
Strive
SCHOOL

/ Multiplicative Models

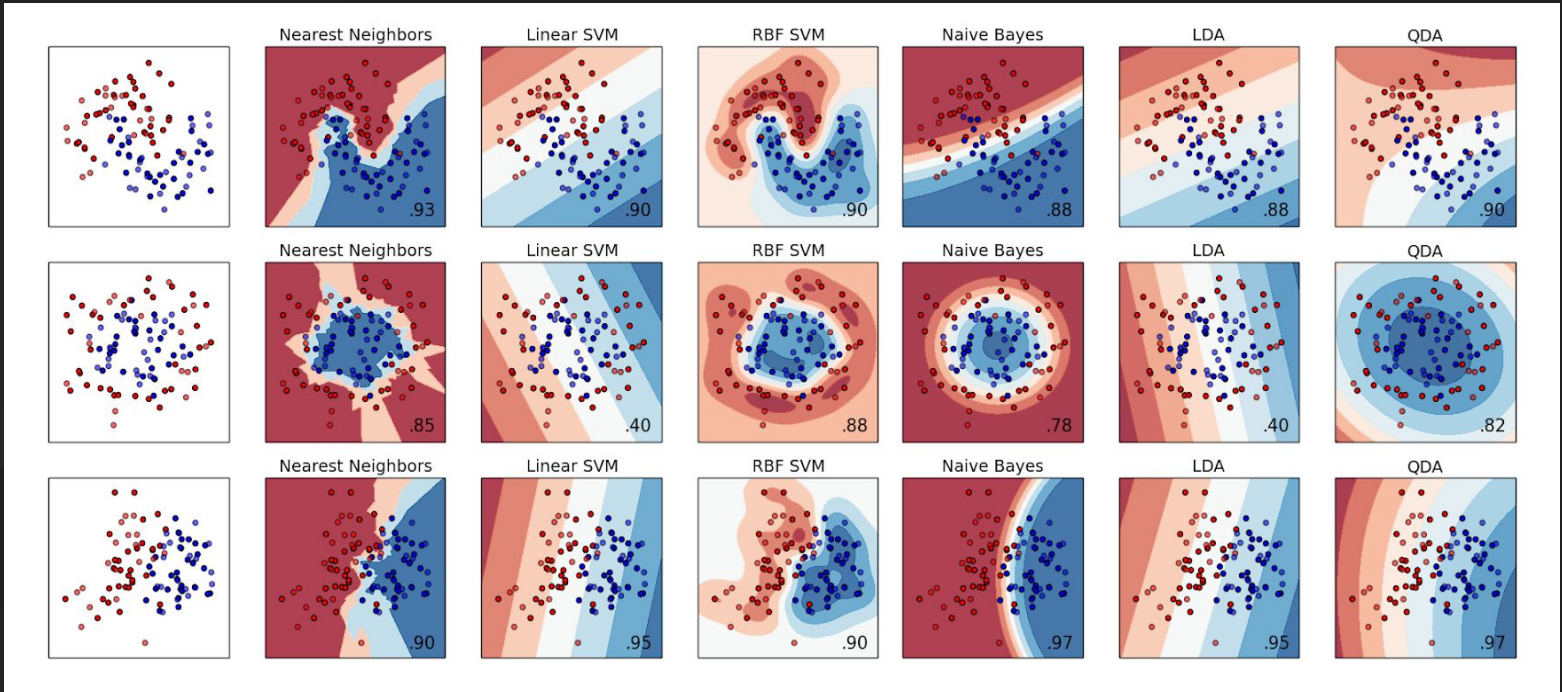
Multiplicative Models

/ This models are based on multiplications to the input variables or based in measuring distances. So all variables needs to be in the same range like a normal distribution. This models are:

- Linear Models (LM)
  - Linear Regression
  - Logistic Regression
  - Ridge
  - Lasso
- Neural Networks (NN)
  - Multi Layer Perceptron (MLP)
  - Tensorflow models
  - Pytorch models
- Generalized Additive Model (GAM)
- Support Vector Machines (SVM)
  - Linear SVM
  - SVM with RBF kernel
- Instance models
  - K-Nearest Neighbors (KNN)
- Bayesian models
  - Naive Bayes (NB)
  - Gaussian Naive Bayes
  - Multinomial Naive Bayes
- Dimensionality Reduction models
  - PCA
  - t-SNE
  - UMAP
- Clustering models
  - K-means

Multiplicative Models

/ Unlike tree models, these models create more curved boundary shapes.



/ Preprocessing

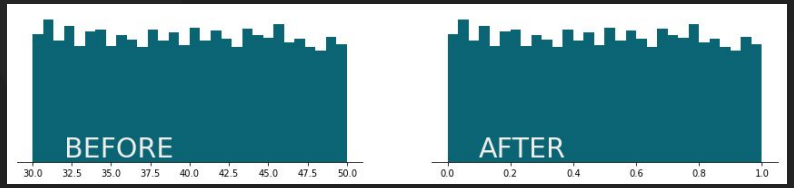
The rule of thumb is:

- Numerical variables: Normalization
- Categorical variables: One Hot Encoding

Numerical Variables

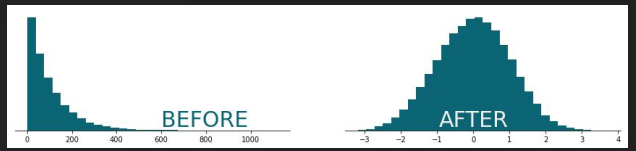
Scaling

- Min-max scaling `MinMaxScaler()`
- Max-abs scaling `MaxAbsScaler()`
- Standard scaling `StandardScaler()`
- Robust scaling `RobustScaler()`



Normalization

- Manually
  - Log
  - Sqrt
- PowerTransformer
  - Box-Cox
  - Yeo-Johnson
- QuantileTransformer
  - (aka GaussRank)



Categorical variables

Ordinal encoding

Categorical Feature		Numeric
Louise	=>	1
Gabriel	=>	2
Emma	=>	3
Adam	=>	4
Alice	=>	5
Raphael	=>	6
Chloe	=>	7
Louis	=>	8
Jeanne	=>	9
Arthur	=>	10

Binary encoding

			Binary Encoded			
Categorical Feature		=	x1	x2	x4	x8
Louise	=>	1	1	0	0	0
Gabriel	=>	2	0	1	0	0
Emma	=>	3	1	1	0	0
Adam	=>	4	0	0	1	0
Alice	=>	5	1	0	1	0
Raphael	=>	6	0	1	1	0
Chloe	=>	7	1	1	1	0
Louis	=>	8	0	0	0	1
Jeanne	=>	9	1	0	0	1
Arthur	=>	10	0	1	0	1

One Hot Encoding (Best option)

Categorical Feature		f1	f2	f3	f4	f5	f6	f7	f8	f9	f10
Louise	=>	1	0	0	0	0	0	0	0	0	0
Gabriel	=>	0	1	0	0	0	0	0	0	0	0
Emma	=>	0	0	1	0	0	0	0	0	0	0
Adam	=>	0	0	0	1	0	0	0	0	0	0
Alice	=>	0	0	0	0	1	0	0	0	0	0
Raphael	=>	0	0	0	0	0	1	0	0	0	0
Chloe	=>	0	0	0	0	0	0	1	0	0	0
Louis	=>	0	0	0	0	0	0	0	1	0	0
Jeanne	=>	0	0	0	0	0	0	0	0	1	0
Arthur	=>	0	0	0	0	0	0	0	0	0	1

Unlike Tree models, One Hot Encoding is usually one of the best encodings for multiplicative models.

/ Q&A

What are your doubts?


Strive
SCHOOL