

Bachelor Thesis

Assessing the Suitability of the Waikato Environment for Knowledge Analysis for Crime Data Mining

Nina Spreitzer

Subject Area: Data Mining

Studienkennzahl: J003/561

Supervisor: Dr. Sabrina Kirrane

Date of Submission: 10.Nov 2020

Department of Information Systems and Operations, Vienna University of Economics and Business, Welthandelsplatz 1, 1020 Vienna, Austria



DEPARTMENT FÜR INFORMATIONS-
VERARBEITUNG UND PROZESS-
MANAGEMENT DEPARTMENT
OF INFORMATION SYSTEMS AND
OPERATIONS

Contents

1	Introduction	8
1.1	Research Questions	8
1.2	Research Method	8
2	Data Mining	9
2.1	Knowledge Discovery in Databases	10
2.2	Data Mining Techniques	10
3	Crime Analysis	13
3.1	Analysing Crime: Motivation and Challenges	15
3.1.1	Operative Crime Analysis	15
3.1.2	Strategic Crime Analysis	16
3.1.3	Challenges of Processing Crime Data	16
3.2	Related Work in Crime Data Mining	17
3.3	Steps of Crime Analysis using Data Mining	18
4	Performing Crime Analysis on Weka	20
4.1	Features of the Weka Explorer	21
4.1.1	Preprocess	22
4.1.2	Classify	25
4.1.3	Cluster	26
4.1.4	Visualization	27
4.2	Data Utilized for Crime Analysis	29
4.3	Usability of the Weka Tool	30
5	Usability Testing	31
5.1	Usability Evaluation Models	32
5.1.1	Effectiveness	33
5.1.2	Efficiency	34
5.1.3	Satisfaction	35
5.2	Evaluation Model for Usability Test of Weka	36
6	Results of Usability Testing	37
6.1	Effectiveness of the Weka Tool	38
6.2	Efficiency of the Weka Tool	39
6.3	Satisfaction with the Weka Tool	40
6.4	Overall Usability of the results	43
6.5	Functionality of the Weka Tool	43

7	Conclusion and Future Work	45
7.1	Conclusion	45
7.2	Future Work	46

List of Figures

1	Steps of Crime Analysis using Data Mining	19
2	Snapshot of Weka GUI Chooser	21
3	Snapshot of Weka Explorer Preprocess	23
4	Snapshot of Weka Explorer Classify	24
5	Snapshot of Weka Explorer Cluster	26
6	Errors per Task during Usability Test	39
7	Time for each Task during Usability Test	40

List of Tables

1	Crime Data Mining - Related Work	18
2	Data Set - Columns Description	28
3	Dataset for Classification	29
4	Dataset for Clustering	30
5	Dimensions of Usability	32
6	Evaluation of Effectiveness	33
7	Evaluation of Efficiency	34
8	Evaluation of Satisfaction	35
9	Proposed Evaluation Model for Weka Usability Test	36
10	Tasks for Weka Usability Test	37
11	Completion Stage of each Task	38
12	Satisfaction with Weka after each Task during Usability Test	41
13	Overall Satisfaction with Weka	42

Abstract

Crime Analysis is a systematic and analytical process that aims at providing information relative to crime patterns and trend correlations. The department *Crime Analysis* of the Austrian Federal Criminal Police Office has the objective of providing investigators and police management with a high-quality basis for decision-making. Crime analysis is a suitable field of applying data mining, due to the large amount of crime data. The objective of applying data mining techniques is extracting interesting information and patterns. The Waikato Environment for Knowledge Analysis (Weka) is an open source software that supports performing data mining techniques. In order to assess the usability of Weka based on ISO 9241-11.

1 Introduction

Crime analysis has, according to Gottlieb et al. [36], the objective to help investigation and management of the police. It improves their work that aims to enhance the quality of our life and our safety standards. Keyvanpour et al. [48] state that crime analysts face the challenge of an increasing amount of crime data and the resulting difficulty of extracting information. Data Mining techniques are a possible solution, as they are, according to McCue et al. [58], automated methods to analyse large datasets. The Department *Crime Analysis* [21] of the Austrian Federal Criminal Police Office is constantly looking for suitable instruments to optimize the performance of their work. The department has expressed an interest in the Waikato Environment for Knowledge Analysis (Weka) [84], which can be used to perform a variety of data mining tasks.

1.1 Research Questions

The objective of this work is to assess the suitability of the Weka data mining tool [84] in the field of data analysis. Derived from this, we construct the following overarching research question, which is further divided into three sub-questions.

Is the Waikato Environment for Knowledge Analysis (Weka) a suitable tool for crime analysis?

- *What is crime analysis?*
- *How can data mining help to analyse crime?*
- *What are the strengths and weaknesses of the Weka data mining tool regarding its usability?*

1.2 Research Method

In order to answer the aforementioned research questions, a review on existing literature and a usability testing is carried out. The test is carried out using a method which evaluates usability according to the dimensions proposed by the International Organisation for Standardisation (ISO) in their ISO 9241-11 report [41]. ISO specifies usability by three factors, namely effectiveness, efficiency and satisfaction. We propose an evaluation model, which defines the attributes and measures to evaluate the degree of effectiveness, efficiency and satisfaction of the Weka tool.

2 Data Mining

Worldwide the total amount of data is constantly increasing. According to Holst [39] the volume of data generated, recorded, copied and utilised worldwide is expected to reach 59 zettabytes in 2020. The growth rate is predicted to be 26 percent per year, which leads to the fact that in 2025 there will be almost three times as much data than in 2020. Witten et al. [83] claim that the gap between the generation of data and the benefits derived from it is also expanding. Further, the authors state that discovering patterns and encapsulating them in theories can be used to forecast what will occur in future situations and to exploit behaviour patterns. This can lead to deeper insights and support better decision making.

Fayyad et al. [32] explain in their work, that a growing trend towards consolidation of distributed and transactional data stores in data warehouses has recently emerged. The authors see this shift as the process of converting a database system from a system whose primary purpose is storage to one whose main use is decision support. Inmon et al. [40] define a data warehouse as *“a subject-oriented, integrated, time-variant, nonvolatile collection of data in support of managements decision making process.”* In response to the new needs of data access, a new work area, known as On-Line Analytical Processing (OLAP), was established. The term OLAP was coined in 1993 by Codd et al. [24]. It includes concepts for the data supply of management and specialist departments, especially in the area of supporting decision making and data analysis. The authors propose that the principal purpose of OLAP systems is to facilitate the query-driven exploration of a data warehouse. This entails the precomputation of aggregating along data "dimensions" in multidimensional data stores. The exploration is issued by the user.

Difficulties occurring in OLAP systems are described by Fayyad et al. [32]. The systems rely on analysts to define and test hypotheses in the data. Thus, the inference and modelling is completely in the hands of the analyst. This results in the query formulation problem. The problem illustrates the difficulty of providing a user access to data when the user does not know how to specify the objectives in relation to a specific query. Often particular patterns occur that can only be described in a query with great difficulty, but are easily recognizable by human analysts. The authors outlined another challenge that results from limited human analysis and visualization capabilities. At a certain volume, humans are no longer able to scale and understand the high dimensions of the data. Thus, it becomes practically impossible for an analyst to go through the different possibilities of projecting the dimensions

or selecting the right subsample. The authors propose data mining as a solution to the aforementioned problems. In contrast to OLAP, data mining techniques offer computer-controlled exploration of data. They justify this proposition by clarifying that data mining enables specification of queries on a much more abstract level and makes it possible to project large volumes of high-dimensional data down to low dimensions and build models within this simplified space. McCue et al. [58] define Data Mining as a systematic analysis of large data sets using automated methods. According to the authors investigating data in this manner, makes it possible to discover new or previously unknown information and at the same time prove or disprove existing hypotheses. Particularly, unique or valuable relationships can be identified between and within the data, which can then be used proactively to categorize or anticipate additional data or subsequent events. Chen et al. [23] describe data mining as the process of extracting interesting information or patterns. That means the information is non-trivial, implicit, previously unknown and potentially useful.

2.1 Knowledge Discovery in Databases

Data mining is often referred as one of the core processes of Knowledge Discovery in Databases (KDD) [33][89]. At the first KDD workshop in 1989 [34] the term KDD was coined to emphasize that knowledge is the objective of any data-driven discovery. Fayyad et al. [33] defines KDD as a “*nontrivial process of identifying valid, novel, potentially useful and ultimately understandable patterns in data.*” A pattern in this context is any expression representing a high-level description of a data subset, including finding structure and fitting a model. KDD being a process implies that it is comprised by many different steps, which repeat iteratively. Non-trivial describes the fact that KDD is not a straightforward computation, it involves inference or search. The retrieved patterns should be valid on a new data set with some degree of certainty and should be novel to the system. Also, patterns need to be potentially useful, which means that user are able to interpret the results, at least after applying some post-processing. Thongsatapornwatana [80] defines two relevant factors to extract hidden information as follows: the data being used must be sufficient and accurate, as well as domain knowledge and experience of specialists is needed during the process.

2.2 Data Mining Techniques

Data mining is a field of study defining various analysis techniques that are useful for their purpose. Kameshwaran et al. [44] describe two learning

approaches in data mining, i.e. supervised learning and unsupervised learning. The authors define supervised learning as an approach in which the investigated variables can be split into two groups, the explanatory variables and dependent variables. The analysis intends to find a relationship between those variables. Supervised learning is referred to as directed data mining. In unsupervised learning, on the other hand, no distinction is made, that means all variables are treated the same way. However, it is not equal to undirected data mining, since there is still some objective. Zhao et al. [89] distinguish between two classes of data mining, e.g. descriptive and prescriptive. Descriptive mining targets to summarize or characterize general properties of data in data repositories. Prescriptive mining, on the contrary, performs inference on current data to make predictions based on historical data. According to Thongsatapornwatana [80] and Zhao et al. [89] data mining techniques can be divided into the following three data mining approaches:

1. Association Rule Mining

This technique is unsupervised and was first introduced by Agrawal et al. [3] in 1993. Zhao et al. [89] declare the purpose of association rule mining is to extract useful correlations, frequent patterns, associations or random structures between sets of elements in transaction databases or other data stores. Thongsatapornwatana [80] explains that the basic association rule consists of two parts. On the left-hand side of the association rule the antecedent (X) and on the right-hand side the consequent (Y). An association rule is an implication expressed as $X \Rightarrow Y$ [3], where antecedent and consequent are disjoint item sets. The rule illustrates that the antecedent implies the consequent [89].

Agrawal et al. [3] express that association rule mining is the art of finding rules that satisfy a predefined minimum of support and confidence from given data. Hence, the important statistical measures for those rules are support and confidence. Zhao et al. [89] define the measurements for association rules as follows: support of an association rule gives the percentage of records containing $X \cup Y$ to the number of all records within the database. Hence, support is often used as a measure if the association rule is interesting. Confidence of an association rule defines the percentage of records containing $X \cup Y$ to the ones that contain X . An interesting association rule $X \Rightarrow Y$ can be generated, only if the percentage exceeds the threshold of confidence.

2. Classification

According to Thongsatapornwatana [80] classification is a model that

assigns objects to a prescribed category. Each record of the input data consists of an attribute set and its label, which specifies the category to which it is assigned. Han et al. [37] explain that the general approach of classification consists of a two-step process, that separates the collected data in two sets, the training data and the testing data. First, a model, the classifier, is constructed based on the training data. This model is used for the description characteristics of a set of data concepts or classes. In the second step the model is used for classification. Here, testing data is being utilized, which is independent of the training data set. That means it is not applied to train the classification model. The purpose of the second step is to examine the performance of the classifier. There are many existing approaches to create such a classification model [37].

Phyu [69] performs a survey of classification methods including decision trees, bayesian network and k-nearest neighbor. Murthy [60] conducted a comprehensive survey on decision trees and specifies decision trees as a method to illustrate the rules retrieved from a dataset by partitioning the data recursively. Decision trees include zero or more internal nodes, as well as one or more leaf nodes. Internal nodes represent decision-points based on the value of the corresponding attribute. Each leaf node is assigned a class label. Phyu [69] defines a bayesian network to be a graphical model that provides probability relationships between a set of variable features. It is a directed acyclic graph and the nodes are in one-to-one correspondence with the features. Arcs represent random influences between features, while the absence of possible arcs indicates conditional dependencies. A survey of neighest neighbor techniques was performed by Bhatia et al. [15]. The authors describe nearest neighbor as a rule that identifies the class of an unknown data point based on its nearest neighbor whose class is already known. Cover et al. [25] propose the k-nearest neighbor in which the value k is used to calculate the nearest neighbors.

3. Clustering

Berkhin [12] defines clustering as a technique that divides data into groups with similar objects. The author clarifies that all objects within one cluster are similar to one another and different from the objects in other groups. Clustering does not require the use of predefined groups, and is hence an unsupervised data analyzing technique. Zhao et al. [89] explain that the similarity between objects is defined by similarity functions.

Rai et al. [72] conducted a survey on clustering techniques in which they distinguish between hierarchical, partitioning and density-based methods. The authors define hierarchical clustering as a method of building a hierarchy of clusters. This means that the data is not partitioned into a particular cluster in a single step. Instead, a series of partitions takes place. This runs from a single cluster containing all objects to n clusters each containing a single object. Kameshwaran et al. [44] distinguish between an agglomerative (bottom-up) and a divisive (top-down) approach for hierarchical clustering.

On the other hand, partitioning clustering methods according to Pujari et al. [71] construct a partition of records into a predefined number of clusters. An iterative optimization paradigm is usually used for this purpose. This means that partitioning is performed at the beginning and is then optimized with an iterative control strategy. Data points are exchanged to see if it improves the quality of the cluster. If no swapping leads to an improvement, a locally optimal partition has been found. The aforementioned authors distinguish between the k-means and the k-medoid algorithm. By using k-means techniques, each cluster is represented by the cluster center of gravity. Whereas, the k-medoid algorithms use objects near the center to represent each cluster. Density-based clustering technique was first proposed by Ester et al. [31]. It relies on density-based notion of clusters and aims to discover clusters in arbitrary shape. The core idea is that for each point of one cluster the neighborhood of a given radius must comprise of at least a minimum number of points. In other words, the density within the neighborhood must exceed a specific threshold value.

Data Mining is applied in many fields. Padhy et al. [68] carried out a survey on data mining applications. To give a number of examples, data mining is applied in healthcare [50][86], education [22][85], manufacturing engineering [38] and market basket analysis [46][82]. Many researchers [9][26][48][56][74][87] propose crime analysis as an appropriate field for applying data mining.

3 Crime Analysis

Crime is defined by Becker et al. [11] as *“an act or the commission of an act that is forbidden by a public law and that makes the offender liable to punishment by that law.”* Crimes encompass a broad spectrum of activities.

The Federal Criminal Police Office of the Austrian Ministry of Interior¹ summarize the different types of crime as: cyber crime, organized crime, violent crime, property crime, economical crime and environmental crime. According to the police crime statistics of Austria 2019 [2] the committed crimes between 2010 and 2019 range from approximately 470,000 to 550,000 each year. The statistic shows that cyber crime and economic crime are increasing year on year, whereas property crime is descending steadily. The overall crime rate in Austria has tended to fall over the last nine years.

Researchers [9][27][74][87] claim that crime is basically unpredictable. However, Brantingham et al. [18] state that hot and cold spots for crime occur. Also, offenders and victims appear repetitively. According to Nath et al. [61] about 50% of the crimes are committed by 10% of the criminals. Therefore, it can be said that crime is not entirely random, but neither does it occur consistently or uniformly [18][27][74][87].

Criminology is defined by Malathi et al. [56] as a field that studies crime and criminal behavior, as well as law enforcement in order to identify crime characteristics. Crime analysis as a part of criminology is, according to the aforementioned authors, a task including the exploration and detection of crimes and the relations to criminals. Osborne et al. [66] express the difficulties of defining the field of crime analysis, because the focus can be on many different issues. Organizations practice crime analysis in different forms, including crime mapping, creating statistics and analyzing reports. Thus there are many possible ways to define crime analysis, such as:

“Crime analysis is an investigation tool, defined as ‘the set of systematic, analytical processes that provide timely, pertinent information about crime patterns and crime-tend correlations’ [30].

“Crime analysis is the breaking up of acts committed in violation of laws into their parts to find out their nature and reporting statements of these finding” [66].

“Crime analysis is a law implementation task which includes an organized analysis that recognizes and determines the pattern of crime” [70].

“Crime analysis is defined as a set of systematic, analytical processes directed at providing timely and pertinent information relative to crime patterns

¹<https://bundeskriminalamt.at>

and trend correlations to assist operational and administrative personnel in planning the deployment of resources for the prevention and suppression of criminal activities, aiding the investigative process, and increasing apprehensions and the clearance of cases” [36].

Comparing those definitions with one another, it can be recognized that Gottlieb et al. [36] combines the most important elements of each one. Hence, we will use his definition of crime analysis for the purposes of this thesis.

3.1 Analysing Crime: Motivation and Challenges

Austria’s Ministry of the Interior dedicates a department entirely to crime analysis. The Department “*Crime Analysis*” (Department II/BK/4)² is in charge of the Criminal Intelligence Service and is thus responsible for crime analysis, crime statistics, information logistics and the criminal psychological service. A portfolio [21] provides an overview of the department’s objectives and key activities. According to the portfolio the mission is to provide investigators and the management of the police force in Austria with a high-quality basis for decision-making through professional analysis. The department’s vision is entitled “*Analysis before Decision*” and presents the aim of improving the quality of police work. It is distinguished between strategic and operational crime analysis. The operational criminal analysis provides the foundation for professional criminal investigation work. In strategic crime analysis, decision bases are supplied to professional police management. The results of both combined are an indispensable basis for effective and efficient use of resources to avoid redundancies and achieve synergies.

3.1.1 Operative Crime Analysis

Operational crime analysis includes the collection of information, the evaluation of this crime data, as well as the analysis and visualization of the results. It aims at short-term police measures such as arrests, seizures, confiscations. Also, it provides a basis for decisions regarding further investigative approaches in complex criminal cases and supports investigating officers. The processed data are mainly personal data, which have been determined from observation, telephone surveillance, call data recapture, eavesdropping or other such activities. The operational analysis is therefore an evaluation that accompanies, supports and initiates investigations.

²<https://bundeskriminalamt.at/101/abteilungen.aspxa4>

3.1.2 Strategic Crime Analysis

Strategic Crime Analysis deals with mid- and long-term questions concerning characteristic patterns of crime and investigates factors of possible future events. It enables decision-makers to set priorities and develop mid- and long-term plans by establishing focal points. Results of a strategic crime analysis can be presented in the form of situation reports. Situation reports are used to present abstract, non-personal overviews of the state of crime in a specific space, at a specific time, in relation to general or specific areas of crime. It provides a description of the crime development including political, economic, social, technological, ecological and legislative aspects. With regard to this development, various trend calculation models are used for resource and deployment planning. Essential elements of such a report are causal research, comprehensible forecasts, threat scenarios and risk analyses. Based on these, preventive or repressive countermeasures can be developed.

3.1.3 Challenges of Processing Crime Data

As previously mentioned information logistics is a subcategory of the department of criminal analysis in the Austrian Ministry of the Interior. According to the department's portfolio [21] it includes the planning, management, implementation and control of data and information flows. Constant core tasks are the preparation of data, their presentation for analytical use, as well as the ongoing support and improvement of the electronic information flow. The goal is to process criminal investigation data into usable information and make it available to potential user. Special working groups are concerned with finding suitable search and analysis technologies. An important aspect is managing the large volume of data.

Malathi et al. [56] describe the major challenges data analysts face today. One of the described difficulties is the increasing amount of information that is collected, stored and analyzed. As the data available are often inconsistent and incomplete, analysis becomes significantly more complex and therefore more time consuming. According to the authors this is followed by an additional challenge of identifying appropriate techniques to make an accurate and efficient analysis that can handle this large quantity of data. The law enforcement needs to utilize the current technology for analysing crime in order to remain on the leading edge in the race against criminals. Amarnathan [6] draws attention to the fact that also criminals become increasingly more technologically sophisticated when committing crimes. Nath [62] proposes utilizing an interdisciplinary approach between computer science

and criminal justice to perform crime analysis. More specifically, data mining is suggested as a technique that can be applied to support analysis on the increasing amount of crime data. Keyvanpur et al. [48] also suggest the application of data mining in crime analysis, due to the high volume of data and the complexity of relationships between the types of data. David et al. [26] affirms this by claiming that clustering and classification based models can be utilized to identify crime patterns and criminals.

3.2 Related Work in Crime Data Mining

Many researchers are addressing the subject of using data mining techniques for crime analysis. A detailed overview of existing work in this field can be found in Table 1. The works are categorized according to the overall purpose they pursue the utilized data mining techniques as well as the underlying algorithm.

Some papers have set the focus on finding similarities in crime data. Al-Janabi et al. [5] apply association rule mining and classification for finding patterns, trends and relationships. Anitha [7] compares three different clustering algorithms to identify patterns of crime against women in India. In addition to clustering, the analysis by Tayal et al. [79] also utilizes classification algorithms to group entries according to their similarity. Kaza et al. [47] attempt to find relation in crime entries using a classification technique for social network analysis. An approach to group crime related emails is adopted in the paper of Sharma et al. [76].

In addition to detecting patterns, several authors apply data mining techniques with the aim of crime prediction. Malathi et al. [56] use association rule mining, clustering and classifications to support solving crimes with crime prediction. Rajeswari et al. [73] utilize clustering and classification techniques to forecast criminality rates against women. Sathyadevan et al. [74] concentrate their prediction not on single cases but attempts to predict regions with high crime rates. The authors make use of association rule mining and classification algorithms. Kondaveeti [51] focuses on the identification of crime trends in geographic data utilizing association rule mining.

Other work targets the detection of potential crime suspects. To do so, Deshmuk et al. [27] compares three classification methods and adopts the best performing one. In comparison, Yu et al. [88] use a clustering technique to achieve this goal.

Table 1: Crime Data Mining - Related Work

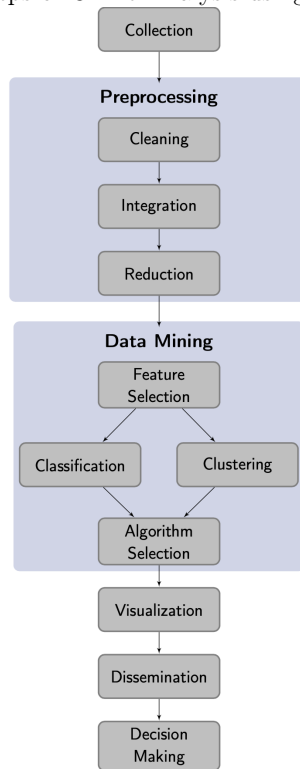
Author	Task	Technique	Algorithm
Al-Janabi et al. [5]	Discovering trends, patterns and relationships	Association rule mining, classification	-
Anitha [7]	Identification of crime patterns	Comparing 3 clustering algorithms	K-means, hierarchical, DBSCAN
Deshmuk et al. [27]	Identification of potential suspects	Comparing 3 classification algorithms	J48, JRip, Naive Bayes
Kaza et al. [47]	Finding relationships with social network analysis	Classification	Cox-Regression
Kondaveeti [51]	Crime trends in geographic data	Association rule mining	-
Malathi et al. [56]	Prediction of crime and help to solve crimes	Association rule mining, Clustering, Classification	Apriori, K-means, Decision Tree
Rajeswari et al. [73]	Prediction of crime against women	Clustering, Classification	K-means, ID3, Decision Tree
Sathyadevan et al. [74]	Prediction of regions with high criminality rate	Association rule mining, classification	Naive Bayes, Apriori, Decision Tree
Sharma et al. [76]	Grouping crime-related emails	Classification	Decision Tree, ID3
Tayal et al. [79]	Grouping instances by similarities	Clustering, Classification	K-means, KNN
Yu et al. [88]	Detecting crime suspects	Clustering	SMMSM

3.3 Steps of Crime Analysis using Data Mining

A workflow for the procedure of applying data mining for the purpose of crime analysis was mapped by many researchers [5][9][26][74][79][80]. Based on the mentioned processes, we propose the consolidated workflow, depicted in Figure 1, by selecting certain steps from the different models and combining them.

The process is initiated by the collection of data, joining to a dataset which is

Figure 1: Steps of Crime Analysis using Data Mining



then used for analysis. The resulting dataset needs to be preprocessed before applying any data mining techniques. Krishnamurthy et al. [52] claim that raw data requires preprocessing, as it is in different formats, collected from different sources. The authors present essential subcategories of this step as data cleaning, data integration and data reduction. With data cleaning missing values are added, noisy data is smoothed, outliers are removed and inconsistency is resolved. Data integration involves the merging of data. In the step of data transformation for standardizing data, data normalization and attribute construction are performed. After applying these steps, better results can be achieved in data mining.

Subsequently data mining techniques are applied. The proposed papers mention the use of clustering and classification for crime analysis. David et al.

[26] suggests the use of clustering to find patterns of crimes. Whereas classification is used for crime prediction. In order to apply the chosen technique one has to choose a specific algorithm for proceeding data mining. Deshmukh et al. [27] point out that a large number of possible algorithms are available, hence it is challenging selecting the appropriate one. The authors provide a list of criteria for selecting the most suitable. The algorithm needs to be simple to understand and easy to interpret. It must provide the intelligence to identify useful data, as well as contribute to a high information gain. After a suitable algorithm is found and applied, the result of the executed data mining techniques needs to be visualized. Thereby the essential knowledge can be extracted. The outcome is then disseminated and can serve as base for decision-making.

The procedure of data mining requires the implementation of sophisticated instruments to enable the execution of the algorithms. A number of previous works utilize the Waikato Environment for Knowledge Analysis (Weka) [84] to perform data mining in the field of crime analysis. Anitha [7] analyzes patterns of collected crime data by applying three clustering algorithms, namely k-means, hierarchical clustering, DBSCAN. Each algorithm is deployed within the Weka tool. Deshmukh et al. [27] compare the performance of three classification algorithms, specifically J48, Naive Bayes and JRip. Also in this paper the Weka software is used for the comparison. Rajeswari et al. [73] apply k-means clustering and ID3 classification with the help of the Weka tool. Tayal et al. [79] utilize the software within the verification process in order to verify the results of k-means clusters. Zubi et al. [90] use Weka together with excel to analyze and preprocess the data.

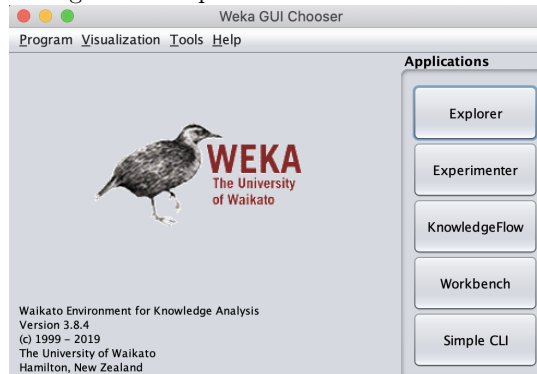
4 Performing Crime Analysis on Weka

The Weka Software [29] is a system which enables users to perform data mining without actual coding. It is developed by the University of Waikato and supported by the New Zealand government. Weka is an open source software released under the GNU General Public License [54], which uses a Java-based³ graphical user interface (GUI). Weka implements data mining algorithms for preprocessing, classification, clustering and association rule mining, as well facilitating visualization. The following software description refers to Weka Version 3-8-4⁴, as it is the latest stable version, and is based on Weka Explorer Tutorial [4], Weka Manuel [17] and the Weka Explorer

³<https://www.java.com/de/>

⁴<https://sourceforge.net/projects/weka/files/weka-3-8/3.8.4/>

Figure 2: Snapshot of Weka GUI Chooser



User Guide [49].

The software's starting page, shown in Figure 2, provides a choice between four different GUIs, namely the Explorer, Experimenter, KnowledgeFlow, Workbench and Simple Command Line (Simple CLI). According to Witten et al. [84] the explorer is used to explore datasets by giving access to all its facilities through menu selection. The authors explain that the Experimenter GUI is designed to give the answer to the question which algorithms and parameters perform best for the given problem. The Experimenter allows different models to be compared automatically by collecting performance statistics and providing significance tests. Bouckaert et al. [17] explain that the KnowledgeFlow GUI presents an interface inspired by data flows. It allows steps to be selected, placed and connected on a canvas, with the aim of forming a knowledge flow to process and analyze data. This interface is able to handle data either in batches or incrementally. The Workbench is described by the authors as an all-in-one application that combines all important Weka-GUIs. Witten et al.[84] refer to the simple CLI as the old-fashion interface, as it only accepts command-lines.

4.1 Features of the Weka Explorer

Witten et al. [84] state that the Explorer is the most popular interface of the Weka software. Kaur et al. [45] even refer to it as the main interface.

By selecting the explorer from the GUI choice, one gets redirected to a separate window. At the very top of the page a bar containing six different

tabs give users the opportunity to jump from one tab to another. The tabs present different functionalities, namely *Preprocess*, *Classify*, *Cluster*, *Associate*, *Select attributes* and *Visualize*. At first only the first tab is active, as it is necessary to open a dataset before one can use the further steps. This can be done within the preprocessing window. Weka allows for files to be opened from the local file system, via a Uniform Resource Locator (URL), from a database or to generate artificial data from data-generators. The Weka tool only supports data files as an Attribute-Relation File Format (ARFF)⁵ or a CSV-format⁶. Once a dataset is opened the five other tabs are activated.

The status box is displayed at the very bottom of each window. Its purpose is to keep the user informed of the current condition of the software. For instance, after opening a data file the status box displays “*Reading from file...*”. Once the data is successfully opened, it changes to “*OK.*”. Next to this box, the Weka software provides a Log Button. After clicking on the button, a separate window pops up and provides a list of previous actions with the corresponding timestamp. On the right hand-side of this button, one can find the status icon of Weka, which is designed as a bird. A moving bird indicates a running process, whereas a sitting bird means that no process is running at this moment. The digit next to the bird shows the number of concurrently running processes.

The following chapters describe the parts of the Weka software that are relevant for the scope of this thesis. The focus lays on the steps of preprocessing, classification, clustering and the visualization of the results.

4.1.1 Preprocess

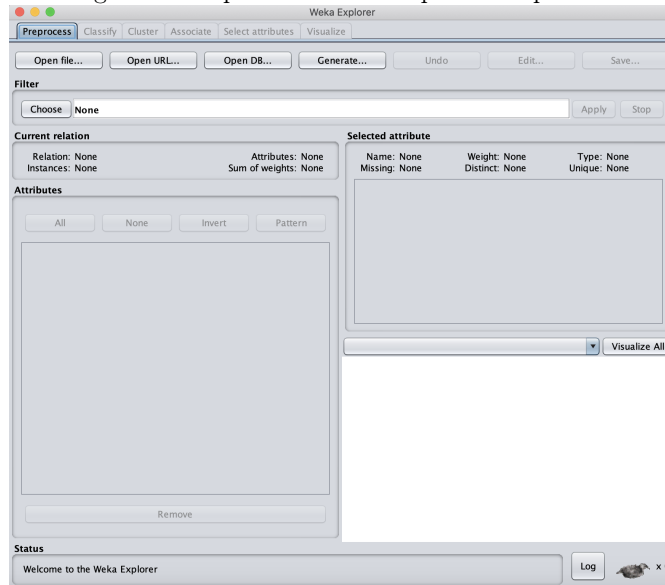
Once data is opened successfully, the user is able to investigate and modify it in various ways within the first tab. A snapshot of this tab is given in Figure 3. On the upper right side one can find an *Undo* and *Save* button in order to revoke all the changes made or to save the modified dataset. The window of preprocessing includes boxes, called Filter, Current relation, Attributes, Selected attribute and a visualization window.

Witten et al. [84] explain that *Current relation* stands for the currently opened data, as it is interpreted as a single relational table in database terminology. Columns are referred to as attributes and rows are called instances. The *Current relation* panel gives general information, such as the name of

⁵<https://www.cs.waikato.ac.nz/ml/weka/arff.html>

⁶<https://json-csv.com/csv>

Figure 3: Snapshot of Weka Explorer Preprocess

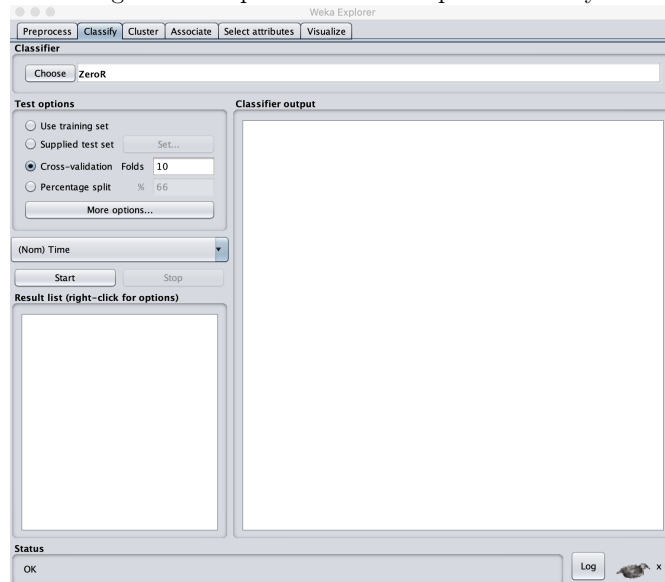


the opened relation, number of instances and number of attributes. The box titled *Attribute* includes four buttons and underneath a table with three columns. The table displays a number that identifies each attribute in the same order as they are specified in the data file, a selection tick box per attribute to choose which attributes are currently present in the relation and the name of the attributes. The attributes can be ticked manually or the four buttons can be utilized to select all, none, invert (boxes ticked become unticked and vice versa) and after patterns (tick attributes after condition).

By clicking on one attribute in the table, the box called *Selected attribute* provides the user with the name and the type of the attribute. It also gives the total number and percentage of missing, distinct, as well as unique values. The Weka software automatically computes basic statistics on each attribute. This statistical information is presented in a table within *Selected attributes*. The table displays the frequency of different values, if a categorical attribute is selected. The minimum, maximum, mean and standard deviation are given for numeric attributes. Below these statistics a visualization field displays the specific attribute as a histogram⁷. The visualization is color-coded ac-

⁷<https://www.dictionary.com/browse/histogram>

Figure 4: Snapshot of Weka Explorer Classify



ording to a chosen *Class* attribute, which can be changed in a drop-down menu above. A button, titled *Visualize all*, opens a separate window to visualize all attributes colored based on the selected *Class* attribute.

At the very top, user can find the *Filter* panel. By clicking on the *Choose-Button*, it gives a list of all filters included in Weka. The filters are grouped in supervised and unsupervised and further differentiate between attribute or instance filters. One can choose from a wide range of filtering options. To just state some examples the filters allow changing numeric to nominal or binary, add values, standardize as well as remove by name. Once a filter is selected, the field next to the button shows it's name and options. By clicking on this field a separate window, called the *GenericObjectEditor* allows for the configuration of the specific filter and provides the user with information about the method. The configuration can either be saved or just applied once. Next to the field in the main window, the *Apply* button initiates the filtering on the attributes that are ticked.

4.1.2 Classify

The second panel *Classify* gives access to classification algorithms. Users can choose from a variety of available learning schemes, such as decision trees, logistic regression, as well as bayes' nets, just to name a few. The window encompasses four fields, namely *Classifier*, *Text options*, *Result list* and *Classifier output*. A snapshot of the interface is provided in Figure 4.

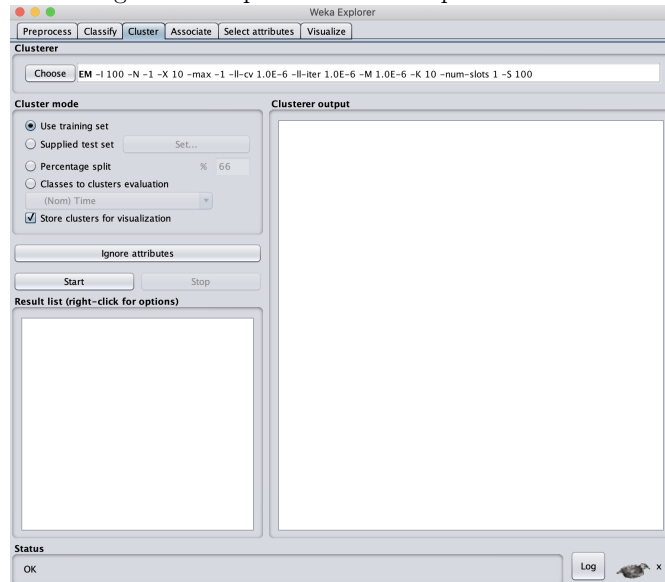
The *Classifier* at the top allows the classification method of choice to be selected. Its functionality are similar to choosing a filter in the *Preprocess* tab. By clicking the button *Choose* a list of the methods is given. The field next to the button provides information about the chosen algorithm and redirects by clicking to the *GenericObjectEditor* for modifying the method. The results of the chosen classifier are tested according to the settings within the *Test option* box. The test mode on which the performance is evaluated differs between using the training set, a supplied test set, cross-validation or percentage split. The button *More options* provides further options on what is included in the output. Below this box is a drop-down option to choose the target attribute for prediction. Beneath this the *Start* and *Stop* buttons are placed to initiate building the model and cancel the procedure.

When the model training operation is complete the Classifier output area is filled with the results in a text-based format. The output includes following sections:

- *Run information*: The algorithm used, the relation name, the number of instances as well as attributes and selected test mode
- *Classifier model*: A textual representation of the produced model
- *Summary*: A list of summarizing statistics on the model accuracy based on the chosen test mode
- *Detailed Accuracy By Class*: A more detailed accuracy per each class
- *Confusion Matrix*: A confusion matrix is defined in the Encyclopedia of Machine Learning [81] as “a two-dimensional matrix, indexed in one dimension by the true class of an object and in the other by the class that the classifier assigns”

In addition to the results, a new entry emerges within the box *Result list* when a training is finished. Each entry in this box represents one run of training a model with its timestamp. By right-clicking on one entry a list

Figure 5: Snapshot of Weka Explorer Cluster



with different options is given. It includes viewing the text-based result in a separate window, saving the result, saving the model, load a pre-defined model, re-evaluate the model with a selected test set or visualize the result in different ways.

4.1.3 Cluster

The cluster methods in the Weka tool are used for finding similar instances and grouping them. The interface of this tab is shown in Figure 4.1.3. The window resembles very much the *Classify* window. It consists of four separate fields, which are *Clusterer*, *Cluster mode*, *Clusterer output*, as well as *Result list*. The functionality of these areas is again comparable to the classification methods. Within *Clusterer* the algorithm can be defined and modified by clicking on the field next to it. The *Cluster mode* specifies the evaluation of the result. Three of the options are the same as in *Classify*, in particular using training set, supplied test set or percentage split. The last option compares the match of the chosen cluster to a defined class in the data. This class attribute can be chosen from a drop-down menu. Underneath this box, the button *Ignore attribute* enables user to choose attributes, which should be ignored while performing clustering.

When the training set is complete a new entry appears in the list of results with equal options as with classification result entries. The *Clusterer output* also displays text-based results with following information:

- *Run information*: Used algorithm, number of clusters, relation name, number of instances and attributes as well as selected test mode
- *Clustering model (full training set)*: Number of iterations, sum of squared error within clusters and center information of each cluster
- *Evaluation on training set*: Number of iterations, sum of squared error within clusters, center information of each cluster, number of instances of each cluster, class assigned to each cluster and number of incorrectly classified instances

4.1.4 Visualization

As we are focusing on visualizing the trained models, we will not describe the functionality of the *Visualize* tab, as this tab refers only to visualizing the attributes and their correlations within 2D plots. The models, on the other hand, can be visualized by right-clicking on one entry within the entry list.

In order to visualize a classification model one can choose between visualizing the classifier errors, tree or graph, the margin curve, threshold curve, as well as cost curve. *Visualizing classifier errors* provides the results of the classification. Instances shown as crosses represent correctly classified data. Squares indicate incorrect classification. *Visualize tree/graph* displays a graphical representation of the structure. For instance, a decision tree is visualized with its roots and nodes. Bouckaert et al. [17] explain the margin and threshold curve in detail. The cost curve visualizes the expected costs, as described by Drummond et al. [28]. The results of the clustering can be visualized by choosing *Visualize Cluster assignments*, which allows visual inspection of the clusters made.

A separate window named *Weka Classifier/Clusterer Visualizer* opens, when one of the aforementioned visualization alternatives is chosen. The window's main area holds a 2D scatter plot, depicting results of the chosen model training. Drop-down options at the very top of the window, allow the user to modify the graph to their needs. The user is able to change the Attribute shown on the X- and Y-axis, as well as the attribute on which the color of

Table 2: Data Set - Columns Description

German	English	Type	Description
PK SIMO	ID	Integer	Official ID for each incident
Bundesland	State	String	State in which crime was committed
Bezirk	District	String	District in which crime was committed
Postleitzahl	zipcode	Float	Zipcode of the district
Tatzeit von	starting time	String	Date and time when crime started
Tatzeit bis	ending time	String	Date and time when crime ended
FK-Abschnitt	Paragraph	Integer	Paragraph of the criminal code
Delikt	Offence	String	Section of the criminal code
Schlagwort	Keyword	String	Keyword to describe type of incident
Geklaert	Solved	Integer	0: crime was solved 1: crime was not solved
Versuch	Attempt	Integer	0: burglar was successful 1: burglar was not successful
Begehung	Crime Commission	String	Code describing how the burglar gained unauthorized access
Gut	Good	String	Code or Codes describing each stolen good
Oertlichkeit	Locality	Float	Code describing which kind of locality

instances are based on. To the right of the plot one can find a series of horizontal strips representing each attribute. Clicking on those elements allows users to switch between different attributes displayed on one of the axes. Hence, the functionality remains the same as for two of the drop-down lists at the top. Above those strips, the visualizer provides a *Jitter* option, which randomly displays all instances, until the user is able to see all points, even if they are concentrating on one spot.

A fourth drop-down provides selection options. This option can be used to modify the shape used for data selection. *Select Instance* is set by default, which means clicking on one instance brings up a separate window with its specific information. User can also select a subset of data points. The subset can be defined by creating either a rectangle, a free-form polygon or building a polyline. The visualization window also includes four buttons to either *Submit/Reset*, *Clear*, *Open* and *Save*. Once a subset of the data is selected, the chosen area turns grey. Clicking the *Submit* button at this point, extracts all instances except the ones within the grey field. By doing this, the same button changes to *Reset*, in order to return to the original graph by undoing all removals. The *Clear* button extracts the selected area without affecting

Table 3: Dataset for Classification

State	District	Solved	Attempt	Keyword	Year	Season	Time
Vienna	Simmering	No	No	Apartment	2012	Winter	Night
Carinthia	Villach	Yes	Yes	Vehicle	2012	Winter	Night
Carinthia	Villach	Yes	Yes	Vehicle	2012	Winter	Night
Carinthia	Villach	Yes	Yes	Vehicle	2012	Winter	Night
Carinthia	Villach	Yes	No	Vehicle	2012	Winter	Night
Carinthia	Villach	Yes	Yes	Vehicle	2012	Winter	Night

the other parts of the graph. The *Open* Button provides the opportunity to open previously saved instances. *Save* enables user to save currently visible instances as an ARFF file.

4.2 Data Utilized for Crime Analysis

The Federal Criminal Police Office provide burglary data for Vienna and Carinthia from 2012 to 2018. It must be mentioned that this data record consists solely of the reported burglaries. Only what reaches the police in an official manner can be documented. This dataset is utilized to perform data mining tasks within the Weka software. The dataset consists of 15 columns and around 156.000 rows. Each row represents one burglary incident and the columns provide detailed information about the crime. For confidentiality reasons we are not permitted to publish the actual dataset. The most important information about the described attributes is given in Table 2, such as original name in German, the English name, datatype and description. The data contain exclusively “*Theft by burglary or with weapons*”[1]. It concentrates on vehicle, basement, house and apartment burglaries.

As indicated in the steps of crime data mining in Section 3.3, the collected data needs to be preprocessed in order to apply data mining techniques reasonably. Accordingly, the given data is converted into appropriate formats. Data cleaning consists of following activities: changing the data type of different columns (*Starting time*, *Ending time*, *Solved*, *Attempt* and *Locality*), abjusting the inconsistency within the date column (e.g. *1900*), as well as excluding inaccurately entries (location *Bregenz* assigned to *Vienna*). Since the clustering and classification technique are used for different purposes, we further prepare two separate datasets.

Table 4: Dataset for Clustering

Year	District	Vehicle	Basement	House	Apartment
2012	Alsergrund	178	73	1	218
2012	Brigittenau	487	310	6	290
2012	Donaustadt	741	379	463	307
2012	Döbling	175	126	98	248
2012	Favoriten	813	639	125	669
2012	Floridsdorf	566	326	289	294

- **Data for Classification**

We apply the technique of decision trees with the algorithm "J48" for our classification tasks. The objective hereby is to classify the entries whether they are clarified or not. Classifying data for this purpose can be used to support the executive's operational management in prioritizing during resolution. To implement this classification method, the different attributes need to be categorical. Therefore we replace numerical values like date and time by nominal ones. We implement this by grouping the time into four different categories, namely morning, noon, evening and night. We do the same with the month, which is categorized into spring, summer, autumn, winter. Additionally we keep the columns "*District*" and "*State*". Table 3 shows the first rows of the data set.

- **Dataset for Clustering**

For clustering we use the k-means algorithm. Based on the dataset provided by Anitha [7], we add up the total count of the four different categories of burglary per district and year. In this file we choose to exclude Carinthia and focus on Vienna, in order to have a more compact dataset. In order to achieve this we sum up each entry to the total number of incidents for each year and district. Table 4 shows an extract from the final data set.

4.3 Usability of the Weka Tool

Researchers [8][57][77] claim Weka as easy to use due it's graphical user interface. Barnum [10] explains the need for a product's usability to be maximized by making it clear that users have too many other applications to choose from. Weka is not the only product that offers these capabilities. In 2019 Nguyen et al. [63] conduct a survey on machine and deep learning frameworks for large-scale data mining. The authors affirm that while

the number of algorithms is growing, also the number of implementations through frameworks and libraries is increasing. Software development in this area is evolving rapidly, with a large number of open source software coming from academia, industry, start-ups, or broader open source communities. As a consequence, we conduct a case study with the aim to evaluate the usability of the software. The degree of usability is evaluated via a usability test with two employees of the Federal Criminal Police Office. Both participants work in the field of crime analysis and hence hold prior knowledge of data mining techniques.

5 Usability Testing

Table 5 gives a comparison on the various perspectives on the usability attributes. Booth [16] suggests in his early study four different attributes, such as attitude, effectiveness, learnability and usefulness. Shackel [75], on the other hand, claims that usability depends on the the number of accomplished tasks, learnability, effectiveness, flexibility and the user’s attitude. According to Joo et al. [43] one of the most cited models in usability engineering is introduced by Nielsen [64] and proposes that usability can be determined by the degree of learnability, efficiency, memorability, error rate (the ease of error recovery) and subjective satisfaction. The International Organization for Standardization (ISO)⁸ assesses usability based on three measurements, namely effectiveness, efficiency and satisfaction. ISO establishes this construct as an international standard, named ISO 9241-11 [41]. Brinck et al. [19] defines usability as a correlation of functional correctness, efficient to use, easy to remember and learn, error tolerant and subjectively pleasing. Oulanov et al. [67] sees efficiency, helpfulness and adaptability as usability criteria. Lee [53] postulates in his usability test study usefulness, effectiveness, satisfaction, supportiveness, as well as intuitiveness as attributes of usability.

Bevan [13] presents the benefits of using standards. According to the author applying the attributes proposed by the ISO provides an appropriate evaluation. Effectiveness and efficiency can easily be related to business objectives by evaluating whether user are able to interact with the system and how long it takes to complete tasks. Satisfaction is relevant, as it motivates repeated usage. Hence, we conduct our evaluation of Weka according to the definition of ISO 9241-11 [41]. Usability is therefore defined as *“the extent to which a system, product or service can be used by specified users to achieve specified*

⁸<https://www.iso.org/home.html>

Table 5: Dimensions of Usability

Attribute	Booth	Brinck et al.	Shackel	Nielsen	ISO 1994	Oulanov	Lee et al.
Attitude	×		×				
Effectiveness	×		×		×		×
Learnability	×	×	×	×			
Usefulness	×						×
Functional Correctness		×					
Efficiency		×		×	×	×	
Memorability		×		×			
Error Tolerent		×					
Satisfaction		×		×	×		×
Accomplished Tasks			×				
Fleixbility			×				
Error Rate				×			
Helpfulness						×	
Adaptability						×	
Supportiveness							×
Intuitiveness							×

goals with effectiveness, efficiency and satisfaction in a specified context of use” [41]. The three attributes are interpreted by the ISO as following:

- Effectiveness presents “the accuracy and completeness with which specified users can achieve specified goals in particular environments”.
- Efficiency measures “the resources expended in relation to the accuracy and completeness of goals achieved”.
- Satisfaction shows “the comfort and acceptability of the work system to its users and other people affected by its use”.

5.1 Usability Evaluation Models

Researchers introduce different ways how the introduced attributes can be evaluated. Moumane et al. [59] explains that parameters like effectiveness and efficiency are objectively assessed through variables. Satisfaction, on the

Table 6: Evaluation of Effectiveness

Evaluation	<i>Jeng</i>	<i>Bevan et al.</i>	<i>Moumane et al.</i>	<i>Georgsson et al.</i>
Correctly completed tasks	×	×		
Data entry time			×	
Degrees of task completion				×
Number of completed tasks	×	×		
Number of errors		×		×
Number of tasks with errors		×		
Response time				
Time to learn and use			×	
Task time			×	
Time to install			×	

other hand, is a subjective attribute and can be evaluated through questionnaires or emotional expressions.

5.1.1 Effectiveness

Jeng [42] proposes an evaluation model for assessing usability, applying the definitions of the ISO 9241-11. The suggested model evaluates effectiveness by assessing whether the system as a whole can provide its functionality and information effectively. This is measured by the number of correctly completed tasks. Bevan et al. [14] introduce a model to assess the three usability indicators. Effectiveness is evaluated by how many tasks are completed, how many objectives are achieved, the number of errors during a task, the number of tasks with errors and the task error intensity. Moumane et al. [59] design another evaluation model based on the ISO standards while performing usability testing for mobile applications. Based on this model effectiveness is measured by the time to learn and use, data entry time, tasks time, response time and time to install. Georgsson et al. [35] state in their paper that ISO measures the attributes in the following manner. Effectiveness is determined by the degree of task completion, applying three categories:

1. Completed with ease: When the task was executed without any help from the test leader.

Table 7: Evaluation of Efficiency

Evaluation	<i>Jeng</i>	<i>Bevan et al.</i>	<i>Moumane et al.</i>	<i>Georgsson et al.</i>
Completion rate			×	
Cost-effectiveness		×		
Need of fatigue		×		
Number of unnecessary actions		×		
Number of errors			×	
Task time	×	×		×
Time efficiency		×		

2. Completed with difficulty: When the subject faced minor difficulties achieving the task and/or requires minor hint to accomplish the task.
3. Failed to complete: The subject did not manage to perform the task, even when minor hints are provided.

According to the authors, the second measurement for effectiveness is the number of errors while performing a task. An error was coded if the user created an error that could not be solved or prevented further progress. All aforementioned effectiveness evaluations methods are summarized in Table 6.

5.1.2 Efficiency

Efficiency is estimated during the evaluation performed by Jeng [42] by the systems ability to retrieve information efficiently and is judged by the time of completing a task. Bevan et al. [14] measure efficiency through the task time, the time efficiency, the cost-effectiveness, the ratio of productive time, number of unnecessary actions and the degree of fatigue. Moumane et al. [59] evaluates the efficiency with the number of error and the completion rate of the tasks. The model proposed by Georgsson et al. [35] measures the efficiency through time of performing or trying to perform the task. Table 7 provides an overview of the different attributes used to measure efficiency.

Table 8: Evaluation of Satisfaction

Evaluation	<i>Jeng</i>	<i>Bevan et al.</i>	<i>Moumane et al.</i>	<i>Georgsson et al.</i>
Information	×			×
Complexity				×
Capability				×
Discretionary Usage		×		
Usability	×		×	×
Error correction	×		×	
Feature Satisfaction		×		
Feature Utilization		×		
Frequency of use				×
Inconsistency				×
Labelling	×			
Learnability			×	×
Organization of information	×		×	
Overall Satisfaction		×	×	
Pleasure		×		
Support				×
Trust		×		
Visualization	×		×	

5.1.3 Satisfaction

The evaluation model proposed by Jeng [42] measures satisfaction with Likert scales [55] and questionnaires. A Likert scale aims to measure attitudes with a five point scale used to express how much the user agrees or disagrees with a particular statement. Based on this model, satisfaction can be evaluated by the rate of ease of use, organization of information, labelling, visual appearance, content and error correction of the system. Bevan et al. [14] estimates the satisfaction with the individual features the overall systems, the discretionary usage of the system, the feature utilization, the proportion of users complaining, the trust and pleasure of the users as well as the physical comfort. A questionnaire, namely QUIS 7.0 [65], is utilized to capture the satisfaction of usage in the evaluation model Moumane et al. [59]. QUIS

Table 9: Proposed Evaluation Model for Weka Usability Test

Attribute	Measurement	Instrument
Effectiveness	Task Completion	Completed with ease Completed with difficulty Failed to complete
	Error Rate	Number of errors for each task
Efficiency	Task Time	Compare task time of tested user
Satisfaction	Ease of use Organization of Information Labelling Visual appearance Content Error Correction	Likert scale after task Post-test questionnaire

7.0 evaluates the satisfaction in six sections, the overall satisfaction, visualization, information, learnability, capability and usability. Georgsson et al. [35] evaluates it with a System Usability Scale (SUS) devised according to Brooke [20]. With an SUS-tool statements are measured on a scale from 0 to 100. The statements include the frequency of use, complexity, ease of use, support, content, inconsistency, learnability. All attributes used to evaluate the satisfaction of the user are outlined in Table 8.

5.2 Evaluation Model for Usability Test of Weka

We propose an evaluation model, given in Table 9, to estimate the usability of the Weka tool. Applying the three usability attributes proposed by ISO 9241-11 [41], the usability is judged by the level of effectiveness, efficiency and satisfaction. The assessment of effectiveness and efficiency is derived from the model applied by Georgsson et al. [35]. Effectiveness is therefore determined by the completion stage and the amount of errors per task. Satisfaction is based on the evaluation method proposed by Jeng et al. [42]. The authors use a pre-test questionnaire to collect demographic data, a Likert scale after each task and a post-test questionnaire. The evaluation of usability after each task is intended to assess the level of current satisfaction. The post-test questionnaire examines satisfaction with the overall performance. It is a combination of Likert scales and open questions. When conducting our usability tests, we neglect the pre-test questionnaire since we do not consider

Table 10: Tasks for Weka Usability Test

No.	Task Description
1	Open the Weka Explorer and load the CSV file "df_clustering".
2	Select the attribute "House", select "District" as "Class" and click on the button to visualize all.
3	Tell me in which district the most burglary into houses take place.
4	Close the tap, go to "Cluster", choose "SimpleKMeans" and change the cluster number from two to four.
5	Choose District as classes to cluster evaluation and start clustering.
6	Load the CSV-file "df_classification".
7	Tell me if the attribute year is numeric or nominal. If necessary, change it to nominal by choosing an unsupervised attribute filter.
8	Change the value of "Attempt" and "Solved" from Numeric to binary.
9	Go to "Classify" and choose "J48" Tree.
10	Classify by the attribute "solved" and start the classification.
11	Visualize the decision tree.

demographic characteristics in evaluating our results. In addition, we adapt the post-test questionnaire proposed by Jeng [42] to suit our needs.

According to Jeng [42] the selected tasks need to be representative of a typical use of the system. Since we evaluate the usability of crime data mining with the Weka tool [29], we design the tasks in a way that illustrates exemplary sequences of steps needed to perform crime data mining. Srivastava [78] describes the individual steps in order to implement clustering and classification in the Weka tool [29]. The chosen tasks are aligned with these steps and presented in Table 10. This test is designed as a usability test and therefore does not include testing the performance and accuracy of the algorithms. Only the usability of the software is examined. The templates for the Weka usability test can be found attached in the appendix. Appendix A includes the tasks to be completed and the scales to rate satisfaction. Appendix B shows the post-test questionnaire.

6 Results of Usability Testing

The conducted tests provide us with some insights into Weka's usability based on the chosen evaluation model. During execution of the task quantitative results are collected. As previously shown the values include level of completion, error rate, task time and degree of satisfaction after each task. This

Table 11: Completion Stage of each Task

Task	Completion					
	with ease		with help		failed	
	A	B	A	B	A	B
Task 1		○	●			
Task 2	●			○		
Task 3				○	●	
Task 4			●	○		
Task 5	●	○				
Task 6	●	○				
Task 7					●	○
Task 8			●			○
Task 9	●	○				
Task 10	●	○				
Task 11				○	●	

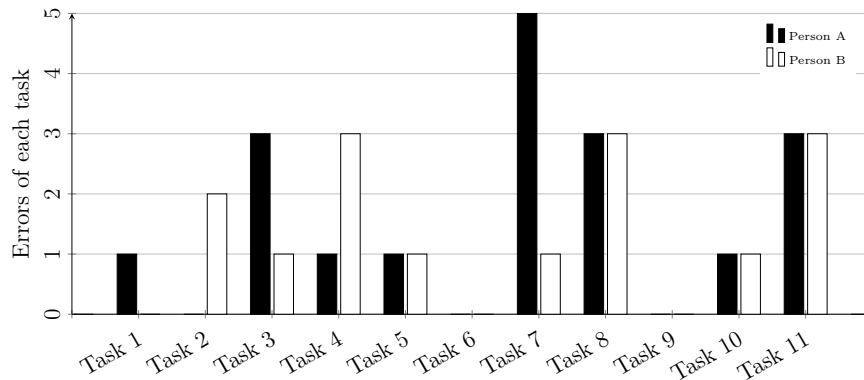
allows the overall experience to be determined, as well as different features to be analyzed and compared. The questionnaire provides quantitative and qualitative answers. With this, the overall satisfaction is evaluated. It also indicates the most and least satisfying features. It is important to consider that this evaluation only covers the requested functions and not the entire scope of the software.

6.1 Effectiveness of the Weka Tool

Two parameters are employed to evaluate the effectiveness of the Weka software, the level of execution and the number of errors made. From the results one can conclude that the two are strongly correlating. The tasks that were performed with ease usually had no errors or at most one. The ones that failed to be executed had the highest error rates. The detailed result of task execution is shown in Table 11, as well as the number of errors in Figure 6.

The main results of the categorizations regarding the effectiveness for different functionalities are as follows: both participants are able to complete task 6, 9 and 10 without any help. Task 6 and 9 are also the only tasks where no error occurred. Hence it can be concluded that the execution of these functions is managed most effectively. Whereas the most ineffective

Figure 6: Errors per Task during Usability Test



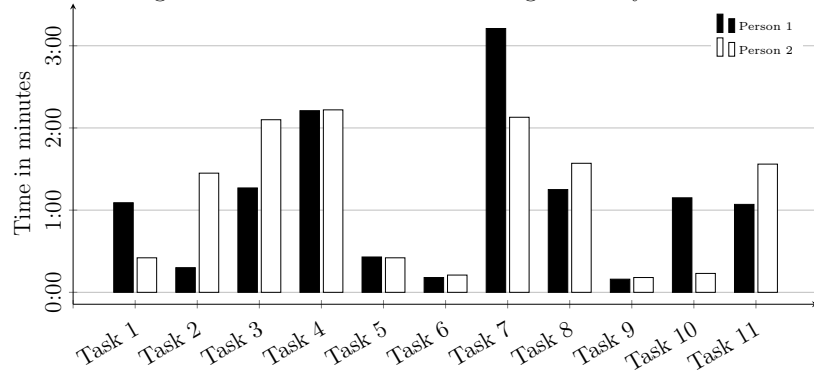
one is task 7, as it is not executable by neither of the participants. Task 8 is almost identical to 7 and can also not be executed by person B and by person A only with additional help. This task also reaches the highest total error rate with 5 errors during one execution by person B. Other ineffective tasks are 3 and 11. These cannot be executed by person A and only with assistance and many errors by person B. Task 4 is also among the tasks with a higher overall error rate and could only be executed by both participants with some assistance.

6.2 Efficiency of the Weka Tool

The efficiency is evaluated by the duration required to complete a task. The time per task is always between 0:15 and 3:20 minutes. The exact times are presented in a Figure 7. It is also clearly visible that the task duration correlates with the degree of completeness as well as the error rate. The tasks with the shortest time span, which are task six and nine, are also the ones with no help needed during execution and no measured error. The efficiency of task 7 is the worst, as the duration of its performance is more than one minute longer than any other recorded time.

One can say that the the greatest part of time per task is spent on searching and experimenting with buttons. The actual execution of the tasks requires a very short time span, as it is only some clicks away. The duration also depends on whether the participant has performed a similar functionality in a previous task or not.

Figure 7: Time for each Task during Usability Test



6.3 Satisfaction with the Weka Tool

In terms of satisfaction, the tasks are largely divided into two groups. The ones which are ranked with the two best ratings and the others which are rated with the two worst ratings. The more satisfactory tasks include Task 1, 5, 6, 9 and 10 whereas Task 3, 4, 8 and 11 score poorly. The two participants do not share the same opinion on Task 2 and Task 7. Person A rates the system after Task 2 with the highest value, whereas Person B grades the satisfaction with the the second lowest value after completing the same task. During task 7, person B is two grades more pleased with the tool than person A. The results of satisfaction after each activity are comparable to the results of of the effectiveness and efficiency testing. Unlike the results obtained for the other attributes, Task 7 was not ranked poorest. Although person A gives the worst rating, person B counteracts this by giving the second highest rating. Satisfaction after Task 5 and 6 are the two best ranked. Both received the best rating from the participants. Task 4 leads to the worst satisfaction result, with the lowest score from person A and the second lowest score from person B. An overview of the ranked scales can be found in Table 12

The overall satisfaction with the software is evaluated with a questionnaire after the execution of all tasks. The questionnaire consists of several questions which can be rated on a scale of 1 to 5, three open questions and two Yes/No questions to assess the total degree of satisfaction. In general, the ease of system use is rated with the second poorest rating. Hence the two participants consider the Weka Tool very difficult to use. Regarding the

Table 12: Satisfaction with Weka after each Task during Usability Test

Task	Easy to use			Difficult to use	
	1	2	3	4	5
Task 1		○	●		
Task 2	●				○
Task 3				●	○
Task 4				○	●
Task 5		○	●		
Task 6	●	○			
Task 7			○		●
Task 8				●	○
Task 9	●	○			
Task 10		●	○		
Task 11				●	○

clarity of organizing information within the software person A is satisfied on a mid-level, whereas person B perceives the information rather as unclear. Both participants consider Weka to be visually unattractive. Participant B highlights the possibility of recovering from mistakes as very difficult. The overall satisfaction with the Weka software is rated second worst from both respondents. In addition, they feel lost during the interaction with the system and don't believe clicking a button leads them to the expected answer. The detailed results of the ratings within the satisfaction questionnaire can be found in Table 13.

The open questions refer to the best and worst features of the software, as well as a missing features that the participants would like Weka to include and if the test person has any further comments on the tool. Person A considers the preprocessing side as the best feature of the Weka tool. The underlying reason for this is that it allows a good overview of data structure to be obtained. In contrast, person B regards the different stages outlined in tabs at the top to be the best element of the software. He appreciates each capability being well separated from each other and not having them all concentrated on a single site. As the worst feature person A mentions the procedure on how to visualize the decision tree. Stating that he does not find it intuitive. Person B on the other hand sees the way algorithms are modified as the worst aspect of Weka. The representation in the field next to

Table 13: Overall Satisfaction with Weka

Easy to use			Difficult to use	
1	2	3	4	5
			<input checked="" type="radio"/>	<input type="radio"/>
Clear Info Organization			Unclear Info Organization	
1	2	3	4	5
		<input checked="" type="radio"/>	<input type="radio"/>	
Clear Terminology			Unclear Terminology	
1	2	3	4	5
	<input checked="" type="radio"/>	<input type="radio"/>		
Visually Attractive			Visually Unattractive	
1	2	3	4	5
			<input type="radio"/>	<input checked="" type="radio"/>
Easy Error Recovery			Uneasy Error Recovery	
1	2	3	4	5
		<input checked="" type="radio"/>		<input type="radio"/>
Overall Satisfied			Overall Unsatisfied	
1	2	3	4	5
			<input checked="" type="radio"/>	<input type="radio"/>
Feeling Lost while Using				
Yes			No	
	<input checked="" type="radio"/>	<input type="radio"/>		
Expect Button to lead to correct answer				
Yes			No	
			<input checked="" type="radio"/>	<input type="radio"/>

Choose does not look like a button to him. Thus without help it is not clear that it can be clicked. The answer to the question what content or feature should be included in the tool is identical for each participant. Both note that more information and help for the different functionalities is required. Only participant B has further comments which are shared at the end of the questionnaire. He comments on the fact that the tool is principally not intuitive and that one has to first learn how to use it with help, videos or similar.

6.4 Overall Usability of the results

Summarizing our findings, there are more tasks that show difficulties during performance compared to tasks that appear to be intuitive during execution. A significant fraction (55%) of the tasks can only be performed with help or not at all. This does not indicate a high degree of usability. In addition, nine out of eleven tasks involve at least one error. The duration also indicates a low level of user-friendliness, as the tasks can be executed in just a few clicks and could therefore be completed in a few seconds. However, the time per task averages at approximately one minute. Both participants often express frustration during the tasks. This is mirrored in their satisfaction, since it is mostly below average.

Looking at the actual activities, most problems are related to the actual identification of the correct fields or buttons with which the task can be executed. For example task 1 shows that it is not clear at first that by default only ARFF files are searched for. In order to select CSV-files, it is necessary to change the datatype to be searched for in a drop down menu. The participants recall this when opening the second file in task 6, hence no problems occur during execution. Another complication during the preprocessing stage is the fact that it was unclear that there is a difference in choosing specific attributes. That is between ticking the box of each attribute and selecting the attribute by simply clicking somewhere in the row. The box is used to apply the filters and the selection is to get detailed information about the attribute and its visualization. Also, the visualization of the individual attributes could also be improved by using more annotations of what is displayed. The most challenging functions are also considered the worst features by the two respondents in the questionnaire. The design of the field with the specifications about the algorithm applied needs to be changed. It is not clear to the user that you there is the possibility to click on it to adjust the algorithm. Additionally, displaying the tested models causes confusion. It is bewildering that these can only be accessed by right clicking on an entry in the results list and are not available under the visualization tab.

6.5 Functionality of the Weka Tool

In addition to the usability test we conducted an informal interview. The focus here is on the suitability of Weka's functionalities in the context of crime data mining, leaving usability out for once. Thoughts and perspectives on this, can be summarized as follows.

The participants see Weka as beneficial for the field of strategic criminal analysis, as it offers a good possibility to discover correlations between attributes. It is also a useful instrument to answer concrete questions and validate models. In addition, the respondents see a great advantage in the fact that it does not require any programming skills. It is also beneficial that there is a desktop application that can be used without internet connection and there is no necessity to download packages before using basic functionalities. The strongest argument for the Weka tool is that it is open source and hence there are no costs involved. Despite the fact that one has to deal with the complexity of usability first, there is no need for an expensive license in order to run the corresponding methods when required.

In terms of functionality the following features are not included in the Weka Explorer but are essential for criminal analysis. It does not offer the possibility to simply extend the dataset nor to work on two datasets in parallel. It also lacks the ability to automatize operations on different records or extended data. These are essential aspects in the field of criminal analysis, as repetitive questions often occur. Furthermore, the implementation of the results in other systems is also an essential functionality. This can be extremely challenging to achieve with Weka, since the data records from the tool are only saved in ARFF format, which is a software-specific standard. Likewise saving a model result is done in a binary form which can only be read by Weka. It needs to be mentioned at this point, that the test and the results refer to the Explorer only. Some of these problems could be handled by using other GUIs or by possible extensions of the Weka tool. These aspects are not considered in this thesis. It is also worth noting that in order to use Weka in a meaningful way, there needs to be a good understanding of each data mining algorithm.

Altogether, the interview concerning the use of Weka's functionality revealed. It is suitable for at least two particular things in the area of crime analysis. First, to examine and explore data, for instance finding correlations. Second, if a particular question exists and the answer can be provided by an algorithm implemented. Also in regard to preprocessing, a concrete objective must be given in order to apply the given functions effectively. A reasonable application for using the software is the validation of models. There is no programming knowledge required in order to apply any algorithm. However, there must exist an understanding of the algorithms, it is necessary to know in advance what can be achieved with specific algorithms and where the differences lie between them.

Weka enables users to gain knowledge about data mining by offering a wide range of choices on their site⁹. The website provides a book¹⁰ that offers an introduction to the field of data mining and includes an appendix with reference to the functionality within the Weka software. They also compile several free online courses¹¹ that teach the usage of Weka. Additionally, an extra site titled WekaWiki¹² exists for providing relevant information about the software and also a Weka Blog¹³ is available in order to share unresolved questions from users.

7 Conclusion and Future Work

This chapter summarizes how we approach assessing the usability of the Weka tool within Crime Data Mining. Furthermore, additional research fields are outlined that can be conducted in this area.

7.1 Conclusion

It can be concluded that crime analysis supports operational and strategic decision making of the police. Data mining can be used to find crime patterns and trend correlations with the aim of deriving interesting information from high volume of crime data. In order to perform suitable clustering or classification techniques, the use of appropriate software is required. The Waikato Environment for Knowledge Analysis (Weka) [84] is one potential tool in this area. The department of crime analysis in the Federal Criminal Police Office of the Austrian Ministry of the Interior was interested in assessing the suitability of Weka for their day to day activities.

In order to assess the usability of the tool we conducted a usability test. The usability dimensions effectiveness, efficiency and satisfaction were adopted from ISO 9241-11 [41]. First we proposed an evaluation model with measurable attributes. The model was based on former literature in this field and was used to perform a usability assessment of the Weka tool. The participants were two employees of the department Crime Analysis. The test was conducted by performing eleven tasks that indicate a representative usage of the tool to apply crime data mining.

⁹<https://www.cs.waikato.ac.nz/ml/weka/index.html>

¹⁰<https://www.cs.waikato.ac.nz/ml/weka/book.html>

¹¹<https://www.cs.waikato.ac.nz/ml/weka/courses.html>

¹²<https://waikato.github.io/weka-wiki/>

¹³<https://waikato.github.io/weka-blog/>

The key results of the overall usability could be concluded as following. Approximately half of the tasks were executed ineffectively as well as inefficiently. The subjective opinions of the tool could also be summarized as difficult to use and visually unattractive. Both participants were rather unsatisfied during usage and the functionality of the tool was not seen as intuitive. The main problem was the design and position of fields and buttons. The participants faced great difficulties by identifying the elements needed to perform operations. Additionally, they needed more guidance and information during usage to achieve better usability.

7.2 Future Work

As this thesis only focused on assessing the usability of the *Weka Explorer*, this area can be extended by evaluating the usability of the alternative GUIs the tool includes, namely *Experimenter*, *KnowledgeFlow*, *Workbench*, *Simple CLI*. An additional area in which it is advisable to go into greater detail is the assessment of the actual functionality and performance of the software. Furthermore, we suggest to examine and compare the different algorithms included in Weka. Another interesting avenue for future work is the identification of appropriate application scenarios for the algorithms within crime data mining.

References

- [1] §129 stgb. <https://www.jusline.at/gesetz/stgb/paragraf/129>. Accessed: 2020-10-04.
- [2] Polizeiliche kriminalstatistik 2019, 2019.
- [3] Rakesh Agrawal, Tomasz Imieliński, and Arun Swami. Mining association rules between sets of items in large databases. In *Proceedings of the 1993 ACM SIGMOD international conference on Management of data*, pages 207–216, 1993.
- [4] Svetlana S Aksenova. Weka explorer tutorial. *School of Engineering and Computer Science California State University*, 2004.
- [5] Kadhim B Swadi Al-Janabi and K Haydar. Crime data analysis using data mining techniques to improve crimes prevention procedures. ICIT, 2010.
- [6] LC Amarnathan. Technological advancement: Implications for crime. *The Indian Police Journal*, 2003.
- [7] A Anitha. Prediction of crime rate using data clustering technique. In *Soft Computing for Problem Solving*, pages 443–454. Springer, 2020.
- [8] Rohit Arora. Comparative analysis of classification algorithms on different datasets using weka. *International Journal of Computer Applications*, 54(13), 2012.
- [9] Shraddha Ramdas Bandekar and C Vijayalakshmi. Design and analysis of machine learning algorithms for the reduction of crime rates in india. *Procedia Computer Science*, 172:122–127, 2020.
- [10] Carol M Barnum. *Usability testing essentials: ready, set... test!* Morgan Kaufmann, 2020.
- [11] Ronald F Becker and Aric W Dutelle. *Criminal investigation*. Jones & Bartlett Learning, 2018.
- [12] Pavel Berkhin. A survey of clustering data mining techniques. In *Grouping multidimensional data*, pages 25–71. Springer, 2006.
- [13] Nigel Bevan. International standards for usability should be more widely used. *Journal of usability studies*, 4(3):106–113, 2009.

- [14] Nigel Bevan, Jim Carter, Jonathan Earchy, Thomas Geis, and Susan Harker. New iso standards for usability, usability reports and usability measures. In *International conference on human-computer interaction*, pages 268–278. Springer, 2016.
- [15] Nitin Bhatia et al. Survey of nearest neighbor techniques. *arXiv preprint arXiv:1007.0085*, 2010.
- [16] Paul A Booth. *An introduction to human-computer interaction*. Psychology Press, 1989.
- [17] Remco R Bouckaert, Eibe Frank, Mark Hall, Richard Kirkby, Peter Reutemann, Alex Seewald, and David Scuse. Weka manual for version 3-9-1. *University of Waikato, Hamilton, New Zealand*, 2016.
- [18] Paul J Brantingham, Patricia L Brantingham, and Martin A Andresen. The geometry of crime and crime pattern theory. *Environmental criminology and crime analysis*, 2, 2017.
- [19] Tom Brinck, John Bunyan, Darren Gergle, Scott D Wood, David Blythe, and Tom McReynolds. *Designing Web sites that work: Usability for the Web*. Morgan Kaufmann, 2002.
- [20] John Brooke. Sus: a ‘quick and dirty’ usability. *Usability evaluation in industry*, page 189, 1996.
- [21] Bundeskriminalamt. Portfolio abteilung ii/bk/4 kriminalanalyse, August 2018. Not publicly available.
- [22] Anita N Chaware. Educational data mining: An emerging trends in education. *International Journal of Advanced Research in Computer Science*, 2(6), 2011.
- [23] Ming-Syan Chen, Jiawei Han, and Philip S. Yu. Data mining: an overview from a database perspective. *IEEE Transactions on Knowledge and data Engineering*, 8(6):866–883, 1996.
- [24] Edgar F Codd. *OLAP, On-Line analytical Processing mit TM/1*. Management Informations Systeme, 1994.
- [25] Thomas Cover and Peter Hart. Nearest neighbor pattern classification. *IEEE transactions on information theory*, 13(1):21–27, 1967.

- [26] H David and A Suruliandi. Survey on crime analysis and prediction using data mining techniques. *Ictact journal on soft computing*, 7(3), 2017.
- [27] SR Deshmukh, Arun S Dalvi, Tushar J Bhalerao, Ajinkya A Dahale, Rahul S Bharati, and Chaitali R Kadam. Crime investigation using data mining. *International Journal of Advanced Research in Computer and Communication Engineering*, 4(3):22–24, 2015.
- [28] Chris Drummond and Robert C Holte. Explicitly representing expected cost: An alternative to roc representation. In *Proceedings of the sixth ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 198–207, 2000.
- [29] Frank Eibe, Mark A Hall, and Ian H Witten. The weka workbench. online appendix for data mining: practical machine learning tools and techniques. In *Morgan Kaufmann*. 2016.
- [30] MN Emig, RO Heck, and M Kravitz. Crime analysis—a selected bibliography. *Washington, DC: US National Criminal Justice Reference Service*, 1980.
- [31] Martin Ester, Hans-Peter Kriegel, Jörg Sander, Xiaowei Xu, et al. A density-based algorithm for discovering clusters in large spatial databases with noise. In *Kdd*, volume 96, pages 226–231, 1996.
- [32] Usama Fayyad. Data mining and knowledge discovery in databases: implications for scientific databases. In *Proceedings. Ninth International Conference on Scientific and Statistical Database Management (Cat. No. 97TB100150)*, pages 2–11. IEEE, 1997.
- [33] Usama Fayyad, Gregory Piatetsky-Shapiro, and Padhraic Smyth. From data mining to knowledge discovery in databases. *AI magazine*, 17(3):37–37, 1996.
- [34] William J Frawley, Gregory Piatetsky-Shapiro, and Christopher J Matheus. Knowledge discovery in databases: An overview. *AI magazine*, 13(3):57–57, 1992.
- [35] Mattias Georgsson and Nancy Staggers. Quantifying usability: an evaluation of a diabetes mhealth system on effectiveness, efficiency, and satisfaction metrics with associated user characteristics. *Journal of the American Medical Informatics Association*, 23(1):5–11, 2016.

- [36] Steven Gottlieb, Sheldon I Arenberg, Raj Singh, et al. *Crime analysis: From first report to final arrest*. Alpha Publishing Montclair, CA, 1994.
- [37] Jiawei Han, Jian Pei, and Micheline Kamber. *Data mining: concepts and techniques*. Elsevier, 2011.
- [38] Jennifer Anne Harding, Muhammad Shahbaz, and A Kusiak. Data mining in manufacturing: a review. 2006.
- [39] Arne Holst. Volume of data/information created worldwide from 2010 to 2024. <https://www.statista.com/statistics/871513/worldwide-data-created/>, July 2020.
- [40] William H Inmon and Richard D Hackathorn. *Using the data warehouse*. Wiley-QED Publishing, 1994.
- [41] Iso 9241-11:1998(en) ergonomic requirements for office work with visual display terminals (vdts) — part 11: Guidance on usability,. Standard, International Organization for Standardization, 1998.
- [42] Judy Jeng. What is usability in the context of the digital library and how can it be measured? *Information technology and libraries*, 24(2):3, 2005.
- [43] Soohyung Joo, Suyu Lin, and Kun Lu. A usability evaluation model for academic library websites: efficiency, effectiveness and learnability. *Journal of Library and Information Studies*, 9(2):11–26, 2011.
- [44] K Kameshwaran and K Malarvizhi. Survey on clustering techniques in data mining. *International Journal of Computer Science and Information Technologies*, 5(2):2272–2276, 2014.
- [45] Harjot Kaur and Prince Verma. Comparative weka analysis of clustering algorithms. *International Journal of Information Technology and Computer Science*, 9(8):56–67, 2017.
- [46] Manpreet Kaur and Shivani Kang. Market basket analysis: Identify the changing trends of market data using association rule mining. *Procedia computer science*, 85(Cms):78–85, 2016.
- [47] Siddharth Kaza, Daning Hu, Homa Atabakhsh, and Hsinchun Chen. Predicting criminal relationships using multivariate survival analysis. In *Proceedings of the 8th annual international conference on Digital government research: bridging disciplines & domains*, pages 290–291, 2007.

- [48] Mohammad Reza Keyvanpour, Mostafa Javideh, and Mohammad Reza Ebrahimi. Detecting and investigating crime by means of data mining: a general crime matching framework. *Procedia Computer Science*, 3:872–880, 2011.
- [49] Richard Kirkby and Eibe Frank. Weka explorer user guide for version 3-4. *University of Weikato*, pages 3–4, 2002.
- [50] Hian Chye Koh, Gerald Tan, et al. Data mining applications in healthcare. *Journal of healthcare information management*, 19(2):65, 2011.
- [51] Anirudh Kondaveeti, Huan Liu, George Runger, and Jeremy Rowe. Extracting geographic knowledge from sensor intervention data using spatial association rules. In *Proceedings 2011 IEEE International Conference on Spatial Data Mining and Geographical Knowledge Services*, pages 127–130. IEEE, 2011.
- [52] Revatthy Krishnamurthy and J Satheesh Kumar. Survey of data mining techniques on crime data analysis. *International Journal of Data Mining Techniques and Applications*, 1(1):47–49, 2012.
- [53] KP Lee. A study on the improvement plan by analyzing user interaction pattern with the riss. *Technical Report KR2004-17*, 2004.
- [54] GNU General Public License. Gnu general public license. *Retrieved December, 25:2014*, 1989.
- [55] Rensis Likert. A technique for the measurement of attitudes. *Archives of psychology*, 1932.
- [56] A Malathi and S Santhosh Baboo. An enhanced algorithm to predict a future crime using data mining. 2011.
- [57] Zdravko Markov and Ingrid Russell. An introduction to the weka data mining system. *ACM SIGCSE Bulletin*, 38(3):367–368, 2006.
- [58] Colleen McCue. *Data mining and predictive analysis: Intelligence gathering and crime analysis*. Butterworth-Heinemann, 2014.
- [59] Karima Moumane, Ali Idri, and Alain Abran. Usability evaluation of mobile applications using iso 9241 and iso 25062 standards. *Springer-Plus*, 5(1):548, 2016.

- [60] Sreerama K Murthy. Automatic construction of decision trees from data: A multi-disciplinary survey. *Data mining and knowledge discovery*, 2(4):345–389, 1998.
- [61] Shyam Varan Nath. Crime pattern detection using data mining. In *2006 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology Workshops*, pages 41–44. IEEE, 2006.
- [62] Shyam Varan Nath. Crime data mining. In *Advances and Innovations in Systems, Computing Sciences and Software Engineering*, pages 405–409. Springer, 2007.
- [63] Giang Nguyen, Stefan Dlugolinsky, Martin Bobák, Viet Tran, Álvaro López García, Ignacio Heredia, Peter Malík, and Ladislav Hluchý. Machine learning and deep learning frameworks and libraries for large-scale data mining: a survey. *Artificial Intelligence Review*, 52(1):77–124, 2019.
- [64] Jakob Nielsen. *Usability engineering*. Morgan Kaufmann, 1993.
- [65] Kent L Norman and College Park University of Maryland. *QUIS 7.0: questionnaire for user interaction satisfaction; ersion 7.0*. University of Maryland, 1998.
- [66] Deborah Osborne and Susan Wernicke. *Introduction to crime analysis: Basic resources for criminal justice practice*. Psychology Press, 2003.
- [67] Alexei Oulanov and Edmund JY Pajarillo. Cuny+ web: usability study of the web-based gui version of the bibliographic database of the city university of new york (cuny). *The Electronic Library*, 2002.
- [68] Neelamadhab Padhy, Dr Mishra, Rasmita Panigrahi, et al. The survey of data mining applications and feature scope. *arXiv preprint arXiv:1211.5723*, 2012.
- [69] Thair Nu Phyu. Survey of classification techniques in data mining. In *Proceedings of the International MultiConference of Engineers and Computer Scientists*, volume 1, 2009.
- [70] S Prabakaran and Shilpa Mitra. Survey of analysis of crime detection techniques using data mining and machine learning. In *Journal of Physics: Conference Series*, volume 1000, page 012046. IOP Publishing, 2018.

- [71] Arun K Pujari, K Rajesh, and D Suresh Reddy. Clustering techniques in data mining—a survey. *IETE Journal of Research*, 47(1-2):19–28, 2001.
- [72] Pradeep Rai and Shubha Singh. A survey of clustering techniques. *International Journal of Computer Applications*, 7(12):1–5, 2010.
- [73] KHS Rajeswari, M Deepthi, and DND Harini. Statistical analysis of crimes concerning women in india using data mining techniques.
- [74] Shiju Sathyadevan et al. Crime analysis and prediction using data mining. In *2014 First International Conference on Networks & Soft Computing (ICNSC2014)*, pages 406–412. IEEE, 2014.
- [75] Brian Shackel and Simon J Richardson. *Human factors for informatics usability*. Cambridge university press, 1991.
- [76] Mugdha Sharma. Z-crime: A data mining tool for the detection of suspicious criminal activities based on decision tree. In *2014 International Conference on Data Mining and Intelligent Computing (ICDMIC)*, pages 1–6. IEEE, 2014.
- [77] Swasti Singhal and Monika Jena. A study on weka tool for data preprocessing, classification and clustering. *International Journal of Innovative technology and exploring engineering (IJITEE)*, 2(6):250–253, 2013.
- [78] Shweta Srivastava. Weka: a tool for data preprocessing, classification, ensemble, clustering and association rule mining. *International Journal of Computer Applications*, 88(10), 2014.
- [79] Devendra Kumar Tayal, Arti Jain, Surbhi Arora, Surbhi Agarwal, Tushar Gupta, and Nikhil Tyagi. Crime detection and criminal identification in india using data mining techniques. *AI & society*, 30(1):117–127, 2015.
- [80] Ubon Thongsatapornwatana. A survey of data mining techniques for analyzing crime patterns. In *2016 Second Asian Conference on Defence Technology (ACDT)*, pages 123–128. IEEE, 2016.
- [81] Kai Ming Ting. *Confusion Matrix*, pages 209–209. Springer US, Boston, MA, 2010.
- [82] Andrej Trnka. Market basket analysis with data mining methods. In *2010 International Conference on Networking and Information Technology*, pages 446–450. IEEE, 2010.

- [83] Ian H. Witten, Eibe Frank, Mark A. Hall, and Christopher J. Pal. Chapter 1 - whatâ™s it all about? In Ian H. Witten, Eibe Frank, Mark A. Hall, and Christopher J. Pal, editors, *Data Mining (Fourth Edition)*, pages 3 – 41. Morgan Kaufmann, fourth edition edition, 2017.
- [84] IH Witten, E Frank, MA Hall, and CJ Pal. *Data mining fourth edition: Practical machine learning tools and techniques*, 2016.
- [85] Surjeet Kumar Yadav, Brijesh Bharadwaj, and Saurabh Pal. Data mining applications: A comparative study for predicting student’s performance. *arXiv preprint arXiv:1202.4815*, 2012.
- [86] Illhoi Yoo, Patricia Alafaireet, Miroslav Marinov, Keila Pena-Hernandez, Rajitha Gopidi, Jia-Fu Chang, and Lei Hua. Data mining in healthcare and biomedicine: a survey of the literature. *Journal of medical systems*, 36(4):2431–2448, 2012.
- [87] Chung-Hsien Yu, Max W Ward, Melissa Morabito, and Wei Ding. Crime forecasting using data mining techniques. In *2011 IEEE 11th international conference on data mining workshops*, pages 779–786. IEEE, 2011.
- [88] Guangzhu Yu, Shihuang Shao, and Bing Luo. Mining crime data by using new similarity measure. In *2008 Second International Conference on Genetic and Evolutionary Computing*, pages 389–392. IEEE, 2008.
- [89] Qiankun Zhao and Sourav S Bhowmick. Association rule mining: A survey. *Nanyang Technological University, Singapore*, page 135, 2003.
- [90] Zakaria Suliman Zubi and Ayman Altaher Mahmmud. Crime data analysis using data mining techniques to improve crimes prevention. *International Journal of Computers*, 8(1):39–45, 2014.

Appendix A

Thank you very much for participating in this experiment. The objective of this test is to evaluate the usability of the Weka tool. I will ask you to perform a series of tasks. Please keep in mind that we are testing the software's usability and not your performance. In this sense, don't worry if you cannot accomplish a particular task. Since we track the time for each task, please clearly speak up when you think you have achieved the task, or you give up managing the task. Afterwards, please rank the statement given under the task from one to five regarding the ease of use of the system.

1. Open the Weka Explorer and load the CSV-file "df.clustering".

1 2 3 4 5
Easy to use Difficult to use

2. Select the attribute "House" and select "District" as "Class" and click on the button to visualize all.

1 2 3 4 5
Easy to use Difficult to use

3. Tell me in which district the most burglary into houses take place

1 2 3 4 5
Easy to use Difficult to use

4. Close the tap, go to "Cluster", choose "SimpleKMeans", change the cluster number from two to four.

1 2 3 4 5
Easy to use Difficult to use

5. Choose "District" as classes to cluster evaluation and start clustering.

1 2 3 4 5
Easy to use Difficult to use

6. Load the second CSV-file "df.classification".

1 2 3 4 5
Easy to use Difficult to use

7. Tell me if year is numeric or nominal. If necessary, change it to nominal by choosing an unsupervised attribute filter.

1 2 3 4 5
Easy to use Difficult to use

8. Change the value type of "Attempt" and "Solved" from numeric to binary.

	1	2	3	4	5
Easy to use					Difficult to use

9. Go to "Classify" and choose "J48" Tree.

	1	2	3	4	5
Easy to use					Difficult to use

10. Classify by the attribute "Solved" and start the classification.

	1	2	3	4	5
Easy to use					Difficult to use

11. Visualize the decision tree.

	1	2	3	4	5
Easy to use					Difficult to use

8. Can you recover from mistakes easily?

1 2 3 4 5
Easy Difficult

9. Your overall reaction to the system.

1 2 3 4 5
Satisfied Unsatisfied

10. Do you feel lost while using the tool?

Yes *No*

11. Is the tool easy to navigate?

Yes *No*

12. When you click a button on the tool, do you expect that the click leads to the correct answer?

Yes *No*

13. Do you have any other comments about the tool?
