MINI PROJECT CANVAS

Title (preliminary): BayesMeetsFloorball

Group members: Olli Haapasalo

Workshop # : SELF

MOTIVATION

The target group of my project is floorball fans. Possibly also coaches and players itself.

The goal is to better understand the possibilities of data-analysis in floorball.

DATA COLLECTION

Data is from www.fliiga.com. Data is obtained by using web scraping.

PREPROCESSING

The goal of the preprocessing pipeline is to make data tidy for analysis and visualization. This includes formatting, naming, uniting.

EXPLORATORY DATA ANALYSIS (EDA)

Look at the data!

First, some calculation is needed, for example average, mean and mode.

The distributions for shooting percentages and shots per game are probably useful too.

VISUALIZATIONS

The visualizations are related to shooting percentages and number of shots. Some re-classifications is probably needed. Histograms for top players are the goal.

LEARNING TASK

(focus on problem definition)

The main idea is to apply Bayes statistics to estimate "real" shooting percentages. This is relevant when data is scarce.

The secondary goal is to use data to predict the best goal scorers.

LEARNING APPROACH

(focus on solution implementation)

We use "Empirical Bayes" to evaluate general shooting percentages in the first phase. The result after combining Beta (prior) and Binomial (likelihood) is Beta (posterior) because of the conjugate prior attribute.

For the predicting we use the Binomial model.

COMMUNICATION OF RESULTS

I will use a blog post or website for communication.

DATA PRIVACY AND ETHICAL CONSIDERATIONS

(if applicable)

There is no need for anonymization nor pseudonymization.

All Data is public without personal information.

ADDED VALUE

The added knowledge from floorball statistics.

LEGEND

WEEK 1: Data collection/preprocessing

WEEK 2: EDA & visualizations

WEEKS 3-4: Predicting $ visualizations