Clustering of Covid-19 Time Series

Wintersemester 2020/2021

Idea

1. Consider the topic of clustering of time series with the application to Covid-19 cases.

Idea

1. Consider the topic of clustering of time series with the application to Covid-19 cases.
2. The daily cases of each country give rise to a time series of Covid-19 cases.

Idea

1. Consider the topic of clustering of time series with the application to Covid-19 cases.
2. The daily cases of each country give rise to a time series of Covid-19 cases.
3. Use clustering algorithms for time-series to find clusters in Covid-19 data by countries and timespans.

Idea

1. Consider the topic of clustering of time series with the application to Covid-19 cases.
2. The daily cases of each country give rise to a time series of Covid-19 cases.
3. Use clustering algorithms for time-series to find clusters in Covid-19 data by countries and timespans.
4. Predict number of future cases using cluster analysis of the results.

Tasks

- Read Covid-19 data into usable data structure [ECD24]

Tasks

- Read Covid-19 data into usable data structure [ECD24]
- Visualize data (e.g. plotly dash)

Tasks

- Read Covid-19 data into usable data structure [ECD24]
- Visualize data (e.g. plotly dash)
- Data preparation

Tasks

- Read Covid-19 data into usable data structure [ECD24]
- Visualize data (e.g. plotly dash)
- Data preparation
  - find and remove outliers (no data at weekend, etc. delayed sending)

Tasks

- Read Covid-19 data into usable data structure [ECD24]
- Visualize data (e.g. plotly dash)
- Data preparation
  - find and remove outliers (no data at weekend, etc. delayed sending)
  - smooth time series (1d convolution, kernels, etc.)

Tasks

- Read Covid-19 data into usable data structure [ECD24]
- Visualize data (e.g. plotly dash)
- Data preparation
  - find and remove outliers (no data at weekend, etc. delayed sending)
  - smooth time series (1d convolution, kernels, etc.)
- Choose and compare different clustering algorithms for time series (see [AAJX19] and tslearn for examples). Does the data preparation help? Which metrics are useful two compare such two time series? Discuss and interpret the results.

Tasks

- Read Covid-19 data into usable data structure [ECD24]
- Visualize data (e.g. plotly dash)
- Data preparation
  - find and remove outliers (no data at weekend, etc. delayed sending)
  - smooth time series (1d convolution, kernels, etc.)
- Choose and compare different clustering algorithms for time series (see [AAJX19] and tslearn for examples). Does the data preparation help? Which metrics are useful two compare such two time series? Discuss and interpret the results.
- Cluster between countries but also different timespans of the same country (use timespans of different lengths).

Tasks

- Read Covid-19 data into usable data structure [ECD24]
- Visualize data (e.g. plotly dash)
- Data preparation
  - find and remove outliers (no data at weekend, etc. delayed sending)
  - smooth time series (1d convolution, kernels, etc.)
- Choose and compare different clustering algorithms for time series (see [AAJX19] and tslearn for examples). Does the data preparation help? Which metrics are useful two compare such two time series? Discuss and interpret the results.
- Cluster between countries but also different timespans of the same country (use timespans of different lengths).
- Compare, analyze and discuss different countries, different trends of infection numbers by their clustering results.

Tasks

- Read Covid-19 data into usable data structure [ECD24]
- Visualize data (e.g. plotly dash)
- Data preparation
  - find and remove outliers (no data at weekend, etc. delayed sending)
  - smooth time series (1d convolution, kernels, etc.)
- Choose and compare different clustering algorithms for time series (see [AAJX19] and tslearn for examples). Does the data preparation help? Which metrics are useful two compare such two time series? Discuss and interpret the results.
- Cluster between countries but also different timespans of the same country (use timespans of different lengths).
- Compare, analyze and discuss different countries, different trends of infection numbers by their clustering results.
- Try to predict the number of future cases using the cluster analysis

Mohammed Ali, Ali Alqahtani, Mark W. Jones, and Xianghua Xie.
Clustering and classification for time series data in visual analytics: A survey.
IEEE Access, 7:181314–181338, 2019.

ECDC.
Daily covid-19 cases.
https://www.ecdc.europa.eu/en/publications-data/download-todays-data-geographic-distribution-covid-19-cases-worldwide, 2020-09-24.