictions and applied to real world situations. It is the application of techniques that are used to conduct productive analytics. The five tasks that these types of software packages are designed for are as follows: (i) Association-Identifying correlations among data and establishing relationships between data that exist together in a given record [9, 10]. (ii) Classification Discovering and sorting data into groups based on similarities of data [6]. Classification is one of the most common applications of data mining. The goal is to build a model to predict future outcomes through classification of database records into a number of predefined classes based on a certain criteria. Some common tools used for classification analysis include neural networks, decisions trees, and if-then-else rules [10]. (iii) Clustering-Finding and visually presenting groups of facts previously unknown or left unnoticed [6]. Heterogeneous data is segmented into a number of homogenous clusters. Common tools used for clustering include neural networks and survival analysis [10]. (iv) Forecasting-Discovering patterns and data that may lead to reasonable predictions [9].

2.2 Machine Learning

Arthur Samuel is a pioneer in the field of machine learning and artificial intelligence. He defines machine learning as a field of study that allows computers to learn without explicit programming [11]. In essence, machine learning is a way for computer systems to learn through examples. There are many machine learning algorithms available to users that can be implemented on data sets. The algorithm has a better understanding of the data set because it has more examples to implement. In the field of data mining, there are five machine learning algorithms for analysis: (i) Classification analysis algorithms-these algorithms use attributes in the data set to predict the value of one or more variables taking discrete values. (ii) Regression analysis algorithms-These algorithms use the properties of the data set to predict the value (e.g. profit and loss) of one or more variables taking continuous values. (3) Segmentation analysis algorithm - divide data into groups or groups with similar attributes.

2.3 Algorithms Selected for Analysis

Random Trees-Aldous [12, 13] discussed scaling limits of various classes of discrete trees conditioned to be large. In the case of a Galton-Watson tree with a finite variance critical offspring distribution and conditioned to have a large number of vertices, he proved that the scaling limit is a continuous random tree called the Brownian CRT. Their main result (Theorem 2.1) stated that the rescaled height function associated with a forest of independent (critical, finite variance) Galton-Watson trees converged in distribution towards reflected Brownian motion on the positive half-line.

In order to derive the Theorem 2.1, they first state a very simple "moderate deviations" Lemma 1.1 for sums of independent random variables.

Lemma 1.1: Let Y_1, Y_2, … be a sequence of i.i.d. real random variables. We assume that there exists a number $\lambda > 0$ such that E $[\exp(\lambda|Y_1|)] < \infty$, and that $E[Y_1] = 0$. Then, for every $\alpha > 0$, we can choose N sufficiently large so that for every $n \geq N$ and $l \in \{1, 2, 3, \ldots, n\}$

$$\mathrm{P}\left[|Y_1 + \cdots Y_l| > n^{\alpha + \frac{1}{2}}\right] \ll e^{-n^{\alpha/2}} \tag{2.1}$$

According to Lemma 1.1, they get the Theorem 2.1 as following:

Theorem 2.1: Let θ_1, θ_2, … be a sequence of independent μ-Galton-Watson trees, and let $(H_n; n \geq 0)$ be the associated height process. Then

$$\frac{1}{\sqrt{p}} H_{[pt]}, t \gg 0 \underset{p\to\infty}{\rightarrow} \left(\frac{2}{\sigma}\gamma_t, t \gg 0\right) \tag{2.2}$$

Where γ is a reflected Brownian motion. The convergence holds in the sense of weak convergence on the Skorokhod space D(R+; R+).

In their papers, they introduce the exit measure from a domain D, which is in a sense uniformly spread over the set of exit points of the Brownian snake paths from D. they then derive the key integral equation (Theorem 2.2) for the Laplace functional of the exit measure. In the particular case when the underlying spatial motion ε is d-dimensional Brownian motion, this quickly leads to the connection between the Brownian snake and the semilinear PDE(Partial differential equation) $\Delta u = u^2$.

Theorem 2.2: Let g be a nonnegative bounded measurable function on E. For every $x \in$ E, set

$$u(x) = \mathrm{N}_x\left(1 - exp - Z^D, g\right), x \in D \tag{2.3}$$

The function u solves the integral equation [20]

$$u(x) + 2\prod\nolimits_x \left(\int_0^T u\varepsilon_s^2 ds\right) = \prod\nolimits_x \left(1_{\{\tau<\infty\}} g(\varepsilon_\tau)\right) \tag{2.4}$$

Random continuous trees can be used to model the genealogy of self-similar fragmentations [14]. The Brownian snake has turned out to be a powerful tool in the study of super-Brownian motion: See the monograph [15] and the references therein.

Since Random Trees was proposed, the algorithm has become a popular and widely used tool for nonparametric regression applications.

***Bayesian Networks*-**A Bayesian network (BN) approximates the joint probability distribution for a multivariate system based on expert knowledge and sampled observations that are assimilated through training [16, 17]. A tractable scoring metric, known as K2, is obtained from $\mathrm{P}(F, T)$ using the assumptions in [6], which include fixed ordering of variables in X:

$$\mathrm{g} = \log\left(\prod\nolimits_{j=1,\ldots,q_j} \frac{(r_i - 1)!}{\left(\overline{N_{ij}} + r_i - 1\right)!} \prod\nolimits_{k=1,\ldots r_i} N_{ijk}!\right) \tag{2.5}$$

where $\mathrm{r_i}$ is the number of possible instantiations of $\mathrm{X_i}$, and $\mathrm{q_i}$ is the number of unique instantiations of pi. N_{ijk} is the number of cases in T. The BN (K, H) represents a factorization of the joint probability over a discrete sample space,

$$\mathrm{p(X)} = p(X_1, \ldots, X_n) = \prod\nolimits_{i=1,\ldots,n} p(X_i|\pi_i) \tag{2.6}$$

for which all probabilities on the right-hand side are given by the CPTs. Therefore, when a variable $\mathrm{X_i}$ is unknown or hidden, Bayes' rule of inference can be used to calculate the posterior probability distribution of Xi given evidence of the set of l variables, that are conditionally dependent on $\mathrm{X_i}$,

$$\mathrm{P}(X_i|\overline{\mu_i}) = \frac{\mathrm{P}(\overline{\mu_i}|X_i)\mathrm{P}(X_i)}{\mathrm{P}(\overline{\mu_i})} \tag{2.7}$$

Bayesian networks are particularly well suited for crime analysis, as they learn from data and use the experience of criminologists to select nodes and node sequencing. The confidence level provided for the criminal files informs the detective about the possible accuracy of each prediction. In addition, BN's graphical structure represents the most important relationship between criminal behavior and crime scene behavior, which may help develop new scientific assumptions about criminal behavior.

***Neural Networks*-**Artificial Neural Networks (ANN) have been developed as generalizations of mathematical models of biological nervous systems. The basic processing elements of neural network are called artificial neurons, or simply neurons or nodes. The neuron pulse is then calculated as the weighted sum of the input signal of the transfer function transformation. The artificial neurons' learning ability can be realized by adjusting the weight according to the selected learning algorithm [12].

Architectures: An ANN consists of a set of processing elements, also known as neurons or nodes, which are interconnected. It can be described as a directed graph in which each node performs a transfer function of the form

$$y_i = f\left(\sum\nolimits_{i=1,\ldots,n} w_{ij}x_j - \theta_i\right)_i \tag{2.8}$$

where y_i is the output of the node i, x_j is the th input to the node, and w_{ij} is the connection weight between nodes i and j. θ_i is the threshold (or bias) of the node. Usually, f_i is nonlinear, such as a heaviside, sigmoid, or Gaussian function.

In (11), each term in the summation only involves one input x_j. High-order ANN's are those that contain high-order nodes, i.e. nodes in which more than one input are involved in some of the terms of the summation. For example, a second-order node can be described as

$$y_i = f_i\left(\sum\nolimits_{j,k=1,\ldots,n} w_{ijk}x_jx_k - \theta_i\right) \tag{2.9}$$

where all the symbols have similar definitions to those in (11).

3 Crime Classification in YD County

As it has been said previously, this project is based on a Project of national social science foundation about the causes and countermeasures of ethnic minority crimes in the county of YD. In this chapter, the principle and formula of Bayesian network, random tree and neural network are given briefly.

3.1 Description of the Problem

In this problem, a training dataset with nearly 35 months of crime reports from all across YD county was provided. This dataset contains all crimes classified in categories, which are the different crime typologies. The main goal of the challenge is to predict these categories of the crimes that occurred.

For the algorithm evaluation, another unlabelled dataset is provided (the test one). It is used to evaluate the algorithm accuracy with new unclassified data.

How Is the Problem Going to be Solved?
In this contest, we will use a different algorithm to get a good result. Each one will be explained, tested and tested, and finally we will see which of them is best for this case.

Cross-validation will be used to validate the model, so the database must be divided into subsets of tests, training, and verification. This division must be layered to ensure that the proportion of the original components is maintained in each segment (the number of crimes per category is the same).

All development and testing is done on a server provided by the university department. In this way, the death penalty can last a whole day without worrying about them, and execution is faster.

Results Submission Format and Evaluation
The data submitted to the contest evaluation must be in a specific format that meets the requirements. To properly evaluate the data, the resulting data set must contain sample

ids that contain a list of all categories and the probability that each sample belongs to each category. Remind the training dataset to label the crime types of all samples (10 different).

Then, instead of predicting which category a given sample belongs to, the output will always be the probability vector.

3.2 Dataset Analysis

Data

The data in this article involves the reported cases of the crime of smuggling, selling, transporting and manufacturing drugs, Theft, Intentional injury, Illegal business crime, Illegal possession of drugs, Rape, Crime of fraud, Gang fighting, manslaughter, Robbery in YD county between the years 2012 and 2015. The summary of the data is as provided in Table 1.

Table 1. Summary statistics of the data set on crime activities.

Crime types	Numbers of crime
The crime of smuggling, selling, transporting and manufacturing drugs	224
Theft	86
Intentional injury	85
Illegal business crime	23
Illegal possession of drugs	18
Rape	18
Crime of fraud	13
Gang fighting	11

The data provides insight into criminal activity, and its research can help reduce crime (protecting communities) and decision-making. Part of the analysis provided can be used to explain the relationship between certain criminal activities.

The data can further be analyzed using other statistical methods like Random Trees, Bayesian, Quasi-neural networks and so on.

From Table 1 and Fig. 3, The crime of smuggling, selling, transporting and manufacturing drugs is the most important crime types, the following is theft, intentional injury and so on.

Data Analysis

The provided dataset has different “features”, each one being of a different relevance. In this chapter we will proceed to analyze this database and extract the useful information out of it. There are 478 samples of crime analysis. These data were collected from 2012-09-01 to 2015-07-21 in YD.

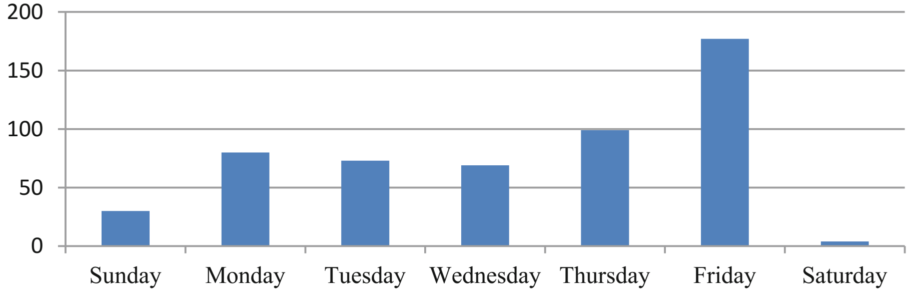


Fig. 1. Number of crimes by day of the week (YD County).

Another interesting analysis is to count the number of crimes that occur every day of the week so that we know if this is relevant information. Figure 1 shows that the day when the offenders choose the most is Friday, and the rest of the week is distributed differently.

We chose the top eight of the crime categories as the basis for this analysis and discussion.

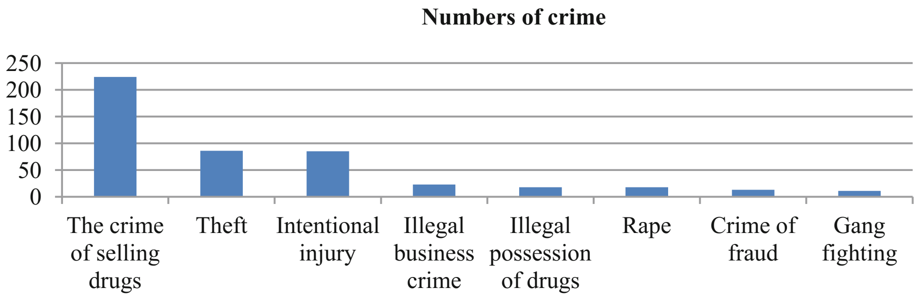


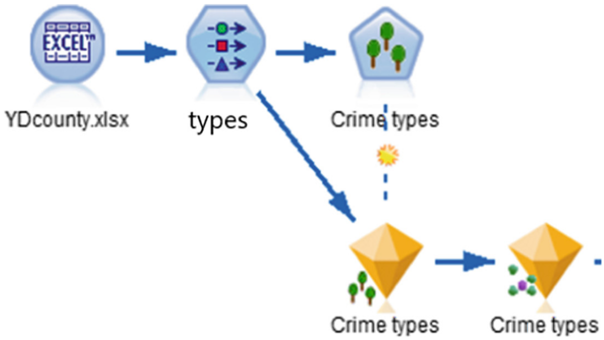
Fig. 2. Numbers of crime types.

From the Fig. 2, we can see that the crime of smuggling, selling, transporting and manufacturing drugs 224; Theft 86; Intentional injury 85; Illegal business crime 23, which are the main parts of crime types.

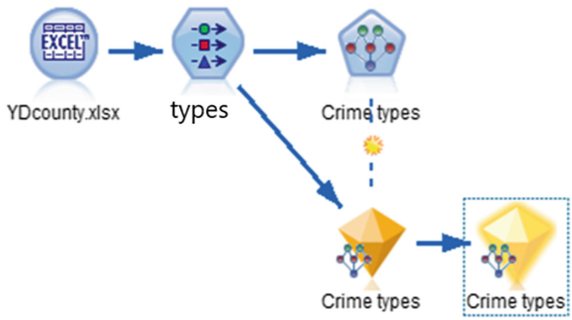
4 Results

The data set selected in this study is a denormalized data set for community and crime. Including the social and economic data from 2012-09-01 to 2015-07-21, the law enforcement data of the YD County People's Procuratorate, and the crime data from 2012-09-01 to 2015-07-21. It also includes 478 cases or criminal cases and 12 total attributes reported from the entire YD County, often referred to as features.

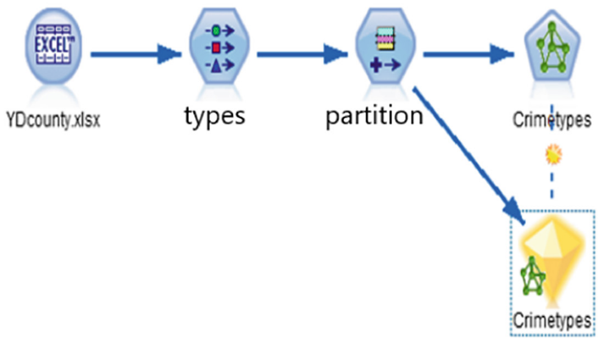
This section describes all the implementation results of the random tree, Bayesian algorithm, and neural network algorithm. The algorithm is run to predict the following characteristics of each data set: he smuggles crimes, sells, transports and manufactures drugs, illegally holds drugs, rapes, steals, murders, robberies, intentional injuries, illegal business crimes, fraud, criminal gang fights.



(a) Random Trees



(b) Bayesian Networks



(C) Neural Networks

Fig. 3. The modeling of machine learning algorithm.

4.1 Modeling Based on Machine Learning Algorithm

In this section, we build three models based on Random Trees, Bayesian, Quasi-neural networks algorithms. Figure 3 shows the modeling of machine learning algorithms.

4.2 Relation Between Sex (or Gender), Ethnic, Age, Education and Crime

The relationship between sex and crime as following: first, the number relationship is as Table 2 and Fig. 4 shows the relationship between sex and crime. From Fig. 4 and Table 2, number of the males more than females.

Table 2. Convictions according to gender for YD.

Year	Males	Females
2012	6	0
2013	169	20
2014	221	16
2015	89	11

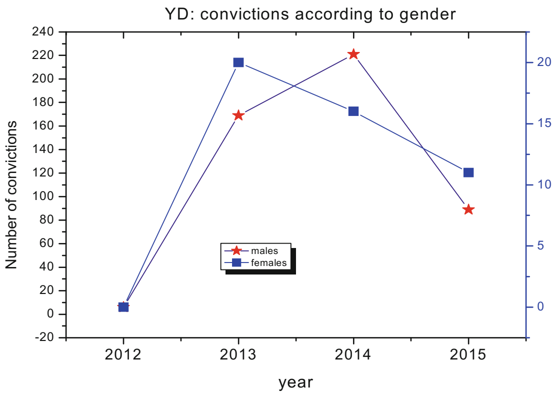


Fig. 4. Relationship between sex and crime.

In the 478 samples, the relationship between age and Crime numbers for drugs See Fig. 5. From Fig. 5, it shows that people for the age from 16 to 35, accounting for about half of the total sample, are the main criminal group.

The relationship between education and Crime numbers for drugs See Fig. 6. From Fig. 6, it shows that people for illiteracy or semi-illiteracy, Primary school and Junior high school, are the main criminal group.

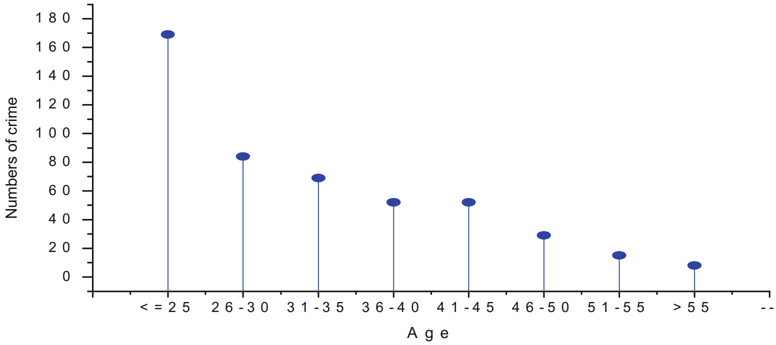


Fig. 5. The relationship between age and Crime numbers for crime types.

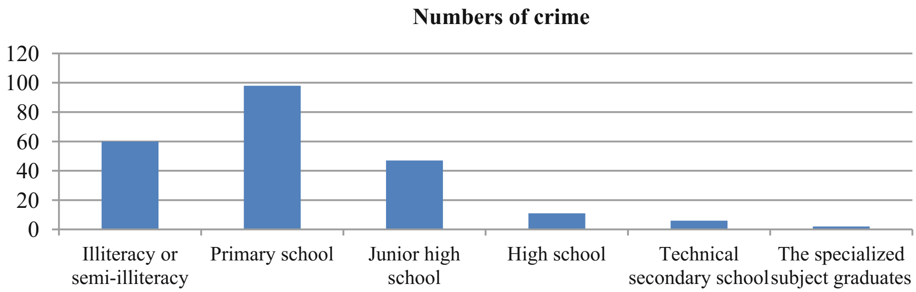


Fig. 6. The relationship between education and crime.

4.3 The Conditional Probability Table for Bayesian Networks

The Conditional probability based on for The crime of smuggling, selling, transporting and manufacturing drugs and Gang fighting for professional, education, gender, ethnic and age are followed as Tables 3, 4, 5, 6, 7 and 8.

Table 3. Conditional probability for crime types.

The crime of smuggling, selling, transporting and manufacturing drugs	Gang fighting
0.95	0.05

Table 4. Conditional probability for gender.

Crime types	Male	Female
The crime of smuggling, selling, transporting and manufacturing drugs	0.86	0.14
Gang fighting	1.00	0.00

Table 5. Conditional probability for professional.

Crime types	Gender	Individual worker	Migrant worker	Farmers	Foreigners	Student	Teacher
The crime of smuggling, selling, transporting and manufacturing drugs	Male	0.01	0.00	0.97	0.02	0.00	0.00
The crime of smuggling, selling, transporting and manufacturing drugs	Female	0.00	0.00	0.97	0.03	0.00	0.00
Gang fighting	Male	0.00	0.00	1.00	0.00	0.00	0.00

Table 6. Conditional probability for education.

Crime types	Gender	High school	Illiteracy or semi-illiteracy	Junior high school	Primary school	Technical secondary school	Specialized subject graduates
The crime of smuggling, selling, transporting and manufacturing drugs	Male	0.05	0.22	0.23	0.46	0.03	0.01
The crime of smuggling, selling, transporting and manufacturing drugs	Female	0.06	0.53	0.09	0.28	0.03	0.00
Gang fighting	Male	0.00	0.00	0.82	0.18	0.00	0.00

Table 7. Conditional probability for ethnic.

Crime types	Gender	bai	Blang	Buyi	dai	Deang	Han	Hui	Lisu	man	miao	Other	wa	yi
The crime of smuggling, selling, transporting and manufacturing drugs	Male	0.00	0.01	0.01	0.02	0.00	0.55	0.02	0.00	0.01	0.02	0.01	0.00	0.39
The crime of smuggling, selling, transporting and manufacturing drugs	Female	0.00	0.00	0.00	0.00	0.00	0.47	0.00	0.00	0.00	0.00	0.03	0.00	0.50
Gang fighting	Male	0.00	0.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Table 8. Conditional probability for age.

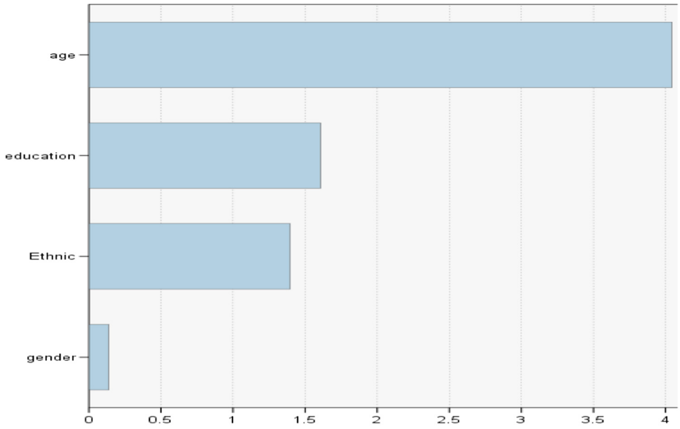
Crime types	Gender	<=24.6	24.6–35.2	35.2–45.8	45.8–56.4	>56.4
The crime of smuggling, selling, transporting and manufacturing drugs	Male	0.22	0.38	0.28	0.10	0.02
The crime of smuggling, selling, transporting and manufacturing drugs	Female	0.12	0.44	0.32	0.12	0.00
Gang fighting	Male	0.82	0.18	0.00	0.00	0.00

4.4 The Comparison of Result for Random Trees, Neural Networks and Bayesian Networks

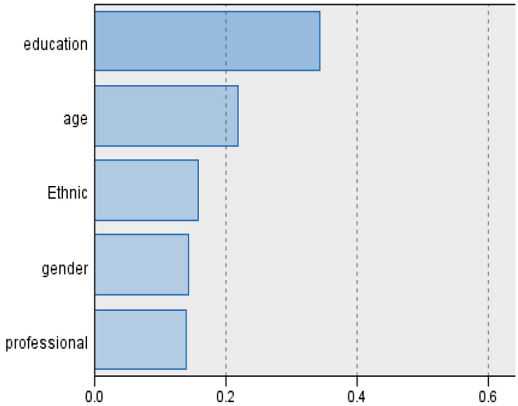
The Comparison of Data Analysis Results Predictive Variable Importance
The comparison of data analysis results predictive variable importance based on different algorithms is as following:
Predictive variable importance of crime types for Random Trees, Bayesian Network and Neural Network See Fig. 7. From Fig. 7, it shows that the age is important variable for Random Trees and Neural Networks, and the education is important variable for Bayesian Networks.

The Comparison of Model Accuracy
The data analysis results based on Random Trees and Bayesian Networks. The comparison of model accuracy based on Random Trees Classification and Bayesian Networks Classification are as follows:

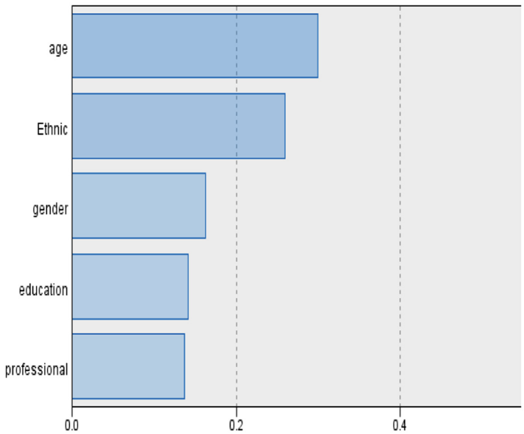
From Fig. 8, it shows that Model accuracy for Random Trees is much higher than Bayesian Networks.



(a) Random Trees



(b) Bayesian Network



(c) Neural Networks

Fig. 7. Predictive variable importance of crime types for Random Trees, Bayesian Network and Neural Networks.

Model information for Random Trees	
The target field	**Crime types**
Model building method	Random Trees Classification
The number of predictive variables entered	5
Model accuracy	0.974
Misclassification rate	0.026
Model information for Bayesian Networks	
The target field	**Crime types**
Model building method	Bayesian Networks Classification
The number of predictive variables entered	5
Model accuracy	0.537
Misclassification rate	0.463

Fig. 8. The comparison of model accuracy.

5 Conclusions and Future Development

In the field of artificial intelligence, machine learning is a very powerful field. If the model is done correctly, the accuracy that some algorithms can achieve can be surprising. Of course, the current and future of intelligent systems are subject to ML and big data analysis. From the above discussion, we can draw the following conclusions:

In the data we selected, the crime rate of was the highest, reaching 46.86%, which was the main crime type in YD county. In fact, it was theft and intentional injury, reaching 17.99% and 17.78% respectively.

For ethnic groups, the top five were han, yi, wa, dai and lang, accounting for 68.43%, 23.43%, 1.88%, 1.67% and 1.25% respectively. From this, the han nationality is the main criminal in the nation.

Through the data collection of the three algorithms, the validity and accuracy of the random tree algorithm in predicting crime data are observed. The performance of the Bayesian network algorithm is relatively poor, probably due to the existence of certain random factors in various crimes and related features (the correlation between the three algorithms is low).

Acknowledgements. This study is supported by scientific research projects of National Social Science Foundation (13CFX038). The authors would like to express their gratitude to the Office of the national social science foundation.

References

1. McClendon, L., Meghanathan, N.: Using machine learning algorithms to analyze crime data. Mach. Learn. Appl. Int. J. **2**(1), 1–2 (2015)
2. Murder. http://www.fbi.gov/ucr/cius2009/offenses/violent_crime/murder_homicide.html
3. Forcible Rape. http://www.fbi.gov/ucr/cius2009/offenses/violent_crime/forcible_rape.html
4. Robbery. http://www.fbi.gov/ucr/cius2009/offenses/violent_crime/robbery.htm

5. Assault. http://www.fbi.gov/ucr/cius2009/offenses/violent_crime/aggravated_assault.html
6. Vaquero Barnadas, M.: Machine learning applied to crime prediction. http://upcommons.upc.edu/bitstream/handle/2117/96580/machine%20learning%20applied%20TO%20crime%20precticion.pdf
7. Ronsivalle, G.B.: Neural and bayesian networks to fight crime: the NBNC meta-model of risk analysis. In: Artificial Neural Networks-Application, pp. 29–42 (2011)
8. Ward, J.T., Ray, J.V., Fox, K.A.: Developed a computer model exploring differences in self-control across sex, race, age, education, and language: considering a bifactor MIMIC model. J. Crim. Justice **56**, 29–42 (2018)
9. Nirkhi, S.M., Dharaskar, R.V., Thakre, V.M.: Data mining: a prospective approach for digital forensics. Int. J. Data Min. Knowl. Manag. Process **2**(6), 41–48 (2012)
10. Ngai, E.W.T., Xiu, L., Chau, D.C.K.: Application of data mining techniques in customer relationship management: a literature review and classification. Expert Syst. Appl. 2592–2602 (2008)
11. McCarthy, J.: Arthur samuel: pioneer in machine learning. AI Mag. **11**(3), 10–11 (1990)
12. Aldous, D.: The continuum random tree I. Ann. Probab. **19**(1), 1–28 (1991)
13. Aldous, D.: The continuum random tree III. Ann. Probab. **21**(1), 248–289 (1993)
14. Haas, B., Miermont, G.: The genealogy of self-similar fragmentations with negative index as a continuum random tree. Electron. J. Probab. **9**, 57–97 (2004)
15. Le Gall, J.F.: Spatial Branching Processes, Random Snakes and Partial Differential Equations. Birkhauser, Boston (1999)
16. Heckerman, D.: A Bayesian approach to learning causal networks. Technical report MSR-TR-95-04, pp. 1–23 (1995)
17. Jensen, F.V.: Bayesian Networks and Decision Graphs. Springer, New York (2001)