

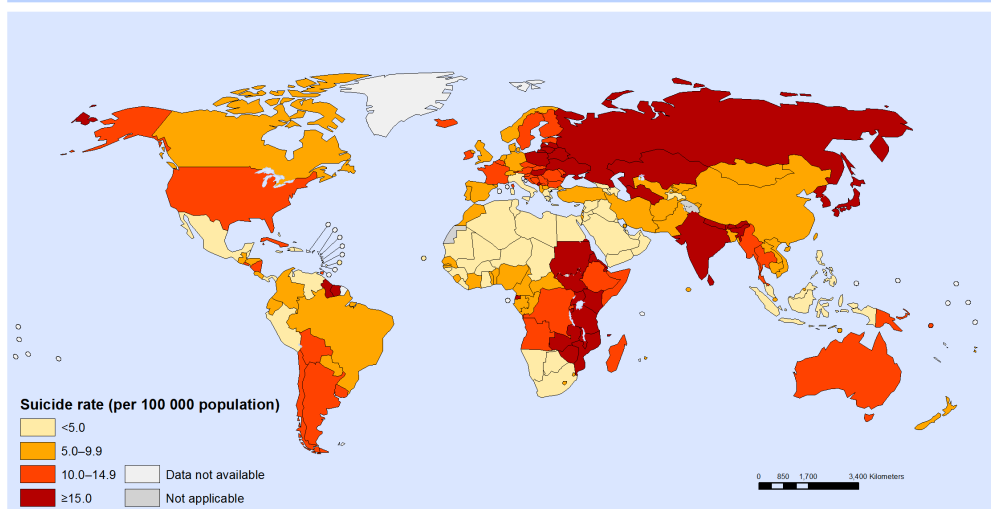
Group 10

Modeling Suicide Rates at the Country Level

Monal Garg (mgarg@sas), Amit Gupta (akgupta@seas), Aashish Jain (aashj99@seas), Moksh Jawa (moksh@seas)

Mentor: Jane Lee (janelee@seas)

Age-standardized suicide rates (per 100 000 population), both sexes, 2012



The boundaries and names shown and the designations used on this map do not imply the expression of any opinion whatsoever on the part of the World Health Organization concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries. Dotted and dashed lines on maps represent approximate border lines for which there may not yet be full agreement.

Data Source: World Health Organization
Map Production: Health Statistics and
Information Systems (HSI)
World Health Organization

World Health
Organization
© WHO 2014. All rights reserved.

Inputs :

Youth Unemployment, Alcohol
Consumption, Life Expectancy,
Country Population, Fertility
Rate

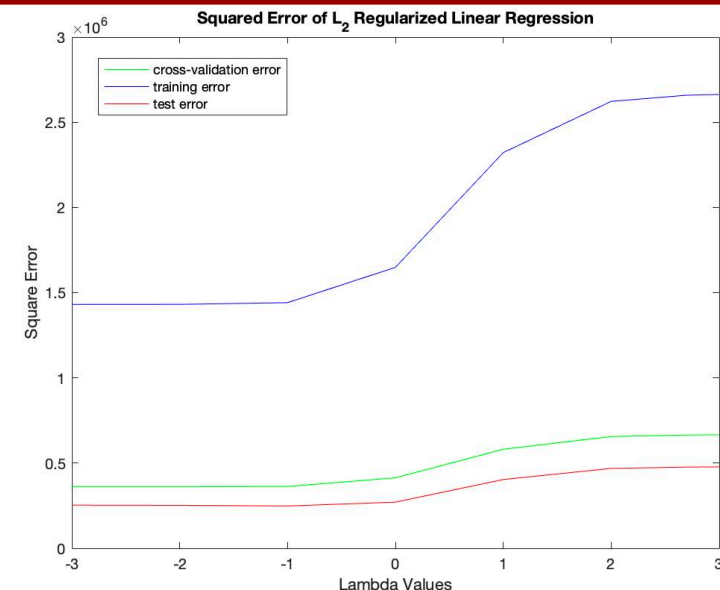
Output :

Number of Suicides per 100k
population (in a year)

Model :

Regression Problem - $R^8 \rightarrow R$

- Countries too distinct for Nearest Neighbor



Nearest Neighbors

- Euclidean Distance
- Minkowski Distance

$$MD_p(\mathbf{q}, \mathbf{x}_i) = \left(\sum_{f \in F} |\mathbf{q}_f - \mathbf{x}_{if}|^p \right)^{\frac{1}{p}}$$

- $k = [1, 2, 5, 10]$
- $q = [1, 2, 3]$



Best Method:

$k = 1, q = 2$

1-NN with Euclidean

Regression

- Least Squares
 - Unregularized
 - L_2 Regularized
- Logistic
 - Unregularized
 - L_2 Regularized
- $\lambda = [10^{-3}, 10^{-2}, 10^{-1}, 1, 10, 100, 500, 1000]$

Best Method:

$\lambda = 10^{-2}$

Logistic, L_2 Regularization

Neural Network

- ReLU
- Sigmoid
- # Iterations = 20000
- Learning Rate = 0.01
- 1 Hidden Layer
- Hidden Layer Size = $[1, 5, 10, 15, 25, 50, 100]$

$$Y_{country, t=2013} \approx Y_{country, t=2014}$$

Best Method:

Sigmoid Loss Function

Hidden Layer with 100

Nodes

- Observable data is best explanation
- Machine learning is **not a must** given data

