

Machine Learning - 1100-MLOENG (Ćwiczenia informatyczne Z-23/24)

[Home](#) > [My courses](#) > [Machine Learning - 1100-MLOENG \(Ćwiczenia informatyczne Z-23/24\)](#) > [Ensemble learning - boosting](#) > [Dummy coding](#)

Dummy coding

A dummy variable is either 1 or 0 and 1 can be represented as either True or False and 0 can be represented as False or True depending upon the user. This variable is used to categorize the characteristic of an observation. For example, a person is either male or female, discipline is either good or bad, etc.

In the caret package we have the function `dummyVars`, which allows **factor variables to be converted into a numerical type**.

```
library(caret)
ex<- data.frame(num = 1:3,
  gender = c("m", "m", "f"),
  animal = c("cat", "dog", "fish"),
  dates = as.Date(c("2023-01-01", "2022-12-31", "2021-01-01")))
d <- dummyVars("~ .", ex, fullRank = T)
str(d)
d2 = as.data.frame(predict(d, newdata=ex))
names(d2)
d2
```

Attention should be paid to the role of the parameter **fullRank**. For fullrank=T we have one variable less than the number of existing categories so as to avoid correlated variables, but also to save space. On the other hand, for fullrank=F we get as many columns as the number of variables in the factors.

```
dd <- dummyVars("~ .", ex, fullRank = F)
dd
dd2 = as.data.frame(predict(dd, newdata=ex))
names(dd2)
dd2
```

Last modified: środa, 17 stycznia 2024, 9:31

Accessibility settings

Przetwarzanie danych osobowych

Platformą administruje Komisja ds. Doskonalenia Dydaktyki wraz z Centrum Informatyki Uniwersytetu Łódzkiego [Więcej](#)

Informacje na temat logowania

Na platformie jest wykorzystywana metoda logowania za pośrednictwem [Centralnego Systemu Logowania](#).

Studentów i pracowników Uniwersytetu Łódzkiego obowiązuje

Deklaracja dostępności

nazwa użytkownika i hasło
wykorzystywane podczas logowania
się do systemu USOSweb.