Babeș-Bolyai University

EXPERIMENTAL RESULTS EXPERIENCE LEVEL CLASSIFICATION

Author: Turcu Ciprian-Stelian

INTRODUCTION

Purpose of the Study:

- Build a custom from scratch model to asses performance on the selected dataset
- Evaluate and compare the custom built model against other more complex models directly accessible through libraries.

Models Used:

- From scratch: Custom ordinal regression model (cmodelv3).
- Library-based: builtinModel and builtinv2.

Key Questions:

- How does our from scratch model and library-based models compare in performance?
- What insights do the results provide about feature importance and improvements?

SCRATCH MODEL (CMODELV3)

Framework:
Built on Ordinal Regression

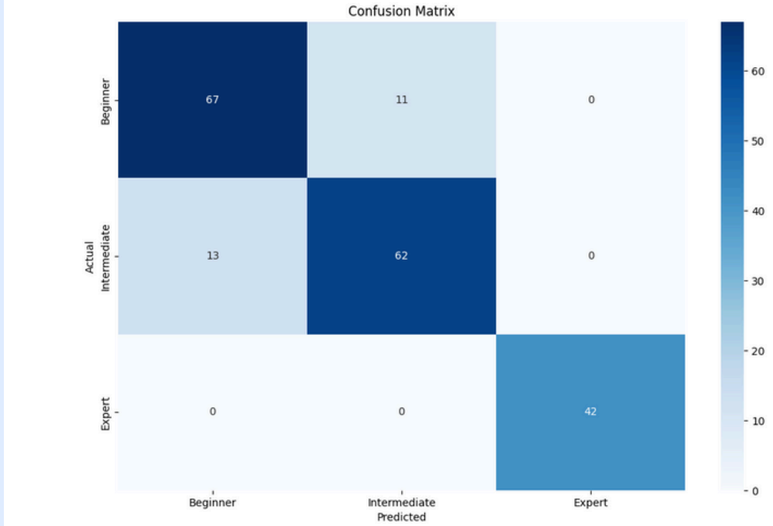
Key Features:
- Cumulative thresholds (θ).
- Negative Log-Likelihood optimization.
- Gradient-based L-BFGS-B algorithm.

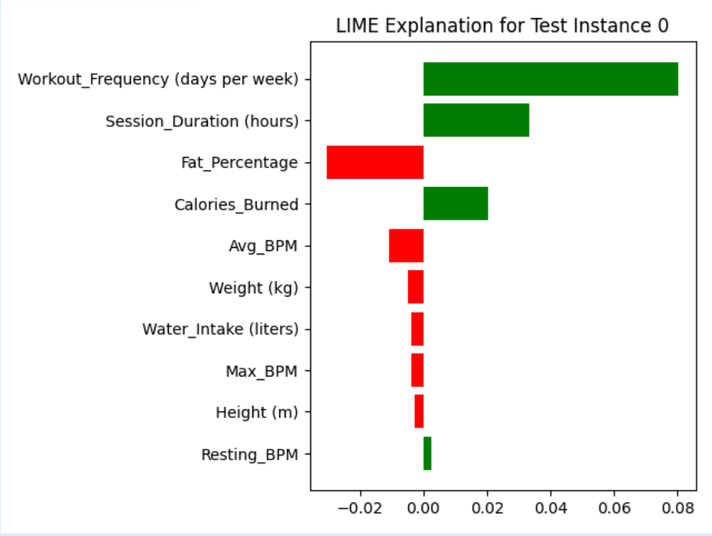
Hyperparameter Tuning:
- Grid Search (Bounds: lb, ub, Standardization).

ANALYSIS RESULTS

Scratch Model (cmodelv3)

- Cross-Validation Accuracy: 0.8725 (CI: ±0.04622).
- Test Accuracy: 0.8769.
- **Metrics**:
  - Precision: 0.8769, Recall: 0.8769, F1-Score: 0.8769.
  - AUC: 0.9721, AUPRC: 0.9534.
- **Confusion Matrix**:
  - High accuracy for all classes, with minor misclassifications between Beginner and Intermediate levels.
- **LIME key features**:
  - **Workout Frequency (days per week)**: The most important feature across all classes.
  - **Fat Percentage**: A critical determinant for class separations.





BUILTINMODEL (LIBRARY BASED)

Framework:

Built on Ordered Logistic Regression (Statsmodels - OrderedModel).

Key Features:

- Log-Likelihood with multiple solver options: bfgs, newton, lbfgs.

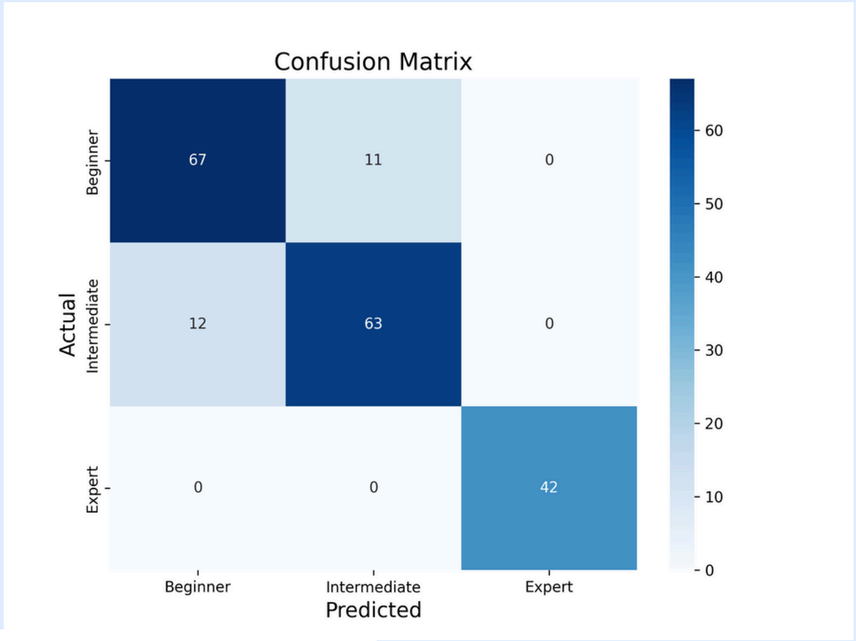
Challanges:

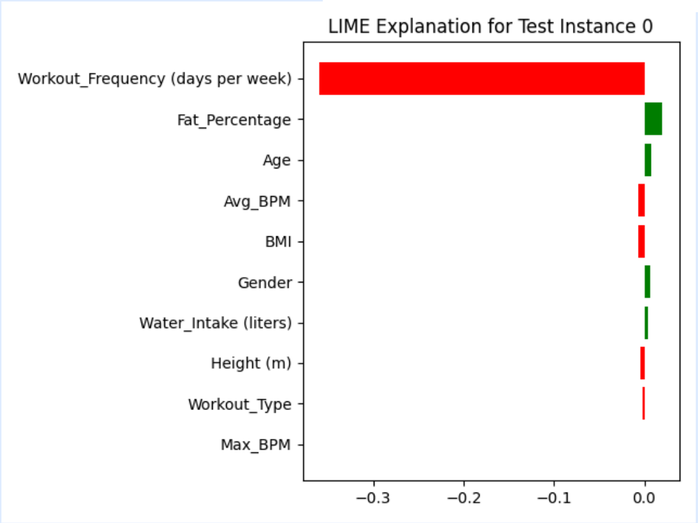
- Numerical issues with the newton solver.

ANALYSIS RESULTS

builtinModel (Library based)

- Cross-Validation Accuracy: 0.8643 (CI: ±0.04527).
- Test Accuracy: 0.8821.
- **Metrics**:
  - Precision: 0.8821, Recall: 0.8821, F1-Score: 0.8820.
- **Confusion Matrix**:
  - High accuracy across all classes, with minor misclassifications between Beginner and Intermediate levels.
- **LIME key features**:
  - **Session Duration (hours):** A dominant factor influencing predictions.
  - **Height (m):** Strongly associated with distinctions between categories.





BUILTINV2 (LIBRARY BASED)

Framework:

Multiclass Logistic Regression (Scikit-learn).

Key Features:

- Regularization Strength (C): Best = 0.01.
- Solver: lbfgs.

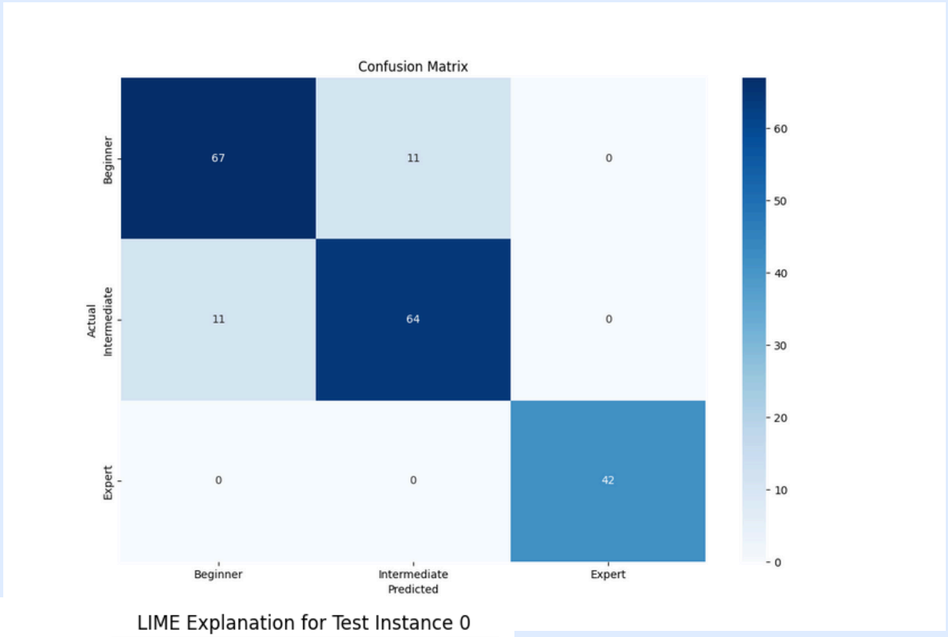
Hyperparameter Tuning:

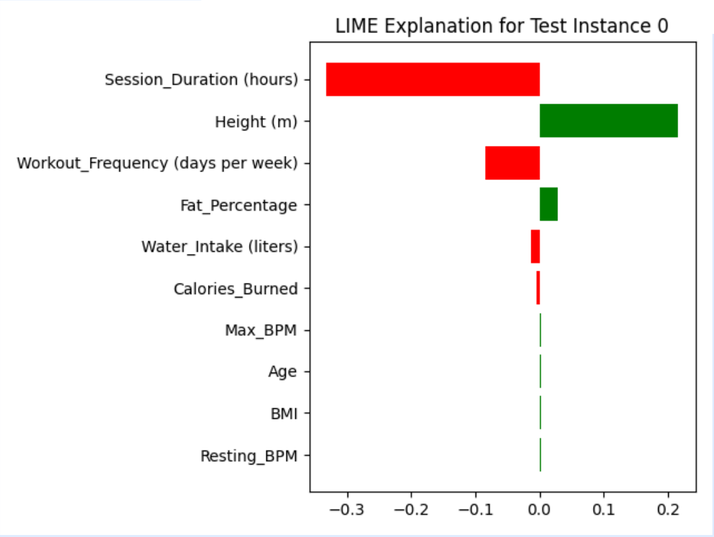
- Grid Search - combinations of C and Solver

ANALYSIS RESULTS

builtinv2 (Library based)

- Cross-Validation Accuracy: 0.8730 (CI: ±0.04440).
- Test Accuracy: 0.8872.
- **Metrics**:
  - Precision, Recall, F1-Score: 0.8872.
  - AUC: 0.9752, AUPRC: 0.9587.
- **Confusion Matrix**:
  - Highest accuracy among all models, with minimal misclassifications.
- **LIME key features**:
  - **Workout Frequency (days per week):** A critical feature for predictions.
  - **Session Duration (hours):** Strongly associated with intermediate and expert classifications.





COMPARATIVE ANALYSIS

Model	CV Acccuracy	Final Test Accuracy	Precision	Recall	F1	AUC	AUPRC
cmodelv3	0.8725	0.8769	0.8769	0.8769	0.8769	0.9721	0.9534
builtinModel	0.8643	0.8821	0.8821	0.8821	0.8820	N/A	N/A
builtinv2	0.8730	0.8872	0.8872	0.8872	0.8872	0.9752	0.9587

CONCLUSIONS & FUTURE WORK

Key Takeaways:

- Best Accuracy: Builtinv2 (0.8872).
- Best Explainability: cmodelv3 (LIME results).

Future Directions:

- Reducing misclassifications.
- Enhancing threshold optimization for scratch models.
- Exploring hybrid models combining custom and library-based approaches.

Babeș-Bolyai University

THANK YOU

Turcu Ciprian-Stelian | 2025