Folia Oeconomica
Acta Universitatis Lodziensis
ISSN 0208-6018 e-ISSN 2353-7663

https://www.czasopisma.uni.lodz.pl/foe/
1(362) 2023
DOI: https://doi.org/10.18778/0208-6018.362.04

Adam Chwila https://orcid.org/0000-0003-4671-4298

University of Economics in Katowice, Katowice, Poland, achwila@gmail.com

The Application of Artificial Intelligence Models in Commercial Banks – Opportunities and Threats

Abstract: One of the main sectors that makes heavy use of the development of advanced computational methods is the banking sector. The goals of our research are as follows: 1) to compare scientific and regulatory approaches to defining artificial intelligence (AI) and machine learning (ML), 2) to propose AI and ML definitions for regulatory purposes that allow us to clearly state if a given method is AI/ML or not, 3) to compare the complex quantitative methods applied in banking in terms of complexity and interpretability in order to provide a clear classification of methods to the interested parties (practitioners and management), 4) to propose a possible approach towards the further development of quantitative methods in the areas of required strict interpretability. Our literature review focuses on the definitions of AI/ML applied by scientists and regulators, as well as the proposals of application of complex quantitative solutions in different banking domains. We propose practical definitions of AI and ML based on the current state of the art and requirements of clarity in the banking industry (a very limited risk appetite regarding regulations non-compliance) and compare quantitative methods applied in different banking domains. For regulatory purposes, we propose general and inclusive definitions of AI and ML which allow for a clear classification of specific methods. In the case of strict requirements towards the interpretability of applied methods, we propose a gradual and controlled increase in the complexity of existing solutions. Therefore, we propose the differentiation of quantitative methods in terms of interpretability and complexity. We also think that the definitions of AI/ML in further regulations should make it possible to clearly say whether particular approaches are AI/ML. Our research is directed to policymakers, practitioners, and executives related to the banking sector.

Keywords: artificial intelligence, machine learning, ethics, banking, AI explainability, regulation

JEL: C10, C40, C50, G10, G21

1. Introduction

The advancing development of machine learning (ML) and artificial intelligence (AI) causes rapid growth of interest in advanced quantitative methods among different economic sectors. By following global trends in the field of digitisation, banks are able to provide products that meet the constantly growing expectations of customers and compete with FinTech companies. Due to the development of the e-commerce segment and the improvement of various services with intuitive applications of online stores, society has very similar expectations regarding the banking services provided, such as speed, ease of use, transparency, and full digitisation. AI methods can be applied in many dimensions of bank activities, such as:

1) market risk area, for example: forecasting of client behaviour or matching maturity profiles of assets and liabilities,
2) pricing of the loan collateral (real estate), pricing of financial instruments,
3) customer service: scanning information from ID cards, recognition of speech and speakers through the construction of chatbots, construction of intelligent application interfaces, adapting the displayed content to the profile and potential needs of the client,
4) automatic detection of frauds and money laundering attempts (as part of the Know Your Customer system), cybersecurity systems,
5) internal ratings-based (IRB) models used to calculate regulatory capital for credit risk,
6) stress-testing,
7) construction of systems that automatically assess creditworthiness, etc.

In terms of the application of Big Data and advanced analytics methods, many of the above-presented dimensions are already being explored to varying degrees by many players in the sector (European Banking Authority, 2020). Big Data can be defined as massive structured and unstructured volumes of data that come from different sources (Sagiroglu, Sinanc, 2013), while advanced analytics is defined as the application of multiple analytic methods that address the diversity of big data in order to obtain expected results (Kaisler et al., 2014). The application of advanced analytics methods is a challenge not only for banks but also for regulators. Banks perceive opportunities for risk reduction and income increase in a smooth adoption of solutions aimed towards AI by application of more complex approaches, acquiring specialist competencies, and development of IT systems. At the same time, applications of AI are connected with threats such as inappropriate usage or avoidance due to a lack of clarity from regulators and reluctance towards AI on the part of managers and society. Our research goals are as follows:

1) We want to compare and evaluate different scientific and regulatory approaches towards defining AI and ML.
2) We want to propose AI and ML definitions for legislative purposes.
3) We want to evaluate different complex quantitative methods applied in banking in terms of complexity and interpretability.
4) We want to propose a possible approach towards further development of quantitative methods in the areas of required strict interpretability.

Our research is targeted towards the following stakeholders:

1) policymakers who define AI and/or ML for regulatory purposes,
2) practitioners who apply complex quantitative methods in different banking domains,
3) executives who make final decisions regarding the application of complex quantitative systems in different banking domains.

The main benefits arising from our study are as follows:

1) a better understanding of the current state of AI/ML definitions proposed by scientists, regulators, and worldwide organisations,
2) knowledge of the application of complex quantitative methods in different banking domains,
3) knowledge regarding the relative comparison of the above-mentioned methods with each other in terms of complexity and interpretability,
4) knowledge about possible approaches regarding the further development of complex quantitative methods applications.

2. Definitions of AI

One of the key aspects in the world of advanced computation methods is the understanding of AI and ML concepts through their definitions. Nevertheless, such a task is not easy due to the following reasons:

1) AI and ML can be differently defined in various industries and contexts,
2) an understanding of AI in society, among managers, scientists, politicians, and legal institutions varies, as there is a lack of coherent AI identity (Monett, Lewis, 2018 after Nilsson, 2010),
3) the consequences of classifying a given method as AI or ML are differently approached.

Nowadays, AI is defined by scientists in the field and by different institutions that focus mainly on the legal state and its implications. We focus on presenting the case from both sides. The state-of-the-art snapshot helps with the understanding of AI by researchers and practitioners. The new view of different legal institutions has an impact on the approach that companies must consider when investing in AI-based solutions. From a scientific perspective, AI definitions have been shaped throughout the 20th and 21st centuries, while legal institutions have been developing their approaches during the last few years. From the science development point of view, the current state of the art is more important, while legal aspects are more important from companies' and banks' point of view, as they must obey regulatory requirements. This section and the following one are structured as follows: first, the state-of-the-art aspects are presented and then the legal institutions' approaches are discussed.

In the first half of the 20th century, the concept of AI appeared unsystematically, for example, in theatres. In 1920, the science fiction play R.U.R. written by Karel Čapek (McCorduck, 2004) introduced for the first time the word *robot*. However, a more systematic approach towards AI began shaping in the middle of the 20th century. In 1950, Alan Turing proposed the so-called Turing Test (Turing, 1950) stating that if a machine could carry a conversation via a teleprinter that cannot be distinguished from a conversation with a human, then that is a sign that a machine was *thinking*. Nevertheless, the Turing Test was criticised as an AI definition (Russel, Norvig, 2021). From our point of view, such a test would have very poor practical applications because it only focuses on a small part of AI research, i.e. chat-bots.

1956 is commonly treated as a year when a modern concept of AI was introduced for the first time in the summer workshop Dartmouth Summer Research Project on Artificial Intelligence (Solomonoff, 1985). The conference proposal was based on a thesis that "[…] every aspect of learning or any other feature of intelligence can in principle be so precisely described that a machine can be made to simulate it" (McCarthy et al., 1955). The thesis influenced the direction of AI development, though it was still lacking in the field of defining AI: its interpretation was dependent on the understanding of what exactly was treated as a "feature of intelligence," which pointed to the discussion if a particular improvement was complex enough in order to be treated as AI.

An idea of AI can be described in the context of the so-called intelligent agent, which has been described in a textbook adopted by over 1500 schools and universities over the world (*Artificial Intelligence: A Modern Approach*, 2022). An intelligent agent can be defined as a "system that receives percepts from the environment and performs actions" (Russel, Norvig, 2021). Such agents can be then classified into different categories (Russel, Norvig, 2021): simple reflex agents, model-based reflex agents, goal-based agents, utility-based agents, and learning agents. From the above-presented list, the two are especially interesting taking into consideration the subject of our article:

1) a simple reflex agent reacts only to current incentives from the environment (it ignores the history of the environment) and acts according to a simple rule: if condition A is fulfilled, then action B is taken,
2) a learning agent can improve its performance based on the actions taken in the environment.

The important aspects of the above-presented definition are as follows: an AI does not necessarily have to learn anything, it acts autonomously in a given environment, and it may be based on a very simple mechanism (for example, a streetlamp that automatically turns on when it is dark or a thermostat). From our point of view, the categorisation of AI agents into clearly defined categories is a very coherent and practical approach: it enables the creation of different follow-ups for each type of AI. However, such a solution has its disadvantages. The category assignment potentially can be very hard in practical applications, especially if the consequences of qualifying a particular solution into one or another category are very different from each other. Another drawback of such an approach is that regulators and worldwide institutions often want to settle on a more simplified definition. Therefore, even if the categorisation of intelligent agents is quite well designed, it may be perceived as too complex and not applicable in everyday practice.

There is a proposal to use the term Computational Intelligence (CI) instead of AI (Poole, Mackworth, Goebel, 1998) and to define it also from an agent perspective, where an agent can be a program that has prior knowledge about the world and history based on which it learns. However, some researchers emphasised that the term "artificial" is not an issue in defining AI, and replacing it with another word would not change much (Wang, 2019).

Indeed, the term "intelligence" is what causes the ongoing discussion regarding the definitions and boundaries of AI. The task of defining intelligence is very hard (Monett, Lewis, 2018 after Kambhampati, 2017), and there is no unified way of intelligence interpretation. A recent survey regarding the AI definition, conducted among over 400 researchers, educators, and developers, concluded that around half of respondents expressed the need for separate definitions of human and machine intelligence, whereas the second half stated that a single definition should be enough (Monett, Lewis, 2018). Among nine definitions, the survey respondents most commonly accepted the following definition of machine intelligence (Monett, Lewis, 2018): "the essence of intelligence is the principle of adapting to the environment while working with insufficient knowledge and resources […] an intelligent system should rely on finite processing capacity, work in real time, open to unexpected tasks, and learn from experience […]" (Wang, 2008). A similar approach is considered in (Wang, 2019), where intelligence is defined as a strategy of problem-solving that is fundamentally different from computation, where computation is defined as a finite and repeatable process that carries out a predetermined algorithm to realise a function that maps input data to output data (Wang, 2019 after Hopcroft, Motwani, Ullman, 2007). While the above-presented proposals take up the difficult task of formulating a coherent definition of AI from the scientific point of view and seem to work well as a dictionary definition, the biggest challenges are related to their practical implications. Real-world problems are often not limited to the construction of an algorithm that optimally solves a given issue, and therefore qualification of many software implementations as AI or computation would be ambiguous, which is an especially important drawback in terms of interpretation of legal regulations.

There are also other definitions that similarly to those above emphasise the aspect of learning, for example, based on (Kaplan, Haenlein, 2019): "AI is a system's ability to correctly interpret external data, to learn from such data, and to use those learnings to achieve specific goals and tasks through flexible adaptation." However, in the case of such a definition, the argument regarding difficulties of practical application remains valid, similarly to earlier proposals. There are also attempts of defining AI as computational cognition (Rapaport, 2020), complementation of Wang definition from 2019 (Yampolskiy, 2020), or differentiation into different AI systems, i.e. Artificial General Intelligence, Strong AI, Weak AI, or Human-level AI (Yampolskiy, Fox, 2012; Wang, 2019; Emmert-Streib, Yli-Harja, Dehmer, 2020). However, such attempts are rather philosophical discussions regarding the classification of AI systems into categories with no clear guideline of affiliation of certain systems into one or another category. To sum things up, in the current state, there is no consensus among researchers regarding a single AI definition. Such a state is rather natural, taking into consideration that research on AI is still rapidly developing and an understanding of underlying concepts is still evolving among researchers and practitioners. However, it has its implications – potential practical approaches to the discussed subject may significantly differ from each other.

This issue is especially important in light of many recent regulatory attempts of defining AI. Because there is no widely accepted AI definition, different regulators may stick to definitions that are more or less vague. In the case of official regulators, companies may be forced to comply their actions with a particular legislation act. In other cases, companies may stick to one of the definitions proposed by worldwide organisations of a different kind to set the direction of the whole organisation. Below, we would like to focus on the most recent points of view of worldwide institutions and evaluate their approaches.

In 2019, the OECD defined AI as a system that fulfils the following set of conditions (OECD, 2019):

1) it is a machine-based system, capable of influencing the environment by producing outcomes for a given set of objectives,
2) a system that is designed to operate with varying levels of autonomy,

3) a system that uses machine or human-based input to perceive the environment, abstract its perceptions into a model, and finally use the model to formulate options for outcomes.

Some clarification of the “varying levels of autonomy” and “model” would be necessary to use such a definition in practice, as it is unclear in the current state if the authors’ intention was to potentially support qualifying most of the existing software into AI or not.

The International Organisation of Standardisation defines AI as “research and development of mechanisms and application of AI systems,” while an AI system is defined as an “engineered system that generates outputs such as content, forecasts, recommendations or decisions for a given set of human-defined objectives” (ISO, 2022). In the case of such a definition, similarly as in the case of the OECD one, most software potentially could be classified as AI, however, it is not clear if that is the ISO intention.

The Alan Touring Institute (Aitken et al., 2022) cites, after Minsky’s statement during a summer workshop at Dartmouth College in 1956, that AI is “science of making computers do things that require intelligence when done by humans” – which is a very broad, vague definition. In the mentioned report, the Alan Touring Institute states after (Buiten, 2019) that there is no single definition of AI to form the basis of regulation. The disadvantage of such an approach is the lack of clarity regarding the term “intelligence,” which has been discussed above.

The National Institute of Standards and Technology in the United States defines AI after The American National Standard Dictionary for Information Technology (The American National Standard Dictionary for Information Technology ANSDIT, 2022) as (NIST, 2019):

1) a branch of computer science devoted to developing data processing systems that performs functions normally associated with human intelligence, such as reasoning, learning, and self-improvement,
2) the capability of a device to perform functions that are normally associated with human intelligence such as reasoning, learning, and self-improvement.

An appeal to human intelligence may be useful in the scientific research and philosophical discussions mentioned above, although it lacks practical application.

Below, we would like to focus on the legal institutions that create regulations that must be obeyed by companies and inhabitants of certain countries or regions.

In April 2021, there was a pioneering proposal to regulate artificial intelligence methods at the European Union level, the so-called Artificial Intelligence Act (European Commission, 2021). Originally, AI was defined as “software that was developed with one or more of the techniques such as: machine learning, statistical approaches, logic/knowledge-based approaches and could, for a given set of human-defined objectives, generate outputs such as content, predictions, recommendations, or decisions influencing the environments they interacted with” (European Commission, 2021). In 2022, during the EU

Member States' debate on the proposal for the Artificial Intelligence Act, the original AI definition was criticised for being too broad and for its lack of ability to distinguish traditional software from AI. In the compromise text proposal, AI was redefined as "a system designed to operate with a certain level of autonomy and that, based on machine and/or human-provided data and inputs, infers how to achieve a given set of human-defined objectives using machine learning and/or logic- and knowledge based approaches, and produces system-generated outputs such as content, predictions, recommendations or decisions, influencing the environments with which the AI system interacts" (European Commission, 2022). The above-presented definition was inspired by the OECD definition, and it was stated that some statistical approaches that were deleted from the changed definition could also be classified as machine learning. Nevertheless, from our point of view, the discussion regarding the differentiation of so-called traditional software from AI is still unavoidable, especially if there is a significant difference in terms of consequences regarding the classification of software as AI/not AI. The above-mentioned approach still lacks clarity, for example, on how to interpret "a certain level of autonomy."

In 2022, the UK laid out a regulatory model for AI which states that due to little consensus regarding the AI definition among different scientists and institutions, the preferred approach of the UK Parliament is to identify core characteristics of AI instead, allowing for the formulation of more detailed definitions within specific domains and sectors by their regulators (Dorries, 2022). It is an approach that is contrary to the EU approach, where there is an intention of creating a uniform AI definition for all sectors. The core characteristics of AI identified by the UK regulation are (Dorries, 2022):

1) adaptiveness of a technology (AI systems often partially operate on the basis of instructions that have not been expressly programmed with human intent, having instead been "learnt" on the basis of a variety of techniques),
2) the autonomy of the technology (AI often demonstrates a high degree of autonomy, operating in dynamic and fast-moving environments by automating complex cognitive tasks).

The above-presented UK Parliament proposal of AI characteristics is intentionally general and broad. Although it enables flexibility, the proposed approach has several drawbacks: "adaptiveness" and "autonomy" can be differently interpreted in different domains. Moreover, the further evolution of AI may require a change of the approach in the future into a single binding definition, which would cause an operational burden for the companies that have adapted to different, individual definitions.

Although a similar approach towards defining AI has not been adopted in the USA yet, the Algorithmic Accountability Act was enacted in 2022 by the Senate and House of Representatives of the USA (Clarke, 2022). Instead of focusing on AI, it sets requirements regarding automated systems implemented by companies in terms of: transparency and informing consumers about the applied automation. Such an approach emphasises only one of the AI aspects which is automation, though it creates more clarity regarding its interpretation and purpose than many other regulations focused on AI in general. Nevertheless, it leaves the issue of AI definition, interpretation, and its consequences in the USA to the future.

There are also attempts to regulate and support the development of AI in Asia. In Japan, the AI Governance was published in 2022 (Ministry of Economy, Trade and Industry in Japan, 2022). In this regulation, AI is associated with the term "Weak AI," particularly with ML. The regulation focuses less on the topic of AI definition and more on general approaches toward governance and development. Therefore, there is also a mentioned earlier disadvantage of such a definition: it lacks clarity for practical implications that rely on qualifying a given system as AI.

In 2022, China released an AI regulation that touches upon the application of algorithms in the online recommendation systems used by companies (Cyberspace Administration of China, 2022) which is very similar to the above-mentioned USA Act. The regulation forbids the practice of offering different prices to consumers based on their personal data and allows them to decline personalised recommendations. Although it does not focus on an AI definition, the regulation concentrates on practical aspects of a certain part of AI and, similarly as in the USA, leaves some room for further legislations.

In the case of AI application by banking regulators, the European Banking Authority uses the following definition of AI: "Artificial intelligence (AI) systems are software (and possibly also hardware) systems designed by humans that, given a complex goal, act in the physical or digital dimension by perceiving their environment through data acquisition, interpreting the collected structured or unstructured data, reasoning on the knowledge, or processing the information, derived from these data and deciding the best action(s) to take to achieve the given goal […] AI includes several approaches and techniques, such as machine learning, machine reasoning and robotics" (European Banking Authority, 2020). In the case of such an approach, similarly to earlier mentioned attempts, there is too much ambiguous software implemented in practice which would be very hard to correctly interpret as AI/not AI.

When it comes to legal regulations, we can distinguish two trends: one that focuses on governance and general AI treatment (for example, the EU and UK approach) and the other one that focuses on practical aspects of clearly defined software (for example, the USA and Chinese approach). Since clarity allows for a further prudent approach to AI, the latter approach seems to be more appealing in the long term. It leaves the discussion on the AI definition in the research domain instead of incorporating the approach into the legal system, which in our eyes is quite practical. However, not defining

AI and ML approaches explicitly leaves the regulatory risk of future legislative changes. The strict character of both the USA and Chinese regulations confirms the existence of such risk.

The above-presented concepts vary, and it can be difficult to identify a particular method as AI or not AI depending on the definition. However, we can perceive the repeated intention of differentiation of traditional software (or computations) from AI. In the context of banking, most applications of computer systems are quantitative with various degrees of human interaction involved and various influences on the decision-making process. The question is how to define AI in banking in order to clearly state if a given tool should be classified as AI. Another question is if it is necessary to define tools and methods in banking as AI or if the same rules should be applied to all the quantitative tools used in banking. Since many above-presented approaches focus on the aspect of learning from data, the relevant aspect of our considerations is the definition of ML. Some of the above-presented definitions are too broad to directly classify ML as an AI system (for example, the definition of the UK Parliament), however, none of the cited above papers excludes ML from the AI scope.

3. Definitions of ML

The term ML is easier to define than AI due to the fact that it is narrower and less abstract in meaning than "intelligence." For ML, we present a less extensive literature review than in the case of AI because most ML definitions are quite similar to each other. As in the case of the AI section, we start with some definitions from the state-of-the-art insight, then definitions from worldwide organisations that shape the understanding of ML globally, and in the end, we present the approach of different regulators.

In the case of ML, the key aspect is how to exactly define whether a machine, computer, or program is learning. Initially, the term ML was defined as a field of study that provided learning capability to computers without being explicitly programmed (Samuel, 1959). In fact, it is a quite general definition which does not explain what is meant by "learning." Another more specific approach is to define ML as follows: a computer program which is said to learn from experience E with respect to some task T and some performance measure P, if its performance on T, as measured by P, improves with experience E (Mitchell, 1997). Mitchell's proposal has been successfully applied in the engineering set-up and has been also referred to as a useful definition quite recently (Alzubi, Nayyar, Kumar, 2018). In our eyes, such a definition is also quite clear and useful from the scientific perspective. In other words, ML can be also perceived as the automated detection of meaningful patterns in data with endowing programs with the ability to learn and adapt (Osisanwo et al., 2017 after Shalev-Shwartz, Ben-David, 2014: 7).

The mentioned definitions are quite close to each other and seem clear and useful from a scientific perspective. The issue with them is connected with their practical applications, namely what is actually meant when stating that "X is an ML algorithm"? In statistical modelling, the equivalent of the word "learn" would be a process of estimation of model parameters. However, after the parameters are estimated, the learning process ends (at least temporarily). That means that when stating "X is an ML algorithm," it actually means that the algorithm has somehow automatically set parameters based on the available data during the parameters estimation process. Model parameters can be estimated with any optimiser or technique. However, in practice, model parameters can be also set by humans due to various reasons, for example, specific expectations regarding future events. And in such cases, the ML algorithm is no longer ML because it lacks the element of automatic learning from data. From the scientific point of view, such divagation is actually quite irrelevant, however, from the legal point of view and in terms of practical application by companies, it may be important.

Let us focus on some worldwide organisations" ways of defining ML. ML can be defined as "a set of techniques that allows machines to learn in an automated manner through patterns and inferences rather than through explicit instructions from a human" (OECD, 2019). The Alan Touring Institute defines ML as "a system that can perform tasks as a result of a learning process that relies on data" (Aitken et al., 2022). In such a case, our point made above is also binding: a given model can be considered as ML if its parameters are somehow automatically obtained from some kind of data. On the other hand, an ML definition strictly referring to the process of parameters estimation is given by the ISO. In the AI standards, ISO defines ML as: a process of optimising model parameters through computational techniques, such that the model's behaviour reflects the data or experience" (ISO, 2022). In the above-mentioned definition, a model is "a physical, mathematical, or otherwise logical representation of a system, entity, phenomenon, process or data," and a model parameter is "an internal variable of a model that affects how it computes its outputs" (ISO, 2022). Interestingly, in the IT governance standards, the ISO defines ML as "a process using algorithms rather than procedural coding that enables learning from existing data in order to predict future outcomes" (ISO, 2017). On the ISO website, there is information that the standard was reviewed in 2022 and the version remains current (ISO, 2017).

Let us focus on regulatory institutions defining ML. In the proposal of the EU act, the following ML definition applies (European Commission, 2022):

"ML approaches focus on the development of systems capable of learning and inferring from data to solve an application problem without being explicitly programmed with a set of step-by-step instructions from input to output. Learning refers to the computational process of optimising from data the parameters of the model, which is a mathematical construct generating an output based on input data."

"ML approaches include, for instance, supervised, unsupervised and reinforcement learning, using a variety of methods including deep learning, statistical techniques for learning and inference (including, for instance, logistic regression, Bayesian estimation) and search and optimisation methods."

Therefore, the EU Commission also takes the straightforward approach, referring to the ML based on the automatic process of parameters estimation. In 2021, the EBA issued a discussion paper on ML application for the Internal Rating-Based (IRB) approach and uses the above-mentioned definition of the ISO IT governance standard in its papers (European Banking Authority, 2020; 2021). Nevertheless, the EBA argues that within the scope of ML there are different methods characterised by different levels of complexity, and that the term ML was introduced only in 1959. At the same time, even though statistical methods created long before 1959, such as linear or logistic regression, are classified as ML based on the initial ML definition (ISO, 2017), the EBA points out the fact that nowadays practitioners often refer to ML as to more advanced and complex models than, for example, logistic and linear regression (European Banking Authority, 2021). Therefore, in the EBA discussion paper, the term ML is applied only to the models that are characterised by a high number of parameters and, therefore, require for their estimation a large volume of (potentially unstructured) data that are able to reflect non-linear relations between the variables (European Banking Authority, 2021). The potential application of the above-mentioned definition in the banking industry faces several challenges:

1) The nonlinearity of dependencies is not always clear.
2) The complexity of the final output, for example, a linear regression model, is not the only source of complexity involved in the process – others include methods used for the optimisation of parameters or search for optimal hyperparameters with cross-validation (Hastie, Tibshirani, Friedman, 2008), etc. – such algorithms can be much more complex than the form of the final model itself.
3) How to define the volume of data based on which the built model is advanced and complex.
4) Such a definition is vulnerable to the passing of time – every year the complexity of newly proposed methods is growing, and therefore the perception of a complex method can change for different parties involved.
5) Such a definition is not in line with other ML and AI classifications presented by the wide range of regulators and institutes cited above. It may mean that from the EBA perspective a given method applied to IRB is not ML, and therefore it is not an AI, but from the perspective of the EU Commission, it is a full-fledged AI. Therefore, further governance of quantitative methods may be quite challenging.

The EBA approach to defining ML is clearly different than other approaches cited above. The above-presented considerations raise similar questions to the ones raised for AI. How to define ML (being a part of AI) in the banking industry, and is it even necessary to define specific tools used in banking as ML? Maybe a given set of rules regarding quantitative methods in the banking industry should be applied to all tools, without any distinction if they are treated by some as simple and by others as complex. Or taking into consideration that banking is understood as a wide range of different domains, perhaps particular guidelines that cover strictly defined areas (such as IRB in credit risk) should be applied in practice. Before answering these questions, let us look at the main concerns behind using advanced quantitative methods.

4. Requirements towards complex quantitative systems

In the general sense, an interest in AI regulation is due to the dynamic increase in complexity of quantitative solutions applied in many areas, including banking. Since banks are institutions of public trust, it is important to cover the potential uncertainty around the applied solutions. The common concerns that are mentioned in the context of complex quantitative system applications are as follows:

1) It should be lawful. This means that the provider of a quantitative system should be able to prove that the proposed solution complies with applicable laws and regulations. The increasing complexity of quantitative solutions may cause difficulties in determining whether a given solution is legally applicable in practice. The compliance with the law is considered (for example, in European Commission, 2019).
2) It should be ethical. Nevertheless, the exact coverage of the meaning of an ethical quantitative system is quite challenging. The commonly mentioned aspects of an ethical system are as follows: a lack of discrimination and stigmatisation (due to age, gender, ethnicity, disability, etc.), respect for privacy, and prevention of physical and mental harm. The above-mentioned considerations may result in specific questions, for example, if setting a rule according to which people above a certain age are not able to get a mortgage loan is ethical. Another example of a similar issue may be connected with the usage of data from social network activity, for example, social network contacts with a bad credit history that affects the creditworthiness score of a given person (Sadok et al., 2022).
3) It should be transparent. Transparency means that the applied solutions are sufficiently documented. This also means that the algorithms behind particular solutions are clear (traceable) and do what they are supposed to do. Such a system should be also possible to be adequately validated.
4) It should be explainable. In this article, transparency, which is mentioned above, means the clarity of the algorithms applied in the process, while explainability means that the output produced by a system can be understood by a human being in a sufficient way – mainly by determining the degree of impact of input information

on the output. Sufficiency of output understanding is dictated by a specific application of a given method, for example, for one domain, the indication of the most impactful variables on the output would be sufficient, while clear correspondence between input and output (including relationships within the input) is required for another.

5) It should be secure. The developed systems should be implemented in such a way that their usage is secure and the probability of being hacked is very low.
6) It should handle data well. This means that the input data should be of good quality and sufficiently protected. The data governance process should be maintained with proper care.
7) It should be supervised by a human being. Such a system should be the subject of a regular monitoring process that is adapted to the nature of the applied solution.
8) It should be accurate. Some level of bias is unavoidable, for example, in the context of forecasting, nevertheless, it should be measurable. The sources of bias must be identified and sufficiently covered. The bias may arise from several factors (NIST, 2019):
   – systemic bias occurs due to procedures and practices of a given institution,
   – human bias that may arise, for example, as an effect of simplified judgment and heuristics,
   – computational bias may be a consequence of nonrepresentative input applied to a given algorithm or wrong handling of outlier data,
   – algorithmic bias can arise, for example, from: over-fitting or under-fitting of a system to the data used to learn patterns or wrong application of a chosen mathematical representation in order to solve a given problem.

The list of the above-presented concerns can be extended or refined in the context of the specific area of application or institution. The computations based on statistics have been present in the banking industry for decades, however, the mentioned above discussion concerning AI and ML application in banking seems to be causing some issues: starting with the differences between the definition of ML in the document published by the EBA (European Banking Authority, 2021) and other authors/institutions (ISO, 2017; 2022; OECD, 2019; Aitken et al., 2022; European Commission, 2022).

5. Comparison of quantitative tools applied in banking areas

Why it is important to have clear AI and ML definitions in banking is connected with the fact that the banking sector is heavily regulated on many levels (national credit law, national and international banking-specific regulators, etc.). Therefore, regulations concerned with AI and ML governance are focused on imposing additional restrictions and requirements on the tools positively classified as AI or ML. We believe that vague definitions of AI and ML should be avoided. The reasons behind such a claim are as follows:

1) when constructing advanced and costly quantitative frameworks in banks, clarity is the key to an appropriate approach over a long period of time,
2) the lack of clear legal foundations may delay the bank's transition towards more advanced solutions,
3) it is possible that the lack of clarity will result in looking for legal loopholes and for such setting of complex quantitative systems that they are not classified as AI/ML.

Therefore, we would like to:

1) firstly, take a look at different banking areas and briefly present possible AI applications which are based on the available literature and the author's experience,
2) secondly, classify the methods applied above, assess their complexity, interpretability and threats/opportunities in terms of practical applications, and compare their usefulness taking into consideration the restrictive regulatory environment.

Table 1 consists of different banking domains along with assigned possible applications of advanced techniques discussed in this paper. The list of areas is compiled based on the author's perspective, therefore some banking categories are wider than others. Please be aware that the specific applications listed next to the extracted domains do not cover all possible AI applications in particular areas. The range of currently developed and applied solutions makes it impossible to include all the existing solutions. However, presented applications show general trends within specific domains. It is worth mentioning that many of the applications are proposed by scientists and that their practical implementations in banks must be adapted to the existing regulations and be resistant to their potential changes. The regulator approaches are very different for countries and geographical regions such as the EU, which has been discussed in sections 2 and 3.

Table 1. Specific banking areas and possible applications of advanced solutions within those areas

Banking area	Advanced methods applications
Risk management: credit risk	For Probability of Default (PD), the applied methods involve logistic regression, Support Vector Machines or logistic regression with random coefficients (Dong, Lai, Yen, 2010). Other proposals involve Naive Bayes, neural networks, the K-Nearest Neighbour classifier, decision tree or random forest models (Wang et al., 2020). Other commonly used credit risk models are Loss Given Default (LGD), which is a share of an asset that is lost when a client defaults, and Exposure at Default (EAD), which is a predicted loss that the bank may incur in the case of client default. For LGD and EAD, the considered approaches involve e.g.: Naive Bayes, linear regression with data transformations, mixture models, neural networks, and logistic regression (Yang, Tkachenko, 2012). The PD, LGD and EAD are the three main parameters needed to calculate economic/regulatory capital for banking institutions under Basel II. For the credit cards scoring system, the bidirectional long short-term memory (LSTM) neural network has been proposed (Ala'raj, Abbod, Majdalawieh, 2021). For credit risk stress testing purposes, least absolute shrinkage and selection operator regression (LASSO) (Chan-Lau, 2017) and Multivariate Adaptive Regression Splines (MARS) (Jacobs Jr., 2018) have been proposed. The real estate price (where real estate is a loan collateral) can be estimated with various ML techniques such as linear regression, convolutional neural networks, and random forest (Potrawa, Tetereva, 2022).
Risk management/ IT: cybersecurity	Cybersecurity can be classified as a part of operational risk presented below, however, we decided to present it as a separate section. The cybersecurity section presented here is based mainly on Sarker et al. (2020). An intrusion detection system (network and software security) can be built with the application of: Support Vector Machines, neural networks (including recurrent neural networks and LSTM), the K-Nearest Neighbour classifier, the K-means algorithm, Naıve Bayes, the decision tree model, the genetic algorithm, and the hidden Markov model. The Support Vector Machines model has been used for DDoS detection (where the DDoS is Distributed Denial of Service, which is an attack made with multiple computers and internet connections meant to make a machine or network inaccessible for intended users). For malicious activities and anomaly detection, neural networks, Adaboost, decision tree models, and Support Vector Machines have been applied. Probabilistic neural networks have been proposed for user authentication with keystroke dynamics, where keystroke dynamics is the typing style of a client (Revett et al., 2007).

Banking area	Advanced methods applications
Risk management: liquidity risk	Artificial neural networks and Bayesian networks can be applied to various liquidity risk processes such as stress tests, simulations, recovery and contingency plan. AI/ML can also improve the Internal Liquidity Adequacy Assessment Process (ILAAP) and Asset Liability Management (ALM) processes (Milojević, Redzepagic, 2021). For a liquidity risk early warning prediction system, LASSO regression, random forest and gradient boosting with decision trees have been proposed (Drudi, Nobili, 2021). For early warning liquidity risk, system neural networks and Bayesian networks have been also proposed (Tavana et al., 2018).
Risk management: operational risk	In the case of operational risk, the area where advanced quantitative techniques are heavily explored, the Know Your Customer (KYC) process is used. The KYC guards the bank against financial fraud (including credit card fraud), money laundering, and terrorist financing. For anti-money laundering, Support Vector Machines (Keyan, Tingting, 2011; Chen, 2020), neural networks (González, Velásquez, 2013), Bayesian networks (Khan et al., 2013), decision trees and random forests (Chen, 2020) have been proposed. For fraud detection systems, Bayesian algorithms, the K-Nearest neighbour, Support Vector Machines and the bagging ensemble classifier based on the decision tree model have been applied (Pun, Lawryshyn, 2012; Dal Pozzolo, 2015; Leo, Sharma, Maddulety, 2019).
Risk management: interest rate risk	The topic connected with both interest rate risk and liquidity risk is matching of maturity profiles of assets and liabilities. Therefore, in the case of assets, loan prepayment models are applied and in the case of liabilities, savings and current accounts churn prediction models are used. In the case of savings accounts, churn prediction neural networks, gradient boosting based on decision trees, the Generalised Linear Model (GLM), Support Vector Machines and random forests have been applied (Verma, 2020). For prepayment modelling, random forest alongside the proportional hazard model (Liang, Lin, 2014), neural networks (Zhang, Teng, Lin, 2019), logistic regression (Zahi, Achchab, 2020), and the gradient boosting classifier based on decision trees (Schultz, Fabozzi, 2021) have been proposed. For future interest rates prediction, the Gaussian mixture model (Kanevski, Timonin, 2010) has been proposed.

Banking area	Advanced methods applications
Risk management: market risk	For investment risk prediction, the Adaboost Support Vector Machine has been proposed (Luo, Metawa, 2019). For the Credit Default Swap (CDS) derivative, the spread approximation method with random forest regression has been proposed (Mercadier, Lardy, 2019). For evaluation of risk premium of commodity futures contracts, LSTM neural networks have been proposed (Rad et al., 2021). Value at Risk (VaR) models are commonly applied in a market risk area, for example, for equity risk. The VaR computes a maximum loss over a given period with an assumed level of confidence. For VaR calculation, an important aspect is the future volatility prediction. For volatility estimation, neural networks and the Generalised Autoregressive Conditional Heteroskedastic (GARCH) model have been proposed (Monfared, Enke, 2014; Zhang et al., 2017). For foreign exchange (FX) risk, the genetic algorithm alongside the LSTM neural network has been proposed (Loh et al., 2022). For derivatives pricing, neural networks and boosted random trees have been proposed (Ye, Zhang, 2019).
Risk management: model risk	AI and ML approaches can be applied during the validation of applied quantitative systems in different banking areas. Advanced ML models can be built as a benchmark for the existing, simpler models. For data quality validation, the outlier detection with ML can be applied, for example, based on the Gaussian Mixture Model, the Dirichlet Process Mixture Model, neural networks, probabilistic principal component analysis (PPCA), Support Vector Machines (Domingues et al., 2018), or the so-called isolation forest model based on the random forest algorithm (Liu, Ting, Zhou, 2008).
Customer experience	Nowadays customer experience in the case of using digital banking and visiting bank stationary branches can be simplified and time-optimised with the application of ML. With convolutional neural networks widely applied for image recognition (Hijazi, Kumar, Rowen, 2015; Liu, 2018), the bank can apply models that significantly shorten time necessary for the existing processes. The image recognition system can be applied, for example, to: a model that automatically reads information from the client ID or a model that verifies if the next tranche of a mortgage can be transferred to the client based on the construction progress documented with photos. In the case of digital banking, a personalised system with transaction categorisation and cash flow prediction can be built with recurrent neural networks (Kotios et al., 2022). The automated credit risk scoring system, offering customised loans to the clients based on their characteristics (every-month cash flows, etc.) can be developed with ML techniques (discussed in the credit risk section above). Advanced chat bots based on a neural networks approach called Natural Language Processing (NLP) (Adamopoulou, Moussiades, 2020) can be developed, for example, to call customers with a reminder of an overdue loan instalment. To automatically propose banking products that a given client would be most interested in, banks can develop a profiling system adapted to digital banking, for example, based on k-means and neural networks algorithms (Dawood, Elfakhrany, Maghraby, 2019).

Source: own elaboration

Most of the solutions proposed for application in given banking areas presented in Table 1 appear repeatedly in various banking areas due to the fact that most of those quantitative tools are quite universal and can be applied to a wide range of problems covered by banking. Table 2 consists of the author's subjective assessment of the above-mentioned methods alongside the threats and opportunities connected with their applications. Table 2 comprises only methods that are:

1) meant to be applied to datasets with observable dependent variables (labelled dataset),
2) meant for classification or regression problems (or both).

Because of the above, optimisation algorithms (such as genetic algorithms) or methods meant to group a given dataset with no dependent variable (unlabelled dataset) into several categories (such as the K-means algorithm) are out of the scope of the proposed categorisation. In the case of Table 2, the author assumes that the evaluation of complexity and interpretability is connected with the situation which is not over-simplified, i.e. at least several auxiliary variables are assumed to be considered and particular methods are applied due to their unique features, which means that, for example, the considered neural network has at least several neurons with activation function(s) or Support Vector Regression is not based on linear kernel. The evaluated methods include:

1) Linear regression based on data with no transformations (Jiang, 2007: 1).
2) Linear models based on transformed data, including multinomial regression, the model based on data pre-processed with PCA or PPCA (Fraser, 1967; Abdi, Williams, 2010).
3) Generalised Linear Models, including logistic regression (Dobson, Barnett, 2018).
4) Regularised linear regression, including LASSO, ridge regression, and elastic net regression (Hoerl, Kennard, 1970; Tibshirani, 1996; Zou, Hastie, 2005).
5) Decision tree for both classification and regression (Breiman et al., 1984).
6) Random forest for both classification and regression (James et al., 2017: 319).
7) Mixed (mixture) models: models with both fixed and random coefficients, including Gaussian mixture models and Drichlet process mixture models (Neal, 2000; Biecek, 2013; Rao, 2013: 96; Krennmair, Schmid, 2022).
8) Multivariate Adaptive Regression Splines (Friedman, 1991).
9) Boosting methods for classification and regression, including Adaboost, Gradient boosting with decision trees, boosting with other models, i.e. linear regression or Support Vector Machines, and the bagging ensemble classifier (Schapire, 2003; Hastie, Tibshirani, Friedman, 2008; Chen, Guestrin, 2016).
10) Markov models, including hidden Markov models and Monte Carlo hidden Markov models (Thrun, Langford, 1998).
11) Neural networks, including recurrent, LSTM, bidirectional LSTM, convolutional, NLP, etc. (Hochreiter, Schmidhuber, 1997; Goodfellow, Bengio, Courville, 2016; Li et al., 2022).

12) Bayesian methods, including Naive Bayes and Bayesian networks (Ren et al., 2009; Mihaljević, Bielza, Larrañaga, 2021).
13) Support Vector Machines, including regression implementation as Support Vector Regression (Vapnik, Levin, Cun, 1994; Smola, Schölkopf, 2004).
14) Ensemble methods: results of several different complex models combined together (Ganaie et al., 2022).
15) Proportional hazard models (Basu, Manning, Mullahy, 2004).
16) K-Nearest Neighbour classifier (Cunningham, Delany, 2021).
17) Autoregressive models, including GARCH (Bauwens, Laurent, Rombouts, 2006).

Table 2. Assessment of advanced quantitative methods in the context of banking

Quantitative method	**Complexity/ Interpretability***	**Opportunities of application in the banking area**	**Threats of application in the banking area**
Linear regression	Very low complexity/ very easy to interpret	Simplicity, understandability of the method by a lot of staff, the short time needed to estimate and interpret the model, the method easy to explain to high management and the regulator	Works well only for linear dependencies between variables, the necessity of testing for the fulfillment of assumptions, the regulator may potentially question such a solution as the one with too weak predictive power
Linear models based on transformed data	Low complexity/ moderate to interpret	Transformations can be tailored to a specific economic theory; the application of transformations may be a great way to enhance existing linear models in banking	In the case of PCA and more complicated transformations, the interpretability, in general, is more difficult, multinomial regression and some transformations may produce unstable predictions

Quantitative method	Complexity/ Interpretability*	Opportunities of application in the banking area	Threats of application in the banking area
Generalised Linear Models	Low complexity/ easy to interpret	Logistic regression is a staple in the case of binary classification, for example, in the case of PD modelling, ease of interpretation, which is especially important in the case of explanation of the reasons for loan rejection to clients, etc.	Using less popular link functions than the logit/probit requires more knowledge and understanding of statistics, the assumption regarding the independence of random variables, and in some cases applying Generalised Linear Models may be questioned as too simple in the final form and at the same time too complex in terms of requirements regarding distribution assumptions, etc.
Regularised linear regression	Very low complexity/ easy to interpret	May enhance the performance of linear regression models at the same time being relatively simple, very helpful if the linear model is the right choice, but the issues with overfitting were detected (too good fit to the training dataset with considerably worse performance in the case of actual predictions)	The existence of hyperparameter requires some serious enhancement of the estimation process, which is also more time-consuming than the standard linear regression estimation process, the choice of regularised regression over non-regularised requires additional documentation
Decision tree	Very low complexity/ very easy to interpret	Simplicity may work very well in the case of simple segmentation tasks, interpretability, easy to explain, visualise, and capture non-linear relationships, may be easily combined with other methods	Inadequate for more complex problems, a single decision tree is very sensitive to the dataset based on which model parameters are estimated, which may result in weak performance on new data, requires additional testing for stability
Random forest	Average complexity/ hard to interpret	Greatly improves stability issues of single decision trees, fast computation, despite being hard to interpret, the algorithm behind random forests is relatively easy to explain	Requires model agnostic methods in order to interpret results and therefore in most cases extensive documentation is essential

Quantitative method	Complexity/ Interpretability*	Opportunities of application in the banking area	Threats of application in the banking area
Mixed (mixture) models	Average complexity/ moderate to interpret	Random effects provide great tools for specific requirements regarding available data and can handle them better than fixed effects, they can be applied to many different practical problems, mixture models can be associated with different models (linear models as well as Generalised Linear Models or even random forests)	Mixture models require additional testing, i.e. regarding assumed distributions of random effects, random effects are not so easy to interpret and explain to higher management as fixed effects, application of random effects to more complex models, for example, to random forests significantly increases time needed for computation
Multivariate Adaptive Regression Splines	Low complexity/ easy to interpret	Elasticity, simple for interpretation, automatic selection of the auxiliary variables for the model, computer implementations of the model are time-efficient	No possibility of explicitly presenting the formulas describing the confidence intervals for the parameters, similarly to a single decision tree, may be inadequate for more complex problems and requires additional testing for stability
Boosting methods for classification and regression	High complexity/ very hard to interpret	Models based on boosting often give very accurate results (many Kaggle competitions were won with the application of boosting methods), existence of very time-effective software implementations, boosting may be applied for different classes of base models, e.g.: linear models, decision trees, etc.	Results are very difficult to interpret (require model agnostic methods), the correct choice of hyperparameters may be difficult, sensitivity to overfitting to training data (in that case, the model fits the training data well but does not work well in the case of actual predictions)
Markov models	Average complexity/ easy to interpret	A wide variety of applications, strong economic background	The method is based on discrete states, which may cause serious technical issues in implementation, the method is superseded often by other more accurate models, frequently requires other models/methods to produce accurate results

Quantitative method	Complexity/ Interpretability*	Opportunities of application in the banking area	Threats of application in the banking area
Neural networks	Very high complexity/ very hard to interpret	Very accurate if done correctly, and the method fits well with many classes of problems, despite being very complex, there are a lot of scientific materials and tutorials available	A very time-consuming method, requires specific knowledge and experience in order to correctly choose network architecture, sensitive to overfitting, requires model agnostic methods for results interpretation and very extensive documentation
Bayesian methods	High complexity/ easy to interpret	Easy to interpret, Bayesian methods provide a convenient setting for a wide range of methods, for example, issues with missing data	May produce misleading results in certain cases, is time-consuming, requires expert knowledge and experience to produce an accurate model that works well in practice
Support Vector Machines	High complexity/ hard to interpret	A computationally time-effective method even with large datasets, many statistical software implementations, despite being a relatively complex method, the number of hyperparameters is relatively small (for example, in comparison with Gradient boosting based on decision trees)	The method difficult to interpret, among other advanced algorithms relatively hard to understand and explain to higher management
Ensemble methods	Very high complexity/ very hard to interpret	May produce very accurate tools if done well, the future-proof method in terms of accuracy	In the case of ensemble of several complicated models, the interpretability may be very hard or even impossible, a very time-consuming method that requires a lot of knowledge and carefulness from its practitioners, due to poor interpretability can be explored only in the areas where interpretability is not required, neither expected

Quantitative method	Complexity/ Interpretability*	Opportunities of application in the banking area	Threats of application in the banking area
Proportional hazard models	Average complexity/ easy to interpret	If the survival time is available, this method is often preferred over the logistic regression approach, easy to interpret, robust (accurate even in the case of no specification of baseline haard)	Hazards should be proportional in order to work well, sometimes the final model may be tricky to implement, and sometimes the hazard ratio may be misleading in interpretation
K-Nearest Neighbour classifier	Low complexity/ easy to interpret	Easy to understand and implement in practice	Time-consuming computations if the dataset is large, sensitivity to redundant data (earlier feature selection advised), in the case of more difficult tasks may be outperformed by more complex methods
Autoregressive models (including GARCH)	Low complexity/ easy to interpret	A simple mechanism, may work well in the case of economic banking data, easy to interpret, may be very effective when there is a lack of suitable auxiliary information	The data autocorrelation should be at a certain level in order to work well, a significant amount of historical data is required in order to estimate parameters accurately

* Interpretability means general clarity of the impact that input has on output of the model.

Source: own elaboration

In cases where strict interpretability is required, higher management may be reluctant to approve changes to the model incorporating more accurate but at the same time potentially less interpretable solutions. In such cases, we propose a gradual transition towards more complicated approaches. For example, in the case of credit risk modelling, it may be impossible to transition from logistic regression to a neural network due to the anxiety of higher management or regulatory barriers. As indicated in Table 2, methods difficult to interpret would be very hard or impossible to apply in the case of strict interpretability requirements. However, in such cases, existing solutions may be enhanced with methods that are easy to interpret. The gradual improvement may be connected, for example, with:

1) the incorporation of random effects into the models with only fixed effects, as in the case of transition from logistic regression to logistic regression with random effects in PD modelling (Dong, Lai, Yen, 2010),

2) the introduction of additional segmentation into the modelling problem, for example, in the case of modelling a given portfolio with linear regression, the improvement may include the division of a given dataset into several sub-portfolios with a decision tree and then the application of separate regularised linear regression to each.

The advantages of a gradual improvement of the existing solutions are as follows:

1) executives who make final decisions may perceive such proposals as safer and accept them more willingly,
2) the risk of unexpected issues (such as loss of social trust and reputation) occurring in the course of implementation is lower than in the case of a sudden change of a simple model into a much more complex model,
3) such models are easier to develop due to the greater use of existing solutions than in the case of a complete paradigm change,
4) it is easy to compare old and enhanced versions of the model in terms of predictions, accuracy, result simulations, etc.,
5) such models are more robust to potential regulatory changes regarding AI treatment.

6. Discussion

In many areas of the banking industry, for example, in the field of credit risk, quantitative solutions that are widely applied are already regulated to some extent. The approaches such as linear or logistic regression are examples of quantitative models that are a base for many quantitative systems applied in practice. The banking industry already aims to become data-driven and comply with all regulations, be ethical and trustworthy. There are differences between certain domains, for example, the credit risk area is generally more strictly and in detail regulated than other areas such as, for example, models that aim to propose to clients a new product in mobile banking applications. Given the current state of quantitative solutions in banking as well as the current approach of many regulators and institutions towards the regulation of AI and ML for the legislative purpose, we propose the following:

– a definition of AI that is general and basically classifies all quantitative systems mentioned in the article as AI: "AI is a system that at certain stages of its process perceives its environment and acts autonomously,"
– a definition of ML that is also general and includes all the quantitative approaches mentioned in the article: "ML is an approach that generates output based on some input data of any sources" (for example, expert opinion parameters or observation of a particular phenomenon).

Due to the inclusive character of the above-presented definitions, basically all quantitative and autonomous processes in the banking industry can be qualified as AI. Therefore, the regulations concerning applied solutions should be applicable to the methods being staples in the industry as well as to the new proposals considered by experts from particular domains. This means that the definition fulfils important regulatory goals, i.e. clarity and ease of practical application.

The above-proposed approach towards legislative definitions would mean that requirements regarding specific AI/ML challenges could be formulated in a specific context, for example, of credit risk. The same effect can be achieved if the AI/ML definition is not a part of the regulation at all (as in the case of the US and Chinese examples discussed in section 2). However, in such cases, the potential threat of future regulatory interpretations of AI/ML remains and the uncertainty is often one of the factors that blocks innovations and changes. Therefore, we argue that a broad definition of AI/ML in the case of regulations is a better solution. In the case of the regulation incorporating our proposal, its scope regarding specific recommendations should either be focused on AI/ML requirements in general (with an intention that most quantitative solutions qualify as such) or/and on requirements orientated towards specific domains (for example, credit risk modelling).

The regulations aiming towards requirements of explainability of AI methods in the credit risk area should be straightforward and should not lead to a situation where, for example, a method classified by the EU as AI is not an AI in the understanding of EBA. Moreover, the general approach towards the AI definition allows for the formulation of different requirements in different banking areas which are not connected with classifying a given method as AI. If our proposal is applied, basically all the present quantitative solutions applied in banking would be classified as AI. At the same time, all the models that are based on data would be classified as ML. The most important advantage of such an approach is regulatory clarity. From a practical perspective, such an approach does not distinguish between more and less complex approaches, and therefore one can argue that it is not necessary to bother the industry with defining the underlying concepts. AI and ML could be also defined in a less broad manner, for example, with the requirement of learning aspect in the case of AI. Nevertheless, that requires a clear definition of what learning from data means. Even if learning from data is defined in a narrow way, it would probably mean that simple linear regression and complex neural network models end up classified in the same category, therefore, e.g.: explainability requirements towards application of computational methods to a given problem should be formulated anyway. If regulators would like to distinguish simple approaches from complex ones, there should be clear instructions to make a cut-off between the AI model and the non-AI model – and such a task is not easy. The discussion over such a distinction is difficult mainly due to the potential black-box character of complex, nonlinear solutions, and the inability of a human to understand, for example, relationships between variables created by a model or to correctly interpret model results.

A significant issue in the discussion over clarity and explainability of the applied solutions is the complexity of the methods that are indirectly used in the process (the methods that search for optimal model parameters/hyperparameters, e.g.: genetic algorithms or algorithms that aim to make model results more explainable). For example, the proposal of simplification of complex models based on explanatory method outcomes (Gosiewska, Kozak, Biecek, 2021) makes the final model parameters more understandable in terms of final outcome interpretation, however, the whole algorithm that leads to the final solution is much more complex than for standard, statistical solutions such as linear or logistic regression because it involves:
1) the development of a complex model (with an optimal parameters/hyperparameters search),
2) the application of a selected explainable method algorithm,
3) an additional algorithm that shifts the results of the explainable method to create a new, simplified model,
4) the testing of the final solution.

Therefore, banking industry regulators should focus on the following aspects of quantitative AI systems:
1) the definition of sufficient explainability of quantitative methods applied in banking,
2) the required level of final result clarity – for example, if it is necessary to point to the most important client features that resulted in refusal to grant a loan and to clearly state their numerical influence on the final outcome.

The required level of model clarity would make it unnecessary to distinguish between more and less complex solutions (AI or not AI). The interpretation of results produced, for example, by linear regression should be relatively easy, while the fulfilment of the same requirements by a neural network could be a challenging task. Therefore, independently of the level of complexity of a given system, if a model developer can fulfil a certain level of algorithm clarity and final outcome interpretation, the proposed solution can be applied to a certain banking domain. Nevertheless, the definition of sufficient explainability level is also not an easy task due to the stochastic nature of many methods proposed to explore the model outcomes interpretation.

A broad approach towards AI and ML can help to create awareness regarding the understanding of AI by society. For example, the statement that in order to explain the outcome of a particular AI system, another AI system has to be created may arouse anxiety, unless such a statement appears along with the information that basically all quantitative systems based on data are classified as AI by regulators. Society is generally concerned by the vision of an AI that: has an unclear goal (black box), gets out of control, and potentially can generate harm that is nonpredictable. An interesting opinion regarding that matter is that instead of identifying systems that produce bad results as essentially mysterious and uncontrolled, we should call the application of an inappropriate technology a highly likely abuse purposely made by the creator of the system – the cultural logic of complex and mysterious technology is often used to justify the close involvement of the AI industry in policymaking and regulation, which further strengthens the market position of large companies and provides the legitimacy of their inclusion in regulatory processes (Cath, 2018; Kroll, 2018).

7. Conclusions

In this research, we have attempted to study AI and ML applications in banking. Firstly, we discussed the definitions of AI and ML applied by scientists, worldwide organisations, and different regulators across the world. The conclusions resulting from our research regarding those definitions are as follows:

1) Most definition emphasise differentation between "traditional software" and AI/ML methods.
2) Scientists focus on philosophical aspects of the correct definition of the term "intelligence," which generally makes the task of clear classification of particular methods as AI quite impossible.
3) Worldwide organisations and regulators generally try to construct a definition that can help practitioners classify particular methods as AI, however, the cut-off point between AI and non-AI methods is often difficult to detect. Some regulators avoid defining AI/ML and instead focus on a regulatory approach towards specific, narrowly defined areas.
4) In the case of ML, the approaches towards its definition on the part of scientists, worldwide organisations, and regulators are, unlike in the case of AI, quite similar. Nevertheless, particular definitions often lack the clarity that would allow for classifying specific methods as ML without any doubts.

Secondly, we have discussed general requirements for quantitative systems, which should be lawful, ethical, secure, transparent, able to handle data well, explainable, accurate, and supervised by human beings.

Thirdly, we have compared proposals for the application of complex quantitative methods in different banking domains. We have also assessed quantitative methods in terms of complexity, interpretability, as well as opportunities and threats of their application in banking. The conclusions arising from our research are as follows:

1) There are a lot of proposals for complex AI/ML solutions to different banking problems. Nevertheless, the requirements regarding the interpretability of applied quantitative methods are dependent on geographical regions and banking domains, therefore the application of specific proposals by the banks can be difficult or impossible.
2) In the cases of strict requirements regarding the interpretability of applied methods, we propose a gradual increase of the complexity of existing methods, for example, with an additional application of methods that were assessed by us as relatively easy to interpret.

For legislative purposes, we have proposed our own definitions of AI and ML that are generally inclusive regarding the practical classification of quantitative methods as AI/ML. The primary advantage of such a proposal is regulatory clarity for practitioners who must adhere to given regulations.

Our conclusions are important because there are a lot of differences between existing approaches toward AI/ML definitions in different environments (scientists/practitioners/regulators), and different understandings of the application of complex quantitative methods by groups interested in the further development of advanced computational methods in banking. Our findings are useful for policymakers in terms of defining AI or ML in future regulations, practitioners in terms of AI/ML applications, and executives in terms of knowledge regarding complex quantitative methods, the possibility of their applications in different banking domains, and their characteristics – to make more accurate and faster decisions.

8. Limitations and future work

The main limitations of the above-presented considerations are that they refer to the industry in general, with no distinction between small banks and large international financial institutions. They also ignore the geographical location aspect. Future work regarding the approach towards increasing complexity of data-driven approaches in the banking industry may focus on several aspects mentioned earlier but not elaborated in detail in the article:
1) an analysis of consequences of differences between existing regulations in different geographics,
2) a proposal of clear guidance over explainability regarding specific applications of AI in different banking domains.

The further development of AI and the increasing complexity of a wide range of systems applied across different industries is expected by societies all over the world. The applied terminology and proposed regulations must be clear in order to make reality more transparent rather than more unclear.

References

Abdi H., Williams L.J. (2010), *Principal Component Analysis*, “Wiley Interdisciplinary Reviews: Computational Statistics”, vol. 2, no. 4, pp. 433–459.

Adamopoulou E., Moussiades L. (2020), *An Overview Of Chatbot Technology*, [in:] *Artificial Intelligence Applications and Innovations: 16th IFIP WG 12.5 International Conference, AIAI 2020, Neos Marmaras, Greece, Proceedings*, Part II , 16, pp. 373–383.

Aitken M., Leslie D., Ostmann F., Pratt J., Margetts H., Dorobantu C. (2022), *Common Regulatory Capacity for AI*, https://doi.org/10.5281/zenodo.6838946 [accessed: 1.11.2022].

Ala’raj M., Abbod M.F., Majdalawieh M. (2021), *Modelling Customers Credit Card Behaviour Using Bidirectional LSTM Neural Networks*, “Journal of Big Data”, vol. 8, no. 1, pp. 1–27.

Alzubi J., Nayyar A., Kumar A. (2018), *Machine Learning from Theory to Algorithms: An Overview*, “Journal of Physics: Conference Series”, vol. 1142, https://doi.org/10.1088/1742-6596/1142/1/012012

Artificial Intelligence: A Modern Approach, 4th Global ed. (2022), https://aima.cs.berkeley.edu/global-index.html [accessed: 11.02.2023].

Basu A., Manning W.G., Mullahy J. (2004), *Comparing Alternative Models: Log Vs Cox Proportional Hazard?*, “Health Economics”, vol. 13, no. 8, pp. 749–765.

Bauwens L., Laurent S., Rombouts J.V. (2006), *Multivariate GARCH models: a Survey*, “Journal of Applied Econometrics”, vol. 21, no. 1, pp. 79–109.

Biecek P. (2013), *Analiza danych z programem R. Modele liniowe z efektami stałymi, losowymi i mieszanymi*, Wydawnictwo Naukowe PWN, Warszawa.

Breiman L., Friedman J.H., Olshen R.A., Stone C.J. (1984), *Classification and Regression Trees*, The Wadsworth and Brooks, Belmond.

Buiten M.C. (2019), *Towards Intelligent Regulation of Artificial Intelligence*, “European Journal of Risk Regulation”, vol. 10, no. 1, pp. 41–59.

Cath C. (2018), *Governing Artificial Intelligence: Ethical, Legal and Technical Opportunities and Challenges*, “Philosophical Transactions of the Royal Society A”, vol. 376, 20180080.

Chan-Lau M.J.A. (2017), *Lasso Regressions and Forecasting Models in Applied Stress Testing*, International Monetary Fund, Washington.

Chen T., Guestrin C. (2016), *XGBoost: A Scalable Tree Boosting System*, [in:] *KDD “16: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, New York, pp. 785–794, https://dl.acm.org/doi/proceedings/10.1145/2939672 [accessed: 12.02.2023].

Chen T.H. (2020), *Do You Know Your Customer? Bank Risk Assessment Based on Machine Learning*, “Applied Soft Computing”, vol. 86, 105779.

Clarke Y.D. (2022), *H.R.6580 – 117th Congress (2021–2022): Algorithmic Accountability Act of 2022*, https://www.congress.gov/bill/117th-congress/house-bill/6580/text [accessed: 12.02.2023].

Cunningham P., Delany S.J. (2021), *K-Nearest Neighbour Classifiers-A Tutorial*, “ACM Computing Surveys (CSUR)”, vol. 54, no. 6, pp. 1–25.

Cyberspace Administration of China (2022), *Provisions on the Management of Algorithmic Recommendations in Internet Information Services*, https://www.chinalawtranslate.com/en/algorithms/ [accessed: 12.02.2023].

Dal Pozzolo A. (2015), *Adaptive Machine Learning for Credit Card Fraud Detection*, Unpublished doctoral dissertation, Université libre de Bruxelles, Faculté des Sciences – Informatique, Bruxelles.

Dawood E.A.E., Elfakhrany E., Maghraby F.A. (2019), *Improve Profiling Bank Customer's Behavior Using Machine Learning*, "IEEE Access", vol. 7, pp. 109320–109327.

Dobson A.J., Barnett A.G. (2018), *An Introduction to Generalized Linear Models*, CRC Press, Boca Raton.

Domingues R., Filippone M., Michiardi P., Zouaoui J. (2018), *A Comparative Evaluation of Outlier Detection Algorithms: Experiments and Analyses*, "Pattern Recognition", vol. 74, pp. 406–421.

Dong G., Lai K.K., Yen J. (2010), *Credit scorecard based on logistic regression with random coefficients*, "Procedia Computer Science", vol. 1, no. 3, pp. 2463–2468.

Dorries N. (2022), *Establishing a Pro-Innovation Approach to Regulating AI. An Overview of The UK's Emerging Approach, Presented to Parliament by The Secretary of State for Digital, Culture, Media and Sport by Command of Her Majesty, Department for Digital, Culture, Media & Sport, UK Parliament Command Paper: CP 728*, London.

Drudi M.L., Nobili S. (2021), *A Liquidity Risk Early Warning Indicator for Italian Banks: a Machine Learning Approach*, Bank of Italy Temi di Discussione (Working Paper) no. 1337.

Emmert-Streib F., Yli-Harja O., Dehmer M. (2020), *Artificial Intelligence: A Clarification of Misconceptions*, "Myths and Desired Status. Frontiers in Artificial Intelligence", no. 3, 524339.

European Banking Authority (2020), *EBA Report on Big Data and Advanced Analytics*, EBA/REP/2020/01, Paris.

European Banking Authority (2021), *EBA Discussion Paper on Machine Learning for IRB Models*, EBA/DP/2021/04, Paris.

European Commission (2019), *Ethics Guidelines for Trustworthy AI. High-Level Expert Group on Artificial Intelligence*, Brussels.

European Commission (2021), *Proposal for a Regulation of The European Parliament and of The Council Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act) and Amending Certain Union Legislative Acts*, Brussels.

European Commission (2022), *Proposal for a Regulation of The European Parliament and of The Council Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act) and Amending Certain Union Legislative Acts – Second Presidency compromise text*, Brussels.

Fraser D.A.S. (1967), *Data Transformations and The Linear Model*, "The Annals of Mathematical Statistics", pp. 1456–1465.

Friedman J.H. (1991), *Multivariate Adaptive Regression Splines*, "Annals of Statistics", vol. 19, no. 1, pp. 1–67.

Ganaie M.A., Hu M., Malik A.K., Tanveer M., Suganthan P.N. (2022), *Ensemble deep learning: A review*, "Engineering Applications of Artificial Intelligence", vol. 115, 105151.

González P.C., Velásquez J.D. (2013), *Characterization and Detection of Taxpayers with False Invoices Using Data Mining Techniques*, "Expert Systems with Applications", vol. 40, no. 5, pp. 1427–1436.

Goodfellow I., Bengio Y., Courville A. (2016), *Deep Learning*, MIT Press, Cambridge.

Gosiewska A., Kozak A., Biecek P. (2021), *Simpler Is Better: Lifting Interpretability-Performance Trade-Off Via Automated Feature Engineering*, "Decision Support Systems", vol. 150, 113556.

Hastie T., Tibshirani R., Friedman J. (2008), *The Elements of Statistical Learning*, Springer Science + Business Media LLC, New York.

Hijazi S., Kumar R., Rowen C. (2015), *Using Convolutional Neural Networks for Image Recognition*, Cadence Design Systems Inc., San Jose.

Hochreiter S., Schmidhuber J. (1997), *Long Short-Term Memory*, "Neural Computation", vol. 9, no. 8, pp. 1735–1780.

Hoerl A.E., Kennard R.W. (1970), *Ridge Regression: Biased Estimation for Nonorthogonal Problems*, "Technometrics", vol. 12, no. 1, pp. 55–67.

Hopcroft J.E., Motwani R., Ullman J.D. (2007), *Introduction to Automata Theory, Languages, and Computation*, Addison-Wesley, Boston.

ISO (2017), *ISO/IEC 38505–1: 2017 Information technology – Governance of IT – Governance of data – Part 1: Application of ISO/IEC 38500 to the governance of data*, https://www.iso.org/standard/56639.html [accessed: 1.11.2022].

ISO (2022), *ISO/IEC 3WD 22989 Information Technology – Artificial Intelligence – Artificial Intelligence Concepts and Terminology*, https://www.iso.org/standard/74296.html [accessed: 1.11.2022].

Jacobs Jr. M. (2018), *The Validation of Machine-Learning Models for The Stress Testing of Credit Risk*, "Journal of Risk Management in Financial Institutions", vol. 11, no. 3, pp. 218–243.

James G., Witten D., Hastie T., Tibshirani R. (2017), *An Introduction to Statistical Learning with Applications in R*, Springer, New York.

Jiang J. (2007), *Linear and Generalized Linear Mixed Models and Their Applications*, Springer Science + Business Media LLC, New York.

Kaisler S.H., Espinosa J.A., Armour F., Money W.H. (2014), *Advanced Analytics – Issues and Challenges in a Global Environment*, [in:] *47th Hawaii International Conference on System Sciences*, Waikoloa, pp. 729–738.

Kambhampati S. (2017), *On the Past and Future of AI. Interviews With Experts in Artificial Intelligence*, https://goo.gl/nspv6y [accessed: 12.02.2023].

Kanevski M.F., Timonin V . (2010), *Machine Learning Analysis and Modeling of Interest Rate Curves*, ESANN, Bruges.

Kaplan A., Haenlein M. (2019), *Siri, Siri, In My Hand: Who's The Fairest in The Land? On The Interpretations, Illustrations, and Implications of Artificial Intelligence*, "Business Horizons", vol. 62, no. 1, pp. 15–25.

Keyan L., Tingtin Y. (2011), *An Improved Support-Vector Network Model for Anti-Money Laundering*, [in:] *Fifth International Conference on Management of e-Commerce and e-Government*, IEEE, Shanghai, pp. 193–196.

Khan N.S., Larik A.S., Rajput Q., Haider S. (2013), *A Bayesian Approach for Suspicious Financial Activity Reporting*, "International Journal of Computers and Applications", vol. 35, no. 4, pp. 181–187.

Kotios D., Makridis G., Fatouros G., Kyriazis D. (2022), *Deep Learning Enhancing Banking Services: A Hybrid Transaction Classification and Cash Flow Prediction Approach*, "Journal of Big Data", vol. 9, no. 1, 100.

Krennmair P., Schmid T. (2022), *Flexible Domain Prediction Using Mixed Effects Random Forests*, "Journal of the Royal Statistical Society: Series C (Applied Statistics)", vol. 71, no. 5, pp. 1865–1894.

Kroll J.A. (2018), *The Fallacy of Inscrutability*, "Philosophical Transactions of the Royal Society A", vol. 376, 20180084.

Leo M., Sharma S., Maddulety K. (2019), *Machine Learning in Banking Risk Management: A Literature Review*, "Risks", vol. 7, no. 1, 29.

Li Z., Liu F., Yang W., Peng S., Zhou J. (2022), *A Survey of Convolutional Neural Networks: Analysis*, "Applications, and Prospects, IEEE Transactions on Neural Networks and Learning Systems", vol. 33, no. 12, pp. 6999–7019.

Liang T.H., Lin J.B. (2014), *A Two-Stage Segment and Prediction Model for Mortgage Prepayment Prediction And Management*, "International Journal of Forecasting", vol. 30, no. 2, pp. 328–343.

Liu F.T., Ting K.M., Zhou Z.H. (2008), *Isolation Forest*, [in:] *Eighth IEEE International Conference on data Mining*, Pisa, pp. 413–422.

Liu Y.H. (2018), *Feature Extraction and Image Recognition with Convolutional Neural Networks*, "Journal of Physics: Conference Series", vol. 1087, no. 6, 062032.

Loh L.K.Y., Kueh H.K., Parikh N.J., Chan H., Ho N.J.H., Chua M.C.H. (2022), *An Ensembling Architecture Incorporating Machine Learning Models and Genetic Algorithm Optimization for Forex Trading*, "FinTech", vol. 1, no. 2, pp. 100–124.

Luo T., Metawa N. (2019), *Intelligent Algorithm of Optimal Investment Model Under Stochastic Interest Rate and Stochastic Volatility*, "Journal of Intelligent & Fuzzy Systems", vol. 37, no. 1, pp. 283–292.

McCarthy J., Minsky M., Rochester N., Shannon C.E. (1955), *A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence*, http://raysolomonoff.com/dartmouth/boxa/dart564props.pdf [accessed: 1.11.2022].

McCorduck P. (2004), *Machines Who Think*, A.K. Peters Ltd., Natick.

Mercadier M., Lardy J.P. (2019), *Credit Spread Approximation and Improvement Using Random Forest Regression*, "European Journal of Operational Research", vol. 277, no. 1, pp. 351–365.

Mihaljević B., Bielza C., Larrañaga P. (2021), *Bayesian Networks for Interpretable Machine Learning and Optimization*, "Neurocomputing", vol. 456, pp. 648–665.

Milojević N., Redzepagic S. (2021), *Prospects of Artificial Intelligence and Machine Learning Application in Banking Risk Management*, "Journal of Central Banking Theory and Practice", vol. 10, no. 3, pp. 41–57.

Ministry of Economy, Trade and Industry in Japan (2022), *Governance Guidelines for Implementation of AI Principles Ver. 1.1" Compiled*, https://www.meti.go.jp/english/press/2022/0128_003.html [accessed: 12.02.2023].

Mitchell T. (1997), *Machine Learning*, McGraw Hill, New York.

Monett D., Lewis C.W.P. (2018), *Getting Clarity by Defining Artificial Intelligence – A Survey*, [in:] V.C. Müller (ed.), *Philosophy and Theory of Artificial Intelligence 2017*, Springer, Berlin, pp. 212–214.

Monfared S.A., Enke D. (2014), *Volatility Forecasting Using a Hybrid Gjr-Garch Neural Network Model*, "Procedia Computer Science", vol. 36, pp. 246–253.

Neal R.M. (2000), *Markov Chain Sampling Methods for Dirichlet Process Mixture Models*, "Journal of Computational and Graphical Statistics", vol. 9, no. 2, pp. 249–265.

Nilsson N.J. (2010), *The Quest for Artificial Intelligence. A History of Ideas and Achievements*, Cambridge University Press, Cambridge.

NIST (2019), *U.S. Leadership in AI: A Plan for Federal Engagement in Developing Technical Standards and Related Tools*, National Institute of Standards and Technology, Tech. Rep, https://www.nist.gov/system/files/documents/2019/08/10/ai_standards_fedengagement_plan_9aug2019.pdf [accessed: 1.11.2022].

OECD (2019), *Scoping the OECD AI principles: Deliberations of the Expert Group on Artificial Intelligence at the OECD (AIGO)*, OECD Publishing, Paris.

Osisanwo F.Y., Akinsola J.E.T., Awodele O., Hinmikaiye J.O., Olakanmi O., Akinjobi J. (2017), *Supervised Machine Learning Algorithms: Classification and Comparison*, "International Journal of Computer Trends and Technology", vol. 48, no. 3, pp. 128–138.

Poole D., Mackworth A., Goebel R. (1998), *Computational Intelligence: A Logical Approach*, Oxford University Press, New York.

Potrawa T., Tetereva A. (2022), *How Much Is the View from The Window Worth? Machine Learning-Driven Hedonic Pricing Model of The Real Estate Market*, "Journal of Business Research", vol. 144, pp. 50–65.

Pun J., Lawryshyn Y. (2012), *Improving Credit Card Fraud Detection Using a Meta-Classification Strategy*, "International Journal of Computer Applications", vol. 56, no. 10, pp. 101–105.

Rad H., Low R.K.Y., Miffre J., Faff R.W. (2021), *The Commodity Risk Premium and Neural Networks*, http://doi.org/10.2139/ssrn.3816170

Rao J.N.K. (2003), *Small Area Estimation*, John Wiley & Sons, New York.

Rapaport W.J. (2020), *What is Artificial Intelligence?*, "Journal of Artificial General Intelligence", vol. 11, no. 2, pp. 52–56.

Ren J., Lee S.D., Chen X., Kao B., Cheng R., Cheung D. (2009), *Naive Bayes Classification of Uncertain Data*, [in:] *ICDM 2009, The Ninth IEEE International Conference on Data Mining*, Miami, pp. 944–949.

Revett K., Gorunescu F., Gorunescu M., Ene M., Magalhaes S., Santos H. (2007), *A Machine Learning Approach to Keystroke Dynamics Based User Authentication*, "International Journal of Electronic Security and Digital Forensics", vol. 1, no. 1, pp. 55–70.

Russell S.J., Norvig P. (2021), *Artificial Intelligence: A Modern Approach*, Pearson, Hoboken.

Sadok H., Sakka F., El Hadi El Maknouzi M. (2022), *Artificial Intelligence and Bank Credit Analysis: A Review*, "Cogent Economics & Finance", vol. 10, no. 1, 2023262.

Sagiroglu S., Sinanc D. (2013), *Big data: A review*, [in:] *2013 International Conference on Collaboration Technologies and Systems (CTS)*, IEEE, San Diego, pp. 42–47, https://doi.org/10.1109/CTS.2013.6567202

Samuel A.L. (1959), *Some Studies in Machine Learning Using the Game of Checkers*, "IBM Journal of Research and Development", vol. 3, no. 3, pp. 210–229.

Sarker I.H., Kayes A.S.M., Badsha S., Alqahtani H., Watters P.N.A. (2020), *Cybersecurity Data Science: an Overview from Machine Learning Perspective*, "Journal of Big Data", vol. 7, pp. 1–29.

Schapire R.E. (2003), *The Boosting Approach to Machine Learning: an Overview*, [in:] D.D. Denison, M.H. Hansen, C.C. Holmes, B. Mallick, B. Yu (eds.), *Nonlinear estimation and classification*, Springer, New York, pp. 149–171.

Schultz G.M., Fabozzi F.J. (2021), *Rise of the Machines: Application of Machine Learning to Mortgage Prepayment Modeling*, "The Journal of Fixed Income", vol. 31, no. 3, pp. 6–19.

Shalev-Shwartz S., Ben-David S. (2014), *Understanding Machine Learning from Theory to Algorithms*, Cambridge University Press, Cambridge.

Smola A.J., Schölkopf B. (2004), *A tutorial on support vector regression*, "Statistics and Computing", vol. 14, no. 3, pp. 199–222, https://doi.org/10.1023/B:STCO.0000035301.49549.88

Solomonoff R.J. (1985), *The Time Scale of Artificial Intelligence; Reflections on Social Effects*, "Human Systems Management", vol. 5, pp. 149–153.

Tavana M., Abtahi A.R., Di Caprio D., Poortarigh M. (2018), *An Artificial Neural Network and Bayesian Network Model for Liquidity Risk Assessment in Banking*, "Neurocomputing", vol. 275, pp. 2525–2554.

The American National Standard Dictionary for Information Technology ANSDIT (2022), *ANSI INCITS 172–220 (R2007) Information Technology (Revision and Redesignation of ANSI X3.172–1996)*, https://webstore.ansi.org/Standards/INCITS/ansiincits1722002r2007 [accessed: 1.11.2022].

Thrun S., Langford J. (1998), *Monte Carlo Hidden Markov Models*, Carnegie-Mellon University, School of Computer Science, Pittsburgh.

Tibshirani R. (1996), *Regression Shrinkage and Selection via the LASSO*, "Journal of the Royal Statistical Society, Series B (Methodological)", vol. 58, no. 1, pp. 267–288.

Turing A.M. (1950), *Computing Machinery and Intelligence*, "Mind a Quarterly Review of Psychology and Philosophy", vol. 236, pp. 433–460.

Vapnik V., Levin E., Cun Y.L. (1994), *Measuring The VC-Dimension of a Learning Machine*, "Neural Computation", vol. 6, no. 5, pp. 851–876.

Verma P. (2020), *Churn Prediction for Savings Bank Customers: A machine Learning Approach*, "Journal of Statistics Applications & Probability", vol. 9, no. 3, pp. 535–547.

Wang P. (2008), *What Do You Mean by "AI"*, "Frontiers in Artificial Intelligence and Applications", vol. 171, no. 1, pp. 362–373.

Wang P. (2019), *On Defining Artificial Intelligence*, "Journal of Artificial General Intelligence", vol. 10, no. 2, pp. 1–37.

Wang Y., Zhang Y., Lu Y., Yu X . (2020), *A Comparative Assessment of Credit Risk Model Based on Machine Learning – a Case Study of Bank Loan Data*, "Procedia Computer Science", vol. 174, pp. 141–149.

Yampolskiy R.V. (2020), *On Defining Differences between Intelligence and Artificial Intelligence*, "Journal of Artificial General Intelligence", vol. 11, no. 2, pp. 68–70.

Yampolskiy R.V., Fox J. (2012), *Artificial General Intelligence and the Human Mental Model*, [in:] A. Eden, J. Søraker, J.H. Moor, E. Steinhart (eds.), *Singularity Hypotheses: A Scientific and Philosophical Assessment*, Springer, Berlin, pp. 129–145.

Yang B.H., Tkachenko M. (2012), *Modeling Exposure at Default and Loss Given Default: Empirical Approaches and Technical Implementation*, "The Journal of Credit Risk", vol. 8, no. 2, 81.

Ye T., Zhang L. (2019), *Derivatives Pricing Via Machine Learning*, Boston University Questrom School of Business Research Paper, no. 3352688.

Zahi S., Achchab B. (2020), *Modeling Car Loan Prepayment Using Supervised Machine Learning*, "Procedia Computer Science", vol. 170, pp. 1128–1133.

Zhang H.G., Su C.W., Song Y., Qiu S., Xiao R., Su F. (2017), *Calculating Value-At-Risk for High-Dimensional Time Series Using a Nonlinear Random Mapping Model*, "Economic Modelling", vol. 67, pp. 355–367.

Zhang J., Teng F., Lin S. (2019), *Agency MBS Prepayment Model Using Neural Networks*, "The Journal of Structured Finance", vol. 24, no. 4, pp. 17–33.

Zou H., Hastie T. (2005), *Regularization and variable selection via the elastic net*, "Journal of The Royal Statistical Society: Series B (Statistical Methodology)", vol. 67, no. 2, pp. 301–320.

Wykorzystanie modeli sztucznej inteligencji w bankach komercyjnych – szanse i zagrożenia

Streszczenie: Jednym z głównych sektorów, które w dużym stopniu wykorzystują rozwój zaawansowanych metod obliczeniowych, jest sektor bankowy. Cele badań przedstawionych w artykule to: 1) porównanie naukowego i regulacyjnego podejścia do definiowania sztucznej inteligencji (AI) i uczenia maszynowego (ML); 2) zaproponowanie definicji AI i ML na potrzeby regulacyjne, które pozwolą jednoznacznie stwierdzić, czy dana metoda jest AI/ML, czy nie; 3) porównanie złożonych metod ilościowych stosowanych w bankowości pod względem złożoności i interpretowalności w celu jasnej klasyfikacji metod dla zainteresowanych stron (praktyków i kadry zarządzającej); 4) zaproponowanie możliwego podejścia do dalszego rozwoju metod ilościowych w obszarach o wymaganej ścisłej interpretowalności. Przegląd literatury koncentruje się na definicjach AI/ML stosowanych przez naukowców i regulatorów oraz propozycjach zastosowania złożonych rozwiązań ilościowych w różnych domenach bankowości. Badania skupione są na proponowaniu praktycznych definicji AI i ML na podstawie aktualnego stanu wiedzy i wymogów przejrzystości w branży bankowej (bardzo ograniczony apetyt na ryzyko, dotyczący niezgodności z regulacjami) oraz na porównaniu metod ilościowych stosowanych w różnych domenach bankowości wraz z ich oceną. Autor proponuje ogólne i inkluzywne definicje AI i ML na potrzeby regulacyjne, które pozwalają jednoznacznie sklasyfikować konkretne metody. W przypadku zaostrzonych wymagań dotyczących interpretowalności stosowanych metod proponuje stopniowe i kontrolowane zwiększanie złożoności istniejących rozwiązań. Z tego powodu przedstawia ocenę metod ilościowych pod względem interpretowalności i złożoności. Autor uważa również, że definicje AI/ML w dalszych regulacjach powinny umożliwiać jednoznaczne zaklasyfikowanie konkretnych podejść jako AI/ML. Badania skierowane są do twórców regulacji, praktyków i kadry zarządzającej związanej z sektorem bankowym.

Słowa kluczowe: sztuczna inteligencja, uczenie maszynowe, etyka, bankowość, wyjaśnialność AI, regulacje

JEL: C10, C40, C50, G10, G21

Received: 2022-12-08; revised: 2023-04-04. Accepted: 2023-06-14