Kodelja, Zdenko

Is machine learning real learning?

CEPS Journal 9 (2019) 3, S. 11-23

Empfohlene Zitierung/ Suggested Citation:
Kodelja, Zdenko: Is machine learning real learning? - In: CEPS Journal 9 (2019) 3, S. 11-23 - URN: urn:nbn:de:0111-pedocs-181330 - DOI: 10.26529/cepsj.709
http://nbn-resolving.org/urn:nbn:de:0111-pedocs-181330
http://dx.doi.org/10.26529/cepsj.709

doi: 10.26529/cepsj.709

Is Machine Learning Real Learning?

Zdenko Kodelja[1]

The question of whether machine learning is real learning is ambiguous, because the term "real learning" can be understood in two different ways. Firstly, it can be understood as learning that actually exists and is, as such, opposed to something that only appears to be learning, or is misleadingly called learning despite being something else, something that is different from learning. Secondly, it can be understood as the highest form of human learning, which presupposes that an agent understands what is learned and acquires new knowledge as a justified true belief. As a result, there are also two opposite answers to the question of whether machine learning is real learning. Some experts in the field of machine learning, which is a subset of artificial intelligence, claim that machine learning is in fact learning and not something else, while some others – including philosophers – reject the claim that machine learning is real learning. For them, real learning means the highest form of human learning. The main purpose of this paper is to present and discuss, very briefly and in a simplifying manner, certain interpretations of human and machine learning, on the one hand, and the problem of real learning, on the other, in order to make it clearer that the answer to the question of whether machine learning is real learning depends on the definition of learning.

Keywords: learning, machine learning, artificial intelligence, philosophy, education

1 Educational Research Institute, Ljubljana, Slovenia; zdenko.kodelja@guest.arnes.si.

Ali je strojno učenje pravo učenje?

Zdenko Kodelja

Vprašanje, ali je strojno učenje pravo učenje, je dvoumno, ker pojem »pravo učenje« lahko razumemo na dva različna načina. Prvič, lahko je razumljeno kot učenje, ki dejansko obstaja in je kot tako nasprotno nečemu, kar se le zdi kot učenje, ali pa je napačno poimenovano kot učenje, čeprav je nekaj drugega; nekaj, kar je drugačno od učenja. Drugič, lahko ga razumemo kot najvišjo obliko človeškega učenja, ki predpostavlja, da agent razume, kar se je naučil, in pridobi novo znanje kot upravičeno resnično verjetje. Posledično obstajata tudi dva nasprotujoča si odgovora na vprašanje, ali je strojno učenje pravo učenje. Nekateri strokovnjaki s področja strojnega učenja, ki je podpodročje umetne inteligence, trdijo, da je strojno učenje dejansko učenje in ne nekaj drugega, medtem ko nekateri drugi – vključno filozofi – zavračajo trditev, da je strojno učenje pravo učenje. Za njih je pravo učenje najvišja oblika človeškega učenja. Glavni namen prispevka je na eni strani na kratko in na preprost način predstaviti in razpravljati o določenih interpretacijah človeškega in strojnega učenja ter na drugi o problemu pravega učenja, zato da bi bilo jasneje vidno, da je odgovor na vprašanje, ali je strojno učenje pravo učenje, odvisen od opredelitve učenja.

Ključne besede: učenje, strojno učenje, umetna inteligenca, filozofija, edukacija

Introduction

There are two opposite answers to the question of whether or not machine learning is real learning. Some experts in the field of machine learning, which is a subset of artificial intelligence, claim that machine learning is in fact learning, while some others reject this claim, arguing that machine learning "is not real learning" (Bringsjord et al., 2018, p. 136). It seems, therefore, that at least one of these claims is wrong. Such a conclusion – which presumes that there is only one truth and that any judgement contrary to it is untrue, since, according to the principle of non-contradiction, two contradictory statements cannot be true at the same time – would be correct only if the term "learning" had the same meaning in both cases. Since it does not have the same meaning, however, the essential question is whether the difference between learning and real learning is the difference between learning and something that is called "learning" although it is in fact something else, or whether it is the difference between two different sorts of learning. It seems that the difference in question is not the difference between learning and non-learning, but rather the difference between machine learning and a specific form of human learning.[2] From the perspective of human learning understood in such a way, machine learning might appear to be something that is not, or is not yet, real learning. However, from the same perspective, animal learning could also be seen as something that is not real learning.[3] Despite this, almost no one argues that animal learning is not learning. In any case, the answer to the aforementioned question of whether machine learning is real learning depends on how we understand learning.

In this paper, the problem of learning will be discussed in the context of philosophy of education. If we agree with the interpretation that philosophy of education is not so much a specific corpus of knowledge, but rather a questioning – a questioning in the sense that it brings into question over and over again all that we know, or believe that we know, about education (Reboul, 1995, p. 3) – then it follows that we have to also ask questions about the concept of "learning" over and over again. Such perpetual questioning is necessary not only because we can otherwise quickly become prisoners of various dogmatisms, but also because this concept, or at least this term, is nowadays used in an entirely new context, that is, in the context of artificial

2 Real learning in the formal sciences is described as "a phenomenon that has been firmly in place in homes and schools since at least Euclid" (Bringsjord et al., p. 136).

3 Conditioning, both classical and instrumental, for instance, is commonly understood as a form of learning in psychology and in education. From the perspective of human learning, however, we can argue that "being conditioned is not a form of learning" (Hamlyn, 1987, p. 179), on condition that we presuppose "that learning must at least involve the acquisition of knowledge through experience and that changes of behaviour due to learning must be the result of knowledge" (ibid., p. 180). Even in this case, however, it seems that the claim that conditioning is not a form of learning is related only to human learning and not to animal learning, which by definition does not include the acquisition of knowledge.

intelligence, or more precisely, the theories and practice of so-called machine learning. Therefore, my intention is not to try to prove or disprove the existence of machine learning as a form of real learning. Even if I would like to do one or the other, I cannot do either, simply because I am not an expert in this field of knowledge. Instead, I will very briefly present and discuss certain interpretations of human and machine learning, on the one hand, and the problem of real learning, on the other, in order, I hope, to make it a little clearer that the answer to the aforementioned question of whether machine learning is real learning depends on the definition of learning.[4]

Human learning

Until now, learning has been discussed in philosophy of education mostly, if not exclusively, as human learning.[5] Although there are a variety of theories and concepts of such learning, for the purpose of this paper it seems to be sufficient to indicate some essential features of human learning as it is predominantly understood in philosophy of education. According to one of its basic philosophical interpretations, learning is the acquisition of knowledge. This interpretation is based on two classical theories of knowledge and mind: empiricism and rationalism.[6] In spite of their differences, rationalists mostly argue that it is possible to

4 Here, the term "definition" means a real definition of learning and not a nominal one, that is, it refers to the real essence of learning and to the meaning of the word "learning".

5 Moreover, at least in the context of Anglo-Saxon philosophy of education, and at least during the last decades of the twentieth century, philosophers of education wrote rather little – in comparison with psychologists – about learning. However, they extensively studied the concept of teaching, which "cannot be understood without some reference to 'learning'" (Peters, 1987, p. 6). In English, this distinction between learning and teaching is obvious both on a terminological and conceptual level. In some other languages, however, this distinction is not so clear, simply because the same word can mean learning and teaching at the same time. "In French, the word apprendre means both 's'instruire' (*learning*, *lernen*) and 'instruire' (teaching, lehren); we learn the algebra, we teach algebra to someone. The ambiguity is itself significant; indeed, there is perhaps no absolute opposition between the one who teaches and the one who is learning; sometimes it is the same person" (Reboul, 1980, p. 9). The same can be said for the Slovenian word "*učiti*", which can also mean both teaching (*učiti*) and learning (*učiti se*).

6 Rationalism and empiricism – which have different forms – are commonly described as opposite philosophical doctrines developed by two groups of philosophers whose most important representatives are: Descartes, Spinoza and Leibniz, on the one hand, and Locke, Berkeley and Hume, on the other. However, these doctrines are not always in conflict, and the philosophies of each and all of these philosophers could not simply be reduced to general descriptions of rationalism and empiricism. The main differences in understanding the mind can be briefly expressed as follows: "For empiricism the mind is, as Locke put it, like a great mirror which passively receives reflections from without, while for rationalism the mind is more active, involved in its own operations" (Hamlyn, 1987, p. 178). In addition to rationalism and empiricism, pragmatism – whose representatives are Peirce, James and Dewey – is sometimes discussed in philosophy of education as a third great philosophical approach to knowledge and mind. According to pragmatism, "the mind is conceived neither as a deep well of necessary truths nor as a blank slate upon which experience writes", but rather, "as a capacity for active generation of ideas whose function it is to resolve the problems posed to an organism by its environment" (Scheffler, 1965, p. 5). Looking from the pragmatic point of view, learning is an active process: "to learn something significant about the world, we must do more than operate logically upon basic truths that appear to us self-evident, and we must go beyond reasonable generalisation of observed phenomenal patterns in our past experience" (ibid., p. 4).

acquire a knowledge of the external world by reason alone. They believe that such knowledge can be acquired *a priori*, that is, independently of sense experience, by intuition and deduction. Empiricists think just the opposite, claiming that all human knowledge – except that of logical relations between our own concepts – is *a posteriori*, because it derives from sense experience.

Empiricists' theories of knowledge and mind have had a significant impact on some very influential theories of learning developed by empirically oriented psychology. According to these theories, "learning embraces any modification of behaviour in an organism as the result of experience, or even as the result of stimulation from the environment" (Hamlyn, 1987, p. 179). This is a very wide definition of learning and it better characterises animal learning than human learning. Essential for human learning is the acquisition of knowledge and understanding,[7] while "many abilities acquired by animals through learning involve little or no understanding of what is involved" (Hamlyn, 1987, p. 179). This basic interpretation of human learning as the acquisition of knowledge has been dominant in philosophy of education. In this context, knowledge means different things, and it is commonly presented and analysed in its two main forms: firstly, as propositional or factual knowledge,[8] and secondly, as practical knowledge.[9] The first form of knowledge is also known as "knowledge-that",[10] and the second as "knowledge-how".

Knowledge-that is knowledge of a fact or truth. It differs from a false belief and a lucky guess. As such, it is usually seen as rationally justified true beliefs.[11] In opposition to knowledge-that, knowledge-how is defined as knowledge of how to do something (Ryle, 1945). However, the open question remains as to whether or not knowledge-how is independent of knowledge-that. Some think that it is independent, others that it is not, and still others that it is independent to a considerable degree. Despite the fact that there is no consensus on the right answer to this

7 In human learning, the acquisition of knowledge is connected with understanding: even at the lowest level of learning, that is, "in rote-learning, it is essential to understand what is going on, and in higher forms of learning understanding is much more important still" (Hamlyn, 1987b, pp. 198–199).

8 They are called so because the object of knowledge is a proposition or a fact.

9 Practical knowledge, as opposed to theoretical knowledge, is not primarily knowledge about something, but rather knowledge about how to do something in practice: to play the piano, for instance. Knowing how to play the piano in theory is not the same as knowing how to play the piano in practice.

10 Some other important sorts of knowledge, such as "knowledge-who", "knowledge-why", "knowledge-where", and so on, are usually understood as particular kinds of "knowledge-that".

11 Knowledge is here explained it this way: *S* knows that *p* if (1) *p* is true; (2) *S* believes that *p*; (3) *S* is justified in believing that *p*. According to this interpretation of knowledge, justified true belief is necessary and sufficient for knowledge. However, some philosophers argue that it is only necessary but not sufficient, because a true belief can be inferred from a justified false belief. Since it is merely lucky that such beliefs are true, they cannot be knowledge (Gettier, 1963, pp. 121–123). In order to avoid this problem, some philosophers (who think that the definition of knowledge as a justified true belief is, in principle, correct) have tried to determine an appropriate fourth condition that would prevent deriving a true belief from a justified false belief.

question, the prevalent interpretation is that knowing-how is in some way distinct from knowing-that, and that, even if knowing how to do something requires relevant knowledge-that, knowing-how cannot be reduced to knowledge of facts or truths or theoretical knowledge about how to do something. This distinction between knowledge-that and knowledge-how is very similar to the distinction between declarative and procedural knowledge as explained in discussions of artificial intelligence and cognitive science. In this context, moreover, declarative knowledge is interpreted as explicit knowledge of particular things (facts, information, concepts, events, processes, their attributes, and their relations to each other), which are stored in the memory, whereas procedural knowledge is the knowledge of how to perform a specific skill or task, or how to operate. As such, it is related to the methods, procedures or operation of computers.

When knowledge is interpreted in this manner, it is obvious that learning as the acquisition of knowledge also has two main forms: learning-that and learning-how. Learning that *p* involves coming to know that *p*, while learning-how to do *p* involves coming to know how to do *p*. In the first case, *p* means a truth or fact; in the second case, it means skills or attainments. However, the supposition that learning-that implies knowing-that is not always true. Therefore, we cannot, as Scheffler emphasises, "generalize … that whenever a person X has learned that Q, he has come to know that Q" (Scheffler, 1965, p. 7). If Q is not true, then someone who has learned that Q, despite believing that Q is true, cannot come to know that Q; he only comes to believe that Q. Consequently, one who has learned that Q comes to know that Q if Q is true and he is able to provide adequate evidence for his belief that Q is true.[12] In the case of learning-how, the consequence of such learning is that one comes to know how to do something, but there is no agreement on what exactly this knowledge is: for some authors, it is some sort of disposition (to perform a specific skill or task), while for others, it is some sort of ability (to perform a specific skill or task). Both learning-that and learning-how are relevant to understanding machine learning, as well.

Machine learning

There are different definitions of machine learning.[13] One of those often mentioned is the definition attributed to A. Samuel, one of the pioneers

12 According to Scheffler, the educational term "learning" is related to the cognitive terms "knowing" and "believing" as follows: "learning that Q involves coming to believe that Q. Under certain further conditions (truth of 'Q' and, for the strong sense of knowing, proper backing of 'Q'), it also involves coming to know that Q" (Scheffler, 1965, p. 13).

13 The term "machine" in the expression "machine learning" refers to an algorithm and not to a piece of equipment, that is, to a computer understood as a piece of hardware.

of machine learning, who in 1959 supposedly defined machine learning as a "field of study that gives computers the ability to learn without being explicitly programmed" (Samuel, 1959, pp. 210–229).[14] In this definition, it is clearly stated that computer programs or special algorithms have the capacity to learn. As such, these programs essentially differ from traditional computer programs, which depend on rules in the form of a code written by human experts in order to instruct computers how to perform a specific task. In other words, traditional programs are not programmed *to learn* to perform a specific task, but only to perform such tasks. In addition, traditional programs are not able to perform tasks better than the human experts who programmed them, while machine-learning algorithms in some domains already can. In fact, a checkers program from 1955 included machine learning and was the first program that was successfully programmed to learn to play a better game of checkers than could "be played by the person who wrote that program" (ibid., p. 210). Since then, computers have defeated the best human players in games such as chess, Go, scrabble, Jeopardy!, and most video and other board games (Bhatnagar et al., 2018, p. 118; Bostrom, 2017, pp. 15–16;). Similar results of rapid progress in machine learning are seen in some other domains, as well. For instance, "deep learning methods – essentially many-layered neural networks – have, thanks to the combination of faster computers, larger data sets, and algorithmic refinements, begun to approach (and in some cases exceed) human performance on many perceptual tasks, including handwriting recognition, image recognition and image captioning, speech recognition, and facial recognition" (Bostrom, 2017, p. 321). Despite these and some other fascinating achievements in particularly narrow domains, however, the fact is that in many domains "we do not yet know how to make computers learn nearly as well as people learn" (Mitchell, 1997, p. 1). The question is whether it is possible at all that computers will learn as well as, or even better than, humans learn. There are two opposite answers to this question. On the one hand, there are those who believe that, in the more or less distant future, there will be algorithms capable of doing all of the tasks that the human brain does, while, on the other hand, there are those who believe just the opposite. In both cases, the answers are in the domain of belief, and we therefore do not know whether or not it will be possible that computers will learn as well as, or even better than, human beings learn. We can therefore only believe or not believe that such computers will exist in the future. We nonetheless know at least that one of these alternative predictions about the existence of such computers in the future is necessarily true, because the disjunction "p or

14 This quotation can be found in many sources, but there is no such definition of machine learning in the cited article.

non-p" is logically necessary. However, one might object that this disjunction is itself not necessary if the predicted statements refer to such computers as future contingent objects (which, as such, must neither be impossible nor inevitable), or if they are contingent statements about future computers and, as such, are neither true nor false, but undetermined. In any case, these problems are not essential for the purpose of this paper, as it simply deals with the question of whether already existing machine learning really learns. Moreover, discussing these problems might even be misleading, because arguing about the possible existence of future machine-learning machines that would learn as well as, or even better than, human beings learn presupposes that computers learn, that machine learning is real learning, which is precisely the presupposition in question. Therefore, such a presupposition is a logical fallacy, a sort of *petitio principii*: we take for granted exactly what is in dispute, or in other words, what is to be proven has already been supposed. We would have made the same mistake if we had assumed that machine learning is already by definition learning, since the word "learn" appears in "machine learning", as well.[15] For the problem is conceptual, not only terminological. If machine learning is real learning, the use of the term "learning" in the expression "machine learning" is correct; if it is not real learning, then the use of the term "learning" is wrong and misleading. To avoid the confusion, some authors write this term in quotation marks when it refers to machine learning. This means that, for them, real learning is human learning and perhaps also animal learning, but not machine learning. The view of some proponents of machine learning is just the opposite, including Samuel's seminal interpretation of machine learning as "programming of a digital computer to behave in a way which, if done by humans or animals, would be described as involving the process of learning" (Samuel, 1959, p. 210).[16] If human and animal learning are the criteria for real learning, then machine learning is, according to this interpretation, real learning. Machine learning is real learning

15 A possible answer to the paper's title "Do Machine-Learning Machines Learn?" is that a machine that machine learns *by definition* learns, since "learn" appears in "machine learn". This, however, assumes at the outset that what is called "machine learning" today *is* real learning, which that is precisely what is in question; hence the *petitio* (Bringsjord et al., 2018, p. 136).

16 In this case, the best features of two sorts of learning were combined in the computer program: rote learning and learning-by-generalisation. Rote learning is a very elementary kind of learning "in which the program only saved all of the board positions encountered during play, together with their complete scores" (Samuel, 1959, p. 214). However, "the program with rote learning soon learned to imitate master play during the opening moves. It was always quite poor during the middle game, but it easily learned how to avoid most of the obvious traps during end-game play and could usually drive on toward a win when left with a piece advantage". In contrast to the rote-learning program, "the program with the generalization procedure has never learned to play in a conventional manner and its openings are apt to be weak. On the other hand, it soon learned to play a good middle game, and with a piece advantage it usually polishes off its opponent in short order" (ibid., p. 221).

for some other experts on this issue, too, including T. Mitchell, who argues that his "definition of learning is broad enough to include most tasks that we would conventionally call 'learning' tasks, as we use the word in everyday language. It is also broad enough to encompass computer programs that improve their performance at some task through experience" (Mitchell, 1997, p. 2). According to his definition, "A computer program is said to learn from experience E with respect to some class of tasks T and performance measure P if its performance at tasks in T, as measured by P, improves with experience E" (ibid.). In the case of the aforementioned checkers program, this definition can be stated as follows: "a computer program that learns to play checkers might improve its performance as measured by its ability to win at the class of tasks involving playing checkers games, through experience obtained by playing games against itself" (ibid.).

Since then, more and more powerful machine learning programs have been created.[17] Not only are they able to play checkers and some other games at a superhuman level, but they also "cover a diverse set of learning tasks, from learning to classify emails as spam, to learning to recognize faces in images, to learning to control robots to achieve targeted goals" (Mitchell, 2017, p. 1); they can recognise human voices, translate text, and so on.

Nevertheless, when searching for the answer to the question of whether machine learning is real learning it is enough if we take into consideration the aforementioned basic machine learning program based on the rote-learning procedure (Samuel, 1959). For if machine learning is in this case real learning (equally as real as real human rote learning), then learning in other more developed machine learning programs, based on higher forms of human learning, is also real learning.

Real learning

The term "real" usually means that something actually exists and is, therefore, not apparent, imaginary, fictitious or merely so-called. For this reason, it seems that the expression "real learning", too, means learning that actually exists and is, as such, opposed to something that only appears to be learning or is misleadingly called learning despite being something else. The claim that

17 They have typically been based on three types of learning algorithms: *supervised learning* (the learning algorithms are trained using labelled examples that contain both the inputs and the desired output); *unsupervised learning* (the learning algorithms must explore input data that has not been labelled, classified or categorised, and find some patterns in the data); *reinforcement learning* (the learning algorithms must, through trial and error, discover which actions in an environment produce the greatest cumulative rewards).

contemporary machine learning is not real learning can also be understood in this way (Bringsjord et al., 2018, p. 136). However, this understanding of real learning seems to be wrong, because Bringsjord, Govindarajulu, Banerjee and Hummel do not argue that machine learning is not learning, but rather that it is not real learning as they conceive it. They define real learning in the context of mathematics and formal sciences by saying that an agent ***a*** has really learned a unary function **f**: N→N "only if

(c1) **a** understands the formal definition Df of **f**,

(c2) can produce both **f** (x) for all x ∈ N, and

(c3) proof of the correctness of what is supplied in (c2)" (ibid.).

Since no existing forms of machine learning satisfy these three conditions, the authors of this very demanding interpretation of real learning conclude that machine learning is not real learning.[18] Moreover, according to Bringsjord and Govindarajulu (2019, p. 2), "most of what is called 'learning' in artificial intelligence today", that is, in machine learning, does not satisfy the necessary conditions for being real learning even when real or genuine learning is defined as acquisition of new knowledge that consists of justified true belief. For real or "genuine learning of φ by an agent, …, must result in the acquisition of knowledge by the agent, and knowledge in turn consists in the holding of three conditions, to wit: (1) the agent must believe that φ holds; (2) must have cogent, expressible, surveyable justification for this belief; and (3) φ must in fact hold" (ibid).

Real learning, understood in such a way, presupposes that an agent (human or artificial) can really learn only if he, she or it has certain abilities – the abilities of reasoning and communication, for instance – on a human level or above. It is not surprising, therefore, that the authors of the discussed definition of real learning find real learning neither in the different forms of machine learning (symbol-based, connectionist, genetic and probabilistic) nor in some forms of learning in psychology and the cognitive sciences (associative learning: classical and instrumental conditioning, representational, observational, statistical, neurocentric and instructional learning).

From the assertion that machine learning is not real learning, however, the conclusion that different forms of machine learning are not learning does

18 In their opinion, this does not mean that current machine learning is not real learning only in the formal sciences; it is absent in creative writing, as well. "In order to learn to be a creative writer one must generate stories, over and over, and learn from the reaction and analysis thereof, and then generate again, and iterate the process. Such learning, which is real learning in creative writing, is not only not happening in machine learning today; it is also hard to imagine it happening in even machine learning of tomorrow" (Bringsjord et al., 2018, pp. 145–146.).

not necessarily follow. They are learning, but on a lower level than real learning. We have already seen that the first elementary form of machine learning was on the level of rote learning, while real learning – understood in the way that the aforementioned authors defined it – is on a much higher level of human learning. Nevertheless, the claim that machine learning is not real learning can also be seen from a different point of view, that is, as similar to the more known and discussed claims that artificial intelligence is not real intelligence or that intelligent machines are not intelligent. In both cases, the truthfulness of each of these claims depends on the criteria for learning and intelligence that have been used.[19]

Conclusion

In discussing the interpretation of learning in the previous section, we saw reasons to doubt whether machine learning is real learning. If we accept the interpretation of real learning that Bringsjord, Govindarajulu, Banerjee and Hummel defend, then we have to agree with their conclusion, as well; namely, that machine learning is not real learning. However, this does not mean that machine learning is not learning. Quite the opposite: it is learning, but until now it has only been on the level of animal and lower forms of human learning. Therefore, machine learning, too, can be seen as real learning in the sense that it is not something different from learning. Of course, the expression "real learning" here has another meaning than in the argumentation of the previously discussed authors who claim that machine learning is not real learning. For them, real learning is not just any kind of learning, but only the highest forms of human learning. Such an interpretation of real learning is not new. We have already seen that some philosophers of education have claimed that lower forms of learning, such as conditioning, are not learning at all. However, such understandings of learning are exceptions to the general rule, that is, they are in opposition to generally accepted interpretations of learning in scientific and philosophical theories. Therefore, arguing that machine learning is real learning is equally justified as claiming that animal and lower forms of human learning are real learning and not something different from learning. It seems that even those who claim that such forms of learning are not real learning do not think that they are not forms of learning at all, but rather that they are not real learning understood as the highest form of human learning

19 The polemics about machine intelligence reveal something interesting; namely, that the criteria for intelligence were changed when artificial intelligence outperformed human intelligence: "we stopped seeing 'winning the chess game' as a sign for intelligence in 1997, when 'Deep Blue' beat chess champion Kasparov" (Danziger, 2018, p. 170).

(which presupposes that an agent understands what is learned and acquires new knowledge as a justified true belief).

Nevertheless, even if machine learning is not real learning, there are three fundamental reasons for studying machine learning. The first one "is to understand the process" of learning itself,[20] the second is "to provide computers with the ability to learn",[21] and the third is to better understand machine learning and artificial intelligence as a constitutive part of the fourth scientific revolution that has been reshaping human reality, changing our self-understanding, our relationships, our society, our education and our employment (Floridi, 2016), and has "had a radical and widespread impact on our moral lives and on contemporary ethical debates" (Floridi, 2015).[22] At least the first and the third of these reasons are important and challenging for philosophy of education, as well.

References

Bhatnagar, S., et al. (2018). Mapping intelligence: Requirements and possibilities. In V. C. Müller (Ed.), *Philosophy and theory of artificial intelligence 2017*. Berlin: Springer.

Bostrom, N. (2017). *Superintelligence*. Oxford, UK: Oxford University Press.

Bringsjord, S., & Govindarajulu, N. S. (2019). Learning *Ex Nihilo*. arXiv:1903.03515v2 [cs.AI].

Bringsjord, S., et al. (2018). Do machine-learning machines learn? In V. C. Müller (Ed.), *Philosophy and theory of artificial intelligence 2017*. Berlin: Springer.

Danziger, S. (2018). Where intelligence lies: Externalist and sociolinguistic perspectives. In V. C. Müller (Ed.), *Philosophy and theory of artificial intelligence 2017*. Berlin: Springer.

Floridi, L. (2015). *The ethics of information*. Oxford, UK: Oxford University Press.

Floridi, L. (2016). *The fourth revolution*. Oxford, UK: Oxford University Press.

Gettier, E. (1963). Is justified true belief knowledge? *Analysis, 23,* 1963/6.

Hamlyn, D. W. (1987). Human learning. In R. S. Peters (Ed.), *Philosophy of education*. Oxford, UK: Oxford University Press.

Hamlyn, D. W. (1987b). Logical and psychological aspects of learning. In R. S. Peters (Ed.), *Philosophy of education*. Oxford, UK: Oxford University Press.

Mitchell, T. (1997). *Machine learning*. New York, NY: McGraw Hill.

20 "By developing computer models of learning, psychologists have attempted to gain an understanding of the way humans learn" (Nath, 2009, p. 34).

21 "Learning research has potential for extending the range of problems to which computers can be applied. In particular, the work in machine learning is important for expert systems development, problem solving, computer vision, speech understanding, conceptual analysis of databases, and intelligent authoring systems" (ibid.).

22 These debates are on a variety of new topics, such as: artificial evil; morality of artificial agents; artificial moral agents with or without mental states; virtual artificial agents; accountability of artificial agents; robotic consciousness and self-consciousness; informational privacy, personal identity and biometrics; hate speech and online harassment; cyber terror and war; virtual community, and so on.

Mitchell, T. (2017). *Machine learning* (Draft of the chapter 14). Retrieved from http://www.cs.cmu.edu/~tom/mlbook/keyIdeas.pdf

Nath, R. (2009). *Philosophy of artificial intelligence*. Boca Raton, FL: Universal Publishers.

Peters, R. S. (1987). Introduction. In R. S. Peters (Ed.), *Philosophy of education*. Oxford, UK: Oxford University Press.

Reboul, O. (1980). *Qu'est-ce qu'apprendre? Pour une philosophie de l'enseignement* [What is learning? For a philosophy of teaching?]. Paris: PUF.

Reboul, O. (1995). *La philosophie de l'éducation* [Philosophy of education]. Paris: PUF.

Ryle, G. (1949). *The concept of mind*. Chicago, IL: University of Chicago Press.

Samuel, A. (1959). Some studies in machine learning using the game of checkers. *IBM Journal of Research and Development, 44*(1), 211–229.

Scheffler, I. (1965). *Conditions of knowledge*. Glenview, IL: Scott, Foresman and Company.

Biographical note

Zdenko Kodelja is a PhD. both in Philosophy and in Education, and is a full-time researcher and a Head of the Centre for Philosophy of Education at the Educational Research Institute, Ljubljana, Slovenia.