About the transformation - categ\_mapping: converts a categorical column into binary variables (one‑hot). Can use n columns or n-1 columns.

# Categorical mapping  
# A categorical attribute with $n$ distinct values can be mapped into $n$ binary (one‑hot) attributes.  
  
# It is also possible to map into $n-1$ binary attributes: the case where all binary attributes are zero represents the last categorical value (not explicit in columns).  
  
# installation   
#install.packages("daltoolbox")  
  
# loading DAL  
library(daltoolbox)

Apply one-hot mapping to the Species column in a data frame.

# dataset for the example   
  
iris <- datasets::iris  
head(iris)

## Sepal.Length Sepal.Width Petal.Length Petal.Width Species  
## 1 5.1 3.5 1.4 0.2 setosa  
## 2 4.9 3.0 1.4 0.2 setosa  
## 3 4.7 3.2 1.3 0.2 setosa  
## 4 4.6 3.1 1.5 0.2 setosa  
## 5 5.0 3.6 1.4 0.2 setosa  
## 6 5.4 3.9 1.7 0.4 setosa

# creating the categorical mapping  
  
cm <- categ\_mapping("Species")  
iris\_cm <- transform(cm, iris)  
print(head(iris\_cm))

## Speciessetosa Speciesversicolor Speciesvirginica  
## 1 1 0 0  
## 2 1 0 0  
## 3 1 0 0  
## 4 1 0 0  
## 5 1 0 0  
## 6 1 0 0

# creating the categorical mapping  
# It can be done from a single column, but it must be a data frame  
  
diris <- iris[,"Species", drop=FALSE]  
head(diris)

## Species  
## 1 setosa  
## 2 setosa  
## 3 setosa  
## 4 setosa  
## 5 setosa  
## 6 setosa

iris\_cm <- transform(cm, diris)  
print(head(iris\_cm))

## Speciessetosa Speciesversicolor Speciesvirginica  
## 1 1 0 0  
## 2 1 0 0  
## 3 1 0 0  
## 4 1 0 0  
## 5 1 0 0  
## 6 1 0 0

References - Bishop, C. M. (2006). Pattern Recognition and Machine Learning. Springer. (1-of-K/one-hot encoding)