Data Fellowship Webinar Series

Data Project Example

01 **Plan**

02 **Data Prep**

03 **Analysis**

Modelling 04

Refine and Compare 05

Communicate and Implement 06

Plan

HOW DO YOU KNOW WHO IS WINNING?

CRICKET
(THE REAL BEAUTIFUL SPORT)

Project Aim

Build a model that predicts whether a team has improved their chances of winning a Twenty20 match through their performance

Project Aim

- Predict the end of innings score
- Convert this score into chance of winning
- Display the outcome with an easy to navigate dashboard

Data Prep

Web Scraping

RESULT

South Group (N), Vitality Blast at Lord's, Jul 18 2019

Essex		164/6
Middlesex	17/20 ov, target 165	166/3

Middlesex won by 7 wickets (with 18 balls remaining)

END OF OVER 15 10 runs

Middx: 139/2 CRR: 9.26 • RRR: 5.20 • Need 26 runs from 30b

AB de Villiers	76 (39)	Matthew Quinn	3-0-20-1
Dawid Malan	39 (32)	Dan Lawrence	1-0-17-0

1st_wickets	2nd_target	2nd_total	2nd_wickets	Date	Day_Night	Night_game	Stadium	bat_first	bat_second	comp	country	country_run_rate	dl	draw_perc	game_id	innings	loss_perc	match_notes
8				2005-06-13	0	0	The Rose Bowl, Southampton	England	Australia	Australia tour of England and Scotland 2005	England	10.7		0.00495628703038367	0	1st Innings	0.54576271974595	England won t
8				2005-06-13	0	0	The Rose Bowl, Southampton	England	Australia	Australia tour of England and Scotland 2005	England	10.7	0	0.00524509232693316	0	1st Innings	0.557972483527951	England won t
8				2005-06-13	0	0	The Rose Bowl, Southampton	England	Australia	Australia tour of England and Scotland 2005	England	10.7	0	0.00549926336932667	0	1st Innings	0.586974338971047	England won t
8				2005-06-13	0	0	The Rose Bowl, Southampton	England	Australia	Australia tour of England and Scotland 2005	England	10.7	0	0.00529417144015744	0	1st Innings	0.574417528963374	England won t
8				2005-06-13	0	0	The Rose Bowl, Southampton	England	Australia	Australia tour of England and Scotland 2005	England	10.7	0	0.00539652764756998	0	1st Innings	0.578403711578506	England won t
8				2005-06-13	0	0	The Rose Bowl, Southampton	England	Australia	Australia tour of England and Scotland 2005	England	10.7	0	0.00547463675213397	0	1st Innings	0.582785476266529	England won t
8				2005-06-13	0	0	The Rose Bowl, Southampton	England	Australia	Australia tour of England and Scotland 2005	England	10.7	0	0.0055209388663048	0	1st Innings	0.61781764023444	England won t
8				2005-06-13	0	0	The Rose Bowl, Southampton	England	Australia	Australia tour of England and Scotland 2005	England	10.7	0	0.00541708854486011	0	1st Innings	0.61875924092059	England won t
8				2005-06-13	0	0	The Rose Bowl, Southampton	England	Australia	Australia tour of England and Scotland 2005	England	10.7	0	0.00566149478708638	0	1st Innings	0.623550905820508	England won t
8				2005-06-13	0	0	The Rose Bowl, Southampton	England	Australia	Australia tour of England and Scotland 2005	England	10.7	0	0.00556413851741254	0	1st Innings	0.58163825895685	England won t
8				2005-06-13	0	0	The Rose Bowl, Southampton	England	Australia	Australia tour of England and Scotland 2005	England	10.7	0	0.0056872857644369	0	1st Innings	0.591710953681639	England won t
8				2005-06-13	0	0	The Rose Bowl, Southampton	England	Australia	Australia tour of England and Scotland 2005	England	10.7	0	0.00583822964220899	0	1st Innings	0.633424163339887	England won t
8				2005-06-13	0	0	The Rose Bowl, Southampton	England	Australia	Australia tour of England and Scotland 2005	England	10.7	0	0.00588284099969437	0	1st Innings	0.656347921227462	England won t
8				2005-06-13	0	0	The Rose Bowl, Southampton	England	Australia	Australia tour of England and Scotland 2005	England	10.7	0	0.00585672583790262	0	1st Innings	0.662042885828931	England won t
8				2005-06-13	0	0	The Rose Bowl, Southampton	England	Australia	Australia tour of England and Scotland 2005	England	10.7	0	0.00602800574523748	0	1st Innings	0.675630708626729	England won t
8				2005-06-13	0	0	The Rose Bowl, Southampton	England	Australia	Australia tour of England and Scotland 2005	England	10.7	0	0.00586665904315945	0	1st Innings	0.670931464714785	England won t
8				2005-06-13	0	0	The Rose Bowl, Southampton	England	Australia	Australia tour of England and Scotland 2005	England	10.7	0	0.0057708497919106	0	1st Innings	0.674767921605666	England won t
8				2005-06-13	0	0	The Rose Bowl, Southampton	England	Australia	Australia tour of England and Scotland 2005	England	10.7	0	0.00582157135273606	0	1st Innings	0.64919358855136	England won t
8				2005-06-13	0	0	The Rose Bowl, Southampton	England	Australia	Australia tour of England and Scotland 2005	England	10.7	0	0.00579580085143644	0	1st Innings	0.655937678999765	England won t
8				2005-06-13	0	0	The Rose Bowl, Southampton	England	Australia	Australia tour of England and Scotland 2005	England	10.7	0	0.00543029855101272	0	1st Innings	0.662369177455531	England won t
8				2005-06-13	0	0	The Rose Bowl, Southampton	England	Australia	Australia tour of England and Scotland 2005	England	10.7		0.00542654449498803	0	1st Innings	0.692104274977702	England won t
	179	79	10	2005-06-13	0	0	The Rose Bowl, Southampton	England	Australia	Australia tour of England and Scotland 2005	England	7.49914118859499		0.00542654449498803	0	2nd Innings	0.30246918052731	England won t
8	179	79	10	2005-06-13	0	0	The Rose Bowl, Southampton	England	Australia	Australia tour of England and Scotland 2005	England	7.49914118859499	0	0.0186720649644342	0	2nd Innings	0.713378407865319	England won t
8	179	79	10	2005-06-13	0	0	The Rose Bowl, Southampton	England	Australia	Australia tour of England and Scotland 2005	England	7.49914118859499	0	0.023171149493238	0	2nd Innings	0.622002994664169	England won t
8	179	79	10	2005-06-13	0	0	The Rose Bowl, Southampton	England	Australia	Australia tour of England and Scotland 2005	England	7.49914118859499	0	0.0138270145093407	0	2nd Innings	0.80435971184471	England won t
8	179	79	10	2005-06-13	0	0	The Rose Bowl, Southampton	England	Australia	Australia tour of England and Scotland 2005	England	7.49914118859499	0	0.00185604628500191	0	2nd Innings	0.990283090535141	England won t
8	179	79	10	2005-06-13	0	0	The Rose Bowl, Southampton	England	Australia	Australia tour of England and Scotland 2005	England	7.49914118859499	0	0.0014744982573397	0	2nd Innings	0.995125278597284	England won t
8	179	79	10	2005-06-13	0	0	The Rose Bowl, Southampton	England	Australia	Australia tour of England and Scotland 2005	England	7.49914118859499	0	0.000258457207732457	0	2nd Innings	0.999686844067467	England won t
8	179	79	10	2005-06-13	0	0	The Rose Bowl, Southampton	England	Australia	Australia tour of England and Scotland 2005	England	7.49914118859499	0	5.54394215495372E-05	0	2nd Innings	0.999922175800969	England won t
8	179	79	10	2005-06-13	0	0	The Rose Bowl, Southampton	England	Australia	Australia tour of England and Scotland 2005	England	7.49914118859499	0	2.46724945551225E-05	0	2nd Innings	0.999959788274768	England won t
8	179	79	10	2005-06-13	0	0	The Rose Bowl, Southampton	England	Australia	Australia tour of England and Scotland 2005	England	7.49914118859499	0	4.64560438908958E-05	0	2nd Innings	0.99994114230435 9	England won t
8	179	79	10	2005-06-13	0	0	The Rose Bowl, Southampton	England	Australia	Australia tour of England and Scotland 2005	England	7.49914118859499	0	0.000207276996351945	0	2nd Innings	0.999577990035256	England won t
8	179	79	10	2005-06-13	0	0	The Rose Bowl, Southampton	England	Australia	Australia tour of England and Scotland 2005	England	7.49914118859499	0	4.4005822555717E-05	0	2nd Innings	0.999948275241261	England won t
8	179	79	10	2005-06-13	0	0	The Rose Bowl, Southampton	England	Australia	Australia tour of England and Scotland 2005	England	7.49914118859499	0	4.96999604987154E-05	0	2nd Innings	0.999947641015313	England won t
8	179	79	10	2005-06-13	0	0	The Rose Bowl, Southampton	England	Australia	Australia tour of England and Scotland 2005	England	7.49914118859499	0	2.14949940710482E-05	0	2nd Innings	0.999977724513442	England won t
8	179	79	10	2005-06-13	0	0	The Rose Bowl, Southampton	England	Australia	Australia tour of England and Scotland 2005	England	7.49914118859499	0	3.14841005546313E-06	0	2nd Innings	0.999996815148695	England won t
	179	79	10	2005-06-13	0	0	The Rose Bowl, Southampton	England	Australia	Australia tour of England and Scotland 2005	England	7.49914118859499		1.28639411007461E-06	0	2nd Innings	0.999998711548021	England won t
10				2005-10-21	1	0	The Wanderers Stadium, Johannesburg	South Africa	New Zealand	New Zealand tour of South Africa 2005/06	South Africa	9.825		0.00984130375513266	1	1st Innings	0.533734809585803	New Zealand
10				2005-10-21	1	0	The Wanderers Stadium, Johannesburg	South Africa	New Zealand	New Zealand tour of South Africa 2005/06	South Africa	9.825	0	0.00853493120043438	1	1st Innings	0.478343079634449	New Zealand

DATA COLLECTION

4157 games used

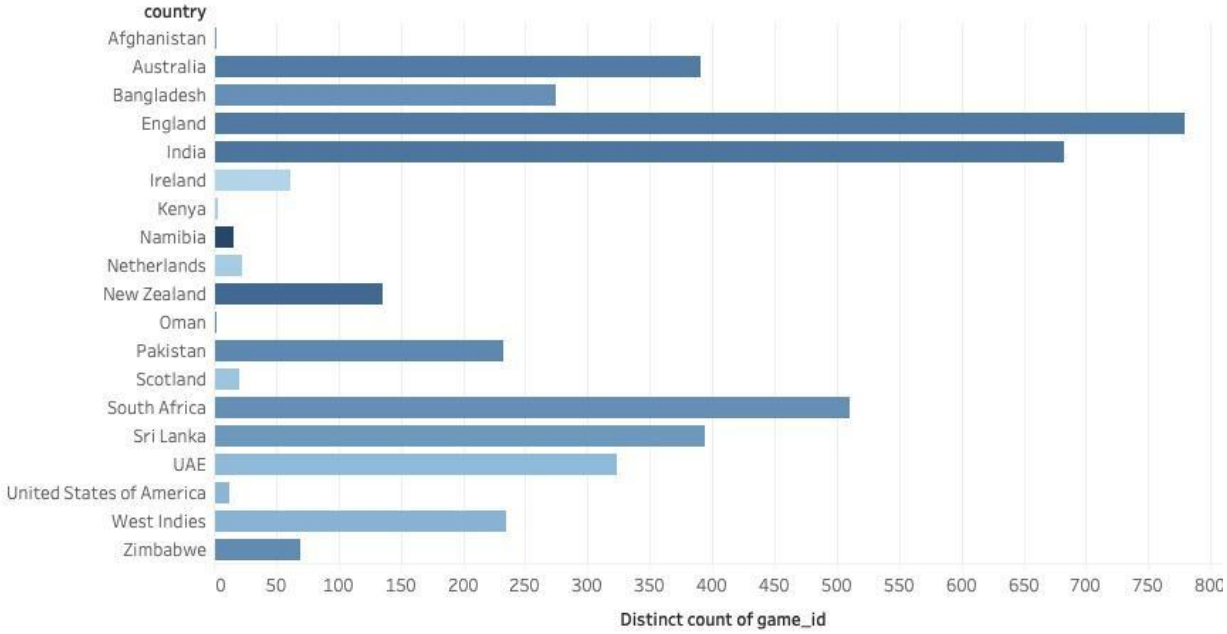
From 19 countries

And 203 Stadiums

country	Frequency
England	779
India	682
South Africa	510
Sri Lanka	394
Australia	391
UAE	323
Bangladesh	275
West Indies	235
Pakistan	232
New Zealand	136
Zimbabwe	69
Ireland	61

Stadium	Frequency
Kennington Oval, London	59
Trent Bridge, Nottingham	58
Edgbaston, Birmingham	56
The Rose Bowl, Southampton	45
Sophia Gardens, Cardiff	44
Riverside Ground, Chester-le-Street	42
County Ground, Hove	41
Headingley, Leeds	40
The Cooper Associates County Ground, Taunton	39
County Ground, Northampton	38
Old Trafford, Manchester	37
County Ground, Chelmsford	36
Grace Road, Leicester	35
County Ground, New Road, Worcester	34
Lord's, London	33

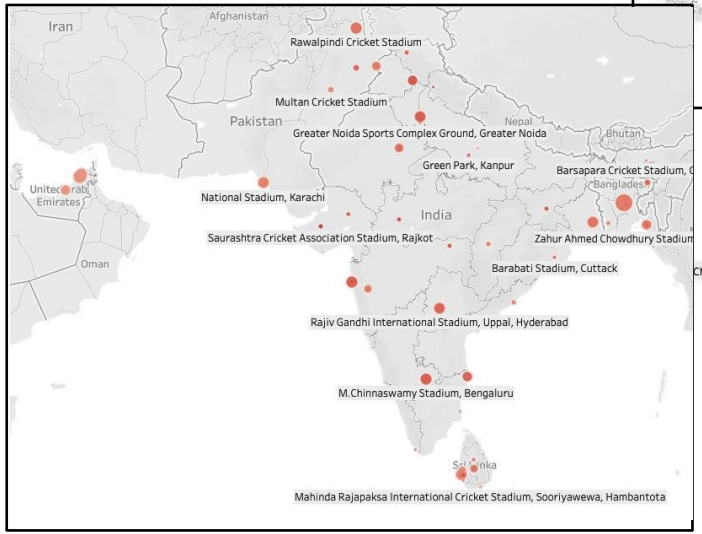
Games Scraped per country

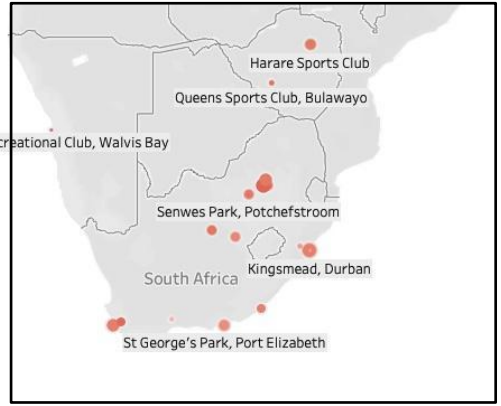


Riverside Ground, Chester-le-Street
North Marine Road Ground, Scarborough
Headingley, Leeds
Old Trafford, Manchester
Queen's Park, Chesterfield
County Ground, Derby
United Kingdom
unty Ground, New Road, Worcester
College Ground, Cheltenham
County Ground, Bristol
Nevill Ground, Tunbridge Wel
County Ground, Hove

Northern Territory
Tony Ireland Stadium, Townsville
Traeger Park, Alice Springs
South Australia
Brisbane Cricket Ground, Woolloongabba, Brisbane
New South Wales
Australian Capital Territory
Adelaide Oval
Manuka Oval, Canberra
Victoria
Docklands Stadium, Melbourne
McLean Park, N
Aurora Stadium, Launceston
Westpac Stadium, Wellir
Bellerive Oval, Hobart
Hagley Oval, Christchurch
Queen's Park, Invercargill

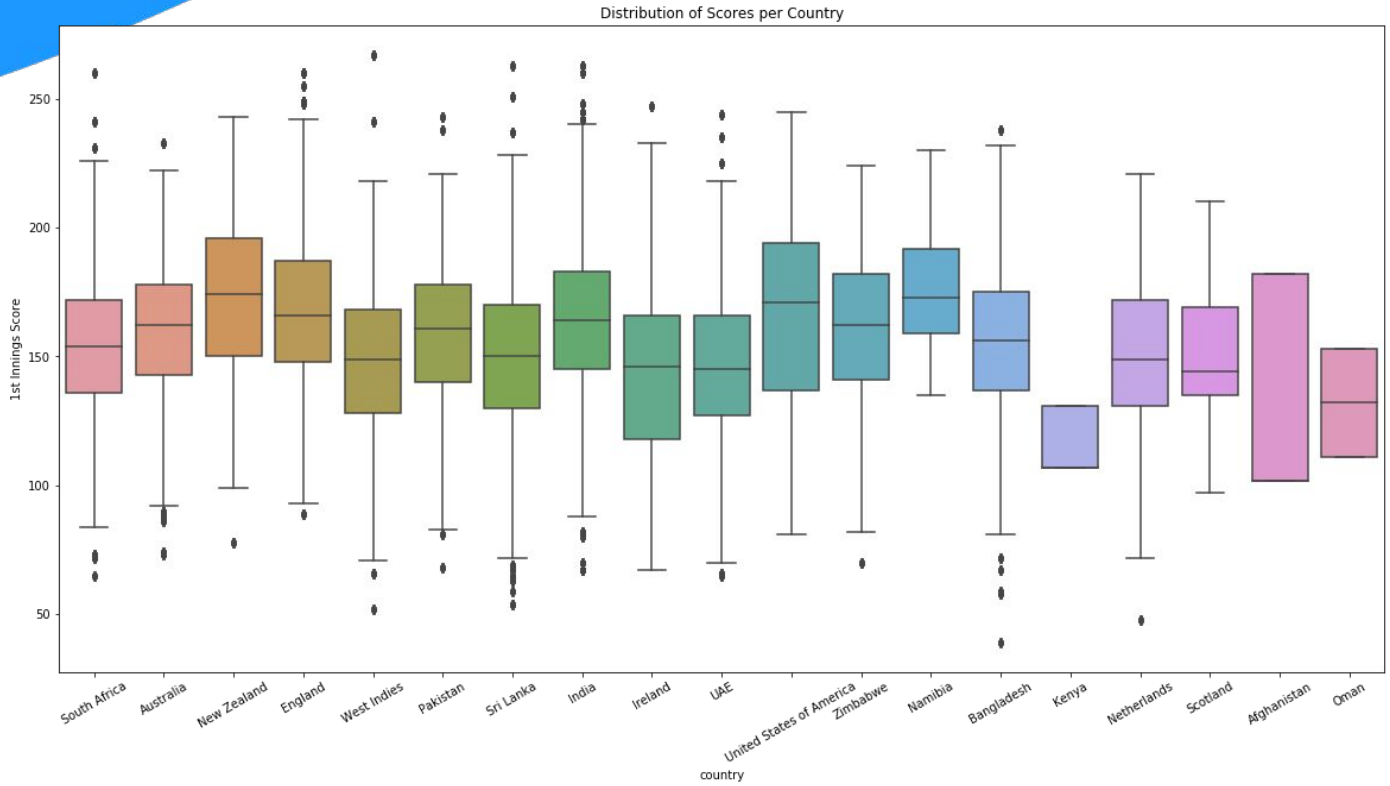
Georgia
White Hill Field, Sandys Parish, Hamilton
Florida
oward Regional Park Stadium Turf Ground, Lauderhill
Cuba
Sabina Park, Kingston, Jamaica
Sir Vivian Richards Stadium, North Sound, Antigua
Nicaragua
Costa Rica
National Cricket Stadium, St George's, Grenada
Queen's Park Oval, Port of Spain, Trinidad
Venezuela
Guyana
Colombia
Providence Stadium, Guyana

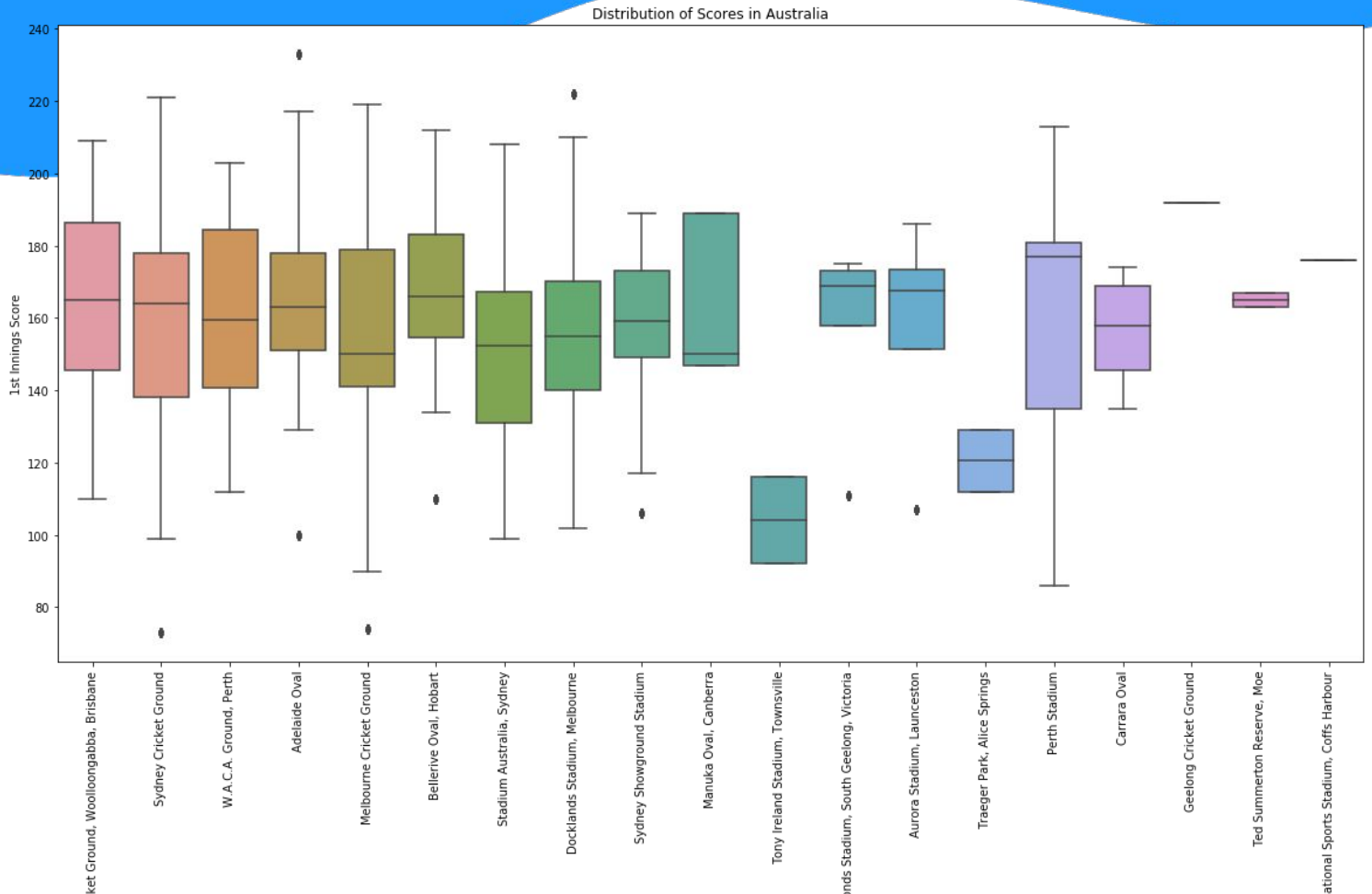

Iran
Afghanistan
Rawalpindi Cricket Stadium
Multan Cricket Stadium
Pakistan
Nepal
Greater Noida Sports Complex Ground, Greater Noida
Bhutan
Green Park, Kanpur
Barsapara Cricket Stadium,
Bangladesh
United Arab Emirates
National Stadium, Karachi
India
Saurashtra Cricket Association Stadium, Rajkot
Zahur Ahmed Chowdhury Stadium
Oman
Barabati Stadium, Cuttack
Rajiv Gandhi International Stadium, Uppal, Hyderabad
M.Chinnaswamy Stadium, Bengaluru
Sri Lanka
Mahinda Rajapaksa International Cricket Stadium, Sooriyawewa, Hambantota


Harare Sports Club
Queens Sports Club, Bulawayo
creational Club, Walvis Bay
Senwes Park, Potchefstroom
Kingsmead, Durban
South Africa
St George's Park, Port Elizabeth

Analysis

EDA AND FEATURE SELECTION

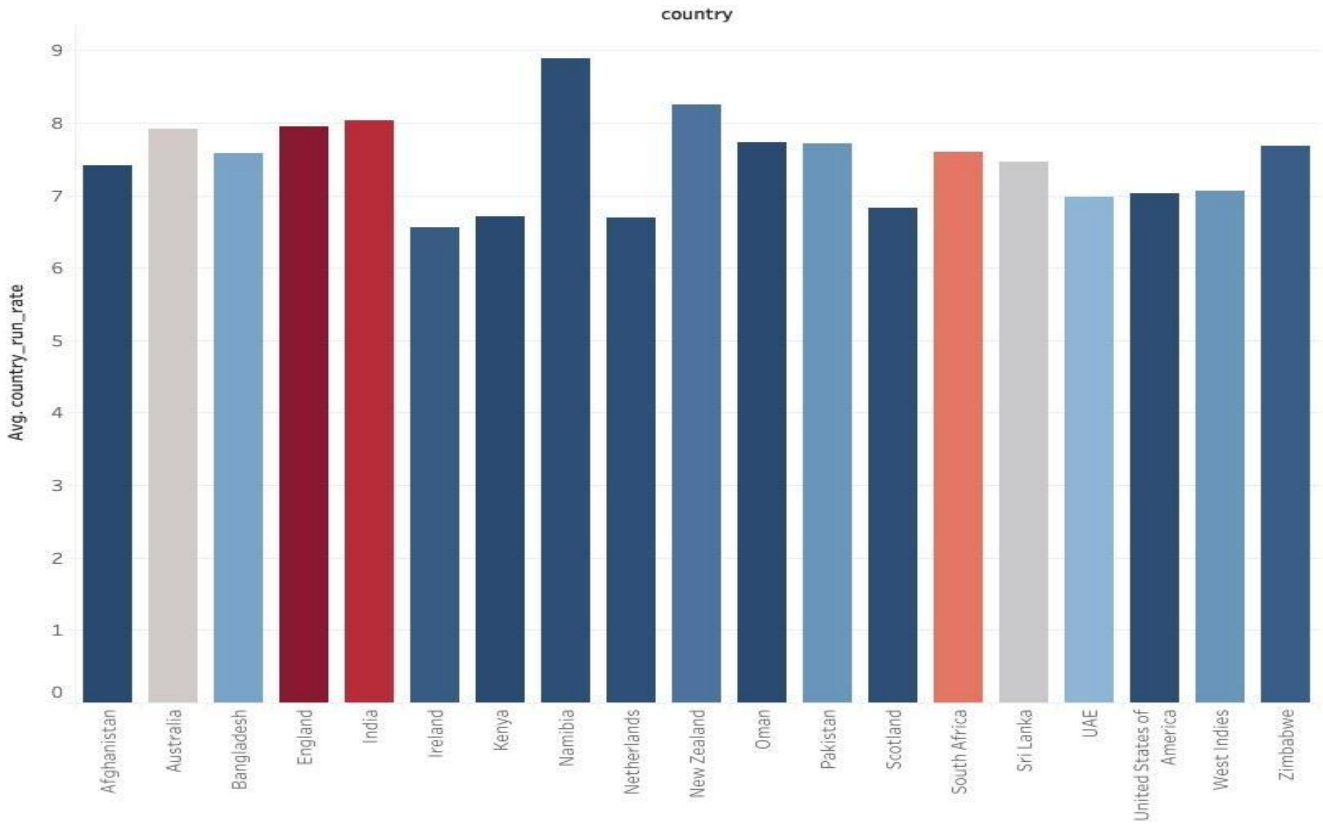



Distribution of Scores in Australia
1st Innings Score
240
220
200
180
160
140
120
100
80
ket Ground, Woolloongabba, Brisbane
Sydney Cricket Ground
W.A.C.A. Ground, Perth
Adelaide Oval
Melbourne Cricket Ground
Bellerive Oval, Hobart
Stadium Australia, Sydney
Docklands Stadium, Melbourne
Sydney Showground Stadium
Manuka Oval, Canberra
Tony Ireland Stadium, Townsville
nds Stadium, South Geelong, Victoria
Aurora Stadium, Launceston
Traeger Park, Alice Springs
Perth Stadium
Carrara Oval
Geelong Cricket Ground
Ted Summerton Reserve, Moe
ational Sports Stadium, Coffs Harbour

LEFT – 1ST INNINGS
RIGHT– 2ND INNINGS

country	run_rate
Namibia	8.896875
New Zealand	8.655490
England	8.336887
India	8.213121
United States of America	8.197674
Zimbabwe	8.043573
Pakistan	8.005620
Australia	8.003973
Bangladesh	7.813804
South Africa	7.702358
Sri Lanka	7.592398

country	run_rate
New Zealand	7.729550
England	7.672362
India	7.581399
Namibia	7.478125
Pakistan	7.295720
Australia	7.292067
United States of America	7.178295
Zimbabwe	7.155410
South Africa	7.047741
Bangladesh	7.025998
Afghanistan	7.000000


country
Avg. country_run_rate
0
1
2
3
4
5
6
7
8
9
Afghanistan
Australia
Bangladesh
England
India
Ireland
Kenya
Namibia
Netherlands
New Zealand
Oman
Pakistan
Scotland
South Africa
Sri Lanka
UAE
United States of America
West Indies
Zimbabwe

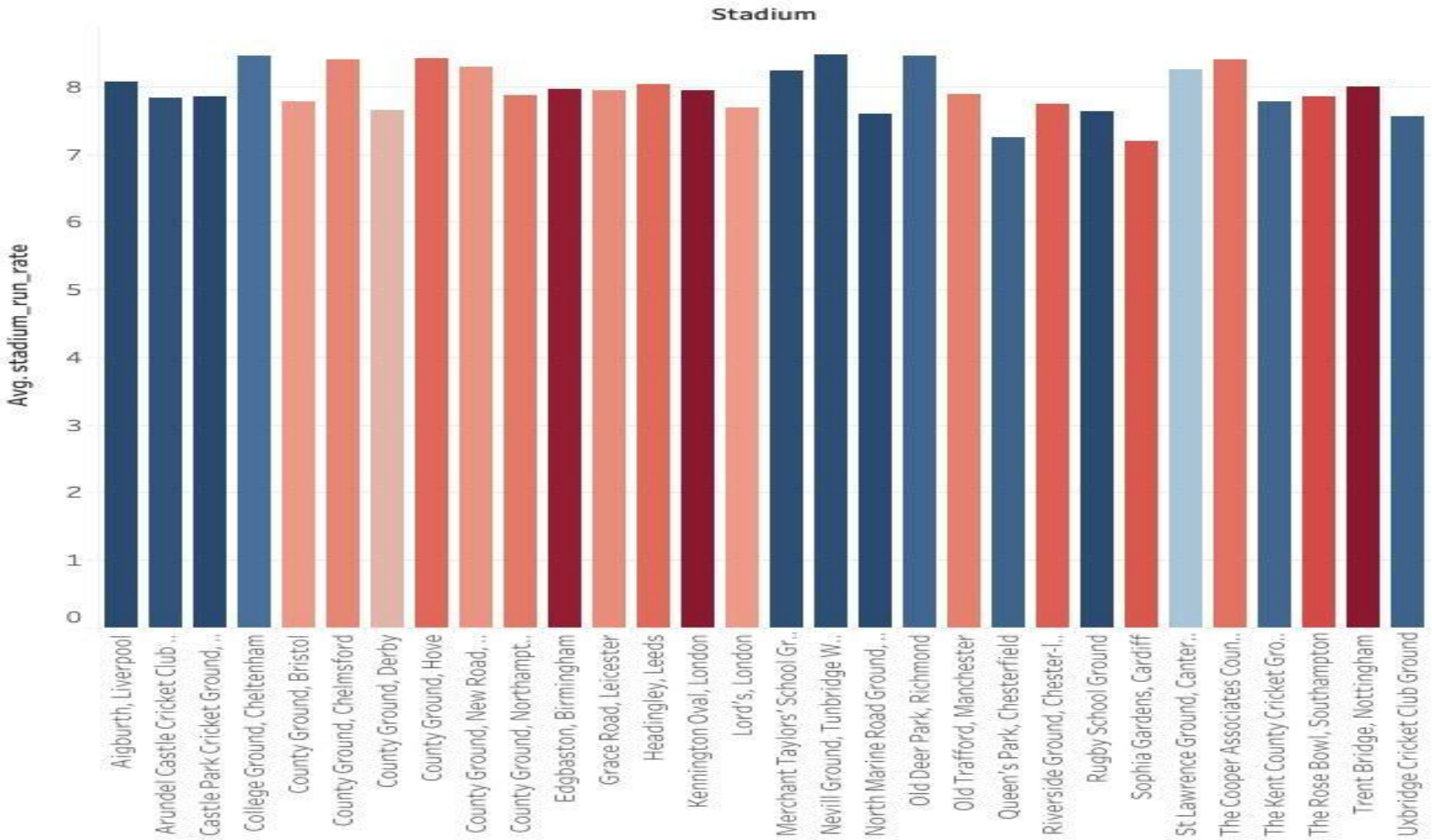
Stadiums by Run Rate

Stadium	stadium_run_rate
Merchant Taylors' School Ground, Northwood	8.892952
St Lawrence Ground, Canterbury	8.873454
County Ground, Hove	8.698080
The Cooper Associates County Ground, Taunton	8.608833
Nevill Ground, Tunbridge Wells	8.549427
County Ground, Chelmsford	8.549327
College Ground, Cheltenham	8.375129
Headingley, Leeds	8.332917
Trent Bridge, Nottingham	8.306841
Old Deer Park, Richmond	8.302117
Arundel Castle Cricket Club Ground	8.253655
County Ground, New Road, Worcester	8.252812

1st Innings

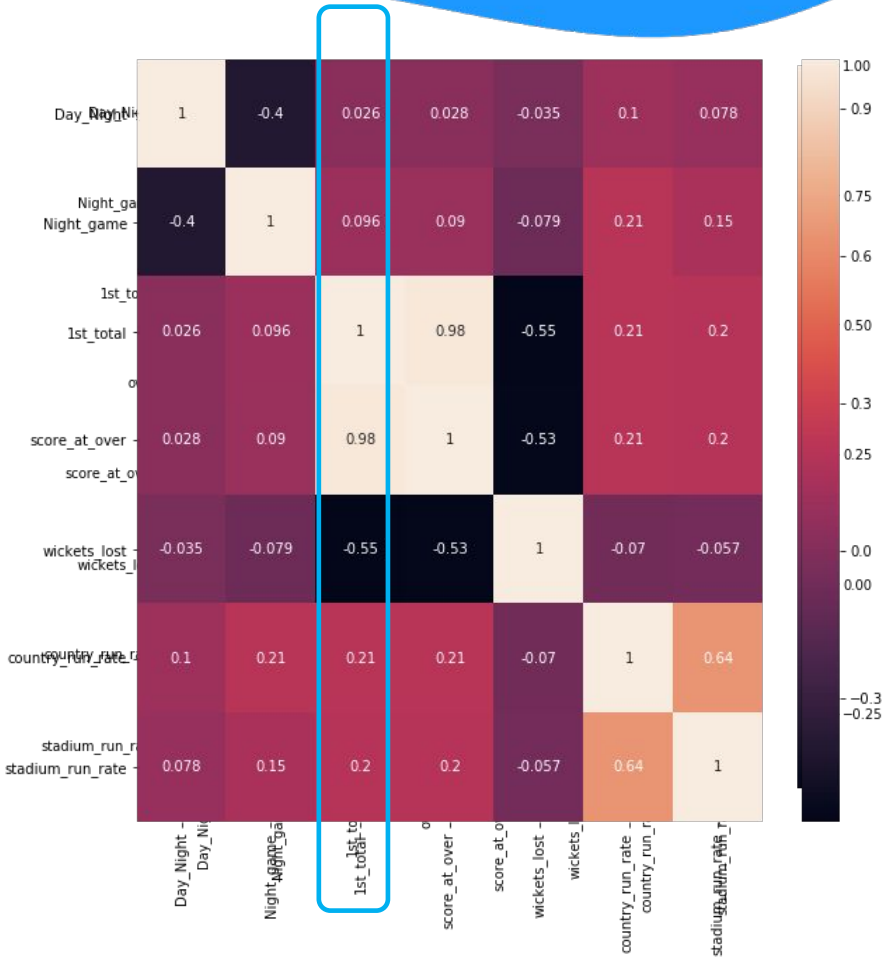
Stadium	stadium_run_rate
Old Deer Park, Richmond	8.623398
College Ground, Cheltenham	8.545952
Nevill Ground, Tunbridge Wells	8.392855
County Ground, New Road, Worcester	8.344737
County Ground, Chelmsford	8.262012
The Cooper Associates County Ground, Taunton	8.174346
County Ground, Hove	8.135564
The Kent County Cricket Ground, Beckenham	7.963038
Aigburth, Liverpool	7.935732
Edgbaston, Birmingham	7.903319
Old Trafford, Manchester	7.876192
County Ground, Bristol	7.874419

2nd Innings


Stadium
Avg.stadium_run_rate
0
1
2
3
4
5
6
7
8
Aigburth, Liverpool
Arundel Castle Cricket Club ..
Castle Park Cricket Ground, ..
College Ground, Cheltenham
County Ground, Bristol
County Ground, Chelmsford
County Ground, Derby
County Ground, Hove
County Ground, New Road, ..
County Ground, Northampt..
Edgbaston, Birmingham
Grace Road, Leicester
Headingley, Leeds
Kennington Oval, London
Lord's, London
Merchant Taylors' School Gr..
Nevill Ground, Tunbridge W..
North Marine Road Ground, ..
Old Deer Park, Richmond
Old Trafford, Manchester
Queen's Park, Chesterfield
Riverside Ground, Chester-l..
Rugby School Ground
Sophia Gardens, Cardiff
St Lawrence Ground, Canter..
The Cooper Associates Coun..
The Kent County Cricket Gro..
The Rose Bowl, Southampton
Trent Bridge, Nottingham
Uxbridge Cricket Club Ground

EDA AND FEATURE SELECTION

Correlations at over 19



Modelling (Part 1)

Features

- Dummied time of day (Day, Day/Night, Night)
- Score at over
- Wickets lost at over
- Country average run rate
- Stadium average run rate

r-squared scores

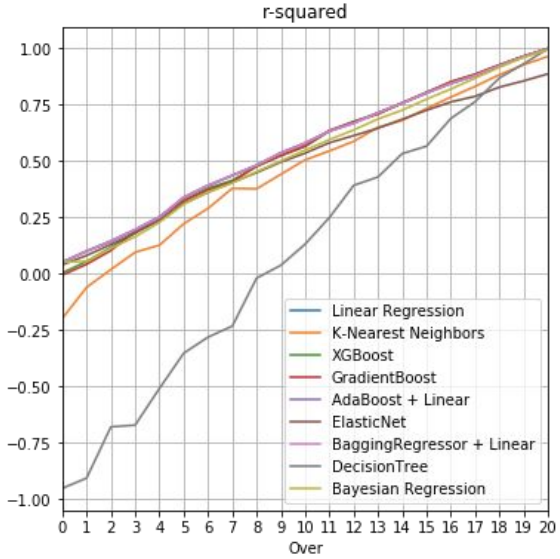
over	Linear Regression	K-Nearest Neighbors	XGBoost	GradientBoost	AdaBoost + Linear	ElasticNet	BaggingRegressor + Linear	DecisionTree	Bayesian Regression
0	0.0519119	-0.198636	0.00393352	-0.00646525	0.0481185	0.0392683	0.0498144	-0.953537	0.0516038
1	0.100582	-0.060764	0.0542073	0.0417531	0.0994976	0.0809709	0.0995572	-0.908029	0.0546928
2	0.145268	0.0184723	0.116387	0.102787	0.14077	0.129252	0.145424	-0.680441	0.112477
3	0.195981	0.0952377	0.180045	0.190448	0.194946	0.181302	0.19586	-0.672487	0.164085
4	0.252232	0.126476	0.251201	0.245352	0.245773	0.235841	0.25216	-0.509567	0.227929
5	0.339332	0.221875	0.322786	0.325385	0.337125	0.312002	0.340039	-0.353228	0.308297
6	0.390946	0.289968	0.379502	0.372212	0.391354	0.361413	0.392015	-0.281966	0.362893
7	0.436672	0.379088	0.414565	0.407826	0.435483	0.406286	0.436515	-0.232676	0.402453
8	0.481279	0.377044	0.474999	0.479301	0.48181	0.448793	0.481394	-0.0200488	0.45319
9	0.53634	0.441326	0.525711	0.523471	0.537135	0.495944	0.535563	0.0374191	0.500965
10	0.579814	0.505571	0.565054	0.567234	0.579811	0.535292	0.580208	0.131167	0.548534
11	0.633193	0.545056	0.633399	0.635847	0.634278	0.581608	0.63224	0.249494	0.595403
12	0.668979	0.587847	0.675676	0.674582	0.669979	0.613218	0.66796	0.392254	0.638392
13	0.714949	0.650061	0.711947	0.70874	0.714862	0.646044	0.714483	0.429644	0.687283
14	0.757539	0.680789	0.758175	0.758403	0.757309	0.686035	0.757625	0.53341	0.726346
15	0.802576	0.733542	0.805921	0.803696	0.802613	0.726074	0.802588	0.566442	0.775168
16	0.845147	0.784399	0.851491	0.852468	0.844676	0.762808	0.84495	0.689432	0.820057
17	0.880202	0.831861	0.88538	0.88543	0.87835	0.787789	0.880177	0.764217	0.868481
18	0.923935	0.883739	0.926952	0.927003	0.923955	0.828549	0.923805	0.869554	0.916663
19	0.964506	0.928275	0.966155	0.966177	0.963891	0.856399	0.964512	0.931741	0.9605

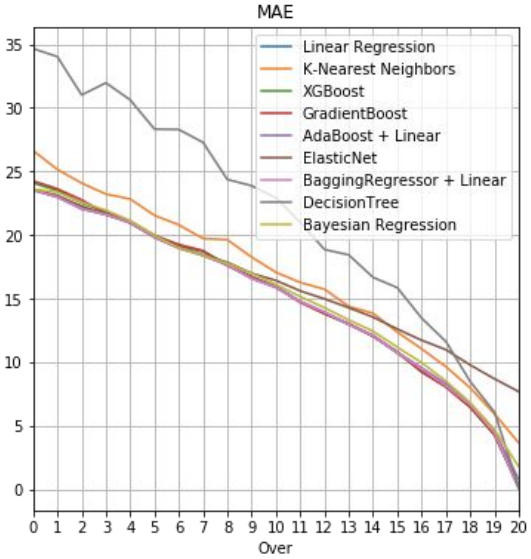
Mean Absolute Error

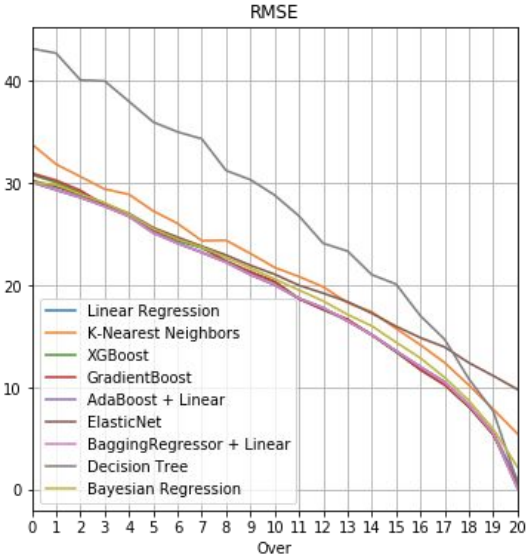
over	Linear Regression	K-Nearest Neighbors	XGBoost	GradientBoost	AdaBoost + Linear	ElasticNet	BaggingRegressor + Linear	DecisionTree	Bayesian Regression
0	23.5209	26.6303	24.0778	24.2415	23.5706	23.5706	23.5353	34.6546	23.5884
1	22.9783	25.1466	23.4862	23.6228	22.9821	23.1077	22.9921	34.0178	23.3663
2	22.0344	24.0736	22.5261	22.7529	22.088	22.2628	22.0399	31.0085	22.5642
3	21.5996	23.2043	21.802	21.6652	21.5918	21.6378	21.6131	31.9588	21.9354
4	20.9567	22.8236	20.9675	21.032	21.0311	20.9384	20.9564	30.6151	21.1289
5	19.8249	21.5324	19.9649	19.9578	19.8209	19.8312	19.8202	28.3168	19.9687
6	19.0024	20.7886	19.1097	19.2463	18.9417	19.0681	18.9864	28.2912	18.9774
7	18.4051	19.7128	18.7096	18.755	18.3832	18.4313	18.4067	27.2713	18.3963
8	17.6419	19.6142	17.6714	17.5798	17.6444	17.8371	17.6478	24.3594	17.7566
9	16.5735	18.2528	16.6957	16.697	16.6052	16.982	16.5653	23.8665	16.9767
10	15.8709	17.0562	15.9432	15.9212	15.8882	16.414	15.8715	22.9091	16.0965
11	14.7333	16.2611	14.7528	14.6835	14.7758	15.597	14.7327	20.9929	15.1541
12	13.9212	15.7259	13.7927	13.8088	13.9237	14.9672	13.9315	18.8523	14.2736
13	12.9497	14.3722	12.977	13.0322	12.9816	14.2715	12.9546	18.4276	13.3107
14	12.074	13.8385	12.0321	12.0264	12.1114	13.5321	12.0778	16.6581	12.4215
15	10.7307	12.3548	10.7302	10.8118	10.7622	12.5937	10.7283	15.8498	11.1686
16	9.46694	11.0401	9.26163	9.21532	9.54574	11.7146	9.46593	13.4564	9.95109
17	8.14205	9.63415	8.06078	8.04539	8.30374	10.9712	8.1393	11.6141	8.50546
18	6.63786	7.9671	6.44861	6.45206	6.65586	9.77077	6.64292	8.46753	6.80774
19	4.42474	5.92807	4.28164	4.27851	4.52867	8.67973	4.43268	6.06725	4.69125

Root Mean Squared Error

over	Linear Regression	K-Nearest Neighbors	XGBoost	GradientBoost	AdaBoost + Linear	ElasticNet	BaggingRegressor + Linear	Decision Tree	Bayesian Regression
0	30.0447	33.7822	30.7955	30.9559	30.1048	30.2444	30.0779	43.1275	30.0458
1	29.3002	31.82	30.0461	30.2433	29.3179	29.6179	29.3169	42.6759	29.8309
2	28.5631	30.6085	29.0416	29.2643	28.6381	28.8294	28.5605	40.0499	28.8991
3	27.7028	29.3871	27.976	27.7979	27.7206	27.9545	27.7049	39.955	28.053
4	26.7161	28.8754	26.7345	26.8388	26.8313	27.0074	26.7174	37.9591	26.9573
5	25.112	27.253	25.4245	25.3757	25.1539	25.6262	25.0986	35.9398	25.5158
6	24.1112	26.0333	24.3366	24.4792	24.1031	24.6888	24.09	34.9807	24.491
7	23.1884	24.3447	23.639	23.7747	23.2129	23.8056	23.1916	34.3016	23.7169
8	22.2514	24.3848	22.3857	22.2938	22.24	22.9376	22.2489	31.2033	22.6886
9	21.0373	23.0924	21.2771	21.3272	21.0193	21.9346	21.0549	30.3116	21.6694
10	20.0268	21.7241	20.3755	20.3243	20.0268	21.0611	20.0174	28.7977	20.6072
11	18.7115	20.8386	18.7063	18.6437	18.6838	19.984	18.7358	26.765	19.5159
12	17.7753	19.8344	17.5946	17.6243	17.7485	19.2142	17.8027	24.0853	18.4452
13	16.495	18.2762	16.5816	16.6737	16.4975	18.3808	16.5084	23.3326	17.126
14	15.1541	17.3879	15.1342	15.127	15.1613	17.2444	15.1514	21.0221	16.0091
15	13.5603	15.7537	13.4449	13.5218	13.559	15.973	13.5599	20.0952	14.4206
16	12.0096	14.1708	11.761	11.7223	12.0279	14.8634	12.0172	17.0078	12.8736
17	10.4806	12.4164	10.2516	10.2494	10.5613	13.9491	10.4817	14.7034	10.9816
18	8.26545	10.2186	8.09987	8.09707	8.26439	12.4092	8.27254	10.8241	8.69794
19	5.53861	7.87336	5.40842	5.40672	5.58644	11.1405	5.53816	7.68077	5.93464

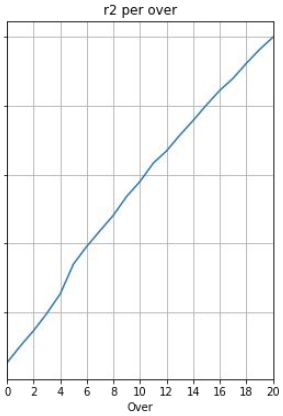

r-squared
Over
Linear Regression
K-Nearest Neighbors
XGBoost
GradientBoost
AdaBoost + Linear
ElasticNet
BaggingRegressor + Linear
DecisionTree
Bayesian Regression

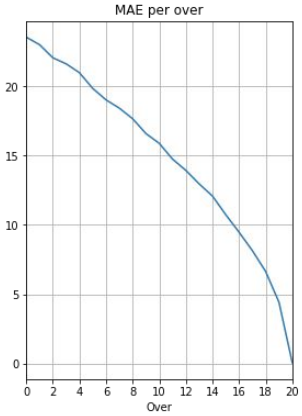

MAE
Over
Linear Regression
K-Nearest Neighbors
XGBoost
GradientBoost
AdaBoost + Linear
ElasticNet
BaggingRegressor + Linear
DecisionTree
Bayesian Regression

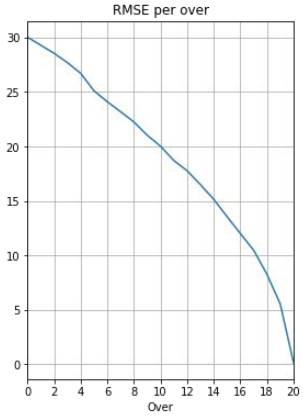

RMSE
Over
Linear Regression
K-Nearest Neighbors
XGBoost
GradientBoost
AdaBoost + Linear
ElasticNet
BaggingRegressor + Linear
Decision Tree
Bayesian Regression

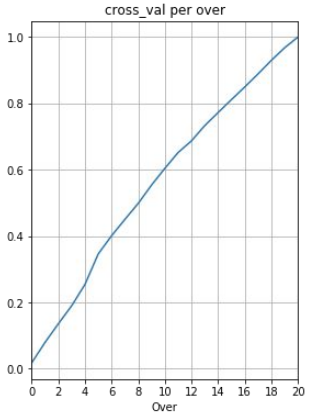
Refine and Compare (Part 1)

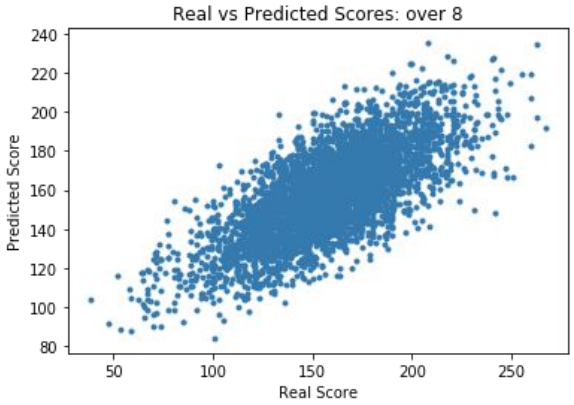
Over	r2	MAE	RMSE	cross_val
0	0.0519119	23.5209	30.0447	-0.0541269
1	0.100582	22.9783	29.3002	0.00383859
2	0.145268	22.0344	28.5631	0.0788545
3	0.195981	21.5996	27.7028	0.154139
4	0.252232	20.9567	26.7161	0.219608
5	0.339332	19.8249	25.112	0.306117
6	0.390946	19.0024	24.1112	0.381231
7	0.436672	18.4051	23.1884	0.427177
8	0.481279	17.6419	22.2514	0.479402
9	0.53634	16.5735	21.0373	0.536908
10	0.579814	15.8709	20.0268	0.584454
11	0.633193	14.7333	18.7115	0.63792
12	0.668979	13.9212	17.7753	0.679991
13	0.714949	12.9497	16.495	0.718691
14	0.757539	12.074	15.1541	0.761736
15	0.802576	10.7307	13.5603	0.805459
16	0.845147	9.46694	12.0096	0.848558
17	0.880202	8.14205	10.4806	0.888172
18	0.923935	6.63786	8.26545	0.92779
19	0.964506	4.42474	5.53861	0.965943
20	0.999978	0.0658222	0.138275	0.999306

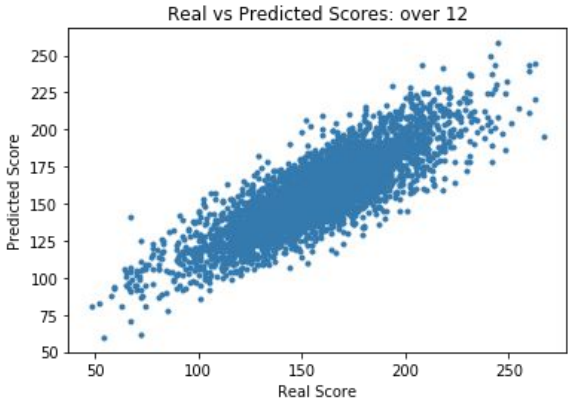


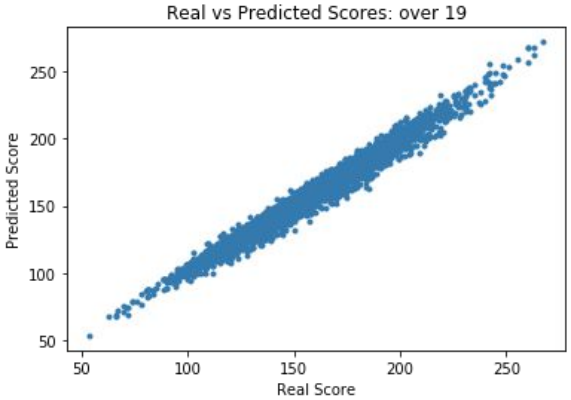


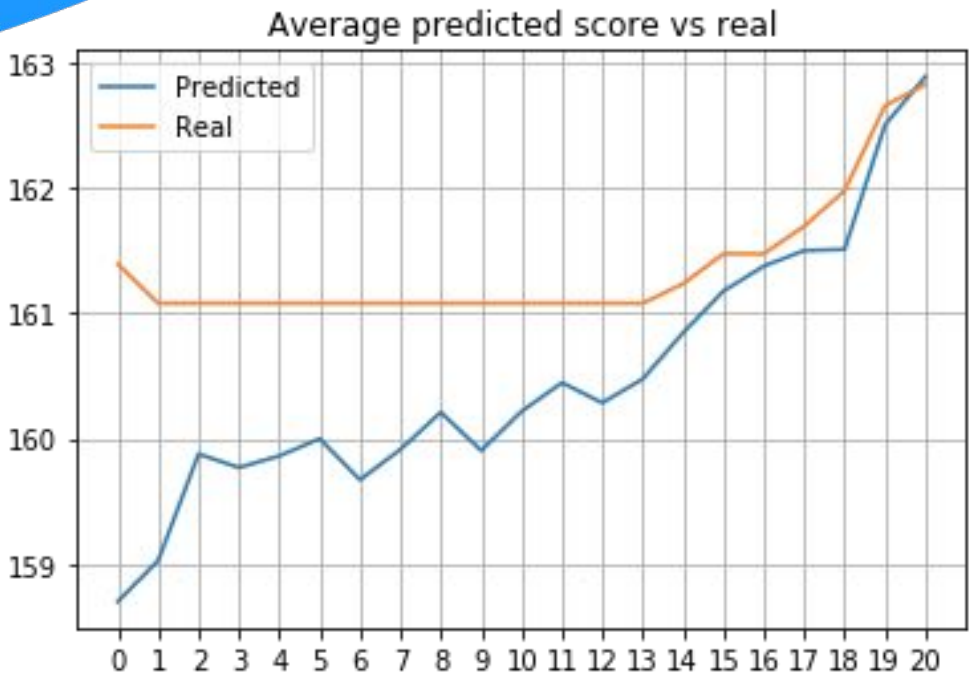




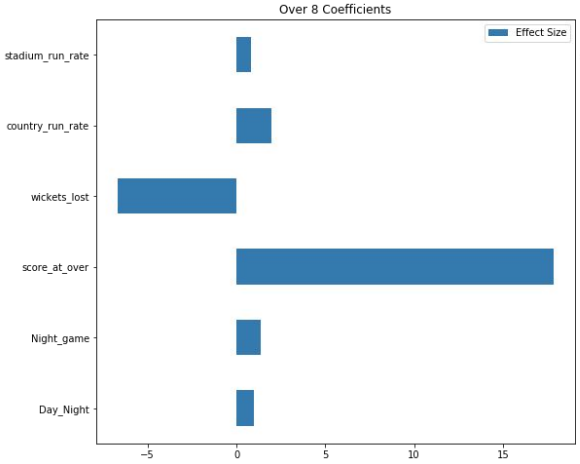

Real vs Predicted Scores: over 8
Predicted Score
Real Score
240
220
200
180
160
140
120
100
80
50
100
150
200
250

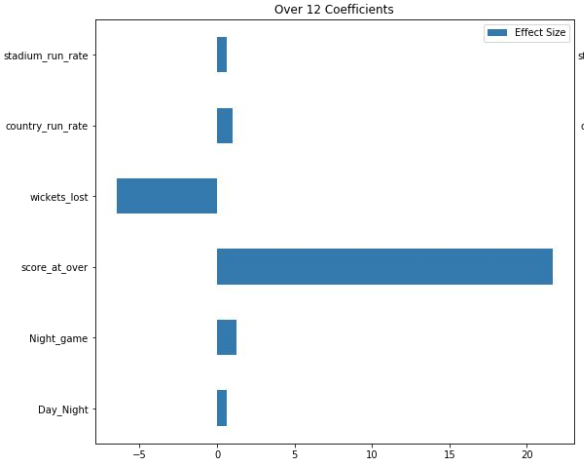

Real vs Predicted Scores: over 12
Predicted Score
Real Score
250
225
200
175
150
125
100
75
50
50
100
150
200
250

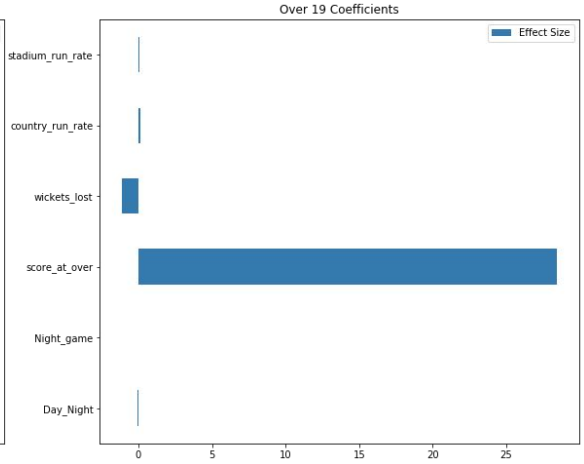

Real vs Predicted Scores: over 19
Predicted Score
Real Score
250
200
150
100
50
50
100
150
200
250


Average predicted score vs real
Predicted
Real
163
162
161
160
159
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20

over	Day_Night	Night_game	score_at_over	wickets_lost	country_run_rate	stadium_run_rate
0	0.69412	1.94366	8.88178e-16	0	4.32325	2.84595
1	1.05213	2.14476	5.08122	-3.57793	3.86395	2.8085
2	1.10256	1.98776	8.15804	-5.17288	3.51244	2.39272
3	0.971542	1.84489	10.0556	-5.84989	3.28712	2.20452
4	1.11668	1.93734	12.0206	-6.2454	3.02911	1.64227
5	0.899527	1.74513	14.1367	-6.81947	2.65059	1.33453
6	1.0257	1.72321	15.5874	-6.87493	2.31548	1.19836
7	1.02222	1.64167	16.533	-7.19064	2.20515	0.90713
8	0.969055	1.38291	17.8408	-6.67215	1.97468	0.804796
9	0.805982	1.42182	18.5608	-7.08826	1.90315	0.750876
10	0.716268	1.1279	19.5866	-7.13091	1.70621	0.700021
11	0.757489	1.17528	20.577	-6.94324	1.20324	0.699231
12	0.628812	1.22631	21.6418	-6.46949	0.988151	0.633939
13	0.507444	1.04526	22.5145	-6.30958	0.839822	0.604622
14	0.245768	0.665571	23.297	-5.88098	0.735098	0.512314
15	0.164211	0.463468	24.2887	-5.24715	0.497863	0.683297
16	0.0429207	0.366857	25.2968	-4.58523	0.338129	0.615641
17	-0.108638	0.21255	26.3241	-3.81248	0.31932	0.407333
18	-0.0887813	0.145977	27.5287	-2.4264	0.264374	0.212233
19	-0.0474078	0.00617813	28.4647	-1.13198	0.138774	0.0928619
20	-0.0158586	-0.00810433	29.1223	0.00880171	0.0341547	0.00611548


Over 8 Coefficients
Effect Size
stadium_run_rate
country_run_rate
wickets_lost
score_at_over
Night_game
Day_Night
−5
0
5
10
15


Over 12 Coefficients
Effect Size
stadium_run_rate
country_run_rate
wickets_lost
score_at_over
Night_game
Day_Night
−5
0
5
10
15
20


Over 19 Coefficients
Effect Size
stadium_run_rate
country_run_rate
wickets_lost
score_at_over
Night_game
Day_Night
0
5
10
15
20
25

Run Rate

$$Run\ Rate = \frac{Total\ Runs\ Scored}{Total\ Overs\ Bowled}$$

Example: If a team scores 63 runs in 9 overs, they have a run rate of 7.

If they stay at this rate, they will score 77 runs in the remaining 11 overs.

This gives a predicted score of 140

Duckworth-Lewis-Stern

Example: If a team has scored 63 runs in 9 overs, losing 3 wickets, by DLS they have 56.7% resources available.

Therefore, 63 runs is equal to the 43.3% of resources used.

Predicted score is 63/0.433 ≈ 145

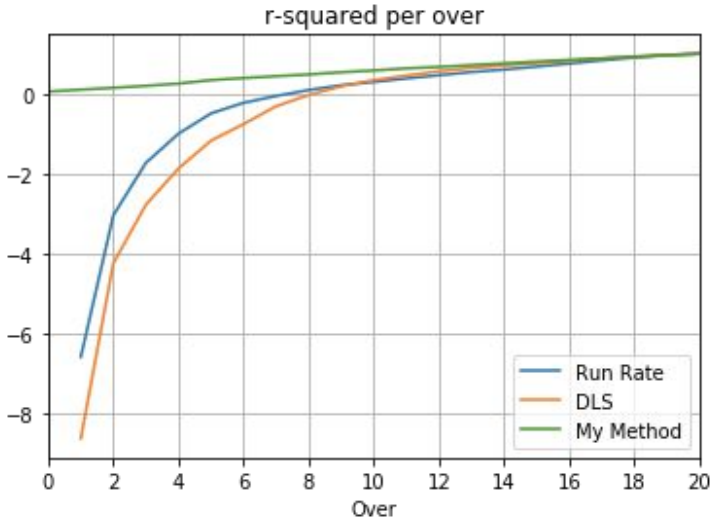
	0	1	2	3	4	5	6	7	8	9	10
0	100.0	96.8	92.6	86.7	78.8	68.2	54.4	37.5	21.3	8.3	0.0
1	96.1	93.3	89.2	83.9	76.7	66.6	53.5	37.3	21.0	8.3	0.0
2	92.2	89.6	85.9	81.1	74.2	65.0	52.7	36.9	21.0	8.3	0.0
3	88.2	85.7	82.5	77.9	71.7	63.3	51.6	36.6	21.0	8.3	0.0
4	84.1	81.8	79.0	74.7	69.1	61.3	50.4	36.2	20.8	8.3	0.0
5	79.9	77.9	75.3	71.6	66.4	59.2	49.1	35.7	20.8	8.3	0.0
6	75.4	73.7	71.4	68.0	63.4	56.9	47.7	35.2	20.8	8.3	0.0
7	71.0	69.4	67.3	64.5	60.4	54.4	46.1	34.5	20.7	8.3	0.0
8	66.4	65.0	63.3	60.6	57.1	51.9	44.3	33.6	20.5	8.3	0.0
9	61.7	60.4	59.0	56.7	53.7	49.1	42.4	32.7	20.3	8.3	0.0
10	56.7	55.8	54.4	52.7	50.0	46.1	40.3	31.6	20.1	8.3	0.0
11	51.8	51.1	49.8	48.4	46.1	42.8	37.8	30.2	19.8	8.3	0.0
12	46.6	45.9	45.1	43.8	42.0	39.4	35.2	28.6	19.3	8.3	0.0
13	41.3	40.8	40.1	39.2	37.8	35.5	32.2	26.9	18.6	8.3	0.0
14	35.9	35.5	35.0	34.3	33.2	31.4	29.0	24.6	17.8	8.1	0.0
15	30.4	30.0	29.7	29.2	28.4	27.2	25.3	22.1	16.6	8.1	0.0
16	24.6	24.4	24.2	23.9	23.3	22.4	21.2	18.9	14.8	8.0	0.0
17	18.7	18.6	18.4	18.2	18.0	17.5	16.8	15.4	12.7	7.4	0.0
18	12.7	12.5	12.5	12.4	12.4	12.0	11.7	11.0	9.7	6.5	0.0
19	6.4	6.4	6.4	6.4	6.4	6.2	6.2	6.0	5.7	4.4	0.0
20	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

COMPARISON

r-squared

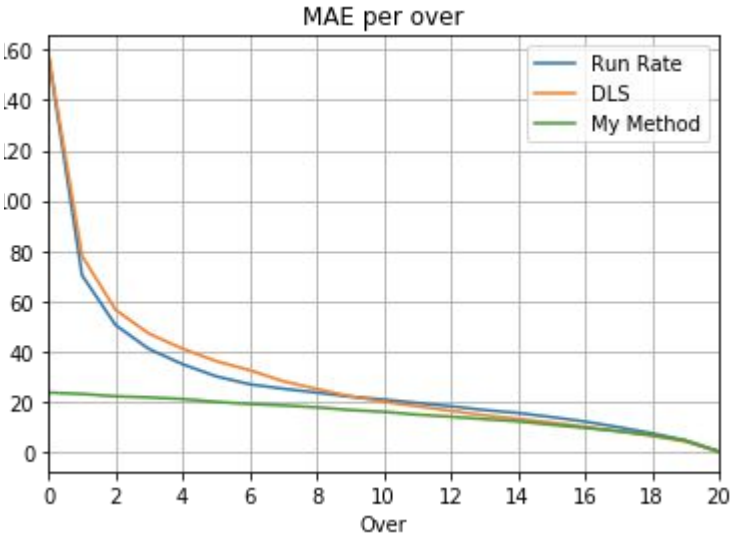
Over	Run Rate	DLS	My Method
0	nan	nan	0.0519119
1	-6.59742	-8.63974	0.100582
2	-3.04566	-4.24825	0.145268
3	-1.73308	-2.78241	0.195981
4	-1.00342	-1.8764	0.252232
5	-0.494712	-1.17731	0.339332
6	-0.229761	-0.762637	0.390946
7	-0.0584694	-0.315596	0.436672
8	0.0908041	-0.0354663	0.481279
9	0.209991	0.18631	0.53634
10	0.293271	0.337606	0.579814
11	0.379819	0.452655	0.633193
12	0.460246	0.559579	0.668979
13	0.540503	0.644455	0.714949
14	0.601973	0.705346	0.757539
15	0.677607	0.764089	0.802576
16	0.753086	0.82219	0.845147
17	0.833083	0.87955	0.880202
18	0.903646	0.927909	0.923935
19	0.962282	0.969572	0.964506
20	1	1	0.999978

ba359e806fd46bdb382fb0413c076...



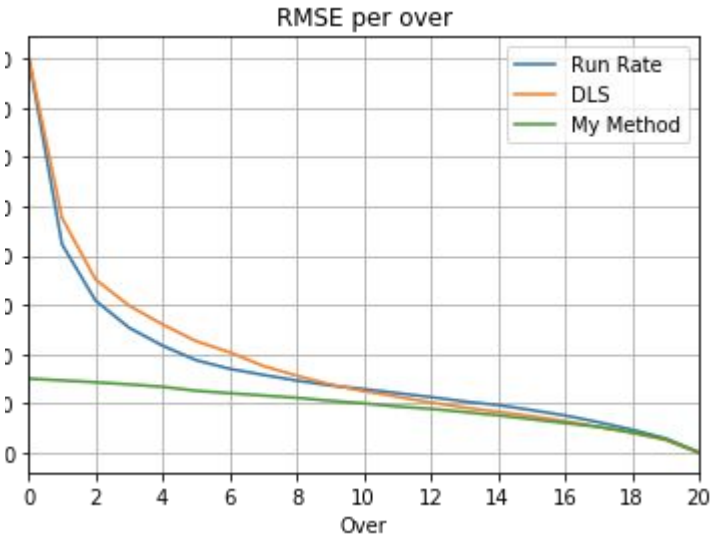
Mean Absolute Error

Over	Run Rate	DLS	My Method
0	157.738	157.738	23.5209
1	70.2623	77.9735	22.9783
2	50.4271	56.4818	22.0344
3	40.862	46.992	21.5996
4	34.8284	40.9933	20.9567
5	30.0139	36.0496	19.8249
6	26.843	32.445	19.0024
7	25.066	28.0631	18.4051
8	23.5185	24.9664	17.6419
9	21.9165	22.0379	16.5735
10	20.7565	19.8193	15.8709
11	19.4005	18.0812	14.7333
12	18.0668	16.2656	13.9212
13	16.5655	14.4681	12.9497
14	15.4081	13.0703	12.074
15	13.7764	11.5705	10.7307
16	11.9792	9.94777	9.46694
17	9.74486	8.15349	8.14205
18	7.32295	6.3003	6.63786
19	4.45788	4.06596	4.42474
20	0	0	0.0658222



Root Mean Squared Error

	Run Rate	DLS	My Method
Over			
0	160.728	160.728	30.0447
1	84.5641	95.2545	29.3002
2	61.7089	70.2846	28.5631
3	50.72	59.6674	27.7028
4	43.4249	52.0329	26.7161
5	37.5087	45.2703	25.112
6	34.0223	40.7319	24.1112
7	31.564	35.1896	23.1884
8	29.2563	31.2218	22.2514
9	27.2713	27.6771	21.0373
10	25.7939	24.9717	20.0268
11	24.1629	22.6997	18.7115
12	22.5418	20.3622	17.7753
13	20.7653	18.2661	16.495
14	19.3058	16.6107	15.1541
15	17.2731	14.7758	13.5603
16	15.082	12.7986	12.0096
17	12.3747	10.5121	10.4806
18	9.35315	8.0903	8.26545
19	5.79713	5.20684	5.53861
20	0	0	0.138275



Modelling (Part 2)

Features

- Dummied time of day (Day, Day/Night, Night)
- Score at over
- Wickets lost at over
- Country average run rate
- Stadium average run rate
- Target Score

r-squared scores

	Linear Regression	K-Nearest Neighbors	XGBoost	GradientBoost	AdaBoost + Linear	ElasticNet	BaggingRegressor + Linear	DecisionTree
over								
0	0.381073	0.27751	0.370535	0.362524	0.300075	0.334855	0.380182	-0.126559
1	0.408461	0.332367	0.406398	0.398686	0.328619	0.355307	0.407569	-0.0812266
2	0.435449	0.341795	0.440104	0.435969	0.352127	0.382099	0.435477	-0.144841
3	0.465583	0.393737	0.477596	0.474484	0.403084	0.413412	0.465424	-0.00552868
4	0.507446	0.431813	0.52933	0.529094	0.464282	0.442821	0.507769	0.145764
5	0.531925	0.473775	0.548079	0.544078	0.500128	0.466125	0.532544	0.0550042
6	0.550301	0.518516	0.584919	0.581955	0.509833	0.482732	0.550908	0.130117
7	0.571545	0.545287	0.613717	0.608032	0.566706	0.502191	0.571033	0.230966
8	0.580771	0.546454	0.630823	0.631822	0.577798	0.513515	0.580799	0.345033
9	0.592392	0.615234	0.658509	0.656701	0.592697	0.527164	0.592362	0.363058
10	0.606007	0.634553	0.692027	0.688724	0.578774	0.53887	0.606004	0.339984
11	0.630982	0.676385	0.717423	0.717685	0.620623	0.56521	0.631303	0.472521
12	0.658426	0.700536	0.755851	0.752395	0.658442	0.594592	0.658018	0.474274
13	0.686241	0.733832	0.788144	0.79068	0.683443	0.621219	0.686524	0.59835
14	0.705554	0.747368	0.808705	0.813613	0.692551	0.643736	0.705569	0.639598
15	0.74214	0.804607	0.84184	0.842619	0.738092	0.683592	0.742026	0.725766
16	0.789402	0.836237	0.880106	0.878677	0.781572	0.72666	0.789629	0.782256
17	0.850117	0.874235	0.918122	0.919487	0.847228	0.781665	0.850547	0.84345
18	0.904074	0.902723	0.945696	0.946242	0.902035	0.819388	0.903794	0.907244
19	0.961032	0.932886	0.974688	0.975308	0.958476	0.860021	0.961132	0.959099

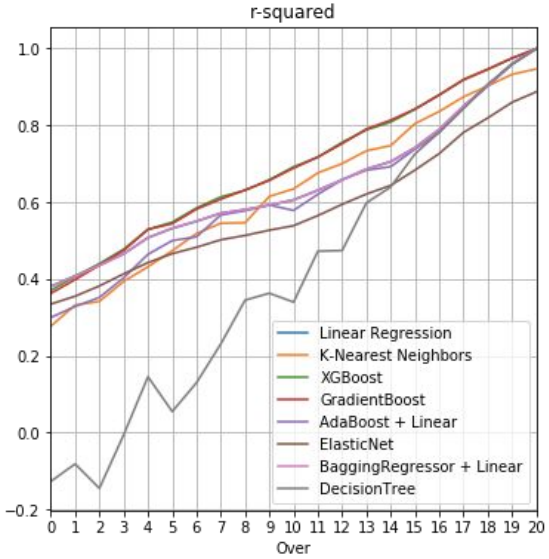
Mean Absolute Error

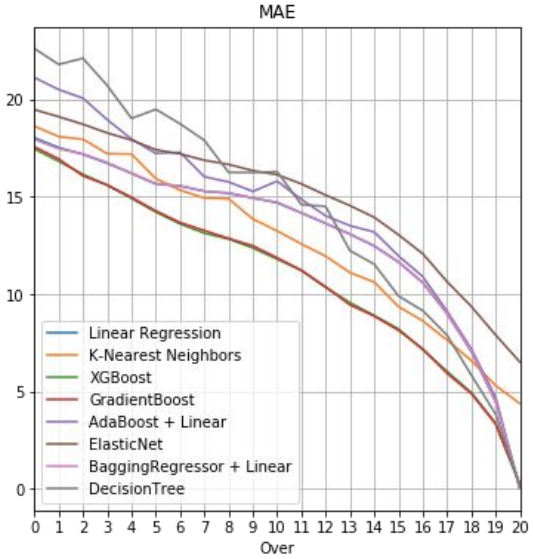
over	Linear Regression	K-Nearest Neighbors	XGBoost	GradientBoost	AdaBoost + Linear	ElasticNet	BaggingRegressor + Linear	DecisionTree
0	18.0284	18.6352	17.4558	17.571	21.124	19.4798	17.9456	22.6151
1	17.5277	18.0957	16.8253	16.9499	20.5104	19.1159	17.4657	21.7983
2	17.1861	17.9537	16.1593	16.0759	20.0656	18.7277	17.2015	22.1151
3	16.7278	17.2193	15.5938	15.6112	18.9637	18.2818	16.7662	20.7173
4	16.2078	17.1966	14.9244	14.9741	17.9663	17.9174	16.2141	19.0298
5	15.6634	15.9341	14.2232	14.2828	17.2334	17.4213	15.6849	19.4957
6	15.5613	15.3477	13.6114	13.6819	17.2865	17.2097	15.5599	18.7528
7	15.2894	14.9401	13.1331	13.2698	16.0335	16.8789	15.2876	17.9048
8	15.1889	14.9063	12.8408	12.8643	15.7639	16.6658	15.2123	16.2578
9	14.9436	13.8603	12.3742	12.4868	15.2787	16.3474	14.937	16.25
10	14.6998	13.2504	11.8005	11.8752	15.808	16.1406	14.7235	16.2937
11	14.173	12.5821	11.2204	11.2296	14.8714	15.6666	14.1894	14.5997
12	13.6373	11.9535	10.3436	10.3823	14.0249	15.0912	13.6306	14.5102
13	13.0882	11.1108	9.5726	9.45537	13.5206	14.5496	13.0809	12.2401
14	12.4772	10.6243	8.90695	8.87641	13.1958	13.9503	12.4931	11.5419
15	11.6551	9.36372	8.21908	8.14783	11.9686	13.0528	11.6539	9.91311
16	10.5953	8.63208	7.15666	7.1887	10.9051	12.0804	10.6025	9.15723
17	9.01637	7.65116	6.04076	5.93993	9.1296	10.647	9.00557	7.89867
18	7.0707	6.59356	4.94635	4.89894	7.19637	9.37489	7.0812	5.84615
19	4.4953	5.31143	3.37303	3.34315	4.70572	7.90956	4.48575	3.85306

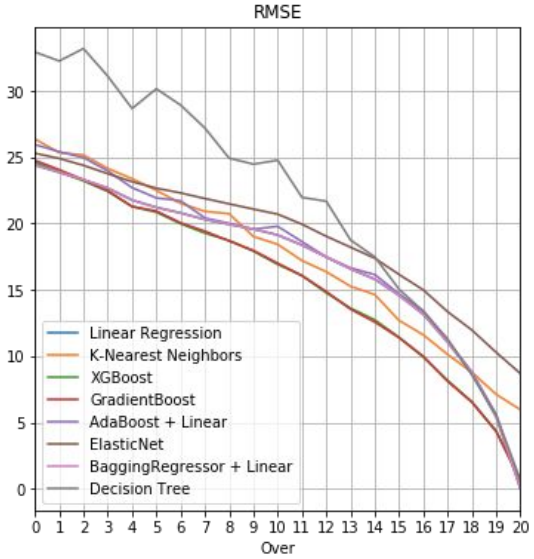
Root Mean Squared Error

	Linear Regression	K-Nearest Neighbors	XGBoost	GradientBoost	AdaBoost + Linear	ElasticNet	BaggingRegressor + Linear	Decision Tree
over								
0	24.4016	26.3642	24.6085	24.7646	25.9493	25.2963	24.4192	32.9212
1	23.8556	25.3436	23.8972	24.0519	25.4146	24.9044	23.8736	32.2521
2	23.3051	25.164	23.2088	23.2944	24.9657	24.3814	23.3045	33.1873
3	22.6746	24.1507	22.4183	22.485	23.9638	23.7556	22.678	31.1026
4	21.7684	23.38	21.2793	21.2846	22.7022	23.1524	21.7612	28.6674
5	21.2206	22.5001	20.8512	20.9433	21.9295	22.6631	21.2065	30.1519
6	20.7998	21.5224	19.9832	20.0544	21.7156	22.3078	20.7858	28.9287
7	20.3026	20.9155	19.2776	19.4189	20.4169	21.8842	20.3147	27.2002
8	19.9422	20.7424	18.7139	18.6886	20.0128	21.4824	19.9416	24.9263
9	19.578	19.0215	17.92	17.9673	19.5707	21.0864	19.5787	24.4736
10	19.1386	18.4323	16.9209	17.0114	19.789	20.7051	19.1387	24.771
11	18.3762	17.2086	16.0805	16.073	18.6323	19.9467	18.3682	21.9702
12	17.4771	16.3644	14.7759	14.8801	17.4767	19.0403	17.4875	21.6823
13	16.5796	15.2705	13.6237	13.5419	16.6534	18.2167	16.5721	18.7586
14	15.8043	14.6392	12.7387	12.5742	16.1495	17.3843	15.8039	17.485
15	14.5872	12.698	11.4242	11.3961	14.7012	16.1586	14.5904	15.0432
16	13.1618	11.6063	9.93082	9.98984	13.4042	14.9947	13.1547	13.3832
17	11.0623	10.1333	8.17623	8.10781	11.1684	13.3516	11.0464	11.3057
18	8.72589	8.78709	6.56531	6.53226	8.81814	11.9733	8.7386	8.58046
19	5.42994	7.12607	4.37629	4.32232	5.6052	10.2914	5.42298	5.563

COMPARISON

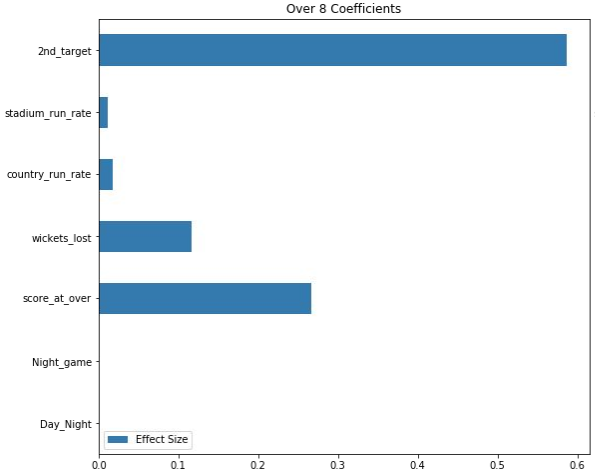

r-squared
Over
Linear Regression
K-Nearest Neighbors
XGBoost
GradientBoost
AdaBoost + Linear
ElasticNet
BaggingRegressor + Linear
DecisionTree

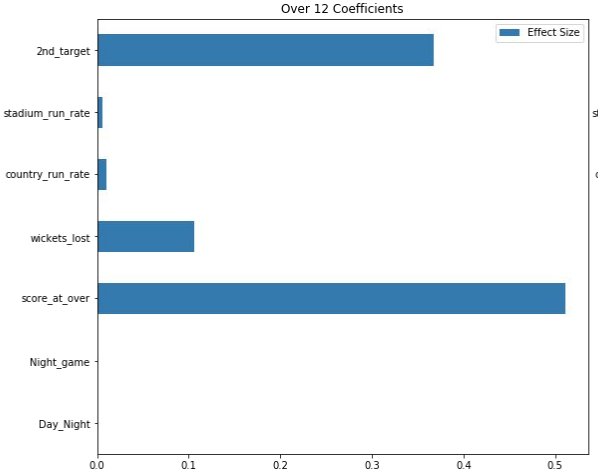

MAE
Over
Linear Regression
K-Nearest Neighbors
XGBoost
GradientBoost
AdaBoost + Linear
ElasticNet
BaggingRegressor + Linear
DecisionTree

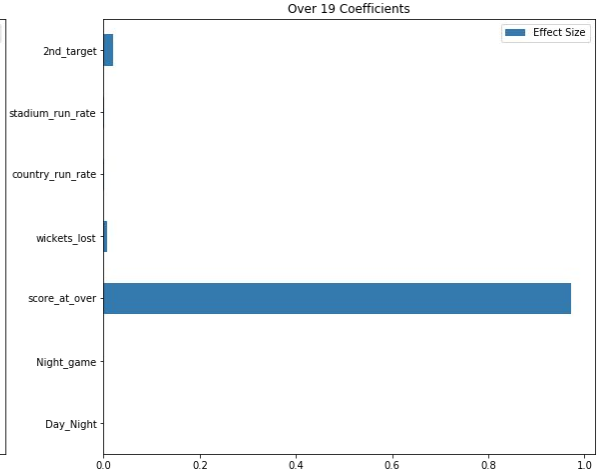

RMSE
Over
Linear Regression
K-Nearest Neighbors
XGBoost
GradientBoost
AdaBoost + Linear
ElasticNet
BaggingRegressor + Linear
Decision Tree

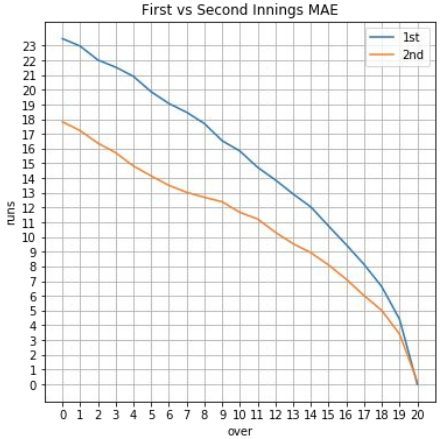
Refine and Compare (Part 2)

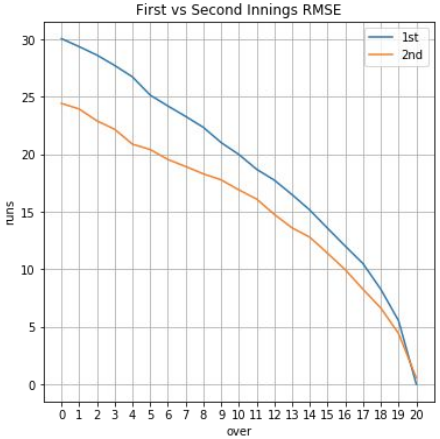
over	Day_Night	Night_game	score_at_over	wickets_lost	country_run_rate	stadium_run_rate	2nd_target
0	0.000515905	0.00366305	0	0	0.0666027	0.0735115	0.855707
1	0.000361786	0.00239526	0.0327476	0.0287161	0.054737	0.0627369	0.818305
2	0.000774107	0.00168262	0.0695413	0.0644746	0.0427719	0.0535572	0.767198
3	0.000265311	0.000816028	0.116858	0.079165	0.0422775	0.0437974	0.716821
4	6.8996e-06	0.000684722	0.151808	0.0921839	0.0311176	0.0363691	0.68783
5	0.000621632	0.000188761	0.181492	0.0961397	0.0266944	0.0292782	0.665585
6	0.000156885	0.00086488	0.223826	0.10295	0.0188972	0.028041	0.625264
7	0.00107639	0.000681915	0.255616	0.110232	0.0183957	0.0129746	0.601024
8	0.000435763	0.000509563	0.267219	0.116003	0.0178126	0.011539	0.586481
9	0.000455818	0.000403322	0.314757	0.106136	0.0185913	0.00943987	0.550217
10	0.000113953	9.53194e-05	0.35502	0.106874	0.0148125	0.00968832	0.513395
11	0.000375615	8.23521e-05	0.446059	0.0988414	0.010571	0.00725587	0.436815
12	0	0.000154102	0.510734	0.105617	0.0102885	0.00574964	0.367457
13	7.08509e-05	5.98294e-05	0.557809	0.0936309	0.00853967	0.00547884	0.334411
14	7.19771e-05	1.86428e-05	0.608765	0.0887662	0.00624179	0.00443017	0.291706
15	2.07057e-05	0.000123535	0.679426	0.075417	0.0037557	0.00268769	0.238569
16	8.41366e-05	0.000141579	0.773059	0.0563737	0.00307137	0.00193559	0.165335
17	0	0.000272576	0.851103	0.0384333	0.00235983	0.00136315	0.106469
18	1.01824e-05	9.57587e-05	0.914043	0.0202393	0.00154057	0.000499153	0.0635721
19	3.04352e-06	3.51671e-05	0.972938	0.00701654	0.000489171	0.0004716	0.0190463

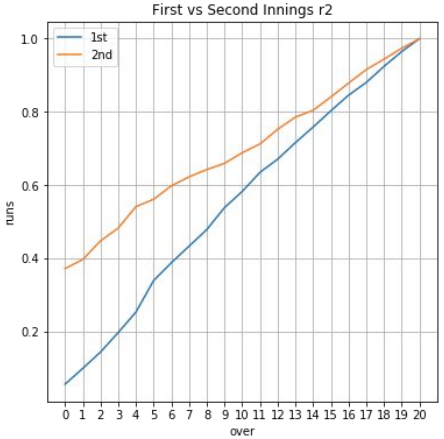

Over 8 Coefficients
2nd_target
stadium_run_rate
country_run_rate
wickets_lost
score_at_over
Night_game
Day_Night
Effect Size
0.0
0.1
0.2
0.3
0.4
0.5
0.6


Over 12 Coefficients
Effect Size
2nd_target
stadium_run_rate
country_run_rate
wickets_lost
score_at_over
Night_game
Day_Night
0.0
0.1
0.2
0.3
0.4
0.5


Over 19 Coefficients
Effect Size
2nd_target
stadium_run_rate
country_run_rate
wickets_lost
score_at_over
Night_game
Day_Night
0.0
0.2
0.4
0.6
0.8
1.0







1ST VS 2ND INNINGS COMPARISON

Modelling (Part 3)

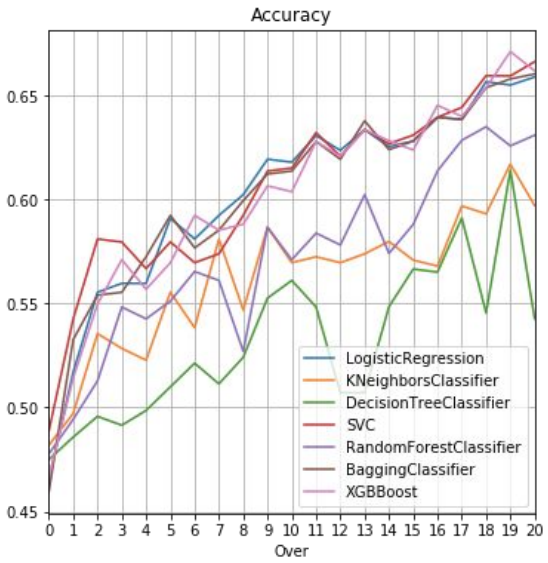
Features

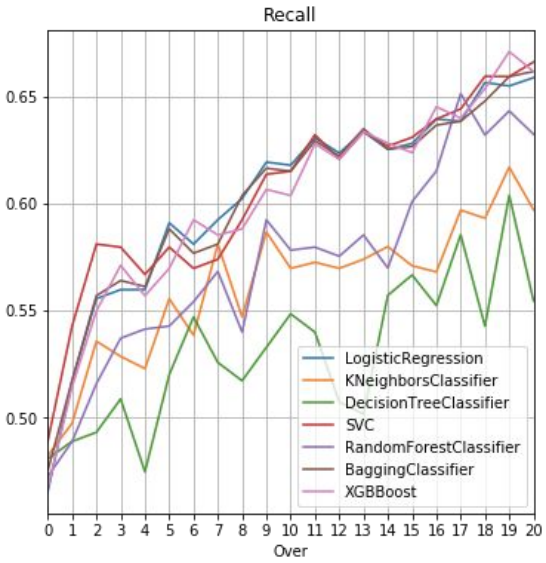
- Dummied time of day (Day, Day/Night, Night)
- Predicted Score
- Country average run rate
- Stadium average run rate
- Target Score (2nd Innings Only)

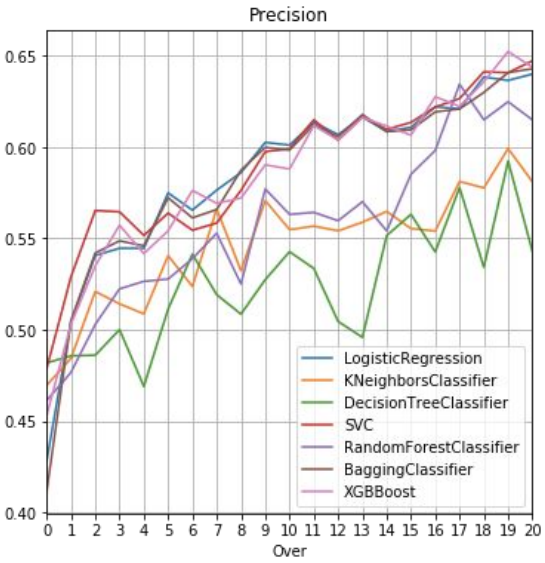
I st Innings

	LogisticRegression	KNeighborsClassifier	DecisionTreeClassifier	SVC	RandomForestClassifier	BaggingClassifier	XGBBoost
0	0.465278	0.481944	0.488889	0.488889	0.472222	0.463889	0.466667
1	0.517045	0.497159	0.494318	0.542614	0.484375	0.53267	0.514205
2	0.555398	0.535511	0.490057	0.580966	0.524148	0.553977	0.549716
3	0.559659	0.528409	0.49858	0.579545	0.535511	0.565341	0.571023
4	0.559659	0.522727	0.485795	0.566761	0.548295	0.56392	0.556818
5	0.590909	0.555398	0.509943	0.579545	0.56392	0.588068	0.569602
6	0.580966	0.538352	0.536932	0.569602	0.553977	0.580966	0.59233
7	0.59233	0.580966	0.508523	0.573864	0.56108	0.588068	0.585227
8	0.602273	0.546875	0.525568	0.59233	0.542614	0.599432	0.588068
9	0.619318	0.586648	0.541193	0.613636	0.59233	0.615057	0.606534
10	0.617898	0.569602	0.56108	0.615057	0.600852	0.619318	0.603693
11	0.630682	0.572443	0.544034	0.632102	0.576705	0.630682	0.627841
12	0.62358	0.569602	0.504261	0.620739	0.580966	0.619318	0.620739
13	0.633523	0.573864	0.505682	0.633523	0.588068	0.630682	0.633523
14	0.625356	0.579772	0.551282	0.626781	0.586895	0.625356	0.628205
15	0.62804	0.570815	0.575107	0.630901	0.579399	0.630901	0.623748
16	0.639485	0.567954	0.562232	0.639485	0.599428	0.639485	0.645207
17	0.638451	0.596844	0.578192	0.644189	0.644189	0.639885	0.639885
18	0.656566	0.593074	0.536797	0.659452	0.62482	0.656566	0.65368
19	0.654971	0.616959	0.612573	0.659357	0.646199	0.656433	0.671053
20	0.658824	0.597059	0.544118	0.666176	0.627941	0.660294	0.661765

COMPARISON

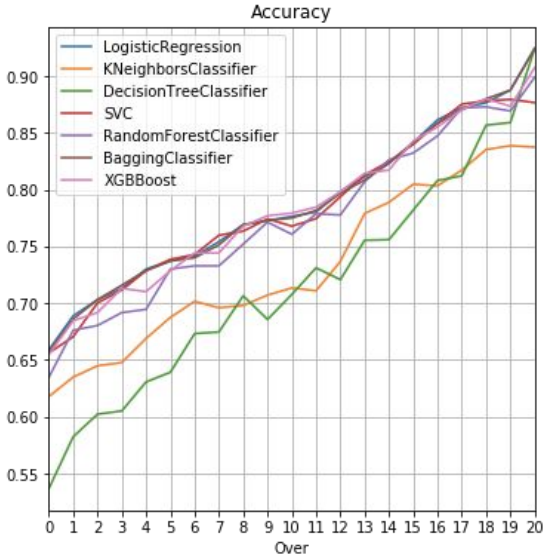

Accuracy
0.65
0.60
0.55
0.50
0.45
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20
Over
LogisticRegression
KNeighborsClassifier
DecisionTreeClassifier
SVC
RandomForestClassifier
BaggingClassifier
XGBBoost

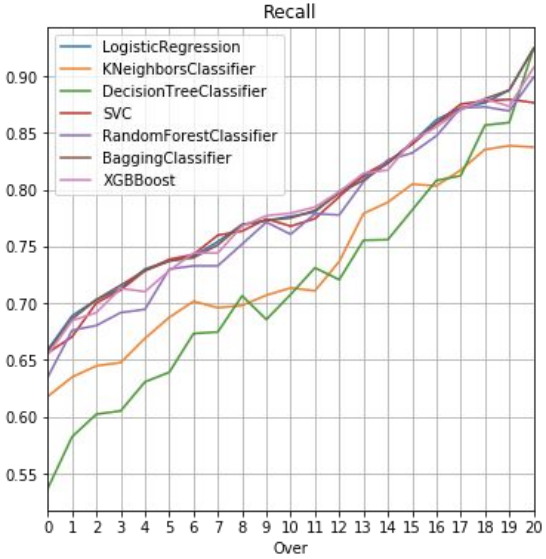

Recall
0.65
0.60
0.55
0.50
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20
Over
LogisticRegression
KNeighborsClassifier
DecisionTreeClassifier
SVC
RandomForestClassifier
BaggingClassifier
XGBBoost

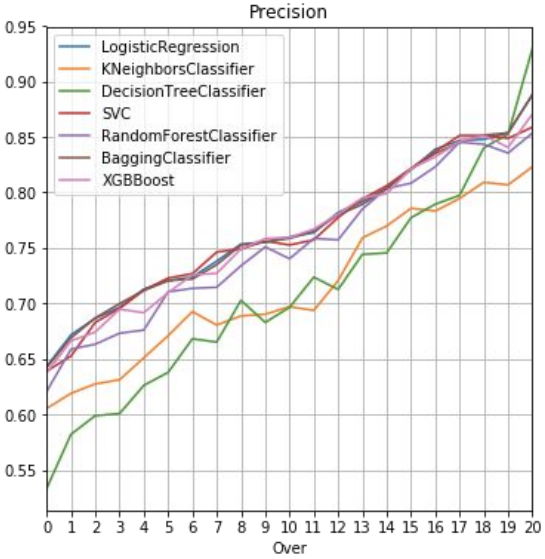

Precision
0.65
0.60
0.55
0.50
0.45
0.40
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20
Over
LogisticRegression
KNeighborsClassifier
DecisionTreeClassifier
SVC
RandomForestClassifier
BaggingClassifier
XGBBoost

2nd Innings

	LogisticRegression	KNeighborsClassifier	DecisionTreeClassifier	SVC	RandomForestClassifier	BaggingClassifier	XGBBoost
0	0.659091	0.617898	0.535511	0.65625	0.640625	0.65767	0.65483
1	0.68892	0.634943	0.59517	0.670455	0.678977	0.68892	0.684659
2	0.703125	0.644886	0.610795	0.700284	0.678977	0.703125	0.691761
3	0.713068	0.647727	0.612216	0.711648	0.698864	0.714489	0.713068
4	0.730114	0.669034	0.629261	0.728693	0.711648	0.724432	0.710227
5	0.737216	0.6875	0.646307	0.738636	0.731534	0.737216	0.728693
6	0.741477	0.701705	0.673295	0.742898	0.737216	0.742898	0.744318
7	0.754261	0.696023	0.649148	0.759943	0.735795	0.754261	0.744318
8	0.769231	0.698006	0.707977	0.763533	0.752137	0.764957	0.767806
9	0.772857	0.707143	0.678571	0.774286	0.764286	0.771429	0.777143
10	0.776504	0.713467	0.697708	0.767908	0.767908	0.775072	0.77937
11	0.780347	0.710983	0.715318	0.774566	0.776012	0.780347	0.784682
12	0.798246	0.736842	0.719298	0.79386	0.785088	0.798246	0.798246
13	0.811487	0.779087	0.759941	0.81296	0.802651	0.811487	0.814433
14	0.823353	0.788922	0.760479	0.82485	0.826347	0.820359	0.817365
15	0.841463	0.804878	0.785061	0.839939	0.839939	0.841463	0.842988
16	0.861635	0.803459	0.803459	0.858491	0.847484	0.860063	0.855346
17	0.872093	0.817276	0.80897	0.875415	0.868771	0.870432	0.870432
18	0.876565	0.83542	0.862254	0.878354	0.872987	0.878354	0.880143
19	0.887755	0.838776	0.861224	0.879592	0.869388	0.885714	0.873469
20	0.92437	0.837535	0.907563	0.876751	0.890756	0.92437	0.907563


Accuracy
LogisticRegression
KNeighborsClassifier
DecisionTreeClassifier
SVC
RandomForestClassifier
BaggingClassifier
XGBBoost
0.90
0.85
0.80
0.75
0.70
0.65
0.60
0.55
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20
Over


Recall
LogisticRegression
KNeighborsClassifier
DecisionTreeClassifier
SVC
RandomForestClassifier
BaggingClassifier
XGBBoost
0.90
0.85
0.80
0.75
0.70
0.65
0.60
0.55
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20
Over


Precision
LogisticRegression
KNeighborsClassifier
DecisionTreeClassifier
SVC
RandomForestClassifier
BaggingClassifier
XGBBoost
0.95
0.90
0.85
0.80
0.75
0.70
0.65
0.60
0.55
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20
Over

Refine and Compare (Part 3)

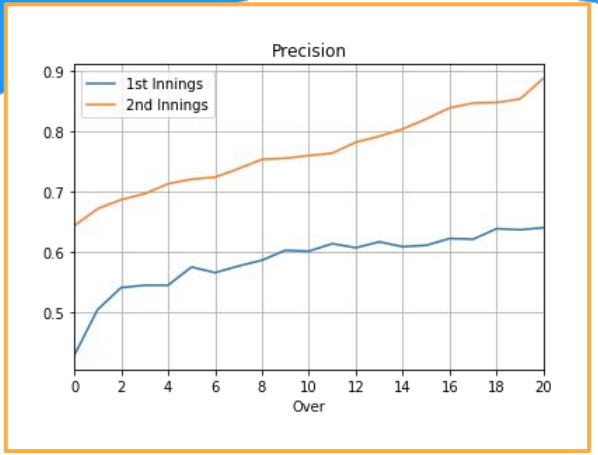
I st Innings

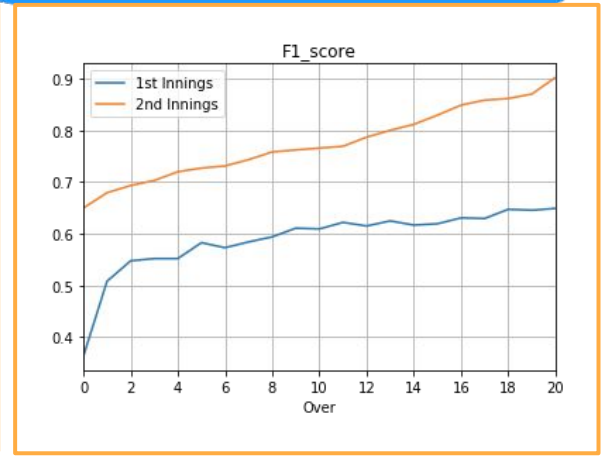
Over	accuracy	Recall	Precision	F1_score	crossval	ROC_AUC_Score	baseline
0	0.465278	0.465278	0.428056	0.363233	0.478685	0.461005	0.500469
1	0.517045	0.517045	0.504385	0.508087	0.526523	0.51603	0.50192
2	0.555398	0.555398	0.540691	0.54764	0.556621	0.546427	0.50192
3	0.559659	0.559659	0.544627	0.552017	0.579295	0.544632	0.50192
4	0.559659	0.559659	0.544627	0.552017	0.580163	0.555234	0.50192
5	0.590909	0.590909	0.574945	0.582811	0.594854	0.573216	0.50192
6	0.580966	0.580966	0.565544	0.573044	0.600196	0.584415	0.50192
7	0.59233	0.59233	0.57648	0.58426	0.602713	0.584981	0.50192
8	0.602273	0.602273	0.586086	0.594054	0.608619	0.595443	0.502041
9	0.619318	0.619318	0.602631	0.610857	0.622234	0.600044	0.502041
10	0.617898	0.617898	0.601155	0.609388	0.622822	0.602979	0.502041
11	0.630682	0.630682	0.613637	0.62204	0.632245	0.610582	0.502041
12	0.62358	0.62358	0.606877	0.615088	0.63446	0.619394	0.501921
13	0.633523	0.633523	0.616647	0.624916	0.63872	0.625045	0.501802
14	0.625356	0.625356	0.608558	0.616816	0.639789	0.631322	0.501083
15	0.62804	0.62804	0.610932	0.619365	0.639695	0.634819	0.500843
16	0.639485	0.639485	0.622205	0.63071	0.645595	0.645938	0.500603
17	0.638451	0.638451	0.620982	0.62958	0.650362	0.648539	0.499516
18	0.656566	0.656566	0.638453	0.64729	0.659409	0.652472	0.497207
19	0.654971	0.654971	0.636611	0.645623	0.661889	0.663249	0.493252
20	0.658824	0.658824	0.640096	0.649082	0.662466	0.667938	0.49379

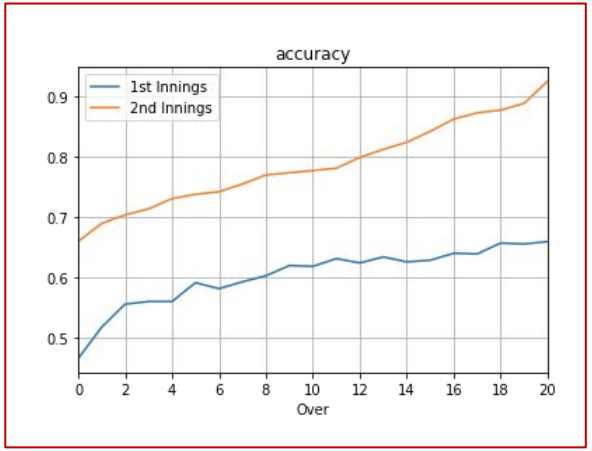
2nd Innings

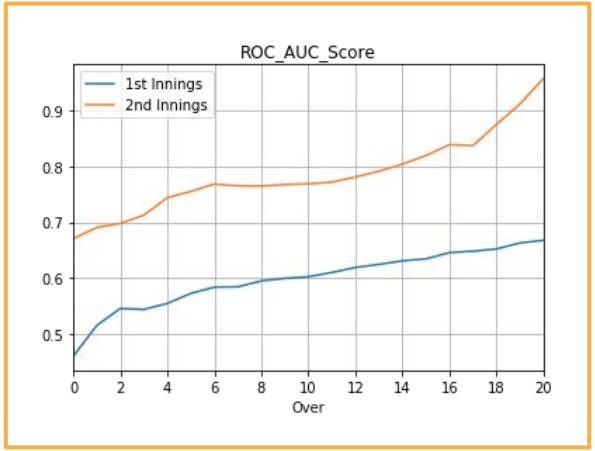
Over	accuracy	Recall	Precision	F1_score	crossval	ROC_AUC_Score	baseline
0	0.659091	0.659091	0.643213	0.649897	0.66214	0.670829	0.504561
1	0.68892	0.68892	0.671554	0.67936	0.692136	0.690652	0.50192
2	0.703125	0.703125	0.686518	0.693318	0.707644	0.697952	0.502041
3	0.713068	0.713068	0.696397	0.703023	0.722229	0.712811	0.50192
4	0.730114	0.730114	0.712823	0.719997	0.7348	0.743717	0.50192
5	0.737216	0.737216	0.720501	0.726996	0.745831	0.755083	0.50192
6	0.741477	0.741477	0.723696	0.73118	0.75627	0.768157	0.501801
7	0.754261	0.754261	0.738006	0.743426	0.764846	0.765071	0.501561
8	0.769231	0.769231	0.753375	0.758265	0.775585	0.76469	0.501202
9	0.772857	0.772857	0.755207	0.76216	0.783634	0.767109	0.500362
10	0.776504	0.776504	0.759752	0.765671	0.788006	0.768696	0.499154
11	0.780347	0.780347	0.763659	0.76923	0.789982	0.771602	0.49745
12	0.798246	0.798246	0.781871	0.786869	0.809816	0.780767	0.494253
13	0.811487	0.811487	0.791716	0.800129	0.818626	0.790978	0.492005
14	0.823353	0.823353	0.803811	0.811723	0.826555	0.803694	0.496252
15	0.841463	0.841463	0.820445	0.829313	0.842717	0.819055	0.50179
16	0.861635	0.861635	0.838844	0.848865	0.857793	0.83824	0.510959
17	0.872093	0.872093	0.84675	0.858429	0.87291	0.836723	0.52198
18	0.876565	0.876565	0.847754	0.861605	0.882034	0.874437	0.535862
19	0.887755	0.887755	0.853622	0.870314	0.901472	0.91108	0.570569
20	0.92437	0.92437	0.888025	0.902321	0.929872	0.956581	0.666667

INNINGS COMPARISON

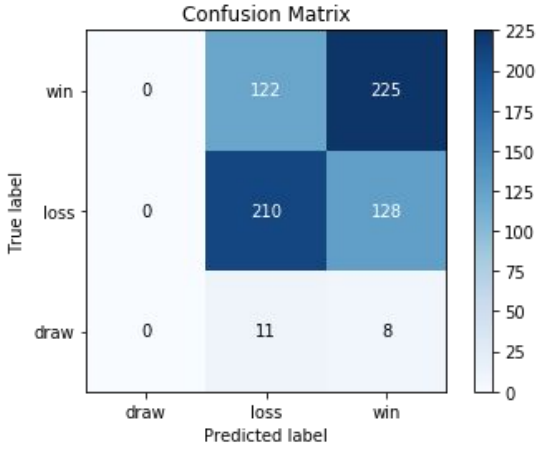




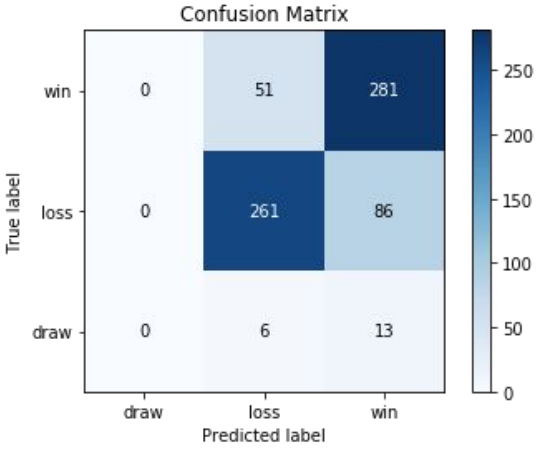




CONFUSION MATRIX

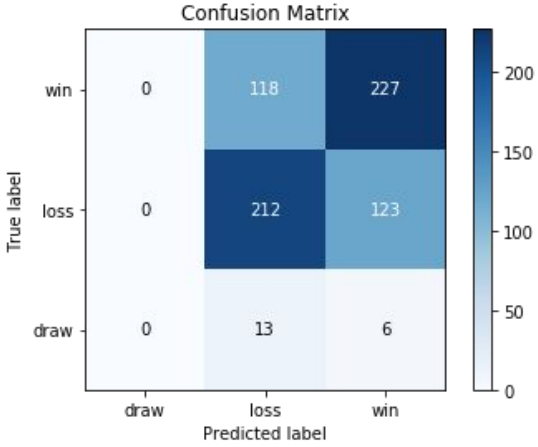


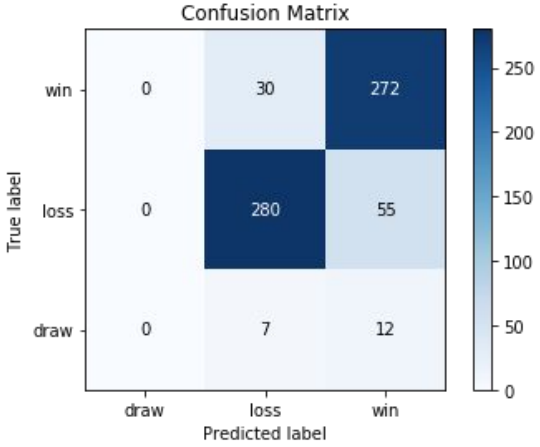
1st Innings (over 10)



2nd Innings (over 10)

CONFUSION MATRIX

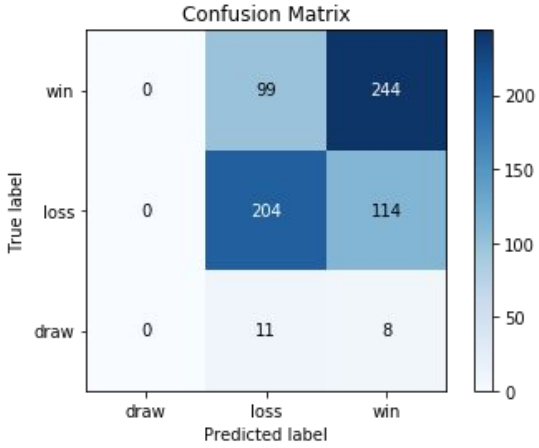




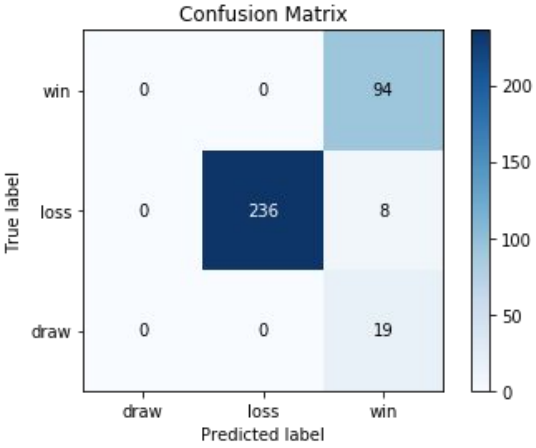
1st Innings (over 15)

2nd Innings (over 15)

CONFUSION MATRIX



1st Innings (over 19)

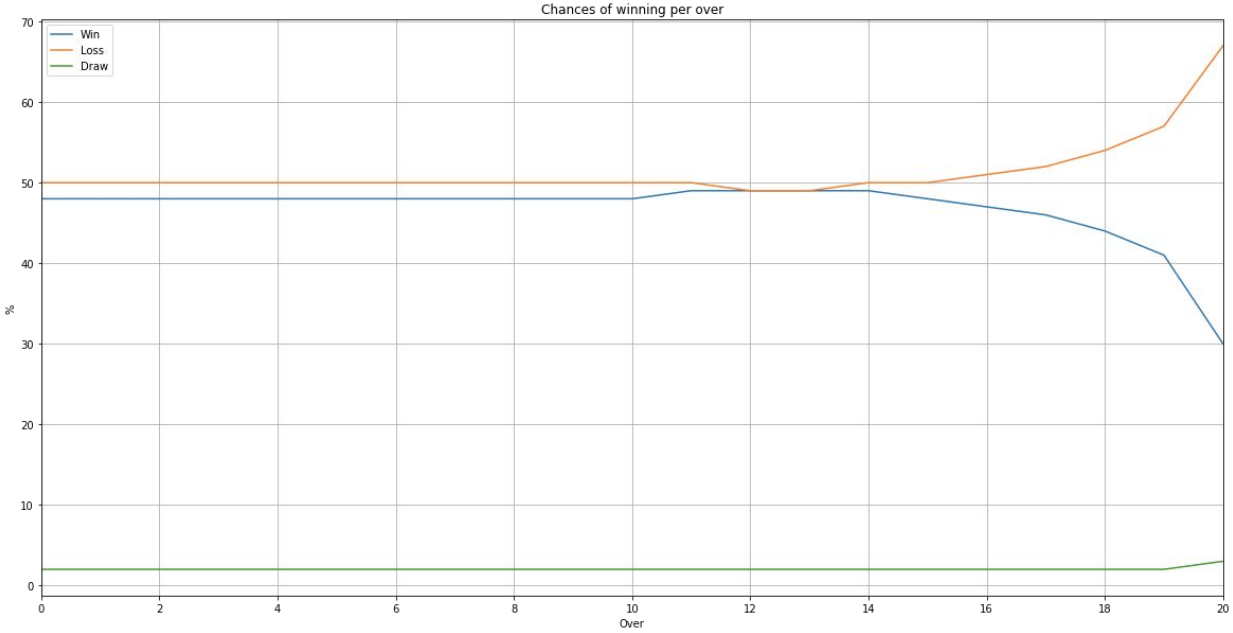


2nd Innings (over 19)

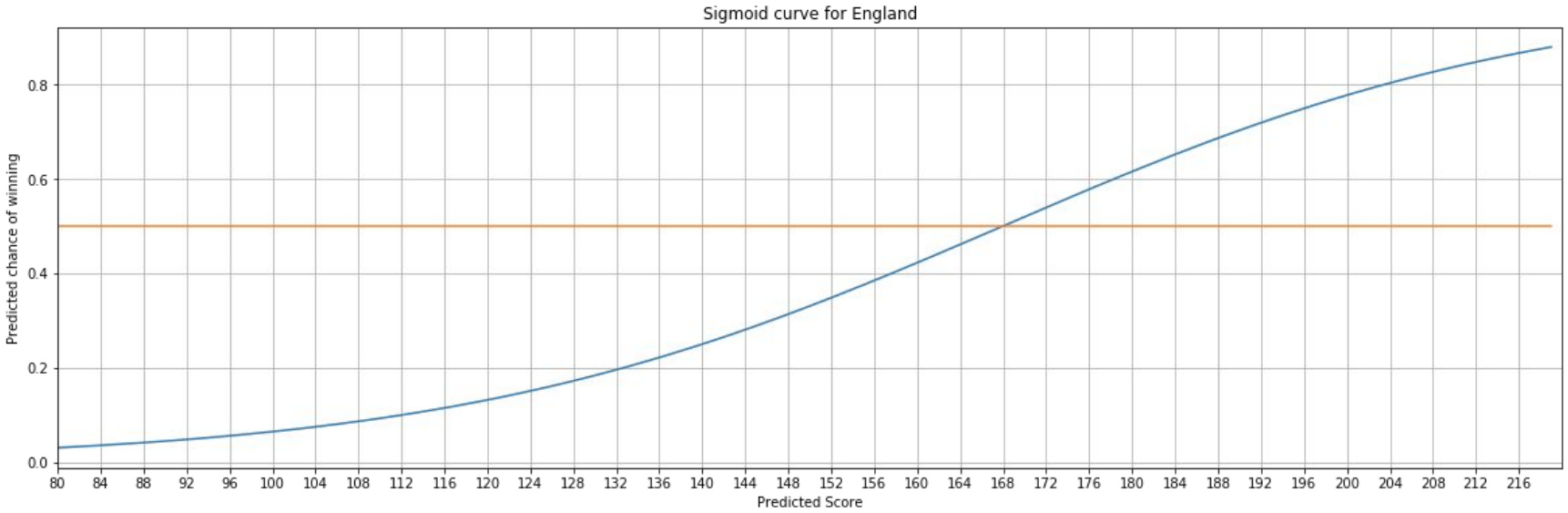
Communicate and Implement

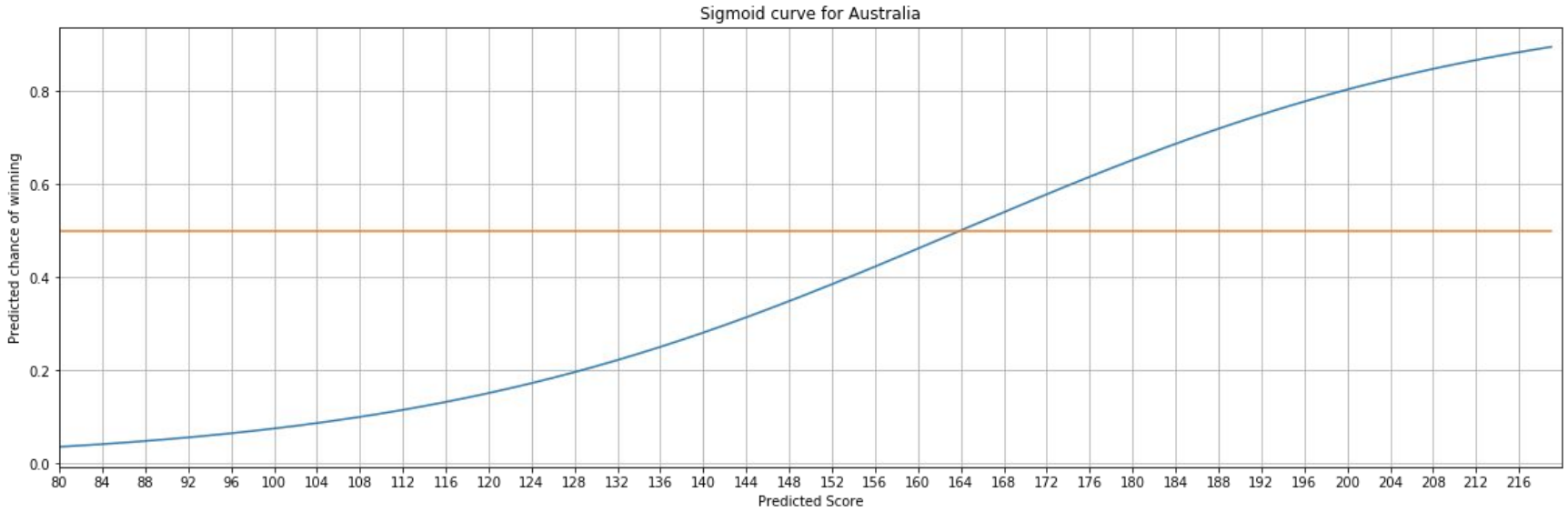
	Win	Loss	Draw
0	48	50	2
1	48	50	2
2	48	50	2
3	48	50	2
4	48	50	2
5	48	50	2
6	48	50	2
7	48	50	2
8	48	50	2
9	48	50	2
10	48	50	2
11	49	50	2
12	49	49	2
13	49	49	2
14	49	50	2
15	48	50	2
16	47	51	2
17	46	52	2
18	44	54	2
19	41	57	2
20	30	67	3

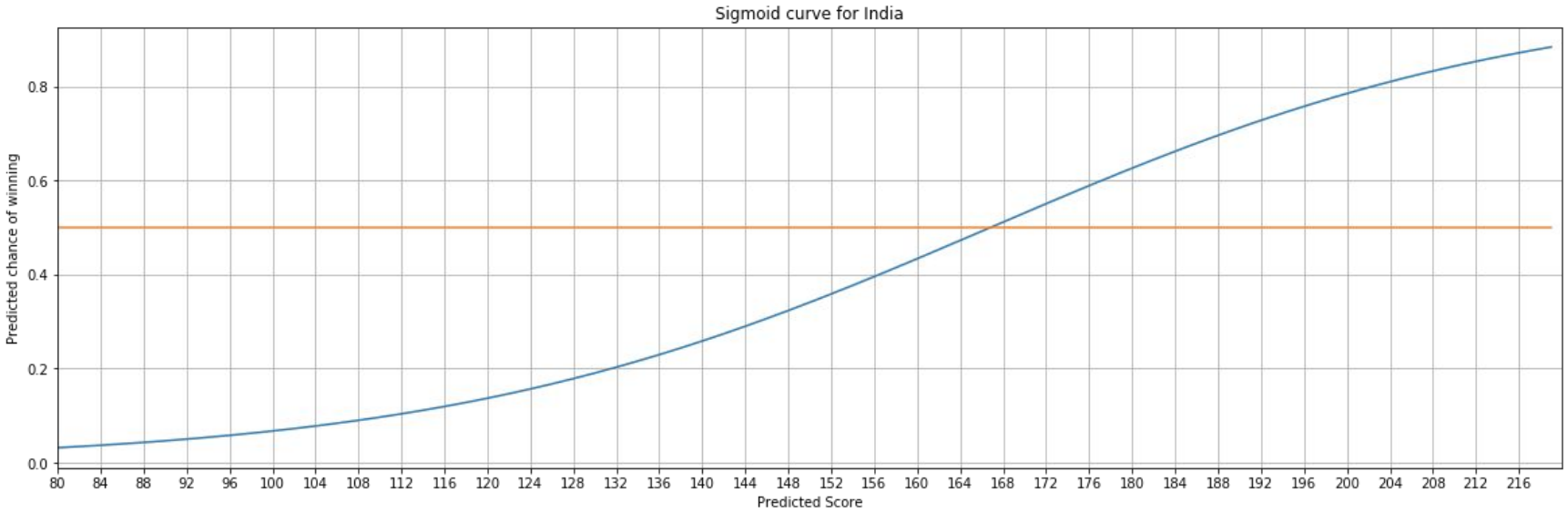
2nd Innings chances of winning per over (observed)

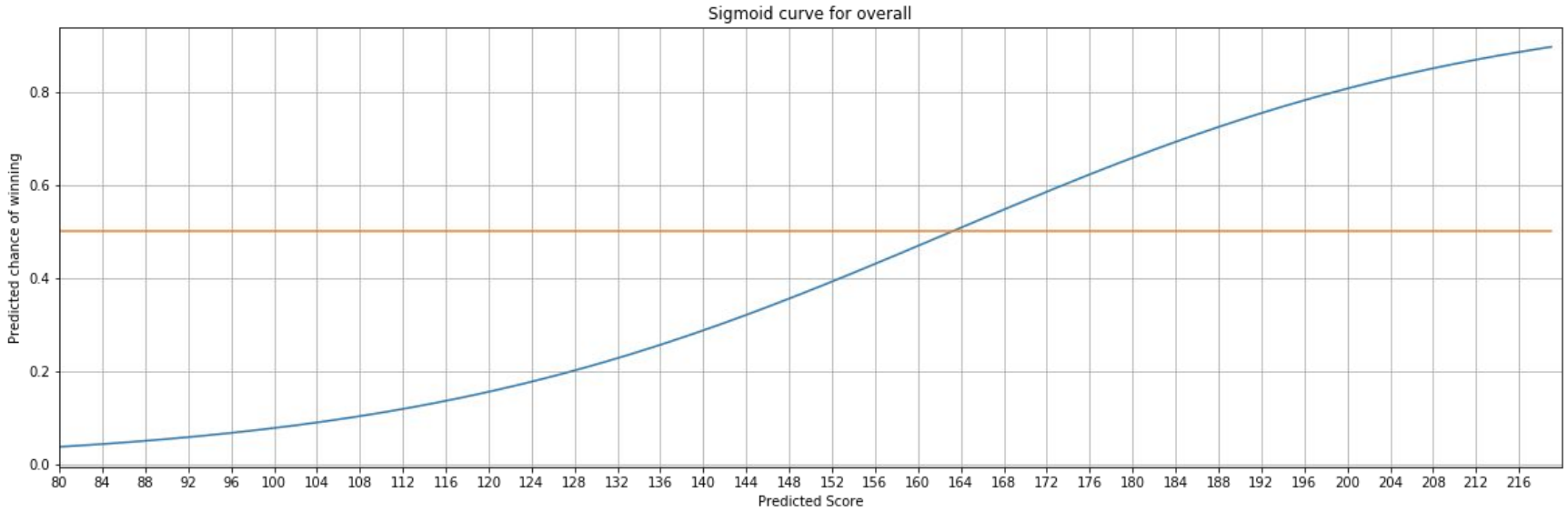


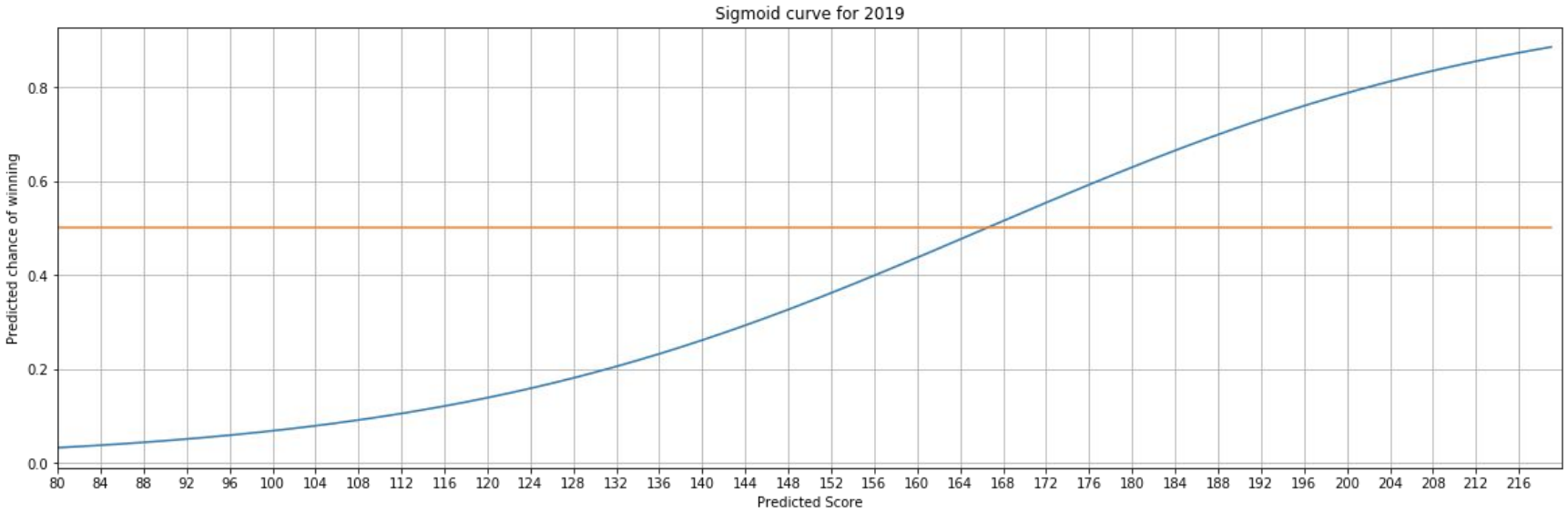
PAR SCORES


Sigmoid curve for England
Predicted chance of winning
0.0
0.2
0.4
0.6
0.8
80 84 88 92 96 100 104 108 112 116 120 124 128 132 136 140 144 148 152 156 160 164 168 172 176 180 184 188 192 196 200 204 208 212 216
Predicted Score

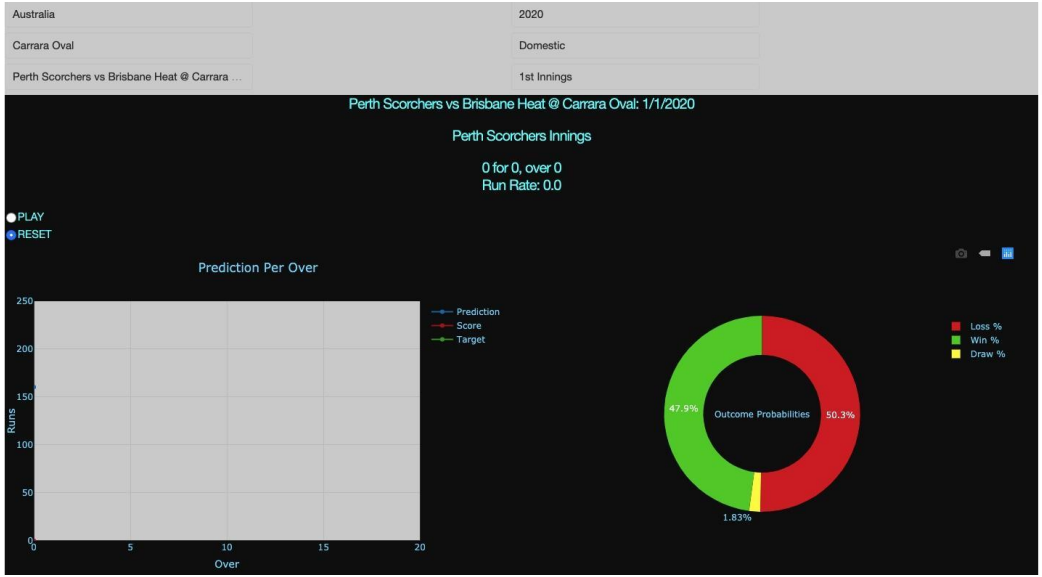

Sigmoid curve for Australia
Predicted chance of winning
0.0
0.2
0.4
0.6
0.8
80 84 88 92 96 100 104 108 112 116 120 124 128 132 136 140 144 148 152 156 160 164 168 172 176 180 184 188 192 196 200 204 208 212 216
Predicted Score


Sigmoid curve for India
Predicted chance of winning
0.0
0.2
0.4
0.6
0.8
80 84 88 92 96 100 104 108 112 116 120 124 128 132 136 140 144 148 152 156 160 164 168 172 176 180 184 188 192 196 200 204 208 212 216
Predicted Score


Sigmoid curve for overall
Predicted chance of winning
0.0
0.2
0.4
0.6
0.8
80 84 88 92 96 100 104 108 112 116 120 124 128 132 136 140 144 148 152 156 160 164 168 172 176 180 184 188 192 196 200 204 208 212 216
Predicted Score


Sigmoid curve for 2019
Predicted chance of winning
0.0
0.2
0.4
0.6
0.8
80 84 88 92 96 100 104 108 112 116 120 124 128 132 136 140 144 148 152 156 160 164 168 172 176 180 184 188 192 196 200 204 208 212 216
Predicted Score

DASHBOARD



[Demonstration]