

PREDICTING POWERLIFTING

DIA
ROBERT REIMANN
TIIT VAINO

CLEANING
THE
DATA

37 COLUMNS
12 COLUMNS

1.42M ROWS
365K ROWS

MINED DATA
REGARDING EVERY
PERSON'S
COMPETITION
HISTORY

THE PROJECT IS BASED ON A
PUBLIC KAGGLE POWERLIFTING
DATASET. THE DATASET
CONTAINS INFO ABOUT A
PERSON'S CHARACTERISTICS AND
COMPETITION RESULT

SELECTING THE
APPROPRIATE
MODEL

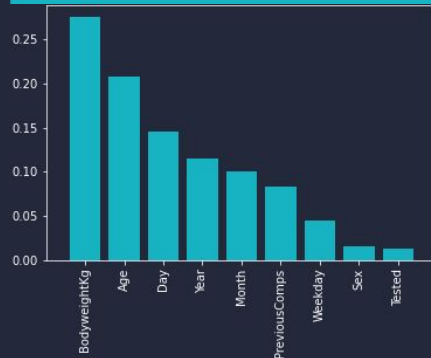
5 DIFFERENT
CLASSIFICATION
MODELS WERE
SUCCESSFUL

RANDOM FOREST
AND EXTRA TREES
CLASSIFIERS WERE
TOP PERFORMERS

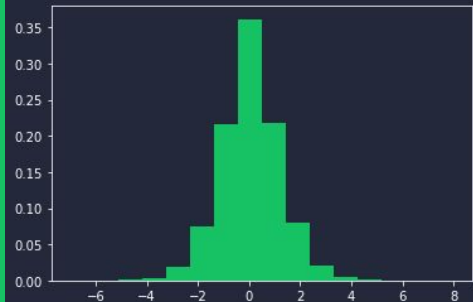
THE MAIN GOAL OF THIS PROJECT
WAS TO TRAIN A MODEL WHICH
COULD PREDICT A PERSON'S
FUTURE COMPETITION RESULT
BASED ON INFO THAT COULD BE
FOUND BEFORE THE COMPETITION

RESULTS

BODYWEIGHT AND AGE WERE THE
BEST PREDICTORS FOR THE
ATHLETE'S PERFORMANCE



ALGORITHM PERFORMANCE



PREDICTION ERROR