

2023

Tutorial - Shodhguru Labs: Optimization and Hyperparameter Tuning for Neural Networks

Kaushik Roy
University of South Carolina - Columbia

Follow this and additional works at: https://scholarcommons.sc.edu/aii_fac_pub



Part of the [Computer Engineering Commons](#), and the [Electrical and Computer Engineering Commons](#)

Publication Info

Preprint version 2023.

© Kaushik Roy, 2023

This Article is brought to you by the Artificial Intelligence Institute at Scholar Commons. It has been accepted for inclusion in Publications by an authorized administrator of Scholar Commons. For more information, please contact digres@mailbox.sc.edu.

Tutorial - Shodhguru Labs: Optimization and Hyperparameter Tuning for Neural Networks

Kaushik Roy

kaushikr@email.sc.edu

AI Institute, University of South Carolina
Columbia, SC, USA

ABSTRACT

Neural networks have emerged as a powerful and versatile class of machine learning models, revolutionizing various fields with their ability to learn complex patterns and make accurate predictions. The performance of neural networks depends significantly on the appropriate choice of hyperparameters, which are critical factors governing their architecture, regularization, and optimization techniques. As the demand for high-performance neural networks grows across diverse applications, the need for efficient optimization and hyperparameter tuning methods becomes paramount. This paper presents a comprehensive exploration of optimization strategies and hyperparameter tuning techniques for neural networks. Neural networks have emerged as a powerful and versatile class of machine learning models, revolutionizing various fields with their ability to learn complex patterns and make accurate predictions. The performance of neural networks depends significantly on the appropriate choice of hyperparameters, which are critical factors governing their architecture, regularization, and optimization techniques. As the demand for high-performance neural networks grows across diverse applications, the need for efficient optimization and hyperparameter tuning methods becomes paramount. This tutorial presents an exploration of optimization strategies and hyperparameter tuning techniques for neural networks using state-of-the-art Python libraries.

1 OVERVIEW OF PYTHON LIBRARIES

Several state-of-the-art Python libraries for hyperparameter tuning and optimization scheduling were widely used in the machine learning community. Some of the popular libraries include:

1. Optuna

Optuna is a powerful hyperparameter optimization framework based on the Bayesian optimization technique. It offers a simple and intuitive interface, allowing users to define a search space for hyperparameters and optimize their models efficiently. Optuna supports various algorithms, including Tree-structured Parzen Estimator (TPE) and CMA-ES, and integrates seamlessly with popular machine learning libraries like TensorFlow, PyTorch, and scikit-learn [1].

2. Hyperopt

Hyperopt is another Bayesian optimization library designed for hyperparameter tuning. It provides various optimization algorithms, such as Random Search, TPE, and Sequential Model-Based Optimization (SMBO). Hyperopt is widely used for automating the search for hyperparameter configurations and has integrations with popular frameworks like Keras, XGBoost, and lightgbm [2].

3. scikit-optimize

Scikit-optimize, also known as skopt, is a versatile library that combines Bayesian optimization and sequential model-based optimization techniques. It offers easy-to-use optimization functions, enabling users to tune hyperparameters efficiently. Skopt can be used alongside popular machine learning libraries like scikit-learn and XGBoost [3].

Tune (Ray Tune):

Tune is part of the Ray project, a general-purpose distributed computing framework. It provides a scalable solution for hyperparameter tuning, and it supports popular algorithms like Random Search, HyperBand, and Population-Based Training (PBT). Ray Tune is especially suitable for distributed hyperparameter tuning and is designed for parallel and cloud-based optimization [4].

Ax Platform:

Developed by Facebook, Ax is a comprehensive platform for managing experiments and hyperparameter tuning. It combines Bayesian optimization, bandit optimization, and multi-fidelity optimization techniques. Ax is particularly well-suited for conducting large-scale experiments and tuning complex models [5].

BoTorch:

BoTorch is another library from Facebook AI, focused on Bayesian optimization for optimizing functions that are expensive to evaluate. It's built on top of PyTorch, making it well-suited for optimizing complex neural network architectures and hyperparameter spaces [6].

Keras Tuner:

Keras Tuner, an official Keras add-on, provides hyperparameter tuning capabilities for models built with Keras. It supports various search algorithms like Random Search, Hyperband, and Bayesian optimization. Keras Tuner offers an easy-to-use API for tuning hyperparameters within Keras models [7].

These libraries offer various optimization techniques, including Bayesian optimization, random search, and more. Depending on your specific use case and the complexity of the models you are working with, one of these libraries should suit your needs for hyperparameter tuning and optimization scheduling in Python.

2 TUTORIAL ORGANIZER BIO

Kaushik Roy: He is a Ph.D. student at the Artificial Intelligence Institute at the University of South Carolina (AIISC). He completed his master's in computer science at Indiana University Bloomington and has worked at UT Dallas's Starling lab. His research interests include statistical relational artificial intelligence, sequential decision-making, knowledge graphs, and reinforcement Learning. His work has been published in reputed Artificial Intelligence-related conferences, journals, symposia, and workshops [8-40]. Please visit his webpage.

3 TUTORIAL AUDIENCE AND PREREQUISITES

This tutorial will benefit academic and industry researchers, professionals, and Machine Learning (ML) practitioners. Only basic knowledge of traditional ML pipelines is required.

REFERENCES

- [1] Takuya Akiba, Shotaro Sano, Toshihiko Yanase, Takeru Ohta, and Masanori Koyama. Optuna: A next-generation hyperparameter optimization framework. In *Proceedings of the 25th ACM SIGKDD international conference on knowledge discovery & data mining*, pages 2623–2631, 2019.
- [2] James Bergstra, Dan Yamins, David D Cox, et al. Hyperopt: A python library for optimizing the hyperparameters of machine learning algorithms. In *Proceedings of the 12th Python in science conference*, volume 13, page 20. Citeseer, 2013.
- [3] Tim Head, Gilles Louppe, MechCoder, Iaroslav Shcherbatyi, et al. scikit-optimize/scikit-optimize: v0. 5.2. *Version v0*, 5, 2018.
- [4] Richard Liaw, Eric Liang, Robert Nishihara, Philipp Moritz, Joseph E Gonzalez, and Ion Stoica. Tune: A research platform for distributed model selection and training. *arXiv preprint arXiv:1807.05118*, 2018.
- [5] Sterling G Baird, Marianne Liu, and Taylor D Sparks. High-dimensional bayesian optimization of 23 hyperparameters over 100 iterations for an attention-based network to predict materials property: A case study on crabnet using ax platform and saasbo. *Computational Materials Science*, 211:111505, 2022.
- [6] Maximilian Balandat, Brian Karrer, Daniel Jiang, Samuel Daulton, Ben Letham, Andrew G Wilson, and Eytan Bakshy. Botorch: A framework for efficient monte-carlo bayesian optimization. *Advances in neural information processing systems*, 33:21524–21538, 2020.
- [7] Mohamad Zaim Awang Pon and Krishna Prakash KK. Hyperparameter tuning of deep learning models in keras. *Sparklinglight Transactions on Artificial Intelligence and Quantum Computing (STAIQC)*, 1(1):36–40, 2021.
- [8] Kaushik Roy, Yuxin Zi, Vignesh Narayanan, Manas Gaur, and Amit Sheth. Knowledge-infused self attention transformers. *arXiv preprint arXiv:2306.13501*, 2023.
- [9] Kaushik Roy, Yuxin Zi, Vignesh Narayanan, Manas Gaur, and Amit Sheth. Ksat: Knowledge-infused self attention transformer—integrating multiple domain-specific contexts. *arXiv preprint arXiv:2210.04307*, 2022.
- [10] Yuxin Zi, Kaushik Roy, Vignesh Narayanan, Manas Gaur, and Amit Sheth. Ierl: Interpretable ensemble representation learning—combining crowdsourced knowledge and distributed semantic representations. *arXiv preprint arXiv:2306.13865*, 2023.
- [11] Kaushik Roy, Yuxin Zi, Manas Gaur, Jinendra Malekar, Qi Zhang, Vignesh Narayanan, and Amit Sheth. Process knowledge-infused learning for clinician-friendly explanations. *arXiv preprint arXiv:2306.09824*, 2023.
- [12] Amit Sheth, Kaushik Roy, and Manas Gaur. Neurosymbolic artificial intelligence (why, what, and how). *IEEE Intelligent Systems*, 38(3):56–62, 2023.
- [13] Revathy Venkataraman, Kaushik Roy, Kanak Raj, Renjith Prasad, Yuxin Zi, Vignesh Narayanan, and Amit Sheth. Cook-gen: Robust generative modeling of cooking actions from recipes. *arXiv preprint arXiv:2306.01805*, 2023.
- [14] Tarun Garg, Kaushik Roy, and Amit Sheth. Can language models capture graph semantics? from graphs to language model and vice-versa. *arXiv preprint arXiv:2206.09259*, 2022.
- [15] Usha Lokala, Kaushik Roy, Utkarshani Jaimini, and Amit Sheth. Acm web conference 2023. 2023.
- [16] Kaushik Roy, Vedant Khandelwal, Raxit Goswami, Nathan Dolbir, Jinendra Malekar, and Amit Sheth. Demo alleviate: Demonstrating artificial intelligence enabled virtual assistance for telehealth: The mental health case. *arXiv preprint arXiv:2304.00025*, 2023.
- [17] Kaushik Roy, Manas Gaur, Misagh Soltani, Vipula Rawte, Ashwin Kalyan, and Amit Sheth. Proknow: Process knowledge for safety constrained and explainable question generation for mental health diagnostic assistance. *Frontiers in big Data*, 5:1056728, 2023.
- [18] Amit Sheth, Manas Gaur, Kaushik Roy, Revathy Venkataraman, and Vedant Khandelwal. Process knowledge-infused ai: Toward user-level explainability, interpretability, and safety. *IEEE Internet Computing*, 26(5):76–84, 2022.
- [19] Adam Tsakalidis, Jenny Chim, Iman Munire Bilal, Ayah Zirikly, Dana Atzil-Slonim, Federico Nanni, Philip Resnik, Manas Gaur, Kaushik Roy, Becky Inkster, et al. Overview of the clpsych 2022 shared task: Capturing moments of change in longitudinal user posts. In *Proceedings of the Eighth Workshop on Computational Linguistics and Clinical Psychology*, pages 184–198, 2022.
- [20] Shrey Gupta, Anmol Agarwal, Manas Gaur, Kaushik Roy, Vignesh Narayanan, Ponnurangam Kumaraguru, and Amit Sheth. Learning to automate follow-up question generation using process knowledge for depression triage on reddit posts. *arXiv preprint arXiv:2205.13884*, 2022.
- [21] Kaushik Roy, Manas Gaur, Qi Zhang, and Amit Sheth. Knowledge-infused reinforcement learning, 2022.
- [22] Kaushik Roy, Usha Lokala, Manas Gaur, and Amit P Sheth. Tutorial: Neuro-symbolic ai for mental healthcare. In *Proceedings of the Second International Conference on AI-ML Systems*, pages 1–3, 2022.
- [23] Vipula Rawte, Megha Chakraborty, Kaushik Roy, Manas Gaur, Keyur Faldu, Prashant Kikani, Hemang Akbari, and Amit Sheth. Tdlr: Top (semantic)-down (syntactic) language representation. *UMBC Faculty Collection*, 2022.
- [24] Sandeep K Chaudhuri, Joshua W Kleppinger, Ritwik Nag, Kaushik Roy, Rojina Panta, Forest Agostinelli, Amit Sheth, Utpal N Roy, Ralph B James, and Krishna C Mandal. A cdzntese gamma spectrometer trained by deep convolutional neural network for radioisotope identification. In *Hard X-Ray, Gamma-Ray, and Neutron Detector Physics XXIII*, volume 11838, page 1183806. SPIE, 2021.
- [25] Amit Sheth, Manas Gaur, Kaushik Roy, and Keyur Faldu. Knowledge-intensive language understanding for explainable ai. *IEEE Internet Computing*, 25(5):19–24, 2021.
- [26] Manas Gaur, Kaushik Roy, Aditya Sharma, Biplav Srivastava, and Amit Sheth. “who can help me?”: Knowledge infused matching of support seekers and support providers during covid-19 on reddit. In *2021 IEEE 9th International Conference on Healthcare Informatics (ICHI)*, pages 265–269. IEEE, 2021.
- [27] Nathan Dolbir, Triyasha Dastidar, and Kaushik Roy. Nlp is not enough—contextualization of user input in chatbots. *arXiv preprint arXiv:2105.06511*, 2021.
- [28] Usha Lokala, Francois Lamy, Triyasha Ghosh Dastidar, Kaushik Roy, Raminta Daniulaityte, Srinivasan Parthasarathy, and Amit Sheth. edarktrends: Harnessing social media trends in substance use disorders for opioid listings on cryptomarket. *arXiv preprint arXiv:2103.15764*, 2021.
- [29] Kaushik Roy, Qi Zhang, Manas Gaur, and Amit Sheth. Knowledge infused policy gradients for adaptive pandemic control. *arXiv preprint arXiv:2102.06245*, 2021.
- [30] Kaushik Roy, Usha Lokala, Vedant Khandelwal, and Amit Sheth. “is depression related to cannabis?”: A knowledge-infused model for entity and relation extraction with limited supervision. *arXiv preprint arXiv:2102.01222*, 2021.
- [31] Usha Lokala, Manas Gaur, Kaushik Roy, and Amit Sheth. Knowledge infused natural language understanding for public health, epidemiology, and substance use. 2021.
- [32] Kaushik Roy, Qi Zhang, Manas Gaur, and Amit Sheth. Knowledge infused policy gradients with upper confidence bound for relational bandits. In *Machine Learning and Knowledge Discovery in Databases. Research Track: European Conference, ECML PKDD 2021, Bilbao, Spain, September 13–17, 2021, Proceedings, Part I 21*, pages 35–50. Springer, 2021.
- [33] Amit Sheth, Kaushik Roy, Manas Gaur, and Usha Lokala. Tutorial on knowledge in-wisdom out-explainable data for ai in cyber social threats and public health. 2021.
- [34] Parth Asawa, Manas Gaur, Kaushik Roy, and Amit Sheth. Covid-19 in spain and india: comparing policy implications by analyzing epidemiological and social media data. *arXiv preprint arXiv:2010.14628*, 2020.
- [35] Srijita Das, Sriraam Natarajan, Kaushik Roy, Ronald Parr, and Kristian Kersting. Fitted q-learning for relational domains. *arXiv preprint arXiv:2006.05595*, 2020.
- [36] Amit Bhat and Kaushik Roy. Optimized knn for sequential execution. In *2016 International Conference on Inventive Computation Technologies (ICICT)*, volume 1, pages 1–6. IEEE, 2016.
- [37] Azra Nasreen, Kaushik Roy, Kunal Roy, and G Shobha. Key frame extraction and foreground modelling using k-means clustering. In *2015 7th International Conference on Computational Intelligence, Communication Systems and Networks*, pages 141–145. IEEE, 2015.
- [38] Amit G Bhat, Kaushik Roy, Prajesh P Anchalia, and HM Jeevith. Design and implementation of a dynamic intelligent traffic control system. In *2015 17th UKSim-AMSS International Conference on Modelling and Simulation (UKSim)*, pages 369–373. IEEE, 2015.
- [39] Prajesh P Anchalia, Kaushik Roy, and Kunal Roy. Two class fisher’s linear discriminant analysis using mapreduce. In *2015 17th UKSim-AMSS International Conference on Modelling and Simulation (UKSim)*, pages 463–469. IEEE, 2015.
- [40] Prajesh P Anchalia and Kaushik Roy. The k-nearest neighbor algorithm using mapreduce paradigm. In *2014 5th International Conference on Intelligent Systems, Modelling and Simulation*, pages 513–518. IEEE, 2014.