

Proposing a Graph Mining Method on Medical Networks

Brayan Keritino, Quio Feri, Lionip Cheler*

Department of Computer Engineering, Open University of Kiev, Kiev, Ukraine

Abstract

Clustering algorithms can be categorized based on their cluster model, as listed above. The following overview will only list the most prominent examples of clustering algorithms, as there are possibly over 100 published clustering algorithms. Not all provide models for their clusters and can thus not easily be categorized. An overview of algorithms explained in Wikipedia can be found in the list of statistics algorithm. The clustering model most closely related to statistics is based on distribution models. Clusters can then easily be defined as objects belonging most likely to the same distribution. A convenient property of this approach is that this closely resembles the way artificial data sets are generated: by sampling random objects from a distribution. Here we proposed a distributed clustering to bag data clustering.

Keywords

Graph Mining, Big Data Mining, Fuzzy, Distributed Dataset, Medical Networks

Received: June 11, 2015 / Accepted: June 28, 2015 / Published online: July 20, 2015

© 2015 The Authors. Published by American Institute of Science. This Open Access article is under the CC BY-NC license.

<http://creativecommons.org/licenses/by-nc/4.0/>

1. Introduction

The manual extraction of patterns from data has occurred for centuries. Early methods of identifying patterns in data include Bayes' theorem (1700s) and regression analysis (1800s). The proliferation, ubiquity and increasing power of computer technology has dramatically increased data collection, storage, and manipulation ability. As data sets have grown in size and complexity, direct "hands-on" data analysis has increasingly been augmented with indirect, automated data processing, aided by other discoveries in computer science, such as neural networks, cluster analysis, genetic algorithms (1950s), decision trees and decision rules (1960s), and support vector machines (1990s). Data mining is the process of applying these methods with the intention of uncovering hidden patterns¹⁻¹⁰ in large data sets. It bridges the gap from applied statistics and artificial intelligence (which usually provide the mathematical background) to database management by exploiting the way data is stored and indexed in databases to execute the actual learning and

discovery algorithms more efficiently, allowing such methods to be applied to ever larger data sets.

The term is a misnomer, because the goal is the extraction of patterns and knowledge from large amount of data, not the extraction of data itself.¹⁻¹² It also is a buzzword¹³ and is also frequently applied to any form of large-scale data or information processing (collection, extraction, warehousing, analysis, and statistics) as well as any application of computer decision support system, including artificial intelligence, machine learning, and business intelligence. The popular book "Data mining: Practical machine learning tools and techniques with Java"¹⁴ (which covers mostly machine learning material) was originally to be named just "Practical machine learning", and the term "data mining" was only added for marketing reasons.^{15, 16} Often the more general terms "(large scale) data analysis", or "analytics" – or when referring to actual methods, artificial intelligence and machine learning-are more appropriate.

The actual data mining task is the automatic or semi-automatic analysis of large quantities of data to extract

* Corresponding author

E-mail address: liohler@outlook.com (L. Cheler)

previously unknown interesting patterns such as groups of data records (cluster analysis), unusual records (anomaly detection) and dependencies (association rule mining). This usually involves using database techniques such as spatial indices. These patterns can then be seen as a kind of summary of the input data, and may be used in further analysis or, for example, in machine learning and predictive analytics. For example, the data mining step might identify multiple groups in the data, which can then be used to obtain more accurate prediction results by a decision support system. Neither the data collection, data preparation, nor result interpretation and reporting are part of the data mining step, but do belong to the overall KDD process as additional steps.

The related terms data dredging, data fishing, and data snooping refer to the use of data mining methods to sample parts of a larger population data set that are (or may be) too small for reliable statistical inferences to be made about the validity of any patterns discovered. These methods can, however, be used in creating new hypotheses to test against the larger data populations.

For the research and evolution, the premier professional body in the field is the association for Computing Machinery's (ACM) Special Interest Group (SIG) on Knowledge Discovery and Data Mining (SIGKDD)^{11, 12} Since 1989 this ACM SIG has hosted an annual international conference and published its proceedings,¹³ and since 1999 it has published a biannual academic journal titled "SIGKDD Explorations".¹⁷⁻²²

2. Process

The Knowledge Discovery in Databases (KDD) process is commonly defined with the stages:

- (1) Selection
- (2) Pre-processing
- (3) Transformation
- (4) Data Mining
- (5) Interpretation/Evaluation.

Polls conducted in 2002, 2004, and 2007 show that the CRISP-DM methodology is the leading methodology used by data miners.^{15, 16, 17, 23-25} The only other data mining standard named in these polls was SEMMA. However, 3-4 times as many people reported using CRISP-DM.

2.1. Pre-processing

Before data mining algorithms can be used, a target data set must be assembled. As data mining can only uncover patterns actually present in the data, the target data set must be large

enough to contain these patterns while remaining concise enough to be mined within an acceptable time limit. A common source for data is a data mart or data warehouse. Pre-processing is essential to analyze the multivariate data sets before data mining. The target set is then cleaned. Data cleaning removes the observations containing noise and those with missing data.

2.2. Data Mining

Data mining involves six common classes of tasks:¹¹

Anomaly detection (Outlier/change/deviation detection) – The identification of unusual data records, that might be interesting or data errors that require further investigation.

Association rule learning (Dependency modelling) – Searches for relationships between variables. For example a supermarket might gather data on customer purchasing habits. Using association rule learning, the supermarket can determine which products are frequently bought together and use this information for marketing purposes. This is sometimes referred to as market basket analysis.

Clustering—is the task of discovering groups and structures in the data that are in some way or another "similar", without using known structures in the data.

Classification—is the task of generalizing known structure to apply to new data. For example, an e-mail program might attempt to classify an e-mail as "legitimate" or as "spam."

Regression—attempts to find a function which models the data with the least error.

Summarization—providing a more compact representation of the data set, including visualization and report generation.

2.3. Results Validation

Data mining can unintentionally be misused, and can then produce results which appear to be significant; but which do not actually predict future behavior and cannot be reproduced on a new sample of data and bear little use. Often this results from investigating too many hypotheses and not performing proper statistical hypothesis testing. A simple version of this problem in machine learning is known as overfitting, but the same problem can arise at different phases of the process and thus a train/test split - when applicable at all - may not be sufficient to prevent this from happening.

3. Clustering and Classification Methods

Clustering algorithms can be categorized based on their cluster model, as listed above. The following overview will only list the most prominent examples of clustering

algorithms, as there are possibly over 100 published clustering algorithms. Not all provide models for their clusters and can thus not easily be categorized. An overview of algorithms explained in Wikipedia can be found in the list of statistics algorithms.

There is no objectively "correct" clustering algorithm, but as it was noted, "clustering is in the eye of the beholder. The most appropriate clustering algorithm for a particular problem often needs to be chosen experimentally, unless there is a mathematical reason to prefer one cluster model over another. It should be noted that an algorithm that is designed for one kind of model has no chance on a data set that contains a radically different kind of model^{4, 5, 7} For example, k-means cannot find non-convex clusters.

3.1. Connectivity Based Clustering

Connectivity based clustering, also known as hierarchical clustering, is based on the core idea of objects being more related to nearby objects than to objects farther away. These algorithms connect "objects" to form "clusters" based on their distance. A cluster can be described largely by the maximum distance needed to connect parts of the cluster. At different distances, different clusters will form, which can be represented using a dendrogram, which explains where the common name "hierarchical clustering" comes from: these algorithms do not provide a single partitioning of the data set, but instead provide an extensive hierarchy of clusters that merge with each other at certain distances. In a dendrogram, the y-axis marks the distance at which the clusters merge, while the objects are placed along the x-axis such that the clusters don't mix.

These methods will not produce a unique partitioning of the data set, but a hierarchy from which the user still needs to choose appropriate clusters. They are not very robust towards outliers, which will either show up as additional clusters or even cause other clusters to merge (known as "chaining phenomenon", in particular with single-linkage clustering). In the general case, the complexity is which makes them too slow for large data sets. For some special cases, optimal efficient methods (of complexity) are known: SLINK²⁴ for single-linkage and CLINK for complete-linkage clustering. In the data mining community these methods are recognized as a theoretical foundation of cluster analysis, but often considered obsolete. They did however provide inspiration for many later methods such as density based clustering (figure 1).

20 clusters extracted, most of which contain single elements, and since linkage clustering does not have a notion of "noise". Many people used this method on the different application.

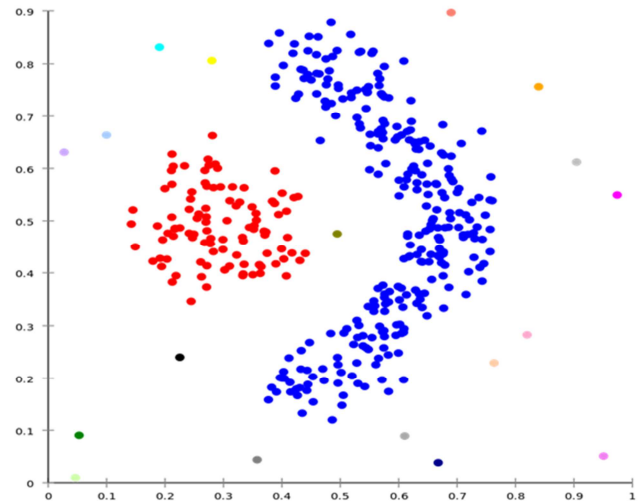


Figure 1. Single-linkage on density-based clusters.

3.2. Centroid-Based Clustering

In centroid-based clustering, clusters are represented by a central vector, which may not necessarily be a member of the data set. When the number of clusters is fixed to k , k-means clustering gives a formal definition as an optimization problem: find the k cluster centers and assign the objects to the nearest cluster center, such that the squared distances from the cluster are minimized.

The optimization problem itself is known to be NP-hard, and thus the common approach is to search only for approximate solutions. A particularly well known approximative method is Lloyd's algorithm⁷ often actually referred to as "k-means algorithm". It does however only find a local optimum, and is commonly run multiple times with different random initializations. Variations of k-means often include such optimizations as choosing the best of multiple runs, but also restricting the centroids to members of the data set (k-medoids), choosing medians (k-medians clustering), choosing the initial centers less randomly (K-means++) or allowing a fuzzy cluster assignment (Fuzzy c-mean).

Most k-means-type algorithms require the number of clusters - k - to be specified in advance, which is considered to be one of the biggest drawbacks of these algorithms. Furthermore, the algorithms prefer clusters of approximately similar size, as they will always assign an object to the nearest centroid. This often leads to incorrectly cut borders in between of clusters (which is not surprising, as the algorithm optimized cluster centers, not cluster borders).

K-means has a number of interesting theoretical properties. On the one hand, it partitions the data space into a structure known as a Voronoi diagram. On the other hand, it is conceptually close to nearest neighbor classification, and as such is popular in machine learning. Third, it can be seen as a

variation of model based classification, and Lloyd's algorithm as a variation of the Expectation-maximization algorithm for this model discussed below (Figure 2).

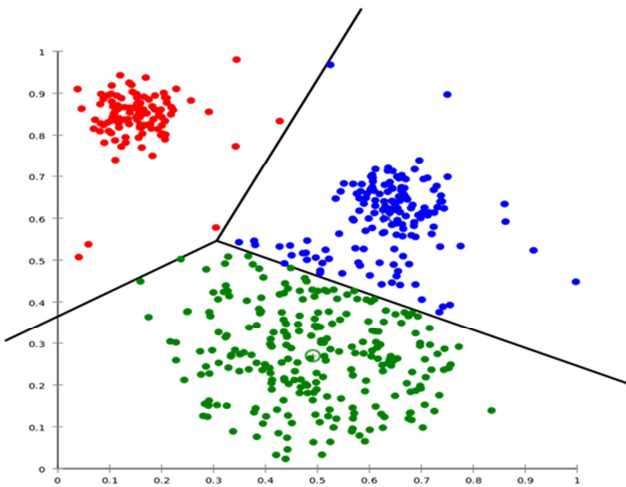


Figure 2. Means separates data into Voronoi-cells, which assumes equal-sized clusters.

3.3. Distribution-Based Clustering

The clustering model most closely related to statistics is based on distribution models. Clusters can then easily be defined as objects belonging most likely to the same distribution. A convenient property of this approach is that this closely resembles the way artificial data sets are generated: by sampling random objects from a distribution.

While the theoretical foundation of these methods is excellent, they suffer from one key problem known as overfitting, unless constraints are put on the model complexity. A more complex model will usually be able to explain the data better, which makes choosing the appropriate model complexity inherently difficult.

One prominent method is known as Gaussian mixture models (using the expectation-maximization algorithm). Here, the data set is usually modelled with a fixed (to avoid overfitting) number of Gaussian distributions that are initialized randomly and whose parameters are iteratively optimized to fit better to the data set. This will converge to a local optimum, so multiple runs may produce different results. In order to obtain a hard clustering, objects are often then assigned to the Gaussian distribution they most likely belong to; for soft clusterings, this is not necessary.

Distribution-based clustering produces complex models for clusters that can capture correlation and dependence between attributes. However, these algorithms put an extra burden on the user: for many real data sets, there may be no concisely defined mathematical model (figure 3). Many people used this method on the different application.

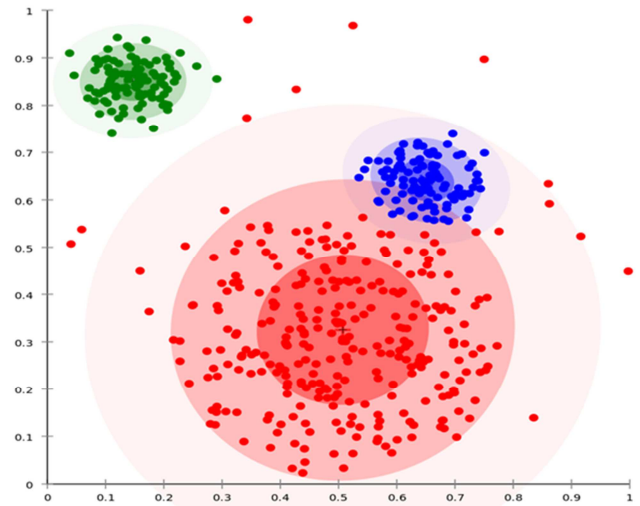


Figure 3. On Gaussian-distributed data, EM works well, since it uses Gaussians for modelling clusters.

3.4. Density-Based Clustering

In density-based clustering⁸ clusters are defined as areas of higher density than the remainder of the data set. Objects in these sparse areas - that are required to separate clusters - are usually considered to be noise and border points.

The most popular⁹ density based clustering method is DBSCAN.¹⁰ In contrast to many newer methods, it features a well-defined cluster model called "density-reachability". Similar to linkage based clustering, it is based on connecting points within certain distance thresholds. However, it only connects points that satisfy a density criterion, in the original variant defined as a minimum number of other objects within this radius. A cluster consists of all density-connected objects (which can form a cluster of an arbitrary shape, in contrast to many other methods) plus all objects that are within these objects' range. Another interesting property of DBSCAN is that its complexity is fairly low - it requires a linear number of range queries on the database - and that it will discover essentially the same results (it is deterministic for core and noise points, but not for border points) in each run, therefore there is no need to run it multiple times. OPTICS^{11, 21, 22,23,24} is a generalization of DBSCAN that removes the need to choose an appropriate value for the range parameter ϵ , and produces a hierarchical result related to that of linkage clustering. DeLi-Clu¹² Density-Link-Clustering combines ideas from single-linkage clustering and OPTICS, eliminating the ϵ parameter entirely and offering performance improvements over OPTICS by using an R-tree index.

The key drawback of DBSCAN and OPTICS is that they expect some kind of density drop to detect cluster borders. Moreover, they cannot detect intrinsic cluster structures which are prevalent in the majority of real life data. A

variation of DBSCAN, EnDBSCAN¹³ efficiently detects such kinds of structures. On data sets with, for example, overlapping Gaussian distributions - a common use case in artificial data - the cluster borders produced by these algorithms will often look arbitrary, because the cluster density decreases continuously. On a data set consisting of mixtures of Gaussians, these algorithms are nearly always outperformed by methods such as EM clustering that are able to precisely model this kind of data.

Mean-shift is a clustering approach where each object is moved to the densest area in its vicinity, based on kernel density estimation. Eventually, objects converge to local maxima of density. Similar to k-means clustering, these "density attractors" can serve as representatives for the data set, but mean-shift can detect arbitrary-shaped clusters similar to DBSCAN. Due to the expensive iterative procedure and density estimation, mean-shift is usually slower than DBSCAN or k-Means (figure4).

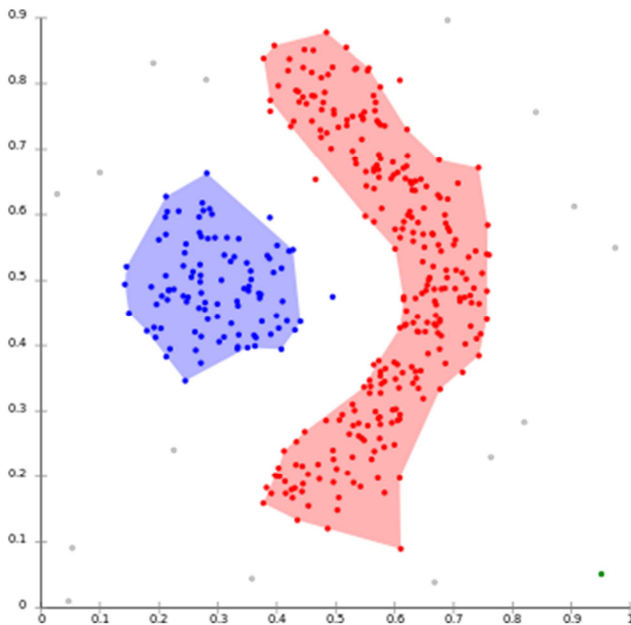


Figure 4. Density-based clustering with DBSCAN.

4. Applications

4.1. Business

In business, data mining is the analysis of historical business activities, stored as static data in data warehouse databases. The goal is to reveal hidden patterns and trends. Data mining software uses advanced pattern recognition algorithms to sift through large amounts of data to assist in discovering previously unknown strategic business information. Examples of what businesses use data mining for include performing market analysis to identify new product bundles, finding the root cause of manufacturing problems, to prevent customer

attrition and acquire new customers, cross-sell to existing customers, and profile customers with more accuracy²².

4.2. Science and Engineering

In recent years, data mining has been used widely in the areas of science and engineering, such as bioinformatics, genetics, medicine, education and electrical power engineering.

In the study of human genetics, sequence mining helps address the important goal of understanding the mapping relationship between the inter-individual variations in human DNA sequence and the variability in disease susceptibility. In simple terms, it aims to find out how the changes in an individual's DNA sequence affects the risks of developing common diseases such as cancer, which is of great importance to improving methods of diagnosing, preventing, and treating these diseases. One data mining method that is used to perform this task is known as multifactor dimensionality reduction.¹³

4.3. Human Rights

Data mining of government records – particularly records of the justice system (i.e., courts, prisons) – enables the discovery of systemic human rights violations in connection to generation and publication of invalid or fraudulent legal records by various government agencies.^{12, 13, 14, 15}

4.4. Medical Data Mining

In 2011, the case of *Sorrell v. IMS Health, Inc.*, decided by the Supreme Court of the United States, ruled that pharmacies may share information with outside companies. This practice was authorized under the 1st Amendment of the Constitution, protecting the "freedom of speech." However, the passage of the Health Information Technology for Economic and Clinical Health Act (HITECH Act) helped to initiate the adoption of the electronic health record (EHR) and supporting technology in the United States.¹⁵ The HITECH Act was signed into law on February 17, 2009 as part of the American Recovery and Reinvestment Act (ARRA) and helped to open the door to medical data mining.¹⁶ Prior to the signing of this law, estimates of only 20% of United States based physician were utilizing electronic patient records.¹⁵ Søren Brunak notes that "the patient record becomes as information-rich as possible" and thereby "maximizes the data mining opportunities."¹⁵ Hence, electronic patient records further expands the possibilities regarding medical data mining thereby opening the door to a vast source of medical data analysis.

4.5. Spatial Data Mining

Spatial data mining is the application of data mining methods to spatial data. The end objective of spatial data mining is to

find patterns in data with respect to geography. So far, data mining and Geographic Information Systems (GIS) have existed as two separate technologies, each with its own methods, traditions, and approaches to visualization and data analysis. Particularly, most contemporary GIS have only very basic spatial analysis functionality. The immense explosion in geographically referenced data occasioned by developments in IT, digital mapping, remote sensing, and the global diffusion of GIS emphasizes the importance of developing data-driven inductive approaches to geographical analysis and modeling.¹¹⁻¹⁹

Data mining offers great potential benefits for GIS-based applied decision-making. Recently, the task of integrating these two technologies has become of critical importance, especially as various public and private sector organizations possessing huge databases with thematic and geographically referenced data begin to realize the huge potential of the information contained therein. Among those organizations are:

- offices requiring analysis or dissemination of geo-referenced statistical data
- public health services searching for explanations of disease clustering
- environmental agencies assessing the impact of changing land-use patterns on climate change
- geo-marketing companies doing customer segmentation based on spatial location.

4.6. Sensor Data Mining

Wireless sensor networks can be used for facilitating the collection of data for spatial data mining for a variety of applications such as air pollution monitoring.¹⁰ A characteristic of such networks is that nearby sensor nodes monitoring an environmental feature typically register similar values. This kind of data redundancy due to the spatial correlation between sensor observations inspires the techniques for in-network data aggregation and mining. By measuring the spatial correlation between data sampled by different sensors, a wide class of specialized algorithms can be developed to develop more efficient spatial data mining algorithms.^{11, 18}

4.7. Visual Data Mining

In the process of turning from analogical into digital, large data sets have been generated, collected, and stored discovering statistical patterns, trends and information which is hidden in data, in order to build predictive patterns. Studies suggest visual data mining is faster and much more intuitive than is traditional data mining.^{52, 53, 54} See also Computer vision.

4.8. Music Data Mining

Data mining techniques, and in particular co-occurrence analysis, has been used to discover relevant similarities among music corpora (radio lists, CD databases) for purposes including classifying music into genres in a more objective manner.⁵

4.9. Surveillance

Data mining has been used by the U.S. government. Programs include the Total Information Awareness (TIA) program, Secure Flight (formerly known as Computer-Assisted Passenger Prescreening System (CAPPS II)), Analysis, Dissemination, Visualization, Insight, Semantic Enhancement (ADVISE),⁶ and the Multi-state Anti-Terrorism Information Exchange (MATRIX).⁵⁷ These programs have been discontinued due to controversy over whether they violate the 4th Amendment to the United States Constitution, although many programs that were formed under them continue to be funded by different organizations or under different names.⁷

In the context of combating terrorism, two particularly plausible methods of data mining are "pattern mining" and "subject-based data mining."

4.10. Pattern Mining

"Pattern mining" is a data mining method that involves finding existing patterns in data. In this context patterns often means association rules. The original motivation for searching association rules came from the desire to analyze supermarket transaction data, that is, to examine customer behavior in terms of the purchased products. For example, an association rule "beer \Rightarrow potato chips (80%)" states that four out of five customers that bought beer also bought potato chips.

In the context of pattern mining as a tool to identify terrorist activity, the National Research Council provides the following definition: "Pattern-based data mining looks for patterns (including anomalous data patterns) that might be associated with terrorist activity — these patterns might be regarded as small signals in a large ocean of noise."^{9, 11} Pattern Mining includes new areas such a Music Information Retrieval (MIR) where patterns seen both in the temporal and non-temporal domains are imported to classical knowledge discovery search methods.

4.11. Subject-Based Data Mining

"Subject-based data mining" is a data mining method involving the search for associations between individuals in data. In the context of combating terrorism, the National Research Council provides the following definition:

"Subject-based data mining uses an initiating individual or other datum that is considered, based on other information, to be of high interest, and the goal is to determine what other persons or financial transactions or movements, etc., are related to that initiating datum."²⁶

4.12. Knowledge Grid

Knowledge discovery "On the Grid" generally refers to conducting knowledge discovery in an open environment using grid computing concepts, allowing users to integrate data from various online data sources, as well make use of remote resources, for executing their data mining tasks. The earliest example was the Discovery Net,^{2,3} developed at Imperial College London, which won the "Most Innovative Data-Intensive Application Award" at the ACM SC02 (Supercomputing 2002) conference and exhibition, based on a demonstration of a fully interactive distributed knowledge discovery application for a bioinformatics application. Other examples include work conducted by researchers at the University of Calabria, who developed a Knowledge Grid architecture for distributed knowledge discovery, based on grid computing^{4,5}.

5. Privacy Concerns and Ethics

While the term "data mining" itself has no ethical implications, it is often associated with the mining of information in relation to peoples' behavior (ethical and otherwise).

The ways in which data mining can be used can in some cases and contexts raise questions regarding privacy, legality, and ethics. In particular, data mining government or commercial data sets for national security or law enforcement purposes, such as in the Total Information Awareness Program or in ADVISE, has raised privacy concerns.

Data mining requires data preparation which can uncover information or patterns which may compromise confidentiality and privacy obligations. A common way for this to occur is through data aggregation. Data aggregation involves combining data together (possibly from various sources) in a way that facilitates analysis (but that also might make identification of private, individual-level data deducible or otherwise apparent). This is not data mining per se, but a result of the preparation of data before – and for the purposes of – the analysis. The threat to an individual's privacy comes into play when the data, once compiled, cause the data miner, or anyone who has access to the newly compiled data set, to be able to identify specific individuals, especially when the data were originally anonymous.

It is recommended that an individual is made aware of the following before data are collected:

- The purpose of the data collection and any (known) data mining projects.
- How the data will be used?
- Who will be able to mine the data and use the data and their derivatives?
- The status of security surrounding access to the data.
- How collected data can be updated?

6. Conclusion

Generally, data mining (sometimes called data or knowledge discovery) is the process of analyzing data from different perspectives and summarizing it into useful information - information that can be used to increase revenue, cuts costs, or both. Data mining software is one of a number of analytical tools for analyzing data. It allows users to analyze data from many different dimensions or angles, categorize it, and summarize the relationships identified. Technically, data mining is the process of finding correlations or patterns among dozens of fields in large relational databases.

References

- [1] L. Hubert, Comparing partitions. *Journal of Classification*, 2, 123-154 (1985).
- [2] A. Bewley, Real-time volume estimation of a dragline payload, *IEEE International Conference on Robotics and Automation, UK*, 1571-1576 (2011).
- [3] H. Alinejad-Rokny, A Method to Avoid Gapped Sequential Patterns in Biological Sequences: Case Study: HIV and Cancer Sequences, *Journal of Neuroscience and Neuroengineering*, 4(1), (2015).
- [4] Z. Deng, and Z. Wang, A New Algorithm for Fast Mining Frequent Itemsets Using N-Lists, *Science China Information Sciences*, 55, 2008-2030 (2012).
- [5] Z. Deng, and Z. Wang, A New Fast Vertical Method for Mining Frequent Patterns, *International Journal of Computational Intelligence Systems*, 3, 733-744 (2010).
- [6] G. Webb, OPUS: An Efficient Admissible Algorithm for Unordered Search, *Journal of Artificial Intelligence Research*, 4, 531-565 (1995).
- [7] Hasanzadeh E., Poyan M., Alinejad-Rokny H. (2012). "Text Clustering on Latent Semantic Indexing with Particle Swarm Optimization (PSO) Algorithm", *International Journal of the Physical Sciences*, 7(1): 116-120.
- [8] D. Bayardo, Constraint-based rule mining in large, dense databases, *Data Mining and Knowledge Discovery*, 4, 217-240 (2013).
- [9] H. Alinejad-Rokny, Proposing on Optimized Homolographic Motif Mining Strategy based on Parallel Computing for Complex Biological Networks, *Journal of Medical Imaging and Health Informatics*, 7, (2015).

- [10] F. Salieb-Aouissi, QuantMiner: A Genetic Algorithm for Mining Quantitative Association Rules, *International Joint Conference on Artificial Intelligence, Paris*, 1035-1040 (2011).
- [11] Esmaeili L., Minaei-Bidgoli B., Alinejad-Rokny H. (2012). "A Hybrid Recommender System for Joining Virtual Communities", *Research Journal of Applied Sciences Engineering and Technology*, 4(5): 500-509.
- [12] S. Azevedo, and Santos, a parallel overview, *In Proceedings of the IADIS European Conference on Data Mining, UK*, 182-185 (2008).
- [13] A. Braha, and T. Elovici, Theory of actionable data mining with application to semiconductor manufacturing control, *International Journal of Production Research*, 47, 234-254 (2009).
- [14] D. Braha, and A. Shmilovici, On the Use of Decision Tree Induction for Discovery of Interactions in a Photolithographic Process, *IEEE Transactions on Semiconductor Manufacturing*, 16, 456-476 (2006).
- [15] H.A. Rowley, S. Baluja, and T. Kanade, Neural network-based face detection, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20, 29-38 (1998).
- [16] Parvin H., Alinejad-Rokny H., Parvin S. (2013). A Classifier Ensemble of Binary Classifier Ensembles, *International Journal of Learning Management Systems*, 1(2): 37-47.
- [17] H. Alinejad-Rokny, H. Pourshaban, A. Goran, M. MirnabiBaboli, "Network Motifs Detection Strategies and Using for Bioinformatic Networks". *Journal of Bionanoscience* 8(5), 353-359 (2014).
- [18] S.Z. Li, L. Zhu, Z.Q. Zhang, A. Blake, H.J. Zhang, and H. Shum, Statistical learning of multi-view face detection, *European Conference on Computer Vision, Paris*, 67-81 (2002).
- [19] Motameni H., Alinejad-Rokny H., Pedram M.M. (2011). "Using Sequential Pattern Mining in Discovery DNA Sequences Contain Gap", *American Journal of Scientific Research*, 14, 72-78.
- [20] Alinejad-Rokny H., Pourshaban H., Goran A., MirnabiBaboli M. (2014). "Network Motifs Detection Strategies and Using for Bioinformatic Networks", *Journal of Bionanoscience*, 8(5): 353-359.
- [21] Alinejad Rokny H., Pedram M.M., Shirgahi H. (2011). "Discovered motifs with using parallel Mprefixspan method", *Scientific Research and Essays*, 6(20): 4220-4226.
- [22] X.R. Chen, and A. Yuille, Detecting and reading text in natural scenes, *IEEE International Conference on Computer Vision and Pattern Recognition, US*, 366-373 (2004).
- [23] MirnabiBaboli M., RashidiBajgan S., Alinejad-Rokny H. (2015). "The Enhancement of Network Lifetime by Mobile Node in OPP Routing of Wireless Sensor Network", *Information Sciences Letters*, 4(1): 1-6.
- [24] K.I. Kim, K. Jung, and J.H. Kim, Texture-based approach for text detection in images using support vector machines and continuously adaptive mean shift algorithm, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(12), 1631-1639 (2003).
- [25] H. Alinejad Rokny, M.M. Pedram, H. Shirgahi, Discovered motifs with using parallel Mprefixspan method, *Scientific Research and Essays*, 6(20), 4220-4226 (2011).
- [26] M. Hadid, K. Pietika, and T. Ahonen, A discriminative feature space for detecting and recognizing faces, *Conference on Computer Vision and Pattern Recognition UK*, 12(3). 797-804 (2004).