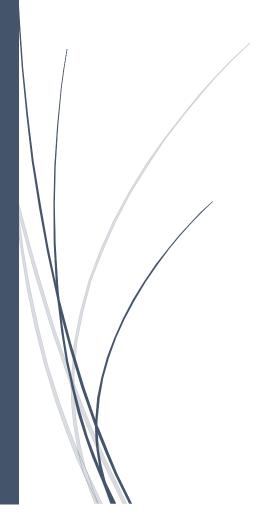# Machine Learning

Project (Final Deliverable)

Libraries Used:

1. pandas
2. numpy
3. matplotlib
4. sklearn
5. torch
6. datetime
7. os

Approach:

- Add the cases date-wise for each country to get the cases time_series for that country
- Remove the initial zeros so that we can start from the day the first confirmed case is reported
- Get the absolute difference of each entry (day-wise) for each country with Pakistan and then sum it. Put it in ascending order.
- The countries coming first will be most closely related to Pakistan (day-wise)
- Convert the cumulative time_series of all countries to new-cases-per-day
- Use sklearn's MinMaxScaler to normalize the top 5 countries most related to Pakistan
- Make a function to convert the normalized countries data to a sequence. The function will return X, y, pred_X, pred_y. X and y will be used to train our model. pred_X and pred_y will be used to predict starting from 10th May 2020 to 26th May 2020 and calculating mean squared error respectively
- Make a RNN Model using relu non-linearity.
- Make another function that keeps predicting and shifts-left the array. The values predicted are for dates from 27th May 2020 to 27th June 2020
- Plot the top 5 countries with Pakistan and the predictions using each countries' data

Model Used:

I used RNN model for the time_series prediction with the paraments:

- num_epochs = 2001
- learning_rate = 0.01
- input_dimension = 1
- hidden_dimension = 200
- num_layers = 2
- output_dimension = 1
- criterion = torch.nn.MSELoss()   # mean squared error

Results and Findings:

- Since the data we had only contained number of cases and no other feature or relation to Pakistan, using other countries data the model poorly predicted Pakistan's future
- However, we can see that the model predicted the future accurately for each country using their own data (10th May 2020 to 27th May 2020)

Prediction (using Pakistan's data)

Prediction (using Belgium's data)

Prediction (using Canada's data)

Prediction (using Romania's data)

Prediction (using Israel's data)

Prediction (using Russia's data)