

Motoric 8.11.

Presentation

- Hot Deck Imputation ?
 - split data into specific columns (e.g. gender)
- feed model and predict for boys
- feed same model and predict for girls

From Jeroen:

- For false negatives / model
 - use Logistic Regression
 - change Threshold
 - calculate recall score
 - Plot it

To Do's:

- clean Data

- convert Birthdate

- "normalize" Format (15.12.1992 / 15-12-1992)

- to date

- imputation

- use a model

- use smote as analysis

- evaluation "plan"

- models

- feed with cleaned data

- feed with imputed data

- median

- mean

- model

- data understanding

- plot missing data

- every column

- Which missing data type do we have?

- MCAR

- MAR

- MNAR