Towards Better Insights: Preliminary Results Of A Machine Learning Model For Fracture Risk Assessment

Ed Sykes[1], Ravi Jain[2], Alex Canales[1], Wei Bin Wang[1], Mohit Deol[1], Jennifer Weldon[2], Rohit Shanker[3], Volodymyr Voytenko[1], Jansen Sullivan[3] , Doug Sauer[3]

[1]Centre for Mobile Innovation, Sheridan College, Oakville, [2]Ontario Osteoporosis Strategy, Osteoporosis Canada, Toronto, [3]Inovex, Oakville, Canada

INTRODUCTION

Background:

- Fracture risk assessments are essential to evaluate and prevent osteoporotic fractures. While it is desirable to collect all relevant data about patients regarding their susceptibility to fractures to make an accurate assessment, in real-world environments osteoporosis patient data can be incorrect, inconsistent, or missing.
- A crucial component to calculating the fracture risk is the Bone Mineral Density (BMD) T-score (used in fracture risk calculators: e.g., FRAX, CAROC). BMD tests are conducted at hospitals or diagnostic imaging clinics at the request of a physician. However, many people don't get tested.

Purpose of the study:

- Create a Machine Learning (ML) model that predicts the BMD T-score and assesses the fracture risk based on relevant patient data features.

METHODS

Sample:

- The Ontario Fracture Liaison Service (FLS) dataset enrolled 26,804 fragility fracture patients from 39 hospital sites from July 2017 to March 2021. Each patient record contains 205 features, from which 29 were selected for our ML research (e.g., age, gender, weight, fracture history, etc.). A small set of de-identified representative data were made available (1000 records).

Analysis:

- The sample data augmented by other Ontario FLS data reports, osteoporosis research articles, and advice from osteoporosis experts were used to create a Synthetic Data Generator and Cleaning Script that was used to process the 29 features found in the dataset.
- The following ML algorithms were explored to compare predicted T-score values to actual values: Linear Regression, Ridge Regression, Lasso Regression, Polynomial (degree 2), Random Forests, Decision Trees, Categorical Boosting, XG Boosting. 80% of the data was used for training and 20% for testing.
- Evaluation metrics:
  - Root Mean Squared Error (RMSE)
  - Mean Absolute Error (MAE)
  - R-Squared Score
  - Mean Squared Error (MSE)

DATA PROCESSING

Data Quality:

- The cleaning process consisted of removing duplicates, eliminating irrelevant data, performing type conversion fixing syntax errors, filling in missing values through mean imputation, and dealing with outliers. This left us with a useable dataset of 772 patient records.

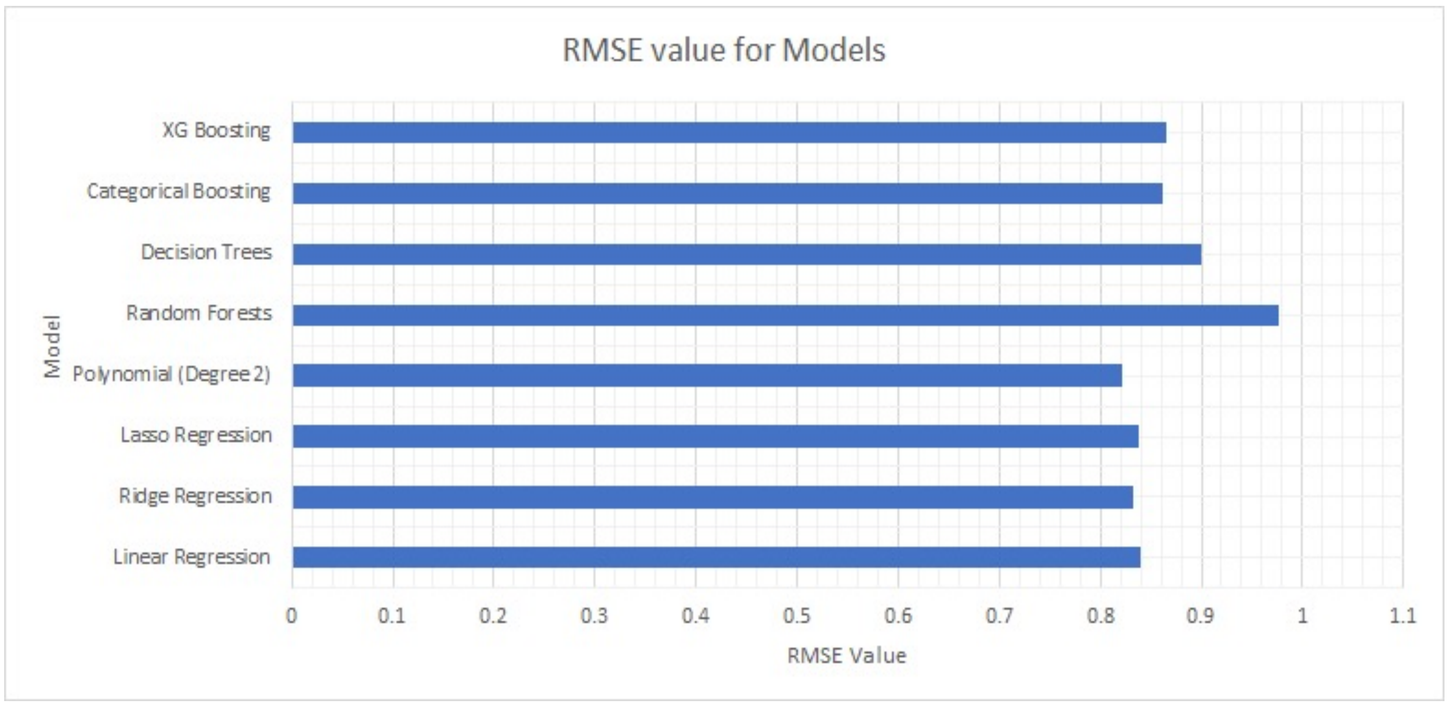
Selecting Features:

- Although 29 features were selected for development, we primarily focused on data features used for CAROC and FRAX calculation. CAROC uses 5 features while FRAX uses 12 features. Although both risk assessment tools incorporate previous fracture as one of the risk factors, they can be used for patients without fracture history.
- Since the Ontario FLS data includes only fragility fracture patients, we chose an additional 17 data features from this unique dataset.

RESULTS

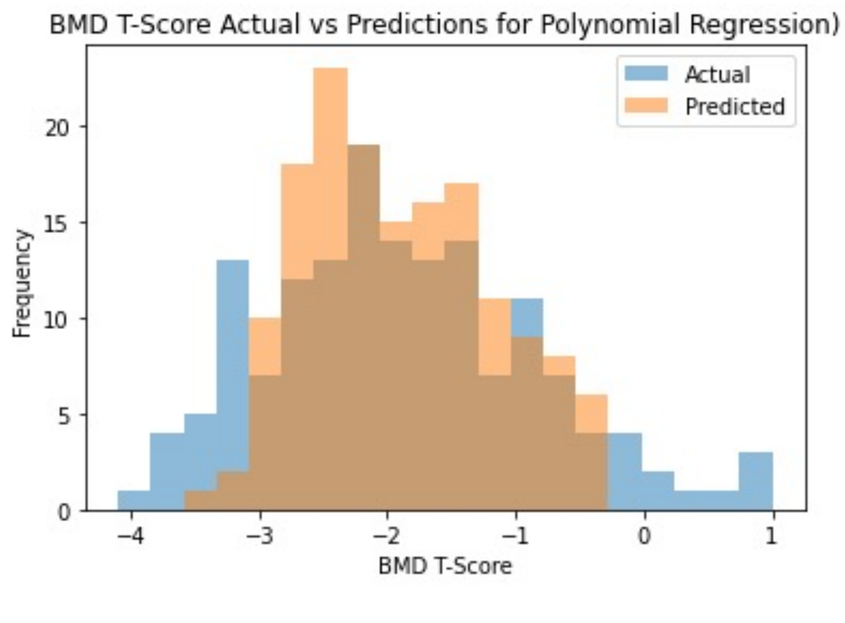
Model	RMSE (SI)	MAE	R-Squared	MSE
Linear Regression	0.839 (0.441)	0.655	0.348	0.704
Ridge Regression	0.833 (0.437)	0.645	0.358	0.693
Lasso Regression	0.8389 (0.440)	0.661	0.348	0.704
Polynomial (Degree 2)	0.822 (0.431)	0.649	0.375	0.675
Random Forests	0.976 (0.512)	0.777	0.118	0.953
Decision Trees	0.900 (0.472)	0.712	0.250	0.810
Categorical Boosting	0.863 (0.433)	0.653	0.368	0.683
XG Boosting	0.865 (0.453)	0.680	0.377	0.748

*Values are rounded up to the nearest thousands. A Standard Scaler was used on 9 features that were found using the Feature Importance methods in Scikit.



PREDICTIONS

- Our best model was the Polynomial Regression Model (degree 2) which performed at an accuracy of 96.1% (± 1.8 of their Actual BMD T-Scores)
  - 52.9% of the predictions were within 0.0 to 0.6 difference from the patients actual BMD T-Scores.
  - 31.6% of the predictions were within 0.6 to 1.2 difference.
  - 11.6% of the predictions were within 1.2 to 1.8 difference and the remaining 3.87% were within 1.8 and 2.4 difference.
- 155 patient records were used for the test set (20% of the 772 usable patient dataset).



CONCLUSIONS

- ❖ Given how most of the Osteoporosis patient data used in this research was categorical, the best performing models were those that focused on balancing continuous, nominal, and ordinal data.
- ❖ We created an ML model that can predict the BMD T-score using relevant patient data collected from the Ontario FLS program that performs at an accuracy of 96.1% (± 1.8 their Actual BMD T-Score) which can be used to do a preliminary assessment of their fracture risk. This model uses limited features available in real-world situations.
- ❖ Our next steps involve improving the Machine Learning models by using larger sample sizes and reducing the dimensionality through feature combinations (e.g., Height and Weight => BMI) or by lowering the weights/influence of certain features such as smoking.

Acknowledgement: This study was supported through funding from NSERC (grant #06227). The Ontario Osteoporosis Strategy is funded by the Ontario Ministry of Health. The views expressed are those of the stakeholders and do not necessarily reflect those of the Ministry. Fracture Screening and Prevention Program (Ontario FLS) is implemented by Osteoporosis Canada.

Sheridan | Centre for Mobile Innovation OSTEOPOROSIS CANADA inovex