Life Expectancy in the US

Khairul Omar | Stuti Singh | Aleksandar Gakovic

from LiveLong Consultancy

Life expectancy by county

> 84
79 - 84
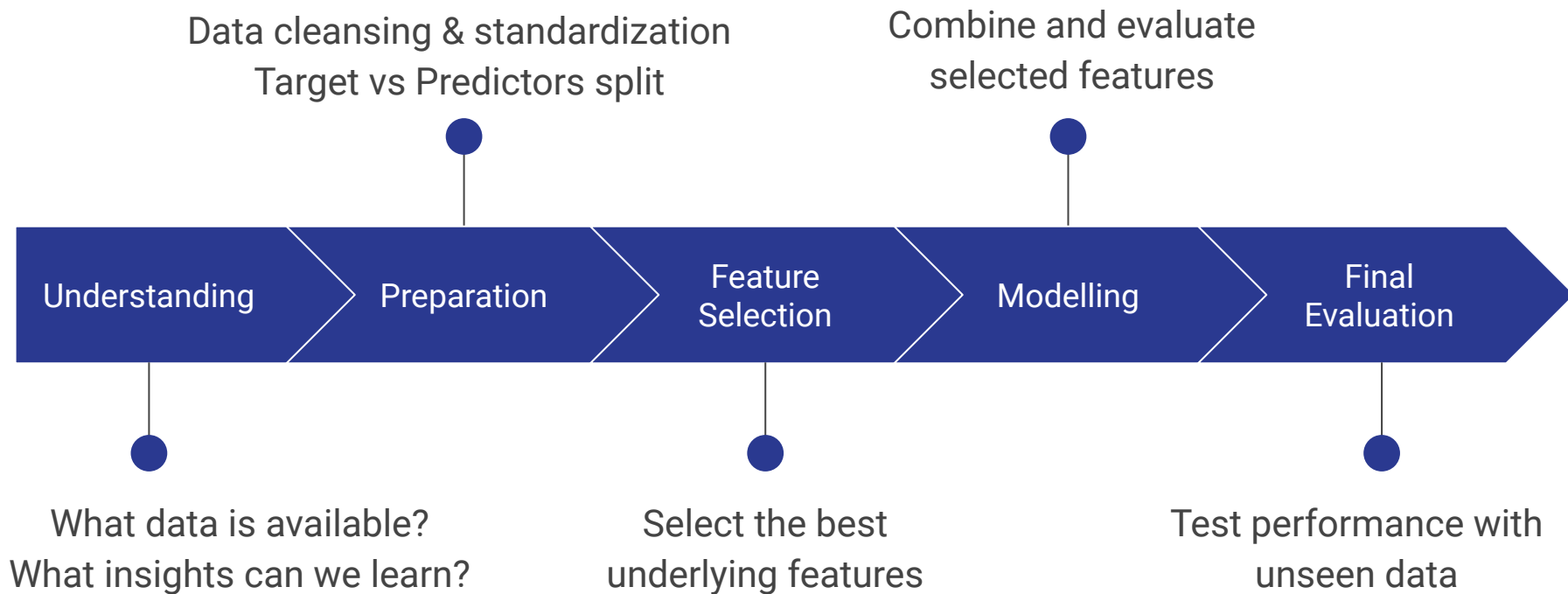77 - 79
75 - 77
70 - 75
< 70

Introduction

- **Problem**: Big differences in life expectancy (20+ yrs) between counties

- **Mission**: to support govt to formulate healthcare policy and budget allocation

- **Objective**: to find the most significant contributing factors to life expectancy
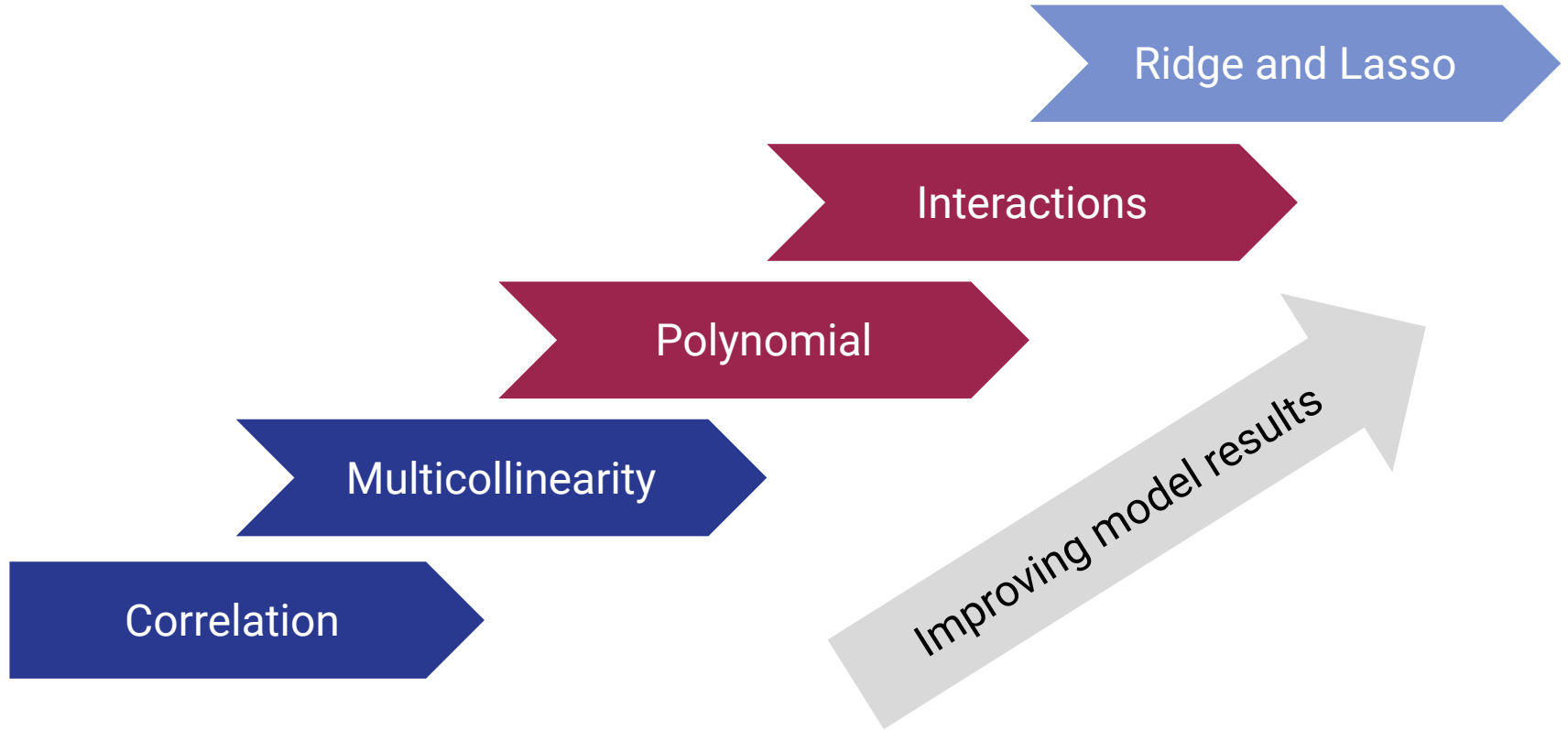
Methodology

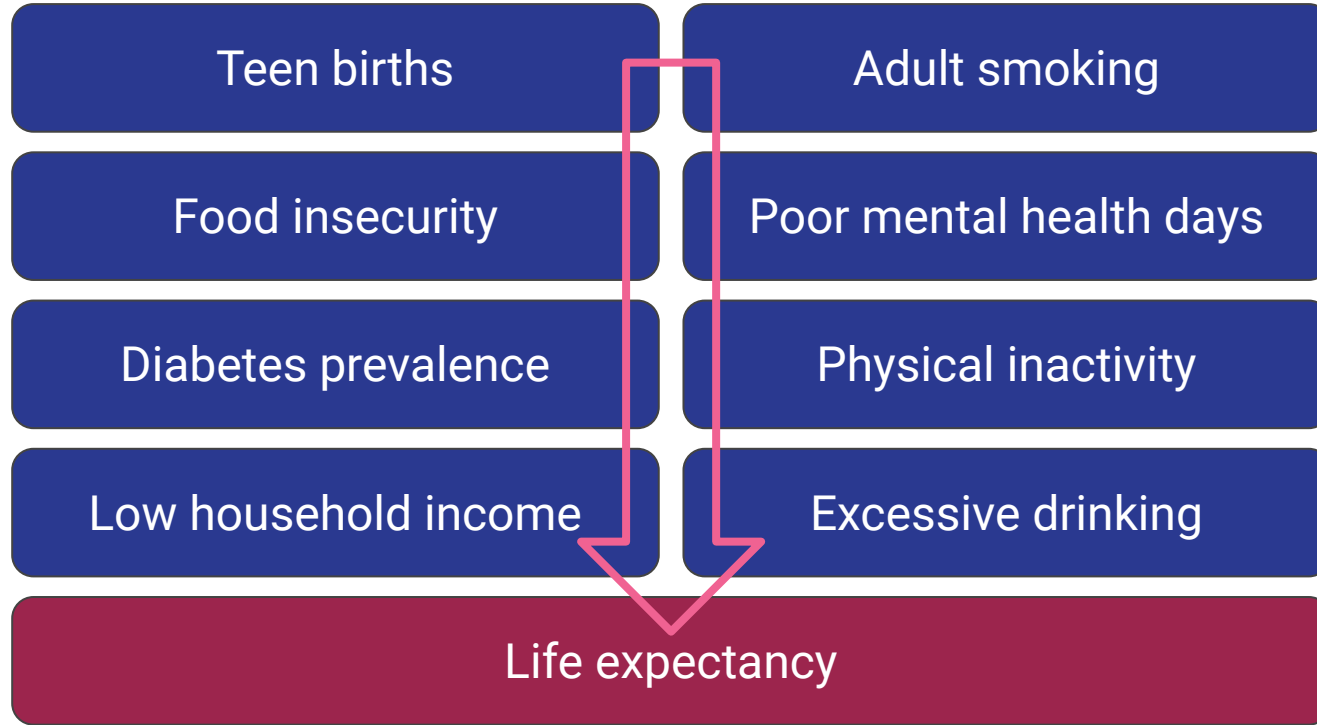Data cleansing & standardization
Target vs Predictors split

Combine and evaluate
selected features

| Understanding | Preparation | Feature Selection | Modelling | Final Evaluation |

What data is available?
What insights can we learn?

Select the best
underlying features

Test performance with
unseen data

Feature selection techniques

Ridge and Lasso

Interactions

Polynomial

Multicollinearity

Correlation

Improving model results

Top factors that reduce life expectancy

Teen births	Adult smoking
Food insecurity	Poor mental health days
Diabetes prevalence	Physical inactivity
Low household income	Excessive drinking

Life expectancy

Model result and validation
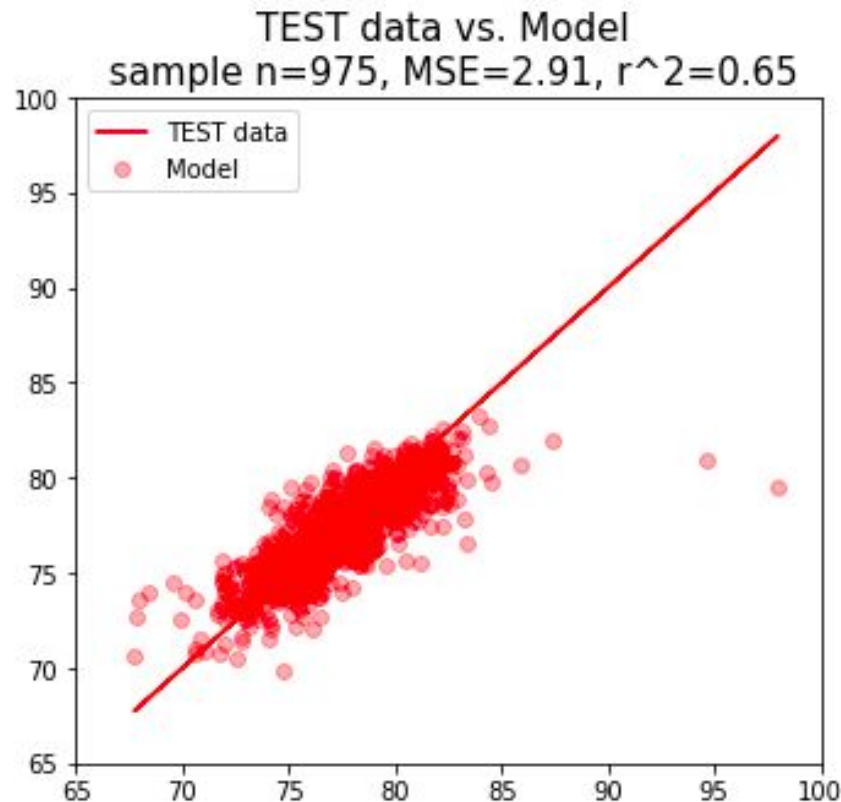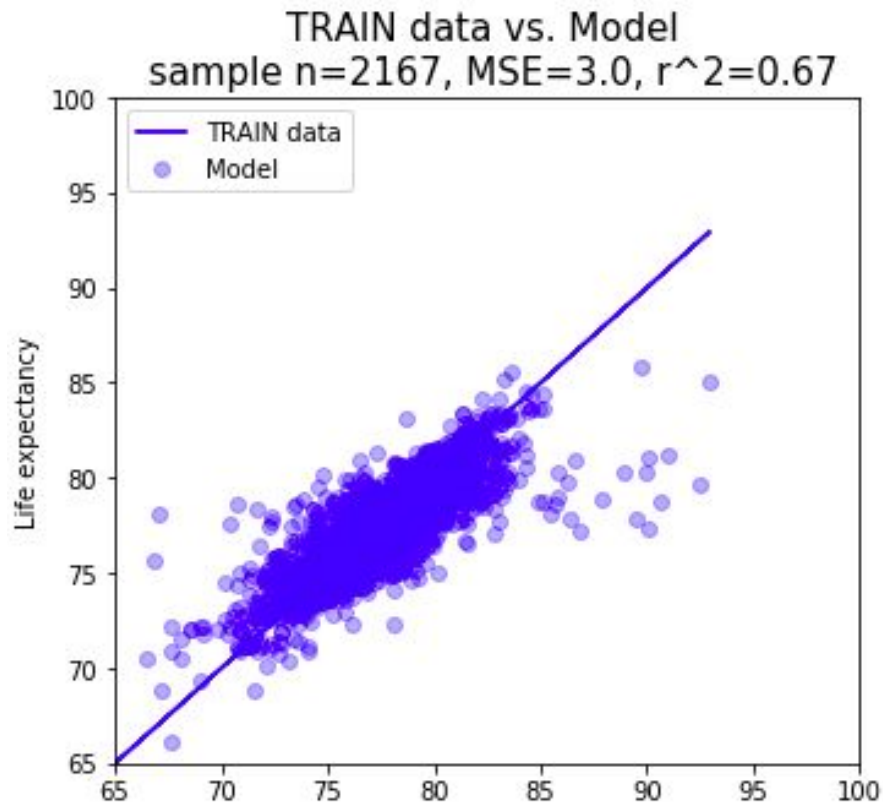
Conclusions

- Teen births, smoking and food insecurity are top contributors to lower life expectancy

- Trade-off between model's accuracy and ability to predict using unseen data input

- Valuable insight for the govt to assist in formulating healthcare policies and budgeting

Q & A