Using Cognitive Skills and Social Attitudes of Ordinary Citizens to Predict Preparedness for Pandemic: Lessons from PIAAC

Chong Ho (Alex) Yu, PH.D., D. Phil. (Azusa Pacific University)  
Jolia Awadallah (Alliant International School of Psychology)  
David Xiao (University of Maryland)  
March 24 2022 International PIAAC Research Conference
Introduction

- Since 2020, the world has been facing a tragic pandemic caused by a small virus that led to the death of millions of people.
- In spite of availability of cutting-edge technologies, the countermeasures launched by the US and Western nations are not as effective as expected.
Lit Review

- However, the relatively poor performance of the USA in public health is not unforeseeable.
- In the past there are many other indicators. In spite of better medical technologies and higher living standards, the life expectancy of Americans is decreasing and adult mortality rates are increasing.
Lit Review Cont..

- Orange line: 2016
- Others: Previous years
- Source: Sterling et al. (2022)
Lit Review Cont..

- Orange line: 2016
- Others: Previous years
- Source: Sterling et al. (2022)
Lit Review Cont..

- Obesity rates are getting worse in the US.
- Source: Sterling et al. (2022)
Rieger & Wang (2021)’s study indicates social trust variables as important predictors of pandemic control.
Moon et al. (2021) shows that skepticism and disposition to conspiracy theories made the pandemic worse.

Distrust the government -> accept conspiracy theories and reduce prosocial behaviors (don’t care about other people).
In 2019 the Global Health Security Index rated many Western developed nations as the most prepared countries for countering pandemic.

USA is Number 1.
Gap Con...

- This prediction is understandable because the USA and European nations have better technologies.
- However, when the COVID19 pandemic broke out, the outcome is surprisingly: Those highly rated countries did not do well!
If the criteria used by the Global Health Security Index cannot predict pandemic control, then what are the really important factors?

Having a few elite scientists and medical researchers is insufficient. Cognitive skills and social mentalities (sense of civil responsibilities) of ordinary citizens, as indicated in the PIAAC survey, might be more important.
Data sources

- **Worldometer**: Pandemic data as of March 16, 2022
- **Programme for International Assessment of Adult Competencies (PIAAC)**: Round 1-3 from 2012 to 2017
  - Cover 37 countries: e.g., USA, UK, Japan, South Korea, Germany...etc.
  - Cognitive domains: Literacy, numeracy, and problem solving
  - Social attitudes: e.g., willingness to do voluntary work, trust other people...etc.
Methodologies and Results

- **Model comparison**: Identify the best overall model.
- **XGBoost**: Identify the most important factors contributing to pandemic control.
- **Data visualization**: Find the relationship between the independent variables and the outcome.
Eight different modeling methods were used to select the best model. XGBoost (Extreme Gradient Boosting) is the best in terms of RASE.

<table>
<thead>
<tr>
<th>Method</th>
<th>N</th>
<th>RASE</th>
</tr>
</thead>
<tbody>
<tr>
<td>XGBoost</td>
<td>37</td>
<td>0.00053</td>
</tr>
<tr>
<td>Fit Least Squares</td>
<td>30</td>
<td>1117.7</td>
</tr>
<tr>
<td>Neural Boosted</td>
<td>20</td>
<td>493.5</td>
</tr>
<tr>
<td>Bootstrap Forest</td>
<td>37</td>
<td>742.7</td>
</tr>
<tr>
<td>Support Vector Machines</td>
<td>30</td>
<td>834.9</td>
</tr>
<tr>
<td>Boosted Tree</td>
<td>37</td>
<td>825.1</td>
</tr>
<tr>
<td>Fit Stepwise</td>
<td>30</td>
<td>1016.4</td>
</tr>
<tr>
<td>Generalized Regression Lasso</td>
<td>37</td>
<td>1176.6</td>
</tr>
</tbody>
</table>
XGBoost

- Boosting consists of a series of partition trees.
- In boosting the subsequent partition tree models are fine-tuned by learning the errors from the previous models.
- No further improvement after 22 iterations.
Variable importance in terms of information gain

**Gain**: The average improvement in information gain by partitioning with the variable
Variable importance in terms of number of splits across all decision trees

**Splits**: The number of splits by this variable across all the partition trees, also known decision trees.

**Cover**: The amount of data covered by the splits involving the variable.

<table>
<thead>
<tr>
<th>Feature</th>
<th>Splits</th>
<th>Gain</th>
<th>Cover</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean(F_Q05a)</td>
<td>296</td>
<td>10393.5</td>
<td>3.16892</td>
</tr>
<tr>
<td>Mean(G_Q01f)</td>
<td>125</td>
<td>8821.47</td>
<td>23.96</td>
</tr>
<tr>
<td>Mean(G_Q01a)</td>
<td>120</td>
<td>7239.56</td>
<td>4.69167</td>
</tr>
<tr>
<td>Mean(PVNUM5)</td>
<td>89</td>
<td>71451.3</td>
<td>36.7416</td>
</tr>
<tr>
<td>Mean(G_Q01e)</td>
<td>62</td>
<td>80528.8</td>
<td>20.3387</td>
</tr>
<tr>
<td>Mean(I_Q04m)</td>
<td>54</td>
<td>262030</td>
<td>24.4074</td>
</tr>
<tr>
<td>Mean(PVPSL5)</td>
<td>47</td>
<td>53592.3</td>
<td>18.4468</td>
</tr>
<tr>
<td>Mean(F_Q05b)</td>
<td>38</td>
<td>12504.8</td>
<td>17.8421</td>
</tr>
<tr>
<td>Mean(I_Q04h)</td>
<td>35</td>
<td>154715</td>
<td>13.9714</td>
</tr>
<tr>
<td>Mean(I_Q07a)</td>
<td>30</td>
<td>802428</td>
<td>18.5</td>
</tr>
<tr>
<td>Mean(G_Q01a_T)</td>
<td>28</td>
<td>25163</td>
<td>11.7857</td>
</tr>
<tr>
<td>Mean(I_Q06a)</td>
<td>23</td>
<td>231796</td>
<td>16.1739</td>
</tr>
<tr>
<td>Mean(H_Q01a)</td>
<td>19</td>
<td>124862</td>
<td>21.5789</td>
</tr>
<tr>
<td>Mean(I_Q05f)</td>
<td>15</td>
<td>7302.67</td>
<td>13.8667</td>
</tr>
<tr>
<td>Mean(I_Q04b)</td>
<td>14</td>
<td>86414.2</td>
<td>15.4286</td>
</tr>
<tr>
<td>Mean(I_Q07b)</td>
<td>12</td>
<td>1444840</td>
<td>31.4167</td>
</tr>
<tr>
<td>Mean(H_Q01e)</td>
<td>11</td>
<td>627.15</td>
<td>11.5455</td>
</tr>
<tr>
<td>Mean(I_Q04l)</td>
<td>10</td>
<td>389.812</td>
<td>12.5</td>
</tr>
<tr>
<td>Mean(H_Q02d)</td>
<td>10</td>
<td>320.995</td>
<td>16.9</td>
</tr>
<tr>
<td>Mean(H_Q03b)</td>
<td>9</td>
<td>95425.2</td>
<td>15.5556</td>
</tr>
</tbody>
</table>
Data visualization

- 1 = strongly agree
- 5 = Strongly disagree
- More agreement to “there are only a few people you can trust completely” -> higher deaths /1M population.
Data visualization

- 1 = strongly agree
- 5 = Strongly disagree
- More agreement to “If you are not careful, other people will take advantage of you” -> higher deaths /1M population
Data visualization

- **1** = Never
- **5** = Everyday
- The pattern is not clear
Data visualization

- 1 = Never
- 5 = Everyday
- No pattern at all
- Although these two variables are identified as important predictors in terms of the number of splits across all decision trees, no discernible pattern is observed.
Higher score = better problem solving skill
Problem solving skill is not highly ranked in terms of information gain, but there is some degree of negative association between problem solving and deaths/1 M population. Literacy and numeracy were examined and they have no relationship with deaths by COVID19.
Conclusion and Discussion

- In line with the literature, variables related to social trust are important predictors of pandemic control in terms of reducing deaths/1 M.
- Between the three key skill sets (literacy, numeracy, and problem solving) measured by PIAAC, problem solving skill is slightly associated with the outcome of pandemic control while the other two are not.
References

Thank You!