



Indian Institute of Technology Madras Zanzibar Campus



## Shaastra Techathon 2024 - AI ML Challenge

### Code Instructions

May 2024

## 1 Project Overview

This document serves as a README for our project submitted to the Techathon in AI/ML hosted by Shaastra IIT Madras. The project focuses on developing a machine learning model to predict COVID-19 related deaths using time-series data.

## 2 Instructions to run the code

To run the `ipynb` notebook and create the submission file (or check the metric), follow these steps:

1. Ensure that the notebook is uploaded to, and opened with Google Colab (this is what we have used, given our PC didn't possess GPU's - kindly use jupyter notebooks if computational resources more powerful than gcolab GPU are available).
2. Upload the datasets (`train_data_covid` and `test_data_covid`) into the `data` directory.
3. Install the necessary modules before importing them, if needed:

```
pandas
numpy
tensorflow
sklearn
kerastuner
```

4. Run the notebook cells sequentially for importing necessary modules and libraries, loading data, preprocessing it and performing feature engg.
5. After the initial steps, either of the following can be chosen, **as per computational resources available**:

- **Fast, low accurate model:**

choose to run the first option labelled 'Sample' model. This is followed by sample submission file creation using the model trained with sample hyper parameters.

**Note:- These values of hyper parameters aren't tuned.**

- **Slow, highly accurate model (RECOMMENDED):**

This is the actual code used to create our best version of submission. To run this, skip the two cells for sample model training and sample file submission and **run the cell labelled 'Model to find the actual best hyper parameter vales'** and subsequently the next cell to create the actual best version of the submissions.

We used bayesian optimization for tuning our model. **The original code used 50 trials.** This number can be changed as per available time/resources.

We have the options in our code because right after when we tuned our hyper parameters, our session had crashed and we couldn't retrieve the values of hyper parameters - Although luckily our submission file had been saved.

**Kindly contact us at [zda23b019@iitmz.ac.in](mailto:zda23b019@iitmz.ac.in) for further queries if at all any.**