

TIME SERIES FORECASTING

1. Time series forecasting

00:00 - 00:06

Great job so far. Now, we'll learn about modeling and time series forecasting.

2. Modeling in data science

00:06 - 00:26

Data scientists and machine learning scientists spend a lot of time building models. Models attempt to represent a real-world process with statistics. At a high level, models define relationships between variables with equations. These definitions are based on statistical assumptions and historical data.

3. Predictive modeling

00:26 - 00:35

Predictive modeling is a sub-category of modeling used for prediction. By modeling a process, we can enter new inputs and see what outcome it outputs.

4. Predictive modeling

00:35 - 00:44

For instance, you can enter a future date in a model of unemployment rate to get a prediction of what the unemployment rate will be next month.

5. Predictive modeling

00:44 - 00:52

The output can be the probability of an outcome, for example, the probability that a tweet is fake.

6. Predictive modeling

00:52 - 01:05

Predictive models can be as simple as a linear equation with an x and y variable to a deep learning algorithm that is uninterpretable by humans. Let's look at using predictive modeling on time series data.

7. Time series data

01:05 - 01:29

Time series is a series of data points sequenced by time. Examples include daily stock and gas prices over the years. Often times, it's in the forms of rates, such monthly unemployment rates or a patient's heart rate during surgery. They can be measurements like CO2 levels or the height of tides recorded regularly over a time period.

8. Plotting time series data

01:29 - 01:45

Let's plot an example. We have time series data of Canadian unemployment rates measured monthly from 1976 to 2015. Time series data is usually plotted as a line graph like this, with time on the x-axis.

9. Seasonality in time series

01:45 - 02:00

Often times when plotting time series, you can find patterns. For example, this plot graphs the average temperature in Boston over three years. Can you figure out the pattern here?

10. Seasonality in time series

02:00 - 02:23

The line peaks during summer months and reaches its lowest during winter months. If we graphed ice cream sales, we'd see a similar pattern. This is called seasonality. Seasonality is when there are repeating patterns related to time such as months and weeks. Another example is spending spikes at the end of the month when people receive a paycheck.

11. Forecasting time series

02:23 - 02:55

Time series data is used in predictive modeling to predict metrics at future dates. We call this forecasting. For example, predicting rainfall next month or the state of traffic and the stock market in a couple of hours to what the population will be in 20 years. We can build predictive models using time series data from the past years or decades to generate predictions. This uses a combination of statistical and machine learning methods. Let's look at an example.

12. Pea prices in Rwanda

02:55 - 03:27

The United Nation provides open data on global food prices. Here we have the price of peas in Rwandan Francs from 2011 and 2016. There's some seasonality here, can you spot it? Prices are lowest around December and January, but peak around August. Some years show a second peak around April. There seems to be a general increase in pea prices annually. Can we forecast what will happen with consideration to this seasonality and price increase?

13. Forecasting pea prices in Rwanda

03:27 - 04:15

Here is the forecast of a predictive model. The blue line depicts the forecast. The seasonality remains and it anticipates a continued increase of pea prices, seen by the higher peaks and lows. There are also two blue areas shown along the forecast. These are confidence intervals, which are extremely useful for evaluating predictions. We see two confidence intervals: 80%

and 95%. The model is 80% sure that the true value will be in the area labeled as 80. Same goes with the area labeled as 95. If we're using this forecast to make big decisions, confidence intervals can help us buffer for the unexpected.

14. Let's practice!

04:15 - 04:20

Ok, it's time for some exercises!