



UNIVERSIDADE FEDERAL DO PARANÁ
SETOR DE CIÊNCIAS EXATAS
DEPARTAMENTO DE ESTATÍSTICA
CURSO DE ESTATÍSTICA

Karen Regina Galvão de Oliveira

Luis Henrique de Oliveira

**Driblando a sorte e marcando um gol:
uma abordagem estatística no futebol**

**CURITIBA
2017**



UNIVERSIDADE FEDERAL DO PARANÁ
SETOR DE CIÊNCIAS EXATAS
DEPARTAMENTO DE ESTATÍSTICA
CURSO DE ESTATÍSTICA

Karen Regina Galvão de Oliveira

Luis Henrique de Oliveira

**Driblando a sorte e marcando um gol:
uma abordagem estatística no futebol**

Trabalho de Conclusão de Curso apresentado à disciplina Laboratório B do Curso de Estatística do Setor de Ciências Exatas da Universidade Federal do Paraná, como exigência parcial para obtenção do grau de Bacharel em Estatística.

Orientador: Prof Eduardo Vargas Ferreira

**CURITIBA
2017**

RESUMO

Para desmistificar grandes clichês do futebol, este trabalho tem como objetivo estudar os fatores relevantes que conjugam na vitória do time e delinear aspectos significativos, sob a ótica estatística, que conduzem ao sucesso nas partidas. Técnicas de estatística não paramétrica, agrupamento de clusters, modelos lineares generalizados e análise de classificação e regressão baseada em árvores foram utilizadas. Os resultados sugerem que: O número de faltas dadas a mandantes e visitantes segue a mesma distribuição, entretanto há um viés favorável aos mandantes na aplicação de cartões amarelos; Apesar de muito comentado pelos técnicos em coletivas, a diferença entre os tipos de campos não interfere no resultado da partida; Na demissão de técnicos, identificou-se a queda de performance nos jogos antecedentes à demissão, e os times que tiveram técnicos demitidos não esboçaram reação imediata; Na predição do resultado de partidas, considerou-se a performance média das equipes e a cotação das apostas para os jogos, chegando em um sistema de decisão com retorno financeiro considerável.

Palavras-chave: Engenharia de características. Estatística no Futebol. Análise preditiva.

Sumário

RESUMO	3
1 INTRODUÇÃO	5
2 MATERIAL E MÉTODOS	7
2.1 Material	7
2.2 Métodos	9
3 RESULTADOS E DISCUSSÃO	11
3.1 Arbitragem	12
3.2 Fatores Externos.....	13
3.3 Equipes	16
3.4 Jogos	19
4 CONSIDERAÇÕES FINAIS.....	22
REFERÊNCIAS.....	23
ANEXOS	25

1 INTRODUÇÃO

O Futebol é considerado um dos esportes mais populares do mundo. Teve sua origem na Inglaterra, em 1863, e aos poucos foi se popularizando mundialmente, culminando com a criação da Federação Internacional de Futebol (FIFA), em 1904, para que campeonatos internacionais fossem realizados oficialmente. Em 1908 passou a ser a segunda modalidade coletiva disputada em jogos olímpicos e em 1930 houve a realização da primeira Copa do Mundo, no Uruguai. No Brasil, o esporte chegou em 1895, quando Charles Miller retornou da Inglaterra trazendo a primeira bola de futebol na bagagem. Pouco a pouco, os campeonatos foram se desenvolvendo, começando com os campeonatos estaduais, seguindo ao âmbito nacional, com o Campeonato Brasileiro da Série A, competição nacional de maior tradição, criada com a proposta de englobar times de todas as regiões do país. Com grandes atletas na história, o futebol brasileiro é reconhecido internacionalmente pela sua qualidade, exportando jogadores para clubes no mundo inteiro. Nos dias atuais, o futebol se tornou muito mais que uma paixão, mas uma indústria que movimenta mais de 500 bilhões de reais anualmente^[1], porém o Brasil representa uma parcela ínfima das receitas mundiais.

Após a revolução na indústria do *baseball* americano, iniciada por Billy Beane em 2002, popularizada com o livro de Lewis^[2], aliado às informações sobre incertezas intrínsecas aos esportes, grandes clubes criaram departamentos de *analytics* para encontrar maneiras de minimizar a sorte atribuída aos resultados das partidas. Com metodologias cada vez mais eficazes na coleta e armazenamento de dados, as análises estatísticas desde então tiveram um papel indispensável na evolução desse esporte, que hoje se tornou um negócio lucrativo.

Ramos e Oliveira^[3] estudaram os gols da Eurocopa de 2004, e através de estatísticas descritivas concluíram que a maioria dos gols ocorreu com a bola em movimento, por atacantes e meio campistas. Claudino et al^[4] estudaram os fatores de estresse e ansiedade pré-competitiva em jogadores Sub-20 e verificaram fatores influentes tanto positiva como negativamente. Ainda no âmbito das questões psicológicas, Verardi et al^[5] analisaram estratégias de enfrentamento de problemas, observando que independentemente da posição de jogo os jogadores utilizam as mesmas estratégias de modo que é importante no treinamento habilitá-los a resolver problemas que surgem no decorrer de uma partida.

Outra abordagem de estudo deve-se a Lee^[6], que utilizou a distribuição Poisson para modelar os gols nas partidas. Baseando-se na mesma técnica, Louzada et al^[7] utilizaram modelos de regressão linear para prever resultados de partida e classificação final da *English Premier League*, na temporada 2008-2009. De modo semelhante, Tavares e Suzuki^[8] estudaram o Campeonato Brasileiro de 2013 obtendo previsões para a classificação final.

Assim como um investidor analisa e escolhe criteriosamente as oportunidades para destinar seus investimentos, a fim de obter o melhor retorno, um clube de futebol deve aplicar seus recursos de forma consciente e fundamentada desde a contratação de jogadores e comissão técnica até a infraestrutura para a categoria de base, reforma do estádio, etc. Sob essa perspectiva, é de fundamental importância a área de *analytics* nos clubes, onde estudos de apoio para decisões da comissão técnica (dentro e fora de campo) são desenvolvidos, minimizando as incertezas que permeiam o futebol.

Por exemplo, a comissão técnica deve saber sobre a escolha da escalação mais adequada, considerando alguns fatores como a formação do time para enfrentar determinado adversário, fase e entrosamento dos jogadores, impactos nas performances devido ao número de partidas disputadas e *break* de temporada, etc. Além disso, é essencial entender se tais escolhas mudam de acordo com outros fatores, como a relação entre o mando de campo e a torcida: será que a torcida é o “12º jogador”, capaz de influenciar na decisão da partida e até mesmo da arbitragem? Neste caso, políticas para atrair público aos jogos são extremamente importantes. Outras questões são igualmente de grande interesse, tais como: será que a fase do time muda imediatamente com a demissão do técnico, de tal forma que valha a pena as rescisões contratuais e todos os custos incumbidos a essa decisão? Ou será que contratações são capazes de resolver problemas em determinados setores do time?

Não são poucas as questões a serem respondidas sobre esse mercado competitivo do futebol. Sendo assim, o objetivo deste trabalho é estudar os fatores relevantes que conjugam na vitória do time de futebol e delinear aspectos significativos, sob a ótica estatística, que conduzem ao sucesso do time nas partidas. O trabalho está dividido da seguinte forma, na Seção 2 são apresentados os dados utilizados e as metodologias de análise, na Seção 3 encontram-se os resultados e na Seção 4 as conclusões e sugestões para estudos futuros.

2 MATERIAL E MÉTODOS

2.1 Material

O conjunto de dados contém estatísticas e informações relevantes de cada partida da temporada 2011-12 da *English Premier League*. Após a unificação dos dados disponibilizados no *GitHub* da *Soccermetrics*^[9], obtidos através do sistema de coleta de dados da Opta, houve a criação de uma base de dados única. Além das variáveis previamente disponibilizadas, criou-se variáveis de interesse para o estudo e a coleta de outras informações, como o valor dos times, em libras, no começo do campeonato^[10] e apostas no resultado de cada partida^[11]. No quadro 1 há a descrição das principais variáveis. Também foram utilizadas algumas das variáveis referentes a performance dos jogadores nas partidas, que foram agrupadas por time, e estão descritas no anexo.

Quadro 1 – Descrição das principais variáveis disponíveis no conjunto de dados

VARIÁVEL	DESCRIÇÃO
Mandante	Time mandante
Visitante	Time visitante
Arbitro	Árbitro da partida
Horario	1 = partida iniciada até as 17hrs 0 = partida iniciada após as 17hrs
Temperatura	Temperatura no início da partida (°C)
Clima	Clima predominante no início da partida
Altitude	Altitude em que se localiza o estádio (m)
Gramado	Tipo de gramado do estádio
Comprimento	Comprimento do campo
Largura	Largura do campo
Classico	1 = Clássico regional 0 = Times de diferentes regiões
Publico	Público total
PcPublico	Percentual de ocupação do estádio
NovoTecnico *	1 = Técnico diferente do que no início da temporada 0 = Mesmo técnico que no início da temporada
Valor *	Valor do time
Gols *	Gols do time
ApostaM	Apostas na vitória do time mandante
ApostaEmpate	Apostas no empate

ApostaV	Apostas na vitória do time visitante
Zagueiros *	Quantidade de zagueiros de ofício em campo do time
Meias *	Quantidade de meias de ofício em campo do time
Atacantes *	Quantidade de atacantes de ofício em campo do time
VitoriaM	1 = Vitória do time mandante
VitoriaV	1 = Vitória do time visitante
Empate	1 = Empate entre as equipes
Campo *	Tipo de gramado do estádio do time
DifCampo	1 = Time jogando em gramado diferente do que seu estádio 0 = Time jogando em mesmo tipo de gramado
PcPublicoAnt	Percentual de público na rodada anterior como mandante
PublicoAnt	Público total na rodada anterior como mandante
VIT_U1J *	1 = Vitória no último jogo do time
VIT_U2J *	Número de vitórias do time nos últimos 2 jogos
VIT_U3J *	Número de vitórias do time nos últimos 3 jogos
VIT_U4J *	Número de vitórias do time nos últimos 4 jogos
VIT_U5J *	Número de vitórias do time nos últimos 5 jogos
VIT_UJ *	Número de vitórias do time até a rodada anterior
VitoriasMandandoM	Número de vitórias do time mandante em casa até a rodada anterior
EmpatesMandandoM	Número de empates do time mandante em casa até a rodada anterior
DerrotasMandandoM	Número de derrotas do time mandante em casa até a rodada anterior
VitoriasVisitandoV	Número de vitórias do time visitante fora de casa até a rodada anterior
EmpatesVisitandoV	Número de empates do time visitante fora de casa até a rodada anterior
DerrotasVisitandoV	Número de derrotas do time visitante fora de casa até a rodada anterior
PasseChave *	Quantidade de passes chaves realizados pelo time
EscanteioChave *	Quantidade de escanteios chave realizados pelo time
FaltaChave *	Quantidade de faltas chave realizados pelo time
LateralChave *	Quantidade de laterais chave realizados pelo time
TiroMetaChave *	Quantidade de tiros de meta chave realizados pelo time
JogadaEnsaaiadaChave *	Quantidade de jogadas ensaiadas chave realizados pelo time

* Variável disponível para o time mandante e visitante

Fonte: Os autores (2017).

O campeonato inglês é disputado de modo similar ao campeonato brasileiro, 20 equipes jogam entre si em 2 turnos, cada vez com o mando de campo para uma equipe. Deste modo temos 10 jogos por rodada e 38 rodadas, totalizando 380 partidas disputadas.

Utilizou-se *software* R, versão 3.2.2 (R CORE TEAM, 2015) para a análise dos dados.

2.2 Métodos

Foram utilizados testes estatísticos paramétricos, não paramétricos e construídos modelos lineares generalizados, com resposta seguindo a distribuição Poisson e Binomial. Foi adotado o nível de 5% de significância. Para análises envolvendo a comparação entre mandante e visitante, como por exemplo o número de faltas, após análises descritivas, a normalidade das variáveis foi analisada através de histogramas e verificada por meio do teste de Shapiro-Wilk^[12].

Cada partida foi considerada uma unidade amostral e, buscando avaliar diferenças entre mandante e visitante, foram feitos testes pareados. Com a violação da suposição de normalidade, utilizamos o teste de Wilcoxon^[12]. Primeiramente são calculadas as diferenças entre os grupos, que são ordenadas crescentemente e então são atribuídos *postos*. O teste de hipóteses consiste em $H_0: \Delta = 0$ vs $H_1: \Delta \neq 0$, em que Δ representa a média das diferenças entre os grupos. A estatística do teste é a soma dos postos de todas as diferenças positivas observadas, sendo o p-valor obtido através da aproximação normal da distribuição desta estatística.

Em alguns casos a variável de interesse foi separada por grupos relevantes (diferentes populações) e, devido à falta de normalidade, para a comparação de seus valores foi utilizado o teste de Wilcoxon-Mann-Whitney^[12]. A hipótese de interesse é equivalente ao teste pareado, tendo por objetivo avaliar se as variáveis possuem a mesma mediana ou se uma tende a ter valores maiores do que a outra. Todas as observações são ordenadas, atribuem-se *postos* e então é feita a soma dos postos por população. A estatística do teste é a soma de postos da população com menor número de observações (n_1), obtendo-se a significância através da distribuição normal desta variável, que possui média dada por $\mu_T = \frac{n_1(n_1+n_2+1)}{2}$ e variância $\sigma_T = \sqrt{\frac{\mu_T \cdot n_2}{6}}$, em que n_2 é o número de observações da população de comparação.

O teste qui-quadrado^[12] foi utilizado para analisar a homogeneidade de distribuição de variáveis categóricas, de acordo com características de interesse, como o resultado da partida. A hipótese testada é que as diferentes categorias têm a mesma proporção de partidas com alguma característica específica. A estatística do teste é dada por $Q_{obs}^2 = \sum_{i=1}^r \sum_{j=1}^s \frac{(O_{ij}-E_{ij})^2}{E_{ij}}$, em que r e s representam a quantidade de classes das variáveis envolvidas, O_{ij} as proporções observadas e

E_{ij} as proporções esperadas, computadas pelo produto dos totais marginais das classes dividido pelo total geral de observações. A significância é obtida comparando o valor Q_{obs}^2 com a distribuição χ^2 com $(r - 1)(s - 1)$ graus de liberdade.

Para verificar a influência de árbitros nas partidas, foi utilizado agrupamento de clusters, dado que não havia interesse em analisá-los individualmente, apenas subgrupos com posicionamento similar. O método de *k-means*^[13] foi utilizado para buscar agrupamentos, de modo que a variação dentro de cada cluster seja tão pequena quanto possível. Em um chute inicial, o algoritmo atribui, aleatoriamente, cada observação em um dos K clusters predefinidos, em seguida são feitas iterações até que os clusters se estabilizem, buscando em cada cluster seu centroide e atribuindo as observações ao cluster mais próximo, com base na menor distância Euclidiana.

As relações entre as variáveis também foram avaliadas com o uso de modelos lineares generalizados^[14]. Estes modelos são definidos por um componente aleatório, representando a variável independente - que segue uma distribuição da família exponencial de dispersão; um componente sistemático, correspondente aos preditores lineares (covariáveis de interesse - x_i) e a função de ligação, com o papel de linearizar a relação entre estes componentes. Primeiramente é verificada a distribuição adequada à variável resposta.

A distribuição binomial é uma distribuição de probabilidades discreta, usada para modelar a contagem de “sucessos” em n realizações independentes e igualmente prováveis de um experimento com dois desfechos possíveis, ou com um número maior de resultados, classificados em duas categorias. Neste caso, a função de ligação logito foi utilizada, e a probabilidade π_i , de ocorrência do evento é dada por $\pi_i = \frac{e^{\beta_0 + \beta_1 x_1 + \dots + \beta_i x_i}}{e^{\beta_0 + \beta_1 x_1 + \dots + \beta_i x_i} + 1}$.

A distribuição de Poisson é utilizada para modelar a contagem de ocorrências de algum evento de interesse em unidades de tempo ou espaço. A função de ligação logarítmica foi utilizada e a média da variável resposta é dada por $\mu_i = \exp^{\beta_0 + \beta_1 x_1 + \dots + \beta_i x_i}$.

Para avaliar a qualidade dos ajustes propostos, foram verificados os resíduos padronizados da *deviance*, uma medida de distância das observações em relação à média estimada, na escala do logaritmo da verossimilhança. Considerando que a falta de normalidade

dos resíduos é notável na modelagem de dados discretos, também foram analisados graficamente os resíduos quantílicos aleatorizados^[15], que apresentam distribuição Normal, independente da distribuição da variável resposta e de sua dispersão.

Também foi utilizado o método de árvore de classificação e regressão (CART)^[13]. A construção da árvore de decisão é feita em dois passos: divide-se o espaço dos preditores em regiões distintas e então para cada observação que caia em determinada região é feita a predição. O objetivo é minimizar a soma de quadrado dos resíduos, de modo que são feitas divisões binárias recursivas, repetidamente, até se obter o ponto de corte tal que a árvore resulte na menor soma de quadrados dos resíduos.

3 RESULTADOS E DISCUSSÃO

Buscando avaliar os fatores de influência no sucesso do time na partida, primeiramente levantamos quatro grandes grupos, conforme a Figura 1. Primeiramente os fatores relacionados a arbitragem foram estudados, seguidos por fatores externos, de equipe e então o jogo propriamente. Duas rodadas foram retiradas das análises para validação dos modelos propostos (representando 5.27% do total de observações). Em seguida, apresentaremos as análises de cada segmento em separado, na ordem: Arbitragem, Fatores Externos, Equipe e finalmente Jogo.

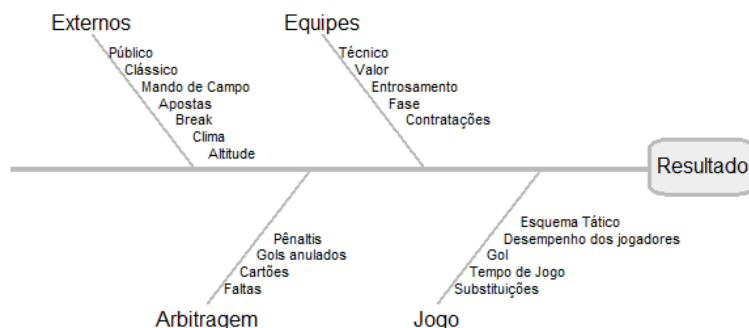


FIGURA 1 – Fatores de influência no resultado dos times na partida

3.1 Arbitragem

Considerando o envolvimento da arbitragem nos resultados das partidas, primeiramente analisamos a distribuição de faltas e cartões para mandantes e visitantes. Após verificar a distribuição destas variáveis, utilizamos o teste de Wilcoxon pareado para analisar a diferença entre mandantes e visitantes. Não há evidências de diferença quanto ao número de faltas concedidas à mandantes e visitantes (p -valor = 0.19), aos cartões vermelhos (p -valor = 0.13) ou aos pênaltis (p -valor = 0.17). Entretanto, observou-se uma diferença significativa entre a aplicação de cartões amarelos para mandantes e visitantes (p -valor < 0.01), sendo a média de cartões amarelos dados aos mandantes (1.33) inferior aos visitantes (1.66). Deste modo, é possível verificar que apesar de as faltas seguirem a mesma distribuição, o mesmo não acontece com os cartões amarelos. Ou seja, há um certo viés favorável ao mandante na aplicação de cartões amarelos.

Para analisar uma possível influência dos árbitros responsáveis pelas partidas, verificamos a distribuição das faltas totais concedidas pelos mesmos, nas suas partidas de atuação. É possível identificar dois grupos de árbitros através da Figura 2, classificados em 2 categorias utilizando o método de *k-means* com base na mediana de faltas.

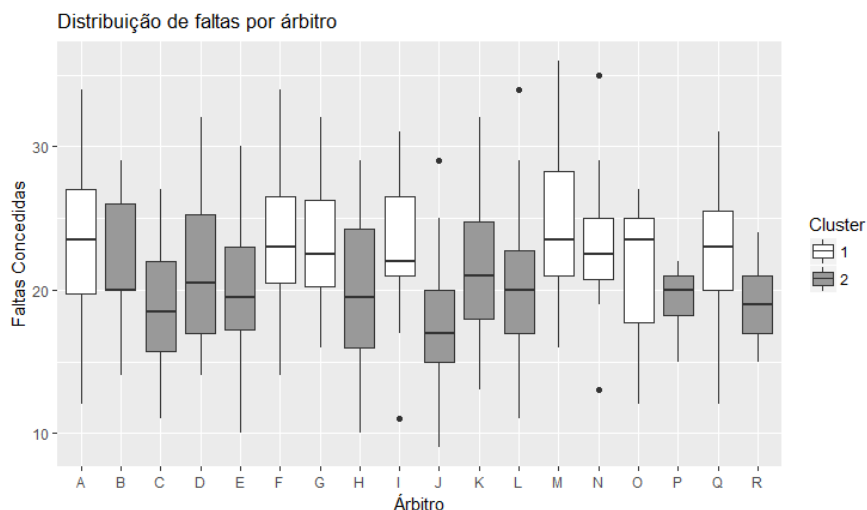


FIGURA 2 – Distribuição das faltas por árbitro, categorizados por cluster

Utilizando o modelo de Poisson, avaliamos o efeito de clássico, público, rodada, valor dos times e arbitragem na distribuição de faltas totais e cartões amarelos para mandante e visitante. Para padronizar a informação de público da partida, utilizamos o percentual de lotação

do estádio, considerando o público pagante e a capacidade do estádio. Foi possível identificar efeito significativo de clássico (p-valor < 0.01) e percentual de público (p-valor<0.01) no número de faltas. Em jogos considerados “clássicos”, quando há uma rivalidade maior entre os times por serem da mesma região, a quantidade de faltas é 1.10 vezes a quantidade de faltas de partidas comuns. A diferença entre os grupos de árbitros não foi significativa para a determinação de cartões aplicados à mandantes ou visitantes, indicando assim que o viés favorável ao mandante se deve principalmente ao mando de campo propriamente dito, não podendo ser explicado pelas variáveis analisadas. O número de cartões demonstrou-se influenciado pela rodada em questão, seguindo a tendência de diminuição de faltas e cartões com o passar das rodadas. Os modelos, detalhados na Tabela 1, foram considerados adequados com base na análise gráfica dos resíduos, disponível no anexo.

TABELA 1 – Coeficientes do modelo Poisson considerando o total de faltas, cartões amarelos aplicados ao mandante e ao visitante

	Faltas Totais			Cartões Amarelos Mandante			Cartões Amarelos Visitante		
	Coeficiente	Erro Padrão	p-valor	Coeficiente	Erro Padrão	p-valor	Coeficiente	Erro Padrão	p-valor
Intercepto	3.329	0.105	< 0.01*	0.862	0.416	0.038*	0.065	0.401	0.871
Clássico	0.099	0.037	< 0.01*	0.051	0.148	0.731	0.026	0.135	0.846
Valor Visitante	0	0.0001	0.861	0.001	0.0004	< 0.01*	0.0001	0.0003	0.694
Valor Mandante	0	0.0001	0.702	0.0001	0.0004	0.898	0.0001	0.0004	0.887
Percentual de Público	-0.331	0.12	< 0.01*	-0.674	0.477	0.158	0.542	0.454	0.232
Árbitro	0.138	0.023	< 0.01*	-0.001	0.092	0.99	0.121	0.082	0.138
Rodada	-0.002	0.001	0.054	-0.009	0.004	0.029*	-0.007	0.004	0.041*

* Estatisticamente significativo a 5%

3.2 Fatores Externos

Buscando avaliar os fatores externos envolvidos, verificamos uma maior proporção de jogos iniciados até as 18hrs (76.67%), 11 estádios com gramado sintético (55%), distribuição aproximadamente normal das temperaturas (média 10.68°C, desvio padrão 5.78°C), 40% dos jogos realizados em altitude inferior a 50m e 75% dos campos com as mesmas dimensões. Nos

casos em que comprimento ou largura do gramado eram diferentes, a proporção é mantida (correlação 0.68). Foram registradas 14 situações climáticas diferentes no início das partidas, porém as situações extremas de interesse, como neve e neblina, ocorreram em apenas 8.6%, de modo que não foi analisada a influência do clima nos cenários de interesse.

Analisando as correlações das variáveis correspondentes a fatores externos com público, rodada, quantidade de gols do mandante e visitante e resultado da partida, foi possível identificar forte correlação entre altitude e as demais variáveis; correlação negativa entre rodada e temperatura (-0.43) pois o campeonato inicia no verão; pequena correlação positiva entre percentual de público e rodada (0.09), temperatura (0.06) e gols da equipe mandante (0.13), conforme mostra o correlograma da Figura 3.

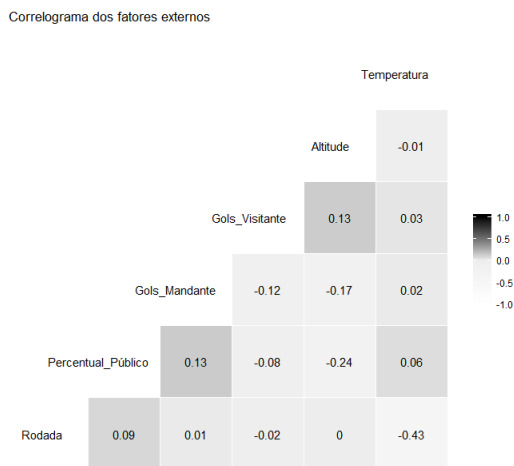


FIGURA 3 – Correlograma das variáveis de fatores externos

A diferença entre o tipo de gramado costuma receber grande destaque por parte de técnicos após as partidas, principalmente quando um time que possui em seu campo gramado natural, disputa uma partida em gramado sintético. Através do teste qui-quadrado verificou-se que a quantidade total de gols por partida é independente do gramado (p-valor = 0.87). De maneira análoga, havendo ou não diferença de gramado entre os times daquela partida, a distribuição do resultado da partida sob a ótica do mandante (vitória, derrota ou empate) é a mesma (p-valor = 0.53), conforme a Figura 4.

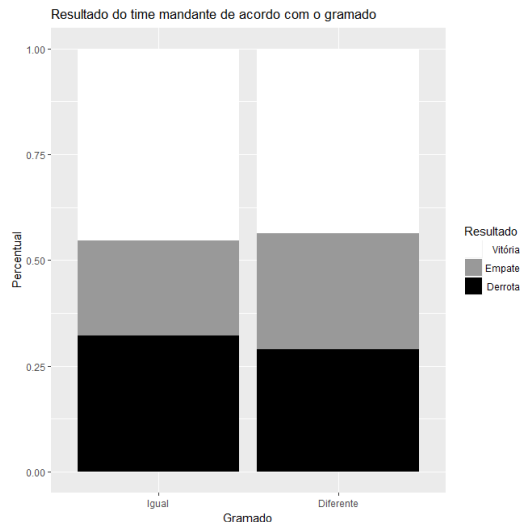


FIGURA 4 – Distribuição do resultado do time mandante de acordo com a diferença de gramado dos times

Outro fator externo de interesse é o público da partida. Primeiramente particionamos nosso estudo em partidas que o mandante havia vencido na rodada anterior ou não, comparando as distribuições do percentual de público nesses cenários, através do teste de Wilcoxon-Mann-Whitney, verificamos que o percentual médio de público quando o mandante venceu o último jogo (0.93) é superior ao percentual médio de público quando perdeu ou empatou (0.90, p-valor < 0.01). Quando analisamos a vitória do mandante, exclusivamente com base no percentual de público, conforme o modelo binomial com link logito da Tabela 2, para um aumento de 3% no público, espera-se um aumento de 14.2% na razão de chances de vitória comparada às chances de derrota. Nesta análise os empates, correspondentes a 25% das observações, foram desconsiderados, demonstrando a relevância da torcida nas partidas.

TABELA 2 – Coeficientes do modelo Binomial considerando a vitória do mandante

	Vitória/Derrota do mandante		
	Coeficiente	Erro Padrão	p-valor
Intercepto	-3.669	1.156	< 0.01*
Percentual de Público	4.429	1.259	< 0.01*

* Estatisticamente significativo a 5%

3.3 Equipes

O entrosamento entre os jogadores é considerado muito importante no desempenho dos times^[16]. Buscando analisar esta relação, verificou-se que o Everton foi o time que mais manteve seu elenco, com a repetição dos 11 jogadores iniciais em apenas 4 rodadas, o que corresponde a aproximadamente 10% do campeonato. Buscando o entrosamento por setores, em média os times iniciaram 7 partidas com os mesmos atacantes, 8 mantendo os zagueiros e 5 jogos com a repetição de meio de campo. Não foi encontrada correlação entre as variáveis de entrosamento e posição final do campeonato ou valor do time.

A relação entre time e técnico também foi estudada. Das 20 equipes que disputaram o campeonato inglês de 2011/12, houve a contratação de um novo técnico em 4 delas: Sunderland, Queens Park Rangers, Wolverhampton Wanderers e Chelsea. A partir da primeira vitória, analisamos a proporção de vitórias nos jogos anteriores até o momento da demissão do técnico, conforme a Figura 5. É possível identificar a queda de performance dos times nas rodadas anteriores à demissão.

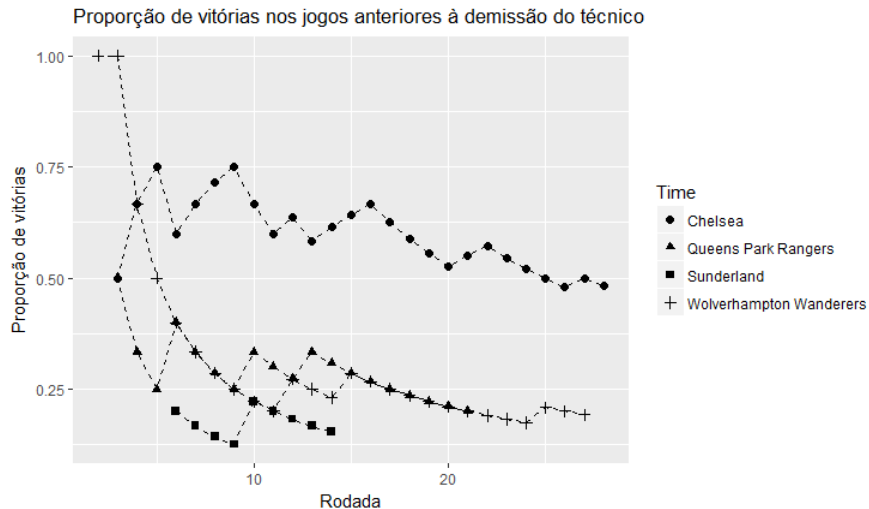


FIGURA 5 – Proporção de vitórias nos jogos anteriores à demissão do técnico

A classificação das equipes, antes e depois da demissão do técnico foi analisada com base na Figura 6. A equipe do Sunderland deixou a 16ª colocação, em que arriscava cair de divisão, terminando o campeonato na 13ª colocação e evitando assim o rebaixamento. Verificamos aqui que a troca de técnicos foi eficiente, por mais que a equipe não tenha conquistado nenhuma vaga em outros campeonatos, ela se manteve na divisão mais importante do futebol inglês.

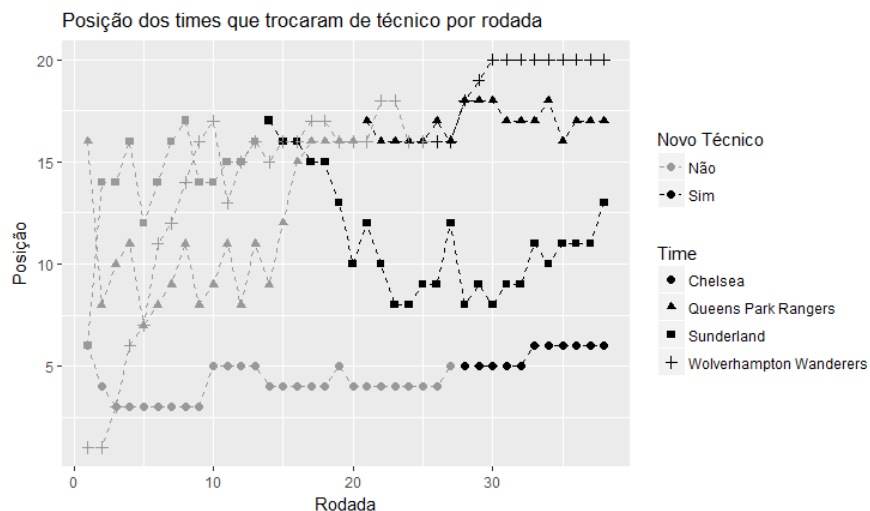


FIGURA 6 – Posição dos times que trocaram de técnico por rodada

A equipe do Queens Park Rangers também estava na 16ª colocação quando teve seu técnico demitido e terminou o campeonato na 17ª posição, sendo rebaixada de divisão. Nos 18 jogos que sucederam a demissão do técnico, a equipe venceu apenas uma das primeiras 8 partidas, caindo para a 18ª colocação. Porém, a equipe venceu 5 dos 10 últimos jogos do campeonato. Essa mudança de performance, no final do campeonato, não foi o bastante para livrar a equipe do rebaixamento, entretanto apontou que, com a adaptação do elenco ao novo técnico, houve uma melhora.

A equipe do Wolverhampton Wanderers também teve seu técnico demitido quando estava na 16ª colocação, beirando a zona de rebaixamento, e terminou o campeonato na última colocação (20ª), caindo de divisão. Ao contrário das demais equipes, esta troca de técnicos não esboçou nenhum resultado positivo, com 13 jogos e nenhuma vitória.

A equipe do Chelsea, por sua vez, teve seu técnico demitido quando estavam na 5ª colocação, brigando por uma vaga na UEFA Champions League, e terminou o campeonato na 6ª colocação. Diferentemente dos outros times, a performance até então era muito boa, porém não o suficiente para conseguir a classificação para a Champions League de 2012/2013. Apesar de o novo técnico ter perdido uma colocação na Premier League, a equipe venceu a Champions League de 2011/12, disputada simultaneamente ao campeonato nacional.

Dentre as quatro equipes que tiveram seus respectivos técnicos demitidos, uma teve melhora imediata de performance e as outras três caíram de colocação. Vale ressaltar que não podemos afirmar que a mudança, ou não, de performance das equipes foi exclusivamente devida à demissão dos técnicos, pois existem diversos outros fatores ligados à performance das equipes, como contusão de jogadores importantes, dificuldade dos últimos jogos, etc. Sabe-se que, apesar de pouco compreendido pelos fãs do esporte, até mesmo os melhores times terão fases ruins, podendo ficar jogos e mais jogos sem vencer, então quem dera aqueles que brigam para não cair de divisão. Fato é que, depois da fase ruim, assumindo que outros fatores permaneçam como antes, a boa fase voltará. Isso é a Regressão à Média. Portanto, os clubes que demitiram seus técnicos durante a fase ruim da equipe podem ter perdido a chance de ver o retorno à fase boa.

Foram analisadas as correlações entre o número de vitórias/empates/derrotas do mandante em casa até a partida em questão, assim como as vitórias/empates/derrotas dos visitantes quando jogaram fora de seu estádio. Estas variáveis foram utilizadas para incorporar o desempenho das equipes e verificar a influência de resultados anteriores nas partidas. De acordo com a Figura 7 é possível identificar que a correlação entre o número de derrotas anteriores do mandante com a vitória do visitante (0.13) é mais acentuada que a correlação entre o número de vitórias anteriores do visitante fora com a sua vitória (0.07).

Correlograma dos fatores de equipe

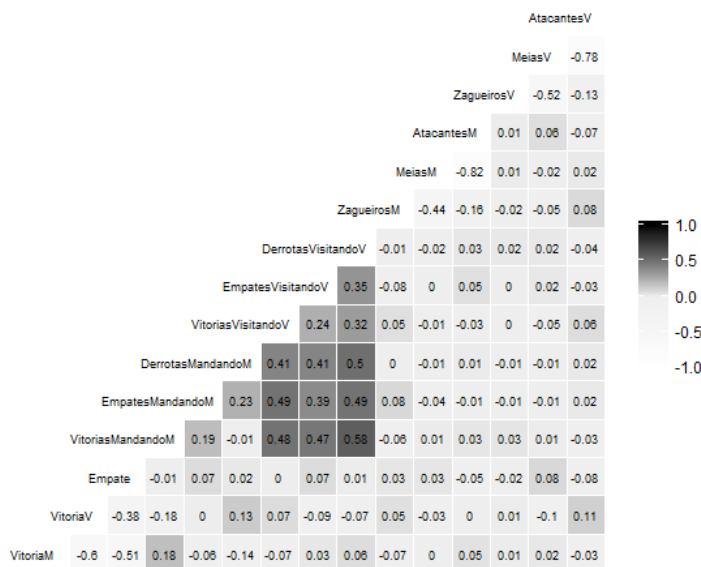


FIGURA 7 – Correlação entre as variáveis de equipe

Outra questão de interesse é a formação tática da equipe. Buscando analisar estratégias de enfrentamento, construímos variáveis representando a quantidade de jogadores por posição (zagueiros, meio campistas e atacantes), com base na posição de ofício dos jogadores iniciais das partidas. A contagem por setor foi necessária pois não havia informação de esquema tático utilizado na base de dados, o que inviabilizou a comparação de esquemas tradicionais utilizados em partidas de futebol^[17], dado que não foi possível verificar a posição na qual o jogador de fato atual. A correlação entre a quantidade de jogadores por setor dos times mandante e visitante é praticamente nula, como mostrado pela Figura 7. Ainda que a posição considerada não tenha sido necessariamente a posição do jogador durante a partida, é possível identificar neste caso como tem sido pouco utilizado o estudo de esquema tático pelas equipes.

Buscando compreender o desempenho das equipes, de maneira geral, tentou-se resumir as 186 variáveis de performance das equipes utilizando análise de componentes principais e análise fatorial. Como a base de dados fornecia a informação para cada jogador atuante na partida, houve a necessidade de agrupá-los por time, devido à grande quantidade de zeros destas variáveis. O uso de componentes principais não surtiu o efeito desejado no conjunto de dados, havendo a necessidade de utilizar mais de 50 componentes para se obter uma variação explicada inferior a 75%. Além disso, com a grande quantidade de colunas tornou-se inviável a análise da relação entre as componentes de maneira gráfica. Quando se utilizou a análise fatorial, não houve convergência do algoritmo. Com isso, optou-se por não utilizar algumas das variáveis de performance no simulador de resultados das partidas descrito na subseção seguinte.

3.4 Jogos

Com o intuito de realizar previsões de resultados das partidas, foram desenvolvidos três modelos logísticos: um para calcular a probabilidade de vitória do time mandante; um para a probabilidade de vitória do time visitante e outro para a probabilidade de empate. Como variáveis preditoras foram utilizadas médias de agrupamentos de variáveis de performance, até o jogo anterior, e também variáveis que apontam a fase do time, até os últimos 5 jogos disputados. Assim, as primeiras 5 rodadas do campeonato não foram utilizadas. Com isso, restaram 310 jogos para desenvolvimento do modelo e 20 jogos para validação.

Para cada modelo, partiu-se do modelo completo, composto pelas variáveis de performance, público e apostas, e então utilizou-se o critério de informação de acaike (AIC) para selecionar o modelo final. Com base na chance de vitória do mandante, empate e vitória do visitante, foi implementada uma regra de decisão do resultado baseada em árvores de classificação e regressão (CART), conforme a Figura 8.

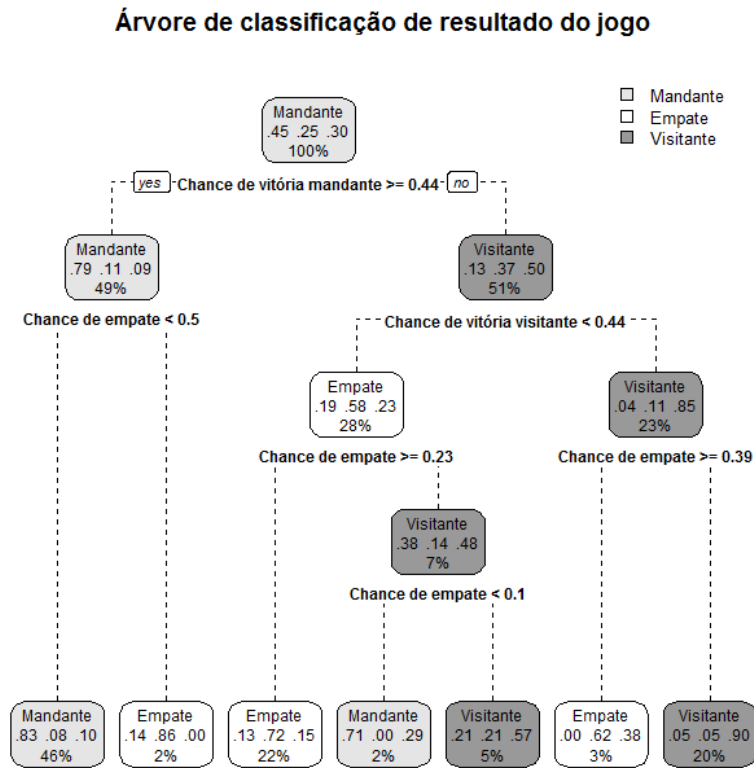


FIGURA 8 – Árvore de classificação de resultado da partida

A regra de classificação adotada foi:

1. Se a chance de vitória do mandante maior que 0.43 e chance de empate menor que 0.5, decide-se por vitória do mandante;
2. Se a chance de vitória do mandante maior que 0.43 e chance de empate maior que 0.5, decide-se por empate;
3. Se a chance de vitória do mandante menor que 0.44, chance de vitória do visitante menor que 0.44 e chance de empate maior que 0.22, decide-se por empate;

4. Se a chance de vitória do mandante menor que 0.44, chance de vitória do visitante menor que 0.44 e chance de empate menor que 0.1, decide-se por vitória do mandante;
5. Se a chance de vitória do mandante menor que 0.44, chance de vitória do visitante menor que 0.44 e chance de empate entre 0.1 e 0.22, decide-se por vitória do visitante;
6. Se a chance de vitória do mandante maior que 0.44, chance de vitória do visitante maior que 0.43 e chance de empate maior que 0.38, decide-se por empate;
7. Se a chance de vitória do mandante maior que 0.44, chance de vitória do visitante maior que 0.43 e chance de empate menor que 0.39, decide-se por vitória visitante.

Após calcular as probabilidades e realizar a predição para os jogos de validação, o acerto de classificação do modelo foi 0.4. Para verificar a qualidade de previsão do modelo proposto, foi feita a comparação com as cotações das apostas. Assim, a aposta que teve o menor valor pago por unidade monetária (u.m) foi considerada como a decisão da previsão do resultado da partida. Isto é, se a aposta pagar 1.1 u.m em caso de vitória do mandante, 5 u.m para empate e 2 u.m para vitória do visitante, consideramos como resultado predito a vitória do time mandante. Aplicando este critério de decisão nos 20 jogos de validação, o acerto de classificação também foi de 0.4. Quando aplicada CART às cotações das apostas, o acerto de classificação caiu para 0.25.

Considerando que o acerto de classificação das predições foi o mesmo, houve a necessidade de considerar outro critério para escolha do melhor modelo. Supondo que fossem realizadas apostas nestas 20 partidas com base no uso dos modelos logísticos ou apenas considerando as apostas, e ainda considerando as cotações pagas, foi possível comparar os dois cenários. No primeiro, haveria um retorno final de 13%, enquanto no segundo, um retorno de -29.3%. Assim, verificou-se que o modelo de árvore de classificação, baseado nas probabilidades de vitória, empate e derrota, dadas pelos modelos logísticos, apesar de possuir um baixo índice de acerto, ainda foi considerado satisfatório, dado o retorno financeiro.

4 CONSIDERAÇÕES FINAIS

Após o levantamento dos principais fatores relevantes que conduzem ao sucesso do time nas partidas, foi possível desmistificar grandes clichês do futebol. Verificou-se que o número de faltas concedidas a mandantes e visitantes segue a mesma distribuição, entretanto há de fato um viés favorável ao mandante na aplicação de cartões amarelos. Este viés não pôde ser atribuído à arbitragem, mesmo com alguns árbitros mais rígidos que outros, no que se refere à marcação de faltas. Além disso, foi possível verificar que o efeito de “clássicos”, proporcionado pela rivalidade entre equipes, gera uma tensão maior dentro de campo: a quantidade de faltas nestas partidas é 10% superior que em jogos comuns. Também foi identificada uma tendência de queda no número de faltas e cartões com o desenvolvimento do campeonato.

No que se refere aos fatores externos, apesar da diferença entre os tipos de gramados receber grande destaque por parte de técnicos, até mesmo na tentativa de justificar derrotas após os jogos, verificou-se que o resultado da partida é independente de haver ou não esta diferença. O percentual médio de público quando o mandante venceu o último jogo é superior à quando o time perdeu ou empatou e esta diferença tem grande influência na chance de vitória do mandante.

Analisando as equipes, houve uma baixa repetição de elenco durante o campeonato. Quanto aos técnicos que foram demitidos, foi identificada uma queda de performance dos times nas rodadas anteriores à sua dispensa. O desempenho dos novos treinadores só surtiu efeito imediato em uma equipe, com as outras três caindo de colocação. É importante ressaltar que não se pode afirmar que a mudança, ou não, de performance dos times foi exclusivamente devida à demissão dos técnicos.

Considerando as predições de resultados, o modelo proposto mostrou-se satisfatório, pois, apesar de possuir um baixo índice de acerto, apresentou um retorno financeiro considerável. Quando comparado com a tomada de decisão baseada nas cotações das apostas o resultado foi ainda mais expressivo. Enquanto em um haveria prejuízo de quase 30%, o outro retorna cerca de 13% de lucro. Driblando a sorte, foi possível identificar a importância de análise de performance histórica das equipes no desfecho das partidas.

REFERÊNCIAS

- [1] OSWALD, V. **Brasil movimentada apenas 2% do mercado da bola.** O Globo. Disponível em: <<https://oglobo.globo.com/brasil/brasil-movimentada-2-do-mercado-da-bola-13026765>>. Acesso em: 18 de junho de 2017.
- [2] LEWIS, MICHAEL. **Moneyball: The art of winning an unfair game.** WW Norton & Company, 2004.
- [3] RAMOS, LUIZ A., E MH OLIVEIRA JÚNIOR. **Futebol: classificação e análise dos gols da Eurocopa 2004.** Revista Brasileira de Futebol 1.1 (2008): 42-48.
- [4] CLAUDINO, JOÃO GUSTAVO DE OLIVEIRA, ET AL. **Análise dos fatores de estresse e ansiedade pré-competitiva em jogadores de futebol de campo da categoria Sub-20.** Revista Brasileira de Futebol 1.1 (2008): 11-19.
- [5] VERARDI, CARLOS EDUARDO LOPES, ET AL. **Estratégias de enfrentamento em jogadores de futebol.** Revista Brasileira de Ciência e Movimento 19.4 (2012): 60-67.
- [6] LEE, ALAN J. **Modeling scores in the Premier League: is Manchester United really the best?** Chance 10.1 (1997): 15-19.
- [7] LOUZADA, FRANCISCO, ADRIANO K. SUZUKI, E LUIS EB SALASAR. **Predicting Match Outcomes in the English Premier League: Which Will Be the Final Rank?** Journal of Data Science 12.2 (2014): 235-254.
- [8] TAVARES, LEANDO E ADRIANO KAMIMURA SUZUKI. **Modelagem estatística para previsão esportiva: uma aplicação no futebol.** Matemática e Estatística em Foco 3.1 (2015): 32-47.
- [9] SOCCERMETRICS RESEARCH. **English Premier League 2011 2012.** Disponível em: <github.com/soccermetrics/project-data/tree/master/english-premier-league-2011-2012>. Acesso em: 27 de julho de 2017.
- [10] TRANSFERMARKET UK. **Clubs of the Premier League 11/12.** Disponível em: <www.transfermarkt.co.uk/premier-league/startseite/wettbewerb/GB1/plus/?saison_id=2011>. Acesso em: 10 de outubro de 2017.
- [11] FOOTBALL DATA RESULTS ODDS TIPSTERS. **Season 2011/2012 Premier League.** Disponível em: <www.football-data.co.uk/mmz4281/1112/E0.csv>. Acesso em: 08 de setembro de 2017.

- [12] BUSSAB, WILTON DE O. E PEDRO A. MORETTIN. **Estatística Básica**. Saraiva, 2010.
- [13] JAMES, GARETH ET AL. **An introduction to statistical learning**. Springer, 2013.
- [14] TURKMAN, MARIA A. A. E SILVA, GIOVANI L. **Modelos Lineares Generalizados: da teoria à prática**. VIII Congresso Anual da Sociedade Portuguesa de Estatística, Lisboa, 2000.
- [15] DUNN, PETER K.; SMYTH, GORDON K. **Randomized quantile residuals**. Journal of Computational and Graphical Statistics, v. 5, n. 3, p. 236-244, 1996.
- [16] ESPN. **Teams such as Leicester, Napoli benefiting from "settled" lineups in title races**. Disponível em: <www.espnfc.com/blog/marcotti-musings/62/post/2835470>. Acesso em: 15 de novembro de 2017.
- [17] TORCEDORES UOL. **Veja como funcionam os principais esquemas táticos do futebol**. Disponível em: <torcedores.uol.com.br/noticias/2014/08/veja-como-funciona-os-principais-esquemas-taticos>. Acesso em: 10 de novembro de 2017.

ANEXOS

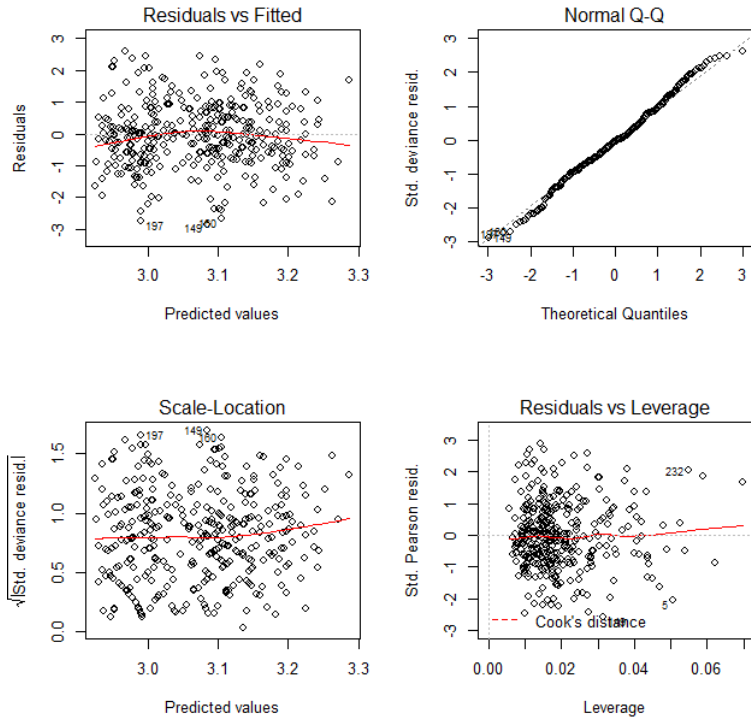


FIGURA 1 – Gráfico de resíduos do modelo de Poisson para total de faltas por partida

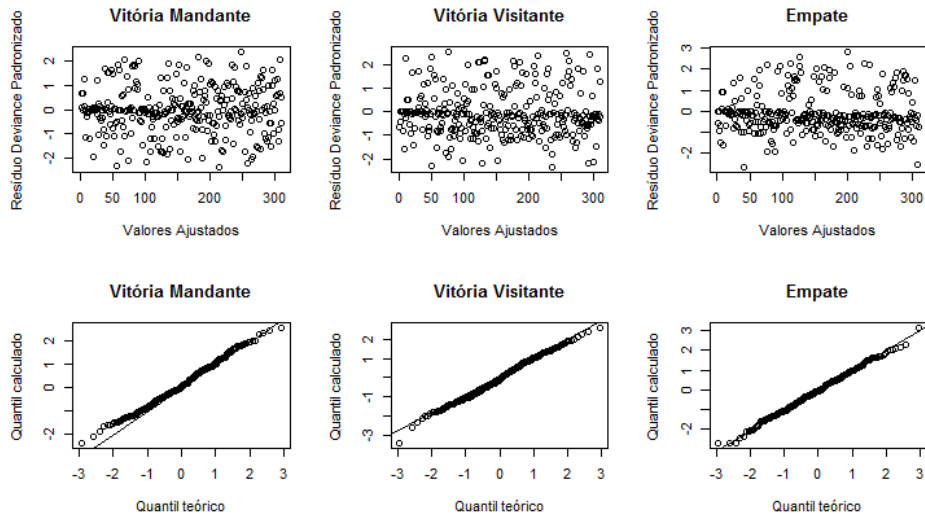


FIGURA 2 – Gráfico de resíduos deviance padronizado e quantil-quantil da normal para os modelos Binomial de vitória do mandante, vitória do visitante e empate

Quadro 1 – Descrição das variáveis de performance durante o jogo disponíveis no conjunto de dados para mandante e visitante

VARIÁVEL	DESCRIÇÃO
SubstitutosOn	Substitutos
SubstitutosOff	Substituídos
PrimeiroGol	Primeiro gol da partida
GolVitoria	Gol da vitoria
ChutesNoGol	Chutes no gol
ChutesFora	Chutes para fora
ChutesBloqueados	Chutes bloqueados
PenaltisAdquiridos	Penaltis adquiridos
GolsPenaltis	Gols de penaltis
PenaltisDefendidos	Penaltis defendidos
PenaltisFora	Penaltis para fora
PenaltisPerdidos	Penaltis perdidos
GolsFaltas	Gols de faltas
FaltasNoGol	Cobranca de faltas que acertaram o gol
FaltasFora	Cobranca de faltas que foram para fora
FaltasBloqueadas*	Cobranca de faltas que foram bloqueadas
GolsPerto	Gols de dentro da área
ChutesPertoNoGol	Chutes de dentro da área que acertaram o gol
ChutesPertoFora	Chutes de dentro da área que foram para fora
ChutesBloqueadosPerto	Chutes de dentro da área que foram bloqueados
GolsLonge	Gols de fora da área
ChutesLongeNoGol	Chutes de fora da área que acertaram o gol
ChutesLongeFora	Chutes de fora da área que foram para fora
ChutesBloqueadosLonge	Chutes de fora da área que foram bloqueados
GolsCabeceios	Gols feitos por cabeceios
CabeceiosNoGol	Cabeceios que acertaram o gol
CabeceiosFora	Cabeceios que foram para fora
CabeceiosBloqueados	Cabeceios que foram bloqueados
GolsEsquerda	Gols com o pé esquerdo
ChuteEsquerdaNoGol	Chutes com o pé esquerdo que acertaram o gol
ChuteEsquerdaFora	Chutes com o pé esquerdo que foram para fora
ChuteEsquerdaBloqueados	Chutes com o pé esquerdo que foram bloqueados
GolsDireita	Gols com o pé direito
ChuteDireitaNoGol	Chutes com o pé direito que acertaram o gol
ChutedireitaFora	Chutes com o pé direito que foram para fora
ChuteDireitaBloqueados	Chutes com o pé direito que foram bloqueados
OutrosGols	Outros tipos de gols
OutrosChutes	Outros tipos de chutes que acertaram o gol
OutrosChutesFora	Outros tipos de chutes que foram para fora

OutrosChutesBloqueados	Outros chutes bloqueados que foram
ChutesSalvosNaLinha	Chutes salvos na linha
ChutesPertoSalvosNaLinha	Chutes de perto salvos na linha
ChutesLongeSalvosNaLinha	Chutes de longe salvos na linha
GolsJogoAberto	Gols em jogada aberta
GolsPosEscanteio	Gols pos cobrança de escanteio
GolsPosLateral	Gols pos cobrança de lateral
GolsFalta	Gols em cobrança de falta
GolsJogEnsaaiadas	Gols em jogada ensaiada
GolsPenaltis	Gols de penaltis
TentativaJogoAberto	Tentativa de gol que acertou o gol em jogada aberta
TentativaPosEscanteio	Tentativa de gol que acertou o gol pos cobrança de escanteio
TentativaPosLateral	Tentativa de gol que acertou o gol pos cobrança de lateral
TentativaFaltas	Tentativa de gol que acertou o gol em cobrança de faltas
TentativaJogEnsaaiadas	Tentativa de gol que acertou o gol em jogada ensaiada
TentativaPenaltis	Tentativa de gol que acertou o gol em pênaltis
TentativaJogoAbertoFora	Tentativa de gol para fora em jogada aberta
TentativaForaPosEscanteio	Tentativa de gol para fora pos cobrança de escanteio
TentativaForaPosLateral	Tentativa de gol para fora pos cobrança de lateral
TentativaFaltasFora	Tentativa de gol para fora em cobrança de faltas
TentativaJogEnsaaiadasFora	Tentativa de gols para fora em jogada ensaiada
TentativaPenaltisFora	Tentativa de gols para fora em penaltis
GolsSubstitutos	Gols feitos por substitutos realizados
Passes_S	Passes certos realizados
Passes_U	Passes errados realizados
Assistencias	Assistencia realizada
PasseExclCruzamentos_S	Passes certos (excluindo cruzamentos)
PasseExclCruzamentos_U	Passes errados (excluindo cruzamentos)
PasseCampoProprio_S	Passes certos na metade defensiva do campo
PasseCampoProprio_U	Passes errados na metade defensiva do campo
PasseCampoAdversario_S	Passes certos na metade ofensiva do campo
PasseCampoAdversario_U	Passes errados na metade ofensiva do campo
PasseDefesa_S	Passes certos na defesa
PasseDefesa_U	Passes errados na defesa
PasseVeio_S	Passes certos no meio de campo
PasseVeio_U	Passes errados no meio de campo
PasseAtaque_S	Passes certos no ataque
PasseAtaque_U	Passes errados no ataque
PasseCurto_S	Passes curtos certos
PasseCurto_U	Passes curtos errados
PasseLongo_S	Passes longos certos
PasseLongo_U	Passes longos errados
PassePrimeira_S	Passes de primeira certos

PassePrimeira_U	Passes de primeira errados
EscanteiosCruzados_S	Escanteios cobrados em forma de cruzamento corretos
EscanteiosCruzados_U	Escanteios cobrados em forma de cruzamentos errados
EscanteiosAdquiridos	Escanteios adquiridos
EscanteiosConcedidos	Escanteios concedidos
Escanteiosarea_S	Escanteios cobrados na área certos
Escanteiosarea_U	Escanteios cobrados na área errados
EscanteiosCurtos	Escanteios curtos
Laterais_S	Laterais cobrados certos
Laterais_U	Laterais cobrados errados
Dribles_S	Dribles certos
Dribles_U	Dribles errados
EscanteiosCruzadosEsquerda_S	Escanteios cobrados em forma de cruzamento pela esquerda certos
EscanteiosCruzadosEsquerda_U	Escanteios cobrados em forma de cruzamento pela esquerda errados
CruzamentosEsquerda_S	Cruzamentos da esquerda certos
CruzamentosEsquerda_U	Cruzamentos da esquerda errados
EscanteiosEsquerda_S	Escanteios da esquerda certos
EscanteiosEsquerda_U	Escanteios da esquerda errados
EscanteiosCruzadosDireita_S	Escanteios cobrados em forma de cruzamento pela direita certos
EscanteiosCruzadosDireita_U	Escanteios cobrados em forma de cruzamento pela direita errados
CruzaVentosDireita_S	Cruzamentos da direita certos
CruzaVentosDireita_U	Cruzamentos da direita errados
EscanteiosDireita_S	Escanteios da direita certos
EscanteiosDireita_U	Escanteios da direita errados
Lancamentos_S	Lancamentos certos
Lancamentos_U	Lancamentos errados
AjeitadaBola_S	Ajeitadas de bola certas
AjeitadaBola_U	Ajeitadas de bola erradas
BolasEnfiadas	Bolas enfiadas
EscanteiosAltos_S	Escanteios altos certos
EscanteiosAlto_U	Escanteios altos errados
CruzamentosAltos_S	Cruzamentos altos certos
CruzamentosAltos_U	Cruzamentos altos errados
Cruzamentos_S	Cruzamentos em jogada aberta certos
Cruzamentos_U	Cruzamentos em jogada aberta errados
Toques	Toques na bola
AssistenciaEscanteio	Assistencia por cobranca de escanteio
AssistenciaFalta	Assistencia por cobranca de falta
AssistenciaLateral	Assistencia por cobranca de lateral
AssistenciaTiroVeta	Assistencia por cobranca de tiro de meta
AssistenciaJogEnsaiaada	Assistencia em jogada ensaiada

Duelos_S	Duelos ganhos
Duelos_U	Duelos perdidos
DuelosAr_S	Duelos aereos ganhos
DuelosAr_U	Duelos aereos perdidos
DuelosChao_S	Duelos no chao ganhos
DuelosChao_U	Duelos no chao perdidos
Desarme_S	Tentativa de desarme bem sucedida
Desarme_U	Tentativa de desarme mal sucedida
DesarmeUltimoHomem	Desarme feito como ultimo homem
BolasSalvas	Jogada de bola para longe
SalvouCabeca	Jogada de bola para longe de cabeca
SalvouOutros	Jogada de bola para longe de outras maneiras
SalvouNaLinha	Jogada de bola para longe quando a bola estava entrando no gol
Bloqueio	Bloqueio de chutes
Interceptacao	Interceptacao
RecuperacaoPosse	Recuperacao de posse de bola
FaltasConcedidas	Faltas concedidas
FaltasConcedidasExcHandballs	Faltas concedidas (excluindo mao na bola)
FaltasAdquiridas	Faltas adquiridas
FaltasPertoAdquiridas	Faltas de perto adquiridas
FaltasLongeAdquiridas	Faltas de longe adquiridas
PenaltisAdquiridos	Penaltis adquiridos
HandballsConcedidos	Faltas concedidas por mao na bola pelo time
PenaltisConcedidos	Penaltis concedidos
Impedimentos	Impedimentos do time
CartaoAmarelo	Cartoes amarelos para o time
CartaoVermelho	Cartoes vermelhos para o time
GolsConcedidos	Gols concedidos
GolsConcedidosPerto	Gols de perto concedidos
GolsConcedidosLonge	Gols de longe concedidos
ChutesDefendidos	Chutes defendidos
ChutesPertoDefendidos	Chutes de perto defendidos
ChutesLongeDefendidos	Chutes de longe defendidos
PenaltisDefendidos	Penaltis defendidos
CatchesGoleiro	Catches feitos pelo goleiro do time
SocosGoleiro	Socos feitos pelo goleiro do time
DropsGoleiro	Drops feitos pelo goleiro do time
EscanteiosSemIntercepGoleiro	Escanteios sem interceptacao do goleiro do time
DistribuicaoBolaGoleiro	Distribuicao de bola realizadas pelo goleiro do time
DistribuicaoBolaGoleiro_S	Distribuicao de bola correta realizada pelo goleiro do time
DistribuicaoBolaGoleiro_U	Distribuicao de bola errada realizada pelo goleiro do time
CleanSheetJogadores	Quantidade de partidas sem sofrer gols dos jogadores do time

CleanSheet	Quantidade de partidas sem sofrer gols do time
ErrosResultandoGol	Erros do time que resultaram em gols
ErrosResultandoChance	Erros do time que resultaram em chance de gol
DuelosPerdidos	Duelos perdidos
ChutesConcedidos	Chutes concedidos
ChutesConcedidosPerto	Chutes de perto concedidos
ChutesConcedidosLonge	Chutes de longe concedidos
PerdaPosse	Perda de posse de bola do time
Desarmado	Desarme efetuado contra o time
GrandesChances	Grandes chances de gol do time
GigantesChances	Chances claras de gol do time
PasseFrente	Passes para frente
PasseTras	Passes para tras
PassesEsquerda	Passes na esquerda do campo
PassesDireita	Passes na direita do campo
Toques_U	Toques na bola errados
Toques_S	Toques na bola certos
TakeOnOverrun	Take On`s Overrun
ToquesAtaque	Toques na bola no ataque
ToquesGrandearea	Toques na bola na grande área
ToquesPeqarea	Toques na bola na pequena área
