Building Machine Learning Weight Prediction Models Using a Binarized Weight Loss and Time Horizon Grid and Optuna Optimiation to Predict Weight Loss at Different Time Intervals

Mansur Shomali, MD, CM[1], Rui Han[1], Abhimanyu Kumbara, MS, MBA[1], Anand Iyer, PhD, MBA[1]
[1]Welldoc, Columbia, Maryland,USA

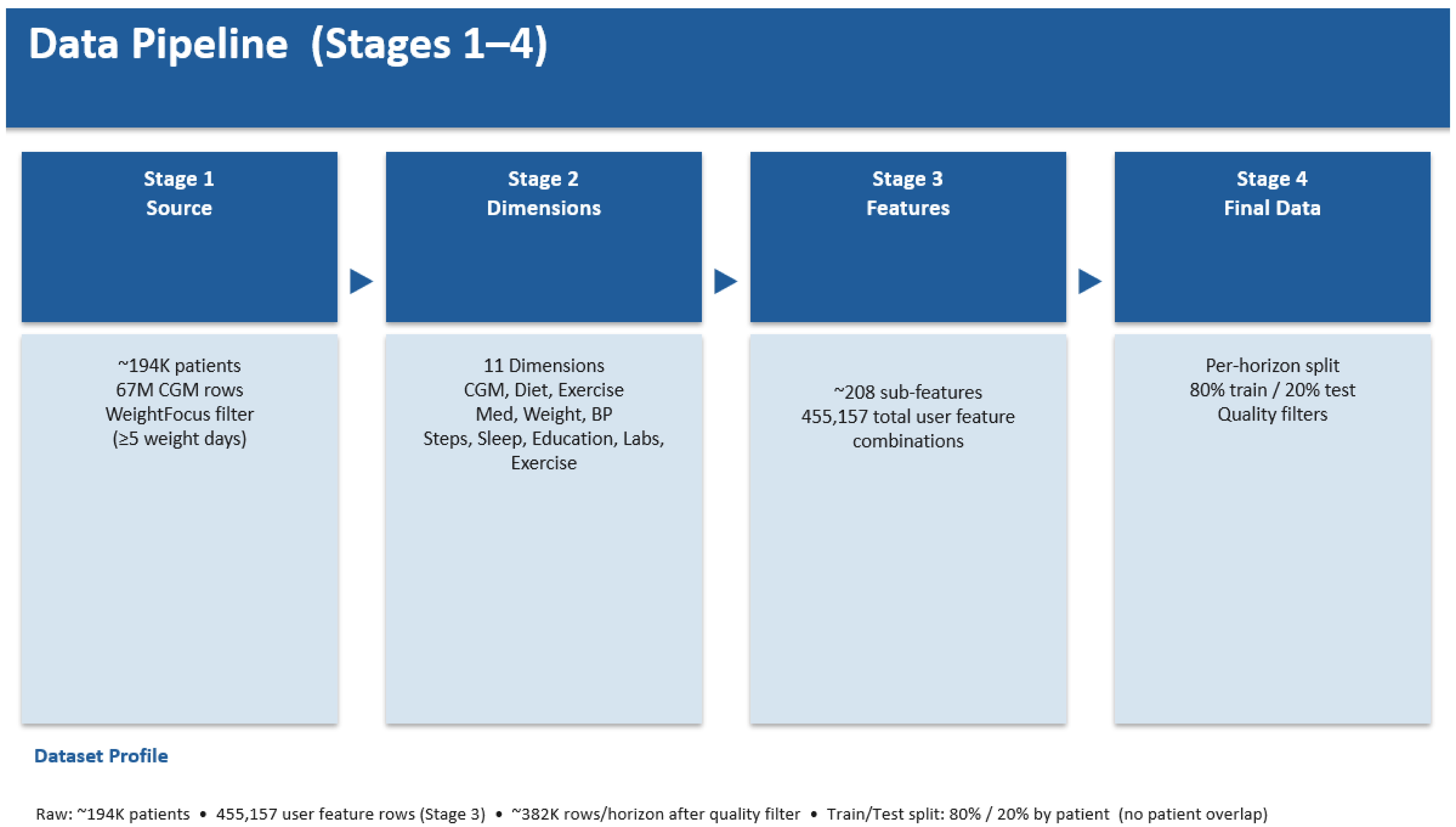
BACKGROUND AND AIMS

Accurately predicting patient weight trajectories is critical for delivering effective health management and targeted interventions.[1-6] This study aimed to predict binary weight-loss outcomes for patients using a comprehensive dataset – including glucose values from continuous glucose monitors (CGM), diet, exercise, medication, blood pressure, and steps – from digital health platform. We evaluated predictions at six forecast horizons, which ranged from 15 days to 90 days from an observation window, in 15-day increments. These horizons were tested against six binary weight-loss thresholds (greater than 2%, 4%, 6%, 8%, 10%, and 12%), resulting in a total of 36 distinct binary labels.

We also aimed to use Optuna, an open-source, python-based hyperparameter optimization framework designed to automate and accelerate the tuning of machine learning models. It utilizes efficient Bayesian optimization (TPE algorithms) to find optimal parameters faster than grid or random search, featuring dynamic search spaces, easy parallelization, and automatic pruning of unpromising trials with data sets.

Additionally, we evaluated the accuracy of our predictive models with varying levels of Optuna optimization for an XGBoost model.

Figure 1:



MATERIALS AND METHODS

We analyzed raw data from approximately 194,000 patients, applying a "WeightFocus" filter to isolate active weight-trackers who had at least five distinct weight-measurement days. This filtering resulted in a final cohort of 80,392 patients and 455,157 case rows. For feature engineering, we extracted 39 feature categories yielding roughly 208 sub-features per case, incorporating historical weight lookbacks, diet, exercise, medication, steps, sleep, demographics, and continuous glucose monitoring (CGM) statistics. Because the vast majority of the cohort (92.8%) had no glucose data and only 0.7% utilized CGM devices, we employed an XGBoost algorithm capable of handling missing values natively. We utilized an MLMultiLabelPredictor framework with Optuna for Bayesian hyperparameter optimization, running 5 trials per label. For each horizon, the dataset was split into 80% for training and 20% for testing, ensuring no patient overlap between the splits. Four stages of the data pipeline, as described above, is summarized in Figure 1 below.

RESULTS

Table 1 below showcases the results. For The models demonstrated strong discriminative power across all combinations, successfully ranking patients by their weight-loss risk. All 36 label-and-horizon combinations exceeded an Area Under the Receiver Operating Characteristic Curve (AUC) of 0.71, achieving an average AUC of 0.818. We found that shorter forecast horizons were significantly more predictable; the 15-day (0.5 month) horizon averaged an AUC of 0.855, compared to 0.788 for the 90-day (3 month) horizon. Furthermore, stricter weight-loss thresholds yielded clearer predictive signals, with the >12% loss threshold averaging an AUC of 0.876 versus 0.750 for the >2% threshold. F1 scores were moderate, ranging from 0.28 to 0.72. This was expected due to significant class imbalance, as large percentage weight losses are rare events (for example, the positive rate for >12% weight loss at 30 days was <0.5%)

Evaluation — AUC (AUROC)

AUC by Horizon × Loss Threshold

Horizon	loss_2%	loss_4%	loss_6%	loss_8%	loss_10%	loss_12%	Avg
WeightAf0p5M	0.712	0.839	0.879	0.903	0.898	0.899	0.855
WeightAf1M	0.738	0.797	0.854	0.872	0.886	0.892	0.840
WeightAf1p5M	0.758	0.735	0.834	0.853	0.863	0.889	0.822
WeightAf2M	0.754	0.786	0.790	0.837	0.869	0.873	0.818
WeightAf2p5M	0.759	0.777	0.766	0.768	0.802	0.859	0.788
WeightAf3M	0.777	0.754	0.758	0.765	0.824	0.850	0.788

Table 1:

CONCLUSIONS

The multi-label XGBoost model can effectively predict and rank patient weight-loss risk, with performance highest at shorter horizons and stricter weight-loss thresholds. While the current baseline model (v0001) successfully handles sparse glucose data, its F1 scores are limited by class imbalance and a constrained hyperparameter search budget of only 5 trials. Future model iterations (v0002) will address these limitations by increasing Optuna trials to 50 per label to improve AUC and F1 scores, alongside implementing in-patient and out-patient data splits and additional evaluation metrics designed to reduce false positive predictions.

REFERENCES

1. Shahabi, F., Battalio, S.L., Pfammatter, A.F. *et al.* A machine-learned model for predicting weight loss success using weight change features early in treatment. *npj Digit. Med.* **7**, 344 (2024). https://doi.org/10.1038/s41746-024-01299-y
2. Teke J, Msiska M, Abisoye O, et a. Artificial intelligence for obesity management: A review of applications, opportunities, and challenges, Obesity Medicine, 2025;58, 100657. https://doi.org/10.1016/j.obmed.2025.100657. (https://www.sciencedirect.com/science/article/pii/S2451847625000776)
3. Fallah-Fini, S., Adam, A., Cheskin, L.J., Bartsch, S.M. and Lee, B.Y. The Additional Costs and Health Effects of a Patient Having Overweight or Obesity: A Computational Model. Obesity 2017;25: 1809-1815. https://doi.org/10.1002/oby.21965
4. MacEwan JP, Chiu K, Ahmad NN, et al. Clinical, economic, and health-related quality of life outcomes in patients with overweight or obesity in the United States: 2016–2018. Obesity Science and Practice 2023 https://doi.org/10.1002/osp4.726
5. Thorpe K, Toles A, Shah B, et al. Weight Loss-Associated Decreases in Medical Care Expenditures for Commercially Insured Patients With Chronic Conditions. *Journal of Occupational and Environmental Medicine* 63(10):p 847-851, October 2021. | *DOI:* 10.1097/JOM.0000000000002296
6. Horstman CM, Ryan DH, Aronne LJ, et al. Return on Investment: Medical Savings of an Employer-Sponsored Digital Intensive Lifestyle Intervention for Weight Loss. Obesity 2021;20(4):654-661. https://doi.org/10.1002/oby.23117