Euro 2016 Predictions Using Team Rating Systems

Jan Lasek

deepsense.io

Machine Learning and Data Mining for Sports Analytics Workshop at ECML PKDD 2016

Introduction

There were two challenges within the Euro 2016 prediction competition

- the match prediction challenge and
- the tournament elimination challenge.

Estimated probabilities for the first challenge were used to generate predictions for the second one.

Match outcome prediction

via team rating systems

(Not only) my approach:

1. estimate team ratings based on historical match data and
2. use them to predict future match outcomes.

$$\textit{Data} \rightarrow \textit{Ratings} \rightarrow \textit{Predictions}$$

Match outcome prediction

via team rating systems

(Not only) my approach:

1. estimate team ratings based on historical match data and
2. use them to predict future match outcomes.

$$Data \rightarrow Ratings \rightarrow Predictions$$

Three rating models were employed:

Match outcome prediction

via team rating systems

(Not only) my approach:

1. estimate team ratings based on historical match data and
2. use them to predict future match outcomes.

$$\textit{Data} \rightarrow \textit{Ratings} \rightarrow \textit{Predictions}$$

Three rating models were employed:

- the ordinal logistic regression model,

Match outcome prediction

via team rating systems

(Not only) my approach:

1. estimate team ratings based on historical match data and
2. use them to predict future match outcomes.

$$Data \rightarrow Ratings \rightarrow Predictions$$

Three rating models were employed:

- the ordinal logistic regression model,
- the Poisson model and

Match outcome prediction

via team rating systems

(Not only) my approach:

1. estimate team ratings based on historical match data and
2. use them to predict future match outcomes.

$$\textit{Data} \rightarrow \textit{Ratings} \rightarrow \textit{Predictions}$$

Three rating models were employed:

- the ordinal logistic regression model,
- the Poisson model and
- the least squares model.

Match outcome prediction

via team rating systems

(Not only) my approach:

1. estimate team ratings based on historical match data and
2. use them to predict future match outcomes.

$$\textit{Data} \rightarrow \textit{Ratings} \rightarrow \textit{Predictions}$$

Three rating models were employed:

- the ordinal logistic regression model,
- the Poisson model and
- the least squares model.

They were combined into an ensemble model.

Match outcome prediction

via team rating systems

(Not only) my approach:

1. estimate team ratings based on historical match data and
2. use them to predict future match outcomes.

$$\textit{Data} \rightarrow \textit{Ratings} \rightarrow \textit{Predictions}$$

Three rating models were employed:

- the ordinal logistic regression model,
- the Poisson model and
- the least squares model.

They were combined into an ensemble model.
The data used were:

- `http://laenderspiel.cmuck.de/` - special **thanks to Christian Muck** for cordially exporting the data
- betting odds from `http://betexplorer.com/`

Ordinal logistic regression model (1)

Under this model, match outcomes - H (home team win), D (draw) and A (away team win) - are linked to team ratings via the following equations

$$\begin{aligned}
\mathbb{P}(H) &= \frac{1}{1+e^{c-(r_i-r_j+h)}}, \\
\mathbb{P}(D) &= \frac{1}{1+e^{-c-(r_i-r_j+h)}} - \frac{1}{1+e^{c-(r_i-r_j+h)}}, \\
\mathbb{P}(A) &= 1 - \frac{1}{1+e^{-c-(r_i-r_j+h)}},
\end{aligned}$$

where $h > 0$ is a parameter accounting for the home team advantage and $c > 0$ in an intercept which governs the draw margin.

Ordinal logistic regression model (2)

Model fitting: the weighted maximum likelihood method with regularization was used:

$$-L(\mathcal{M}|\mathbf{r}, h, c) + \lambda \cdot \left(\frac{1}{2}(1-\gamma)\|\mathbf{r}\|_2^2 + \gamma\|\mathbf{r}\|_1 \right),$$

where $\mathcal{M}$ is a dataset of matches and the likelihood function has a form

Ordinal logistic regression model (2)

Model fitting: the weighted maximum likelihood method with regularization was used:

$$-L(\mathcal{M}|\mathbf{r}, h, c) + \lambda \cdot \left(\frac{1}{2}(1-\gamma)\|\mathbf{r}\|_2^2 + \gamma\|\mathbf{r}\|_1 \right),$$

where $\mathcal{M}$ is a dataset of matches and the likelihood function has a form

$$L(\mathcal{M}|\mathbf{r}, h, c) = \frac{1}{|\mathcal{M}|} \sum_{m \in \mathcal{M}} \phi(m) \cdot \log \mathbb{P}(o_m),$$

where:

- $\mathbb{P}(o_m)$ equal to the probability of the actual outcome of a match m attributed by the model and
- $\phi(m)$ being a weighting function depending both on time and match type (e.g., friendly game or World Cup finals match).

Poisson model (1)

The assumption here is that the goals scored by a team can be modelled as a Poisson distributed variable.

Given the attacking and defensive skills (model's parameters) of teams i and j, a_i, a_j and d_i, d_j, respectively, the rates of Poisson variables for a home team i and visiting team j, λ and μ respectively, are modelled as:

$$\lambda = c + h + a_i - d_j,$$

$$\mu = c + a_j - d_i.$$

Poisson model (2)

Under this model, the probability of a score x to y is a product of two individual Poisson variables with rates λ and μ respectively and equal to

$$\frac{\lambda^x \cdot e^{-\lambda}}{x!} \cdot \frac{\mu^y \cdot e^{-\mu}}{y!}.$$

The model's parameters are estimated using the weighted maximum likelihood method with regularization.

Least squares model

The least squares model assumes that the difference $s_i - s_j$ in the scores produced by the teams corresponds to the difference in their ratings

$$s_i - s_j = r_i - r_j + h.$$

Again, h is a correction for the home team advantage.

Tuning the predictive power (1)

In the competition, the accuracy was evaluated using logarithmic loss (*logloss*)

$$\frac{1}{m}\sum_{i=1}^{m}\log \mathbb{P}(o_m).$$

The parameters of the ratings systems are optimized for

- World Cup finals held between 1994 and 2010 (5 tournaments),
- UEFA European Championships 1996-2008 (4) and
- Copa America finals 1999-2011 (5).

This amounts for a set of 562 matches.

Tuning the predictive power (2)

Finally, the predictions are evaluated against 2014 World Cup finals, 2012 UEFA European Championships and 2015 Copa America.

Table : Evaluation of the final test set (112 matches).

Method	**Logloss**	**Accuracy**
Bookmakers	0.9726	52%
Ensemble	0.9950	56%
Least squares	0.9985	55%
Poisson	0.9991	55%
Ordinal regression	1.0002	52%
FIFA Women World Rankings	1.0060	50%
EloRatings.net	1.0189	51%
Random guess	1.0986	33%

Challenge I - Match outcome prediction

The final submission was an ensemble of the three discussed models obtained by averaging. In the contest the solution yielded 1.0776 logloss and 41% accuracy.

The probabilities generated for the first challenge were used for simulating tournament outcome 1.000.000 times in a Monte Carlo experiment. Based on the simulations, the probabilities of advancing a given stage were estimated.

Challenge II - Tournament elimination

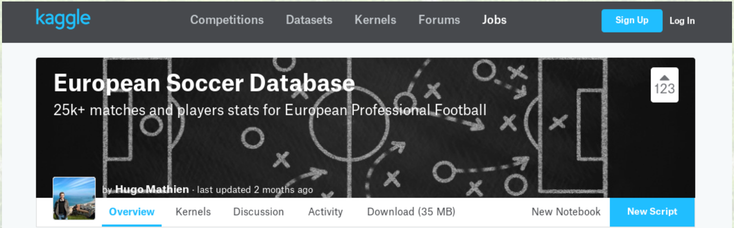
Table : Estimated probabilities of advancing past a given stage.

Team	Group stage	Quarterfinal	Semifinal	Final	Champions
France	98.01%	82.6%	67.71%	51.21%	37.55%
Spain	92.60%	72.24%	51.11%	33.95%	19.08%
Germany	94.71%	70.41%	45.99%	24.88%	13.21%
England	93.52%	67.5%	40.87%	22.25%	10.40%
Belgium	84.38%	48.2%	26.10%	11.51%	4.55%
Portugal	**91.37%**	**54.70%**	**26.31%**	**12.09%**	**4.42%**
Italy	72.43%	33.38%	14.83%	5.26%	1.55%
Ukraine	76.81%	37.05%	15.5%	5.53%	1.52%
Croatia	66.00%	31.92%	14.65%	5.27%	1.50%
Russia	75.34%	37.84%	13.07%	4.29%	1.14%
Turkey	61.90%	27.97%	12.07%	4.00%	1.05%
Switzerland	69.98%	30.49%	11.80%	3.97%	0.88%
Poland	67.40%	26.58%	9.35%	2.77%	0.60%
Sweden	57.89%	20.76%	7.45%	2.11%	0.47%
Romania	62.64%	23.82%	8.07%	2.35%	0.45%
Austria	71.63%	27.01%	7.46%	2.07%	0.43%
Slovakia	63.66%	25.57%	6.96%	1.79%	0.37%
Republic of Ireland	54.68%	18.64%	6.38%	1.72%	0.35%
Czech Republic	46.28%	16.19%	5.60%	1.44%	0.29%
Hungary	56.86%	16.08%	3.37%	0.69%	0.11%
Iceland	47.81%	11.32%	2.02%	0.36%	0.05%
Albania	31.46%	6.62%	1.26%	0.19%	0.02%
Wales	34.29%	7.98%	1.19%	0.16%	0.02%
Northern Ireland	28.32%	5.11%	0.88%	0.13%	0.01%

Can we do better?

How to obtain a model with a better predictive power?

- apply methods for improving a model efficacy, e.g., bagging
- use more data on, for example, the players and their skills
- ...



https://www.kaggle.com/hugomathien/soccer

That is all!

Many thanks to

- The organizers for hosting such an exciting competition

That is all!

Many thanks to

- The organizers for hosting such an exciting competition
- The competitors themselves

That is all!

Many thanks to

- The organizers for hosting such an exciting competition
- The competitors themselves
- and you for your attention!