Indice

Elenco delle figure

Capitolo 1

Data Understanding

1.1 Data understanding

1.2 Data semantics

1.3 Distribution of the variables and statistics

1.4 Assessing data quality

1.5 Variable transformations

1.6 Pairwise correlations and eventual elimination od redundant variable

Capitolo 2

Clustering analysis

2.1 Clustering Analysis by K-means

2.1.1 Choice of attributes and distance function

2.1.2 Identification of the best value of k

2.1.3 Characterization of the obtained clusters by using both analysis of the k centroids and comparison of the distribution of variables within the clusters and that in the whole dataset

Capitolo 3

Association Rules

Capitolo 4

Classification