



Article

Optimizing Neural Networks for Chemical Reaction Prediction: Insights from Methylene Blue Reduction Reactions

Ivan Malashin , Vadim Tynchenko * , Andrei Gantimurov, Vladimir Nelyub and Aleksei Borodulin

Artificial Intelligence Technology Scientific and Education Center, Bauman Moscow State Technical University, 105005 Moscow, Russia; i.malashin@emtc.ru (I.M.); agantimurov@emtc.ru (A.G.); vladimir.nelub@emtc.ru (V.N.); alexey.borodulin@emtc.ru (A.B.)

* Correspondence: vadimond@mail.ru

Abstract: This paper offers a thorough investigation of hyperparameter tuning for neural network architectures using datasets encompassing various combinations of Methylene Blue (MB) Reduction by Ascorbic Acid (AA) reactions with different solvents and concentrations. The aim is to predict coefficients of decay plots for MB absorbance, shedding light on the complex dynamics of chemical reactions. Our findings reveal that the optimal model, determined through our investigation, consists of five hidden layers, each with sixteen neurons and employing the Swish activation function. This model yields an NMSE of 0.05, 0.03, and 0.04 for predicting the coefficients A, B, and C, respectively, in the exponential decay equation $A + B \cdot e^{-x/C}$. These findings contribute to the realm of drug design based on machine learning, providing valuable insights into optimizing chemical reaction predictions.

Keywords: drug design; chemical reaction prediction; hyperparameter tuning; neural network architecture; methylene blue reduction; machine learning in chemistry



Citation: Malashin, I.; Tynchenko, V.; Gantimurov, A.; Nelyub, V.; Borodulin, A. Optimizing Neural Networks for Chemical Reaction Prediction: Insights from Methylene Blue Reduction Reactions. *Int. J. Mol. Sci.* **2024**, *25*, 3860. <https://doi.org/10.3390/ijms25073860>

Academic Editor: Hao Lin

Received: 10 March 2024

Revised: 24 March 2024

Accepted: 28 March 2024

Published: 29 March 2024



Copyright: © 2024 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

The kinetic assessment of chemical reactions is pivotal in elucidating reaction mechanisms and optimizing processes across diverse domains of chemistry. Among these, the reduction of Methylene Blue (MB) by ascorbic acid (AA) [1,2] stands as a notable reaction, owing to its widespread applications in analytical and industrial contexts, attributed to its sensitivity and specificity. MB is a water-soluble cationic dye molecule [3], readily reduced to leucomethylene blue, and oxidized back to its blue form. The reduction reaction of MB is actively researched by scientists worldwide [1,4,5], reflecting its significance in various fields of study. Despite challenges in employing it as a clock reaction, its kinetics remain a valuable teaching tool, measured spectrophotometrically at $\lambda = 665$ nm, offering insights into rate law dependencies on MB, ascorbic acid, and other concentrations.

The reduction of MB has been scrutinized and extensively discussed in academic literature. The optimal conditions for decolorizing methylene blue using ascorbic acid were determined by Weiland et al. [6] by investigating the wavelength and intensity of illumination. Apples and potatoes were halved, immersed in a methylene blue solution, and exposed to light. Decolorization was observed, revealing the tissue's reducing capacity, with apples showing greatest effect around the core and under the skin, while potatoes exhibited less decolorization, primarily in a band near the skin.

Seno et al. [7] proposed that the reduction of MB with L-ascorbic acid in aqueous surfactant solutions is accelerated by hexadecyltrimethylammonium bromide (HTAB) [8] but not by tetrabutylammonium bromide. Micellar effects, affected by pH and inhibited by KCl, suggest changes in substrate dissociation and product binding to micelles.

MB reduction to leucomethylene blue (MBH) by ascorbic acid in acetonitrile is studied by Hallock et al. [9] using cavity ring-down spectroscopy (CRDS) [10]. This technique allows for precise monitoring of concentration changes on a microsecond timescale,

which was previously unattainable. CRDS utilizes reflective mirrors to form an optical cavity, and changes in light intensity are recorded as the absorber affects the decay time. This method offers significantly increased sensitivity compared to traditional absorption measurements.

The impact of the anionic surfactant sodium dodecyl sulphate (SDS) pre-micellar clusters on the electron transfer reaction between MB and AA in dilute acid conditions was explored by Sen et al. [11]. The reaction, exhibiting first-order kinetics with respect to both MB and AA, involves uncatalyzed and H^+ -catalyzed paths. Iodide ion was found to accelerate the reaction rate.

Simple classroom experiments, like clock reactions, offer accessible avenues to grasp advanced concepts. For instance, Snehalatha et al. [12] demonstrated MB and L-ascorbic acid clock reaction to undergraduates, showcasing kinetics principles with visible color changes. This approach contrasts with the classic "Blue Bottle" experiment and has been expanded to include studies on persulfate-iodide and bromination reactions, elucidating factors like ionic strength and substituent effects on reaction rates.

Recent literature has seen a surge in studies investigating the properties and behaviors of MB through the lens of machine learning (ML) algorithms. Kooh et al. [13] employ supervised ML algorithms to model MB dye adsorption by *Azolla pinnata* [14], aiming for accurate predictions of adsorption capacity across various conditions. SVR-RBF [15] emerges as the top-performing algorithm, achieving an R value of 0.994 with minimal error.

As global regulations tighten on water discharge color, the need for affordable dye removal solutions grows. Marzban et al. [16] investigates MB removal using sodium alginate-kaolin beads, assessing various parameters' impact on efficiency. Characterization and modeling techniques reveal an artificial neural network (ANN) [17] as the most accurate predictor. The developed model aids in optimizing treatment processes, achieving efficient removal of basic dyes. Adsorption kinetics [18] followed a pseudo-second-order model, and encapsulating kaolin in sodium alginate significantly enhanced removal efficiency.

Asfaram et al. [19] investigates zinc sulfide nanoparticles with activated carbon (ZnSNPs-AC) for MB adsorption. Characterization methods and modeling techniques, including RSM [20], ANN, and least squares-support vector machine (LS-SVM), were employed. LS-SVM [21] and ANN models outperformed central composite design (CCD) [22] in predicting MB adsorption efficiency.

Investigations into the adsorptive removal of MB dye from water using different parts of an *Abelmoschus esculentus* (lady's finger) [23] with processed seed powder (LFSP) are described by Nayak et al. [24]. Batch studies evaluated biosorption performance under varying conditions, revealing pseudo-second-order kinetics and intra-particle diffusion as significant factors. Langmuir and Temkin isotherms provided the best fit. Thermodynamic analysis indicated spontaneous and endothermic processes. RSM combined with ANN showed RSM's dominance in both data fitting and estimation capabilities.

Water contamination from artificial dyes in domestic and industrial wastewater poses a significant environmental challenge, necessitating effective adsorbents for dye removal. Soleimani et al. [25] addresses water contamination by utilizing calcium alginate hydrogels reinforced with cellulose nanocrystals (CA/CNC) [25] as an effective and environmentally friendly adsorbent for MB removal. Characterization confirmed CA/CNC's suitability, with adsorption kinetics following the pseudo-first-order model and equilibrium data fitting the Langmuir isotherm model. RSM and an artificial neural network integrated with the whale optimization algorithm (ANN-WOA) [26] were applied for analysis, with ANN-WOA demonstrating superior predictive accuracy. These findings highlight CA/CNC hydrogels as promising bio-adsorbents for addressing MB contamination in water.

Mahmoodi et al. [27] introduces a novel approach by coupling an artificial neural network (ANN) with particle swarm optimization (PSO [28]) to predict MB adsorption by a starch-based superadsorbent. The superadsorbent, synthesized with acrylic acid and acrylamide polymers, was functionalized with catecholamine groups via dopamine polymerization. Comparing the performance, ANN-PSO showed superior accuracy compared

to RSM, and the prediction model for adsorption capacity was exemplified through a set of explicit equations arranged in a matrix format.

This paper introduces a novel approach by coupling an artificial neural network (ANN) with particle swarm optimization (PSO) to predict methylene blue (MB) adsorption by a starch-based superadsorbent. The superadsorbent, synthesized with acrylic acid and acrylamide polymers, was functionalized with catecholamine groups via dopamine polymerization. ANN-PSO outperformed response surface methodology (RSM) in accuracy, with the best ANN topology identified as 3-7-1. The ANN-PSO model yielded lower root-mean-square error (22.46), higher correlation coefficient (0.99), and normalized standard deviation (16.83) compared to RSM (82.89, 0.98, and 65.41, respectively).

Alibak et al. [29] employ cascade correlation neural network (CCNN) [30] modeling to simulate the removal of MB using a bio-based hydrogel. The CCNN, trained with the Levenberg-Marquardt algorithm, predicts MB uptake based on factors such as adsorbent type, temperature, initial dye concentration, pH, and contact time. Results show excellent agreement with experimental data, confirming that MB uptake intensifies with temperature, contact time, pH, and initial dye concentration, while the hydrogel reinforced with bentonite/GO nanoparticles proves to be the most effective adsorbent.

In the landscape of chemistry and engineering, leveraging machine learning (ML) to forecast the coefficients of the rate law governing the interactions involving MB marks a significant advancement. Our study undertakes a thorough examination of the kinetic characteristics associated with the reduction of MB by ascorbic acid within the super diluted solutions (SDSs) of MB, water, and ascorbic acid at various concentrations. By scrutinizing the reaction kinetics across these diverse solvent matrices, our aim is to elucidate the nuanced correlation between solvent composition and reaction rate, ultimately identifying the most sustainable model for predicting rate law coefficients through neural network hyperparameters tuning [31].

2. Results

2.1. Spectroscopic Analysis Results

Based on the conducted measurements, a total of 220 files were obtained, each containing at least 10 spectra corresponding to different combinations of reactions with varying concentrations of mixed substances. Figure 1 illustrates examples of absorption intensity spectra over time for various reactions with differing concentrations.

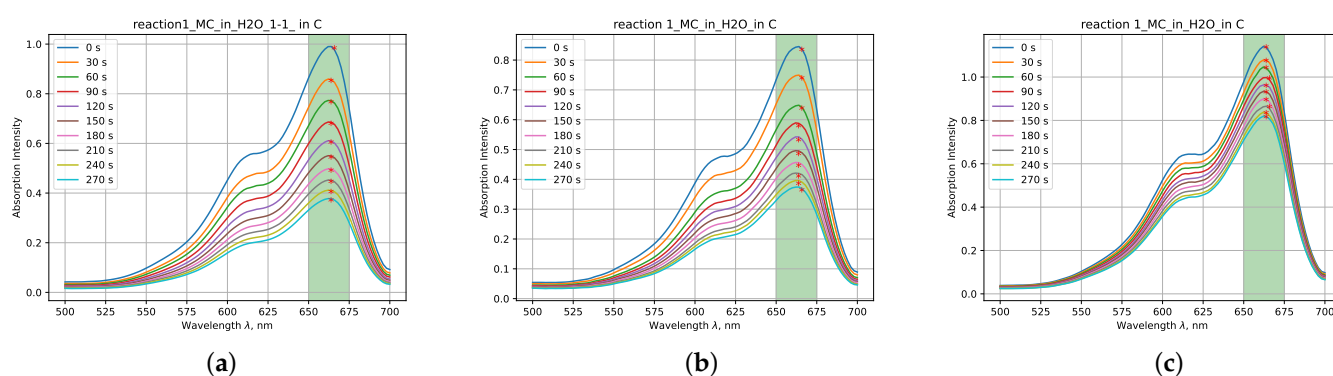


Figure 1. Examples of absorption intensity spectra for reaction MB+H₂O (a–c) with varying volumetric ratios of mixed substances at 1:1, 9:1, and 99:1, respectively.

2.2. Exponential Fitting

As previously mentioned, these spectra were fitted with an exponential function. In some cases, the fitting process failed due to the limited number of measurements, resulting in the library's inability to capture the decaying trend accurately. Such profiles were excluded from further analysis.

Consequently, after filtering, a dataset for training was obtained, consisting of 180 records of conducted reactions. In some instances, the same reaction was conducted multiple times (specifically, six times). Examples of such reactions are illustrated in Figure 2.

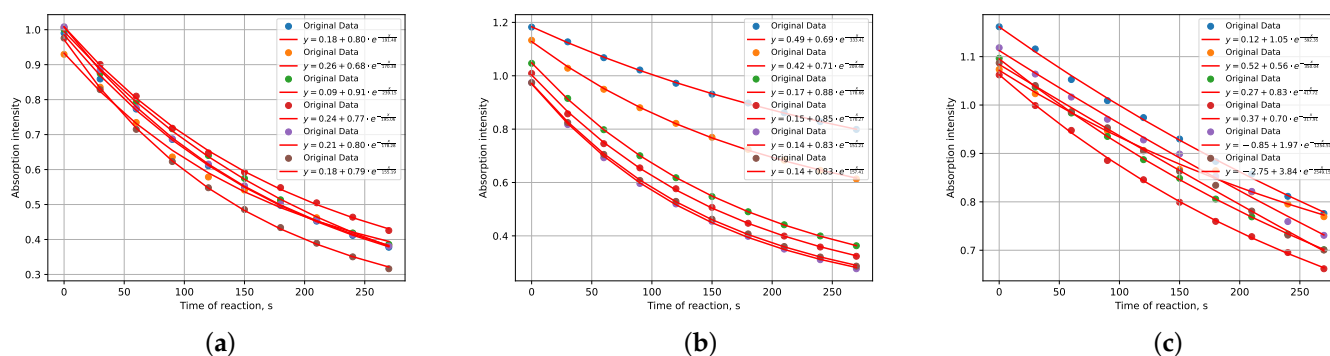


Figure 2. Examples of combined decay profiles with original data of maximum absorption intensity for reactions (a–c) MB+H₂O, conducted six times with varying volumetric ratios of mixed substances at 1:1, 9:1, and 99:1, respectively. (each point represents the maximum absorbance intensity at $\lambda = 665$ nm).

2.3. Hyperparameters Optimization Pipeline

To enhance the efficiency of predicting coefficients A, B, and C in decay profiles, a Python code was developed to fine-tune the architecture [32] and the hyperparameters of a neural network model tailored for regression tasks. The grid search method explores a range of parameters to optimize the neural network architecture. These parameters include the number of hidden layers [33] (1 to 5), the number of neurons per layer [34] (16 to 256), activation functions (ReLU [35], Sigmoid [36], Tanh [37], Leaky ReLU [38], ELU [39], and Swish [40]), training epochs (50 or 100), and batch sizes [41] (16 or 32). Each combination of these parameters is tested to identify the optimal configuration for the neural network model.

For each parameter combination, a neural network model is instantiated with the specified architecture and hyperparameters. The model is trained on a training dataset and evaluated on a separate test dataset. During evaluation, the Normalized Mean Squared Error (NMSE) [42,43] is calculated for each target variable (A, B, C) to assess the model's predictive performance.

Subsequently, the NMSE scores obtained for each parameter combination are stored, along with the corresponding architecture and hyperparameters. These results are then saved in an Excel file, with columns representing the architecture and hyperparameters used for training, as well as the NMSE scores for each target variable.

Throughout this process, the grid search [44] seeks to identify the optimal combination of model architecture and hyperparameters that minimize prediction error for the given regression task. This systematic approach facilitates the exploration of a wide range of model configurations, leading to informed decisions regarding the most effective neural network architecture for the predictive task at hand. Figure 3 illustrates the results of NMSE scores for predicting the coefficients A, B, and C.

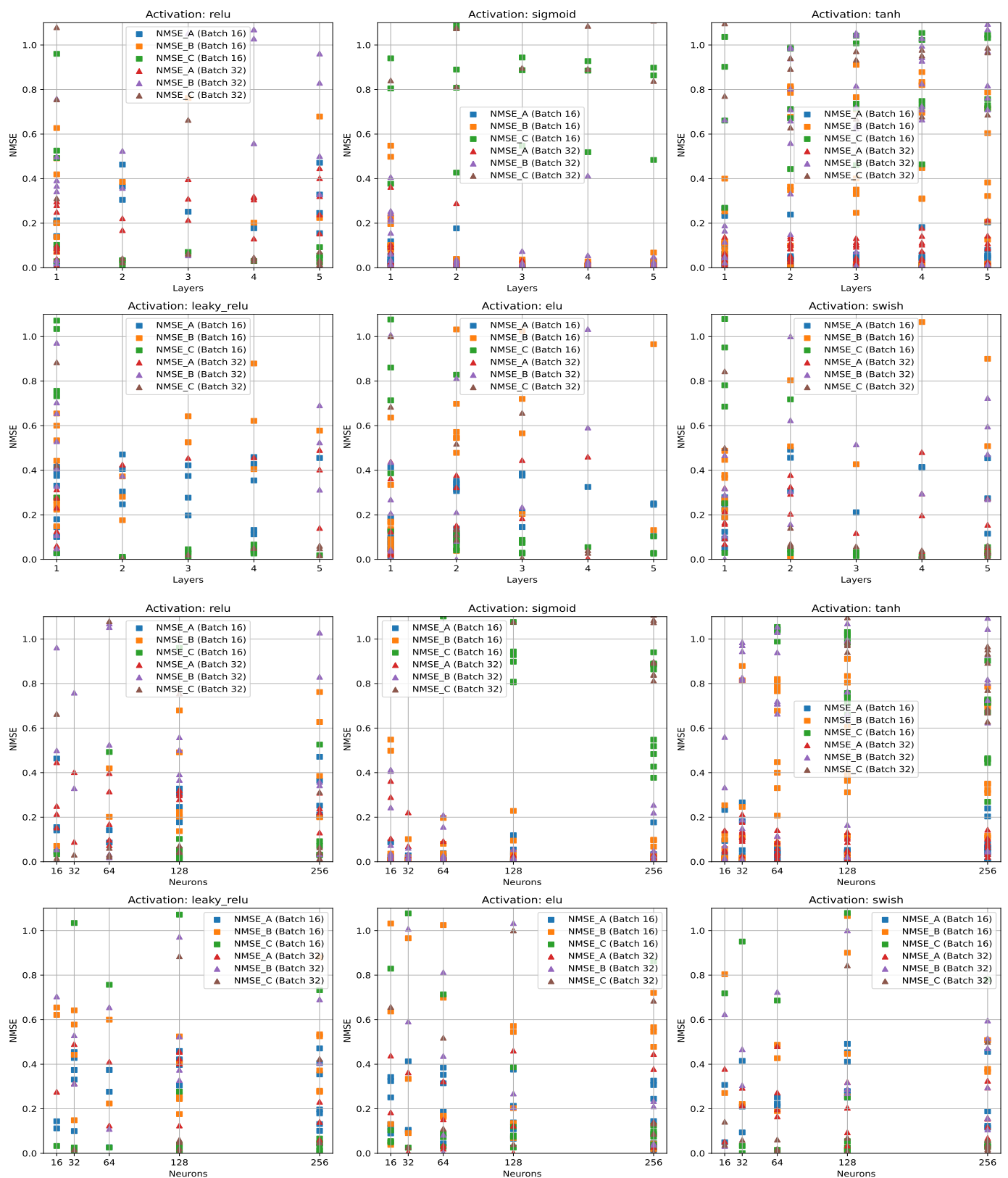


Figure 3. NMSE scores representing coefficients A, B, and C in approximating decay intensity profiles, varying across neural network hyperparameters, are visualized in the plane of NMSE against the number of layers and neurons.

In the course of our study, where we systematically explored parameter combinations using a grid search on our dataset, we identified optimal models with the lowest normalized mean squared error (NMSE) values. The following models demonstrated the best performance:

- A model with 1 hidden layer containing 256 neurons, utilizing the ELU activation function, and trained for 100 epochs. For this model, the NMSE values for coefficients A, B, and C were 0.011391, 0.019457, and 0.124704, respectively.
- A model with 1 hidden layer containing 64 neurons, utilizing the ELU activation function, and trained for 100 epochs. For this model, the NMSE values for coefficients A, B, and C were 0.152417, 0.083142, and 0.110623, respectively.
- A model with 5 hidden layers containing 16 neurons each, utilizing the Swish activation function, and trained for 100 epochs. For this model, the NMSE values for coefficients A, B, and C were 0.053186, 0.033048, and 0.037663, respectively.

All these models were trained using a batch size of 32 and 100 epochs. The performance of each model was evaluated using cross-validation, resulting in variance in the NMSE metrics. Other models exhibiting low NMSE are presented in Table 1.

Table 1. Best Models Performance Summary.

Hidden Layers	Neurons	Activation	Epochs	Batch Size	NMSE_A	NMSE_B	NMSE_C
1	256	ELU	100	16	0.01 ± 0.005	0.02 ± 0.008	0.12 ± 0.03
2	64	ELU	100	32	0.15 ± 0.02	0.08 ± 0.01	0.11 ± 0.02
2	128	Leaky ReLU	100	16	0.25 ± 0.03	0.18 ± 0.02	0.01 ± 0.005
2	256	ELU	100	32	0.14 ± 0.02	0.21 ± 0.03	0.09 ± 0.015
3	128	ELU	100	16	0.21 ± 0.02	0.20 ± 0.03	0.03 ± 0.01
4	128	ReLU	50	16	0.18 ± 0.02	0.20 ± 0.03	0.03 ± 0.01
5	16	Swish	100	32	0.05 ± 0.01	0.03 ± 0.005	0.04 ± 0.008
5	128	ReLU	100	16	0.25 ± 0.03	0.22 ± 0.02	0.06 ± 0.01

3. Discussion

3.1. Optimization of Neural Network Hyperparameters for Chemical Reaction Prediction

The optimization of hyperparameters for neural network architectures is critical in achieving optimal performance for datasets related to chemical reactions, which encompass information on various volumetric compositions of reacting substances. This study underscores the importance of carefully selecting and fine-tuning the architectural parameters of neural networks to accurately model and predict the behavior of complex chemical systems. By systematically exploring different combinations of layers, neurons, activation functions, epochs, and batch sizes through techniques like grid search, we can identify configurations that yield the lowest errors and the highest predictive accuracy.

Recent advances in artificial neural networks, particularly in deep learning, have demonstrated their efficacy in drug discovery and toxicology research. The study [45] explores various hyper-parameter configurations for deep neural networks and finds that optimized models outperform traditional cheminformatics methods, with non-linear models performing well under low noise levels and Naïve Bayes models excelling at higher noise levels.

Computational simulations offer promise in rational material design, but their high computational cost limits their ability to efficiently explore material phase space. The study [46] employs machine learning to predict a critical parameter in catalyst performance, the reaction barrier, advancing the computational prediction of catalytic activity.

A demonstration of the effectiveness of transformer-based models in automatically inferring reaction classes from simple text-based representations of chemical reactions is shown in [47]. Additionally, the authors showed that the learned representations serve as reaction fingerprints, capturing nuanced differences between reaction classes better than

traditional methods. These findings facilitate a deeper understanding of chemical reaction spaces, as illustrated by an interactive reaction atlas offering visual clustering and similarity searching capabilities.

3.2. Exploring Architectural Choices and Model Performance

The results obtained from this optimization process provide valuable insights into the interplay between architectural choices and model performance [48]. For instance, we observed that certain configurations, such as networks with a higher number of layers and neurons, tend to capture more intricate patterns in the data but may also be prone to overfitting, especially with limited training data. On the other hand, simpler architectures with fewer layers and neurons may generalize better to unseen data but could potentially overlook subtle nuances in the dataset.

Furthermore, the choice of activation functions plays a crucial role in shaping the non-linear relationships between input features and output predictions [49]. By experimenting with a variety of activation functions like ReLU, sigmoid, tanh, leaky ReLU, ELU, and swish, researchers can gain deeper insights into how different activation functions influence the learning dynamics and convergence of neural networks.

Additionally, the impact of training parameters such as the number of epochs and batch size on model convergence and computational efficiency cannot be understated. Fine-tuning these parameters based on the dataset size, complexity, and computational resources available is essential to strike a balance between training time and model performance.

Overall, the findings from this study pave the way for future research endeavors aimed at further refining neural network architectures for chemical reaction datasets. Potential avenues for future exploration include the investigation of more advanced neural network architectures such as recurrent neural networks (RNN) [50] or attention mechanisms, ensemble methods for combining multiple neural network models, and the incorporation of domain-specific knowledge into model design through techniques like transfer learning or neural architecture search. Moreover, exploring the applicability of neural networks in conjunction with other machine learning techniques like support vector machines (SVM) [51] or random forests (RF) [52] could offer complementary insights and enhance predictive capabilities for chemical reaction datasets.

4. Materials and Methods

4.1. Preparation of Chemical Reactions

The study focuses on investigating the kinetic properties of the chemical reaction involving the reduction of methylene blue by ascorbic acid in various solvent systems. These solvent systems encompass super dilute solutions (SDS) of methylene blue, ascorbic acid, and water, as well as purified water. Rigorous characterization procedures will be employed to ensure the uniformity and purity of each solvent system prior to experimentation. Subsequently, four distinct reactions will be conducted as follows:

1. MB (dissolved in H₂O) + AA (dissolved in H₂O) → Products
2. MB (dissolved in a SDS of H₂O) + AA (dissolved in a SDS of H₂O) → Products
3. MB (dissolved in a SDS of MB) + AA (dissolved in a SDS of AA) → Products
4. MB (dissolved in a SDS of AA) + AA (dissolved in a SDS of MB) → Products

The subsequent task entailed assessing the kinetic characteristics (reaction rate of neutralization or decolorization) of the model reaction upon mixing its components with varying concentrations of super-dilute solutions (SDSs). To achieve this, methylene blue was pre-mixed with the following:

1. SDS MB / SDS water / water (in a volume ratio of 1:1)
2. SDS MB / SDS water / water (in a volume ratio of 9:1)
3. SDS MB / SDS water / water (in a volume ratio of 99:1)

A solution of MB is prepared in the designated solvent and transferred into a cuvette with a volume of 2 mL. The cuvette is then placed in the cuvette compartment of the

spectrophotometer, and the absorption spectrum of the solution is recorded in the range of 500 to 700 nm. The measurements obtained on the spectrophotometer PB 2201 (SOLAR, Belarus, Minsk) involved recording at least 10 consecutive readings with a 30 s interval between each. Subsequently, the maximum absorbance for each of the curves was identified, resulting in first-order decay plots for MB+ absorbance at 665 nm.

4.2. Processing of Results

A Python script was developed to automate the processing of data obtained from the spectrophotometer. This script facilitated the conversion of cumbersome .txt files into more manageable .xlsx files, each storing spectral curves at various time points, with a new curve generated every 30 s. The script also automatically detected the maximum of each spectrum, typically located around 665 nm as theorized. Upon detection, it created .xlsx files for the decay of the maximum absorbance. Using the lmfit library, the script then performed curve fitting using an exponential decay function of the form

$$A + B \cdot e^{-x/C} \quad (1)$$

The resulting fits were visualized alongside the original data, providing insight into the decay behavior. Consequently, a dataset was compiled containing filenames corresponding to reaction experiments and coefficients obtained from curve fitting.

The dataset was constructed by recording the volumes of substances involved in each reaction, where each row represented a distinct reaction scenario. For instance, if the initial mixing volume for substances was 2000 mL and the volumetric ratios of MB and AA were 99:1, then 99 and 1 were entered into the respective columns for MB and AA. Similarly, if SDS of AA and H₂O were added in a 9:1 ratio, 9 and 1 were recorded in the corresponding AA_SDS and H₂O columns. Additionally, an extra solution was consistently introduced in a volume of 120 mL. Up to six substances could be included: MB, SDS MB, AA, SDS AA, H₂O, and SDS H₂O. An attempt was made to incorporate temperature data during solution measurements and to indicate the influence of sunlight exposure, but these additions did not contribute to model improvement. Further columns, denoted as MB+, AA+, MB_SDS+, AA_SDS+, H₂O+, and H₂O_SDS+, were appended to include additional substances.

Figure 4 illustrates the proposed neural network architecture for predicting coefficients A, B, and C in chemical reactions. The diagram showcases the input parameters of the model, allowing for variation in the number of neurons and hidden layers. Despite these variations, the model focuses on predicting the three coefficients, providing a streamlined approach to reaction prediction.

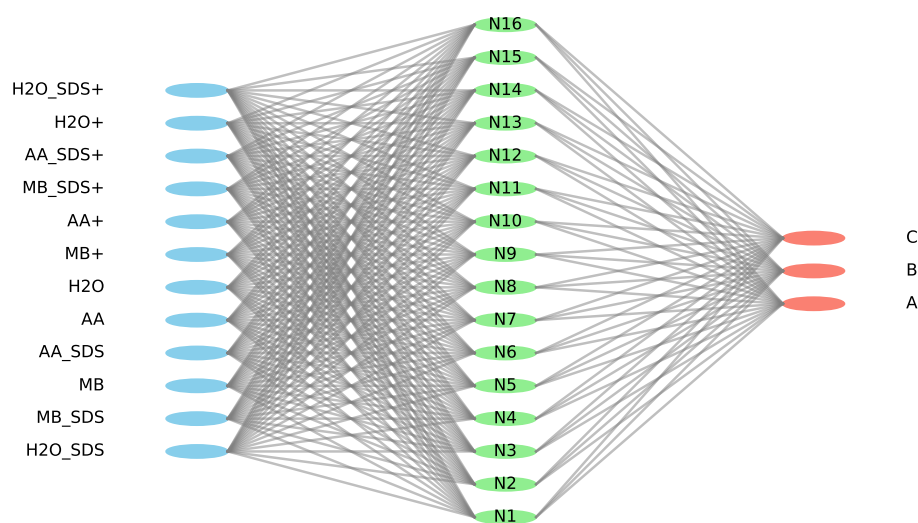


Figure 4. One of the possible Neural Network Architectures for Predicting Coefficients A, B, and C: a variant with sixteen neurons and one hidden layer.

5. Conclusions

In this study, we investigated the optimization of hyperparameters for the architecture of a neural network for analyzing data related to chemical reactions. Our dataset contains information on various volumetric compositions of substances involved in the reactions. The primary aim of our investigation was to ascertain the optimal parameters of the neural network for accurately predicting the profiles of exponential decay coefficients in chemical reactions.

In the process of hyperparameter optimization, we used the grid search method, exploring different combinations of parameters such as the number of layers, the number of neurons per layer, activation functions, the number of training epochs, and batch size. This approach allowed us to systematically investigate various configurations of neural networks and evaluate their performance on our dataset.

The results of our study showed that the choice of optimal hyperparameters significantly influences the performance of the neural network. Our investigation unveiled the significant impact of activation functions on enhancing the efficiency of the neural network. We found that for certain types of data and tasks, specific activation functions may be more suitable than others. For example, the Swish function performed best for certain types of chemical data, while ReLU was preferable for others.

Overall, our study confirmed the importance of selecting the right hyperparameters to ensure the optimal performance of neural networks in analyzing chemical data. Our findings can be valuable for researchers and engineers working in the fields of chemistry and machine learning and can serve as a basis for further research in this area.

Author Contributions: Conceptualization, I.M. and V.T.; Data curation, I.M. and V.N.; Formal analysis, V.T., A.G. and A.B.; Funding acquisition, A.G., V.N. and A.B.; Investigation, A.G.; Methodology, I.M. and V.T.; Project administration, A.G., V.N. and A.B.; Resources, I.M., V.T., V.N. and A.B.; Software, I.M.; Supervision, V.T., A.G., V.N. and A.B.; Validation, I.M., V.T. and A.B.; Visualization, I.M., A.G. and V.N.; Writing—original draft, I.M. and V.T.; Writing—review and editing, A.G., V.N. and A.B. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The code and datasets for this research are openly accessible on our Git repository: ChemReactOpt, <https://github.com/catauggie/ChemReactOpt>, accessed on 24 March 2024. This centralized resource promotes transparency and facilitates reproducibility in scientific research. We invite the scientific community to explore, scrutinize, and build upon our work, fostering collective progress in the field.

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

References

1. Ghaedi, M.; Nasab, A.G.; Khodadoust, S.; Rajabi, M.; Azizian, S. Application of activated carbon as adsorbents for efficient removal of methylene blue: Kinetics and equilibrium study. *J. Ind. Eng. Chem.* **2014**, *20*, 2317–2324.
2. Mowry, S.; Ogren, P.J. Kinetics of methylene blue reduction by ascorbic acid. *J. Chem. Educ.* **1999**, *76*, 970.
3. Ayad, M.M.; El-Nasr, A.A. Adsorption of cationic dye (methylene blue) from water using polyaniline nanotubes base. *J. Phys. Chem. C* **2010**, *114*, 14377–14383.
4. Fan, S.; Wang, Y.; Wang, Z.; Tang, J.; Tang, J.; Li, X. Removal of methylene blue from aqueous solution by sewage sludge-derived biochar: Adsorption kinetics, equilibrium, thermodynamics and mechanism. *J. Environ. Chem. Eng.* **2017**, *5*, 601–611.
5. Mouni, L.; Belkhir, L.; Bollinger, J.C.; Bouzaza, A.; Assadi, A.; Tirri, A.; Dahmoune, F.; Madani, K.; Remini, H. Removal of Methylene Blue from aqueous solutions by adsorption on Kaolin: Kinetic and equilibrium studies. *Appl. Clay Sci.* **2018**, *153*, 38–45.
6. Weiland, E.; et al. Reduction of methylene blue by ascorbic acid in light. *Biochem. Ztschr.* **1944**, *317*, 168–170.
7. Seno, M.; Kousaka, K.; Kise, H. Reduction of methylene blue with L-ascorbic acid or L-cysteine in micellar systems. *Bull. Chem. Soc. Jpn.* **1979**, *52*, 2970–2974.
8. Ramimoghadam, D.; Hussein, M.Z.B.; Taufiq-Yap, Y.H. The effect of sodium dodecyl sulfate (SDS) and cetyltrimethylammonium bromide (CTAB) on the properties of ZnO synthesized by hydrothermal method. *Int. J. Mol. Sci.* **2012**, *13*, 13275–13293.

9. Hallock, A.J.; Berman, E.S.; Zare, R.N. Ultratrace kinetic measurements of the reduction of methylene blue. *J. Am. Chem. Soc.* **2003**, *125*, 1158–1159.
10. Bellido-Pedraza, C.M.; Calatrava, V.; Llamas, A.; Fernandez, E.; Sanz-Luque, E.; Galvan, A. Nitrous oxide emissions from nitrite are highly dependent on nitrate reductase in the microalga *Chlamydomonas reinhardtii*. *Int. J. Mol. Sci.* **2022**, *23*, 9412.
11. Sen, P.K.; Mukherjee, P.; Pal, B. Effects of pre-micelles of anionic surfactant SDS on the electron transfer reaction between methylene blue and ascorbic acid. *J. Mol. Liq.* **2016**, *224*, 472–479.
12. Snehalatha, T.; Rajanna, K.; Saiprakash, P. Methylene blue-ascorbic acid: An undergraduate experiment in kinetics. *J. Chem. Educ.* **1997**, *74*, 228.
13. Kooh, M.R.R.; Thotagamuge, R.; Chau, Y.F.C.; Mahadi, A.H.; Lim, C.M. Machine learning approaches to predict adsorption capacity of *Azolla pinnata* in the removal of methylene blue. *J. Taiwan Inst. Chem. Eng.* **2022**, *132*, 104134.
14. Cannavò, S.; Bertoldi, A.; Valeri, M.C.; Damiani, F.; Reale, L.; Brillì, F.; Paolucci, F. Impact of High Light Intensity and Low Temperature on the Growth and Phenylpropanoid Profile of *Azolla filiculoides*. *Int. J. Mol. Sci.* **2023**, *24*, 8554.
15. Izonin, I.; Tkachenko, R.; Holoven, R.; Shavarskyi, M.; Bukin, S.; Shevchuk, I. Multistage SVR-RBF-based model for heart rate prediction of individuals. In Proceedings of the International Conference of Artificial Intelligence, Medical Engineering, Education, Wuhan, China 19–21 August 2022; pp. 211–220.
16. Marzban, N.; Moheb, A.; Filonenko, S.; Hosseini, S.H.; Nouri, M.J.; Libra, J.A.; Farru, G. Intelligent modeling and experimental study on methylene blue adsorption by sodium alginate-kaolin beads. *Int. J. Biol. Macromol.* **2021**, *186*, 79–91.
17. Cherkasov, A. Inductive QSAR descriptors. Distinguishing compounds with antibacterial activity by artificial neural networks. *Int. J. Mol. Sci.* **2005**, *6*, 63–86.
18. Fisher, K.S.; Vreugdenhil, A.J. Adsorption of chromium (VI) using an activated carbon derived from petroleum coke feedstock. *Int. J. Mol. Sci.* **2022**, *23*, 16172.
19. Asfaram, A.; Ghaedi, M.; Azghandi, M.A.; Goudarzi, A.; Dastkhoo, M. Statistical experimental design, least squares-support vector machine (LS-SVM) and artificial neural network (ANN) methods for modeling the facilitated adsorption of methylene blue dye. *RSC Adv.* **2016**, *6*, 40502–40516.
20. Kalantari, K.; Ahmad, M.B.; Masoumi, H.R.F.; Shamelì, K.; Basri, M.; Khandanlou, R. Rapid adsorption of heavy metals by Fe₃O₄/talc nanocomposite and optimization study using response surface methodology. *Int. J. Mol. Sci.* **2014**, *15*, 12913–12927.
21. Bao, Y.; Kong, W.; Liu, F.; Qiu, Z.; He, Y. Detection of glutamic acid in oilseed rape leaves using near infrared spectroscopy and the least squares-support vector machine. *Int. J. Mol. Sci.* **2012**, *13*, 14106–14114.
22. Luna, J.M.; Rufino, R.D.; Albuquerque, C.D.; Sarubbo, L.A.; Campos-Takaki, G.M. Economic optimized medium for tensio-active agent production by *Candida sphaerica* UCP0995 and application in the removal of hydrophobic contaminant from sand. *Int. J. Mol. Sci.* **2011**, *12*, 2463–2476.
23. Chowdhury, N.S.; Jamaly, S.; Farjana, F.; Begum, N.; Zenat, E.A. A review on ethnomedicinal, pharmacological, phytochemical and pharmaceutical profile of lady's finger (*Abelmoschus esculentus* L.) plant. *Pharmacol. Pharm.* **2019**, *10*, 94–108.
24. Nayak, A.K.; Pal, A. Green and efficient biosorptive removal of methylene blue by *Abelmoschus esculentus* seed: Process optimization and multi-variate modeling. *J. Environ. Manag.* **2017**, *200*, 145–159.
25. Soleimani, S.; Heydari, A.; Fattahi, M.; Motamedisade, A. Calcium alginate hydrogels reinforced with cellulose nanocrystals for methylene blue adsorption: Synthesis, characterization, and modelling. *Ind. Crop. Prod.* **2023**, *192*, 115999.
26. Mirjalili, S.; Lewis, A. The whale optimization algorithm. *Adv. Eng. Softw.* **2016**, *95*, 51–67.
27. Mahmoodi-Babolan, N.; Heydari, A.; Nematollahzadeh, A. Removal of methylene blue via bioinspired catecholamine/starch superadsorbent and the efficiency prediction by response surface methodology and artificial neural network-particle swarm optimization. *Bioresour. Technol.* **2019**, *294*, 122084.
28. Wang, D.; Tan, D.; Liu, L. Particle swarm optimization algorithm: An overview. *Soft Comput.* **2018**, *22*, 387–408.
29. Alibak, A.H.; Khodarahmi, M.; Fayyazsanavi, P.; Alizadeh, S.M.; Hadi, A.J.; Aminzadehsarikhanbeglou, E. Simulation the adsorption capacity of polyvinyl alcohol/carboxymethyl cellulose based hydrogels towards methylene blue in aqueous solutions using cascade correlation neural network (CCNN) technique. *J. Clean. Prod.* **2022**, *337*, 130509.
30. Mohamed, S.A.E.M.; Mohamed, M.H.; Farghally, M.F. A new cascade-correlation growing deep learning neural network algorithm. *Algorithms* **2021**, *14*, 158.
31. Safarik, J.; Jalowiczor, J.; Gresak, E.; Rozhon, J. Genetic algorithm for automatic tuning of neural network hyperparameters. In Proceedings of the Autonomous Systems: Sensors, Vehicles, Security, and the Internet of Everything, SPIE, Orlando, FL, USA, 16–18 April 2018; Volume 10643; pp. 168–174.
32. Kandel, I.; Castelli, M. How deeply to fine-tune a convolutional neural network: A case study using a histopathology dataset. *Appl. Sci.* **2020**, *10*, 3359.
33. Stathakis, D. How many hidden layers and nodes? *Int. J. Remote Sens.* **2009**, *30*, 2133–2147.
34. Alvarez, J.M.; Salzman, M. Learning the number of neurons in deep networks. *Adv. Neural Inf. Process. Syst.* **2016**, *29*, 1–9.
35. Daubechies, I.; DeVore, R.; Foucart, S.; Hanin, B.; Petrova, G. Nonlinear approximation and (deep) ReLU networks. *Constr. Approx.* **2022**, *55*, 127–172.
36. Yin, X.; Goudriaan, J.; Lantinga, E.A.; Vos, J.; Spiertz, H.J. A flexible sigmoid function of determinate growth. *Ann. Bot.* **2003**, *91*, 361–371.

37. Kalman, B.L.; Kwasny, S.C. Why tanh: Choosing a sigmoidal function. In Proceedings of the [Proceedings 1992] IJCNN International Joint Conference on Neural Networks, Baltimore, MD, USA, 7–11 June 1992; Volume 4; pp. 578–581.
38. Xu, J.; Li, Z.; Du, B.; Zhang, M.; Liu, J. Reluplex made more practical: Leaky ReLU. In Proceedings of the 2020 IEEE Symposium on Computers and communications (ISCC), Rennes, France, 7–10 July 2020; pp. 1–7.
39. Rönnerberg, J.; Lunner, T.; Zekveld, A.; Sörqvist, P.; Danielsson, H.; Lyxell, B.; Dahlström, Ö.; Signoret, C.; Stenfelt, S.; Pichora-Fuller, M.K.; et al. The Ease of Language Understanding (ELU) model: Theoretical, empirical, and clinical advances. *Front. Syst. Neurosci.* **2013**, *7*, 31.
40. Oliver, N.; Smith, G.; Thakkar, C.; Surendran, A.C. SWISH: Semantic analysis of window titles and switching history. In Proceedings of the 11th International Conference on Intelligent User Interfaces, Sydney, Australia, 29 January–1 February 2006; pp. 194–201.
41. Schmeiser, B. Batch size effects in the analysis of simulation output. *Oper. Res.* **1982**, *30*, 556–568.
42. Händel, P. Understanding normalized mean squared error in power amplifier linearization. *IEEE Microw. Wirel. Components Lett.* **2018**, *28*, 1047–1049.
43. Wang, Z.; Bovik, A.C. Mean squared error: Love it or leave it? A new look at signal fidelity measures. *IEEE Signal Process. Mag.* **2009**, *26*, 98–117.
44. Liashchynskiy, P.; Liashchynskiy, P. Grid search, random search, genetic algorithm: A big comparison for NAS. *arXiv* **2019**, arXiv:1912.06059.
45. Koutsoukas, A.; Monaghan, K.J.; Li, X.; Huan, J. Deep-learning: Investigating deep neural networks hyper-parameters and comparison of performance to shallow methods for modeling bioactivity data. *J. Cheminform.* **2017**, *9*, 1–13.
46. Singh, A.R.; Rohr, B.A.; Gauthier, J.A.; Nørskov, J.K. Predicting chemical reaction barriers with a machine learning model. *Catal. Lett.* **2019**, *149*, 2347–2354.
47. Schwaller, P.; Probst, D.; Vaucher, A.C.; Nair, V.H.; Kreutter, D.; Laino, T.; Reymond, J.L. Mapping the space of chemical reactions using attention-based neural networks. *Nat. Mach. Intell.* **2021**, *3*, 144–152.
48. Doelling, M.; Nasrollahi, F. Building performance modeling in non-simplified architectural design. In Proceedings of the 30th CAAD Curriculum, Taipei, Taiwan, 27–30 June 2011; Volume 1, pp. 97–106.
49. Dubey, S.R.; Singh, S.K.; Chaudhuri, B.B. Activation functions in deep learning: A comprehensive survey and benchmark. *Neurocomputing* **2022**, *503*, 92–108.
50. Grossberg, S. Recurrent neural networks. *Scholarpedia* **2013**, *8*, 1888.
51. Steinwart, I.; Christmann, A. *Support Vector Machines*; Springer Science & Business Media: Berlin/Heidelberg, Germany, 2008.
52. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.