Machine Learning Course | Arabic

Data Preprocessing

Level - 01

Link to Lecture on Youtube

Terminologies

Features / Independent variables

Target / Dependent variable

Samples / Points

	weight	height	drinks alcohol	healthy
0	112	181	0	0
1	123	165	1	1
2	176	167	1	1
3	145	154	1	1
4	198	181	0	0
5	211	202	1	0
6	145	201	1	1
7	181	153	1	1
8	90	142	0	1
9	101	169	1	1

Introduction

What are the problems that the data include?

- Missing Values.
- Duplicates.

- String, computers aren't work with strings.
- Outliers.
- Non useful features.
- Highly correlation features.

	weight	height	drinks alcohol	healthy
0	112	181	0	0
1	123	165	1	1
2	176	167	1	1
3	145	154	1	1
4	198	181	0	0
5	211	202	1	0
6	145	201	1	1
7	181	153	1	1
8	90	142	0	1
9	101	169	1	1

Introduction

What is data preprocessing?

Data preprocessing is the process of **handling** and **cleaning up** all the previously mentioned issues.

Introduction

Why descriptive statistics is important?

- It can summarize the data and give us an intuition about what the data is talking about.
- No need to read data line by line to understand it.

Table of contents

Table of contents

SLIDESMANIA.COM

Thank You!

Do you have any questions?

Write them in the comments

hozaifazaki99@gmail.com