Co-funded by the
Erasmus+ Programme
of the European Union

MIND COURSE SUPPORT

LECTURE 7
Digitalization and Industry 4.0
Data Technology for enabling digitalization of the manufacturing sector and Industry 4.0

This project has been funded with support from the European Commission. This publication [communication] reflects the views only of the authors, and the Commission cannot be held responsible for any use which may be made of the information contained therein.

Contents

This project has been funded with support from the European Commission. This publication [communication] reflects the views only of the authors, and the Commission cannot be held responsible for any use which may be made of the information contained therein.

1. OBJECTIVES OF THE LECTURE

The objectives of this course are getting knowledge on the concepts of data technology for enabling digitalization of the manufacturing sector and Industry 4.0. Specifically, the topics that will be covered within the lecture are big data, intelligent analytics and blockchain technology. In today's digitalized world, and especially in modern industrial environments, every performed process is detected and registered by the large amount of collected data. With the introduction of the Industry 4.0 concept in production, digitalized business systems with a continuous inflow of a large amount of information have been created. Due to the importance of digitization within Industry 4.0 and the techniques for collecting, storing, and processing this information, this course will discuss the concept of Big Data and Intelligent Analytics that can effectively treat that information. Finally, in order to consider the aspects of security and manipulation of collected data using the latest digital approaches, blockchain technology will be presented and the basic ways of functioning will be explained in the third part of this course.

Objectives of the course are:

- Introduction to the Big data concepts and characteristics
- Understanding different technologies for acquiring, analyzing, and processing data
- Introduction to the Blockchain technology
- Understanding fundamental Blockchain features: security, decentralization, mining, hash functions, privacy and authentication
- Gaining the knowledge of machine learning types commonly applied for analytics
- Understanding the requirements for providing optimal analytical environment
- Introduction to descriptive, predictive and prescriptive analytics
- Presentation of real-world applications in the domains of Big data, Blockchains and Machine learning powered analytics

This project has been funded with support from the European Commission. This publication [communication] reflects the views only of the authors, and the Commission cannot be held responsible for any use which may be made of the information contained therein.

2 Big Data

2.1 Introduction to Big Data concept

Report by research company International Data Corporation (IDC) claims that between 2012 and 2020 the amount of information in the digital universe has grown by 35 trillion gigabytes. This huge amount of data is contributed by the digitization of most processes, the existence of different social networks, blogs, online markets, the use of different types of sensors incorporated into wearables, etc. For example, according to the Facebook official report at the beginning of 2020, this platform has an average of 2.60 billion monthly active users and 1.73 billion daily active users, is available in 101 languages with over 300,000 users helping in translation, users share over 100 billion messages on a daily basis, 350 million photos are uploaded every day, every 20 minutes 1 million links are shared, 20 million friend requests and 3 million messages are sent, 55 million status updates are made every day, etc [1].

No one can deny that the Internet has changed the way businesses operate, the functioning of the government, education and lifestyle of people around the world. Earlier, organizations used transaction-processing systems that inherently used Relational Data Base Management Systems (RDBMS) and simple data analysis techniques like Structured Query Language (SQL) for their day-to-day operations that helped them in their decision making and planning. Nowadays, we have concept of Big Data whose size, diversity, distribution and / or time sensitivity imply the use of new technological and analytical architectures in order to achieve additional market value [2]. Based on this fact, the comprehensive definition of the Big Data concept can be given [3]:

"*Big Data is a term encompassing the use of techniques to capture, process, analyse and visualize potentially large datasets in a reasonable timeframe not accessible to standard IT technologies. By extension, the platform, tools and software used for this purpose are collectively called Big Data technologies*".

This project has been funded with support from the European Commission. This publication [communication] reflects the views only of the authors, and the Commission cannot be held responsible for any use which may be made of the information contained therein.

Since the term Big Data is actually generic, it is difficult to determine when it was first used and who used it. Some sources state that it was introduced by John Mashey back in the mid-1990s, who was a leading expert at company *Silicon Graphics* [4]. According to the other source, Francis X. Diebold, an economist at the University of Pennsylvania, published the first scientific paper and reference in the field of Big Data [5]. The literature also states that Roger Magoulas of O'Reilly media first used this concept in the IT in 2005 [6].

However, most of the existing definitions of Big Data concept have in common the use of three "Vs" in some form, which represents the initial letters of words: Volume, Variety, and Velocity [7]. Often, two more "Vs" are added that refer to Veracity and Value [8].

2.2 Big data characteristics

Big Data is characterized by the following dimensions:

Volume

Volume refers to the magnitude of the data that is being generated and collected. It is increasing at a faster rate from terabytes to petabytes (1024 terabytes). In the past, an excessive amount of data created problems with storage, but with today's prices of memory devices, that is no longer a problem. However, other problems arise, including determining the importance of certain data in a large crowd. Many factors contribute to the increase in the volume of data, and the most important of them are:

- The number of data based on transactions is growing: in the field of banking, insurance, medical services, communications (mobile), etc.
- The number of unstructured data coming from social media is growing.
- The amount of data read from sensors (thermal, electromagnetic, mechanical, chemical, optical, ion radiation sensors, etc.) and similar devices (traffic counters, GPS devices, scanners at retail cash registers, etc.) has increased.
- Extensive use of cloud technology.
- Introduction of intelligent devices in households and measuring devices (electricity, gas consumption, etc. - smart meters).

Velocity

This project has been funded with support from the European Commission. This publication [communication] reflects the views only of the authors, and the Commission cannot be held responsible for any use which may be made of the information contained therein.

Velocity refers to the rate of generation of data, i.e. to the time necessary to obtain a recommendation for certain action. The reasons why this dimension is imperative in modern business are clear:

- Primarily because of a competitive match: it is necessary to identify the problem, to recognize the opportunity before others. Sometimes it's a matter of seconds, even milliseconds.
- Data have a very short shelf life. They quickly become obsolete and no longer represent a competitive advantage.

This only confirms that the data must be collected, processed and analysed practically in real time in order to gain insight into the essence of the data as soon as possible. This would help to retain the customers as well as to increase the service level.

Variety

Big data can come in multiple forms, including structured and non-structured data (Figure 1) [9]. Contrary to the most of the traditional data analysis performed by organizations, most of the Big Data is unstructured or semi-structured in nature, which requires different techniques and tools to process and analyze them.

Unstructured data is basically information that either does not have a predefined data model and / or does not fit well into a traditional database: text, pdf document, video, images, audio, geospatial data, internet data, slick streams, log files, etc.

"Quasi" structured data is textual data that is given in a non-standard format and, as such, can be formatted, which requires a lot of knowledge, tools and time. A typical example of this data type is web click stream data, which may contain certain inconsistencies, primarily in format and content.

This project has been funded with support from the European Commission. This publication [communication] reflects the views only of the authors, and the Commission cannot be held responsible for any use which may be made of the information contained therein.

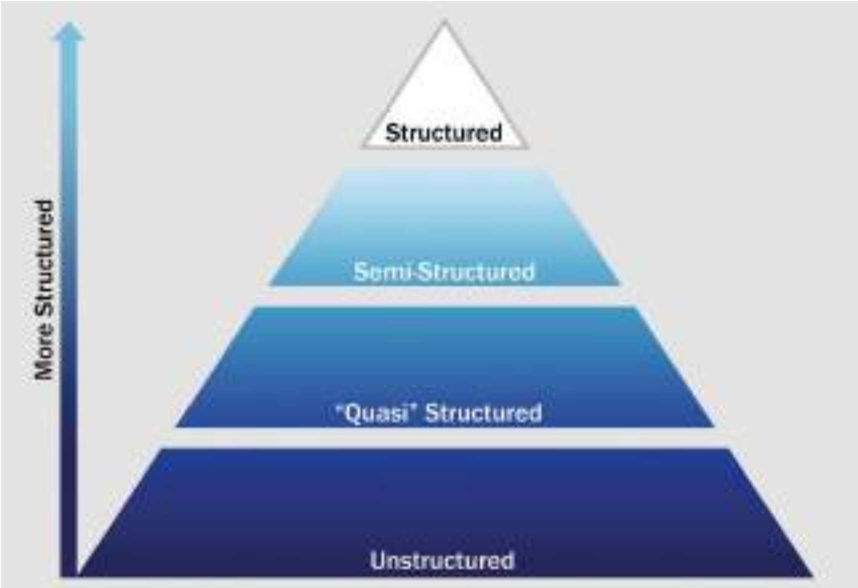


Figure 1. Data structures

Semi-structured data are used to describe structured data that do not fit into the formal structure of the data model. These data do not contain tags that separate semantic elements, do not have a common structure, have the ability to implement a hierarchy within the data, and involve multiple ways of representing the same type of data. An XML (eXtensible Markup Language) programming language similar to HTML and developed by the W3C (World Wide Web Consortium) is used to represent semi-structured data in order to overcome the limitations of HTML. XML is designed as a programming language for describing data - implying a description of the data, not their appearance.

Structured data have a clearly defined type, format, and structure [10]. This type of data is most often stored in company databases and / or data warehouses. Compared to the "traditional" statistical methodology, structured data can be classified into the category of metric or numerical variables whose processing, analysis and interpretation is very precisely defined and relatively simple.

It is important to emphasize that the growth of all types of data is evident, with the difference that the generation of structured data follows a linear trend, in contrast to unstructured data whose growth is exponential.

This project has been funded with support from the European Commission. This publication [communication] reflects the views only of the authors, and the Commission cannot be held responsible for any use which may be made of the information contained therein.

Veracity

Having in mind that data exists in different forms and are collected from a multitude of sources, controlling the accuracy, reliability and/or reliability of data is a challenge for Big Data scientists and researchers. Social networks, for example, have introduced the hash tags option (#) and data are collected in the form of abbreviations, some data also contain typos or colloquial speech. Big Data analytics enables work with such a data structure, and most often the amount of available data compensates the shortcomings of quality and accuracy.

Value

This characteristic implies the possibility of converting data into value, i.e. profit. Data in its original form is unusable and they have to be analysed to discover very high value. For this analysis some of the sophisticated mathematical and statistical methods and techniques, as well as artificial intelligence are used. More details about these technologies will be discussed in the following section.

2.3 ***Technologies for collecting, pre-processing and analysing Big Data***

Big data technologies support both the ability to collect large amounts of data, and ability to understand that data in order to extract some value. As an analytical tool, the value chain shown in Figure 2 [11], can be applied to information flows to understand the value creation of data technology. In a Data Value Chain, information flow is described as a series of steps needed to generate value and useful insights from data.

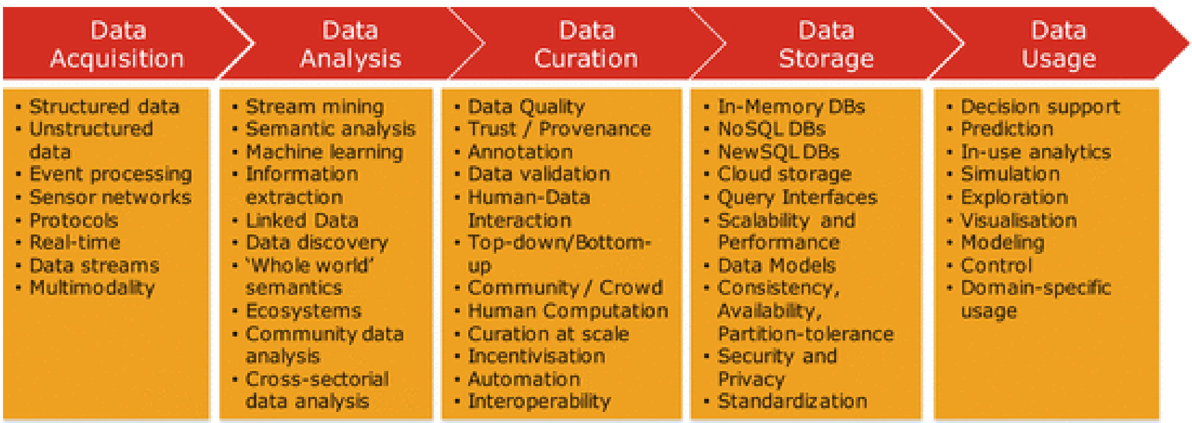


Figure 2. The Big Data Value Chain

This project has been funded with support from the European Commission. This publication [communication] reflects the views only of the authors, and the Commission cannot be held responsible for any use which may be made of the information contained therein.

Practically speaking, the main idea of organizations that have access to big data collections is to explore these data and extract useful information or knowledge for their future decision making. The real breakthrough in Big Data technology happened when companies like Google and Yahoo translated this idea into concrete products such as MapReduce, Big Table and Hadoop.

MapReduce

MapReduce is a solution introduced by Google as a way to efficiently execute a set of functions over vast amounts of data in a serial way. The component "map" distributes a programming problem or task to a large number of systems and manages task setup (that involves a balanced workload) and error recovery. After the distributed processing is completed, another function named "reduce" is called, which joins all the elements back together to provide the result. One example of MapReduce usage could be the task of determining how many pages of a book are written in each of some 50 different languages.

Big Table

Big Table is also a solution developed by Google, as a distributed data storage system designed to manage highly scalable structured data. The data are organized into tables with rows and columns. Unlike the traditional relational database model, the Big Table is a thinned, distributed, and permanently sorted multidimensional map. The Big Table is designed to store large amounts of data on regular servers. The Big Table maps two arbitrary strings (a row-related key and a column-related key) and a time instant into some bound bit string. The Big Table is designed to be able to go up to the petabyte level, working on over a hundred thousand machines that allow easy addition of new machines to the system and their immediate inclusion in a way that does not require any reconfiguration or interruption of the system.

Hadoop

Hadoop is an open-source framework, developed by Yahoo and Google, that is written in Java and it provides cross-platform support. Over half of the Fortune 50 companies use Hadoop including Amazon Web services, Hortonworks, IBM, Intel, Microsoft, Facebook, etc. Hadoop allows applications based on the MapReduce system to run on large clusters of

This project has been funded with support from the European Commission. This publication [communication] reflects the views only of the authors, and the Commission cannot be held responsible for any use which may be made of the information contained therein.

ordinary hardware, as well as to parallelize data processing using nodes to increase computational speed and decrease response. Hadoop consists of two main components, a highly scalable distributed file system that supports the amount of data measured in petabytes, while the other component is the MapReduce system.

In addition, there are many tools that help data scientists to process data and analyze them. Many new languages, frameworks and data storage technologies have emerged that supports handling of big data.

R: R is an open source programming language and environment for statistical calculations and data visualization. R is an implementation of a language S developed by John Chambers with colleagues at *Bell Laboratories te Robert Gentleman* from the University of Auckland, New Zealand. It has an effective data handling and storage facility because it supports statistical methods for linear and nonlinear modelling, classical statistical tests, time series analysis, and clustering. It is easily expandable via packages covering a very wide range of modern statistics, data munging, data mining and machine learning algorithms.

Python: Python is a high-level object-oriented programming language supported by Windows, Linux and Mac platforms. It was realized by the Dutchman Guido van Rosum in the 1980s. The name of the language comes from the British comedy group Monty Python. The most important revisions are: Python 1.0 (1991), Python 2.0 (2000), Python 3.0 (2008). Its simplicity, efficiency, elegance and open source have led to the huge popularity. NumPy, Scikit, and Pandas support some of the popular packages for machine learning and data mining for data pre-processing, computing and modelling. NumPy is a python library used for working with arrays, linear algebra, Fourier transform, and matrices. Scikit is a free software machine learning library which supports various classification, regression and clustering algorithms including support vector machines, random forests, gradient boosting. Pandas help in data manipulation and analysis. In particular, it offers data structures and operations for manipulating numerical tables and time series.

Scala: Scala (scalable language) is an open source programming language running on the Java Virtual Machine which supports both object-oriented paradigm and functional programming. While object-oriented programming combines data structures with the actions you want to perform on them, functional programming keeps both separate. As a result, every value is an object, and every operator is a method, while you can pass functions around as variables. Furthermore, Scala is very concise and one word can replace several loops. Scala's libraries cover the machine learning, data analysis, data visualization, and NLP, which make this language suitable for handling big data.

2.4 ***Applications of Big Data analytics***

At the beginning, few sectors like Telecommunication, Retail and Finance have been early adopters of big data analytics [12]. Today, there is no sector which is not touched by the concept of big data analytics.

Healthcare

The healthcare industry historically has generated large amounts of data, driven by different sources such as electronic patient record, clinical decision support system including medical imaging, clinical data, data sensed from different sensors for vital signs rates, etc. The integration of clinical, public health and behavioural data collected from different platforms helps to develop a robust medical treatment system, which can understand patterns and trends within these data. In that way, big data analytics has the potential to improve care, save lives and provide lower costs of healthcare. For example, Medical researchers can use large amounts of data on treatment plans and recovery rates of cancer patients in order to find trends and treatments that have the highest rates of success in the real world. For example, researchers can examine tumor samples in biobanks that are linked up with patient treatment records. Using this data, researchers can see things like how certain mutations and cancer proteins interact with different treatments and find trends that will lead to better patient outcomes.

Telecommunication

Today, only one global telecommunications company collects billions of detailed call records per day from 120 different systems and stores each for at least nine months. Communications data is mainly derived from: data collected from customer relationship management, billing systems, and terminal self-registration platforms; data in access network

This project has been funded with support from the European Commission. This publication [communication] reflects the views only of the authors, and the Commission cannot be held responsible for any use which may be made of the information contained therein.

and core network (mobile signalling, DPI, M2M data, etc.); data in operators own Internet applications: online business hall data, palm business office data, wing payment data, etc. These data can be very useful for further improvements in the communication level and attracting new customers, or for providing these data to the other sectors like police, military, marketing agencies, etc. One example is the so called targeted marketing, which improves the adoption of mobile services, reduces churn, thus, increasing the revenue of mobile service providers. These providers analyze a number of factors such as demographic data (gender, age, marital status, and language preferences), customer preferences, household structure and usage details to model the customer preferences and offer a relevant personalized service to them.

Sales and marketing

A recent study by DataMeer found customer analytics dominate big data use in sales and marketing departments. Marketing analytics helps the organizations to evaluate their marketing performance, to analyze the consumer behaviour and their purchasing patterns, to analyze the marketing trends which would aid in modifying the marketing strategies like the positioning of advertisements in a webpage, implementation of dynamic pricing and offering personalized products. For example, McKinsey found that biopharma companies typically spend 20% to 30% of their revenues on selling, general, and administrative. If these companies could more accurately align their selling and go-to-market strategies with regions and territories that had the greatest sales potential, go-to-market costs would be immediately reduced.

Government

Big data and analytics can be successfully applied to any public-sector to provide measurable outcomes, including actions against different nationwide issue, response to a local disaster, protection against the loss of sensitive information or intellectual property, or simply making government to improve existing citizen services. Policymakers are using satellite imagery, cell phone data and more to produce alternative economic indicators for new – and

This project has been funded with support from the European Commission. This publication [communication] reflects the views only of the authors, and the Commission cannot be held responsible for any use which may be made of the information contained therein.

real-time – policy insights. Policy makers use satellite imagery, data from mobile phones and other sources to create alternative economic indicators for new insights into policies and in real time. Most of the produced data by Government are unstructured and text-intensive, so, as such, they need some special algorithms (machine learning) and text analytics supports for analysis. These supporting technologies can effectively automate systematic, multilevel checks on insurance, procurement and tax records to mark applications that require further examination, or to identify deficiency which should be solved. Big data techniques are transforming processes to detect fraud, collusion and money laundering through the analysis of procurement and financial information. Algorithms are used to track procurement records, invoices, bank information, certificates of origin and other data to identify fraud, collusion and shell companies. Police and military forces use big data analytics to make better decisions about their actions. For example, basic information such as crime type and location can help officers make smarter decisions on patrol.

3. Analytics and Intelligence

Industry 4.0 technologies' essential components are analytical tools that possess tremendous power to analyze and process big data acquired from modern industrial processes and the overall production and business cycles [13]. The data can be collected from connected production elements: machines, programmable controllers, components, interfaces, customer specifications, etc. These and many more connected features share information and cooperate in the same environment, creating large quantities of raw, unprepared data. Advanced analytics is enabled nowadays by utilizing mostly intelligent methods to manipulate and extract useful insights from collected data. In this section, an approach of using machine learning (ML) techniques, as one of the most popular modern tools for intelligent analytics, will be introduced, and its analytical powers will be described.

This project has been funded with support from the European Commission. This publication [communication] reflects the views only of the authors, and the Commission cannot be held responsible for any use which may be made of the information contained therein.

Figure. 3 Analytical environment within Industry 4.0

Machine learning, as a branch of artificial intelligence (AI) [14], can be described shortly as a predictive analytical tool that focuses on past events to generate reliable forward-looking, and autonomous decisions. In most cases, ML operates in real-time, learns by using prepared training sets, and make meaningful recommendations and predictions from vast amounts of data. Models can adapt and acquire new knowledge without human supervision, learn from past iterations, and produce accurate outcomes through time. This iterative learning makes intelligent systems "smarter" and more reliable, improving their processing potential that can be observed while finding hidden insights, historical trends, and relationships, etc.

3.1.ML training techniques for advanced analytics

Utilizing ML within Industry 4.0 environment can effectively provide a meaningful value out of the collected data, unify the data, and make desirable predictions and business decisions. However, if a specific ML or AI model should be adequately used, the appropriate training type must be selected and applied first. There are four main types of learning algorithms [15] applicable in Industry 4.0 analytics: supervised, semi-supervised, unsupervised, and reinforcement learning.

This project has been funded with support from the European Commission. This publication [communication] reflects the views only of the authors, and the Commission cannot be held responsible for any use which may be made of the information contained therein.

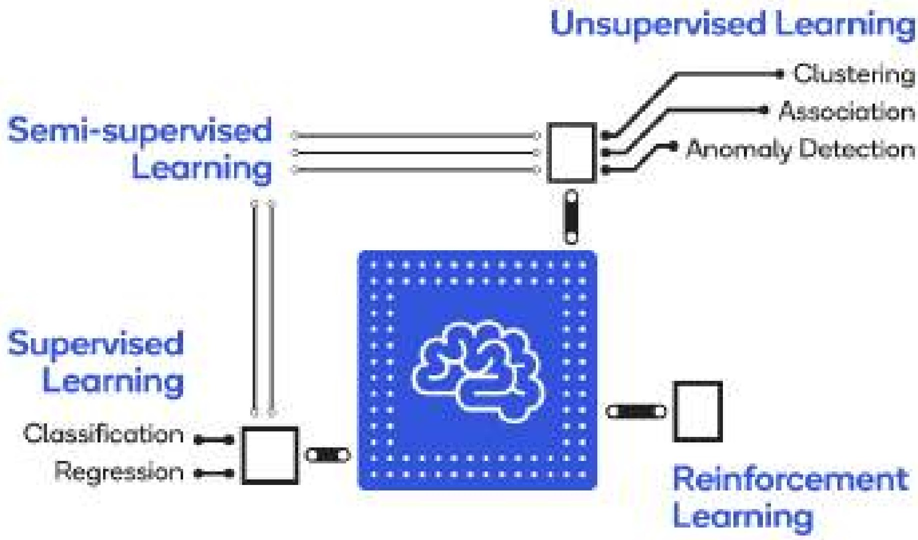


Figure. 4 Types of learning

- **Supervised learning** represents training types for which should be provided input/output pairs of data within a training set. The learning process is performed in a way to achieve an accurate generalization of training data, and then the general rules of behaviors of the specific model can be established. The most common supervised algorithms are: Random forest, Decision Tree, Linear Regression, Logistic Regression, Naïve Bayes, Support Vector Machines, Neural Networks.
- **Semi-supervised learning** is based on providing a small amount of labeled data and a significantly larger quantity of unlabeled data. This learning type requires less computational and time costs than supervised learning models and could be the best choice when there are expectations that the input data will unpredictably change over time. The most common semi-supervised algorithms are: Generative models, Low-density separation, Graph-based methods.
- **Unsupervised learning** is based on processing unlabeled data. Pre-defined and desirable outputs do not exist, and such systems try to examine the data and find some meaningful structures, hidden knowledge, and valuable patterns. The most common unsupervised algorithms are: Hierarchical clustering, K-means clustering, K-modes, Gaussian Mixture Model.

This project has been funded with support from the European Commission. This publication [communication] reflects the views only of the authors, and the Commission cannot be held responsible for any use which may be made of the information contained therein.

- **Reinforcement learning** is a useful method for training intelligent systems when placed in volatile and interactive environments. Such systems are challenged with tasks; they give response actions, and feedback information is provided in the form of rewards and punishments. If the system provides an adequate response to the task, a reward signal will be introduced to the algorithm, while if the action is wrong, the punishment signal will be generated. The most common reinforcement learning algorithms are: Q-learning, Deep Q Network, SARSA, Monte Carlo.

Described ML training types represent the base of algorithms that can be effectively used for solving various analytical problem. The following picture graphically presents for which purposes specific ML training types should be selected in accordance with the properties of a given problem.

The first question that should be answered based on Fig. 5 is if the collected data set requires a reduction of dimensions and significant data preprocessing. If the answer is yes, some of the unsupervised learning techniques for dimension reduction should be used (Principal component analysis, Singular Value Decomposition, Latent Dirichlet Analysis, etc.). If the dimensionality reduction is not the goal, it should be determined if desired output responses within the data exist or not. If the target variables are known, supervised learning techniques should be used. Once again, the choice should be made: if the task is to predict numeric values (regression) or provide a classification of features. When this choice is made, depending on the goal of greater importance – the speed of the model's calculations or achieved accuracy, an appropriate supervised learning algorithm should be selected. Finally, if the target variables are not included within the data set, clustering (unsupervised learning techniques) should be used for treating the data. When the ML types are defined, it is good to know what are further requirements for providing suitable digital environment for intelligent analytical processes.

This project has been funded with support from the European Commission. This publication [communication] reflects the views only of the authors, and the Commission cannot be held responsible for any use which may be made of the information contained therein.

Figure. 5 Selection procedure of optimal ML algorithm

3.2.Requirements for providing optimal digital analytical environment

The benefits of using AI and ML solutions in digitalized industrial processes [16] are numerous. Traditional rule-based systems perform well in static situations but they can have problems or be complicated for design when the data is constantly or rapidly changing. In these cases, a rule-based system should be constantly updated with new rules and features, in order to follow new data tendencies. However, by using ML models, this requirement can be easily eliminated and the development time reduced, finally implying reduced production and maintenance costs. However, to fulfill all the requirements within Industry 4.0 to establish an analytically suitable environment for the implementation of intelligent algorithms, the following phases should be configured and covered properly:

- Established connectivity via communication protocols, gateway devices and routing between PLCs, computers, and other digital controllers and base stations
- On-line data recording via sensors and historical data collection
- Data aggregation, preparation and formatting into suitable data formats (csv for example)

This project has been funded with support from the European Commission. This publication [communication] reflects the views only of the authors, and the Commission cannot be held responsible for any use which may be made of the information contained therein.

ML approaches are also convenient from the perspective of using different types of data within one system. A high volume of insights can be easily extracted both from structured and unstructured data - by unifying data across different platforms and different sources. Another benefit - scalability, is provided by the parallelism of ML algorithms which results in fast processing and effective generation of output values. ML applications can be easily scaled to multiple machines instead of a single one, additionally providing fewer computation costs and faster responses.

3.3.Descriptive, predictive and prescriptive analytics

In general, all analytical options can be categorized into three different groups [17]: descriptive, predictive, and prescriptive analytical approaches.

Descriptive analytics provides insights into the past. It summarizes raw data, describes it, and generate outcomes that are interpretable by humans. Descriptive-analytic technology is valuable from the perspective of learning from past behaviors and understanding how these behaviors can make an influence on future events and outcomes. It is good to highlight that the majority of the regularly used statistical methods also belong to this category (averages, percent changes, deviations). Examples of descriptive analytics are historical insights regarding production, finances, sales, customers, etc.

Predictive analytics utilizes forecasting techniques and statistical models to understand and predict future events and outcomes. These analytical tools are based on probability theory and provide information with insights based on collected data and estimate the likelihood of future outcomes. They can combine historical data to identify patterns and capture relationships between various data sets by utilizing statistical models. Companies use predictive analytics to forecast behaviors of customers and production demands, to identify purchasing trends, or to produce a credit score (to determine if a customer will make future credit payments on time or not).

This project has been funded with support from the European Commission. This publication [communication] reflects the views only of the authors, and the Commission cannot be held responsible for any use which may be made of the information contained therein.

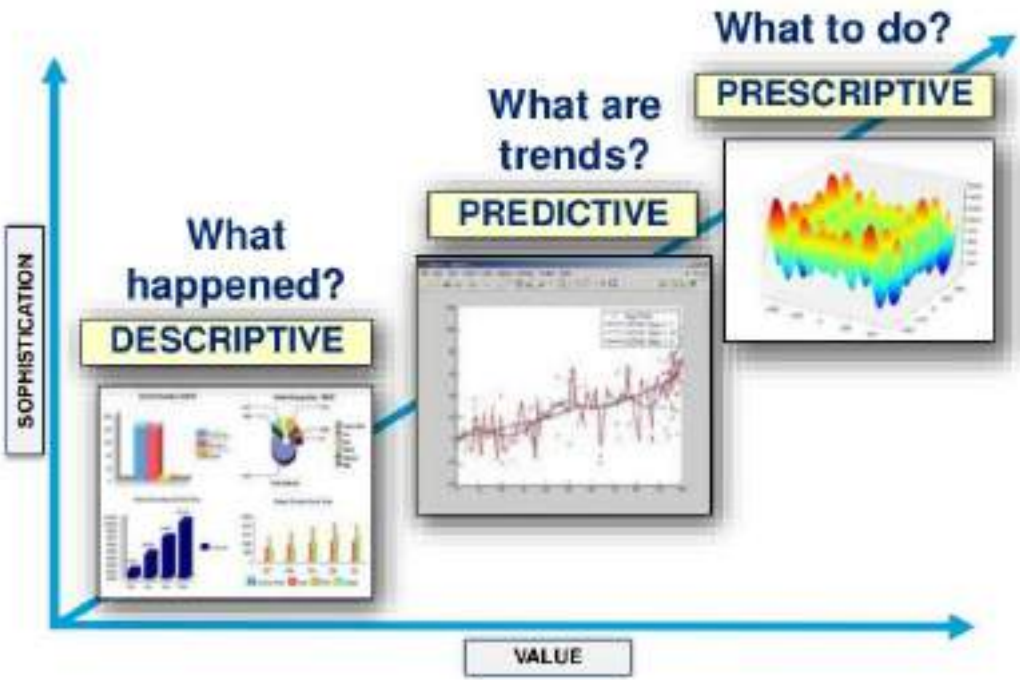


Figure. 6 Descriptive, predictive and prescriptive analytical approaches

Prescriptive analytics allows the prescription of different possible actions and providing meaningful advices. It quantifies the effects of future decisions and generates advice on possible outcomes. All that is happening before actually making the decisions. In some cases, prescriptive tools can even predict why some events will occur, and provide recommendations related to actions that will take advantage of generated predictions. Multiple features can be predicted at the same time by using multiple techniques as are computational modeling procedures, machine learning algorithms, and pre-defined business rules. Like previous types of analytics, prescriptive analytics can also use historical and real-time data, transactional data, structured and unstructured big data, etc. However, besides all the benefits it can bring, prescriptive analytics is complex for realization and maintenance, so these tools are not used commonly on a regular basis. Still, larger companies use them for optimization processes and scheduling tasks, as well as for monitoring of supply chains.

This project has been funded with support from the European Commission. This publication [communication] reflects the views only of the authors, and the Commission cannot be held responsible for any use which may be made of the information contained therein.

3.4.Real-world ML applications and analytical approaches in Industry 4.0

- The most important utilization of intelligent algorithms within the industry sector is **the digitalization of production processes** [18], their full automatization, real-time monitoring, and creating intelligent autonomous decision-making systems that can generate reliable and optimal control and production decisions. **Smart manufacturing** is also defined by the potential for adaptive manufacturing and making preventive measures. ML tools applied for different types of manufacturing data analysis can give engineers information for fine-tuning process parameters and improving manufacturing performances.
- **Predictive quality.** Manufacturers constantly try to reduce production losses and prevent inefficiencies of production processes. For the purpose of obtaining desirable production quality, Industrial AI and machine learning algorithms could be used to reveal hidden causes of production losses faced on a regular basis. In most cases for this kind of ML utilization, supervised learning algorithms are used to identify patterns and trends in data. When a production problem is identified, automated alerts and recommendations can be generated with the purpose to inform engineers of an imminent problem.

This project has been funded with support from the European Commission. This publication [communication] reflects the views only of the authors, and the Commission cannot be held responsible for any use which may be made of the information contained therein.

- **Predictive Maintenance.** Predictive maintenance uses supervised or unsupervised machine learning algorithms to analyze and predict the next failures of a system and alert operators to perform required maintenance procedures with the purpose to prevent failures. Predictive ML analytics can significantly reduce maintenance costs and even eliminate the requirements for scheduled downtime in multiple cases. Today's intelligent predictive maintenance tools are so sophisticated that they make it possible to inform technicians which tools should be used during a specific maintenance process, which procedure to follow and what specific components should be inspected and replaced if needed. Predictive analytics successfully prevents secondary damage of components and equipment that finally leads to a longer remaining useful life of systems.
- **Market adaption.** ML analytics can be very effectively used for optimizing the supply chains of operations and provide useful information on how to better respond to market changes. ML algorithm in this case takes into account patterns that usually include date information, location, economic and macroeconomic features, etc. Intelligent approaches in market analytics can be used to optimize different business and production processes, management plans, work efficiency of employees, consumption of energy and raw resources, and finally, to make better financial decisions.
- **Image recognition** [19], as one of the most popular features of using artificial intelligence and deep neural networks, can help in verifying the quality of produced mechanical parts or other products on a production line. High-quality cameras in addition to such an intelligent image recognition structure can perform visual analysis and find defect parts more easily and more accurately than any human supervision.
- AI can help manufacturers to make **personalized offers** to each customer, where the customer can participate during the production process and manually select additional features for his products on-line, and unrestricted number of times, without the requirement for any communication with human operators. Before Industry 4.0, a customer had the possibility only to order a product in a defined time frame, with strict

This project has been funded with support from the European Commission. This publication [communication] reflects the views only of the authors, and the Commission cannot be held responsible for any use which may be made of the information contained therein.

deadlines and not significant freedom to change his mind on something or to request personalized production. With industry 4.0 and achieved digitalization, everything is changing.

- By digitalizing and automating routine processes and making manufacturing tasks smart and autonomous, operators and engineers could focus on **more creative and value-added activities** to actively contribute to the future progress of a company.
- Machine learning applications can also be useful for **detecting frauds** and 3rd parties with suspicious tendencies within a system and addressing **cyber-security issues**. Pattern recognition and learned habits of honest clients can be used for checking if the examined activity is suspicious/malicious or regular one. ML for fraud detection establishes itself as a regular tool in current Industrial 4.0 environments and within digitalized accompanying systems.

4. Blockchain

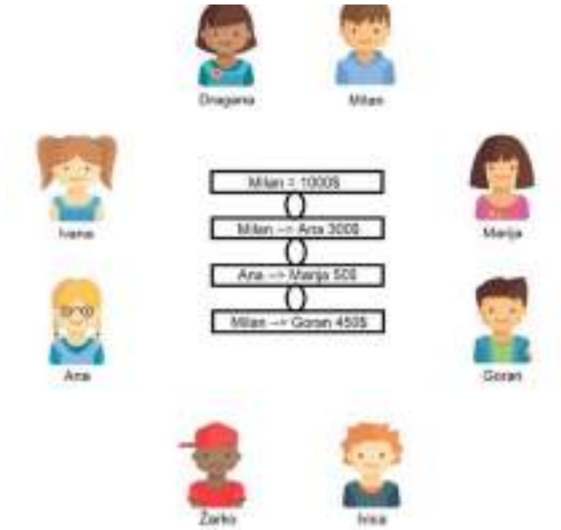
The basic concepts of blockchain technology were developed in Satoshi Nakamota's original 2008 paper entitled “Bitcoin: A Peer-to-Peer Electronic Cash System”. By the way, no one today knows who Satoshi is, because he never revealed his identity to the public. Blockchain [20] concept will be explained on the example of Bitcoin [21], but it should be emphasized that crypto currencies are only one application of blockchain technology.

The blockchain is practically one main financial book - a ledger in which it is reliably written who has how much money and who has given whom how much money in the past (if it is a blockchain for cryptocurrency). This book – ledger is open and public, which means that anyone can see it at any time (all transactions that have ever taken place) and that new transactions can always be added to its end, but the records already existing in the book cannot be changed.

Let there be 8 people in the system: Dragana, Milan, Marija, Goran, Ivica, Žarko, Ana, and Ivana. Let Milan have $1,000 at the start. We enter this information as the first block of our ledger. Now Milan wants to transfer $300 to Ana. We enter that transaction as the next block

This project has been funded with support from the European Commission. This publication [communication] reflects the views only of the authors, and the Commission cannot be held responsible for any use which may be made of the information contained therein.

and bind it to the previous one. Let two more transactions occur after that. First, Ana sends $50 to Maria, and then Milan sends $450 to Goran. We also write these transactions in blocks and bind them into a chain with the previous blocks.



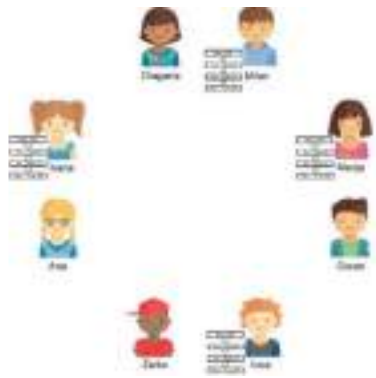
We can see that a blockchain simply represents a chain list of all transactions (chain of blocks) that have ever occurred within a single currency. Each cryptocurrency has its own blockchain - one large file with recorded transactions. For example, Bitcoin blockchain is a file that currently is around 300 GB large, in which all transactions between people that have ever happened related to Bitcoin are recorded.

4.1.Security and decentralization

Blockchain is revolutionary because of its safety [22]. Namely, once entered the chain, the transaction can no longer be deleted or changed in any way. Security is guaranteed in two ways - by connections (links) between blocks (more on that later) and by the fact that the general ledger is not centralized (as in the banks) but the complete blockchain file can be stored by anyone who wants it (distributed ledger). These volunteers are called nodes. Each of the nodes keeps a fresh copy of the file of this register of transactions. Every time a new transaction occurs, it is instantly updated with each of the nodes. There is no "central

This project has been funded with support from the European Commission. This publication [communication] reflects the views only of the authors, and the Commission cannot be held responsible for any use which may be made of the information contained therein.

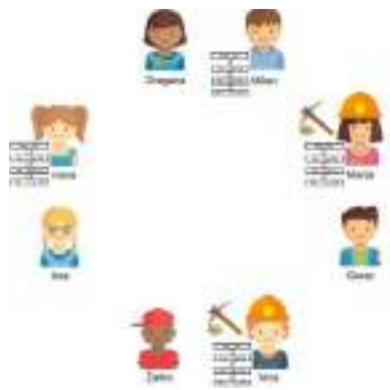
authority" because the blockchain is decentralized. In such a way, the intermediary (the trusted third party) is avoided, as well as its commissions, and the transfer of money is almost instantaneous. Not to mention that our data in the bank can be hacked (someone can steal our card data) or damaged, or the bank can go bankrupt...



The question now is how to ensure that all the copies of the registry are identical (synchronized) at all times on each computer, i.e. that all computers have the same version of the ledger.

4.2.Mining

The role of miners [23] is to confirm and validate transactions. Each node can apply to be a miner.



This project has been funded with support from the European Commission. This publication [communication] reflects the views only of the authors, and the Commission cannot be held responsible for any use which may be made of the information contained therein.

Let Goran want to send $100 to Ana. Goran initiates the transaction and his computer sends a notification of the desired transaction to all miners in the network (Marija and Ivica). At that point, it is an invalidated transaction, and, as such, cannot be entered into the ledger.

The miners first check whether Goran really sent a request for a given transaction, which will be discussed later (subtitle Privacy and authentication). After verifying the authenticity of the request, the miners also check if Goran has enough funds for the transaction, by going through the complete blockchain chain, checking all the transactions that Goran has ever had, and calculating if there is enough money available. So, with a blockchain, you cannot simply check the "account balance", but someone's balance is always recalculated, based on all the transactions he has ever had. This makes the blockchain extra secure. For someone to forcibly change someone's "available balance", they would have to change literally all the transactions from the past, which is practically impossible (we will see why later).

After these verifications, a race starts among the miners - which miner will be the first to validate the requested transaction and bind it into the blockchain, i.e. to add it as the next block in the chain. What does that actually mean?

The last validated block and the new block that the miner wants to write into the chain together form a cryptographic entity with only one part missing to be valid (will be explained in the part on SHA256). That fraction is a number called Nonce and there is no way for it to be calculated. The only way to find this number is for the computer to guess - to generate random numbers and test whether they "match". You need to try an incredible amount of numbers before finding the right one. So, the only way to find this number is for thousands of computers around the world to work on it at the same time - until someone guesses the number. On average, that takes about ten minutes. The first miner in the world to guess this number immediately sends a message to all other miners and other nodes that he has found the right number and what it is. Then all the other miners check if it is the correct number and, if so, stop searching further, enter the transaction block in their version of the chain and move on to looking for a random number for the next transaction, which arrived in the meantime. In this way, the register is always synchronized and identical for all the nodes.

4.3.Cryptographic hash function

A blockchain is a series of consecutive blocks interconnected by a cryptographic function. A cryptographic hash function [24] is a mathematical algorithm that maps data of arbitrary size into a series of bits of fixed length (hash) designed to act as a one-way function, i.e. a function that cannot be inverted. This means that it is not possible to get the initial - input data from the hash in any other way than by brute force search, i.e. by trying all possible input values until a matching hash output is obtained.

The basic elements of a block in a blockchain are the block number, transaction data, the hash of the previous block and the hash of the current block (the hash length is shortened in the figure due to space):

For instance, Bitcoin uses a hash function called SHA256 (Secure Hash Algorithm 2 made by the USA security agency NSA). This function for any data set generates a 64-digit string of hexadecimal numbers that acts as a fingerprint for the given input data. The SHA256 function is deterministic - it always produces the same output hash for the same input data, and the slightest change in input produces a completely different output hash. There is a total of:

$16^{64}=2^{256}$=115.792.089.237.316.195.423.570.985.008.687.907.853.269.984.665.640.564. 039.457.584.007.913.129.639.936 (78 digits) of possible hash values. Functioning of SHA256 can be checked at https://passwordsgenerator.net/sha256-hash-generator/. For instance, two similar inputs, different only in the size of first letter - Marko and marko produce completely different hash values:

This project has been funded with support from the European Commission. This publication [communication] reflects the views only of the authors, and the Commission cannot be held responsible for any use which may be made of the information contained therein.

48CF6C91C1D1EB864E1D7D6D38E65E6C02249083ACA96740BA82F2F123A18177 and 8C5FAF36CE0DAE48351F5E09C5133FDADDCF52D9BAF4369DB027766A12C1742F.

In our example, the hash of the current block is obtained when we enter into SHA256 function the block number, data stored in the block and the hash of the previous block, i.e.:

SHA256 (Block Number, Data, Previous Block Hash) ---> Hash

This shows how the blocks are connected - not only does the current block reference the hash of the previous block, but that hash directly affects the hash value of the current block. Therefore, if anyone messes up with the data in a block (for example, wants to write that some money has been transferred to him), it will spoil not only the hash of that block, but also the hashes of all subsequent blocks. So, it is impossible to change the information in the blockchain, ever. If a hacker changes one block, the hash changes, he has to change the hash of the next block, why does he need time, and during that time others have already added new blocks to the end of the chain... The only way to succeed with the fraud (to make his chain longest and therefore only valid chain) is to catch up with the others by having more than 50% of the total processing power in the world.

What is the role of the miner here?

Blocks in a blockchain have another field called Nonce (from *number used only once*):

This project has been funded with support from the European Commission. This publication [communication] reflects the views only of the authors, and the Commission cannot be held responsible for any use which may be made of the information contained therein.

Nonce is an integer number that together with the block number, data and previous hash becomes the input for the SHA256 function and calculation of the hash of the current block:

SHA256 (Block Number, Nonce, Data, Previous Block Hash) ---> Hash

Unlike the other components of the block, Nonce was created to be completely under our control. By changing it, we can change the hash of the current block without affecting other data within the block. Each time we select a different value for Nonce, we get a different value for the hash block.

There is a total of 16^{64} possible SHA256 cryptographic hash values, but not all of them are valid, because every two weeks the Bitcoin network defines a minimum value for a hash. Any value above will be rejected and below will be accepted. The current minimum value starts with 18 zeros which means there are 16^{46} valid hash values (64-18=46). This means that the probability that a randomly selected hash is valid is: $16^{46}/16^{64}=16^{-18}=$ 0.000000000000000000002%. This is the probability that some Nonce value will generate a valid hash for a given block. The miners are competing precisely over who will be the first to find Nonce that generates a valid hash for the current block. The miner who finds it first gets the right to bind the block in the chain and gets a reward.

The number of zeros at the beginning (minimum hash value) is determined every two weeks so as to ensure that on average each new block is added in approximately 10 minutes, taking into account the current computational power of all miners in the network. The more miners, the more zeros (making finding the right number more difficult). It is part of the monetary policy of Bitcoin algorithm to control the total amount of money in circulation.

This project has been funded with support from the European Commission. This publication [communication] reflects the views only of the authors, and the Commission cannot be held responsible for any use which may be made of the information contained therein.

So, what millions of miners are constantly doing is spinning values for Nonce in the hope of being the first to get a valid hash for the next block. When a valid hash is found, a block is added to the chain and a new race begins for the next block and the next prize.

All the important current info related to the Bitcoin can be found at: https://www.blockchain.com/explorer. It can be noted that the transactions are in Bitcoins, not in dollars, as in our example, and that each block does not contain only one transaction, but about 2,000 of them that have accumulated in approximately last ten minutes.

4.4.Rewarding the miners

The first miner in the world who manages to validate the transaction is rewarded - in crypto money, i.e. he receives a certain number of bitcoins. This reward currently amounts to 6.25 Bitcoins (initially it was 50BTC) or about $71k (the current value of Bitcoin is $11,388). These bitcoins are practically generated from nothing at the moment when the miner discovers the desired number. The bitcoin system itself has been designed from day one in such a way so that new bitcoins are created exclusively in this way - by mining. The award for miners is always known in advance, and it is halved every 4 years (for every 210,000 blocks). Nakamoto initially limited the algorithm so that the total number of bitcoins that can be created is 21 million. To date, about 88.2% of this projected amount of bitcoin has been generated, i.e. about 18.5 million. The more bitcoin in circulation, the smaller is the reward, so inflation is prevented. When the miners create 21 million bitcoins, this reward will be zero , and the miners will only earn from tips in transactions. Tips are an alternative way for miners to earn. Namely, the person who intends to execute the transaction can predetermine a small amount in Bitcoins as a reward for the miner who first validates the block with his transaction. In that way, he also motivates the miners to include his transaction in the current block they are processing.

The whole mining process looks a lot like a lottery. That is why people buy expensive computers that can check hash functions faster, which increases their chance of being the first in the world to validate a transaction, and thus earn more money. Of course, these powerful

This project has been funded with support from the European Commission. This publication [communication] reflects the views only of the authors, and the Commission cannot be held responsible for any use which may be made of the information contained therein.

computers consume a huge amount of electricity, so the miner must pay a large electricity bill for his earnings. The problem is even bigger with the so-called mining farms that are most often built in those countries that have low electricity prices. Currently, all bitcoin miners in the world consume electricity twice as much as whole Serbia.

Figure 14 Bitcoin farms in Mongolia

The chance to hit a winning number with a computer intended for mining today is about one in a million, i.e. it takes more than 30 years, on average, whereby it is necessary to pay the electricity bill during the whole period. That is the reason why almost no one mines bitcoins on their own anymore. Instead, miners join their computers into groups of computers that work together, so-called mining pools. This increases the chance of winning, but at the same time reduces the reward because it is now shared equally with all the computers in that pool.

Solving the described cryptographic problem - finding a Nonce number that gives the appropriate hash function of the block and passing that number and block to other nodes is otherwise called Proof of Work (PoW). This is proof to us that someone struggled to come up with the appropriate hash. Due to its large power consumption, the PoW mechanism is harmful to the human environment [25] and rewards people with more money and better computer equipment. It also leads to pooling and centralization instead of decentralization which was the basic idea of blockchain. Today, three largest pools can join and gain over 50%

This project has been funded with support from the European Commission. This publication [communication] reflects the views only of the authors, and the Commission cannot be held responsible for any use which may be made of the information contained therein.

of processing power, which allows them to take over the Bitcoin chain and start approving fraudulent transactions.

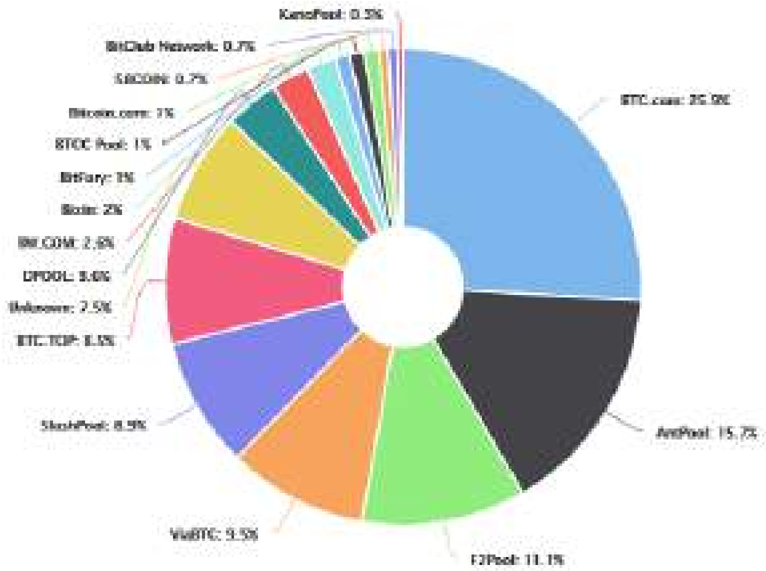


Figure 15 Bitcoin mining pools

That is why an alternative called Proof of Stake (PoS) was developed. With this mechanism, there is no situation where everyone tries to solve a cryptographic problem, but one node is randomly selected to validate the next block. Instead of miners, here we have validators, and instead of mining new blocks, we have forging (minting) of blocks. Validators are not chosen completely randomly, but the probability of their selection is determined according to a certain deposit in the money they have invested. Here, the rich are favoured linearly, while with the PoW mechanism, the more electricity they buy, the better the price from the manufacturer they obtain. Security with PoS is ensured by validators losing their pledge if they validate an incorrect block.

4.5.Privacy and authentication

So far, we have assumed that the blockchain contains information such as:

Goran forwarded $100 to Ana

This project has been funded with support from the European Commission. This publication [communication] reflects the views only of the authors, and the Commission cannot be held responsible for any use which may be made of the information contained therein.

Ivica forwarded $15 to Milan

In reality, the records are as in the following figure:

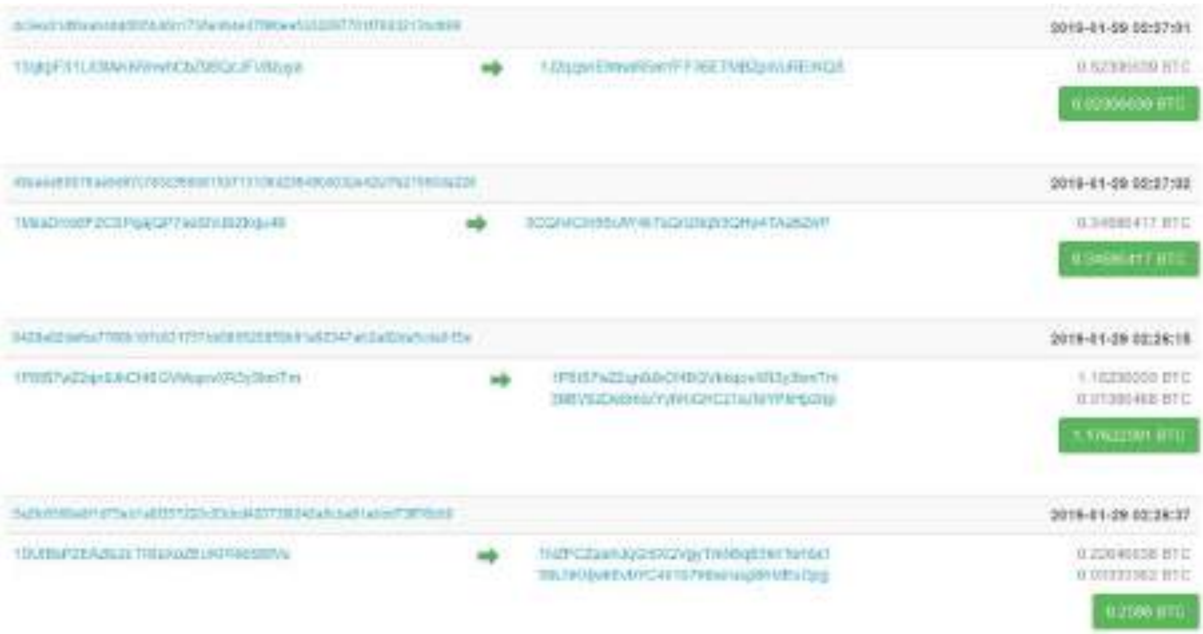


Figure 16 Block transactions

The real people are actually hiding behind these numbers, which are called addresses. The public can only see that some number code has sent money to another number code (similar to mail addresses). Each user, when entering the system for the first time, gets one long number called a private key that only he knows. The user can save it somewhere or keep it on a USB drive. If someone finds out your private key, they can steal all your money by transferring it to themselves. If you lose that key, all the money is permanently lost, and no one can access it anymore.

When you want to send or receive money, your computer generates another number called address or public key, based on your private key. Likewise, the recipient's computer generates another address based on his private key. Your address and the recipient's address are then sent to the miners for verification along with the amount being transferred. The public can see the addresses, but no one can use your address to find out your private key, and of course, there is no chance of finding out who you are.

This project has been funded with support from the European Commission. This publication [communication] reflects the views only of the authors, and the Commission cannot be held responsible for any use which may be made of the information contained therein.

Thus, a digital signature is a function of the content of the message (transaction) and the private key:

Digital signature = function (message content, private key)

If the content of the message changes, the digital signature is no longer valid. This disables the possibility of copying digital signature to another message. On the other hand, the signature is verified using the public key of the sender of the transaction:

Verification(message content, digital signature, public key) = True/False

The option of holding the private key on a piece of paper or USB does not seem very reliable. That is why there are specialized sites, i.e. applications that serve to help handle bitcoins. When you open an account (analogous to opening a bank account) on these sites (digital wallet), they create this private key for you, which even you cannot know, for security reasons. Instead, you get an address, which you can then use to send and receive bitcoin, also from the app itself.

4.6.Other applications (other than cryptocurrencies)

After the initial success of bitcoin, people began to think about the other possible applications of the blockchain technology [26]. They quickly realized that this is much broader than simple money transfers and that technology can be used wherever we have data and transactions of any kind, property information, areas where there is a high risk of fraud or corruption, sectors where there are intermediaries who take commissions and slow down activities. In all alternative applications, the only difference from Bitcoin is in the content of the blocks; instead of data on bitcoin transactions, some other data is entered in the blocks.

Recently, the so-called smart contracts are especially experimented with - they are like the normal contracts but in digital format. These are simple programs stored in blockchain blocks that are executed automatically if the certain conditions are met (usually some predefined bitcoin transfer is performed). Example - Kickstarter - crowdfunding platform: people with ideas present their idea, determine the amount of money they need for development and start collecting money from individuals who believe in their vision. Both sides trust a third party (Kickstarter) that in the case the amount is reached, they will transfer the money to the inventors or in case of failure, return the money to the investors. Smart contracts kick the middleman out of the game. The great advantage of smart contracts is their immutability (no one can cheat with already concluded contracts) and distribution (everyone can monitor the fulfilment of contract conditions). The most important platform based on the implementation of smart contracts is Ethereum and it was created in 2015. The platform should not be confused with a cryptocurrency, built on that platform, called Ether.

Examples of blockchains currently being experimented with:

- car register that prevents frauds with used cars. From time to time, the car sends mileage data that is entered into the blockchain together with the large services (automatically from a car mechanic), participations in major accidents (from the police), changes of ownership...

- real estate register, cadastres of change of ownership over an apartment, house, land ...

- monitoring of food or industrial products from the field or factory to the customer in the market; blockchain does not have to be human made, it can be created and read by sensors, robots, or programs

- voting, elections - registration and verification of voters and counting of votes

- hospital records of the patient

- copyright and payment to performers e.g. for songs listened to, intermediary applications that take a monthly membership fee are avoided - this can be done through smart contracts that record what we have listened to and proportionally distribute our membership fee to artists according to merit

- production and distribution of electricity without intermediaries

This project has been funded with support from the European Commission. This publication [communication] reflects the views only of the authors, and the Commission cannot be held responsible for any use which may be made of the information contained therein.

5. REFERENCES

[1] Facebook report (2020), avaialble at https://investor.fb.com/investor-news/press-release-details/2020/Facebook-Reports-Second-Quarter-2020-Results/default.aspx

[2] N. Marz, J. Warren (2014), "*Big Data: Principles and best practices of scalable realtime data systems*", Manning Publications Co., USA.

[3] NESSI White paper (2012), "*Big data: A new world of opportunities*", available at http://www.nessi-europe.com/Files/Private/NESSI_WhitePaper_BigData.pdf

[4] J. Mashey (1998), "*Big Data and the next wave of InfraStress*", avalailable at https://static.usenix.org/event/usenix99/invited_talks/mashey.pdf

[5] L. Reichlin, M. Watson (2003), "*Big Data: Dynamic factor models for macroeconomic measurement and forecasting*", Advances in Economics and Econometrics, Theory and Applications, Eighth World Congres, pp 115-122.

[6] E. Ularu, F. Puican, A. Apostu, M. Velicanu (2012), "*Perspectives on Big Data and Big Data analytics*", Database Systems Journal, vol. 3, no. 4, pp. 3-14.

[7] C. Baru, M. Bhandarkar, R. Nambiar, M. Poess, T. Rabl (2013), "*Benchmarking Big Data systems and the bigdata top100 list*". Big Data, vol. 1, no. 1, pp. 60-64.

[8] R. Higdon, W. Haynes, L. Stanberry, E. Stewart, G. Yandl, C. Howard, E. Kolker (2013), "*Unraveling the complexities of life sciences data*". Big Data, vol. 1, no. 1, pp. 42-50.

[9] D. Dietrich (2015). "*Data science and big data analytics: Discovering, analyzing, visualizing and presenting data*", John Wiley & Sons, USA.

[10] R. Turban, R. Rainer, R. Potter (2009), "*Introduction to information systems*", Data Status, Serbia.

[11] E. Curry (2016), "*The big data value chain: definitions, concepts, and theoretical approaches*", New Horizons for a Data-Driven Economy, pp. 29-37.

[12] R. Villars, C. Olofson, M. Eastwood (2011), "*Big data: What it is and why you should care*", White paper, vol. 14, pp. 1-14.

[13] A. Sharma, H. Pandey (2020), "*Big Data and Analytics in Industry 4.0*", In: Nayyar A., Kumar A. (eds) A Roadmap to Industry 4.0: Smart Production, Sharp Business and Sustainable

This project has been funded with support from the European Commission. This publication [communication] reflects the views only of the authors, and the Commission cannot be held responsible for any use which may be made of the information contained therein.

Development. Advances in Science, Technology & Innovation (IEREK Interdisciplinary Series for Sustainable Development). Springer, Cham.

[14] E. S. Brunette, R. C. Flemmer and C. L. Flemmer (2009) "*A review of artificial intelligence*", 2009 4th International Conference on Autonomous Robots and Agents, Wellington, 2009, pp. 385-392.

[15] J. van Leeuwen (2004), "*Approaches in Machine Learning*" In: Verhaegh W.F.J., Aarts E., Korst J. (eds) Algorithms in Ambient Intelligence. Philips Research, vol 2. Springer, Dordrecht.

[16] B. Li, B. Hou, B., W. Yu (2017) "Applications of artificial intelligence in intelligent manufacturing: a review". *Frontiers Inf Technol Electronic Eng*, 18, pp. 86–96 (2017).

[17] J. Hlaváč and J. Štefanovič (2020) "*Machine Learning and Business Intelligence or from Descriptive Analytics to Predictive Analytics*," ***Cybernetics & Informatics (K&I)***, Velke Karlovice, Czech Republic, 2020, pp. 1-4.

[18] P. Zheng, H. Wang, Z. Sang, "*Smart manufacturing systems for Industry 4.0: Conceptual framework, scenarios, and future perspectives*". *Front. Mech. Eng.* **13,** 137–150 (2018).

[19] X. Li, Y. Shi (2018), "*Computer Vision Imaging Based on Artificial Intelligence*", ***International Conference on Virtual Reality and Intelligent Systems (ICVRIS)***, Changsha, 2018, pp. 22-25.

[20] D. Tapscott, A. Tapscott (2016), "*Blockchain Revolution"*, Portfolio.

[21] A. Antonopoulos (2014) "*Mastering Bitcoin: Unlocking Digital Cryptocurrencies"*, O'Reilly Media.

[22] K. K. R. Choo, A. Dehghantanha, R. M. Parizi (2020), "*Blockchain Cybersecurity, Trust and Privacy"*, Springer.

[23] A. Norman (2017), "*Cryptocurrency Mining"*, CreateSpace Independent Publishing Platform.

[24] J. Hoffstein, J. Pipher, J. H. Silverman (2014), "*An Introduction to Mathematical Cryptography*", Springer.

[25] A. de Vries (2018), "*Bitcoin's Growing Energy Problem"*, Joule, Vol. 2, No. 5, pp 801-805.

[26] A. Bahga, V. Madisetti (2017), "*Blockchain Applications: A Hands-On Approach"*, VPT.

This project has been funded with support from the European Commission. This publication [communication] reflects the views only of the authors, and the Commission cannot be held responsible for any use which may be made of the information contained therein.