International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING

ISSN:2147-6799 www.ijisae.org Original Research Paper

Research on the Philosophy of Science Oriented to Deep Learning under the Ethical Dilemma

Jie Zhou[1], Ratna Roshida Ab.Razak[2]

Submitted: 11/08/2022 **Accepted:** 21/11/2022

Abstract: The rapid growth in technology has resulted in the technical up-gradation of works in every field. In this study, we are going to research the philosophy of science-oriented to deep learning under the ethical dilemma The philosophy of science is concerned with the value and use of scientific knowledge and its underlying assumptions, techniques, and implications. An ethical dilemma arises when a person is forced to choose between two courses of action, neither of which is morally permissible. Random Forest Algorithm is used in this research to perform regression and classification tasks. It is found that the Random Forest Algorithm outperforms other algorithms in classification problems.

Keywords: *Deep learning, Ethical dilemma, Philosophy of science, Random forest algorithm*

1. Introduction

In terms of the number of scholars and academic societies dedicated to studying and discussing the philosophy of science, the field is growing. At first glance, these tendencies might be explained by the rapid development of science at the same time, the massive investments made in scientific research, and the lasting cultural impact of scientific paradigms in the late twentieth and early twenty-first century. [1] However, much study in the philosophy of science is conducted in what amounts to complete isolation from actual scientific practise. The Society for Philosophy of Science in Practice (SPSP) was founded with the goal of advancing the philosophical investigation of “science in practise,” which its founders described as “scientific praxis” and “the functioning of science in practical spheres of life.” With a few exceptions, such as recent writing on models, experimentation, and measurement that has engaged in thorough examination of scientific procedures in pursuit of philosophical questions, concern about practise has tended to fall outside the mainstream of Anglophone analytic philosophy of science. [2] To remedy this, SPSP was founded to promote in-depth investigations of scientific practise that also consider philosophical and ethical concerns.

Rather than focusing on how scientific concepts apply to the real world, as many conventional approaches to the philosophy of science do, [3] many contemporary philosophers of science are interested in scientific practise. However, the conventional wisdom in the field of social studies of technology and science has been to focus on scientific investigation as a human creation, often ignoring its impact on the global economy. Investigating not only the theories and results produced by scientists, but also the procedures by which they arrive at these conclusions, [4] is necessary for those interested in the assumptions and techniques underpinning the sciences. Both perspectives are valid, but they only show us a part of the scientific picture since they leave out important perspectives or methods. In addition, one of the most valuable lessons from the history of science is the significance of studying scientific methodology within the framework of its historical development. To understand these mechanisms, one must look beyond the surface of previously published research and theory. [5] The Society for Philosophical Study of Science (SPSP) is committed to creating a philosophy of science that actively integrates theory, practise, and the wider world. Defining “practise” is a prerequisite to grasping the significance of the SPSP approach. Deliberate, methodical, and repeated action with the specific goal of improving performance in a certain setting is what is meant by “practise”. [6] Studies of certain practises are thus crucial because they reveal the types of behaviours linked to and required for the advancement of knowledge in a given field. Philosophers have proposed recasting discussions of epistemological notions as actions, such as “truth,” “reality,” “belief,” “certainty,” “observation,” “explanation,” “justification,” “evidence,” and so on. Instead than focusing on theoretical or abstract worries about what counts as good scientific evidence, we might instead explore different (and often contradicting) ways of gathering and weighing data. [7] By gaining insight into the motivations of scientists, we are better able to evaluate not only the epistemological hurdles they face, but also the values, norms, and goals that drive their pursuit of knowledge. Thus, we can no longer accept as self-evident or beyond doubt the philosophical and ontological principles on which such behaviours are based. Focusing on practise helps link the field of philosophy of science with its key topics because the field’s preferred approach to the topic is predominantly epistemological, highly theoretical, and frequently overlooks the ramifications of the sciences as practised. [8] This study aimed at evaluating the philosophy of science oriented to deep learning under the ethical dilemma.

Faculty of Human Ecology, University Putra Malaysia, Malaysia
[1.] *Email: zhoujiezjtx@163.com*
[2]. *Email: ratna_razak@upm.edu.my*

2. Related Work

The combination of philosophical inquiry with a detailed examination of both past and present scientific practises is central to our approach to the philosophy of science. This sense of urgency is a compelling case for bringing in experts from the "front lines" of science and practise, including practitioners, researchers, and policymakers. [9] The purpose is not only to get knowledge from the professionals' viewpoints and experiences, but also to urge them to consider the rationale behind and consequences of their own acts. Many people hold out hope that philosophers of science will finally recognise the field for what it is. In the context of deep learning, the ethical questions raised by the (algorithmic) epistemic opacity conundrum are practical and functional, and involve the issue of epistemic opacity. This condition is exemplified by the epistemic opacity conundrum. [10] For deep learning and comparable AI systems to serve the purposes we have in mind for them to serve, their epistemical opaqueness to us must grow, which hinders interpretability, communicability, and transparency. The operating factors and architecture that give deep learning its strength and variety that make it so valuable are also the reasons why it can be unpredictable and harmful. [11]

Humans require a methodology for addressing the ethical concerns associated with epistemic opacity and making up for the lack of transparency at all levels of abstraction and explanation, beginning with the systems' source code and design and progressing through their higher-level behaviour and operation. [12] In this paper, the author argues that normative ethical goals should be incorporated into the process of training a deep learning neural network. [13] The end goal is to improve the algorithm in various ways, including teaching it to make moral judgments. The cost function, which evaluates the network's performance after each iteration, would thus enforce the normative ethical imperative. Finding the finest normative ethics, if they even exist, is a contentious topic on which people cannot agree. This, however, does not indicate that one normative ethical theory is superior to another when applied to AI. One type of normative ethical framework might be more suitable for use in neural network training than others due to its intrinsic features.

The literature recommends less rule-rigid, less bivalently alethic, and less distinct normative ethical frameworks in this case. Ethical systems that can anticipate and respond to ethical dilemmas may prove very effective. [14] As we saw in the introduction, deep learning systems require patience as they train towards optimal responses and produce illogical and hard to understand patterns of thought. [15] In high-stakes situations like natural disasters, industrial accidents, or medical emergencies, this could cause serious complications. In addition, if the appropriate normative ethical imperatives are input into the hidden layers of neurons, the deep learning system's ability to perform quite sophisticated decision making with its long chains of trained decision modules may be leveraged to great practical ethical advantage. Perhaps the best solution to a challenging situation can be found using this method. The increasing demand in the market can be met with this.

No one will comprehend the reasoning behind the curriculum or how it functions. Is there a compelling reason for people to care if its routine and reliable use as a medical diagnostic system or robotic transportation system improves patient outcomes and saves lives? When it comes down to it, people just don't, [16] in the author's opinion. If the processes are functioning as expected, the so-called algorithmic black-box opacity is not an issue (although the prevalence of self-modifying code may provide researchers and ethicists with some pause). Furthermore, qualitative virtue ethical goals (represented by a later layer of neurons) may be employed in conjunction with a mixed rule-consequentialist model (represented by an earlier layer of neurons). [17] Needles, biopsies, psychometric testing, radiation therapy, anti-psychotics, and other drugs with severe negative side effects may be less likely to effectively override the "don't injure humans" law inside a medical diagnostic system. The Asimov rule might exist in theory, but even if it did, it wouldn't mean much in practise. [18] Existing deep learning systems present significant normative and practical ethical challenges. It will likely be to our advantage to discover how to successfully train them so that we are ready for more strong implementations of artificial general intelligence capabilities in the near future. However, there is no existing research found on this topic so we conducted this study.

3. Materials and Methods

The philosophy of science is concerned with the value and use of scientific knowledge and its underlying assumptions, techniques, and implications. For instance, when examining whether or not scientific findings represent a study of truth, this field of research overlaps with metaphysics, ontology, and epistemology. Although these fundamental concerns of research are of interest to a wide range of scientists and thinkers, many also focus on challenges in particular scientific fields. Using the most recent scientific research, some scientists and philosophers have come to some interesting conclusions concerning the nature of reality. An ethical dilemma arises when a person is forced to choose between two courses of action, neither of which is morally permissible. Although humans face a variety of moral and ethical challenges every day, the vast majority of these problems have straightforward answers. Standards of ethics provide a moral compass by which individuals and institutions can make judgments about right and evil. It is possible for businesses to be required to adopt internal ethical standards developed by firms or professional groups.

When it comes down to it, a deep learning system is just a neural network with three or more layers. These neural networks are an effort to create artificial intelligence that can "learn" procedural tasks from big data sets in the same way as a human brain can. Although a single-layer neural network may still generate a rough estimate, it is possible to improve accuracy by adding hidden layers. [19] Self-driving cars rely heavily on the science of deep learning, which enables them to recognise road signs and tell the difference between people and lampposts. It is crucial for voice-enabled functions in electronic gadgets including smartphones, tablets, TVs, and hands-free audio systems. In deep learning, a computer model is taught to make instantaneous classifications based on inputs such as images, text, or audio. In some cases, DL models can even outperform humans in terms of precision. For model training, we use a huge dataset of labelled data and a multi-layered neural network structure.

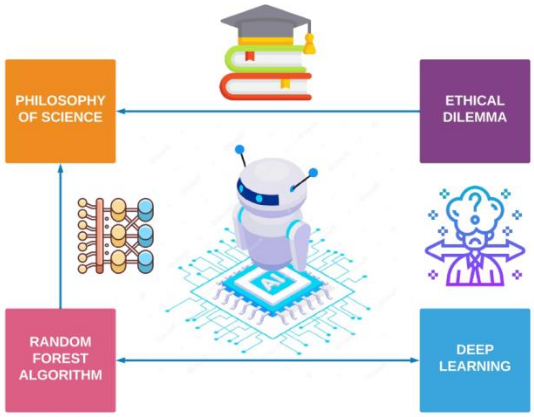

Figure 1. Illustration of the philosophy of science-oriented to DL under the ethical dilemma.

As a type of supervised machine learning algorithm, random forest is frequently applied to problems of classification and regression. Decision trees are built from the collected data, with the majority rule being used for classification and an average for regression. Frequently employed for both classification and regression tasks, random forest is a type of supervised machine learning algorithm. The system uses the majority vote for classification and the average for regression after constructing decision trees from multiple samples. One of the many things that sets the Random Forest Algorithm apart is its ability to deal with data sets that contain both continuous and categorical variables, for use in things like regression and classification. It excels at classifying tasks when other algorithms fail. It extract information from vast volumes of visual data on human philosophy of science directed to deep learning under the ethical dilemma training attitudes and behaviours, as well as to analyse philosophy of science oriented to deep learning under the ethical dilemma actions. Regardless of the fact that research and engineering are progressing at breakneck speeds while data transmission volumes are increasing, extracting behavioural science data from enormous video data sets has become an important task in a range of industries. The footage captured by intelligent surveillance cameras may be instantaneously configured and evaluated. Human characteristics might be detected within real time, ensuring the reliability and completeness of security alerts. As either a result, animal behaviour identification has both philosophical and practical implications, and has become a research topic in a diverse variety of fields. When images can be categorized based on frame and time, the recognition system has become a classification issue.

On the other hand, multi-category classification is far more common. There appear to be a few alternatives here:

a Classification methods are also used, however non-linear and non-soft max regression analysis have also been used. A non - linear and non-categorization is already communicated as $k(i)1,2,\ldots,a$ with an n-category aggregate. So, for such test dataset u, Equation (1) provides the classification possibility addressed in soft and was categorization.

$$M_\vartheta\left(u^{(i)}\right) = \sum \begin{bmatrix} A(k^{(i)} = 1|u^{(i)};\vartheta) \\ A(k^{(i)} = 2|u^{(i)};\vartheta) \\ \cdots \\ A(k^{(i)} = a|u^{(i)};\vartheta) \end{bmatrix} = \sum \frac{1}{\sum_{j=1}^{a} \rho^{\vartheta_j^H k^{(i)}}} \begin{bmatrix} \rho^{\vartheta_1^H u^{(i)}} \\ \rho^{\vartheta_2^H u^{(i)}} \\ \cdots \\ \rho^{\vartheta_a^H u^{(i)}} \end{bmatrix} \quad (1)$$

The ϑidentifies some designer's parameters, that are also permitted by a -line framework. As shown in (2), each dividing line can be viewed as a ρclassification feature for a single category.

$$\rho = \sum_{j=1}^{a} \frac{1}{\rho^{\vartheta_j^H k^{(i)}}} \begin{bmatrix} \rho_1^H \\ \rho_2^H \\ \vdots \\ \rho_a^H \end{bmatrix} + \sum_u M_\vartheta\left(u^{(i)}\right) \; (2)$$

It receive high seems to be the extracting equation stated in (3) when a probability is H normalised in $\frac{1}{\sum_{j=1}^{a} \rho^{\vartheta_j^H k^{(i)}}}$.

$$G(\vartheta) = -\frac{1}{R}\left[\sum_{i=1}^{R}\sum_{j=1}^{a} 1\{k^{(j)} = j\} \sum_{j=1}^{u} \ln \frac{\vartheta_j^H u^{(i)}}{\sum_{j=1}^{a} \rho^{\vartheta_j^H u^{(i)}}}\right] \quad (3)$$

With such an emotive purpose, the R valuation requirements apply. Afterwards when, a soft maximum connect using $\vartheta_j^H u^{(i)}$is used to aggregate the scenarios in a $\sum_{j=1}^{a} \rho^{\vartheta_j^H u^{(i)}}$ category. Equation (4) calculates the probability of u being assigned to one of the j groups.

$$\ln A\left(k^{(i)} = j|g^{(i)};\vartheta\right) = \sum \frac{\vartheta_j^H u^{(i)}}{\sum_{j=1}^{a} \rho^{\vartheta_j^H u^{(i)}}} \quad (4)$$

Equation (4) depicts its $g^{(i)}$extract generalisation of regression $\ln A(k^{(i)} = j|u^{(i)};\vartheta$ analysis. The $k^{(i)} = j$ representation regarding similarities among optimum solutions is seen in the (5)

$$G(\vartheta) = -\frac{1}{R}\left[\sum_{i=1}^{R}\sum_{j=1}^{a} 1\{k^{(i)} = j\} \sum_{j=1}^{k} \ln A(k^{(i)} = j|u^{(i)};\vartheta\right] \quad (5)$$

Similarly, by using an iterative optimization process that includes correct analysis, all optimization techniques in just this equation can be minimized. As a result, (6) illustrates how to estimate a different form of efficiency formalism.

$$\Delta_{\vartheta_j} G(\vartheta) = -\frac{1}{R}\sum_{i=1}^{R} u^{(i)}\left(1\{k^{(i)} = j\}\right) - \sum_{i=1}^{R}\left(A\left(k^{(i)}\right) = j|u^{(i)};\vartheta\right) \quad (6)$$

Equation (6) shows that $\Delta_{\vartheta_j} G(\vartheta)$ is a variables, and that its f^{th} $\frac{\varphi J(\vartheta)}{\varphi \vartheta_{jl}}$ appears to be any f^{th} classification of a currency exchange function. Its equation is used in regression model, but it is also modified recursively to solve the optimization problems, as seen above. Since each analytical solution parameter is stripped with same percentage, the total importance of such a failed functionality doesn't vary greatly, signaling that the parameter must no longer become the only answer.

This dataset was created for the Philosophy Data Project, and it was used to develop the site's features. As a former philosophy instructor who is now a data scientist, I thought it would be fun to apply data science tools to philosophy history. The first intention was to use the data to create a categorization model. After all, a book of philosophy is an attempt to organise one's thoughts on the universe in a methodical manner. Using data from the history of philosophy to classify texts would allow us to classify people's perspectives on the world. Whereas some programmes concentrate on sentiment analysis, we concentrate on philosophical or ideological analysis in this project. There are no limits to what we can do with someone's worldview once we grasp it - from advertising to political campaigning to self-exploration and therapy.

4. Results and Discussion

The graph that follows, A non-linear and non-categorization is already communicated as$k(i)1,2,\ldots,a$ with an n-category aggregate. So, (1) illustrates the classification possibility again for test dataset u depending on this to derive in Fig. 2, which depicts that scientific method oriented to machine learning underneath the ethical dilemma training teaching performance. The performance of effective instruction is evaluated that use the Random Forest Algorithm, Support Vector Machine, and the Fuzzy Set Model, all of which are based on reluctance. The calculation is completed by combining the students' results. The Random Forest Algorithm is a suggested approach with such a lower percentage at the early phases of new technology.However, its algorithm was able to achieve results that have been comparable with those of a Support Vector Machine at a later stage

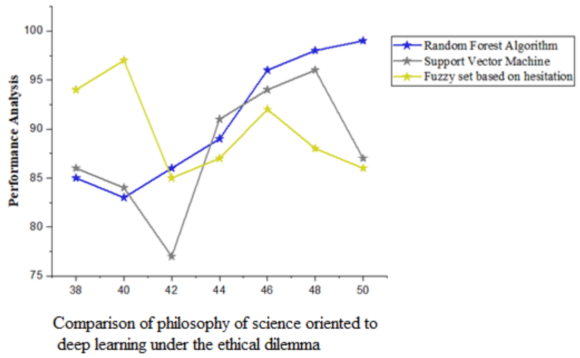

Figure 2. Comparison of measurement the philosophy of science oriented to deep learning under the ethical dilemma.

The findings of the fuzzier are different from those of the random forest as well as svm classification approaches. Also with random forest technique, this comparison graph shows how continual knowledge and education can enhance students' but also teachers' performance. The mathematical form of the learning analysis is displayed. Table 1 reveals it for an average evaluation weight of 38, using random forest technique may have provided worse learning accuracy than that of other algorithms. Furthermore, compared to the other two approaches with varying outcomes, the random forest algorithm improves accuracy less consistently.Eventually, the proposed methodology achieved a 98% accuracy, which is a maximum gain of 90% over the Support vector machine and an 18% growth over the Fuzzy system model.

Table 1. Result analysis for the accuracy in philosophy of science oriented to deep learning under the ethical dilemma

Evaluation of ethical dilemma	Random forest algorithm (%)	Support vector machine (%)	Fuzzy set based on hesitation (%)
38	87	87	95
40	85	86	98
42	84	78	86
44	90	93	88
46	98	95	93
48	97	97	89
50	99	88	87

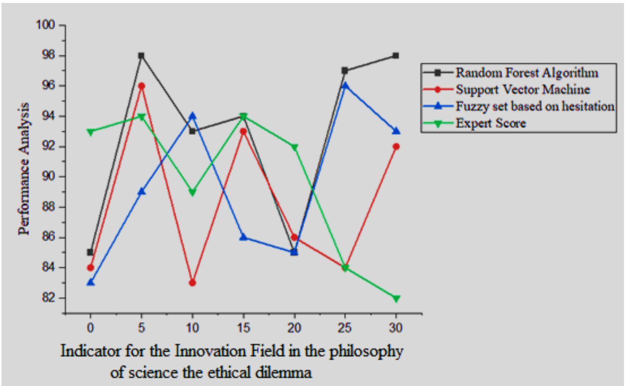

Figure 3. Comparison result analysis for indicator for the innovation field in the philosophy of science the ethical dilemma.

The evaluation process of a scientific method geared to AI learning underneath the ethical dilemma education courses is given alongside the teaching - learning activities, and also the analysis is published with its $g^{(i)}$extract generalisation of regression. The analysis is shown in (4) $\ln A(k^{(i)} = j|u^{(i)};\vartheta$. In Fig. 2, the $k^{(i)} = j$ expression for similarity across ideal shows up as Equation. When compared to expert evaluation, the features extraction evaluation using the random forest method outperformed several assessments. The graph's peaks show the assessment value for a specific assessment for students (refer Fig. 3).

Table 2. Comparison result analysis for the innovation field in the philosophy of science the ethical dilemma using different algorithm

Evaluation of ethical dilemma	Random forest algorithm (%)	Support vector machine (%)	Fuzzy set based on hesitation (%)	Expert score
0 – 5	87	85	86	94
5 – 10	99	97	90	95
10 – 15	94	86	95	88
15 – 20	95	95	88	95
20 – 25	87	87	87	94
25 – 30	98	85	98	87
30 – 35	99	94	95	85

Table 2 provides a quantitative measure of the illustration. This proposed random forest method produced varying results during the learning process, but it has demonstrated steady and enhanced performance at a later stage, according to this table. The proposed system beat the current Support vector machine, fuzzy set depending on hesitation, and expert score, respectively, by 7%, 6%, and 17% toward the final regarded evaluation weight category of 30–35.

5. Conclusion

Because of the exponential development of technology, all kinds of works have undergone technical improvement. Here, we investigate the philosophical questions raised by the ethical conundrum of applying deep learning. The philosophy of science is concerned with the value and use of scientific knowledge and its underlying assumptions, techniques, and implications. An ethical dilemma arises when a person is forced to choose between two courses of action, neither of which is morally permissible. In this study, the Random Forest Algorithm was employed to do both regression and classification analyses. The Random Forest Algorithm is superior to other algorithms when tasked with solving classification issues. The proposed algorithm has provided an accuracy of 99%.

References

[1] A. Daly et al., "Artificial intelligence, governance and ethics: global perspectives", *SSRN Electron J.*, 2019.

[2] P. De Sutter, Automated decision-making processes: ensuring consumer protection, and free movement of goods and services, 2020.

[3] J. Cowls, "Deciding how to decide: six key questions for reducing AI's democratic deficit," in The 2019 Yearbook of the Digital Ethics Lab, C. Burr, S. Milano, eds. Springer International Publishing, Cham, pp. 101–116, 2020.

[4] T. Dragičević, P. Wheeler, and F. Blaabjerg, "Artificial intelligence aided automated design for reliability of power electronic systems," *IEEE Trans Power Electron*, vol. 34, pp. 7161–7171, 2019.

[5] D. Greene, A. L. Hoffmann, L. Stark, "Better, nicer, clearer, fairer: a critical assessment of the movement for ethical artificial intelligence and machine learning." in Proceedings of the 52nd Hawaii International Conference on System Sciences, 2019.

[6] A. Jobi, M. Ienca, E. Vayena, "Artificial intelligence: The global landscape of ethics guidelines," *Nat. Mach. Intell.*, vol. 1, pp. 389–399, 2019.

[7] B. Mittelstadt, "Principles alone cannot guarantee ethical AI," *Nat. Mach. Intell.*, vol. 1, pp. 501–507, 2019.

[8] M. Brundage, S. Avin, J. Wang, H. Belfield, G. Krueger, G. Hadfield, H. Khlaaf et al., Toward trustworthy AI development: mechanisms for supporting verifiable claims. ArXiv:2004.07213 [Cs], 2020.

[9] J. Cowls, A. Tsamados, M. Taddeo, L. Floridi, "A definition, benchmark and database of AI for social good initiatives," *Nat. Mach. Intell.*, 2021.

[10] G. Ras, M. van Gerven, P. Haselager, Explanation methods in deep learning: users, values, concerns and challenges. ArXiv:1803.07517, 2018. http://arxiv.org/abs/1803.07517. Accessed 24 Aug 2020.

[11] I. Rahwan, "Society-in-the-loop: programming the algorithmic social contract," *Ethics Inf. Technol.*, vol. 20, no. 1, pp. 5–14, 2018.

[12] H. Roberts, J. Cowls, J. Morley, M. Taddeo, V. Wang, and L. Floridi, "The Chinese approach to artificial intelligence: An analysis of policy, ethics, and regulation," *AI Soc.*, vol. 36, pp. 59–77, 2020.

[13] A. Erasmus, T. D. Brunet, and E. Fisher, "What is interpretability?" *Philos. Technol.*, vol. 34, pp. 833–862, 2020.

[14] L. H. Gilpin, D. Bau, B. Z. Yuan, A. Bajwa, M. Specter, and L. Kagal, "Explaining explanations: An overview of interpretability of machine learning." in IEEE 5th international conference on data science and advanced analytics (DSAA), pp. 80–89, 2018. arXiv:1806.00069.

[15] I. Goodfellow, Y. Bengio, and A. Courville, Deep learning: MIT Press, 2016.

[16] J. Jebeile, V. Lam, and T. Räz, "Understanding climate change with statistical downscaling and machine learning," *Synthese*, vol. 199, pp. 1877–1897, 2021.

[17] A. Schubbach, "Judging machines: Philosophical aspects of deep learning," *Synthese*, vol. 198, pp. 1807–1827, 2019.

[18] R. Shwartz-Ziv and N. Tishby, Opening the black box of deep neural networks via information, 2017. ArXiv:1703.00810.

[19] M. Reyes, R. Meier, S. Pereira, C. A. Silva, F. Dahlweid, H. von Tengg-Kobligk, R. M. Summers, and R. Wiest, "On the Interpretability of Artificial Intelligence in Radiology: Challenges and Opportunities," *Radiol. Artif. Intell.*, vol. 2, no. 3, 2020.