iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 12 **Issue:** IV **Month of publication:** April 2024

DOI: https://doi.org/10.22214/ijraset.2024.60597

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 12 Issue IV Apr 2024- Available at www.ijraset.com

MLCompare: Platform to Compare Machine Learning Models

Pal Sonam Vinod[1], Wagh Shraddha Balasaheb[2], Yadav Mandar Pandurang[3], Yadgude Samrudhi Ravindra[4], Dr. S F Sayyad[5]
[1, 2, 3, 4, 5]*Computer Department, AISSMS College of Engineering, Savitribai Phule Pune University*

Abstract: ***The field of machine learning (ML) continues to advance, the demand for efficient and effective methods for model development and selection becomes increasingly pressing. Automated Machine Learning (AutoML) platforms have emerged as powerful tools to address this need by automating the process of model selection, hyperparameter tuning, and feature engineering. By integrating a wide array of ML techniques and leveraging cloud computing resources, the platform enables researchers and practitioners to efficiently explore and compare the performance of various algorithms across diverse datasets. Key features include a user-friendly interface, a robust evaluation framework encompassing traditional metrics and advanced techniques, and parallel processing capabilities. Through experiments on benchmark datasets from multiple domains, the platform demonstrates its ability to streamline algorithm selection and provide valuable insights into algorithm performance. This contribution advances the field of AutoML by facilitating informed decision-making and accelerating innovation in ML applications.***

Keywords: AutoML, machine learning algorithms, comparative analysis, automated model selection, hyperparameter tuning, feature engineering

I. INTRODUCTION

In the field of machine learning (ML), the evaluation and comparison of models is a critical process which ensures optimal performance and informed decision-making. However, these tasks often demand significant time and effort, involving manual steps such as selecting target variables, cleaning data, and choosing appropriate models. To streamline this process, we introduce an innovative AutoML (Automated Machine Learning) platform tailored for comparative analysis of ML models. Our platform offers two distinct systems: a manual system and an automated system.

A. Manual System

The manual system empowers users with hands-on control over the testing and evaluation process. Researchers and practitioners can carefully select target variables, clean data, and choose models for comparison, allowing for customized experimentation and detailed insights into model performance. This system caters to those who value flexibility and seek a deep understanding of the intricacies involved in ML model evaluation.

B. Automated System

The automated system innovates ML model evaluation by automating repetitive tasks such as data preprocessing, model selection, and performance evaluation. By leveraging predefined workflows, this system accelerates experimentation while ensuring consistency and reproducibility. Researchers benefit from time efficiency and scalability, enabling them to focus on higher-level tasks such as result interpretation and hypothesis generation.

II. BACKGROUND

Machine learning (ML) has emerged as a powerful tool for data analysis and decision-making across various industries. As the demand for ML solutions continues to rise, there is a growing need for efficient methods to evaluate and compare the performance of ML models. Traditionally, this process involves manual intervention at various stages, including data preprocessing, feature selection, model training, and performance evaluation. However, manual methods are often time-consuming, labour-intensive, and prone to human error. To address these challenges, automated machine learning (AutoML) has gained popularity as a promising approach to streamline the model development and evaluation process. AutoML algorithms automate key tasks such as feature engineering, model selection, hyperparameter tuning, and performance optimization, reducing the burden on human analysts and accelerating the time-to-insight.

©IJRASET: All Rights are Reserved

Despite the benefits of AutoML, there remains a need for platforms that facilitate comparative analysis of ML models. Researchers and practitioners often seek to explore multiple models and algorithms to identify the most suitable solution for their specific task or problem domain. Additionally, they may wish to compare the performance of different models under various experimental conditions or datasets. In response to these needs, our research introduces an AutoML platform specifically designed for comparative analysis of ML models. This platform offers both manual and automated systems, providing users with the flexibility to choose the approach that best suits their requirements. By combining the advantages of manual control with the efficiency of automation, our platform aims to empower researchers and practitioners with comprehensive tools for informed decision-making in the field of machine learning.

III. METHODOLOGIS

Throughout the project phases we have structured into several distinct phases, each contributing to the systematic development and implementation of the AutoML platform customized for comparative analysis of ML models. We initiate our methodology by undertaking the design phase, wherein we carefully conceptualize the architecture and features of the AutoML platform. This involves in-depth exploration sessions to define user interfaces that prioritize intuitive interactions, as well as the backend infrastructure required to support data processing, model training, and evaluation. Additionally, we focus on the integration of essential libraries and frameworks to ensure compatibility and efficiency throughout the platform.

Following the design phase, we transition to the data preparation phase, recognizing the critical importance of high-quality data. Here, we establish protocols for data ingestion, cleaning, and preprocessing. Our objective is to ensure data integrity and consistency across diverse datasets and experimental scenarios, laying a solid foundation for reliable model evaluation.

With a well-prepared dataset in hand, we proceed to implement the manual system within the AutoML platform. This involves the development of user-friendly interfaces that empower researchers and practitioners to exercise manual control over key aspects of the experimentation process. Users are granted the flexibility to handpick target variables, customize data preprocessing steps, and arrange a selection of models for comparative analysis, thus promoting a deeper understanding of the underlying mechanisms driving model performance. Simultaneously, we embark on the development of the automated system, aimed at streamlining the model development and evaluation process. Leveraging state-of-the-art AutoML algorithms and techniques, we automate tasks such as feature engineering, model selection, hyperparameter tuning, and performance evaluation. By harnessing the power of automation, we seek to expedite experimentation while ensuring consistency and reproducibility across different experimental setups.

As both the manual and automated systems take shape, we proceed to integrate them seamlessly into the AutoML platform. Thorough testing is conducted at each integration point to validate functionality, stability, and reliability. Benchmark datasets and real-world use cases are employed to assess platform performance comprehensively, with a keen focus on identifying and addressing any potential bottlenecks or issues. User engagement forms a central pillar of our methodology, as we actively solicit feedback from researchers and practitioners throughout the development process. Through user evaluations and feedback sessions, we gather insights into platform usability, effectiveness, and areas for improvement. This iterative feedback loop informs subsequent refinements and enhancements, ensuring that the platform aligns closely with the evolving needs and expectations of its user base.

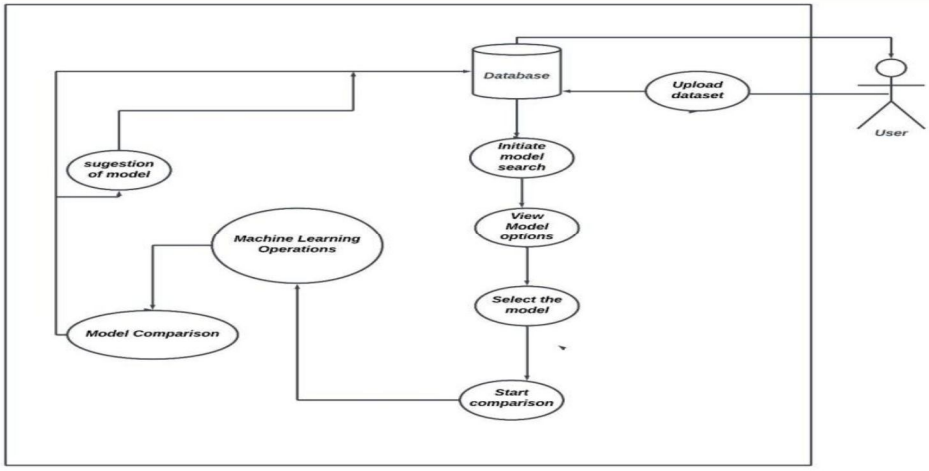


Figure 1: Use Case Diagram

IV. RESULTS

The implementation of the AutoML platform states promising outcomes, empowering users with robust tools for comparative analysis of ML models. Through the manual system, researchers and practitioners gained hands-on control over experimentation, allowing for tailored selection of target variables, data preprocessing steps, and model choices. This approach facilitated a deeper understanding of model performance and underlying mechanisms. In parallel, the automated system streamlined the model development process, accelerating experimentation and ensuring consistency and reproducibility across different experimental setups. Detailed integration and testing confirmed the platform's functionality, stability, and reliability, with benchmark datasets and real-world use cases validating its performance comprehensively. User feedback played a pivotal role in refining platform usability and effectiveness, resulting in iterative improvements. Ultimately, comprehensive documentation and deployment made the platform accessible to the ML community, fostering collaboration and innovation in the field.

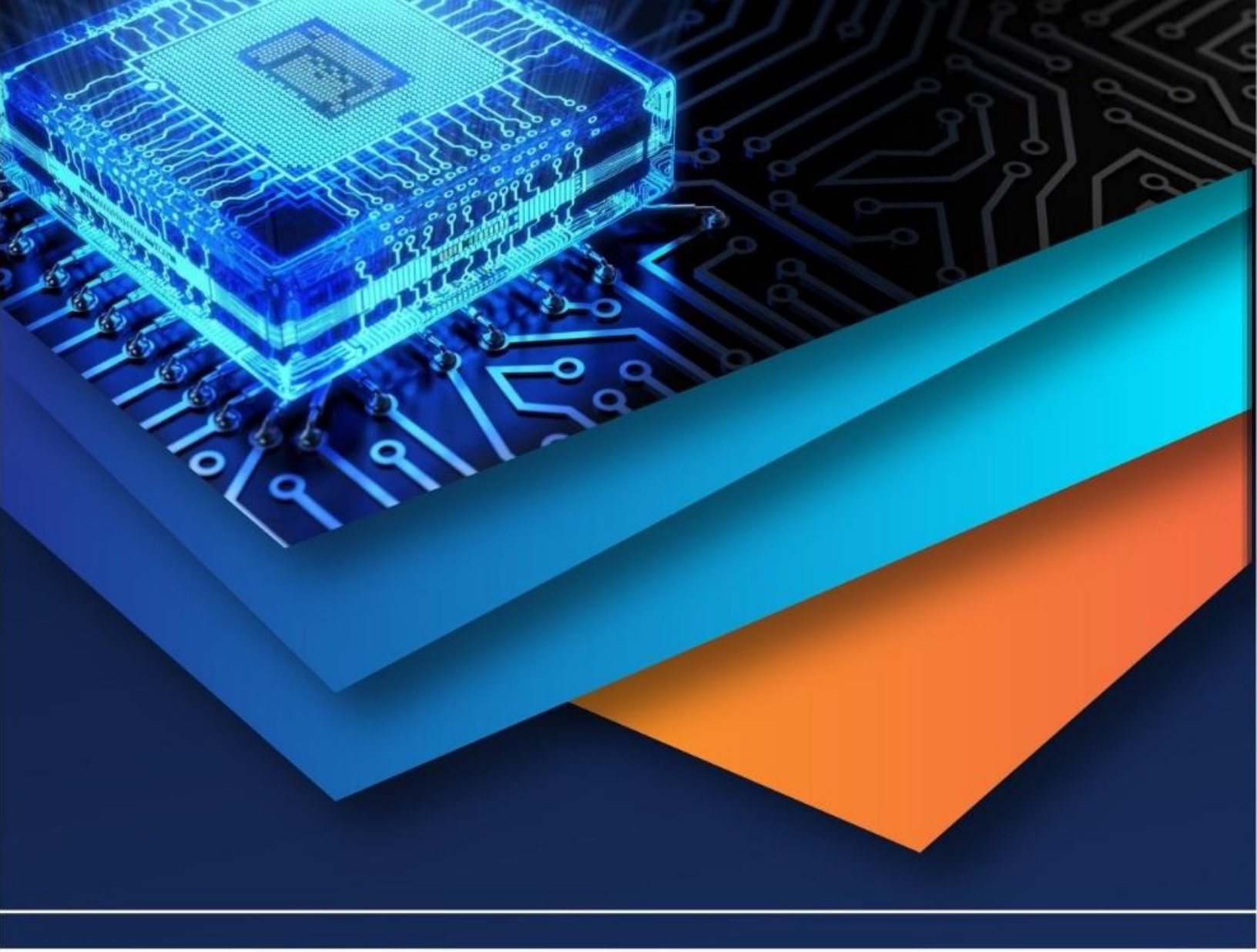
V. CONCLUSION

In conclusion, the development and implementation of the AutoML platform for comparative analysis of ML models represent a significant advancement in the field of machine learning. The platform offers a versatile toolkit that caters to the diverse needs of researchers and practitioners, combining the flexibility of manual control with the efficiency of automation. Through rigorous testing and user feedback, we have ensured that the platform meets high standards of functionality, reliability, and usability. By democratizing access to sophisticated ML model evaluation techniques, the platform has the potential to drive innovation and accelerate progress in various domains. Moving forward, continued refinement and expansion of the platform will further enhance its capabilities and impact, ultimately advancing the state-of-the-art in machine learning research and applications.

REFERENCES

[1] "Automated Machine Learning: Methods, Systems, Challenges" by Hutter et al. (2019), in the Journal of Machine Learning Research.
[2] "PyCaret: An open-source low-code machine learning library in Python" by Moez Ali, Moez Ali, Abhinav Sharma, Abhinav Sharma, Himanshu Sharma in ArXiv, 2020
[3] "Auto-sklearn: Automated Machine Learning with scikit-learn" by Matthias Feurer, Katharina Eggensperger, Stefan Falkner, Marius Lindauer, Frank Hutter in NeurIPS, 2015
[4] "Hyperparameter tuning and AutoML for vision tasks with scikit-learn and PyCaret" by Fábio Silva, Felipe Soares, Victor Hugo Silva, Rodrigo Silva in arXiv, 2021
[5] "AutoML: A Survey of the State-of-the-Art" by He et al. (2019), in the Knowledge and Information Systems journal.
[6] "Efficient and Robust Automated Machine Learning" by Feurer et al. (2015), presented at the International Conference on Machine Learning (ICML).
[7] "Practical Automated Machine Learning for the Analyst" by Olson et al. (2016), presented at the ACM SIGKDD Conference on Knowledge Discovery and Data Mining.
[8] "Automated Machine Learning: Methods, Systems, Challenges" edited by Hutter et al. (2020), published by Springer.
[9] "Python Machine Learning" by Raschka and Mirjalili (2019), published by Packt.
[10] "Auto-sklearn: Automated Machine Learning with Scikit-learn" by Feurer et al. (2015), a technical report from the University of Freiburg.
[11] "ML-Plan: Automated Machine Learning via Hierarchical Planning" by Mohr et al. (2018), a technical report from the University of Waikato.
[12] TensorFlow AutoML documentation: Official documentation from TensorFlow, providing insights into their AutoML capabilities.
[13] H2O AutoML documentation: Official documentation from H2O.ai, offering information about their AutoML platform.
[14] “On Hyperparameter Optimization of Machine Learning Algorithms: Theory and Practice" by Bergstra and Bengio (2012), in the Journal of Machine Learning Research.
[15] "Automated Machine Learning in Action" by Emmanuel Ameisen (2020), published by Manning Publications

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)