Machine Learning na previsão de resultados em partidas de futebol

Trabalhos correlacionados

Artigo: Predictive analysis and modelling football results using machine learning approach for English Premier League

Ano: 2018

Department of Computer Science and Engineering, School of Engineering Sciences and Technology, Jamia Hamdard, New Delhi, India

Atributos

Últimos k chutes no gol, quantidade de gols, escanteios, taxa de ataque, taxa de defesa, últimas vitórias do time, etc

Atributos para times mandantes e visitantes

Tabela de Atributos

Feature name	Feature abbreviation	Class ca
Home form	HForm	Class A
Away form	AForm	Class A
Home streak	HSt	Class A
Away streak	ASt	Class A
Past k home shots on target	HSTKPP	Class A
Past k away shots on target	ASTKPP	Class A
Past k home goals	HGKPP	Class A
Past k away goals	AGKPP	Class A
Past k home corners	HCKPP	Class A
Past k away corners	ACKPP	Class A
Home attack rating	HAttack	Class A
Away attack rating	AAttack	Class A
Home midfield rating	HMidField	Class A
Away midfield rating	AMidField	Class A
Home defence rating	HDefence	Class A
Away defence rating	ADefense	Class A
Home overall rating	HOverall	Class A
Away overall rating	AOverall	Class A
Home goal difference	HTGD	Class A
Away goal difference	ATGD	Class A
Home weighted streak	HStWeighted	Class A
Away weighted streak	AStWeighted	Class A
Form differential	FormDifferential	Class B
Streak differential	StDifferential	Class B
Past k shots on target differential	STKPP	Class B
Past k goals differential	GKPP	Class B

Fonte: https://www.bbc.com/sport/football/premier-league/table

Artigo: Football Match Prediction with Tree Based Model Classification

Ano: 2019

Computer Science Department, BINUS Graduate Program-Master in Computer Science, Bina Nusantara University Jakarta, 11480, Indonesia

Tabela de atributos

Fonte: football-data.co.uk

Features	Data Scale	Description
FTR	Nominal	Full-time result. Used as the label of the model
HTHG	Ratio	Number of goal for the home team in half time
HTAG	Ratio	Number of goal for away team in half time
HS	Ratio	Number of shot done by the home team
AS	Ratio	Number of shot done by away team
HST	Ratio	Number of shot on target is done by the home team
AST	Ratio	Number of shot on target is done by the home team
HF	Ratio	Number of fouls done by the home team
AF	Ratio	Number of fouls done by away team
HC	Ratio	Number of corners obtained by the home team
AC	Ratio	Number of corners obtained by away team
HY	Ratio	Number of yellow cards obtained by the home team
AY	Ratio	Number of yellow cards obtained by away team
HR	Ratio	Number of red cards obtained by the home team
AR	Ratio	Number of red cards obtained by away team

Tese de mestrado: PREDICTION OF THE FOOTBALL MATCH RESULTS WITH USING MACHINE LEARNING ALGORITHMS

Ano: 2019

COMPUTER ENGINEERING DEPARTMENT, ÇANKAYA UNIVERSITY, Turkey

Atributos

Home Team, Away Team, FTHG (Full Time Home Team Goals), FTAG (Full Time Away Team Goals), FTR (Full Time Result), HTHG (Half Time Home Team Goals), HTAG (Half Time Away Team Goals), HTR (Half Time Result), HS (Home Team Shots), AS (Away Team Shots), HST (Home Team Shots on Target), AST (Away Team Shots on Target), HF (Home Team Fouls Committed), AF (Away Team Fouls Committed), HC (Home Team Corners), AC (Away Team Corners), HY (Home Team Yellow Cards), AY (Away Team Yellow Cards), HR (Home Team Red Cards), AR (Away Team Red Cards), LMH (Last Match Home Team Result) and LMA (Last Match Away Team Result)

Fonte: https://datahub.io/sports-data/spanish-la-liga

Técnicas usadas

1. Predictive analysis and modelling football results using machine learning approach for English Premier League

Gaussian naive Bayes
Support vector machine
Random forest
Gradient boosting

Resultados e comparações

Gaussian naïve Bayes results.

(a) Confusion matrix

	Predicted win	Predicted loss	Predicted draw
Actual win	197	46	47
Actual loss	57	96	36
Actual draw	72	45	44

(b) Precision–recall table

	Precision	Recall	F1-score
Home wins	0.62	0.66	0.63
Away wins	0.51	0.51	0.51
Draws	0.33	0.25	0.29

Resultados e comparações

Linear SVM results.

(a) Confusion matrix

	Predicted win	Predicted loss	Predicted draw
Actual win	254	36	0
Actual loss	96	93	0
Actual draw	122	39	0

(b) Precision–recall table

	Precision	Recall	F1-score
Home wins	0.54	0.88	0.67
Away wins	0.55	0.49	0.52
Draws	0.0	0.0	0.0

Random forest results.

(a) Confusion matrix

	Predicted win	Predicted loss	Predicted draw
Actual win	225	36	29
Actual loss	64	101	24
Actual draw	86	40	35

(b) Precision–recall table

	Precision	Recall	F1-score
Home wins	0.60	0.78	0.68
Away wins	0.57	0.53	0.55
Draws	0.40	0.22	0.28

Table 5
RBF SVM results.

(a) Confusion matrix

	Predicted win	Predicted loss	Predicted draw
Actual win	238	37	15
Actual loss	84	96	9
Actual draw	110	36	15

(b) Precision–recall table

	Precision	Recall	F1-score
Home wins	0.55	0.80	0.67
Away wins	0.57	0.51	0.54
Draws	0.39	0.09	0.15

Table 7
Gradient boosting results.

(a) Confusion matrix

	Predicted win	Predicted loss	Predicted draw
Actual win	222	37	31
Actual loss	58	99	32
Actual draw	88	31	42

(b) Precision–recall table

	Precision	Recall	F1-score
Home wins	0.60	0.77	0.67
Away wins	0.59	0.52	0.54
Draws	0.40	0.26	0.31

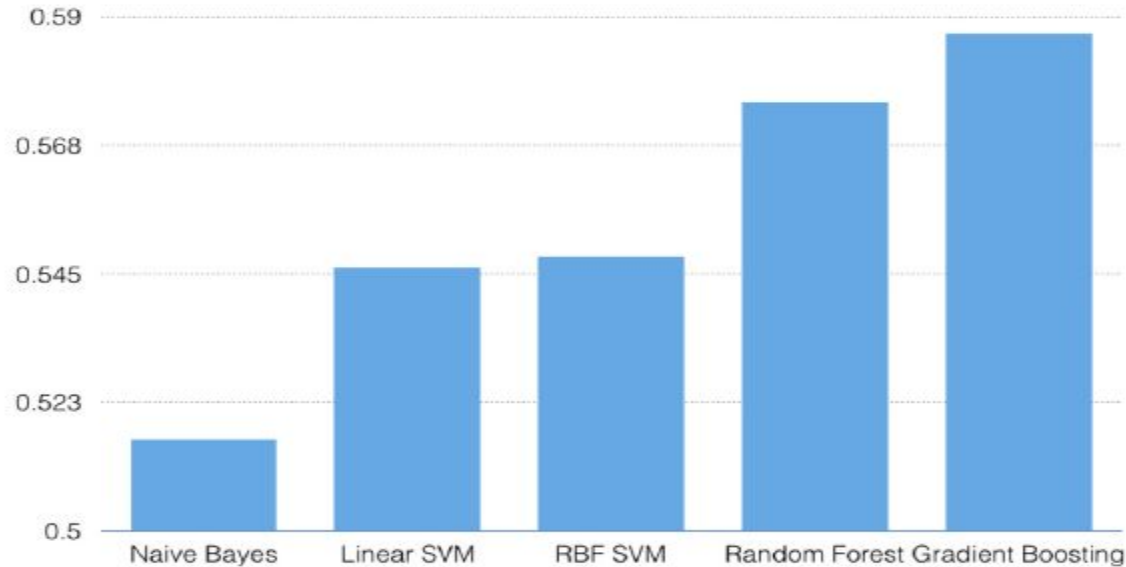# Resultados e comparações

Fig. 5. Mean test accuracy scores of the different machine learning models.

Técnicas usadas

2. Football Match Prediction with Tree Based Model Classification

C5.0
Random Forest
Gradient Boosting

Resultados e comparações

Evaluation Metrics	C5.0	Random Forest	XGBoost
Sensitivity	0.5837	0.6545	0.6345
Specificity	0.8078	0.8381	0.8295
Precision	0.6053	0.6652	0.6477
Recall	0.5837	0.6545	0.6345
F1	0.5506	0.6546	0.6348
Kappa	0.4258	0.5123	0.4867
Accuracy	0.6487	0.6855	0.6789
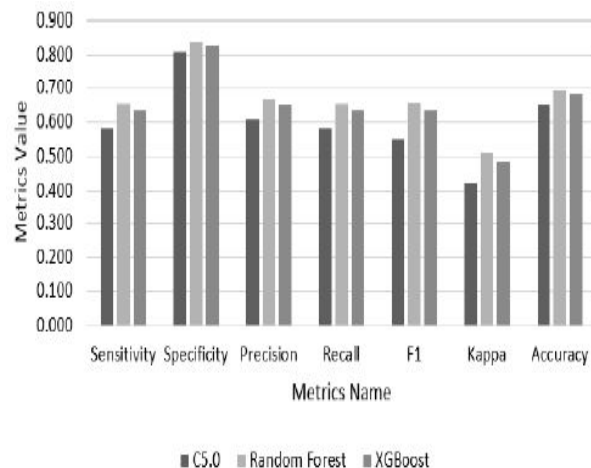
Técnicas usadas

3. PREDICTION OF THE FOOTBALL MATCH RESULTS WITH USING
MACHINE LEARNING ALGORITHMS

Support Vector Machine Classifier
K-NN Classifier
Artificial Neural Network Classifier

Resultados e comparações

Table 19: Confusion Matrix table of Backward Elimination Results with SVM Linear Kernel

A	B	C	<-- classified as
276	38	44	a = D
45	669	0	b = H
55	1	392	c = A

Average	TP Rate	FP Rate	Precision	Recall	F-Measure	Class
	0.771	0.086	0.734	0.771	0.752	D
	0.937	0.048	0.945	0.937	0.941	H
	0.875	0.041	0.899	0.875	0.887	A
Weighted Avg.	0.880	0.055	0.882	0.880	0.881	

Resultados e comparações

Table 36: Confusion Matrix table of Backward Elimination Results with K-NN

A	B	C	<-- classified as
128	139	91	a = D
72	609	33	b = H
86	45	317	c = A

Table 35: Detailed Accuracy by Class Table of Backward Elimination Results with K-NN

Average	TP Rate	FP Rate	Precision	Recall	F-Measure	Class
	0.358	0.136	0.448	0.358	0.398	D
	0.853	0.228	0.768	0.853	0.808	H
	0.708	0.116	0.719	0.708	0.713	A
Weighted Avg.	0.693	0.173	0.678	0.693	0.683	

Resultados e comparações

Table 46: Backward Elimination Results with ANN Confusion Matrix Table

A	B	c	<-- classified as
206	84	68	a = D
21	693	0	b = H
54	0	394	c = A

Table 45: Backward Elimination Results with ANN Detailed Accuracy by Class Table

Average	TP Rate	FP Rate	Precision	Recall	F-Measure	Class
	0.575	0.065	0.733	0.575	0.645	D
	0.971	0.104	0.892	0.971	0.930	H
	0.879	0.063	0.853	0.879	0.866	A
Weighted Avg.	0.851	0.083	0.843	0.851	0.844	

Resultados e comparações
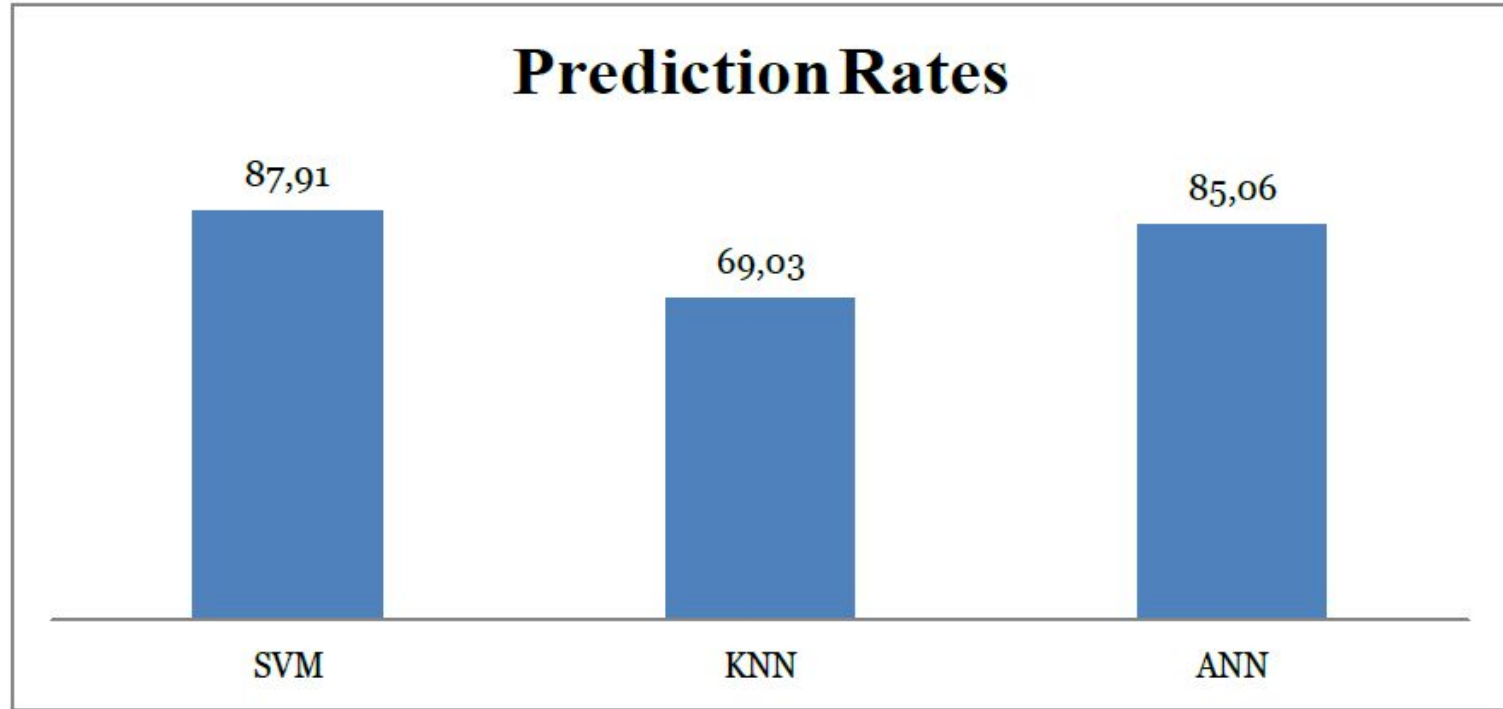
Conclusão

Artigos	Ligas usadas	Técnicas Usadas	Melhores Técnicas de cada artigo	Acurácia
1º artigo	Premier League	Gaussian Naïve Bayes, SVM, Floresta Aleatória e Gradient Boosting	Gradient Boosting e Floresta Aleatória	Aprox. 0.5672 e 0.5641
2º artigo	Premier League	C5.0, Floresta Aleatória e Gradient Boosting	Floresta Aleatória e Gradient Boosting	0.6855 e 0.6789
3º artigo	La Liga	SVM, KNN e Rede Neural Artificial	SVM e Rede Neural Artificial	0.8791 e 0.8506
Comparação Final	La Liga	-	SVM	0.8791

Comparativo Técnicas de Machine Learning