Advancement Report on
Fuzzy Time Series Analysis of the Covid-19 Epidemic

Team Members:

1. Mohammad Minhazul Alam (1935365650)
2. S.M. Shahadat Hossain (1935190650)
3. Md. Romman Riyadh Shishir (2016652650)
4. Md. Abdul Motaleb (1935369650)
5. Sabrina Yeasmin (1935074650)
6. Sadman Hasan (2016160050)
7. Eumna Huda (1925395050)

Fuzzy Time Series Analysis of Covid-19: Our Initial Project Goal

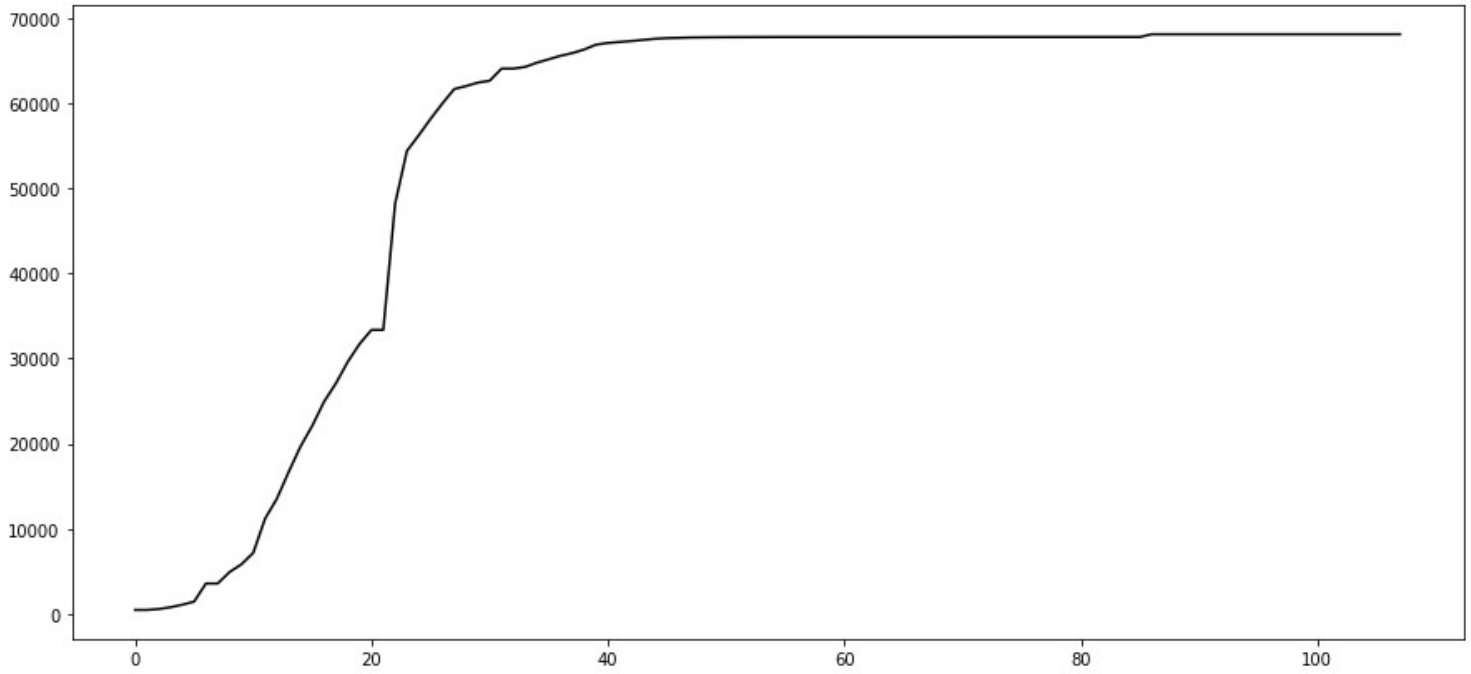
1. Fuzzy Time Series (FTS) Analysis using up-to-date dataset containing US, EU, and Asian data of the Covid-19 infection
2. Using different algorithms of FTS and comparing the result with other Machine Learning algorithms
3. Comparing the data of different countries with different demography, weather and economic status to find pattern
4. Helping the policy makers by providing with a usable prediction model

Fuzzy Time Series Analysis of Covid-19: Datasets - Introduction

1. COVID-19 Data Repository by Johns Hopkins CSSE
   - https://github.com/CSSEGISandData/COVID-19
   - Time series data of state/country wise number of confirmed cases, recovered cases and death
2. Uncover COVID-19 challenge dataset
   1. https://www.kaggle.com/roche-data-science-coalition/uncover
   2. Includes Apple & Google Mobility Trends data, Infection spread data, Test data, country health ranking, etc. of US, Canada, and some other countries
   3. Will be helpful for analyzing exogenous variables and data normalization

Fuzzy Time Series Analysis of Covid-19: Initial Data Exploration

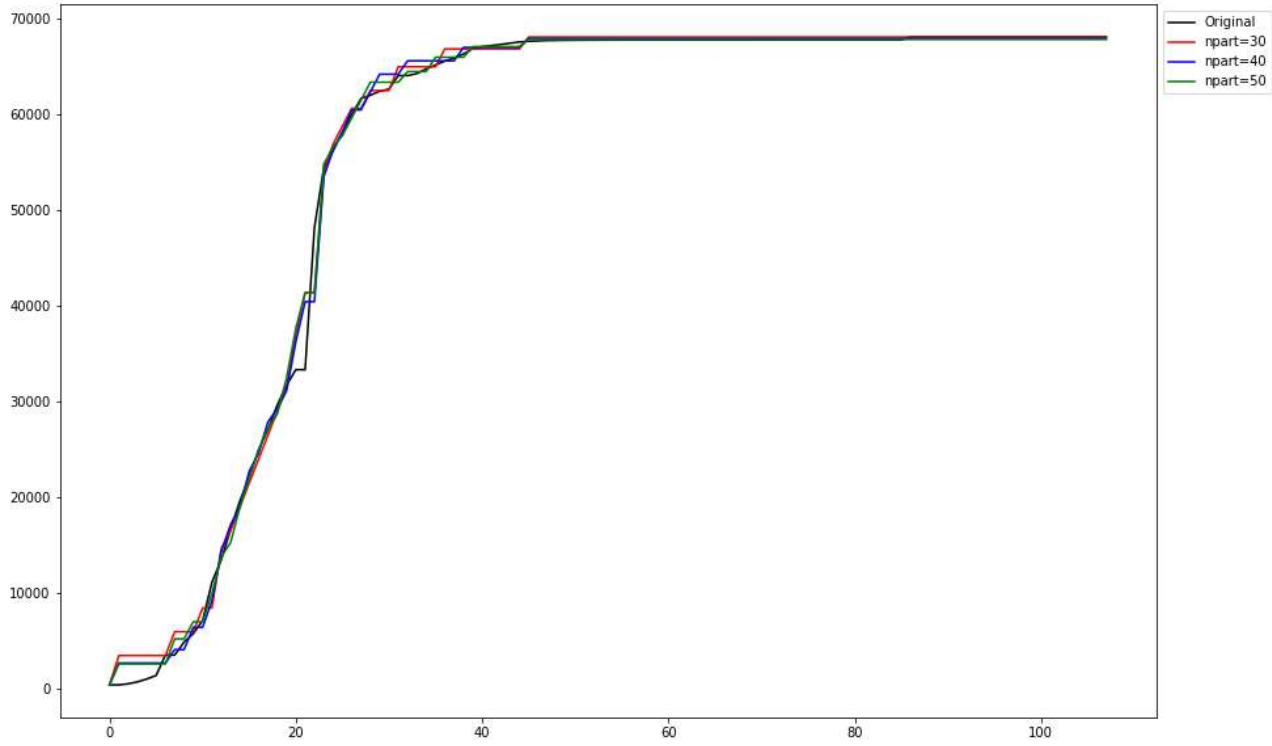
- Cumulative confirmed cases of Hubei China
- January 22 – May 8, 2020 (day wise)



Fuzzy Time Series Analysis of Covid-19: Initial Data Exploration

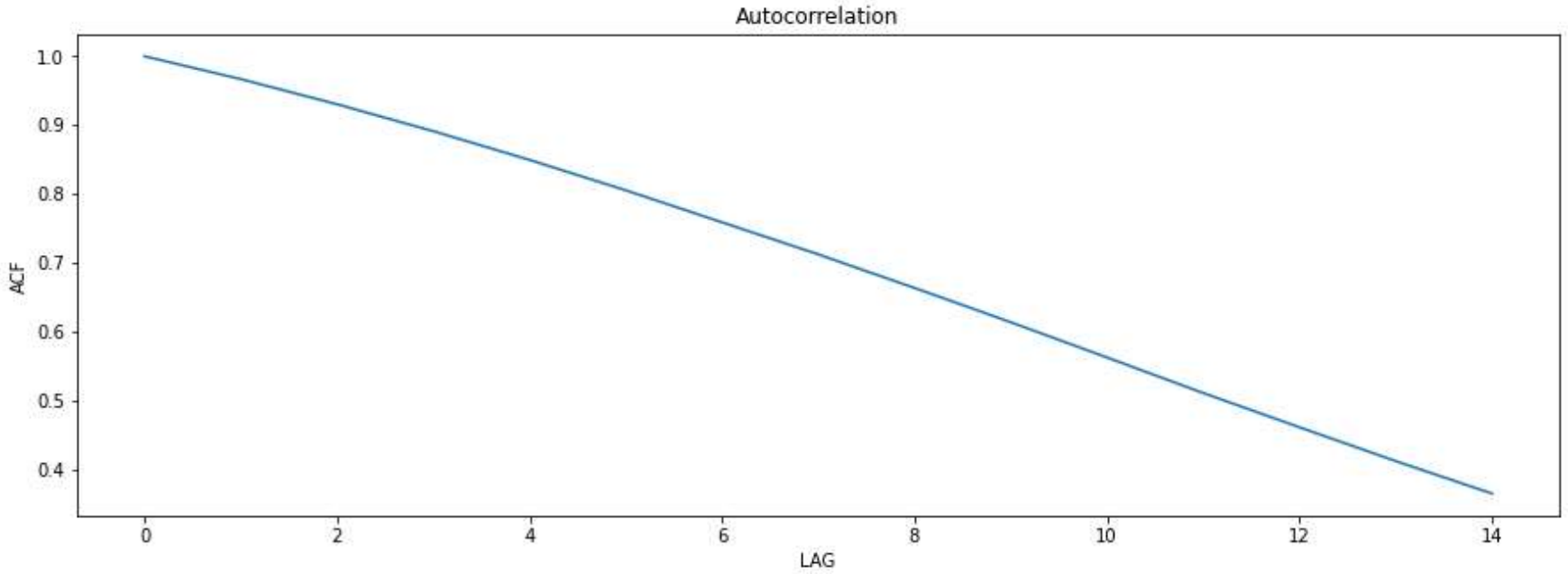
- Effect of Partition Number on FTS prediction of 1 day look ahead (Hubei, China data)

	Partitions	RMSE	MAPE	U
0	30	1346.29	20.72	0.74
1	40	1232.67	15.13	0.67
2	50	1239.25	14.64	0.68



Fuzzy Time Series Analysis of Covid-19: Initial Data Exploration

- Calculating Lag using ACF (Hubei, China data)



Fuzzy Time Series Analysis of Covid-19: Initial Data Exploration

- As seen from ACF analysis and RMSE - order is natural and more than six lags may not be necessary
- the number of data is small (per region) so we can use Probabilistic Weighted High Order FTS and let the model choose lag weights

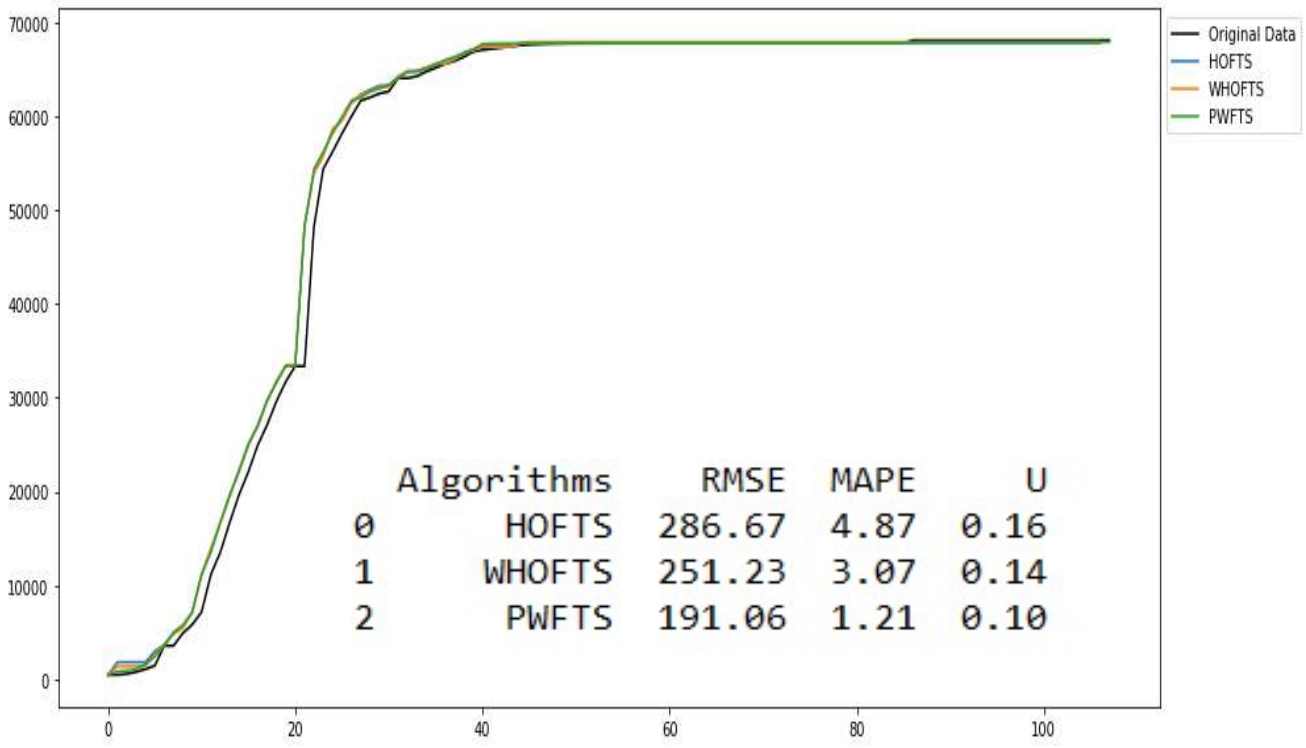
	Order	RMSE	MAPE	U
0	1	1232.67	15.13	0.67
1	2	750.30	8.88	0.41
2	3	587.67	5.57	0.32
3	4	539.62	3.46	0.29
4	5	519.55	2.19	0.28
5	6	505.62	1.12	0.27

Fuzzy Time Series Analysis of Covid-19: FTS Algorithms Used

- Algorithms used for prediction:
  - High Order Fuzzy Time Series (HOFTS)
  - Weighted High Order Fuzzy Time Series (WHOFTS)
  - Probabilistic Fuzzy Time Series (PWOFTS)
- Partition used: 40 – 80
- Order used: 2 - 6

Fuzzy Time Series Analysis of Covid-19: FTS Algorithms Used

- Result of various algorithms -



	Algorithms	RMSE	MAPE	U
0	HOFTS	286.67	4.87	0.16
1	WHOFTS	251.23	3.07	0.14
2	PWFTS	191.06	1.21	0.10

Fuzzy Time Series Analysis of Covid-19: Next Target

Dataset Normalization & Exogenous variable detection

- Different country/state has different extraneous ***conditions***
- ***Conditions*** include weather, healthcare situation, social awareness levels, population density, etc.
- Some of these variables will be used to normalize the data (both x and y axes)
- Some others will be used as Exogenous variables if enough correlation found
- Multi-variable FTS will be used to model these data and tested

Fuzzy Time Series Analysis of Covid-19: Next Target

- We shall compare the result of FTS with Machine Learning models like PNN+cf, LSTM, etc.
- Result will be compared with the Fuzzy counterparts

Fuzzy Time Series Analysis of Covid-19: Our Codebase

- https://github.com/minhazul-alam/Fuzzy_Systems

Thank You