



Data Incubator Fellowship Interview Pitch Presentation

November 19, 2020

By: Ephrem Tadesse

Title: COVID-19 Ethiopia Timeseries Data Analysis
and Forecasting

Introduction

- **COVID-19**

- ✓ Coronavirus disease has widely spread all over the world since the beginning of 2020

- **Ethiopia**

- ✓ Largest populous country in Africa next to Nigeria with limited healthcare facilities

- **Data analysis techniques**

- ✓ Various ds and ml tools such as Pandas, NumPy, Scikit Learn and statsModels.

- **Timeseries data**

- ✓ A time-stamped dataset is sequence of data points indexed in time order

- **Forecasting models**

- ✓ ARIMA, regression models, and sequential deep learning models

Motivation

- *Its remembered that, following Covid-19 pandemic schools and other government and private organizations has been forced to stop their regular activities.*
- In Africa, mainly in Ethiopia available healthcare infostructures are limited in proportion to available population size
- Most of the organizations has no culture of building informative datasets; thus making decisions and planning is done without evidence.
- *Allocating resources and planning to create the new normal is difficult due lack of proper projections of the COVID-19 cases, recovery and death rate.*
- From my previous research experience I am passionate about building, extracting, and manipulating timeseries data in structured and unstructured format.
- To do so, this project has been proposed to tackle all the points above

Main features of the project

- Timeseries dataset construction from the scratch
- Data munging and wrangling
- Explanatory data analysis
- Forecasting the rate of cases, recovery and deaths

Significance of the project

- This project mainly helps government offices to allocate resources and make plans to create the new normal during and after COVID-19 pandemic
- Develop culture of building informative datasets and making decisions based on evidence or actual experience in Ethiopia.
- Moreover, it helps the scientific community to make comparison between other countries and deduce conclusion or to draw other relevant hypothesis.

Target customers of the project

- Direct beneficiaries/stakeholders
 - Ministry of Health in Ethiopia, public health institute in Ethiopia
 - Healthcare organizations(Hospitals, clinics, Medical Laboratories)
- Indirect beneficiaries
 - Other government and private offices in Ethiopia

Challenges and opportunities

- ***Challenges***

- Premature dataset for this particular task to apply other sequential models such as LSTM networks.
- Since the dataset is collected from Ethiopia ministry of health daily reports; whenever the internet has disconnected we suffer to collect because of lack of willingness and collaboration from stakeholders.

- ***Opportunities***

- Easily understandable domains to get insight and interpretation about the dataset without the help of domain experts or third party requests.
- The government will utilize the final result of this project to made a change for policy design for similar incidents

KPI of the project

- Interactive user interface for end users
- Easily understandable and interpretable visualization tools and techniques
- The future is not yet known, that is why we entertain based on our previous experience or evidence of data; thus, whenever the estimation is not deviating in certain extent from the actual circumstance will be considered as KPI's.

End of presentation

Thank You!

- Author profile: <https://ephremta.github.io/>
- Project sharable link: <https://github.com/ephremta/Covid-19-Ethiopia>