Data mining algorithms advancing deep machine learning in engineering industrial control system programs [1]

Dr Lalamani Budeli

Abstract

The emergence of machine learning which enables a system to learn from data rather than through explicit programming allows industrial control systems to improve their complex control performance. Machine learning requires that the right set of data be applied to a learning process and big data can help improve the accuracy of machine-learning models possible to virtualize data so it can be stored most efficiently and cost-effectively. Electrical and computer engineers work at the forefront of technological innovation, contributing to the design, development, testing, and manufacturing processes for new generations of devices and equipment. As these professionals strive for innovation, their pursuits may overlap with the rapidly expanding applications for artificial intelligence.

Data mining is a computational technique or process of discovering patterns in large data sets and values involving machine learning, mathematical, statistics, and database system. We can compare both algorithms based on those data set records and find the best classification algorithms. Data mining solves the problem by analyzing a large amount of available data by providing useful patterns and rules using some classification methods.

Historically, most machinery and engineering components used in manufacturing and the operation of power plants, water and wastewater plants, transport industries, and other critical infrastructures were dumb, and those that were computerized typically used proprietary protocols. The networks they belonged to were air-gapped and protected from the outside world. This has changed over the years and components of today's ICSs are often connected directly or indirectly to the internet.

Keywords: Data mining, machine learning, industrial control systems

Introduction

Industrial control system (ICS) is a general term used to describe the integration of hardware and software with network connectivity to support critical infrastructure. ICS technologies include, but are not limited to, supervisory control and data acquisition (SCADA) and distributed control systems (DCS), industrial automation and control systems (IACS), programmable logic controllers (PLCs), programmable automation controllers (PACs), remote terminal units (RTUs), control servers, intelligent electronic devices (IEDs) and sensors.

[1] How to cite this paper: Budeli, L. (2020). Data mining algorithms advancing deep machine learning in engineering industrial control system programs, *PM World Journal*, Vol. IX, Issue XII, December.

Recent progress in areas like machine learning and natural language processing have affected almost every industry and area of scientific research, including engineering. Machine learning and electrical engineering professionals leverage AI to build and optimize systems and also provide AI technology with new data inputs for interpretation. Besides, harnessing artificial intelligence's potential may reveal chances to boost system performance while addressing problems more efficiently: AI could be used to automatically flag errors or performance degradation so that engineers can fix problems sooner. Electrical and computer engineering leaders have opportunities to realign how their organizations manage daily operations and grow over time.

Research hypothesis

The following are the research hypothesis

1. Data mining techniques can improve the machine learning algorithm of the industrial control system
2. Data mining and machine learning projects success will improve industrial control systems intelligence, creating a competitive advantage for the firms avoiding machine damage.

Control theory in engineering

According to Dullerud and Paganini (2013-12)**,** control theory is an interdisciplinary branch of engineering and mathematics that deals with the behavior of dynamical systems with inputs, and how their behavior is modified by feedback. Tin and Poon (2005-17) said that the typical objective of control theory is to control a system called the plant, so that its output follows a desired control signal, called the reference, which may be a fixed or changing value. The figure below indicates a closed-loop control system.

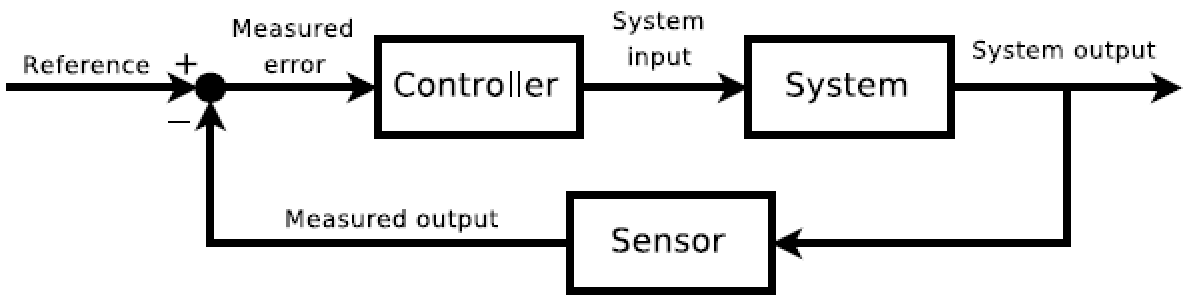

Figure 1 Industrial control systems (Prett and García ,2013-287)

According to Lyshevski (2012-234), a controller is designed to monitors the output and compares it with the reference. Prett and García(2013-289) whispered that the difference between actual and desired output, called the error signal, is applied as feedback to the input of the system, to bring the actual output closer to the reference. Zelikin and Borisov (2012-110). indicated that the important topic studied in control theory is stability (whether the output will converge to the reference value or oscillate about it), controllability and observability. Harjunkoski, Nyström and Horch (2009-1913) held that the transfer function, also known as the system function or network function, is a mathematical representation of the relation between the input and output based on the differential equations describing the system. Papadopoulos, Tanzman, Baker, Belliardi, Dube and Schneider Automation Inc (2000-22) whispered that the central idea of these control systems is the feedback loop, the controller affects the system output, which in turn is measured and fed back to the controller. According to Liu and Daley (2001-1197), the control system can be classified as:

- Linear versus nonlinear control theory - Linear control theory applies to systems made of devices that obey the superposition principle, which means roughly that the output is proportional to the input. Nonlinear control theory covers a wider class of systems that do not obey the superposition principle and applies to more real-world systems because all real control systems are nonlinear.
- Frequency domain versus time domain – in the frequency domain, the values of the state variables, the mathematical variables representing the system's input, output and feedback are represented as functions of frequency. Using time-domain state-space representation, the values of the state variables are represented as functions of time.
- SISO vs MIMO - Control systems can be divided into different categories depending on the number of inputs and outputs. Single-input single-output (SISO) is the simplest and most common type, in which one output is controlled by one control signal. Multiple-input multiple-output (MIMO) is found in more complicated systems.

Venugopal and Bernstein (1996-1075) indicated that the modern control theory utilizes the time-domain state-space representation, a mathematical model of a physical system as a set of input, output and state variables related by first-order differential equations. Aoki (2013-256) thought that to abstract from the number of inputs, outputs and states, the variables are expressed as vectors and the differential and algebraic equations are written in matrix form (the latter only being possible when the dynamical system is linear).

Process control concepts

According to Valenzuela et al (2005-309), Industrial control system (ICS) is a collective term used to describe different types of control systems and associated instrumentation, which includes the devices, systems, networks, and controls used to operate and/or automate industrial processes. Malhotra and Singh (2011-41) supposed that depending on the industry, each ICS functions differently and are built to electronically manage tasks efficiently. Seborg, Mellichamp, Edgar and Doyle III (2010-288) believed that today the devices and protocols used in an ICS are used in nearly every industrial sector and critical infrastructure such as the manufacturing, transportation, energy, and water treatment industries. Raab, Ferdowsi et al (2011-4) held that there are several types of ICSs, the most common of which are Supervisory Control and Data Acquisition (SCADA) systems, and Distributed Control Systems (DCS). Odgaard and Stoustrup

(2014-1225) assumed that local operations are often controlled by so-called Field Devices that receive supervisory commands from remote stations.

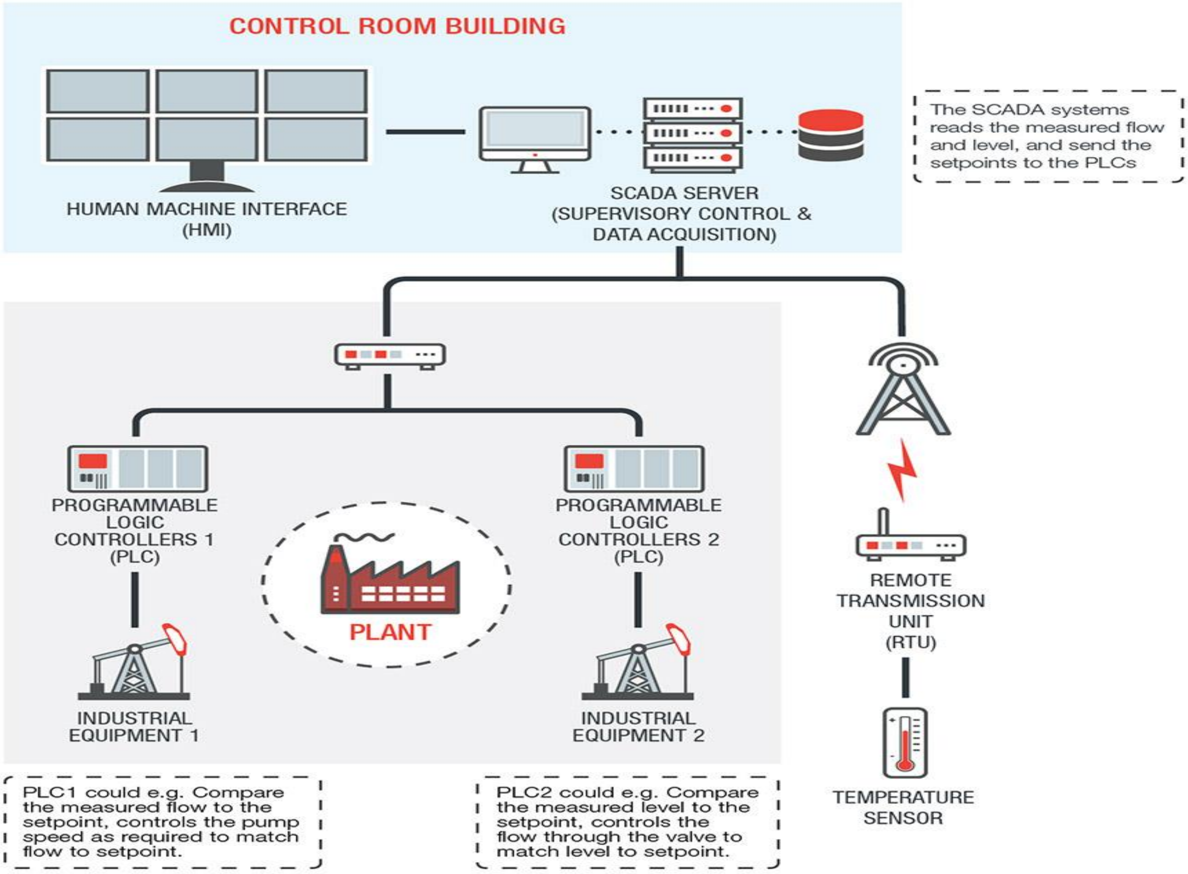

Figure 2 SCADA systems (Trendmicro (2018/2019) SCADA system and network protection [online]. https://documents.trendmicro.com/images/tex/articles/ics-system.jpg (4 August 2020)

Types of Industrial Control Systems

Supervisory Control and Data Acquisition (SCADA)

According to Ahmed et al (2012-233), SCADA is not a system that can provide full control. Instead, its capabilities are focused on providing control at the supervisory level. SCADA systems are composed of devices (generally Programmable Logic Controllers (PLC) or other commercial hardware modules) that are distributed in various locations. Barr (2004-17) indicated that SCADA systems can acquire and transmit data, and are integrated with a Human Machine Interface (HMI) that provides centralized monitoring and control for numerous process inputs and outputs.

Distributed Control System (DCS)

Deiningeret al (2010-5) supposed that each DCS uses a centralized supervisory control loop to manage multiple local controllers or devices that are part of the overall production process. This gives industries the ability to quickly access production and operation data. And by using multiple devices within the production process, a DCS can reduce the impact of a single fault on the overall system. Barenji, Barenji and Hashemipour (2014-1784) indicated that a DCS is also commonly used in industries such as manufacturing, electric power generation, chemical manufacturing, oil refineries, and water and wastewater treatment.

Information Technology and Operational Technology (IT and OT)

Han et al (2015-2509) alleged that the convergence of IT and OT provides enterprises greater integration and visibility of the supply chain– which include their critical assets, logistics, plans, and operation processes. Having a good view of the supply chain helps organizations remain competitive. Voba et el (2017-150) suspected that on the flip side, the convergence of OT and IT allows easier access to these two components that are targets of cybercriminals. In many organizations, OT infrastructure is at best poorly protected against cyber attacks.

Programmable Logic Controller (PLC)

Liu et al (2014-27) thought that in a SCADA system, a PLC provides the same functionality as Remote Terminal Units (RTU). In DCS, PLCs are used as local controllers within a supervisory control scheme. PLCs are also implemented as primary components in smaller control system configurations. Bayindir and Cetinceviz (2011-325) whispered that a PLC has many input terminals, through which it interprets signal high and low logical states from sensors and switches. It also has many output terminals, through which it outputs high and low signals to power lights, solenoids, contactors, small motors, and other devices lending themselves to on/off control.

Remote Terminal Unit (RTU)

Jusoh et all (2013-572) purported that a remote terminal unit (RTU) is a microprocessor-controlled electronic device that interfaces objects in the physical world to a distributed control system or SCADA (supervisory control and data acquisition system) by transmitting telemetry data to the system and employing messages to control the connected objects from the supervisory system. Aamir et al (2013-1010) apprehended that automation technology is rapidly advancing, while the growing demands of utility owners for more cost-effective control systems must be met. Idachaba and Ogunrinde (2012-158) believed that new communication techniques, devices, and standard protocol interfaces, combined with the immense computing power of today's hardware components, open the way to new concepts in the automation industry

Control Loop

McMillan (2014-23) believed that a controller seeks to maintain the measured process variable (PV) at a set point (SP) despite unmeasured disturbances (D). The major components of a control system include a sensor, a controller and a final control element. Jelali (2012-157) understood that every control loop consists of hardware such as PLCs and actuators. McMillan (2014-17) assumed that the control loop interprets signals from sensors, control valves, breakers, switches, motors, and other similar devices. The variables measured by these sensors are then transmitted to the controller to carry out a task and/or complete a process.

Human Machine Interface (HMI)

Shi et al (2019-356) understood that HMI includes any device or software that allows you to interact with a machine. This can be as simple and ubiquitous as the traditional single-touch display mounted on a machine or as technologically advanced as a multi-touch-enabled control panel or even connected mobile technology such as smartphones and smartwatches. Meinherz et al (2012-8) thought that HMIcan also display status information and historical data gathered by the devices in the ICS environment. It is also used to monitor and configure setpoints, control algorithms, and adjust and establish parameters in the controllers.

Remote diagnostics and maintenance

Hessmeret al (2012-14) supposed that a major element of reliability and availability will be linked to touchless maintenance and operation. Internet of things (IoT) devices and gateways will need to operate for long periods without physical (hands-on) maintenance or technical support to resolve problems. Grubic (2014-31) whispered that the only alternatives to hands-on servicing are either self-diagnostics and repair or remote diagnostics and resolution by the system itself or by human operators. Hessmeret al (2012-17) believed that remote diagnostics and management will be a combination of automated and semiautomated capabilities that are supplied by multiple, coordinated parts of the IoT service.

Control server

Batni et all (2011-11) believed that organizations can securely deploy, design, and maintain advanced computing infrastructure by incorporating multiple supervisory control applications into one easy-to-maintain virtualized server. Control Server contains the highest capacity plant SCADA and outcome optimizing platform. Senga and Toshiba Corp (2011-2) purported that the control server uses server-class hardware to integrate the features associated with engineering and operator workstations, historians, and advanced communication gateways with new optimizing capabilities. Watanabe and NEC Corp (2016-22) indicated that a control server hosts the DCS or PLC supervisory control software and communicates with lower-level control devices.

Intelligent Electronic Device (IED)

Bradetich et all (2013-3) specified that a smart device capable of acquiring data, communicating with other devices, and performing local processing and control. The use of IEDs in control systems like SCADA and DCS allows for controls at the local level to be done automatically. Intelligent electronic devices (IEDs) have been deployed extensively in power automation systems recently, and the shift from RTUs to IEDs is evident due to the integration and interoperability features of the IEDs. Spanier et al (2015-4) signposted that IED brings a relay panel with many single-function electromechanical relays, control switches, extensive wiring, and much more into a single box.

Data Historian

Huang et al (2014-22) indicated that a data historian is a centralized database for logging all process information within an ICS environment and then exporting data to the corporate IS. McGreevy et al (2010-17) believed that the data gathered is then used for process analysis, statistical process control, and enterprise-level planning. The figure below shows the data historian on a network.

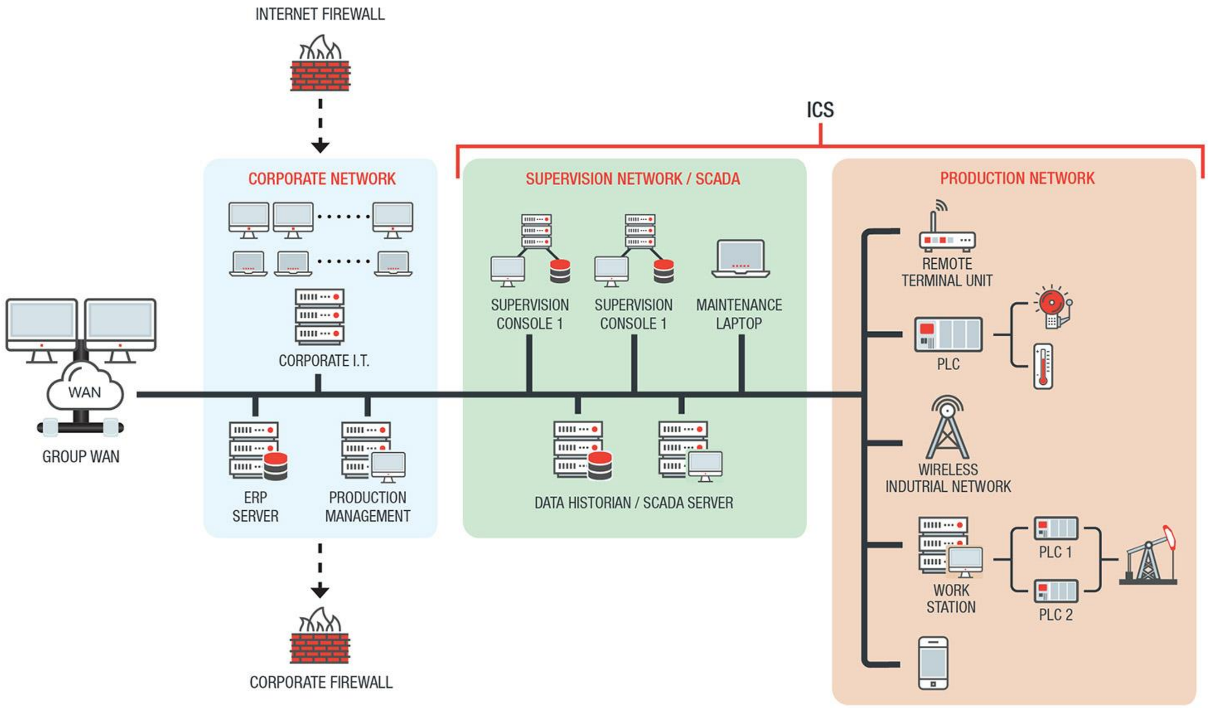

Figure 3 Data historian on a network. Trendmicro (2018/2019) SCADA system and network protection [online]. https://documents.trendmicro.com/images/tex/articles/ics-system.jpg (4 August 2020)

Data mining algorithms

According to McGreevy et al (2010-2), data mining is the process of finding patterns and repetitions in large datasets and is a field of computer science. Data mining techniques and algorithms are being extensively used in Artificial Intelligence and Machine learning. Cichosz (2014-274) invented that the need for data mining algorithms in today's world of big data is increasingly required because large databases with many terabytes are becoming a norm. There are many algorithms but let’s discuss the top 10 in the data mining algorithms list.

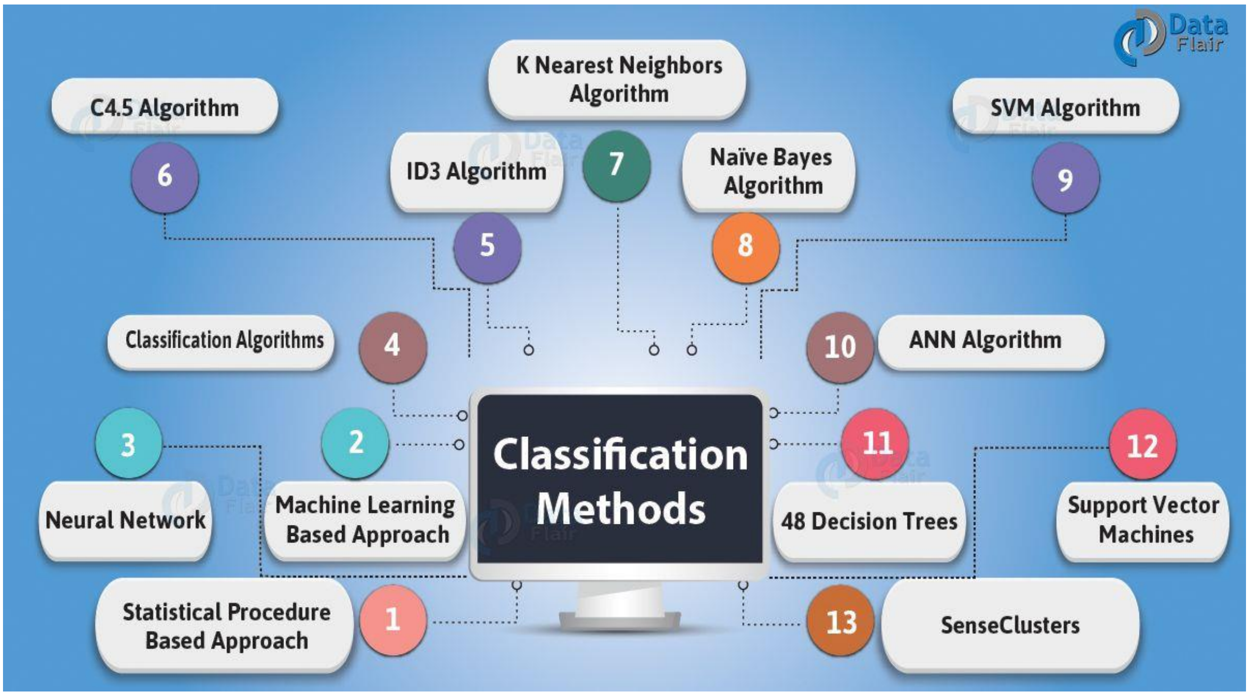

Figure 4 Machine learning classifications methods. Data fair (2018/2019) Machine learning classifications method [online]. https://d2h0cx97tjks2p.cloudfront.net/blogs/wp-content/uploads/sites/2/2018/02/Classification-Methods-01.jpg (7 August 2020)

C4.5 Algorithm

Hssina et al (2014-13) whispered that C4.5 is one of the top data mining algorithms and was developed by Ross Quinlan, and is used to generate a classifier in the form of a decision tree from a set of data that has already been classified. Classifier here refers to a data mining tool that takes data that we need to classify and tries to predict the class of new data.

Adhatrao et al (2013-7) specified that every data point will have its attributes. The decision tree created by C4.5 poses a question about the value of an attribute and depending on those values, the new data gets classified. Hssina et al (2014-17) signposted that the training dataset is labeled with lasses making C4.5 a supervised learning algorithm. Decision trees are always easy to interpret and explain making C4.5 fast and popular compared to other data mining algorithms.

K-mean Algorithm

Yadav and Sharma (2013-2974) understood that one of the most common clustering algorithms. k-means works by creating a k number of groups from a set of objects based on the similarity between objects. It may not be guaranteed that group members will be exactly similar, but group members will be more similar as compared to non-group members. Shah and Singh (2012-435) indicated that as per standard implementations, k-means is an unsupervised learning algorithm as it learns the cluster on its own without any external information.

Support Vector Machines

Rauf et el (2012-960) understood that in terms of tasks, the support vector machine (SVM) works similar to C4.5 algorithm except that SVM doesn't use any decision trees at all. SVM learns the datasets and defines a hyperplane to classify data into two classes. Singh and Bhatia (2011-718) assumed that a hyperplane is an equation for a line that looks something like "y = mx + b". SVM exaggerates to project your data to higher dimensions. Once projected, SVM defined the best hyperplane to separate the data into the two classes.

Apriori Algorithm

According to Li et al (2012-4), the apriori algorithm works by learning association rules. Association rules are a data mining technique that is used for learning correlations between variables in a database. Once the association rules are learned, it is applied to a database containing a large number of transactions. Al-Maolegi and Arkok (2014-22) directed that the apriori algorithm is used for discovering interesting patterns and mutual relationships and hence is treated as an unsupervised learning approach. Singh et al (2013-23) designated that this algorithm is highly efficient, it consumes a lot of memory, utilizes a lot of disk space and takes a lot of time.

Expectation-Maximization Algorithm

Tzoreff and Weiss (2017-35) assumed that Expectation-Maximization (EM) is used as a clustering algorithm, just like the k-means algorithm for knowledge discovery. EM algorithms work in iterations to optimize the chances of seeing observed data. Next, it estimates the parameters of the statistical model with unobserved variables, thereby generating some observed data. Sammaknejad et al (2019-138) directed that the Expectation-Maximization (EM) algorithm is again unsupervised learning since we are using it without providing any labeled class information

PageRank Algorithm

Coppola et al (2019-292) rumored that PageRank is commonly used by search engines like Google. It is a link analysis algorithm that determines the relative importance of an object linked within a network of objects. Link analysis is a type of network analysis that explores the associations among objects. Google search uses this algorithm by understanding the backlinks between web pages. Agryzkov et al (2016-19) granted that it is one of the methods Google uses to determine the relative importance of a webpage and rank it higher on google search engine.

The PageRank trademark is proprietary to Google and the PageRank algorithm is patented by Stanford University. PageRank is treated as an unsupervised learning approach as it determines the relative importance just by considering the links and doesn't require any other inputs.

Adaboost Algorithm

Sun et al (2016-93) signposted that AdaBoost is a boosting algorithm used to construct a classifier. A classifier is a data mining tool that takes data predicts the class of the data based on inputs. A boosting algorithm is an ensemble learning algorithm that runs multiple learning algorithms and combines them. Boosting algorithms take a group of weak learners and combine them to make a single strong learner. A weak learner classifies data with less accuracy. The best example of a weak algorithm is the decision stump algorithm which is a one-step decision tree. Adaboost is perfect supervised learning as it works in iterations and each iteration, it trains the weaker learners with the labeled dataset. Adaboost is a simple and pretty straightforward algorithm to implement. Tang et al (2020-352) exhibited that after the user specifies the number of rounds, each successive AdaBoost iteration redefines the weights for each of the best learners. This makes Adaboost a super elegant way to auto-tune a classifier. Adaboost is flexible, versatile and elegant as it can incorporate most learning algorithms and can take on a large variety of data.

kNN Algorithm

Zhang et al (2018-46) anticipated that kNN is a lazy learning algorithm used as a classification algorithm. A lazy learner will not do anything much during the training process except for storing the training data. Lazy learners start classifying only when new unlabeled data is given as an input. Deng et al (2016-147) assumed that C4.5, SVN and Adaboost, on the other hand, are eager learners that start to build the classification model during training itself. Pandey and Jain (2017-36) point out that since kNN is given a labeled training dataset, it is treated as a supervised learning algorithm.

Naive Bayes Algorithm

Feng et al (2016-6) specified that Naive Bayes is not a single algorithm though it can be seen working efficiently as a single algorithm. Naive Bayes is a bunch of classification algorithms put together. Ma et al (2016-314) expected that the assumption used by the family of algorithms is that every feature of the data being classified is independent of all other features that are given in the class. Naive Bayes is provided with a labeled training dataset to construct the tables. Suresh and Dillibabu (2018-3145) agreed that it is treated as a supervised learning algorithm.

CART Algorithm

Sang at el (2019-12) apprehended that CART stands for classification and regression trees. It is a decision tree learning algorithm that gives either regression or classification trees as an output. In CART, the decision tree nodes will have precisely 2 branches. Just like C4.5, CART is also a classifier. Li (2017-40) believed that the regression or classification tree model is constructed by using a labeled training dataset provided by the user. Hence it is treated as a supervised learning technique

Machine learning

Raschka and Mirjalili (2017-14) supposed that machines are by nature not intelligent. Biamonte et al (2017-198) signposted that Initially, machines were designed to perform specific tasks, such as running on the railway, controlling the traffic flow, digging deep holes, traveling into the space, and shooting at moving objects.

Machine learning vs humans

Yao et al (2018-4) detailed that the fundamental difference between humans and machines in performing their work is intelligence. Arnaldo and Veeramachaneni (2019-41) thought that the human brain receives data gathered by the five senses: vision, hearing, smell, taste, and tactility. Ansari et al (2018-119) held that these gathered data are sent to the human brain via the neural system for perception and taking action. Gil et al (2019-620) assumed that a machine cannot deal with the gathered data intelligently. It cannot analyze data for classification, benefit from previous experiences, and store the new experiences to the memory units that is, machines do not learn from experience. Fong et al (2018-8) suspected that machine learning is a branch of artificial intelligence that aims at enabling machines to perform their jobs skillfully by using intelligent software.

Machine learning techniques

According to Ucci et el (2019-144), machine learning is a natural outgrowth of the intersection of Computer Science and Statistics.

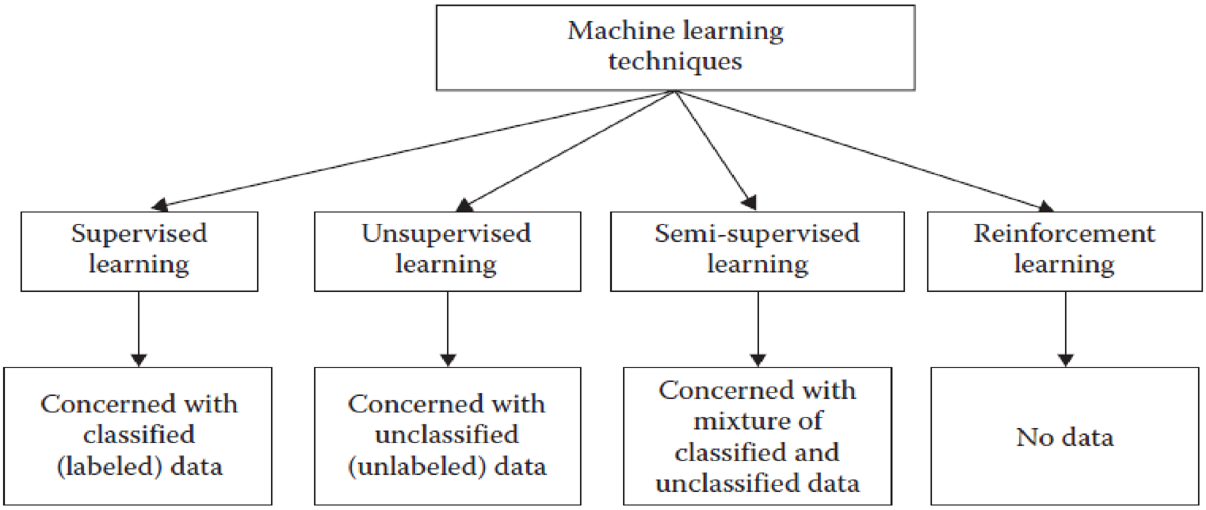

Figure 5 Different machine learning techniques and their requirements (Mohssen et el, 2016-7)

- Supervised Learning – In supervised learning, the target is to infer a function or mapping from training data that is labeled. The training data consist of input vector X and output vector Y of labels or tags.
- Unsupervised Learning - In unsupervised learning, we lack supervisors or training data.
- Semi-Supervised Learning -In this type of learning, the given data are a mixture of classified and unclassified data. This combination of labeled and unlabeled data is used to generate an appropriate model for the classification of data.
- Reinforcement Learning - The reinforcement learning method aims at using observations
- Gathered from the interaction with the environment to take actions that would maximize the reward or minimize the risk.

Machine learning in engineering

Rahman et al (2019-4) suspected that the adoption of machine learning in engineering has been especially valuable for expanding the horizons of signal processing. Kim et al (2020-7) purported that these systems function efficiently to increase the accuracy and subjective quality when sound, images, and other inputs are transmitted. Zheng and Casari (2018-118) believed that the machine learning algorithms make it possible to model signals, detect meaningful patterns, develop useful inferences, and make highly precise adjustments to signal output. Zheng and Casari (2018-162) assumed that in turn, signal processing techniques can also be used to improve the data fed into machine learning systems. Rahman et al (2019-7) thought that by cutting out much of the noise that would otherwise be included in these inputs, engineers achieve cleaner results in the performance of Internet-of-Things devices and other AI-enabled systems. Kim et al (2020-8) agreed that the Department of Electrical and Computer Engineering at MSU demonstrates the innovative, life-changing possibilities that can come from applying AI to investigations in signal processing. Rahman et al (2019-9) suspected that multidisciplinary researchers synthesize concepts from electrical and computer engineering, artificial intelligence and other fields to simulate the way biological eyes process visual information. Varshney (2016-4) apprehended that these efforts serve to deepen our understanding of how our senses function while leading to greater capabilities for visual prosthetics, brain-computer interfaces, motion sensors, and computer vision algorithms. Kim et al (2020-9) rumored that with the ever-increasing amounts of data becoming available there is good reason to believe that smart data analysis will become even more pervasive as a necessary ingredient for technological progress.

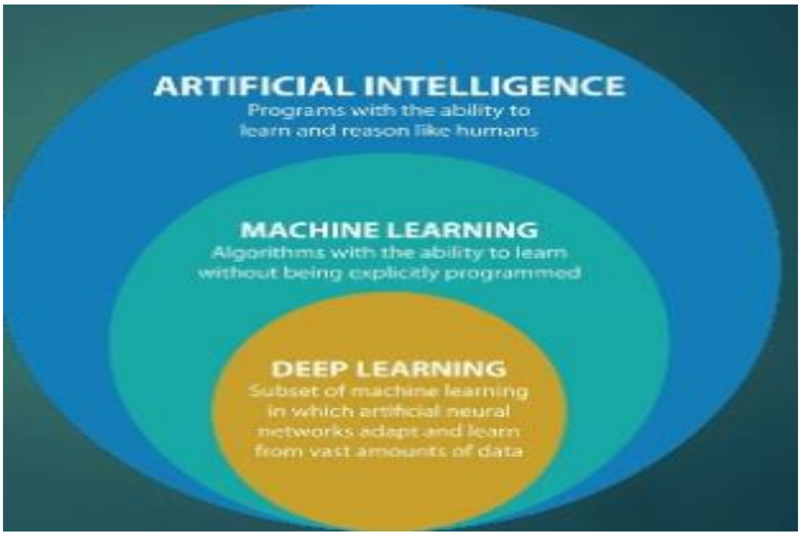

Figure 6 Artificial intelligence, machine learning and deep learning. Agility (2019/2020) all the solution you need to engage your customer everywhere [online]. https://www.argility.com/argility-ecosystem-solutions/iot/machine-learning-deep-learning (9 August 11, 2020)

Zheng and Casari (2018-157) rumored that a lot of research is going on in the field of data mining algorithms. Varshney (2016-5) signposted that these algorithms are used to extract knowledge out of data to devise new and innovative strategies. This applies to the machine-learning process as well. Kim et al (2020-6) expected that research institutions use mining to improve their managerial standard and make decision-making effective.

Artificial intelligence

According to Shukla et al (2019-527), since the earliest days of computing, Computer expert and scientists have been captivated by the idea of creating a machine capable of replicating the human brain and an analogy that says a human brain is like a computer was invented. According to Salehi and Burgueno (2018-188), the human brain and its capabilities go far beyond human thinking. Feldt et al (2018-39) said that academics and scientist has serious debate surrounding whether consciousness can be separated from advanced intelligence and the lack of understanding how consciousness arises in the human brain.

Salehi and Burgueno (2018-182) assumed that the artificial intelligence (AI) used in the engineering sector combines both software and hardware components. Think of the robots on a car assembly line and the software that controls them. Feldt et al (2018-34) indicated that they are in themselves quite impressive feats of engineering, but are they intelligent? Shukla et al (2019-521) supposed that the intensification of artificial intelligence capacities permits us to develop machines capable of performing ever more complicated manufacturing, and even design, tasks. Feldt et al (2018-41) supposed that it used to be thought that the analogy of the human brain is like a computer ran deep. However, we now know that the picture is much more complicated, the way that the brain works goes beyond a simple computer.

Shukla et al (2019-522) suspected that with the massive amounts of data being produced by the current Big Data Era, we're bound to see innovations that we can't even fathom yet, and potentially as soon as in the next ten years. Salehi and Burgueno (2018-187) engaged that the analogy to deep learning is that the rocket engine is the deep learning model and the fuel is the huge amount of data we can feed to these algorithms.

Case study: Blast furnace control system

Developing a control strategy for a blast furnace control is based on black box principles. The control system has two levels. At the basic level, classical control approaches are used. The process level is a combination of model-based and artificial intelligence (AI) approaches. Developed models are used in real-time for furnace state estimation and prediction as well as for decision support.

Introduction

The efficiency of the control systems is based on local dynamic optimization. Blast furnace process control is one of the main contributors to the successful blast furnace operation and belongs to the key factors of its economic effectivity. Blast furnace process complexity in combination with growing demand for effectivity and reducing environmental impact has necessitated a change in the process control strategy. In the past, static calculations based on black-box principles were used to predetermine some fundamental set points with limited feed-back from the process. This type of control can give acceptable results only when the process has a small deviation from the stationary operating point. Because of process instability, small disturbances can cause significant deviation from the operating point which requires setpoint correction for which are AI methods generally used.

The core determination of the process control system is to reach specified pig iron composition tap, temperature at the tap, and volume. The outputs from the process control system are set points of:

- Charging material composition
- Coke ratio
- Charging material distribution
- Blast wind volume
- Blast wind
- Temperature
- Blast wind moisture
- Oxygen injection
- Fuel injection
- Tapping time and tapping intensity.

The setpoints are unwavering to fulfill process requirements (productivity and quality) with minimal energy and material costs.
Background

In a blast furnace, Stoichiometric models are used for the heat and material balance taking into account chemical reactions during heating and melting. Feed-back signals were used from the process to allow for the closing material and heat balance. This was achieved by measuring blast furnace gas flow, temperature and composition which allows the calculation of the oxidation/reduction process. Based on the data for required pig iron grade the associated charge and heat and mass balance calculation is performed. The calculation is based on a stoichiometric model with consideration to field factors such as wind temperature and predicted heat losses by furnace gas, alkali content etc. The calculation aims at reaching the requested melting intensity, pig iron composition and temperature requested slag basicity. Further are calculated wind parameters, optimal flame temperature, optimal heat and gas utilization as well as economic characteristics. The process status determination is based on information retrieved from field instrumentation. The following models were developed:

- Material distribution and position of particular material zones
- Gas distribution
- Temperature distribution
- Thermal state of the furnace and its parts
- Pig iron and slag chemical composition in the dropping zone
- Hearth liquid level
- Geometry of material zones (LTTR, cohesive zone, dropping zone, dead man)
- Shaft geometry (scaffolds) furnace

The predictive model makes real-time simulations based on actual data about the furnace inputs. Vertically, each zone consists of one or more charges, horizontally, the furnace is divided into 8 rings, 8 segments and 64 elements. The modeled processes are: gas flow, material flow, thermal process, chemical process, physical process and geometrical process. The furnace state is determined for each element with heat and material balance.

The calculation is based on information retrieved from the laboratory and field instrumentation allowing the closing of the balances. In the model are used the following basic chemical reactions.

- Water vaporization
- Carbon oxides dissociation
- Indirect reduction of iron oxides by carbon monoxide
- Indirect reduction of iron oxides by hydrogen
- Direct reduction
- Boudouard's reaction

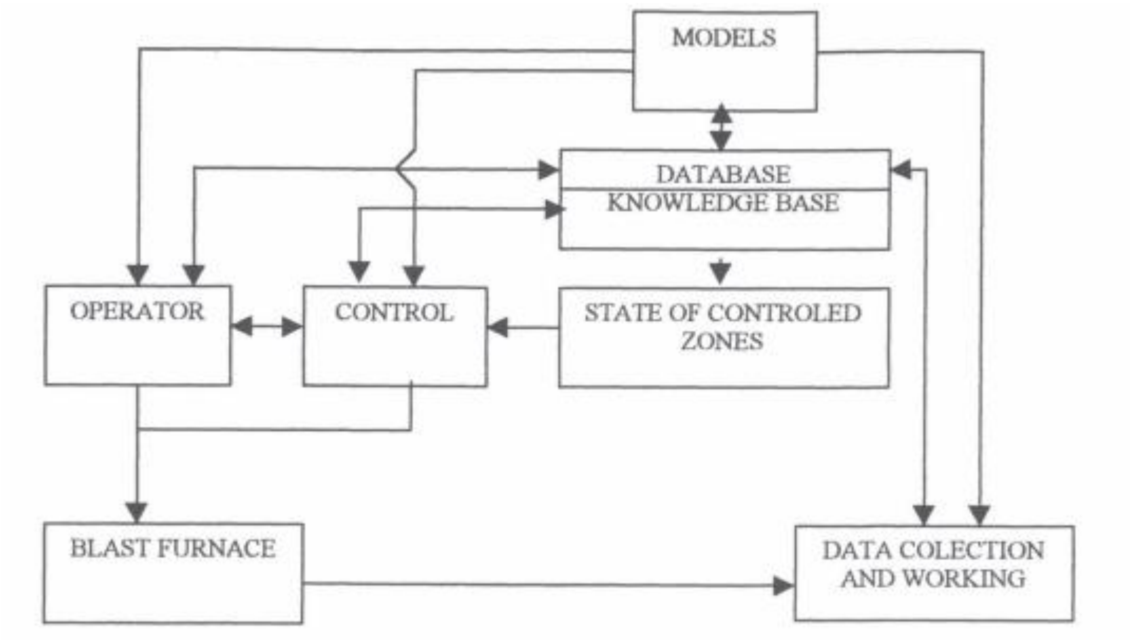

Figure 7 Blast furnace decomposition for dynamic model (Kostial et all, 2000 – 252)

Evaluation of the Case

With variation in the charge composition and various operating practices, the dynamics of the process have to be taken into account providing recalculation of set points and immediate feedback in real-time. One precondition for the dynamic process control was the development of the sensors and measuring techniques which give real-time information about the process state. The evaluation of dynamic models makes it possible to go from process supervision to proactive real-time control. Different approaches reflect specific situation and control philosophy

Proposed Solution/Changes

Machine learning analyzes large amounts of data automatically to identify errors, component deterioration, and poor process optimization using algorithms, neural networks that allow one to group/exclude data, detect oscillations in commands to devices, predict signal behaviors and train neural networks to make them more precise. The control system operates on three decision levels; strategic, tactical, and operative. On the strategic level, charging material composition is determined. The decision is based and static model simulations. On the tactical level, energy inputs and charge distribution to reach the desired (optimal) process trajectory characterized by heat balance and temperature distribution are determined. For the decision predictive model is used. Operative decisions are for the correction of process disturbances detected by dynamic models. Because gas flow distribution has the main influence on process behavior, it is the correction of the distribution and liquid levels which can be effectively used to behavior.

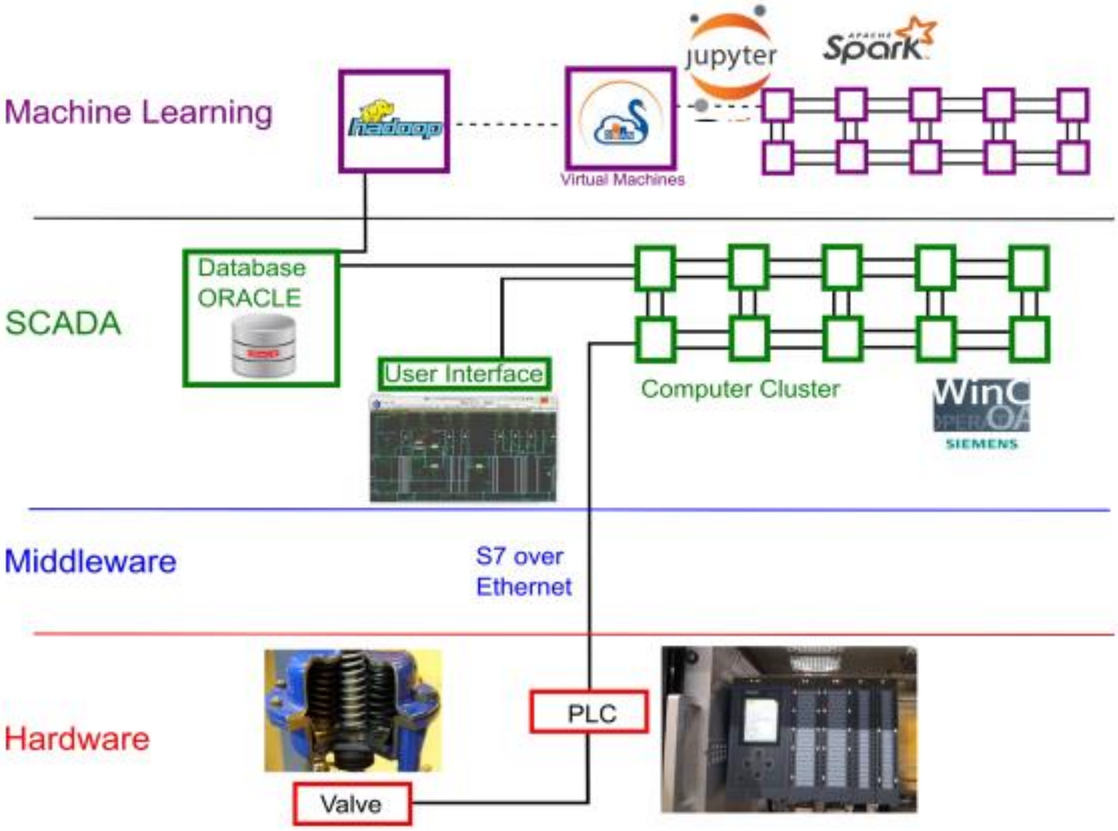

Figure 8 Machine learning integration to industrial control system network

P-chart

The P charts are used to monitor the proportion of nonconforming units of a blast furnace process based on samples taken from the process at given times (hours, shifts, days, weeks, months, etc.). The initial series of samples is used to estimate the proportion nonconforming of a process which is then used to produce control limits for the proportions. Once the control limits have been established for the P chart, these limits may be used to monitor the proportion nonconforming of the process going forward. Using cryogenic data analysis (Machine learning), The signals are collected for all field equipment, data clean-up is performed, and the data is visually pre-analyzed. The machine learning algorithm is allowed to perform its functions after the number of groups has been set and the starting conditions have been specified. The chart below indicates the proportion of occurrences for a blast furnace control system.

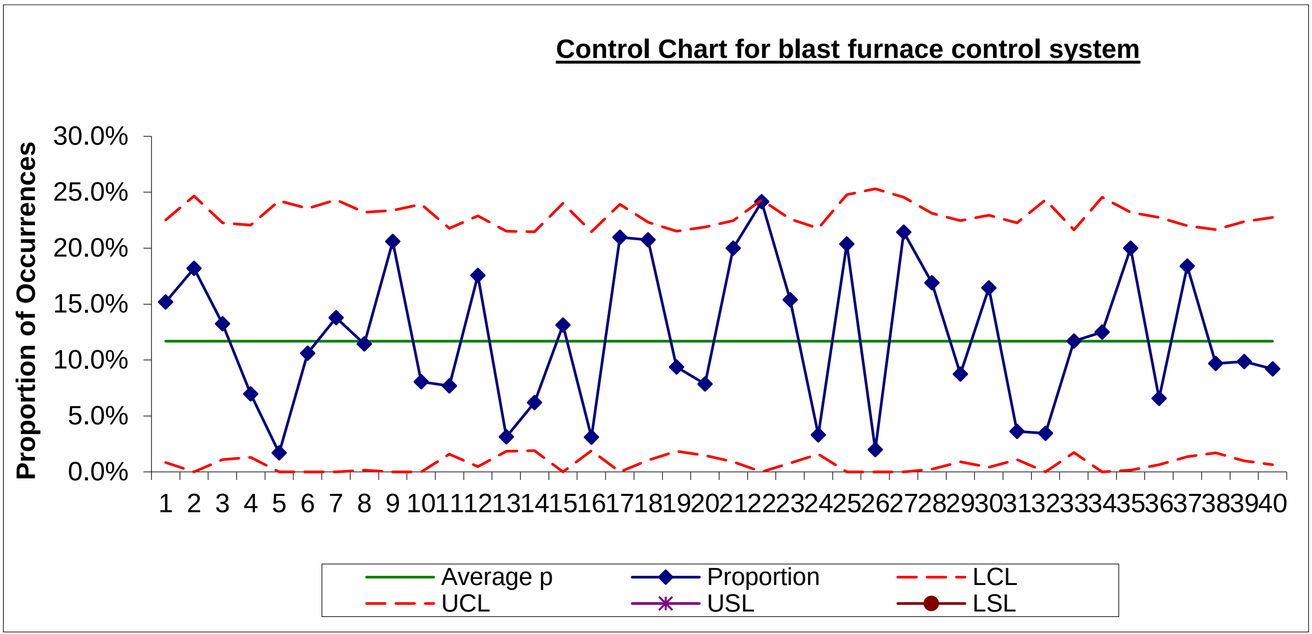

Figure 9 P Control Chart for blast furnace control system

C-chart

The C chart is used to monitor process stability over time to identify and correct instabilities in a process. The special causes that contribute to the unusually high number of readings should be identified. The figure below shows the C- chart for the blast furnace control system.

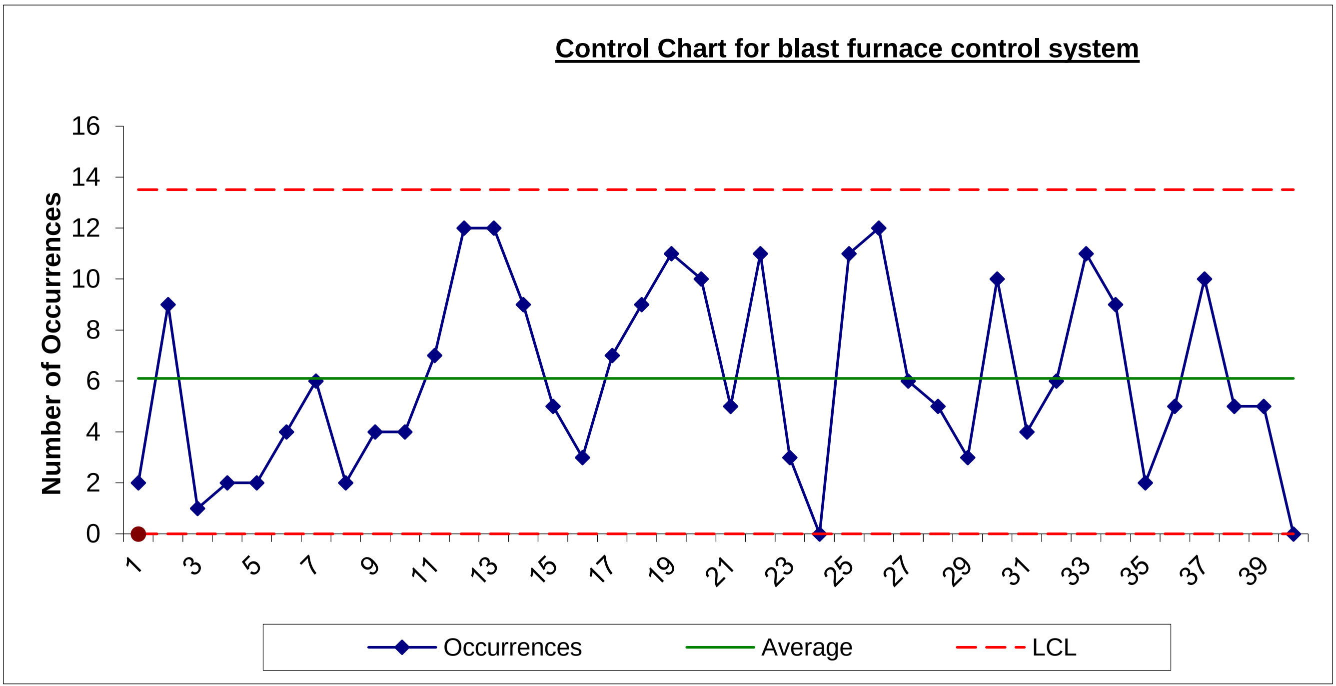

Figure 10 C Control Chart for blast furnace control system

U-chart

U control charts are used when the opportunity for defects to occur is large but the number that occurs is small. U and the c control charts use the Poisson distribution to model the results. Four conditions must be met for this to be correct.

1. The counts must be discrete.
2. The counts must occur in a well-defined region of space or time.
3. The counts are independent of each other and the likelihood of a count is proportional to the size of the area of opportunity. This means that the probability of finding a bubble on a plastic sheet is not related to which part of the plastic sheet is selected.
4. The counts are rare compared to the opportunity. For example, the opportunity for customer complaints to occur is large. However, the number that occurs is small.

U control charts can be used when the above conditions for counting data are met. The c control chart can be used in the area of opportunity does not change from subgroup to subgroup. U control chart can be used if the subgroup size (the area of opportunity) changes from subgroup to subgroup. This area of opportunity is often referred to as an inspection unit. Suppose you monitor pumps in a large plant for leaks. An inspection unit can be one item, such as one pump, or it can be multiple items such as 10 pumps. If the inspection unit is 10 pumps, the subgroup size, n, indicates how many tens were inspected. For example, if 42 pumps were inspected, the subgroup size, n, would be 4.2 inspection units.

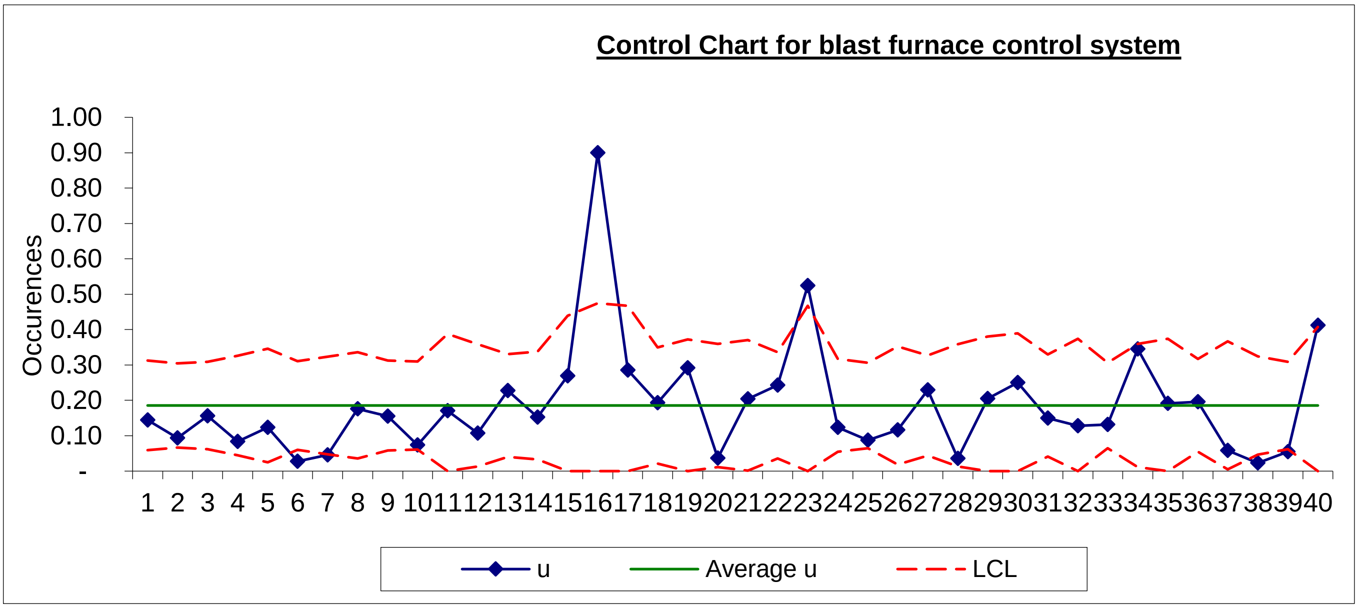

Figure 11 U Control Chart for blast furnace control system

Machine learning concepts: Mean and Standard Deviation

The Mean (applied to signals) : the average value of the sampled signal.

In (1) : Mean = 0
For elements in X :
Mean = Mean + elements
Mean = mean/N
Print (Mean)
4.9

The Standard Deviation: the difference between a value and the mean of the signal.

In (2) = Varience = 0
For element in x :
Varience = Varience + (element – mean) ** 2
Varience = Varience /(N-1)
Sd = Math.sgrt (Varience)
Print (varience)
Print (sd)
7.098
3.097

Machine Learning Algorithm: kmeans

The k-means Machine Learning algorithm groups data samples (represented by points) depending on their values so that the area of each cluster (group) is minimal and the number of samples contained in it is the largest possible.

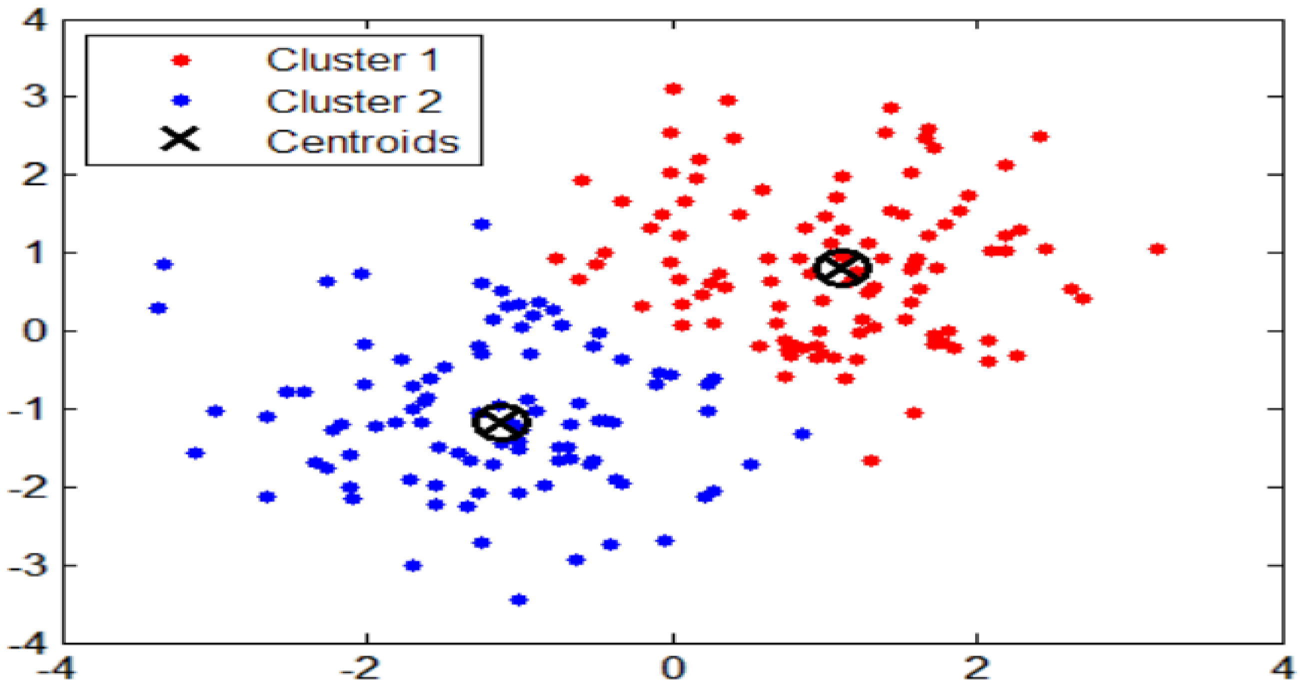

Figure 12 Machine Learning algorithm groups data samples

Findings

This research used a case-based methodology to investigate how data mining algorithms can advance machine learning in industrial control system projects. The research findings for each research hypotheses as outlined in the research hypothesis, are analyzed in the following paragraph

Hypothesis 1

H_0 : Data mining techniques can not improve the deep machine learning algorithm of the industrial control system.

H_1 : Data mining techniques can improve the deep machine learning algorithm of the industrial control system.

Data mining and analytics have played an important role in knowledge discovery and decision making in the process industry over the past several decades. As a computational engine to data mining and analytics, machine learning serves as basic tools for information extraction, data pattern recognition and predictions. From the perspective of machine learning, the case study results indicated that data mining and analytics applications in the process industry will assist in process control and control systems intelligence. Data mining techniques will reduce the level of industrial programming required in blast furnaces which include alarms management, error management and faults management. This requires that calibration data equip the industrial control system to decide for any situation without any intervention of a person (Unsupervised learning).

From the above results H_1 : Is accepted - 1. Data mining techniques can improve the deep machine learning algorithm of the industrial control system.

Hypothesis 2

H_0 : Data mining and machine learning projects success will not improve industrial control systems intelligence, creating a competitive advantage for the firms avoiding machine damage.

H_1 : Data mining and machine learning projects success will improve industrial control systems intelligence, creating a competitive advantage for the firms avoiding machine damage.

Developments, expansions and upgrades in the field of industrial information technology have introduced a large number of new threats to industries. This real-time data demonstrates whether additional attention to the system is needed and requirements in terms of resources, performance, or improvements are being met. Industrial Control Systems (ICS) are also recognized as attractive targets for threat actors. While these networks were generally thought to be more secure due to lack of connection to the outside world of the corporate network or on the internet, now it is not the same case and attackers have managed to compromise them and steal valuable production data. Some of the most effective tools in combating these threats are the emerging techniques in artificial intelligence. By combining these threats with real-time data monitoring along with orchestration and automated response, AI analytics solutions are proving their best possible desirable outcome when compared to legacy systems and human-intervention driven response times. There are a large number of tools used in AI, including search and mathematical optimization, logic, methods based on probability, and many others. AI in control engineering is often not about simulating human intelligence. We can learn something about how to make machines solve problems by observing system and process problems, but most work in intelligent control involves studying real problems in the world. This will reduce organization operational and maintenance costs because only 50% of the maintenance staff will be required while machine damage (Property damage) can be reduced by 80% as a result of improved control system intelligence which allows controllers to decide for any condition unsupervised.

From the above results $\mathbf{H_1}$: Is accepted - Data mining and machine learning projects success will improve industrial control systems intelligence, creating a competitive advantage for the firms avoiding machine damage.

Conclusion

By using data mining algorithms and deep machine learning, process control of the blast furnace system operating conditions can be optimized, hot metal quality can be improved, and energy consumption and operational costs can be reduced. Modern expert process control systems continuously monitor certain parameters in the blast furnace and by using various process models to calculate and diagnose process disturbances to suggest or take, corrective actions such as modification of the rate of reducing agents or changes to the burden distribution. It was discovered that machine deep learning in blast furnace control system can achieve the following:

- Detect the category data kinds belong to by classification of algorithms
- An anomaly detection algorithm will assist to identify what is bizarre
- Use regression algorithm to find out the quality and quantities data
- Use clustering algorithm to find how data is arranged
- Use reinforcement learning algorithm to determine what should be done next
- Use deep machine learning in control and feedback system
- Use deep machine learning in cybersecurity
- Use deep machine learning in load balancing

The control system has basic and process levels. The basic level includes process measurement and setpoints control while the process level includes furnace state estimation and prediction, process analysis and decision. The dream of machines appearing as smart as humans is still far from being realized. In general, a smart machine is an intelligent system that uses equipment such as sensors, RFID, a Wi-Fi, or communications link to receive data and interpret it to make decisions. They use machine learning algorithms to accomplish tasks usually performed by humans in an order to enhance efficiency and productivity. Computer engineers provided the ever-more-powerful machines that make AI applications possible. Control theory deals with designing devices that act optimally based on feedback from the environment. Initially, the mathematical tools of control theory were quite different from AI, but the fields are coming closer together.

Recommendations

For designers, builders, and users of process control systems to take advantage of this emerging technology, it will become increasingly important to understand what constitutes AI, the capabilities and limitations of various AI technologies, and how to determine where, and whether, an AI solution should be attempted. It is recommended that:

- Future engineering researchers must develop models to identify engineering problems that require AI solutions and areas where they are not required. This can assist organizations aiming to use AI to resolve some of their problems.
- The specific machine learning algorithm or commination must be found, to develop a plug and play models to be used to resolve industrial problems

Biography

Aamir, M., Poncela, J., Uqaili, M.A., Chowdhry, B.S. and Khan, N.A., 2013. Optimal design of remote terminal unit (RTU) for wireless SCADA system for energy management. Wireless personal communications, 69(3), pp.999-1012.

Adhatrao, K., Gaykar, A., Dhawan, A., Jha, R. and Honrao, V., 2013. Predicting students' performance using ID3 and C4. 5 classification algorithms. arXiv preprint arXiv:1310.2071.

Agryzkov, T., Tortosa, L. and Vicent, J.F., 2016. New highlights and a new centrality measure based on the Adapted PageRank Algorithm for urban networks. Applied Mathematics and Computation, 291, pp.14-29

Al-Maolegi, M. and Arkok, B., 2014. An improved apriori algorithm for association rules. arXiv preprint arXiv:1403.3948.

Ahmed, I., Naedele, M., Schatz, B., Sasaki, R. and West, A., 2012. Supervisory control and data acquisition (SCADA).

Aoki, M., 2013. State space modeling of time series. Springer Science & Business Media.

Arnaldo, I. and Veeramachaneni, K., 2019. The Holy Grail of" Systems for Machine Learning" Teaming humans and machine learning for detecting cyber threats. ACM SIGKDD Explorations Newsletter, 21(2), pp.39-47.

Ansari, F., Erol, S. and Sihn, W., 2018. Rethinking human-machine learning in Industry 4.0: How does the paradigm shift treat the role of human learning? Procedia manufacturing, 23, pp.117-122.

Batni, R.P. and Sharma, R., Alcatel Lucent SAS, 2011. Control server that manages resource servers for selected balance of load. U.S. Patent 8,051,171.

Bayindir, R. and Cetinceviz, Y., 2011. A water pumping control system with a programmable logic controller (PLC) and industrial wireless modules for industrial plants—An experimental setup. ISA transactions, 50(2), pp.321-328.

Barenji, R.V., Barenji, A.V. and Hashemipour, M., 2014. A multi-agent RFID-enabled distributed control system for a flexible manufacturing shop. The International Journal of Advanced Manufacturing Technology, 71(9-12), pp.1773-1791.

Barr, D., 2004. Supervisory control and data acquisition (SCADA). Systems National Communications System, Communication Technologies.

Besser, D.A. and Matan, S., XSLENT Energy Tech LLC, 2012. Power extractor with control loop. U.S. Patent 8,212,399.

Biamonte, J., Wittek, P., Pancotti, N., Rebentrost, P., Wiebe, N. and Lloyd, S., 2017. Quantum machine learning. Nature, 549(7671), pp.195-202.

Bradetich, R., Dearien, J.A., Grussling, B.J. and Remaley, G., Schweitzer Engineering Laboratories Inc, 2013. Local intelligent electronic device (IED) rendering templates over limited bandwidth communication link to manage remote IED. U.S. Patent 8,578,012.

Cichosz, P., 2014. Data mining algorithms: explained using R. John Wiley & Sons
Coppola, M., Guo, J., Gill, E. and De Croon, G.C., 2019. The PageRank algorithm as a method to optimize swarm behavior through local analysis. Swarm Intelligence, 13(3-4), pp.277-319.

Deininger, T.A., Hogan, M. and Xu, Y., Siemens Industry Inc, 2010. System and method for batch process control with diverse distributed control system protocols. U.S. Patent 7,747,416.

Deng, Z., Zhu, X., Cheng, D., Zong, M. and Zhang, S., 2016. Efficient kNN classification algorithm for big data. Neurocomputing, 195, pp.143-148.

Dimiduk, D.M., Holm, E.A. and Niezgoda, S.R., 2018. Perspectives on the impact of machine learning, deep learning, and artificial intelligence on materials, processes, and structures engineering. Integrating Materials and Manufacturing Innovation, 7(3), pp.157-172.

Doyle, J.C., Francis, B.A. and Tannenbaum, A.R., 2013. Feedback control theory. Courier Corporation.

Dullerud, G.E. and Paganini, F., 2013. A course in robust control theory: a convex approach (Vol. 36). Springer Science & Business Media.

Feldt, R., de Oliveira Neto, F.G. and Torkar, R., 2018, May. Ways of applying artificial intelligence in software engineering. In 2018 IEEE/ACM 6th International Workshop on Realizing Artificial Intelligence Synergies in Software Engineering (RAISE) (pp. 35-41). IEEE.

Feng, W., Sun, J., Zhang, L., Cao, C. and Yang, Q., 2016, December. A support vector machine based naive Bayes algorithm for spam filtering. In 2016 IEEE 35th International Performance Computing and Communications Conference (IPCCC) (pp. 1-8). IEEE.

Fong, R.C., Scheirer, W.J. and Cox, D.D., 2018. Using human brain activity to guide machine learning. Scientific reports, 8(1), pp.1-10.

Gil, Y., Honaker, J., Gupta, S., Ma, Y., D'Orazio, V., Garijo, D., Gadewar, S., Yang, Q. and Jahanshad, N., 2019, March. Towards human-guided machine learning. In Proceedings of the 24th International Conference on Intelligent User Interfaces (pp. 614-624).

Grubic, T., 2014. Servitization and remote monitoring technology. Journal of Manufacturing Technology Management.

Harjunkoski, I., Nyström, R. and Horch, A., 2009. Integration of scheduling and control—Theory or practice?. Computers & Chemical Engineering, 33(12), pp.1909-1918.

Han, J., Susilo, W., Mu, Y., Au, M.H. and Cao, J., 2015. AAC-OT: accountable oblivious transfer with access control. IEEE Transactions on Information Forensics and Security, 10(12), pp.2502-2514.

Hessmer, R., Todorov, I.A., Hadrich, M. and Ross, L.D., Schneider Electric Systems USA Inc, 2012. Method and system for performing remote diagnostics on a process data access. U.S. Patent 8,095,632.

Huang, S., Chen, Y., Chen, X., Liu, K., Xu, X., Wang, C., Brown, K. and Halilovic, I., 2014, June. The next generation operational data historian for iot based on informix. In Proceedings of the 2014 ACM SIGMOD international conference on Management of data (pp. 169-176).

Hssina, B., Merbouha, A., Ezzikouri, H. and Erritali, M., 2014. A comparative study of decision tree ID3 and C4. 5. International Journal of Advanced Computer Science and Applications, 4(2), pp.13-19.

Idachaba, F.E. and Ogunrinde, A., 2012. Review of Remote Terminal Unit (RTU) and Gateways for Digital Oilfield delpoyments. IJACSA) International Journal of Advanced Computer Science and Applications,, 3(8), pp.157-160.

Isaac, E.S., Morey Corp, 2014. Method and system for remote diagnostics of vessels and watercrafts. U.S. Patent 8,798,847.

Jelali, M., 2012. Control performance management in industrial automation: assessment, diagnosis and improvement of control loop performance. Springer Science & Business Media.

Jusoh, W.W., Hanafiah, M.M., Ghani, M.A. and Raman, S.H., 2013, June. Remote terminal unit (RTU) hardware design and implementation efficient in different application. In 2013 IEEE 7th International Power Engineering and Optimization Conference (PEOCO) (pp. 570-573). IEEE.

Kim, G.B., Kim, W.J., Kim, H.U. and Lee, S.Y., 2020. Machine learning applications in systems metabolic engineering. Current opinion in biotechnology, 64, pp.1-9.

Liu, J.J., Wang, H.J., Xu, M. and Zhu, Y.H., Siemens AG, 2014. Module for programmable logic controller (PLC). U.S. Patent Application 29/464,579

Liu, G.P. and Daley, S., 2001. Optimal-tuning PID control for industrial systems. Control Engineering Practice, 9(11), pp.1185-1194.

Li, N., Zeng, L., He, Q. and Shi, Z., 2012, August. Parallel implementation of apriori algorithm based on mapreduce. In 2012 13th ACIS international conference on software engineering, artificial intelligence, networking and parallel/distributed computing (pp. 236-241). IEEE.

Li, M., 2017, November. Application of CART decision tree combined with PCA algorithm in intrusion detection. In 2017 8th IEEE International Conference on Software Engineering and Service Science (ICSESS) (pp. 38-41). IEEE.

Lyshevski, S.E., 2012. Control systems theory with engineering applications. Springer Science & Business Media.

Ma, Y., Liang, S., Chen, X. and Jia, C., 2016, July. The approach to detect abnormal access behavior based on naive bayes algorithm. In 2016 10th International Conference on Innovative Mobile and Internet Services in Ubiquitous Computing (IMIS) (pp. 313-315). IEEE.

Malhotra, R., Singh, N. and Singh, Y., 2011. Genetic algorithms: Concepts, design for optimization of process controllers. Computer and Information Science, 4(2), p.39.

Meinherz, C., Brockman, C.M., Fooks, E.I., Stein, M.N., Hardegger, M., Chen, W.J., Berner, R., Galera, R., Black, R.M., Merz, R. and Nair, S., 2012. Time of flight (tof) human machine interface (hmi). U.S. Patent Application 12/904,471.

McGreevy, R.J., Sadowski, D.A., Herbst, R.J. and Hartman, J.R., Rockwell Automation Technologies Inc, 2010. Distributed historian architecture. U.S. Patent 7,831,317.

McMillan, G.K., 2014. Tuning and control loop performance. Momentum Press.

Odgaard, P.F. and Stoustrup, J., 2014. A benchmark evaluation of fault tolerant wind turbine control concepts. IEEE Transactions on Control Systems Technology, 23(3), pp.1221-1228.

Rauf, A., Sheeba, S.M., Khusro, S. and Javed, H., 2012. Enhanced k-mean clustering algorithm to reduce number of iterations and time complexity. Middle-East Journal of Scientific Research, 12(7), pp.959-963

Raschka, S. and Mirjalili, V., 2017. Python machine learning. Packt Publishing Ltd.

Rahman, M.S., Rivera, E., Khomh, F., Guéhéneuc, Y.G. and Lehnert, B., 2019. Machine learning software engineering in practice: An industrial case study. arXiv preprint arXiv:1906.07154.

Pandey, A. and Jain, A., 2017. Comparative analysis of KNN algorithm using various normalization techniques. International Journal of Computer Network and Information Security, 9(11), p.36

Papadopoulos, A.D., Tanzman, A., Baker Jr, R.A., Belliardi, R.G. and Dube, D.J., Schneider Automation Inc, 2000. System for remotely accessing an industrial control system over a commercial communications network. U.S. Patent 6,061,603.

Prett, D.M. and García, C.E., 2013. Fundamental process control: Butterworths series in chemical engineering. Butterworth-Heinemann

Sammaknejad, N., Zhao, Y. and Huang, B., 2019. A review of the expectation maximization algorithm in data-driven process identification. Journal of process control, 73, pp.123-136
Salehi, H. and Burgueno, R., 2018. Emerging artificial intelligence methods in structural engineering. Engineering structures, 171, pp.170-189.

Sang, X., Guo, Q., Wu, X., Fu, Y., Xie, T., He, C. and Zang, J., 2019. Intensity and stationarity analysis of land use change based on CART algorithm. Scientific Reports, 9(1), pp.1-12.

Seborg, D.E., Mellichamp, D.A., Edgar, T.F. and Doyle III, F.J., 2010. Process dynamics and control. John Wiley & Sons.

Senga, E., Toshiba Corp, 2011. Stream delivery system, call control server, and stream delivery control method. U.S. Patent Application 12/970,769.

Singh, R.V. and Bhatia, M.S., 2011, June. Data clustering with modified K-means algorithm. In 2011 International Conference on Recent Trends in Information Technology (ICRTIT) (pp. 717-721). IEEE.

Singh, J., Ram, H. and Sodhi, D.J., 2013. Improving efficiency of apriori algorithm using transaction reduction. International Journal of Scientific and Research Publications, 3(1), pp.1-4.

Sun, B., Chen, S., Wang, J. and Chen, H., 2016. A robust multi-class AdaBoost algorithm for mislabeled noisy data. Knowledge-Based Systems, 102, pp.87-102.

Suresh, K. and Dillibabu, R., 2018. Designing a machine learning based software risk assessment model using Naïve Bayes algorithm. TAGA Journal, 14, pp.3141-3147.

Shah, S. and Singh, M., 2012, May. Comparison of a time efficient modified K-mean algorithm with K-mean and K-medoid algorithm. In 2012 International Conference on Communication Systems and Network Technologies (pp. 435-437). IEEE.

Shi, Q., Zhang, Z., Chen, T. and Lee, C., 2019. Minimalist and multi-functional human machine interface (HMI) using a flexible wearable triboelectric patch. Nano Energy, 62, pp.355-366.

Shukla, A.K., Janmaijaya, M., Abraham, A. and Muhuri, P.K., 2019. Engineering applications of artificial intelligence: A bibliometric analysis of 30 years (1988–2018). Engineering Applications of Artificial Intelligence, 85, pp.517-532.

Spanier, J., Wang, W. and da Silva, D.S., Electro Industries Gauge Technology, 2015. Intelligent electronic device for receiving and sending data at high speeds over a network. U.S. Patent 9,080,894.

Tang, D., Tang, L., Dai, R., Chen, J., Li, X. and Rodrigues, J.J., 2020. Mf-adaboost: Ldos attack detection based on multi-features and improved adaboost. Future Generation Computer Systems, 106, pp.347-359.

Tin, C. and Poon, C.S., 2005. Internal models in sensorimotor integration: perspectives from adaptive control theory. Journal of Neural Engineering, 2(3), p.S147.

Tzoreff, E. and Weiss, A.J., 2017. Expectation-maximization algorithm for direct position determination. Signal processing, 133, pp.32-39.

Valenzuela, L., Zarza, E., Berenguel, M. and Camacho, E.F., 2005. Control concepts for direct steam generation in parabolic troughs. Solar Energy, 78(2), pp.301-311.

Varshney, K.R., 2016, January. Engineering safety in machine learning. In 2016 Information Theory and Applications Workshop (ITA) (pp. 1-5). IEEE.

Venugopal, R. and Bernstein, D.S., 1996, September. State space modeling and active control of slosh. In Proceeding of the 1996 IEEE International Conference on Control Applications IEEE International Conference on Control Applications held together with IEEE International Symposium on Intelligent Contro (pp. 1072-1077). IEEE.

VoBa, S. and Ureche, O.T., Microsoft Technology Licensing LLC, 2017. Distribution control and tracking mechanism of virtual machine appliances. U.S. Patent 9,703,586.

Ucci, D., Aniello, L. and Baldoni, R., 2019. Survey of machine learning techniques for malware analysis. Computers & Security, 81, pp.123-147.

Watanabe, Y., NEC Corp, 2016. Communication control system, control server, forwarding node, communication control method, and communication control program. U.S. Patent 9,379,975.

Yadav, J. and Sharma, M., 2013. A Review of K-mean Algorithm. International journal of engineering trends and technology, 4(7), pp.2972-2976.

Yao, Q., Wang, M., Chen, Y., Dai, W., Yi-Qi, H., Yu-Feng, L., Wei-Wei, T., Qiang, Y. and Yang, Y., 2018. Taking human out of learning applications: A survey on automated machine learning. arXiv preprint arXiv:1810.13306.

Zelikin, M.I. and Borisov, V.F., 2012. Theory of chattering control: with applications to astronautics, robotics, economics, and engineering. Springer Science & Business Media
Zhang, S., Cheng, D., Deng, Z., Zong, M. and Deng, X., 2018. A novel kNN algorithm with data-driven k parameter computation. Pattern Recognition Letters, 109, pp.44-54.

Zheng, A. and Casari, A., 2018. Feature engineering for machine learning: principles and techniques for data scientists. " O'Reilly Media, Inc.".

About the Author

Dr Lalamani Budeli

South Africa

Dr Lalamani Budeli obtained his degree in Electrical Engineering at the Vaal University (VUT), BSc honours in Engineering Technology Management at University of Pretoria (UP), Master in engineering development and Management at North West University (NWU), Master of business administration at Regent Business School (RBS) and a Doctor of Philosophy in Engineering Development and Management at North West University (NWU), Potchefstroom, South Africa. Currently, he is a managing director of BLIT, an engineering, research, and project management company based in South Africa.

His research interests include project portfolio management, agile project management, plant life cycle management, advanced systems analytics, project early warning system, and the use of artificial intelligence in project management. Currently, he is spending most of the time on research that is looking at the development of system and application that uses the latest technology like blockchain, internet of things (IoT), Big data, and artificial intelligence. Lalamani Budeli can be contacted at budelil@blit.co.za.