TECHNOLOGIES AND TOOLS USED FOR PREDICTION OF CHARACTERISTICS OF ACTIVITY OF INDUSTRIAL ENTERPRISES

Katherina M. Gritsyuk, Vera I. Gritsyuk

National Technical University "Kharkiv Polytechnic Institute“,
Kharkiv Institute of State University of Telecommunications
e-mail: prichkoel12@gmail.com
Ukraine

Abstract: Technologies of Industry 4.0 lead to changes in industry, society. Industry 4.0 – automation of industrial enterprises when modern technologies are used. Data analytics is playing a significant role in the development of industry and others spheres of life’s work of humans. Among use cases of predictive analytics for industry there is predictive maintenance that allows ensure the maximum level of reliability. Thanks to the development of the Internet of Things, Industrial Internet of Things, Big Data analytics, Machine Learning technologies, it became possible predict different characteristics with great accuracy, control them.

Key words: Data analytics, Big Data, Internet of Things, Machine Learning, predictive maintenance.

1. INTRODUCTION

In our time the data produced by humans and machines has reached high levels. Everyday large amounts of data, received from different sources and with different formats, are created and made available to organizations. The development of Big Data (BD) technologies and analytics, Internet of Things (IoT) has led to their greater use for industrial automation. At this time, the concept of Industry 4.0 (I4.0), which refers to the era of the fourth industrial revolution, appeared.

I4.0 has changed the approaches in the work of industry and significantly increases the efficiency of production processes.

The strategy of many countries is using of the computing technologies such as “BD analytics”, the technologies of Artificial Intelligence (AI). The apply of Data analytics, Machine Learning (ML), deep learning to problems in human health, physics, materials science, maintenance and others becomes an important tasks of our time.

Data analytics is playing a significant role in the development of smart industry. Data analytics application are generally divided into three groups: descriptive, predictive and prescriptive. Prediction of events that can occur in the future is a function of predictive analytics thanks to the development of predictive model. Predictive analytics can help improve customer relationships, operational processes and others. Capabilities of predictive analytics are unlimited: there are use cases for manufacturing, for retail, for healthcare and others. Among use cases of predictive analytics for industry there is predictive maintenance (PdM), for retail – pricing and others.

Any breakdown of machine or equipment may lead to disruption for the supply chain, so maintenance aims to reduce the number of failures occurred during production [1]. PdM is one of the most priority topics in maintenance. The data such as current consumption, voltage, vibration and others are collected and transmitted through an IoT gateway and other sources to a centralized storage, where the factory managers see them on a monitor in real-time.

The purpose of this work is: researching of different technologies, tools for predictions, control in industrial enterprises; researching of PdM, technologies, tools, which are applied for PdM; to make different predictions of indices of activity of different interprises, to find coefficients of the correlation R, coefficients of determination R^2 of obtained models and to find the model which is most adequate to experimental data.

Data analytics, ML and other technologies help predict different characteristics, control them, predict times of equipment failure and repair equipment with minimal downtime, i.e. increase income in systems using machines.

2. INDUSTRY 4.0

I4.0 is taking place in our time. Now we live in the time of the end of the third digital revolution, which began in the second half of the last century. Its specific traits are the development of information and communication technologies, automation and robotization of production processes. The specific traits of I4.0 are fully automated production facilities where controlling of all processes are carried out in real time and taking into account changing external conditions. Cyberphysical systems (CPS) are creating. They have a relationship and coordination between computing and physical resources. Computing resources control physical processes and make decentralized decisions. They are able to unite in one network, interact in real time, self-adjust and self-learn. An important role is played by IoT which provide communications between personnel and machines, machine and machine. Enterprises create products in accordance with the requirements of the individual customer, optimizing the cost of production.

The concept of digitalization, so popular over the past few years, refers to the stack of the 4th industrial revolution - the massive introduction of CPS in production and other areas of human life, work and leisure. The term Industry 4.0 got its name in 2011 from one of 10 projects of the state strategy for technological development of Germany until 2020. The German government in 2011 have brought into the world a new theme called Industry 4.0 - the fourth industrial revolution [2, 3] . I4.0 aim is to increase level of automatization achieving a higher level of operational productivity and efficiency [3], connecting the physical and virtual world. It will bring computerization and inter-connection into the industry. I4.0 can be understand as Cyber-Physical Systems production [4]. It is based on knowledge and data integration and it can be summed as an manufacturing process. This process is connected with algorithms, BD and high technologies - the IoT and Internet of Services (IoS), Industrial Automation, Cybersecurity (CS), Cloud Computing (CC) or Intelligent Robotics [5]. I4.0 is built not on one technology, but on combining data, tools and processes from different application areas in order to reduce overall costs, reduce risks and increase efficiency using CPS. I4.0 based on the following key technologies: BD and analytics; IoT; Cloud Computing; methods and tools of AI, including ML; virtual and augmented reality; 3D printing; quantum computing; blockchain. I4.0 means the universal digitalization of industrial production. I4.0 includes reasonable production and industrial use of AI. Predictive production maintenance is the element of I4.0.

The created CFS have many advantages. The real-time feedback loop, implemented by sensors, minimizes errors, provides control of the production line and makes it possible to take the necessary measures to eliminate all potential problems. The high speed of data access provides the making of more competent and informed decisions in real time. And this list of advantages goes on. The key advantages of I4.0 include: mass wares of production with individual characteristics; increasing of security; quality improvement; productivity increasing. The new forms of interaction between man and machine, machine and machine (M2M), for example, sensor interfaces and augmented reality system, appear; the methods to transfer digital commands to the physical world, for example, advanced robotics and 3D printing, improve. Supercomputers, smart homes, smart cities appear.

3. PREDICTIVE MAINTENANCE

I4.0 fundamentally changed the approach in work of the industrial business. One of the key requirements is to reduce equipment maintenance costs by increasing its efficiency, reducing downtime and increasing the productivity of technical specialists.

Untimely maintenance of industrial equipment is not only direct losses from breakdowns in the form of new parts costs. It is also a reduction in production volumes, costs of reorganization and re-planning of processes. In addition, disruption of supply plans, reputation losses among customers and partners, which lead to even greater costs in the future.

Currently, several methods of equipment maintenance have been formed, among them there are three main: reactive, planned, predictive [1].

The first type is the maintenance of equipment after its failure. This approach is justified when servicing simple, inexpensive equipment, when there is a reservation, and replacement will be cheaper than repair work to restore equipment.

The second type of maintenance is planned preventive maintenance of equipment in accordance with the regulations. In this case, maintenance is carried out in accordance with the manufacturer's recommendations at regular intervals. Equipment maintenance work is carried out with a certain period, which is determined by the methods of statistical analysis. When maintenance is on the regulations, the ability to use the manufacturer's warranty is not lost. But it turns out that at least 50% of all maintenance on the regulation are performed without their actual need.

The third type of maintenance is predictive [6, 7], also known as maintenance by actual technical condition. In this type of maintenance, the state of the equipment is controlled continuously or periodically. Depending on the results obtained, a prediction is made of the technical condition of the equipment and maintenance programs are formed. PdM systems are able to predict the state of the system based on the current state of the equipment and determine the necessary maintenance measures [8]. This improves maintenance efficiency and system productivity and reduces maintenance costs.

The I4.0 methodology offers the approach – remote control and monitoring (RCM) and PdM. PdM technology is based on maintenance methodology on the basis of reliability. PdM allows make not scheduled repairs, but when it becomes necessary. Due to this, it is possible, on the one hand, not to spend money and time on scheduled maintenance of equipment that can work normally a few more months without repair, and on the other hand, the probability of unplanned downtime caused by an unexpected breakdown is reduced.

This is achieved by: collection of data of the technical condition of equipment and their preprocessing; early fault detection; predicting the time of failure; maintenance planning; optimization of resources allocated for equipment maintenance. PdM is a method of predicting of failure by analyzing production data to identify patterns and predict issues before they happen. Most of all planned maintenance activities are ineffective. PdM has become a leading I4.0 use case for manufacturers and asset managers.

The following base components are required for PdM: sensors – sensors for data collection are installed on the physical product or machine; data communication – the communication system that allows data to flow between sensors, the monitored asset and the central data store, may be it is IoT facilities and other sources; central data store – the central data store in which asset data and business data are stored, processed and analyzed; predictive analytics – predictive analytics algorithms are applied to the data to identify patterns and generate insights; root cause analysis – data analysis tools are used by maintenance and process engineers to determine the corrective action to be performed [9].

As the IIoT develops due to installing on the equipment the various sensors, data about its technical condition can be collected not periodically, but continuously, without interrupting of the equipment exploitation. Timely detection of even small deviations in operating parameters will allow to quickly take measures to ensure the normal operation of the equipment. BD technologies will allow you to predict the time of failure with high accuracy.

Predictive analytics is based on statistical methods. It is important to understand that the predictive analytics system is related to BD too, AI, ML.

Key advantages of the PdM system: increased efficiency – analytics-driven approach improve overall equipment effectiveness (OEE) by reducing unnecessary maintenance, increase asset life and enable analysis of causes of a system failure; prevention unexpected failures; increased production safety; reduction of the number of accidents with negative impact on the environment; formation of an optimal set of spare parts and materials; reduced downtime; growth of income-manufacturers get profit when they propose customers services which based on analytics: PdM dashboards, optimized maintenance schedules or a technician dispatch service [10].

Data from IoT devices is not the only source. The ability to combine data from a large variety of sources for the most accurate predictions is the advantage of predictive maintenance. Other data sources might include: data from programmable controllers; manual data from human inspection; geographical data; manufacturing execution systems; static data, like manufacturer service recommendations for each asset; equipment usage history data; building management systems; external data from APIs, like weather; parts composition.

Any source of data can augment IoT data and be used to build and test a PdM model. It is necessary to determine which data sources are the best indicators of failure, wear, or breakdown and add or remove features to improve the final model.

Industrial AI can be applied to PdM in the manufacturing industry. We are in the beginning of using this technology, but there are already many advantages from industrial AI. It offers a variety of techniques to analyze the great amounts of data collected from the manufacturing process for creating the models. These techniques are called ML algorithms.

ML, deep learning are suitable for PdM data processing because the increasing number of sensor data streams makes manual monitoring and analysis impossible. PdM is already widely used in the production of aircraft and helicopters, and this method has already shown its advantage over traditional planned maintenance in practice, the use of predictive maintenance is actively spreading with the growing popularity of electric vehicles. Large oil platform operators use PdM to reduce downtime for pumping equipment. The initial data to the PdM system usually comes from the IoT systems, equipment which fitted out with the sensors that transmit information to the system for PdM, etc. BD specialists then use ML technology to creating models. This provides the planning of the existence functions of the researched object, taking into account the BD obtained in time of the researching of the object. PdM uses in follows spheres: railway; manufacturing; oil and gas industry and others.

4. BIG DATA AND INTERNET OF THINGS

Every day every person leaves a lot of traces. We sent a request to a search engine, walked along the street with a smartphone equipped with a GPS module, used the navigator, made a purchase in a store using a credit card, download music or installed the application - any of these actions generates a stream of information. If consider the number of people living on Earth, we understand, that a lot of information is accumulating. Even more data are produced by machines, whose working is based on digital technology or suppose the digitization of physical or chemical processes, as, for example, this happens at petrochemical enterprises. There is no generally accepted definition for the concept of “Big Data”. Sometimes this concept is means only the volumes of heterogeneous and quickly incoming digital information (over 100 GB per day) that cannot be processed with traditional tools. The term BD appeared in 2008 with the help of editor of “Nature” Clifford Lynch. Technologies for working with BD are a whole complex of various tools, approaches and methods of working with information.

Big Data (BD) is characterized by parameters that are denoted as 3V - according to the first letters of the English words volume, velocity and variety. For BD, the values of these parameters are high, and with time they only become higher. With the volume everything is obvious: the volume of BD is large and it is constantly growing. Over time two more “V” were added to three “V”: veracity and value (some call others “V”). Everything is not clear with veracity: with the growth of volumes and speed of data receipt, their quality and veracity are more difficult to control. So BD is not just a lot of data, but a quantity that has already gone to a new, previously inaccessible quality. Recently BD has found application in many such areas as smart homes, increasing the income of banks and retail chains, increasing the efficiency of a wide variety of industries.

Internet of Things. Already in 2018 the volume of world data traffic may exceed two zettabytes, and this growth is not associated with the use of the network by human users, but with an increase in the number of devices, machines and mechanisms connected to it, interacting both with a person and with each other. This is the so-called Internet of Things. The term "Internet of things" appeared in 1999, it was formulated by the American researcher Kevin Ashton. And the first thing that was connected to the Internet was presented in 1990: at the Interop exhibition a toaster connected via the Internet was shown. According to Cisco estimates, in 2009 the number of devices connected to the Internet exceeded the population of the Earth. At that moment the “Internet of people” became the

“Internet of things.” Now the number of devices connected to the network is about 10 milliards.

The segment of the IIoT is actively growing, where units and whole production facilities, transported goods and office equipment are combined into a single information space. The IIoT provides full machine interaction (M2M). It is a means of implementing I4.0, because this system combines software, computer networks, production facilities with integrated sensors in a single cycle for collecting and exchanging data.

5. ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING

Some definitions of AI are presented below. For the first time the term AI was used by computer scientist John McCarthy in 1956. “AI is the science and engineering of making intelligent machines, especially intelligent computer programs” [11]. M. Haenlein and A. Kaplan said, that AI is the ability of a system to correctly interpret external data, learn from such data and use the received knowledge to achieve specific goals and tasks using flexible adaptation [12]. AI - the property of intelligent systems to perform creative functions that are traditionally considered the prerogative of a person [13]. In Fig. 1 the relationship between AI and ML is shown.

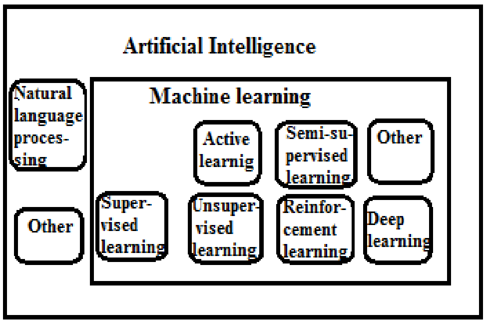


Fig. 1. The relationship between AI and ML

AI technology has reached true popularity today during increasing of data volumes, improvement of algorithms, optimization of computing power and data storage. The ability of AI to independently receive knowledge in the process of work is closely related to the problems of ML. This direction has been central from the beginning of the development of AI.

5.1. Machine Learning

The ML algorithms must analyze input (historical or a training set of data) and output data. A machine monitoring system includes different input factors from temperature to pressure and engine speed. The system predicts when a breakdown will occur. ML is a subset of AI that has been booming over the last decades, interfacing between theoretical computer science, statistics, and numerical methods. The most successful is the term "restoration of dependencies from empirical data" [14]. The main task of ML is to identify hidden patterns in the aggregate of the same type of objects based on a priori assumptions about the character of these patterns. The terms “Machine Learning” and “Artificial

Intelligence” are often confused with each other. ML is part of AI. ML is sometimes confused with predictive analytics (or predictive modeling). ML can be used for predictive modeling, but this is just one type of ML, and its application is wider than predictive modeling. ML is an AI subsection that explores methods that allow computers to improve their performance based on received experience. T. M. Mitchell gave the definition of the term “training”: It is said that a computer program is learning on the basis of experience E with respect to a some class of problems T and quality measure P if the quality of the solution of problems from T, measured on the basis of P, improves with getting of experience E [15]. Already in 1957 the first neural network model was proposed that realizing ML algorithms similar to modern. Currently a variety of ML systems are being developed for using in such technologies as the IoT, the IIoT, the concept “smart” city, when creating unmanned vehicles and many others. Siri from Apple, M from Facebook and Amazon Echo were created with the help of ML [16] . The scope of ML application is constantly expanding. Tasks in science, manufacturing, business, transport, healthcare of forecasting, control and decision-making often transfer to learning on precedents.

5.2. Formulation of the ML problem

Let X - a quantity of objects descriptions, Y - a quantity of admissible answers. There is an unknown target dependence – a representation $y^* : X \to Y$ whose values are known only on the objects of the finite training data set $X^m = \{(x_1, y_1),...,(x_m, y_m)\}$. It is required to construct an algorithm $a : X \to Y$ that would approximate the unknown target dependence both on the elements of the data set and on the whole quantity X. A pair $(x_i, y_i) \in X \times Y$ is called observation or precedent. Learning algorithms operate not with the objects themselves, but with their descriptions. The most common is a feature description. With this approach, the object x is represented as a vector $x = (x_1, x_2,..., x_m)$, where $x_j \in Q_j \left(j = 1,2,...m \right)$. In this way $X = Q_1 \times Q_2 \times ... \times Q_m$. A component x_j is called the j-th attribute or property, or feature of an object x.

It is also said that the algorithm should have the ability to generalize empirical facts, or derive general knowledge (regularity, dependence) from particular facts (observations, precedents). How to choose the algorithm that will satisfy us from the predictive model? For this, the so-called loss function is introduced. Empirical risk is the average error of an algorithm on the training set. The method of minimizing empirical risk (Empirical Risk Minimization, ERM) is a general approach to solving a wide class of training problems by precedents, primarily learning tasks with a teacher, including classification and regression tasks.

Loss Functions. A loss function is introduced $L(y, y')$ that characterizes the deviation of the response $y = a(x)$ from the correct response $y' = y^*(x)$ on a random object $x \in X$. Typical loss function selection: in classification task $L(y, y') = [y' \neq y]$, the indicator function is 1 if the condition is satisfied, 0- if it is not satisfied; in regression task $L(y, y') = (y' - y)^2$.

A functional is introduced that characterizes the average error (empirical risk) of an algorithm a on a random data set X^m

$$Q(a, X^m) = \frac{1}{m} \sum_{i=1}^{m} L(a(x_i), y^*(x_i)) \qquad (1)$$

The ERM is one of the most common approaches to learning algorithms on precedents. It consists in finding an algorithm in a given model of algorithms $A = \{a : X \to Y\}$ that minimizes the average error on the training data set

$$a = \arg\min_{a \in A} Q(a, X^m) \qquad (2)$$

Unfortunately, the small value of the functional on the training data set does not guarantee that the constructed algorithm will well restore the target dependency all over the space X. There is a danger of over-fitting or retraining when an attempt is made to describe specific data more accurately than the noise level in the data and the error of the model itself allows in principle.

5.3. Types and algorithms of Machine Learning

Standard types of ML tasks are following:

- Supervised learning (learning with a teacher) is the most general case. This learning method is used in cases where there is a large amount of data with markers (tags, labels). Each precedent is a pair of "object, response." It is required to find the functional dependency of the responses on the descriptions of the objects and build an algorithm that receives a description of the object at the input and produces a response at the output.

- The classification task is different in that the set of valid responses is finite. They are called the labels of classes.
- The regression task is different in that the valid response is a real number or a numerical vector.
- The task of ranking (learning to rank) is different in that the responses must be obtained right away on the quantity of objects, and then sort them by the values of the responses.

- Unsupervised learning (learning without a teacher). The task of the machine is to find the connections between the separate data, identify patterns, select templates, organize data or describe their structure, perform data classification.

- The task of clustering is to group objects into clusters using data on pairwise similarity of objects.

- The task of finding association rules (association rules learning). The initial data are presented in the form of feature descriptions. It is required to find such sets of features, and such values of these features that are especially often (not accidentally often) found in feature descriptions of objects.
- The task of dimensionality reduction is according to the initial features, using the conversion functions, go to the smallest number of new features without losing any essential information about the objects of data set.
- The task of filling in the missing values is to replace the missing values in the matrix of objects – attributes with their predicted values.

- Semi-supervised learning takes an intermediate position between learning with and without a teacher. The responses are known only on a part of precedents.

- Transductive learning. The final training set of precedents is given. It is required to make predictions on these partial data for the test data set.

- Reinforcement learning. The role of the objects is played by the pairs "situation, accepted decision", the responses are the values of the functional, characterizing the correctness of the accepted decisions (environmental reaction).

- Dynamic learning (online learning) can be either learning with a teacher or without a teacher. Precedents flow. It is required to immediately make a decision on each precedent and at the same time finish learning the model of dependency taking into account new precedents.

- Active learning is different in that the learner has the opportunity to independently assign the next precedent, which will become known.

- Neural networks and deep learning. At the end of the 20th century more attention began to be paid to artificial neural networks (ANNs). ANN is a system of connected and interacting artificial neurons based on simple processors. Usually neurons are located in the artificial network by levels, or layers. Input, or neurons of the first level, receive data from outside. Then neurons process impulses and transmite them to the neurons of the next level, or hidden. And so on, until impulses reach the output level. The scheme of a neural network, where: x1 ... xn are input parameters; Input- input layer; Hidden-hidden layer, is shown in Fig. 2. Deep learning applied to complex ANNs.

Tasks that can be solved by ML are following: forecasting: demand, sales, stock filling, equipment loading, further development of the enterprise, etc.; detection: trends, hidden relationships, anomalies, repeatable elements, etc.; recognition: photo, video, audio content, fraud attempts, lie, internal threats, external attacks on the security system, etc.; automation: the work of operators in online chats, telephone operators, etc.; classification: analysis of customers, clients and their segmentation according to various parameters; clustering; chat bot development.

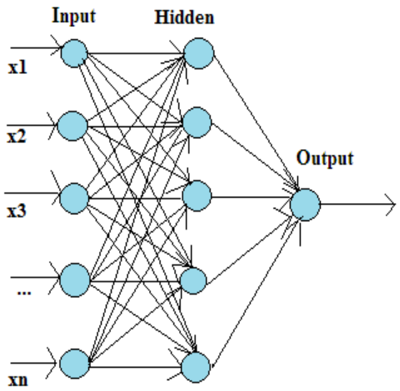


Fig. 2. Neural network scheme

Some classes of ML algorithms are presented below. Regression algorithms are used for statistical analysis. Regression algorithms can evaluate the strength of relation between variables in a data set. Regression analysis can be useful for predicting the future values of data based on historical values. Regularization is used to avoid the problem of over-fitting such a technique to modify models. It is possible to apply regularization to any ML model

The Bayes method (Naive Bayes, NB) refers to probabilistic classification methods [17, 18]. The advantages of the method are as follows: high speed of work, support for incremental learning, simple implementation of the algorithm in the form of a program, easy interpretability of the results of the algorithm. The Bayes method has disadvantages.

The method of k-Nearest Neighbors (KNN) refers to metric methods and is considered the simplest classifier [19]. An object is assigned to the class that is most common among the neighbors of this element. The disadvantages include: insufficient productivity in real tasks, because the number of neighbors used for classification will be big enough; difficulty in finding suitable weights and determining which features are necessary for classification; dependence on the chosen metric of distance between examples.

The Support Vector Machine (SVM) is a linear classification method, currently called one of the best methods [20]. The potential disadvantages of the support vector machine are as follows: the impossibility of calibration the probability of hitting the specific class, suitable only for solving problems with two classes, the model parameters are difficult to interpret.

The Decision Trees method refers to logical classification methods [21]. A decision tree is an acyclic graph used to classify documents described by a set of characteristics.

Dimension reduction helps systems remove data that are unnecessary for analysis. The algorithms are used to remove needless data, outliers, and other non-useful data.

Neural networks are actively used in connection with the appearance of large volumes of data and large computational capabilities [22]. Their efficiency is enough high. Deep learning is being used now in different applications.

Some ML tasks are presented below.

The **regression model** is presented: $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + ... \beta_m x_m + \varepsilon$, where ε - the error of the empirical value from the “true” line , the coefficients $\beta_k\, (k = 0,1,...m)$ - the regression parameters or the coefficients of regression. Then the forecasted value $\hat{y}$ of value y is presented:

$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2 x_2 + ... + \hat{\beta}_m x_m$. In the matrix form it is obtained: $Y = X\beta + \varepsilon$.

For find the vector β RSS, residual sum of squares, should be minimal: $RSS = (Y - X\beta)^T (Y - X\beta) = \varepsilon^T \varepsilon \to \min$.

Coefficients $\hat{\beta}$: $\hat{\beta} = \arg\min_{\beta} \sum_{i=1}^{n} \left(y_i - \beta_0 - \sum_{j=1}^{p} x_{ij} \beta_j \right)^2$. $\hat{\beta} = \left(X^T X \right)^{-1} X^T Y$

In **ridge regression**: $\hat{\beta}^{rid} = \arg\min_{\beta} \left\{ \sum_{i=1}^{n} \left(y_i - \beta_0 - \sum_{j=1}^{p} x_{ij} \beta_j \right)^2 + \lambda \sum_{j=1}^{p} \beta_j^2 \right\}$. It is obtained $\hat{\beta}^{rid} = \left(X^T X + \lambda I \right)^{-1} X^T Y$.

In **Lasso regression**: $\hat{\beta}^{las} = \arg\min_{\beta} \left\{ \frac{1}{2} \sum_{i=1}^{n} \left(y_i - \beta_0 - \sum_{j=1}^{p} x_{ij} \beta_j \right)^2 + \sum_{j=1}^{p} \left| \beta_j \right| \right\}$.

The **logistic regression** model: $\log\left(\frac{\mu(x)}{1 - \mu(x)} \right) = \alpha + \beta_1 X_1 + ... + \beta_p X_p$, where binary value $\mu(x) = \Pr(Y = 1 | X)$. It is necessary to find the maximum of log-likelihood (in the two-class case): $l(\beta) = \sum_{i=1}^{n} \left\{ y_i \beta^T x_i - \log\left(1 + e^{\beta^T x_i} \right) \right\}$. $\beta = \{\beta_{10}, \beta_1\}$.

MARS is multivariate adaptive regression splines. MARS model uses piecewise basis functions in the form $(x - t)_+$ and $(t - x)_+$. It is obtained:

$(x - t)_+ = \begin{cases} x - t, \; if \;\; x \succ t, \\ 0, \; otherwise, \end{cases}$ and $(x - t)_+ = \begin{cases} t - x, \; if \;\; x \prec t, \\ 0, \; otherwise \end{cases}$.

The collection of basis function is: $\gamma = \left\{ (X_j - t)_+ , (t - X_j)_+ \right\}_{\substack{t \in \{x_{1j}, x_{2j}, ... x_{nj}\} \\ j = 1,2...p}}$.

The MARS model has the form: $f(X) = \beta_0 + \sum_{m=1}^{M} \beta_m h_m(X)$, Each $h_m(X)$ is a function in γ or combination of one or two such functions. β_m are find from minimizing the RSS from standard linear regression.

6. EXPERIMENTS AND RESULTS

Let's consider the predictions of the cost of sales, net income of different interprises. It is regression analysis applied for this. Operating costs and costs of sales of PJSC "Wimm-Bill-Dann Ukraine" (Kyiv region) are presented in Table 1.

Table 1. Operating costs and costs of sales of PJSC "Wimm-Bill-Dann Ukraine"

Years / Costs, uah	2009	2010	2011	2012	2013	2014	2015
Material costs	453272	522895	711787	660252	734600	793749	875404
Labor costs	57151	72426	57270	14584	75836	73355	79290
Deduction for social needs	18142	24824	29688	33973	26012	24784	27260
Amortization	25836	27746	35646	53193	64143	34535	61679
Other opera-ting expenses	68878	159658	184341	287002	139879	107783	111041
The cost of sales	514860	768845	830025	887780	988363	982500	1106766

Source: data of PJSK "Wimm-Bill-Dann Ukraine"

The regression equation describing the relationship between the cost of sales and operating costs: material costs, labor costs, deductions for social needs, amortization, other operating expenses, - has the form:

$$y = a_0 + a_1 x_1 + a_2 x_2 + a_3 x_3 + a_4 x_4 + a_5 x_5 \qquad (3)$$

After the analysis had carried out the equation is obtained [23]:

$$y = 8027{,}264971 + 1{,}000955 x_1 + 2{,}8936104 x_3 + 2{,}40821339 x_5 \qquad (4)$$

The coefficient of determination R^2: $R^2 = \sum_{i=1}^{n} (\hat{y}_i - \bar{y})^2 / \sum_{i=1}^{n} (y_i - \bar{y})^2$.

The multiple coefficient of the correlation R is 0,9566841, the coefficient of determination R^2 is 0,91524451, which indicates that the model can be considered adequate to experimental data and based on this model it is possible to conduct forecasting of the cost of sales of this enterprise.

Let's consider the prediction of net income of the enterprise PJSK "Evrocement-Ukraine" (Kharkiv region). The output data are presented in Table 2. The analysis is carried out, and the regression equation is obtained:

$$y = 83192{,}21 + 1{,}07089 x \qquad (5)$$

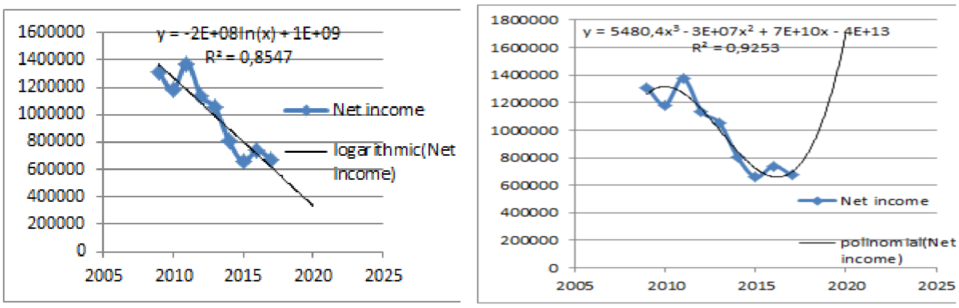
where y-Net income, x-costs of sales.

Table 2. Net income and the costs of sales of PJSK "Evrocement-Ukraine"

Years / Costs, uah	2009	2010	2011	2012	2013	2014	2015	2016	2017
Net Income	1306463	1182794	1374123	1134838	1053014	804662	659392	738484	673427
Costs of sales	10677948	1115900	1199375	921214	936110	690623	579533	568392	557981

Source: data of PJSK "Evrocement-Ukraine"

Let's build the dependence of net incomes from the years. The prediction of net income of the enterprise during 2018-2020 is obtained. It is presented on Fig. 3: a), b).



a) b)

Fig. 3. Predicted level of net income of PJSK "Evrocement-Ukraine"
a) logarithmic trend line, b) polynomial trend line.

Analysis of this solutions show that the most adequate to experimental data is a model which presented the polynomial trend line due to analysis of coefficients R^2. It is obtained: for polynomial trend line $R^2 = 0{,}9253$, for logarithmic trend line - $R^2 = 0{,}8547$.

The obtained regression models and trends allow predict seeking characteristics of activity of enterprises during several years ahead.

The managers then analyze and make decisions about the project, changing of some parameters for increasing of enterprises income.

7. CONCLUSIONS

The using of I4.0 technologies such as IoT, BD analytics, ML will increase productivity, reduce costs and reduce customer service time. CPS arises in which embedded computers and networks check and control physical processes, with feedback loops where physical processes influence on computation and back to front. The new forms of interaction between man and machine, machine and machine appear; the methods to transfer digital commands to the physical world improve. As presented in this work, increasing the productivity of industrial enterprises, reducing downtime is achieved through the using of predictive maintenance in enterprises - one from use cases of predictive analytics for industry. Large amounts of data received from sensors, IIoT devices and other sources are used. Such technologies as BD analytics, ML and deep learning, help build models with which it is possible to get highly accurate predictions about different characteristics of activity of enterprises, the failure of equipment and others. Then, managers make decisions about changing of some parameters of enterprise activity, repairing equipment or it parts and others.

The predictions of indices of activity of different enterprises are executed with the help of regression analysis, the model which most adequate to experimental data is found. Understanding the principles of ML, Data analytics will help in building models that allow get high precision predictions about different characteristics, equipment failures and others and then control them. This will cause an increasing of production efficiency, an increasing of enterprise income, etc.

REFERENCES

[1] Lee, C. K.., Ng, K. H. , Cao, Yi. Big Data Analytics for Predictive Maintenance Strategies. *Computers&Operations Research*, 2017, http://doi.org/10.4018/978-1-5225-0956-1.ch004.

[2] Wagner, T., Herrmann, C., Thiede S. *Industry 4.0 Impacts on Lean Production Systems, Procedia CIRP 63*, 2017, pp. 125–131. http://doi.org/10.1016/j.procir.2017.02.041.

[3] Lu, Y. Industry 4.0: A survey on technologies, applications and open research issues, *Journal of Industrial Information Integr*ation, 2017, pp, 1–10. http://doi.org/ 10.1016/j.jii.2017.04.005.

[4] Weyer, S., Schmitt, M., Ohmer, M., Gorecky, D. Towards Industry 4.0 – Standardization as the crucial challenge for highly modular, multi-vendor production systems, *IFAC-Papers OnLine*, **3** (vol. 48), 2015, pp. 579-584, https://doi.org/10.1016/j.ifacol.2015.06.143.

[5] Peruzzini, M., Grandi, F., Pellicciari, M. Benchmarking of Tools for User Experience Analysis in Industry 4.0, *Procedia Manufacturing*, 11, 2017, pp. 806–813, https://doi.org /10.1016/j.promfg.2017.07.182.

[6] Rawi, Zaid. Machinery Predictive Analytics. *SPE Intelligent Energy Conference and Exhibition* , March 23–25, Utrecht, Netherlands, 2010.

[7] Stone, P. Introducing Predictive Analytics: Opportunities. *Digital Energy conference and exhibition*. April 11–12, Houston, Texas, USA, 2007, https://doi.org/10.2118 /106865-MS, 2007.

[8] Vlasov, A. I., Yudin, A. V., Salmina, M. A., Shakhnov, V. A., Usov, K. A. Design Methods of Teaching the Development of Internet of Things Components with Considering Predictive Maintenance on the Basis of Mechatronic Devices. *International Journal of Applied Engineering Research*. **20** (vol. 12), 2017, pp. 9390–9396.

[9] http://www.seebo.com.

[10] http://www.faststreamtech.com.

[11] McCarty, John. What is Artificial Intelligence, Stanford University, 2007.

[12] http://www.sciencedirect.com

[13] Averkin, A. N., Gaaze-Rapoport, M. G., Pospelov D. A. Tolkovyj slovar' po iskusstvennomu intellektu.- M.: Radio i svyaz, 1992, 256 p.

[14] Vorontsov, K.V. Teoriya nadezhnosti obucheniya po pretsendentam. M., 2010. 316 P.

[15] Mitchell, T. Machine Learning. McGraw Hill. p. 2. ISBN 978-0-07-042807-2, 1997.

[16] Lewis-Kraus G. Prologue: You Are You Have Read. *The New York Times Magazine*, 2016.

[17] Tang, B. et al. A Bayesian classification approach using class-specific features for text categorization. *IEEE Transactions on Knowledge and Data Engineering*, **6** (vol. 28), 2016, pp. 1602–1606.

[18] Yoo, J. Y., Yang, D. Classification scheme of unstructured text document using TF-IDF and naive bayes classifier. *Advanced Science and Technology Letters*, vol. 3, 2015, pp. 263–266.

[19] Bijalwan, V. et al. KNN based machine learning approach for text and document mining. *International Journal of Database Theory and Application*, **1** (vol. 7), 2014, pp. 61–70.

[20] Lilleberg, J., Zhu, Y., Zhang, Y. Support vector machines and word2vec for text classification with semantic features., *2015 IEEE 14th International Conference on Cognitive Informatics & Cognitive Computing*, 2015, pp. 136–140.

[21] Pliakos, K., Geurts, P., Vens, C. Global multi-output decision trees for interaction prediction. *Machine Learning*, 2018, pp. 1257–1281.

[22] Abadi, M. et al. TensorFlow: A System for Large-Scale Machine Learning. *OSDI'16*, vol. 16, 2016, pp. 265–283.

[23] Gritsyuk, K. Forecasting of production cost and other indices of activity of industrial enterprise. *Technological audit and production reserves*. **3** (vol. 2), 2017, pp. 47-52. http://doi.org/10.15587/2312-8372.2017.103150.

Information about authors:

Katerina M. GRITSYUK, Ph.D., National Technical University "Kharkiv Polytechnic Institute", Kharkiv, Ukraine. She is lecturer and researcher. Her scientific interests include: regression analysis, artificial intelligence, knowledge-based systems, machine learning, prediction of different indices.

Vera I. GRITSYUK, Ph.D., Associate Professor. Kharkiv Institute of State University of Telecommunications, Ukraine. Number of publications in Ukrainian editions is 47 and number in foreign indexed editions is 9. Her scientific interests include: regression analysis, modeling, prediction, stochastic control systems.

Manuscript received on 07 September 2020
Revised manuscript received on 08 October 2020