

Datathon 2020 Final Round

Use Case

1. Predict vaccine distribution based on the COVID Spread. You may use the implementations that have been done for the Datathon Round 1. Distribution requirement needs to be predicted to the following levels.

- 1.Province
- 2.District
- 3.Age
- 4.Profession

You may present your findings as a report or dashboard.

2. Predict impact to the main industries in Sri Lanka due to COVID. You can consider the main industries as follows.

- Agriculture
- Apparel
- IT Services
- Tourism and Hospitality
- Telecommunication
- Transportation
- Banking and Finance
- Retail
- Food Production & Distribution

Select three industries from above and showcase,

1.How the industry has been affected due to Covid. You may present your findings as a report or dashboard.

2.Predict the performances for the next six months out of the selected industries.

Resources

- You may use the following data sources and any other relevant data sources available. (You may use dummy data for the models where relevant data is not available)

<http://www.statistics.gov.lk>

<https://knoema.com/atlas/Sri-Lanka/datasets>

Evaluation Criteria

Duration for the Viva Session - 15 minutes(7 minutes to present, 8 minutes Q&A)

- Ideation & Presentation (5)
- Approach (25)
 - Approaches used to tackle the problem
 - What type of machine learning approaches used predictions generated
- Novelty of the idea (10)
- Implementation (25)

Must have:

- Data sources and Data Pipeline
- Model implementation
- Model evaluation and accuracy
- Feature engineering and data preprocessing

Nice to have:

- Data warehousing and ETL process
- Flow orchestration
- Deployment process

All the data sources, source codes/scripts need to be submitted at the end of the competition

- Visualization (20)
 - Report or Dashboard
- Q&A (15)