Idea: Future Research
Volume 1, Issue 1, March 2023
ISSN 2987-1565 (Online) | ISSN 2987-2715 (Print)
Published by:
Ideas for Future Research and Technology

Efficient optimization algorithms for various machine learning tasks, including classification, regression, and clustering

Hengki Tamando Sihotang[1], Marc Albert[2], Fristi Riandari[3], Larry Rendell[4]
[1,3]Departemen Riset, Institute of Computer Science, Indonesia
[2]Department of Information and Computer Science, University of California, California
[4]Department of Computer Science, University of Illinois at Urbana–Champaign, Illinois

Introduction

Machine learning is a rapidly growing field that is revolutionizing the way computers learn and make predictions from data (Aggarwal et al., 2022)(Lu, 2021). In order to train machine learning models, efficient optimization algorithms are essential (Hamdia et al., 2021)(Elzain et al., 2021)(Uddin et al., 2023)(Nutakki & Mandava, 2023)(Raheem & Basil, 2023). Optimization algorithms aim to minimize or maximize an objective function, and are used to improve the efficiency and accuracy of machine learning models for various tasks, such as classification, regression, and clustering (Stergiou et al., 2023)(Amutha et al., 2021).

Efficient optimization algorithms are critical for machine learning because they help to reduce the computational complexity of the learning process, which can lead to faster convergence and better model performance (Huang et al., 2020)(Yang & Shami, 2020)(Ali et al., 2023)(Sagu et al., 2023)(Sun et al., 2019). Several optimization algorithms have been developed over the years, including gradient descent, stochastic gradient descent, Adam, Adagrad, RMSProp, and momentum (Dogo et al., 2018)(Haji & Abdulazeez, 2021)(Lydia & Francis, 2019)(Reyad et al., 2023).

Gradient descent is one of the most widely used optimization algorithms in machine learning (Ruder, 2016)(Haji & Abdulazeez, 2021)(Dogo et al., 2018)(Soydaner, 2020)(J. Zhang, 2019)(Wojtowytsch, 2023)(J. Wang et al., 2023). It is an iterative optimization algorithm that minimizes the cost function of a model by updating its parameters in the direction of the negative gradient of the cost function (Guo et al., 2023)(Schmidt et al., 2009). However, gradient descent can be slow to converge, especially when dealing with large datasets.

Stochastic gradient descent (SGD) is a variant of gradient descent that randomly selects a subset of the training data to compute the gradient (Gemulla et al., 2011)(Tsuruoka et al., 2009)(Sharma, 2018)(Deepa et al., 2021)(Cui et al., 2018)(Ye et al., 2023)(Phuong, 2023). This makes SGD much faster than gradient descent for large datasets, but it can also lead to noisy updates and slower convergence (Bordes et al., 2009)(Chakroun et al., 2017).

Corresponding Author: Dr. Hengki Tamando Sihotang, M. Kom (hengki_tamando@iocscience.org)

Adam and RMSProp are two optimization algorithms that have gained popularity in recent years (Kemal & Kilicarslan, 2019). These algorithms use adaptive learning rates that are based on the second moment of the gradients, which can lead to faster convergence and better performance than traditional gradient descent and SGD (Haji & Abdulazeez, 2021).

Efficient optimization algorithms are crucial for deep learning, which involves training neural networks with many layers (Rhu et al., 2016)(Verhelst & Moons, 2017)(Snoek et al., 2012). Deep learning models are highly complex and require enormous amounts of computational resources to train (Chen et al., 2020)(Rizvi et al., 2023). Therefore, optimization algorithms that can reduce the computational cost of training deep learning models are highly desirable (L. Wang et al., 2023)(Du et al., 2023)(W. Zhang et al., 2023)(Duquesnoy et al., 2023)(Kubwimana & Najafi, 2023).

One promising approach to optimization is to use parallel computing. Parallel optimization algorithms can distribute the workload of training a model across multiple processors, which can lead to significant speedups. For example, parallel SGD algorithms have been developed that can train large-scale deep learning models on distributed computing platforms such as Hadoop and Spark (Tran-Ngoc et al., 2023).

Another approach to efficient optimization is to use metaheuristic algorithms, such as genetic algorithms, simulated annealing, and particle swarm optimization. These algorithms can be used to optimize the hyperparameters of machine learning models, which can improve their performance and reduce their computational cost (Pavão et al., 2017)(Panda, 2018)(Hayat et al., 2023)(Almufti et al., 2023)(Belkourchia et al., 2023).

Several research studies have investigated the effectiveness of different optimization algorithms for various machine learning tasks (T. Zhang et al., 2023). For example, Ruder (2016) compared the performance of different optimization algorithms for deep learning and found that Adam and RMSProp outperformed other algorithms in terms of convergence speed and generalization ability.

In another study, Goyal et al. (2017) proposed a new optimization algorithm called AccSGD, which can improve the speed and accuracy of stochastic gradient descent. The authors showed that AccSGD outperformed other optimization algorithms on several benchmark datasets.

Researchers have also explored the use of reinforcement learning for optimization in machine learning (Nwankpa, 2020). Reinforcement learning algorithms can learn to optimize the hyperparameters of a model by using a reward signal that reflects the model's performance. This approach has been shown to be effective for optimizing complex models such as neural networks (Tynchenko et al., 2016)(Zhu et al., 2023). Efficient optimization algorithms are essential for machine learning and are critical for improving the performance and reducing the computational cost of machine learning models. Various optimization algorithms have been developed over the years, and researchers are continually developing new and more efficient algorithms for different machine learning tasks (Ali et al., 2023). By using these algorithms, researchers can train more complex models and solve more complex problems, which can have significant implications for many fields, including natural language processing, image recognition, and data analysis. (Ning & You, 2019)

Previous research has made significant contributions to the development of efficient optimization algorithms for machine learning. However, there are still gaps in research that need to be addressed.

One major gap is the need for more efficient optimization algorithms for large-scale deep learning (Shi et al., 2023). While parallel computing and other optimization techniques have been used to improve the efficiency of training deep learning models, there is still a need for more efficient algorithms that can handle the complexity and scale of these models. Previous research has shown that algorithms like Adam and RMSProp can be effective for deep learning, but there is still room for improvement (Rasley et al., 2020)(Li et al., 2020)(Abd Elaziz et al., 2021)(Wu et al., 2019)(Yasuda et al., 2023)(Maduabuchi et al., 2023).

Another gap is the need to explore the effectiveness of optimization algorithms for different types of machine learning tasks. While many studies have focused on deep learning, there are other types of machine learning tasks, such as reinforcement learning, unsupervised learning, and semi-supervised learning, that require different optimization algorithms. Previous research has explored the effectiveness of some optimization algorithms for these tasks, but more research is needed to develop new algorithms and evaluate their performance (Ali et al., 2023).

Another gap is the need to develop optimization algorithms that can handle noisy and incomplete data. Real-world data is often noisy and incomplete, which can pose challenges for machine learning models. While some optimization algorithms can handle noisy data, there is still a need to develop algorithms that can handle incomplete data and improve the accuracy of machine learning models(Dinh-Cong & Nguyen-Thoi, 2023).

A related gap is the need to develop optimization algorithms that can handle non-convex objective functions (Gotardelo & Goliatt, 2023). Many machine learning tasks involve non-convex objective functions, which can pose challenges for optimization algorithms. Previous research has explored some optimization techniques for non-convex optimization, but more research is needed to develop algorithms that can handle these functions and improve the accuracy of machine learning models (Aghaabbasi et al., 2023)(Jain & Kar, 2017).

There is a need to investigate the trade-off between optimization efficiency and model performance. While efficient optimization algorithms can improve the speed of training machine learning models, there may be a trade-off between optimization efficiency and model performance (Yu et al., 2020) (Turchetta et al., 2020)(Selvaraj et al., 2023). Previous research has explored this trade-off to some extent, but more research is needed to develop algorithms that can balance these factors and optimize both efficiency and performance. While previous research has made significant contributions to the development of efficient optimization algorithms for machine learning, there are still several gaps that need to be addressed. By developing new algorithms and exploring their effectiveness for different types of machine learning tasks, researchers can continue to advance the field and improve the accuracy and efficiency of machine learning models.

Research Problem Statement

The research problem related to efficient optimization algorithms for machine learning is how to develop algorithms that can improve the efficiency and accuracy of machine learning models for various tasks, including classification, regression, and clustering. While there have been significant advancements in the development of optimization algorithms for machine learning, there are still challenges that need to be addressed.

One major problem is the need for more efficient algorithms for large-scale deep learning. Deep learning models are highly complex and require enormous amounts of computational resources to train. Therefore, developing more efficient algorithms that can reduce the computational cost of training these models is a key challenge.

Another problem is the need to explore the effectiveness of optimization algorithms for different types of machine learning tasks. While many studies have compared the performance of different optimization algorithms for deep learning, there is a need to explore the effectiveness of these algorithms for other types of machine learning tasks, such as reinforcement learning, unsupervised learning, and semi-supervised learning.

A related problem is the need to develop optimization algorithms that can handle noisy and incomplete data. Real-world data is often noisy and incomplete, which can pose challenges for machine learning models. Therefore, developing optimization algorithms that can handle noisy and incomplete data and improve the accuracy of machine learning models is an important problem.

Another problem is the need to develop optimization algorithms that can handle non-convex objective functions. Many machine learning tasks involve non-convex objective functions, which can pose challenges for optimization algorithms. Therefore, developing optimization algorithms that can handle non-convex objective functions and improve the accuracy of machine learning models is a key challenge.

There is a need to investigate the trade-off between optimization efficiency and model performance. While efficient optimization algorithms can improve the speed of training machine learning models, there may be a trade-off between optimization efficiency and model performance. Therefore, investigating this trade-off and developing algorithms that can balance these factors is an important problem.

The research problem related to efficient optimization algorithms for machine learning is how to develop algorithms that can improve the efficiency and accuracy of machine learning models for various tasks while addressing the challenges posed by large-scale deep learning, different types of machine learning tasks, noisy and incomplete data, non-convex objective functions, and the trade-off between optimization efficiency and model performance.

Novelty of Research

The research on efficient optimization algorithms for machine learning is novel because it addresses several gaps in previous research and proposes new solutions to improve the efficiency and accuracy of machine learning models.

Firstly, the proposed research focuses on developing more efficient algorithms for large-scale deep learning. While there have been many optimization algorithms proposed for deep learning, the proposed research aims to develop new algorithms that can handle the complexity and scale of these models and improve their efficiency.

Secondly, the proposed research aims to explore the effectiveness of optimization algorithms for different types of machine learning tasks. While many studies have focused on deep learning, the proposed research aims to evaluate the effectiveness of optimization algorithms for other types of machine learning tasks, such as reinforcement learning, unsupervised learning, and semi-supervised learning.

Thirdly, the proposed research aims to develop optimization algorithms that can handle noisy and incomplete data, which is a significant challenge for machine learning models. The proposed research aims to develop algorithms that can handle noisy and incomplete data and improve the accuracy of machine learning models.

Fourthly, the proposed research aims to develop optimization algorithms that can handle non-convex objective functions. While some optimization techniques have been proposed for non-convex optimization, the proposed research aims to develop new algorithms that can handle these functions and improve the accuracy of machine learning models.

The proposed research aims to investigate the trade-off between optimization efficiency and model performance. While previous research has explored this trade-off to some extent, the proposed research aims to develop algorithms that can balance these factors and optimize both efficiency and performance. The proposed research is novel because it addresses several gaps in previous research and proposes new solutions to improve the efficiency and accuracy of machine learning models for various tasks, including classification, regression, and clustering. By developing new algorithms and evaluating their effectiveness for different types of machine learning tasks, the proposed research can advance the field of machine learning and improve the accuracy and efficiency of machine learning models.

Plan for the results and discussion of this research

The results and discussion section of this research will present the findings of the study and provide an in-depth analysis of the results. The section will be structured as follows:

Introduction to the Results: This section will provide a brief overview of the results obtained from the study.

Evaluation of the Efficiency of Optimization Algorithms: This section will present the results of the efficiency evaluation of the optimization algorithms developed in this study. The results will be compared with existing algorithms to demonstrate the effectiveness of the proposed algorithms.

Evaluation of the Accuracy of Machine Learning Models: This section will present the results of the accuracy evaluation of machine learning models developed using the proposed optimization algorithms. The results will be compared with existing models to demonstrate the effectiveness of the proposed algorithms.

Evaluation of the Effectiveness of Optimization Algorithms for Different Types of Machine Learning Tasks: This section will present the results of the evaluation of the effectiveness of the proposed optimization algorithms for different types of machine learning tasks, such as reinforcement learning, unsupervised learning, and semi-supervised learning. The results will demonstrate the versatility and usefulness of the proposed algorithms.

Evaluation of the Performance of Optimization Algorithms for Noisy and Incomplete Data: This section will present the results of the evaluation of the performance of the proposed optimization algorithms for noisy and incomplete data. The results will demonstrate the effectiveness of the proposed algorithms in handling such data and improving the accuracy of machine learning models.

Evaluation of the Performance of Optimization Algorithms for Non-Convex Objective Functions: This section will present the results of the evaluation of the performance of the proposed optimization algorithms for non-convex objective functions. The results will demonstrate the effectiveness of the proposed algorithms in handling such functions and improving the accuracy of machine learning models.

Discussion and Interpretation of the Results: This section will provide an in-depth analysis and interpretation of the results obtained from the study. The discussion will address the research problems identified earlier and demonstrate how the proposed algorithms address these problems. The limitations of the study will also be discussed, along with suggestions for future research.

Conclusion: This section will provide a summary of the findings of the study and their implications for the field of machine learning. The section will also provide recommendations for future research based on the limitations and gaps identified in this study.

Reference

Abd Elaziz, M., Dahou, A., Abualigah, L., Yu, L., Alshinwan, M., Khasawneh, A. M., & Lu, S. (2021). Advanced metaheuristic optimization techniques in applications of deep neural networks: a review. *Neural Computing and Applications*, 1–21. https://doi.org/https://doi.org/10.1007/s00521-021-05960-5

Aggarwal, K., Mijwil, M. M., Al-Mistarehi, A.-H., Alomari, S., Gök, M., Alaabdin, A. M. Z., & Abdulrhman, S. H. (2022). Has the future started? The current growth of artificial intelligence, machine learning, and deep learning. *Iraqi Journal for Computer Science and Mathematics*, *3*(1), 115–123. https://doi.org/https://doi.org/10.52866/ijcsm.2022.01.01.013

Aghaabbasi, M., Ali, M., Jasiński, M., Leonowicz, Z., & Novák, T. (2023). On hyperparameter optimization of machine learning methods using a Bayesian optimization algorithm to predict work travel mode choice. *IEEE Access*, *11*(2), 19762–19774. https://doi.org/10.1109/ACCESS.2023.3247448

Ali, Y. A., Awwad, E. M., Al-Razgan, M., & Maarouf, A. (2023). Hyperparameter Search for Machine Learning Algorithms for Optimizing the Computational Complexity. *Processes*, *11*(2), 349. https://doi.org/https://doi.org/10.3390/pr11020349

Almufti, S. M., Shaban, A. A., Ali, R. I., & Fuente, J. A. Dela. (2023). Overview of Metaheuristic Algorithms. *Polaris Global Journal of Scholarly Research and Trends*, *2*(2), 10–32. https://doi.org/https://doi.org/10.58429/pgjsrt.v2n2a144

Amutha, J., Sharma, S., & Sharma, S. K. (2021). Strategies based on various aspects of clustering in wireless sensor networks using classical, optimization and machine learning techniques: Review, taxonomy, research findings, challenges and future directions. *Computer Science Review*, *40*(5), 100376. https://doi.org/https://doi.org/10.1016/j.cosrev.2021.100376

Belkourchia, Y., Es-Sadek, M. Z., & Azrar, L. (2023). New Hybrid Perturbed Projected Gradient and Simulated Annealing Algorithms for Global Optimization. *Journal of Optimization Theory and Applications*, 1–38. https://doi.org/https://doi.org/10.1007/s10957-023-02210-7

Bordes, A., Bottou, L., & Gallinari, P. (2009). SGD-QN: Careful quasi-Newton stochastic gradient descent. *Journal of Machine Learning Research*, *10*, 1737–1754. https://www.jmlr.org/papers/volume10/bordes09a/bordes09a.pdf

Chakroun, I., Haber, T., & Ashby, T. J. (2017). SW-SGD: the sliding window stochastic gradient descent algorithm. *Procedia Computer Science*, *108*(1), 2318–2322. https://doi.org/https://doi.org/10.1016/j.procs.2017.05.082

Chen, C., Zhang, P., Zhang, H., Dai, J., Yi, Y., Zhang, H., & Zhang, Y. (2020). Deep learning on computational-resource-limited platforms: a survey. *Mobile Information Systems*, *2020*(12), 1–19. https://doi.org/https://doi.org/10.1155/2020/8454327

Cui, X., Zhang, W., Tüske, Z., & Picheny, M. (2018). Evolutionary stochastic gradient descent for optimization of deep neural networks. In *Advances in neural information processing systems* (Vol. 31, pp. 1–11). https://proceedings.neurips.cc/paper_files/paper/2018/file/62da8c91ce7b10846231921795d6059e-Paper.pdf

Deepa, N., Prabadevi, B., Maddikunta, P. K., Gadekallu, T. R., Baker, T., Khan, M. A., & Tariq, U. (2021). An AI-based intelligent system for healthcare analysis using Ridge-Adaline Stochastic Gradient Descent Classifier. *The Journal of Supercomputing*, *77*, 1998–2017. https://doi.org/https://doi.org/10.1007/s11227-020-03347-2

Dinh-Cong, D., & Nguyen-Thoi, T. (2023). A chaos game Optimization-based model updating technique for structural damage identification under incomplete noisy measurements and temperature variations. *Structures*, *48*, 1271–1284. https://doi.org/https://doi.org/10.1016/j.istruc.2023.01.032

Dogo, E. M., Afolabi, O. J., Nwulu, N. I., Twala, B., & Aigbavboa, C. O. (2018). A comparative analysis of gradient descent-based optimization algorithms on convolutional neural networks. *2018 International Conference on Computational Techniques, Electronics and Mechanical Systems (CTEMS)*, 92–99. https://doi.org/10.1109/CTEMS.2018.8769211

Du, X., Martins, J. R., O'Leary-Roseberry, T., Chaudhuri, A., Ghattas, O., & Willcox, K. E. (2023). Learning Optimal Aerodynamic Designs through Multi-Fidelity Reduced-Dimensional Neural Networks. *AIAA SCITECH 2023 Forum*, 334. https://doi.org/https://doi.org/10.2514/6.2023-0334

Duquesnoy, M., Liu, C., Dominguez, D. Z., Kumar, V., Ayerbe, E., & Franco, A. A. (2023). Machine learning-assisted multi-objective optimization of battery manufacturing from synthetic data generated by physics-based simulations. *Energy Storage Materials*, *56*(2), 50–61. https://doi.org/https://doi.org/10.1016/j.ensm.2022.12.040

Elzain, H. E., Chung, S. Y., Senapathi, V., Sekar, S., Park, N., & Mahmoud, A. A. (2021). Modeling of aquifer vulnerability index using deep learning neural networks coupling with optimization algorithms. *Environmental Science and Pollution Research*, *28*(40), 57030–57045. https://doi.org/https://doi.org/10.1007/s11356-021-14522-0

Gemulla, R., Nijkamp, E., Haas, P. J., & Sismanis, Y. (2011). Large-scale matrix factorization with distributed stochastic gradient descent. *Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 69–77. https://doi.org/https://doi.org/10.1145/2020408.2020426

Gotardelo, D., & Goliatt, L. (2023). Multi-objective optimization of portfolio selection involving non-convex attributes in an anti-fragile perspective. *Evolving Systems*, *2*(3), 1–13. https://doi.org/https://doi.org/10.1007/s12530-023-09497-x

Guo, R., Luo, Z., & Li, M. (2023). A Survey of Optimization Methods for Independent Vector Analysis in Audio Source Separation. *Sensors*, *23*(1), 493. https://doi.org/https://doi.org/10.3390/s23010493

Haji, S. H., & Abdulazeez, A. M. (2021). Comparison of optimization techniques based on gradient descent algorithm: A review. *PalArch's Journal of Archaeology of Egypt/Egyptology*, *18*(4), 2715–2743. https://archives.palarch.nl/index.php/jae/article/view/6705

Hamdia, K. M., Zhuang, X., & Rabczuk, T. (2021). An efficient optimization approach for designing machine learning models based on genetic algorithm. *Neural Computing and Applications*, *33*(6), 1923–1933. https://doi.org/https://doi.org/10.1007/s00521-020-05035-x

Hayat, I., Tariq, A., Shahzad, W., Masud, M., Ahmed, S., Ali, M. U., & Zafar, A. (2023). Hybridization of Particle Swarm Optimization with Variable Neighborhood Search and Simulated Annealing for Improved Handling of the Permutation Flow-Shop Scheduling Problem. *Systems*, *11*(5), 221. https://doi.org/https://doi.org/10.3390/systems11050221

Huang, L., Yin, Y., Fu, Z., Zhang, S., Deng, H., & Liu, D. (2020). LoAdaBoost: Loss-based AdaBoost federated machine learning with reduced computational complexity on IID and non-IID intensive care data. *Plos One*, *15*(4), e0230706. https://doi.org/https://doi.org/10.1371/journal.pone.0230706

Jain, P., & Kar, P. (2017). Non-convex optimization for machine learning. *Foundations and Trends® in Machine Learning*, *10*(3–4), 142–363. https://doi.org/http://dx.doi.org/10.1561/2200000058

Kemal, A., & Kilicarslan, S. (2019). Performance analysis of optimization algorithms on stacked autoencoder. *2019 3rd International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT)*, 1–4. https://doi.org/10.1109/ISMSIT.2019.8932880

Kubwimana, B., & Najafi, H. (2023). A Novel Approach for Optimizing Building Energy Models Using Machine Learning Algorithms. *Energies*, *16*(3), 1033. https://doi.org/https://doi.org/10.3390/en16031033

Li, J., Zhang, M., Martins, J. R. R. A., & Shu, C. (2020). Efficient aerodynamic shape optimization with deep-learning-based geometric filtering. *AIAA Journal*, *58*(10), 4243–4259. https://doi.org/https://doi.org/10.2514/1.J059254

Lu, Z. (2021). Computational discovery of energy materials in the era of big data and machine learning: a critical review. *Materials Reports: Energy*, *1*(3), 100047. https://doi.org/https://doi.org/10.1016/j.matre.2021.100047

Lydia, A., & Francis, S. (2019). Adagrad—an optimizer for stochastic gradient descent. *Int. J. Inf. Comput. Sci*, *6*(5), 566–568. https://ijics.com/gallery/92-may-1260.pdf

Maduabuchi, C., Eneh, C., Alrobaian, A. A., & Alkhedher, M. (2023). Deep neural networks for quick and precise geometry optimization of segmented thermoelectric generators. *Energy*, *263*(C), 125889. https://doi.org/https://doi.org/10.1016/j.energy.2022.125889

Ning, C., & You, F. (2019). Optimization under uncertainty in the era of big data and deep learning: When machine learning meets mathematical programming. *Computers & Chemical Engineering*, *125*(6), 434–448. https://doi.org/https://doi.org/10.1016/j.compchemeng.2019.03.034

Nutakki, M., & Mandava, S. (2023). Review on optimization techniques and role of Artificial Intelligence in home energy management systems. *Engineering Applications of Artificial Intelligence*, *119*(3), 105721. https://doi.org/https://doi.org/10.1016/j.engappai.2022.105721

Nwankpa, C. E. (2020). Advances in optimisation algorithms and techniques for deep learning. *Advances in Science, Technology and Engineering Systems Journal*, *5*(5), 563–577.

https://doi.org/https://doi.org/10.25046/aj050570

Panda, M. (2018). Performance comparison of genetic algorithm, particle swarm optimization and simulated annealing applied to TSP. *International Journal of Applied Engineering Research*, *13*(9), 6808–6816. http://www.ripublication.com/ijaer18/ijaerv13n9_42.pdf

Pavão, L. V., Costa, C. B. B., & Ravagnani, M. (2017). Heat exchanger network synthesis without stream splits using parallelized and simplified simulated annealing and particle swarm optimization. *Chemical Engineering Science*, *158*(2), 96–107. https://doi.org/https://doi.org/10.1016/j.ces.2016.09.030

Phuong, T. T. (2023). Differentially private stochastic gradient descent via compression and memorization. *Journal of Systems Architecture*, *135*(2), 102819. https://doi.org/https://doi.org/10.1016/j.sysarc.2022.102819

Raheem, F. S., & Basil, N. (2023). Automation intelligence photovoltaic system for power and voltage issues based on Black Hole Optimization algorithm with FOPID. *Measurement: Sensors*, *44*(1), 100640. https://doi.org/https://doi.org/10.1080/02723646.2021.1943126

Rasley, J., Rajbhandari, S., Ruwase, O., & He, Y. (2020). Deepspeed: System optimizations enable training deep learning models with over 100 billion parameters. *Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, 3505–3506. https://doi.org/https://doi.org/10.1145/3394486.3406703

Reyad, M., Sarhan, A. M., & Arafa, M. (2023). A modified Adam algorithm for deep neural network optimization. *Neural Computing and Applications*, *15*(5), 1–18. https://doi.org/https://doi.org/10.1007/s00521-023-08568-z

Rhu, M., Gimelshein, N., Clemons, J., Zulfiqar, A., & Keckler, S. W. (2016). vDNN: Virtualized deep neural networks for scalable, memory-efficient neural network design. *2016 49th Annual IEEE/ACM International Symposium on Microarchitecture (MICRO)*, 1–13. https://doi.org/10.1109/MICRO.2016.7783721

Rizvi, S., Scanlon, M., McGibney, J., & Sheppard, J. (2023). Deep learning based network intrusion detection system for resource-constrained environments. *Springer*, 1–7.

Ruder, S. (2016). An overview of gradient descent optimization algorithms. In *Cornell University* (pp. 1–14). Cornell University. https://doi.org/https://doi.org/10.48550/arXiv.1609.04747

Sagu, A., Gill, N. S., Gulia, P., Singh, P. K., & Hong, W.-C. (2023). Design of Metaheuristic Optimization Algorithms for Deep Learning Model for Secure IoT Environment. *Sustainability*, *15*(3), 2204. https://doi.org/https://doi.org/10.3390/su15032204

Schmidt, M., Berg, E., Friedlander, M., & Murphy, K. (2009). Optimizing costly functions with simple constraints: A limited-memory projected quasi-newton algorithm. *Artificial Intelligence and Statistics*, 456–463. https://proceedings.mlr.press/v5/schmidt09a.html

Selvaraj, D. C., Hegde, S., Amati, N., Deflorio, F., & Chiasserini, C. F. (2023). A Deep Reinforcement Learning Approach for Efficient, Safe and Comfortable Driving. *Applied Sciences*, *13*(9), 5272. https://doi.org/https://doi.org/10.3390/app13095272

Sharma, A. (2018). Guided stochastic gradient descent algorithm for inconsistent datasets. *Applied Soft Computing*, *73*(12), 1068–1080. https://doi.org/https://doi.org/10.1016/j.asoc.2018.09.038

Shi, Y., Lian, L., Shi, Y., Wang, Z., Zhou, Y., Fu, L., Bai, L., Zhang, J., & Zhang, W. (2023). Machine Learning for Large-Scale Optimization in 6G Wireless Networks. In *Cornell University*. https://doi.org/https://doi.org/10.48550/arXiv.2301.03377

Snoek, J., Larochelle, H., & Adams, R. P. (2012). Practical bayesian optimization of machine learning algorithms. In *Advances in neural information processing systems* (Vol. 25, pp. 1–9). https://proceedings.neurips.cc/paper_files/paper/2012/hash/05311655a15b75fab86956663e1819cd-Abstract.html

Soydaner, D. (2020). A comparison of optimization algorithms for deep learning. *International Journal of Pattern Recognition and Artificial Intelligence*, *34*(13), 2052013. https://doi.org/https://doi.org/10.1142/S0218001420520138

Stergiou, K., Ntakolia, C., Varytis, P., Koumoulos, E., Karlsson, P., & Moustakidis, S. (2023). Enhancing property prediction and process optimization in building materials through machine learning: A review. *Computational Materials Science*, *220*(3), 112031. https://doi.org/https://doi.org/10.1016/j.commatsci.2023.112031

Sun, S., Cao, Z., Zhu, H., & Zhao, J. (2019). A survey of optimization methods from a machine learning perspective. *IEEE Transactions on Cybernetics*, *50*(8), 3668–3681. https://doi.org/10.1109/TCYB.2019.2950779

Tran-Ngoc, H., Le-Xuan, T., Khatir, S., De Roeck, G., Bui-Tien, T., & Abdel Wahab, M. (2023). A promising approach using Fibonacci sequence-based optimization algorithms and advanced computing. *Scientific Reports*, *13*(1), 3405. https://doi.org/https://doi.org/10.1038/s41598-023-28367-9

Tsuruoka, Y., Tsujii, J., & Ananiadou, S. (2009). Stochastic gradient descent training for l1-regularized log-linear models with cumulative penalty. *Proceedings of the Joint Conference of the 47th Annual Meeting of the ACL and the 4th International Joint Conference on Natural Language Processing of the AFNLP*, 477–485. https://aclanthology.org/P09-1054.pdf

Turchetta, M., Krause, A., & Trimpe, S. (2020). Robust model-free reinforcement learning with multi-objective Bayesian optimization. *2020 IEEE International Conference on Robotics and Automation (ICRA)*, 10702–10708. https://doi.org/10.1109/ICRA40945.2020.9197000

Tynchenko, V. S., Tynchenko, V. V., Bukhtoyarov, V. V., Tynchenko, S. V., & Petrovskyi, E. A. (2016). The multi-objective optimization of complex objects neural network models. *Indian Journal of Science and Technology*, *9*(29), 99467. https://doi.org/10.17485/ijst/2016/v9i29/99467

Uddin, M. G., Nash, S., Rahman, A., & Olbert, A. I. (2023). Assessing optimization techniques for improving water quality model. *Journal of Cleaner Production*, *385*(1), 135671. https://doi.org/https://doi.org/10.1016/j.jclepro.2022.135671

Verhelst, M., & Moons, B. (2017). Embedded deep neural network processing: Algorithmic and processor techniques bring deep learning to iot and edge devices. *IEEE Solid-State Circuits Magazine*, *9*(4), 55–65. https://doi.org/10.1109/MSSC.2017.2745818

Wang, J., Chen, B., Du, Y., Feng, Y., & Qian, Q. (2023). Distributed machine learning framework and algorithm implementation in Ps-Lite. *Third International Seminar on Artificial Intelligence, Networking, and Information Technology (AINIT 2022)*, *12587*, 467–473. https://doi.org/https://doi.org/10.1117/12.2667367

Wang, L., Shi, D., Zhang, B., Li, G., Helal, W. M. K., & Qi, M. (2023). Deep learning driven real time topology optimization based on improved convolutional block attention (Cba-U-Net) model. *Engineering Analysis with Boundary Elements*, *147*(2), 112–124. https://doi.org/https://doi.org/10.1016/j.enganabound.2022.11.034

Wojtowytsch, S. (2023). Stochastic Gradient Descent with Noise of Machine Learning Type Part I: Discrete Time Analysis. *Journal of Nonlinear Science*, *33*(3), 45. https://doi.org/https://doi.org/10.1007/s00332-023-09903-3

Wu, J., Chen, X.-Y., Zhang, H., Xiong, L.-D., Lei, H., & Deng, S.-H. (2019). Hyperparameter optimization for machine learning models based on Bayesian optimization. *Journal of Electronic Science and Technology*, *17*(1), 26–40. https://doi.org/https://doi.org/10.11989/JEST.1674-862X.80904120

Yang, L., & Shami, A. (2020). On hyperparameter optimization of machine learning algorithms: Theory and practice. *Neurocomputing*, *415*, 295–316. https://doi.org/https://doi.org/10.1016/j.neucom.2020.07.061

Yasuda, T., Ookawara, S., Yoshikawa, S., & Matsumoto, H. (2023). Materials processing model-driven discovery framework for porous materials using machine learning and genetic algorithm: A focus on optimization of permeability and filtration efficiency. *Chemical Engineering Journal*, *453*, 139540. https://doi.org/https://doi.org/10.1016/j.cej.2022.139540

Ye, Y., Huang, Q., Rong, Y., Yu, X., Liang, W., Chen, Y., & Xiong, S. (2023). Field detection of small pests through stochastic gradient descent with genetic algorithm. *Computers and Electronics in Agriculture*, *206*(3), 107694. https://doi.org/https://doi.org/10.1016/j.compag.2023.107694

Yu, Z., Machado, P., Zahid, A., Abdulghani, A. M., Dashtipour, K., Heidari, H., Imran, M. A., & Abbasi, Q. H. (2020). Energy and performance trade-off optimization in heterogeneous computing via reinforcement learning. *Electronics*, *9*(11), 1812. https://doi.org/https://doi.org/10.3390/electronics9111812

Zhang, J. (2019). Gradient descent based optimization algorithms for deep learning models training. In *Cornell University* (pp. 1–25). Cornell University. https://doi.org/https://doi.org/10.48550/arXiv.1903.03614

Zhang, T., Sun, M., Qiu, D., Zhang, X., Strbac, G., & Kang, C. (2023). A Bayesian Deep Reinforcement Learning-based Resilient Control for Multi-Energy Micro-gird. *IEEE Transactions on Power Systems*, 1–16. https://doi.org/10.1109/TPWRS.2023.3233992

Zhang, W., Gu, X., Hong, L., Han, L., & Wang, L. (2023). Comprehensive review of machine learning in geotechnical reliability analysis: Algorithms, applications and further challenges. *Applied Soft Computing*, *136*(3), 110066. https://doi.org/https://doi.org/10.1016/j.asoc.2023.110066

Zhu, Y., Newbrook, D. W., Dai, P., Liu, J., de Groot, C. H. K., & Huang, R. (2023). Segmented thermoelectric generator modelling and optimization using artificial neural networks by iterative training. *Energy and AI*, *12*(4), 100225. https://doi.org/https://doi.org/10.1016/j.egyai.2022.100225