INTRODUCTION

In the rapidly evolving field of machine learning, developing and optimizing predictive models often requires extensive knowledge of algorithms, parameter tuning, and feature engineering. This process can be time-consuming and complex, especially for users who are new to machine learning or working with large and high-dimensional datasets. To address these challenges, automl (automated machine learning) frameworks have emerged as a powerful solution to automate key stages of the ml pipeline, including data preprocessing, model selection, hyperparameter optimization, and ensemble building.

One of the most advanced and user-friendly automl tools available today is autogluon, an open-source library developed by amazon web services (aws). Autogluon simplifies the model training process while delivering competitive performance across a wide range of supervised learning tasks such as classification, regression, and object detection. By automatically training multiple models and stacking them into an optimized ensemble, autogluon can achieve high accuracy with minimal manual intervention.

ABSTRACT

The increasing demand for efficient and accurate machine learning (ML) models across various domains has led to the development of Automated Machine Learning (automl) tools that simplify and accelerate the model-building process. This project explores the use of autogluon, an open-source automl framework developed by Amazon Web Services (AWS), to automate key aspects of the ML pipeline. Autogluon provides an intuitive interface that enables users to perform data preprocessing, feature selection, model training, hyperparameter tuning, and ensemble optimization with minimal coding effort.

In this study, the autogluon framework is applied to a real-world dataset to evaluate its performance in generating high-quality predictive models. The project workflow involves importing and preparing the dataset, automatically training multiple models, and selecting the best-performing ensemble based on accuracy and efficiency metrics. The experiment demonstrates autogluon's ability to handle complex datasets and deliver competitive results compared to traditional manual ML approaches.

CONCLUSION

The implementation of autogluon in this project effectively demonstrates the power and practicality of Automated Machine Learning (automl) in modern data science workflows. By automating crucial stages such as data preprocessing, model selection, hyperparameter tuning, and ensemble construction, autogluon simplifies the process of building accurate and efficient machine learning models. The results obtained from this study show that autogluon can deliver highly competitive performance with minimal manual intervention, making it an excellent tool for both beginners and experienced data scientists.

Through this project, it becomes evident that autogluon not only saves time and computational resources but also enhances model reliability and reproducibility. Its built-in ensemble learning strategy ensures that the final model is robust and generalizes well to unseen data. Moreover, the framework's flexibility to handle diverse datasets and problem types—ranging from classification to regression—underscores its versatility in practical applications.