Evolving Backpropagation: Fundamentals, Enhancements and Applications in Deep Neural Networks

Name: Yugam Amish Patel

Student Number: 7937959

COMP 3190: Introduction to Artificial Intelligence

Instructor: Dr. John Anderson

Abstract

Backpropagation is an essential algorithm for training deep neural networks, and it plays an important role in the advancement of artificial intelligence. This paper presents a comprehensive analysis of backpropagation, starting from its fundamental principles to innovative enhancements and applications in deep neural networks. The study begins by establishing the theoretical underpinnings of the backpropagation algorithm and progresses to explore significant enhancements such as Fahlman's sigmoidal modification, which mitigates the 'stuck units' issue, and the dynamic learning rate adjustments introduced by methods like Delta-Bar-Delta and SuperSAB. These enhancements are contextualized within their mathematical frameworks and practical implications for improving neural network training efficacy. Furthermore, the paper discusses the application of backpropagation in diverse fields, specifically highlighting its role in pharmacokinetics for drug metabolism prediction and environmental science for soil bioremediation. In the pharmacokinetic domain, the paper examines a case study on Rosuvastatin, demonstrating the BPANN model's predictive strength and its implications for personalized medicine. In environmental health, the application of BPANNs in predicting the bioremediation outcomes for PAH-contaminated soils illustrates how artificial intelligence can bridge the gap between ecological restoration and public health. Additionally, the paper addresses the problematic issues of overtraining and the architectural design of backpropagation networks, providing insights into frequency analysis and optimization strategies to combat overfitting. This research not only underscores the versatility and adaptability of backpropagation in modern computational tasks but also projects the continuous evolution of this algorithm as a critical tool for future advancements in neural network applications.

1. Introduction of Neural networks and backpropagation

Neural networks are a cornerstone of artificial intelligence, powering machines to recognize patterns, make decisions, and solve problems across different domains [Hecht-Nielsen, 1992]. From identifying images to translating languages, these networks mimic the way humans think and learn, making tasks that once seemed exclusively human, achievable by computers [Abraham, 2005]. At the heart of these systems is supervised learning, a method where the network is fed data and correct answers, learning to make predictions and improve over time. This process relies on robust learning algorithms, among which backpropagation stands out for its ability to refine these complex neural networks, making them smarter with each iteration [Hecht-Nielsen, 1992].

In the architecture of neural networks, the backpropagation algorithm emerges as a pivotal mechanism for refining the system's learning capabilities in supervised settings. Described as a distributed information processing structure, each node within these networks is responsible for a particular component of the overall processing, similar to the neurons within the human brain [Hecht-Nielsen, 1992]. Backpropagation helps these nodes modify their weights, which is an important step in lowering the network's total prediction error [Hecht-Nielsen, 1992]. The adjustments are made by calculating gradients—a measure of how much a change in each weight affects the final output error. This is vital for the network to learn accurately from the input data provided. Continuous research and innovation in backpropagation are critical because they address and overcome issues such as vanishing gradients, which

occur when gradients shrink as they propagate back through the network, making it impossible for the network to learn [Hecht-Nielsen, 1992]. Backpropagation innovations aim to increase the efficiency of this learning process and produce more accurate models, which are required to address the growing complexity of problems that neural networks are expected to tackle.

2. Fundamentals and Mechanisms of Backpropagation

Neural network learning is essentially based on the backpropagation algorithm, which uses differential calculus concepts, namely the chain rule $\frac{dy}{dx} = \frac{dy}{du} \cdot \frac{du}{dx}$ to calculate the gradients. The significance of these gradients lies in their ability to show how sensitive the error is to changes in the network's weights, which are symbolically represented by ΔW, where W stands for weights [Hecht-Nielsen, 1992]. This methodology requires a thorough understanding of the topology of the error function, Φ(E), where E represents the error, rather than relying only on the rote application of calculus. Hence, backpropagation conceptualises learning as an optimisation problem, aiming to identify the global minimum (E_{min}) of the error surface, which signifies the point of least error or the zenith of the network's performance. This pursuit involves navigating through a multidimensional space, encapsulated by the error surface, to find this point of minimal error [Chauvin & Rumelhart, 2013].

The modification of weights during the backpropagation process is intricately influenced by the learning rate, α, and the magnitude of the gradient, ‖

$\nabla E \|$, each constituting a strategic step towards the minimization of the loss function, L. This delicate equilibrium aims to enhance the network's predictive accuracy by ensuring it does not overshoot the minimal error point, thus optimising performance [Hecht-Nielsen, 1992]. Chauvin and Rumelhart (2013) elucidate that this method bears resemblance to the task of fitting a multidimensional curve to data points in statistical analysis, a technique known as nonlinear regression. Through the adjustment of weights, denoted as ΔW, in a manner that reduces the discrepancy between predicted outputs and actual targets, the backpropagation algorithm adeptly facilitates this fitting process.

Additionally, the architecture of backpropagation underscores the importance of selecting appropriate activation functions, for instance, the sigmoid function, $\sigma(x) = \frac{1}{1+e^{-x}}$, thereby augmenting its capabilities for nonlinear modelling [Hecht-Nielsen, 1992]. A prominent obstacle encountered by the algorithm is the vanishing gradient problem, which becomes evident in deeply layered networks, leading to gradients that diminish in magnitude as they are propagated backwards through the network, consequently impeding the model's learning efficiency [Chauvin & Rumelhart, 2013]. In response to these challenges, continuous advancements and the application of sophisticated mathematical strategies are being pursued to ensure that backpropagation remains a viable and effective method for neural network training.

3. Innovations and Advancements: Refining the Backpropagation Algorithm

3.1 Fahlman's Sigmoidal Modification: Addressing the Challenge of 'Stuck Units'

In the realm of neural network training, the backpropagation algorithm has been subject to a variety of enhancements aimed at overcoming inherent challenges. One such challenge is the issue of 'stuck units', a phenomenon that occurs when the gradient of the sigmoidal transfer function asymptotically approaches zero, thereby leading to minimal weight updates and impeding the learning process [Wythoff, 1993]. This can result in what is commonly referred to as 'local minima', though Fahlman (1988) posits that these are often merely instances of stuck units where the network ceases to learn due to negligible error signal propagation. Fahlman's insightful modification introduces a constant term to the derivative of the sigmoid function, ensuring that the gradient never vanishes entirely [Fahlman, 1988]. This constant, typically ranging between 0.01 to 0.1, acts as a preventative measure against the derivative approaching zero, thus mitigating the risk of stuck units and facilitating more robust learning dynamics [Wythoff, 1993]. Empirical studies have underscored the efficacy of this modification, revealing not just an enhancement in the reliability of convergence but also, in many cases, an acceleration of the learning process [Wythoff, 1993].

By incorporating this subtle yet powerful tweak into the backpropagation algorithm, Fahlman not only addressed a fundamental limitation but also paved the way for a deeper understanding of neural network behaviour during training. This enhancement underscores the intricate balance required

between mathematical theory and practical implementation in the development of artificial neural networks.

3.2 Dynamic Learning Rate Adjustments: The Delta-Bar-Delta and SuperSAB Approaches

Dynamic adjustment of the learning rate emerges as a critical improvement in the backpropagation algorithm, addressing the inefficiencies associated with a static learning rate. The innovation of the Delta-Bar-Delta rule introduced a method in which each weight has its own stepsize parameter, dynamically adjusted as training progresses [Wythoff, 1993]. This approach alleviates the burdensome task of manually setting learning parameters and instead employs feedback from the error surface itself to guide these adjustments. The SuperSAB algorithm, a variant of the Delta-Bar-Delta rule proposed by Tollenaere (1990), further refines this process. It amplifies the stepsize if the gradient's sign remains consistent, promoting faster convergence, and conversely reduces it if a step results in a gradient sign inversion, thereby avoiding the pitfalls of drastic oscillations across the error surface [Wythoff, 1993].

These dynamic learning rate methods have been transformative, with empirical evidence suggesting a potential reduction in learning times by an order of magnitude in certain scenarios [Wythoff, 1993]. However, the caveat lies in the susceptibility to local minima, as aggressive adjustments could potentially undermine the escape from such suboptimal points.

Notwithstanding these limitations, the Delta-Bar-Delta and SuperSAB represent a leap forward in optimizing the backpropagation process, providing a more nuanced and responsive learning pathway that contrasts sharply with the rigidity of its predecessor.

3.3 Leveraging Classical Optimization: From Steepest Descent to Conjugate Gradients

The incorporation of classical optimization techniques into the backpropagation framework represents a significant shift from heuristic to more principled approaches in neural network training. Steepest descent, one of the earliest methods employed, follows the gradient of the error surface to minimize loss. However, its application is often marred by inefficiency, particularly in high-dimensional spaces, leading to slow convergence and susceptibility to local minima [Wythoff, 1993]. This paved the way for the adoption of conjugate gradient methods, which introduce a refinement by employing second-order information to dictate the search direction, effectively reducing the tendency to 'zig-zag' and thus accelerating the convergence process [Wythoff, 1993].

This evolution from steepest descent to conjugate gradients is not just a transition in techniques but also a reflection of the growing sophistication in understanding the error landscape of neural networks. My analysis suggests that while conjugate gradients have markedly improved the efficiency of backpropagation, the method's dependency on the accurate estimation of

the error surface curvature can be a limitation, particularly in complex networks with intricate loss surfaces. The ongoing challenge is to refine these methods further, potentially by integrating adaptive heuristics that can better navigate the intricacies of high-dimensional error landscapes encountered in deep learning.

3.4 Generalizing Error Functions: Minkowski-r Backpropagation and Its Implications

The generalization of error functions in backpropagation is exemplified by the introduction of the Minkowski-r error metric, which extends the flexibility of the learning algorithm to accommodate various types of data distributions [Chauvin & Rumelhart, 2013]. By incorporating the Minkowski-r norm, the backpropagation algorithm is no longer confined to the Euclidean norm, thereby allowing for the adjustment of the error metric to better suit the underlying data distribution characteristics. The standard backpropagation algorithm utilizes the Minkowski error with $r = 2$, which corresponds to the Euclidean distance, a choice that may not be optimal for all data types [Chauvin & Rumelhart, 2013]. The general form of the Minkowski-r backpropagation is represented as $\frac{\partial E}{\partial w_{hi}} = (y_i - \hat{y}_i)^r (y_i - \hat{y}_i)(1 - y_i)^{r-1}$, enables a variable influence on the learning process depending on the choice of r [Chauvin & Rumelhart, 2013].

$$\frac{\partial E}{\partial w_{hi}} = \frac{\partial E}{\partial y_i} \frac{\partial y_i}{\partial x_i} \frac{\partial x_i}{\partial w_{hi}}, \tag{4}$$

$$\frac{\partial E}{\partial w_{hi}} = (y_i - \hat{y}_i) y_i (1 - y_i) y_h. \tag{5}$$

$$E = \frac{1}{r} \sum_i (y_i - \hat{y}_i)^r. \tag{6}$$

$$\frac{\partial E}{\partial w_{hi}} = (y_i - \hat{y}_i)^{r-1} y_i (1 - y_i) y_h. \tag{7}$$

$$w_{hi}(n + 1) = \eta \frac{\partial E}{\partial w_{hi}} + w_{hi}(n). \tag{8}$$

The Minkowski-r approach offers a methodological advancement that aligns with the need for a more nuanced understanding of the error landscape in neural network training. A lower value of r, for instance, diminishes the weight of large deviations, which can be beneficial in cases where the feature vectors do not follow a Gaussian distribution. Conversely, a larger value of r emphasizes these deviations, which can be advantageous in domains where outlier identification is crucial. My analysis, building upon the foundational work of Chauvin and Rumelhart, proposes that selecting an appropriate r value is not merely a technical choice but a strategic decision that reflects the desired sensitivity of the model to the characteristics of the input data. Thus, Minkowski-r backpropagation stands out as a versatile tool in the arsenal of neural network training techniques, offering a customizable approach to error minimization that can be tailored to diverse problem domains. This adaptability is illustrated in Figure 3 [Chauvin & Rumelhart, 2013], where a nonlinearly separable problem— the "mesh" problem— is

resolved using backpropagation with different values of r, demonstrating the algorithm's capacity to handle intricate decision boundaries effectively.

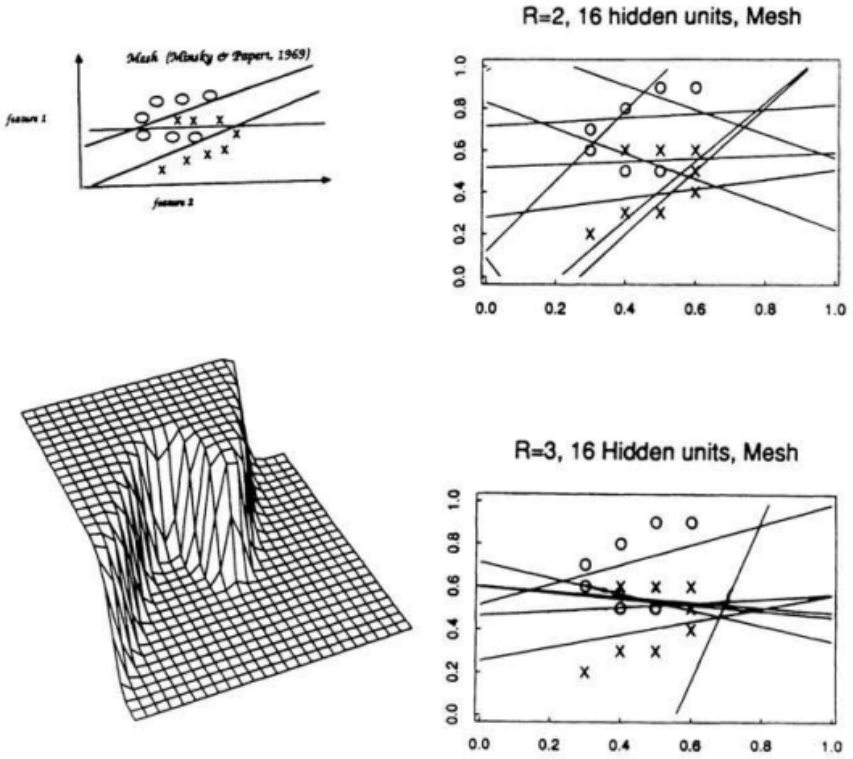

Figure 3. A nonlinearly separable problem, the "mesh" problem solved: (a) typical back-propagation solution (b) and Minkowski-*r* solution with $r = 3$.

4. Overcoming Backpropagation Challenges

4.1 Overtraining and Frequency Response of Nodes

Within the realm of neural network training, overtraining emerges as a formidable challenge, often leading to a model's performance degradation when exposed to new, unseen data. This phenomenon, characterized by a network's excessive fitting to the training data, can be mitigated by monitoring mean squared error (MSE) on an independent validation set [Wythoff, 1993]. Chauvin (1992) described networks as initially capturing the dominant patterns within data, akin to the DC component in signal

processing, before transitioning to finer, often noisy patterns, comparable to higher-frequency signals. This observation gives rise to an analogy with frequency responses in signal processing, suggesting a network's learning trajectory mirrors a movement from low to high-frequency fitting [Wythoff, 1993].

To understand this frequency behaviour, one may consider the gradient's magnitude along the weight vector, mathematically conceptualized as the derivative of the sigmoid function output concerning the input - a relationship expressed as $\delta o/\delta x_i = -o(1-o)w_i = LW_i$ [Wythoff, 1993, Eq. 17]. The node's frequency, proportional to one-quarter the length of its weight vector, thus offers a measure of the node's responsiveness to input variations, visualized as a 'hyper-cone' in the weight space [Wythoff, 1993, Eq. 19]. Insight into these dynamics provides a strategic angle to controlling a network's generalization capability, aligning frequency responses with the complexity of the function being modelled.

$$\frac{\partial o}{\partial x_i} = -o(1-o)w_i = uw_i \qquad (17)$$

$$\sqrt{\sum_{i=1}^{n}\left(\frac{\partial o}{\partial x_i}\right)^2} = \sqrt{\sum_{i=1}^{n}(uw_i)^2} = u\sqrt{\sum w_i^2} \qquad (18)$$

$$u\sqrt{\sum w_i^2} = -o(1-o)\sqrt{\sum w_i^2} \qquad (19)$$

In practical terms, this analysis posits that by constraining the weight vectors' lengths, and consequently their frequency response, one may tailor the network to the intricacies of the target function, potentially averting overfitting and promoting better generalization [Wythoff, 1993]. It invites a nuanced consideration of weight initialization and adjustments throughout the training process, emphasizing the importance of a strategic approach to network design and training, balancing the learning of underlying data patterns with the preservation of the model's ability to generalize.

4.2 Architectural Optimization in Neural Networks

Determining the optimal architecture for a neural network is a complex yet essential task, particularly in the context of backpropagation networks. The architectural design must be capable of capturing the essence of the data without succumbing to overfitting—where the model learns the training data too well, including its noise and anomalies [Wythoff, 1993]. This dilemma often manifests in the design phase, where the choice between a minimalistic and an overly complex network can dictate the success of the model.

To address this, various methods have been employed, with some starting from a simplistic network and incrementally adding complexity, while others begin with a larger network, systematically pruning it down. These methods, though seemingly opposite, converge on the same goal: to sculpt an architecture that balances complexity with the ability to generalize from

training to unseen data. Weight decay, as described by Weigend et al. (1993), emerges as a nuanced technique in this context, penalizing the cost function for the use of excessive network components, thereby encouraging a leaner, more efficient architecture (Wythoff, 1993, Eq. 20).

$$E = E_0 + \gamma \sum_{i=1}^{n} \sum_{j=1}^{m} \frac{w_{ij}^2 / w_0^2}{1 + w_{ij}^2 / w_0^2} \qquad (20)$$

This process of pruning can be likened to the careful sculpting of a statue, where the removal of extraneous material reveals the intended form. In neural networks, this involves cutting away weights and connections that do not contribute significantly to the model's predictive power, enhancing the network's ability to perform well on new data. Such techniques not only optimize the computational efficiency but also reduce the model's propensity to overfit, ensuring that the final architecture is both robust and reliable.

5. Real-World Impact of Backpropagation: Pharmacokinetics and Bioremediation

5.1 BPANNs in Pharmacokinetics: Case Study of Rosuvastatin

5.1.1 Enhancing Predictive Accuracy in Drug Metabolism

The advent of backpropagation artificial neural networks (BPANNs) has heralded a new era in the field of pharmacokinetics, particularly in the prediction of drug metabolism and pharmacokinetic parameters. Rosuvastatin, a widely used statin with notable interindividual variability in plasma concentration and pharmacokinetic parameters,

exemplifies the complexity of predicting drug behaviour within the human body [Xu et al., 2020]. Traditional pharmacokinetic modelling often grapples with the challenge of accounting for the myriad of variables influencing drug metabolism, including genetic factors, age, body mass index (BMI), and organ function. BPANNs, with their capacity to learn and adapt from data, offer a sophisticated alternative, enabling the capture of nonlinear relationships and interactions that are not readily apparent through conventional methods [Xu et al., 2020].

In a ground-breaking study, Xu et al. (2020) constructed a BPANN model to predict the plasma concentration of Rosuvastatin in healthy subjects, demonstrating the model's superior predictive capability over multiple linear regression analysis and nonlinear mixed effects modelling. This advancement is particularly significant given the critical role of pharmacokinetics in drug development and regulatory approval, where accurate predictions can dramatically impact therapeutic efficacy and patient safety. The application of BPANNs in this context not only showcases the potential for enhanced precision in drug metabolism predictions but also underscores the importance of adopting machine learning techniques in the pharmaceutical sciences. By integrating BPANNs into pharmacokinetic studies, researchers can navigate the complex biological landscape with greater accuracy, potentially revolutionizing the field of drug

development and personalized medicine. This shift towards data-driven, adaptive models like BPANNs reflects a broader trend in healthcare and biomedical research, where the fusion of computational power and biological understanding opens new frontiers in the quest for optimized therapeutic interventions.

5.1.2 Bridging Biological Variability and Pharmacological Safety

The utilization of BPANNs extends beyond mere predictive accuracy, serving as a bridge between the biological variability inherent in human populations and the paramount concern of pharmacological safety. The study by Xu et al. (2020) on Rosuvastatin not only highlights the BPANN's capability to manage the interindividual variability affected by race, age, BMI, and organ function but also emphasizes the model's role in enhancing drug safety. This variability, which significantly influences drug metabolism and efficacy, poses a considerable challenge in ensuring the safe use of medications across diverse patient populations. By accurately predicting plasma concentrations and pharmacokinetic parameters, BPANNs facilitate a more informed decision-making process in clinical settings, tailoring drug dosages to achieve optimal therapeutic outcomes while minimizing adverse effects [Xu et al., 2020].

Furthermore, the incorporation of BPANNs into pharmacokinetic research mirrors the shift towards precision medicine, where

treatments can be customized to individual patient profiles. This approach is crucial in the context of drugs like Rosuvastatin, where the therapeutic window is narrow, and the consequences of under or overdosing can have significant clinical implications. Through the sophisticated analysis of complex datasets, BPANNs enable researchers and clinicians to navigate the multifaceted landscape of human biology, offering a path towards safer, more effective drug therapy.

5.2 BPANNs for Environmental Health: PAH Bioremediation

5.2.1 Pioneering Soil Decontamination Through Neural Networks

The application of backpropagation artificial neural networks (BPANNs) extends into the environmental sector, offering groundbreaking approaches to addressing soil contamination. Polycyclic aromatic hydrocarbons (PAHs), a group of organic contaminants known for their persistence and carcinogenic properties, present significant challenges for environmental health. Traditional methods for analysing and remediating PAH-contaminated soils often fall short due to the complex nature of these pollutants and their interactions with environmental factors. However, the study by Olawoyin (2016) demonstrates how BPANNs can be effectively utilized to predict the bioremediation outcomes of PAH-contaminated soil, showcasing the model's ability to assimilate and analyse vast datasets to forecast contaminant behaviour and degradation

pathways. The strength of BPANNs in this context lies in their flexibility and learning capacity, which allow for the modelling of complex, nonlinear relationships between various soil properties and the efficacy of bioremediation strategies. By training on historical data, BPANNs can identify patterns and predict outcomes with high accuracy, providing invaluable insights into the most effective approaches for reducing PAH concentrations in contaminated sites [Olawoyin, 2016]. This capability not only enhances the efficiency of remediation efforts but also contributes to a deeper understanding of the dynamics governing soil decontamination processes.

The integration of BPANNs into environmental science signifies a paradigm shift towards more intelligent, data-driven strategies for managing pollution and restoring ecosystems. As environmental challenges become increasingly complex, the adoption of advanced computational tools like BPANNs offers a promising path forward, bridging the gap between theoretical models and practical, on-the-ground solutions for soil decontamination and public health protection.

5.2.2 From Soil Remediation to Public Health: The ANN Advantage

The application of BPANNs in the field of environmental health, particularly through the lens of PAH bioremediation, illuminates a direct pathway to enhancing public health outcomes. Olawoyin's

(2016) exploration of neural networks in predicting the efficacy of soil decontamination techniques underscores a vital connection: cleaner environments lead to healthier populations. By accurately modelling the degradation pathways of PAHs in contaminated soils, BPANNs contribute not only to the rehabilitation of ecosystems but also to the reduction of human exposure to harmful carcinogens. This direct impact on public health is amplified by the capacity of BPANNs to process and learn from complex environmental data, enabling more targeted and effective remediation strategies that align with both ecological and human health objectives.

The implications of this technology extend beyond the immediate benefits of soil decontamination. As PAHs are known to contribute to a range of serious health issues, including cancer, the role of BPANNs in mitigating these risks cannot be overstated. The nuanced understanding and predictive power provided by neural networks offer a new frontier in environmental management, where decisions are informed by a comprehensive analysis of potential outcomes and their impacts on public health [Olawoyin, 2016]. Moreover, the deployment of BPANNs in environmental health research exemplifies a broader trend towards leveraging artificial intelligence to solve complex public health challenges. This intersection of environmental science and public health, facilitated by advanced computational models, represents a forward-thinking approach to ensuring the well-

being of communities worldwide. As the capabilities of neural networks continue to evolve, their potential to transform our approach to environmental management and public health protection becomes increasingly evident.

6. Conclusion

We conclude our wide trip of backpropagation within neural networks with a remarkable narrative of growth and adaptation. Backpropagation's algorithmic delicacy has enabled substantial advancements in neural networks, the foundation of modern artificial intelligence. This learning algorithm, which has its roots in the calculus of variations, has been continuously modified to address issues such as vanishing gradients and improving the learning process for improved prediction performance [Hecht-Nielsen, 1992; Chauvin & Rumelhart, 2013].

Significant advancements in the field have been made to the backpropagation algorithm, especially with the addition of dynamic learning rates and Fahlman's modification [Wythoff, 1993]. These developments improve neural networks' ability to handle difficult, practical tasks in addition to addressing intrinsic computational difficulties. The transformational power of neural networks is demonstrated by the application spectrum where these theoretical advances are being translated, most notably in pharmacokinetics and environmental health. This has been proved by Xu et al. (2020) with their precise pharmacokinetic parameter prediction, and Olawoyin (2016) has shown

how BPANNs can be used to manage environmental toxins and enhance public health outcomes. The integration of theoretical advances in backpropagation with their real-world applications captures the central objective of AI research, which is to develop intelligent systems that complement and expand human knowledge. In addition to being an academic endeavour, the fusion of advanced mathematical methods with knowledge of intricate biological and environmental systems is a critical step toward using AI to solve problems in the real world.

As we look forward, the continuous interplay between innovation in backpropagation algorithms and their applications heralds a future where AI is seamlessly integrated into the fabric of daily life. The promise of AI, as evidenced by the progression of backpropagation, is not just in the creation of systems that can learn but also in the development of solutions that are as diverse as they are impactful, aligning with the multifaceted needs of society.

References

[Abraham, 2005] Abraham, A., 2005. Artificial neural networks. *Handbook of measuring system design*.

[Chauvin & Rumelhart, 2013] Chauvin, Y. and Rumelhart, D.E., 2013. *Backpropagation: theory, architectures, and applications*. Psychology press.

[Hecht-Nielsen, 1992] Hecht-Nielsen, R., 1992. Theory of the backpropagation neural network. In *Neural networks for perception* (pp. 65-93). Academic Press.

[Olawoyin, 2016] Olawoyin, R., 2016. Application of backpropagation artificial neural network prediction model for the PAH bioremediation of polluted soil. *Chemosphere*, *161*, pp.145-150.

[Weigend, Rumelhart and Huberman, 1991] Weigend, A.S., Rumelhart, D.E. and Huberman, B.A., 1991. Back-propagation, weight-elimination and time series prediction. In *Connectionist models* (pp. 105-116). Morgan kaufmann.

[Wythoff, 1993] Wythoff, B.J., 1993. Backpropagation neural networks: a tutorial. *Chemometrics and Intelligent Laboratory Systems*, *18*(2), pp.115-155.

[Xu et al., 2020] Xu, Y., Lou, H., Chen, J., Jiang, B., Yang, D., Hu, Y. and Ruan, Z., 2020. Application of a backpropagation artificial neural network in predicting plasma concentration and pharmacokinetic parameters of oral single-dose rosuvastatin in healthy subjects. *Clinical pharmacology in drug development*, *9*(7), pp.867-875.