EY Techathon 2021

#iSolve4aBillion Challenge

AI/ML

Help save lives and reduce the chain of transmission with AI enabled prioritization of vaccine delivery.
As the country gears to unveil, register and approve its choice of vaccine candidates against COVID-19, the first set of population expected to be immunized are likely to be the frontline health workers, doctors, citizens over 50 years of age and citizens with co-morbidities. The challenge for the healthcare administrators will be to prioritize this population, considering the limited availability of vaccines initially and at the same time, limiting the spread of COVID-19.

Solution

Provide a predictive model to help make informed decisions/insights around epidemiological and vaccine supply circumstances and priority group segments that need to be recommended for vaccination. The proposed AI algorithm/solution should have a user interface (e.g. web enabled) to interact with data for visualization, analysis as well as support for API integration and containerization (K8s/ Docker are encouraged).

Prioritize vaccine delivery

Human wellbeing

Reduce societal and economic disruption by containing transmission, reducing severe disease and death, or a combination of these strategies.

Equal respect

Treat the interests of all individuals and groups with equal consideration as allocation and priority-setting decisions are being taken and implemented.

National equity

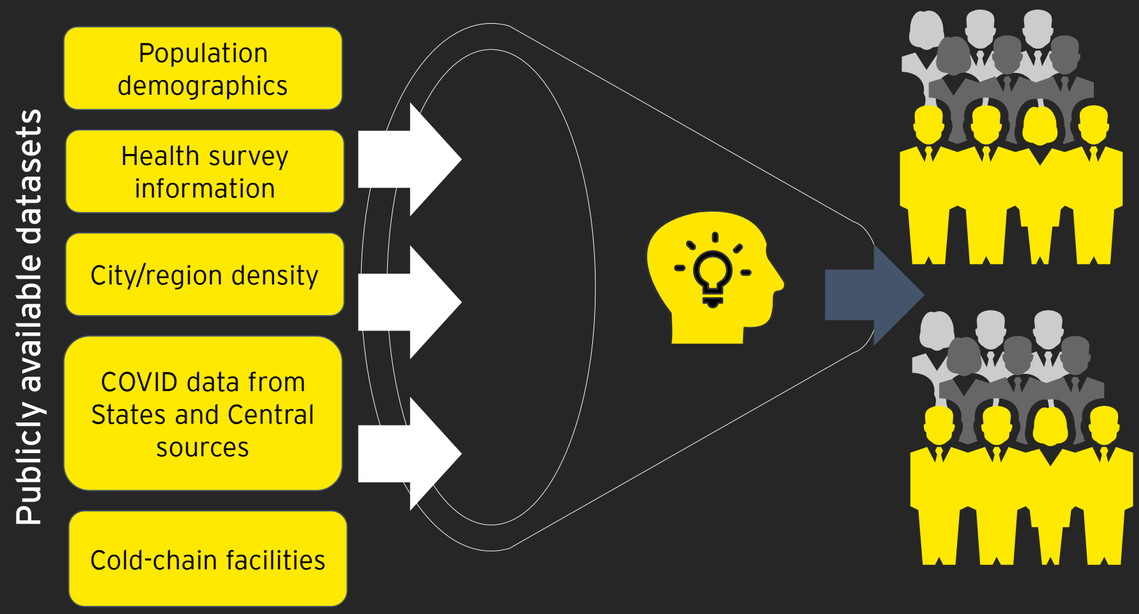
Ensure that vaccine prioritization accounts for the vulnerabilities, risks and needs of groups who, because of underlying societal, economic (low and middle income groups), geographic or biomedical factors are at risk of experiencing greater burdens from the pandemic.

Reciprocity

Protect those who bear significant additional risks and burdens of COVID-19 to safeguard the welfare of others, including health and other essential workers.

Legitimacy

Employ best available scientific evidence, expertise and significant engagement with relevant stakeholders for vaccine prioritization between various groups using transparent, accountable and unbiased processes to engender deserved trust in prioritization decisions.



EY
Building a better working world

Solution approach

1. **Gather population demographics related data from public sources** — Population density, electoral rolls, data for comorbidities across States and districts etc.
2. **Gather COVID-19 cases related data from public sources** — Gather State and city-wise data on COVID-19 risk clusters, containment zones, day-wise COVID +ve trends and testing results.
3. **Perform exploratory data analysis** — Perform a uni-variate and bi-variate analysis and find the co-relation between the attributes. Eliminate bias in data.
4. **Perform feature engineering and dimensionality reduction** — Determine feature selection and feature creation based on principles defined.
5. **Create baseline Machine Learning models** — Run the data on multiple simple clustering models (K-means, GMM) with hyper parameters.
6. **Compare multiple partition-based clustering models** — Compare multiple robust models to improve the evaluation metrics - Silhouette coefficient and Elbow Curve.
7. **Evaluate a ranking algorithm and make submissions** — Evaluate a ranking algorithm to identify prioritization metrics.
8. **Build interactive dashboards** — Represent the results in an interactive dashboard with multiple categorizations based on population risk and societal and economic disruption.

Technology and tools
Python, Java or any programming language like Julia, Haskell etc., EY's proprietary AI platform – ASpace, web application framework

Reference data sets
Google's Datasets Search Engine, Kaggle Datasets. Amazon Datasets (Registry of Open Data on AWS)

EY will open up our proprietary AI platform ASpace to develop these solutions to a limited number of participants on a first come first serve basis

Expected results and sample visualizations

Expected results

Populations with significantly elevated risks of severe disease or death

- Older adults defined by age-based risk
- Groups with comorbidities
- Sociodemographic groups at disproportionately higher risk of severe disease or death

Populations with significantly elevated risk of being infected

- Health workers at high or very high risk
- Employment categories unable to physically distance
- Groups living in dense urban neighbourhoods
- Groups living in multigenerational households

Reduce societal and economic disruption:

- School-aged children to minimize disruption of education and socioemotional development
- Workers in non-essential but economically critical sectors, particularly in occupations that do not permit remote work or physical distancing while working

Building a better working world