A Hybrid Hierarchical Machine Learning Framework for Optimized Predictive Analytics in Bio-Ethanol Production

B. Siva Naga Kishore
Student, M.Tech, Department of Computer Science, ANITS(Autonomous), Sangivalasa, Bheemunipatnam, Vishakhapatnam, Andhra Pradesh, India

Ch.K.Rupesh Kumar
Assistant Professor, Department of Computer Science, ANITS(Autonomous), Sangivalasa, Bheemunipatnam, Vishakhapatnam, Andhra Pradesh, India

Abstract:

Machine learning has great potential in bio-production industries, and in the bio-ethanol industry in particular, where traditional models have under-represented these challenges due to their inability to deal with instability at process-level and multi-scale systems. To address these challenges, this study presents a Hybrid Hierarchical Machine Learning Framework, which combines a global root knowledge model with factory-specific knowledge branch models. This two-tiered architecture not only guarantees generalizability across diverse production sites but also allows for adaptability to localized variations, improving everything from accuracy to scalability. The root-level model learns high-level patterns with an ensemble of models, while the branch-level models consist of integrated learning components to learn complex interactions between features specific to each factory. We validate the framework on a real-world dataset of 86,754 records and achieve superior predictive performance in various metrics such as R^2, MSE, MAE, and SMAPE. Results show that algorithms such as XGBoost and Gradient Boosting are better than classic algorithms globally as well as locally. Additionally, the system allows for real-time adaptability and extensibility, which can handle dynamic inputs like environmental or economic factors. Thus, the framework can serve as an advanced decision-support tool for stakeholders keen on maximizing bio-ethanol yield, minimizing resource wastage, and ensuring production aligns with sustainability objectives. This work is a first step in several future directions, including deployment in real time with IoT integration, policy-based modeling of scenarios, and applications to other sectors of renewable energy. Thus, it is established in this study as a versatile foundation for data-driven innovation in the production of green energy.

Keywords:

Hybrid Hierarchical Model, Bio-Ethanol Production, Real Time Prediction, Multi Objective Optimization, Ensemble Learning, Root-Branch Framework.