MLOps HW3 Summary
Luke Schwenke & Aaron Chan
November 7, 2023

Best Model

- XGBoost

Data Insights from AutoML run

- Each run shows:
    i. MSE, RMSE, R2, MAE
    ii. Start Time and Status
    iii. Hyperparameters
    iv. Feature importance (SHAP values)
    v. Artifacts

What are the reported top 5 features?

- Weight
- Age
- Height
- Schedule
- Experience

What are the top 3 models per validation score when using

- all features
    1. XGBoost 17,261
    2. XGBoost 17,591
    3. XGBoost 17,625
- only the top features (if you have to choose a number put in 3)
    1. MLP 21,347
    2. MLP 21,411
    3. MLP 21,421

What are the top 3 models per speed when using

Calculation: Modified On - Created On (speed of training)

Fastest with all variables:

- Experiment-1699239190542VebawbOG-015-bd157986-aws-training-job
    i. MSE = 19,141
- Experiment-1699239190542VebawbOG-092-3b5ac107-aws-training-job
    i. MSE = 19,349
- Experiment-1699239190542VebawbOG-093-34f08be3-aws-training-job

    i. MSE = 17,811

Fastest using top 3 variables:

- Experiment-1699244384780brcLmG23-018-d3c9e40a-aws-training-job
    i. MSE = 22,063
- Experiment-1699244384780brcLmG23-086-f92cf166-aws-training-job
    i. MSE = 21,739
- Experiment-1699244384780brcLmG23-059-be51fd13-aws-training-job
    i. MSE = 25,351

How do the top models compare to your previously developed model (assignments 1 and 2) in terms of validation score and speed?

- Homework #1: MSE = 42,058 | 3-Layer Neural Network
- Homework #2 (Neptune.ai): MSE = 40,231 | 4-Layer Neural Network
- Homework #3 (AutoML): MSE = 17,261 | XGBoost
    i. Speed was slower for HW3 as AWS SageMaker iterated through many model combinations and hyperparameters

Is your platform AutoML no-code/low-code/full-code and why?

The platform No-code but AWS SageMaker has the option to be all 3. You can visually construct the pipeline, have AWS generate the code and adjust, or you can fully write the code yourself.