

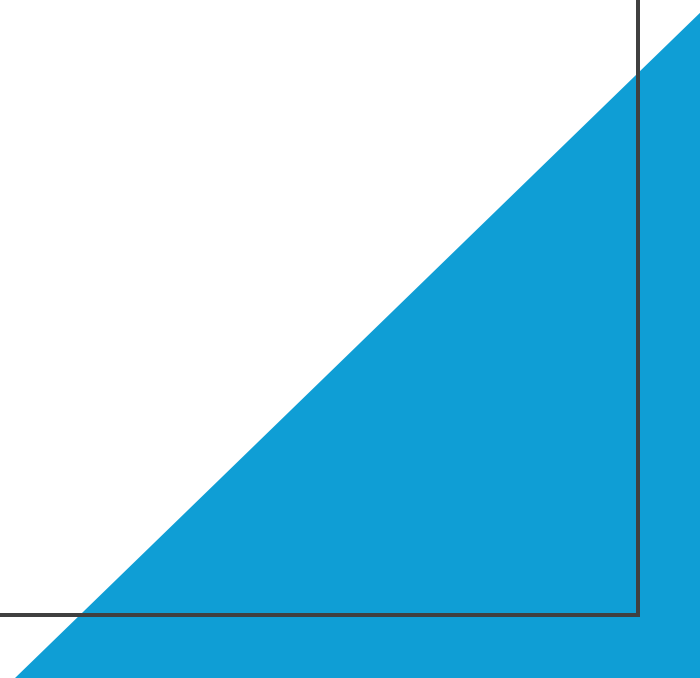
Project 1

Advanced Machine Learning

Michał Ciasnocha

Jerzy Kraszewski

Filip Kołodziejczyk



Datasets

Small datasets (at most 10 features)

- NHANES Age Prediction (senior/non-senior)
- Date Fruit Classification (ajwa/medjool)
- Fertility (normal/altered)

Large datasets (more than 10 features)

- Mice Down Syndrome (control/ts65Dn)
- Mushroom (edible/poisonous)
- Room Occupancy Estimation (≤ 1 / > 1)
- Taiwanese Bankruptcy Prediction (non-bankrupt/bankrupt)
- Estimation of Obesity Levels (non-obese/obese)
- Wine Color (red/white)

Choice criteria:

- goal: binary classification, numerical features (or transformable)
- nice to have: original and interesting

Preprocessing

- Transforming target into binary
- Removing co-linear features (where $r > \text{threshold}$)
- One-hot encoding categorical nominal features
- Handling missing values – by dropping data or simple imputation

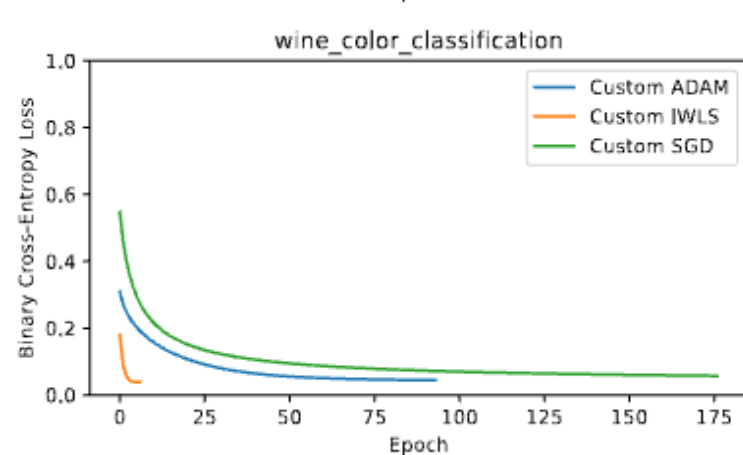
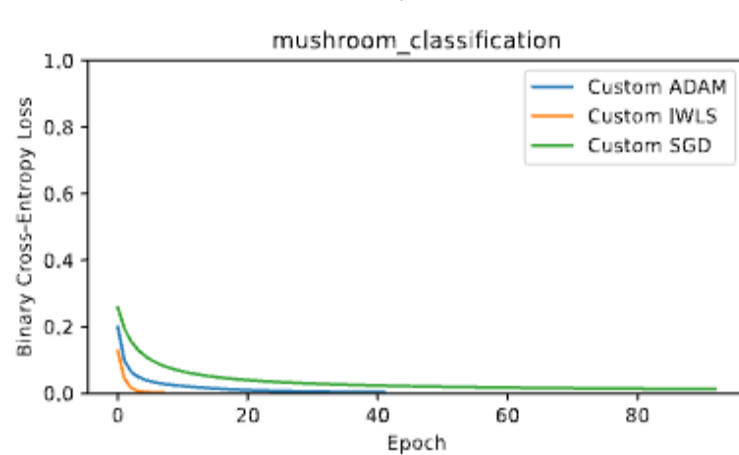
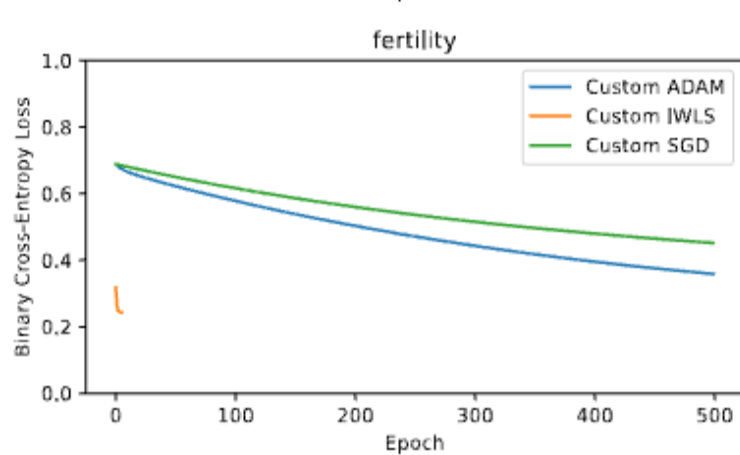
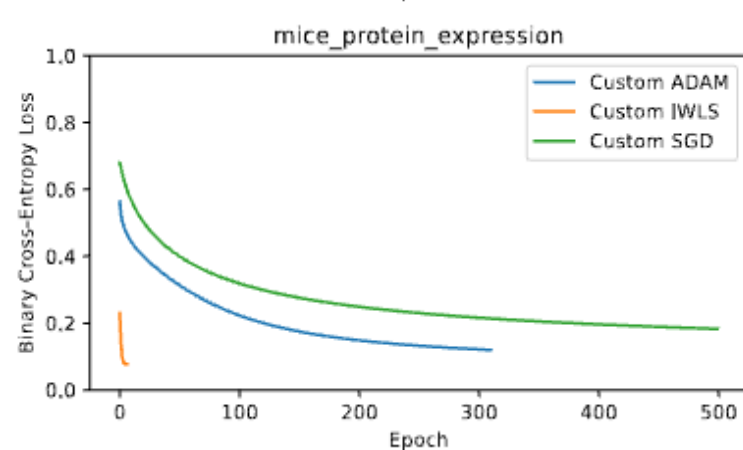
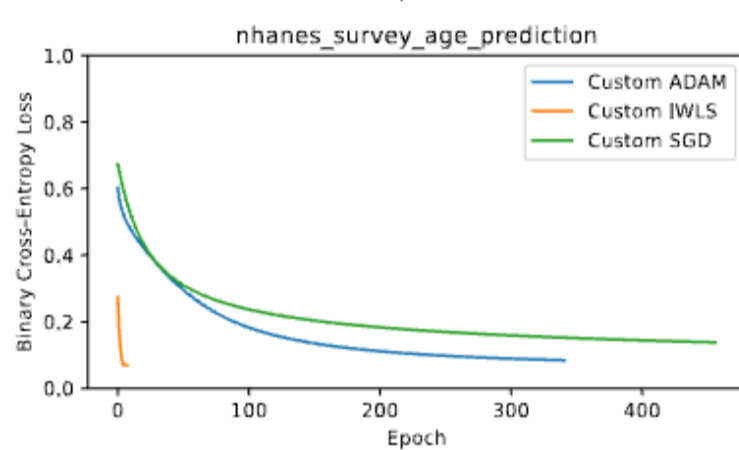
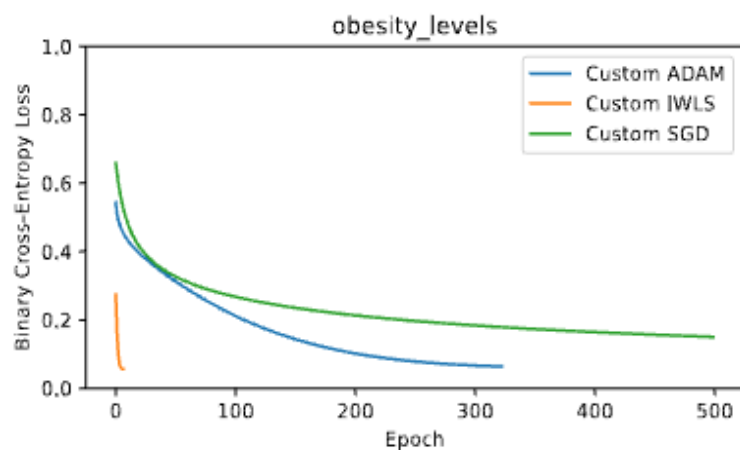
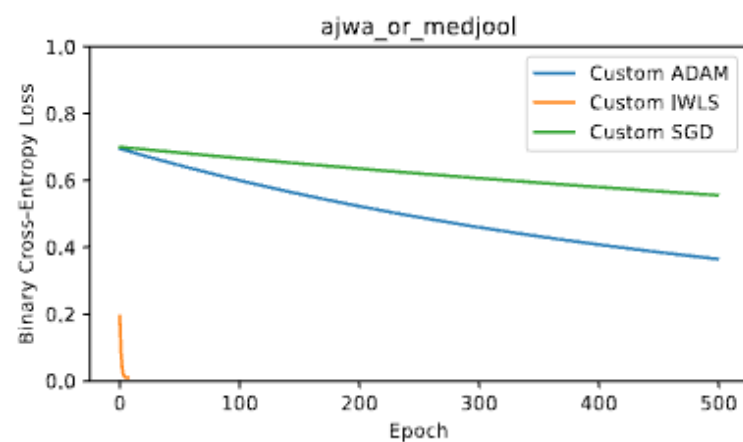
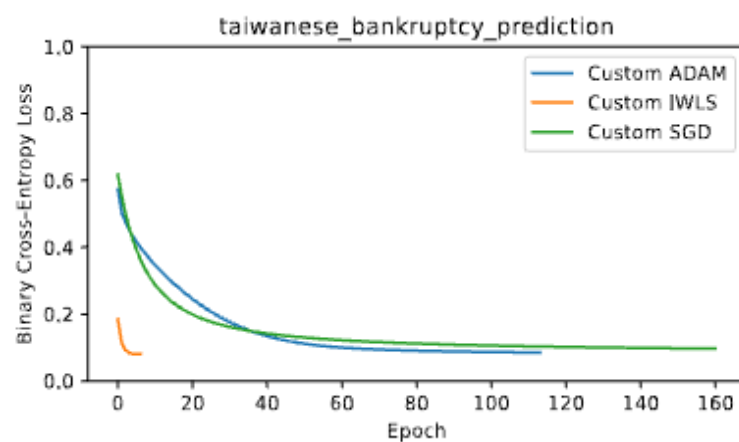
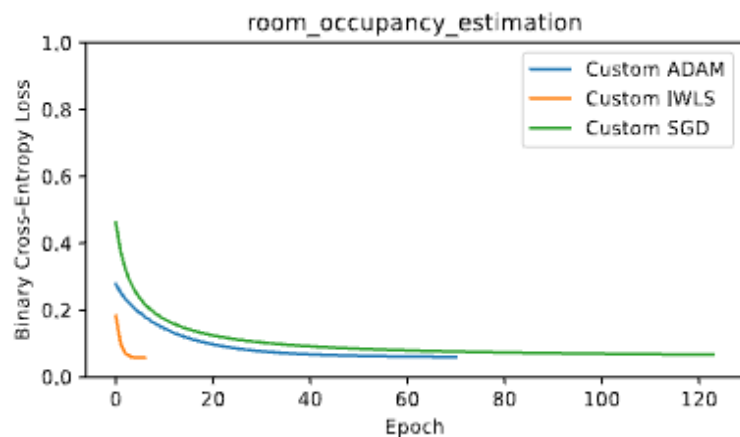
Binary Cross-Entropy

- Measures the difference between two probability distributions: the actual distribution (true labels) and the predicted distribution (predicted probabilities by the model).
- Chosen over the log-likelihood function for convergence and stop property as it is directly related to the classification objective and has normalized values (range from 0 to 1).

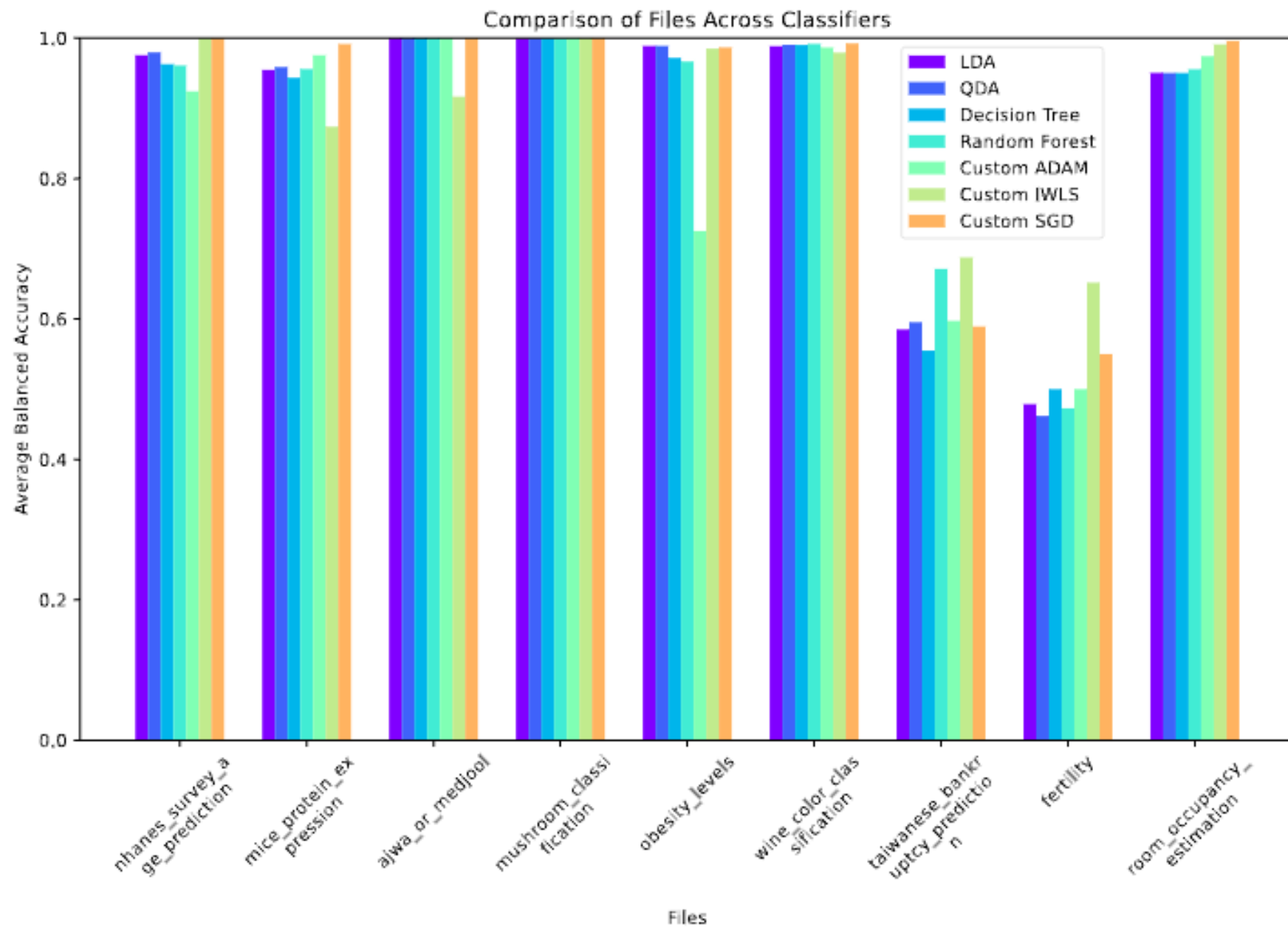
Methodology

- Stop property: BCE with tolerance set at 0.0001 (stops when the change is smaller than that)
- Performance measure: Balanced Accuracy to consider imbalanced classes (predefined function used)
- Custom linear regression to be compliant with optimizers, using sigmoid as activation function
- Optimizers based on base class Optimizer as a modular approach

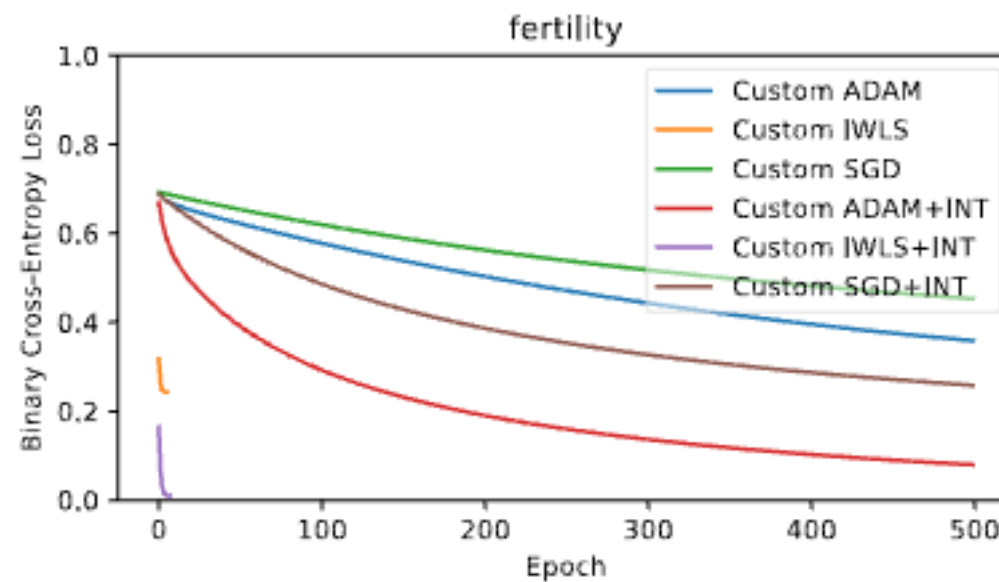
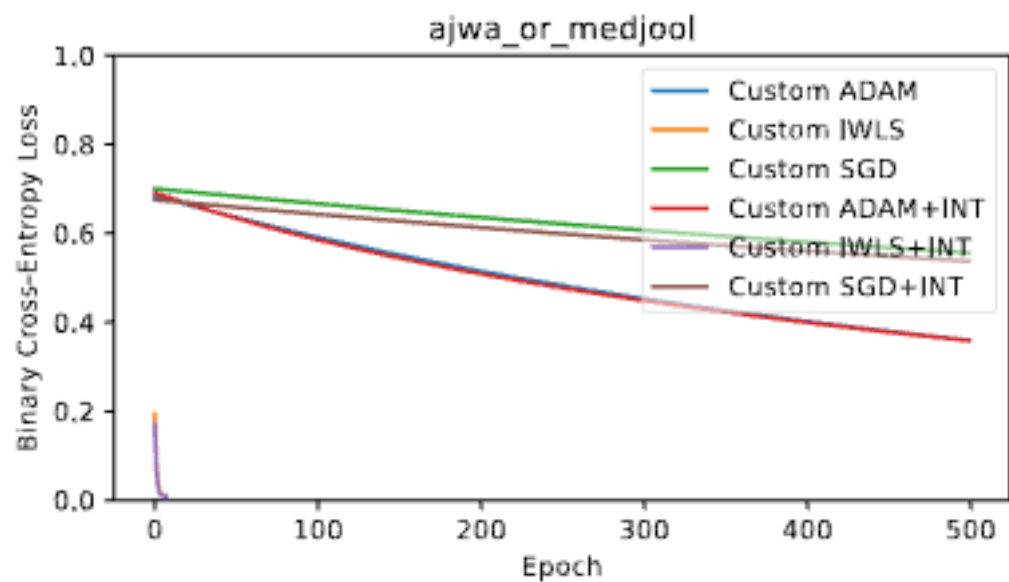
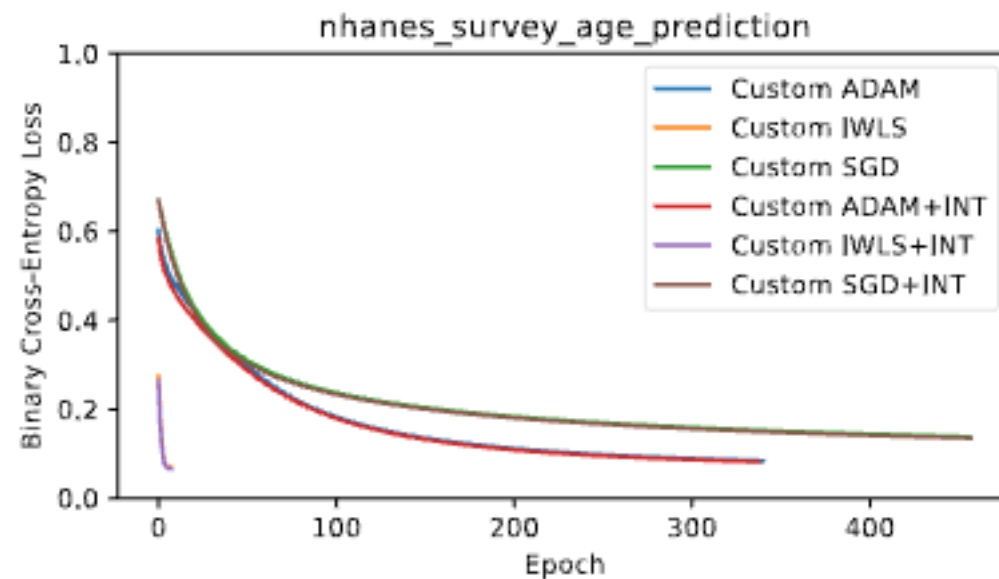
Convergence



Accuracy



Interaction (convergence)



Interaction (accuracy)

