https://www.udemy.com/course/machinelearning/learn/lecture/19346436#questions/11801136

Back to All Questions

Feature Scaling

1

Nikitha · Lecture 104 · 4 years ago

I dint understand why we don't require feature scaling for Regression algorithms? Can someone please explain this?

5 replies Follow replies

Yogeshwar 2

4 years ago

we always require feature scaling...

its just that some classes like DecisionTreeRegressor already have the mechanism built into it for feature scaling

while others like svr dont

Akhil 3

4 years ago

feature scaling is done when there is a large marginal difference between two of your columns. Suppose in X , in col 1 there is a cust id of 18753246246 and in col 2 there is age 19 . Here there is large diff b/n two column values . if you scale them their values come into the range of -1 to 1 .

Nikitha 0

4 years ago

Thanks alot

Course content

Section 1: Welcome to the course! Here we will help you get started in the best...
6 / 6 | 19min

Section 2: ---------- Part 1: Data Preprocessing ----------
4 / 4 | 10min

Section 3: Data Preprocessing in Python
24 / 24 | 1hr 32min

Section 4: Data Preprocessing in R
11 / 11 | 42min

Section 5: ---------- Part 2: Regression ----------
1 / 1 | 1min

Section 6: Simple Linear Regression
17 / 17 | 1hr 12min

Section 7: Multiple Linear Regression
18 / 26 | 2hr 16min

Section 8: Polynomial Regression

4 years ago

Thanks alot

SuperDataScience — Instructor

0

4 years ago

Hi, I'm Mango - I'm a bot, and I'm sorry i'm not able to answer that question as i'm still learning. One of our TA's will respond shortly to assist and please feel free to follow up with us!

Tin

12

4 years ago

Hi nischita,

In other regression models there are coefficients which compensate the difference in scaling between variables. You don't need feature scaling because It's the nature of Linear or Multiple regression themselves, not because the mechanism does it automatically.

SVR algorithm doesn't include coefficient so you need to implement feature scaling to eliminate the difference manually.

And the example of Cust ID vs Cust age is also not appropriate because basically data mining models have nothing to do with Cust ID. It's not an variable contributing to the model and no data analyst will include Cust ID in their model.

Thanks

H0 Add reply

Course content

Section 1: Welcome to the course! Here we will help you get started in the best...
6 / 6 | 19min

Section 2: ---------- Part 1: Data Preprocessing ----------
4 / 4 | 10min

Section 3: Data Preprocessing in Python
24 / 24 | 1hr 32min

Section 4: Data Preprocessing in R
11 / 11 | 42min

Section 5: ---------- Part 2: Regression ----------
1 / 1 | 1min

Section 6: Simple Linear Regression
17 / 17 | 1hr 12min

Section 7: Multiple Linear Regression
18 / 26 | 2hr 16min

Section 8: Polynomial Regression