Prudential ADS Technical Audit

Rohan Mukerji & Amando Xu

Background

- Prudential Life Insurance Assessment

- Kaggle Automated Decision System: Ridge Regression

| Variable	Description
Id	A unique identifier associated with an application.
Product_Info_1-7	A set of normalized variables relating to the product applied for
Ins_Age	Normalized age of applicant
Ht	Normalized height of applicant
Wt	Normalized weight of applicant
BMI	Normalized BMI of applicant

Implementation

- Data Pre-processing
    - Plotting
    - Fillna
    - Dummies
- Ridge Regression Experimentation
- Model Running
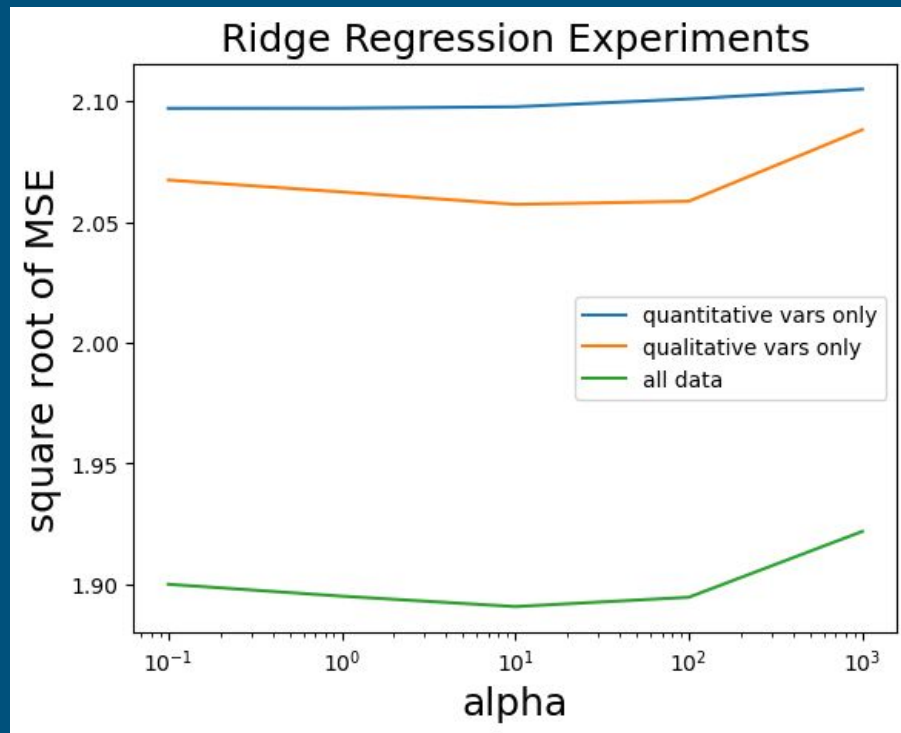
Outcomes (RMSE Sub-populations)

Age	Height	Weight	BMI
Young Adults: 1.67583	Short: 1.8751	Underweight: 1.8613	Underweight: 1.8737
Middle-aged: 1.82752	Medium: 1.8997	Normal-weight: 1.9312	Normal: 1.9269
Seniors: 2.22613	Tall: 1.9214	Overweight: 1.8692	Overweight: 1.9723
			Obese: 1.8160

Outcomes (Class Imbalances)

Age	Height	Weight	BMI
Young Adults: 0.465627	Short: 0.423144	Underweight: 0.593045	Underweight: 0.591530
Middle-aged: 0.321466	Medium: 0.313247	Normal-weight: 0.325355	Normal: 0.440901
Seniors: 0.201872	Tall: 0.239127	Overweight: 0.250118	Overweight: 0.261165
			Obese: 0.263925

Outcomes (Class Imbalances)

Age	Height	Weight	BMI
Young Adults: 0.465627	Short: 0.423144	Underweight: 0.593045	Underweight: 0.591530
Middle-aged: 0.321466	Medium: 0.313247	Normal-weight: 0.325355	Normal: 0.440901
Seniors: 0.201872	Tall: 0.239127	Overweight: 0.250118	Overweight: 0.261165
			Obese: 0.263925

Outcomes (Demographic Parity)

Age	Height	Weight	BMI
Young Adults: 0.579	Short: 0.364	Underweight: 1.022	Underweight: 1.075
Middle-aged: 0.079	Medium: 0.092	Normal-weight: 0.016	Normal: 1.219
Seniors: 0.781	Tall: 0.256	Overweight: 1.082	Overweight: 0.355
			Obese: 0.213

Outcomes (SHAP)

- Leverages game theory
- Provides local explanations for individual instances
- Assesses model stability, robustness and performance

Expected outcomes:

- Gain insights into model's decision-making process
- Identify potential biases
- Improve model performance

Summary

- Relatively high RMSE, all over 1, meaning a larger deviation between predictions from the actual values
- Favorable for stakeholders prioritizing bias & but not fairness
- Single Ridge Regression model limits robustness
- Recommendation: Enhance model selection complexity & access data with multiple models instead of one

Thank you