

TEAM FC:

Transfer Recommendations with Machine Learning

by David VanHeeswijk

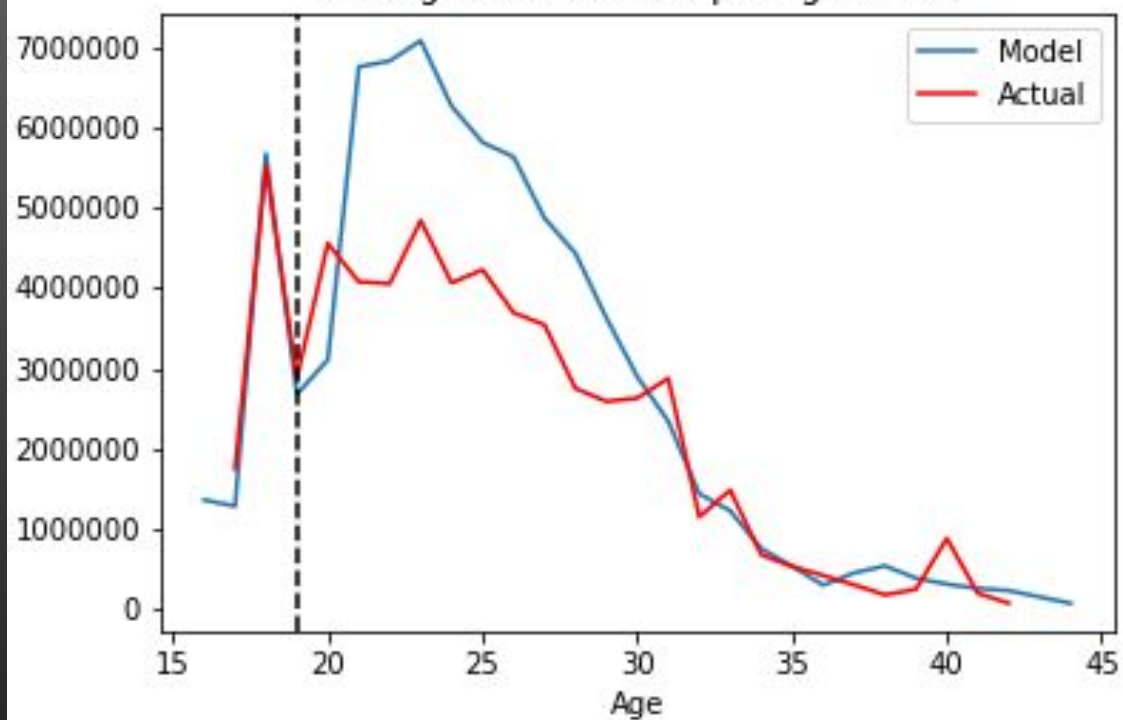
Club Goals for Transfer Policy

- 26 years or younger only
- Limited Transfer Budget
 - \$40 Million Max
- At most 1 Non European player
 - Look to maximize sell on value

What leagues and ages are best to target scouting?

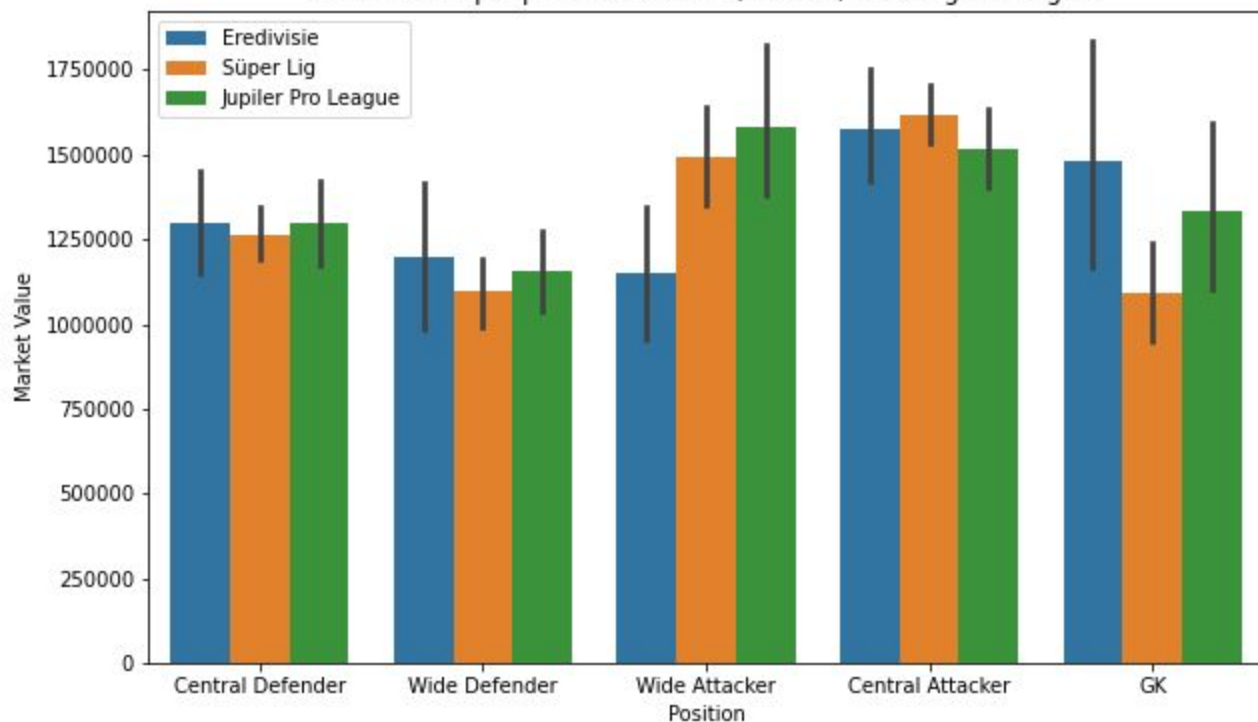
19 and 20 year old players are cheaper
than 18 year olds and 21+

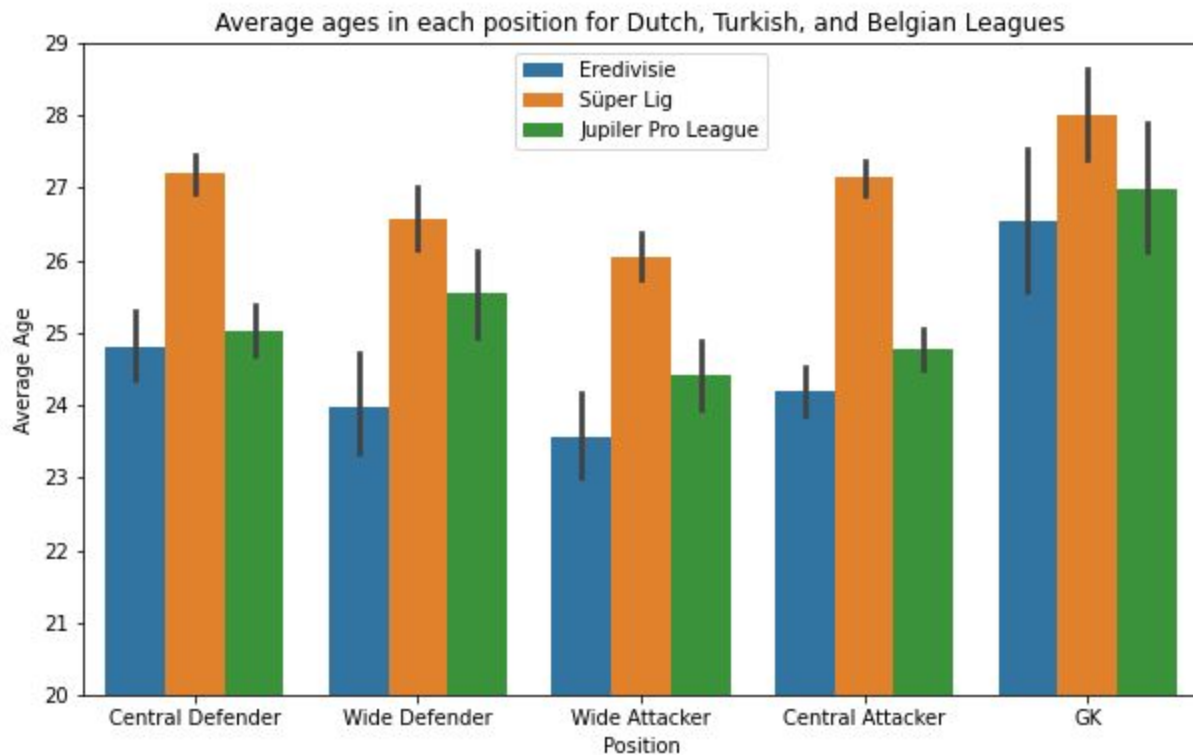
Average market value per age: 2018

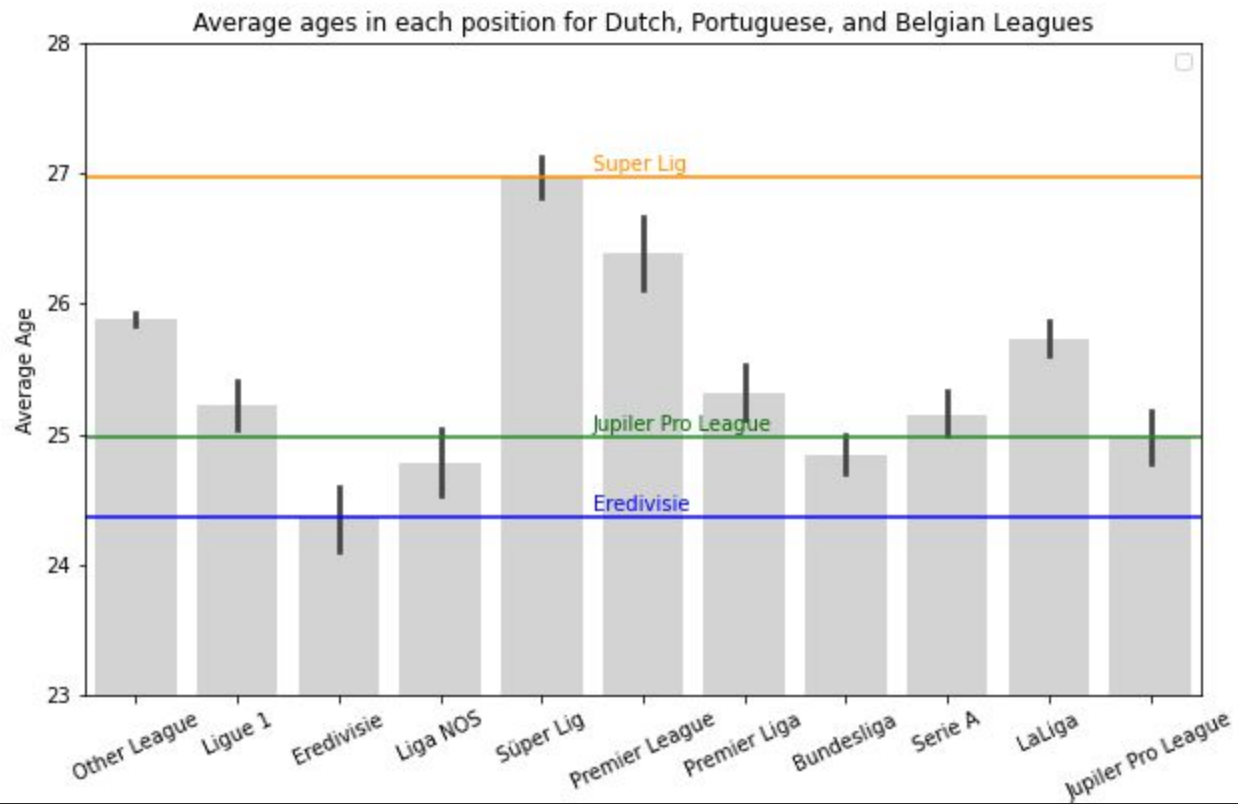


Eredivisie and Jupiler Leagues produce
best players for value

TIER 3: Value per position for Dutch, Turkish, and Belgian Leagues





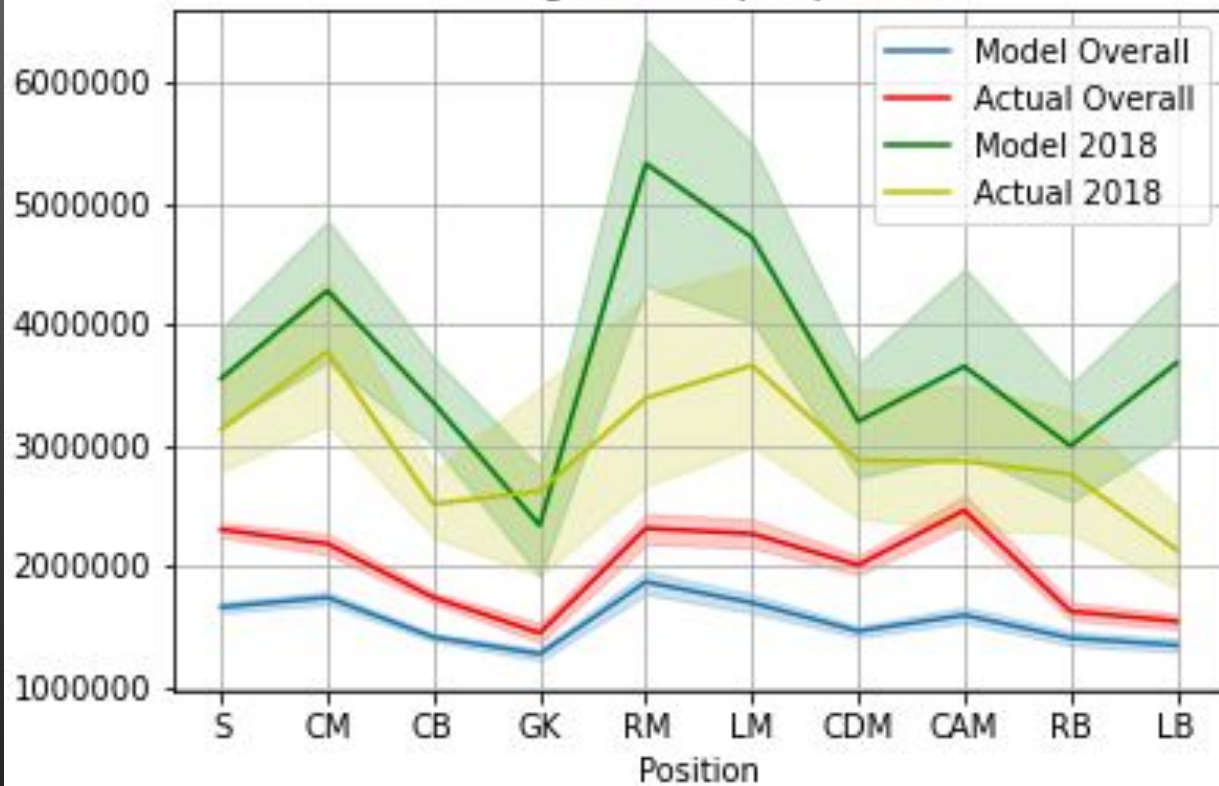


What is the cheapest position to target players within the top league?

Attacking minded right
wing backs

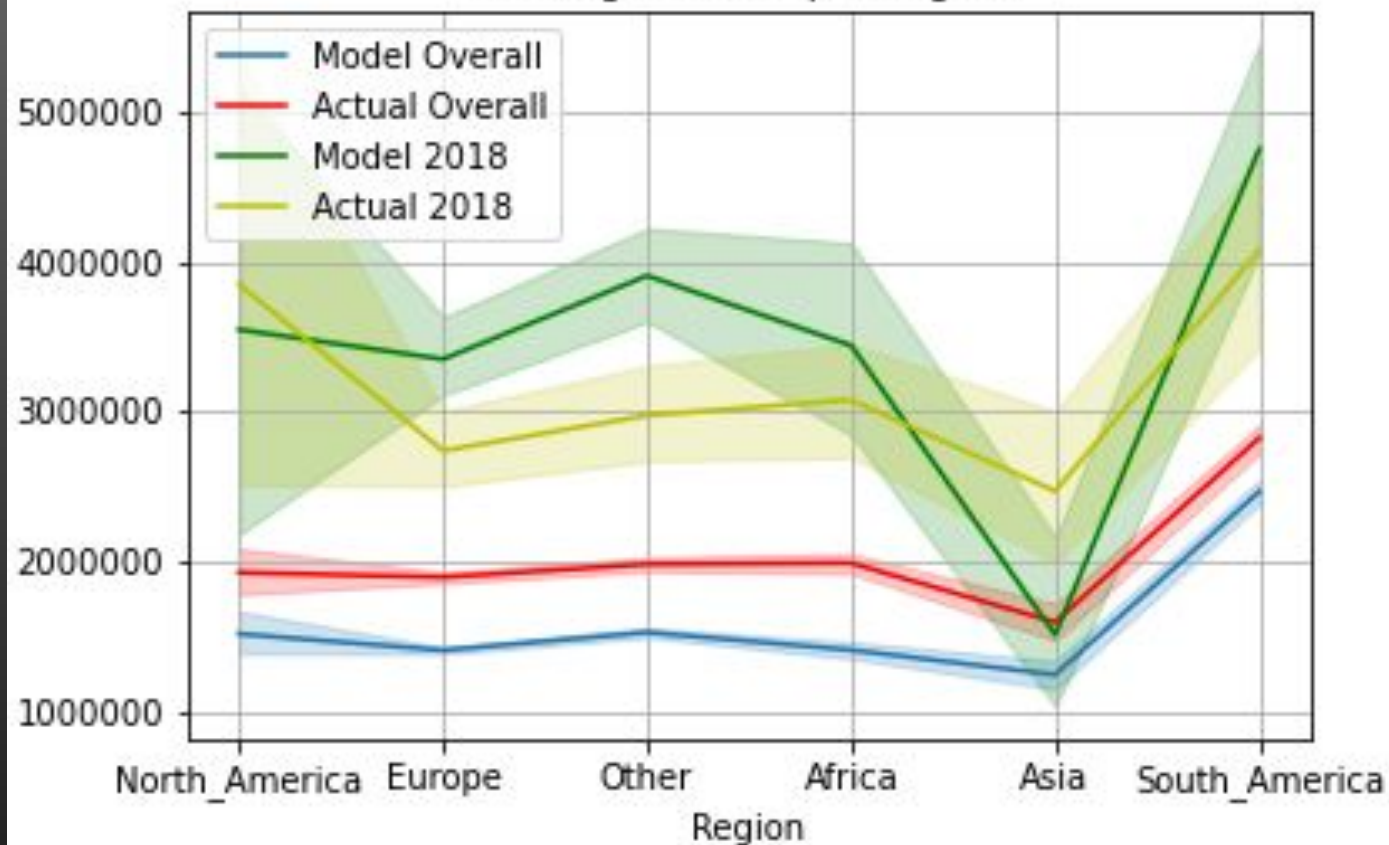
Attacking minded center
defensive midfielders

Average values per position



What region should we target outside of Europe to maximize transfer profit?

Average values per region



Support for our findings

DATA

[European Football Transfers Dataset | Kaggle](#)

- **Transfers**
- **Player Stats (simple)**
- **Player Info**
- **Club Info**
- **League Info**

Features Chosen

Player Info

- Age (in a given season)
- Nationality
- Height
- Field Position
- *Market Value*

Club or League Info

- Fees (if player was transferred)
- League of player (broken into tiers)
- Season of play

Player Stats

- Goal Contributions (G + A)
- Total Minutes (Season)
- Minutes per game

Model Selection

Trained and Tested 3 models

- Random Forest Regressor
- Gradient Boosting Regressor
- K Means Regressor

Also Trained and tested with
modified market values (discrete)

| Model | R2 Score |
|-------------------------------|-----------------|
| Random Forest Continuous | 0.725304 |
| Gradient Boost | 0.682265 |
| K Means Regressor | 0.394440 |
| <i>Random Forest Discrete</i> | <i>0.722457</i> |

Model Hyperparameters

| Parameter name | Parameter Value |
|---------------------|-----------------|
| 'n_estimators' | 800 |
| 'min_samples_split' | 2 |
| 'min_samples_leaf' | 1 |
| 'max_features' | 'auto' |
| 'max_depth' | 100 |
| 'bootstrap' | True |

Future Improvements

Add more features with player Stats

Reduce to only current data

Include features for non playing info that affects value

Use cloud based computing to assist in the hyperparameter optimization