

Review

Artificial Neural Networks for Pyrolysis, Thermal Analysis, and Thermokinetic Studies: The Status Quo

Nikita V. Muravyev ^{1,*}, Giorgio Luciano ², Heitor Luiz Ornaghi, Jr. ³, Roman Svoboda ⁴ and Sergey Vyazovkin ^{5,*}

¹ N.N. Semenov Federal Research Center for Chemical Physics, Russian Academy of Sciences, 4 Kosygina Str., 119991 Moscow, Russia

² CNR, Istituto di Scienze e Tecnologie Chimiche “Giulio Natta”, Via De Marini 6, 16149 Genova, Italy; giorgio.luciano@ge.ismac.cnr.it

³ Department of Materials, Federal University for Latin American Integration (UNILA), Silvio Américo Sasdelli Avenue, 1842, Foz do Iguacu-Paraná 85866-000, Brazil; hlornagj@ucs.br

⁴ Department of Physical Chemistry, Faculty of Chemical Technology, University of Pardubice, Studentská 95, CZ-53210 Pardubice, Czech Republic; roman.svoboda@upce.cz

⁵ Department of Chemistry, University of Alabama at Birmingham, 901 S. 14th Street, Birmingham, AL 35294, USA

* Correspondence: n.v.muravyev@ya.ru (N.V.M.); vyazovkin@uab.edu (S.V.)

Abstract: Artificial neural networks (ANNs) are a method of machine learning (ML) that is now widely used in physics, chemistry, and material science. ANN can learn from data to identify nonlinear trends and give accurate predictions. ML methods, and ANNs in particular, have already demonstrated their worth in solving various chemical engineering problems, but applications in pyrolysis, thermal analysis, and, especially, thermokinetic studies are still in an initiatory stage. The present article gives a critical overview and summary of the available literature on applying ANNs in the field of pyrolysis, thermal analysis, and thermokinetic studies. More than 100 papers from these research areas are surveyed. Some approaches from the broad field of chemical engineering are discussed as the venues for possible transfer to the field of pyrolysis and thermal analysis studies in general. It is stressed that the current thermokinetic applications of ANNs are yet to evolve significantly to reach the capabilities of the existing isoconversional and model-fitting methods.

Keywords: artificial neural networks; conversion degree; kinetics; machine learning; pyrolysis; thermal analysis



Citation: Muravyev, N.V.; Luciano, G.; Ornaghi, H.L., Jr.; Svoboda, R.; Vyazovkin, S. Artificial Neural Networks for Pyrolysis, Thermal Analysis, and Thermokinetic Studies: The Status Quo. *Molecules* **2021**, *26*, 3727. <https://doi.org/10.3390/molecules26123727>

Academic Editor: Teobald Kupka

Received: 30 May 2021

Accepted: 15 June 2021

Published: 18 June 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Machine learning has found a number of applications in chemistry [1–3] and material science [4–6]. Among the supervised machine learning methods, the artificial neural networks (ANN) are most popular. Artificial neural networks are the models inspired by biological neural systems such as the human brain [7]. ANN represents the oriented graphs with nodes called by the same analogy with brain as “neurons”. The key features of the neural network are its topology (number of layers and neurons in it) and the strength of connections between the neurons (defined by mathematical weights). The benefits of ANN include the capability to approximate any continuous non-linear function, and high tolerance to noisy or missing data [8,9]. Thanks to these properties, the neural networks have found their use in vastly different fields, e.g., image classification [10], predicting combustion instability [11], or impact sensitivity of energetic materials [12]. Although ANNs have been extensively applied in chemical engineering since the 1990s [13–15], their usage in thermal analysis has commenced later [9,16], and most of the studies emerged only recently [17–21]. The present review aims to summarize such studies and to assess some future prospects.

Thermal analysis (TA) is defined as “the study of the relationship between a sample property and its temperature as the sample is heated or cooled in a controlled manner” [22]. In this review, we focus mainly on applications of neural networks to thermal analysis and thermokinetic studies. Additionally, as one can see from the literature survey, typical applications involve pyrolysis as a thermal process considered most commonly by researchers. To keep the scope of review concise, the works on prediction of yield of pyrolysis product (regarded as a sort of TA experiment) are included, while reports on the selectivity and kinetics of pyrolysis gas products quantified with ANNs [23–25] are not covered in this review.

The most common methods of thermal analysis are differential scanning calorimetry (DSC) and thermogravimetric analysis (TGA) normally applied under linear rising temperature or constant temperature conditions. DSC produces a differential signal, i.e., the heat flow data, that is generally assumed to be proportional to the conversion rate $d\alpha/dt$. This assumption ignores the thermal inertia term in the heat flow that is a potential source of a systematic error in kinetic evaluations [26]. However, this error tends to be negligible when using smaller masses and slower heating rates [27]. In turn, TGA data, representing the mass change of the sample, is an integral signal, and its normalized change yields the conversion degree α . The resulting conversion degree plotted against temperature and time is also known as the transformation–time–temperature (TTT) dataset. With this data array, the thermokinetic analysis is then performed. The procedure relies on the following basic equation [28]:

$$\frac{d\alpha}{dt} = k(T)f(\alpha), \quad (1)$$

The multiplicand in Equation (1), i.e., the rate constant, usually assumes an Arrhenius form $k(T) = A \exp(-E_a/RT)$ with E_a being the activation energy and A the preexponential factor. The multiplier in Equation (1) is the reaction model. Various approaches have been proposed for thermokinetic analysis, the most effective being the model-fitting [29,30] and the isoconversional [31,32] techniques. The obtained kinetic parameters generally serve for obtaining mechanistic insights into the processes [33,34], or for predicting their thermal behavior [35].

From the conceptual point of view, the traditional thermokinetic analysis, i.e., Equation (1) with the kinetic triplet $(A, E_a, f(\alpha))$, is a mechanistic, phenomenological, or parametric modeling. The models of this type are known as white-box models, in contrast to the data-driven or nonparametric black-box modelling (Figure 1). The papers discussed in the present review are primarily concerned with the latter approach. To give a simple example of a black box model, let us imagine a neural network that is trained to output the conversion rate data once fed with temperature and time values. This ANN is readily applicable and does not require knowledge of the process kinetics. If carefully designed and properly trained, the network offers good descriptive accuracy within the domain of the supplied process data. Main limitations of the black-box models are the uncertain interpretation [36] and poor generalization ability [37]. To combine the benefits of white- and black-box approaches, hybrid semi-parametric modeling is proposed [38] (Figure 1). The application of hybrid models will be addressed in detail in Section 4 of this review. It is worth mentioning that some of the further-discussed ANNs are trained on synthetic data generated with Equation (1) or give as an output the kinetic parameters of Equation (1). Such models cannot be regarded as authentic black-box models because they either implicitly or explicitly utilize a phenomenological model.

White-box models (phenomenological, parametric)	Gray-box models (hybrid, semi-parametric)	Black-box models (data-driven, nonparametric)
--	--	--

Figure 1. Three conceptual approaches to chemical engineering modeling.

The present review briefly surveys the conceptual details of ANNs, and summarizes their applications in pyrolysis, thermal analysis, and kinetic studies. A strong emphasis is put on the kinetic applications as this is a quickly growing field that needs special attention for proper development. Figure 2 schematically shows the blocks that represent major topics identified via analysis of the literature. The article is roughly structured according to these topics that are presented as the article sections. Note that the divisions between the sections are rather conditional and that the topics covered oftentimes overlap with other chemical engineering or chemometric problems. The review is intended to reflect the existing state of affairs in the area of the aforementioned applications and to provide certain critique, and to highlight some desired directions for future research.

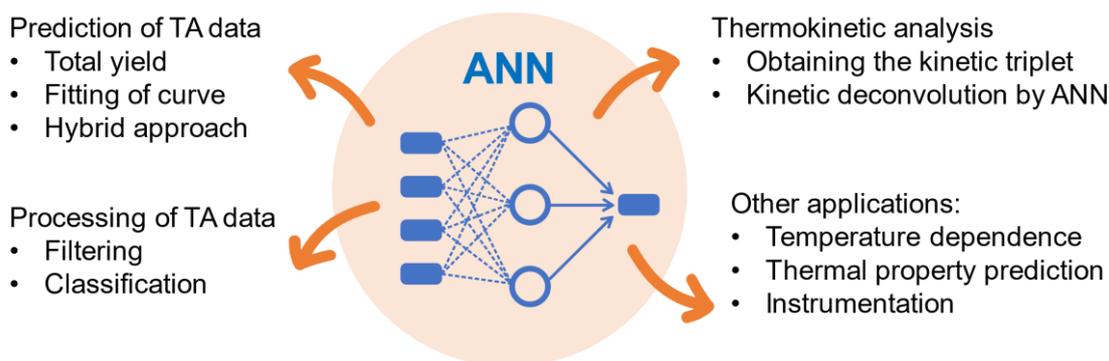


Figure 2. Logical scheme of the present review. Several aspects of using artificial neural networks (ANN) for thermal analysis (TA) studies are indicated to be discussed in main text.

2. Technical Details behind ANNs

Before discussing general aspects of the ANNs implementation, we should recognize that there is no single mathematical tool that resolves all TA-specific problems. ANN is only one of several machine learning methods and its accuracy for specific applications is hardly assessed a priori (this situation is known in mathematics as a “no free lunch” theorem [39]). However, for some applications the ANN model has been compared with other machine learning approaches, e.g., random forest [40], gradient boosting machine [40], and it has been found that ANN performs better.

Several types of ANNs are used in chemical engineering studies (e.g., see review by Ali et al. [41]). Nonetheless, almost all papers considered in the present review utilize the same type of ANN known as feed-forward neural network, which is briefly described below. More practical details are readily available from monographic sources [42,43].

ANN is a parallel processing system of interconnected computational elements, called neurons (Figure 3). The neurons are arranged in several layers—input, hidden, and outer layers. The topology of ANN is encoded numerically as follows, e.g., 25-15-10 denotes the neural network with 25 neurons in the input layer, 15 neurons in the hidden layer, and 10 neurons in the output layer. Note that most studies reviewed adopt the architecture with a single hidden layer, rarely two hidden layers are used [20,44], and only for exceptionally complex thermal profiles the authors has had to introduce even more hidden layers [45]. Connections between the neurons are displayed as lines in Figure 2, each connection has its own strength value called weight. Output of the neuron is determined by each layer’s transfer functions (activation function). While in most studies the hyperparameters (parameters used to control the learning process) and parameters of ANN are tuned manually, there are some techniques to make the procedure more rigorous and human-independent. In particular, the differential evolution algorithm can be utilized to optimize the number of hidden layers, number of neurons in it, weights, biases, and activation functions [46].

Let us consider ANN in more detail. The output of i -th neuron in $(L + 1)$ -th layer is denoted as y_i^{L+1} . This value is comprised of a specific constant value called bias b_i^{L+1} , and

N_L signals from previous layers, amplified or weakened by the corresponding coefficients, called weights w_{ij}^{L+1} [19]:

$$y_i^{L+1} = f\left(\sum_{j=1}^{N_L} w_{ij}^{L+1} y_j^L + b_i^{L+1}\right). \quad (2)$$

Among a broad variety of the activation functions [42,43], those of sigmoidal type $f(u) = \text{Logsig}(u) = 1/(1 + e^{-u})$, symmetric sigmoidal form $f(u) = \text{Tansig}(u) = -1 + 2/(1 + e^{-2u})$ have been used most frequently. Thus, the ANN model transforms the input vector (x_1, x_2, \dots, x_i) to the output vector (y_1, y_2, \dots, y_j) . Before using, ANN is trained on the $[(x_1, y_1), \dots, (x_n, y_n)]$ data, by modifying the weights of connections between the neurons to minimize the prediction error (supervised learning). Although more efficient and modern algorithms (e.g., Stochastic Gradient Descent or Adaptive Moment Estimation) are used broadly in ANN training, the application considered in the present review appear to be limited primarily to the Levenberg–Marquardt or error backpropagation algorithms. The key problem with the optimization is that it can converge to local minima. Some modifications for the common procedures are constantly proposed to improve the computational time and robustness. For example, ant colony optimization (ACO) algorithm is suggested [47] to improve the convergence of backpropagation technique, and it is shown that it slightly improves the performance of ANN [48]. To evaluate how well ANN predicts in the domains not used for training, the data are usually split into the training and validation sets. The split ratio should generally depend on the quality and structure of the data. However, our analysis of the literature concerned with the present review topic suggests that the typical ratio for sizes of the training and validation sets is $\sim 70\%$ and $\sim 30\%$, respectively. The realistic confidence limits for ANN predictions can be generated as described elsewhere [49]. Ensemble learning techniques, like stacking of several ANNs, boosting or bagging, have been suggested as a viable methods to further improve the generalization ability [50].

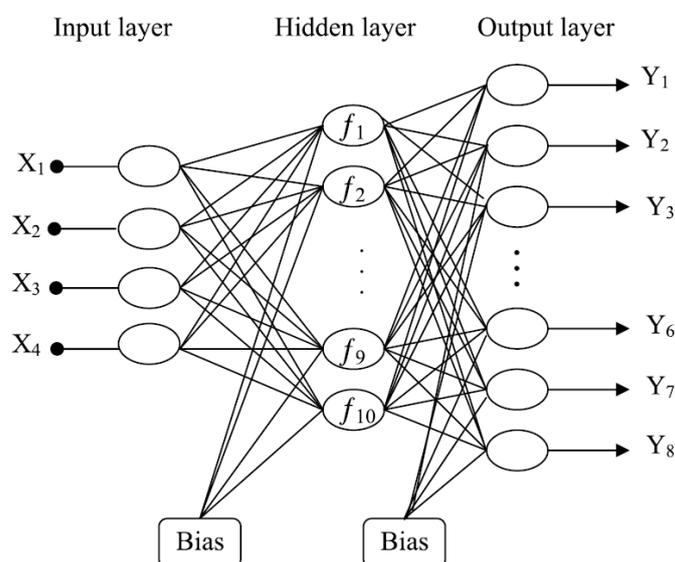


Figure 3. Scheme of the artificial neural network with indicated layers, weights, functions [51]. Reprinted with permission from Taylor & Francis.

Optimal ANN architecture is defined by learning quality and generalization ability. Both parameters rely on the data, which should be representative and big enough in size. However, as the studies considered below mostly involve the time-consuming and labor-intensive measurements, the size of available data is usually small, especially when compared with typical applications of machine learning. Just to give an example of typical number of learning data, 50,000 images are used to train ANN to recognize the handwritten

numbers from 1 to 10 [52]. For the thermal behavior data, the amount of input data is defined by a particular experimental setup and sampling frequency, but usually is far less than the number above. To give an estimate of how much data is needed, we notice that the theoretical estimates suggest that the number of learning patterns necessary to attain the sufficient generalization ability should be ten times larger than the total number of weights in ANN [53,54]. The problem of sparse data can be alleviated by the generation of the artificial raw thermal analysis data [21], as discussed in the following section.

An important aspect of the application of ANNs to the TA-generated data is the representation of the input data. TA data used for thermokinetic analysis are usually in the form of several sets of three-dimensional data (conversion, temperature, time). These data should be transformed to decrease the dimensionality of the information being fed to the neurons of the input layer. The simplest way of doing this is to limit input to a single experiment and input one of the dependent parameters, e.g., time for isothermal runs [17], and train ANN to output the conversion data (note, that utilization of data corresponding only to a single run is not recommended from the kinetic point of view [28], although in some specific cases the approach could work). Raw thermal analysis signal can be considered as the output, while the temperature is the independent variable and each heating rate is considered separately [20,55]. A set of temperature values corresponding to the specific conversion degrees is used as inputs, and the single row of input vector is formed from several experiments performed at different heating rates [21,56]. This thin-out procedure should retain the kinetic information, i.e., have considerable number of points, and contain the regions with critical information [9,16] (peak, inflexion, etc.).

Some more advanced methods for mapping the multi-dimensional thermal analysis data for the input layer of neural network still need to be developed. Muravyev and Pivkina [21] propose using principal component analysis [57–59] or introducing additional hidden layers for this purpose. To the best of authors' knowledge, no more advanced methods of introduction of TA data to ANN have been proposed so far. Possibly, some methods from other fields can be borrowed, e.g., that employed for spectral dimensionality reduction in spectroscopy and signal analysis [60].

Finally, reporting of the optimized architecture of the ANN model is strongly recommended. Only a few examples [19,61,62] are found in the reviewed literature that provide the network biases and weights. As a guiding principle, one must provide enough detail on the ANN architecture, so that others can reproduce the computations. This principle is commonly accepted in reporting experimental studies. It is only prudent to follow it in publishing ANN work.

3. Prediction of Conversion Data (Single Value, Whole Curve)

There is a number of papers [63–66] on the application of ANNs for obtaining the total conversion (e.g., mass loss by TGA) using experimental conditions as an input. One of the first papers on the topic has been by Carsky and Kuwornoo [67] who used ANN to predict the main quantities for coal pyrolysis: tar, volatiles, and char yields. Table 1 shows the input parameters, which those authors have compiled from the literature on pyrolysis of several varieties of coals in fluidized bed reactor, spouted bed reactor, hot-rod, and wire-mesh reactors (since the experiments comprise the heating and isothermal hold steps, we consider them as TA data). An important aspect of the study is that the input parameters are in most cases not distributed uniformly or continuously. Thus, the authors have used, additionally, the fuzzy modeling to transform the particle holding time. For other variables, e.g., particle size, a qualitative classification (fine, coarse) was used instead of unevenly distributed numerical values. The optimized neural network (47-16-3) has described well the experimental tar, volatiles, and char yields, as is shown by root mean square errors of 4.2, 4.9, and 6.4%, respectively.

Table 1. The list of the parameters used to model the coal pyrolysis, and some details on their incorporation to ANN [67]. Reprinted with permission from Elsevier.

Parameter	Format	Transformation	Input/Output
Particle holding time	Fuzzy class membership (s)	Logical (0/1)	Input
Particle holding time	Fuzzy class membership (min)	Logical (0/1)	Input
Freeboard residence time of volatiles	Numerical (s)	Linear, Inverse function, Inverse fourth root function, Fuzzy center (0) function	Input
Temperature	Numerical (°C)	Linear, Tanh, Square root function	Input
Time of preheating of particles	Numerical (s)	Linear, Inverse fourth root function, Fuzzy center (0) function, Inverse function	Input
Rate of heating	Numerical (K/s)	Linear, Inverse fourth root function, Inverse function	Input
Secondary reaction in freeboard	Qualitative (none, small, medium, great)	Enumerated strings: small, medium, great	Input
Secondary reaction in bed	Qualitative (small, medium, great)	Enumerated strings: small, great	Input
Reactor type	String (flb, spb, hr, wm) ¹	Enumerated strings: flb, spb, hr, wm	Input
Particle size	Qualitative (fine, coarse)	Enumerated strings: fine, coarse	Input
Concentration of coal particles	Qualitative (low, medium, high)	Enumerated strings: low, medium, high	Input
Bed depth	Qualitative (shallow, deep, medium, mono-layer, mono-layer-shallow)	Enumerated strings: shallow, deep, medium, mono-layer, mono-layer-shallow	Input
Sample size	Numerical (g of coal)	Linear, Inverse fourth root function, Inverse function	Input
Coal: wt% ash (d.b.) ²	Numerical (% db)	Linear, Inverse fourth root function, Inverse function	Input
Coal: wt% volatiles (d.a.f.) ³	Numerical (% daf)	Linear, Fourth power function, Exponential	Input
Carrier gas velocity	Numerical (m/s)	Linear, Inverse fourth root function, Inverse function	Input
Yield of tar	Numerical (wt%)	Linear	Output
Yield of volatiles	Numerical (wt%)	Inverse function	Output
Yield of char	Numerical (wt%)	Inverse function	Output

¹ Designation of reactor types: flb = fluidized bed reactor, spb = spouted bed reactor, hr = hot-rod fixed bed reactor, wm = wire-mesh reactor. ² d.b.—dry basis value (except all moisture). ³ d.a.f.—dry ash free value (except all moisture and ash).

Thus, originally, this approach has been proposed to derive a single final value of conversion. In more recent studies, the idea has been extended to employ ANNs for predicting the conversion values at various moments, i.e., a TA signal itself. Such studies differ drastically in experimental systems (*Tithonia diversifolia* weed biomass [63], low-density polyethylene [64], high-density polyethylene [68], safflower seed press cake [65], durian rinds [69], rape straw [70], coal gangue and peanut shell [71], pet coke [72], sewage sludge and peanut shell [73], sewage sludge and coffee grounds [74], vegetable fibers [75], rice husk and sewage sludge [45], sludge, watermelon rind, corncob, and eucalyptus [76], *Sargassum* sp. seaweed [77], cotton cocoon shell, tea waste, and olive husk [66], mechanoactivated coals [78], cattle manure [79], lignocellulosic forest residue and olive oil residue [80], cotton cocoon shell, fabricated tea waste, olive husk, and hazelnut shell [81]) and in some minor details, but the general concept remains the same. Thus, we illustrate it with the

study by Kataki et al. [63], who used a neural network model (4-14-1) to predict the product yield for the pyrolysis of dried weed biomass. The four input parameters have been the pyrolysis temperature (during the 30 min isothermal segment), heating rate used to attain this temperature, nitrogen flow rate, and particle size of the specimen. Thirty pyrolysis experiments with varying experimental conditions have been designed and performed, and the target output (yield of the bio-oil) has been determined. The optimized ANN has demonstrated a better accuracy than the response surface method used as an alternative. Finally, ANN has been used to design the experimental conditions for maximizing the output quantity, and the result has agreed closely with the neural network prediction. Although the inner relationships of ANN are not known, some insights can be gained by looking at the relative importance of the input descriptors. Figure 4 shows that the pyrolysis temperature is the main factor (more than 30% of relative importance) followed by heating rate (25%), flow rate (22%), and particle size (18%). The discussed approach can be considered as a sort of experimental design and can be definitely used for industrial thermal processes to optimize the single target property (e.g., yield of product).

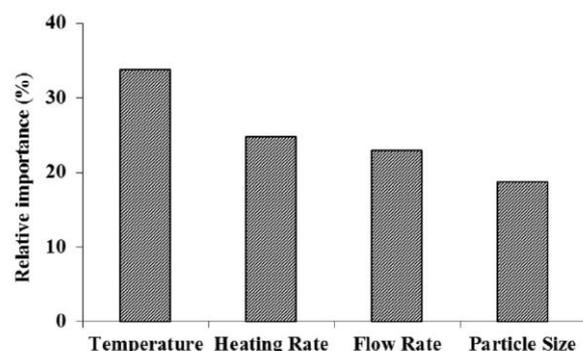


Figure 4. The example of relative importance analysis for input descriptors of trained neural network model [63]. Reprinted by permission from Springer Nature.

The idea from the above-mentioned studies can be expanded further. ANN can be trained to predict the conversion degree (or the primary TA signal) based on temperature and time. This approach can be classified as a nonlinear fit of the TA data, because no extrapolation is performed, and all data are within the experimental domain space. The graphical representation of the workflow from the study by Bezerra et al. [20] is shown in Figure 5a. It is rather a general approach that predicts the thermal signal from a principal descriptor (temperature) and another parameter, which defines the specific nonisothermal measurement, the heating rate. Figure 5b,c show the good quality fit via neural network models applied to complex thermally-induced processes. Figure 5b illustrates the experimental TGA signal for nonisothermal pyrolysis of a carbon reinforced fiber composite along with its fit via the ANN model (2-21-21-1). Figure 5c represents the isothermal data for iron oxide reduction process [19] fitted with a neural network (3-20-1). In both cases, ANN offers an excellent fit of experimental data.

As one may notice, the ANN used to obtain the data shown in Figure 5b contains two hidden layers, whereas the one used to fit the data from Figure 5c contains only a single hidden layer of neurons. The selection of complexity of the ANN topology is, thus, forced by the data. This observation is in line with the findings by Cepeliogullar et al. [44]. They have studied the pyrolysis of refuse-derived fuel and analyzed in detail the impact of the ANN architecture parameters. Due to the complex thermal behavior of the investigated highly heterogeneous samples, only after introduction of the second hidden layer of neurons an adequate fitting quality is achieved. In addition to the number of neurons in both hidden layers, the authors considered various activation functions. For the optimal neural network, the prediction ability was tested by comparison with direct experiment not used in the model development. The predicted mass loss curve for heating rate of 25 K min^{-1} closely agrees with the experiment. Note, however, that ANN has been trained

with data acquired at 5, 10, 15, 20, 30, 40, and 50 K min⁻¹, thus, this prediction is still within that domain space and it is, in fact, a sort of interpolation.

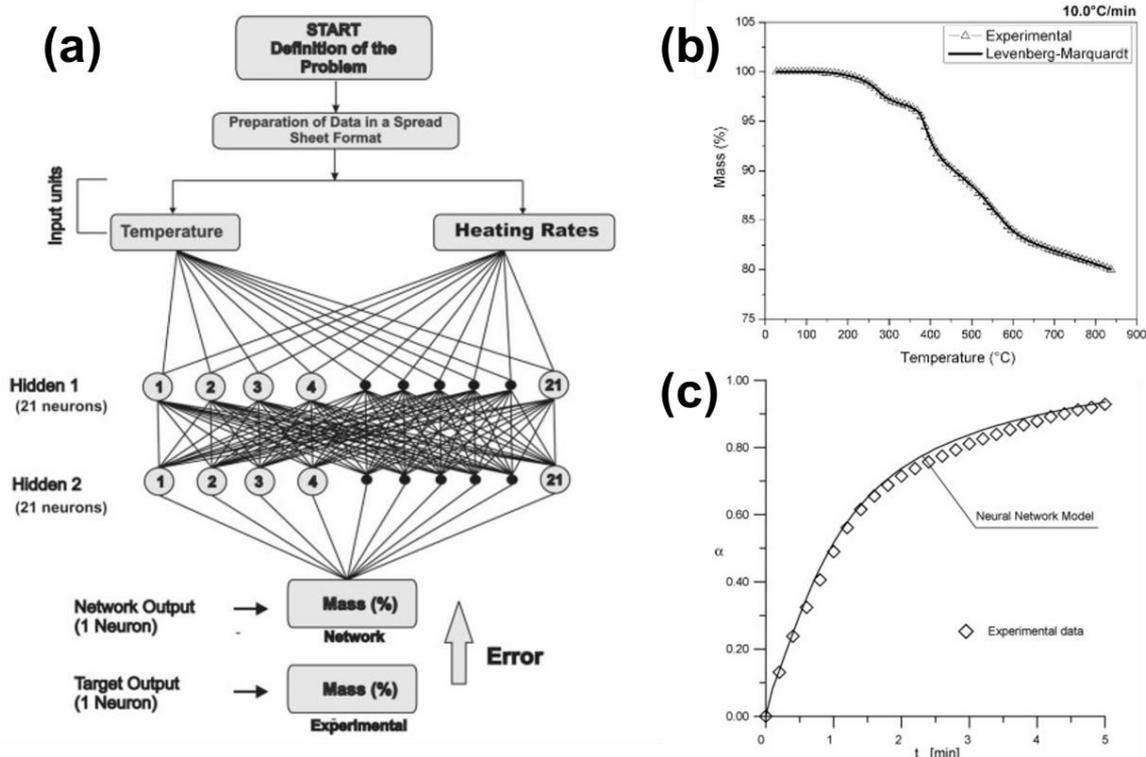


Figure 5. Typical approach to thermal signal fit with ANN model: (a) the workflow [20], and examples of use (b,c). Nonisothermal data (b) corresponds to the pyrolysis of carbon reinforced carbon composite (heating rate of 10 K min⁻¹) [20]. Isothermal data (c) represents the iron oxide reduction ($T = 880$ °C, environmental gas 90% N₂ + 10% CO) [19]. Reprinted with permission from Elsevier.

Burgaz et al. [82] have applied the neural network models to directly predict several thermal analysis signals (DSC, TGA, and dynamic mechanical analysis, DMA) for poly(ethylene oxide)/clay nanocomposites. The input parameters are the amount of clay in the sample and temperature, other details are summarized in Table 2. Authors note that the amount of data necessary to attain good performance is signal-dependent, i.e., it is smaller for DMA than for the TGA and DSC signals. The higher complexity of the mass loss and heat flow signals also forces the use of two hidden layers instead of only one for DMA.

Table 2. Details of the neural network models for prediction of various thermal signals for poly(ethylene oxide)/clay nanocomposites [82]. Reprinted with permission from Elsevier.

ANN and Training Parameters	DMA	DMA	TGA	DSC
Input data	Temperature and clay wt%			
Output data	E'	tanδ (E''/E')	Weight loss	Heat flow
Training data	193	193	406	565
Testing data	47	47	45	141
Number of hidden layers	1	1	2	2
Activation functions	Tansig-Purelin	Tansig-Purelin	Tansig-Logsig-Purelin	Tansig-Logsig-Purelin
Number of neurons in hidden layer	9	9	4–3	5–4
Number of iterations	288	152	48	65
Regression (R2)	0.9997	0.9982	0.99994	0.99985

From the research discussed in this section, we can conclude that ANNs can successfully fit various thermal signals. The topology of the ANN model and other parameters depend on the amount of learning data, the selected input parameters, but most profoundly on the inherent complexity of data. The examples available so far do not include some testing on distant extrapolation accuracy of the neural networks output. Research in this direction would certainly be of value. In all, due to its excellent interpolation accuracy, this particular approach (conversion prediction from the experimental conditions) already has a plenty of industrial applications.

4. Advancements in Prediction of Conversion Data with ANNs (Moving Window, Hybrid Models)

It appears that chemical engineering has several approaches to the ANN applications that can be transferred to pyrolysis and thermal analysis problems. Note that publications discussed in the previous section do not use the conversion data as input, but using such input in ANNs with moving time window is highly beneficial. Figure 6 illustrates schematically the concept [54]: the input data include the temperature (pressure) data and the concentration values from previous time steps and the output data are the concentration of the reaction products at this particular time period. The depth of the time shift and the sampling time become important parameters of the model. The neural model with moving window outperforms the traditional neural network as shown on the data for hydrogenation of 2,4-dinitrotoluene in stirred tank reactor [54]. The main drawback of this kind of ANNs is the increased number of parameters that necessitates large quantities of input data and results in considerable computational times [83].

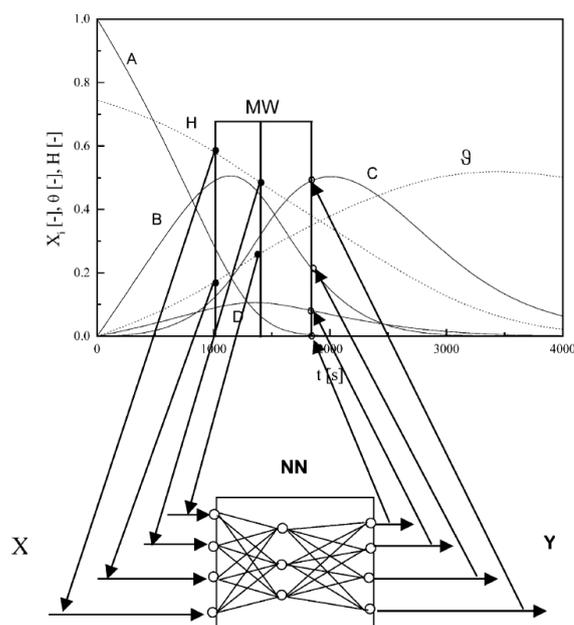


Figure 6. Modeling of the reactor by neural networks with moving time window [54]. X is the vector of input data, Y is vector of output data. Reprinted with permission from Elsevier.

The utilization of neural networks primarily for prediction of the conversion profiles can be considered as a sort of black-box approach. It has no need of knowledge of the process fundamentals. However, alternatively to this fully black-box approach, the hybrid approach is proposed [38], being a combination of traditional modeling with description by ANN of the unknown part of the process (e.g., the reaction kinetics). It is deemed that the latter technique is beneficial as it affords some degree of the generalization of the results [54]. We are not aware of using this approach to the specific applications (i.e., to pyrolysis, thermal analysis, and thermokinetic studies) considered in the present review, thus, we make an excursus in chemical engineering to introduce the basic idea of the method.

To exemplify the hybrid approach, let us consider the study of heterogeneous oxidation of 2-octanol with nitric acid by Molga et al. [84]. The chemical reaction was performed in the reaction calorimeter (RC1, Mettler-Toledo) and during the experiment, a small amount of samples were extracted and analyzed chemically. The adopted reaction mechanism comprises the oxidation of 2-octanol (A) to 2-octanone (P) by action of nitrosonium ion (B), and the subsequent oxidation with formation of carboxylic acids (X):



The neural network works here as a kinetic model: it gets the temperature and concentrations of A, P, X, B as inputs and outputs the reactions rates for two reactions, R_1, R_2 . ANN is then used in combination with traditional mass balance and heat balance equations, e.g.:

$$\frac{dn_P}{dt} = (R_1 - R_2)(n_{A,0} + n_{N,0}), \quad (4)$$

(n_i is molar amount of product i in the system, subscript N stands for the nitric acid, a source of nitrosonium ion). That is, the hybrid model is built, where the system of equations is solved with the kinetic data coming from the trained ANN (Figure 7).

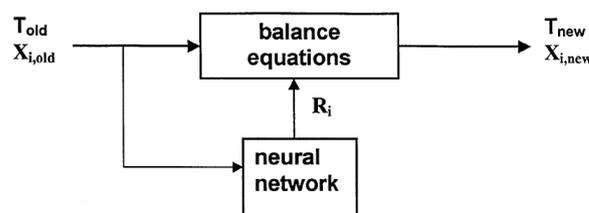


Figure 7. Schematic diagram of the hybrid model of chemical reactor [84]. Reprinted with permission from Elsevier.

This hybrid approach is well developed in chemical engineering [85], but it has not been yet applied systematically to the problems considered in the present review. One of the examples of its application is the study by Guo et al. [86], where the hybrid model was built for gas production rates in the course of coal pyrolysis. More research with this approach would be welcome. For example, it should be useful for purposes of thermal behavior prediction [35]. The conversion rate term will be supplied by ANN from input T_i, α_{i-1} and then the output $d\alpha/dt|_i$ value will enter the conventional phenomenological model, thermal balance equations for storage condition [87]:

$$m_s c_{p,s} \frac{dT_s}{dt} + m_s c_{p,c} \frac{dT_c}{dt} = US(T_{env} - T_c) + m_s \Delta H_d \frac{d\alpha}{dt}, \quad (5)$$

where ΔH_d is the enthalpy of decomposition reaction, U —the heat transfer coefficient, S —the area of a contact surface between sample and container, c_p, m , and T stand for the isobaric heat capacity, mass, and temperature, correspondingly. The subscripts s, c, env . refer to the sample, container, and environment parameters, respectively.

There are other ways of combining traditional modeling with that by neural network [85,88]. An example is a study by Curteanu et al. [46] where the conversion and molar weights (α, MW) for the free radical polymerization of styrene are obtained by various models as the functions of time, temperature, and concentration of the initiator. Specifically, the authors consider four situations: (1) serial, the process parameters are an input to the traditional model and its output (α, MW) is sent to the ANN model that gives the final values of (α, MW); (2) parallel, both, the traditional model and ANN are fed with the process parameters and give the (α, MW) values, these values are averaged to give the final result; (3) serial-parallel, the ANN corrects the output by traditional model; (4) the traditional model output is used up to the time of gel effect appearance, then the output of the ANN model takes over. The latter situation shows the best accuracy of

prediction, and the approach is probably transferable to other processes with the change of reaction mechanism.

5. Processing of the Raw Thermal Analysis Data (Filtering, Classification)

First, we consider possible applications of ANNs for treatment and analysis of the raw thermal analysis data. Sbirrazzuoli et al. [9,16] have proposed using the neural network models for filtering and deconvolution of the DSC data. Caloric thermoanalytical signals have been generated by the Joule effect. The respective electric pulse has been programmed to certain thermodynamic and kinetic parameters. ANN has been trained on the resulting DSC data setting as target values the expected parameters. In this manner, the transfer function of the calorimeter can be obtained and used for deconvolution of DSC data. Moreover, the results have shown high tolerance of ANNs to a noise introduced in the simulated DSC data.

Another important aspect of the ANNs application to the raw TA data is their ability to recognize some regularities in the data. Recently, Cruz et al. [40] have proposed to employ ANNs for detecting adulteration in milk. That is, the DSC curves for samples with varying content of additives (starch, formaldehyde, whey, urea) show certain differences, and the general idea is to train the neural network to detect these undesired additives by analyzing the DSC data. The DSC measurements have comprised the cooling stage (from 25 °C to −40 °C at 10 K min^{−1} rate) followed by the heating step up to 180 °C at 10 K min^{−1}. Typical DSC curves obtained in this manner are shown in Figure 8. The observed thermal events were analyzed in terms of the onset and peak temperatures and heat effects. Table 3 compiles the results for pure milk (the control sample) and milk with different amount of the starch additive, the subscripts *c*, *m*, and *b* correspond to the crystallization, melting, and boiling processes.

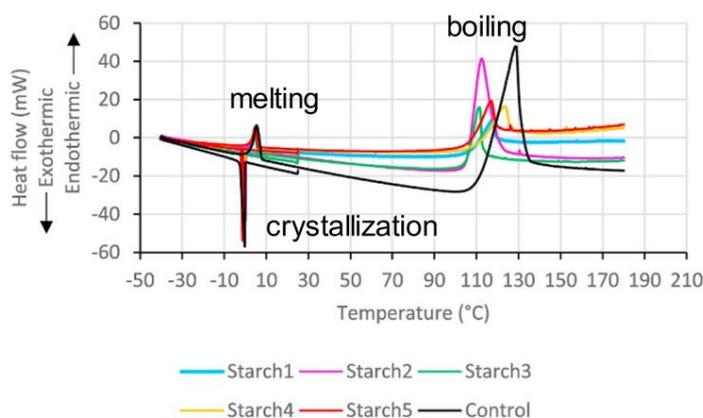


Figure 8. DSC signals for milk samples adulterated with various amount of starch [40]. Thermal processes that take place during cooling and subsequent heating are indicated. Reprinted with permission from Elsevier.

As seen from Table 3, the measured quantities are all altered by the additive presence, and the main purpose of that study is to train the neural network to identify the adulterated sample. This is a chemometric problem of classification, i.e., ANN should output 1 for the correct additive type and 0 for another additive types. The ANN (9-3-5) has been used. The input data are nine measured thermal properties (first column in Table 1), and the outputs are five sample classes (control, starch, formaldehyde, whey, urea). Apart from ANN, other machine learning methods have been tested, namely, the random forest and gradient boosting machine (GBM). Excellent prediction capacity (100%) is accomplished for ANN and GBM methods. Interestingly, these two top-performing methods reveal different values of the input variables importance, e.g., the onset of melting T_{om} is deemed not important by ANN (only 14%), whereas it is one of most important predictors by GBM (84%). Overall, the application of ANN for preliminary analysis of TA results is

very promising. This example of the classification problem can be easily extended to other problems of analytical and industrial relevance.

Table 3. Thermal properties derived from the DSC data for various milk samples [40]. Reprinted with permission from Elsevier.

Parameters	Control	Starch1	Starch2	Starch3	Starch4	Starch5
T_{oc} (°C)	−0.430	0.035	0.065	−0.480	−0.670	−0.535
T_{pc} (°C)	−0.400	−0.290	−0.035	−0.795	−0.535	−0.740
ΔH_c (J g ^{−1})	−253.50	−233.430	−247.197	−244.578	−243.339	−244.377
T_{om} (°C)	−0.385	−0.645	−0.535	−1.140	−1.925	−1.230
T_{pm} (°C)	5.315	4.600	5.665	5.495	5.350	6.510
ΔH_m (J g ^{−1})	220.833	263.704	236.707	199.669	217.466	226.643
T_{ob} (°C)	104.505	103.955	105.645	105.400	107.025	163.480
T_{pb} (°C)	123.330	118.615	131.570	123.805	121.935	168.880
ΔH_b (J g ^{−1})	3577.5	870.990	4313.429	1416.971	1675.230	844.814

Notes: The results show only the mean value, the standard deviation values can be found in the original paper. Control sample = pure milk without additives, samples called Starch 1, 2, 3, 4, 5 = 1, 2, 3, 4, 5 g L^{−1} milk. Subscripts *o* and *p* denote the onset and peak values, subscripts *c*, *m*, and *b* correspond to crystallization, melting, and boiling processes.

Wesolowski et al. [89] have suggested utilizing neural networks for compatibility assessment. This study appears to be the only example of the Kohonen network [90] (also known as self-organizing map, SOM) being applied directly to the TA data. In short, this type of network has no hidden layer and serves to map the input data into a two-dimensional grid. SOMs differ also in the training method; they apply unsupervised competitive learning. Figure 9b,c gives examples of the output topological maps. In their work [89], the authors looked at the thermogravimetric profiles of caffeine mixtures with several excipients, e.g., cellulose and glycine. First, the correlation matrix between the temperatures at certain mass loss values is built. Several temperature values at specific mass loss percentages (T_5 , T_{20} , T_{35} , T_{55} , T_{70}) that show strong correlation are selected and serve as inputs in SOM (Figure 9a). The optimal map size was found to be 5 × 5. Figure 9b,c shows the results of calculations for two typical mixtures. The highlighted neuron for the particular sample is the one that satisfies best the selection criteria. The interpretation of SOM maps is based on the assumption that the samples with close compositions should nest in allied regions. Thus, for caffeine/cellulose, the mixtures 7:3, 9:1 lie close to the neat caffeine. Analogously, the system with excess of cellulose activates the neurons in the proximity of that for pure cellulose. Another mixture, caffeine with glycine, exhibits a much different behavior (Figure 9c), thus, it is deemed incompatible. More mixtures were investigated in this manner, and the outcome of SOM were in agreement with other experimental techniques (DSC, FTIR). The proposed approach using SOMs outperforms the two other chemometric methods—the cluster analysis and the principal component analysis. Both these methods are not capable of resolving the mixtures adequately, when given the same characteristic temperatures as input.

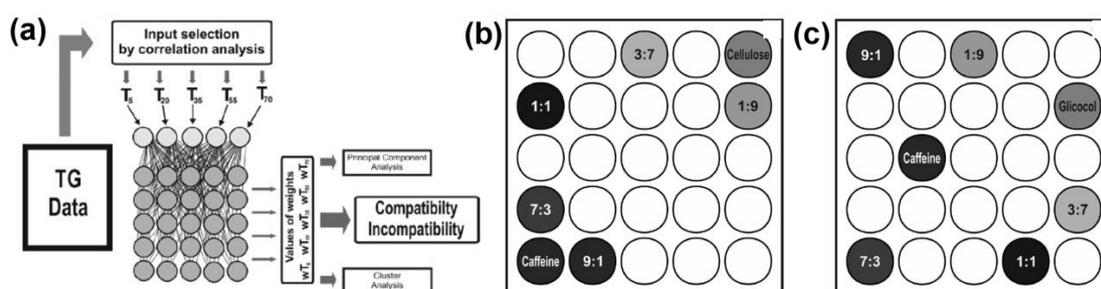


Figure 9. ANN application in the compatibility assessment [89]: (a) Schematic of the proposed workflow and mapping with ANN of caffeine mixtures with cellulose (b) and glycine (c), mixture ratios are indicated, i.e., 9:1, 7:3, 1:1, 3:7, 1:9. Reprinted with permission from Elsevier.

6. Thermokinetic Analysis with Neural Networks

In this section, we discuss several applications of ANNs relevant to thermokinetic analysis. It should be immediately noted that some publications [17,91–93] on using ANNs for kinetic analysis deal with analysis of single-heating rate data or a single isothermal measurement. This approach is not recommended by the Kinetics Committee of the International Confederation for Thermal Analysis and Calorimetry (ICTAC) [28], as it does not allow for properly discriminating between the reaction models and, thus, for obtaining reliable kinetic parameters, as discussed further. Nevertheless, some of these works are discussed below to illustrate the development and current status of the ANNs applications in thermokinetics.

One of the benefits of the ANNs use for kinetic studies [9] is that they could bypass an explicit usage of Equation (1). This statement needs some clarification. For instance, some of the examples discussed in Section 3 show that one can predict the conversion degree as function of temperature and time without invoking Equation (1) and determining the kinetics triplets at all. However, such predictions are yet to be thoroughly verified to prove their general validity. On the other hand, evaluation of the kinetic triplets with the aid of ANN is possible without explicitly separating the $k(T)$ and $f(\alpha)$ terms as well as making any assumptions about their mathematical form. Some evaluations of this type [21,94,95] have already been thoroughly tested in terms of predictions outside of the data space where the ANN is trained. However, such evaluations are based on the neural networks trained on the data simulated according to Equation (1), so that the knowledge of Equation (1) implicitly propagates into predictions.

Gauri et al. [15] compare the performance of the mechanistic kinetic modeling with the neural network. They consider a heterogeneous reaction of various limestones with SO_2 . Two reaction models have been proposed in literature for this type of reaction: the shrinking unreacted core model and the distributed pore model. As an example, we will consider only the simpler one, the shrinking unreacted core model:

$$\frac{d\alpha}{dt} = \frac{3K_s}{\frac{1}{(1-\alpha)^{2/3}} + \gamma \left(\frac{1}{(1-\alpha)^{1/3}} - \frac{1}{(1+P\alpha)^{1/3}} \right)}, \quad (6)$$

where $K_s = (k_s C_s)/(R_0 N_0)$, $\gamma = (R_0 k_s)/D_e$, $P = N_0(V_p - V_r)$ with k_s standing for the rate constant of surface reaction, N_0 —initial concentration of the solid reactant, C_s —the surface concentration of gas, R_0 —the initial radius of the grain, V_p and V_r —the molar volume of solid product and reactant, D_e —the diffusivity of Ca^{2+} ions. The more complex, distributed, pore model, additionally uses the pore properties that have to be evaluated in separate structural measurements. Obviously, the application of these mechanistic models is not straightforward. Moreover, there are some simplifying assumptions used in their derivation. On the contrary, the neural network model does not need these additional parameters and relies only on the TA data. Comparison of the fit performance by the two mechanistic models and by ANN shows that the latter is most accurate.

Another feature of using ANNs is noticed when analyzing a heterogeneous process of iron oxide reduction [19]. A complicating factor, along with a consecutive chemical reaction ($\text{Fe}_2\text{O}_3 \rightarrow \text{Fe}_3\text{O}_4 \rightarrow \text{FeO}$), is the structural changes of the interface throughout the reaction (sintering, cracking etc.). As a result, the conversion degree shows non-monotonic behavior, i.e., the reduction rate is delayed at temperatures about 750 °C (Figure 10). Authors note that this effect was found by other workers. The benefit of the ANN model is that it describes all the data, including this effect, whereas the single traditional mechanistic model fails to do that. Consideration of various reducing gas compositions shows that ANN also permits fitting the reaction course both in the kinetic and diffusion regimes. Accomplishing the same with the mechanistic models requires using two different models.

An idea of using the neural networks primarily for determination of the kinetic triplets has been put forward by Sbirrazzuoli et al. [9] and further developed by Conesa et al. [56], Muravyev and Pivkina [21], and other more recent workers [61,94–97]. As the approach

develops, the complexity of the reaction domain has been increasing. The initiatory work [9] has employed for training and testing the second-order reaction models with activation energy $74\text{--}80\text{ kJ mol}^{-1}$, preexponential factor $\ln(A, \text{s}^{-1}) = 18\text{--}20$. Yet, the recent study [21] has considered ten reaction models with a large span of the activation energy and preexponential factor values (Figure 11).

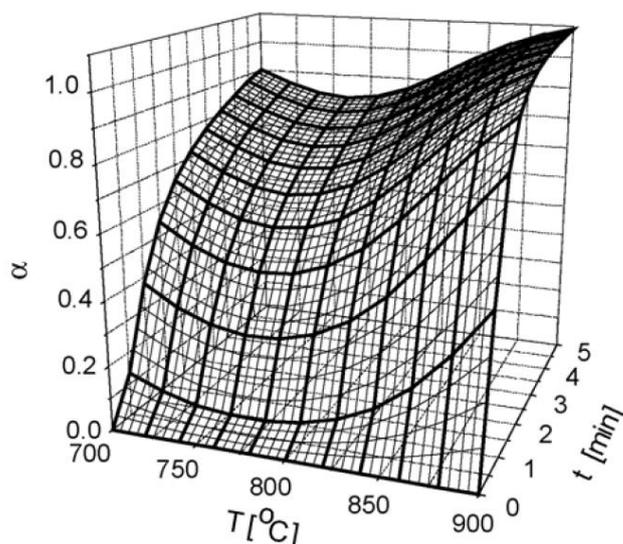


Figure 10. ANN (3-20-1) representation of the conversion degree for Fe_2O_3 reduction in $90\%\text{N}_2 + 5\%\text{CO} + 5\%\text{H}_2$ gas flow [19]. Reprinted with permission from Elsevier.

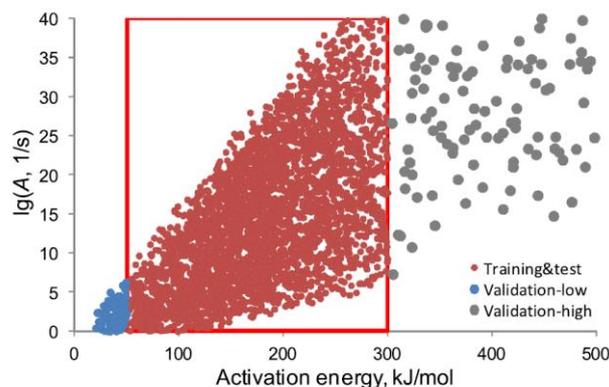


Figure 11. The plot populated with the (preexponential factor, activation energy) pairs used for training and testing of ANNs (red points), and validation of the results (gray and blue points) [21]. Reprinted with permission from Elsevier.

The general idea of the approach is illustrated by Figure 12. The input vector is a set of sparse TA data (e.g., 20 points [56] or 49 points [21] for each run) at several heating rates. The output variables are the activation energy E_a , preexponential factor A , and reaction order n . Conesa et al. [56] also employ ANN to evaluate the kinetic triplets when using the pyrolysis data of lignin, cellulose, and polyethylene as input. These triplets when inserted in Equation (1) give rise to the TGA data that match closely the experimental mass loss profiles. Thus, the neural network here is utilized as an alternative to traditional nonlinear optimization procedure.

Aghbashlo et al. [96,97] adopt the same approach but utilize a more advanced mathematical apparatus. The authors first use the principal component analysis to reduce the number of input variables from six (C, H, N, O, S content, and heating rate) to three. Then these transformed input data are fed into an adaptive neuro-fuzzy inference system (ANFIS) to produce the output kinetic parameters E_a , A , n . ANFIS [98] represents a

combination of neural network model with a fuzzy logic (FL) system. Additionally, the authors use the genetic algorithm (GA) to optimize the ANFIS parameters. The developed model results in the kinetic parameters that agree well with those obtained by traditional optimization procedure for five various biomass feedstocks (bamboo, *Phyllanthus emblica* kernel, *Musa balbisiana*, cattle manure, peanut shell). Although the computational approach from the study (ANN+FL+GA) is rather advanced, the major limitation is the constrained training that is limited to single heating rate and reaction order model.

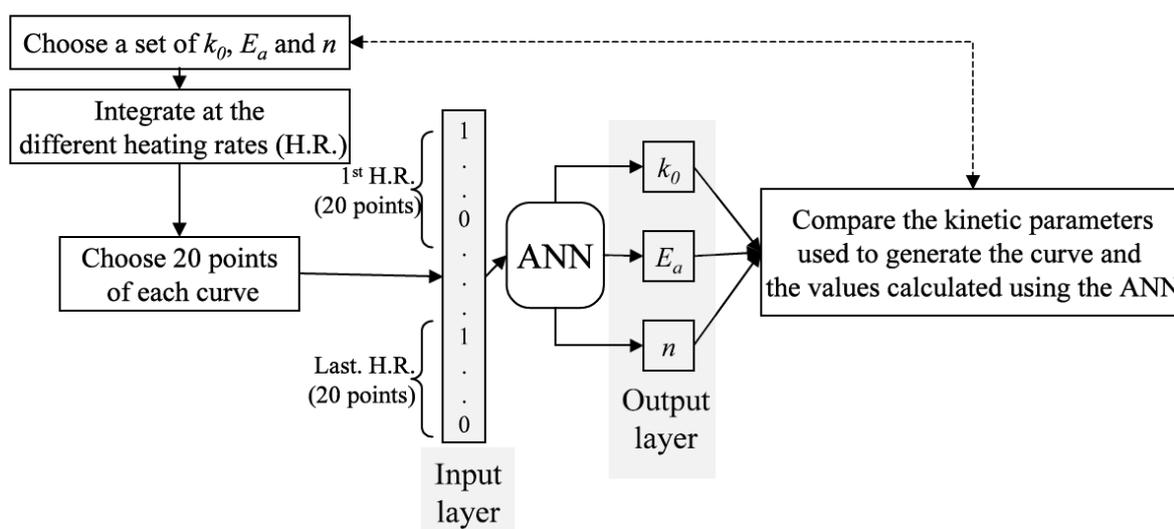


Figure 12. The framework of thermokinetic analysis with ANN proposed in Ref. [56]. Reprinted with permission from Elsevier.

Muravyev and Pivkina [21] extend the above-described approach [56] by considering a broad range of the activation energy and preexponential factor pairs (Figure 11). Additionally, instead of n -th order reaction model, 10 other reaction models are taken into account. Table 4 summarizes the results of their study. It is found that the best accuracy (determined as ANN performance out of the domain space used for training and testing, Figure 11) is achieved when the specific ANN is trained to determine only one component of the kinetic triplet (E_a , A , $f(\alpha)$). Additionally, it is shown that the suggested approach confirms some previously established facts: (1) failure of the kinetic analysis based on a single heating rate data [28]; (2) enhanced capability of controlled rate thermal analysis [99] to discriminate among the reaction models. Conceptually, since this approach involves the synthetic data as input, it paves the way to the reinforcement learning instead of the supervised learning for neural network. That is, one routine generates thermal analysis profiles for known kinetic parameters, while another (ANN) uses this large amount of data for training.

Huang et al. [94] have adopted the same approach, but further increased the number of reaction types to 17, and utilized a more advanced mathematical tool, the general regression neural network. Most importantly, the authors employ datasets of simulated nonisothermal experiments at several different heating rates. They stress that only at three heating rates one can achieve an acceptable recognition accuracy for the reaction model and that the accuracy improves significantly when the number of experiments (heating rates) used in ANN training is five. The relative error in the output preexponential factor and activation energy values is as low as 4% for the simulated data. Additionally, the authors applied the developed neural network to the experimental data for $\text{Li}_4\text{Ti}_5\text{O}_{12}/\text{C}$ high-temperature reaction. The ANN output agrees well with the literature results [100] of these data analysis accomplished with isoconversional and model-fitting kinetic approaches. It is noteworthy that according to the isoconversional analysis, the activation energy of the process is more or less constant throughout the reaction, thus, indicating the single-step reaction kinetics.

Table 4. ANN details and performance of models developed for kinetic analysis in Ref. [21]. Reprinted with permission from Elsevier.

#	Input	N_i	Output	N_o	Cases	N_h	$P, \%$		AC, %
	Parameters		Parameters				Training	Test	
1	$T(\alpha)$ B = 5 K/min	49	Reaction model	6	3278	15	99.1	98.2	88
2	$T(\alpha)$ B = 5 K/min	49	E_a	1	6106	32	93.8	92.2	62
3	$T(\alpha)$ $\beta = 0.5$, 20 K/min	98	Reaction model	6	199	18	99.8	99.4	100
4	$T(\alpha)$ $\beta = 0.5$, 20 K/min	98	E_a	1	147	14	99.9	99.9	98
5	$T(\alpha)$ $\beta = 0.5$, 20 K/min	98	lgA	1	139	10	99.9	99.9	98
6	$T(\alpha)$ $\beta = 0.5$, 20 K/min	98	E_a , lgA	2	147	13	99.9	99.9	98(E), 92(A)
7 ¹	$T(\alpha)$ $\beta = 0.5$, 20 K/min	50	Reaction model	10	1109	20	99.8	99.4	100
8 ¹	$T(\alpha)$ $\beta = 0.5, 5,$ 20 K/min	75	Reaction model	10	483	25	99.8	99.1	100
9	$T(\alpha)$ $c = 1 \cdot 10^{-3} \text{ min}^{-1}$	49	Reaction model	6	851	27	90.1	84.3	65
10 ²	$T(\alpha)$ $c = 1 \cdot 10^{-3} \text{ min}^{-1}$	49	Reaction model	6	678	15	99.9	99.9	100
11 ²	$T(\alpha)$ $c = 1 \cdot 10^{-3} \text{ min}^{-1}$	49	Reaction model, E_a , lgA	8	678	21	99.6	99.2	81 (E), 44(A), 100 (R)

Notes: N_i , N_h and N_o are the number of neurons in input, hidden, and output layer of network, P and AC —the performance and accuracy of the respective ANN. ¹ For these models, the reaction types (B1, D2, D3, F2, L2, F1, P2, R3, A2, A3) were used, and the step for conversion degree $\Delta\alpha = 0.05$ was applied instead of $\Delta\alpha = 0.02$. ² For these models, the reaction types (F1, A2, A3, D1, D3, B1) were used instead of (F1, R2, A3, F2, D3, B1).

Kuang and Xu [95] have implemented the most advanced mathematical approach to this problem, i.e., the one-dimensional convolutional neural network (CNN). As in the previously considered work [94], the single neural network was utilized to estimate the whole kinetic triplet. The application of the trained CNN model for estimating the preexponential factor and activation energy has resulted in less than 3% error in case of using a dataset of three heating rates for training. Remarkably, the authors have used noisy input data to compare the CNN model with the one used in a previous study, the feedforward neural network (more specifically, multi linear perceptron, MLP). Figure 13 displays that with increasing the amplitude of the Gaussian noise, the performance of MLP starts decreasing, which does not apply to the CNN model. Obviously, the convolutional neural network shows superior robustness toward the noisy data. Finally, the model has been applied to the pine wood pyrolysis data. The details of calculations (e.g., the isoconversional plots) are not provided, but apparently all the discussed kinetic data are associated with the first mass loss stage (corresponds to ~85% of total mass loss). The kinetic parameters obtained with the CNN agree well with those computed by two isoconversional methods, Kissinger–Akahira–Sunose and Flynn–Wall–Ozawa methods.

It should be mentioned that ANN does not always yield the results that agree well with the isoconversional ones. These two approaches have been compared as applied to the kinetics in the system lumefantrine/molecularly imprinted polymer [101]. The isoconversional method has demonstrated that in the range of conversions from 0.05 to 0.90, the activation energy rises from roughly 80 to 150 kJ mol⁻¹. This is a sure sign of the kinetics being complex, i.e., to involve at least two steps with significantly different activation energies [102]. Yet, the ANN estimates yield the activation energy whose values fluctuate in the range of approximately 160–200 kJ mol⁻¹ depending on the heating rate used in analysis [101]. Of course, one cannot expect ANNs to reveal the multi-step nature of the process kinetic as long as the networks are trained on single-step kinetics. Likewise, the usage of single heating rate ANN analysis and training is a faulty approach to kinetic calculations as already demonstrated in a previous study [28].

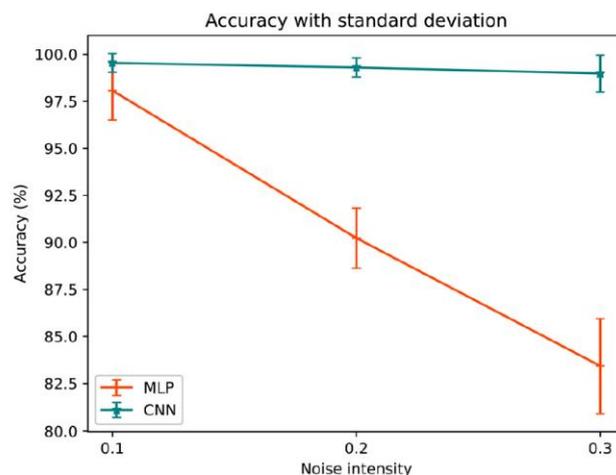


Figure 13. Comparison of the accuracy of convolutional neural network (CNN) and multilinear perceptron (MLP) models as a function of the introduced noise in the input data [95]. Reprinted with permission from Elsevier.

A principally different attempt to perform kinetic analysis with ANNs has been made in a series of papers by Sebastiao and coworkers. The general idea is to utilize the neurons in the form of reaction model functions. The authors apply a similar approach for the description of isothermal processes of widely differing materials, i.e., rhodium (II) acetate [17,18], efavirenz and lamivudine [93], thalidomide [103], polyurethane [104], hybrid of carbon nanotubes with poly(3-hexylthiophene) [105], composite $\text{Li}_4\text{SiO}_4/\text{MgO}$ [106], calcium trihydrate/furosemide system [107], and lumefantrine/molecularly imprinted polymer [101]. To show the basics of this method, we will follow the first paper [18], where the ANN with a single neuron in the hidden layer is considered (Figure 14). The neural network is fed with the time data from an isothermal measurement and outputs the conversion degree value. Given the weights and biases, as shown in Figure 13, and the sigmoidal activation function, the output value is:

$$o = \frac{w_{32}}{1 + e^{-(w_{21}i + w_{20})}} + w_{30}. \quad (7)$$

Equation (7) can be compared with the best mechanistic kinetic model of Prout and Tompkins [108]:

$$\frac{d\alpha}{dt} = k\alpha(1 - \alpha) \rightarrow \ln \frac{\alpha}{1 - \alpha} = kt + C \rightarrow \alpha = \frac{1}{1 + e^{-(kt+C)}}. \quad (8)$$

Using the similarity of the Equations (7) and (8), the logarithm of the optimized weight values w_{21} for each run is plotted against the reciprocal temperature to give the activation energy value. A larger number of the adjustable parameters in Equation (7) as compared to the traditional Equation (8) results in several times smaller residual error for the model.

The idea of using the activation functions to mimic the mechanistic reaction models is further advanced in another study by Sebastiao et al. [17]. More neurons have been added to the hidden layer, but their activation functions are set in an unusual, for ANNs, way. These activation functions have been selected to represent various kinetic models, e.g.:

$$[-\ln(1 - \alpha)]^{1/m} = kt + C, \quad (9)$$

for Kolmogorov–Johnson–Mehl–Avrami–Erofeev nucleation-growth reaction [108], or:

$$1 - \frac{2\alpha}{3} - (1 - \alpha)^{\frac{2}{3}} = kt + C, \quad (10)$$

for Ginstling–Brounshtein three-dimensional diffusion model [108]. That is, first the ideal reaction model (e.g., one of Equations (8)–(10)) is optimized on the experimental data, then the resulting parameters (k , C) are set as weight and bias for the respective neuron (e.g., as w_{21} and w_{20} for neuron designated as “2” in Figure 14a). Then, instead of the usual training of ANN using the back-propagation algorithm, the Levenberg–Marquardt optimization is performed with weights and biases for the hidden layer (i.e., w_{21} and w_{20}) fixed, but adjusting the interconnection weights in the output layer (e.g., w_{32} in Figure 14a, note, that here no biases like w_{30} are considered for the outer neuron). From these weights the relative contribution of the individual reaction models to the whole process can be obtained. By way of example, Figure 14b represents such contributions for several isothermal experiments on the thermal decomposition of lamivudine [93]. Here, the three-dimensional diffusion model (Equation (8)) provides the main contribution to the process. Overall, the above discussed approach can be called as kinetic deconvolution by neural network.

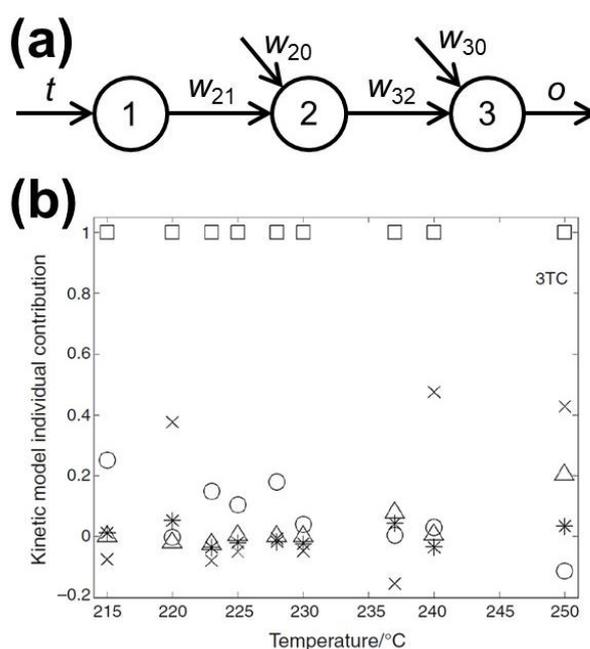


Figure 14. Kinetic analysis with ANNs comprising the activation functions for hidden layer neurons having the mathematical form of ideal reaction types. (a) The scheme of simple ANN model used in Ref. [18]. (b) Relative contributions of kinetic models in the course of thermal decomposition of lamivudine, squares depict three-dimensional diffusion model, crosses—linear contraction model, circles—second-order KJMAE nucleation-growth, triangles—Prout–Tompkins, asterisk—first-order reaction [93]. Reprinted by permission from Springer Nature.

As seen from above, the kinetic applications of ANN still are in their early stages of development. There remains much to be done for these applications to match the capabilities of the well-established isoconversional and model-fitting techniques [28,102]. In regard to this, there are two important principles that must be followed by any computational technique so that it can produce reliable kinetic triplets. First, in its computations, a technique must use simultaneously multiple heating rates or, more generally, multiple temperature programs [28]. For kinetic applications of ANN, it means that the network training sets must include data obtained at more than one heating rate. A failure to produce reliable kinetic triplets has already been demonstrated [21] for ANN trained on the data sets obtained at single heating rates. This problem is yet another manifestation of the fundamental failure of the kinetic techniques based on single heating rate [109]. In the case of the traditional isoconversional and model-fitting techniques, reliable kinetic evaluations require one to use simultaneously no less than 3–4 heating rates. The same principle should be adhered to when training and using ANNs for kinetic applications.

Second, a computational technique must be capable of detecting and treating multi-step kinetics [102]. The reality of thermally stimulated processes in the condensed phase is such that they typically include more than one kinetic step. However, the literature analysis indicates that, so far, the ANNs used for kinetic purposes have been trained exclusively on single-step kinetics. Such restricted training limits drastically the application area of ANNs. To remove this important limitation, proper strategies for multi-step kinetics training ought to be developed. In the meantime, it seems prudent to support the usage of the ANNs trained on single-step kinetics by verifying whether the kinetics studied is actually a single-step one. The latter is readily confirmed as the absence of any significant variation in the isoconversional activation energy [102].

7. Other Applications Relevant to Thermal Analysis Studies

This section gives a short overview of papers that do not directly focus on thermal analysis data or thermokinetic parameters, but that can potentially find some applications in this field. First, neural networks are used to interpolate and extrapolate the temperature dependency of desired property. Thus, Wunderlich et al. [13] have proposed using ANNs for prediction of the heat capacity (c_p) data. Note that this property is normally measured with the thermal analysis method of DSC, and it is of importance to interpolate and extrapolate the results to particular values of temperature. The experimental heat capacity values for 15 common polymers have been taken at every 10 K over the 110–360 K temperature range. These values serve as inputs for neural network (25-15-10), and the output was c_p at $T = 10, 20, \dots, 100$ K. The reported errors of trained ANN in the c_p data for five polymers from test set are less than 0.8%. Thus, the accuracy of this extrapolation procedure is better than typical experimental errors (1–5%). Another group of applications comprises predictions of thermal properties with ANNs. This topic is vast and is not considered here in detail. It worth mentioning, though, that artificial intelligence methods have been applied for prediction of thermal conductivity [110,111], vapor pressure [112], flammability [113], density [114], thermal stability [115], and mechanical properties [116]. Third group of applications includes some developments in experimental devices that can be potentially applied in future thermal analyzers and relevant instruments. For instance, Gao et al. [117] discuss an optimal iterative controller for nonlinear processes based on ANN model (note, that this topic has been widely studied, e.g., [118,119]). The authors demonstrate that the proposed system offers a better temperature control as compared to a standard PID controller. This method could be used in thermal analysis equipment where the temperature control is critical, e.g., in accelerating rate calorimeters [120].

8. Conclusions

Artificial neural networks (ANNs) are a system of powerful tools for solving data analysis problems in a broad variety of applications. Regarding the applications considered in the present review, they appear to be in an initiatory stage and use primarily the feed-forward neural networks as a computational tool. Beyond ANNs, some other artificial intelligence approaches have been sporadically applied to the TA problems and are worth mentioning. They include the expert analysis of thermogravimetric data [121], genetic approach [122] for determining kinetic parameters, modeling with adaptive neuro-fuzzy inference system [96,123,124] to predict the mass loss data, extreme gradient boosting algorithm [125] for product yield evaluation. A wider application of machine learning in these areas is expected as the computational tools become more readily available.

Our review indicates that the most popular ANN applications are associated with predicting pyrolysis product conversion. This popularity is driven by an obvious practical importance of the problem. Its solution requires carefully accounting for the TA data complexity via employing ANNs of an increasingly complex topology. In general, the performance of ANN models is quite good when assessed within the domain close to that used for the model development. However, evaluation of the performance outside such

domain is almost absent in the literature. Studies on the extrapolation capabilities of the ANN models should certainly be encouraged.

Neural network models can also be applied to analyze the TA data directly as shown in examples of utilizing ANNs for identifying milk samples and probing compatibility of drug excipients. These applications can be readily extended to a large variety of multi-component materials, whose characteristic TA parameters change discretely with composition. Such applications are of obvious relevance to composition optimization and quality control. This type of usage of ANNs in assaying the TA data is definitely worthy of further exploration, especially in combination with some approaches from chemometrics, spectroscopy, and other fields such as filtering, baseline correction, deconvolution, classification, and so on.

In the area of the kinetic applications of ANNs, one should clearly differentiate between two groups of studies: (i) those that use single heating rate or single isothermal run for deriving the Arrhenius parameters; and (ii) those that derive the Arrhenius parameters via simultaneously using multiple heating rates or multiple temperature programs in general. The results of the first group must be treated with great caution because it is now established that single temperature program data are fundamentally incapable of yielding reliable kinetic parameters regardless of the computational technique used. A failure of such approach has already been demonstrated by comparing ANNs employing single and multiple heating rates. Only the multiple heating rates approach has proved to be correct, and this must be kept in mind while developing further kinetic applications of ANNs. Another crucial trait that must be implemented in future developments is the ability to recognize and treat the multi-step kinetics.

Overall, the ANN applications present a growing trend in the area covered by this review. ANN's provide newer more sophisticated and flexible mathematical tools that are poised to accomplish a greater level of detail in the process description. Nevertheless, a greater level of detail does not necessarily translate to qualitatively deeper levels of understanding. Indeed, the latter are the true milestones of scientific progress. Only time will tell if ANNs spurred any considerable progress in the field under review. On the other hand, one must realize that no progress can ever be accomplished without trying new approaches and tools.

Author Contributions: Conceptualization, N.V.M., H.L.O.J., R.S. and G.L.; investigation, H.L.O.J., R.S. and G.L.—original draft preparation, N.V.M.; writing—review and editing, S.V. All authors have read and agreed to the published version of the manuscript.

Funding: N.V.M. acknowledges the core funding from the Russian Federal Ministry of Science and Higher Education (0082-2018-0002, AAAA-A18-118031490034-6).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

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

References

1. Wilbraham, L.; Mehr, S.H.M.; Cronin, L. Digitizing Chemistry Using the Chemical Processing Unit: From Synthesis to Discovery. *Acc. Chem. Res.* **2021**, *54*, 253–262. [[CrossRef](#)]
2. Walters, W.P.; Barzilay, R. Applications of Deep Learning in Molecule Generation and Molecular Property Prediction. *Acc. Chem. Res.* **2021**, *54*, 263–270. [[CrossRef](#)] [[PubMed](#)]
3. Tkatchenko, A. Machine learning for chemical discovery. *Nat. Commun.* **2020**, *11*, 4125. [[CrossRef](#)]
4. Yamada, H.; Liu, C.; Wu, S.; Koyama, Y.; Ju, S.; Shiomi, J.; Morikawa, J.; Yoshida, R. Predicting Materials Properties with Little Data Using Shotgun Transfer Learning. *ACS Cent. Sci.* **2019**, *5*, 1717–1730. [[CrossRef](#)] [[PubMed](#)]
5. Sieniutycz, S.; Szwast, Z. Neural Networks—A Review of Applications. In *Optimizing Thermal, Chemical, and Environmental Systems*; Elsevier: Amsterdam, The Netherlands, 2018; pp. 109–120; ISBN 978-0-12-813582-2.
6. Xu, A.; Chang, H.; Xu, Y.; Li, R.; Li, X.; Zhao, Y. Applying artificial neural networks (ANNs) to solve solid waste-related issues: A critical review. *Waste Manag.* **2021**, *124*, 385–402. [[CrossRef](#)]
7. McCulloch, W.S.; Pitts, W. A logical calculus of the ideas immanent in nervous activity. *Bull. Math. Biol.* **1943**, *5*, 115–133. [[CrossRef](#)]

8. Cybenko, G. Approximation by superpositions of a sigmoidal function. *Math. Control. Signals Syst.* **1989**, *2*, 303–314. [[CrossRef](#)]
9. Sbirrazzuoli, N.; Brunel, D. Computational neural networks for mapping calorimetric data: Application of feed-forward neural networks to kinetic parameters determination and signals filtering. *Neural Comput. Appl.* **1997**, *5*, 20–32. [[CrossRef](#)]
10. Egmont-Petersen, M.; de Ridder, D.; Handels, H. Image processing with neural networks—A review. *Pattern Recognit.* **2002**, *35*, 2279–2301. [[CrossRef](#)]
11. Cammarata, L.; Fichera, A.; Pagano, A. Neural prediction of combustion instability. *Appl. Energy* **2002**, *72*, 513–528. [[CrossRef](#)]
12. Cho, S.; No, K.; Goh, E.; Kim, J.; Shin, J.; Joo, Y.; Seong, S. Optimization of Neural Networks Architecture for Impact Sensitivity of Energetic Molecules. *Bull. Korean Chem. Soc.* **2005**, *26*, 399–408. [[CrossRef](#)]
13. Darsey, J.A.; Noid, D.W.; Wunderlich, B.; Tsoukalas, L. Neural-net extrapolations of heat capacities of polymers to low temperatures. *Die Makromol. Chem. Rapid Commun.* **1991**, *12*, 325–330. [[CrossRef](#)]
14. Ventura, S.; Silva, M.; Perez-Bendito, D.; Hervas, C. Multicomponent Kinetic Determinations Using Artificial Neural Networks. *Anal. Chem.* **1995**, *67*, 4458–4461. [[CrossRef](#)] [[PubMed](#)]
15. Bandyopadhyay, J.K.; Annamalai, S.; Gauri, K.L. Application of artificial neural networks in modeling limestone–SO₂ reaction. *AIChE J.* **1996**, *42*, 2295–2302. [[CrossRef](#)]
16. Sbirrazzuoli, N.; Brunel, D.; Elegant, L. Neural networks for kinetic parameters determination, signal filtering and deconvolution in thermal analysis. *J. Therm. Anal. Calorim.* **1997**, *49*, 1553–1564. [[CrossRef](#)]
17. Sebastião, R.C.O.; Braga, J.P.; Yoshida, M.I. Competition between kinetic models in thermal decomposition: Analysis by artificial neural network. *Thermochim. Acta* **2004**, *412*, 107–111. [[CrossRef](#)]
18. Sebastiao, R.C.O.; Braga, J.P.; Yoshida, M.I. Artificial neural network applied to solid state thermal decomposition. *J. Therm. Anal. Calorim.* **2003**, *74*, 811–818. [[CrossRef](#)]
19. Wiltowski, T.; Piotrowski, K.; Lorethova, H.; Stonawski, L.; Mondal, K.; Lalvani, S.B. Neural network approximation of iron oxide reduction process. *Chem. Eng. Process. Process Intensif.* **2005**, *44*, 775–783. [[CrossRef](#)]
20. Bezerra, E.M.; Bento, M.S.; Rocco, J.A.F.F.; Iha, K.; Lourenço, V.L.; Pardini, L.C. Artificial neural network (ANN) prediction of kinetic parameters of (CRFC) composites. *Comput. Mater. Sci.* **2008**, *44*, 656–663. [[CrossRef](#)]
21. Muravyev, N.V.; Pivkina, A.N. New concept of thermokinetic analysis with artificial neural networks. *Thermochim. Acta* **2016**, *637*, 69–73. [[CrossRef](#)]
22. Cavalheiro, É.T.G. Thermal Analysis. In *Reference Module in Chemistry, Molecular Sciences and Chemical Engineering*; Elsevier: Amsterdam, The Netherlands, 2018; ISBN 978-0-12-409547-2.
23. Sun, Y.; Liu, L.; Wang, Q.; Yang, X.; Tu, X. Pyrolysis products from industrial waste biomass based on a neural network model. *J. Anal. Appl. Pyrolysis* **2016**, *120*, 94–102. [[CrossRef](#)]
24. Hough, B.R.; Beck, D.A.; Schwartz, D.T.; Pfaendtner, J. Application of machine learning to pyrolysis reaction networks: Reducing model solution time to enable process optimization. *Comput. Chem. Eng.* **2017**, *104*, 56–63. [[CrossRef](#)]
25. Hua, F.; Fang, Z.; Qiu, T. Application of convolutional neural networks to large-scale naphtha pyrolysis kinetic modeling. *Chin. J. Chem. Eng.* **2018**, *26*, 2562–2572. [[CrossRef](#)]
26. Šesták, J. Ignoring heat inertia impairs accuracy of determination of activation energy in thermal analysis. *Int. J. Chem. Kinet.* **2019**, *51*, 74–80. [[CrossRef](#)]
27. Vyazovkin, S. How much is the accuracy of activation energy affected by ignoring thermal inertia? *Int. J. Chem. Kinet.* **2020**, *52*, 23–28. [[CrossRef](#)]
28. Vyazovkin, S.; Burnham, A.K.; Criado, J.M.; Perez-Maqueda, L.A.; Popescu, C.; Sbirrazzuoli, N. ICTAC Kinetics Committee recommendations for performing kinetic computations on thermal analysis data. *Thermochim. Acta* **2011**, *520*, 1–19. [[CrossRef](#)]
29. Burnham, A.K.; Dinh, L.N. A comparison of isoconversional and model-fitting approaches to kinetic parameter estimation and application predictions. *J. Therm. Anal. Calorim.* **2007**, *89*, 479–490. [[CrossRef](#)]
30. Muravyev, N.V.; Melnikov, I.N.; Monogarov, K.A.; Kuchurov, I.V.; Pivkina, A.N. The power of model-fitting kinetic analysis applied to complex thermal decomposition of explosives: Reconciling the kinetics of bicyclo-HMX thermolysis in solid state and solution. *J. Therm. Anal. Calorim.* **2021**. [[CrossRef](#)]
31. Vyazovkin, S. Modern Isoconversional Kinetics: From Misconceptions to Advances. In *Handbook of Thermal Analysis and Calorimetry*; Elsevier: Amsterdam, The Netherlands, 2018; Volume 6, pp. 131–172; ISBN 978-0-444-64062-8.
32. Sbirrazzuoli, N. Model-free isothermal and nonisothermal predictions using advanced isoconversional methods. *Thermochim. Acta* **2021**, *697*, 178855. [[CrossRef](#)]
33. Kiselev, V.G.; Muravyev, N.V.; Monogarov, K.A.; Griбанov, P.S.; Asachenko, A.F.; Fomenkov, I.V.; Goldsmith, C.F.; Pivkina, A.N.; Gritsan, N.P.; Asachenko, A.F. Toward reliable characterization of energetic materials: Interplay of theory and thermal analysis in the study of the thermal stability of tetranitroacetimidic acid (TNA). *Phys. Chem. Chem. Phys.* **2018**, *20*, 29285–29298. [[CrossRef](#)]
34. Koga, N. Physico-Geometric Approach to the Kinetics of Overlapping Solid-State Reactions. In *Handbook of Thermal Analysis and Calorimetry*; Elsevier: Amsterdam, The Netherlands, 2018; Volume 6, pp. 213–251; ISBN 978-0-444-64062-8.
35. Opfermann, J.; Hädrich, W. Prediction of the thermal response of hazardous materials during storage using an improved technique. *Thermochim. Acta* **1995**, *263*, 29–50. [[CrossRef](#)]
36. Rudin, C. Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. *Nat. Mach. Intell.* **2019**, *1*, 206–215. [[CrossRef](#)]

37. Novak, R.; Bahri, Y.; Abolafia, D.A.; Pennington, J.; Sohl-Dickstein, J. Sensitivity and Generalization in Neural Networks: An Empirical Study. *arXiv* **2018**, arXiv:1802.08760.
38. Psychogios, D.C.; Ungar, L.H. A hybrid neural network-first principles approach to process modeling. *AIChE J.* **1992**, *38*, 1499–1511. [[CrossRef](#)]
39. Wolpert, D.H. The Lack of a Priori Distinctions between Learning Algorithms. *Neural Comput.* **1996**, *8*, 1341–1390. [[CrossRef](#)]
40. Farah, J.S.; Cavalcanti, R.N.; Guimarães, J.T.; Balthazar, C.F.; Coimbra, P.T.; Pimentel, T.C.; Esmerino, E.A.; Duarte, M.C.K.; Freitas, M.Q.; Granato, D.; et al. Differential scanning calorimetry coupled with machine learning technique: An effective approach to determine the milk authenticity. *Food Control* **2021**, *121*, 107585. [[CrossRef](#)]
41. Ali, J.M.; Hussain, M.; Tade, M.O.; Zhang, J. Artificial Intelligence techniques applied as estimator in chemical process—A literature survey. *Expert Syst. Appl.* **2015**, *42*, 5915–5931. [[CrossRef](#)]
42. Da Silva, I.N.; Hernane Spatti, D.; Andrade Flauzino, R.; Liboni, L.H.B.; dos Reis Alves, S.F. *Artificial Neural Networks. A Practical Course*; Springer: Cham, Switzerland, 2017; ISBN 978-3-319-43161-1.
43. Graupe, D. *Principles of Artificial Neural Networks: Basic Designs to Deep Learning*; World Scientific: Singapore, 2019; ISBN 9789811201226.
44. Çepeliogullar, Ö.; Mutlu, I.; Yaman, S.; Haykiri-Acma, H. A study to predict pyrolytic behaviors of refuse-derived fuel (RDF): Artificial neural network application. *J. Anal. Appl. Pyrolysis* **2016**, *122*, 84–94. [[CrossRef](#)]
45. Naqvi, S.R.; Hameed, Z.; Tariq, R.; Taqvi, S.A.; Ali, I.; Niazi, M.B.; Noor, T.; Hussain, A.; Iqbal, N.; Shahbaz, M. Synergistic effect on co-pyrolysis of rice husk and sewage sludge by thermal behavior, kinetics, thermodynamic parameters and artificial neural network. *Waste Manag.* **2019**, *85*, 131–140. [[CrossRef](#)]
46. Ghiba, L.; Drăgoi, E.N.; Curteanu, S. Neural network-based hybrid models developed for free radical polymerization of styrene. *Polym. Eng. Sci.* **2021**, *61*, 716–730. [[CrossRef](#)]
47. Liu, Y.-P.; Wu, M.-G.; Qian, J.-X. Evolving Neural Networks Using the Hybrid of Ant Colony Optimization and BP Algorithms. In *Advances in Neural Networks—ISNN 2006*; Wang, J., Yi, Z., Zurada, J.M., Lu, B.-L., Yin, H., Eds.; Lecture Notes in Computer Science; Springer: Berlin/Heidelberg, Germany, 2006; Volume 3971, pp. 714–722; ISBN 978-3-540-34439-1.
48. Liu, Y.; Wu, M.; Qian, J. Predicting coal ash fusion temperature based on its chemical composition using ACO-BP neural network. *Thermochim. Acta* **2007**, *454*, 64–68. [[CrossRef](#)]
49. Shao, R.; Martin, E.; Zhang, J.; Morris, A. Confidence bounds for neural network representations. *Comput. Chem. Eng.* **1997**, *21*, S1173–S1178. [[CrossRef](#)]
50. Sridhar, D.V.; Seagrave, R.C.; Bartlett, E.B. Process modeling using stacked neural networks. *AIChE J.* **1996**, *42*, 2529–2539. [[CrossRef](#)]
51. Niaei, A.; Towfighi, J.; Khataee, A.R.; Rostamizadeh, K. The Use of ANN and the Mathematical Model for Prediction of the Main Product Yields in the Thermal Cracking of Naphtha. *Pet. Sci. Technol.* **2007**, *25*, 967–982. [[CrossRef](#)]
52. Nielsen, M. Chapter 1: Using Neural Nets to Recognize Handwritten Digits. In *Neural Networks and Deep Learning*; Determination Press: San Francisco, CA, USA, 2019.
53. Vapnik, V.N. *The Nature of Statistical Learning Theory*; Springer: New York, NY, USA, 2000; ISBN 978-1-4419-3160-3.
54. Molga, E. Neural network approach to support modelling of chemical reactors: Problems, resolutions, criteria of application. *Chem. Eng. Process. Process Intensif.* **2003**, *42*, 675–695. [[CrossRef](#)]
55. Agnol, L.D.; Ornaghi, H.L., Jr.; Monticeli, F.; Dias, F.T.G.; Bianchi, O. Polyurethanes synthesized with polyols of distinct molar masses: Use of the artificial neural network for prediction of degree of polymerization. *Polym. Eng. Sci.* **2021**, *61*, 1810–1818. [[CrossRef](#)]
56. Conesa, J.A.; Caballero, J.; Labarta, J.A. Artificial neural network for modelling thermal decompositions. *J. Anal. Appl. Pyrolysis* **2004**, *71*, 343–352. [[CrossRef](#)]
57. Ventura, S.; Silva, M.; Perez-Bendito, D.; Hervas, C. Artificial Neural Networks for Estimation of Kinetic Analytical Parameters. *Anal. Chem.* **1995**, *67*, 1521–1525. [[CrossRef](#)]
58. Casier, B.; Carniato, S.; Miteva, T.; Capron, N.; Sisourat, N. Using principal component analysis for neural network high-dimensional potential energy surface. *J. Chem. Phys.* **2020**, *152*, 234103. [[CrossRef](#)]
59. Ravi Kumar, G.; Nagamani, K.; Anjan Babu, G. A Framework of Dimensionality Reduction Utilizing PCA for Neural Network Prediction. In *Advances in Data Science and Management*; Borah, S., Emilia Balas, V., Polkowski, Z., Eds.; Lecture Notes on Data Engineering and Communications Technologies; Springer: Singapore, 2020; Volume 37, pp. 173–180; ISBN 9789811509773.
60. Strange, H.; Zwiggelaar, R. Spectral Dimensionality Reduction. In *An Introduction to Distance Geometry applied to Molecular Geometry*; Springer: Berlin/Heidelberg, Germany, 2014; pp. 7–22.
61. Sunphorka, S.; Chalermssinsuwan, B.; Piumsomboon, P. Artificial neural network model for the prediction of kinetic parameters of biomass pyrolysis from its constituents. *Fuel* **2017**, *193*, 142–158. [[CrossRef](#)]
62. Zhu, Q.; Jones, J.; Williams, A.; Thomas, K. The predictions of coal/char combustion rate using an artificial neural network approach. *Fuel* **1999**, *78*, 1755–1762. [[CrossRef](#)]
63. Bhuyan, N.; Narzari, R.; Baruah, S.M.B.; Kataki, R. Comparative assessment of artificial neural network and response surface methodology for evaluation of the predictive capability on bio-oil yield of *Tithonia diversifolia* pyrolysis. *Biomass Convers. Biorefinery* **2020**. [[CrossRef](#)]
64. Dubdub, I.; Al-Yaari, M. Pyrolysis of Low Density Polyethylene: Kinetic Study Using TGA Data and ANN Prediction. *Polymers* **2020**, *12*, 891. [[CrossRef](#)]

65. Angin, D.; Tiryaki, A.E. Application of response surface methodology and artificial neural network on pyrolysis of safflower seed press cake. *Energy Sources Part A Recover. Util. Environ. Eff.* **2016**, *38*, 1055–1061. [CrossRef]
66. Karaci, A.; Caglar, A.; Aydinli, B.; Pekol, S. The pyrolysis process verification of hydrogen rich gas (H-rG) production by artificial neural network (ANN). *Int. J. Hydrogen Energy* **2016**, *41*, 4570–4578. [CrossRef]
67. Carsky, M.; Kuwornoo, D.K. Neural network modelling of coal pyrolysis. *Fuel* **2001**, *80*, 1021–1027. [CrossRef]
68. Al-Yaari, M.; Dubdub, I. Application of Artificial Neural Networks to Predict the Catalytic Pyrolysis of HDPE Using Non-Isothermal TGA Data. *Polymers* **2020**, *12*, 1813. [CrossRef]
69. Arumugasamy, S.K.; Selvarajoo, A. Feedforward Neural Network Modeling of Biomass Pyrolysis Process for Biochar Production. *Chem. Eng. Trans.* **2015**, *45*, 1681–1686. [CrossRef]
70. Li, X.H.; Fan, Y.S.; Cai, Y.X.; Zhao, W.D.; Yin, H.Y. Optimization of Biomass Vacuum Pyrolysis Process Based on GRNN. *Appl. Mech. Mater.* **2013**, *411–414*, 3016–3022. [CrossRef]
71. Bi, H.; Wang, C.; Lin, Q.; Jiang, X.; Jiang, C.; Bao, L. Combustion behavior, kinetics, gas emission characteristics and artificial neural network modeling of coal gangue and biomass via TG-FTIR. *Energy* **2020**, *213*, 118790. [CrossRef]
72. Sathiya Prabhakaran, S.P.; Swaminathan, G.; Joshi, V.V. Thermogravimetric analysis of hazardous waste: Pet-coke, by kinetic models and Artificial neural network modeling. *Fuel* **2021**, *287*, 119470. [CrossRef]
73. Bi, H.; Wang, C.; Jiang, X.; Jiang, C.; Bao, L.; Lin, Q. Prediction of mass loss for sewage sludge-peanut shell blends in thermogravimetric experiments using artificial neural networks. *Energy Sources Part A Recover. Util. Environ. Eff.* **2020**, 1–14. [CrossRef]
74. Chen, J.; Xie, C.; Liu, J.; He, Y.; Xie, W.; Zhang, X.; Chang, K.; Kuo, J.; Sun, J.; Zheng, L.; et al. Co-combustion of sewage sludge and coffee grounds under increased O₂/CO₂ atmospheres: Thermodynamic characteristics, kinetics and artificial neural network modeling. *Bioresour. Technol.* **2018**, *250*, 230–238. [CrossRef]
75. Monticeli, F.M.; Neves, R.M.; Júnior, H.L.O. Using an artificial neural network (ANN) for prediction of thermal degradation from kinetics parameters of vegetable fibers. *Cellulose* **2021**, *28*, 1961–1971. [CrossRef]
76. Gu, C.; Wang, X.; Song, Q.; Li, H.; Qiao, Y. Prediction of gas-liquid-solid product distribution after solid waste pyrolysis process based on artificial neural network model. *Int. J. Energy Res.* **2021**, 6707. [CrossRef]
77. Saleem, M.; Ali, I. Machine Learning Based Prediction of Pyrolytic Conversion for Red Sea Seaweed. In Proceedings of the 7th International Conference on Biological, Chemical & Environmental Sciences (BCES-2017), Budapest, Hungary, 6–7 September 2017.
78. Abdurakipov, S.S.; Butakov, E.B.; Burdukov, A.P.; Kuznetsov, A.V.; Chernova, G.V. Using an Artificial Neural Network to Simulate the Complete Burnout of Mechanoactivated Coal. *Combust. Explos. Shock. Waves* **2019**, *55*, 697–701. [CrossRef]
79. Cao, H.; Xin, Y.; Yuan, Q. Prediction of biochar yield from cattle manure pyrolysis via least squares support vector machine intelligent approach. *Bioresour. Technol.* **2016**, *202*, 158–164. [CrossRef] [PubMed]
80. Çepeliogullar, Ö.; Mutlu, I.; Yaman, S.; Haykiri-Acma, H. Activation energy prediction of biomass wastes based on different neural network topologies. *Fuel* **2018**, *220*, 535–545. [CrossRef]
81. Aydinli, B.; Caglar, A.; Pekol, S.; Karaci, A. The prediction of potential energy and matter production from biomass pyrolysis with artificial neural network. *Energy Explor. Exploit.* **2017**, *35*, 698–712. [CrossRef]
82. Burgaz, E.; Yazici, M.; Kapusuz, M.; Alisir, S.H.; Ozcan, H. Prediction of thermal stability, crystallinity and thermomechanical properties of poly(ethylene oxide)/clay nanocomposites with artificial neural networks. *Thermochim. Acta* **2014**, *575*, 159–166. [CrossRef]
83. Himmelblau, D.M. Applications of artificial neural networks in chemical engineering. *Korean J. Chem. Eng.* **2000**, *17*, 373–392. [CrossRef]
84. Molga, E.J.; van Woezik, B.A.A.; Westerterp, K.R. Neural networks for modelling of chemical reaction systems with complex kinetics: Oxidation of 2-octanol with nitric acid. *Chem. Eng. Process. Process Intensif.* **2000**, *39*, 323–334. [CrossRef]
85. Von Stosch, M.; Oliveira, R.; Peres, J.; de Azevedo, S.F. Hybrid semi-parametric modeling in process systems engineering: Past, present and future. *Comput. Chem. Eng.* **2014**, *60*, 86–101. [CrossRef]
86. Bing, G. Modelling coal gasification with a hybrid neural network. *Fuel* **1997**, *76*, 1159–1164. [CrossRef]
87. Roduit, B.; Borgeat, C.; Berger, B.; Folly, P.; Alonso, B.; Aebischer, J.N. The prediction of thermal stability of self-reactive chemicals: From milligrams to tons. *J. Therm. Anal. Calorim.* **2005**, *80*, 91–102. [CrossRef]
88. Agarwal, M. Combining neural and conventional paradigms for modelling, prediction and control. *Int. J. Syst. Sci.* **1997**, *28*, 65–81. [CrossRef]
89. Rojek, B.; Suchacz, B.; Wesolowski, M. Artificial neural networks as a supporting tool for compatibility study based on thermogravimetric data. *Thermochim. Acta* **2018**, *659*, 222–231. [CrossRef]
90. Kohonen, T. *Self-Organizing Maps*; Springer Series in Information Sciences; Springer: Berlin/Heidelberg, Germany, 2001; Volume 30; ISBN 978-3-540-67921-9.
91. Strzelczak, A. The application of artificial neural networks (ANN) for the denaturation of meat proteins—The kinetic analysis method. *Acta Sci. Pol. Technol. Aliment.* **2015**, *18*, 87–96. [CrossRef]
92. Charde, S.J. Degradation Kinetics of Polycarbonate Composites: Kinetic Parameters and Artificial Neural Network. *Chem. Biochem. Eng. Q.* **2018**, *32*, 151–165. [CrossRef]
93. Ferreira, B.D.L.; Araujo, B.C.R.; Sebastião, R.C.O.; Yoshida, M.I.; Mussel, W.N.; Fialho, S.; Barbosa, J. Kinetic study of anti-HIV drugs by thermal decomposition analysis: A multilayer artificial neural network propose. *J. Therm. Anal. Calorim.* **2017**, *127*, 577–585. [CrossRef]

94. Huang, Y.W.; Chen, M.Q.; Li, Q.H. Artificial neural network model for the evaluation of chemical kinetics in thermally induced solid-state reaction. *J. Therm. Anal. Calorim.* **2019**, *138*, 451–460. [[CrossRef](#)]
95. Kuang, D.; Xu, B. Predicting kinetic triplets using a 1d convolutional neural network. *Thermochim. Acta* **2018**, *669*, 8–15. [[CrossRef](#)]
96. Aghbashlo, M.; Almasi, F.; Jafari, A.; Nadian, M.H.; Soltanian, S.; Lam, S.S.; Tabatabaei, M. Describing biomass pyrolysis kinetics using a generic hybrid intelligent model: A critical stage in sustainable waste-oriented biorefineries. *Renew. Energy* **2021**, *170*, 81–91. [[CrossRef](#)]
97. Aghbashlo, M.; Tabatabaei, M.; Nadian, M.H.; Davoodnia, V.; Soltanian, S. Prognostication of lignocellulosic biomass pyrolysis behavior using ANFIS model tuned by PSO algorithm. *Fuel* **2019**, *253*, 189–198. [[CrossRef](#)]
98. Jang, J.-S. ANFIS: Adaptive-network-based fuzzy inference system. *IEEE Trans. Syst. Man Cybern.* **1993**, *23*, 665–685. [[CrossRef](#)]
99. Rouquerol, J. Sample-Controlled Thermal Analysis. In *Reference Module in Chemistry, Molecular Sciences and Chemical Engineering*; Elsevier: Amsterdam, The Netherlands, 2018; p. 9780124095472144000.
100. Hu, X.; Lin, Z.; Yang, K.; Deng, Z. Kinetic Analysis of One-Step Solid-State Reaction for $\text{Li}_4\text{Ti}_5\text{O}_{12}/\text{C}$. *J. Phys. Chem. A* **2011**, *115*, 13413–13419. [[CrossRef](#)] [[PubMed](#)]
101. De Freitas-Marques, M.B.; Araujo, B.C.R.; Da Silva, P.H.R.; Fernandes, C.; Mussel, W.D.N.; Sebastião, R.D.C.D.O.; Yoshida, M.I. Multilayer perceptron network and Vyazovkin method applied to the non-isothermal kinetic study of the interaction between lumefantrine and molecularly imprinted polymer. *J. Therm. Anal. Calorim.* **2020**. [[CrossRef](#)]
102. Vyazovkin, S.; Burnham, A.K.; Favergeon, L.; Koga, N.; Moukhina, E.; Pérez-Maqueda, L.A.; Sbirrazzuoli, N. ICTAC Kinetics Committee recommendations for analysis of multi-step kinetics. *Thermochim. Acta* **2020**, *689*, 178597. [[CrossRef](#)]
103. Ferreira, B.D.L.; Araújo, N.R.S.; Ligório, R.F.; Pujatti, F.J.P.; Mussel, W.N.; Yoshida, M.I.; Sebastião, R.C.O. Kinetic thermal decomposition studies of thalidomide under non-isothermal and isothermal conditions. *J. Therm. Anal. Calorim.* **2018**, *134*, 773–782. [[CrossRef](#)]
104. Ferreira, B.D.D.L.; Araújo, N.R.; Ligório, R.F.; Pujatti, F.J.; Yoshida, M.I.; Sebastião, R.C. Comparative kinetic study of automotive polyurethane degradation in non-isothermal and isothermal conditions using artificial neural network. *Thermochim. Acta* **2018**, *666*, 116–123. [[CrossRef](#)]
105. Ferreira, L.D.L.; Medeiros, F.S.; Araujo, B.C.; Gomes, M.S.; Rocco, M.L.M.; Sebastião, R.C.; Calado, H.D. Kinetic study of MWCNT and MWCNT@P3HT hybrid thermal decomposition under isothermal and non-isothermal conditions using the artificial neural network and isoconversional methods. *Thermochim. Acta* **2019**, *676*, 145–154. [[CrossRef](#)]
106. Vieira, S.S.; Paz, G.M.; Araujo, B.C.; Lago, R.M.; Sebastião, R.C. Use of neural network to analyze the kinetics of CO_2 absorption in $\text{Li}_4\text{SiO}_4/\text{MgO}$ composites from TG experimental data. *Thermochim. Acta* **2020**, *689*, 178628. [[CrossRef](#)]
107. Marques, M.B.D.F.; De Araujo, B.C.R.; Sebastião, R.D.C.D.O.; Mussel, W.D.N.; Yoshida, M.I. Kinetics study and Hirshfeld surface analysis for atorvastatin calcium trihydrate and furosemide system. *Thermochim. Acta* **2019**, *682*, 178408. [[CrossRef](#)]
108. Khawam, A.; Flanagan, D.R. Solid-State Kinetic Models: Basics and Mathematical Fundamentals. *J. Phys. Chem. B* **2006**, *110*, 17315–17328. [[CrossRef](#)] [[PubMed](#)]
109. Vyazovkin, S.; Wight, C.A. Model-free and model-fitting approaches to kinetic analysis of isothermal and nonisothermal data. *Thermochim. Acta* **1999**, *340–341*, 53–68. [[CrossRef](#)]
110. Rahman, M.S.; Rashid, M.; Hussain, M. Thermal conductivity prediction of foods by Neural Network and Fuzzy (ANFIS) modeling techniques. *Food Bioprod. Process.* **2012**, *90*, 333–340. [[CrossRef](#)]
111. Ramezanizadeh, M.; Nazari, M.A.; Ahmadi, M.H.; Lorenzini, G.; Pop, I. A review on the applications of intelligence methods in predicting thermal conductivity of nanofluids. *J. Therm. Anal. Calorim.* **2019**, *138*, 827–843. [[CrossRef](#)]
112. Lazzús, J.A. Prediction of solid vapor pressures for organic and inorganic compounds using a neural network. *Thermochim. Acta* **2009**, *489*, 53–62. [[CrossRef](#)]
113. Asante-Okyere, S.; Xu, Q.; Mensah, R.A.; Jin, C.; Ziggah, Y.Y. Generalized regression and feed forward back propagation neural networks in modelling flammability characteristics of polymethyl methacrylate (PMMA). *Thermochim. Acta* **2018**, *667*, 79–92. [[CrossRef](#)]
114. Shahbaz, K.; Baroutian, S.; Mjalli, F.S.; Hashim, M.A.; AlNashef, I.M. Densities of ammonium and phosphonium based deep eutectic solvents: Prediction using artificial intelligence and group contribution techniques. *Thermochim. Acta* **2012**, *527*, 59–66. [[CrossRef](#)]
115. Lisa, G.; Wilson, D.A.; Curteanu, S.; Lisa, C.; Piuleac, C.-G.; Bulacovschi, V. Ferrocene derivatives thermostability prediction using neural networks and genetic algorithms. *Thermochim. Acta* **2011**, *521*, 26–36. [[CrossRef](#)]
116. Kopal, I.; Harničárová, M.; Valíček, J.; Kušnerová, M. Modeling the Temperature Dependence of Dynamic Mechanical Properties and Visco-Elastic Behavior of Thermoplastic Polyurethane Using Artificial Neural Network. *Polymers* **2017**, *9*, 519. [[CrossRef](#)] [[PubMed](#)]
117. Gao, F.; Wang, F.; Li, M. Neural network-based optimal iterative controller for nonlinear processes. *Can. J. Chem. Eng.* **2000**, *78*, 363–370. [[CrossRef](#)]
118. Chen, L.; Ge, B.; de Almeida, A.T. Self-tuning PID Temperature Controller Based on Flexible Neural Network. In *Advances in Neural Networks—ISNN 2007*; Liu, D., Fei, S., Hou, Z.-G., Zhang, H., Sun, C., Eds.; Lecture Notes in Computer Science; Springer: Berlin/Heidelberg, Germany, 2007; Volume 4491, pp. 138–147. ISBN 978-3-540-72382-0.
119. Khalid, M.; Omatu, S. A neural network controller for a temperature control system. *IEEE Control Syst.* **1992**, *12*, 58–64. [[CrossRef](#)]
120. Townsend, D.I.; Tou, J.C. Thermal hazard evaluation by an accelerating rate calorimeter. *Thermochim. Acta* **1980**, *37*, 1–30. [[CrossRef](#)]

121. Voga, G.P.; Belchior, J.C. An approach for interpreting thermogravimetric profiles using artificial intelligence. *Thermochim. Acta* **2007**, *452*, 140–148. [[CrossRef](#)]
122. Voga, G.P.; Coelho, M.G.; De Lima, G.M.; Belchior, J.C. Experimental and Theoretical Studies of the Thermal Behavior of Titanium Dioxide–SnO₂ Based Composites. *J. Phys. Chem. A* **2011**, *115*, 2719–2726. [[CrossRef](#)]
123. Lerkkasemsan, N.; Achenie, L.E. Pyrolysis of biomass—Fuzzy modeling. *Renew. Energy* **2014**, *66*, 747–758. [[CrossRef](#)]
124. Fazilat, H.; Akhlaghi, S.; Shiri, M.E.; Sharif, A. Predicting thermal degradation kinetics of nylon6/feather keratin blends using artificial intelligence techniques. *Polymer* **2012**, *53*, 2255–2264. [[CrossRef](#)]
125. Pathy, A.; Meher, S.; Balasubramanian, P. Predicting algal biochar yield using eXtreme Gradient Boosting (XGB) algorithm of machine learning methods. *Algal Res.* **2020**, *50*, 102006. [[CrossRef](#)]