Aula M2A38 DATA PREPARATION AND FEATURE SELECTION.

Leitura complementar:

- Data Preparation: What is Data Preparation for Machine Learning?
- Ordinal and One-Hot Encodings for Categorical Data
- Ways To Handle Categorical Data With Implementation
- Categorical Variables
- Data Preparation and Feature Engineering in ML
- What Is Data Preparation in a Machine Learning Project
- What's Data Science Pipeline?
- scikit-learn Machine Learning in Python
- What is Data Preparation?
- The Definitive Guide for Data Preparation that Beginners should read
- How to Prepare your Data
- Six Steps to Master Machine Learning with Data Preparation
- 4 Distance Measures for Machine Learning
- Importance of Distance Metrics in Machine Learning Modelling
- Tutorial: Linear Regression with Stochastic Gradient Descent
- How to Use StandardScaler and MinMaxScaler Transforms in Python
- How to Save and Reuse Data Preparation Objects in Scikit-Learn
- Preprocessing with sklearn: a complete and comprehensive guide
- Python Sklearn Data Preprocessing for Data Science
- 7 Steps to Mastering Data Preparation for Machine Learning with Python — 2019 Edition
- How to Transform Data to Better Fit The Normal Distribution
- How To Prepare Your Data For Machine Learning in Python with Scikit-Learn
- Compare the effect of different scalers on data with outliers
- sklearn.preprocessing.MinMaxScaler

- `.fit()`
- `.transform()`
- What and why behind fit_transform() and transform() in scikit-learn!
- Feature Scaling — Effect Of Different Scikit-Learn Scalers: Deep Dive
- Compare different scalers on data with outliers
- How and why to Standardize your data: A python tutorial
- sklearn.preprocessing.StandardScaler
- Scale, Standardize, or Normalize with Scikit-Learn
- StandardScaler and Normalization with code and graph
- Feature Transformation
- sklearn.preprocessing.RobustScaler
- .transform()
- How to Scale Data With Outliers for Machine Learning
- Hands-On PCA Data Preprocessing Series. Part I: Scaling Transformers
- Why is scaling required in KNN and K-Means?
- All about Categorical Variable Encoding
- .get_dummies()
- How to One Hot Encode Sequence Data in Python
- Preprocessing: OneHotEncoder() vs pandas.get_dummies
- Machine Learning — Feature Encoding with OneHotEncoder (inside a Pipeline)
- One Hot Encoding in Scikit-Learn
- One-Hot-Encoding, Multicollinearity and the Dummy Variable Trap
- The Curse of Dimensionality
- sklearn.preprocessing.OrdinalEncoder
- Types of Categorical Data Encoding Schemes
- CountFrequencyEncoder()
- How to Encode Categorical Data
- Target Encoder

- Extending Target Encoding
- Target Encoding and Bayesian Target Encoding
- Target Encoding Vs. One-hot Encoding with Simple Examples
- How to Use the ColumnTransformer for Data Preparation
- Use ColumnTransformer in SciKit instead of LabelEncoding and OneHotEncoding for data preprocessing in Machine Learning
- sklearn.compose.ColumnTransformer
- Column Transformer with Mixed Types
- .pipeline()
- 6.3. Preprocessing data
-
    -
-