



UNIVERSIDADE FEDERAL DE MINAS GERAIS
ESCOLA DE ENGENHARIA

FELIPE MACHADO DOS SANTOS

**REDES NEURAS ARTIFICIAIS PARA PREVISÃO DE
RESULTADOS NO CAMPEONATO BRASILEIRO DE
FUTEBOL**

BELO HORIZONTE

2018

FELIPE MACHADO DOS SANTOS

**REDES NEURAIS ARTIFICIAIS PARA PREVISÃO DE
RESULTADOS NO CAMPEONATO BRASILEIRO DE
FUTEBOL**

Monografia apresentada ao Departamento de Engenharia Mecânica da Escola de Engenharia da Universidade Federal de Minas Gerais como requisito parcial para obtenção do título de Bacharel em Engenharia Mecânica.

Orientador:

Prof. Dr. Eduardo José Lima II

BELO HORIZONTE

2018

Agradecimentos

Agradeço principalmente ao professor Eduardo José Lima II pelo apoio e orientação durante o projeto. Agradeço aos examinadores Adolfo Guerra e Fernanda Lima pela disponibilidade de avaliar este trabalho. Por fim, agradeço ao menino Ney.

Resumo

As redes neurais artificiais são sistemas que, para fenômenos bem comportados, são capazes de obter alto nível de generalização e assim prever resultados para dados de entrada futuros depois de passar por um processo de treinamento com amostras e dados históricos. O desafio proposto nesse trabalho é utilizar as redes neurais artificiais para previsão de fenômenos altamente aleatórios, como é o caso de resultados de uma partida de futebol. Existem inúmeras variáveis que podem influenciar no resultado de um jogo e não podem ser facilmente detectadas e quantificadas, limitando a taxa de acerto obtida pelas redes neurais. A qualidade do resultado das redes neurais será avaliado pelo retorno de investimento que esta teria ao fazer apostas em sites de apostas esportivas. Dessa maneira, mesmo que as taxas de acerto da rede sejam baixas, elas podem ser altas o suficiente para gerar lucros ao realizar apostas, indicando que existe valor no método. Após diversas simulações e alteração de parâmetros da rede, chegou-se a conclusão de que o método é promissor, mas precisa ser melhor dominado para ter sucesso. Alguns anos apresentam resultados de retorno de investimento entre 10-15%, mas devido a aleatoriedade e alta variabilidade dos resultados, esses mesmos anos podem apresentar resultados de prejuízo, quando parâmetros diferentes são utilizados. Como é impossível prever quais parâmetros levarão a melhores resultados em um ano futuro, o método é inconsistente. Caso o método seja aperfeiçoado, com novos dados de entrada e estratégias de aposta que tenham uma relação mais significativa com o resultado, é possível que sejam gerados resultados positivos de forma consistente. Esse trabalho mostra que é possível obter resultados realmente bons – o que falta é apenas entender o caminho para atingí-los consistentemente.

Palavras-chave: Redes Neurais Artificiais; Previsão; Classificação; Apostas esportivas; Campeonato Brasileiro de Futebol

Abstract

Artificial neural networks are systems that, for well behaved phenomena, can obtain a high level of generalization and therefore predict results for future data, after going through a process of training using historic data. The challenge proposed on this work is to utilize the artificial neural networks to try to predict highly random phenomena, such as the results of a soccer match. There are a vast amount of variables that can influence the result of a game, but cannot be easily detected and quantified, limiting the prediction accuracy obtained by the neural networks. The quality of the networks will be defined by the return on investment that it achieves by betting on sports betting websites. This way, even if the prediction accuracy rate is low, they can still be high enough to generate profits on the bets, indicating that there is value on the method. After many simulations and changes in the networks parameters, the conclusion is that the method is promising, but still needs refinement to be successful. In some years, results of return on investment are between 10-15%, but due to the high randomness and variability of the results, these same years can present loss when different parameters are used. Since it is impossible to predict which parameters will lead to better results in a future year, the method is inconsistent. If the method is refined, with new input data and betting strategies that have a more significant impact on the result, it is highly likely that the results generated by this method will be consistently positive. This work shows that it is possible to find really good results – what is missing is just understanding a way to reach them consistently.

Keywords: Artificial Neural Networks; Prediction; Classification; Sports Betting; Brazilian Soccer Championship

Lista de Figuras

| | |
|---|----|
| Figura 1: Componentes básicos de um neurônio biológico | 10 |
| Figura 2: esquema representativo de um neurônio utilizado em redes neurais artificiais | 12 |
| Figura 3: Tipos de função de ativação | 13 |
| Figura 4: Esquema da estrutura de uma rede neural artificial com diversas camadas | 14 |
| Figura 5: Comparativo entre redes sub-treinadas, bem treinadas e treinadas em excesso. | 19 |
| Figura 6: modelo representativo do diagrama de confusão e legenda explicativa dos campos | 28 |
| Figura 7: diagrama de confusão retirado do Matlab® | 28 |
| Figura 8: Exemplo de gráfico de distribuição de apostas e taxa de acerto por resultado | 29 |
| Figura 9: exemplo de gráfico de retorno de investimento | 33 |
| Figura 10: listagem de amostras a serem utilizadas para treinamento | 59 |
| Figura 11: diagrama de confusão obtido pelo método estatístico | 60 |
| Figura 12: diagrama de confusão obtido pelo método estatístico para os anos 2014-2017 | 62 |
| Figura 13: diagrama de confusão obtido pelas redes neurais com setup básico e estratégia de aposta bet on all..... | 65 |
| Figura 14: diagrama de confusão obtido pelas redes neurais com setup básico e estratégia de aposta bet if minimum..... | 66 |
| Figura 15: diagrama de confusão obtido pelas redes neurais com setup básico e estratégia de aposta bet if value | 67 |
| Figura 16: faixas de ROI obtidas pelas redes neurais com setup básico nas três estratégias de aposta para os anos teste de 2014-2017 | 69 |
| Figura 17: distribuição de apostas e taxa de acerto por tipo de resultado com setup básico e variação do parâmetro mod2min nos anos teste de 2014-2017..... | 72 |
| Figura 18: faixas de ROI obtidas pelas redes neurais com setup básico e variação do parâmetro mod2min nos anos teste de 2014-2017 com o método bet if minimum.... | 73 |

| | |
|--|----|
| Figura 19: distribuição de apostas e taxa de acerto por tipo de resultado com setup básico e variação do parâmetro mod3min nos anos teste de 2014-2017..... | 74 |
| Figura 20: faixas de ROI obtidas pelas redes neurais com setup básico nas três estratégias de aposta para os anos teste de 2014-2017 com o método bet if value . | 75 |
| Figura 21: distribuição de apostas e taxa de acerto por tipo de resultado com setup básico para cada um dos anos teste entre 2014-2017, para o caso bet on all..... | 76 |
| Figura 22: faixas de ROI obtidas pelas redes neurais com setup básico em cada ano teste entre 2014-2017 para cada método de aposta | 77 |
| Figura 23: ROI mediano obtido pelas redes neurais com cada método de aposta e para cada ano entre 2014-2017 e valor obtido pelo método estatístico | 80 |
| Figura 24: distribuição de apostas e taxa de acerto por tipo de resultado com setup básico para diferentes quantidades de anos utilizados como treinamento, para o caso bet on all, e valores de PMV médio | 81 |
| Figura 25: faixas de ROI obtidas pelas redes neurais com setup básico e diferentes quantidades de anos utilizados como amostras de treinamento | 82 |
| Figura 26: distribuição de apostas com setup básico para um número de neuronios na camada escondida entre 5 e 2000 para o caso bet on all (escala variável) | 83 |
| Figura 27: taxa de acerto total e PMV com setup básico para um número de neuronios na camada escondida entre 5 e 2000 para o caso bet on all (escala variável)..... | 86 |
| Figura 28: faixas de ROI obtidas pelas redes neurais com setup básico e diferentes quantidades de neurônios na camada escondida para o método de aposta bet on all | 87 |
| Figura 29: faixas de ROI obtidas pelas redes neurais com setup básico e diferentes quantidades de neurônios na camada escondida para o método de aposta bet if minimum..... | 87 |
| Figura 30: faixas de ROI obtidas pelas redes neurais com setup básico e diferentes quantidades de neurônios na camada escondida para o método de aposta bet if value | 88 |
| Figura 31: pontuações obtidas por cada grupo de entradas com cada ano entre 2014-2017 utilizado como amostras de teste..... | 91 |
| Figura 32: pontuações obtidas por cada grupo de entradas com cada ano entre 2014-2017 utilizado como amostras de teste - combinado..... | 92 |

| | |
|--|----|
| Figura 33: pontuações médias obtidas pelas redes neurais para cada quantidade de entradas utilizadas na análise | 94 |
| Figura 34: comparativo entre as pontuações obtidas por cada grupo de entrada e identificação dos grupos de entradas selecionados e descartados para a Análise 9.95 | |
| Figura 35: pontuações obtidas por cada entrada individual com cada ano entre 2014-2017 utilizado como amostras de teste | 96 |
| Figura 36: pontuações obtidas por cada entrada individual entre 2014-2017..... | 97 |
| Figura 37: pontuações combinadas obtidas por cada entrada individual ao longo dos 4 anos de teste | 98 |
| Figura 38: retornos de investimento para 2017 em função do ROI mínimo utilizado na validação de uma rede neural gerada | 99 |

Lista de Tabelas

| | |
|---|----|
| Tabela 1: grupos de variáveis de entrada, variáveis de entrada e descrição das variáveis de entrada a serem utilizadas..... | 23 |
| Tabela 2: Dados de entrada para cálculo do exemplo de resultado de uma rede..... | 34 |
| Tabela 3: exemplo de cálculo para pontuação na situação definida na Tabela 2 | 34 |
| Tabela 4: siglas utilizadas na apresentação dos resultados | 35 |
| Tabela 5: resumo das análises a serem realizadas..... | 36 |
| Tabela 6: variáveis, fórmulas e probabilidades de ocorrência de cada quantidade de gols entre 0 e 5 para o time mandante..... | 40 |
| Tabela 7: variáveis, fórmulas e probabilidades de ocorrência de cada quantidade de gols entre 0 e 5 para o time mandante..... | 40 |
| Tabela 8: possíveis placares em que cada equipe fazem 5 gols ou menos..... | 40 |
| Tabela 9: cálculo da probabilidade de ocorrência de cada um dos resultados apresentados na Tabela 8..... | 41 |
| Tabela 10: probabilidades de ocorrência de cada um dos resultados apresentados na Tabela 8..... | 41 |
| Tabela 11: cálculo das probabilidades de ocorrência dos três possíveis resultados . | 42 |
| Tabela 12: entradas e saídas utilizadas pelo programa da Análise 2..... | 43 |
| Tabela 13: distribuição de apostas e taxa de acerto por tipo de resultado para o método estatístico | 60 |
| Tabela 14: taxa de ocorrência real e taxa de previsão do resultado para o método estatístico | 61 |
| Tabela 15: distribuição de apostas e taxa de acerto por tipo de resultado para o método estatístico, apenas para os anos de 2014-2017..... | 62 |
| Tabela 16: distribuição de apostas, taxa de acerto por tipo de resultado e taxa de acerto para o método estatístico em cada ano entre 2014-2017 | 62 |
| Tabela 17: ROI obtido utilizando o método estatístico em cada ano enter 2014-2017 | 63 |
| Tabela 18: distribuição de ocorrência real de resultados nas temporadas 2007 - 2017 | 64 |
| Tabela 19: distribuição de apostas e taxa de acerto por tipo de resultado para a rede neural com setup básico e estratégia de apostas bet on all..... | 65 |

| | |
|--|----|
| Tabela 20: distribuição de apostas e taxa de acerto por tipo de resultado para a rede neural com setup básico e estratégia de apostas bet if minimum | 67 |
| Tabela 21: distribuição de apostas e taxa de acerto por tipo de resultado para a rede neural com setup básico e estratégia de apostas bet if value..... | 68 |
| Tabela 22: resumo geral da taxa de acerto geral, distribuição de apostas, taxa de acerto por tipo de aposta, taxa de ocorrência real, taxa de previsão de resultado, payout médio em caso de vitória e retorno de investimento para as três estratégias de aposta e método estatístico | 70 |
| Tabela 23: número de jogos apostados, distribuição de apostas e taxas de acerto para a estratégia de apostas bet if minimum em função do valor de mod2min | 71 |
| Tabela 24: número de jogos apostados, distribuição de apostas, taxas de acerto de apostas para a estratégia de apostas bet if value em função do valor de mod3min.. | 73 |
| Tabela 25: jogos apostados, distribuição de apostas, taxas de acerto e payout médio em caso de vitória pela estratégia bet if minimum em cada ano de teste entre 2014-2017 | 78 |
| Tabela 26: relação entre taxa de acerto total e payout médio em caso de vitória para cada ano teste entre 2014-2017 | 79 |
| Tabela 27: jogos apostados, taxas de vitória, payout médio em caso de vitória e ROI obtidos pela estratégia mod3min em 2017..... | 79 |
| Tabela 28: proporção de apostas em função do número de neurônios na camada escondida para alguns pontos de interesse | 84 |
| Tabela 29: taxas de acerto em função do número de neurônios na camada escondida | 85 |
| Tabela 30: relação entre siglas utilizadas e os grupos de variáveis com uma descrição e número de entradas contidas em cada grupo | 90 |
| Tabela 31: piores e melhores resultados obtidos em cada ano por diferentes setups diferindo pelos grupos de entradas utilizados | 93 |
| Tabela 32: piores e melhores resultados obtidos em cada ano por diferentes setups | 97 |

Lista de Equações

| | |
|--|----|
| Equação 1: saída de um neurônio j do tipo McCulloch e Pitts com n entradas, w_{ij} como peso entre cada conexão do neurônio j com i entradas x_i e valor de bias b | 12 |
| Equação 2: função sigmóide..... | 18 |
| Equação 3: derivada da função sigmóide..... | 18 |
| Equação 4: equação para cálculo da força de ataque do time mandante | 38 |
| Equação 5: equação para cálculo da força de ataque do time visitante..... | 38 |
| Equação 6: equação para cálculo da força de defesa do time mandante..... | 38 |
| Equação 7: equação para cálculo da força de defesa do time visitante..... | 39 |
| Equação 8: cálculo da quantidade de gols esperadas para o time mandante | 39 |
| Equação 9: cálculo da quantidade de gols esperadas para o time visitante | 39 |
| Equação 10: probabilidade de k ocorrências de um evento com ocorrência esperada λ pelo método de Poisson | 39 |

Lista de Símbolos e Abreviações

BIM – bet if minimum

BIV – bet if value

BOA – Bet on all

GA – Gols feitos pelo time visitante

GC – Gols contra

GH – Gols feitos pelo time da casa

GP – Gols pró

PMV – Payout médio em caso de vitória

RNA – Rede Neural Artificial

ROI – Retorno de investimento

V1 – Vitória do mandante

V2 - Empate

V3 – Vitória do visitante

Sumário

| | | |
|----------|--|------------|
| 1 | Introdução | 1 |
| 2 | Revisão Bibliográfica | 3 |
| 2.1 | Apostas e apostas esportivas | 3 |
| 2.2 | Redes Neurais Artificiais | 7 |
| 2.3 | Inteligência artificial em apostas esportivas | 20 |
| 3 | Metodologia..... | 22 |
| 3.1 | Construção das entradas de treinamento | 22 |
| 3.2 | Treinamento das redes neurais artificiais..... | 26 |
| 3.3 | Resultados das redes neurais artificiais | 27 |
| 3.4 | Análises a serem realizadas | 36 |
| 3.4.1 | Análise 1 – Eficiência do método estatístico..... | 38 |
| 3.4.2 | Análise 2 – redes neurais com <i>setup</i> básico | 43 |
| 3.4.3 | Análise 3 – Estratégia de aposta BIM | 49 |
| 3.4.4 | Análise 4 – Estratégia de aposta BIV | 49 |
| 3.4.5 | Análise 5 – Variação de resultado ano a ano..... | 49 |
| 3.4.6 | Análise 6 – necessidade de dados de entrada retroativos | 50 |
| 3.4.7 | Análise 7 – influência do número de neurônios na camada escondida | 50 |
| 3.4.8 | Análise 8 – direcionamento das variáveis de entrada mais importantes | 50 |
| 3.4.9 | Análise 9 – Influência individual de variáveis de entrada..... | 57 |
| 3.4.10 | Análise 10 – Direcionamento estratégico das redes | 57 |
| 4 | Resultados | 59 |
| 4.1 | Levantamento de entradas | 59 |
| 4.2 | Resultado das análises propostas | 60 |
| 4.2.1 | Eficiência do método estatístico | 60 |
| 4.2.2 | Redes neurais com <i>setup</i> básico | 65 |
| 4.2.3 | Estratégia de aposta BIM..... | 71 |
| 4.2.4 | Estratégia de aposta BIV | 73 |
| 4.2.5 | Variação de resultado ano a ano | 75 |
| 4.2.6 | Necessidade de dados de entrada retroativos | 80 |
| 4.2.7 | Influência do número de neurônios na camada escondida | 82 |
| 4.2.8 | Direcionamento das variáveis de entrada mais importantes..... | 89 |
| 4.2.9 | Influência individual de variáveis de entrada | 96 |
| 4.2.10 | Direcionamento estratégico das redes | 99 |
| 5 | Conclusões..... | 101 |
| 6 | Sugestões de Trabalhos Futuros | 103 |
| 7 | Bibliografia..... | 104 |

1 Introdução

As apostas são fenômenos presentes durante grande parte da história da humanidade, e estão relacionadas a convicção do apostador de que consegue prever os resultados de um evento mais precisamente do que outras pessoas. Existe uma conotação negativa quanto ao termo em consequência da atenção da mídia aos jogos de azar – jogos desenhados de forma a garantir a derrota do apostador e a vitória do estabelecimento. Entretanto, em ocasiões onde as apostas são feitas entre os jogadores e apenas intermediadas pela casa, as apostas são apenas uma maneira de premiar aqueles com um conhecimento mais aprofundado sobre o evento em relação aos outros jogadores. Por mais que haja um fator sorte envolvido, apostadores melhor preparados tendem a gerar resultados positivos a longo prazo (Feustel & Howard, 2010).

Historicamente, as apostas eram feitas por um indivíduo a partir de seu conhecimento natural e dados estatísticos. Modelos matemáticos e computacionais não eram utilizados até o início dos anos 2000, pois o processo de decisão envolvido é altamente complexo e entendia-se que era necessário um raciocínio mais profundo do que os computadores conseguiam produzir naquela época.

Entretanto, com o advento da inteligência artificial e os algoritmos de machine learning, computadores estão ganhando cada vez mais a capacidade de tomar decisões por processos de pensamento paralelos ao que acontece nos cérebros de animais. Esses métodos são capazes de, a partir do treinamento de um sistema computacional com exemplos passados, fazer com que ele defina os resultados esperados para um evento com dados de entrada que nunca antes viu.

Para fenômenos bem comportados, as redes neurais atingem taxas de acerto enormes, como o reconhecimento de mais de 90% dos caracteres em um texto escrito a mão (Dash & Nayak, 2012). Eles são os mesmos métodos utilizados por serviços como o Spotify e o Netflix para oferecer sugestões de conteúdo para seus usuários de forma tão assertiva. Dessa maneira, surge-se a ideia de utilizar as redes neurais artificiais para previsão de resultados tidos como altamente aleatórios – a previsão do resultado de uma partida de futebol.

O objetivo deste trabalho é observar a capacidade das Redes Neurais Artificiais (RNA) de prever resultados de partidas do Campeonato Brasileiro de Futebol. A qualidade do modelo será medida ao fazer uma simulação de apostas em sites de aposta e observar o retorno financeiro obtido. Caso o retorno seja positivo, pode-se concluir que a rede neural consegue prever os resultados de uma maneira acima da média comparado à média dos métodos atualmente utilizados, uma vez que os lucros em apostas são uma consequência de se realizar apostas melhor do que a média de todos os outros apostadores.

Na verdade, as apostas em jogos de futebol serão utilizadas apenas como um meio. O objetivo final é a investigação do comportamento das redes neurais em eventos altamente aleatórios, onde um grande número de variáveis pode influenciar o resultado, e apenas uma pequena parte delas pode ser quantificada ou conectada com o resultado de alguma maneira. Deseja-se observar os resultados das variações de diversos parâmetros na configuração das Redes Neurais para obter mais informações sobre como utilizá-los.

Por fim, este trabalho almeja aumentar o entendimento dos discentes da Engenharia Mecânica da UFMG, bem como de outros cursos e de outras universidades, da grande aplicabilidade e eficiência das redes neurais para resolver diversos tipos de problema. De maneira a aumentar esse estímulo, isso será feito utilizando as redes neurais para atacar um problema que se relaciona com uma das maiores fontes de entretenimento do país – o futebol -o que pode levar a um número maior de alunos se interessando pelo tema e, por fim, aprendendo sobre um conceito que pode futuramente ser utilizado para resolver desafios encontrados em suas carreiras.

2 Revisão Bibliográfica

2.1 Apostas e apostas esportivas

Uma rápida pesquisa em relação a jogos de azar e apostas e seus efeitos pessoais e morais não retornam resultados otimistas. Dentre resultados mostrando problemas éticos da oferta de jogos de azar para pessoas com tendências de criar vícios por esses jogos, além de condições psicológicas nocivas associadas ao vício em apostas, encontra-se o decreto-lei Nº 9.215, de 30 de abril de 1946, quando o então presidente Eurico Dutra proibiu a prática de jogos de azar em todo o país, alegando que são desalinhados com a tradição moral e religiosa do povo brasileiro.

Um jogo é definido como “de azar” quando é desenhado para garantir a vitória do estabelecimento que o promove no longo prazo. Uma máquina caça-níquel, por exemplo, possui um complexo algoritmo que gera resultados a cada jogo de modo a sempre garantir para casa uma quantidade de lucro pré-estabelecida em seu programa. Após uma grande quantidade de jogos feitos por uma grande quantidade de indivíduos, é uma certeza que a máquina irá convergir para gerar o lucro programado pelo estabelecimento. O programação da máquina é tão complexa que consegue garantir os lucros e incentivar os jogadores a continuar jogando por mais tempo, promovendo falsa sensação de confiança com algumas vitórias pontuais e gerando alguns casos de vitórias para algumas pessoas para que estas espalhem para outras que é possível ganhar, atraindo ainda mais jogadores para perder seu dinheiro (Schüll, 2014).

Em geral, isso ocorre quando o jogador está jogando diretamente contra o estabelecimento, em jogos como o blackjack, roleta e a já mencionada máquina caça-níquel. Algumas apostas, entretanto, não são realizadas contra o estabelecimento, mas contra outros jogadores. Esse é o caso de jogos como o poker e as apostas esportivas (Blackwood, 2006).

Ao contrário da percepção comum, ao se fazer em uma aposta esportiva, o “inimigo” não é a casa de apostas. Para entender esse fato, vamos definir o conceito de payout em uma aposta. Payout para um evento é o valor recebido de volta pelo apostador que acerta a ocorrência daquele evento. Em geral, ele é definido por uma razão do tipo 3:1. Um jogador aposta uma quantidade x e, caso erre, sairá daquele

jogo com um resultado de $-x$. Entretanto, caso ganhe, ele terá dado x para a casa, porém recebido $3x$ de volta, saindo com um resultado líquido de $2x$. Um payout geralmente é definido pelo inverso da probabilidade de ocorrência daquele evento. Assim, em um jogo de cara ou coroa, o payout justo para cada evento seria de 2:1.

Uma casa de apostas eficiente não é aquela que escolhe payouts que maximiza o seu lucro. O lucro de uma casa de apostas é garantida pela margem que ela pratica em seus payouts. No caso do jogo de cara ou coroa, o que acontece na realidade é, devida a aplicação da margem da casa, os payouts são adequados para, por exemplo 1,9:1. Nesse cenário, a casa de apostas garante o seu lucro caso a quantidade a ser paga para o grupo vencedor seja menor que a quantidade recebida nas apostas pelos indivíduos dos dois grupos. No cara ou coroa, como o payout é igual entre os eventos, se R\$1.000,00 estão apostados em cada evento, a casa recebe R\$2.000,00, e deve reparar apenas R\$1.900,00 ao indivíduo vencedor. (Boyle, 2006)

Caso os payouts sejam desequilibrados, como no caso de se tirar 1 ou 2 em um dado de 6 lados (payouts justos de 3:1 para ocorrência 1 ou 2 e 1,5:1 para ocorrência de 3, 4, 5 ou 6), a casa deve garantir que existem R\$2.000,00 apostados em números de 3 a 6, e R\$1.000,00 apostados nos números 1 ou 2. Ao ajustar o payout ligeiramente para baixo, como por exemplo 2,8:1 e 1,4:1, existem dois cenários: o resultado é o número 1 ou 2, onde a casa recebe R\$3.000,00 e devolve R\$2.800,00 para o grupo vencedor; ou o resultado é um número de 3 a 6, quando a casa recebe R\$3.000,00 e devolve igualmente R\$2.800,00 para o grupo vencedor. Em conclusão, a estratégia de uma casa de aposta é apenas de praticar uma margem de lucro no payout, denominada house edge, e garantir que os dois lados da conta estejam equilibrados.

Imagine uma aposta de análise complexa, como a aposta em qual candidato será o vencedor das eleições dos Estados Unidos da América. Em um determinado momento, os payouts estão de 1,5:1 para o candidato A e 2,5:1 para o candidato B. Perceba que isso significa que as probabilidades deveriam ser de 66,7% para o candidato A e 40% para o candidato B. As probabilidades somam 106,7%, indicando que a casa já está praticando margem nos payouts, com uma house edge de 6,7%. Imagine que um escândalo político aconteça com o candidato A e muito dinheiro é apostado no candidato B. O que a casa de apostas faz não é uma extrema análise do

cenário político e como o escândalo deve afetar exatamente as probabilidades de vitória de cada um. Com essa nova tendência de apostas em B, a casa propositalmente altera os payouts para, por exemplo, 2,1:1 para A e 1,7:1 para B. Isso incentiva pessoas a novamente apostarem em A, reequilibrando as contas. O algoritmo do site de apostas é responsável não pelo cálculo real das probabilidades, mas sim pela reação ao comportamento da população que está apostando no site, gerando combinações de payouts que tendem a equilibrar as contas. Agora, as probabilidades são de 47,6% para A e 58% para B, somando 106,4% e ainda gerando margem de lucro para a casa.

Dessa maneira, os payouts finais não são ativamente definidos pelas análises de probabilidade da casa, mas são uma consequência do comportamento dos próprios apostadores, e podem não refletir a probabilidade real dos eventos.

Em suma, as apostas esportivas não são por natureza entre o jogador e a casa, mas entre todos os jogadores que competem naquela aposta. A falsa ideologia de que a casa de apostas possui um enorme time de profissionais altamente qualificados com conhecimento de cada minúcia de cada equipe, e informações dos bastidores, para selecionar payouts que extraiam o máximo de dinheiro do público é incorreta: ter um time tão vasto e especializado seria muito caro, além de que não poderiam facilmente competir com todos os entusiastas do planeta.

A maneira mais barata de ter o melhor retorno financeiro é ter um time pequeno e ferramentas que apenas se preocupem em balancear as tendências de aposta do público e lucrar em cada partida, e então oferecer o máximo de partidas, ligas e esportes diferentes. Hoje em dia, sites oferecem apostas em mais de 20 esportes, incluindo corrida de cavalos, dardos, badminton, surf e esportes eletrônicos; além disso, esportes clássicos como o futebol tem uma cobertura imensa: é possível apostar em jogos da série D do brasileirão e no Campeonato Amazonense. Garantidamente, o site de apostas não faz a contratação de especialistas em campeonatos de cricket no interior da Malásia – mas apenas oferecem essa opção para que o máximo de jogadores batalhem entre si, enquanto a casa retira sua fatia. Com uma equipe reduzida, pequenos lucros a cada partida e o máximo de partidas cobertas possível, obtem-se o resultado máximo (Feustel & Howard, 2010).

Apenas para deixar claro: existe sim um time de estatísticos e analistas que geram payouts iniciais para o que eles esperam de um jogo, pois as apostas devem partir de algum payout inicial. Entretanto, não precisam estar muito exatos, pois o grande mecanismo que garante o lucro da casa acontece com as mudanças dinâmicas de payouts, que são guiados pelo comportamento dos jogadores. Isso significa que um jogo de apostas esportivas é um jogo que pode ser vencido – basta que o apostador estime as probabilidades de ocorrência dos eventos melhor do que um pouco acima da média do resto dos apostadores – acima da média o suficiente para pagar a fatia de lucro da casa.

Hoje em dia, apostadores, em sua maioria, ainda fazem suas escolhas baseadas em intuição, estatísticas, notícias e conhecimentos do jogo. A partir de seu conhecimento sobre as equipes e jogadores envolvidos, estatísticas que mostram diversas características sobre o time (média de gols, resultados dentro e fora de casa, últimos resultados, etc.), notícias (estado físico dos atletas, clima no treinamento, ajustes no time titular anunciados durante a semana), somados a uma intuição ao acompanhar o esporte por muito tempo, essas pessoas conseguem fazer apostas com valor. Eles possuem resultados positivos não por sorte, mas por habilidade: o desnível entre seu conhecimento e o conhecimento de um apostador comum faz com que eles leiam melhor as probabilidades do que a média dos apostadores, encontrando apostas de valor. Em outras palavras, as apostas esportivas são um jogo que premia inteligência e informação de um usuário, o que é muito similar com operações no mercado financeiro.

Além disso, diversos sites geram valor ao contratarem um time de analistas experientes sobre determinado esporte, que fazem suas análises e então disponibilizam para os usuários inscritos naquele site poderem fazer suas apostas, em troca de uma mensalidade – a dinâmica de apostas realmente é muito parecida com o mercado financeiro. Em sua essência, a operação no mercado financeiro é um jogo de apostas.

Existem diversos modelos de apostas que podem ser utilizados como comparativos para o modelo a ser desenvolvido. Três desses modelos são o Prediction Markets (análise dos mercados de aposta), Tipsters (apostas segundo especialistas que seguem seu conhecimento e intuição) e Betting odds (análises dos

payouts para um mesmo evento em diferentes fontes e casas de aposta). O curioso é que, mesmo nos melhores casos destes três métodos, a taxa de acerto dos resultados gira em torno de 52-54% e a taxa de retorno financeiro não costuma ser maior do que 8% (Spann & Skiera, 2009). Isso significa que muitos profissionais acertam próximo a metade das vezes, para ter um retorno de investimento positivo de 8%. Esses valores podem ser utilizados para estabelecer uma barra com o método aqui desenvolvido.

Como forma de melhorar suas taxas de previsão, modelos matemáticos foram criados para tentar realizar apostas calculadas a partir de dados estatísticos. Um desses modelos é o modelo estatístico que, a partir de dados de média de gols feitos e sofridos por aquele time dentro e fora de casa no passado, calcula-se quais as probababilidades de cada placar no presente. Esse método será utilizado como base para comparação com o método das Redes Neurais no presente trabalho, e será melhor explicado na seção 3.4.1.

Antes de aprendermos sobre as Redes Neurais Artificiais, temos um aviso importante: Para o caso de apostas esportivas em sites online, o apostador não precisa se preocupar se está violando o decreto-lei Nº 9.215, de 30 de abril de 1946. Isso porque os sites de aposta não estão localizados no Brasil e, portanto, a lei não se aplica.

2.2 Redes Neurais Artificiais

Redes Neurais Artificiais (RNAs) são modelos computacionais que recebem seu nome por existirem várias correlações entre a maneira que esses sistemas encontram soluções e a maneira biológica utilizada por cérebros animais. Estes sistemas são programados para aprender progressivamente a partir de exemplos prévios, gerando um aprendizado sobre aquele fenômeno, e então realizar decisões em situações futuras, com informações limitadas, a partir do aprendizado gerado por essas experiências anteriores. (Braga, Ludermir, & Carvalho, 2000).

Embora haja discussão quanto ao grau de similaridade entre o sistema computacional e o órgão biológico – com alguns autores defendendo extrema semelhança, a ponto de sugerir a aplicação das RNAs para modelar e estudar o comportamento animal –, a totalidade dos autores tendem a concordar que existe

similaridade suficiente para se traçarem paralelos, e especialmente para ajudar na compreensão da base teórica e o funcionamento de uma RNA a partir do exemplo biológico.

O cérebro biológico é produto de um longo processo evolutivo que é marcado pelo tradeoff entre custo e benefício. O grande número de 100 bilhões de neurônios e 1 quadrilhão de conexões entre estes no cérebro humano com certeza contribuem com a complexidade de pensamentos, emoções e decisões diárias. Este órgão representa 2% de nosso peso corporal, mas consome 20% da nossa energia (Drubach, 1999). Um cérebro maior seria mais poderoso, mas o aumento de consumo de energia poderia ser proibitivo. Além disso, o tamanho de crânio aumentaria e poderia impossibilitar o processo de parto, o mais traumático dentre os mamíferos: os humanos são a única espécie da classe que majoritariamente requerem apoio de outros indivíduos durante o parto, principalmente devido ao dilema obstétrico – a evolução guiou a espécie ao mesmo tempo para ter uma pelvis menor, facilitando o processo de locomoção, mas estreitando o canal de parto, e para aumento do cérebro e portanto do crânio e cabeça. (Posner, 2013)

Assim, o resultado das forças evolutivas é um cérebro extramamente eficiente: é estimado que o cérebro humano pode guardar $3,5 \times 10^{12}$ bytes e tem uma velocidade de processamento de $2,2 \times 10^9$ megaflops (um milhão de operações de ponto flutuante por segundo) (Fischetti, 2011). O K computer, o computador mais rápido do planeta em 2011, pode guardar 30×10^{12} (8,57x o cérebro humano) bytes e tem uma velocidade de processamento de $8,2 \times 10^9$ megaflops (3,7x o cérebro humano) (Kopytoff, 2011). Entretanto, existe um pequeno detalhe: o consumo do K computer é de 9,9 milhões de W, enquanto o consumo de um cérebro humano é estimado em 20W, o equivalente ao consumo de uma impressora caseira (Sinha, 2015). O cérebro precisa de quase 58 000 vezes menos energia para armazenar dados.

Até pouco tempo, o poder das soluções computacionais estava na força bruta. Os computadores não necessitam de pequeno consumo para sobrevivência e podem ser feitos o quão grande desejarmos – isso possibilita a procura de soluções por pura e simples tentativa de todos os casos ou a realização de bilhões de cálculos com um

supercomputador para chegar a uma conclusão. Essa estratégia funcionava bem para uma grande variedade de problemas, mas possui fortes limitações. Por mais que seja possível refinar o programa para guiar a solução com menos tempo, o método de solução é muitas vezes ineficiente ou incapaz.

O cérebro humano é capaz de aprender, se moldar e adaptar. E é esse dinamismo que deixa ele eficiente e capaz de generalizar, enquanto o computador apenas segue ordens, é engessado aos limites de seu programa e tenta soluções seguindo uma guia que tem que ser pré-definida pelo programador. O cérebro possui um mecanismo de funcionamento, mas tem grande adaptabilidade e consegue tirar conclusões sobre coisas nunca vistas antes, com informações restritas e sem precisar de ajuda.

O principal atrativo da solução de problemas através do machine learning, um ramo da inteligência artificial que tem as redes neurais artificiais como possível abordagem, é fazer uma mímica desse processo. A partir de um conjunto reduzido de exemplos, a rede aprende com os dados iniciais e extrapola essas informações para fornecer respostas coerentes para dados desconhecidos. Ela extrai indiretamente parâmetros de correlação implícitos entre os dados de entrada e saída que descrevem o fenômeno a ser analisado.

De forma bem objetiva, a rede neural é uma “caixa preta” que aprende a dinâmica sobre um processo e lhe fornece quais devem ser as saídas esperadas para um conjunto de entradas fornecidas, sem falar qual foi sua estratégia de solução e sem explicar como chegou aquela conclusão – o programador nunca conseguirá descobrir qual a estratégia utilizada pela rede, mas apenas confia na sua assertividade, dada uma alta taxa de acerto nas previsões.

As redes neurais artificiais, assim como redes neurais biológicas, são compostas por unidades básicas, chamadas de neurônios. O neurônio biológico é dividido em três partes: dendritos, corpo celular e axônio. Os dendritos recebem os sinais oriundos de outros neurônios e os conduz para o corpo celular, onde estes sinais são processados e um novo sinal é produzido, sendo transmitido pelo axônio até as extremidades do neurônio. As zonas ativas de contato entre o axônio de um neurônio e os dendritos de outro são chamadas de sinapse e são por meio destas que eles se conectam e se

comunicam. A Figura 1 ilustra um neurônio biológico com suas partes denominadas; os escritos entre parênteses abaixo dos nomes das estruturas representam os equivalentes na estrutura de uma rede neural.

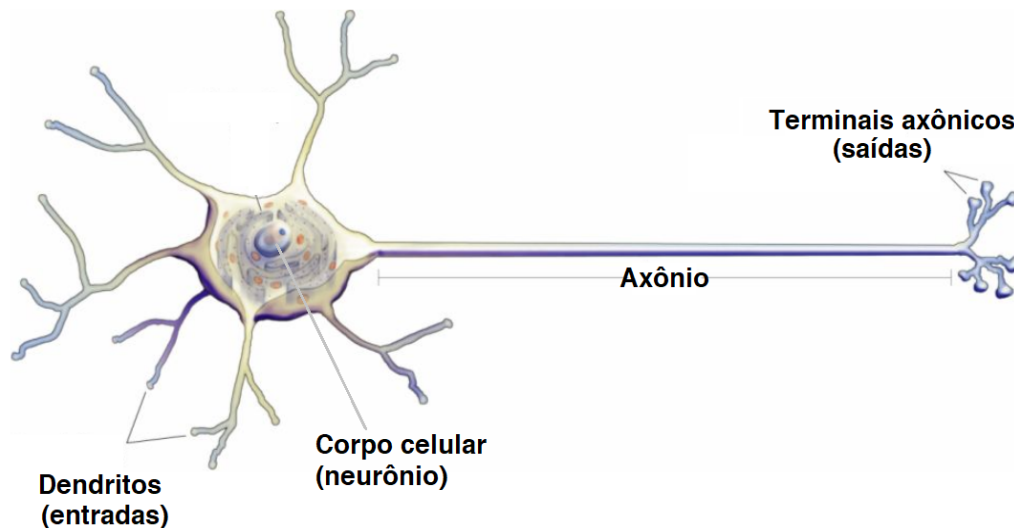


Figura 1: Componentes básicos de um neurônio biológico

Fonte:

<https://www.memorangapp.com/flashcards/104120/The+Nervous+System+I%3A+Neuron+Structure/> - Acessado em 04/06/2018

A transmissão de sinais por um neurônio ocorre somente quando a soma dos impulsos elétricos que recebe pelas conexões com outros neurônios, intermediadas pelos neurotransmissores, ultrapassa seu limiar de excitação, que por sua vez produz outro impulso propagado pelo axônio para neurônios seguintes; dessa maneira, o neurônio tem o comportamento denominado all-or-none: caso a excitação seja suficiente, ele se ativa com força completa; caso contrário, não há nenhuma ativação (Levitan & Kaczmarek, 1991). Esse princípio é importante para evitar que pequenos estímulos desencadeiem reações desnecessárias. Por exemplo, sistemas nervosos no coração não podem ativar e enviar um impulso para causar uma batida do coração em qualquer ocasião, pois isso tornaria o sistema extremamente sensível a irregular, em resposta à estímulos externos (Host, 2000). As sinapses, portanto, atuam como válvulas que controlam a transmissão de sinais ou o fluxo de informação pela rede de neurônios.

O objetivo de uma rede neural artificial é simples: a partir de um conjunto de entradas, prever qual deve ser o valor das saídas. Existem tipos de redes neurais com objetivos diferentes. Uma rede neural do tipo fitting é utilizada para gerar uma função que relaciona entradas e saídas de um problema com números contínuos. Porém, nesse trabalho, iremos utilizar uma rede neural de classificação. Esse tipo de rede tem o objetivo de, a partir de um conjunto de entradas, decidir uma categoria para posicionar aquela amostra, dentre categorias pré-definidas (Bishop, Christopher, 1995). Assim, é fácil entender como o problema de jogos de futebol tratado nesse trabalho pode ser resolvido por esse tipo de rede: a partir de determinadas entradas, a rede deve decidir se aquele jogo terá vitória do mandante, empate ou vitória do visitante, e então posicionar aquele jogo em uma dessas categorias.

Para entender como a rede atinge esse objetivo, iremos passar por algumas etapas: entender como funciona cada neurônio, como a rede organizada e, por fim, como é seu processo de aprendizado.

O primeiro modelo de neurônio artificial foi proposto por McCulloch e Pitts (1943) baseado em um algoritmo denominado de Threshold Logic. Esse neurônio “j” recebe n entradas, denominadas x_1, x_2, \dots, x_n , como se fossem os dentritos biológicos. Cada dentrito representa uma conexão com outro neurônio, e cada conexão possui um peso definido e denominado $w_{1j}, w_{2j}, \dots, w_{nj}$. O primeiro índice se refere ao número do neurônio que envia o sinal, enquanto o segundo índice se refere ao número do neurônio que recebe o sinal. Cada entrada recebida é então ponderada por esses pesos, tornando-se sinais de valor $x_i w_i$. Dessa forma, os pesos são uma medida do quão fortemente aqueles neurônios devem se conectar.

Em seguida, o neurônio possui uma função de transferência que, na forma mais básica, proposta por McCulloch e Pitts (1943), trata-se da soma entre os produtos entre sinais de entrada e pesos das conexões $\sum x_i w_{ij}$. Então, o resultado passa por uma função de transferência, que novamente, na forma mais básica, trata-se de uma função do tipo degrau. Caso o valor resultante da função de transferência seja maior que um determinado valor denominado threshold (T), o neurônio dispara por completo (saídas com valor 1); caso contrário, nenhum sinal é emitido (saídas com valor 0) – semelhante ao processo all-or-none de um neurônio biológico.

Por fim, o neurônio produz uma ou mais saídas. No modelo de McCulloch e Pitts, o neurônio produzia uma saída, que poderia ser comparada a um axônio. Entretanto, nada impede que um neurônio emita várias saídas, que podem se conectar e servir de entrada para outros neurônios – nesse caso, as saídas seriam melhor representadas pelos terminais axônicos.

Matematicamente, a saída de um neurônio de McCulloch e Pitts é dada pela Equação 1. Esses processos ocorrem “dentro” do neurônio, no que seria análogo ao corpo celular. O valor de threshold (T) é normalmente substituído pelo termo bias, sendo que $bias = -T$.

$$\begin{cases} f(x) = 1, & \text{se } \sum_{i=1}^n x_i w_{ij} + b > 0 \\ f(x) = 0, & \text{se } \sum_{i=1}^n x_i w_{ij} + b < 0 \end{cases}$$

Equação 1: saída de um neurônio j do tipo McCulloch e Pitts com n entradas, w_{ij} como peso entre cada conexão do neurônio j com i entradas x_i e valor de bias b

A Figura 2 apresenta um modelo da estrutura de um neurônio discutida nos parágrafos anteriores.

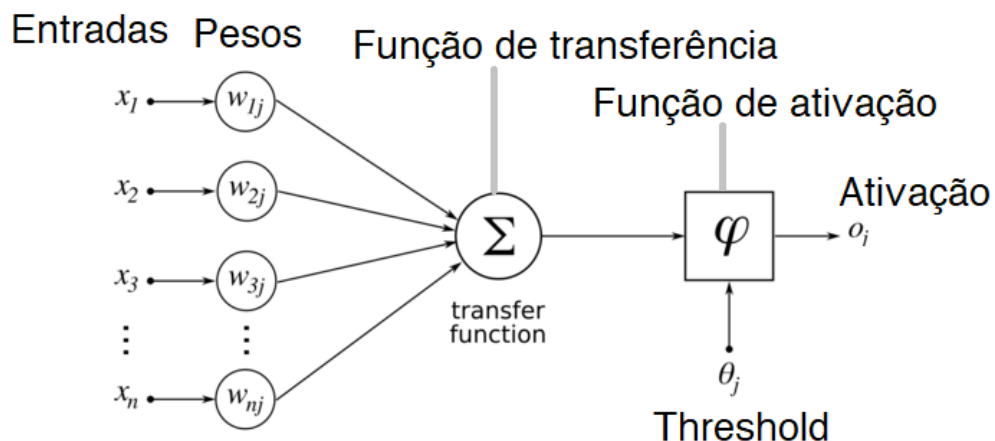


Figura 2: esquema representativo de um neurônio utilizado em redes neurais artificiais

A escolha da função de ativação a ser utilizada pelos neurônios tem um importante impacto no funcionamento da rede neural. A Figura 3 apresenta algumas funções de ativação utilizadas.

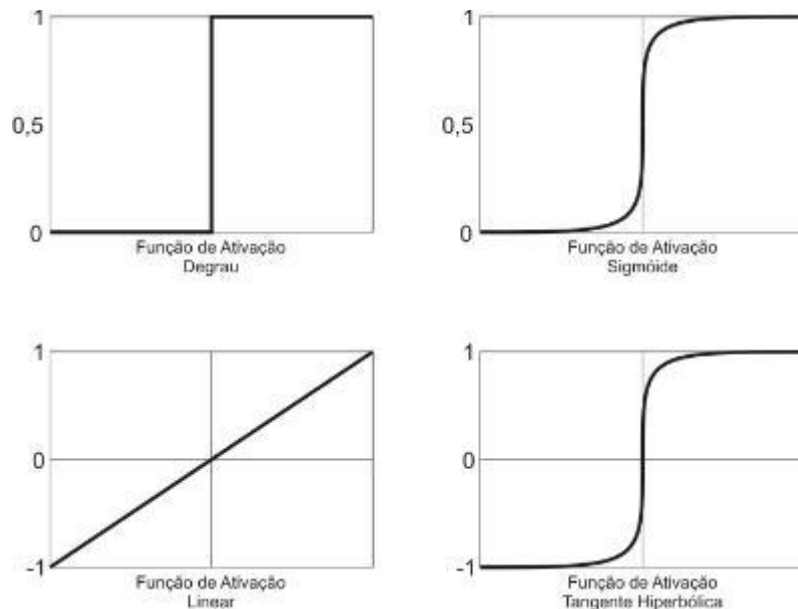


Figura 3: Tipos de função de ativação

Fonte: <http://www.ebah.pt/content/ABAAe14sAH/mestrado-producao-rn?part=3> -
acessado em 04/06/2018

A função do tipo degrau semelhante a um neurônio biológico: caso a soma das entradas no neurônio atinja um valor mínimo, o neurônio envia uma resposta máxima (1). Caso contrário, o neurônio gera uma saída nula que servirá de entrada para os próximos neurônios a que estão conectados. Este tipo de função possui uma função importante: caso a rede neural tenha a função de entregar saídas em forma de verdadeiro ou falso, ela pode utilizar neurônios com função tipo degrau. Eles podem ter um papel muito importante nas redes neurais de classificação, uma vez que uma saída do tipo verdadeiro ou falso para cada possível categoria pode ser utilizado para posicionar uma amostra em uma determinada categoria. No nosso caso, poderíamos utilizar 3 neurônios para as saídas: quando o primeiro é ativado e retorna 1 e os dos outros não são ativados e retornam 0, esse jogo pode ser posicionado na categoria vitória do mandante. Outros tipos de função possuem suas vantagens e desvantagens específicas, que serão explicadas adiante.

Em segundo lugar, essas unidades funcionais são conectadas entre si para formar uma rede mais complexa, como representado na Figura 4.

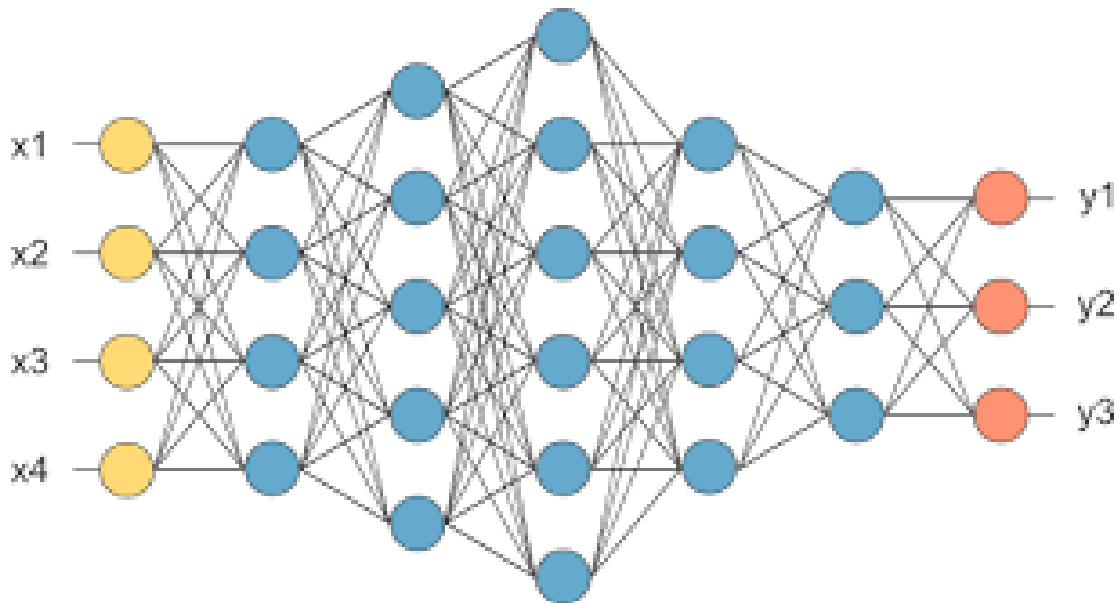


Figura 4: Esquema da estrutura de uma rede neural artificial com diversas camadas

(Retirado de <https://www.neuraldesigner.com/>, acessado em 05/06/2018)

As camadas se diferem em camadas de entrada, camadas escondidas/intermediárias e camadas de saída. A camada de entrada, representada pelos neurônios em amarelo, possuem os valores das entradas fornecidas pelo problema. Então, existe uma sequência de camadas escondidas, representadas pelos neurônios em azul. O número de camadas escondidas é um parâmetro de escolha de quem opera a rede. Entretanto, para problemas de classificação, uma camada escondida é suficiente para resolver o problema, dado uma quantidade ideal de neurônios nessa camada (Bishop, 1996). Os neurônios de saída estão coloridos em laranja e devem ter como saída o valor previsto a partir das entradas.

Nessa rede, cada neurônio recebe um número de entradas igual ao número de neurônios da camada anterior. Essas entradas tem o valor do neurônio que envia essa entrada, multiplicado por um peso da conexão entre esses dois neurônios. Então, o neurônio obtém a soma dos produtos entre todas as entradas e respectivos pesos, soma o valor de seu bias interno, e chega a um resultado. Esse resultado é passado pela função de ativação, e gera uma saída. Por fim, o valor dessa saída será enviado para todos os neurônios do próximo nível.

O dimensionamento da rede – quantidade de camadas intermediárias e neurônios por camada - é realizado por meio de refinamentos sucessivos a partir de uma topologia inicial definida em função de uma inspeção prévia nos dados e na complexidade do problema. Este refino é empírico e iterativo. Inicia-se com um número de neurônios baixo, e iterativamente obtém-se resultados com um número crescente, até que atinja-se um erro mínimo, que então começa a crescer com o aumento do número de neurônios. (Braga, Ludermir, & Carvalho, 2000).

Imagine uma rede neural cuja estrutura é semelhante a estrutura da rede neural representada na Figura 4. Ao criar essa rede neural, definimos os bias de todos os neurônios, bem como os pesos de cada conexão entre eles. Imagine que todos esses valores são gerados de forma aleatória. Assim, ao inserir os dados de entrada na rede, o resultado das saídas será computado após todas as operações camada a camada, neurônio a neurônio. O resultado será um conjunto de saídas com valores totalmente aleatórios.

Imagine que então seja possível ajustar esses pesos e bias, ligeiramente, de maneira guiada, de forma que agora as saídas ainda estão bem incorretas, mas estão ligeiramente mais corretas. Se, em cada pequeno ajuste desse, as saídas estiverem ligeiramente mais perto do resultado, prosseguiremos com os ajustes na mesma direção. Caso os resultados da saída piorem, tentaremos outro caminho. Após muitas e muitas iterações e checagens de melhoria do resultado, a rede neural ajustou seus pesos e bias tantas vezes que agora, ao receber um entrada, suas saídas são bem próximas do que acontece na realidade, se não exatas. Essa é a mágica do processo de aprendizado.

O processo de aprendizado de uma rede neural também tem fortes paralelos com a mente humana. Os 10 bilhões de neurônios de quadrilhões de sinapses estão continuamente se ativando e se reconectando, principalmente nos primeiros anos de vida, quando há maior neuroplasticidade (Doidge, 2007).

Imagine um ser humano há 100 000 anos atrás, quando a sociedade ainda era composta por caçadores-coletores. Um indivíduo daquele tempo pode ter se deparado com uma flor de certo tamanho, coloração, cheiro e forma. Ao ingerir essa planta, ele teve uma reação adversa no seu organismo. Nesse momento, os neurônios em seu

cérebro relacionados com a visão (um centro para reconhecimento de forma, um para reconhecimento de cor, etc.), relacionados ao olfato e os relacionados a reações fisiológicas negativas, irão reorganizar suas conexões. Da próxima vez que uma flor que possua as mesmas características for encontrada, os sinais gerados pelos centros que processam sentidos, através de reações extremamente complexas, irão estimular partes do cérebro relacionadas com ansiedade, nojo, descontentamento, falta de fome, entre outros. Nesse momento, o indivíduo aprendeu que aquela planta não deve ser ingerida. Esse processo simplificado é na verdade infinitamente mais complexo do que o descrito, que possui função apenas ilustrativa.

Uma das maiores vantagens do aprendizado humano é a capacidade de chegar a conclusão com um número restrito de entradas e incerteza em volta delas. Imagine agora que o mesmo homem pré-histórico do exemplo acima encontra um pequeno animal de coloração vermelha e amarela. Ao ter aprendido com exemplos e informações passadas que animais com tal coloração tendem a ser tóxicos/venenosos, ele evita esse novo animal, mesmo que nunca houvesse o visto antes. Isso também ocorre nas redes neurais, e é entendido como a habilidade de generalização da rede.

Após utilizar vários dados como treinamento e para ajustar seus pesos e bias, ela é capaz de receber valores de entrada nunca antes vistos e obter algum valor para a saída que, caso a generalização seja boa, o fenômeno seja bem comportado e as entradas realmente estejam correlacionadas de maneira significativa com as saídas, será próxima do valor real desejado, mesmo que a relação entre as entradas e saídas não seja facilmente descrita matematicamente (Haykin, 1994).

O grande poder das redes neurais em frente a outras soluções matemáticas está justamente nessa habilidade. Para atingir boa capacidade de generalização, a rede neural deve ter disponível um conjunto de dados que apresentam entradas reais e saídas reais que já aconteceram, que serão utilizadas para seu treinamento. Caso o treinamento seja bem sucedido, ela será capaz de prever saídas para entradas futuras a serem analisadas. No caso discutido nesse trabalho, isso significa que a rede irá precisar dos dados de entrada e resultados reais ocorridos em várias partidas antigas do campeonato, para que tente aprender a dinâmica escondida entre essas

entradas e saídas e, com sorte, ter uma alta taxa de acerto na sua previsão nos dados de teste.

O processo de aprendizado de uma rede neural é denominado treinamento. Seu objetivo é, iterativamente, ajustar os pesos e bias de modo que as saídas geradas fiquem cada vez mais perto do valor real. Os ajustes são feitos após cada amostra de entrada e possuem um direcionamento determinado pelo método de treinamento. Diversos métodos para treinamento foram desenvolvidos, mas todos possuem um princípio básico: selecionar uma amostra, inserir as entradas dessa amostra nos seus neurônios de entrada, obter saídas a partir das interações entre seus neurônios de acordo com os valores de pesos e bias e, por fim, comparar as saídas obtidas com os valores reais. A partir da diferença entre as saídas reais e saídas calculadas, um sinal de erro é gerado, que então é utilizado para reajustar os pesos e bias de todos os seus neurônios de uma forma guiada pelo princípio desse método de treinamento. O erro é tido como uma função objetivo que deve ser minimizada. Quando o erro é minimizado e confirma-se que ele permanece baixo para várias amostras de entrada, atinge-se a generalização.

Um dos principais algoritmos de aprendizado supervisionado utilizado é o de Backpropagation. Este algoritmo se baseia no cálculo da contribuição individual de cada neurônio no erro da rede por meio da minimização da função de custo (função objetivo) do estado momentâneo da rede através do ajuste dos pesos dos neurônios da rede.

O treinamento ocorre em duas fases, e cada fase percorre a rede em um sentido. Estas duas fases são chamadas de fase forward e fase backward. A fase forward é simplesmente utilizada para determinar a função de custo do estado atual da rede e consiste na simples definição da saída da rede para um conjunto de dados de entrada de uma amostra. Já a fase backward utiliza as informações adquiridas na fase anterior para ajustar os pesos das conexões dos neurônios e os bias desses neurônios, com base no gradiente da função de custo adquirida. A função gradiente é aquela que dá a direção que causa o maior incremento possível no valor de uma grandeza com a variação das variáveis de entrada, teoricamente causando uma convergência mais rápida.

O nome Backpropagation surge, portanto, do algoritmo se basear na retropropagação do erro para realizar os ajustes de pesos das camadas intermediárias. Essas fases se repetem até que o erro da rede seja menor que um valor de tolerância pré-definido. Quando isso acontece, diz-se que o aprendizado foi concluído.

Nesse momento, é mais fácil entender a utilidade de outras funções de ativação como a função sigmóide. A função sigmóide é apresentada na Equação 2.

$$\sigma(x) = \frac{1}{1 + e^{-x}}$$

Equação 2: função sigmóide

Como os pesos e bias são ajustados de acordo com a função gradiente, que por sua vez é obtida por meio da derivação de funções, a função sigmóide se torna uma ótima candidata, uma vez que ela é facilmente derivável. Sua derivada é dada pela Equação 3.

$$\frac{d}{dx}\sigma(x) = \sigma(x) \times (1 - \sigma(x))$$

Equação 3: derivada da função sigmóide

Além disso, o processo de aprendizado deve, pouco a pouco, convergir para um erro mínimo. Esse processo é mais bem comportado quando as funções de ativação possuem variações suaves ao longo de suas curvas: uma pequena variação nas entradas gera uma pequena variação nas saídas, que podem ser avaliadas pouco a pouco para chegar a uma conclusão. Imagine uma rede cujas camadas intermediárias são compostas por funções degrau: uma pequena variação nas entradas de um neurônio causariam uma abrupta mudança na sua saída de 0 para 1 ou vice versa, gerando um baque no sinal de saída. Além disso, a função deve ser monotonamente crescente, para que um aumento no valor da entrada sempre cause um aumento no valor da saída. Se o comportamento se invertesse, o método de treinamento seria confuso. A função sigmóide possui essas características, bem como a função de ativação tangente hiperbólica, ambas representadas na Figura 3.

Tendo concluído o aprendizado, espera-se que a rede tenha uma boa capacidade de generalização, ou seja, uma boa interpolação não-linear sobre os dados de entrada. No entanto, uma rede pode ser treinada de forma insuficiente, ou de forma exagerada. Quando é treinada de forma insuficiente, ela esgota as amostras de treinamento e ainda não consegue obter uma boa generalização do fenômeno. Quando ela é treinada de forma exagerada, ela se torna uma especialista para aqueles dados de entrada, e é capaz de acertar com perfeição todas as saídas para eles. Entretanto, quando novas saídas são fornecidas com ligeiras diferenças, a rede possui um resultado ruim, porque tenta forçar aquele dado a se encaixar numa estratégia que é ótima apenas para os dados de treinamento.

Assim, têm-se o problema de saber qual o tamanho do conjunto de dados suficiente para se treinar uma rede para que ela forneça uma boa capacidade de generalização. A Figura 5 mostra um exemplo comparativo de uma rede sub-treinada (não treinada o suficiente para generalizar), uma bem treinada e outra treinada em excesso.

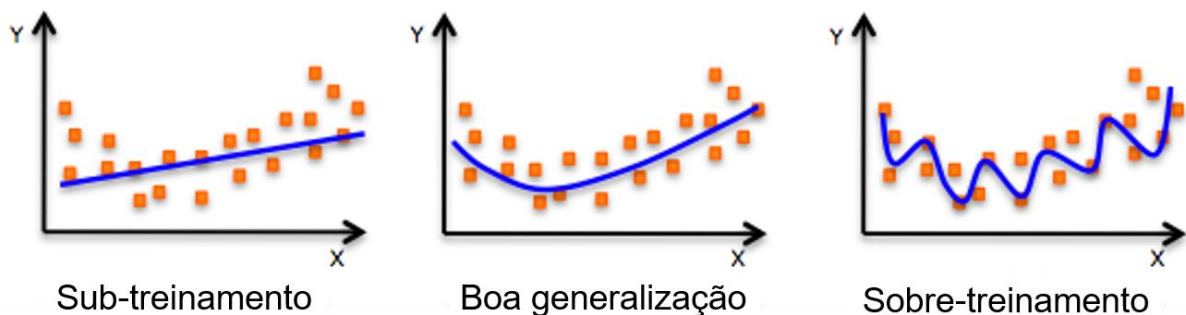


Figura 5: Comparativo entre redes sub-treinadas, bem treinadas e treinadas em excesso.

(Fonte: <https://cs.stackexchange.com/questions/75327/why-is-deep-learning-hyped-despite-bad-vc-dimension> - Acessado em: 04/05/2018)

Para evitar o fenômeno de sobre-treinamento, é utilizada uma técnica de treinamento chamada Método de Treinamento com Parada Antecipada. Este método utiliza, além do erro quadrado médio de cada época de treinamento da rede, uma validação cruzada, em que os dados de treinamento são divididos em dois subconjuntos. O primeiro subconjunto é o de treinamento, compreendendo a grande

maioria dos dados, que é usado para treinar a rede de maneira usual, a fim de ajustar os pesos sinápticos e bias da rede. O segundo subconjunto é o de validação. O erro de validação é monitorado durante o processo de treinamento. O erro de validação normalmente decresce durante a fase inicial de treinamento, assim como o erro de treinamento. Entretanto, quando a rede começa a ser excessivamente ajustada aos dados de treinamento, o erro de validação tipicamente começa a aumentar. Quando o erro de validação aumenta para um número específico de iterações, o treinamento é interrompido e os pesos e bias correspondentes ao menor valor do erro de validação são restaurados.

Por fim, existe um tipo de camada de saída de redes neurais que se aplicam bem ao problema de classificação. Como alternativa a utilização apenas de neurônios de saída com função do tipo de degrau, que apenas posicionam uma amostra em uma dada categoria, é possível utilizar uma camada de saída com funções do tipo softmax, ou função exponencial normalizada. Por passar por processo de normalização, a soma entre os elementos do vetor de saída resulta em 1, permitindo que seus elementos estejam ligados a probabilidade de ocorrência. Agora, as saídas da rede neural indicam qual é a probabilidade que ela considera que uma determinada amostra pertença a uma determinada categoria. A classificação é dada a partir da saída de maior probabilidade, mas a informação sobre a probabilidade em si pode enriquecer a análise. Ela pode ser usada, por exemplo, para alimentar uma estratégia de aposta onde a rede neural só aposta em um determinado evento quando considera que tem uma chance mínima de ocorrer, aumentando sua certeza sobre o resultado. Em problemas de classificação menos aleatórios, a probabilidade da categoria correta seria calculada acima de 99%. Entretanto, como é esperado que o fenômeno de partidas de futebol seja altamente aleatório, essas probabilidades podem ser informações muito importantes.

2.3 Inteligência artificial em apostas esportivas

Atualmente, algumas empresas já estudam soluções computacionais para problemas como previsão de apostas esportivas, ganhadores de óscar, resultados de eleições, tendências de preços de ações, entre outros.

A startup Stratagem desenha um modelo similar ao desenvolvido nesse trabalho, porém não divulga exatamente os parâmetros e dados utilizados. A proposta é que, a partir da extração de vários dados estatísticos de mais de 100 mil partidas, ela consiga construir uma rede neural artificial com resultados consistentemente positivos. Além disso, a empresa empodera a taxa de acerto de suas previsões com um time de especialistas que supervisiona e guia o aprendizado da rede, levando a ainda melhores resultados. Assim, a abordagem da empresa é uma previsão com base em um modelo colaborativo homem-máquina. Em Julho de 2017, a taxa de acerto do modelo girava em torno de 50%, o que era mais do que o necessário para gerar lucros, segundo o criador Andreas Koukorinis (Vincent, 2017).

Além disso, uma startup chamada Unanimous A.I. A tem uma abordagem diferente: ela se baseia na inteligência de colméias para desenhar um sistema que é capaz de criar inteligência de colméias nos seres humanos, aumentar o QI de um grupo de indivíduos e prever com uma acuracidade melhor que os especialistas o resultado de óscars e resultados de futebol americano. Em um episódio, a equipe preveu a ordem de chegada dos 4 primeiros cavalos na corrida de Kentucky Derby, gerando ganhos de \$542 praa cada \$1 apostado. (Ryan, 2018).

Nessa abordagem, uma comunidade de indivíduos deve decidir sobre um evento que possui alguns resultados possíveis. Cada indivíduo possui um ímã com que puxa uma bola central em direção a resposta que acredita – quanto mais certo estiver de sua escolha, pode aproximar o ímã da bola central para atraí-la com mais força. Assim, em tempo real, os jogadores conseguem ver como os outros jogadores estão decidindo e, tomando uma decisão em conjunto, conseguem taxas de acerto muito superior a indivíduos ou a grupos de indivíduos escolhendo separadamente. A taxa de acerto na previsão de resultados na NFL foi de 68% contra 47% de um grupo de indivíduos (Rosenberg & Baltaxe, 2016), e ela previu o prêmio do óscar para 15 de 16 (94%) categorias analisadas, enquanto especialistas de outras revistas acertaram em 88% (Variety), 81% (LA Times) e 69% (Rolling Stone) (Unanimous Ai, 2018).

3 Metodologia

3.1 Construção das entradas de treinamento

Para treinamento da rede neural, devemos obter uma lista de pontos de treinamento que consiste em um vetor $a \times b$, sendo a o número de pontos de treinamento, e b o número de variáveis de interesse, dado por $b = e + s$, sendo e o número de entradas e s o número de saídas do modelo.

Quanto as saídas da rede neural, como trata-se de uma RNA de classificação dentre três possíveis resultados, teremos sempre 3 variáveis de saída. Para os dados de treinamento, temos as saídas y_0 , y_1 e y_2 , tais que o vetor $[y_0 \ y_1 \ y_2]$ representam um resultado, sendo que $[1 \ 0 \ 0]$ representa vitória do mandante, $[0 \ 1 \ 0]$ representa empate e $[0 \ 0 \ 1]$ representa vitória do visitante.

Quanto as entradas, elas serão tantas quantas desejamos utilizar no modelo. Para o dado estudo, serão levantados 7 grupos de variáveis de entrada, sendo elas:

- **Variáveis de entrada estatísticas**, referentes as entradas dos modelos estatísticos, que contem as médias de gols feitos e sofridos por cada time jogando em casa/fora de casa. Essas variáveis podem ser importantes por poderem indicar uma tendência para o placar da partida, que está diretamente relacionado com o resultado do jogo
- **Variáveis de entrada de forma atual**, que contem dados de aproveitamento dos times no campeonato, nas últimas 5 rodadas e jogando dentro/fora de casa. Essas variáveis podem ser importantes por trazerem informações do desempenho da equipe ao longo do campeonato atual, o que pode influenciar as chances de vitória de cada time
- **Variáveis de entrada de ofensividade**, que contem dados de finalizações totais e acuracidade das finalizações de cada time, em média, por jogo. Essas variáveis podem ser importantes, pois o resultado é definido pela quantidade de gols feitos por cada time, que está relacionada com a quantidade de oportunidades de se fazerem um gol, representadas aqui pelas finalizações totais e número de chutes em gol.
- **Variáveis de entrada de características do plantel**, com informações de idade, peso e altura média dos jogadores dos times envolvidos. Essas variáveis

podem ser importantes pois a forma e vigor físico dos jogadores podem influenciar sua condição de jogo, agilidade, estamina, dentre outros fatores que podem afetar o resultado

- **Variáveis de entrada de disciplina**, com informações sobre o número de faltas cometidas e cartões recebidos pelos jogadores de cada equipe, em média, por jogo. Essas variáveis podem ser importantes pois o comportamento de disciplina pode afetar o resultado de um jogo, uma vez que times mais agressivos podem ter mais jogadores pendurados que evitarão fazer faltas, bem como o recebimento de um cartão vermelho em uma partida diminui a tendência de um time fazer gol, bem como aumenta as chances de sofrer mais gols
- **Variáveis de entrada de aspectos financeiros**, que contem dados da receita anual e valor de mercado dos jogadores da equipe. Esses dados podem ser importantes pois o valor de mercado dos jogadores tende a correlacionar com a qualidade dos jogadores do time, partes chave que influenciam o resultado de uma partida. A receita anual está correlacionada com a capacidade do time de aumentar ou reduzir o valor de mercado dos jogadores da equipe, bem como ter uma saúde financeira para pagar os jogadores em dia e incentivá-los a terem bom desempenho.

Uma lista de todas as variáveis levantadas é apresentada e descrita na Tabela 1.

Tabela 1: grupos de variáveis de entrada, variáveis de entrada e descrição das variáveis de entrada a serem utilizadas

| | | |
|-------------------------------------|--------|--|
| Variáveis de entrada estatísticas | AVGGPH | Média de gols feitos pelo time mandante jogando em casa* |
| | AVGGPA | Média de gols feitos pelo time visitante jogando fora de casa* |
| | AVGGCH | Média de gols sofridos pelo time mandante jogando em casa* |
| | AVGGCA | Média de gols sofridos pelo time visitante jogando fora de casa* |
| Variáveis de entrada de forma atual | APH | Aproveitamento do time mandante no campeonato |
| | APA | Aproveitamento do time visitante no campeonato |
| | AP5H | Aproveitamento do time mandante nas últimas 5 rodadas |
| | AP5A | Aproveitamento do time visitante nas últimas 5 rodadas |
| | APHH | Aproveitamento do time mandante jogando em casa |
| | APAA | Aproveitamento do time visitante jogando em casa |

| | | |
|--|---------|--|
| Variáveis de entrada de ofensividade | AVGSOGH | Média de chutes a gol por jogo do time mandante* |
| | AVGSOGA | Média de chutes a gol por jogo do time visitante* |
| | AVGSH | Média de finalizações por jogo do time mandante* |
| | AVGSA | Média de finalizações por jogo do time visitante* |
| Variáveis de entrada de características do plantel | AVGHEIH | Altura média dos jogadores do time mandante |
| | AVGHEIA | Altura média dos jogadores do time visitante |
| | AVGAGEH | Idade média dos jogadores do time mandante |
| | AVGAGEA | Idade média dos jogadores do time visitante |
| | AVGWEIH | Peso médio dos jogadores do time mandante |
| | AVGWEIA | Peso médio dos jogadores do time visitante |
| Variáveis de entrada de disciplina | AVGFAUH | Média de faltas cometidas por jogo do time mandante* |
| | AVGFAUA | Média de faltas cometidas por jogo do time visitante* |
| | AVGYCH | Média de cartões amarelos por jogo do time mandante* |
| | AVGYCA | Média de cartões amarelos por jogo do time visitante* |
| | AVGRCH | Média de cartões vermelhos por jogo do time mandante* |
| | AVGRCA | Média de cartões vermelhos por jogo do time visitante* |
| Variáveis de entrada de estilo de jogo | PASSUCH | Porcentagem de acerto de passes do time mandante* |
| | PASSUCA | Porcentagem de acerto de passes do time visitante* |
| | POSSH | Porcentagem de posse de bola do time mandante* |
| | POSSA | Porcentagem de posse de bola do time visitante* |
| Variáveis de entrada de aspectos financeiros | RECTH | Receita anual do time mandante |
| | RECTA | Receita anual do time visitante |
| | VALOH | Valor de mercado dos jogadores do time mandante |
| | VALOA | Valor de mercado dos jogadores do time visitante |

Todas as variáveis marcadas com um * referem-se aos números obtidos pela equipe na última temporada disputada. As demais variáveis são referentes aos dados da equipe no início da temporada a ser estudada, ou durante ela.

Um modelo será construído que receberá os times e resultados dos 380 jogos do campeonato dos anos de 2006 a 2017. Com os resultados finais de cada ano, é possível montar uma tabela constando todas as variáveis de entrada estatísticas a serem utilizadas para cada time no ano $y+1$. Além disso, a cada rodada serão calculados as variáveis de entrada de forma atual para serem utilizadas nos jogos seguintes.

As variáveis de entrada de ofensividade, características do plantel, disciplina, estilo de jogo e aspectos financeiros serão retiradas de fontes externas sobre estatísticas dos campeonatos, sendo elas:

<http://www.futebol365.pt/competicao/182/estatisticas/?edition=2016>: dados sobre características físicas do plantel

<https://www.whoscored.com/Regions/31/Tournaments/95/Seasons/3753/Stages/7479/TeamStatistics/Brazil-Brasileir%C3%A3o-2013>: dados sobre ofensividade, disciplina e estilo de jogo

https://www.transfermarkt.co.uk/campeonato-brasileiro-serie-a/startseite/wettbewerb/BRA1/plus/?saison_id=2013: valores dos jogadores das equipes

https://pt.m.wikipedia.org/wiki/Lista_de_faturamento_dos_clubes_de_futebol_brasileiro: faturamento das equipes

<https://epoca.globo.com/esporte/epoca-esporte-clube/noticia/2017/05/financas-da-chapecoense-apesar-da-tragedia-equipe-catarinense-continua-solida.html>: faturamento (Chapecoense)

<http://www.avai.com.br/novo/patrimonio/balanco-patrimonial/>: faturamento (Avaí)

<http://www.correiodopovo.com.br/blogs/hiltormombach/2016/07/46325/faturamento-dos-clubes-de-futebol-cresce-15-em-2015-aponta-itu-bba/>

<https://globoesporte.globo.com/futebol/times/america-mg/noticia/america-mg-divulga-lucro-de-quase-r-10-milhoes-em-2016.ghtml>: faturamento (América-MG)

<http://www.santacruzpe.com.br/balanco-financeiro-de-2016/>: faturamento (Santa Cruz)

<http://globoesporte.globo.com/sp/campinas-e-regiao/futebol/times/ponte-preta/noticia/2015/11/orcamento-do-futebol-da-ponte-para-2016-ganha-injecao-de-r-8-milhoes.html>: Faturamento (Ponte Preta)

<http://www.criciumaec.com.br/upload/financeiro/2013-demonstrativo-financeiro-2013-20170622094742.pdf>: Faturamento (Criciúma)

Todos acessados em Maio/2018

3.2 Treinamento das redes neurais artificiais

Todas as redes neurais foram criadas utilizando a ferramenta Neural Network do Matlab® na versão R2017a. A biblioteca oferece diferentes soluções baseadas em redes neurais, sendo que cada solução é mais adequada a um tipo de problema estudado.

Para o estudo em questão, a ferramenta utilizada é chamada Pattern Recognition app, que realiza a construção de uma rede neural de classificação. Redes neurais de classificação recebem um número desejado de entradas sobre um evento, e então o posiciona em um de “n” grupos de acordo com o que acredita ter mais afinidade.

Após treinada, pontos de teste podem ser inseridos na rede neural para, de acordo com sua capacidade de generalização do problema, obter em qual grupo a rede neural acredita que aquele evento mais se encaixa. Nesse momento, uma rede neural de classificação dará como saída um vetor também na forma $1 \times n$, porém agora cada um dos elementos do vetor será o grau de afinidade detectado pela rede neural com aquele determinado grupo, que também pode ser entendido como a probabilidade que a rede neural acredita que aquele evento se encaixe em determinado grupo. Assim, a soma entre os 3 elementos do vetor deverá resultar em 1. Por exemplo, um vetor de saída $[0,64 \ 0,20 \ 0,16]$ sugere que a rede neural acredita haver 64% de chance de vitória do time mandante, 20% de chance de um empate e 16% de chance da vitória do time visitante.

O software utiliza uma rede neural com uma camada escondida e uma camada de outputs com um número de neurônios igual ao número de grupos a classificar. Uma camada escondida é suficiente para se obter bons resultados em problemas de classificação, como discutido na revisão.

Os neurônios na camada escondida têm função sigmoideal, que possui uma derivada de fácil obtenção e permite um treinamento mais rápido (Haykin, 1994), e a influência do número de neurônios será determinada a partir da rodagem da rede com diferentes tamanhos de rede e observando as diferenças nos resultados obtidos.

Já os neurônios na camada de saída com função softmax, ou função exponencial normalizada. Por passar por processo de normalização, a soma entre os elementos do vetor de saída resulta em 1, permitindo que seus elementos estejam ligados a probabilidade de ocorrência.

O método de treinamento utilizado é o scaled conjugate gradient backpropagation. O modelo básico de backpropagation discutido na revisão bibliográfica ajusta os pesos das conexões na direção negativa do gradiente, que causa a maior variação da saída para uma variação da entrada. Entretanto, acontece que essa direção é a que apresenta os passos mais largos em direção a convergência, mas não necessariamente o caminho mais rápido. Em algoritmos de gradiente conjugado, desenvolvidos por (Møller, 1993), a procura é feita ao longo de direções conjugadas ao gradiente, produzindo convergência mais rápida (Edoras, 2002).

A ferramenta permite distribuir os dados de entrada entre os grupos de treinamento, validação e teste de acordo com porcentagens do número total de entradas fornecidas. A divisão de dados sugerida é no formato 70% para treinamento, 15% para validação e 15% para teste. Como os dados de teste estarão separados dos dados de treinamento e validação nesse trabalho, será utilizado o número mínimo de dados para teste, de 5%. Assim, o treinamento receberá 75% das amostras e a validação 20% das amostras, que são aleatoriamente selecionadas.

3.3 Resultados das redes neurais artificiais

Os resultados obtidos pelas redes neurais serão avaliados por meio de três gráficos principais:

- a) Diagrama de confusão
- b) Distribuição de apostas e taxa de acerto
- c) Retorno de investimento

Para validar a qualidade do resultado, será estudado o diagrama de confusão obtido passando as amostras de teste pela rede neural obtida após treinamento. O diagrama de confusão é uma matriz “n” x “n” cujas colunas representam as saídas que realmente ocorreram para aquele evento, enquanto as linhas representam as saídas obtidas pela rede neural. Como tem-se 3 possíveis saídas para o problema em questão, será obtida uma matriz 3x3. Cada um dos quadrados possui o número de amostras e/ou porcentagem em relação ao total de amostras que atendem àquelas duas condições. A Figura 6 explica o funcionamento do diagrama.

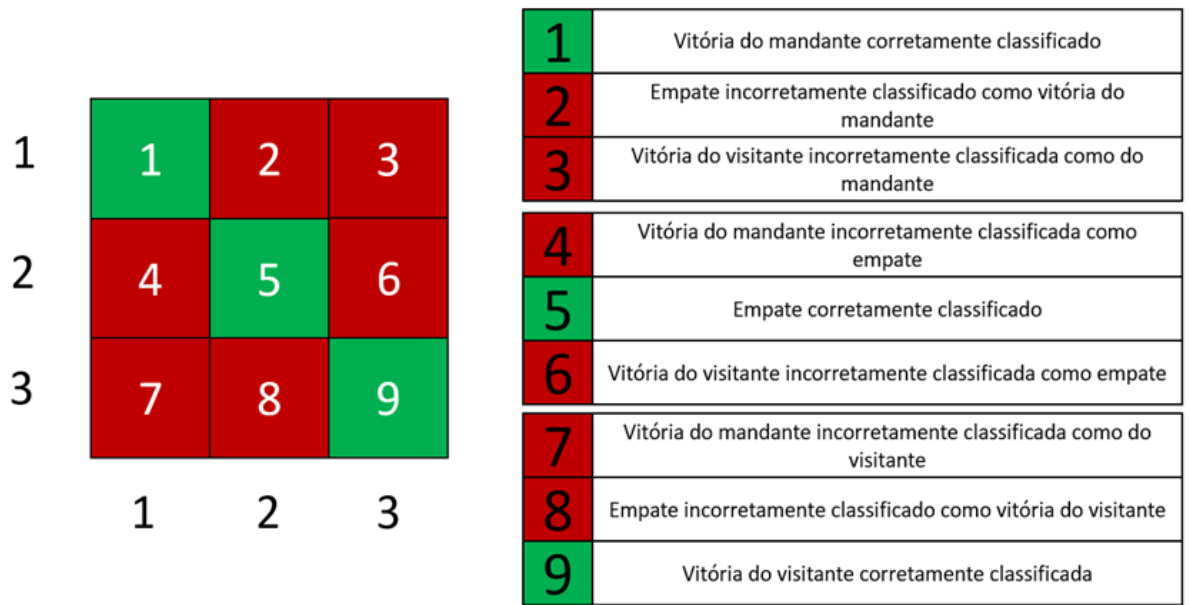


Figura 6: modelo representativo do diagrama de confusão e legenda explicativa dos campos

A Figura 7 trás um exemplo retirado do próprio Matlab®:

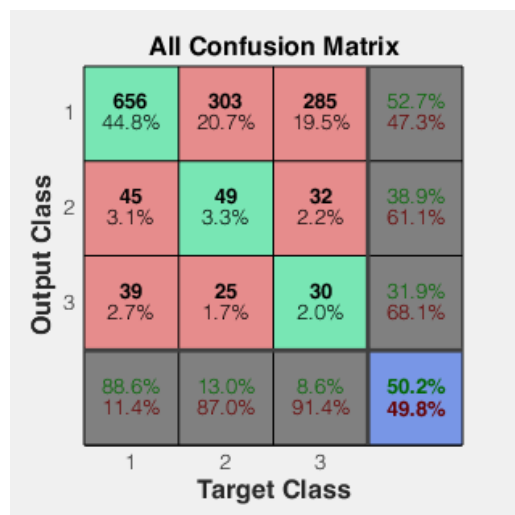


Figura 7: diagrama de confusão retirado do Matlab®

O diagrama fornecido pelo software trás alguns dados adicionais, localizados nas linhas e colunas destacadas em cinza escuro e azul. As células em cinza escuro representam a taxa de acerto e erro da rede neural naquela determinada linha ou coluna. A célula em azul representa a taxa de acerto e erro da rede neural para todos os dados. Esse diagrama mostra que:

- 656 jogos foram corretamente classificados como vitória do mandante;
- 303 jogos foram classificados como vitória do mandante quando na verdade houve um empate;
- 285 jogos foram classificados como vitória do mandante quando na verdade houve um empate;
- Dos jogos classificados como vitória do mandante, 52,7% foram corretamente classificados;
- 45 jogos foram incorretamente classificados como empate quando na verdade houve vitória do mandante;
- Dos jogos onde realmente ocorreram vitória do mandante, 88,6% foram corretamente classificados pela rede neural;
- Dos jogos onde realmente ocorreu vitória do visitante, apenas 8,6% foram corretamente classificados pela rede neural;
- A taxa de acerto total foi de 50,2%. Dos 1 464 jogos analisados, 735 foram corretamente classificados.

O diagrama de confusão trás informações importantes para análise e é o principal método de avaliação de resultado para redes de classificação. Entretanto, para que os resultados fiquem mais visuais, será utilizado um gráfico de distribuição de apostas e taxa de acerto. Um exemplo desse gráfico está apresentado na Figura 8.

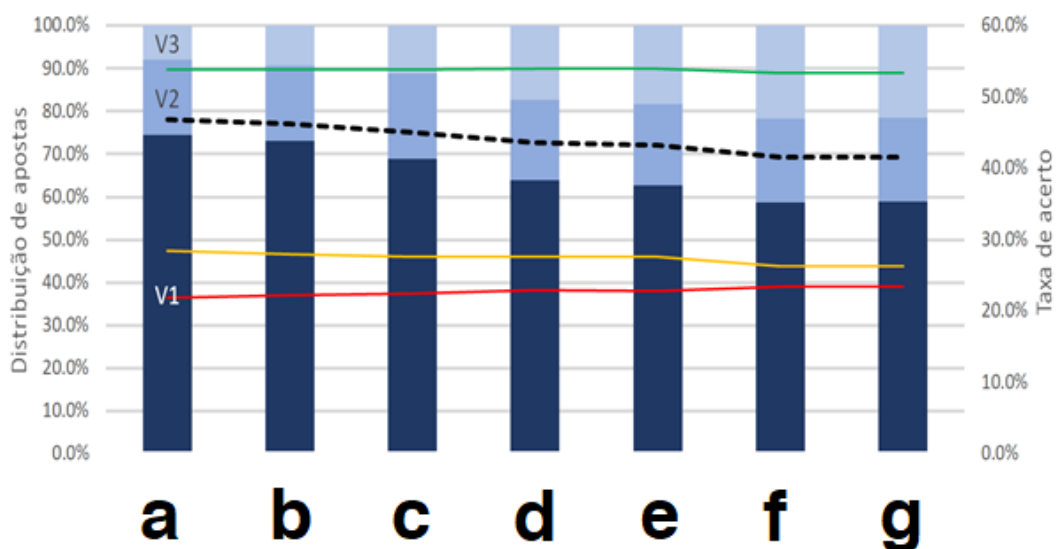


Figura 8: Exemplo de gráfico de distribuição de apostas e taxa de acerto por resultado

A leitura do gráfico é feita da seguinte maneira:

- As barras em três cores apresentam a distribuição das apostas da rede neural para aquele setup. O trecho mais escuro (V1) apresenta a porcentagem de apostas no resultado vitória do mandante, o trecho em tom de azul intermediário (V2) a porcentagem de apostas no empate e, por fim, o trecho em azul mais claro (V3) a porcentagem de apostas na vitória do visitante. Cada barra completa representa 100% das apostas e cada trecho representa a porcentagem do tipo de aposta representado. Por exemplo, a configuração “c” tem ~68% de apostas na vitória do mandante, 22% no empate e 10% na vitória do visitante.
- As linhas coloridas representam a taxa de acerto em cada tipo de resultado, cujo valor é representado no eixo y à direita. A linha verde representa a taxa de acerto em vitória do mandante, a linha amarela em empate e a linha vermelha em vitórias do visitante. Por fim, a linha pontilhada em preto representa a taxa geral de acerto. Nesse exemplo, a configuração “c” acerta ~56% das suas apostas em vitória do mandante, ~28% em empate, ~24% para vitória do visitante e tem ~45% de acerto geral.

Os gráficos anteriores podem ser utilizado para entender o desempenho da rede neural em taxa de acerto pura. Entretanto, a análise pode ter diferentes objetivos. A taxa pura de acertos não é de muita utilidade real, e serve mais como uma curiosidade. No presente trabalho, será explorado como fator de sucesso da rede neural um fator econômico, ou seja, o sucesso financeiro que seria obtido ao realizar apostas em sites de apostas esportivas seguindo as orientações fornecidas pela rede neural. Nesse caso, a taxa de acertos pode indicar, mas não garante, o sucesso da rede. Isso porque o objetivo passa a ser não mais acertar o máximo possível de jogos, mas acertar jogos cujas cotações ofereçam um retorno maior do que o dinheiro perdido nos jogos incorretamente apostados.

Dessa maneira, os dados de entrada serão divididos em dois grupos. A princípio, serão utilizados os dados de 2007, 2008, 2009, 2010, 2011, 2012 e 2013 para treinamento e 2014, 2015, 2016 e 2017 para teste. O primeiro grupo irá treinar a rede

neural, que depois apostará nos jogos do segundo grupo conforme as diretrizes adquiridas.

Três estratégias de aposta serão testadas:

- Bet on all (BOA), onde a rede neural aposta uma quantia fixa de R\$10,00 em todos os jogos possíveis no resultado que encontrar maior probabilidade de ocorrer
- Bet if minimum (BIM), onde a rede neural aposta uma quantia fixa de R\$10,00 em todos os jogos desde que tenha um mínimo de confiança que aquele resultado irá ocorrer. Esse mínimo de confiança pode ser variado para entender seu impacto nos resultados
- Bet if value (BIV), onde a rede neural aposta em todos os jogos uma quantia fixa de R\$10,00 em que calcula que a probabilidade do resultado mais provável ocorrer é maior que a probabilidade cujo site de aposta acredita ser a real por uma margem mínima definida, que pode ser alterada. Essa aposta é chamada de “aposta em valor”, pois o valor em apostas esportivas encontra-se em determinar com melhor acuracidade que um evento tem mais chance de ocorrer do que a casa de apostas acredita. Nesse caso, a casa de apostas está pagando um valor maior do que deveria para quem aposta naquele caso, e esse valor pode ser coletado. Uma aposta de valor terá resultados positivos caso o modelo construído seja melhor que as previsões da casa de aposta em uma magnitude superior a da margem de lucro da casa.

Um modelo pode ser considerado melhor caso obtenha um retorno financeiro mais alto e mais consistente durante os anos de teste. É esperado que a análise a ser realizada trará resultados bem aleatórios e, portanto, a consistência é importante. O resultado de uma partida de futebol deve sim estar relacionado com a média de gols de cada time, sua fase atual no campeonato e ao número de chutes ao gol por partida, mas existem uma quantidade enorme de fatores que podem afetar o resultado e não podem, em tempo hábil, ser devidamente detectados e quantificados. Alguns exemplos dentre muitos outros são: qualidade do descanso da equipe, estado de saúde de familiares dos jogadores, condição do gramado, erros de arbitragem, influência da torcida, condições meteorológicas, etc.

Dessa maneira, as entradas escolhidas provavelmente não são suficientes para prever os resultados com boa acuracidade, e redes neurais com as mesmas entradas, mas treinadas de forma diferente podem chegar a conclusões diferentes sobre um mesmo resultado. Este é mais um argumento que incentiva a análise do resultado da rede neural por meio do resultado financeiro das apostas. Um resultado positivo indica que você está prevendo os resultados não perfeitamente – esse não é o objetivo do presente estudo – mas sim acima da média das outras pessoas, indicando presença de valor.

Para determinar a confiabilidade de um setup, cada combinação de elementos de setup (tipos de entradas, número de neurônios, anos analisados, etc) serão mantidas e diversas redes neurais serão construídas a partir delas com pesos e bias aleatoriamente iniciados em seus neurônios. Isso significa que a coleção de redes neurais similares obterão diferentes resultados de retorno financeiro em apostas e índice de acerto dos resultados. Após obter todos esses valores, será obtido o valor mediano do resultado, bem como uma faixa de confiabilidade. A faixa de confiabilidade contém o mesmo número de resultados acima e abaixo da mediana, porém descartando os 10% resultados mais distantes para cima, bem como os 10% resultados mais distantes para baixo. Isso significa que a faixa de confiabilidade começa com o resultado do setup que representa o 10º percentil dos resultados, e acaba no resultado que representa o 90º percentil dos resultados. Uma faixa de confiabilidade mais estreita significa que aquela combinação de elementos de setup tende a apresentar resultados que fogem menos ao valor da mediana. A confiabilidade e consistência dos resultados apresentados por um setup é muito importante: um setup que gera resultados entre 2 e 6%, de retorno de investimento, com mediana de 4%, é considerado melhor do que um setup que gera resultados entre -2% e 10%, com mediana também em 4%.

Um exemplo de gráfico que apresenta retornos de investimento para diferentes setups está apresentado na Figura 9.

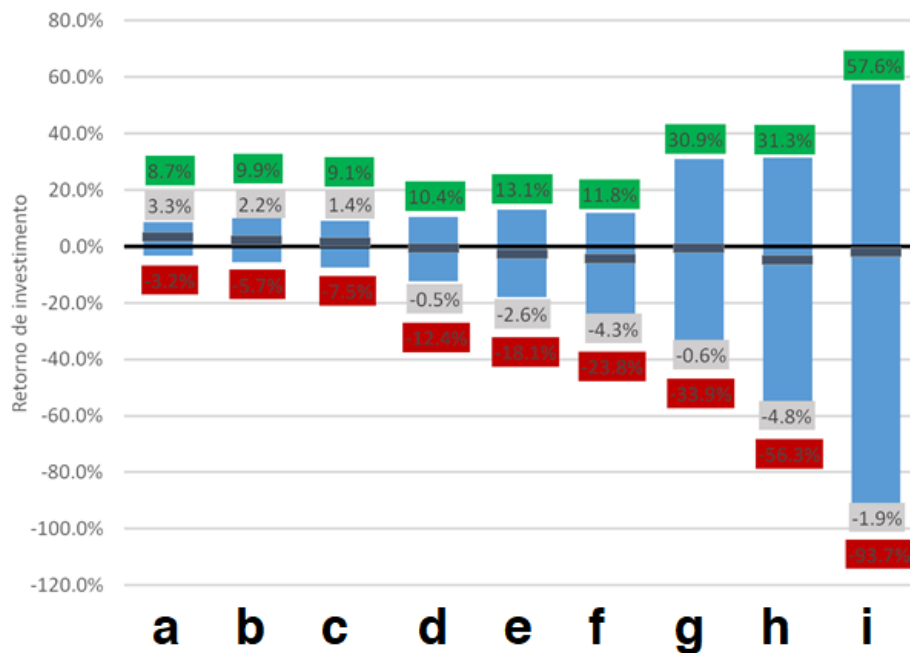


Figura 9: exemplo de gráfico de retorno de investimento

A leitura do gráfico é feita da seguinte maneira:

- As faixas em azul representam os valores de retorno de investimento entre o 10º percentil e o 90º percentil, ou seja, 80% dos valores obtidos, sendo 40% acima da mediana e 40% abaixo da mediana. A mediana é representada pela barra horizontal preta dentro de uma faixa;
- As caixas verdes representam o valor superior da faixa (90º percentil), as caixas vermelhas o valor inferior da faixa (10º percentil) e as caixas em cinza apresentam o valor da mediana, sendo colocadas junto a caixa verde quando positivas e junto a caixa vermelha quando negativas.

Com o objetivo de comparar os resultados obtidos por diferentes setups, o resultado de um setup de rede neural será qualificado como melhor ou pior com base em uma função pontuação. A pontuação de um setup vai de 0 a 100 pontos, sendo 33% atribuídos ao fator confiabilidade e 67% ao fator retorno de investimento. A nota mínima de retorno de investimento é 0 e é atribuída ao pior resultado de ROI entre as redes comparadas, enquanto a nota máxima é 67 e atribuída ao melhor resultado de ROI observado. A pontuação de redes com ROI intermediário varia linearmente com sua posição entre os menor e maior valor. Para o fator confiabilidade, a pior nota, 0, é dada para a maior diferença entre maior e menor valor de ROI dentre as redes geradas

por aquele setup na faixa de confiabilidade (10º percentil até 90º percentil). O valor máximo é dado para a faixa de confiabilidade mais estreita e valores intermediários seguem raciocínio análogo ao da pontuação de ROI. A Tabela 2 define um exemplo de cálculo da pontuação de uma rede neural dentro de um setup:

Tabela 2: Dados de entrada para cálculo do exemplo de resultado de uma rede

| Grandeza | Valor |
|--|-------|
| ROI mediano máximo entre as redes neurais geradas pelos diferentes setups testados | 9,5% |
| ROI mediano mínimo entre as redes neurais geradas pelos diferentes setups testados | 3,3% |
| Variação máxima entre extremos da faixa de confiabilidade do resultado de um setup dentre todos os setups testados (menos confiável) | 15% |
| Variação mínima entre extremos da faixa de confiabilidade do resultado de um setup dentre todos os setups testados (mais confiável) | 5% |
| ROI mediano do resultado obtido com um setup x | 4,5% |
| Superior da faixa de confiabilidade do resultado obtido com um setup x | 2,5% |
| Inferior da faixa de confiabilidade do resultado obtido com um setup x | 8,0% |

A Tabela 3 representa o cálculo da pontuação obtida para o setup x:

Tabela 3: exemplo de cálculo para pontuação na situação definida na Tabela 2

| Pontuação | Cálculo | Valor |
|----------------|--|-------|
| ROI | $67 * \frac{4,5\% - 3,3\%}{9,5\% - 3,3\%}$ | 27,92 |
| Confiabilidade | $33 * \frac{15\% - (8,0\% - 2,5\%)}{15\% - 5\%}$ | 31,35 |
| Total | Pts. ROI + Pts. Confiabilidade | 59,27 |

Por fim, algumas siglas vão ser utilizadas para representar grandezas envolvidas na análise. Essas siglas estão apresentadas na Tabela 4: siglas utilizadas na apresentação dos resultados.

Tabela 4: siglas utilizadas na apresentação dos resultados

| Tipo de grandeza | Sigla | Grandeza representada |
|-----------------------|--------|---|
| Proporção de apostas | V1 | Proporção de apostas em vitória do mandante |
| | V2 | Proporção de apostas em empate |
| | V3 | Proporção de apostas em vitória do visitante |
| Taxas de acerto | WV1 | Taxa de acerto em apostas em vitória do mandante |
| | WV2 | Taxa de acerto em apostas em empate |
| | WV3 | Taxa de acerto em apostas em vitória do visitante |
| | WT | Taxa de acerto em todas as apostas |
| Estratégias de aposta | BOA | Bet on all – aposta em todos os jogos |
| | BIM | Bet if minimum – aposta apenas caso haja uma certeza mínima |
| | BIV | Bet if value - aposta apenas caso encontre discrepância entre probabilidade calculada e probabilidade observada |
| Financeiro | ROI 10 | Retorno de investimento do limite inferior da faixa de confiabilidade (10º percentil) |
| | ROI m | Retorno de investimento mediano |
| | ROI 90 | Retorno de investimento do limite superior da faixa de confiabilidade (90º percentil) |
| Outros | PMV | Payout médio em caso de vitória representa o retorno de investimento médio nas apostas que a rede acerta. Ele está relacionado com o quão difícil o resultado previsto era de acontecer. Se uma rede acertou 5 apostas com cotações 2,3; 2,5; 2,5; 1,5 e 1,7 o PMV é de 1,1 (lembre-se que o ROI em um payout de 2,3 é na verdade 1,3, pois “1” foi apostado) |

3.4 Análises propostas

O presente trabalho possui como sub-objetivos analisar itens como eficácia das redes neurais e como se compara com o método estatístico e como variam os resultados das redes neurais com diferentes setups. Assim, para chegar a essas conclusões de forma mais esclarecedora, é proposta a sequência de análises passo a passo cujas descrições e objetivos estão descritos de forma resumida na Tabela 5 e apresentados de forma detalhada nos subitens a seguir.

Tabela 5: resumo das análises a serem realizadas

| Análise | Descrição do procedimento | Objetivo |
|---------|---|--|
| 1 | Levantamento de amostras, criação de uma planilha que simula o método estatístico aplicado, observação dos resultados | Entender a eficácia do método estatístico |
| 2 | Levantamento de amostras, programação em ambiente MATLAB para obter o resultado da rede neural apenas utilizando as entradas básicas utilizadas pelo método estatístico. Utilização de um setup básico. | Comparar a eficácia da rede neural com o método estatístico |
| 3 | Variação de setup: probabilidade mínima para realizar aposta na estratégia 2 (BIM) | Entender como utilizar a estratégia BIM para melhores resultados |
| 4 | Variação de setup: variação de probabilidade mínima para realizar aposta na estratégia 3 (BIV) | Entender como utilizar a estratégia BIV para melhores resultados |
| 5 | Variação de setup: resultados obtidos para cada ano de teste | Entender como os resultados variam ano a ano e os fatores que causam essa variação |

| | | |
|----|--|--|
| 6 | Variação de setup: resultados obtidos com diferentes quantidades de entradas | Entender a quantidade mínima de entradas para obter um resultado já satisfatório, bem como a quantidade após a qual introduzir mais entradas não ajuda no problema |
| 7 | Variação de setup: resultados obtidos com diferentes quantidades de neurônios na camada escondida | Entender como os resultados variam com a quantidade de unidades funcionais básicas da rede neural, o neurônio |
| 8 | Variação de setup: resultados obtidos com diferentes grupos de variáveis de entrada e comparação para entender os melhores grupos de entrada com base na função pontuação | Pré-selecionar quais as entradas que podem gerar melhores resultados pela rede neural, bem como as entradas que atrapalham ou são irrelevantes |
| 9 | Variação de setup: resultados obtidos com diferentes combinações das variáveis de entrada dos melhores grupos de entrada para entender as melhores entradas com base na função pontuação | Entender quais as entradas que geram melhores resultados, bem como as entradas que atrapalham ou são irrelevantes |
| 10 | Direcionamento estratégico: os resultados de ROI das redes neurais variam muito em torno da mediana. Além dos anos de treinamento, serão inseridos anos de filtragem, que aprovarão a rede neural obtida pelo treinamento apenas caso ela tenha um bom resultado nas suas amostras, e então liberando-a para examinar os dados de análise. | Entender se os anos de filtragem conseguem direcionar as redes neurais para a direção correta da previsão de resultados de futebol, selecionando apenas as com melhor potencial para gerar melhores resultados nos anos de teste |

3.4.1 Análise 1 – Eficiência do método estatístico

Após o levantamento de dados descritos no item 3.2, uma planilha será construída para simular o método estatístico.

O método estatístico objetiva entender as probabilidades dos três possíveis resultados de uma partida – vitória do mandante, empate e vitória do visitante. As entradas utilizadas para tal são a força de ataque do time mandante jogando em casa (ATKH), força de ataque do time visitante jogando fora de casa (ATKA), força de defesa do time mandante jogando em casa (DEFH) e força de defesa do time visitante jogando fora de casa, (DEFA) calculadas a partir dos desempenhos de cada time no campeonato do ano anterior. O processo segue quatro etapas: cálculo das forças de ataque e defesa por time, determinação do placar mais provável, cálculo das probabilidades de todas as possibilidades de placar e por fim cálculo das probabilidades dos três possíveis resultados.

Para calcular a força de ataque do time mandante, basta dividir o número de gols feitos em casa por esse time na temporada passada pelo número de partidas que jogou em casa, obtendo a média de gols em casa daquele time na temporada passada. Por fim, divide-se esse número pela média da média de gols em casa de todos os outros times do campeonato na temporada passada. As outras três grandezas de entrada seguem o mesmo padrão. As Equações 4, 5, 6 e 7 apresentam os cálculos dessas variáveis.

$$ATKH = \frac{\text{Gols pró em casa do mandante na temporada passada}}{\text{Jogos em casa do mandante na temporada passada}} \div \text{média de gols pró em casa por equipe na temporada passada}$$

Equação 4: equação para cálculo da força de ataque do time mandante

$$ATKA = \frac{\text{Gols pró fora de casa do visitante na temporada passada}}{\text{Jogos fora de casa do visitante na temporada passada}} \div \text{média de gols pró fora de casa por equipe na temporada passada}$$

Equação 5: equação para cálculo da força de ataque do time visitante

$$DEFH = \frac{\text{Gols contra dentro de casa do mandante na temporada passada}}{\text{Jogos dentro de casa do mandante e na temporada passada}} \div \text{média de gols contra dentro de casa por equipe na temporada passada}$$

Equação 6: equação para cálculo da força de defesa do time mandante

$$DEFA = \frac{\text{Gols contra fora de casa do visitante na temporada passada}}{\text{Jogos fora de casa do visitante na temporada passada}} \div \text{m\u00e9dia de gols contra fora de casa por equipe na temporada passada}$$

Equa\u00e7\u00e3o 7: equa\u00e7\u00e3o para c\u00e1lculo da for\u00e7a de defesa do time mandante

Em seguida, o placar esperado da partida \u00e9 dado calculando o prov\u00e1vel n\u00famero de gols que cada equipe far\u00e1, sendo GH o n\u00famero de gols feitos pelo mandante e GA o n\u00famero de gols feitos pelo visitante, mostrados nas Equa\u00e7\u00f5es 8 e 9.

$$GH = ATKH \times DEFA \times \text{M\u00e9dia de gols dos mandantes jogando em casa na temporada passada}$$

Equa\u00e7\u00e3o 8: c\u00e1lculo da quantidade de gols esperadas para o time mandante

$$GA = ATKA \times DEFH \times \text{M\u00e9dia de gols dos visitantes jogando fora de casa na temporada passada}$$

Equa\u00e7\u00e3o 9: c\u00e1lculo da quantidade de gols esperadas para o time visitante

Nesse momento, esperamos que o time da casa fa\u00e7a GH gols e o time mandante fa\u00e7a GA gols. Entretanto, os valores obtidos ser\u00e3o provavelmente n\u00e3o inteiros, e \u00e9 imposs\u00edvel que esse cen\u00e1rio aconte\u00e7a. Na verdade, deseja-se encontrar a probabilidade de que cada poss\u00edvel cen\u00e1rio aconte\u00e7a, ou seja: dado que espera-se que uma equipe fa\u00e7a "x" gols, encontrar qual \u00e9 a probabilidade de que ela fa\u00e7a 0, 1, 2, 3, 4 ou 5 gols. Exclui-se uma quantidade de gols de 6 para cima uma vez que a probabilidade de ocorr\u00eancia torna-se muito pequena e, portanto, desprez\u00edvel.

Esse processo \u00e9 feito utilizando-se uma distribui\u00e7\u00e3o de Poisson. Neste caso, a probabilidade de uma equipe fazer k gols dado que o n\u00famero esperado de gols \u00e9 λ \u00e9 dada pela Equa\u00e7\u00e3o 10.

$$f(k, \lambda) = \frac{e^{-\lambda} \lambda^k}{k!}$$

Equa\u00e7\u00e3o 10: probabilidade de k ocorr\u00eancias de um evento com ocorr\u00eancia esperada λ pelo m\u00e9todo de Poisson

Para uma explica\u00e7\u00e3o mais clara, suponhamos que ap\u00f3s os c\u00e1lculos iniciais encontramos um n\u00famero de gols esperados para o mandante (GH) de 1,7 e para o visitante (GA) de 0,9. A probabilidade de cada time fazer um certo n\u00famero de gols no intervalo estudado \u00e9 apresentada nas Tabelas 6 e 7.

Tabela 6: variáveis, fórmulas e probabilidades de ocorrência de cada quantidade de gols entre 0 e 5 para o time mandante

| # gols | 0 | 1 | 2 | 3 | 4 | 5 |
|---------------|-------------------------------|-------------------------------|-------------------------------|-------------------------------|-------------------------------|-------------------------------|
| Variável | GH_0 | GH_1 | GH_2 | GH_3 | GH_4 | GH_5 |
| Fórmula | $\frac{e^{-1,7} * 1,7^0}{0!}$ | $\frac{e^{-1,7} * 1,7^1}{1!}$ | $\frac{e^{-1,7} * 1,7^2}{2!}$ | $\frac{e^{-1,7} * 1,7^3}{3!}$ | $\frac{e^{-1,7} * 1,7^4}{4!}$ | $\frac{e^{-1,7} * 1,7^5}{5!}$ |
| Probabilidade | 0,1827 | 0,3106 | 0,2640 | 0,1496 | 0,0636 | 0,0216 |

Tabela 7: variáveis, fórmulas e probabilidades de ocorrência de cada quantidade de gols entre 0 e 5 para o time mandante

| # gols | 0 | 1 | 2 | 3 | 4 | 5 |
|---------------|-------------------------------|-------------------------------|-------------------------------|-------------------------------|-------------------------------|-------------------------------|
| Variável | GA_0 | GA_1 | GA_2 | GA_3 | GA_4 | GA_5 |
| Fórmula | $\frac{e^{-0,9} * 0,9^0}{0!}$ | $\frac{e^{-0,9} * 0,9^1}{1!}$ | $\frac{e^{-0,9} * 0,9^2}{2!}$ | $\frac{e^{-0,9} * 0,9^3}{3!}$ | $\frac{e^{-0,9} * 0,9^4}{4!}$ | $\frac{e^{-0,9} * 0,9^5}{5!}$ |
| Probabilidade | 0,4066 | 0,3659 | 0,1647 | 0,0494 | 0,0111 | 0,0020 |

A probabilidade de cada placar é dado pelo produto de cada evento independente. Se um resultado é apresentado no formato m-n, onde m é o número de gols do mandante e n o número de gols do visitante, todos os resultados possíveis podem ser representados em uma tabela cujas colunas representam gols do mandante e linhas representam gols do visitante, como feito na Tabela 8.

Tabela 8: possíveis placares em que cada equipe fazem 5 gols ou menos

| | 0 | 1 | 2 | 3 | 4 | 5 |
|---|-----|-----|-----|-----|-----|-----|
| 0 | 0-0 | 1-0 | 2-0 | 3-0 | 4-0 | 5-0 |
| 1 | 0-1 | 1-1 | 2-1 | 3-1 | 4-1 | 5-1 |
| 2 | 0-2 | 1-2 | 2-2 | 3-2 | 4-2 | 5-2 |
| 3 | 0-3 | 1-3 | 2-3 | 3-3 | 4-3 | 5-3 |
| 4 | 0-4 | 1-4 | 2-4 | 3-4 | 4-4 | 5-4 |
| 5 | 0-5 | 1-5 | 2-5 | 3-5 | 4-5 | 5-5 |

A probabilidade de cada um desses placares acontecerem é dado pelo produto entre a probabilidade de cada time fazer a quantidade de gols, representadas na Tabela 9.

Tabela 9: cálculo da probabilidade de ocorrência de cada um dos resultados apresentados na Tabela 8

| | 0 | 1 | 2 | 3 | 4 | 5 |
|---|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|
| 0 | $GH_0 \cdot GA_0$ | $GH_1 \cdot GA_0$ | $GH_2 \cdot GA_0$ | $GH_3 \cdot GA_0$ | $GH_4 \cdot GA_0$ | $GH_5 \cdot GA_0$ |
| 1 | $GH_0 \cdot GA_1$ | $GH_1 \cdot GA_1$ | $GH_2 \cdot GA_1$ | $GH_3 \cdot GA_1$ | $GH_4 \cdot GA_1$ | $GH_5 \cdot GA_1$ |
| 2 | $GH_0 \cdot GA_2$ | $GH_1 \cdot GA_2$ | $GH_2 \cdot GA_2$ | $GH_3 \cdot GA_2$ | $GH_4 \cdot GA_2$ | $GH_5 \cdot GA_2$ |
| 3 | $GH_0 \cdot GA_3$ | $GH_1 \cdot GA_3$ | $GH_2 \cdot GA_3$ | $GH_3 \cdot GA_3$ | $GH_4 \cdot GA_3$ | $GH_5 \cdot GA_3$ |
| 4 | $GH_0 \cdot GA_4$ | $GH_1 \cdot GA_4$ | $GH_2 \cdot GA_4$ | $GH_3 \cdot GA_4$ | $GH_4 \cdot GA_4$ | $GH_5 \cdot GA_4$ |
| 5 | $GH_0 \cdot GA_5$ | $GH_1 \cdot GA_5$ | $GH_2 \cdot GA_5$ | $GH_3 \cdot GA_5$ | $GH_4 \cdot GA_5$ | $GH_5 \cdot GA_5$ |

Para o nosso exemplo, os resultados estão apresentados na Tabela 10.

Tabela 10: probabilidades de ocorrência de cada um dos resultados apresentados na Tabela 8

| | 0 | 1 | 2 | 3 | 4 | 5 |
|---|-------|-------|-------|-------|-------|-------|
| 0 | 0,074 | 0,126 | 0,107 | 0,061 | 0,026 | 0,009 |
| 1 | 0,067 | 0,114 | 0,097 | 0,055 | 0,023 | 0,008 |
| 2 | 0,030 | 0,051 | 0,043 | 0,025 | 0,010 | 0,004 |
| 3 | 0,009 | 0,015 | 0,013 | 0,007 | 0,003 | 0,001 |
| 4 | 0,002 | 0,003 | 0,003 | 0,002 | 0,001 | 0,000 |
| 5 | 0,000 | 0,001 | 0,001 | 0,000 | 0,000 | 0,000 |

Por fim, se $P(GH_m GA_n)$ representa a probabilidade do placar onde o mandante faz m gols e o visitante faz n gols, encontramos a probabilidade dos três resultados possíveis (vitória do mandante, empate ou vitória do visitante) pela soma das probabilidades referentes a combinações de placares que configuram esses resultados, conforme descrito na Tabela 11.

Tabela 11: cálculo das probabilidades de ocorrência dos três possíveis resultados

| | |
|----------------------|--|
| Vitória do mandante | $\sum P(GH_m \times GA_n)$, para todo caso em que $m > n$ |
| Empate | $\sum P(GH_m \times GA_n)$, para todo caso em que $m = n$ |
| Vitória do visitante | $\sum P(GH_m \times GA_n)$, para todo caso em que $m < n$ |

No nosso exemplo, teríamos 55,46% de probabilidade de vitória do mandante (casos em verde na Tabela 10), 23,95% de probabilidade de empate (em laranja) e 19,75% de probabilidade de vitória do visitante (em vermelho). Nesse caso, seguindo o método estatístico, diríamos que o resultado da partida seria vitória do mandante.

Para um dado ano y do Campeonato Brasileiro de Futebol, é possível levantar os dados de entrada para 16 dos 20 times participantes. Isso porque 4 dos times presentes no ano y do campeonato são provenientes do campeonato da série B do ano $y-1$, e suas médias de gols fora e dentro de casa no campeonato da série B não pode ser utilizados e comparados contra o das outras 16 equipes que possuem médias de gols fora e dentro de casa no campeonato da série A. Os jogos que envolvem as equipes provenientes da série B do campeonato passado não poderão ser utilizados para o estudo estatístico, de modo que, em cada ano, 240 dos 380 jogos envolvem apenas dois times que disputaram o campeonato no ano passado.

Serão levantados dados de entrada e resultados dos jogos de onze campeonatos (entre 2007 e 2017), totalizando teoricamente 2640 jogos na amostra. Então, para encontrar a taxa de acerto do modelo, basta calcular a porcentagem de jogos em que o resultado com a maior probabilidade calculada pelo modelo realmente aconteceu.

3.4.2 Análise 2 – redes neurais com setup básico

Após procedimento descrito no item 3.2, um código de MATLAB foi projetado para obter o resultado de um dado setup de rede neural a ser definido na entrada. Suas entradas e saídas estão descritas na Tabela 12.

Tabela 12: entradas e saídas utilizadas pelo programa da Análise 2

| Entradas | |
|-----------|--|
| x | Variáveis de entrada das amostras de treinamento e validação |
| y | Variáveis de saída das amostras de treinamento e validação |
| a | Variáveis de entrada das amostras de teste |
| b | Variáveis de saída das amostras de teste |
| Betodds | Cotações para vitória do mandante, empate e vitória do visitante obtida pela média das cotações entre diversos sites de aposta |
| Betamt | Quantidade apostada por aposta – padrão: R\$10,00; esse valor não afeta o resultado de retorno de investimento, é apenas um parâmetro utilitário |
| Neurons | Número de neurônios na camada escondida |
| Tries | Número de redes geradas no mesmo setup para obter um valor mediano e uma faixa de confiabilidade |
| Mod2min | Probabilidade mínima de ocorrência do evento para aposta na estratégia 2 (BIM) |
| Mod3min | Disparidade mínima de probabilidade de ocorrência do evento definido como mais provável pela rede neural em relação a probabilidade definida pelo site de aposta, definido pela estratégia 3 (BIV) |
| prct | Porcentagem de valores de ROI a serem descartados tanto no extremo superior como no extremo inferior para construção do intervalo de confiança |
| Saídas | |
| resmodres | Matriz 3x14 com uma coletânea de saídas de interesse. Cada linha representa uma estratégia de aposta, e as colunas representam, em ordem: número de neurônios na camada escondida (1 coluna), ROI do 10º percentil (1 coluna), , ROI mediano (1 coluna), ROI do 90º percentil (1 coluna), valores da matriz de confusão (9 colunas), cotação média nas apostas ganhas (1 coluna) |

Os valores de payout históricos a serem incluídos na matriz betodds foram retirados do site <http://www.oddsportal.com/results/#soccer>, acessado em Março/2018.

O setup básico a ser utilizado é:

- 4 variáveis de entrada básicas utilizadas no modelo estatístico (AVGGPA, AVGGPH, AVGGCH, AVGGCA) para definir “x” e “a”
- 7 anos de amostras para treinamento e validação (2007-2013) para definir “x” e “y”
- 4 anos de amostras para teste (2014-2017) para definir “a” e “b”
- 10 para betamt para definir a aposta fixa
- 50 neurônios na camada escondida para definir “neurons”
- 1000 redes geradas na configuração para definir “tries”
- 50% de probabilidade para aposta mínima na estratégia BIM para definir “mod2min”
- 10% de variação entre probabilidade calculada pela rede e apresentada pela casa de apostas para estratégia 3 para definir “mod3min”
- 10 para porcentagem de extremos de ROI descartados em cada extremo, definindo “prct”

O programa utilizado está transcrito no Código 1. A explicação do passo a passo do código não é objeto deste trabalho, mas pequenos comentários foram incluídos para auxiliar a compreensão.

Código 1: cálculo do diagrama de confusão e resultado financeiro para os três métodos de aposta para uma determinada combinação de setup

```
function [resmodres] = Analise1(x,y,betodds,betamt,neurons,tries,mod2min,mod3min,prct)

%Contador de tempo
tic
toci = 0;
%Inicializar matrizes que serão utilizadas para montar o resultado
%p indica a quantidade de partidas apostadas por tipo de resultado em cada estratégia
%w indica a quantidade de partidas vencidas por tipo de resultado em cada estratégia
%tw indica a taxa geral de acertos em cada estratégia
%ab indica o payout medio em caso de vitória
%cf indica os 9 componentes da matriz de confusão
```

```

p1=zeros(tries,3);
w1=zeros(tries,3);
tw1=zeros(tries,1);
ab1=zeros(tries,1);
p2=zeros(tries,3);
w2=zeros(tries,3);
tw2=zeros(tries,1);
ab2=zeros(tries,1);
p3=zeros(tries,3);
w3=zeros(tries,3);
tw3=zeros(tries,1);
ab3=zeros(tries,1);
cf1 = zeros(tries,9);
cf2 = zeros(tries,9);
cf3 = zeros(tries,9);

%Obter as probabilidades de ocorrência previstas pelo site de aposta a
%partir das cotações dos jogos
betprob = 1./betodds;
for temp = 1:size(betodds,1)
    sumtemp = sum(betprob(temp,:));
    betprob(temp,1) = betprob(temp,1)/sumtemp;
    betprob(temp,2) = betprob(temp,2)/sumtemp;
    betprob(temp,3) = betprob(temp,3)/sumtemp;
end

%Inicio das iterações. Cada iteração constroi uma rede neural e obtem os
%resultados para ela

for j = 1:tries
    %Geração da rede neural e treinamento a partir dos dados de treinamento
    x=x.';
    y=y.';
    a=a;
    b=b;
    net = patternnet(neurons,'trainscg','crossentropy');
    net = train(net,x,y);
    %Geração das previsões
    k = net(a.);
    nnet.guis.closeAllViews();
    k=k.';
    x=x.';
    y=y.';
    out = size(k,1);
    %Zerar variáveis acumulativas entre cada iteração
    cashbet1(j) = 0;
    cashwon1(j) = 0;
    cashbet2(j) = 0;
    cashwon2(j) = 0;

```

```

cashbet3(j) = 0;
cashwon3(j) = 0;
%Encontrar resultado previsto e resultado ocorrido
for i = 1:out
    off = find(k(i,:) == max(k(i,:)));
    offr = find(b(i,:) == max(b(i,:)));
    knorm(i,:)=0;
    knorm(i,off)=1;
%A Zerar variáveis de aposta em cada estratégia de aposta
    bet(1)=0;
    bet(2)=0;
    bet(3)=0;
%A estratégia 1 (BOA) aposta todas as vezes
    bet(1)=betamt;
    cashbet1 = cashbet1+bet(1);
    p1(j,off) = p1(j,off)+1;
%A estratégia 2 (BIF) aposta caso a rede entenda que o evento tem uma
%probabilidade mínima de ocorrer definida pela entrada mod2min
    if(k(i,off) > mod2min)
        bet(2)=betamt;
        cashbet2 = cashbet2+bet(2);
        p2(j,off) = p2(j,off)+1;
    end
%A estratégia 3 (BIV) aposta caso a rede entenda que o evento tem uma
%probabilidade de ocorrer maior que a probabilidade que o site de apostas
%julga ser a correta em no mínimo uma quantidade definida pela entrada
%mod3min, detectando uma aposta de valor
    if (((k(i,off) - betprob(i,off))) > mod3min)
        bet(3)=betamt;
        cashbet3 = cashbet3+bet(3);
        p3(j,off) = p3(j,off)+1;
    end
%Checagem da aposta: premiar caso correta, penalizar caso incorreta e
%contagem dos resultados
    if (isequal(off,offr))
        %Contabiliza o retorno financeiro no acerto
        cashwon1(j) = cashwon1(j) + (betodds(i,off)-1)*bet(1);
        cashwon2(j) = cashwon2(j) + (betodds(i,off)-1)*bet(2);
        cashwon3(j) = cashwon3(j) + (betodds(i,off)-1)*bet(3);
        %Contabiliza taxa de acerto de jogos
        w1(j,off) = w1(j,off) + 1;
        w2(j,off) = w2(j,off) + (bet(2)~=0);
        w3(j,off) = w3(j,off) + (bet(3)~=0);
        %Contabiliza dados para montar a matriz de confusão
        cf1(j,1+4*(off-1)) = cf1(j,1+4*(off-1)) + 1;
        cf2(j,1+4*(off-1)) = cf2(j,1+4*(off-1)) + 1*(bet(2)~=0);
        cf3(j,1+4*(off-1)) = cf3(j,1+4*(off-1)) + 1*(bet(3)~=0);
        %Contabiliza dados de payout da aposta
        ab1(j,1) = ab1(j,1) + betodds(i,off)-1;

```

```

ab2(j,1) = ab2(j,1) + (betodds(i,off)-1)*(bet(2)~=0);
ab3(j,1) = ab3(j,1) + (betodds(i,off)-1)*(bet(3)~=0);
else
%Contabiliza o retorno financeiro no erro
cashwon1(j) = cashwon1(j) - bet(1);
cashwon2(j) = cashwon2(j) - bet(2)*(bet(2)~=0);
cashwon3(j) = cashwon3(j) - bet(3)*(bet(3)~=0);
%Contabiliza dados para montar a matriz de confusão
cf1(j,3*(off-1)+offr) = cf1(j,3*(off-1)+offr) + 1;
cf2(j,3*(off-1)+offr) = cf2(j,3*(off-1)+offr) + 1*(bet(2)~=0);
cf3(j,3*(off-1)+offr) = cf3(j,3*(off-1)+offr) + 1*(bet(3)~=0);
end
end
%Montagem da matriz de retorno financeiro
roi1(j)=cashwon1(j)/cashbet1(j);
roi2(j)=cashwon2(j)/cashbet2(j);
roi3(j)=cashwon3(j)/cashbet3(j);
%Montagem da matriz de payout médio
ab1(j) = ab1(j)/sum(w1(j,1:3));
ab2(j) = ab2(j)/sum(w2(j,1:3));
ab3(j) = ab3(j)/sum(w3(j,1:3));
%Montagem da matriz de vitórias em geral
tw1(j) = (sum(w1(j,1:3)))/(sum(p1(j,1:3)));
tw2(j) = (sum(w2(j,1:3)))/(sum(p2(j,1:3)));
tw3(j) = (sum(w3(j,1:3)))/(sum(p3(j,1:3)));
for temp = 1:3
%Montagem da matriz de vitórias por possibilidade de resultado
w1(j,temp)=w1(j,temp)/p1(j,temp);
w2(j,temp)=w2(j,temp)/p2(j,temp);
w3(j,temp)=w3(j,temp)/p3(j,temp);
end
tempsum(1)=sum(p1(j,1:3));
tempsum(2)=sum(p2(j,1:3));
tempsum(3)=sum(p3(j,1:3));
for temp = 1:3
%Montagem da matriz de distribuição de apostas entre os possíveis
%resultados
p1(j,temp)=p1(j,temp)/tempsum(1);
p2(j,temp)=p2(j,temp)/tempsum(2);
p3(j,temp)=p3(j,temp)/tempsum(3);
end

%Mostrador de previsão de tempo restante
if (toci==0)
toci = 1;
toci = toci+1;
timelp = toc;
end
timelf = ceil(timelp*(tries-j));

```

```

timelfh = (timelf-mod(timelf,3600))/3600;
timelf = timelf-timelfh*3600;
timelfm = (timelf-mod(timelf,60))/60;
timelf = timelf-timelfm*60;
%Mensagem utilizada para monitorar o andamento da operação
if(mod(j,10))==0
disp(strcat('NEURONS:',string(neurons)," ",string((j/tries)*100),'% complete'))
disp(strcat("Estimated time left: ",string(timelfh),"h:",string(timelfm),"m:",string(timelf),"s"));
end
end
%Ajustes em matrizes
roi1=roi1.';
roi2=roi2.';
roi3=roi3.';
%Montagem das matrizes finais com os dados desejados para cada estratégia
%de aposta
resmod1 = horzcat(roi1,p1,w1,tw1,cf1,ab1);
resmod2 = horzcat(roi2,p2,w2,tw2,cf2,ab2);
resmod3 = horzcat(roi3,p3,w3,tw3,cf3,ab3);
for temp = 1:18
    %Obtenção dos valores médios das grandezas analisadas dentre todas as
    %redes geradas
    resmod1(tries+1,temp)= nanmean(resmod1(1:tries,temp));
    resmod2(tries+1,temp)= nanmean(resmod2(1:tries,temp));
    resmod3(tries+1,temp)= nanmean(resmod3(1:tries,temp));

end
%Montagem das matrizes de saída com todos os dados de saída de interesse:
%número de neurônios utilizados, 10o percentil, 50o percentil (mediana),
%90o percentil e dados da matriz de confusão
resmodres(1,1) = neurons;
resmodres(1,2) = prctile(resmod1(1:tries,1),prct);
resmodres(1,3) = prctile(resmod1(1:tries,1),50);
resmodres(1,4) = prctile(resmod1(1:tries,1),100-prct);
resmodres(1,5:13) = resmod1(tries+1,9:17);
resmodres(1,14) = resmod1(tries+1,18);
resmodres(2,1) = neurons;
resmodres(2,2) = prctile(resmod2(1:tries,1),prct);
resmodres(2,3) = prctile(resmod2(1:tries,1),50);
resmodres(2,4) = prctile(resmod2(1:tries,1),100-prct);
resmodres(2,5:13) = resmod2(tries+1,9:17);
resmodres(2,14) = resmod2(tries+1,18);
resmodres(3,1) = neurons;
resmodres(3,2) = prctile(resmod3(1:tries,1),prct);
resmodres(3,3) = prctile(resmod3(1:tries,1),50);
resmodres(3,4) = prctile(resmod3(1:tries,1),100-prct);
resmodres(3,5:13) = resmod3(tries+1,9:17);
resmodres(3,14) = resmod3(tries+1,18);
end

```

O resultado esperado pela Análise 2 pode então ser obtido ao carregar as matrizes x, y, a, b e betodds, e rodar a linha de código representada abaixo:

```
[resmodres] = Analise2(x,y,a,b,betodds,10,50,1000,0.5,0.1)
```

3.4.3 Análise 3 – Estratégia de aposta BIM

Para entender como explorar a estratégia bet if minimum, o mesmo setup básico será rodado, variando a entrada mod2min entre 0,5 e 0,82, com incrementos de 0,04. Valores abaixo de 0,5 não são considerados pois garantem que o resultado escolhido é o mais provável, e valores acima de 0,82 não são considerados, pois as ocorrências de partidas com uma probabilidade de um resultado alto são muito pontuais e o resultado fica exageradamente aleatório.

3.4.4 Análise 4 – Estratégia de aposta BIV

Para entender como explorar a estratégia bet if value, o mesmo setup básico será rodado, variando a entrada mod3min entre 0,0 e 0,25, com incrementos de 0,025. Valores abaixo de 0,0 indicam que a rede neural acredita que aquela aposta não gera valor, e valores acima de 0,25 não são considerados, pois as ocorrências de partidas onde uma disparidade de valor tão grande são muito pontuais e o resultado fica exageradamente aleatório.

3.4.5 Análise 5 – Variação de resultado ano a ano

Para definir se o resultado da rede neural é confiável ano após ano, e conseqüentemente que ela pode ser usada para prever satisfatoriamente uma temporada futura, é necessário entender se os resultados obtidos em anos diferentes variam muito, e quais os fatores que motivam essa variação. Para cada ano, serão buscadas explicações factuais sobre aquela temporada para justificar os resultados obtidos.

De maneira prática, basta recortar a parte das matrizes a, b e betodds que representam o ano desejado.

3.4.6 Análise 6 – necessidade de dados de entrada retroativos

É importante entender a necessidade de uma grande quantidade de entradas retroativas, pois isso está diretamente relacionado a facilidade de enriquecimento do modelo. A riqueza do modelo está em levantar uma boa quantidade de entradas consideradas relevantes para o problema de temporadas passadas, o que pode se tornar mais complicado se forem necessários muitos anos retroativos, atingindo anos onde dados estatísticos não eram gerados com tamanha frequência e variedade.

Em termos práticos, as matrizes x e y serão recortadas de forma a cada vez retirar um ano de análise, começando pelos mais antigos. Assim, o modelo será rodado com as entradas básicas porém contendo 7, 6, 5, 4, 3, 2 e 1 anos de amostras de entrada.

3.4.7 Análise 7 – influência do número de neurônios na camada escondida

Os neurônios são as unidades funcionais no modelo e a quantidade de neurônios é fator determinante no resultado da rede. Como descrito na revisão bibliográfica, neurônios a menos geram perda de generalização, enquanto neurônios a mais podem tornar a rede especialista nos dados de treinamento, mas falha nos dados de teste.

Para determinar sua influência no resultado, o setup básico será rodado, variando a variável “neurons” entre 5 e 2000, com incrementos de 5 para valores entre 5 e 50, 10 para valores entre 50 e 200, 20 para valores entre 200 e 500, 50 para valores entre 500 e 1000 e por fim o valor de 2000. Um importante objeto de observação é a estratégia de apostas desenvolvida pela rede neural de acordo com o número de recursos disponíveis.

3.4.8 Análise 8 – direcionamento das variáveis de entrada mais importantes

O grande avanço do método de redes neurais para o método estatístico é que este último tem uma restrição da natureza das entradas. Como ele é baseado na previsão do número de gols a serem feitos por cada time, e não existem relações claras e equacionadas do número de gols com outras grandezas, as grandezas de

entrada só podem ter uma única natureza: serem relacionadas ao número de gols feitos. As redes neurais aceitam entradas de qualquer natureza no processo de classificação.

Em conjunto com o número de neurônios, os dados de entrada em cada amostra são recursos disponíveis para a rede neural, mas funcionam de forma inversa. Mais dados de entrada para menos neurônios na camada escondida geram uma sobrecarga dos últimos que não são capazes de atingir generalização satisfatória. Poucos dados de entrada para muitos neurônios pode gerar especialização.

Enquanto os neurônios são recursos internos possuídos pela rede, as entradas são os recursos que ela recebe como informação para trabalhar. Poucas entradas podem não ser suficientes para realizar uma boa previsão pelo simples fato de serem muito pouco responsáveis pelo resultado final. Uma situação análoga seria tentar escolher um sofá para sua sala de estar dentre 10 opções sabendo apenas a cor deles. Por mais que seja possível fazer a melhor escolha dentre as opções existentes, dificilmente será difícil garantir a satisfação de todos os habitantes da casa sem informações sobre a forma, conforto, tamanho, entre outros.

Como descrito na seção 3.4, espera-se um resultado muito aleatório. Nesse caso, espera-se ser possível oferecer entradas para a rede neural obter resultados bons o suficiente, o que provavelmente será longe de conseguir prever os resultados de uma partida com consistência. Não é necessário levantar todas as entradas existentes que afetam o resultado, mas uma quantidade suficiente para garantir esse retorno mínimo. Além disso, uma entrada só possui real valor quando possui influência em gerar melhores resultados. Caso esteja presente majoritariamente em resultados piores, a entrada pode estar sendo contra-produtiva, esgotando recursos da rede neural que poderiam estar sendo dedicados a entradas mais significativas e guiando o resultado para a direção incorreta.

Com isso em mente, as entradas precisam ser categorizadas entre melhores e piores. Isso será realizado com auxílio da pontuação dos setups de redes neurais. Todas as combinações de variáveis de entrada serão testadas, gerando setups com determinados resultados e pontuações. Cada setup gera pontos para todas as

variáveis de entrada que foram utilizadas para sua construção. Ao somar as pontuações obtidas por todas as variáveis de entrada, é possível entender quais variáveis estão relacionadas a melhores e piores resultados. Se dentre 10 redes, os 5 melhores resultados não possuem uma variável y , enquanto os 5 piores possuem a variável y , esta deve estar afetando a rede de alguma maneira negativa e deve ser descartada.

Entretanto, será necessário uma fase de direcionamento com pré-seleção dos grupos de variáveis mais promissores. Isso se dá pelo fato de que temos 34 variáveis de entrada mapeadas, que serão tomadas como 17 duplas de variáveis. O número de combinações de inclusão ou não de 17 duplas de variáveis na análise, utilizando pelo menos uma dupla, é de $2^{17} - 1$, ou 131071. Como cada setup gerado deve ser simulado pelo menos 100 vezes, levando em média 0,5s por simulação, seriam necessários 76 dias de simulação, tornando o processo proibitivo.

Para contornar esse problema, serão primeiramente estudados a influência dos grupos de variáveis. Com 7 grupos, o número de combinações geradas cai para 127, e o programa consegue obter respostas em horas.

Devido a dificuldade de encontrar dados de anos mais antigos, as 34 entradas só foram mapeadas para os anos de 2014-2017. Dessa maneira, serão utilizados 3 anos de treinamento e validação, e um ano para teste. Serão feitas 4 simulações, sendo cada uma com um ano de teste diferente.

É importante que os grupos de entrada com resultados positivos tenham resultados positivos ao longo de todos os anos analisados para que seja considerado consistente. Dessa maneira, a análise será feita separadamente ano a ano. Se os resultados forem diferentes a cada ano, esse método não é consistente, pois nunca será possível prever qual combinação de grupos de entrada será especialmente bom para um ano que ainda está ocorrendo.

Devemos ressaltar que essa análise possui alguns perigos: as variáveis dentro de um mesmo grupo de variáveis são sempre testadas juntas, de forma que caso uma variável do grupo seja benéfica a análise e outra seja maléfica, a variável benéfica ficará prejudicada. Entretanto, a análise será utilizada apenas para direcionamento, e as entradas serão testadas individualmente na Análise 9.

O código adaptado da Análise 2 para obter o resultado de todos os grupos de entrada está transcrito no quadro Código 2.

Código 2: Cálculo dos resultados obtidos para todas as combinações de grupos de entrada

```
function [resmodres] = Analise8(x,y,betodds,betamt,neurons,tries,prct,mininputs)

%Valores de entrada
yearinputcount = [209 209 207 208];
%Escolha dos grupos de entrada a serem analisados (1 = sim, 0 = não)
GOALS = 1;
FASE = 1;
OFF = 1;
DISC = 1;
ESTL = 1;
FIS = 1;
VAL = 1;
%Inicialização das combinações
inputs = [GOALS FASE OFF DISC ESTL FIS VAL];
actualinputs = 0;
inputindex = zeros(1,size(actualinputs,2));
%Determinação da posição dos grupos de entradas
count=1;
for temp = 1:size(inputs,2)
    if (inputs(temp)==1)
        inputindex(1,count) = temp;
        count=count+1;
    end
end
%Determinação da quantidade de grupos de entrada analisados
actualinputs = sum(inputs==1);
maxiter = zeros(1,actualinputs);
maxiter(1,1) = 1;
for temp = 2:actualinputs
    maxiter(1,temp) = maxiter(1,temp-1)*2+1;
end
%Geração de todas as combinações de grupos de entradas
inputgroupsall = sortrows(de2bi(0:maxiter(end)));
inputgroupsall(1,:) = [];
%Descartar todas as combinações que não possuem um mínimo de grupos de
%entradas
for temp = 1:size(inputgroupsall,1)
    if (sum (inputgroupsall(temp-(maxiter(end)-size(inputgroupsall,1)),:)) < mininputs)
        inputgroupsall(temp-(maxiter(end)-size(inputgroupsall,1)),:) = [];
    end
end
end
```

```

%Iniciação
iteration = 1;
inputgroups = inputgroupsall(iteration,:);
resmodmax(1:tries+1,1:8)=-1;
%Separar as matrizes entre treinamento e teste
off = find(yeartest(1,:) == max(yeartest(1,:)));
a = x(1+(sum(yearinputcount(1:off))-yearinputcount(off)):sum(yearinputcount(1:off)),:);
b = y(1+(sum(yearinputcount(1:off))-yearinputcount(off)):sum(yearinputcount(1:off)),:);
betodds = betodds(1+(sum(yearinputcount(1:off))-yearinputcount(off)):sum(yearinputcount(1:off)),:);
x(1+(sum(yearinputcount(1:off))-yearinputcount(off)):sum(yearinputcount(1:off)),:) = [];
y(1+(sum(yearinputcount(1:off))-yearinputcount(off)):sum(yearinputcount(1:off)),:) = [];
%Guardar as matrizes originais de entrada
origx = x;
origa = a;
tic;
while iteration <= length(inputgroupsall)
cashbet1=0;
cashwon1=0;
%Montar as matrizes de entrada para treinamento e teste baseado na combinação de grupos
%de entradas a ser testado
mati1=zeros(0,size(x,2));
mati2=zeros(0,size(a,2));
for i = 1:size(inputgroups,2)
    if inputgroups(i) == 1
        mati1 = horzcat(mati1,x(1:size(x,1),(1+sum(inputamt(1:i))-inputamt(1)):sum(inputamt(1:i))-inputamt(1)+inputamt(i))));
        mati2 = horzcat(mati2,a(1:size(a,1),(1+sum(inputamt(1:i))-inputamt(1)):sum(inputamt(1:i))-inputamt(1)+inputamt(i))));
    end
end
x=mati1;
a=mati2;
%Contabilização do resultado. A seguinte parte é semelhante à Analise2.
%Contabiliza apenas o resultado do método BOA.
p1=zeros(tries,3);
w1=zeros(tries,3);
tw1=zeros(tries,1);
roi1=zeros(tries,1);
ab1 = zeros(tries,9);
cf1 = zeros(tries,9);
betprob = 1./betodds;
for temp = 1:size(betodds,1)
    sumtemp = sum(betprob(temp,:));
    betprob(temp,1) = betprob(temp,1)/sumtemp;
    betprob(temp,2) = betprob(temp,2)/sumtemp;
    betprob(temp,3) = betprob(temp,3)/sumtemp;
end

for j = 1:tries

```

```

x=x.';
y=y.';
a=a;
b=b;
net = patternnet(neurons,'trainscg','crossentropy');
net = train(net,x,y);
view(net);
k = net(a.);
nnet.guis.closeAllWindows();
k=k.';
x=x.';
y=y.';

out = size(k,1);

cashbet1(j) = 0;
cashwon1(j) = 0;
for i = 1:out
    off = find(k(i,:) == max(k(i,:)));
    offr = find(b(i,:) == max(b(i,:)));
    knorm(i,:)=0;
    knorm(i,off)=1;

    bet(1)=0;
    bet(1)=betamt;
    cashbet1 = cashbet1+bet(1);
    p1(j,off) = p1(j,off)+1;

    if (isequal(off,offr))
        cashwon1(j) = cashwon1(j) + (betodds(i,off)-1)*bet(1);
        w1(j,off) = w1(j,off) + 1;
        cf1(j,1+4*(off-1)) = cf1(j,1+4*(off-1)) + 1;
        ab1(j,1) = ab1(j,1) + betodds(i,off)-1;
    else
        cashwon1(j) = cashwon1(j) - bet(1);
        cf1(j,3*(off-1)+offr) = cf1(j,3*(off-1)+offr) + 1;
    end
end
end
%Montagem do resultado
roi1(j)=cashwon1(j)/cashbet1(j);
ab1(j) = ab1(j)/sum(w1(j,1:3));
tw1(j) = (sum(w1(j,1:3)))/(sum(p1(j,1:3)));
for temp = 1:3
    w1(j,temp)=w1(j,temp)/p1(j,temp);
end
tempsum(1)=sum(p1(j,1:3));
for temp = 1:3
    p1(j,temp)=p1(j,temp)/tempsum(1);
end
end

```

```

disp(strcat(string(find(yeartest==1)+2013)," ", "NEURONS: ",string(neurons)," ", 'COMBINACAO ', " ",string(iteration)," ", 'DE
', " ",string(length(inputgroupsall)), " ", 'PROGRESSO NA ITERACAO ', " ",string((j/tries)*100), '%'))
end
resmod1 = horzcat(roi1,p1,w1,tw1,cf1,ab1);

for temp = 1:size(resmod1,2)
    resmod1(tries+1,temp) = mean(resmod1(1:tries,temp));
end
%Resultados obtidos pelo setup
resmodres(iteration,1) = neurons;
resmodres(iteration,2) = prctile(resmod1(1:tries,1),prct);
resmodres(iteration,3) = prctile(resmod1(1:tries,1),50);
resmodres(iteration,4) = prctile(resmod1(1:tries,1),100-prct);
resmodres(iteration,5:13) = resmod1(tries+1,9:17);
resmodres(iteration,14) = resmod1(tries+1,18);
inputshow = zeros(1,size(inputs,2));
count = 1;
for temp = 1:size(inputs,2)
    inputshow(inputindex(count)) = inputgroups(count);
    if(count<actualinputs)
        count=count+1;
    end
end
%Informações do setup utilizado para aquele resultado
resmodres(iteration,15:15+size(inputs,2)-1) = inputshow(1:size(inputs,2));
%Iniciação da próxima combinação de entradas
iteration = iteration + 1;
if (iteration<=length(inputgroupsall))
    inputgroups = inputgroupsall(iteration,:);
end
resmod1(tries+1,:) = [];
%Restauração das matrizes de entradas originais
x=origx;
a=origa;
%Temporizador
timelp = round(toc);
timelf = timelp*(length(inputgroupsall)/iteration-1);
timelfh = (timelf-mod(timelf,3600))/3600;
timelf = timelf-timelfh*3600;
timelfm = (timelf-mod(timelf,60))/60;
timelf = timelf-timelfm*60;
disp(strcat("Setup finished. Estimated time left: ",string(timelfh),"h:",string(timelfm),"m:",string(timelf),"s"));
end

%Posiciona os setups por ordem de resultado crescente após o fim da análise
[~,idx] = sort(resmodres(:,3));
resmodres = resmodres(idx,:);

end

```

3.4.9 Análise 9 – Influência individual de variáveis de entrada

Com uma pré-seleção, as variáveis de entrada dos grupos relacionados a melhores resultados na combinação dos 4 anos serão testadas individualmente.

Mais uma vez, o número de combinações pode ser proibitivo. Grupos de variáveis de entrada possuem 2 ou 3 duplas, e mesmo dispensando 3 grupos ainda é possível obter, por exemplo, 11 variáveis de interesse, o que ainda gera mais de um dia de análise. Uma condição será incluída para auxiliar nesse problema: o programa irá testar apenas combinações que utilizam um número mínimo de pares de variáveis. Agora, mesmo que seja necessário testar 11 pares de variáveis, mas com cada combinação contendo pelo menos 6 pares, existem um total de $2^{11} - \sum_{i=0}^5 \binom{11}{i} = 1024$, levando metade do tempo inicial. O código utilizado para fazer o teste da combinação de entradas individuais é análoga ao código utilizado na Análise 8, porém o parâmetro `mininputs` é ajustado para permitir uma duração aceitável de análise.

3.4.10 Análise 10 – Direcionamento estratégico das redes

Como os resultados de uma partida de futebol são resultados muito aleatórios, os resultados de todas as redes geradas para um mesmo setup podem variar drasticamente. Estas redes podem apresentar algumas características semelhantes de tal forma que talvez seja possível dizer com antecedência se esta é uma rede que modela bem os resultados de uma partida de futebol ou não. As redes neurais são estruturas obscuras que, a partir do processo de treinamento, geram valores de bias e pesos para seus neurônios e conexões de tal forma que não fica claro qual é a estratégia ou o caminho adotado para que ela atinja sua capacidade de generalização. É possível que a rede neural não consiga entender a dinâmica dos dados de treinamento o suficiente, e por isso geram redes tão diferentes entre si quando alimentadas com os mesmos dados. Ela está tentando fazer o seu melhor para entender a relação entre os dados, gerando redes que captam melhor a essência do esporte e redes que captam pior essa essência, sem saber a real diferença entre elas.

Dessa maneira, é possível que alguns dados de validação sejam capazes de filtrar redes neurais que captaram bem a essência do esporte. Uma rede gerada após treinamento deverá tentar prever os resultados nos anos de validação e, apenas caso

obtenha bons resultados, será liberada para prever o resultado dos jogos de teste. Espera-se que deixando passar apenas aquelas que possuem bons resultados nos dados de validação, os resultados nos testes melhorem. Esse processo é igual ao processo utilizado pela própria rede neural para fazer as paralizações por validação, que servem para parar o treinamento quando a rede já é considerada boa o suficiente, e então evitar treinamento excessivo. Entretanto, esses testes de validação por categorização não são exatamente o mesmo que testes de validação pelo resultado financeiro, que é o objetivo último dessa análise. O treinamento por validação feito pelo MATLAB tem o intuito apenas de garantir uma taxa de acerto nos jogos. Entretanto, uma maior taxa de acerto pode não garantir um melhor resultado financeiro, dependendo de quais partidas ela acerta ou erra, pois pode estar acertando mais em partidas com menos retorno.

Além dos anos de treinamento de 2007-2013, os anos de 2014-2016 serão utilizados para validação para aprovar as redes que finalmente irão prever os resultados de 2017, que serão então comparados ao resultado do ano de 2017 sem processo de direcionamento. Se o processo de validação for benéfico e consistente, ele deve gerar melhoria nos resultados em qualquer ano singular, caso contrário, o método não seria consistente.

Foi possível obter e organizar, com sucesso, 11 temporadas com as 4 entradas estatísticas, e as últimas 4 temporadas com todas as 34 entradas propostas.

4.2 Resultado das análises propostas

4.2.1 Eficiência do método estatístico

Os dados dos campeonatos de 2006 – 2016 foram levantados e utilizados para obter os resultados obtidos pelo método estatístico em apostas entre 2007 – 2017. Ao todo, foram testados 2639 resultados (11 campeonatos, 240 jogos por campeonato, exceto em 2016, quando o jogo Atlético-MG x Chapecoense foi cancelado devido ao triste acidente envolvendo o avião da Chapecoense). O diagrama de confusão obtido está apresentado Figura 11.

| | R1 | R2 | R3 | |
|----|--------|--------|--------|--------|
| T1 | 43.80% | 22.36% | 20.20% | 50.72% |
| T2 | 2.24% | 1.25% | 1.17% | 26.82% |
| T3 | 4.02% | 2.46% | 2.50% | 27.84% |
| | 87.50% | 4.79% | 10.47% | 47.55% |

Figura 11: diagrama de confusão obtido pelo método estatístico

Em geral, o modelo foi capaz de acertar os resultados em 47,56% dos casos. A Tabela 13 apresenta o diagrama de confusão condensado, mostrando a estratégia de distribuição de apostas e a taxa de acerto em cada uma.

Tabela 13: distribuição de apostas e taxa de acerto por tipo de resultado para o método estatístico

| Resultado | Proporção de apostas | Taxa de acerto nas apostas |
|----------------------|----------------------|----------------------------|
| Vitória do mandante | 83,4% | 50,7% |
| Empate | 4,7% | 26,8% |
| Vitória do visitante | 9,0% | 27,9% |

A taxa de acerto nos palpites de vitória do mandante é de 50,72%, alinhada com a porcentagem de vezes que um time de casa realmente ganha no campeonato brasileiro que, entre 2007 e 2017, foi de exatamente 50,0%. Isso pode significar que o método estatístico não apresenta uma estratégia vencedora de apostas, uma vez que o alinhamento entre a taxa de acerto dos jogos em casa e a taxa de ocorrência real destes pode significar que está fazendo apostas aleatórias e, então após um número grande suficiente de jogos, os resultados tendem a convergir para a média real das ocorrências.

O problema principal está nas taxas de acerto dos outros resultados: 26,82% para empates e 27,84% para vitória do visitante. O modelo parece extremamente enviesado para apostas em casa: um resultado que realmente teve vitória do mandante deve ter sido categorizado pela rede como vitória do mandante, já que o método exagera na quantidade dessas apostas, e por isso possui 87,51% de previsão desse resultado. Entretanto, quase metade desses jogos não se provam como vitórias do mandante, derrubando a taxa de acerto de V1. Em paralelo, diversos outros jogos que deveriam ser categorizados como empate ou vitória do visitante são categorizados como vitória do mandante, também gerando erros. A Tabela 14 mostra a taxa de previsão do modelo dos resultados que realmente ocorreram.

Tabela 14: taxa de ocorrência real e taxa de previsão do resultado para o método estatístico

| Resultado | Taxa de ocorrência dos resultados | Taxa de previsão média dos resultados |
|----------------------|-----------------------------------|---------------------------------------|
| Vitória do mandante | 50% | 87,5% |
| Empate | 26% | 4,8% |
| Vitória do visitante | 24% | 10,5% |

O diagrama de confusão e tabela de distribuição de apostas e taxa de acerto para os anos de 2014 à 2017 estão representados pela Figura 12 e Tabela 16. É importante isolar os dados destes anos para poder fazer uma comparação com os resultados da RNA, que irá gerar previsões apenas entre 2014 e 2017.

| | R1 | R2 | R3 | |
|----|--------|--------|--------|--------|
| T1 | 45.42% | 20.63% | 18.85% | 53.50% |
| T2 | 4.17% | 2.71% | 2.29% | 29.55% |
| T3 | 2.71% | 1.35% | 1.88% | 31.58% |
| | 86.85% | 10.97% | 8.14% | 50.00% |

Figura 12: diagrama de confusão obtido pelo método estatístico para os anos 2014-2017

Tabela 15: distribuição de apostas e taxa de acerto por tipo de resultado para o método estatístico, apenas para os anos de 2014-2017

| Resultado | Proporção de apostas | Taxa de acerto nas apostas |
|----------------------|----------------------|----------------------------|
| Vitória do mandante | 84,9% | 53,50% |
| Empate | 9,2% | 29,6% |
| Vitória do visitante | 5,9% | 31,6% |

A Tabela 16 faz um breakdown da distribuição de apostas e taxa de acerto por resultado para cada um dos anos de teste.

Tabela 16: distribuição de apostas, taxa de acerto por tipo de resultado e taxa de acerto para o método estatístico em cada ano entre 2014-2017

| Ano | Distribuição de apostas (V1/V2/V3, em %) | Taxa de acerto por resultado (WV1/WV2/WV3, em %) | Taxa de acerto geral (WT, em %) |
|------|--|--|---------------------------------|
| 2014 | 80,4/10,4/9,2 | 56,0/24,0/41,0 | 51,25% |
| 2015 | 83,8/13,7/2,5 | 58,4/27,3/16,7 | 53,1% |
| 2016 | 82,4/10,9/6,7 | 55,3/42,3/18,8 | 51,46% |
| 2017 | 92,9/1,7/5,4 | 45,3/0,0/38,5 | 44,2% |

Em destaque está a variabilidade dos resultados ao longo dos anos. Em geral, a taxa de acertos em casa é a maior e está um pouco acima da ocorrência real desse resultado para cada ano. Entretanto, para o resultado de empate, há taxas de acerto por volta de 40%, 25% e, no ano de 2017, a rede não conseguiu acertar nenhum empate. Quanto a vitória do visitante, alguns resultados ficaram por volta de 20% e outros por volta de 40%. A variabilidade de taxas de acerto é bem alta.

Uma simulação foi realizada para entender qual seria o retorno de investimento em apostas entre os anos de 2014 e 2017 utilizando-se o modelo estatístico. É realizada uma aposta de R\$10,00 por cada partida. Os resultados estão resumidos na Tabela 17.

Tabela 17: ROI obtido utilizando o método estatístico em cada ano enter 2014-2017

| Ano | Total apostado (R\$) | Lucros (R\$) | ROI |
|------|----------------------|--------------|--------|
| 2014 | 2.400 | 199,2 | 8,3% |
| 2015 | 2.400 | 165,60 | 6,9% |
| 2016 | 2.390 | 129,06 | 5,4% |
| 2017 | 2.400 | -300,00 | -12,5% |

O método estatístico possui um lucro combinado de R\$193,86 a partir de R\$9.590,00, atingindo um retorno de investimento (ROI) combinado de +2,02% ao longo dos quatro anos. Ela obtém resultados positivos para os anos de 2014, 2015 e 2016, mas apresenta resultados bem negativos para o ano de 2017, onde realmente teve uma taxa de acerto bem inferior. Este evento ocorre pois o ano de 2017 apresentou uma taxa de vitórias do time mandante anormal de apenas 44%, comparadas à 53% tanto em 2015 como 2016. Em contrapartida, esse foi o ano que o método estatístico mais priorizou a vitória dos mandantes, culminando em resultados financeiros ruins. A Tabela 18 apresenta a taxa de ocorrência de cada resultado entre os anos de 2007 e 2017.

Tabela 18: distribuição de ocorrência real de resultados nas temporadas 2007 - 2017

| Ano | Vitória do mandante | Empate | Vitória do visitante |
|-------|---------------------|--------|----------------------|
| 2007 | 51% | 24% | 26% |
| 2008 | 55% | 25% | 20% |
| 2009 | 51% | 27% | 22% |
| 2010 | 47% | 31% | 22% |
| 2011 | 48% | 28% | 24% |
| 2012 | 48% | 28% | 24% |
| 2013 | 48% | 28% | 23% |
| 2014 | 52% | 24% | 24% |
| 2015 | 53% | 24% | 23% |
| 2016 | 53% | 25% | 22% |
| 2017 | 44% | 27% | 29% |
| Total | 50% | 26% | 24% |

Dessa maneira, o método tende a ter resultados bons em anos como o de 2008, 2014, 2015 e 2016, enquanto tende a ter resultados ruins em anos como 2011, 2012, 2013 e especialmente 2017. Como aos anos de 2007-2012 tem uma taxa mais baixa de vitória do mandante (em média, 49,7%) do que os anos de 2013-2017 (50,5%), a taxa total de acerto considerando apenas 2014-2017 é superior a taxa de acerto considerando 2007-2017 – 50,5% contra 47,55%, mostrando a dependência deste método a ocorrência de mais ou menos resultados do tipo V1 na temporada.

4.2.2 Redes neurais com setup básico

O diagrama de confusão obtido pela estratégia Bet on all, que aposta em todos os jogos da amostra de teste, está apresentado na Figura 13.

| | R1 | R2 | R3 | |
|----|--------|--------|--------|--------|
| T1 | 37.45% | 16.11% | 15.79% | 54.00% |
| T2 | 9.71% | 5.73% | 5.05% | 27.96% |
| T3 | 5.19% | 2.77% | 2.21% | 21.71% |
| | 71.54% | 23.28% | 9.58% | 45.38% |

Figura 13: diagrama de confusão obtido pelas redes neurais com setup básico e estratégia de aposta bet on all

A taxa de acerto geral é de 45,4%, menor do que no método estatístico, de 50,0%. A Tabela 19 mostra de forma condensada a divisão entre apostas e a taxa de acerto em cada caso.

Tabela 19: distribuição de apostas e taxa de acerto por tipo de resultado para a rede neural com setup básico e estratégia de apostas bet on all

| Resultado | Proporção de apostas | Taxa de acerto nas apostas |
|----------------------|----------------------|----------------------------|
| Vitória do mandante | 69,3 | 54,0% |
| Empate | 20,5 | 28,0% |
| Vitória do visitante | 10,2 | 21,7% |

Nessa estratégia, percebemos que a rede neural também segue a estratégia de priorizar suas previsões na vitória do mandante, porém com menor intensidade: ela aposta na vitória do mandante 69,3% das vezes, contra 84,9% no método estatístico. A distribuição e taxa de acerto nas apostas são diferentes, e as principais diferenças são: a rede neural, em relação ao método estatístico, aposta aproximadamente 20% a

menos em V1, e dobra o número de apostas em V2 e V3. A taxa de acerto cai em todos os casos, especialmente no caso V3.

Em termos de resultado financeiro, a rede neural pode ainda haver um resultado melhor. Acertos em jogos de vitória do visitante tem um retorno maior do que na vitória do mandante. Esses efeitos podem se equilibrar e gerar um resultado financeiro maior. Entretanto, olharemos o desempenho de taxa de acerto da rede nas outras estratégias antes de analisar os retornos financeiros.

A Figura 14 apresenta o diagrama de confusão para a estratégia Bet if minimum:

| | R1 | R2 | R3 | |
|----|--------|--------|--------|--------|
| T1 | 42.10% | 17.38% | 16.80% | 55.19% |
| T2 | 7.79% | 4.94% | 4.22% | 29.17% |
| T3 | 3.33% | 2.00% | 1.43% | 21.09% |
| | 79.10% | 20.32% | 6.35% | 48.47% |

Figura 14: diagrama de confusão obtido pelas redes neurais com setup básico e estratégia de aposta bet if minimum

A taxa de acerto sobe de 45,4 para 48,5% em relação ao método BOA, e a rede aposta em apenas 429 dos 959 jogos disponíveis, pois apenas estes 44,7% atingem a probabilidade mínima do evento mais previsível de 50%, que é sua condição de aposta. Essa estratégia volta a intensificar as apostas na vitória do mandante, em 76,3 % dos casos, evitando apostar na vitória do visitante (6,8% dos casos). A Tabela 20 trás a proporção de apostas em cada resultado e a taxa de acerto dessas apostas.

Tabela 20: distribuição de apostas e taxa de acerto por tipo de resultado para a rede neural com setup básico e estratégia de apostas bet if minimum

| Resultado | Proporção de apostas | Taxa de acerto nas apostas |
|----------------------|----------------------|----------------------------|
| Vitória do mandante | 76,3% | 55,2% |
| Empate | 16,7% | 29,2% |
| Vitória do visitante | 6,8% | 21,1% |

Os 429 jogos apostados possuem um perfil ligeiramente diferente dos 959 jogos totais. Por mais que a diferença seja bem pequena, os mandantes realmente ganham mais vezes nesses jogos (53,3% contra 52,3% de todos), e os visitantes realmente ganham menos vezes nesses jogos (22,4% contra 23,0% de todos), o que pode indicar que a imposição de uma certeza mínima para aposta realmente afeta positivamente a taxa de acerto.

Por fim, a Figura 15 trás o diagrama de confusão para a estratégia Bet if value (BIV).

| | R1 | R2 | R3 | |
|----|--------|--------|--------|--------|
| T1 | 20.12% | 9.74% | 13.23% | 46.69% |
| T2 | 18.72% | 11.03% | 9.53% | 28.08% |
| T3 | 9.40% | 4.91% | 3.33% | 18.86% |
| | 41.70% | 42.94% | 12.75% | 34.47% |

Figura 15: diagrama de confusão obtido pelas redes neurais com setup básico e estratégia de aposta bet if value

A taxa de acerto cai substancialmente, de 45,4% (BOA) e 48,5% (BIM) para 34,5%. Isso porque a rede neural encontra valor, ou seja, acredita que um evento na verdade tem pelo menos 10% a mais chance de ocorrer do que o que dizem as

cotações do site de aposta, no caso de empates e vitórias do visitante. Agora, ela passa a apostar na vitória do mandante em menos da metade das ocasiões. A proporção de apostas e taxa de acerto por tipo de resultado está apresentado na Tabela 21.

Tabela 21: distribuição de apostas e taxa de acerto por tipo de resultado para a rede neural com setup básico e estratégia de apostas bet if value

| Resultado | Proporção de apostas (%) | Taxa de acerto nas apostas (%) |
|----------------------|--------------------------|--------------------------------|
| Vitória do mandante | 43,1 | 46,7 |
| Empate | 39,3 | 28,1 |
| Vitória do visitante | 18,9 | 18,9 |

Valor é encontrado pela rede em apenas 451 dos 959 jogos (47,0%). A taxa de acerto em todos os casos cai: para V1, é de 46,7% (contra 54,0% BOA e 55,2% BIM), para V2 é de 28,1% (contra 28,0% BOA e 29,2% BIM) e para V3 é de 18,9% (contra 21,7% BOA e 21,1% BIM). Mais uma vez, isso não quer dizer que essa estratégia sempre terá um resultado pior que as duas: ela pode ser capaz de prever vitórias do visitante mais improváveis do que o comum, receber um maior payout nessas instâncias e superar o efeito negativo da menor taxa de acerto.

Todas as estratégias de aposta possuem uma taxa de acerto menor do que o modelo estatístico. O resultado financeiro obtido por cada estratégia de aposta está mostrada na **Error! Reference source not found.** Cada barra azul representa a faixa de resultados de ROI que compreendem os 80% resultados intermediários obtidos por todas as redes rodadas. O traço preto centralizado nas faixas azuis representam o ROI mediano, destacado nos textos brancos. A linha vertical vermelha representa o ponto de breakeven (ROI = 0,0%) e a linha vertical verde representa o retorno obtido pelo método estatístico.

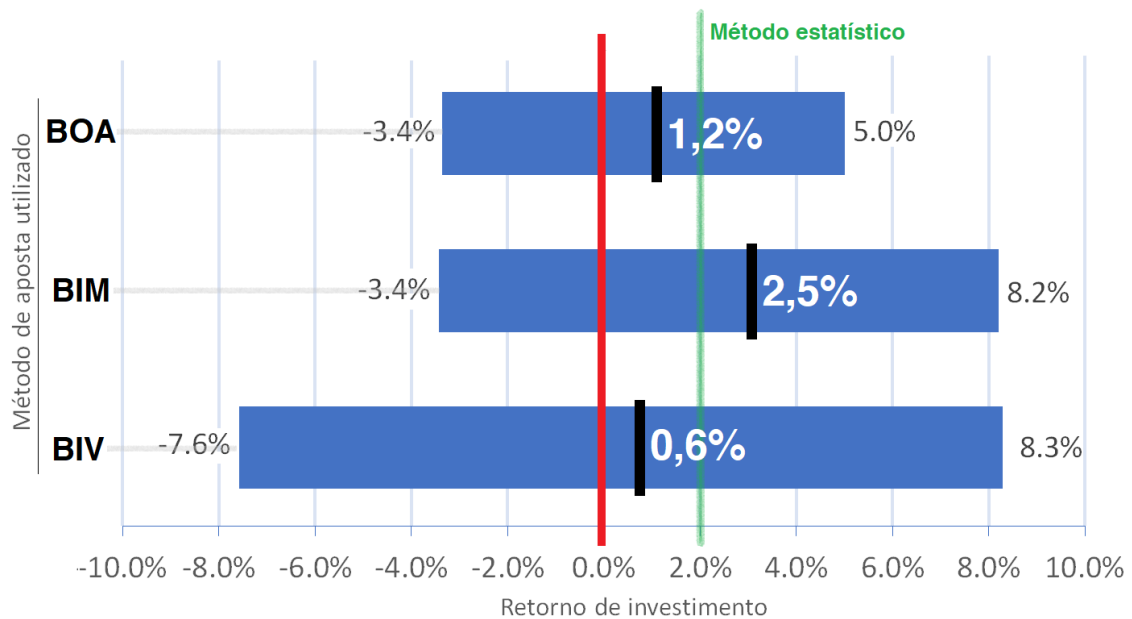


Figura 16: faixas de ROI obtidas pelas redes neurais com setup básico nas três estratégias de aposta para os anos teste de 2014-2017

Nesse momento, percebemos a grande aleatoriedade envolvida no problema. No caso da aposta simples BOA, das 1 000 redes geradas, as 800 redes intermediárias tem um resultado que varia entre -3,4% e +5,0%, o que pode ser a diferença entre um método lucrativo e um método que gera prejuízo. Entretanto, é possível que, depois de todos os ajustes testados na próxima análise, nós cheguemos em um setup que gera lucros consistentemente. Apenas o método BIM conseguiu ter um resultado mediano superior ao gerado pelo método estatístico. Os resultados positivos podem chegar à aproximadamente 8,0%, o que indicaria um ótimo investimento anual. Entretanto, a alta variabilidade presente seria um fator crucial para a não utilização do método.

Para o setup básico e os anos de 2014-2017, a melhor estratégia de aposta é a BIM, seguida pela BOA e então BIV. Entretanto, é importante lembrar que isso não é uma regra, e que não necessariamente a ordem dos métodos por resultado será a mesma.

Em relação a estratégia básica bet on all, a estratégia bet if minimum apresenta um PMV menor (1,11 contra 1,29), pois tende a apostar mais em eventos V1. Entretanto, sua maior taxa de acerto culmina em um melhor resultado geral. A estratégia bet if value possui uma taxa de acerto bem menor do que as outras, mas o

ROI médio em caso de vitória pula para 1,92, fazendo seu resultado mediano não ficar tão para trás das outras, mesmo com uma taxa de acerto geral tão baixa.

A Tabela 22 trás um grande compilado comparativo entre as taxa de acerto geral, perfil de distribuições de apostas, taxas de acerto por resultado, taxa de ocorrência real, taxa de previsão do modelo e retorno de investimento para as três estratégias de aposta, bem como para o método estatístico (apenas anos 2014-2017). A linha retorno de investimento apresenta o ROI mediano no caso das redes neurais e ROI único para o caso do método estatístico; as linhas de 10º percentil e 90º percentil só se aplicam as redes neurais.

Tabela 22: resumo geral da taxa de acerto geral, distribuição de apostas, taxa de acerto por tipo de aposta, taxa de ocorrência real, taxa de previsão de resultado, payout médio em caso de vitória e retorno de investimento para as três estratégias de aposta e método estatístico

| Grandeza | BOA | BIM | BIV | Método estatístico |
|-----------------------------|------|------|------|--------------------|
| Taxa de acerto geral (%) | 45,4 | 48,5 | 34,5 | 50,0 |
| Apostas em V1 (%) | 69,3 | 76,3 | 43,1 | 84,9 |
| Apostas em V2 (%) | 20,5 | 16,7 | 39,3 | 9,2 |
| Apostas em V3 (%) | 10,2 | 6,8 | 18,9 | 5,9 |
| Acerto em V1 (%) | 54,0 | 55,2 | 46,7 | 53,5 |
| Acerto em V2 (%) | 28,0 | 29,2 | 28,1 | 29,6 |
| Acerto em V3 (%) | 21,7 | 21,1 | 18,9 | 31,6 |
| Ocorrência real V1 (%) | 52,3 | 53,3 | 48,2 | |
| Ocorrência real V2 (%) | 24,7 | 24,3 | 25,7 | 24,7 |
| Ocorrência real V3 (%) | 23,0 | 22,4 | 26,1 | 23,0 |
| Taxa de previsão V1 (%) | 71,5 | 79,1 | 41,7 | 86,9 |
| Taxa de previsão V2 (%) | 23,3 | 20,3 | 42,9 | 11,0 |
| Taxa de previsão V3 (%) | 9,6 | 6,4 | 12,8 | 8,1 |
| PMV | 1,29 | 1,11 | 1,92 | 1,04 |
| ROI 10º percentil (%) | -3,4 | -3,4 | -7,6 | - |
| Retorno de investimento (%) | +1,2 | +2,5 | +0,6 | +2,0 |
| ROI 90º percentil (%) | +5,0 | +8,2 | +8,3 | - |

4.2.3 Estratégia de aposta BIM

A Tabela 23 trás os resultados do setup básico com a variação do parâmetro mod2min entre 0,50 e 0,82 em incrementos de 0,04.

Tabela 23: número de jogos apostados, distribuição de apostas e taxas de acerto para a estratégia de apostas bet if minimum em função do valor de mod2min

| Mod2min | N | V1/V2/V3 (%) | WV1/WV2/WV 3 (%) | WT (%) |
|---------|--------|-----------------|------------------------|-----------|
| 0,50 | 424,57 | 76,4/16,8/6,8 | 55,4/29,4/21,1 | 48,7 |
| 0,54 | 285,92 | 76,7/16,7/6,6 | 55,6/29,0/20,8 | 48,9 |
| 0,58 | 191,83 | 75,3/18,1/6,7 | 56,1/28,6/21,5 | 48,8 |
| 0,62 | 127,22 | 73,0/20,0/7,0 | 55,9/28,5/18,2 | 47,8 |
| 0,66 | 85,50 | 69,3/22,5/8,2 | 55,3/28,8/18,7 | 46,3 |
| 0,70 | 53,70 | 62,6/28,9/8,6 | 55,1/28,9/18,4 | 44,4 |
| 0,74 | 35,01 | 59,7/29,7/10,6 | 56,0/30,0/15,1 | 40,0 |
| 0,78 | 16,91 | 52,0/36,2/11,8 | 56,1/29,6/19,0 | 42,1 |
| 0,82 | 11,63 | 54,1/40,6/5,3 | 52,9/28,8/12,9 | 40,0 |

Uma melhor visualização da variação desses números pode ser visualizada na

Figura 17.

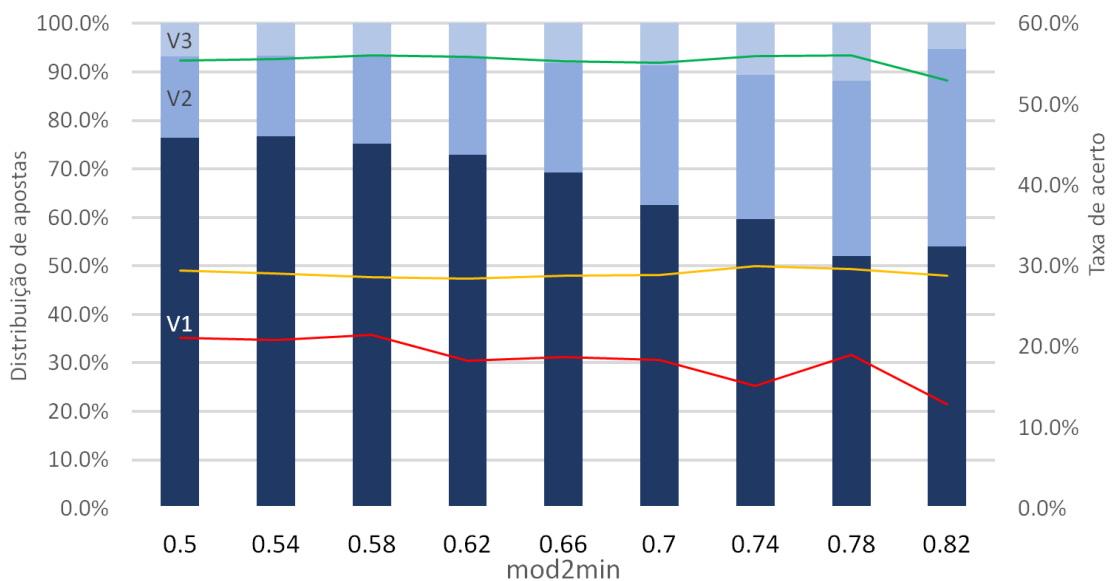


Figura 17: distribuição de apostas e taxa de acerto por tipo de resultado com setup básico e variação do parâmetro mod2min nos anos teste de 2014-2017

O comportamento na rede com o aumento no parâmetro é de selecionar cada vez menos jogos para apostar, uma vez que só aposta em jogos que possui muita certeza. Ela julga que um resultado tem mais de chance de 50% de ocorrer em 44% dos casos, mais de 58% em 20% dos casos e mais de 70% de ocorrer em apenas 5,5%. Mesmo que o campeonato brasileiro seja imprevisível, essa não parece ser uma certeza suficiente.

Entretanto, a rede não parece estar convergindo para a direção correta, uma vez que quando aposta com mais certeza, diminui sua taxa total de acerto. Ela pode não estar selecionando jogos que realmente consegue prever melhor, mas ela pode estar tendo mais certeza de resultados errados. É como se ela estivesse se tornando alguém de opinião mais forte, porém não necessariamente correta, e está sendo penalizada por isso. A taxa de acerto em apostas de cada tipo não muda significativamente, com apenas um decréscimo notável na taxa de acerto em vitórias do visitante. Não alterando a taxa de acerto por tipo de resultado e apostando mais em resultados menos prováveis, a taxa de acerto total cai.

A Figura 18 mostra os resultados de retorno de investimento obtidos.

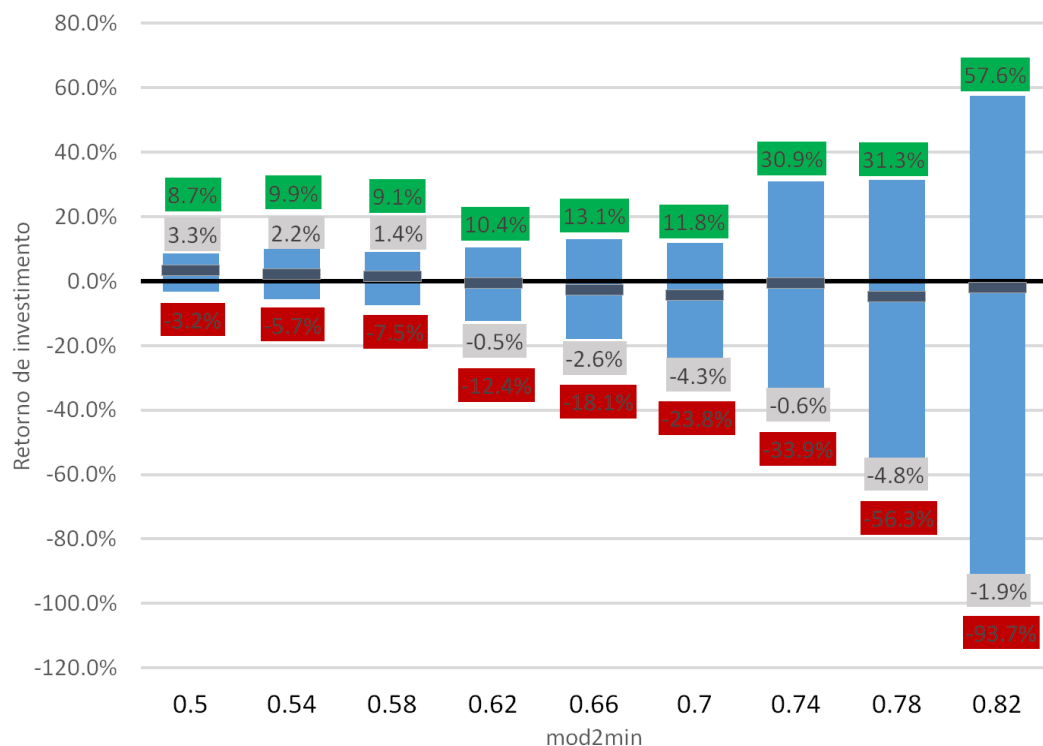


Figura 18: faixas de ROI obtidas pelas redes neurais com setup básico e variação do parâmetro mod2min nos anos teste de 2014-2017 com o método bet if minimum

O aumento do parâmetro mod2min faz a rede apostar apenas com mais certeza. Isso se traduz em resultados medianos que variam pouco em torno de 0,0%, mas que ficam cada vez mais incertos. Isso pode se dar ao fato de que as redes que obtiverem um ajuste correto e apostar com confiança tendem a obter resultados bem melhores do que as que se ajustam incorretamente e apostam com falsa confiança.

4.2.4 Estratégia de aposta BIV

A Tabela 24 trás os resultados do setup básico com a variação do parâmetro mod3min entre 0,0 e 0,25 em incrementos de 0,025.

Tabela 24: número de jogos apostados, distribuição de apostas, taxas de acerto de apostas para a estratégia de apostas bet if value em função do valor de mod3min.

| Mod3min | N | V1/V2/V3 (%) | WV1/WV2/WV3 (%) | WT (%) |
|---------|-------|-----------------|--------------------|-----------|
| 0,0% | 668.7 | 57,0/29,7/13,3 | 48,5/27,8/21,1 | 38,7 |
| 2,5% | 626.0 | 53,3/33,3/13,4 | 48,1/28,0/21,0 | 37,8 |
| 5,0% | 569.0 | 50,3/33,5/16,2 | 47,4/27,8/20,5 | 36,4 |
| 7,5% | 512.6 | 46,9/36,9/16,1 | 46,9/27,9/19,6 | 35,5 |
| 10,0% | 448.7 | 43,1/39,2/17,7 | 47,0/28,0/19,2 | 34,6 |
| 12,5% | 378.9 | 41,2/41,2/17,7 | 46,8/27,8/18,2 | 33,9 |
| 15,0% | 323.4 | 37,5/41,6/20,9 | 46,5/27,8/17,7 | 32,7 |
| 17,5% | 260.0 | 36,8/41,9/21,3 | 46,1/28,5/17,2 | 32,6 |
| 20,0% | 212.7 | 32,2/42,6/25,3 | 45,4/28,4/16,8 | 30,9 |
| 22,5% | 168.2 | 31,5/43,9/24,6 | 45,0/28,6/16,2 | 30,8 |
| 25,0% | 135.0 | 30,0/41,8/28,2 | 44,0/28,8/16,3 | 29,8 |

Uma melhor visualização da variação desses números é apresentada na Figura 19.

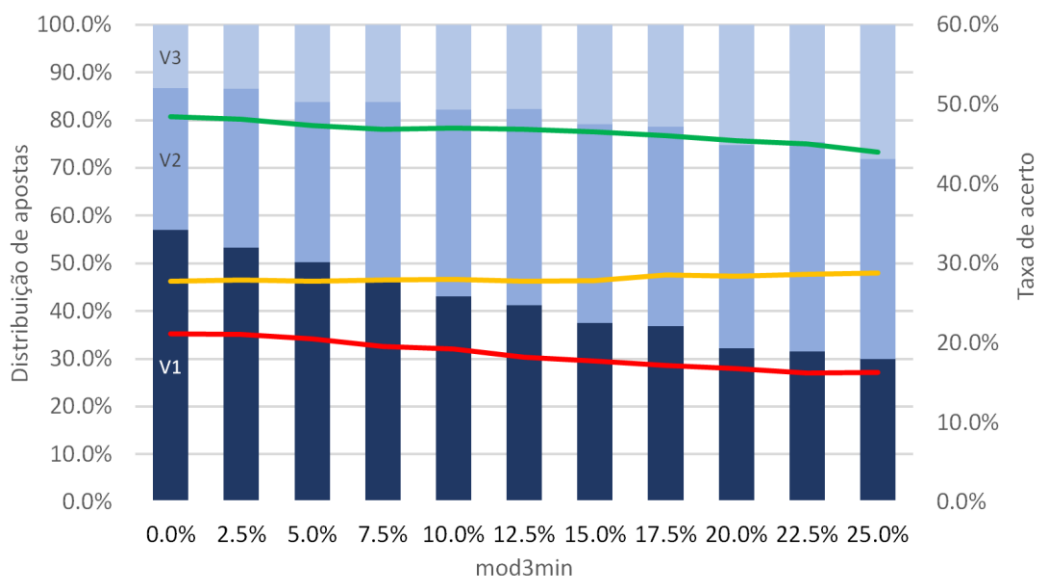


Figura 19: distribuição de apostas e taxa de acerto por tipo de resultado com setup básico e variação do parâmetro mod3min nos anos teste de 2014-2017

Para um aumento de mod3min, a rede neural também encontra cada vez menos casos para apostar. Isso porque casos onde a rede neural conclui que existe tanta discrepância entre seu resultado e as previsões dos payouts da casa de aposta vão diminuindo. Ela discorda da probabilidade do site de aposta em mais de 5% em 62% dos casos, mais de 10% em 47% dos casos e mais de 20% em 14% dos casos. Estes resultados não são condizentes, uma vez que valor pode sim ser encontrado em apostas, mas a discordância entre a previsão da rede e a previsão de todos os usuários do site de apostas não pode ser tão discrepante a ponto de concordar no erro superior a 10 pontos percentuais em quase metade das vezes.

Neste caso, a rede neural entende que existe mais valor em empates e vitórias do visitante. Isso faz sentido, uma vez que o resultado mais esperado pelas casas de aposta é a vitória do mandante, e a rede neural pode estar chegando a conclusão de que o jogo não será bem assim. O mandante pode sim tender a ter vantagem, mas a rede neural acredita que essa vantagem é menor do que o valor de payout da casa de aposta diz. As taxas de acerto mantem-se quase constantes ou com leve queda. A taxa de acerto geral cai de 38,7% para 29,8% com a variação de mod3min de 0% até 25%, mas o payout médio em caso de vitória aumenta consideravelmente, de 1,60 para 2,30.

A variação do resultado financeiro está apresentada na Figura 20.

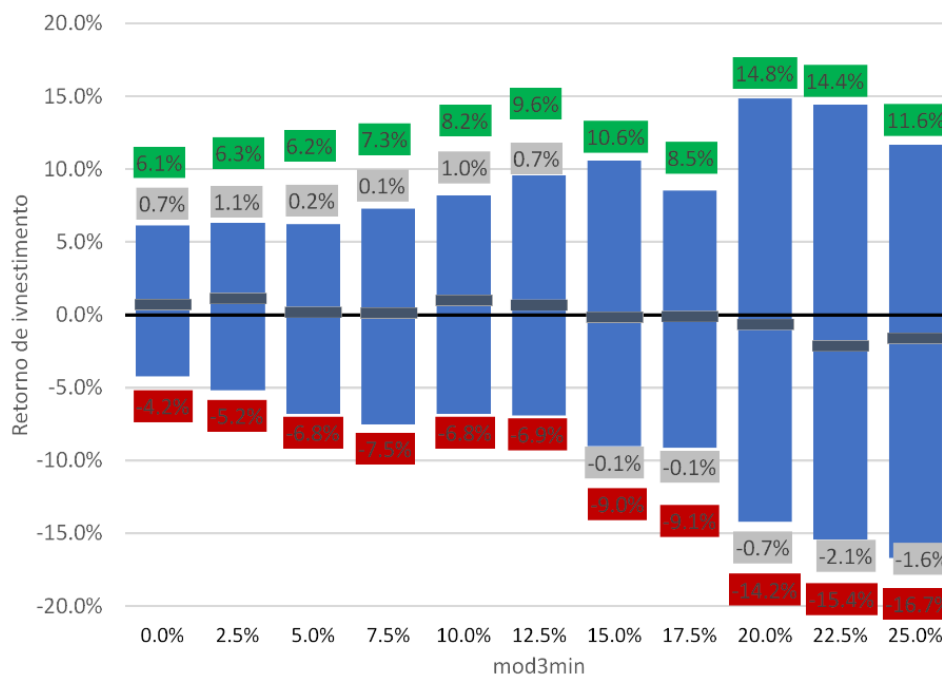


Figura 20: faixas de ROI obtidas pelas redes neurais com setup básico nas três estratégias de aposta para os anos teste de 2014-2017 com o método bet if value

Igualmente ao aumento do parâmetro mod2min, a rede neural tende a manter o ROI mediano enquanto aumenta a variabilidade entre os valores, porém em graus menores. Ocorre uma tendência de queda no ROI mediano para valores muito elevados. Isso também mostra que a rede pode estar se tornando superconfiante no fato de que está conseguindo realizar previsões melhores do que as probabilidades das casas de apostas dizem, mas isso está intensificando seu resultado tanto para o lado positivo quanto para o negativo. Em geral, o método BIV apresenta maior incerteza.

4.2.5 Variação de resultado ano a ano

A Figura 21 mostra de forma visual a distribuição das apostas e taxa de acerto por tipo de resultado em função dos anos utilizados como treinamento, para o caso bet on all. As linhas representam WV1, WV2, WV3 nas cores verde, laranja e vermelha. Por fim, as linhas em pontilhado representam a taxa real de ocorrência do resultado, obedecendo ao esquema de cores das outras linhas.

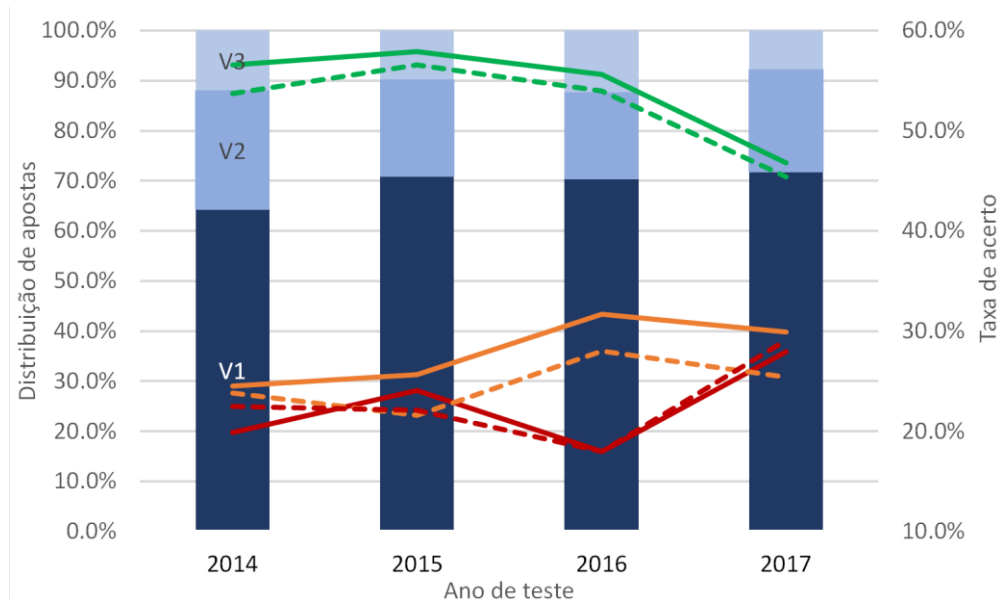


Figura 21: distribuição de apostas e taxa de acerto por tipo de resultado com setup básico para cada um dos anos teste entre 2014-2017, para o caso bet on all

Esta análise deixa clara que a rede neural não está guiando suas escolhas, mas sim está fazendo escolhas de forma altamente aleatória de modo que suas taxas de acerto final acabam seguindo a tendência das taxas reais de ocorrência. A rede neural consegue acertar um pouco mais do que na seleção totalmente aleatória, já que as linhas preenchidas estão em geral deslocadas para cima das linhas pontilhadas da mesma cor, exceto em 2015, quando acaba possuindo a tendência inversa de acertar a vitória do visitante em relação ao que realmente ocorreu.

A Figura 22 mostra os valores de retorno de investimento para cada ano e para cada estratégia de aposta. Os valores acima e abaixo das faixas de ROI são os valores dos limites superior e inferior baseados no 90º e 10º percentis, e o valor acima da barra centralizada é o valor do ROI mediano. Os círculos vermelhos que demarcam alguns anos significam que aquela estratégia utilizada é que possui melhor resultado para o ano demarcado.

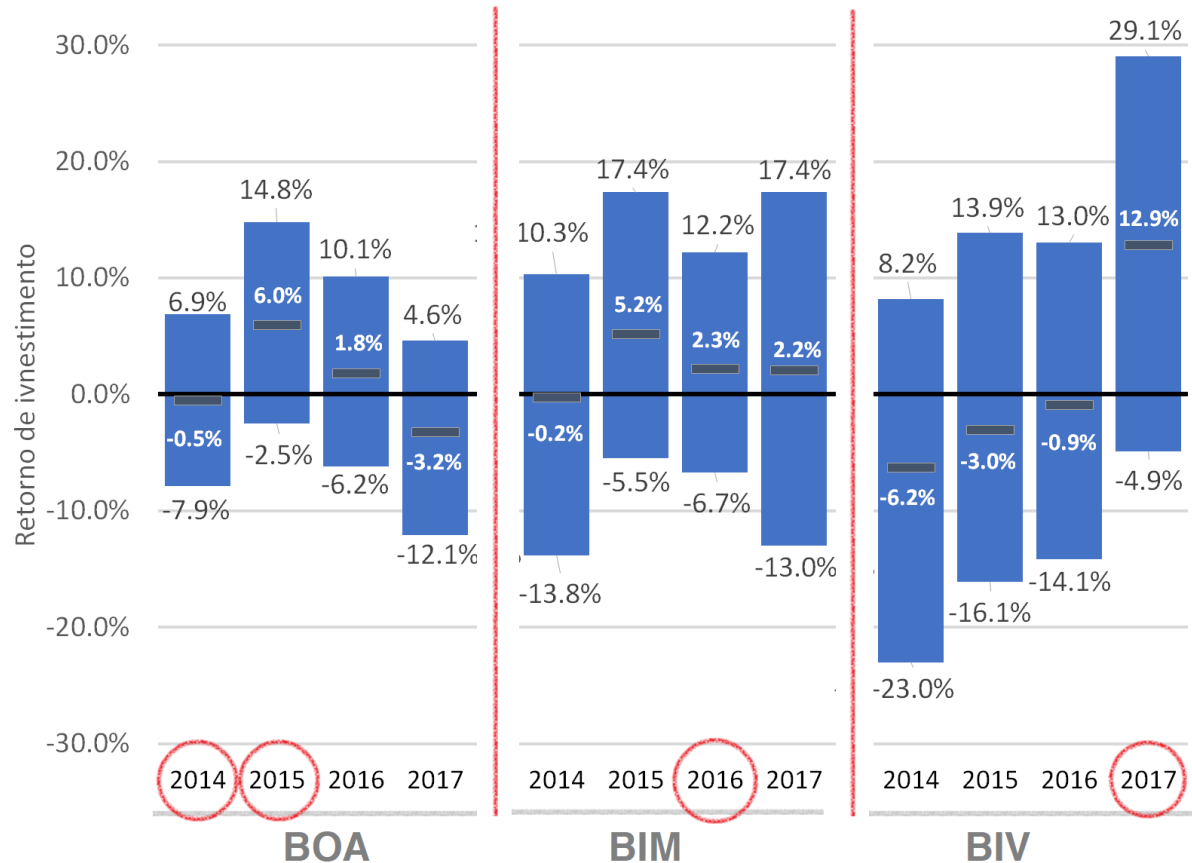


Figura 22: faixas de ROI obtidas pelas redes neurais com setup básico em cada ano teste entre 2014-2017 para cada método de aposta

Como a rede neural tende a apostar mais em vitórias do mandante, era de se esperar que tivesse um melhor resultado nos anos onde esse resultado realmente ocorreu. Analisando o caso básico bet on all, essa observação se confirma parcialmente. As medianas dos resultados em cada ano realmente estão na mesma ordem da ordem de taxas de ocorrência de vitória do mandante: 2015, 2016, 2014 e 2017, sendo que 2015 tem um valor consideravelmente maior (56%) e 2017 tem um valor consideravelmente menor (45%). Os resultados do método BOA em 2015 e 2016 são positivos, indicando que um pequeno deslocamento da taxa de acerto em um tipo de resultado em relação a curva de taxa de ocorrência real (Figura 21) já melhora bastante os resultados financeiros.

A estratégia bet on all obtém resultado positivo em 2 casos e resultado negativo em 2 casos, com uma variação mediana e é a melhor estratégia dentre as estudadas para ter apostado em 2014 e 2015. A estratégia bet if minimum possui 3 resultados

medianos positivos e um negativo, em 2014. Estranhamente, o pior resultado não foi obtido no ano de 2017, que possui ocorrências contrárias a tendência da rede neural. A Tabela 25 apresenta a alocação de apostas e taxa de acerto para o caso BIM.

Tabela 25: jogos apostados, distribuição de apostas, taxas de acerto e payout médio em caso de vitória pela estratégia bet if minimum em cada ano de teste entre 2014-2017

| Ano | N | Distribuição de apostas (%) | Taxa de acerto de apostas | Taxa de acerto total | PMV |
|------|-------|-----------------------------|---------------------------|----------------------|------|
| 2014 | 98,7 | 65,9/22,8/11,3 | 57,4/26,9/17,1 | 45,9 | 1,16 |
| 2015 | 116,6 | 76,5/18,0/5,4 | 58,2/24,5/30,3 | 50,6 | 1,09 |
| 2016 | 117,3 | 81,0/11,6/7,4 | 55,1/33,3/15,9 | 49,7 | 1,06 |
| 2017 | 97,5 | 81,2/14,4/4,4 | 50,7/33,7/28,1 | 47,3 | 1,15 |

Não parece haver um padrão que justifique esse acontecimento. O método apresenta um resultado positivo em para 2017, o ano que menos teve vitórias do mandante, mesmo sendo o ano em que mais aposta no resultado vitória do mandante. Entretanto, tem uma boa combinação de acerto em empate e derrota suficiente para obter uma combinação Taxa de acerto e PMV com resultados positivos. O ano de 2014 tem resultados negativos por possuir um PMV bom em comparação e 2017, mas uma taxa de acerto insuficiente.

Por fim, o método bet if value apresenta resultados negativos e com alta variação entre 2014-2016, mas, novamente de forma surpreendente, possui um resultado extremamente elevado para o ano de 2017, podendo chegar a um resultado máximo quase recorde de 29,1%, enquanto o resultado mínimo não é altamente negativo, de -4,9%. Como normalmente esse método faz, ele evita apostas na vitória do mandante e prioriza em empates, seguido de vitória do visitante, por encontrar mais valor nesses casos. A relação de taxa de acerto e payout médio por ano é apresentada na Tabela 26.

Tabela 26: relação entre taxa de acerto total e payout médio em caso de vitória para cada ano teste entre 2014-2017

| Ano | Taxa de acerto total | PMV |
|------|----------------------|------|
| 2014 | 30,2% | 2,06 |
| 2015 | 34,2% | 1,88 |
| 2016 | 36,4% | 1,73 |
| 2017 | 37,3% | 2,01 |

Em 2014, um PMV de 2,06 não compensa uma WT de 30,2%. Entretanto, no caso de 2017, uma taxa de acerto de 37,3% com um PMV de 2,01 gera excelentes resultados. Dado a ocorrência desse resultado inesperado, foi feita uma análise do retorno de investimento da estratégia BIV para 3 valores de mod3min, que está apresentada na Tabela 27.

Tabela 27: jogos apostados, taxas de vitória, payout médio em caso de vitória e ROI obtidos pela estratégia mod3min em 2017

| Mod3min | N | WT | PMV | ROI 10° | ROI M | ROI 90° |
|---------|-----|-------|------|---------|-------|---------|
| 0,0 | 161 | 38,9% | 1,68 | -7,5% | 3,4% | 16,9% |
| 0,1 | 107 | 37,3% | 2,01 | -4,9% | 12,9% | 29,1% |
| 0,2 | 41 | 22,7% | 2,32 | -15,8% | 16,4% | 50,4% |

Não só a estratégia é a melhor a ser utilizada no ano com a menor ocorrência de vitórias do mandante, mas ela possui gera melhores resultados conforme o seu parâmetro de intensidade aumenta. A rede fica mais confiante que existe valor naquela aposta e acaba tendo melhores retornos como consequência. O lado negativo é que como a rede realiza apostas em média em apenas 41 dos 240 jogos, a variabilidade do resultado é alta.

A Figura 23 trás uma comparação dos valores de retorno de investimento dos três métodos para cada ano e para o método estatístico. Para as redes neurais, os valores de ROI são apresentados como a mediana.

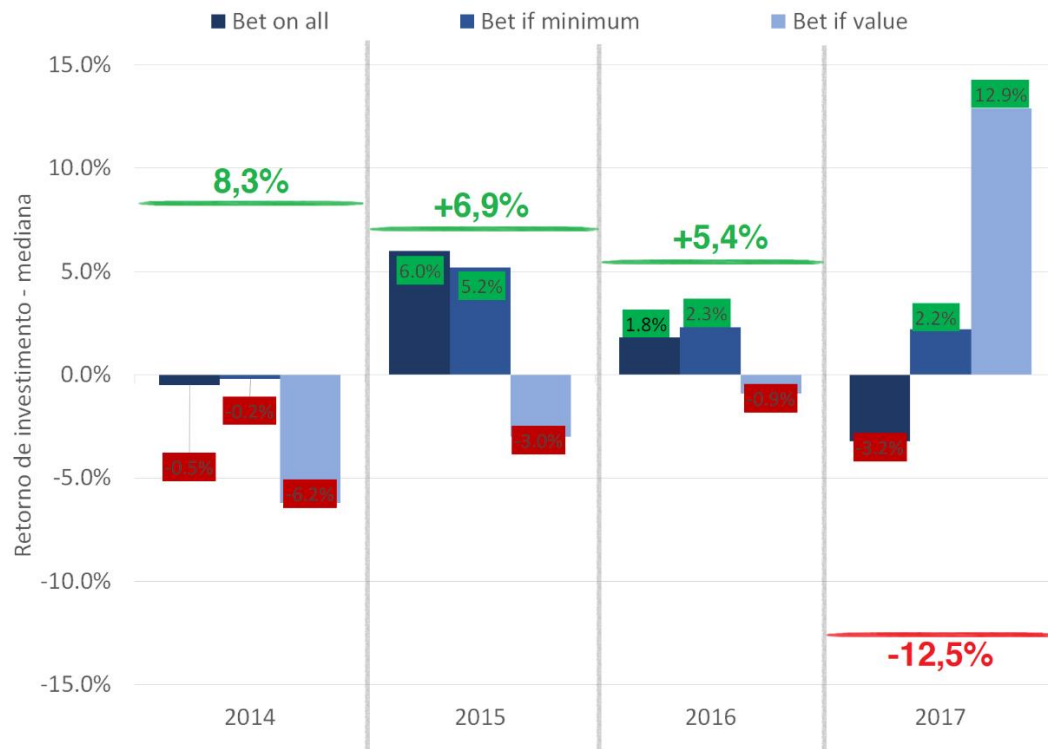


Figura 23: ROI mediano obtido pelas redes neurais com cada método de aposta e para cada ano entre 2014-2017 e valor obtido pelo método estatístico

O método estatístico leva a melhores resultados e sem variabilidade para os anos 2014-2016, e a um resultado pior em 2017. Entretanto, vale ressaltar que, como veremos, a rede neural com outros ajuste apresenta resultados melhores, e que as redes neurais possuem uma faixa de valores que em todos os casos podem superar o método estatístico.

4.2.6 Necessidade de dados de entrada retroativos

A Figura 24 mostra de forma visual a distribuição das apostas e taxa de acerto por tipo de resultado em função dos anos utilizados como treinamento. Começa-se utilizando 7 anos, entre 2007-2013, que fornecem 240 entradas por ano, gerando 1.680 entradas. Então, são utilizados cada vez um ano a menos, descartando anos mais distantes.

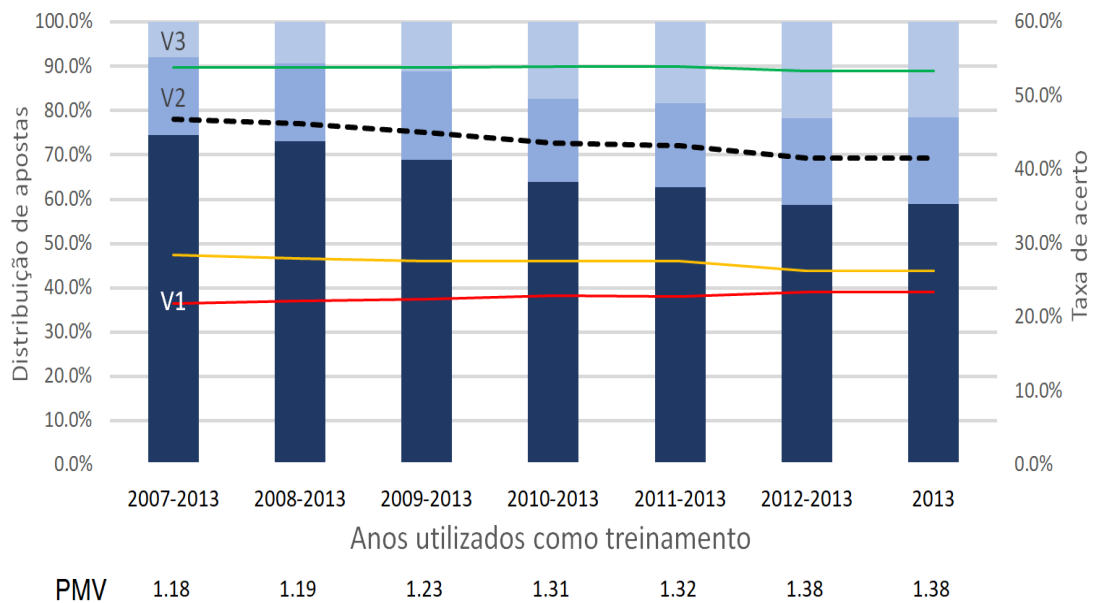


Figura 24: distribuição de apostas e taxa de acerto por tipo de resultado com setup básico para diferentes quantidades de anos utilizados como treinamento, para o caso bet on all, e valores de PMV médio

Quanto menos resultados retroativos utilizados, mostrando insistentemente que o resultado V1 é muito comum, menos enviesada a rede neural fica em apostar em V1. Com os 7 anos de entrada, ela aposta em V1 74,5% das vezes, V2 17,4% das vezes e V3 8,0%. Com apenas o ano de 2013 como entrada, a rede neural tende a apostar mais próximo da distribuição que realmente ocorreu no ano de treinamento: aposta em V1 58,9% das vezes, V2 19,5% das vezes e V3 21,6% das vezes, sendo que a taxa de ocorrência nos dados de treinamento real foi de 46,7%, 30% e 23,3%. As taxas de acerto não sofrem mudanças significativas: WV1 é constante, WV2 possui uma queda de 28,5% para 26,3% e WV3 sobe de 21,9% para 23,5%. Mais uma vez, a rede consegue distribuir os resultados mais próximo do que realmente ocorre, mas sem conseguir selecionar bem qual qual partida colocar em cada categoria. Ocorre um tradeoff entre uma queda de WT com aumento de PMV.

A Figura 25 mostra os resultados de retorno de investimento ao se variar o número de anos que geram amostras de entrada para a rede neural.

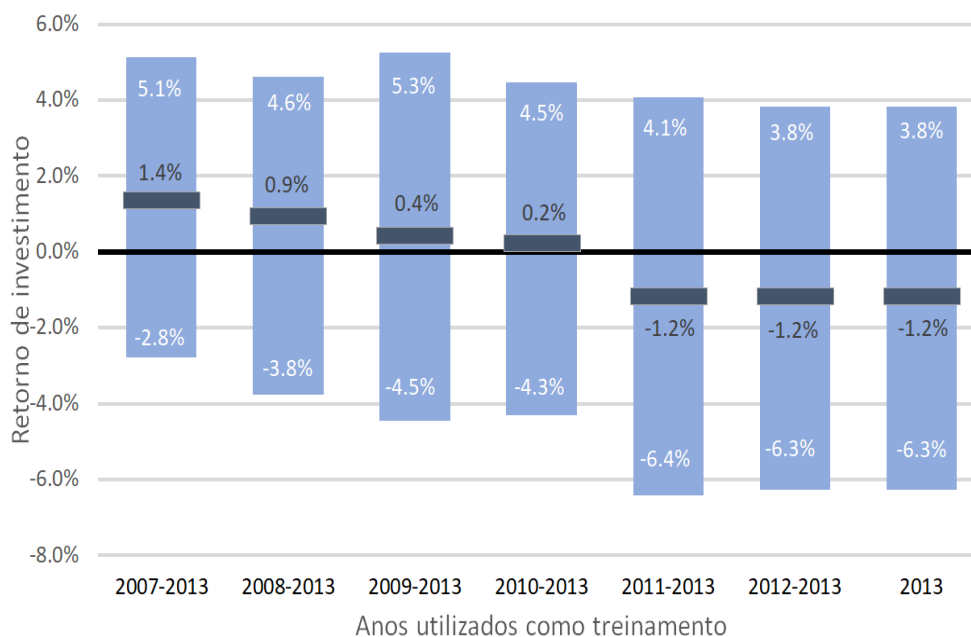


Figura 25: faixas de ROI obtidas pelas redes neurais com setup básico e diferentes quantidades de anos utilizados como amostras de treinamento

Realmente, a utilização de mais amostras melhora o resultado médio das redes neurais, tanto em valor mediano quanto em confiabilidade. Por mais que a distribuição das apostas fique mais condizente com a realidade, o ganho de payout médio em caso de vitória não é suficiente para compensar a queda na taxa de acerto total. Isso indica que a utilização de mais dados anteriores adicionam a análise, e para cada nova entrada pensada como importante para análise, deve ser possível levantá-la por pelo menos 4 anos anteriores. Como o resultado continua melhorando significativamente com 4, 5, 6 e 7 anos utilizados, é possível realizar uma análise com as 4 entradas estatísticas para campeonatos bem antigos, de forma a entender se o resultado continua a aumentar – o que pode ser feito numa análise futura.

4.2.7 Influência do número de neurônios na camada escondida

Ao se variar o número de neurônios na camada escondida observa-se que a rede neural tende a mudar sua estratégia de apostas, apostando cada vez menos na vitória do visitante (V1) e cada vez mais no empate (V2) e na vitória do mandante (V3), conforme mostrado na Figura 26. A rede neural tende a se arriscar mais em resultados

menos prováveis com a presença de mais neurônios. No gráfico, conforme descrito na metodologia, o incremento do número de neurônios vai ficando maior quanto mais neurônios estão sendo analisados, e as linhas em cinza separam zonas onde o incremento do número de neurônios é diferente, o que significa que o gráfico não está em escala no eixo x. O incremento está sinalizado na parte superior da cada faixa. O eixo y representa a proporção acumulada das previsões em cada caso, somando 100%.

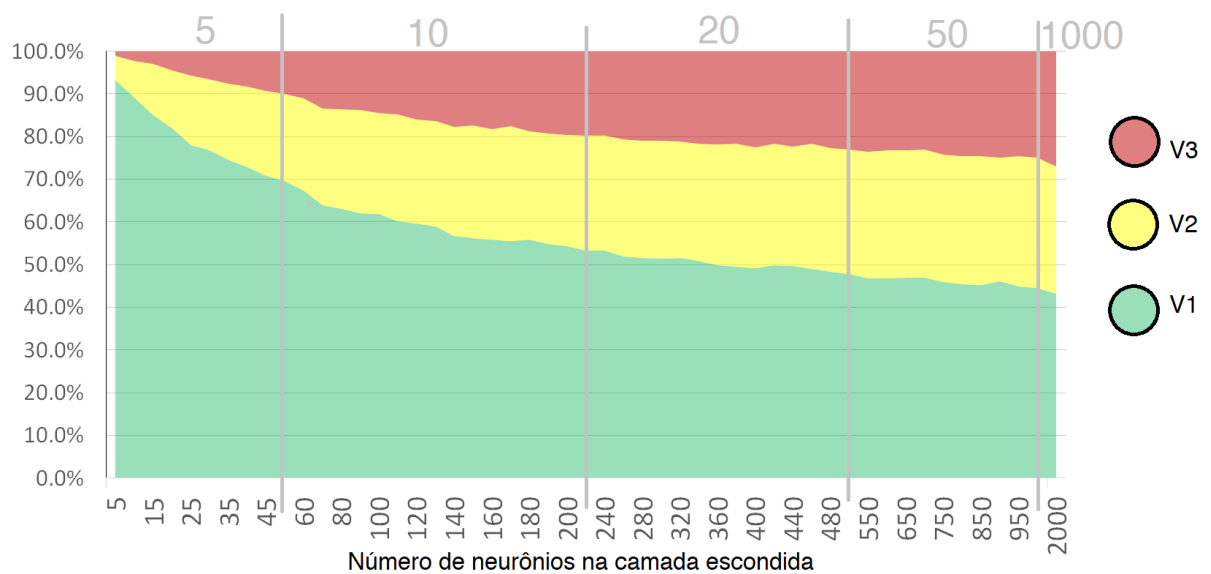


Figura 26: distribuição de apostas com setup básico para um número de neurônios na camada escondida entre 5 e 2000 para o caso bet on all (escala variável)

Com poucos neurônios, a rede tende a simplesmente ser simplista e apostar na vitória do mandante quase em sua totalidade, pois a rede percebe pelas amostras que, na dúvida e com as informações limitadas que apenas 5 neurônios conseguem absorver, deve apostar no mandante. Com 5 neurônios, ela aposta em V1 93,3% das vezes; com 50 neurônios, esse número cai abaixo de 70%, então para 60% com 110 neurônios, 50% para 360 neurônios e até 44,5% quando existem 1000 neurônios. Alguns pontos de interesse estão apresentados na Tabela 28.

Tabela 28: proporção de apostas em função do número de neurônios na camada escondida para alguns pontos de interesse

| Número de neurônios na camada escondida | Proporção V1 | Proporção V2 | Proporção V3 |
|---|--------------|--------------|--------------|
| 10 | 93.3% | 5.6% | 1.1% |
| 25 | 78.0% | 16.3% | 5.7% |
| 50 | 69.6% | 20.4% | 10.0% |
| 100 | 61.9% | 23.7% | 14.5% |
| 150 | 56.2% | 26.5% | 17.3% |
| 200 | 54.3% | 26.2% | 19.5% |
| 300 | 51.5% | 27.6% | 20.9% |
| 400 | 49.1% | 28.3% | 22.6% |
| 500 | 47.7% | 29.3% | 23.0% |
| 600 | 46.8% | 30.1% | 23.2% |
| 700 | 47.0% | 30.0% | 23.0% |
| 800 | 45.5% | 30.0% | 24.5% |
| 900 | 46.1% | 28.9% | 24.9% |
| 1000 | 44.5% | 30.6% | 24.9% |
| 2000 | 43.2% | 29.9% | 26.9% |

Ao apresentar mais neurônios, a rede fica mais próxima da distribuição histórica real entre os resultados (por volta de 50% para V1 e 25% para V2 e V3). Entretanto, existe um problema. O aumento do número de neurônios faz com que a rede aposte em uma quantidade mais real de V2 e V3, mas ela não consegue aumentar sua taxa de acerto nesses eventos. Na verdade, a taxa de acerto não tem incremento significativo em nenhum caso. A rede aprende a ser menos enviesada quanto a apostar apenas no mandante, mas não consegue selecionar melhor em quais casos deveria fazer isso ou não. A média de acertos em V1 gira em torno de 53%, para V2 de 26% e para V3 em torno de 23%, o que é extremamente alinhado com os valores das ocorrências em média no campeonato brasileiro. A rede neural está acertando numa taxa baseada na estatística, como se apenas se dedicasse a entender como

distribuir as apostas, porém o faz de forma aleatória, e as taxas de acerto tendem a ser exatamente as taxas de ocorrência dos resultados ou um pouco acima. A Tabela 29 mostra a taxa de acerto em cada cenário para os mesmos pontos de interesse.

Tabela 29: taxas de acerto em função do número de neurônios na camada escondida

| Número de neurônios na camada escondida | Taxa de acerto em V1 | Taxa de acerto em V2 | Taxa de acerto em V3 | Taxa de acerto total |
|---|----------------------|----------------------|----------------------|----------------------|
| 10 | 52,8% | 28,3% | 23,8% | 51,1% |
| 25 | 53,8% | 28,5% | 21,3% | 47,8% |
| 50 | 54,1% | 28,0% | 22,7% | 45,6% |
| 100 | 53,6% | 26,4% | 22,1% | 42,6% |
| 150 | 53,8% | 25,8% | 22,5% | 40,9% |
| 200 | 53,9% | 25,7% | 22,7% | 40,4% |
| 300 | 53,5% | 25,2% | 22,4% | 39,2% |
| 400 | 53,3% | 24,9% | 22,5% | 38,3% |
| 500 | 53,1% | 25,0% | 22,0% | 37,7% |
| 600 | 53,3% | 25,2% | 22,1% | 37,6% |
| 700 | 53,5% | 25,2% | 22,0% | 37,8% |
| 800 | 53,0% | 24,5% | 21,9% | 36,8% |
| 900 | 53,3% | 25,2% | 22,0% | 37,4% |
| 1000 | 52,9% | 24,9% | 21,9% | 36,6% |
| 2000 | 53,4% | 25,4% | 22,7% | 36,8% |

Na verdade, a taxa de acerto em caso de empate tende a cair com o aumento do número de neurônios, enquanto as taxas em V1 e V3 são estáveis. Ao se arriscar mais, a taxa de acerto total cai de aproximadamente 1 acerto a cada 2 jogos para aproximadamente 1 acerto a cada 3 jogos.

Ao se analisar o retorno de investimento em cada caso, existem duas forças se equilibrando: a rede neural acerta uma quantidade absoluta menor de jogos,

entretanto os jogos que ela acerta possuem um maior payout, tendo em vista que tendem a ser mais improváveis (V2 e V3). A Figura 27 mostra como a taxa de acerto (eixo à esquerda) e o payout médio por aposta vencida (eixo à direita) variam com o número de neurônios na camada escondida.

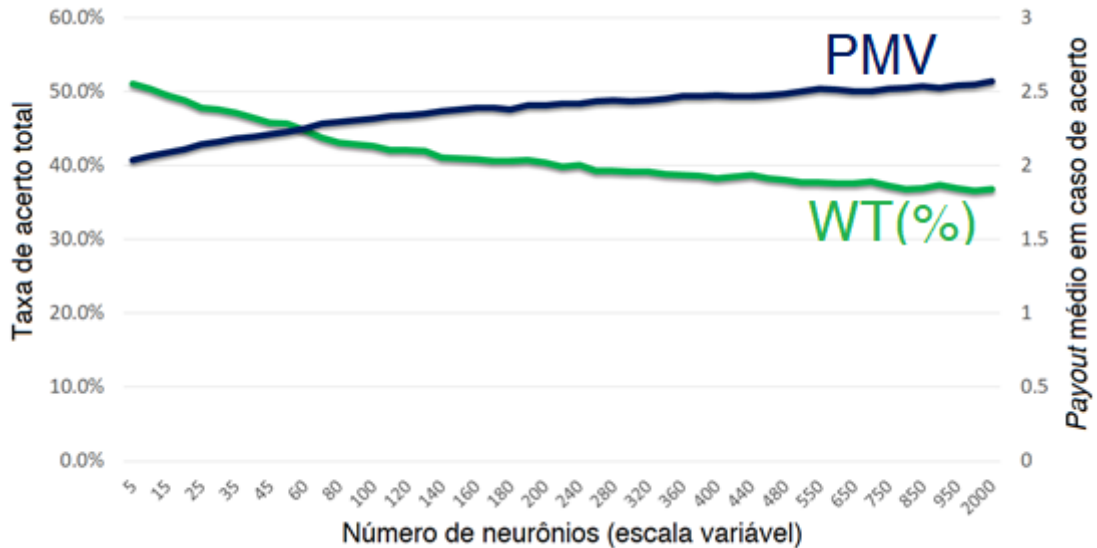


Figura 27: taxa de acerto total e PMV com setup básico para um número de neuronios na camada escondida entre 5 e 2000 para o caso bet on all (escala variável)

Entretanto, ao analisar os resultados de retorno de investimento, conclui-se que o aumento do retorno médio em caso de acerto tem magnitude maior do que a queda de taxa de acerto, gerando resultados cada vez piores. As Figuras 28, 29 e 30 mostram as faixas contendo os retornos de investimento obtidos por cada setup de número de neurônios na camada escondida para as estratégias de aposta bet on all, bet if minimum e bet if value, respectivamente. O limite inferior representa o 10º percentil e o limite superior representa o 90º, enquanto as barras horizontais representam o valor mediano.

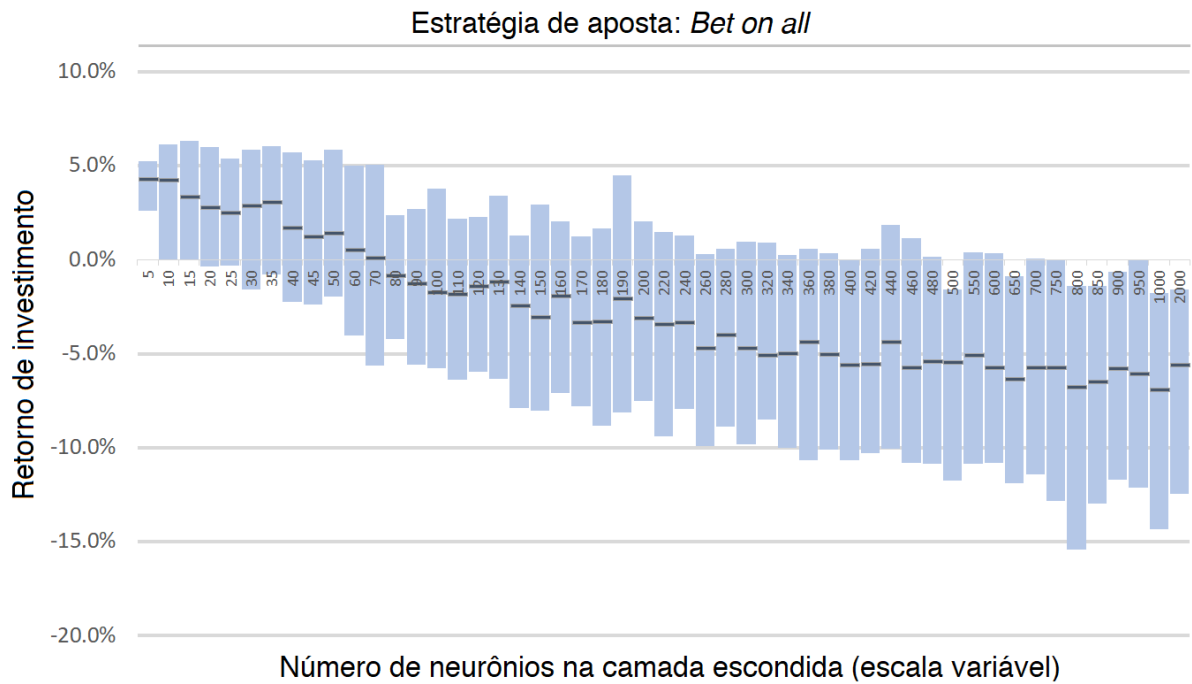


Figura 28: faixas de ROI obtidas pelas redes neurais com setup básico e diferentes quantidades de neurônios na camada escondida para o método de aposta bet on all

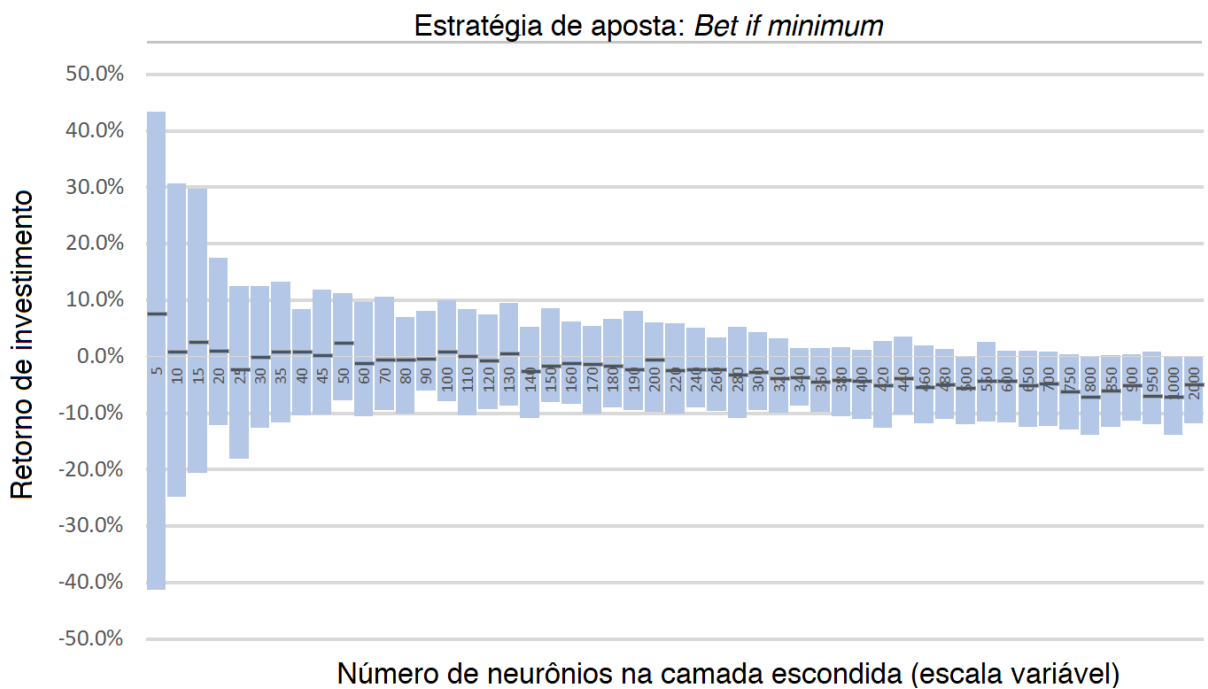


Figura 29: faixas de ROI obtidas pelas redes neurais com setup básico e diferentes quantidades de neurônios na camada escondida para o método de aposta bet if minimum

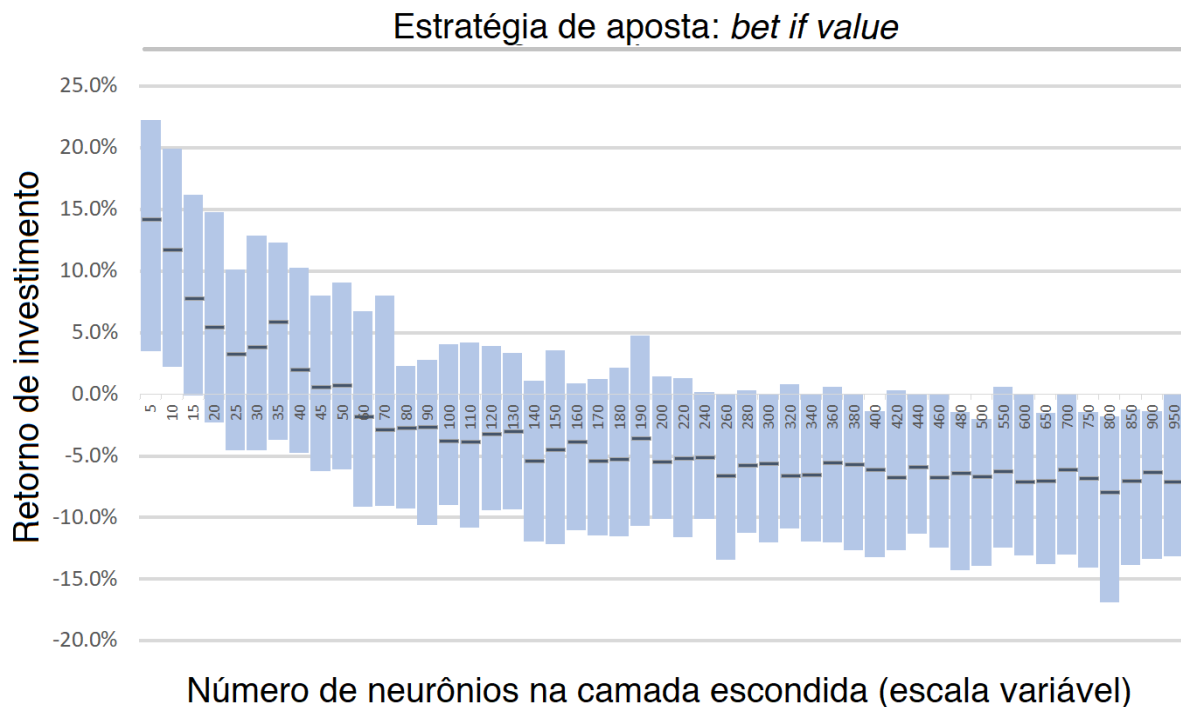


Figura 30: faixas de ROI obtidas pelas redes neurais com setup básico e diferentes quantidades de neurônios na camada escondida para o método de aposta *bet if value*

O aumento do número de neurônios tem efeitos tanto semelhantes quanto diferentes em cada caso. Em todos os casos, o resultado não melhora. Para o caso BOA, a influência negativa é clara: o resultado piora tanto em termos absolutos e fica mais incerto. Resultados com a mediana positiva ocorrem 5 e 70 neurônios, caem até por volta de -5% até 300 neurônios, e então ficam com a mediana estável, mas a incerteza cresce ligeiramente.

No caso BIM, os resultados são interessantes e bem distintos do primeiro caso. Após alguns setups iniciais com uma variação imprevisível, um maior número de neurônios não piora significativamente o resultado mediano, apenas diminui drasticamente sua incerteza. Com 5 neurônios, o resultado das redes varia entre -40% e 44%, com uma mediana positiva de aproximadamente 8%, enquanto com 100 neurônios, varia entre -7% até 10%. Dessa maneira, é possível obter resultados incríveis com a rede, mas também resultados bem negativos. Como a mediana está deslocada para a parte superior, a média é mais próxima ao limite superior, porém acredita-se que a maioria dos apostadores não confiaria nessa técnica. Com um

aumento do número de neurônios, a mediana cai e fica abaixo de 0% já com 30 neurônios.

No caso BIV, a rede possui resultados ótimos com poucos neurônios. Até 50 unidades funcionais, os resultados tem mediana positivas e podem chegar a até 25% de retorno de investimento. Porém, mais uma vez, a partir de 50 neurônios, os resultados ficam majoritariamente negativos.

Nos três casos, o aumento do número de neurônios causa uma piora do resultado geral. Esta análise aponta que, sempre que possível, deve se utilizar o mínimo de neurônios possíveis na análise. Entretanto, esta estratégia não será seguida durante este trabalho. Utilizará-se número de neurônios limitados para evitar resultados cada vez piores, mas com um limite mínimo. Isso porque utilizar um número de neurônios tendendo a 0 significa se aproximar da estratégia de apostar sempre no mandante. Essa estratégia pode entregar retornos positivos em diversos campeonatos, mas o objetivo não é empregar redes neurais para tender a uma estratégia tão simplista. Um número de neurônios de pelo menos 50 será utilizado nas análises, pois existe a hipótese de que um maior número de neurônios poderá melhorar a análise quando forem introduzidas mais entradas no modelo.

4.2.8 Direcionamento das variáveis de entrada mais importantes

Espera-se que, caso exista uma dinâmica escondida que relacionam os 7 diferentes grupos de entrada (média de gols, aproveitamento, ofensividade, disciplina, estilo de jogo, características físicas do plantel e aspectos financeiros) com diferentes resultados, os mesmos grupos de entradas serão benéficos para os resultados de todos os anos. Caso um grupo de entradas seja especialmente benéfico em alguns anos mas totalmente improdutivo em outros, entende-se que novamente a rede não consegue deduzir uma dinâmica existente nesse fenômeno aleatório, e realiza decisões majoritariamente aleatórias. Para melhorar a rede neural, procuram-se dados de entrada que consigam gerar uma melhoria consistente, mesmo que pequena, nos resultados da rede.

A correlação entre grupo de entradas e siglas apresentada na Tabela 30 é utilizada para identificar os grupos, descritos de forma sucinta.

Tabela 30: relação entre siglas utilizadas e os grupos de variáveis com uma descrição e número de entradas contidas em cada grupo

| Sigla | Nome do grupo de variáveis | Descrição geral | N de entradas |
|-------|--|---|---------------|
| GOLS | Variáveis de entrada estatística | Média de gols feitos e sofridos dentro/fora de casa na temporada passada | 4 |
| FASE | Variáveis de entrada de fase atual | Aproveitamentos dos times no campeonato, últimas rodadas e dentro/fora de casa na temporada passada | 6 |
| OFNS | Variáveis de entrada de ofensividade | Média de chutes e chutes a gol por jogo na temporada passada | 4 |
| DISC | Variáveis de entrada de disciplina | Média de faltas e cartões por jogo na temporada passada | 6 |
| ESTL | Variáveis de entrada de estilo de jogo | Média de posse de bola e acerto de passes por jogo na temporada passada | 4 |
| FISI | Variáveis de entrada de características físicas do plantel | Média de altura, peso e idade do plantel | 6 |
| FINA | Variáveis de entrada de aspectos financeiros | Valor de mercado do time e receita do último ano | 4 |

As Figuras 31 e 32 apresentam o resultado comparativo entre a quantidade de pontos obtida por cada grupo de entradas, conforme descrito no item 3.3. Os resultados são apresentados de forma individual para cada ano, e então em forma sobreposta para comparação dos formatos em geral. Como os valores são normalizados, o gráfico tem melhor efeito comparativo e não apresenta valor nos

eixos: a linha externa representa o maior valor e a linha interna representa o menor valor. Como os valores para cada grupo de entradas nos anos individuais são definidos pela comparação com valores de outras entradas apenas nesse ano, e no caso do gráfico de grupos de entradas combinados de todos os anos os valores são comparados com os valores de todos os anos, os gráficos mudarão de formato entre as duas modalidades de apresentação, porém mantendo a ordem de importância de cada grupo.

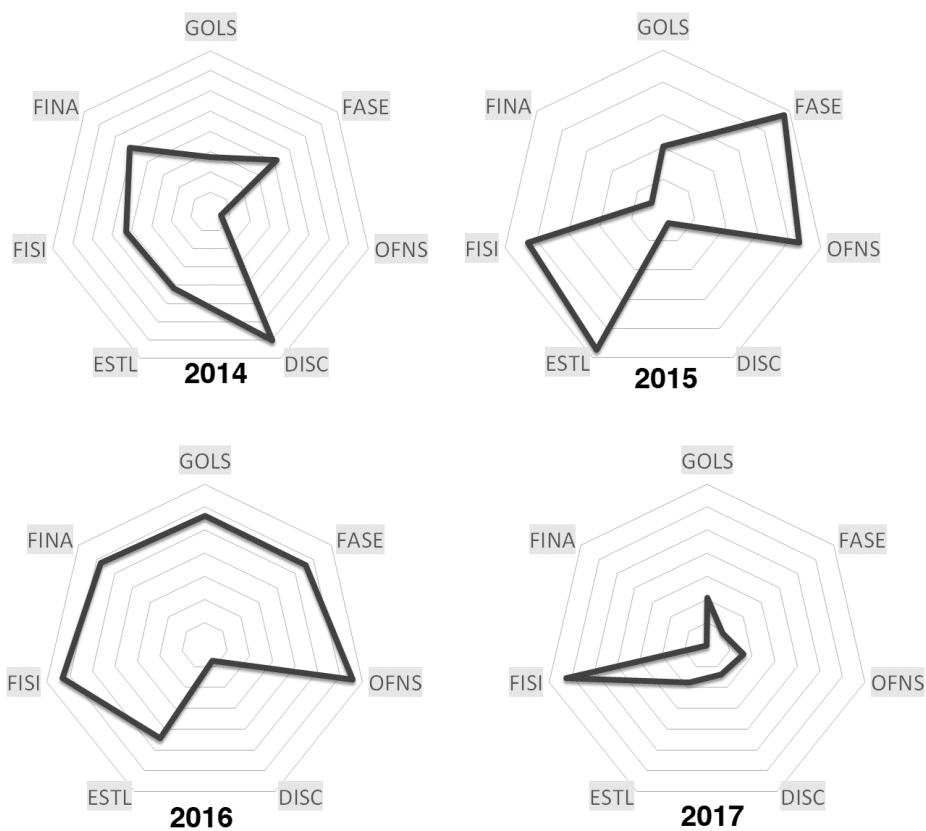


Figura 31: pontuações obtidas por cada grupo de entradas com cada ano entre 2014-2017 utilizado como amostras de teste

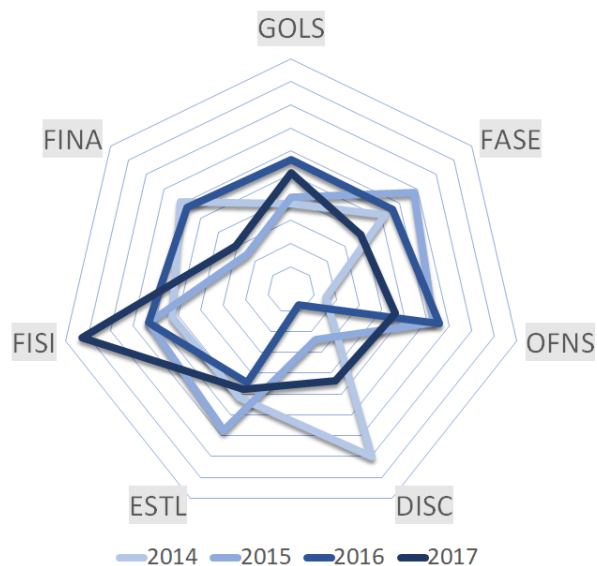


Figura 32: pontuações obtidas por cada grupo de entradas com cada ano entre 2014-2017 utilizado como amostras de teste - combinado

Como podemos observar, existe uma grande variabilidade ano a ano, indicando que os grupos de entradas não oferecem ganhos consistentes ao modelo. No ano de 2014, o gráfico apresenta disciplina como um forte parâmetro e ofensividade e estilo de jogo como parâmetros fracos. Nos anos de 2015 e 2016, essa tendência se inverte: disciplina passa a ser um fator quase nulo, enquanto ofensividade e estilo de jogo se destacam. Esses anos também destacam aspectos físicos e fase atual, porém discordam na importância da média de gols e aspectos financeiros. Por fim, o ano de 2017 apresenta o fator de características físicas como extremamente importante, enquanto todos os outros fatores são secundários.

Um aspecto surpreendente foi a falta de importância da média de gols feitos/sofridos dentro e fora de casa, mesmo sendo o aspecto relacionado mais diretamente ao resultado de uma partida. Mais uma vez, os resultados de cada grupo podem não ter significado, pois a variância entre desempenho dos grupos ano a ano indica aleatoriedade.

Os resultados financeiros variam drasticamente com o grupo de entradas utilizado em cada ano. A Tabela 31 mostra os melhores (+) e piores resultados (-), em conjunto com a combinação de entradas utilizada, para cada ano. As entradas

destacadas em verde compuseram os melhores setups para cada ano, enquanto as entradas destacadas em vermelho compuseram os piores setups.

Tabela 31: piores e melhores resultados obtidos em cada ano por diferentes setups diferenciando pelos grupos de entradas utilizados

| Ano | +/- | ROI 10° (%) | ROI 50° (%) | ROI 90° (%) | GOLS | FASE | OFNS | DISC | ESTL | FISI | FINA |
|------|-----|-------------------|-------------------|-------------------|------|------|------|------|------|------|------|
| 2014 | + | -0,1 | +8,8 | +19,5 | | | | | | | |
| | - | -16,7 | -7,7 | +4,4 | | | | | | | |
| 2015 | + | +4,1 | +9,0 | +13,4 | | | | | | | |
| | - | -10,5 | -0,9 | +9,3 | | | | | | | |
| 2016 | + | -5,4 | +8,2 | +12,9 | | | | | | | |
| | - | -14,7 | -6,6 | +2,8 | | | | | | | |
| 2017 | + | -17,2 | -0,04 | +8,9 | | | | | | | |
| | - | -26,6 | -17,5 | -8,0 | | | | | | | |

Em geral, as entradas tem um efeito semelhante em setups com resultados semelhantes. Isso significa que se o grupo de entradas GOLS está presente no pior setup para o ano de 2014, ele estará também presente na maioria dos outros resultados ruins para aquele ano, bem como se a variável OFNS é presente no melhor setup para o ano de 2015, ela estará também presente na maioria dos outros resultados bons para aquele ano.

A Tabela 31 mais uma vez confirma a alta aleatoriedade da análise. Em cada ano, diferentes grupos estão presentes nas melhores escolhas, com a repetição apenas de ofensividade e aspectos físicos. Todos os grupos são ruins em pelo menos um ano. Outro aspecto está fazendo efeito no resultado: os resultados ruins estão acompanhando setups onde são usadas um número maior de entradas. A Figura 33 mostra a pontuação média obtida por setups de rede neural em função do número de entradas utilizadas,

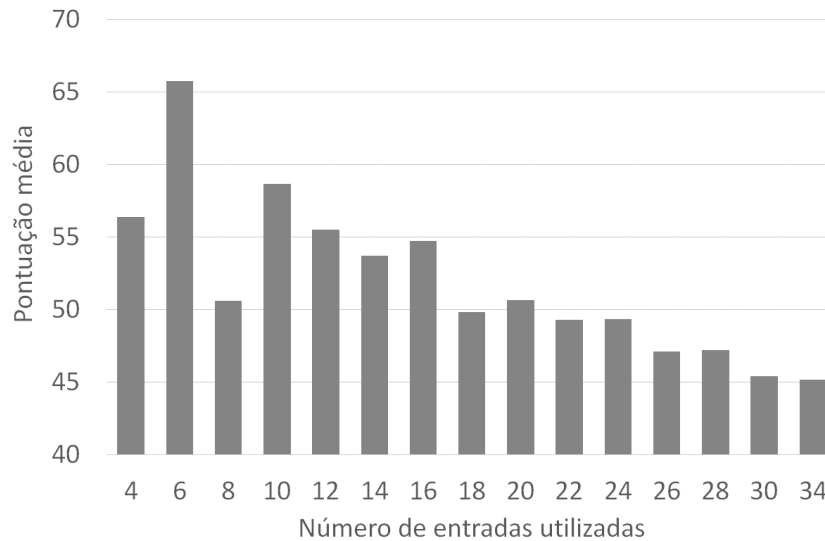


Figura 33: pontuações médias obtidas pelas redes neurais para cada quantidade de entradas utilizadas na análise

As informações confirmam as suspeitas. Isso faz sentido, pois mais entradas com o mesmo número de neurônios podem estar causando uma sobrecarga nestes. Deve ser necessária uma análise onde o número de neurônios é variado conforme o número de entradas utilizadas; para cada setup de entradas, deve ser realizada uma varredura de resultado vs número de neurônios utilizados, para se atingir um ponto máximo e entender se ele é consistente.

Para selecionar os grupos mais importantes, as pontuações ano a ano foram combinadas. As pontuações de cada grupo combinadas ao longo dos quatro anos está representada de forma comparativa na Figura 34 – não há valores quantitativos no eixo y pois a análise é meramente comparativa. Os grupos destacados em verde foram selecionados com resultados positivos, os grupos destacados em cinza foram selecionados com resultados negativos, e os grupos destacados em vermelho foram eliminados com resultados negativos. Ao todo eliminaram-se 10 entradas dentre as 34, sobrando 24 que podem ser resolvidas em tempo hábil.

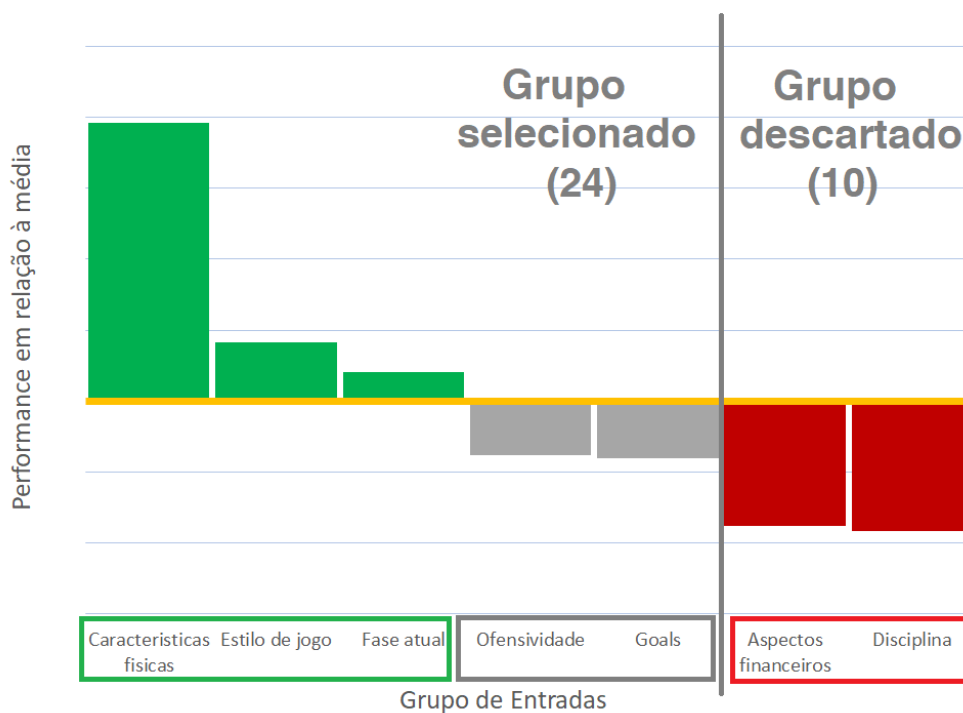


Figura 34: comparativo entre as pontuações obtidas por cada grupo de entrada e identificação dos grupos de entradas selecionados e descartados para a Análise 9.

Dessa maneira, os grupos de entrada selecionados para análise individual e os pares de entradas contidos neles são:

- Características físicas: altura (HEI), peso (WEI) e idade média (AGE) do plantel
- Estilo de jogo: acuracidade dos passes (PASS) e posse de bola (POSS)
- Fase atual: aproveitamentos no campeonato inteiro (AP), nas últimas 5 rodadas (AP5) e jogando fora/dentro de casa (APH)
- Ofensividade: número de finalizações (S) e finalizações no gol (SOG)
- Estatísticas: média de gols por feitos/sofridos pelo time mandante/visitante (GP e GC)

4.2.9 Influência individual de variáveis de entrada

As Figuras 35 e 36 apresentam o resultado comparativo entre a quantidade de pontos obtida por cada entrada individual, conforme descrito no item 3.3. Os resultados são apresentados de forma análoga aos resultados de grupos de entradas.

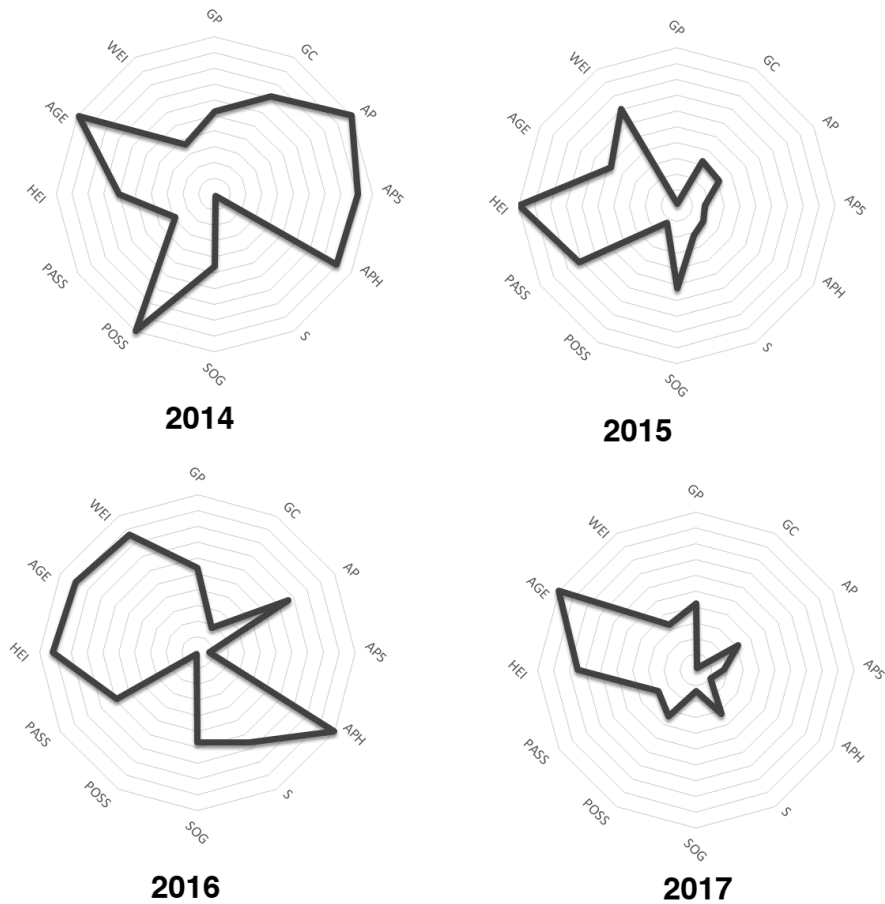


Figura 35: pontuações obtidas por cada entrada individual com cada ano entre 2014-2017 utilizado como amostras de teste

Mais uma vez, a rede apresenta resultados considerados ótimos (+4,5% mediano em 2014, +9,0% mediano em 2015, +5,3% mediano em 2016) mas também apresenta resultados bem ruins (-14,0% mediano em 2017, -1,3% mediano em 2016 e -7,1% mediano em 2014). Dessa maneira, o método não é confiável pois não é possível saber previamente quais entradas deveria ser escolhidas naquele ano. Dessa vez, as redes parecem não favorecer um setup com o mínimo possível de entradas.

Como a rede neural funciona como uma caixa preta que não nos conta seu processo de solução, é muito difícil traçar qual estratégia ela está seguindo para os melhores resultados daquele ano e, provavelmente, ela está encontrando uma dinâmica implícita nos dados de entrada que na verdade não fazem sentido real, e são só uma coincidência. Ou seja, o grupo de entradas não está dando bons resultados por serem entradas suficientes para uma melhor previsão ou funcionarem bem juntas, mas sim porque dentro tantas opções de entradas testadas, algumas devem apresentar coincidentemente valores que levam a rede a tomar conclusões mais corretas, mas nem um pouco consistentes ou de fácil explicação.

Em geral, as entradas com uma melhor quantidade de pontos obtidos são idade média do time, altura média do time e aproveitamentos ao longo do campeonato e em casa/fora. Mais uma vez, é uma conclusão com pouco fundamento, pois imagine-se que apenas as características físicas do time, por si só, não teriam os maiores graus de relação com o resultado. As pontuações obtidas por cada entrada estão apresentadas na Figura 37.

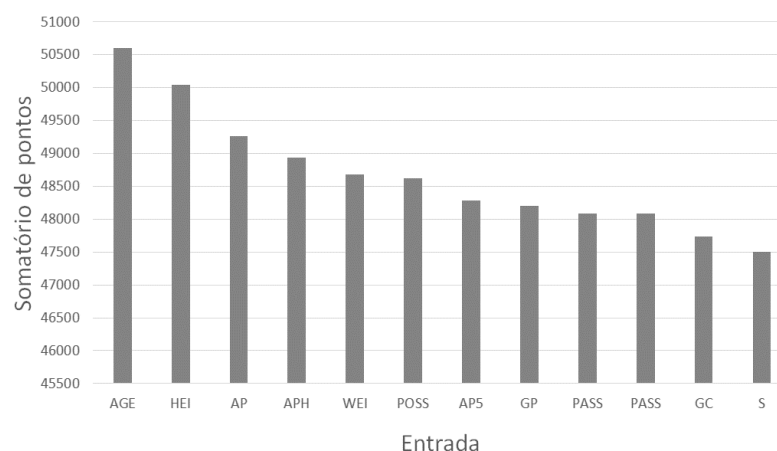


Figura 37: pontuações combinadas obtidas por cada entrada individual ao longo dos 4 anos de teste

4.2.10 Direcionamento estratégico das redes

Após treinamento com os dados de 2007-2013, as redes neurais foram então testadas nos anos de 2014-2016 e, apenas caso obtessem um ROI mínimo com a estratégia bet on all, utilizadas para prever o resultado de 2017. A Figura 38 representa as faixas de ROI obtidas pelo setup básico, com 3 anos de validação, sem imposição de validação (azul escuro) e com imposição de validação com ROI mínimo de 0 até 7%, com incrementos de 1% (azul claro), com ano teste de 2017.

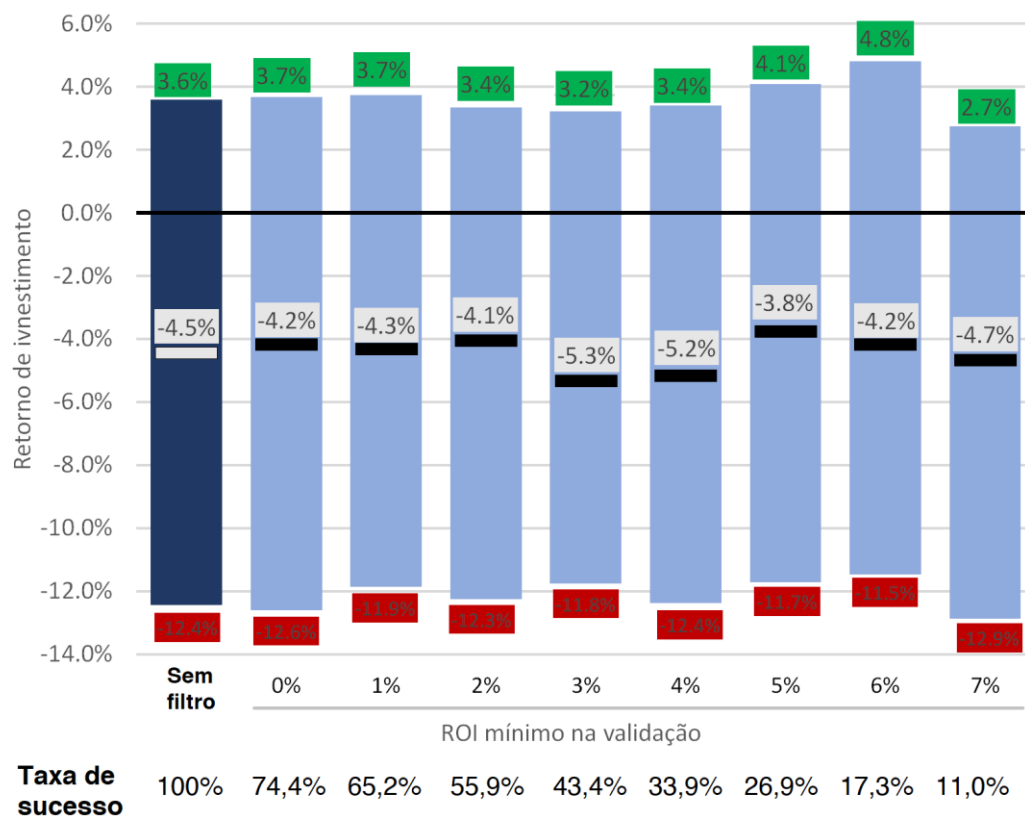


Figura 38: retornos de investimento para 2017 em função do ROI mínimo utilizado na validação de uma rede neural gerada

Infelizmente, a estratégia não parece fazer efeito. Há uma leve melhoria para valores de ROI mínimo na validação de 5% e 6%, mas o comportamento não continua para 7%, mais uma vez gerando dificuldade em entender qual valor poderia ser usado para otimizar o modelo.

Esperava-se que, com a validação, fossem escolhidas redes que perceberam alguma dinâmica escondida especial que melhor descreveria o fenômeno de uma partida de futebol, mas, como não se aplicou ao ano de 2017, não é consistente o

suficiente. Além disso, com o ano de 2017 teve um comportamento diferente de vitórias do mandante, a validação por meio de anos que tiveram maior vitória dos mandantes podem estar escolhendo redes menos adequadas para o ano de 2017. Mas, novamente, nunca é possível saber se um ano terá mais ou menos vitórias do mandante, e essa estratégia não poderia ser utilizada na prática.

Outro entrave dessa estratégia é a taxa de sucesso, representada na parte de baixo do gráfico. A medida que o ROI mínimo na validação aumenta, mais e mais redes treinadas são descartadas e não prosseguem para gerar um resultado no ano teste. Isso significa que, para gerar 1 000 redes de teste para avaliar a confiabilidade de certo resultado, seriam necessárias mais e mais redes, deixando o tempo de treinamento bem alto. Para o caso de ROI mínimo 7%, 89% das redes são rejeitadas, e são necessários realizar 9 090 treinamentos

5 Conclusões

O método estatístico, ao analisar os 4 anos de teste entre 2014-2017, obtem uma taxa de acerto de 50% nos resultados e um retorno de investimento de +2,02% sobre o dinheiro apostado. Para um método simples e que utiliza como entradas dados de anos passados, mesmo com as grandes mudanças de elenco de uma temporada para outra, esse é um resultado bem satisfatório.

As redes neurais artificiais apresentaram bons resultados financeiros para alguns anos analisados. Entretanto, como o fenômeno modelado é altamente aleatório, várias redes neurais geradas a partir do mesmo setup apresentam resultados diferentes, em vez de convergir para uma só estratégia. Dessa maneira, um setup deve ser qualificado quanto ao seu retorno de investimento médio e também em relação a sua consistência. Esses valores mudam drasticamente com o setup escolhido: número de neurônios, parâmetros para estratégias de aposta, ano analisado, entradas utilizadas, etc.

Com a utilização de entradas corretas, o modelo consegue chegar a resultados de retorno de investimento tão altos quanto 10-15% ao ano. Essa taxa de retorno é considerada altíssima em comparação a taxa básica de juros SELIC que estava em 6,50% no dia 15/05/2018. O problema do método das redes neurais é justamente sua inconsistência: os valores para um mesmo setup podem variar de um mínimo de -5% até um máximo de 20% em um determinado caso, o que torna o método proibitivo.

Assim como o método estatístico, as redes neurais possuem um certo viés em apostar na vitória do mandante. O método estatístico chega a apostar 90% das vezes nesse resultado, enquanto a rede varia entre 50-90% dependendo dos parâmetros utilizados, principalmente o número de neurônios. Quando o número de neurônios tende a 0, ela tende a apostar todos os jogos na vitória do mandante. Curiosamente, esse estado é o que costuma gerar os maiores lucros para 2014-2016 que tiveram altas taxas de vitória do mandante, mas gera resultados bem ruins no ano de 2017, quando apenas 45% dos jogos foram vencidos pelos mandantes, contra uma média de 53% nesses outros anos.

As diferentes estratégias de aposta possuem resultados melhores ou piores de acordo com o ano e parâmetros utilizados. Em geral, a estratégia bet on all é menos variável mas tem retornos menores. A estratégia bet if minimum possui uma enorme incerteza com poucos neurônios, com resultados variando entre -90% e +50%. A estratégia bet if value em geral não gera bons resultados, mas curiosamente foi a que gerou melhores resultados no ano de mais difícil previsão, 2017, com ganhos chegando a +12%, contra -12% do método estatístico.

O problema do modelo estudado é justamente esse: por mais que existe, nunca é possível prever qual combinação de entrada, qual estratégia de aposta e qual número de neurônios garantirá o melhor resultado para um dado ano, antes que ele aconteça, tornando o método inconsistente. Em cada ano, uma combinação de entradas parece ser a melhor, quando no ano seguinte essas mesmas entradas podem ser vistas como piores. Um número de neurônios baixo gera melhores lucros, mas tende a virar a estratégia “apostar sempre no mandante”, que não necessita de inteligência artificial para ser seguida.

Utilizando-se um número bem baixo de neurônios na camada escondida, a rede neural é melhor que o método estatístico em todos os anos, o que indica uma superioridade pelo menos perante a esse método. Mas, de qualquer maneira, espera-se que as redes neurais consigam atingir resultados cada vez melhores, uma vez que já existem pesquisas e empresas apostando nas redes neurais como forma de realizar apostas lucrativas. Essas apostas podem ser feitas não só em esportes, mas em outros eventos de previsão e até no mercado de ações.

Para isso, é necessário que sejam levantadas entradas que possui real relevância na geração do resultado. As entradas de média de gols, aproveitamento no campeonato, características físicas, ofensivas, estilo de jogo da equipe, valor da equipe e disciplina não foram capazes de gerar bons resultados consistentemente. Deve-se enriquecer o modelo com novas entradas que possam ter mais valor, e algumas ideias estão apresentadas nas ideias de trabalhos futuros.

Em resumo, o método das redes neurais deixa uma mensagem que pode sim gerar ótimos resultados, mas primeiro deve ser melhor dominado, e mais análises precisam ser feitas.

6 Sugestões de Trabalhos Futuros

As sugestões de trabalhos futuros giram em torno, principalmente, do enriquecimento do modelo proposto. O modelo pode ser enriquecido por:

Levantamento de dados de entrada mais assertivos, como informações do plantel, inclusão de jogadores com contusão ou suspenso, dados estatísticos da mesma temporada, avaliação de cada equipe conforme sua força de goleiro, defesa, meio de campo e ataque, ou até utilização das notas dos jogadores no Cartola FC como entradas para definir a força do plantel. Diversas entradas podem ser imaginadas, e devem ser testadas caso não exista uma dificuldade muito grande em encontrar esses dados. Pode-se utilizar também previsões de outros profissionais e sites de sugestão de aposta, bem como a comparação entre os payouts praticados por diferentes sites para um mesmo jogo.

Levantamento de novas estratégias de aposta, como um sistema que altera a quantidade a ser apostada de acordo com a certeza que a rede neural encontra naquele resultado.

Levantamento de novas estratégias de validação, para tentar filtrar melhor as redes neurais com má capacidade de generalização a partir de amostras de validação, entendendo se a rede neural gerada está fazendo apostas que fazem sentido: está distribuindo corretamente as apostas conforme a ocorrência histórica? Está apostando corretamente em jogos onde há um claro favorito? Etc.

7 Bibliografia

- Bishop, C. M. (1996). *Neural Networks for Pattern Recognition*. Clarendon Press.
- Blackwood, K. (2006). *Casino Gambling For Dummies*. For Dummies.
- Boyle, M. (2006). *The Art of Bookmaking: How to Compile Odds for Any Sporting Event*. High Stakes.
- Braga, A., Ludermir, T., & Carvalho, A. C. (2000). *Redes Neurais Artificiais: Teoria e Aplicações*. Rio de Janeiro: LTC.
- Dash, T., & Nayak, T. (2012). English Character Recognition using Artificial Neural. *Proceedings of National Conference on AIRES-2012*. Andhra University.
- Doidge, N. (2007). *The Brain That Changes Itself*. Viking Press.
- Drubach, D. (1999). *Brain Explained*. Pearson Education.
- Edoras. (2002). *Conjugate Gradient Algorithms*. Fonte: Neural Network Toolbox: <https://edoras.sdsu.edu/doc/matlab/toolbox/nnet/backpr58.html>
- Feustel, E., & Howard, G. (2010). *Conquering Risk: Attacking Vegas and Wall Street*. George S. Howard.
- Fischetti, M. (1 de Novembro de 2011). *Scientific American*. Fonte: Scientific American: <https://www.scientificamerican.com/article/computers-vs-brains/>
- Haykin, S. (1994). *Neural networks: A comprehensive foundation*. Macmillan Coll Div.
- Host, G. J. (2000). *The Nervous System and the Heart*. Humana Press.
- Kopytoff, V. G. (19 de Junho de 2011). *New york Times*. Fonte: New york Times: <https://www.nytimes.com/2011/06/20/technology/20computer.html>
- Levitan, I. B., & Kaczmarek, L. K. (1991). *The Neuron: cell and molecular biology*. Oxford: Oxford University Press.
- Møller, M. (1993). A Scaled Conjugate Gradient Algorithm For Fast Supervised Learning. *Neural Networks Vol. 6*, 525-533.
- Posner, G. (2013). *Human Labor and Birth*. McGraw Hill Education.

- Rosenberg, L., & Baltaxe, D. (2016). Crowds vs Swarms, a Comparison of Intelligence. *2016 Swarm/Human Blended Intelligence Workshop (SHBI)*. IEEE.
- Ryan, K. J. (14 de Junho de 2018). *INC.COM*. Fonte: INC.COM: <https://www.inc.com/kevin-j-ryan/unanimous-ai-swarm-intelligence-makes-startlingly-accurate-predictions.html>
- Schüll, N. (2014). *Addiction by Design: Machine Gambling in Las Vegas*. Princeton University Press.
- Sinha, S. (16 de Outubro de 2015). *Inkjet wholesale*. Fonte: Inkjet wholesale: <https://blog.inkjetwholesale.com.au/printer-cartridge-cost-saving-tips/how-to-calculate-your-printer-power-consumption/>
- Spann, M., & Skiera, B. (2009). Sports Forecasting: A Comparison of the Forecast Accuracy of Prediction Markets, Betting Odds and Tipsters. *Journal of Forecasting*, 55-72.
- Unanimous Ai. (3 de Março de 2018). *Unanimous Ai*. Fonte: Unanimous Ai: <https://unanimous.ai/oscars-2018/>
- Unanimous AI. (14 de Junho de 2018). *Unanimous AI*. Fonte: Unanimous AI: <https://unanimous.ai/world-cup/>
- Vincent, J. (6 de Julho de 2017). *The verge*. Fonte: The verge: <https://www.theverge.com/2017/7/6/15923784/ai-predict-sport-betting-gambling-stratagem>