A General View of Big Data and Machine Learning

Özhan Görçün and Hande Küçükönder

Abstract Nowadays, winds of digitalization have continuing blow in the world. No one cannot ignore these kinds of changes and is continuing to cause dramatic changes in our lives. Big data is one of the crucial parts of these changes. Nowadays, almost millions of people have become the main element providing the data flow. Furthermore, they are continuing to perform it voluntarily and willingly. As a result, companies, industries, and other stakeholders have started to encounter a very huge volume of data about their customers. However, most of the data are not be useful because they are not structured. This chapter is organized into two parts. In the first part, the big data concept was explained and its important elements were discussed. In the second stage, machine learning and its elements were presented information on some computational tools used for converting the unstructured data to structured data.

Keywords Big data · Machine learning · ANN · Genetic algorithms · Decision trees

4.1 Introduction: Big Data

Big Data is one of the most important parts of digitalized world and industry 4.0 process. In recent decades it has been crucial not only for industries but also individual users. Big data which has many components such as users, cloud systems, data processors, and so on is a term defining a system hosting a very huge volume of structured and unstructured data. Although it can host a huge volume of data, the most important issue is the quality of these data since most of the unstructured data may not be useful for industries and users. Therefore, converting these kinds of data

Ö. Görçün (✉)
Gelişim University, Istanbul, Turkey

H. Küçükönder
Department of Numerical Methods, Faculty of Economics and Administration, Bartin University, Bartın, Turkey
e-mail: hkucukonder@bartin.edu.tr

İ. İyigün and Ö. F. Görçün (eds.), *Logistics 4.0 and Future of Supply Chains*,
Accounting, Finance, Sustainability, Governance & Fraud: Theory and Application,
https://doi.org/10.1007/978-981-16-5644-6_4

to structured data is a very crucial task for experts who are responsible to develop these systems as it is required for forecasting and evaluating when decision-makers take strategic decisions.

Nowadays, using big data has become very important for industries, companies, and other stakeholders to survive in a highly competitive environment. These actors want to know how the customers are thinking as well as what is important for them. As a very useful tool, big data has the potential to provide wider information about customers even their special days and daily activities, pleasures, and so on. Moreover, there are many users who provide data both about themselves and about many things for this system and it continues to increases each passing day at the present. Therefore, the data volume which flows to the system continuing to ever-increased. Today, it is possible to see the elements of the information society in all areas of life. Most people now have a smartphone in their pocket, most people have a computer at home and units that manage information technology in the back offices of all companies. However, the information itself is not so visible. However, only half a century after computers entered human life, the amount of information began to be collected in such a way as to gain a meaningful and special quality. Today, not only the amount of information has increased, but also the speed of access to information has increased. The quantitative change brought with its qualitative change. The collection of data to form a meaningful whole took place first in the field of astronomy and genetics. The concept of big data was first used in these areas, and then this concept started to be used for each area. Big data has begun to show itself in all areas of our lives. For example; an Internet search engine from Google, we find big data in every area from diagnosis and treatment of diseases to online shopping. (Mayer-et al. 2013).

According to some reports on this issue, 90% of the data used in the world has been created only in the last two years, as the information has rapidly become obsolete and technology has changed (Url1,2020). On the other hand, there are some disadvantages of big data. It is strongly expected that consisting of very large and complex datasets that are difficult to process using existing database management tools or traditional data processing applications; it causes difficulties in data acquisition, improvement, storage, searching, sharing, transfer, analysis, and visualization stages (Yılmaz et al. 2017).

As a result, although some disadvantages, Big data is started use in various fields of the economy in recent years. According to Sağıroğlu and Koç (2017), it is used in almost all industries from health to logistics. Big data consists of five components defined as 5 V and the 5 V are presented in Fig. 4.1 (Url 2,2020).

4.1.1 Volume

The volume of data is directly proportional to the size of the data and is generally expressed in various data measurement units such as gigabytes, terabytes, and bytes. as well as the quantity of generated and stored data, the size of the data determines the value and potential insight, and whether it can be considered big data or not.

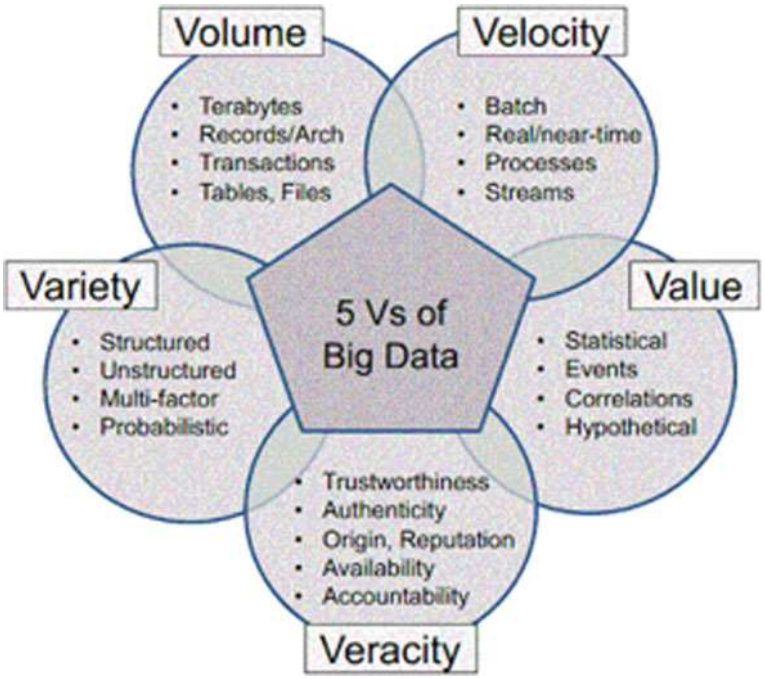

Fig. 4.1 5 V of big data system (Url 2 2020)

However, the storage of high amounts of data is among the issues that need attention. The said increases, new data processing, storage, integration, and archiving makes use of technologies necessary. But new like cloud technology with the proliferation of technologies, the storage of data can be easily and data can be used at any time (Dülger 2015).

4.1.2 Variety

The type and nature of the data help people who analyze it to effectively use the resulting insight. Not only from text, but Big data can also be drawn from images, audio, video, experts, and sources related to this issue state that the data may be either structured or unstructured or semi-structured and unfortunately, these three data types have the possibility of being replaced and only 20 percent of them are structured (Dülger 2015). The diversity of data is due to the diversity of data sources. The produced data can be seen in structural, semi-structural, or non-structural formats. Data in relational databases to structural data, semi-structural data can be given as XML, and JSON format data as non-structural data, as well as audio, video, text files. If the data is used together in these different structures, this diversity in the

data emerges as a big problem. Many new big data specific to this data, especially in the convert-load (ETL) operations of the data technologies are becoming imperative (Sağıroğlu and Koç 2017).

4.1.3 Velocity

The data is in constant motion. In this context, the analysis of data flow has become an important issue for data scientists (Cyganek et al. 2016). Data can be produced at variable speeds. While large volumes of static data create a big data problem, especially phone or IoT devices that we frequently use in daily life, sensor data produced by various machines and similar data sources produce data very quickly. The fact that the data produced in a flow can be analyzed and managed in real-time is another problem of big data. (Sağıroğlu and Koç 2017). Therefore, data that can be produced faster increases the processing speed of the data where it is needed and makes significant contributions to the diversity of data.

4.1.4 Value

Analysis of the developments and innovations that may occur in the coming period, preparing for the changes that will occur, and making decisions regarding the results will be possible thanks to the value provided by the big data. Data that does not produce value has no value. Data is as valuable as the value it creates. (Venkatram and Geetha, 2017). Data that is not quickly grasped is only trash in the real business world. If they serve no purpose, they are digital waste (Spann 2017).

4.1.5 Validity

Analyzes to be made on data that cannot be confirmed for accuracy reduce the real value. There are technologies that can measure the validity of big data and the accuracy of the data. Existing data to the above properties if it has, this data is traditional value it is very difficult to produce. So, this data has a big data perspective the need to be handled and analyzed with big data technologies has arisen (Sağıroğlu and Koç 2017).

4.2 Big Data and Machine Learning

As mentioned above, without data processing processes, big data is a non-functional system. Converting data from unstructured to structured is a crucial and critical task for being able to make comprehensive analyses. For this purpose, there are many tools used to obtain meaningful information and data. Furthermore, these techniques, i.e., Artificial Neural Networks, Genetic Algorithms, and so on are ever developed each passing day, and new techniques are presented to the industries and scientific world.

4.2.1 Machine Learning

Computer technology developing in recent years has brought along collecting and processing the very huge and various types of data. As a result, big data has been very crucial for a new technological world. By contrary, the very important part of data which flow into the systems is not structured data and decision-makers who are in all side of lives such as business, public authorities and so on need to structured data to make right and suitable decisions. Therefore, they require a systematic and applicable tool for obtaining meaningful and applicable data. Based on these requirements the machine learning has become a very strong and popular research area.

In 1959, machine learning suggested by Arthur Samuel is one of the parts of artificial intelligence (Wei et al. 2019). Machine learning has a very crucial role in solving very complicated problems includes modeling and computer-based algorithms which can self-learning based on information, can make self-assessment, and embracing the patterns (Atalay and Çelik 2017). The machine learning consists of two parts. First is to solve a problem with the help of an algorithm and self-learning the model, making meaningful evaluations based on this self-learning (Şahinarslan 2019). Within this perspective, in order to solve problems, which require some functions such as forecasting, classification, and grouping, using some novel machine learning techniques (i.e., Decision trees, ANN, Bayesian networks, genetic algorithms, k the nearest neighbor algorithm (KNN)) have become common techniques as alternatives to the traditional techniques. The main procedure of the machine learning which has very strong relations with some fields such as artificial intelligence, data mining, pattern recognition, and bioinformatics can be demonstrated in six steps as shown in Fig. 4.2 (Wei et al. 2019; Zhang et al. 2020).

As is seen in Fig. 4.2, the first step of the process starts with collecting appropriate data from databases. In this step, the selection of the proper dataset is crucial and critical task for decision-makers considering the type of data, data size, quality, and form (Wei et al. 2019). Therefore, it is required to data transfer from appropriate databases by users for using more qualified data. In addition to that, the form of data can be prepared in different types considering the algorithm used in this process. The next step called future engineering includes operations of extrapolating the suitable features from raw data. Also, remain step of this process consists of some

Fig. 4.2 The main procedure of machine learning approaches (Wei et al. 2019)

applications such as the selection of the proper algorithm, forming test and validation sets, modeling, and examination of validation of forecasting (Wei et al. 2019; Zhang et al. 2020). Some fields such as engineering approaches, cognitive simulation, and theoretical analysis (Şahan 2020) have been effective in developing and forming process of the machine learning which uses in many fields (i.e., medical diagnosis, education, biology, engineering, and so on.).

4.2.2 Artificial Neural Network (ANN)

Artificial Neural Network (ANN) is related to a biological neural structure of humans and it is a mathematical model which is developed by inspiring the working principles of the human brain and is effective for examining the complicate and nonlinear relations (Mjalli et al. 2007). There are four different components such as dendrite, soma, axon, and snaps. When their functions are examined briefly, snaps are spaces providing transmissions electric signals among nerve cells. Soma is the section that is to collect and process these signals. By processing these signals, nerve cells form their own signals and these formed signals are transmitted by axons to the dendrites having a view as tree branches. Also, dendrites have functions to transmit these signals to the snaps and data transfer between two cells can be provided by neurotransmitters which are in snaps (Öztemel 2003).

Artificial neural cells (process) exist in the artificial neural networks (ANN) similar to the biological nerve cells. Each process element includes five different elements such as inputs, weights, the sum function, activation function, and outputs (Aladağ 2019). Inputs collect information and data from outside and the significance and impacts of the information are determined with the help of weights. The sum function provides to compute the net input value. Using the weighting sum is a common approach for purposing to calculate the net input transmitted to the network (each input multiplied by their own weight and calculate the sum of these values)

the function of activation symbolizes the functions providing to produce outputs by processing the inputs (Öztemel 2003).

Artificial Neural Network (ANN) formed many artificial neural cells, consists of many processing elements connected parallelly to each other (Maind and Wankar 2014). There are different three layers such as the input layer, interlayers, and output layers in a neural network whose processor is an artificial neuron. While the section that entering data related to the problem is the input layer, the section processing these data is a hidden layer. The section that transforms the processing information to the outputs is defined as the output layer (Ocak and Şeker 2013; Özşahin and Singer 2019). In Fig. 4.3, an ANN example, which shows these layers as is presented as diagrammatic (Maind and Wankar 2014).

ANN is a very useful tool applied to solve very complicated and nonlinear problems. It has the abilities of self-learning and generalization by considering the data and information without any theoretical assumptions or a rule base (Mjalli et al. 2007). Also, determining the fitting function is not required and it can be accepted as another advantage of the ANN. While some methods used for solving the problem provide a range of parameters that are useful for understanding and evaluating the problem, ANN cannot assess the weights of the connection for the present. Because of that, modeling of ANN is defined as a closed black box (Mjalli et al. 2007; Maind and Wankar 2014).

4.2.3 Genetic Algorithms (GA)

Genetic Algorithms (GA) is a search technique based on natural selection principles and transferring the biological and genetic development processes of lives to the computer environment (Nabiyev 2012). GA described under the mainframe of evolutional algorithms was developed by inspiring the evolution theory of Darwin. Genetic Algorithms, which was firstly introduced by J.H. Holland who are a phycologist and expert of computer sciences, is an intuitional and randomized method (Emel and Taşkın 2002; Nabiyev 2012).

GA can provide effective solutions to real-life problems has been commonly used in the literature for solving the various combinatorial optimization problems. These works can be summarized as vehicle routing (Aydemir and Karagül 2020; Ghannadpour and Zandiyeh 2020), project scheduling (Özköse and Gencer 2019) transportation (Yakıcı Ayan 2008), traveling-salesman problem (Şahin and Karagül 2019), location problems (Gülsün et al. 2009). GA has some differences than other optimization algorithms and these differences can be described as (İşçi and Korukoğlu 2003): GA deal with coded parameters set instead of single parameter and it performs to start making a search with a group of points instead of a single point. Another difference is to focus on the target function directly and it uses the stochastic passage rules. An example of a basic search procedure of the GA is presented in Fig. 4.4 (Gholamia et al. 2014).

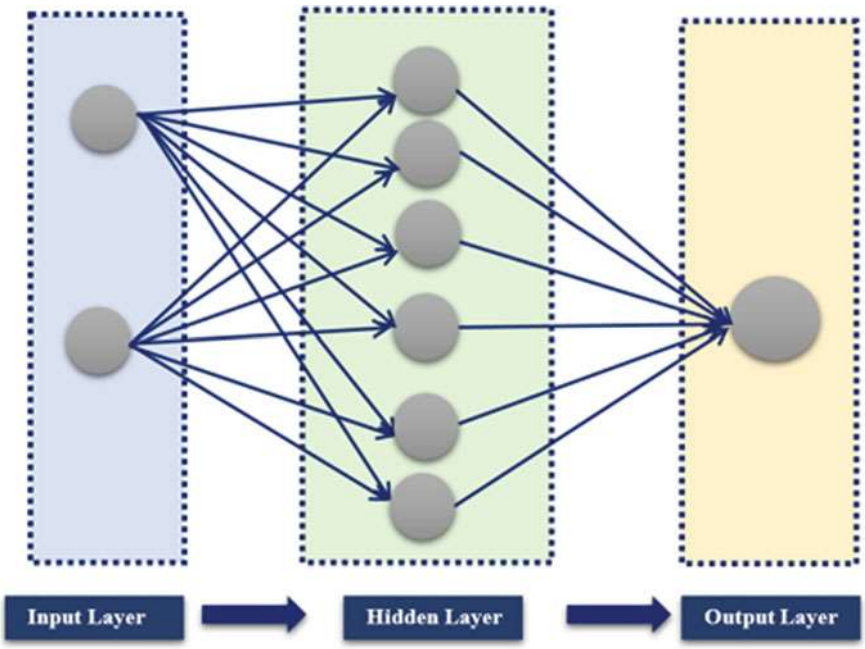

Fig. 4.3 An example of ANN (Maind and Wankar 2014)

When the operational procedures of GA given in Fig. 4.4 are examined, the operation of genetic representation of the problem which is handled is carried out with the help of genetic concepts (genes and chromosomes). After the process of coding is completed, first population examples are produced by forming an initial population for the initial implementation step. Each individual is individual has chromosomes set and the initial population can be produced in different types by using a randomized or special algorithm. In the next step, the target function is determined for individuals in the populations, and the implementation of genetic operations and the selection of the individuals are performed. In this stage, operators of reproduction, crossing, and mutation are used (Gümüşoğlu et al. 2013; Qureshi et al. 2006). The performances of the individuals in the new population after their validation values are calculated based on the validation function. These operations are continued until reach to finalization criterion determined at the beginning and when iteration is reached to this criterion the optimal solution is presented by finalizing the algorithm (Emel and Taşkın 2002; İşçi and Korukoğlu 2003; Gholamia et al. 2014).

4.2.4 Decision Trees (DT)

Decision trees are a diagram, which is viewed as a tree and repetitive diversification of a dataset to sub-groups by classifying (dichotomous classifications) (Namazkhan et al.2020). Understanding and evaluating the learning rules (if–then rules) in decision trees are very easy. In addition to that, costs are very low and integration between databases is possible and it has not non-flexible assumptions (Çalış Boyacı et al. 2014). Decision trees, which can self-learning from data with the help of the method of generalization, were introduced by J. Ross Quinlan in year 1960 (Wei et al.. 2019). In addition, algorithms of the decision trees which are accepted as a data mining technique are an effective tool for presenting complex problems in graphical form by means understandable and lean. The decision trees that have a definitive structure (continuously or categoric) are defined as classifying tree, others which can be used for solving the regression types of problems are regression tree (Alan, 2014). Within this perspective, a decision tree consists of three sections such as node, branch, and leaf. An example of a decision tree is presented in Fig. 4.5 (Aytekin et al. 2018).

When Fig. 4.5 is examined, nodes represent each self-feature. While the root is on the top, the leaf is below the tree. Branches are between roots and leaves. When a decision tree is built, the process starts with the classification of data that is in the root and asking the questions. This respective operation continues until obtaining nodes having no branches or obtaining the leaves. When any dividing is out of the question, the algorithmic process is finalized. Whether gain the generalization ability by the obtained decision tree is determined by using test data. In this process, a new dataset is used and test data is tested in the root section of the tree. Then it is progressed to the next node. This process is finished when it is reached to the leaf of the tree (Kavzoğlu and Çölkesen, 2010; Aytekin et al. 2018).

In this context, some algorithms (i.e., CHAID, CART, ID3, C4.5, and so on) are used in decision trees (Alan 2014; Aytekin et al. 2018). Obtaining the optimum

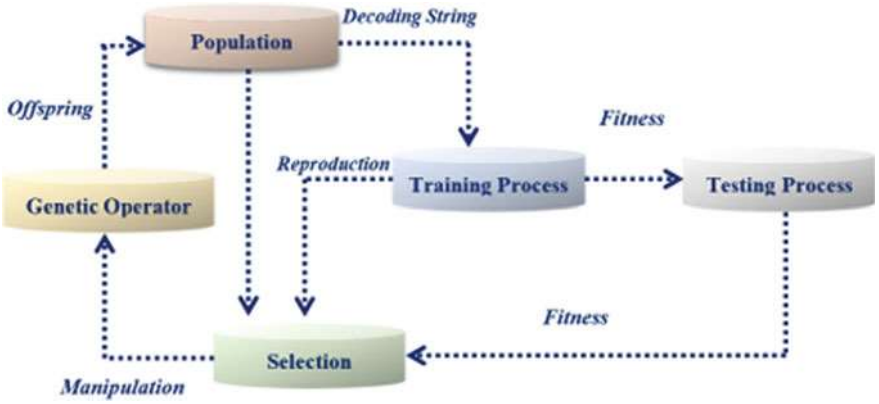

Fig. 4.4 The flowchart of search procedure of GA (Gholamia et al. 2014)

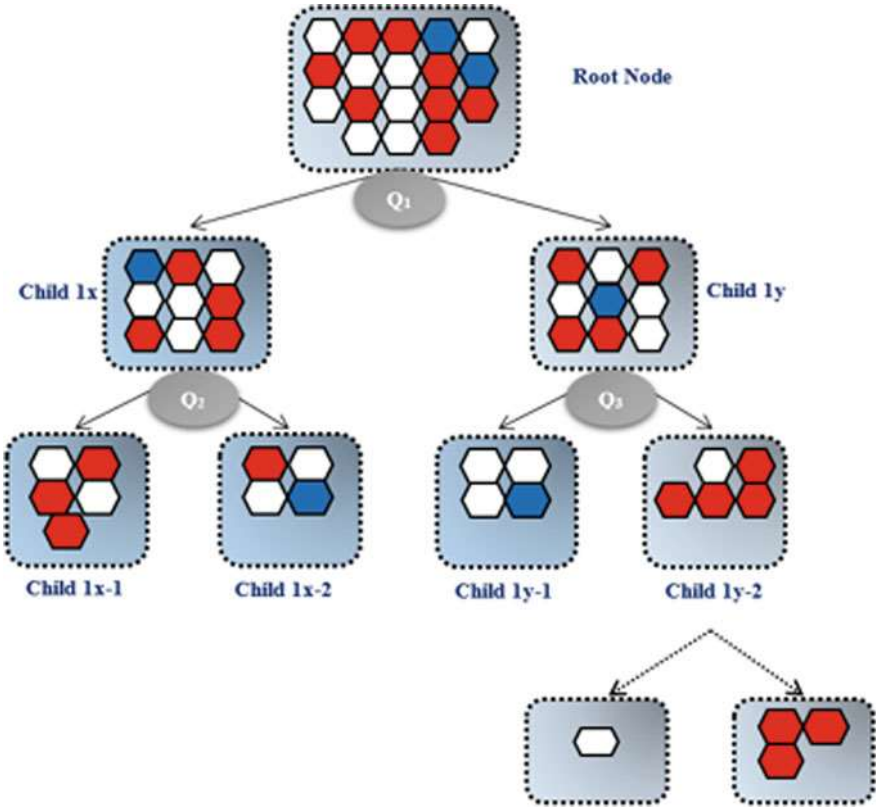

Fig. 4.5 A basic decision trees (Aytekin et al., 2018)

structure of the tree which can help to reduce to a minimum of generalization faults is one of the main aims in this process. Therefore, forming a small-sized tree can help to obtain more successful results. some measurements suggested in the literature are commonly used. Information gaining, ki-square test, entropy, and Gini index are examples (Kavzoğlu and Çölkesen 2010).

4.3 Results and Discussions

In recent years, data has become a strategical weapon for industries and companies. in the past century, oil was the most important factor for companies and nations and data have supplanted oil at the present. Actors having and managing data have become determinative of the rules and conditions of the highly competitive business environment and obtaining and being able to use data have become a sole and exclusive remedy to survive in the highly competitive for companies and other actors. As

the competition hardens in global markets, businesses learn the way to control the increasing data flow, learn about big data processes, and transform the most appropriate data into information by filtering infinitely large data. Simultaneous tracking of what the manufacturer will produce and what the customer will demand over big data will be managed by big data systems. As a result, big data has increased to a size that cannot be managed with traditional processes right now. therefore, mathematical models and computational tools and may be needed to solve this problem for companies and industries to solve these kinds of problems. Finally, we are going through a new era, as one of the fundamental parts of the Big Data Industry 4.0 process, it is clearly seen that it will be the fundamental building block of digital businesses and the digital economy in the coming years.

References

Alan MA (2014) Karar Ağaçlarıyla Öğrenci Verilerinin Sınıflandırılması. Atatürk Üniversitesi İktisadi Ve İdari Bilimler Dergisi 28(4):101–112

Aladağ ÇH (2019) Architecture selection in neural networks by statistical and machine learning. Orient J Comput Sci Technol 12(3):76–89

Atalay M, Çelik E (2017) Büyük Veri Analizinde Yapay Zekâ ve Makine Öğrenmesi Uygulamaları. Mehmet Akif Ersoy Üniversitesi Sosyal Bilimler Enstitüsü Dergisi 9(22):155–172

Aydemir E, Karagül K (2020) Solving a periodic capacitated vehicle routing problem using simulated annealing algorithm for a manufacturing company. Braz J Oper Prod Manag 17(1):1–13

Aytekin Ç, Sütcü CS, Özfidan U (2018) Text classification via decision trees algorithm customer comments case. J Int Soc Res 11(55):782–792

Cyganek B, Graña M, Krawczyk B, Kasprzak A, Porwik P, Walkowiak K, Woźniak M (2016) A survey of big data issues in electronic health record analysis. Appl Artif Intell 30(6):497–520

Çalış Boyacı A, Kayapınar Kaya S, Çetinyokuş T (2014) Veri Madenciliğinde Karar Ağacı Algoritmaları ile Bilgisayar ve İnternet Güvenliği Üzerine Bir Uygulama. Endüstri Mühendisliği Dergisi 25(3–4):2–19

Dülger Ü (2015) Stratejik Büyük Veri Yönetiminin Yatırımlar Üzerindeki Etkileri. Yayınlanmamış Yüksek Lisans Tezi, İstanbul Üniversitesi Fen Bilimleri Enstitüsü

Emel AG, Taşkın Ç (2002) Genetik Algoritmalar ve Uygulama Alanları. Uludağ Üniversitesi İİBF Dergisi 21(1):129–152

Ghannadpour SF, Zandiyeh F. (2020) An adapted multi-objective genetic algorithm for solving the cash in transit vehicle routing problem with vulnerability estimation for risk quantification. Engineering applications of artificial intelligence, vol 96

Gholamia R, Moradzadehb A, Malekic S, Amiric S, Hanachid J (2014) Applications of artificial intelligence methods in prediction of permeability in hydrocarbon reservoirs. J Petrol Sci Eng 122(2014):643–656

Gülsün B, Tuzkaya G, Duman C (2009) Genetik Algoritmalar ile Tesis Yerleşimi Tasarımı ve Bir Uygulama. Doğuş Üniversitesi Dergisi 10(1):73–87

Gümüşoğlu Ş, Erboy N, Özdağoğlu G (2013) Siparişe Dayalı Üretim için Ürün Gruplarının Oluşturulmasında Genetik Algoritma Tabanlı Bir Yaklaşım. Yönetim Ve Ekonomi Dergisi 20(2):259–284

İşçi Ö, Korukoğlu S (2003) Genetik Algoritma Yaklaşımı ve Yöneylem Araştırmasında Bir Uygulama. Yönetim Ve Ekonomi 10(2):191–208

Kavzoğlu T, Çölkesen İ (2010) Karar ağaçları ile Uydu Görüntülerinin Sınıflandırılması Kocaeli Örneği. Harita Teknolojileri Elektronik Dergisi 2(1):36–45

Maind SB, Wankar P (2014) Research paper on basic of artificial neural network. Int J Recent Innov Trends Comput Commun 2(1):96–100. ISSN: 2321-8169
Mayer-Schonberger V, Cukier K (2013) Big data: a revolution that will transform how we live, work and think. Houghton Mifflin Harcourt, Boston, Massachusetts.
Mjalli FS, Al-Asheh S, Alfadala HE (2007) Use of artificial neural network black-box modeling for the prediction of wastewater treatment plants performance. J Environ Manag 83(3):329–338
Nabiyev VV (2012) Yapay Zekâ. Ankara: Seçkin Yayıncılık
Namazkhan M, Albers C, Steg L (2020) A decision tree method for explaining household gas consumption: the role of building characteristics, socio-demographic variables, psychological factors and household behaviour. Renew Sustain Energy Rev 119:109542
Ocak İ, Şeker ŞE (2013) Calculation of surface settlements caused by EPBM tunneling using artificial neural network SVM and Gaussian processes. Environ Earth Sci 70(3):1263–1276
Özköse H, Gencer C (2019) Proje Planlama ve Çizelgelemede Genetik Algoritma Tabanlı Bir Yöntem ile Kritik Yolun Proje Tamamlanma Zamanının tespiti ve Zaman Maliyet Analizi. Bartın Üniversitesi İktisadi Ve İdari Bilimler Fakültesi Dergisi 10(20):278–300
Özşahin Ş, Singer H (2019) Odun Yüzey Pürüzlülüğü Tahmininde Bir Yapay Sinir Ağı Modelinin Kullanılması. Düzce Üniversitesi Bilim Ve Teknoloji Dergisi 7(3):1764–1777
Öztemel E (2003) Yapay Sinir Ağları. İstanbul: Papatya Yayıncılık
Qureshi SA, Mirza SM, Arif M (2006) Fitness function evaluation for image reconstruction using binary genetic algorithm for parallel ray transmission tomography, emerging technologies. In: ICET'06 international conference. Islamabad, Pakistan, pp 196–201
Sağıroğlu Ş, Koç, O(eds) (2017) Büyük Veri ve Açık Veri Analitiği, Yöntemler ve Uygulamalar. Ankara: Grafiker Yayınevi
Spann T (2017) The physics of big data. Erişim: 22.05.2020. https://dzone.com/articles/the-physics-of-big-data/
Şahan AN (2020) Stratejik Yönetim Perspektifinden Sigortacılık Sektöründe Makine Öğrenmesi Algoritmaları ile Anomali tespiti. Yüksek Lisans Tezi, İstanbul Teknik Üniversitesi Fen Bilimleri Enstitüsü
Şahin Y, Karagül K (2019) Gezgin Satıcı Probleminin Melez Akışkan Genetik Algoritma MAGA Kullanarak Çözümü. Pamukkale Üniversitesi Mühendislik Bilimleri Dergisi 25(1):106–114
Şahinarslan FV (2019) Makine Öğrenmesi Algoritmaları ile Nüfus Tahmini: Türkiye Örneği. Yüksek Lisans Tezi, İstanbul Teknik Üniversitesi Sosyal Bilimler Enstitüsü.
URL 1: http://www.accaglobal.com. Access Date: 22.05.2020
URL 2: https://andressilvaa.tumblr.com/post/87206443764/big-data-refers-to-5vs-volume. Access Date: 22.05.2020
Venkatram K, Geetha MA (2017) Review on big data & analytics–concepts, philosophy, process, and applications. Cybernetics Inform Technol 17(2):3–27.
Wei J, Chu X, Sun X-Y, Xu K, Deng H.-X, Chen J, Wei Z, Lei M (2019) Machine learning in materials science. InfoMat 1:338–358. https://doi.org/10.1002/inf2.12028
Yakıcı Ayan T (2008) Sabit Maliyetli Ulaştırma Problemi için Bir Genetik Algoritma. Gazi Üniversitesi İktisadi Ve İdari Bilimler Fakültesi Dergisi 10(1):97–116
Yılmaz B, Bülbül S, Atik M (2017) Büyük Verinin Big Data Muhasebe Üzerindeki Etkisi ve Muhasebeye Sağladığı Katkıların İncelenmesi. Kara Harp Okulu Bilim Dergisi 27(1):79–112
Zhang L, He M, Shao S (2020) Machine learning for halide perovskite materials. Nano Energy 78:105380

Özhan Görçün graduated from Department of Public Administration at Abant İzzet Baysal University in 2002, Logistics and Supply Chain Management Department at Institute of Science in 2009. In 2015, Özhan started Doctorate Program on Business Administration in Istanbul Aydın University. He started his business career in the logistics industry and also taught in the Logistics program at Bahçeşehir University. He is head of Logistics Program of Adıgüzel Vocational

School. He published and contributed to many national and international books, articles and notifications published.

Hande Kucukonder (Ph.D.) is a Vice-Dean of Economics and Administrative Sciences at the Bartin University. She gives many courses in statistical data analysis, Artificial Neural Networks (ANN), Data Mining, Decision-Support Systems and so on. in addition to those, she studies about numerical analysis, statistics, multi-criteria decision-making, and data mining. She got a master's degree in the field of Biometry and Genetics and completed doctoral studies about the same area. she has numerous scientific studies about Artificial Neural Networks (ANN), Data Mining, Multi-criteria decision-making methods, statistical data analysis.