Daily Cumulative Confirmed Cases Prediction of Covid-19 in Italy

Xiaofo Geng
Anqi Xiong
Lulu Zha

Contents

01

Background

🧰 Background

December 2019, following Wuhan reports of a cluster of cases of new viral pneumonia, this epidemic of a novel coronavirus illness has become a global threat. 11th February, the World Health Organization(WHO) named it as COVID-19.

02

Dataset

🩺 Dataset

The original data set used in this project is provided by DingXiangYuan, which is collected by COVID-19 Infection Data Real-time Crawler. By plotting some graphs, we could get a general knowledge of the data set.
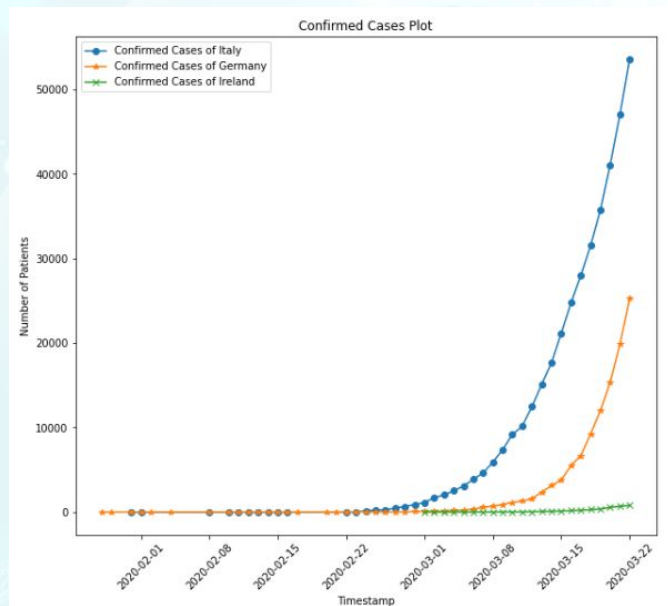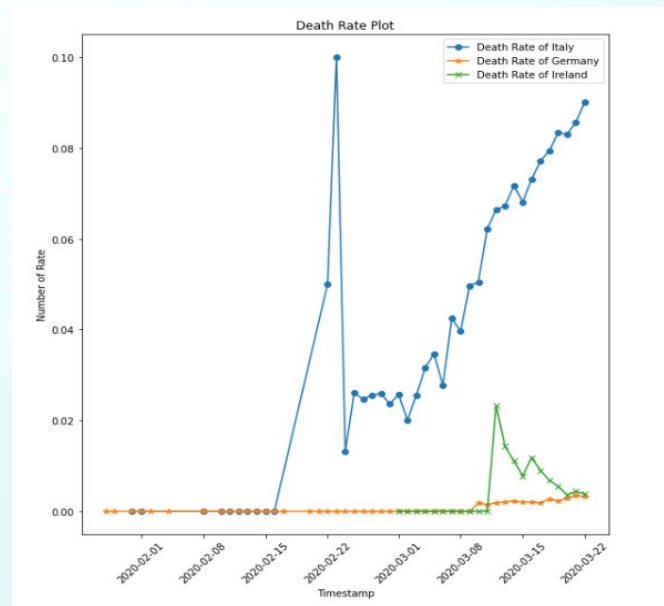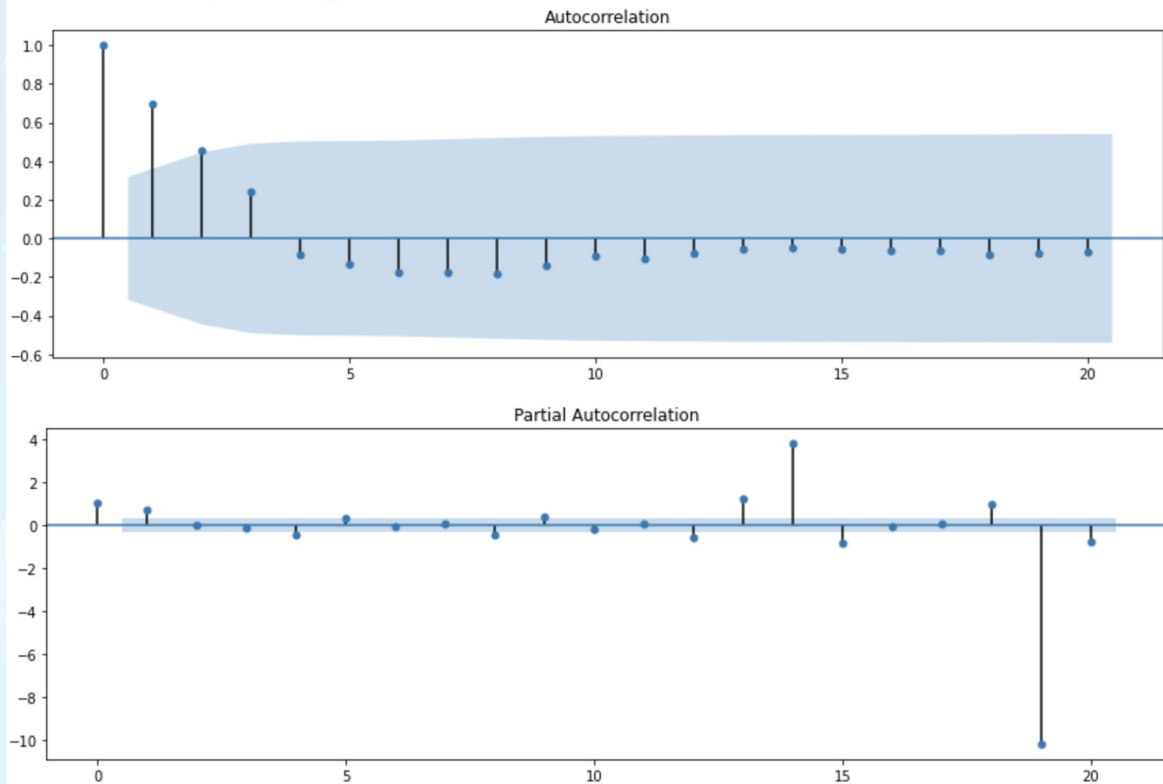

Fig.1

Fig.2

03

Approch

📷 Model Selection

- ARIMA was chosen in this task

- ARIMA(1,2,2) were estimated by ACF and PACF

- AIC and BIC were used to help model selection

🧰 Model Evaluation - Residuals

- No autocorrelation between the residual values (Figure 1)

- The mean of the residual is 0.0117 - very close to 0
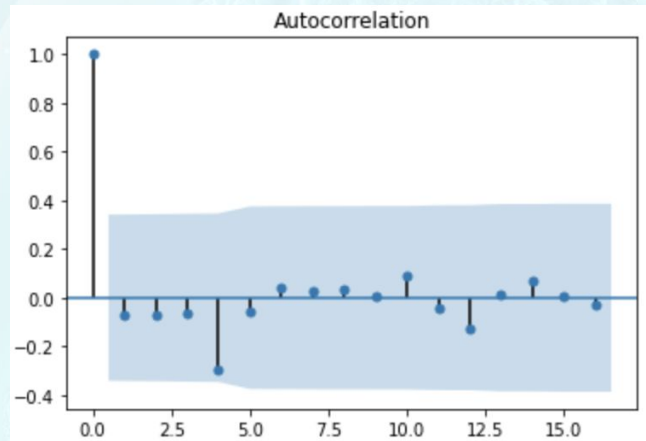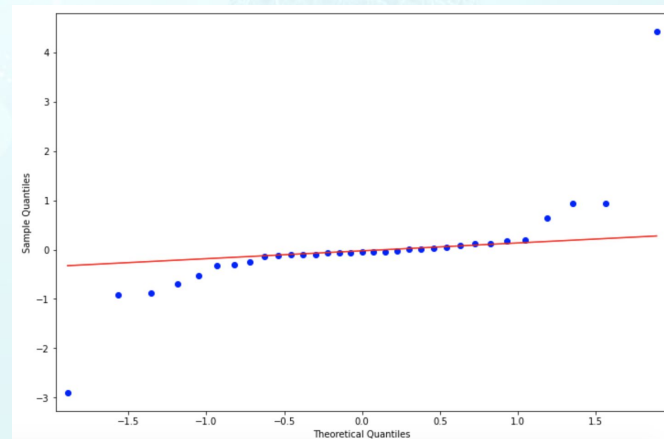
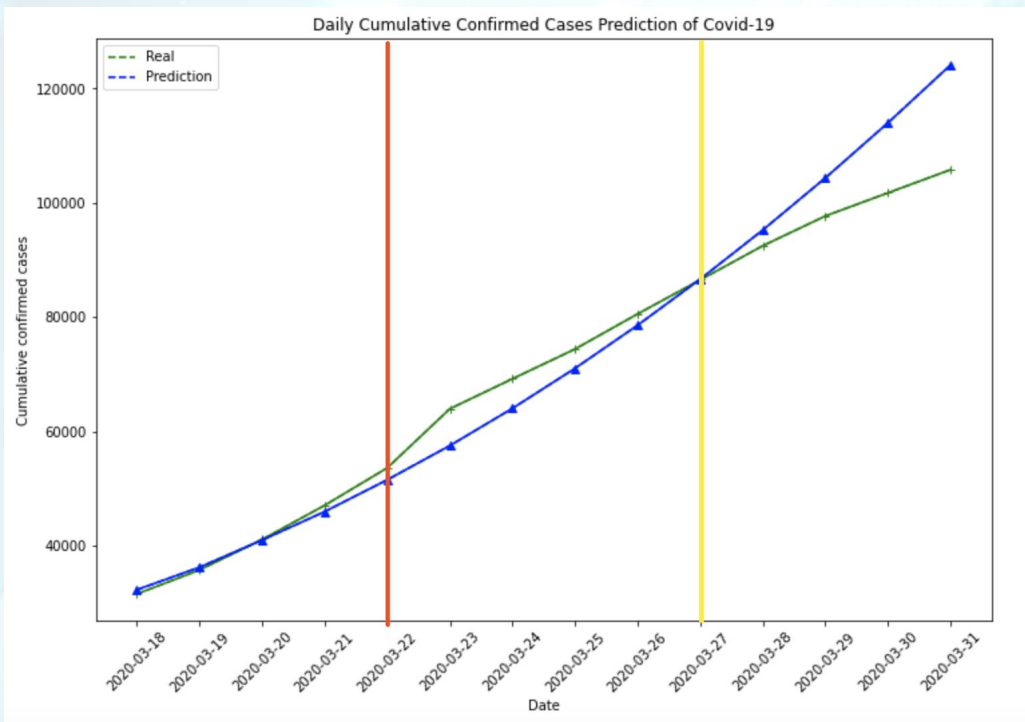- Normally distributed (Figure 2)

Figure 1

Figure 2

04

Results

🏥 Result

Daily Cumulative Confirmed Cases Prediction of Covid-19

- Outperformed in testset

- Performed well within 5 days

- Deviation gradually becomes large out of 5 days

05

Limitations

Limitations

- Limited by the size of dataset

- More external objective factors should be take into account (lockdown, capacity of healthcare resources, etc...)

Members	tasks
Anqi, Xiaofo	Data pre-processing
Anqi, Xiaofo, Lulu	Data visualization
Anqi, Lulu	ARIMA build and result analysis
Xiaofo	Random Forest model build but failed

Question?