Mental Health in the U.S.

by:
Song Ying
Mate Pocs

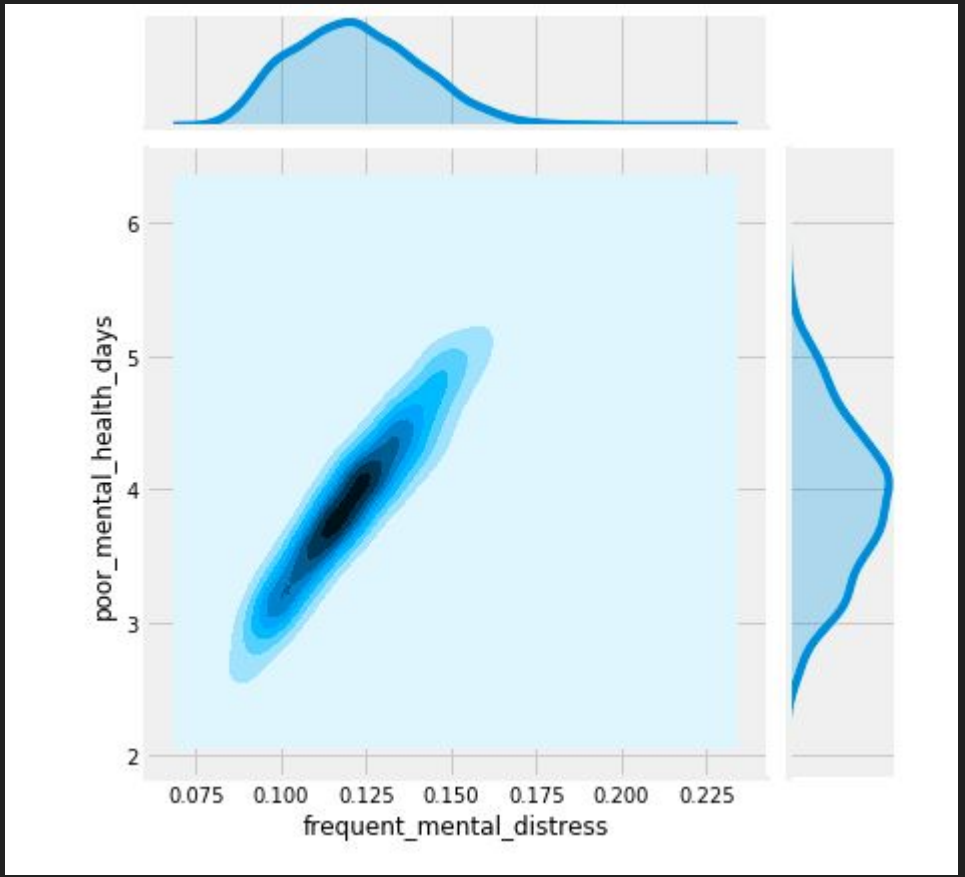
Introduction

Data Source:
2019 County Health Rankings, published by University of Wisconsin

Question: Which economic / demographic / environmental factors influence mental health?

Data Preparation Steps

Baseline Model,
K-Fold validation R^2: 0.76

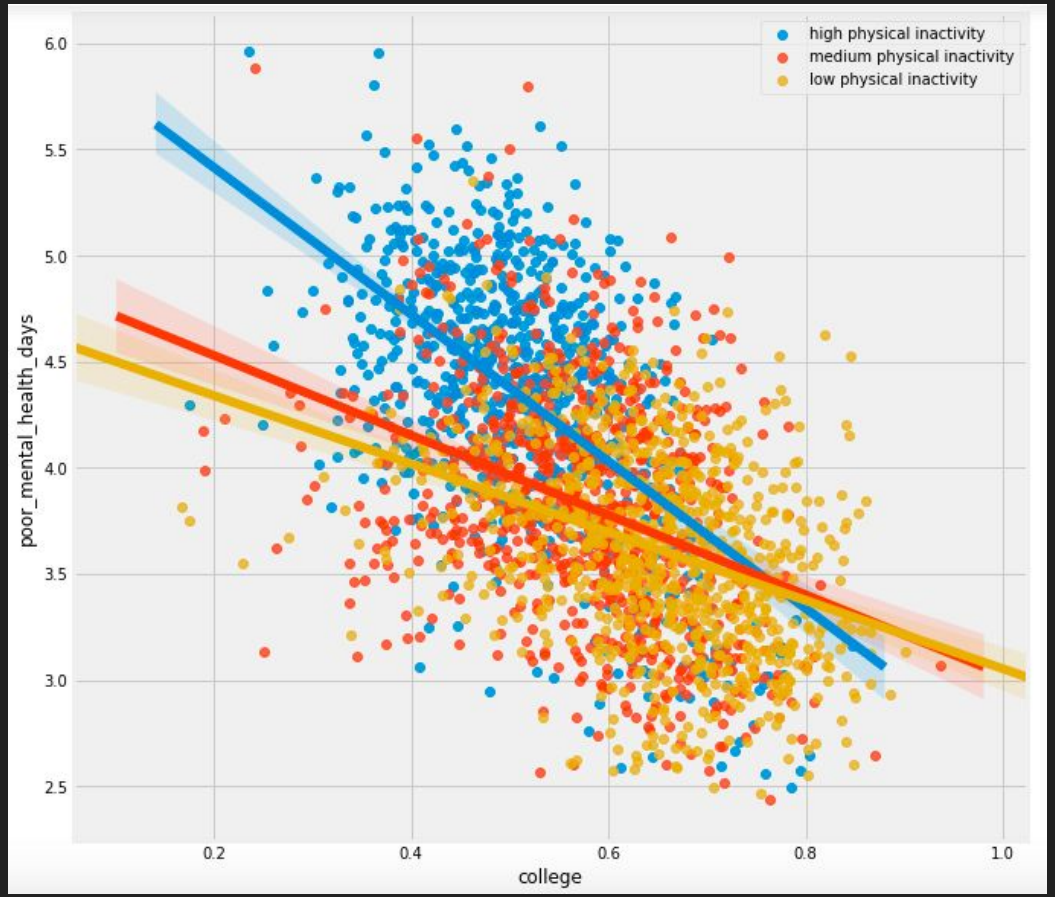


Interactions

A number of variables showed high levels of interaction with others:

- physical inactivity
- insufficient sleep
- excessive drinking
- college

Increases baseline R^2 by 4.1%

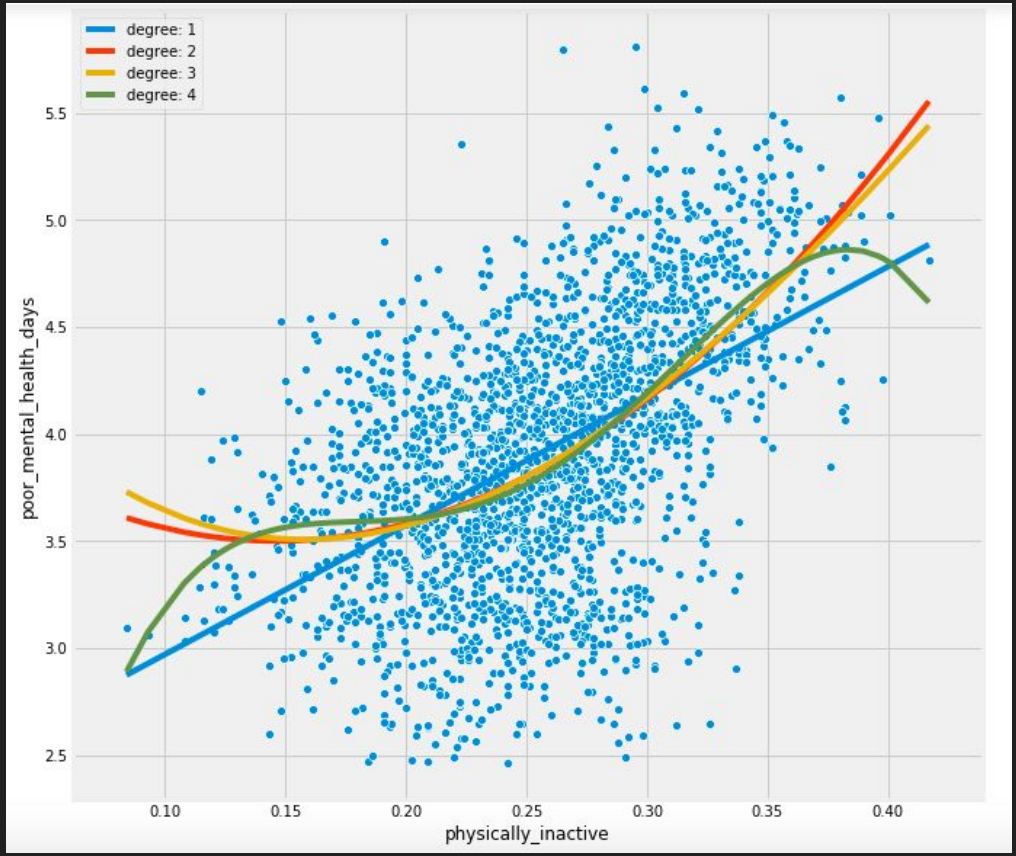


Polynomial Features

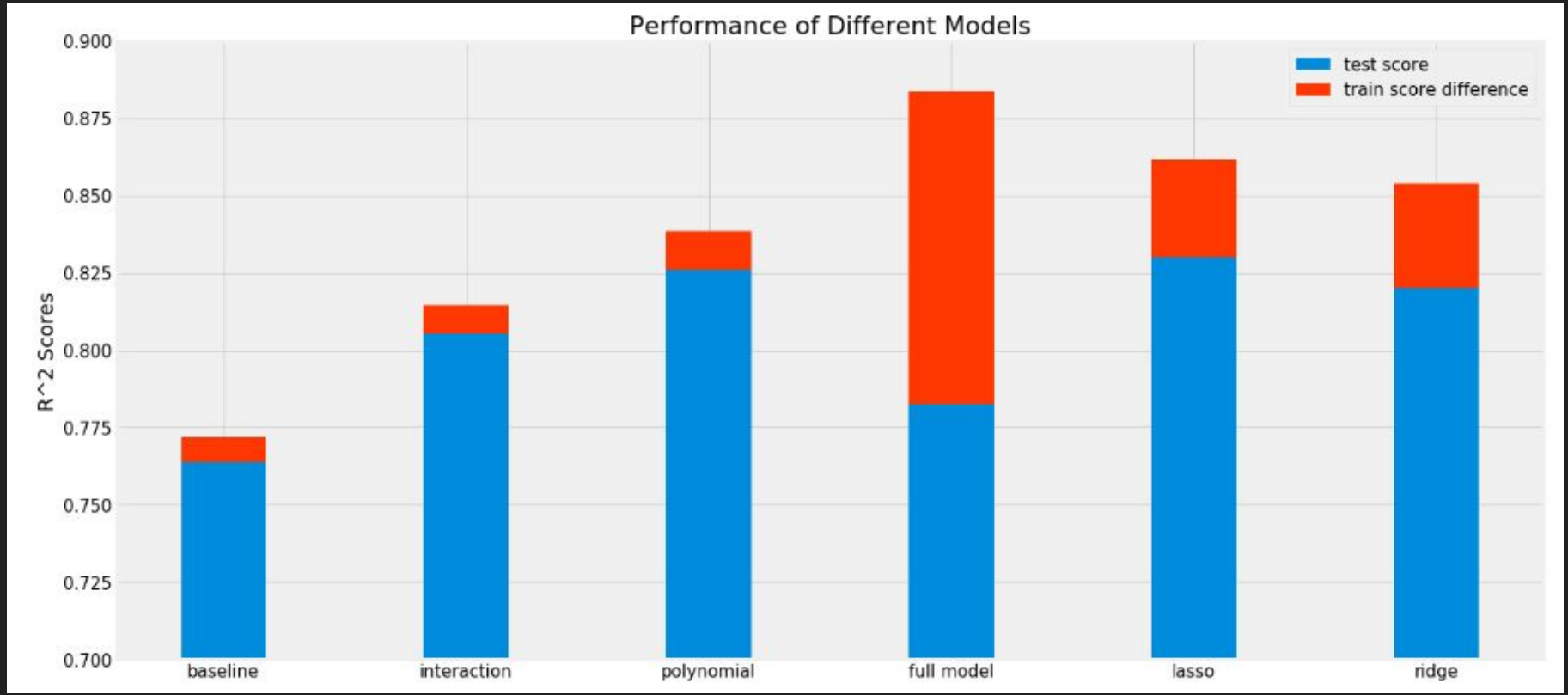
Adding higher degree versions of the variables also increases the performance of the model:

- insufficient sleep
- smokers
- physical inactivity

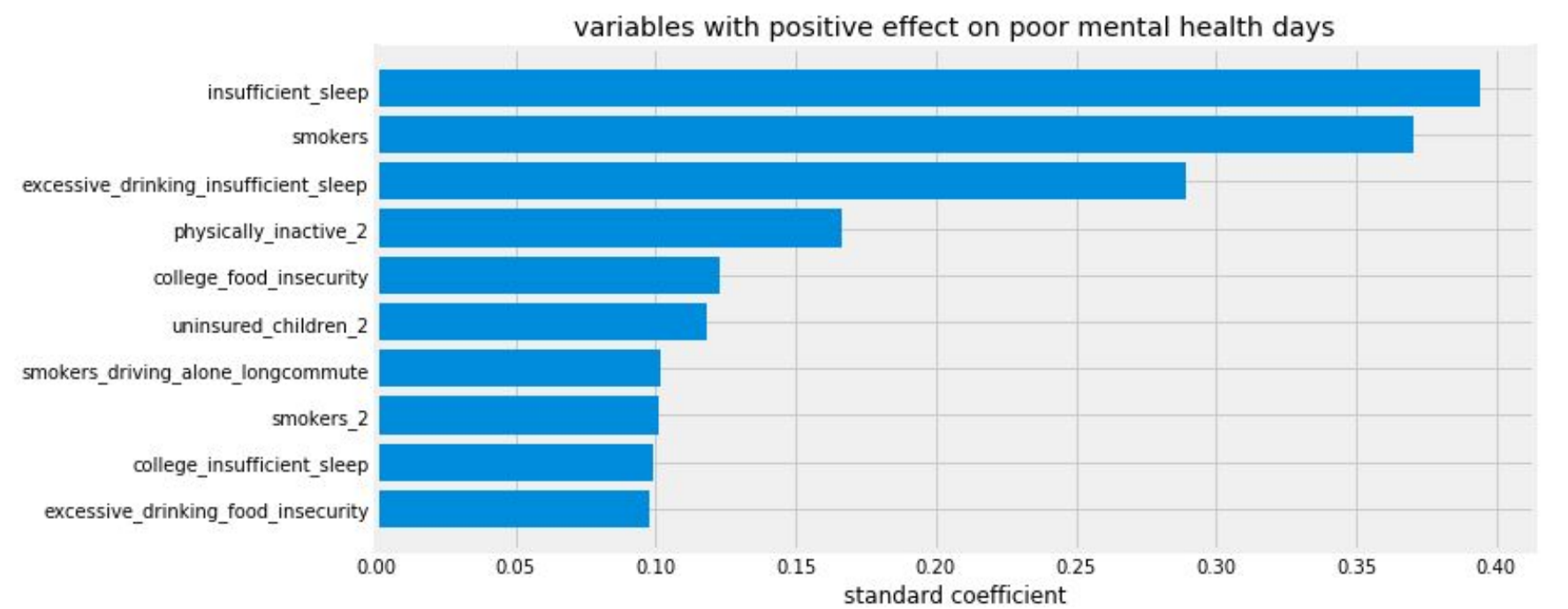
Increases baseline R^2 by 6.2%



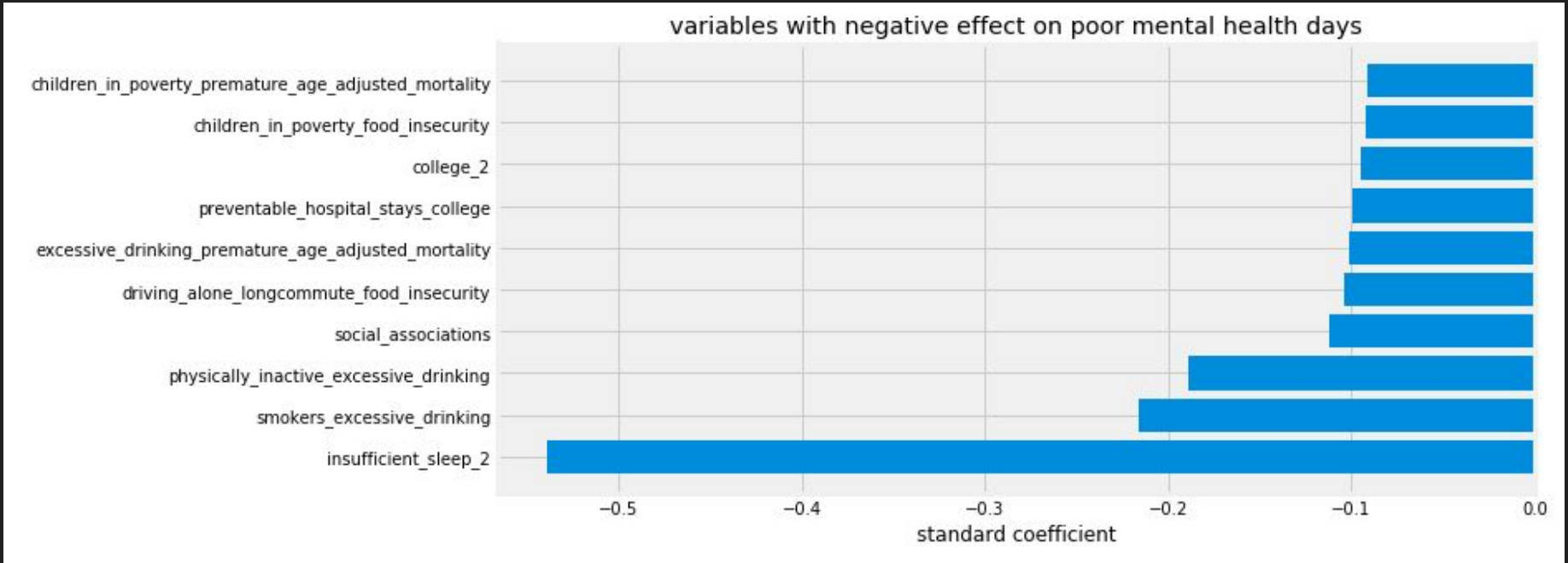
Comparison of Models



Model Insights - Variables Weakening Mental Health



Model Insights - Variables Improving Mental Health



Actionable Insights

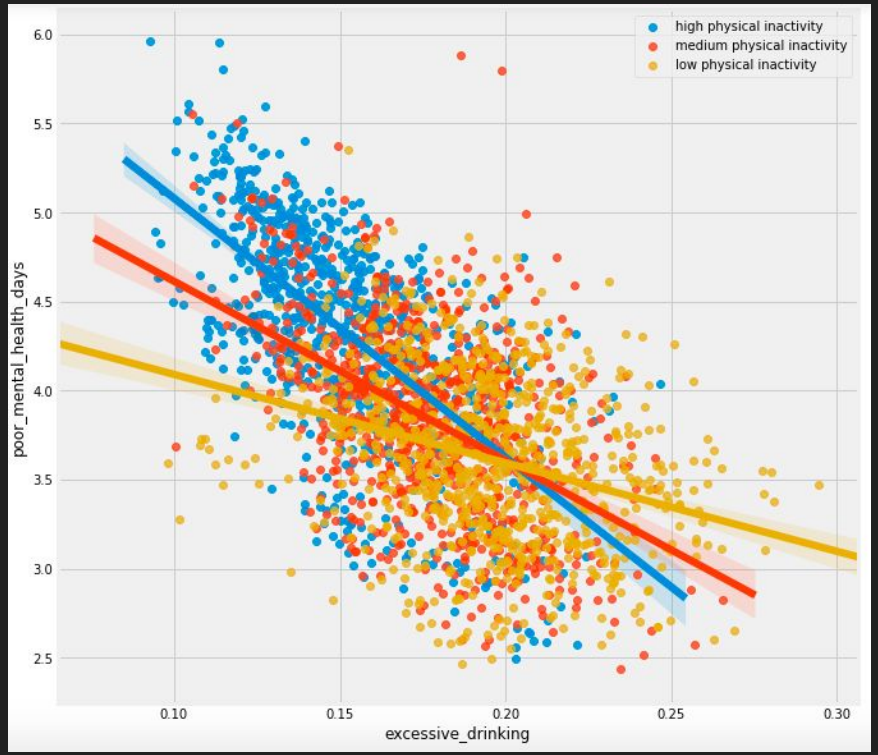
1. Insufficient sleep
    - Good predictor for future mental health issues
    - Mental health can be improved via awareness of this factor
    - Improve sleep quality with external factors, e.g. work-life balance

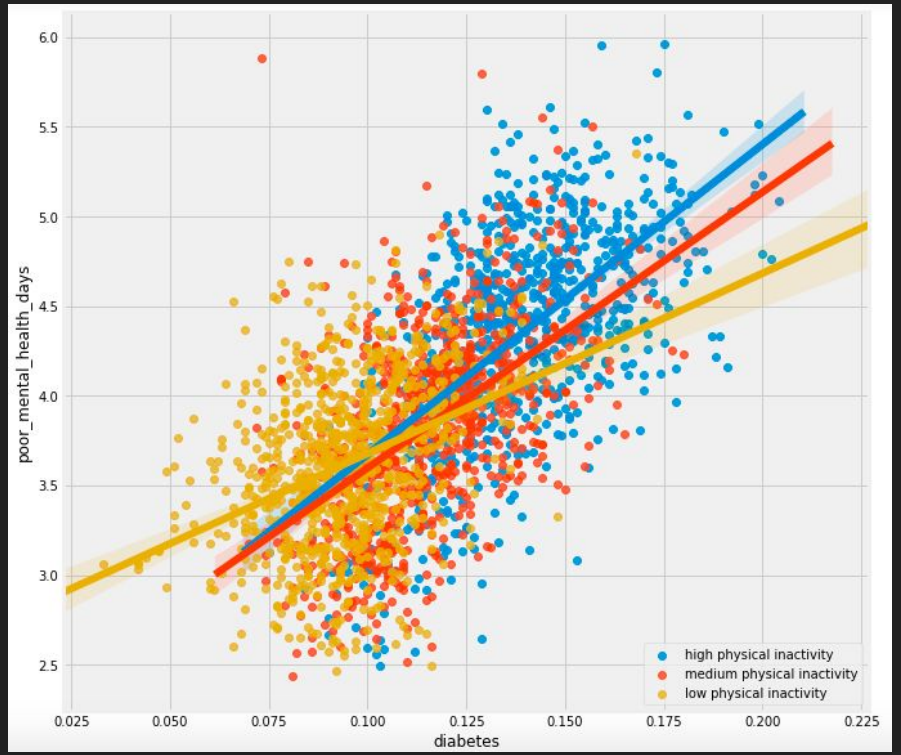
2. Physical inactivity
    - Promote more physical activity by improving access to sports facilities

3. Smoking
    - This fact could improve the effects of anti-smoking advertisements

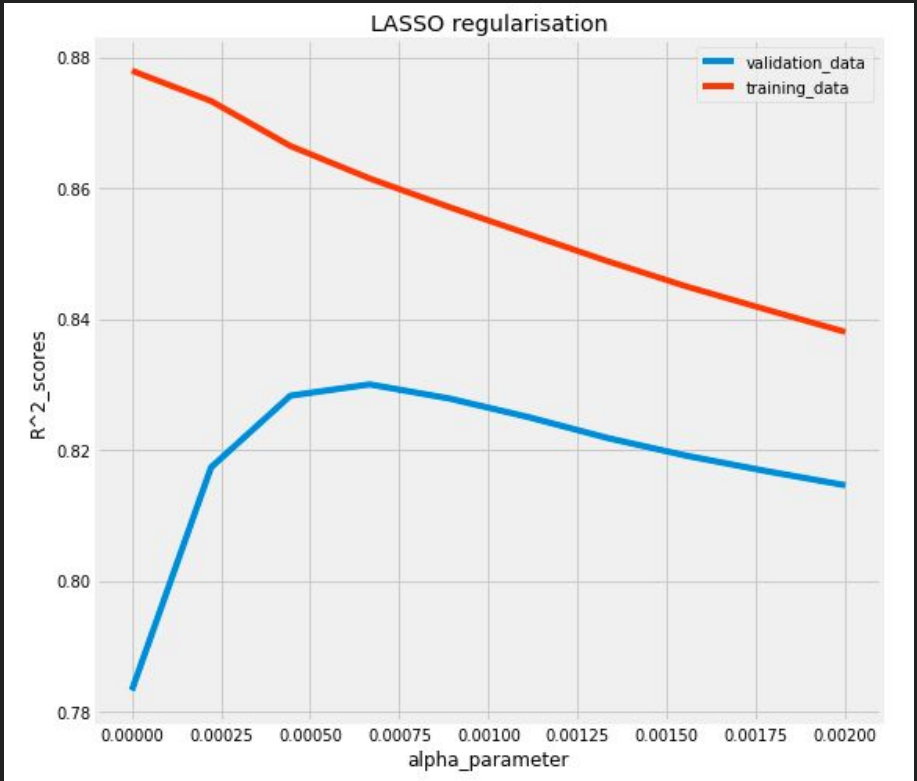
Thank you!
Questions?

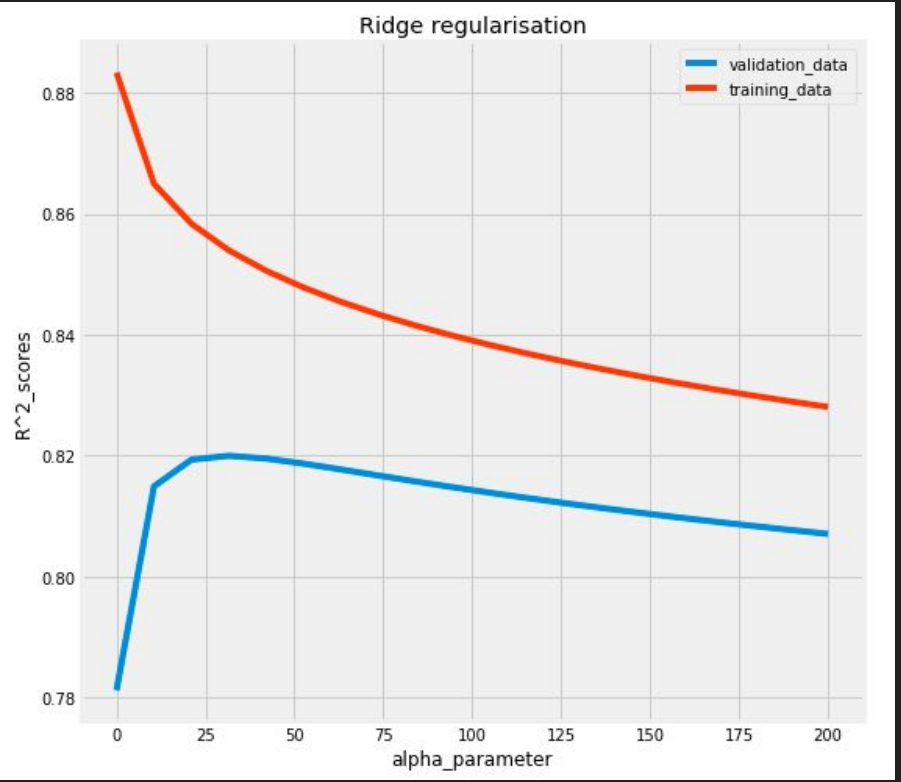
Technical Appendix - Interactions





Technical Appendix - Lasso and Ridge Methods





Technical Appendix - Final Model

1. 2nd degree PolyFeatures transformation of the base data
2. Scaled using StandardScaler
3. LASSO regularisation with alpha = 0.00067
4. Performance:
    a. 0.858 R^2 on train data set
    b. 0.850 R^2 on test data set

Technical Appendix - Test Data Residuals