Review

Combining Fractional Derivatives and Machine Learning: A Review

Sebastian Raubitzek [1,]*[ID]**, Kevin Mallinger** [2][ID] **and Thomas Neubauer** [2][ID]

1   Data Science Research Unit, TU Wien, Favoritenstrasse 9-11/194, 1040 Vienna, Austria
2   SBA Research gGmbh, Floragasse 7, 1040 Vienna, Austria
*   Correspondence: sebastian.raubitzek@tuwien.ac.at

Abstract: Fractional calculus has gained a lot of attention in the last couple of years. Researchers have discovered that processes in various fields follow fractional dynamics rather than ordinary integer-ordered dynamics, meaning that the corresponding differential equations feature non-integer valued derivatives. There are several arguments for why this is the case, one of which is that fractional derivatives inherit spatiotemporal memory and/or the ability to express complex naturally occurring phenomena. Another popular topic nowadays is machine learning, i.e., learning behavior and patterns from historical data. In our ever-changing world with ever-increasing amounts of data, machine learning is a powerful tool for data analysis, problem-solving, modeling, and prediction. It has provided many further insights and discoveries in various scientific disciplines. As these two modern-day topics hold a lot of potential for combined approaches in terms of describing complex dynamics, this article review combines approaches from fractional derivatives and machine learning from the past, puts them into context, and thus provides a list of possible combined approaches and the corresponding techniques. Note, however, that this article does not deal with neural networks, as there is already extensive literature on neural networks and fractional calculus. We sorted past combined approaches from the literature into three categories, i.e., preprocessing, machine learning and fractional dynamics, and optimization. The contributions of fractional derivatives to machine learning are manifold as they provide powerful preprocessing and feature augmentation techniques, can improve physically informed machine learning, and are capable of improving hyperparameter optimization. Thus, this article serves to motivate researchers dealing with data-based problems, to be specific machine learning practitioners, to adopt new tools, and enhance their existing approaches.

Keywords: fractional derivative; fractional calculus; machine learning; artificial intelligence; complexity; regression analysis

1. Introduction

Machine learning and fractional calculus are tools capable of dealing with and describing complex real-life phenomena and its relation to its inherent nonlinear properties.

Whereas machine learning dynamically learns complex behavior from data in most cases, the framework of fractional calculus was previously used to describe complex phenomena by statically modeling them. The issue of a fractional derivative first came up in a letter from de l'Hopital to Leibniz in 1695, i.e., what one would obtain from a derivative of non-integer order n, e.g., $n = \frac{1}{2}$. Leibniz's famous response to this day was: "It will lead to a paradox from which one day useful consequences will be drawn" [1].

Nowadays, both frameworks feature a multitude of applications in various disciplines.

The applications of fractional calculus range from physics, wherein one can describe anomalous diffusion in complex liquids or frequency-dependent acoustic wave propagation, to, e.g., image processing, where one can use fractional derivatives as filters or to enhance images [2]. Another area of research in which fractional calculus is used to describe various phenomena is that of the environmental sciences. Here, fractional calculus is applied to interpret hydrological cycles or analyze underground water chemical

reactions [2,3]. As discussed by Du et al. [4], fractional calculus introduces memory into modeling processes, which in turn is required for specific approaches in, e.g., biology or psychology. Thus, the researchers suggested interpreting the fractional order of a process as a degree of memory of this process [4].

This article aimed to find overlaps between machine learning and fractional calculus, specifically fractional derivatives. Furthermore, we sought to discuss these overlaps or combined applications and put them into a bigger context and finally, give some recommendations on using fractional derivatives to enhance machine learning approaches. At this point, we need to clarify that, when employing the popular term machine learning and narrowing down the scope of this article, we only talk about supervised non-neural-network machine learning applications. The reason for not including neural networks is that there already is an excellent article on the combined approaches of neural networks and fractional calculus. Thus, we are not going to discuss this topic again but instead, refer the reader to the article by Viera et al. [5]. Still, we want to highlight two articles from this review to motivate the topic. Khan et al. [6], presented a fractional backpropagation algorithm for training neural networks. Furthermore, the work in Ref. [7] showed the applicability of a particle swarm optimization algorithm with a fractional order neural network.

We also want to clarify that the scope of this article is to provide additional tools from fractional calculus to the experienced machine learning practitioner rather than introducing people familiar with fractional calculus to the realm of machine learning. Thus, the outcome of this article is a focused list of possible combined applications of machine learning and fractional derivatives and, based on the discussed literature, a discussion on the context of the combination of these two tools as well as recommendations for future research.

This article is structured as follows:

Section 2 describes how this review was performed, the employed search, and the exclusion criteria. Section 3 briefly introduces fractional calculus and describes the different types of fractional derivatives. Section 4 introduces supervised machine learning. Then, we present the results of our literature review in Section 5. We discuss our findings and put them into context in Section 6 and conclude this article in Section 7. We further added Appendix A, which provides a table containing all reviewed articles.

2. Methodology

For this review, we performed an online search on (google scholar) https://scholar.google.com/. We searched for the keywords listed in the following list:

- Machine learning fractional calculus
- Machine learning fractional derivatives
- Support vector fractional calculus
- Support vector fractional derivatives
- Decision tree fractional calculus
- Decision tree fractional derivative
- Random forest fractional calculus
- Random forest fractional derivative
- XGBoost fractional calculus
- XGBoost fractional derivative
- Regression fractional calculus
- Regression fractional derivative
- Ridge regression fractional calculus
- Ridge regression fractional derivative
- Lasso regression fractional calculus
- Lasso regression fractional derivative
- Logistic regression fractional calculus
- Logistic regression fractional derivative
- Naive Bayes fractional calculus
- Naive Bayes fractional derivative
- KNN fractional calculus
- KNN fractional derivative
- Nearest neighbor fractional calculus
- Nearest neighbor fractional derivative
- LightGBM fractional calculus
- LightGBM fractional derivative
- Extreme learning fractional calculus
- Extreme learning fractional derivative
- Extreme boosting fractional calculus
- Extreme boosting fractional derivative

We acknowledge that this list is incomplete; further, it can never be a complete list. This is because of the somewhat recently popularized blurry use of the term machine learning and lingo and the ever-changing landscape of new algorithms and ideas in the field. Here, we want to refer to two homepages [8,9], which are both glossaries that keep up with the fast-changing machine learning terminology. Thus, the previous list contains popular machine learning algorithms, prevalent machine learning, and fractional calculus keywords, and occasionally some more niche keywords, which are included because of the authors' domain knowledge.

Furthermore, the authors added some articles that did not appear in the mentioned search queries but were deemed to be a relevant addition to this article based on the authors' domain knowledge.

As we usually found articles fitting our exclusion criteria within the first ten results, we prioritized our search and analysis to the first 20 results for each search query.

As already mentioned, we employed a range of exclusion criteria to keep this review focused and to reduce the number of featured articles. Thus, we used the following exclusion criteria:

- We excluded articles primarily targeting and/or only targeting neural networks. This was done because there already exists an excellent review of neural networks and fractional calculus [5]. Nonetheless, we will peripherally mention neural network applications or feature articles that include neural networks along with other machine learning algorithms.
- We excluded articles dealing with control problems. This criterion was employed to keep this review focused, as including control problems would have unnecessarily blown up the discussed methodology and made it harder to identify key takeaways for hybrid applications of fractional calculus and machine learning.
- We discarded everything related to grey models. Grey models do not coincide with the classic supervised learning approaches we are looking for in this review.
- We excluded publications that dealt with fractional integrals rather than with fractional derivatives.
- We excluded publications that featured fractional complexity measures, e.g., fractional entropy measures. Nonetheless, we briefly mention them in our discussion if they provide valuable insights into hybrid applications.
- We excluded unsupervised and reinforcement learning approaches. Thus, we are only dealing with supervised learning.

3. Fractional Derivatives

Fractional derivatives are a generalization of integer-ordered derivatives such that they allow for derivatives of the order of $n = \frac{1}{2}$ instead of only integer-valued derivatives, i.e., $n = 1, 2, 3, \ldots$.

Contrary to integer order derivatives, they do not behave locally, i.e., the fractional derivative of a function $f(t)$ depends on all values of the same function, thus inducing memory about every point of a signal into the derivative.

Furthermore, there are many ways to define a fractional derivative; thus, many researchers proposed their version of the fractional derivative, sometimes explicitly tailored to match a particular problem/application. To further illustrate this idea and to give the reader a grasp on the multitude of possible fractional derivatives, we give the list from https://en.wikipedia.org/wiki/Fractional_calculus (Wikipedia) [10]: the listed fractional derivatives are the Grünwald–Letnikov derivative, the Sonin–Letnikov derivative, the Liouville derivative, the Caputo derivative, the Hadamard derivative, the Marchaud derivative, the Riesz derivative, the Miller–Ross derivative, the Weyl derivative, the Erdélyi–Kober derivative, the Coimbra derivative, the Katugampola derivative, the Hilfer derivative, the Davidson derivative, the Chen derivative, the Caputo-Fabrizio derivative, and the Atangana–Baleanu derivative.

Because of this arbitrariness of choosing a fractional derivative, we stick to those fractional derivatives relevant in our literature review, briefly explain them, and reference sources for a discussion on them. Thus, the discussed fractional derivatives are the Grünwald–Letnikov, the Reisz, the Caputo, and the Riemann–Liouville fractional derivatives. Furthermore, we need to differentiate between left- and right-sided derivatives for many of these fractional derivatives. The following list is based on the reviews by de Oliveira et al. and Aslan [11,12], which we refer the reader to for an in-depth discussion of the topic. Furthermore, for a discussion on the discrete version of the fractional derivatives, we refer

to [13,14]. The featured applications combining fractional derivatives and machine learning use the following list of fractional derivatives:

- **The Grünwald–Letnikov fractional derivative**

$$\text{left-sided: } D[f(x)]^{\alpha}_{a^+} = \lim_{h \to 0} \frac{1}{h^{\alpha}} \sum_{k=0}^{\lfloor n \rfloor} (-1)^k \frac{\Gamma(\alpha+1)f(x-kh)}{\Gamma(k+1)\Gamma(\alpha-k+1)}, \quad nh = x - a$$

$$\text{right-sided: } D[f(x)]^{\alpha}_{b^-} = \lim_{h \to 0} \frac{1}{h^{\alpha}} \sum_{k=0}^{\lfloor n \rfloor} (-1)^k \frac{\Gamma(\alpha+1)f(x-kh)}{\Gamma(k+1)\Gamma(\alpha-k+1)}, \quad nh = b - x \tag{1}$$

- **The Caputo Fractional Derivative**

$$\text{left-sided: } D^{\alpha}_{a^+}[f(x)] = \frac{1}{\Gamma(n-\alpha)} \int_a^x (x-\zeta)^{n-\alpha-1} \frac{d^n}{d\zeta^n}[f(\zeta)]d\zeta, \quad x \geq a$$

$$\text{right-sided: } D^{\alpha}_{b^-}[f(x)] = \frac{1}{\Gamma(n-\alpha)} \int_x^b (x-\zeta)^{n-\alpha-1} \frac{d^n}{d\zeta^n}[f(\zeta)]d\zeta, \quad x \leq b \tag{2}$$

- **The Riemann–Liouville fractional derivative**

$$\text{left-sided: } D^{\alpha}_{a^+}[f(x)] = \frac{1}{\Gamma(n-\alpha)} \frac{d^n}{dx^n} \int_a^x (x-\zeta)^{n-\alpha-1} f(\zeta)d\zeta, \quad x \geq a$$

$$\text{right-sided: } D^{\alpha}_{b^-}[f(x)] = \frac{1}{\Gamma(n-\alpha)} \frac{d^n}{dx^n} \int_x^b (x-\zeta)^{n-\alpha-1} f(\zeta)d\zeta, \quad x \leq b \tag{3}$$

- **The Riesz Fractional Derivative**

$$D^{\alpha}_x[f(x)] = -\frac{1}{2\cos\left(\frac{\alpha\pi}{2}\right)} \frac{1}{\Gamma(\alpha)} \frac{d^n}{dx^n} \left(\int_{-\infty}^x (x-\zeta)^{n-\alpha-1} f(\zeta)d\zeta + \int_x^{\infty} (\zeta-x)^{n-\alpha-1} f(\zeta)d\zeta \right) \tag{4}$$

Remark 1.

- *α denotes the order of the fractional derivative. Whereas $\alpha \in \mathbb{C}: \mathbb{R} \in (n-1, n]$, $n \in \mathbb{N}$. Thus, n defines the integer proximity of α. $\mathbb{R}(\cdot)$ denotes the real part of a complex number. Furthermore, as defined here, alpha can be complex. However, every single reviewed article (see Appendix A for the summary table) featuring a combined application of fractional derivatives and machine learning uses a real or at least rational value for α.*
- *$n-1$ and n as $n-1 \leq \mathbb{R}(\alpha) < n$.*
- *$[a, b]$ is a finite interval in $\mathbb{R}$ and $k \in \mathbb{N}$. Furthermore, $f(0) \equiv f(0^+) - f(0^-)$.*
- *ζ is an auxiliary variable used for integration.*
- *$\Gamma(\cdot)$ is the gamma function [15].*
- *$\lfloor \cdot \rfloor$ denotes the floor function, i.e., $\lfloor x \rfloor \max\{z \in \mathbb{Z}: z \leq x\}$*
- *The previously mentioned "memory" is induced by the summation over k in the case of the Grünwald–Letnikov derivative and by the integration over ζ for the other mentioned derivatives.*

4. Supervised Machine Learning

Supervised machine learning is the practice of using a set of input variables, such as blood metabolite, gene expression, or environmental conditions, to predict a quantitative output, such as diseases in humans or crop yields [16]. Thus, aiming to categorize data from prior information via an algorithm trained from labeled data [17,18].

There are numerous supervised learning algorithms, such as linear regression, support vector machines, decision trees, and random forests, to name a few.

One can further partition supervised machine learning into classification and regression tasks. In classification tasks, the algorithm predicts categorical variables such as types

of diseases. In contrast, in regression, the algorithm predicts continuous numerical outputs such as wheat yields in, e.g., Hg/Ha.

We will not further discuss supervised learning for several reasons. First, this review aimed to educate machine learning practitioners about the possible applications of fractional derivatives, thus assuming that the reader is somehow familiar with supervised learning. Second, supervised learning and machine learning, in general, constitute a popular multidisciplinary area of research nowadays. Therefore, it suffers from a fast-changing and ever-increasing landscape of algorithms, terminology, and approaches. Third, there is a vast amount of literature and tutorials available that we cannot cover here. Nonetheless, we want to recommend the books by Jason Brownlee for a tutorial-heavy introduction to the topic [19–21].

5. Results: Combined Approaches of Fractional Derivatives and Machine Learning

Given the previously defined methodology (Section 2), we found 60 relevant publications. We then separated this list of publications in this section into three main categories on how to combine machine learning and fractional derivatives, i.e., preprocessing, machine learning and fractional modeling and optimization. In the following, we will elaborate on these categories separately and further divide them into subcategories wherever necessary.

5.1. Preprocessing

In this section, we discuss the applications of fractional derivatives for data preprocessing in machine learning applications. Part of this discussion is data pre-treatment in general, additional features based on fractional derivatives, and feature extraction based on fractional derivatives. These ideas, though distinct, often overlap, especially when discussing additional features and feature extraction. All these techniques have in common that they are applied before the actual machine learning, e.g., the training process. Thus, we named this Section *preprocessing*.

We sort relevant publications by their type of application, e.g., fractional derivatives used for preprocessing in spectroscopy. In this process, we focus on three categories of applications, i.e., spectroscopy, biomedical and engineering applications.

5.1.1. Spectroscopy

An important area of research for the combined approaches of fractional derivatives and machine learning is spectroscopy. Here, fractional derivatives are used as a preprocessing step to enhance the spectral data and thus, improve the accuracy of the machine learning algorithm. We can again differ between two main categories, i.e., visual and near-infrared (Vis-NIR) spectroscopy and hyperspectral analysis.

Regarding visual and near-infrared spectroscopy, we find a range of agricultural applications. In [22], the researchers estimate organic matter content in arid soil from Vis-NIR spectral data. Here, the Grünwald–Letnikov fractional derivative is used as a pretreatment for the spectral data, i.e., the fractional derivative is applied to the obtained spectral data to improve the accuracy of the employed machine learning algorithms, random forest, and partial least squares regression. Similar to this approach is the approach in [23], where fractional derivatives are again used as a pretreatment to augment the Vis-NIR data to estimate the soil organic content. Here, additionally to the previously used algorithms, the researchers also used memory-based learning.

The same idea, i.e., to augment Vis-NIR data using the G-L derivative to improve machine learning estimates, is used in both [24,25] to estimate the nitrogen content of rubber tree cultivations and in [26] for cotton farming. The employed machine learning algorithms range from partial least squares regression, extreme learning machines, convolutional neural networks, and support vector machines.

For another application in agriculture, Bhadra et al. employed the Grünwald–Letnikov fractional derivative to augment Vis-NIR data to quantify the leaf chlorophyll in sorghum

using various machine learning algorithms such as, again, partial least squares regression, random forest, support vector machines, and extreme learning machines [27].

Urbanization and industrialization are putting a lot of stress on the environment and further polluting agricultural land with, e.g., various metals. Here [28–30] provided ideas to identify zinc lead and other heavy metals from Vis-NIR spectra by, again, employing data augmentation based on the Grünwald–Letnikov fractional derivative to improve the accuracy of machine learning algorithms such as random forest, XGBoost, extreme learning machines, support vector machines, and partial least squares regression.

Finally, the Grünwald–Letnikov fractional derivative is also used to improve the discussed Vis-NIR data to improve the estimation of soil salinity, soil salt, and other water-soluble ions in [31,32]. Here, again, the researchers used the Grünwald–Letnikov fractional derivative to augment the spectral data and/or obtain spectral indices. Again, various ML algorithms were employed, such as partial least squares regression, random forests, and extreme learning machines.

When it comes to hyperspectral data, we find similar approaches to the discussed Vis-NIR spectral applications, i.e., the spectral data were preprocessed using the Grünwald–Letnikov fractional derivative and afterward a measured observable such as, e.g., soil organic matter is predicted using a machine learning algorithm such as, e.g., partial least squares regression [33]. In [34], estimates on top soil organic carbon were given using a Grünwald–Letnikov fractional derivative and random forest.

Furthermore, the (soil) salt content can be estimated using hyperspectral data and a range of ML algorithms, as performed in [31,35,36]. Another agricultural application is to assess the nitrogen content in apple tree canopies via support vector machines, random forest algorithms, and fractional-derivative-augmented hyperspectral data [37].

Cheng et al. predicted the photosynthetic pigments in apple leaves using fractional derivatives to augment the data and range of machine learning algorithms, i.e., support vector machines, k-nearest neighbor, random forest, and neural networks [38]. In [39], it is shown that a similar approach using a different algorithm, XGBoost, can be used to estimate the moisture content of agricultural soil. Furthermore, in [40], XGBoost, LightGBM, random forests, and other algorithms were used to predict the soil electrical conductivity from hyperspectral data using fractional-derivative-augmented data. Although related, but not necessarily an agricultural application, the use of augmented hyperspectral data to estimate the clay content in desert soils using partial least squares regression was discussed in [41].

Furthermore, these ideas of using fractional derivatives to preprocess hyperspectral data to improve ML approaches are not restricted to only using the remotely sensed data of soils. These ideas were applied to water in both [42,43]. In [42], fractional derivatives were used to improve the prediction of total nitrogen in water using XGBoost and random forest, whereas in [43], a similar approach was used for estimating the total suspended matter in water using random forest.

5.1.2. Biomedical Applications

We also find various applications when it comes to the biomedical applications of fractional derivatives combined with ML.

Starting with the classification of EEG signals, we find three publications that combine machine learning and fractional derivatives. Ref. [44] proposed to the classification of electroencephalography (EEG) signals into ictal and seizure-free signals by employing a fractional linear prediction from a fractional order system to model the signals, and/or parts of the signal, and to then use the thus-obtained parameters as features for a support vector machine classification approach. Similar to this is the approach taken by AAruni et al. [45], wherein EEG signals were modeled using the transfer function of a fractional order system. In contrast to the previous publication, the researchers used the error and signal energy as features for support vector machines to differentiate between ictal and seizure-free signals. The third publication dealing with this subject, i.e., modeling EEG signals using fractional transfer functions and classifying between different categories of EEG signals, e.g., seizure

and seizure-free, is [46]. Here, similarly to the two previously referenced publications, the fractional modeling transfer function was used to obtain the error and signal energy to be fed into a k-nearest neighbor classifier.

Furthermore, similar approaches can be used to analyze electrocardiography (ECG) signals. Again, in [47], a fractional order transfer function was applied to the QRS complex signal (from ECG measurements) to obtain six parameters which are then fed into a k-nearest neighbor classifier for person identification. Ref. [48] described an approach to do the same to obtain five parameters from a QRS complex signal to differentiate between healthy and three types of arrhythmic signals using a k-nearest neighbor classifier.

Given the results shown in [49], one can use the same approach, e.g., fractional transfer functions in combination with a k-nearest neighbor classifier, to analyze the respiratory properties of humans. In this process, three model parameters are used as features to classify respiratory diseases.

In [50], the researchers used the Grünwald–Letnikov fractional derivative to preprocess online handwriting data. Afterward, these preprocessed data were fed into a support vector machine and random forest algorithms to detect/classify Parkinson's disease from a person's handwriting.

In contrast to all previously listed publications in this section, in [51], a discrete fractional mask was used to preprocess the CT images and to improve the identification of tumors using support vector machines, random forest, J48, and simple cart machine learning models.

5.1.3. Engineering

Apart from the two previously discussed categories, we found three publications that did not fit into spectroscopical or biomedical applications. Instead, we consider these publications to be engineering approaches.

In [52], the researchers used histogram peak distributions and the Grünwald–Letnikov fractional derivative to preprocess the images of solar panels to then identify defective solder joints via extreme learning machines.

Ref. [53] used the Reisz fractional derivative to preprocess spectral data obtained from a fiber Bragg grating sensor. Afterwards, these data were fed into three machine learning algorithms, i.e., random forest, linear regression, decision trees, and a multi-layer perceptron, to detect the temperature peaks on solar panels from the FBG spectrum.

In [54], the researchers identified construction equipment activity from image/video data using a combination of a fractional derivative and a range of machine learning algorithms. The employed machine learning algorithms were random forest, neural networks (a multi-layer perceptron), and support vector machines. The fractional derivative employed was the Riemann–Liouville fractional derivative, and it was used to preprocess the images, which were fed into the algorithms at a later step.

5.2. *Machine Learning and Fractional Dynamics*

In this section, we discuss publications where fractional dynamics and the corresponding modeling via fractional differential equations has been used in combination with machine learning. There are several ways to do this, e.g., fractional modeling and after employing machine learning for a regression or classification using features obtained from the modeling approach or using machine learning to improve the identification of fractional models/equations from data.

Although the publications by [44–49] have already been categorized as preprocessing, they also serve as a hybrid machine learning and fractional dynamics approach. Specifically, as the authors employed the transfer function from a fractional order model, first a a signal was modeled to obtain the functions parameters or the corresponding error and signal energy of the model as features for classification.

When identifying the correct transfer function and the underlying dynamics, i.e., with the corresponding fractional differential operators, we found an excellent exemplary

application in [55]. Here, the employed machine learning technique was a Gaussian process regression, whereas the unknown equation parameters are treated as the hyperparameters of a physics-informed Gaussian process kernel. This framework was used to identify a linear space-fractional differential equations from data and its applicability is shown for stock market data, e.g., S&P 500 stock data.

In [56], similar ideas were employed to model general damping systems. Here, k-means was employed to first differ between viscous and hysteretic damping. Next, the viscous damping dynamics were recovered using a linear regression approach, and afterwards, the hysteretic damping was identified using a sparse regression approach. This framework also allows for fractional-order models, and its applicability was shown for, e.g., a fractional-order viscoelastic damper.

We further found an application for machine learning to identify fractional-order population growth models from data, as performed in [57]. The researchers developed a least squares support vector machine with an orthogonal kernel for simulating this equation. This orthogonal kernel consists of fractional rational Legendre functions.

In [58], the researchers employed ridge regression as a numerical boundary treatment for the fractional KdV-Burgers equation. Another application for identifying fractional models was given by [59], where the researchers used a Gaussian process regression to identify a fractional chaotic model in finance from data. Furthermore, as the extrapolation of the thus-identified model does not perform well, the researchers suggested employing a neural network to improve prediction accuracy.

In [60], both a fractional order model and a first-order resistor–capacitor model were employed to describe the electrical behavior of battery cells with external short circuit faults. Afterwards, a random forest classifier was fed with features from the so-obtained fractional order model to classify whether a battery cell incurs leakage.

5.3. Optimization

In this section, we differ between derivative-free and gradient-based methods within the optimization context. This is also to say that the derivative-free algorithms, e.g., particle swarm algorithms, still employ fractional derivatives in the form of, e.g., fractional velocities. The featured derivative-free algorithms include bio-inspired algorithms such as, e.g., the fractional bee colony algorithm and the fractional particle swarm algorithm.

5.3.1. Fractional Gradient-Based Optimization

When discussing the optimization, the first big topic is gradient-based algorithms. Here, the basic idea is to alter the gradient-based approach by switching to a fractional gradient. One benefit of doing so is that these algorithms do not fall into the trap of the local minimum problem as much as regular gradient-based algorithms (find a good citation here).

Regarding the actual applications where fractional derivatives are used for optimization in combination with machine learning algorithms, we find a range of fractional gradient descent applications. We want to point out that all the discussed combined approaches employ the Caputo-fractional derivative but use different algorithms and datasets. In [61], the researchers used ridge regression in combination with fractional gradient descent on the testing environment and datasets provided by [62]. The same testing environment is used in [63]. However, in this publication, the researchers used logistic regression instead of ridge regression. In both cases, the researchers reported improvements to basic (non-fractional) implementations.

Two more applications of fractional gradient descent are found in [64,65]. In these publications, the researchers combined fractional gradient descent with support vector machines for two classic datasets, i.e., the Iris and the Rainfall datasets.

Contrary to the previous modeling approaches are the ideas presented in [66]; here, a multilinear regression is modified using a fractional derivative based on the input features to the linear regression. Thus, obtaining the derivative and, consequently, the fractional

derivative from the minimum of the slope means setting it to zero. This idea is used to predict the gross domestic product in Romania. It is shown that models not considering fractional derivatives can be outperformed using the presented concepts. Awadalla et al. [67] presented a similar approach where the same idea is used for fractional linear and fractional squared regression, i.e., the derivatives to derive the minimum are fractionalized. This approach also shows the applicability of the proposed ideas to standard datasets.

5.3.2. Fractional Gradient-Free Optimization

Particle swarm optimization algorithms are some of the most popular gradient-free optimization techniques. In this section, we further differed between regular and Darwinian particle swarm algorithms. All particle swarm algorithms feature a velocity component, which is an attempt to make the algorithms fractional, is changed to a non-integer derivative-based velocity [68].

Apart from gradient-based algorithms, we find a range of bio-inspired gradient-free algorithms. The following discussion includes a fractional particle swarm, fractional Darwinian particle swarm, and fractional bee colony algorithms.

We are starting with the work of Chou et al. [69], which presents a fractional order particle swarm optimization to improve the classification of heart diseases via an XGBoost classifier. Furthermore [70] presented an improved fractional particle swarm optimization algorithm and shows its applicability to optimize support vector machine and K-Means algorithms. As used in the same domain, the researchers used a dataset where the task is to classify heart diseases.

An extension on the regular particle swarm optimization is the Darwinian particle swarm optimization algorithm, introduced in 2005 by Tillet et al. [71]. Couceiro et al. [72] introduced a fractional version of this algorithm. Ghamisi et al. later used these ideas in combination with random forest for classifying spectral images [73]. It is used in [74] again for spectral images, i.e., for the multilevel segmentation of spectral images later to be processed by support vector machines. Of course, we could also list these tasks in preprocessing, but since particle swarm is a classical optimization technique, we decided to use them in this category, i.e., optimization. Another application is discussed by Wang et al. in [75], where the researchers used a fractional Darwinian particle swarm optimization algorithm to optimize the feature selection for extreme learning machines. Another example to use an optimization procedure for preprocessing is found in [76], where the researchers used a fractional order Darwinian particle swarm algorithm to delineate cancer lesions. In a later step, these CT scan images are classified using decision trees. In both [77,78], the fractional-order Darwinian particle swarm is employed to segment MRI brain scans, which are then analyzed for strokes using random forest and support vector machines. Two further biomedical applications for fractional order Darwinian particle swarm are found in [79,80]. In both publications, the researchers used the discussed optimization algorithm to segment medical images, i.e., for detecting brain tumors and foot ulcers, respectively. Both publications then employ a supervised ML algorithm for classification, whereas the former uses Naive Bayes and a tree classifier, and the latter a fuzzy c-means classifier.

A different approach for optimizing machine learning approaches was taken in [81], where the researchers predicted the water quality using a fractional order artificial bee colony algorithm and decision trees. Another example for combining fractional order artificial bee colony algorithms and ML is given in [82] for predicting the nonperforming loans using a nonlinear regression model.

6. Discussion

Here, we summarize and discuss the performed literature analysis. We found a total of 60 publications that fit the criteria defined in Section 2. In the previous section, we separated all publications into three categories, i.e., preprocessing (Section 5.1), optimization (Section 5.3), and machine learning and fractional dynamics. We further sub-sectioned preprocessing and optimization because we found a variety of publications targeting similar

problems in each of those categories. For preprocessing, we found three prominent types of applications, i.e., spectroscopy, biomedical and engineering applications. Contrary to that, we sorted all publications fitting into optimization into two categories. In contrast to preprocessing, we did not choose the categories with respect to the application but regarding the employed technique, whether it was a classic but then fractionalized gradient-based optimization technique or a gradient-free method.

A table containing all publications from Section 5 is given in Appendix A.

6.1. Preprocessing

The results for preprocessing show that a majority of the analyzed publications fall into the realm of spectroscopy. Here, researchers employed fractional derivatives to alter spectral data at hand to improve the machine learning algorithm's accuracy. Overall, all of these approaches have in common that there is one or more *good* orders of fractional derivatives that enhance the data-based learning approach. We want to highlight that tree-based classifiers, e.g., random forest and XGBoost, are very common and perform well for these applications. We want to point out that Ref. [42] used fractional derivatives to preprocess hyperspectral data, which are then used to estimate the total nitrogen content in water via different machine learning algorithms, e.g., random forest.

Regarding biomedical applications, we find that here, similarly to preprocessing spectral data, researchers used fractional derivatives to preprocess their medical images, as done in [51] for CT images. Another example is to preprocess EEG or ECG signals, used to classify diseases or malfunctions as done in [46,48].

When it comes to engineering applications, we again found similar ideas, e.g., preprocessing images and spectral data to identify the malfunctions of solar panels or detect the status of construction equipment as performed in [53,54].

Overall, fractional derivatives can improve the accuracy of machine learning algorithms when employed for preprocessing signals, images, or spectral data. Furthermore, we find that there is a *sweet spot*, i.e., one or more orders of fractional derivatives that can significantly improve the accuracy of the employed algorithm. We further want to point out that, although fractional order preprocessing enhances the accuracy of machine learning applications, it also introduces a new hyperparameter, the order of the fractional derivative. Thus, it does not simplify the task at hand but instead provides a slight increase in accuracy at the cost of optimizing another hyperparameter.

Furthermore, we found that every application in which an actual fractional derivative was used for preprocessing involved the Grünwald–Letnikov fractional derivative.

Furthermore, we considered both data manipulation and feature extraction as preprocessing. Here, one can differ between manipulating the original data using a fractional derivative (e.g., [42]) and extracting features, i.e., error energy and parameters from the original data (e.g., [48]).

The authors believed that the preprocessing step using fractional derivatives is especially useful for non-neural network machine learning algorithms. The reason for this is that, given a deep neural network, the shallow layers of a neural network can make preprocessing obsolete as a neural network is capable of learning a huge variety of data manipulations. Thus, we suggest future research to test this hypothesis. Specifically, one needs to test whether a neural network can perform any preprocessing based on fractional derivatives.

Finally, an addition noteworthy to this discussion on fractional derivatives and preprocessing is the fractional measures of (signal) complexity, which can be used as a preprocessing step for machine learning applications. This is performed in [83], in which the researchers employed a fractional entropy measure to improve the fault detection for train doors.

6.2. Machine Learning and Fractional Dynamics

Regarding possible combinations of fractional dynamics and machine learning, we find that machine learning can be used to identify fractional order behavior from data, as performed in [55] or [59]. Overall, this is the category where we found the least amount of publications. This might be due to the fact that the fractional order modeling of dynamics requires profound knowledge in both mathematics and machine learning. Furthermore, we want to emphasize this category for its potential in the future. We suggest that researchers use, e.g., fractional orthogonal kernels in kernel-based machine learning algorithms as, e.g., support vector machines or Gaussian process regression to model systems from data [84].

6.3. Optimization

Regarding optimization, we split our review into the two main sections, i.e., gradient-based and gradient-free methods. The gradient-based optimization techniques feature gradient descent approaches where the integer-order gradient is replaced with a fractional-order gradient. For the gradient-free techniques, the idea is to replace, e.g., the velocity of a particle swarm algorithm with a fractional velocity.

Here, we want to point out two exemplary applications by Hapsari et al. [64,65] on well-known test-datasets, i.e., the Iris and the rainfall dataset, for the gradient-based optimization of machine learning algorithms.

For the gradient-free optimization algorithms, we highlight the work performed by Li et al. [70], where the researchers present an improved fractional particle swarm optimization algorithm to improve a support vector machine and k-means-based classification of heart diseases.

6.4. Bringing It All Together

This review's motivation and the initial assumption is that combined approaches of fractional derivatives and machine learning improve machine learning tasks. We identified a total of three categories from the literature within our frame of keywords and criteria. The applications include processing, optimizing, and modeling fractional dynamics via machine learning. We discussed all of our findings in the previous sections. This section, however, brings all our results together and puts them into a bigger context. Here, we are taking into account the findings of the work of Niu et al. from 2021 [85], where the researchers proposed a triangle consisting of machine learning, fractional calculus, and the renormalization group. We depicted this triangle in Figure 1. The triangle suggests overlaps and further shows where potential combined approaches of machine learning and fractional calculus can be beneficial. Furthermore, we consider this triangle a closed loop; thus, future research might employ all of the mentioned tools, i.e., machine learning, fractional calculus, and the renormalization group, to tackle complex problems and describe complex systems. Thus, we describe our findings in this context and consider the results of another article by Niu et al. [86], entitled *"Why Do Big Data and Machine Learning Entail the Fractional Dynamics?"* in which the researchers further discuss ideas on the overlap between machine learning and fractional calculus, specifically for big data.

The following paragraphs discuss the overlaps shown by Niu et al., as depicted in Figure 1, compare them to the results of our literature review, and further provide evidence for a link between fractional calculus, machine learning, and the renormalization group.
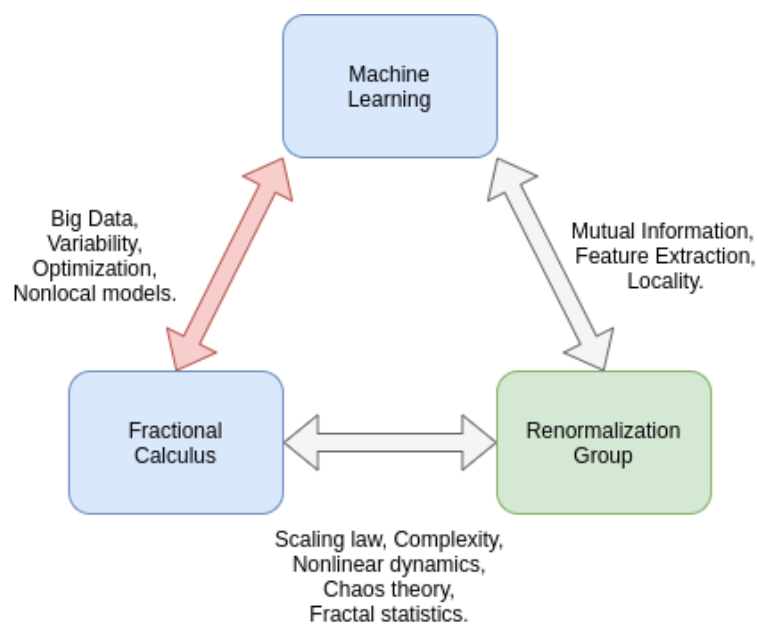
Figure 1. The triangle of machine learning, fractional calculus and the renormalization group as introduced by [85]. The current article focuses on the connection between machine learning and fractional calculus, i.e., the red double-arrow connecting the two blue boxes. This triangle is taken from [85], but we adapted it to better fit this review.

6.4.1. Optimization

We start by discussing the most obvious of the mentioned overlaps/connections, i.e., optimization. Fractional derivatives and, subsequently, fractional calculus is a powerful framework that can enhance optimization tasks, as fractionalized optimization algorithms are less prone to getting stuck in local minima/optima [87]. Thus, it is an upgrade to standard integer-derivative-based optimization tools for optimizing machine learning models' hyperparameters. These optimization tools are also valuable for preprocessing. Here, we want to mention [77,78], these are exemplary applications in which this is used to segment MRI brain scans, later to be applied as an input for a machine learning algorithm to detect strokes.

6.4.2. Variability

Given the results of our literature review, we can link several publications and the corresponding approaches to data variability, thus, providing evidence that fractional calculus and machine learning combined are tools capable of dealing with variability in data.

In [22], for example, the researchers employed fractional derivatives to preprocess Vis-NIR spectral data later to be used in a machine learning task to estimate organic matter content in arid soil. Similar preprocessing approaches are collected in Section 5.1.1. Thus, all of these articles deal with the spatial variability of the obtained spectral data. Furthermore, we can interpret these approaches as coping with climate variability as well, as they take into account or describe the components and interactions of the corresponding climate system.

We also find several publications where fractional derivatives and machine learning can deal with heart rate variability. This means that, in both [47,48], the researchers are using ideas related to fractional derivatives to, e.g., classify arrhythmic ECG signal. This also refers to human variability in general. We also refer to the work by Mucha et al. [50], where the researchers classify Parkinson's disease based on online handwriting data, which were preprocessed using fractional derivatives.

According to the work of Niu et al. [86], variability is the central aspect of big data where employing fractional calculus and machine learning can be beneficial. Thus, we

recommend considering approaches combining fractional derivatives and machine learning when dealing with variability in data and further recommend considering these for big data approaches.As the topic of big data would unnecessarily complicate this discussion, we refer the interested reader to [86,88].

6.4.3. Nonlocal Models

Section 5.2 provides evidence for applying fractional calculus together with machine learning for nonlocal models. This means that all models based on fractional calculus are considered to be nonlocal because of the derivatives' inherent *memory*. Here, we want to point out the work by [55], which provides a machine learning framework for identifying linear space-fractional differential equations from data.

6.4.4. Renormalization Group

The renormalization group refers to a formal apparatus capable of dealing with problems involving many length scales. The basic idea is to tackle the problems in steps, i.e., one step at each length scale. For critical phenomena, the strategy is to carry out statistical averages over thermal fluctuations on all size scales. Thus, the renormalization group approach integrates all the fluctuations in sequence. This means starting on an atomic scale and moving step-by-step to larger scales until fluctuations on all scales are averaged out. The reason for doing so is that theorists have difficulties describing phenomena that involve many coupled degrees of freedom. For example, it takes many variables to characterize a turbulent flow or the state of a fluid near the critical point. Here, Kenneth and Wilson found the self-similarity of dynamic systems near critical points, which can be described using the renormalization group. Thus, the renormalization group is considered a powerful tool for studying the scaling properties of physical phenomena through its treatment of the scaling properties of complex and chaotic dynamics [89,90].

Thus, we want to discuss our findings in a bigger context together with renormalization group methods. We list and discuss evidence from our literature review for the proposed overlaps between fractional derivatives and machine learning with the renormalization group. Thus, can we find evidence for the keywords depicted in Figure 1? This means that the scaling law, complexity, nonlinear dynamics, fractal statistics, mutual information, feature extraction, and locality.

Again, we start our discussion with the most obvious, i.e., feature extraction. Regarding the triangle in Figure 1 [85], the researchers argued that, e.g., deep neural networks and their inherent feature extraction process, i.e., feeding and abstracting the input data from one layer to the next, learn to ignore irrelevant features while keeping relevant ones, which increases the efficiency of the learning process. The researchers further argue that this is similar to renormalization group methods, where only the relevant features of a physical system are extracted by integrating over short-scaled degrees of freedom to describe the system at larger scales. Their assumption is based on the findings of [91], which provides a mapping between the variational renormalization group and deep learning, whereas these findings are based on the Ising model. Furthermore, they discuss that the renormalization group is usually applied to physical systems with many symmetries. Contrary to that, deep learning is typically applied to data with a limited structure. Thus, the connection between the renormalization group and, in this case, deep neural networks, might not be generalizable to machine learning applications at large. Nonetheless, when talking about feature extraction, one can argue that fractional derivatives do just that in, e.g., [47,48]. These articles describe the feature extraction such that new parameters, e.g., the error energy of a signal when being modeled using a fractional transfer function, are found. This means that smaller-scaled features, i.e., the dataset itself, are not fed into the machine learning algorithm. Furthermore, we recommend using the fractional measures of complexity as another fractional calculus-based technique for feature extraction [83].

The researchers [85] discussed locality as the similarity of grouping spins in the black-spin renormalization group. Furthermore, they pointed out that this is similar to the

shallow layers of a deep neural network in which the neurons are locally connected, e.g., without strong bonds, as described in [92]. Though this applies to deep neural networks, we did not find evidence in the reviewed publications. The reason here is that fractional derivatives behave nonlocally, i.e., spatio-temporal fractional derivatives have a certain degree of memory in both space and time.

When dealing with complexity, we need to consider that there is no unified definition for complexity yet. However, the researchers [85,90] argued that the overlap between fractional calculus and renormalization group methods might give us this definition and provide a framework to deal with complexity and thus answer the following two questions:

1. How can we characterize complexity?
2. What method should be used to analyze complexity to better understand real-world phenomena?

We consider finding answers to these questions idealistic. Thus, we need to weaken the statement to "*partially* answer this questions". With much certainty, even when combined, fractional calculus, the renormalization group, and machine learning will only partially answer these questions, primarily because there are many different perspectives on the complexity of, e.g., data. Nonetheless, this triangle of powerful tools might expand definitions and find some definitions and characterizations for particular problems. To give an (in the authors' opinion) reasonable example: this triangle might provide insights into the capability of machine learning to predict multifractal, chaotic, or stochastic time series data by characterizing a datasets memory and which time series can or cannot be learned by different machine learning algorithms.

Nonetheless, did we find evidence for complexity in our list of publications? We indeed found that there are phenomena that are better described using fractional calculus rather than integer-valued calculus, as described in [55]. Further in this work, the researchers used fractional calculus and machine learning to describe, e.g., financial time series data, e.g., the S&P500 index, which is widely accepted to be a complex phenomenon [93–95]. Furthermore, in [96], the researchers used renormalization group methods for analyzing the S&P 500 index. Thus, we conclude that we found evidence for this connection. Additionally, in [93], the researchers also found evidence for two universal scaling models/laws in finance. Thus, we found a connection and a potential research area for all three of the discussed techniques, i.e., the renormalization group, fractional calculus, and machine learning in financial time series analysis. Nonetheless, the discussed work of Gulian et al. [55] deals with linear space fractional differential equations and thus does not provide the framework to deal with nonlinear dynamics.

However, nonlinear dynamics (and subsequently chaos) are discussed in [82], where a fractional bee colony optimization algorithm is used to find the optimal parameters for a model which describes the nonlinear behavior of nonperforming loans. We also need to mention the work performed by Wang et al. [59], where fractional models for finance are parametrized using Gaussian process regression and predicted using recurrent neural networks. Thus, given this evidence, we suggest similar approaches to employ the renormalization group methodology in future research.

However, we did not find the applications of mutual information in this review. In [85], the researchers referred to the work by Koch-Janusz et al. [97] for the applicability of mutual information for both machine learning and the renormalization group. The researchers present a machine learning procedure to perform renormalization group tasks, thus, the reported connection. Here, the mutual information is used as an objective or target function where one of the employed neural networks aims to maximize the mutual information between two random distributions to deal with the reduction in degrees of freedom as the renormalization group would do. However, we did not find applications of mutual information for fractional derivatives and machine learning in our literature review, so we recommend future research to discuss the topic. For example, to what degree can machine learning algorithms be used to perform renormalization group tasks? Furthermore, are there non-model applications where these ideas can be used?

Furthermore, we did not find approaches dealing with scaling laws in the featured list of publications. Scaling laws, also known as power laws, are a description of the scale invariance of natural phenomena. Some notable examples according to [90] are, e.g., Zipf's inverse power law which describes the relative frequency of word usage within a language [98].

Furthermore, finally, we also did not find an application of fractal statistics or fractals. However, given the numerous evidence for fractal behavior and statistics in various fields, e.g., in hydrology [99], rock mechanics [100] or finance [101], we still aim to provide a link. Here, we consider the previous discussion on the potential applicability of the triangle of machine learning, fractional calculus, and renormalization group on financial data. Again, we found evidence for fractal statistics present in economic data, i.e., the S&P 500 index [101]. Thus, we recommend future research for analyzing stock market data to employ ideas from fractional calculus, machine learning, and the renormalization group and subsequently to analyze the fractal behavior of these datasets. Furthermore, we believe that these ideas might be capable of analyzing, learning, and predicting a variety of complex real-life datasets, e.g., environmental data.

6.5. Problems

The last part of this discussion is to name two problems we encountered whilst conducting this literature review. This section should serve as a guideline on what to avoid and how to improve the future research field.

Sometimes the fractional derivative is not discussed sufficiently: we found literature that used fractional derivatives without stating the exact approach, e.g., [40]. However, as stated and referenced in [34], we assumed that the Grünwald–Letnikov fractional derivative is the employed fractional derivative.

We thus urge researchers to be specific about the employed fractional derivative, the corresponding discretization, and/or the employed software package to improve the transparency and reproducibility of their work. We further point out that hardly any article discussing the combined applications of machine learning and fractional derivatives listed in this review shows the employed discretization.

Incorrect usage of keywords: Some researchers assigned the term *fractional calculus* and/or *fractional derivatives* to their articles and describe the fractional calculus framework. However, in the end, they used a curve-fitting approach with fractional polynomial exponents and no fractional derivatives and/or calculus.

7. Conclusions

This review analyzes the combined applications of fractional derivatives and supervised non-neural network approaches. Our research shows that supervised machine learning and fractional derivatives are valuable tools that can be combined to, e.g., improve a machine learning algorithm's accuracy. We found a total of three types of combined applications, i.e., preprocessing, modeling fractional dynamics via machine learning, and optimization.

We further found that most preprocessing applications are spectroscopical applications where a fractional derivative of a specific order is applied to the spectral data to improve the accuracy of the employed algorithm.

For optimization, we found two categories, gradient-based and gradient-free optimization algorithms. For gradient-based optimization, one can replace an integer-order gradient with a fractional-order gradient, thus obtaining a fractional gradient descent optimization algorithm. For the gradient-free optimization, the velocity in, e.g., a particle swarm algorithm, can be replaced by a fractional-derivative-based velocity. Hereby, one obtains a fractional order particle swarm optimization algorithm.

For the third type of combined applications, i.e., using a machine learning to model fractional dynamics, we found that machine learning and fractional derivatives, and subsequently fractional calculus are both techniques that deal with complex real-life phenomena,

addressing the problem of complexity from different angles. Furthermore, this field improves kernel-based machine learning algorithms, such as Gaussian process regression and support vector machines, by providing kernels based on fractional derivatives to introduce memory into the modeling approach.

When bringing all our results together, we find a third tool conceptually linked to both, i.e., the renormalization group, which is a powerful approach for dealing with scaling laws, complex phenomena, complex dynamics, and chaos [102]. Thus, these three techniques, machine learning fractional calculus, and the renormalization group, form a triangle as proposed by Niu et al. [85]. Our review provides evidence for the proposed keywords, the corresponding connections, and thus the established links between machine learning, fractional calculus and the renormalization group.

In the author's opinion, it will still take some time until the renormalization group will be part of mainstream computer science in combination with machine learning and fractional calculus. However, it is important to make today's researchers, especially machine learning practitioners, aware and familiar with these topics as (supervised) machine learning is becoming popular in science in general. This means that, wherever there are available data, scientists are eager to find modern machine learning approaches to analyze, describe, and predict complex phenomena directly from data. Thus, machine learning is changing science overall [103,104].

Furthermore, as pointed out in [105], machine learning is sometimes considered an oracle that can learn any task from observations, i.e., predict chemical reactions and particle physics experiments, but not necessarily giving the practitioner scientist a deeper understanding of the process at hand. Thus, one might ask how and whether machine learning can contribute to our scientific understanding. The researchers further point out the multidisciplinary aspect of machine learning in general and that it will, because of this, undoubtedly keep increasing in popularity and can increase our scientific understanding of the studied processes. Thus, we propose that fractional calculus and the renormalization group will provide powerful tools in this context. Together with machine learning, these tools are able to bridge gaps between separated disciplines by improving numerical analysis, providing insights into complex phenomena, and by unveiling the hidden (or macroscopic; in the case of the renormalization group) mechanics governing these phenomena.

Regarding future research combining fractional derivatives and machine learning, we recommend testing whether neural networks can perform any kind of data manipulation that fractional derivatives can perform. This might be useful as it might avoid pointless contributions on the topic, i.e., fractional-derivative-based preprocessing for neural networks.

Furthermore, future research combining fractional derivatives, the renormalization group, and machine learning might give valuable insights into the learnability of machine learning algorithms and the corresponding underlying dynamics of the data. Specifically, one could connect ideas from the renormalization to a fractional derivative preprocessing while analyzing the performance of the employed machine learning algorithm. This means that finding the sweet spot of mandatory degrees of freedom and the corresponding order of fractional derivatives might provide insights about the memory of the studied data/system. Thus, providing the "minimal" representation of the data/system by optimal performance.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A. Summary Table

Publication	Employed Fractional Derivative	Employed Machine Learning Algorithms	Research Area	Category	Comment
[22]	Grünwald–Letnikov	Random Forest, PLSR	Agriculture	Preprocessing	Vis-NIRSpectroscopy, Organic Matter Content
[24]	Grünwald–Letnikov	PLSR	Agriculture	Preprocessing	Vis-NIRSpectroscopy, Nitrogen Concentration
[31]	Grünwald–Letnikov	Random Forest, PLSR	Agriculture	Preprocessing	Vis-NIRSpectroscopy, Soil Salinity
[28]	Grünwald–Letnikov	Random Forest	Agriculture	Preprocessing	Vis-NIRSpectroscopy, Lead and Zinc Concentration
[23]	Grünwald–Letnikov	Memory Based Learning, PLSR, Random Forest	Agriculture	Preprocessing	Vis-NIRSpectroscopy, Soil Organic Matter
[26]	Grünwald–Letnikov	SVM	Agriculture	Preprocessing	Vis-NIRSpectroscopy, Total Nitrogen Content
[27]	Grünwald–Letnikov	SVM, Random Forest PLSR, ELM	Agriculture	Preprocessing	Vis-NIRSpectroscopy, Leaf Chlorophyll Concentration
[32]	Grünwald–Letnikov	ELM	Agriculture	Preprocessing	Vis-NIRSpectroscopy, Soil Salt, Water-Soluble Ions
[25]	Grünwald–Letnikov	SVM, ELM CNN, PLSR	Agriculture	Preprocessing	Vis-NIRSpectroscopy, Nitrogen Content
[29]	Grünwald–Letnikov	PLSR, GRNN	Agriculture	Preprocessing	Vis-NIRSpectroscopy, Soil Heavy Metal Estimation
[30]	Grünwald–Letnikov	SVM, Ridge Regression, PLSR, Random ForestXGBoost, ELM	Agriculture	Preprocessing	Vis-NIRSpectroscopy, Soil Heavy Metal Estimation
[31]	Grünwald–Letnikov	PLSR	Agriculture	Preprocessing	HyperspectralSpectroscopy, Salt Content
[41]	Grünwald–Letnikov	PLSR	Agriculture	Preprocessing	HyperspectralSpectroscopy, Soil Clay Content
[33]	Grünwald–Letnikov	PLSR	Agriculture	Preprocessing	HyperspectralSpectroscopy, Soil Organic Matter
[35]	Grünwald–Letnikov	SVM	Agriculture	Preprocessing	HyperspectralSpectroscopy, Soil Salt Content
[34]	Grünwald–Letnikov	Random Forest	Agriculture	Preprocessing	HyperspectralSpectroscopy, Top Soil Organic Carbon
[37]	Grünwald–Letnikov	Random Forest, SVM	Agriculture	Preprocessing	HyperspectralSpectroscopy, Canopy Nitrogen Content
[36]	Grünwald–Letnikov	PLSR	Agriculture	Preprocessing	HyperspectralSpectroscopy, Soil Total Salt Content
[42]	Grünwald–Letnikov	Random Forest, XGBoost	Environmental	Preprocessing	HyperspectralSpectroscopy, Total Nitrogen in Water
[39]	Grünwald–Letnikov	XGBoost	Agriculture	Preprocessing	HyperspectralSpectroscopy, Soil Moisture Content
[38]	Grünwald–Letnikov	SVM, KNN	Agriculture	Preprocessing	HyperspectralSpectroscopy, Leaf Photosynthetic Pigment Content
[43]	Grünwald–Letnikov	Random Forest	Environmental	Preprocessing	HyperspectralSpectroscopy, Total Suspended Matter in Water
[40]	Grünwald–Letnikov	XGBoost, LightGBM, Random Forest, Extremely Randomized Trees, Decision Trees, Ridge Regression	Agriculture	Preprocessing	HyperspectralSpectroscopy, Soil Electrical Conductivity
[44]	Fractional Order Modeling	SVM	Biomedical Application	Preprocessing	EEG Signal Analysis, Seizure (Ictal) Detection
[45]	Fractional Order Model=ling	SVM	Biomedical Application	Preprocessing	EEG Signal Analysis, Seizure(Ictal) Detection
[47]	Fractional Order Modeling	KNN	Biomedical Application	Preprocessing	ECG Signal Analysis, Arrhythmia Detection
[48]	Fractional Order Modeling	KNN	Biomedical Application	Preprocessing	ECG Signal Analysis, Arrhythmia Detection
[51]	Discrete Fractional Mask	SVM, RF, Simple Cart, J48	Biomedical Application	Preprocessing	CT Image Signal Analysis, Liver Tumor Detection
[49]	Fractional Order Modeling	KNN	Biomedical Application	Preprocessing	Respiratory Impedance Analysis, Disease Detection
[50]	Grünwald– Letnikov	RN	Biomedical Application	Preprocessing	Handwriting Analysis, Parkinson Disease Detection
[46]	Grünwald– Letnikov	KNN	Biomedical Application	Preprocessing	EEG Signal Analysis, Abnormality Detection
[106]	Discrete Fractional Derivative	RF, XGBoost	Biomedical Application	Preprocessing	Hemoglobin Analysis, Diabetes Detection
[107]	Discrete Fractional Derivative	SVM, Naive Bayes, Random Forest AdaBoost, Bagging	Biomedical Application	Preprocessing	Glucose and Insulin Analysis, Diabetes Detection
[52]	Grünwald– Letnikov	ELM, MLSR	Engineering	Preprocessing	Solar Panel Analysis, Defective Solder Joint Detection
[53]	Reisz	ELM, MLSR	Engineering	Preprocessing	Photovoltaic Panel Temperature Monitoring, Peak Detection
[54]	Riemann–Liouville	RF	Engineering	Preprocessing	Activity Recognition of Construction Equipment
[55]	Fractional Differential Equations	Gaussian Process Regression	Physics	Fractional Dynamics	Modeling of Fractional Dynamics from Data
[56]	Fractional Differential Equations, Caputo Derivative	K-Means, Linear Regression, Sparse Regression	Physics	Fractional Dynamics	Modeling Damping from Data
[57]	Fractional Differential Equations, Caputo Derivative	SVM	Applied Mathematics	Fractional Dynamics	Solving Fractional Differential Equations
[58]	Fractional Differential Equations, Caputo Derivative	Ridge Regression	Applied Mathematics	Fractional Dynamics	Modeling Boundaries of Fractional Differential Equations
[59]	Fractional Differential Equations, , Caputo Derivative	Gaussian Process Regression, RNN	Finance & Economy	Fractional Dynamics	Modeling a Fractional Order Financial System
[60]	Fractional Differential Equations, Grünwald–Letnikov Derivative	RF	Engineering	Fractional Dynamics	Modeling and Analyzing the Fractional Behavior of Batteries
[61]	Caputo	Ridge Regression	Regression in General	Optimization	Fractional Gradient Descent on Testing Environment
[63]	Caputo	Logistic Regression	Regression in General	Optimization	Fractional Gradient Descent on Testing Environment
[64]	Caputo	SVM	Classification in General	Optimization	Fractional Gradient Descent Rainfall Data
[65]	Caputo	SVM	Classification in General	Optimization	Fractional Gradient Descent Iris and Rainfall Data
[66]	Caputo	Fractional Linear Regression	Regression Analysis in Finance	Optimization	Multi Linear Regression with Fractional Derivatives Applied to the Romanian GDP
[67]	Caputo	Fractional Linear Regression, Fractional Quadratic Regression	Regression Analysis in General	Optimization	Multi Regression with Fractional Derivatives
[69]	Grünwald–Letnikov	XGBoost	Biomedical Application	Optimization	Classification of Heart Diseases via XGBoost and fractional PSO

Publication	Employed Fractional Derivative	Employed Machine Learning Algorithms	Research Area	Category	Comment
[70]	Grünwald–Letnikov	SVM, K-Means	Biomedical Application	Optimization	Optimizing K-Means and SVM via Improved Fractional PSO
[73]	Grünwald–Letnikov	Random Forest	Spectroscopy	Optimization, Preprocessing	Optimizing a RF Classification for Spectral Bands
[74]	Grünwald–Letnikov	SVM	Spectroscopy	Optimization, Preprocessing	Multilevel Image Segmentation for SVM Classification
[75]	Grünwald–Letnikov	ELM, SVM, Relief,	Regression in General	Optimization, Preprocessing	Optimizing Features for a range of Algorithms and Datasets using FODPSO
[76]	Grünwald–Letnikov	Decision Trees	Biomedical Application	Optimization, Preprocessing	Optimizing Liver cancer detection from CT scans using Decision Trees and FODPSO
[77]	Grünwald–Letnikov	SVM, RF	Biomedical Application	Optimization, Preprocessing	Optimizing Stroke Detection from MRI Brain Scans Using Random Forest and FODPSO
[78]	Grünwald–Letnikov	SVM, RF	Biomedical Application	Optimization, Preprocessing	Optimizing Stroke Detection from MRI Brain Scans Using Random Forest and FODPSO
[79]	Grünwald–Letnikov	Naive Bayes Hoeffding Tree Classifier	Biomedical Application	Optimization, Preprocessing	Optimizing Foot Ulcer Detection from MRI Brain Scans Using Random Forest and FODPSO
[80]	Grünwald–Letnikov	Fuzzy C-Means	Biomedical Application	Optimization, Preprocessing	Optimizing brain tumor detection from Brain MRI Scans Using Fuzzy C-Means and FODPSO
[81]	Grünwald–Letnikov	Decision Trees	Water Management	Optimization	Predicting Water Quality Using Decision Trees and Fractional Artificial Bee Colony Optimization
[82]	Grünwald–Letnikov	Nonlinear Regression	Finance & Economy	Optimization	Predicting Non-Performing loans Using FOABCO and Nonlinear Regression

References

1. West, B.J. *Tomorrow's Science*; CRC Press: Boca Raton, FL, USA, 2015. [CrossRef]
2. Sun, H.; Zhang, Y.; Baleanu, D.; Chen, W.; Chen, Y. A new collection of real world applications of fractional calculus in science and engineering. *Commun. Nonlinear Sci. Numer. Simul.* **2018**, *64*, 213–231. [CrossRef]
3. Zhang, Y.; Sun, H.; Stowell, H.H.; Zayernouri, M.; Hansen, S.E. A review of applications of fractional calculus in Earth system dynamics. *Chaos Solitons Fractals* **2017**, *102*, 29–46. [CrossRef]
4. Du, M.; Wang, Z.; Hu, H. Measuring memory with the order of fractional derivative. *Sci. Rep.* **2013**, *3*, 3431. [CrossRef] [PubMed]
5. Viera-Martin, E.; Gómez-Aguilar, J.F.; Solís-Pérez, J.E.; Hernández-Pérez, J.A.; Escobar-Jiménez, R.F. Artificial neural networks: a practical review of applications involving fractional calculus. *Eur. Phys. J. Spec. Top.* **2022**, *231*, 2059–2095. [CrossRef]
6. Khan, S.; Ahmad, J.; Naseem, I.; Moinuddin, M. A Novel Fractional Gradient-Based Learning Algorithm for Recurrent Neural Networks. *Circuits Syst. Signal Process.* **2018**, *37*, 593–612. [CrossRef]
7. Aslipour, Z.; Yazdizadeh, A. Identification of wind turbine using fractional order dynamic neural network and optimization algorithm. *Int. J. Eng.* **2020**, *33*, 277–284.
8. Unity Technologies. *AI and Machine Learning, Explained*; Unity Technologies: San Francisco, CA, USA, 2022.
9. Google Developers. *Machine Learning Glossary*; 2022.
10. Wikipedia. Fractional Calculus—Wikipedia, The Free Encyclopedia. 2022. Available online: http://en.wikipedia.org/w/index.php?title=Fractional%20calculus&oldid=1124332647 (accessed on 6 December 2022).
11. de Oliveira, E.C.; Tenreiro Machado, J.A. A Review of Definitions for Fractional Derivatives and Integral. *Math. Probl. Eng.* **2014**, *2014*, 238459. [CrossRef]
12. Aslan, İ. An analytic approach to a class of fractional differential-difference equations of rational type via symbolic computation. *Math. Methods Appl. Sci.* **2015**, *38*, 27–36. [CrossRef]
13. Sagayaraj, M.; Selvam, A.G. Discrete Fractional Calculus: Definitions and Applications. *Int. J. Pure Eng. Math.* **2014**, *2*, 93–102.
14. Goodrich, C.; Peterson, A.C. *Discrete Fractional Calculus*; Springer International Publishing: Cham, Switzerland, 2015. [CrossRef]
15. Artin, E. *The Gamma Function*, dover ed.; Dover Books on Mathematics; Dover Publications, Inc.: Mineola, NY, USA, 2015.
16. Mohri, M.; Rostamizadeh, A.; Talwalkar, A. *Foundations of Machine Learning*; The MIT Press: Cambridge, MA, USA, 2012.
17. Bzdok, D.; Krzywinski, M.; Altman, N. Machine learning: Supervised methods. *Nat. Methods* **2018**, *15*, 5–6. [CrossRef]
18. Singh, A.; Thakur, N.; Sharma, A. A review of supervised machine learning algorithms. In Proceedings of the 2016 3rd International Conference on Computing for Sustainable Global Development (INDIACom), New Delhi, India, 16–18 March 2016; pp. 1310–1315.
19. Brownlee, J. *Basics for Linear Algebra for Machine Learning—Discover the Mathematical Language of Data in Python*, 1.1 ed.; ZSCC: NoCitationData[s0]; Jason Brownlee, Machine Learning Mastery: Melbourne, Australia, 2018.
20. Brownlee, J. *Master Machine Learning Algorithms*, ebook ed.; Machine Learning Mastery: Melbourne, Australia, 2016.
21. Brownlee, J. *Machine Learning Mastery with Python*, 1st ed.; Machine Learning Mastery: Melbourne, Australia, 2016.
22. Wang, J.; Tiyip, T.; Ding, J.; Zhang, D.; Liu, W.; Wang, F. Quantitative Estimation of Organic Matter Content in Arid Soil Using Vis-NIR Spectroscopy Preprocessed by Fractional Derivative. *J. Spectrosc.* **2017**, *2017*, 1375158. [CrossRef]
23. Hong, Y.; Chen, S.; Liu, Y.; Zhang, Y.; Yu, L.; Chen, Y.; Liu, Y.; Cheng, H.; Liu, Y. Combination of fractional order derivative and memory-based learning algorithm to improve the estimation accuracy of soil organic matter by visible and near-infrared spectroscopy. *Catena* **2019**, *174*, 104–116. [CrossRef]
24. Chen, K.; Li, C.; Tang, R. Estimation of the nitrogen concentration of rubber tree using fractional calculus augmented NIR spectra. *Ind. Crop. Prod.* **2017**, *108*, 831–839. [CrossRef]

25. Hu, W.; Hu, W.; Tang, R.; Li, C.; Zhou, T.; Chen, J.; Chen, K.; Chen, K. Fractional order modeling and recognition of nitrogen content level of rubber tree foliage. *J. Near Infrared Spectrosc.* **2021**, *29*, 42–52. [CrossRef]

26. Abulaiti, Y.; Sawut, M.; Maimaitiaili, B.; Chunyue, M. A possible fractional order derivative and optimized spectral indices for assessing total nitrogen content in cotton. *Comput. Electron. Agric.* **2020**, *171*, 105275. [CrossRef]

27. Bhadra, S.; Sagan, V.; Maimaitijiang, M.; Maimaitiyiming, M.; Newcomb, M.; Shakoor, N.; Mockler, T.C. Quantifying Leaf Chlorophyll Concentration of Sorghum from Hyperspectral Data Using Derivative Calculus and Machine Learning. *Remote Sens.* **2020**, *12*, 2082. [CrossRef]

28. Hong, Y.; Shen, R.; Cheng, H.; Chen, Y.; Zhang, Y.; Liu, Y.; Zhou, M.; Yu, L.; Liu, Y.; Liu, Y. Estimating lead and zinc concentrations in peri-urban agricultural soils through reflectance spectroscopy: Effects of fractional-order derivative and random forest. *Sci. Total Environ.* **2019**, *651*, 1969–1982. [CrossRef]

29. Xu, X.; Chen, S.; Ren, L.; Han, C.; Lv, D.; Zhang, Y.; Ai, F. Estimation of Heavy Metals in Agricultural Soils Using Vis-NIR Spectroscopy with Fractional-Order Derivative and Generalized Regression Neural Network. *Remote Sens.* **2021**, *13*, 2718. [CrossRef]

30. Chen, L.; Lai, J.; Tan, K.; Wang, X.; Chen, Y.; Ding, J. Development of a soil heavy metal estimation method based on a spectral index: Combining fractional-order derivative pretreatment and the absorption mechanism. *Sci. Total Environ.* **2022**, *813*, 151882. [CrossRef]

31. Zhang, D.; Tiyip, T.; Ding, J.; Zhang, F.; Nurmemet, I.; Kelimu, A.; Wang, J. Quantitative Estimating Salt Content of Saline Soil Using Laboratory Hyperspectral Data Treated by Fractional Derivative. *J. Spectrosc.* **2016**, *2016*, 1081674. [CrossRef]

32. Lao, C.; Chen, J.; Zhang, Z.; Chen, Y.; Ma, Y.; Chen, H.; Gu, X.; Ning, J.; Jin, J.; Li, X. Predicting the contents of soil salt and major water-soluble ions with fractional-order derivative spectral indices and variable selection. *Comput. Electron. Agric.* **2021**, *182*, 106031. [CrossRef]

33. Xu, X.; Chen, S.; Xu, Z.; Yu, Y.; Zhang, S.; Dai, R. Exploring Appropriate Preprocessing Techniques for Hyperspectral Soil Organic Matter Content Estimation in Black Soil Area. *Remote Sens.* **2020**, *12*, 3765. [CrossRef]

34. Hong, Y.; Guo, L.; Chen, S.; Linderman, M.; Mouazen, A.M.; Yu, L.; Chen, Y.; Liu, Y.; Liu, Y.; Cheng, H.; et al. Exploring the potential of airborne hyperspectral image for estimating topsoil organic carbon: Effects of fractional-order derivative and optimal band combination algorithm. *Geoderma* **2020**, *365*, 114228. [CrossRef]

35. Wang, Z.; Zhang, X.; Zhang, F.; weng Chan, N.; Liu, S.; Deng, L. Estimation of soil salt content using machine learning techniques based on remote-sensing fractional derivatives, a case study in the Ebinur Lake Wetland National Nature Reserve, Northwest China—ScienceDirect. *Ecol. Indic.* **2020**, *119*, 106869. [CrossRef]

36. Tian, A.; Zhao, J.; Tang, B.; Zhu, D.; Fu, C.; Xiong, H. Hyperspectral Prediction of Soil Total Salt Content by Different Disturbance Degree under a Fractional-Order Differential Model with Differing Spectral Transformations. *Remote Sens.* **2021**, *13*, 4283. [CrossRef]

37. Peng, Y.; Zhu, X.; Xiong, J.; Yu, R.; Liu, T.; Jiang, Y.; Yang, G. Estimation of Nitrogen Content on Apple Tree Canopy through Red-Edge Parameters from Fractional-Order Differential Operators using Hyperspectral Reflectance. *J. Indian Soc. Remote Sens.* **2021**, *49*, 377–392. [CrossRef]

38. Cheng, J.; Yang, G.; Xu, W.; Feng, H.; Han, S.; Liu, M.; Zhao, F.; Zhu, Y.; Zhao, Y.; Wu, B.; et al. Improving the Estimation of Apple Leaf Photosynthetic Pigment Content Using Fractional Derivatives and Machine Learning. *Agronomy* **2022**, *12*, 1497. [CrossRef]

39. Ge, X.; Ding, J.; Jin, X.; Wang, J.; Chen, X.; Li, X.; Liu, J.; Xie, B. Estimating Agricultural Soil Moisture Content through UAV-Based Hyperspectral Images in the Arid Region. *Remote Sens.* **2021**, *13*, 1562. [CrossRef]

40. Jia, P.; Zhang, J.; He, W.; Hu, Y.; Zeng, R.; Zamanian, K.; Jia, K.; Zhao, X. Combination of Hyperspectral and Machine Learning to Invert Soil Electrical Conductivity. *Remote Sens.* **2022**, *14*, 2602. [CrossRef]

41. Wang, J.; Tiyip, T.; Ding, J.; Zhang, D.; Liu, W.; Wang, F.; Tashpolat, N. Desert soil clay content estimation using reflectance spectroscopy preprocessed by fractional derivative. *PLoS ONE* **2017**, *12*, e0184836. [CrossRef]

42. Liu, J.; Ding, J.; Ge, X.; Wang, J. Evaluation of Total Nitrogen in Water via Airborne Hyperspectral Data: Potential of Fractional Order Discretization Algorithm and Discrete Wavelet Transform Analysis. *Remote Sens.* **2021**, *13*, 4643. [CrossRef]

43. Wang, X.; Song, K.; Liu, G.; Wen, Z.; Shang, Y.; Du, J. Development of total suspended matter prediction in waters using fractional-order derivative spectra. *J. Environ. Manag.* **2022**, *302*, 113958. [CrossRef] [PubMed]

44. Joshi, V.; Pachori, R.B.; Vijesh, A. Classification of ictal and seizure-free EEG signals using fractional linear prediction. *Biomed. Signal Process. Control* **2014**, *9*, 1–5. [CrossRef]

45. Aaruni, V.C.; Harsha, A.; Joseph, L.A. Classification of EEG signals using fractional calculus and wavelet support vector machine. In Proceedings of the 2015 IEEE International Conference on Signal Processing, Informatics, Communication and Energy Systems (SPICES), Kozhikode, India, 19–21 February 2015; pp. 1–5. [CrossRef]

46. Dhar, P.; Malakar, P.; Ghosh, D.; Roy, P.; Das, S. Fractional Linear Prediction Technique for EEG signals classification. In Proceedings of the 2019 International Conference on Intelligent Computing and Control Systems (ICCS), Madurai, India, 15–17 May 2019; pp. 261–265. [CrossRef]

47. Assadi, I.; Charef, A.; Belgacem, N.; Nait-Ali, A.; Bensouici, T. QRS complex based human identification. In Proceedings of the 2015 IEEE International Conference on Signal and Image Processing Applications (ICSIPA), Kuala Lumpur, Malaysia, 19–21 October 2015; pp. 248–252. [CrossRef]

48. Assadi, I.; Charef, A.; Bensouici, T.; Belgacem, N. Arrhythmias discrimination based on fractional order system and KNN classifier. In Proceedings of the 2nd IET International Conference on Intelligent Signal Processing 2015 (ISP), London, UK, 1–2 December 2015; p. 6. [CrossRef]

49. Assadi, I.; Charef, A.; Copot, D.; De Keyser, R.; Bensouici, T.; Ionescu, C. Evaluation of respiratory properties by means of fractional order models. *Biomed. Signal Process. Control* **2017**, *34*, 206–213. [CrossRef]

50. Mucha, J.; Mekyska, J.; Faundez-Zanuy, M.; Lopez-De-Ipina, K.; Zvoncak, V.; Galaz, Z.; Kiska, T.; Smekal, Z.; Brabenec, L.; Rektorova, I. Advanced Parkinson's Disease Dysgraphia Analysis Based on Fractional Derivatives of Online Handwriting. In Proceedings of the 2018 10th International Congress on Ultra Modern Telecommunications and Control Systems and Workshops (ICUMT), Moscow, Russia, 5–9 November 2018; pp. 1–6. [CrossRef]

51. Ghatwary, N.; Ahmed, A.; Jalab, H. Liver CT enhancement using Fractional Differentiation and Integration. In Proceedings of the World Congress on Engineering 2016 (WCE 2016), London, UK, 29 June–1 July 2016.

52. Liu, H.; Zhang, C.; Huang, D. Extreme Learning Machine and Moving Least Square Regression Based Solar Panel Vision Inspection. *J. Electr. Comput. Eng.* **2017**, *2017*, 7406568. [CrossRef]

53. Dhanalakshmi, S.; Nandini, P.; Rakshit, S.; Rawat, P.; Narayanamoorthi, R.; Kumar, R.; Senthil, R. Fiber Bragg grating sensor-based temperature monitoring of solar photovoltaic panels using machine learning algorithms. *Opt. Fiber Technol.* **2022**, *69*, 102831. [CrossRef]

54. Langroodi, A.K.; Vahdatikhaki, F.; Doree, A. Activity recognition of construction equipment using fractional random forest. *Autom. Constr.* **2021**, *122*, 103465. [CrossRef]

55. Gulian, M.; Raissi, M.; Perdikaris, P.; Karniadakis, G. Machine Learning of Space-Fractional Differential Equations. *SIAM J. Sci. Comput.* **2019**. [CrossRef]

56. Guo, J.; Wang, L.; Fukuda, I.; Ikago, K. Data-driven modeling of general damping systems by k-means clustering and two-stage regression. *Mech. Syst. Signal Process.* **2022**, *167*, 108572. [CrossRef]

57. Parand, K.; Aghaei, A.A.; Jani, M.; Ghodsi, A. Parallel LS-SVM for the numerical simulation of fractional Volterra's population model. *Alex. Eng. J.* **2021**, *60*, 5637–5647. [CrossRef]

58. Guan, X.; Zhang, Q.; Tang, S. Numerical boundary treatment for shock propagation in the fractional KdV-Burgers equation. *Comput. Mech.* **2022**, *69*, 201–212. [CrossRef]

59. Wang, B.; Liu, J.; Alassafi, M.O.; Alsaadi, F.E.; Jahanshahi, H.; Bekiros, S. Intelligent parameter identification and prediction of variable time fractional derivative and application in a symmetric chaotic financial system. *Chaos Solitons Fractals* **2022**, *154*, 111590. [CrossRef]

60. Yang, R.; Xiong, R.; He, H.; Chen, Z. A fractional-order model-based battery external short circuit fault diagnosis approach for all-climate electric vehicles application. *J. Clean. Prod.* **2018**, *187*, 950–959. [CrossRef]

61. Huang, F.; Li, D.; Xu, J.; Wu, Y.; Xing, Y.; Yang, Z. Ridge Regression Based on Gradient Descent Method with Memory Dependent Derivative. In Proceedings of the 2020 IEEE 11th International Conference on Software Engineering and Service Science (ICSESS), Beijing, China, 16–18 October 2020; pp. 463–467. [CrossRef]

62. Li, D.; Zhu, D. An affine scaling interior trust-region method combining with nonmonotone line search filter technique for linear inequality constrained minimization. *Int. J. Comput. Math.* **2018**, *95*, 1494–1526.

63. Wang, Y.; Li, D.; Xu, X.; Jia, Q.; Yang, Z.; Nai, W.; Sun, Y. Logistic Regression with Variable Fractional Gradient Descent Method. In Proceedings of the 2020 IEEE 9th Joint International Information Technology and Artificial Intelligence Conference (ITAIC), Chongqing, China, 11–13 December 2020; Volume 9, pp. 1925–1928. [CrossRef]

64. Hapsari, D.P.; Utoyo, I.; Purnami, S.W. Fractional Gradient Descent Optimizer for Linear Classifier Support Vector Machine. In Proceedings of the 2020 Third International Conference on Vocational Education and Electrical Engineering (ICVEE), Surabaya, Indonesia, 3–4 October 2020; pp. 1–5. [CrossRef]

65. Hapsari, D.P.; Utoyo, I.; Purnami, S.W. Support Vector Machine optimization with fractional gradient descent for data classification. *J. Appl. Sci. Manag. Eng. Technol.* **2021**, *2*, 1–6. [CrossRef]

66. Badík, A.; Feĥan, M. Applying fractional calculus to analyze final consumption and gross investment influence on GDP. *J. Appl. Math. Stat. Inform.* **2021**, *17*, 65–72. [CrossRef]

67. Awadalla, M.; Noupoue, Y.Y.Y.; Tandogdu, Y.; Abuasbeh, K. Regression Coefficient Derivation via Fractional Calculus Framework. *J. Math.* **2022**, *2022*, 1144296. [CrossRef]

68. Couceiro, M.; Ghamisi, P. *Fractional Order Darwinian Particle Swarm Optimization*; Springer International Publishing: Berlin/Heidelberg, Germany, 2016. [CrossRef]

69. Chou, F.I.; Huang, T.H.; Yang, P.Y.; Lin, C.H.; Lin, T.C.; Ho, W.H.; Chou, J.H. Controllability of Fractional-Order Particle Swarm Optimizer and Its Application in the Classification of Heart Disease. *Appl. Sci.* **2021**, *11*, 11517. [CrossRef]

70. Li, J.; Zhao, C. Improvement and Application of Fractional Particle Swarm Optimization Algorithm. *Math. Probl. Eng.* **2022**, *2022*, 5885235. [CrossRef]

71. Tillett, J.; Rao, T.; Sahin, F.; Rao, R. Darwinian Particle Swarm Optimization. In Proceedings of the Indian International Conference on Artificial Intelligence, Pune, India, 20–22 December 2005; pp. 1474–1487.

72. Couceiro, M.S.; Rocha, R.P.; Ferreira, N.M.F.; Machado, J.A.T. Introducing the fractional-order Darwinian PSO. *Signal Image Video Process.* **2012**, *6*, 343–350. [CrossRef]

73. Ghamisi, P.; Couceiro, M.S.; Benediktsson, J.A. Classification of hyperspectral images with binary fractional order Darwinian PSO and random forests. In *Proceedings of the Image and Signal Processing for Remote Sensing XIX*; Bruzzone, L., Ed.; International Society for Optics and Photonics, SPIE: San Francisco, CA, USA, 2013; Volume 8892, p. 88920S. [CrossRef]

74. Ghamisi, P.; Couceiro, M.S.; Martins, F.M.L.; Benediktsson, J.A. Multilevel Image Segmentation Based on Fractional-Order Darwinian Particle Swarm Optimization. *IEEE Trans. Geosci. Remote Sens.* **2014**, *52*, 2382–2394. [CrossRef]

75. Wang, Y.Y.; Zhang, H.; Qiu, C.H.; Xia, S.R. A Novel Feature Selection Method Based on Extreme Learning Machine and Fractional-Order Darwinian PSO. *Comput. Intell. Neurosci.* **2018**, *2018*, 5078268. [CrossRef] [PubMed]

76. Das, A.; Panda, S.S.; Sabut, S. Detection of Liver Cancer using Optimized Techniques in CT Scan Images. In Proceedings of the 2018 International Conference on Applied Electromagnetics, Signal Processing and Communication (AESPC), Bhubaneswar, India, 22–24 October 2018; Volume 1, pp. 1–5. [CrossRef]

77. Subudhi, A.; Acharya, U.R.; Dash, M.; Jena, S.; Sabut, S. Automated approach for detection of ischemic stroke using Delaunay Triangulation in brain MRI images. *Comput. Biol. Med.* **2018**, *103*, 116–129. [CrossRef] [PubMed]

78. Subudhi, A.; Dash, M.; Sabut, S. Automated segmentation and classification of brain stroke using expectation-maximization and random forest classifier. *Biocybern. Biomed. Eng.* **2020**, *40*, 277–289. [CrossRef]

79. Naveen, J.; Selvam, S.; Selvam, B. FO-DPSO Algorithm for Segmentation and Detection of Diabetic Mellitus for Ulcers. *Int. J. Image Graph.* **2022**, 2240011. [CrossRef]

80. Nalini, U.; Rani, D.N.U. NOVEL BRAIN TUMOR SEGMENTATION USING FUZZY C-MEANS WITH FRACTIONAL ORDER DARWINIAN PARTICLE SWARM OPTIMIZATION. *Int. J. Early Child. Spec. Educ. (INT-JECSE)* **2022**, *14*, 1418–1426. [CrossRef]

81. Chandanapalli, S.B.; Reddy, E.S.; Lakshmi, D.R. DFTDT: Distributed functional tangent decision tree for aqua status prediction in wireless sensor networks. *Int. J. Mach. Learn. Cybern.* **2018**, *9*, 1419–1434. [CrossRef]

82. Ahmadi, F.; Pourmahmood Aghababa, M.; Kalbkhani, H. Nonlinear Regression Model Based on Fractional Bee Colony Algorithm for Loan Time Series. *J. Inf. Syst. Telecommun. (JIST)* **2022**, *2*, 141. [CrossRef]

83. Sun, Y.; Cao, Y.; Li, P. Fault diagnosis for train plug door using weighted fractional wavelet packet decomposition energy entropy. *Accid. Anal. Prev.* **2022**, *166*, 106549. [CrossRef]

84. *Learning with Fractional Orthogonal Kernel Classifiers in Support Vector Machines*; Springer International: Cham, Switzerland, 2015.

85. Niu, H.; Chen, Y.; Guo, L.; West, B.J. A New Triangle: Fractional Calculus, Renormalization Group, and Machine Learning. In Proceedings of the 17th IEEE/ASME International Conference on Mechatronic and Embedded Systems and Applications (MESA), Auckland, New Zealand, 29–31 August 2021; Volume 7. [CrossRef]

86. Niu, H.; Chen, Y.; West, B.J. Why Do Big Data and Machine Learning Entail the Fractional Dynamics? *Entropy* **2021**, *23*, 297. [CrossRef]

87. Yousri, D.; AbdelAty, A.M.; Al-qaness, M.A.; Ewees, A.A.; Radwan, A.G.; Abd Elaziz, M. Discrete fractional-order Caputo method to overcome trapping in local optima: Manta Ray Foraging Optimizer as a case study. *Expert Syst. Appl.* **2022**, *192*, 116355. [CrossRef]

88. Khan, N.; Alsaqer, M.; Shah, H.; Badsha, G.; Abbasi, A.A.; Salehian, S. The 10 Vs, Issues and Challenges of Big Data. *In Proceedings of the 2018 International Conference on Big Data and Education (ICBDE '18)*; Association for Computing Machinery: New York, NY, USA, 2018; pp. 52–56. [CrossRef]

89. Wilson, K.G. The renormalization group: Critical phenomena and the Kondo problem. *Rev. Mod. Phys.* **1975**, *47*, 773–840. [CrossRef]

90. Guo, L.; Chen, Y.; Shi, S.; West, B.J. Renormalization group and fractional calculus methods in a complex world: A review. *Fract. Calc. Appl. Anal.* **2021**, *24*, 5–53. [CrossRef]

91. Mehta, P.; Schwab, D.J. An exact mapping between the Variational Renormalization Group and Deep Learning. *arXiv* 2014, arXiv:1410.3831.

92. Lin, H.W.; Tegmark, M.; Rolnick, D. Why Does Deep and Cheap Learning Work So Well? *J. Stat. Phys.* **2017**, *168*, 1223–1247. [CrossRef]

93. Stanley, H.E.; Amaral, L.A.N.; Buldyrev, S.V.; Gopikrishnan, P.; Plerou, V.; Salinger, M.A. Self-organized complexity in economics and finance. *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 2561–2565.

94. Park, J.B.; Lee, J.W.; Yang, J.S.; Jo, H.H.; Moon, H.T. Complexity analysis of the stock market. *Phys. A Stat. Mech. Appl.* **2007**, *379*, 179–187. [CrossRef]

95. Dominique, C.R.; Solis, L.E.R. Short-term Dependence in Time Series as an Index of Complexity: Example from the S&P-500 Index. *Int. Bus. Res.* **2012**, *5*, 38. [CrossRef]

96. Zhou, W.X.; Sornette, D. Renormalization group analysis of the 2000–2002 anti-bubble in the US S&P500 index: explanation of the hierarchy of five crashes and prediction. *Phys. A Stat. Mech. Appl.* **2003**, *330*, 584–604. [CrossRef]

97. Koch-Janusz, M.; Ringel, Z. Mutual information, neural networks and the renormalization group. *Nat. Phys.* **2018**, *14*, 578–582. [CrossRef]

98. Newman, M. Power laws, Pareto distributions and Zipf's law. *Contemp. Phys.* **2005**, *46*, 323–351. [CrossRef]

99. Molz, F.J.; Rajaram, H.; Lu, S. Stochastic fractal-based models of heterogeneity in subsurface hydrology: Origins, applications, limitations, and future research questions. *Rev. Geophys.* **2004**, *42*. [CrossRef]

100. Xie, H. *Fractals in Rock Mechanics*; CRC Press: London, UK, 2020. [CrossRef]

101. Ku, S.; Lee, C.; Chang, W.; Song, J.W. Fractal structure in the S&P500: A correlation-based threshold network approach. *Chaos Solitons Fractals* **2020**, *137*, 109848. [CrossRef]

102. Zaslavsky, G.M.; Zaslavsky, G.M. *Hamiltonian Chaos and Fractional Dynamics*; Oxford University Press: Oxford, UK; New York, NY, USA, 2004.

103. AI Is Changing How We Do Science. Get a Glimpse. Available online: https://www.science.org/content/article/ai-changing-how-we-do-science-get-glimpse (accessed on 6 December 2022).

104. Leeming, J. How AI is helping the natural sciences. *Nature* **2021**, *598*, S5–S7. [CrossRef]

105. Krenn, M.; Pollice, R.; Guo, S.Y.; Aldeghi, M.; Cervera-Lierta, A.; Friederich, P.; dos Passos Gomes, G.; Häse, F.; Jinich, A.; Nigam, A.; et al. On scientific understanding with artificial intelligence. *Nat. Rev. Phys.* **2022**, *4*, 761–769. [CrossRef]

106. Islam, M.S.; Qaraqe, M.K.; Belhaouari, S.B. Early Prediction of Hemoglobin Alc: A novel Framework for better Diabetes Management. In Proceedings of the 2020 IEEE Symposium Series on Computational Intelligence (SSCI), Canberra, Australia, 1–4 December 2020; pp. 542–547. [CrossRef]

107. Islam, M.S.; Qaraqe, M.K.; Belhaouari, S.B.; Abdul-Ghani, M.A. Advanced Techniques for Predicting the Future Progression of Type 2 Diabetes. *IEEE Access* **2020**, *8*, 120537–120547. [CrossRef]