Proposed Solution

* First, we ensure that only relevant features and data are used for diabetes management and employ **minMaxScaler** normalization to ensure uniform feature scaling.
* We address **class labeling imbalance** by dividing the dataset into two classes and balancing the sample dataset using **under-sampling and stratified sampling.**
* Next, we utilize appropriate encoding methods and aggregation strategies for diabetic data. The effectiveness of different gradient boosting models have been applied to a multitude of research problems.
* In this study we will be using the original **GBM** algorithm as a baseline and three newer variants: **XGBoost, LightGBM, and CatBoost**.
* Next, we leverage various classifiers, including **Support Vector Classifier, Random Forest Classifier, Gradient Boosting Classifier, XGBoost Classifier, and LightGBM Classifier**, to predict rehospitalization and ER visit outcomes; and introduce a **Sequential Model with Dense Layers and ReLU activation functions.**
* Finally, we Employ **Binary Cross-Entropy as the loss function, and the Adam optimizer** to fine-tune the model using clinical characteristics selected from demographic, condition, observation, measurement, and drug records provided in the N3C.

Results

Initial results show some progress on the Diabetes Kaggle dataset, as outlined in Table~\ref{tab:Kaggle}.

* + The results indicate that the GBDT model, implemented with LightGBM, exhibits higher reliability and accuracy in diabetes prediction compared to LR, highlighting the potential of machine learning for developing reliable prediction models in diabetes prevention \cite{GBT-Diabetes}.

Future Work (really current work tho)

* In the next step, we propose to utilize deep learning and GBDT models in tabular data analysis. Our proposed approach for predictive modeling in diabetes management has the potential to significantly improve patient care and reduce healthcare costs.
* By proactively identifying diabetic individuals at risk of rehospitalization and ER visits, we can enable early interventions and personalized care strategies, ultimately improving patient outcomes and quality of life.