Question 1: Expected goals

These goal rates were found by using gradient boosted trees under the assumption of the independence on the opponent.

Here is a table of our results:

Country	Expected Goals	ELO	Age	Height	History	Manager
Croatia	1.56	1953	27.7	185	0.5	0
Belgium	1.48	1986	26.0	184	0.0	2
Netherlands	1.37	1968	26.7	184	2.5	1
France	1.36	2085	26.1	184	4.0	2
Italy	1.31	1956	26.3	185	4.5	2
Germany	1.12	1921	28.4	184	7.0	2
England	1.11	1999	26.6	182	0.5	0
Portugal	1.09	2013	26.8	181	1.5	1
Austria	0.92	1857	26.0	182	0.0	2
Slovenia	0.89	1739	25.8	184	0.0	0
Denmark	0.89	1822	26.8	184	1.0	0
Hungary	0.87	1842	26.7	183	0.0	0
Ukraine	0.81	1863	25.1	182	0.0	0
Poland	0.80	1721	27.9	184	0.0	0
Scotland	0.75	1776	27.5	183	0.0	0
Serbia	0.73	1788	27.2	185	0.0	0
Turkey	0.72	1750	25.3	183	0.0	1
Spain	0.68	2019	26.9	182	3.0	1
Slovakia	0.67	1655	27.6	183	0.0	1
Georgia	0.67	1650	27.8	183	0.0	0
Czech Republic	0.67	1772	25.5	185	1.5	0
Switzerland	0.63	1804	28.2	182	0.0	0
Romania	0.58	1664	26.9	182	0.0	0
Albania	0.50	1613	26.6	182	0.0	0

Table 1: λ rates predicted by GBT. It is assumed that the goals scored by A are independent of opponent B and vice-versa. Croatia and Belgium are the top2.

As suggested in (Schauberge and Grall , 2018) ,they can be used as the λ parameters in a Poisson distribution:

$$\text{goals}_i \sim \text{Poi}(\lambda_i), \quad i \in \text{countries}$$

If one adopts this Poisson model, one can answer questions such as:

1. What is the probability of a team scoring (one or more goals) ? It is simply

$$1 - P[k = 0] = 1 - e^{-\lambda}$$

2. What is the probability that the combined goals of both teams are lower than 3 ? The combined goals follow

$$G \sim \text{Poi}(\lambda_A) + \text{Poi}(\lambda_B) = \text{Poi}(\lambda_A + \lambda_A)$$

. And the probability is

$$P[G < 3] = P[G = 0] + P[G = 1] + P[G = 2] = \sum_{k=0}^{2} \frac{(\lambda_A + \lambda_B)^k e^{-(\lambda_A + \lambda_B)}}{k!}$$

3. and various other similar questions...

Remark

If ones attempts the use of the previous Poisson model with $\lambda = \lambda_{GBT}$, the BTTS, O2.5 (both teams to score , over total 2.5 goals) odds tend to be overly conservative , being much higher odds than the ones from betting companies.

This "poor performance" is because a double violation of independence: (i) the assumption of independence-of-the-rates on the opponent is violated (in fact any probability distribution that uses these rates a parameter would fail because the rates must change as a function of the opponent also). (ii) The independence-of-rates in the usual "Poisson sense", i.e. that the rates are fixed in time, independent of the previous goals.

As an example, the Poisson with our GBT-rates predicts:

CountryA	**CountryB**	**BTTS**	**O2.5**	**U2.5**
Belgium	Slovakia	2.65	2.75	1.57
Italy	Albania	3.48	3.68	1.37

while the bookmakers give:

CountryA	**CountryB**	**BTTS**	**O2.5**	**U2.5**
Belgium	Slovakia	1.95	2.12	1.80
Italy	Albania	2.4	2	1.85

Our odds relating to the goals scored are not too realistic in the sense of the assumptions (i),(ii) above. Nevertheless they are reasonably accurate and we will present them (BTTS,O2.5,U2.5) for all of the group matches.

To get a better picture one would need to retrain the GBT as a function of both the (x_A, x_B) and correct for symmetry.

Question 2: The Group Stages

For each group (A,...,F), for each match in that group we present **our odds** of :

1. Country 1 winning, noted by **1**
2. A draw, **X**
3. Country 2 winning, **2**
4. Both teams scoring , **BTTS**
5. The total of goals being strictly smaller than 3 , **under2.5**
6. The total of goals being bigger or equal to 3 , **over2.5**

Note: here the odds are just the inverses of probabilities. Eg: $p = \frac{1}{3} \implies \text{odds} = 3$.

For each group we also present our predicted probabilities of the 24 permutations of the top3 finishes. These probabilities were obtained by Monte Carlo simulations on the 1-X-2 probabilities found by the RF in combination with the goal difference from the GBT- Poisson.

Remark

The 1,X,2 odds were predicted using RF (with the symmetry correction) and hold no assumptions except the ones inherent to the classification problem.

Group A

CountryA	CountryB	1	X	2	BTTS	over2.5	under2.5
Germany	Scotland	2.21	3.97	3.39	2.81	3.47	1.40
Hungary	Switzerland	1.91	7.46	2.93	3.68	5.23	1.24
Germany	Hungary	2.23	3.86	3.41	2.55	3.12	1.47
Scotland	Switzerland	2.08	6.89	2.67	4.05	6.18	1.19
Switzerland	Germany	3.26	4.61	2.10	3.18	3.91	1.34
Scotland	Hungary	2.61	8.49	2.00	3.26	4.51	1.29

Table 2: Our odds for group A

CountryA	CountryB	1	X	2	BTTS	over2.5	under2.5
Germany	Scotland	1.29	6.00	13.00	2.50	1.75	2.10
Hungary	Switzerland	3.50	3.40	2.20	1.85	2.10	1.85
Germany	Hungary	1.36	5.00	9.50	2.15	1.77	2.10
Scotland	Switzerland	3.40	3.60	2.15	1.75	1.91	1.91

Table 3: The bookmakers' odds for group A. The odds for the last two matches are unavailable.

Permutation	Odds
Germany-Scotland-Hungary	11.67
Germany-Hungary-Scotland	12.89
Hungary-Germany-Scotland	14.20
Germany-Hungary-Switzerland	14.84
Hungary-Germany-Switzerland	17.65
Scotland-Germany-Hungary	19.01
Germany-Scotland-Switzerland	19.19
Hungary-Switzerland-Germany	22.73
Germany-Switzerland-Hungary	22.83
Germany-Switzerland-Scotland	23.04
Hungary-Scotland-Germany	24.69
Scotland-Hungary-Germany	25.13
Switzerland-Germany-Hungary	28.57
Scotland-Germany-Switzerland	29.33
Switzerland-Germany-Scotland	30.21
Scotland-Switzerland-Germany	32.36
Scotland-Hungary-Switzerland	33.00
Switzerland-Hungary-Germany	34.60
Hungary-Scotland-Switzerland	35.21
Hungary-Switzerland-Scotland	40.49
Switzerland-Scotland-Germany	45.05
Switzerland-Scotland-Hungary	49.26
Scotland-Switzerland-Hungary	49.50
Switzerland-Hungary-Scotland	55.87

Table 4: Our odds for all the possible top3 permutations in group A

Group B

CountryA	CountryB	1	X	2	BTTS	over2.5	under2.5
Spain	Croatia	3.64	5.05	1.90	2.57	2.58	1.63
Italy	Albania	1.49	5.67	6.50	3.48	3.68	1.37
Croatia	Albania	1.74	5.28	4.24	3.22	2.94	1.51
Spain	Italy	2.63	5.26	2.32	2.78	3.12	1.47
Albania	Spain	7.50	3.43	1.74	5.15	8.61	1.13
Croatia	Italy	2.21	6.36	2.57	1.73	1.83	2.21

Table 5: Our odds for group B

CountryA	CountryB	1	X	2	BTTS	over2.5	under2.5
Spain	Croatia	1.75	3.80	5.00	1.90	2.00	1.91
Italy	Albania	1.38	4.75	11.00	2.40	2.00	1.85
Croatia	Albania	1.57	4.00	6.50	2.10	2.10	1.77
Spain	Italy	2.30	3.40	3.30	1.85	2.10	1.77

Table 6: Bookmakers' odds for group B

Permutation	Odds
	7.47
	8.36
	9.31
	11.09
	12.06
	16.41
	17.61
	21.01
	23.09
	23.75
	31.35
	32.05
	43.48
	44.05
	46.51
	48.31
	66.67
	89.29
	125.00
	138.89
	161.29
	175.44
	178.57
	181.82

Table 7: Our odds for all the possible top3 permutations in group B

Group C

CountryA	CountryB	1	X	2	BTTS	over2.5	under2.5
Slovenia	Denmark	2.87	4.49	2.33	2.88	3.79	1.36
Serbia	England	4.22	3.85	1.99	2.88	3.57	1.39
Slovenia	Serbia	1.87	4.94	3.81	3.28	4.51	1.29
Denmark	England	3.35	5.37	1.94	2.53	3.09	1.48
England	Slovenia	1.77	3.88	5.67	2.53	3.09	1.48
Denmark	Serbia	2.16	5.89	2.72	3.28	4.51	1.29

Table 8: Our odds for group C

CountryA	CountryB	1	X	2	BTTS	over2.5	under2.5
Slovenia	Denmark	5.00	3.70	1.80	2.00	2.20	1.82
Serbia	England	7.50	4.60	1.50	1.95	1.80	2.14
Slovenia	Serbia	4.10	3.60	2.00	1.83	2.00	1.80
Denmark	England	6.00	4.00	1.65	2.00	2.00	1.83

Table 9: Bookmakers' odds for group C

Permutation	Odds
	9.46
	10.98
	13.59
	13.74
	14.76
	18.73
	19.88
	20.88
	23.47
	23.75
	24.75
	27.10
	29.59
	31.25
	34.84
	35.84
	36.50
	40.49
	44.25
	49.02
	55.56
	60.61
	73.53

Table 10: Our odds for all the possible top3 permutations in group C

Group D

CountryA	CountryB	1	X	2	BTTS	over2.5	under2.5
Poland	Netherlands	4.55	6.63	1.59	2.43	2.71	1.59
Austria	France	5.29	2.97	2.11	2.24	2.51	1.66
Poland	Austria	3.09	5.09	2.08	3.02	4.03	1.33
Netherlands	France	2.61	5.05	2.39	1.80	1.95	2.06
Netherlands	Austria	1.94	5.86	3.18	2.23	2.49	1.67
France	Poland	1.75	3.80	6.02	2.44	2.73	1.58

Table 11: Our odds for group D

CountryA	CountryB	1	X	2	BTTS	over2.5	under2.5
Poland	Netherlands	5.15	4.10	1.70	1.91	1.91	1.92
Austria	France	7.00	4.75	1.50	1.80	1.67	2.28
Poland	Austria	3.50	3.60	2.15	1.70	1.83	2.00
Netherlands	France	4.75	3.90	1.80	1.77	2.85	1.83

Table 12: Bookmakers' odds for group D

Permutation	Odds
	9.52
	9.52
	11.54
	11.55
	12.05
	14.57
	21.05
	21.51
	23.75
	23.98
	25.51
	29.41
	30.12
	35.09
	38.61
	45.25
	51.81
	62.89
	65.79
	80.00
	80.65
	91.74
	94.34
	131.58

Table 13: Our odds for all the possible top3 permutations in group D

Group E

CountryA	CountryB	1	X	2	BTTS	over2.5	under2.5
Romania	Ukraine	4.45	3.53	2.03	4.09	6.09	1.20
Belgium	Slovakia	1.64	5.71	4.62	2.65	2.75	1.57
Slovakia	Ukraine	3.39	3.43	2.42	3.69	5.37	1.23
Belgium	Romania	1.42	7.00	6.64	2.94	2.94	1.51
Slovakia	Romania	1.92	11.79	2.54	4.65	7.60	1.15
Ukraine	Belgium	5.08	5.69	1.59	2.33	2.49	1.67

Table 14: Our odds for group E

CountryA	CountryB	1	X	2	BTTS	over2.5	under2.5
Romania	Ukraine	3.60	3.60	2.15	2.00	2.32	1.67
Belgium	Slovakia	1.50	4.50	7.50	1.95	1.80	2.12
Slovakia	Ukraine	4.33	3.40	2.00	2.00	2.20	1.68
Belgium	Romania	1.53	4.33	7.00	1.95	1.80	2.00

Table 15: Bookmakers' odds for group E

Permutation	Odds
	6.82
	8.36
	8.98
	11.30
	12.42
	15.92
	17.39
	20.83
	22.57
	34.01
	35.84
	36.10
	39.22
	47.62
	51.55
	55.56
	85.47
	93.46
	108.70
	129.87
	121.95
	142.86
	181.82

Table 16: Our odds for all the possible top3 permutations in group E

Group F

CountryA	CountryB	1	X	2	BTTS	O2.5	U2.5
Turkey	Georgia	2.18	5.39	2.82	3.99	6.09	1.20
Portugal	Czech Republic	1.68	4.96	4.90	3.09	3.87	1.35
Georgia	Czech Republic	2.62	4.16	2.65	4.19	6.57	1.18
Turkey	Portugal	5.46	4.24	1.72	2.94	3.68	1.37
Georgia	Portugal	4.86	4.58	1.74	3.09	3.87	1.35
Czech Republic	Turkey	3.36	3.96	2.22	3.99	6.09	1.20

Table 17: Our odds for group F

CountryA	CountryB	1	X	2	BTTS	O2.5	U2.5
Turkey	Georgia	1.80	3.80	5.00	2.01	2.20	1.75
Portugal	Czech Republic	1.55	4.20	7.00	1.95	1.91	2.00
Georgia	Czech Republic	5.00	3.65	1.85	1.95	2.05	1.80
Turkey	Portugal	5.50	4.00	1.70	1.91	1.91	1.92

Table 18: Bookmakers' odds for group F

Permutation	Odds
Portugal-Turkey-Georgia	8.50
Portugal-Turkey-Czech Republic	9.33
Portugal-Georgia-Turkey	11.17
Portugal-Czech Republic-Turkey	12.32
Portugal-Georgia-Czech Republic	14.14
Portugal-Czech Republic-Georgia	15.60
Turkey-Portugal-Georgia	16.64
Turkey-Portugal-Czech Republic	18.05
Georgia-Portugal-Turkey	20.75
Georgia-Portugal-Czech Republic	25.38
Turkey-Georgia-Portugal	30.30
Czech Republic-Portugal-Turkey	30.58
Georgia-Turkey-Portugal	40.32
Czech Republic-Turkey-Portugal	42.19
Turkey-Czech Republic-Portugal	43.29
Czech Republic-Portugal-Georgia	43.48
Czech Republic-Georgia-Portugal	56.50
Georgia-Czech Republic-Portugal	62.50
Turkey-Georgia-Czech Republic	63.69
Turkey-Czech Republic-Georgia	69.93
Czech Republic-Turkey-Georgia	78.74
Georgia-Czech Republic-Turkey	84.75
Georgia-Turkey-Czech Republic	96.15
Czech Republic-Georgia-Turkey	131.58

Table 19: Our odds for all the possible top3 permutations in group F

Question 3: Knockout stage (with extra-time and penalties)

In addition to the normal time predictions (GBT for goal rates, "corrected" RF for the 1-X-2 probs.) **we also made extra-time and penalty predictions** (again with the same models: GBT for extratime goal rates,"corrected" RF for extra-time 1-X-2 probs. and 1-2 penalty probs).

For extra time : the models were tuned/trained on a dataset containing about 20 extra-time matches from the past EURO competitions.

For penalties: the models were tuned/trained on a dataset containing penalty shootout data from world cups (since 1982, the year they were first introduced in world cups) and from past EURO cups (since 1976 , the year they were first introduced into EURO). This gives about 60 total samples.

As for the **knockout stage pairings**, for each group: we take the teams with most probable top3 permutation. Not all 3rd placed teams get to go into the knockouts (there is a separate selection for them based on goal difference , goals scored , etc...).For the 3rd team selection: we choose the ones that have the highest ELO , i.e. **for the 3rd place teams: the highest ELO teams go to the knockouts.**

That gives us:

Date	Match	Teams
30 June	1	Croatia v Hungary
29 June	2	Germany v Slovenia
1 July	3	Portugal v Denmark
1 July	4	France v Ukraine
2 July	5	Belgium v Spain
2 July	6	Netherlands v Turkey
30 June	7	England v Austria
29 June	8	Scotland v Italy

Table 20: Round of 16 Matches

We will present the **normal time, extra time and penalty odds** for matches of the round of 16, alongside with the goal rates for normal and extra-time.

Normal time

CountryA	CountryB	1	X	2	BTTS	over2.5	under2.5
Croatia	Hungary	2.28	5.55	2.62	2.18	2.28	1.78
Germany	Slovenia	1.85	4.37	4.34	2.52	3.07	1.48
Portugal	Denmark	1.95	6.25	3.06	2.56	3.15	1.47
France	Ukraine	2.17	3.02	4.83	2.42	2.71	1.59
Belgium	Spain	3.12	2.88	3.01	2.62	2.73	1.58
Netherlands	Turkey	1.62	7.69	3.97	2.61	2.88	1.53
England	Austria	2.23	3.62	3.62	2.48	3.02	1.50
Scotland	Italy	3.64	5.48	1.84	2.60	2.94	1.51

Table 21: Our normal time odds for the knockout matches

Extra time

Country	Extra time λ	ELO	Age	Height	History	Manager
Germany	0.48	1921	28.4	184	7.0	2
Scotland	0.49	1776	27.5	183	0.0	0
Hungary	0.48	1842	26.7	183	0.0	0
Spain	0.50	2019	26.9	182	3.0	1
Croatia	0.48	1953	27.7	185	0.5	0
Italy	0.48	1956	26.3	185	4.5	2
Slovenia	0.49	1739	25.8	184	0.0	0
Denmark	0.48	1822	26.8	184	1.0	0
England	0.50	1999	26.6	182	0.5	0
Netherlands	0.49	1968	26.7	184	2.5	1
Austria	0.50	1857	26.0	182	0.0	2
France	0.48	2085	26.1	184	4.0	2
Belgium	0.49	1986	26.0	184	0.0	2
Ukraine	0.51	1863	25.1	182	0.0	0
Turkey	0.49	1750	25.3	183	0.0	1
Portugal	0.50	2013	26.8	181	1.5	1

Table 22: Our extra time GBT goal rates. All of the countries have $\lambda \approx 0.50$, suggesting that extra-time tends to be low goal for everyone. For extra time it seems reasonable to assume (i) (independence of rates on the opponent) since everyone plays cautiously. Assumption (ii) (the Poisson constant-rates-in-time) still fails , but could be assumed to hold until one team scores.

CountryA	CountryB	1	X	2
Croatia	Hungary	5.78	2.24	2.62
Germany	Slovenia	2.45	5.08	2.53
Portugal	Denmark	1.70	3.06	11.90
France	Ukraine	35.71	1.61	2.85
Belgium	Spain	22.73	1.84	2.42
Netherlands	Turkey	6.49	2.09	2.72
England	Austria	6.67	1.49	5.56
Scotland	Italy	2.55	2.51	4.78

Table 23: Our odds for extra time "win-draw-lose". Overall draws are most probable , with the RF sometimes favoring "underdogs" (eg: giving France almost no probability to win against Ukraine). This is a consequence of the limited dataset (only 20 matches) . Because of this these odds should not be taken too seriously. The (BTTS, over2.5 , under2.5) odds $= (6.50, 12.5, 1.10)$ for each match , since the rates are always $\lambda = 0.5$ (this is assuming a Poisson model, like before).

Penalties

CountryA	CountryB	KeeperHeightA	KeeperHeightB	1	2
Croatia	Hungary	188	191	3.77	1.36
Germany	Slovenia	193	188	2.33	1.75
Portugal	Denmark	186	189	1.80	2.25
France	Ukraine	191	191	1.34	3.92
Belgium	Spain	200	190	2.15	1.87
Netherlands	Turkey	203	198	2.30	1.77
England	Austria	185	184	1.59	2.70
Scotland	Italy	196	196	1.52	2.94

Table 24: 1-2 probabilities calculated using RF. For this penalty model there were only four covariates: ELO and keeper height (in cm) of both teams. The GBT penalty rates were not calculated since the (i) assumption is clearly violated because of the structure of the penalty shootout. The odds for the first and last match are somewhat bizarre. One could argue that according to the dataset and the RF trained on it , if an underdog reaches penalties against a clear favorite , the underdog becomes the favorite. But most likely this is an "inaccuracy" due to the limited dataset.

1 Conclusion - On what to bet ?

All of the state-of-the-art models get beaten by the bookmakers. In other words : a model is yet to be invented which can beat the betting companies in terms of predictive power.

(Most likely betting companies also use the most powerful models in combination with observed betting data.)

That being said , we will use the predictions of our models, GBT-Poisson and corrected RF trained on the previous EUROs with our feature selection , and try to "beat the house". That is ,look for differences in the bookmakers odds and our odds and bet on matches where we could potentially profit.

Here is one possible betting ticket :

CountryA	CountryB	Bet	Our odds	Bookies
Hungary	Switzerland	1	1.91	3.50
Spain	Croatia	X2	1.37	2.16
Slovenia	Serbia	1	1.87	4.10
Poland	Austria	under2.5	1.33	2.00
Belgium	Romania	1	1.42	1.53
Georgia	Czech Republic	X1	1.61	2.1
		TOTAL:	14.9	199

Table 25: A possible "beat the house" bet using one match from from each of the six groups. "X1/X2" means draw or home/away win with the odds are calculated as (here for X2) $\frac{1}{\frac{1}{o_2}+\frac{1}{o_X}}$ i.e. the inverse of $p_{X2} = p_2 + p_X$.

A 5 euro bet on this would give a 1000 euros.