A Presentation 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)

Time Series Analysis - Overview

"Time series is a sequence taken at successive equally spaced points in time…"

$$X(t) = \alpha 0 \cdot X(t-1) + \alpha 1 \cdot X(t-2) + \varepsilon$$

- Example: heights of ocean tides, temperature of an area, stock price of an asset, etc.

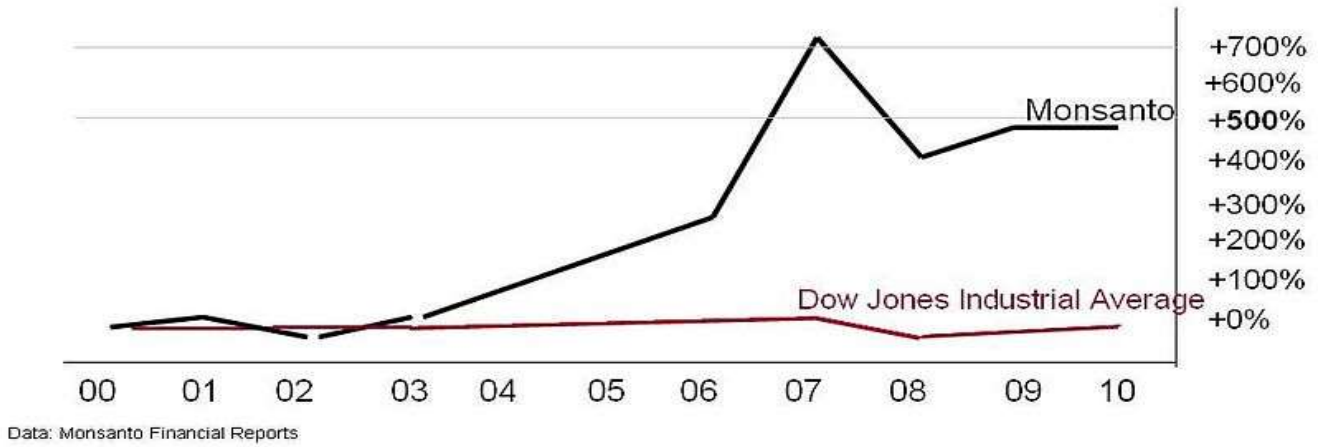

Figure: Monsanto Stock Price: 2000 - 2010

Time Series Analysis - Components

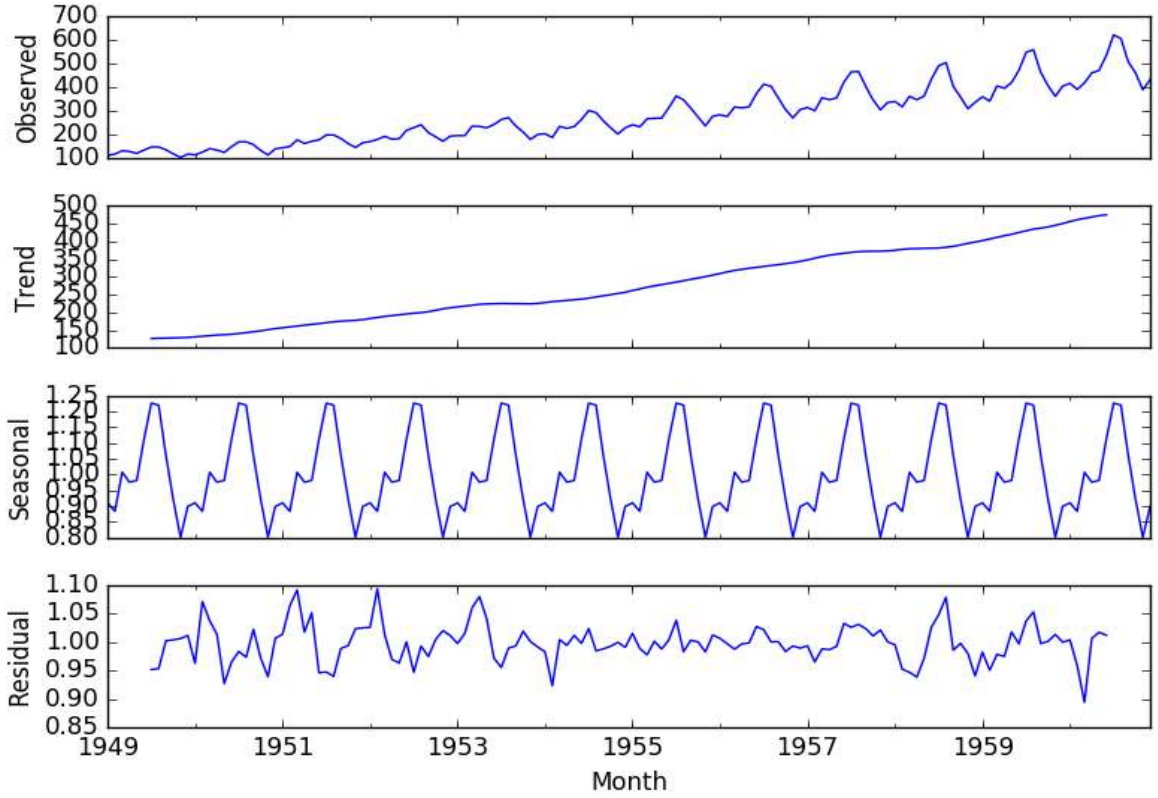

Figure: Components of a Time Series

Source: https://machinelearningmastery.com/decompose-time-series-data-trend-seasonality/

Time Series Analysis – Use

- Prediction of infectious disease spread and effectiveness of countermeasures
- Econometrics, Mathematical Finance
- Forecasting – Weather, Air quality, Temperature, etc.
- Clustering, Anomaly Detection in Machine Learning, Data Mining
- Simulation Model Analysis

Modeling Time Series – Statistically

- Moving Average
- Exponential Smoothing
- Double/Triple Exponential Smoothing
- Seasonal Autoregressive Integrate Moving Average (SARIMA)
- Available Libraries: Prophet (Python & R), Statsmodels (Python)

Modeling Time Series – Fuzzy approach

- Divide the Universe of Discourse from Time Series into partitions (Fuzzy sets)
- Extract if-then rules from Time Series patterns
- Compute new value/prediction
- Defuzzify the result
- Available Libraries: PyFTS (Python)

Modeling Time Series – Fuzzy approach (contd.)

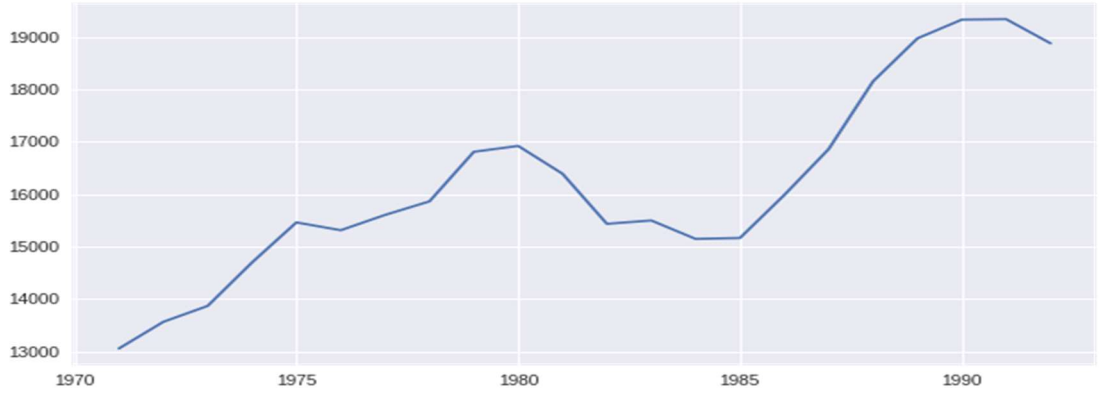

Figure: example time series data

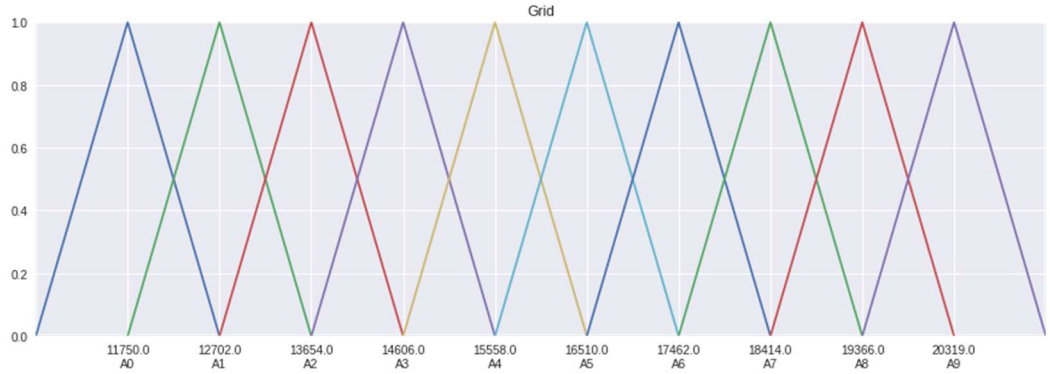

Figure: Fuzzy Sets

Modeling Time Series – Fuzzy approach (contd.)

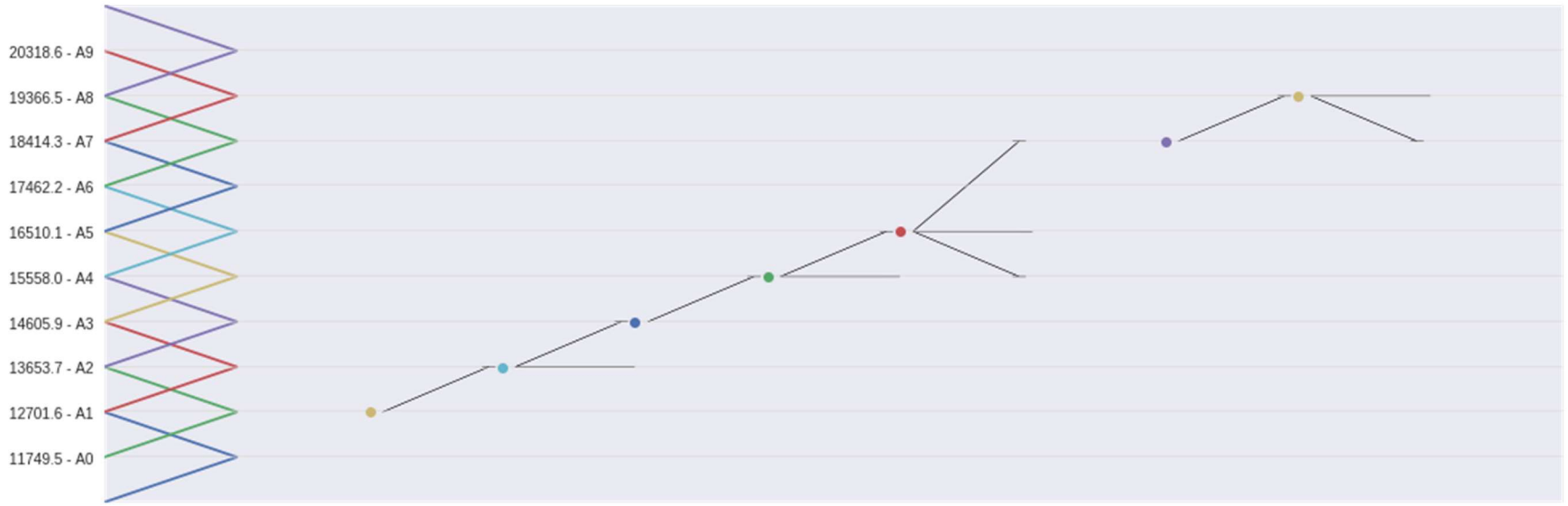

Figure: Generated If-Then rules

Images generated using PyFTS

Fuzzy Time Series Analysis – Parameters

- Number of Fuzzy Sets
  - Affects under-fitting and over-fitting
- Partition type: Grid, Entropy, Cluster
- Membership function
  - Triangular, Gaussian, Trapezoidal, etc.
- Order of the function
  - Number of past values needed to predict the next value
  - Index of the past values
  - Usually calculated using Auto Correlation Function (ACF)
- Methods:
  - Weighted vs. Weightless
  - Mono-variate vs. Multi-variate

Fuzzy Time Series Analysis of Covid-19: Related Publication

"Finding an Accurate Early Forecasting Model from Small Dataset: A Case of 2019-nCoV Novel Coronavirus Outbreak"

- Published – Feb. 7, 2020 on the International Journal of Interactive Multimedia and Artificial Intelligence
- Authors -
  - Simon James Fong - Department of Computer and Information Science, University of Macau, Macau SAR (China)
  - Gloria Li - DACC Laboratory, Zhuhai Institutes of Advanced Technology of the Chinese Academy of Sciences (China)
  - Nilanjan Dey - Department of Information Technology, Techno India College of Technology (India)
  - Rubén González Crespo - Universidad Internacional de La Rioja, Logroño (Spain)
  - Enrique Herrera-Viedma - University of Granada (Spain)

Finding an Accurate Early Forecasting Model from Small Dataset: A Case of 2019-nCoV Novel Coronavirus Outbreak

- Proposed a novel methodology for applying data mining algorithms to a limited data scenario applicable at the start of a new epidemic
  - Group of Optimized and Multi Source Selection (GROOMS)
- Algorithms used for forecasting:
  - A number of Machine Learning models
  - A statistical analysis technique – ARIMA
  - Some simple Data Analytics techniques
  - Used Polynomial Neural Network with Corrective Feedback (PNN+*cf*)
- Compared the result of different algorithms

Finding an Accurate Early Forecasting Model from Small Dataset: A Case of 2019-nCoV Novel Coronavirus Outbreak

- Source of Data:
  - Archive of Chinese Health Authority
  - 21st Jan – 3rd February

Finding an Accurate Early Forecasting Model from Small Dataset: A Case of 2019-nCoV Novel Coronavirus Outbreak

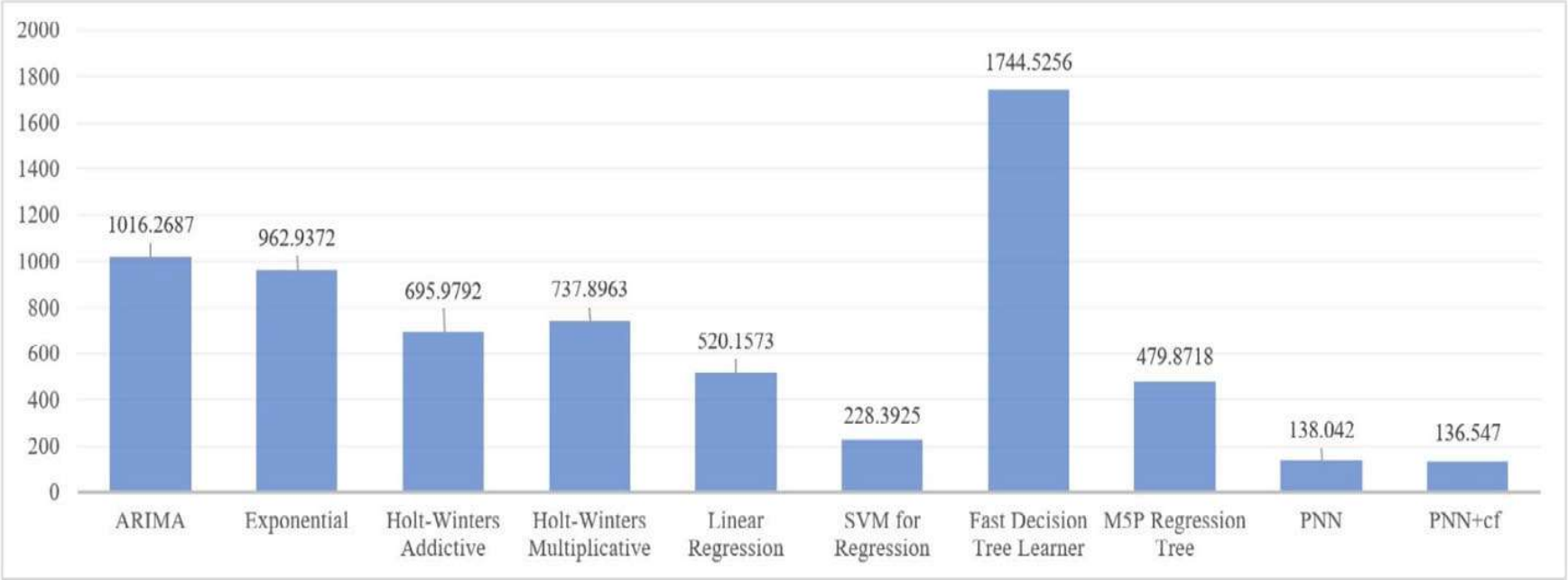

Figure: RMSE comparison of different algorithms

Finding an Accurate Early Forecasting Model from Small Dataset: A Case of 2019-nCoV Novel Coronavirus Outbreak

- Outcome of the paper:
  - A new data augmentation method called GROOMS is proposed
  - Forecasting done with several Machine Learning, Data Mining and Statistical techniques
  - PNN+*cf* is found to have the lowest RMSE for the used dataset

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

1. Use of more up-to-date and larger dataset containing US, EU, and Asian data of Covid-19 infection
2. Using different variations of Fuzzy Time Series (FTS) technique and comparing the results with the other algorithms
3. Comparing the outcome of different countries with different demography, weather and economic status

Fuzzy Time Series Analysis of Covid-19: Usability of the result

- Better understanding of the Covid-19 infection spread from up-to-date data
- Demography and weather effects on the infection, if any
- Will help the governments better prepare for the future condition

Thank You