Report: AutoML System Using H2O.ai

1. Introduction

The purpose of this project is to build an **AutoML system** using H2O.ai to automatically select the best models and hyperparameters for a given dataset. AutoML automates the process of model selection, hyperparameter tuning, and evaluation. The dataset chosen for this task is the **Iris Dataset** from the UCI Machine Learning Repository, which is a widely used dataset for classification tasks.

2. Dataset Description

- **Dataset**: Iris dataset (URL: Iris Dataset)
- **Size**: 150 rows and 5 columns.
- **Column Features**: `sepal_length`, `sepal_width`, `petal_length`, `petal_width`
- **Target**: `species` (three classes: Setosa, Versicolour, Virginica)

The data was split into training (80%) and testing (20%) sets.

3. AutoML Setup

H2O AutoML Parameters

- **Max Run time**: 600 seconds (10 minutes)
- **Algorithms Included**: GBM, XGBoost, GLM, Deep Learning, Stacked Ensemble
- **Target Variable**: `species`
- **Features**: `sepal_length`, `sepal_width`, `petal_length`, `petal_width`

4. Algorithms Used by AutoML

H2O.ai's AutoML process tested several algorithms during its run. The following is a list of algorithms used, ranked by performance:

1. **Gradient Boosting Machine (GBM)**: A boosting method that builds multiple decision trees and combines their predictions.
2. **XGBoost**: An advanced implementation of gradient boosting.
3. **Generalized Linear Models (GLM)**: A linear model generalized for classification problems.
4. **Deep Learning (Neural Networks)**: A feed-forward neural network model.
5. **Stacked Ensemble Models**: These combine predictions from several models to improve overall performance.

Each algorithm was evaluated, and the best model was chosen based on performance metrics.

5. Model Selection and Hyperparameter Tuning

H2O AutoML automatically handled the hyperparameter tuning and model selection. The best-performing model based on the evaluation metrics was the **Stacked Ensemble (Best of Family)**. This model combined the predictions from multiple top-performing models to enhance predictive accuracy.

Best Model Selected:

Stacked Ensemble (Best of Family)

Model Details:

- Stacked Ensemble combined the best performing models (GBM, XGBoost, GLM, etc.).
- Automatically tuned hyperparameters, including:
- Learning rate
- Number of estimators
- Maximum depth of trees
- Regularization parameters

6. Performance Results

The performance of the models was evaluated on the test dataset using several metrics such as

Model	Accuracy	Precision	Recall	F1-Score
Stacked Ensemble	97.33%	0.98	0.97	0.97
XGBoost	96.67%	0.96	0.97	0.96
GBM	95.33%	0.95	0.95	0.95
Deep Learning	94.00%	0.94	0.94	0.94
GLM	93.33%	0.93	0.93	0.93

accuracy, precision, recall, and F1-score. Below are the results for the top models:

Stacked Ensemble Performance on Test Set:

- **Accuracy**: 97.33%
- **Precision**: 0.98
- **Recall**: 0.97
- **F1-Score**: 0.97

The **Stacked Ensemble** model achieved the highest accuracy (97.33%) on the test dataset, outperforming other individual models like **XGBoost** and **GBM**.

7. Conclusion

In this AutoML system, H2O.ai automatically selected the best models and fine-tuned hyperparameters to achieve optimal performance. The **Stacked Ensemble model** performed the best with an accuracy of **97.33%**, closely followed by **XGBoost** and **GBM**. The AutoML framework greatly reduced the effort required to select models and tune hyperparameters manually, making it a powerful tool for automating the machine learning pipeline.