

Trabalho de Conclusão de Curso

Engenharia de Computação

SELEÇÃO DE FATORES DE RISCO DE LESÕES EM ATLETAS, ATRAVÉS DA EXTRAÇÃO DE CONHECIMENTO DOS PESOS DE UMA REDE NEURAL

Autor: Havana Diogo Alves

Orientador: Mêuser Jorge Valença

Co-Orientador: Wellington Pinheiro dos Santos



UNIVERSIDADE
DE PERNAMBUCO

HAVANA DIOGO ALVES

**SELEÇÃO DE FATORES DE RISCO
DE LESÕES EM ATLETAS, ATRAVÉS
DA EXTRAÇÃO DE CONHECIMENTO
DOS PESOS DE UMA REDE NEURAL**

Monografia apresentada como requisito parcial para obtenção do diploma de Bacharel em Engenharia da Computação pela Escola Politécnica de Pernambuco – Universidade de Pernambuco.

Recife, junho de 2010.

A todos que acreditam na beleza de seus sonhos!

Agradecimentos

Aos meus pais, Cleonice Diogo Alves e Marcos Diogo Ramos, pelo incentivo irrestrito, por serem meu porto seguro, por toda a educação e por todo o amor. Aos meus irmãos Tamires Diogo Alves e Marcos Antônio Diogo Alves.

A toda a minha família, em especial minha avó Geni Ramos Barbosa e minha tia Marta Diogo de Lima, pelos bons momentos e pelos bons sentimentos compartilhados.

A meu namorado Edson Bruno Andrade da Silva, por ser paciente e me confortar e me fazer sorrir nos momentos de maior estresse e cansaço.

A todos os meus amigos, pelo apoio e pelo companheirismo.

Aos Professores Mêuser Jorge Valença e Wellington Pinheiro dos Santos, pela grande atuação e importância na minha vida acadêmica.

A minha amiga e companheira de projetos, Liliane Alves do Nascimento, que compartilhou comigo as noites mal dormidas para finalização dos trabalhos da graduação.

Por fim, agradeço a todos os meus professores pela boa formação que recebi.

Obrigado a todos, por compartilharem meus momentos com alegria!

Resumo

O atletismo brasileiro vem crescendo de forma significativa nos últimos anos, conquistando, assim, posições no *ranking* mundial. O aumento na intensidade do treinamento dos atletas foi fator relevante para esse feito. Como conseqüência do aumento da demanda de exercícios cada vez mais modernos e competitivos, houve um crescimento no risco do aparecimento de lesões, o que é motivo de apreensão para atletas e treinadores, pois interrompem o processo evolutivo do treinamento. Muitos são os fatores que podem contribuir para o surgimento de tais lesões, e saber qual a relevância de cada um deles é de grande importância para os profissionais ligados ao esporte, pois sabendo os mecanismos principais do lesionamento é possível tomar medidas preventivas. Na área de saúde, métodos estatísticos tradicionais são comumente usados no estudo dos mais diversos casos, porém vale ressaltar que tais métodos detectam apenas características lineares dos dados. Como o mundo real é repleto de fenômenos não-lineares, torna-se necessária a utilização de técnicas que capturem esta não-linearidade. Para estes casos, as Redes Neurais Artificiais têm mostrado resultados bastante satisfatórios. Este trabalho tem como objetivo utilizar algoritmos de seleção de variáveis que extraem conhecimento dos pesos e sensibilidades de uma rede neural para obter a contribuição de cada fator no surgimento de lesões em atletas. Com isso, espera-se encontrar uma alternativa mais precisa para se entender os mecanismos favorecedores de lesões, podendo dessa forma, ajudar na determinação de medidas preventivas.

Abstract

The Brazilian athletics has grown significantly in recent years, by gaining thus positions in the world ranking. The increased intensity of training of the athletes was a relevant factor in this achievement. As a result of increased demand of exercise increasingly modern and competitive, there was a growth of risk of developing injuries, which worries athletes and coaches, for interrupting the evolutionary process of training. There are many factors that can contribute to the appearance of such injuries, and knowing the relevance of each one is very important to the professionals linked to the sports, because knowing the main mechanisms of the injury, they can take preventive steps. In the field of Health, traditional statistical methods are usually used for studying various cases, but it is noteworthy that such methods only detect linear features of the data. As the real world is full of non-linear phenomena, it becomes necessary to use techniques that capture this nonlinearity. For these cases neural networks have shown good results. This study aims to use variable selection algorithms that extract knowledge from weights and sensibilities to obtain the contribution of each factor in the development of injuries in athletes. We expect to find more accurate alternative to understand the mechanisms that favor injuries and may thus help in determining preventive steps.

Sumário

ÍNDICE DE FIGURAS	IX
ÍNDICE DE TABELAS	XI
TABELA DE SÍMBOLOS E SIGLAS	XII
1 INTRODUÇÃO	13
1.1 Motivação e Problema	13
1.2 Objetivos	14
1.2.1 Objetivo Geral	14
1.2.2 Objetivos Específicos	14
1.3 Estrutura da Monografia	14
2 REVISÃO BIBLIOGRÁFICA	16
2.1 Lesões na Prática de Exercícios	16
2.1.1 Definição de Lesão e seus Tipos	17
2.1.2 Lesões no Atletismo e Fatores que contribuem para o seu surgimento	19
2.2 Redes Neurais Artificiais	21
2.2.1 O Neurônio Biológico	21
2.2.2 O Neurônio Artificial	22
2.2.3 A Rede <i>MultiLayer Perceptron</i> (MLP)	25
2.3 Critério de Parada: validação cruzada	28
2.4 Verificação	29
2.5 Seleção de Variáveis	29
2.5.1 Algoritmo de Garson e o de Garson Modificado	30
2.5.2 Análise de Sensibilidades	32
3 METODOLOGIA	34

3.1	Previsão e Classificação com Rede Neural MLP	34
3.1.1	Implementação e Configuração da Rede	34
3.1.2	Pré-processamento dos dados	36
3.1.3	Validação do algoritmo da RNA MLP e base utilizada	37
3.2	Seleção de Variáveis	38
3.2.1	Experimentos com os algoritmos de Seleção	38
3.3	Estudo da contribuição dos fatores no surgimento de lesões em atletas	40
4	RESULTADOS	44
4.1	Resultados da RNA MLP.	44
4.2	Resultados dos algoritmos de seleção de variáveis	45
4.2.1	Resultados obtidos com as equações	45
4.2.2	Resultados com as bases de dados	49
4.3	Resultado da contribuição dos fatores no surgimento de lesões em atletas	53
5	CONCLUSÃO E TRABALHOS FUTUROS	56
	REFERÊNCIAS	59

Índice de Figuras

Figura 1.	Ruptura total do ligamento, caracterizando uma entorse de 3° grau...	18
Figura 2.	Exemplo de lesão muscular.....	19
Figura 3.	Representação de uma célula nervosa e a direção do impulso nervoso..	22
Figura 4.	Neurônio artificial proposto por McCulloch e Pitts.....	22
Figura 5.	Neurônio artificial com <i>bias</i>	24
Figura 6.	Rede Perceptron.....	25
Figura 7.	Rede MLP com uma camada de entrada, duas intermediárias e uma de saída	26
Figura 8.	Validação cruzada: ponto de parada.	28
Figura 9.	Gráfico da distribuição das classes de vinhos.	40
Figura 10.	Gráfico da distribuição dos Atletas nas cinco categorias pesquisadas.	41
Figura 11.	Distribuição dos atletas por faixa etária. Oberva-se maior quantidade de atletas na faixa dos 16 a 18, e no lado a quantidade de atletas com mais de 25 anos diminui bastante.	42
Figura 12.	Gráfico da Quantidade de Atletas X Horas de Treino Diário.....	43
Figura 13.	Tempo de prática de modalidade em anos.....	43
Figura 14.	Gráfico da convergência do treinamento com a base das espécies-íris. O eixo vertical indica o valor do Erro Médio Quadrático (EMQ) para o ciclo correspondente no eixo horizontal.	45

- Figura 15.** Gráficos da seleção de variáveis para o polinômio multivariado. (a) Gráfico da importância para cada variável obtido com os algoritmos de Garson original e modificado. (b) Gráfico dos valores de importância obtidos com a análise de sensibilidades.....46
- Figura 16.** Gráficos da seleção de variáveis para o polinômio de Kolmogorov-Gabor. (a) Gráfico da importância para cada variável obtido com os algoritmos de Garson original e modificado. (b) Gráfico dos valores de importância obtidos com a análise de sensibilidades.48
- Figura 17.** Gráficos da seleção de variáveis para a função com interação complexa. (a) Gráfico da importância para cada variável obtido com os algoritmos de Garson original e modificado. (b) Gráfico dos valores de importância obtidos com a análise de sensibilidades.....49
- Figura 18.** Gráfico de seleção de variáveis para a base de espécies-íris resultante do algoritmo de Garson.50
- Figura 19.** Gráfico de seleção de variáveis para a base de espécies-íris resultante da análise de sensibilidades.....51
- Figura 20.** Gráfico de seleção de variáveis para a base de classes de vinhos resultante do algoritmo de Garson modificado.....52
- Figura 21.** Gráfico de seleção de variáveis para a base de classes de vinhos resultante da análise de sensibilidades.....53
- Figura 22.** Gráfico da contribuição dos fatores no favorecimento de lesões em atletas. 55

Índice de Tabelas

Tabela 1.	Exemplo de valores para os pesos das conexões de uma rede com 3 entradas, três neurônios intermediários e uma saída.....	30
Tabela 2.	Parâmetros configuráveis para a Execução da Rede	34
Tabela 3.	Pequena amostra de dados da base das Espécies-Íris. Esta base deve ser normalizada e passar por um processo de binarização.....	37
Tabela 4.	Atributos da Base de Dados da Elite do Atletismo Brasileiro	41
Tabela 5.	Configuração dos parâmetros para a obtenção do melhor resultado de classificação das espécies-íris.	44
Tabela 6.	Resultados encontrados pelos algoritmos de seleção de variáveis para a o polinômio multivariado. As linhas com fundo colorido contém as variáveis que fazem parte da equação.....	46
Tabela 7.	Resultados encontrados pelos algoritmos de seleção de variáveis para o polinômio de Kolmogorov - Gabor. As linhas com fundo colorido contém as variáveis que fazem parte da equação.	47
Tabela 8.	Resultados encontrados pelos algoritmos de Garson modificado e de análise de sensibilidades para a base de espécies-íris.....	50
Tabela 9.	Resultados encontrados pelos algoritmos de Garson modificado e de análise de sensibilidades para a base de classes de vinhos.....	52
Tabela 10.	Importância dos fatores para o surgimento de lesões.	54
Tabela 11.	Distribuição da quantidade de atletas lesionados por modalidade.	57

Tabela de Símbolos e Siglas

NAIRS - *National Athletic Injury/Illness Reporting System*

RNA – Rede Neural Artificial

MLP – *MultiLayer Perceptron*

CI – Computação Inteligente

LD – Lesões Desportivas

SAG – Síndrome da Adaptação Geral

IC – Inteligência Computacional

ADALINE – ADAPtative LINear Element

EMQ– Erro Médio Quadrático

EMA – Erro Médio Absoluto

EPMA – Erro Percentual Médio Absoluto

IMC – Índice de Massa Corporal

OMS – Organização Mundial de Saúde

1 Introdução

Esse capítulo se inicia descrevendo o problema e a motivação do desenvolvimento deste trabalho. Posteriormente são expostos seus principais objetivos. Por fim, ele é encerrado mostrando o conteúdo abordado nos capítulos seguintes.

1.1 Motivação e Problema

O Sistema de Registro Nacional de Lesões Atléticas dos Estados Unidos (NAIRS) define lesão como: “acontecimento que limita a participação do atleta por no mínimo um dia após sua ocorrência”. Estudos recentes afirmam que as lesões trazem como consequência o afastamento de treinos e jogos por um determinado tempo, ocorrendo assim prejuízos tanto ao atleta como às equipes [1].

A frequência de lesões pode variar de acordo com o tempo de prática esportiva e o número de horas de treino por semana, sendo consideradas fator principal de afastamento dos atletas, influenciando seu desempenho e possivelmente seu psicológico pela demora ao retorno das práticas esportivas, afetando a equipe como um todo [2]

De acordo com Pastre *et al.* [3], há um aumento de lesões simultaneamente com a crescente demanda da prática de exercícios, tanto recreacionais como competitivos, gerando preocupações aos treinadores e atletas, pois interrompem o processo evolutivo de adaptações sistemáticas impostas pelo treinamento. O mesmo autor ainda cita que as lesões esportivas são resultantes de uma interação de fatores de risco, sendo eles intrínsecos ou extrínsecos. Os fatores intrínsecos estão relacionados à idade, sexo, condição física, desenvolvimento motor, alimentação e fatores psicológicos. Fatores extrínsecos se associam à especificidade técnica de cada modalidade, tipo de equipamento usado, organização e cargas de treino e competição, condições climáticas ou a combinação destes.

Leite [4] classifica as lesões em duas categorias básicas, sendo elas traumáticas ou por excesso de uso. O mesmo autor ainda descreve três categorias de mecanismo

de lesão relacionada com esportes: uso excessivo também denominado *overtraining*, contato direto e insuficiência de partes moles. Estes mecanismos de lesão juntamente com a causa situacional e fatores de risco devem ser compreendidos por médicos, fisioterapeutas e educadores físicos para prevenção, diagnóstico e tratamento dos atletas.

1.2 Objetivos

1.2.1 Objetivo Geral

- Obter a contribuição de cada fator no favorecimento de lesões em atletas.

1.2.2 Objetivos Específicos

- Implementar uma RNA MLP com algoritmo *Backpropagation*;
- Testar o desempenho da rede implementada.
- Implementar algoritmos de seleção de variáveis que utilizam os pesos e sensibilidades resultantes do treinamento da rede.
- Comparar as técnicas de seleção.
- Escolher a técnica de seleção que rendeu resultados mais satisfatórios, para obter a importância dos atributos da base de dados dos atletas.

1.3 Estrutura da Monografia

O Capítulo 2 reúne o conteúdo teórico necessário para se compreender o trabalho proposto, a saber: lesões na prática de exercícios, RNAs, Critério de Parada, Verificação e Seleção de Variáveis. Em seguida, o Capítulo 3 mostra todas as abordagens desenvolvidas para atingir o objetivo geral do projeto, tais como o uso de RNA e de algoritmos de seleção de variáveis. Logo depois, o Capítulo 4 mostra todos os resultados obtidos a partir das aplicações e simulações das técnicas desenvolvidas. Por

fim, o Capítulo 5 resume o trabalho final, reúne dificuldades encontradas, discussões geradas e as conclusões obtidas, além de melhorias e trabalhos futuros.

2 Revisão Bibliográfica

Este capítulo descreve todo conteúdo teórico que foi utilizado na tentativa de resolver o problema descrito no capítulo 1. A seção 2.1 aborda alguns conceitos sobre lesões e suas causas. Em seguida, a seção 2.2 discorre sobre RNAs, uma técnica de CI que vem mostrando ser uma alternativa mais precisa na predição e classificação em problemas do mundo real (problemas não lineares). Já a seção 2.3 fala sobre validação cruzada, que é uma técnica usada como critério de parada da rede. Na seção 2.4 são apresentadas algumas métricas para avaliar o desempenho do treinamento da rede. E, por fim, a seção 2.5 traz três técnicas de seleção de variáveis, as quais serão o foco deste trabalho.

2.1 Lesões na Prática de Exercícios

Nos últimos anos, o número de praticantes de atividade física cresceu vertiginosamente. Segundo Kettunen *et al.* [5], o aumento da demanda de exercícios modernos e competitivos provocou o aumento simultâneo no risco de lesões, causando preocupações tanto para os atletas, quanto para os profissionais responsáveis pelo condicionamento físico desses, pois interrompem o processo evolutivo de adaptações sistemáticas impostas pelo treinamento.

As ocorrências das lesões desportivas (LD), possivelmente, são resultado de exercícios realizados de maneira extenuante e, ainda, inadvertida ou inapropriadamente, sendo subestimadas a prevalência e incidência destes episódios devido à ausência de notificação em todo o universo esportivo, seja na iniciação das modalidades ou em altos níveis de *performance*.

Embora não se conheça com clareza os mecanismos que podem proporcionar as lesões musculares, possivelmente estão relacionadas com as concentrações dos íons cálcio e com a própria mecânica do movimento. De qualquer forma as lesões musculares esqueléticas (alterações morfofuncionais) são uma constante na vida dos indivíduos que realizam alguma atividade física e podem ser mais freqüentes naquelas atividades, onde há um grande número de ações excêntricas [6].

O conhecimento da causa situacional, mecanismo de lesão, fatores de risco, entre outros aspectos, podem auxiliar os profissionais da área de educação física ou fisioterapia no processo de prevenção, diagnóstico e tratamento desses agravos, particularmente no atletismo, caracterizando, portanto, importante contribuição para as ciências da saúde e do esporte.

2.1.1 Definição de Lesão e seus Tipos

Uma lesão é caracterizada pela alteração ou deformidade tecidual, que pode atingir vários tipos e níveis de tecidos.

No caso da população atlética, as lesões envolvem mais comumente o sistema muscoesquelético e, mais raramente o sistema nervoso. As lesões primárias são quase sempre descritas na medicina esportiva como sendo de natureza crônica ou aguda, resultantes de forças macrotraumáticas ou microtraumáticas. As lesões microtraumáticas ocorrem em decorrência do trauma agudo e produzem dor e incapacidade imediatas. As lesões macrotraumáticas incluem fraturas, luxações, subluxações, entorses, distensões e contusões. As lesões microtraumáticas são geralmente denominadas lesões por excesso de uso, ou *overuse*, e são resultantes da sobrecarga repetitiva ou de uma mecânica gestual motora incorreta, relacionada ao treinamento contínuo ou à competição. As lesões microtraumáticas incluem tendinite, tenossinovite, bursite, etc [7].

Os tipos de lesões são classificados de acordo com o tecido a qual elas atingem, podendo, portanto, ser ligamentares, ósseas, musculares ou articulares.

As **lesões ligamentares** são comumente relacionadas à entorse ligamentar. Uma entorse envolve dano a um ligamento que fornece suporte a uma articulação. O ligamento é uma faixa de tecido rígido, relativamente inelástico, que liga um osso a outro.

Se estresses forem aplicados a uma articulação que forcem o movimento além de seus limites ou planos de movimento normais, é provável que ocorra lesão ao ligamento. A gravidade do dano ao ligamento é classificada de diferentes maneiras; entretanto, o sistema mais comumente usado envolve três classes (graus) de entorse ligamentares:

- **Entorse de 1º Grau:** Existe algum estiramento ou talvez ruptura das fibras ligamentares, com pouca ou nenhuma instabilidade articular. Dor branda, pouco edema e rigidez articular podem ser observados.
- **Entorse de 2º Grau:** Existe certa ruptura e separação das fibras ligamentares e instabilidade moderada da articulação. Dor de moderada a aguda; edema e rigidez articular devem ser esperados.
- **Entorse de 3º Grau:** Existe ruptura total do ligamento, manifestada primariamente por grande instabilidade articular. A figura 1 ilustra um caso de entorse de 3º grau.

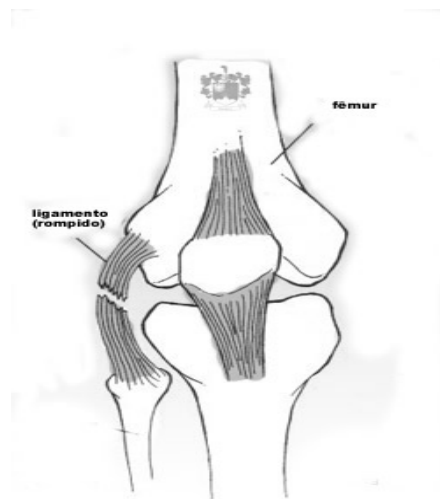


Figura 1. Ruptura total do ligamento, caracterizando uma entorse de 3º grau.

As **lesões ósseas** são caracterizadas geralmente por fraturas, que são lesões extremamente comuns entre a população atlética. Podem ser classificadas, de modo geral, como abertas ou fechadas. A fratura fechada envolve pouco ou nenhum deslocamento dos ossos e, portanto, pouca ou nenhuma ruptura do tecido mole. A fratura aberta, por outro lado, envolve deslocamento suficiente das extremidades fraturadas para que o osso rompa de fato as camadas cutâneas e abra caminho para a pele.

As **lesões articulares** são muito relacionadas a danos à cartilagem. A **osteoartrose** é um distúrbio degenerativo do osso e da cartilagem na articulação e em torno dela.

Outros tipos de lesões que envolvem as articulações são a luxação e a subluxação, que são, respectivamente, o afastamento de duas superfícies articulares, mantendo-se afastadas no primeiro caso, e voltando a posição inicial no segundo.

Em atletas e praticantes de atividades físicas regulares, algumas articulações podem ser mais suscetíveis a uma resposta parecida com a osteoartrose. A proporção do peso corporal em repouso sobre a articulação, a distensão da unidade musculotendinosa e qualquer força externa importante aplicada sobre a articulação são fatores de predisposição. Uma mecânica articular alterada, causada por frouxidão ou por traumas anteriores, são também fatores a serem considerados.

- De todos os tipos de lesões, as **musculares**, são as mais freqüentes entre atletas. Na figura 2 temos ilustrado um caso de lesão muscular.

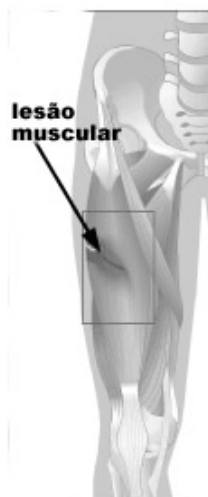


Figura 2. Exemplo de lesão muscular.

2.1.2 Lesões no Atletismo e Fatores que contribuem para o seu surgimento

Dentre as práticas esportivas, o atletismo se destaca pela diversidade de modalidades, cada qual caracterizada pela presença de condições específicas de treinamento e presença de elementos básicos, como correr, saltar, lançar ou arremessar, que são observados nos demais esportes com suas respectivas adaptações [8].

Um dos princípios que norteiam o treinamento esportivo é a Síndrome da Adaptação Geral (SAG). A SAG prevê que o corpo humano deve apresentar três formas de reação ao estresse. A primeira é conhecida como fase de alarme e caracteriza-se

por uma reação de choque e contrachoque; a segunda, e que interessa para fins de performance humana, é conhecida como fase de resistência, quando o organismo reage de tal modo a produzir adaptações para que possa resistir a novas demandas impostas pelo estímulo; e a terceira é conhecida como fase de exaustão, quando há saturação da capacidade de reagir a determinado estímulo [9].

O problema está na busca de alcançar ou até ultrapassar limites que treinadores e atletas tanto necessitam, e que por vezes, ultrapassam a fase de adaptação, resultando em exaustão. Embora existam parâmetros para controlar os estímulos aplicados, há grande dificuldade em articular as diferentes formas de manifestação do estresse de treinamento, as quais podem ser divididas em: bioquímica ou metabólica; física ou estrutural e mental ou psicológica, segundo Sharkey [10].

No caso das lesões musculares as causas mais comuns são: excesso de treinamento; falta de controle nas tensões de exercícios e alongamentos; gestual motor (técnica) indevido nos exercícios e alongamentos; carência de exercícios de alongamento compensatórios após os exercícios físicos; excesso de força e insuficiência de flexibilidade, ou fraqueza com muita flexibilidade; excesso de exercícios, tanto de força quanto de alongamento, em músculos fracos, particularmente naqueles que suportam estruturas de apoio; excesso de exercícios de força isoladamente em grupos musculares com encurtamento; dispensa de aquecimento antes do treinamento e retorno ao treinamento antes da cura total de uma lesão.

Para a instalação de agravos, entre amplas possibilidades de fatores causais, os métodos de treinamento adotados contribuem significativamente para a ocorrência desses eventos. Adicionalmente, para se isolar o efeito dos fatores causais referidos, também são identificados entre outros, os quais se dividem em intrínsecos como idade, sexo ou biótipo, e extrínsecos como clima, piso de prática esportiva ou evento praticado [11 - 13].

2.2 Redes Neurais Artificiais

É uma técnica utilizada em resolução de problemas característicos da Inteligência Computacional (IC), tais como, previsão, classificação, reconhecimento de padrões e clusterização inspirada no funcionamento do cérebro humano. As RNAs foram criadas não para imitar o funcionamento do cérebro, mas para ser um modelo de resolução de problemas não algorítmicos. Assim, devemos entender a unidade funcional básica cerebral, o neurônio.

2.2.1 O Neurônio Biológico

O neurônio é constituído de um corpo celular, dendrito e um axônio. O dendrito é uma estrutura do tipo ramo, com sinapses para a recepção de mensagens eletroquímicas. O axônio é uma projeção longa que leva impulsos a partir do corpo celular [14].

Um impulso nervoso é a transmissão de uma alteração elétrica ao longo da membrana do neurônio a partir do ponto em que ele foi estimulado. A direção normal do impulso no organismo é do corpo celular para o axônio.

Esse impulso nervoso, ou potencial de ação, é uma alteração brusca e rápida da diferença de potencial transmembrana. Normalmente a membrana do neurônio é polarizada em repouso, sendo que o potencial é negativo (-70 mV). O potencial de ação consiste de uma redução rápida da negatividade da membrana até 0mV e inversão deste potencial até valores de cerca de +30mV, seguido de um retorno também rápido até valores um pouco mais negativos que o potencial de repouso de -70mV.

O nível crítico de despolarização que deve ser atravessado a fim de desencadear um potencial de ação é chamado limiar. Os potenciais de ação são causados pela despolarização da membrana além do limiar (Lei do Tudo ou Nada).

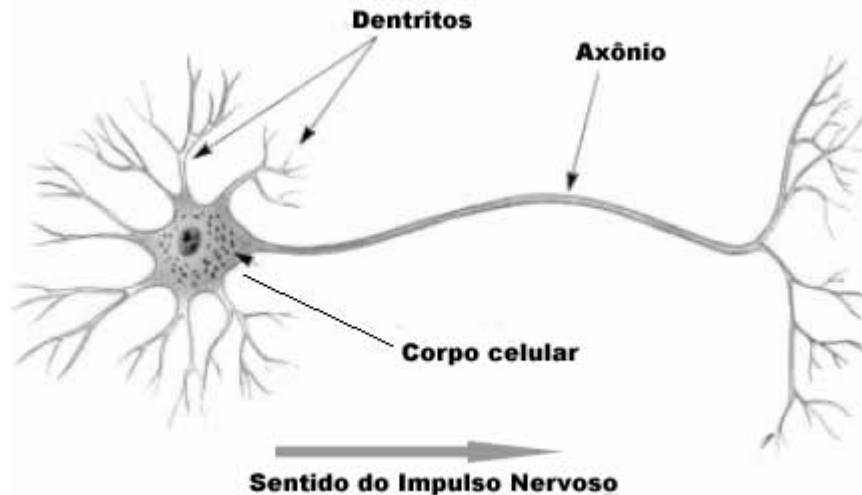


Figura 3. Representação de uma célula nervosa com a indicação da direção do impulso nervoso.

2.2.2 O Neurônio Artificial

O neurônio artificial é um modelo de representação matemática de um neurônio biológico. A primeira proposta de um neurônio artificial foi feita por McCulloch e Pitts em 1943 (Figura 4) [15].

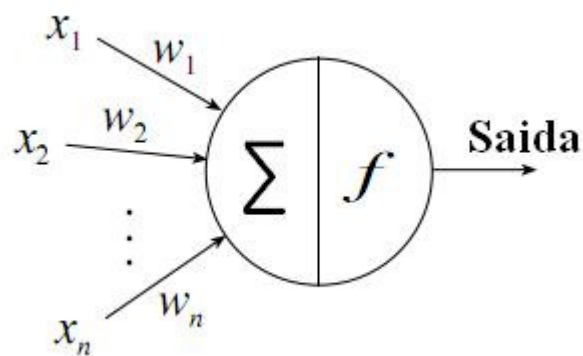


Figura 4. Neurônio artificial proposto por McCulloch e Pitts.

O modelo proposto é uma forma bastante simples de representação do neurônio biológico, no qual utiliza uma regra de propagação e uma função de ativação. Neste modelo a função de ativação é a função degrau (Equação 2.1), assim a emissão de sinal tem por base um limiar, analogamente à lei do Tudo ou Nada.

$$y = \begin{cases} 1, & net \geq 0 \\ 0, & net < 0 \end{cases} \quad (2.1)$$

Na Figura 4, x_1, x_2, \dots, x_n são os valores de entrada do neurônio e w_1, w_2, \dots, w_n são os pesos atribuídos a cada entrada. Durante a propagação do sinal os valores de cada entrada são multiplicados pelos respectivos pesos e o somatório dos produtos obtidos menos o limiar resulta na entrada líquida (net) (Equação 2.2) é utilizado como parâmetro para a função de ativação $f(net)$, que é o valor de saída do neurônio.

$$net_i = \sum_{j=0}^n w_{ij} x_j - \theta \quad (2.2)$$

, onde n é igual à quantidade neurônios na camada de entrada e θ é o limiar.

Para os algoritmos de aprendizado é mais adequado que a consideração do limiar seja implícito. Para isso, como alternativa de representação do neurônio, duas modificações são feitas:

- adiciona-se uma nova entrada de valor fixo $x_0 = +1$;
- acrescenta-se um novo peso sináptico $w_{i0} = -\theta$.

Essas modificações resultam em uma nova forma de cálculo da entrada líquida, representada pela equação 2.3. A nova entrada com valor fixo igual a 1 é chamada *bias*. A arquitetura do neurônio também muda como mostra a figura 5.

$$net_i = \sum_{j=0}^n w_{ij} x_j \quad (2.3)$$

O objetivo do neurônio é o de que os pesos sejam escolhidos de forma que ao se entrar com valores, o valor da saída seja o mais próximo possível do valor desejado.

A junção de vários neurônios artificiais forma uma Rede Neural.

Em 1958, Frank Rosenblatt propôs o que é o modelo de Redes Neurais mais simples, onde várias unidades de processamento estão ligados a uma única saída, como (Figura 6). Esta rede, chamada de rede *Perceptron*, mesmo simples possui

capacidade de aprendizado. A rede *perceptron* também utiliza a função degrau como função de ativação.

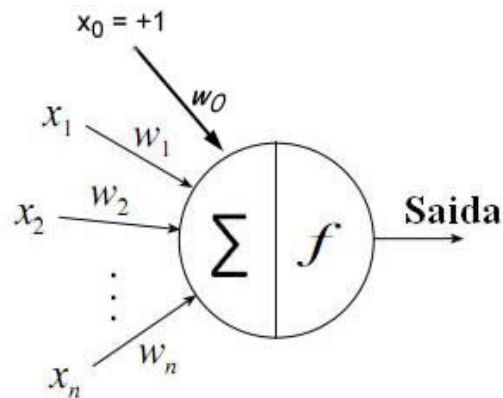


Figura 5. Neurônio artificial com *bias*.

Apesar de possuir capacidade de aprendizado, a rede *perceptron* possui limitações:

- Por possuir saídas com valores discretos, tem dificuldade em trabalhar com valores que possuam ruídos;
- Não possui capacidade de classificar dados não-linearmente separáveis.

Widrow e Hoff apresentaram um novo modelo de RNA chamada de ADALINE (ADAPtative LINear Element), no qual possibilitou a utilização de uma função contínua na saída da rede.

O processamento do sinal na rede ADALINE ocorre da mesma forma que na rede *Perceptron*, porém a diferença está no cálculo da saída, pois é possível utilizar várias funções de ativação, das quais as mais utilizadas são Linear (equação 2.4), Sigmoidal Logística (equação 2.5) e Tangente Hiperbólica (equação 2.6).

Tanto a rede *perceptron* quanto a ADALINE precisam ser treinadas para que os pesos de seus neurônios sejam ajustados.

$$y = f(\text{net}) = \text{net} \quad (2.4)$$

$$y = f(\text{net}) = \frac{1}{1 + e^{-\text{net}}} \quad (2.5)$$

$$y = f(\text{net}) = \frac{e^{\text{net}} - e^{-\text{net}}}{e^{\text{net}} + e^{-\text{net}}} \quad (2.6)$$

A utilização de RNAs com apenas duas camadas, uma de entrada e outra de saída, não era suficiente para resolver alguns tipos de problemas mais complexos. A introdução de mais uma camada foi necessária, dando origem à rede MLP, ou *MultiLayer Perceptron*.

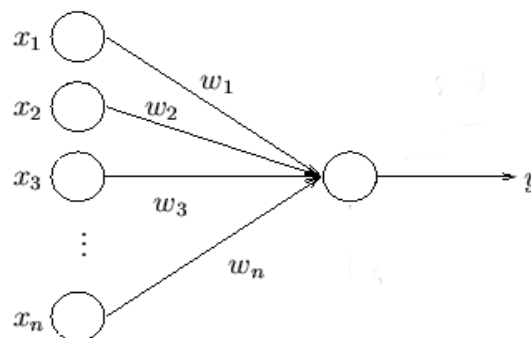


Figura 6. Rede *Perceptron*.

2.2.3 A Rede *MultiLayer Perceptron* (MLP)

Para se caracterizar como uma rede MLP, esta deve possuir no mínimo três camadas, uma de entrada, uma ou mais camadas intermediárias e uma camada de saída (Figura 7).

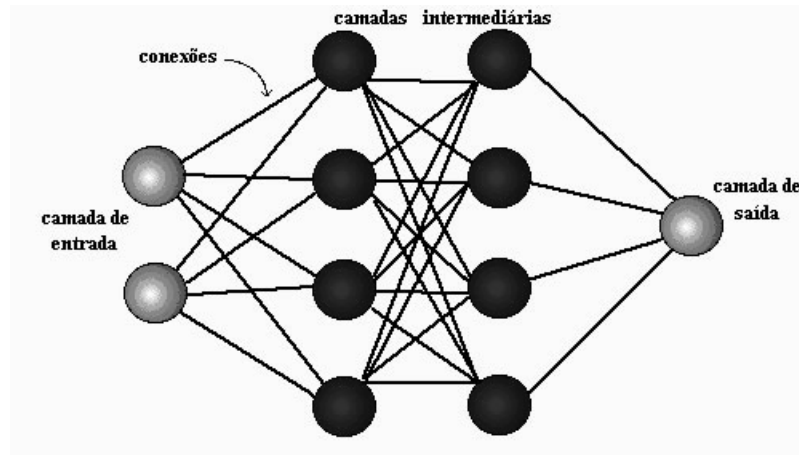


Figura 7. Rede MLP com uma camada de entrada, duas intermediárias e uma de saída.

O treinamento de uma rede MLP é bem mais complexo que o de redes sem camada intermediária. O algoritmo tradicionalmente utilizado para esse fim é conhecido como *backpropagation*, que tem esse nome pelo fato de que durante o treinamento há a propagação e retropropagação do sinal. Sendo assim o *backpropagation* é realizado em duas fases: a fase *Forward* (propagação do sinal) e a fase *Backward* (retropropagação do sinal).

Os valores de entrada líquida dos neurônios da camada intermediária são utilizados como parâmetros pela função de ativação escolhida para o treinamento (Equação 2.6).

$$y^1 = f_1(\text{net}_1^1) \quad (2.6)$$

O resultado da função de ativação para cada neurônio intermediário será então utilizado para o cálculo da entrada líquida dos neurônio da camada de saída, e da mesma forma que a camada intermediária a função de ativação é calculada para cada um (Equações 2.7 e 2.8). Para cada neurônio de saída calcula-se o erro de saída o qual é a diferença entre o valor desejado e o valor calculado, como mostra a equação 2.9.

$$\text{net}_i^2 = \sum_{j=0}^n w_{ij} x_j \quad (2.7)$$

$$y^2 = f_1(\text{net}_1^2) \quad (2.8)$$

$$e_i = (d_i - y_i) \quad (2.9)$$

A fase *Backward* é aquela em que os pesos são ajustados. O ajuste é feito utilizando-se o valor da sensibilidade dos neurônios. Para os neurônios da camada de saída a sensibilidade é calculada pelo produto entre o erro de saída e a derivada da função de ativação (Equação 2.10) e para os neurônios da camada escondida, o cálculo é realizado utilizando-se a equação 2.11, onde no lugar do erro de saída utiliza-se o somatório do produto entre a sensibilidade dos neurônios de saída e seus respectivos pesos.

$$\delta_j = (d_j - y_j) f'(\text{net}_j) \quad (2.10)$$

$$\delta_j = f'(\text{net}_j) \sum \delta_{jw_{ij}} \quad (2.11)$$

A fórmula geral que atribui um novo valor a um determinado peso pode ser vista na Equação 2.12, onde: $w_{ij}(\text{novo})$ é o novo valor atribuído ao i -ésimo peso do neurônio j , $w_{ij}(\text{antigo})$ é o valor corrente do i -ésimo peso do neurônio j , α é a taxa de aprendizado, x_i é o sinal de entrada do i -ésimo neurônio, β é a taxa de *momentum*.

$$w_{ij}(\text{novo}) = w_{ij}(\text{antigo}) + \alpha (d_o - y_o) x_j y_i y'_i \quad (2.12)$$

A taxa de aprendizado (α) é uma constante de proporcionalidade. Quanto maior for essa constante, maior será a mudança nos pesos, aumentando a velocidade do aprendizado, o que pode levar a uma oscilação do modelo na superfície de erro. Seus valores devem estar no intervalo [0,1].

O *momentum* (β) é a constante que determina o efeito das mudanças passadas dos pesos na direção atual do movimento no espaço de pesos. Este valor foi adicionado

de forma empírica e tem por objetivo ajudar o algoritmo backpropagation a escapar de mínimos locais sem ter que se aumentar a taxa de aprendizagem.

2.3 Critério de Parada: validação cruzada

Para que a rede não seja treinada de forma exaustiva, é necessária a aplicação de algum critério de parada. Na grande maioria dos casos a técnica de validação cruzada resulta em treinamentos satisfatórios.

Durante a fase de validação, a qual ocorre a cada final de ciclo de treinamento, parte da base que foi dividida (25%) na fase de pré-processamento é utilizada para o teste de parada. Como o conjunto de validação cruzada nunca é utilizado para corrigir os pesos, eles são sempre inéditos para a rede. Dessa maneira, enquanto o erro de validação cruzada estiver diminuindo, a rede continua generalizando.

A partir do momento que o erro de validação cruzada começa a aumentar seguidamente, mesmo que simultaneamente o erro de treinamento diminua, a rede pode ter deixado de aprender e passou a decorar os padrões de treinamento. Neste momento o treinamento deve ser parado (Figura 8). Posteriormente, a rede é submetida ao conjunto de teste ou verificação que será abordado na seção 2.4.

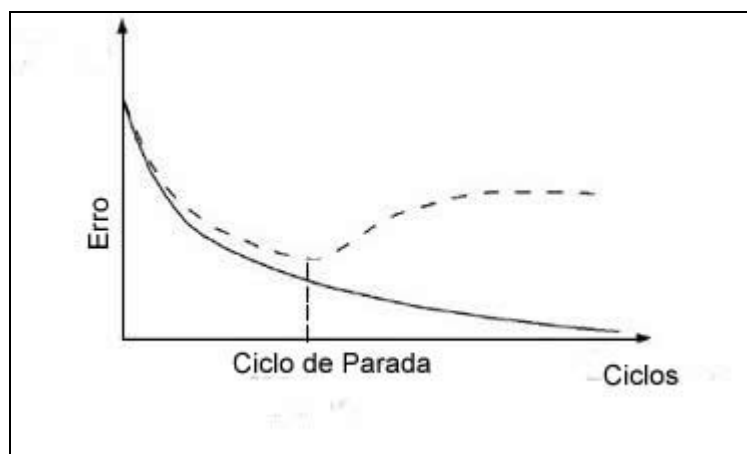


Figura 8. Validação cruzada: ponto de parada.

2.4 Verificação

A fase de verificação tem a função de indicar se o treinamento da rede foi satisfatório, ou seja, para avaliar o desempenho do modelo ajustado.

Diversas métricas podem ser utilizadas para avaliação do desempenho de uma rede. A escolha de cada uma delas depende do problema que está sendo abordado.

Para problemas de classificação é mais comumente usada a taxa de acerto, que é a quantidade de acertos obtidos dividido pelo total de registros que estão sendo usados na verificação.

Em problemas de previsão, várias outras métricas podem ser utilizadas como o Erro Médio Quadrático (EMQ), o Erro Médio Absoluto, o Erro Percentual Médio Absoluto (EPMA), entre outros. O cálculos do EMQ, EMA e EPMA são apresentados, respectivamente, pelas equações 2.13, 2.14 e 2.15, onde : Z_p é o valor previsto; Z_o é o valor obtido pela rede; e N é a quantidade de registros do conjunto de verificação.

$$EMQ = \frac{1}{N} \sum_{p=1}^N (Z_p - Z_o)^2 \quad (2.13)$$

$$EMA = \frac{1}{N} \left[\sum_{p=1}^N |Z_p - Z_o| \right] \quad (2.14)$$

$$EPMA = \frac{1}{N} \left[\frac{\sum_{p=1}^N |Z_p - Z_o|}{Z_o} \right] \quad (2.15)$$

2.5 Seleção de Variáveis

Em um problema de classificação ou predição, a resolução desse pode ser dificultada pela forma de apresentação dos dados. Em alguns casos a quantidade de dados, bem como a quantidade de atributos existentes pode aumentar significativamente o tempo de

treinamento da rede ou até torná-lo inviável. Saber quais atributos são importantes para o bom treinamento e quais deles são irrelevantes é bastante importante para evitar situações de alta dimensionalidade.

Para medir a importância de um dado atributo em um conjunto de dados podemos calcular a sua contribuição relativa e/ou o perfil de contribuição. Diversos métodos já foram propostos e vêm sendo utilizados com esse fim como o algoritmo de Garson [16] em 1991, que posteriormente foi modificado por Goh [17] em 1995 e o algoritmo de Lek et al. [18] em 1996. Em 2007 Valença e Ludemir [19] propuseram o método de Análise das Sensibilidades.

Neste trabalho serão abordados os algoritmos de Garson e o de Goh, além da metodologia da Análise de Sensibilidades.

2.5.1 Algoritmo de Garson e o de Garson Modificado

O algoritmo original proposto por Garson envolve essencialmente o particionamento dos pesos das conexões entre a camada escondida e a de saída de cada neurônio intermediário em componentes associados com cada neurônio de entrada.

Considere uma rede neural com três neurônios de entrada, três neurônios intermediários e um neurônio de saída. A tabela 1 mostra os pesos correspondentes a cada conexão da rede.

Tabela 1. Exemplo de valores para os pesos das conexões de uma rede com 3 entradas, três neurônios intermediários e uma saída.

Neurônios / Pesos	Entrada 1	Entrada 2	Entrada 3	Saída
Escondido 1	-1.67624	3.29022	1.32466	4.57857
Escondido 2	-0.51874	-0.22921	-0.25526	-0.48815
Escondido 3	-1.75691	-1.44702	0.58286	-2.65221

O processo de obtenção dos valores de importância de cada atributo ocorre como segue:

(1) Para cada neurônio intermediário i , o valor absoluto do peso da conexão entre este neurônio e um de saída é multiplicado pelo valor absoluto do peso da conexão entre o mesmo neurônio escondido e um neurônio de entrada. Este cálculo deve ser feito para todos os j – ésimos neurônios da camada de entrada. Assim para o dado exemplo os produtos P_{ij} a seguir são obtidos:

$$P_{11} = 1.67624 \times 4.57857; P_{12} = 3.29022 \times 4.57857; P_{13} = 1.32466 \times 4.57857$$

$$P_{21} = 0.51874 \times 0.48815; P_{22} = 0.22921 \times 0.48815; P_{23} = 0.25526 \times 0.48815$$

$$P_{31} = 1.75691 \times 2.65221; P_{32} = 1.44702 \times 2.65221; P_{33} = 0.58286 \times 2.65221$$

(2) Para cada neurônio escondido, divide-se P_{ij} pela soma de todos os P_{ij} para cada neurônio de entrada, obtendo Q_{ij} . Por exemplo, para o neurônio escondido 1, $Q_{11} = P_{11}/(P_{11} + P_{12} + P_{13})$. Assim:

$$Q_{11} = 0.26645; Q_{12} = 0.522994; Q_{13} = 0.210560$$

$$Q_{21} = 0.517081; Q_{22} = 0.228478; Q_{23} = 0.254441$$

$$Q_{31} = 0.645489; Q_{32} = 0.341388; Q_{33} = 0.013123$$

(3) Para cada neurônio de entrada, os valores de Q_{ij} são somados obtendo-se S_j . Como exemplo, $S_1 = Q_{11} + Q_{21} + Q_{31} + Q_{41}$

(4) Dividindo-se cada valor de S_j pela soma de todos os valores de S_j , obtemos a importância relativa R para cada variável. Assim, expressando em porcentagem: $R(\%) = (S_1 \times 100) / (S_1 + S_2 + S_3 + S_4)$

O algoritmo de Garson modificado por Goh difere do original por não passar pelo processo **(1)** mostrado acima. O cálculo dos Q_{ij} é feito sem os produtos dos pesos. Essa alteração tem dado resultados muito semelhantes aos do algoritmo original, com a vantagem de eliminar um passo, tornando a execução mais simples e rápida.

2.5.2 Análise de Sensibilidades

Por utilizar os valores absolutos dos pesos das conexões, tanto o algoritmo original de Garson quanto o de Goh não permitem uma análise da direção das modificações ocorridas na variável de saída quando ocorre alteração nas variáveis de entrada.

O método de Análise de Sensibilidades permite eliminar algumas variáveis de entrada consideradas desnecessárias, além de se determinar a importância de cada uma dessas variáveis em relação à saída.

Toma-se por base as derivadas parciais da variável de saída em relação aos pesos das conexões. Estas derivadas são na verdade as sensibilidades dos neurônios de saída que foram calculadas no treinamento da rede.

A equação (2.16) representa o cálculo da sensibilidade para cada variável de entrada, onde: Sen_{jL} é a sensibilidade para cada variável j ($j = 1, \dots, N_{inp}$) de entrada com relação à saída para um dado exemplo L , que varia de 1 a N , sendo N o número total de exemplos de treinamento; N_{hid} é a quantidade de neurônios da camada escondida e N_{out} é o número de neurônios da camada de saída; e_i é o erro para o i -ésimo neurônio da camada de saída; w_{ik} são os pesos das conexões entre a camada de saída e a camada escondida; w_{kj} são os pesos das conexões entre a camada escondida e a camada de entrada; $f'(net_i)$ é a derivada da função de ativação dos neurônios da camada de saída e $f'(net_k)$ é a derivada da função de ativação dos neurônios da camada escondida.

$$Sen_{jL} = \sum_{k=1}^{N_{hid}} w_{kj} f'(net_k) \cdot \sum_{i=1}^{N_{out}} w_{ik} f'(net_i) \cdot e_i \quad (2.16)$$

Na equação 2.17 podemos observar que $f'(net_i) \cdot e_i$ é a fórmula para o cálculo da sensibilidade δ_i para um neurônio de saída, e que $f'(net_k) \cdot \sum_{i=1}^{N_{out}} w_{ik} f'(net_i) \cdot e_i$ resulta na sensibilidade δ_k de um neurônio intermediário. Assim podemos escrever a equação 2.17.

$$Sen_{jL} = \sum_{k=1}^{N_{hid}} w_{kj} \cdot \delta_k \quad (2.17)$$

A equação 2.18 é utilizada para se calcular a contribuição ou importância $cont_j(\%)$ de cada variável de entrada com relação à variável de saída.

$$cont_j(\%) = \frac{\sum_{L=1}^N Sen_{jL}^2}{\sum_{j=1}^{N_{inp}} \sum_{L=1}^N Sen_{jL}^2} \quad (2.18)$$

Através dos resultados obtidos podemos observar duas características: a primeira é o perfil das variações da saída para pequenas mudanças em cada variável de entrada e a segunda é a classificação da contribuição relativa de cada variável para a saída da rede.

3 Metodologia

Este capítulo descreve o processo de desenvolvimento e aplicação das teorias apresentadas no Capítulo 2, em cima do problema que o projeto corrente se propôs a resolver. Inicialmente, na Seção 3.1, a implementação e configuração necessária para a simulação de uma Rede Neural são descritas na sub-seção 3.1.1, enquanto na sub-seção 3.1.2 é abordado o pré-processamento dos dados. A seção 3.2 discorre sobre as técnicas de Seleção de Variáveis explanadas na seção 2.4. Na seção 3.3 relata as fases de testes de validação dos algoritmos da RNA MLP e das técnicas de seleção de variáveis.

A seção 3.4, enfim, mostra a utilização das técnicas implementadas e validadas até então, no problema de lesões em atletas.

3.1 Previsão e Classificação com Rede Neural MLP

3.1.1 Implementação e Configuração da Rede

Para realizar as tarefas de classificação ou previsão, as quais foram utilizadas neste trabalho, foi implementado o algoritmo de uma Rede Neural MLP com *Backpropagation* (Seção 2.2.3), sendo a linguagem Java escolhida para tal propósito.

Para a efetiva execução do treinamento alguns parâmetros devem ser configurados, seja automaticamente, ou manualmente. Abaixo a Tabela 1 informa os parâmetros configuráveis do programa, indicando quais são automáticos ou manuais.

Tabela 2. Parâmetros configuráveis para a Execução da Rede

Inicializados Manualmente	Inicializados Automaticamente
Base de Dados.	<i>Momentum.</i>
Taxa de Aprendizado.	Função de Ativação.

Número mínimo e número máximo de ciclos.	Número de neurônios na camada de entrada e na camada de saída.
Objetivo do uso da rede: classificação ou para predição.	Pesos dos neurônios.
Número de Neurônios na camada escondida. (apenas uma camada).	Número de Neurônios na camada escondida. (apenas uma camada).

Ao se escolher a base de dados que será utilizada no treinamento, a quantidade de neurônios na camada de entrada e de saída são calculados de acordo com a quantidade de atributos e a quantidade de saídas inseridas no arquivo. Através dessas quantidades, o número de neurônios na camada escondida é calculado pela equação 3.1.

$$N_{hid} = \frac{N_{out} + N_{inp}}{2} \quad (3.1)$$

, onde N_{hid} é o número de neurônios da camada escondida; N_{out} é o número de neurônios da camada de entrada e N_{inp} representa a quantidade de neurônios da camada de entrada. Caso o número de neurônios da camada intermediária seja considerado insuficiente, a quantidade pode ser configurada manualmente.

Na maioria dos casos uma taxa de aprendizado igual a 0,05 rendia bons resultados no treinamento.

Os números mínimo e máximo de ciclos servem apenas para controlar o tempo de treinamento. A determinação desses valores é puramente empírica, podendo variar para cada caso.

A função de ativação escolhida foi a sigmoideal logística (Equação 2.4), pois esta vem se mostrando ser eficiente para a maioria dos casos de classificação e previsão com Redes MLP.

Indicar qual o objetivo do uso da rede, podendo ser para classificação ou previsão, serve para indicar qual tipo de teste de verificação será realizado. Assim, caso o problema seja de classificação a verificação será feita pela quantidade de acertos obtidos nesta fase. Se for um problema de previsão será calculado o erro médio quadrático (EMQ).

Os pesos das conexões são inicializados por uma função que gera valores aleatórios.

3.1.2 Pré-processamento dos dados

Ao se escolher o arquivo de dados este tem seus registros divididos de forma aleatória para serem utilizados nas fases de execução da rede, seguindo a distribuição de 50% para a fase de treinamento, 25% para a fase de validação cruzada e os outros 25% restantes para a fase de verificação.

Em geral os dados de entrada estão em intervalos de variação bastante distintos de variação de modo que a modelagem da rede ficará prejudicada, pois esta pode interpretar valores mais altos como de maior importância [20]. Nesses casos torna-se necessária a normalização dos dados para evitar distorções. Das diversas técnicas de normalização, a mais usada e que geralmente é suficiente, é a transformação linear, cuja fórmula está representada na equação (3.2).

$$y = (b - a) \frac{x_i - x_{\min}}{x_{\max} - x_{\min}} + a \quad (3.2)$$

, onde y é o valor normalizado; x_i é o valor original; x_{\min} e x_{\max} são, respectivamente, o valor mínimo e o valor máximo do conjunto de dados para determinado atributo; a é o limite inferior e b é o limite superior.

Esta normalização faz com que todos os atributos fiquem com valores no intervalo da função de ativação.

Para os experimentos deste trabalho foram utilizados como limite inferior e superior, respectivamente, 0.1 e 0.9. Estes valores são estabelecidos de forma a se evitar saturação da rede (valores muito próximo a zero ou um).

Além do problema da variabilidade dos dados, os diferentes tipos de dados devem ser tratados. No caso de existirem dados nominais, a técnica mais comumente utilizada é a da binarização, que consiste em atribuir números binários aos valores nominais.

A tabela 3 mostra um exemplo para o conjunto de dados da base das espécies-íris. Como se vê há atributos com diferentes intervalos de variação, além de um dos atributos ser nominal.

Tabela 3. Pequena amostra de dados da base das Espécies-Íris. Esta base deve ser normalizada e passar por um processo de binarização.

<i>Sepal Length</i>	<i>Sepal Width</i>	<i>Petal Length</i>	<i>Petal Width</i>	<i>Classe</i>
5.1	3.5	1.4	0.2	Iris-setosa
4.9	3.0	1.4	0.2	Iris-setosa
4.6	3.4	1.4	0.3	Iris-setosa
5.0	3.4	1.5	0.2	Iris-setosa
5.6	2.7	4.2	1.3	Iris-versicolor
5.7	3.0	4.2	1.2	Iris-versicolor
5.7	2.9	4.2	1.3	Iris-versicolor
6.2	2.9	4.3	1.3	Iris-versicolor
6.3	3.3	6.0	2.5	Iris-virginica
5.8	2.7	5.1	1.9	Iris-virginica
7.1	3.0	5.9	2.1	Iris-virginica
6.3	2.9	5.6	1.8	Iris-virginica
4.9	2.5	4.5	1.7	Iris-virginica

Ao término da execução, os pesos e as sensibilidades obtidos na fase treinamento são gravados em dois arquivos distintos, os quais serão utilizados pelo módulo de seleção de variáveis.

3.1.3 Validação do algoritmo da RNA MLP e base utilizada

Para testar o algoritmo de Rede MLP com *Backpropagation* foi utilizada a base de dados das espécies Íris, obtida no repositório da UCI *Machine Learning* [21], a qual foi escolhida por já se ter um conhecimento prévio das características dessa base.

A base de Espécies-Íris é composta por 150 registros. Cada registro tem 4 atributos: *sepal length*, *sepal width*, *petal length* e *petal width* e uma classe que pode ser *Íris-setosa*, *Íris-virginica* ou *Íris-versicolor*. A distribuição é de 50 registros pertencentes a

cada classe. Para que o algoritmo seja considerado com funcionamento adequado espera-se que a taxa de acerto após o treinamento seja de no mínimo 97%, sendo esse o valor obtido pelo algoritmo de Redes Neurais do software Weka [22].

3.2 Seleção de Variáveis

Após a execução satisfatória do treinamento da rede, podemos dar início à fase de seleção das variáveis.

Os algoritmos aqui abordados foram o de Garson original, o de Garson Modificado (Goh) e o da Análise de Sensibilidades.

O algoritmo de Garson original e o de Garson Modificado utilizam os pesos que foram gravados na fase de treinamento da rede. Já o algoritmo de Análise de Sensibilidades utiliza tanto os pesos, como também as sensibilidades.

Para proceder com a seleção de variáveis seguindo um dos três métodos abordados na seção 2.5, o algoritmo captura os valores dos pesos e/ou das sensibilidades resultantes da fase de treinamento e que foram gravados em arquivos separados. Os devidos cálculos são feitos e os valores das contribuições/importâncias para cada atributo é exibido.

3.2.1 Experimentos com os algoritmos de Seleção

Cada algoritmo de seleção foi submetido a testes. Nesta fase espera-se encontrar o algoritmo que possa dar melhores resultados de seleção. Os experimentos foram realizados utilizando-se, primeiramente, equações e, logo após, duas bases de dados obtidas no site da UCI *Machine Learning*.

- **Utilizando Equações.**

Para testar os algoritmos de seleção foram utilizados o polinômio multivariado, o polinômio de Kolmogorov – Gabor (equação 3.4) e a função com interação complexa (equação 3.5).

$$f(x_1, x_2, x_3, x_4, x_5) = 2 + 3x_1 x_2 + 4x_3 x_4 x_5 \quad (3.3)$$

$$f(x_1, x_2) = \frac{x_1^4 - 16x_1^2 + 5x_1}{2} + \frac{x_2^4 - 16x_2^2 + 5x_2}{2} \quad (3.4)$$

$$f(x_1, x_2) = 1,9 \left(1,35 + e^{x_1} \text{seno} \left(13(x_1 - 0,6)^2 \right) e^{-x_2} \text{seno} (7x_2) \right) \quad (3.5)$$

A validação consiste em gerar para cada função as variáveis que fazem parte desta e variáveis que não têm qualquer influência sobre o resultado da função. Ou seja, para o polinômio de Kolmogorov-Gabor, por exemplo, foram gerados 500 valores aleatórios para as duas variáveis pertencentes ao polinômio, calculamos o valor da função para cada conjunto de valores e depois geramos outras cinco variáveis que não possuem qualquer influência sobre o polinômio.

Para as funções em 3.3 e 3.5 foram gerados também 500 valores aleatórios para as duas variáveis pertencentes a cada uma destas funções. Procedeu-se o cálculo do valor de cada uma delas para cada conjunto de valores e depois outras variáveis não influentes foram geradas.

Para estar correto, espera-se que o algoritmo de seleção atribua maior importância às variáveis que determinam o valor da função e a mínima importância possível para as outras variáveis.

- Base de Dados das Espécies-Íris e de Vinhos

A mesma base utilizada na sub-seção anterior serviu para a validação dos algoritmos de seleção, e além dela foi utilizada a base de classificação de vinhos também obtida no repositório da UCI *Machine Learning*.

A base de classificação de vinhos contém dados da análise química de alguns tipos da bebida os quais são classificados em três classes. Os atributos são de um total de 13: *Malic_acid*, *Flavonoids*, *Nonflavonoid_phenols*, *Ash*, *Proline*, *Alcalinity_of_ash*, *Magnesium*, *Alcohol*, *Total_phenols*, *OD*, *Proanthocyanins*, *Proanthocyanins* e *Hue*. A figura 9 contém o gráfico da distribuição dos registros de vinhos nas três classes.

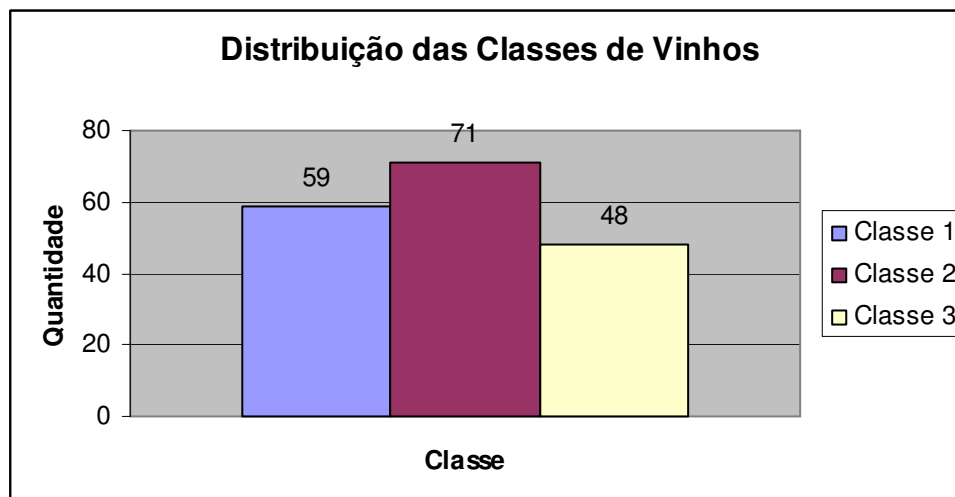


Figura 9. Gráfico da distribuição das classes de vinhos.

Para cada uma das duas bases foi realizado o treinamento da RNA. Após essa fase, cada um dos algoritmos de seleção foi executado. De acordo com os resultados obtidos, novos treinamentos foram realizados, sendo retirados do conjunto de dados aqueles atributos que foram considerados menos relevantes pelo algoritmo.

O que se espera do algoritmo de seleção é que após um novo treinamento, com a nova configuração dos atributos, durante a verificação se obtenha uma taxa de acertos, no mínimo, próxima ao que se conseguiu utilizando todas as entradas dos dados.

3.3 Estudo da contribuição dos fatores no surgimento de lesões em atletas

A base de dados que será utilizada foi resultado da pesquisa de campo realizada por Pastre. Consiste em 215 registros de atletas da elite brasileira, sendo estes distribuídos em cinco diferentes categorias: velocistas, fundistas, nadador categoria borboleta, nadador categoria peito e nadador categoria costas. A distribuição entre essas modalidades é exibida na figura 10. Cada registro possui 9 atributos. A tabela 5 exibe os atributos de cada registro.

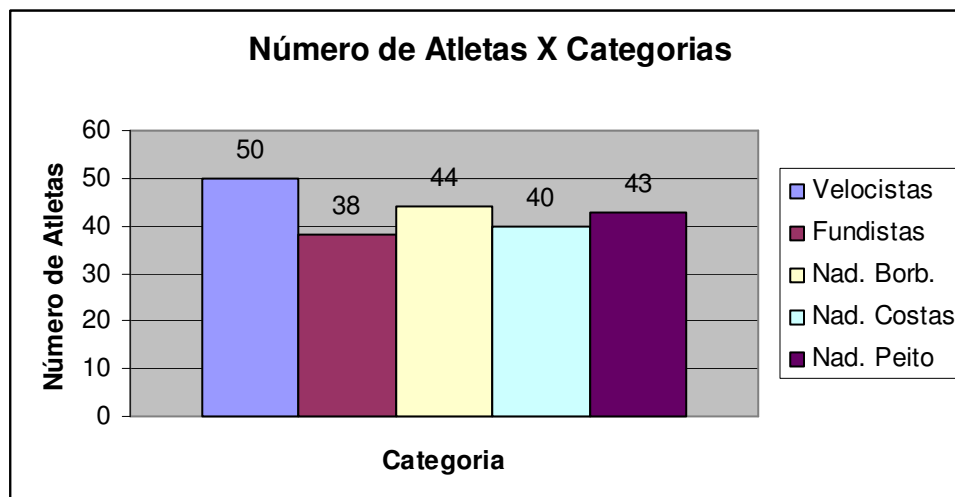


Figura 10. Gráfico da distribuição dos Atletas nas cinco categorias pesquisadas.

Tabela 4. Atributos da Base de Dados da Elite do Atletismo Brasileiro

Nome do Atributo	Descrição
Idade	Na base os valores para esse atributo vão de 16 a 34.
Sexo	Feminino – 1; Masculino - 2
Índice de Massa Corporal - IMC	É o quociente do peso sobre a altura ao quadrado.
Peso	Peso do atleta em quilos.
Estatura	Altura do atleta em metros.
Tempo de prática	Tempo em que o atleta pratica a modalidade, em anos.
Categoria	Modalidade esportiva praticada pelo atleta
Frequência de Ocorrência	Frequência da Ocorrência de Lesões durante o treinamento
Horas	Horas de treino por dia
Lesão	Refere-se à existência de lesão. (1 ou 0)

Dos 215 atletas podemos observar algumas características:

- 80 são do sexo feminino e 135 do sexo masculino;
- têm idades estão entre 16 e 34 anos;
- a média de tempo de prática da modalidade em anos é de 13 anos.
- a média de horas de prática por dia é de 2,77 horas;
- 16 atletas têm IMC > 25, o que de acordo com a Organização Mundial de Saúde (OMS) indica sobrepeso [24].

As figuras 11, 12 e 13 ilustram, respectivamente a distribuição de atletas por faixas etárias, por tempo de treino diário em horas, e por tempo de prática do esporte em anos.

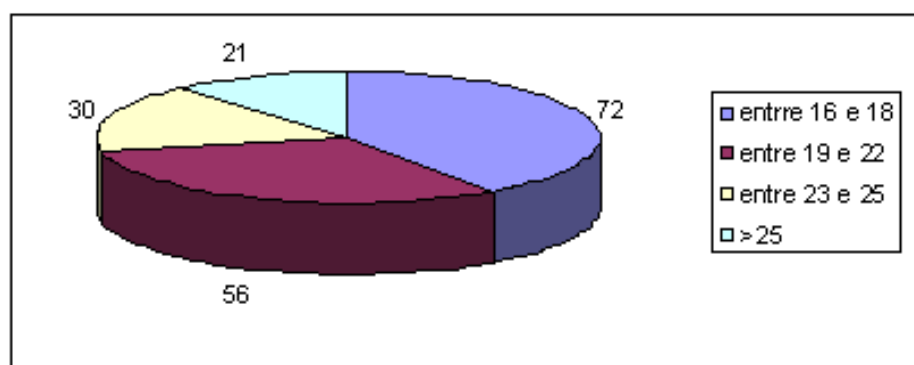


Figura 11. Distribuição dos atletas por faixa etária. Observa-se maior quantidade de atletas na faixa dos 16 a 18, enquanto que a quantidade de atletas com mais de 25 anos diminui bastante.

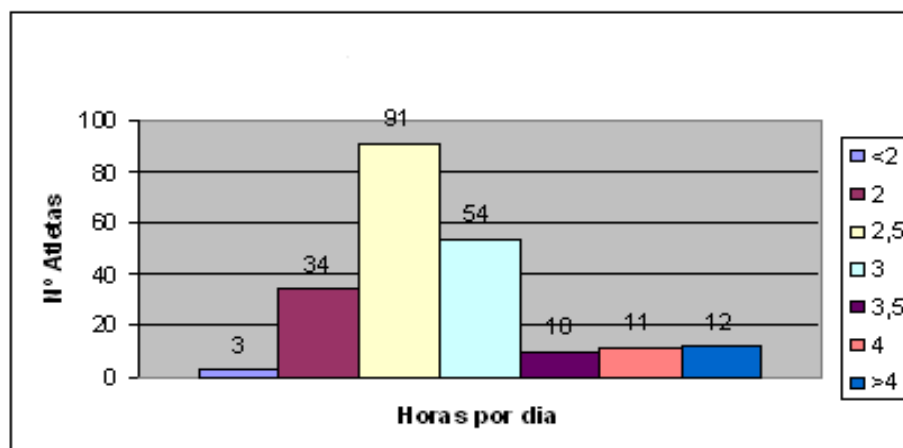


Figura 12. Gráfico da Quantidade de Atletas X Horas de Treino Diário.

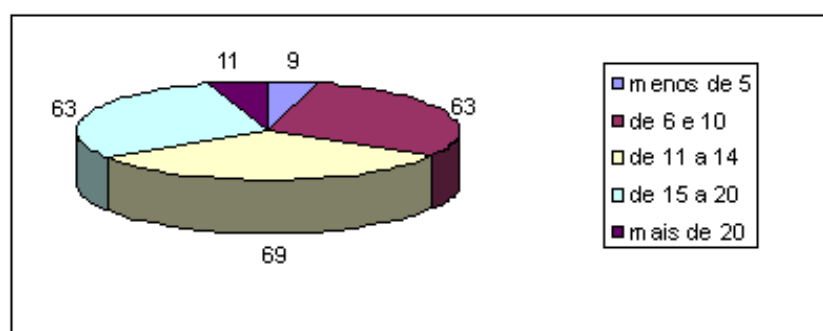


Figura 13. Tempo de prática de modalidade em anos.

Tendo esses dados em mãos, utilizaremos o algoritmo de seleção testado na seção 3.3 que mostrou o melhores resultados. Após esta ação, espera-se encontrar os fatores que são mais relevantes no surgimento de lesões em atletas.

4 Resultados

Este capítulo mostra os resultados obtidos a partir da aplicação da metodologia explicada no Capítulo 3.

4.1 Resultados da RNA MLP.

A avaliação do algoritmo da Rede Neural utilizou duas métricas, Erro Médio Quadrático para previsão e taxa de acertos para classificação (seção 2.4).

Como explanado na seção 3.1, a base utilizada para este teste foi a base das espécies-íris.

Durante os testes os valores da taxa de aprendizado e as quantidades mínima e máxima de ciclos foram alterados. Após diversas tentativas a rede convergiu de forma a encontrar como o melhor resultado 98,5% de acerto durante a verificação. A configuração dos parâmetros que resultou nesta taxa é a que está na tabela 5. A figura 14 mostra o gráfico de convergência do treinamento. O fim do treinamento aconteceu no ciclo 209.

Tabela 5. Configuração dos parâmetros para a obtenção do melhor resultado de classificação das espécies-íris.

Taxa de Aprendizado	0.05
Nº Mínimo de ciclos	200
Nº Máximo de ciclos	900
Quantidade de Neurônios na camada escondida	3 (ajustado manualmente)
Objetivo de uso da Rede	Classificação

Validação Cruzada

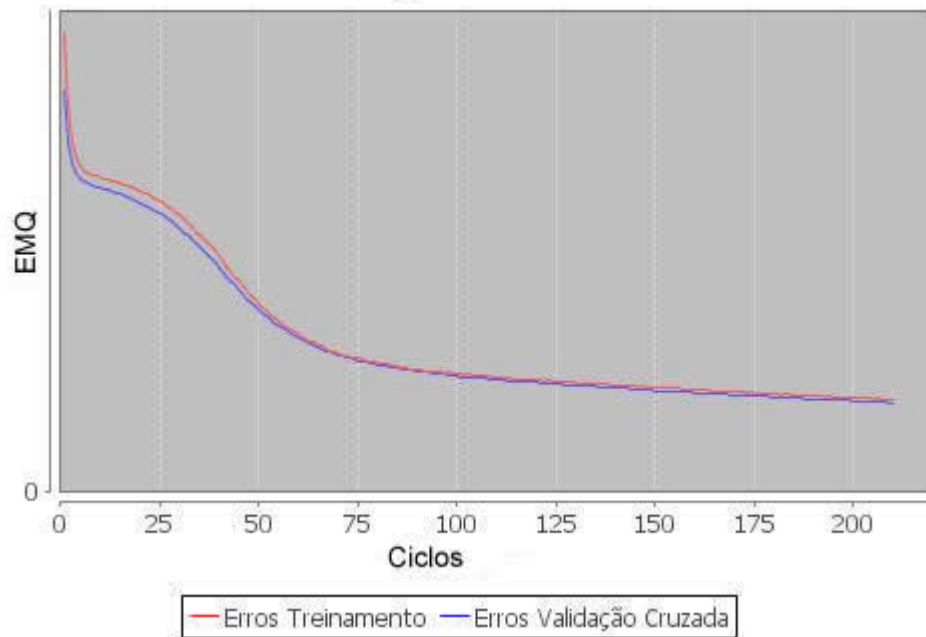


Figura 14. Gráfico da convergência do treinamento com a base das espécies-íris. O eixo vertical indica o valor do Erro Médio Quadrático (EMQ) para o ciclo correspondente no eixo horizontal.

Com o resultado obtido, o algoritmo da RNA MLP foi considerado satisfatório. Assim, foi possível prosseguir com os experimentos utilizando a rede implementada.

4.2 Resultados dos algoritmos de seleção de variáveis

Nesta fase são mostrados os resultados encontrados utilizando os algoritmos de seleção de variáveis explanados no capítulo 3.

4.2.1 Resultados obtidos com as equações

- O Polinômio Multivariado

Para este caso o melhor resultado de treinamento foi encontrado com uma taxa de aprendizagem de 0.05, mínimo de 300 e máximo de 900 ciclos. Foram utilizados 5

neurônios na camada escondida, quantidade obtida através da equação 3.1. O EMQ resultante foi de 0.00426.

Os valores de importância relativa atribuídos por cada algoritmo às entradas da rede são exibidos na tabela 6. É possível observar que o resultado para os algoritmos de Garson original e modificado são idênticos. Assim, o gráfico da figura 15a representa os dados para esses dois algoritmos de Garson, enquanto que o gráfico na figura 15b é formado à partir dos valores encontrados pela análise de sensibilidades.

Tabela 6. Resultados encontrados pelos algoritmos de seleção de variáveis para a o polinômio multivariado. As linhas com fundo cinza contém as variáveis que fazem parte da equação.

Importância Relativa(%)			
Variável	Garson Original	Garson Modificado	Análise de Sensibilidades
x1	16,71	16,71	35,69
x2	16,51	16,51	30,65
x3	12,83	12,83	12,14
x4	11,21	11,21	9,96
x5	11,48	11,48	10,93
x6	6,68	6,68	0,0041
x7	6,91	6,91	0
x8	7,38	7,38	0,545
x9	6,64	6,64	0,03
x10	3,61	3,61	0,04

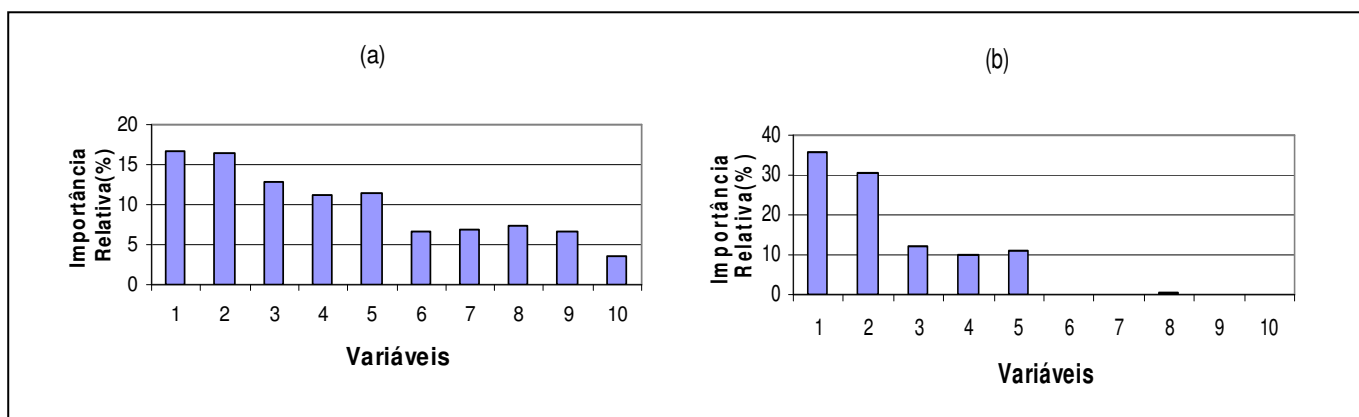


Figura 15. Gráficos da seleção de variáveis para o polinômio multivariado. (a) Gráfico da importância para cada variável obtido com os algoritmos de Garson original e modificado. (b) Gráfico dos valores de importância obtidos com a análise de sensibilidades.

Os dados mostram que todos os algoritmos atribuíram maior importância aos atributos que realmente fazem parte da equação, sendo que a análise de sensibilidades conseguiu melhor distinção entre as variáveis contidas na equação e as não contidas.

- Polinômio de Kolmogorov – Gabor

O treinamento envolvendo o polinômio de Kolmogorov – Gabor teve como melhor resultado EMQ igual a 0.00498, conseguido com a seguinte configuração: taxa de aprendizado igual a 0.05; mínimo de 350 ciclos e máximo de 900 ciclos; 5 neurônios na camada escondida, configurados manualmente.

A tabela 7 mostra os valores de importância obtidos por cada algoritmo de seleção. Novamente observa-se que os valores obtidos pelo algoritmo original de Garson são idênticos aos encontrados pelo algoritmo modificado, valores os quais são ilustrados pelo gráfico da figura 16a. Na figura 16b temos o gráfico dos valores obtidos pela análise de sensibilidades.

Tabela 7. Resultados encontrados pelos algoritmos de seleção de variáveis para o polinômio de Kolmogorov - Gabor. As linhas com fundo cinza contêm as variáveis que fazem parte da equação.

Importância Relativa(%)			
Variáveis	Garson Original	Garson Modificado	Análise de Sensibilidades
x1	48,04	48,04	84,86
x2	44,31	44,31	16,00
x3	5,51	5,51	0,00005
x4	2,15	2,15	0,135

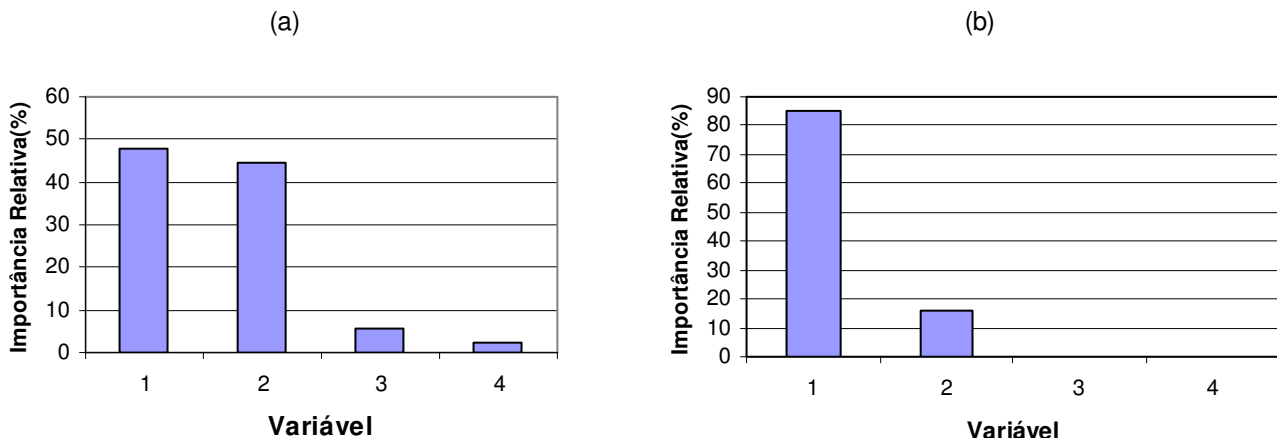


Figura 16. Gráficos da seleção de variáveis para o polinômio de Kolmogorov-Gabor. (a) Gráfico da importância para cada variável obtido com os algoritmos de Garson original e modificado. (b) Gráfico dos valores de importância obtidos com a análise de sensibilidades.

Os três algoritmos atribuíram maior importância às variáveis que faziam parte do polinômio. Enquanto a análise de sensibilidades atribuiu valores próximos de zero às variáveis que não estavam na equação e atribuiu à variável $x1$ importância muito maior que à variável $x2$, ambos os algoritmos de Garson encontraram um certo equilíbrio entre $x1$ e $x2$.

- Função com interação complexa

Em relação a esta função, o melhor resultado de treinamento foi um EMQ de 0,0175, para taxa de aprendizagem igual a 0,06, mínimo de 450 ciclos e máximo de 900. Foram utilizados 4 neurônios na camada escondida, quantidade configurada manualmente. O treinamento foi interrompido na 472º ciclo.

As figuras 17a e 17b ilustram, respectivamente os valores de importância atribuídos a cada variável pelos algoritmos de Garson e pela análise de sensibilidades.

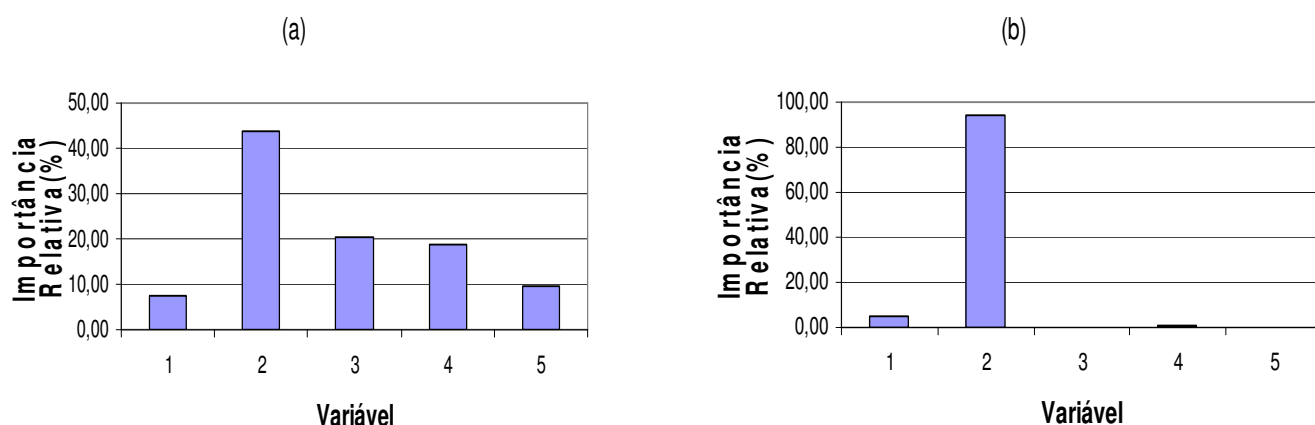


Figura 17. Gráficos da seleção de variáveis para a função com interação complexa. (a) Gráfico da importância para cada variável obtido com os algoritmos de Garson original e modificado. (b) Gráfico dos valores de importância obtidos com a análise de sensibilidades.

Também neste caso, a análise de sensibilidades distinguiu melhor as variáveis contidas na equação.

4.2.2 Resultados com as bases de dados

Nesta seção são apresentados os resultados conseguidos com as bases de dados das espécies-íris e das classes de vinhos.

Como os resultados obtidos pelo algoritmo de Garson original e o modificado foram idênticos para as equações da subseção 4.2.1, os experimentos seguintes foram realizados apenas com o algoritmo de Garson modificado e com o algoritmo de análise de sensibilidades.

- Base das Espécies-Íris

Na seção 4.1 foi visto que o melhor resultado de treinamento para esta base foi de 98,5%.

Os resultados obtidos com o algoritmo de Garson modificado e com a análise de sensibilidades estão inseridos na tabela 8.

Ambos os algoritmos encontraram um certo equilíbrio na importância relativa dos atributos *sepalwidth* e *petallength*. Podemos destacar a diferença na importância relativa atribuída por cada algoritmo aos atributos *sepalwidth* e *petalwidth*. Enquanto a análise de sensibilidades atribuiu importância 0 ao atributo *petalwidth* o algoritmo de Garson atribuiu 36,44%, sendo assim o atributo de maior importância.

Tabela 8. Resultados encontrados pelos algoritmos de Garson modificado e de análise de sensibilidades para a base de espécies-íris.

Atributo	Importância Relativa (%)	
	Garson modificado	Análise de sensibilidades
<i>sepalwidth</i>	7,85	4,35
<i>sepalwidth</i>	29,29	42,05
<i>petallength</i>	26,41	53,59
<i>petalwidth</i>	36,44	0

As figuras 18 e 19 ilustram, respectivamente, os gráficos gerados com os valores de importância relativa obtidos com o algoritmo de Garson modificado e com a análise de sensibilidades.

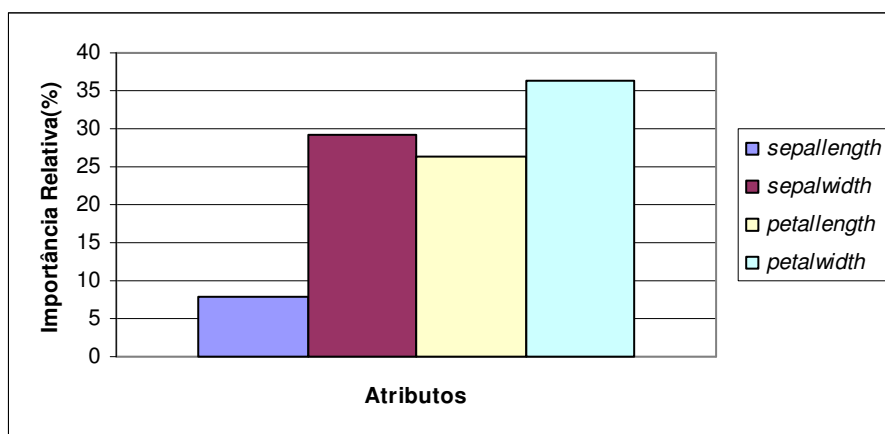


Figura 18. Gráfico de seleção de variáveis para a base de espécies-íris resultante do algoritmo de Garson.

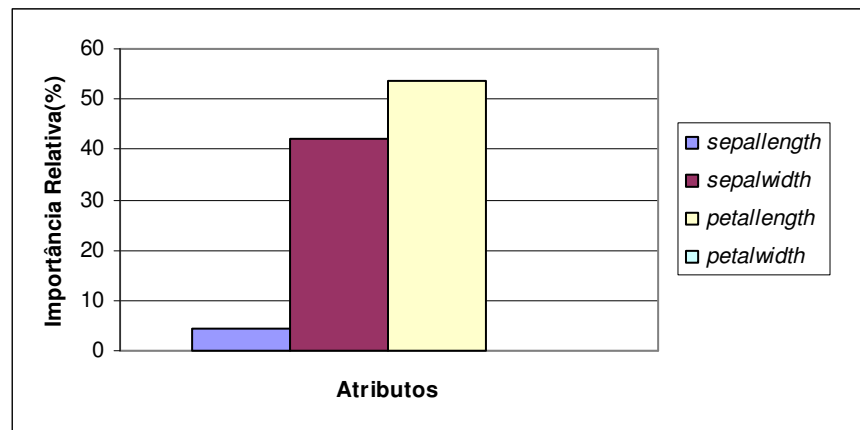


Figura 19. Gráfico de seleção de variáveis para a base de espécies-íris resultante da análise de sensibilidades.

Para comparar os resultados com os dois métodos de seleção, procedeu-se ao treinamento da rede excluindo os atributos de menor valor de importância para cada caso. Assim, a RNA MLP foi executada para a base sem o atributo *sepallength*, sendo este o atributo com menor valor de importância obtido pelo algoritmo de Garson. e para a base sem o *petalwidth* que é o atributo com menor valor de importância para a análise de sensibilidades.

Retirando o atributo *petalwidth* da base de dados, o melhor resultado obtido foi de 95% de taxa de acerto. A rede foi configurada com taxa de aprendizagem de 0.05; 4 neurônios na camada escondida, ajustados manualmente; mínimo de 350 e máximo de 900 ciclos. O treinamento foi parado no ciclo 364.

O melhor resultado obtido para a base sem o atributo *sepallength* também foi igual a 95%. A taxa de aprendizagem utilizada foi de 0.05; mínimo de 350 e máximo de 900 ciclos; 4 neurônios na camada escondida. O treinamento foi interrompido no 367º ciclo.

- Base das Classes de Vinhos

O melhor resultado de treinamento foi de 89% de taxa de acerto, para uma taxa de aprendizado de 0.06, mínimo de 400 ciclos e máximo de 900. Foram utilizados 7 neurônios na camada escondida, quantidade resultante a equação 3.1.

As tabelas 9 e 10, respectivamente, mostram os valores de importância atribuídos a cada variável pelo algoritmo modificado de Garson e pela análise de sensibilidades. As figuras 20 e 21 exibem os gráficos para estes dados.

Tabela 9. Resultados encontrados pelos algoritmos de Garson modificado e de análise de sensibilidades para a base de classes de vinhos.

Atributos	Importância Relativa(%)	
	Garson Modificado	Análise de Sensibilidades
<i>Malic_acid</i>	6,97	1,88
<i>Flavonoids</i>	7,90	0,30
<i>Nonflavonoid_phenols</i>	2,93	6,00
<i>Ash</i>	7,19	14,22
<i>Proline</i>	10,72	38,28
<i>Alcalinity_of_ash</i>	6,48	2,41
<i>Magnesium</i>	3,94	0,13
<i>Alcohol</i>	10,39	25,01
<i>Total_phenols</i>	6,37	0,24
<i>OD</i>	10,82	0,37
<i>Proanthocyanins</i>	5,63	3,13
<i>Color_Intensity</i>	11,29	7,99
<i>Hue</i>	9,37	0

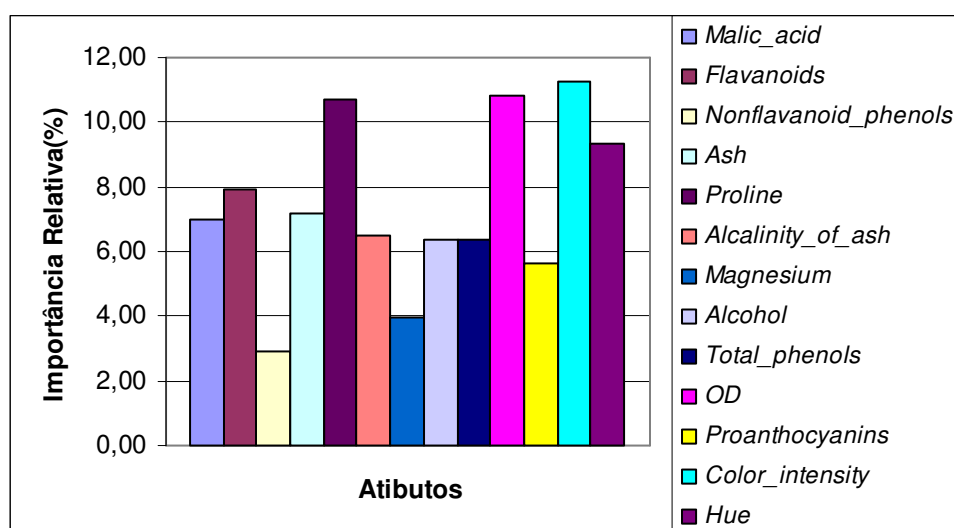


Figura 20. Gráfico de seleção de variáveis para a base de classes de vinhos resultante do algoritmo de Garson modificado.

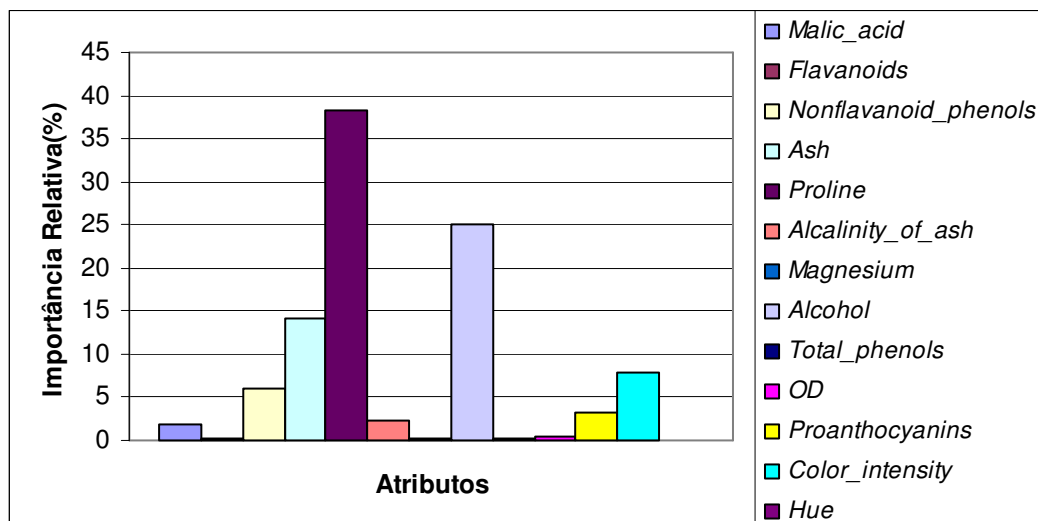


Figura 21. Gráfico de seleção de variáveis para a base de classes de vinhos resultante da análise de sensibilidades.

É possível notar que o algoritmo de Garson não descartou qualquer atributo, tendo apenas dois deles importância menor que 6%. Já na análise de sensibilidades 4 atributos tiveram importâncias próximas de zero.

4.3 Resultado da contribuição dos fatores no surgimento de lesões em atletas

Depois de uma análise sobre os resultados obtidos com os algoritmos de seleção aqui apresentados, ficou decidida a utilização do método de análise de sensibilidades, por dois motivos:

1. É possível fazer uma análise na direção das modificações ocorridas na variável de saída em relação às alterações nas entradas, pois não são utilizados os valores absolutos dos pesos das conexões, como fazem os algoritmos de Garson.
2. Nos testes com as equações, este algoritmo mostrou ser melhor em destacar as variáveis realmente importantes para cada equação.

Assim, para determinar o valor da contribuição de cada fator apresentado no favorecimento de lesões em atletas, procedeu-se o treinamento com a base de dados

referida na seção 3.3 e posteriormente aplicou-se o algoritmo de análise de sensibilidades ao resultado do treinamento.

Por ser um problema de certa complexidade, uma base de dados com uma significativa quantidade de registros seria o ideal para se obter uma taxa de acerto satisfatória. Porém, para este caso há uma grande dificuldade em se conseguir dados em boa quantidade e que estejam disponíveis para experimentos. Segundo Pastre, as investigações sobre lesões nos esportes encontram grandes dificuldades devido a aspectos como critérios de catalogação do conceito de lesão, a descrição de sua incidência, inclusão e descrição dos agravos nos protocolos de estudo, bem como a diversidade dos grupos de desportistas em cada modalidade. Além disso, não parece ser considerada importante a percepção do próprio atleta sobre sua lesão, o que de fato poderia auxiliar em futuros processos de prevenção.

Dessa forma, foi necessário realizar diversos treinamentos, testando vários valores para os parâmetros. Após alguns testes, o melhor resultado obtido foi uma taxa de acerto de 61%. A configuração da rede foi a seguinte: taxa de aprendizagem igual a 0.07; mínimo de 400 ciclos e máximo de 1000 ciclos; quantidade de neurônios da camada escondida igual a 8, calculado pela equação 3.1.

Os valores de importância atribuídos aos fatores são mostrados na tabela 10. A figura 21 mostra o gráfico formado por estes valores.

Tabela 10. Importância dos fatores para o surgimento de lesões.

Fator	Importância Relativa(%)
Idade	14,42
Sexo	13,3
Peso	4,86
Estatutura	10,86
IMC	1,52
T. Prática	36,34
Modalidade	0,39
Horas	18,33
Freqüência de Ocorr.	0

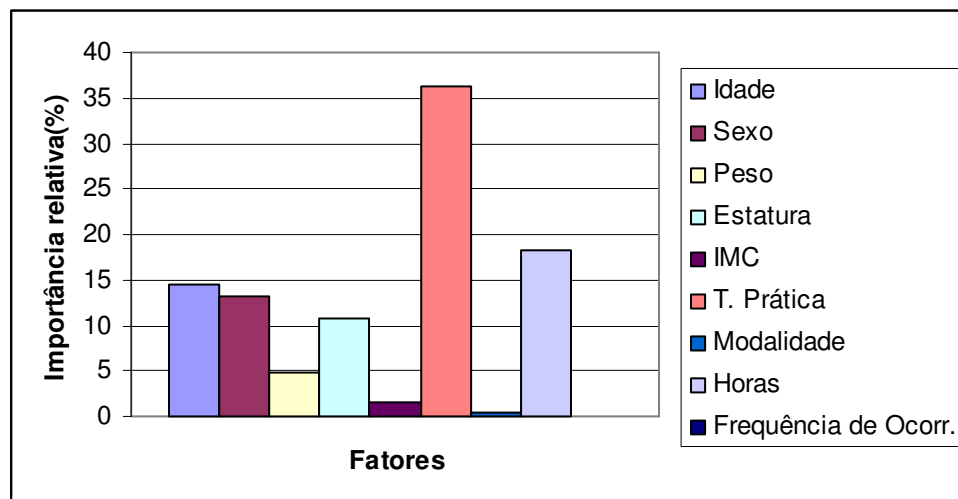


Figura 22. Gráfico da contribuição dos fatores no favorecimento de lesões em atletas.

5 Conclusão e Trabalhos Futuros

O principal objetivo desta monografia foi utilizar a técnica de rede neural, associada a métodos de seleção de variáveis para identificar a contribuição de determinados fatores no surgimento de lesões em atletas. O uso de algoritmos de seleção baseados nos pesos e sensibilidades resultantes do treinamento de uma rede neural mostra ser uma alternativa aos métodos estatísticos tradicionais, comumente utilizados pelos profissionais da área de saúde, e que apenas capturam características lineares.

Para tanto, uma RNA MLP treinada com o algoritmo *Backpropagation* foi implementada e testada utilizando a base de dados das espécies-íris. Logo após, três algoritmos de seleção foram implementados e testados a fim de se encontrar o que melhor se adequasse ao problema. Estes foram o algoritmo de Garson, o algoritmo de Garson Modificado e o algoritmo de análise de sensibilidades.

Todos os algoritmos de seleção aqui apresentados foram experimentados através da aplicação em equações e em duas bases de dados. As equações foram: o polinômio multivariado, o polinômio de Kolmogorov – Gabor e a função com interação complexa. As bases de dados foram: a base das espécies-íris e a base de classes de vinhos, ambas obtidas através do site da *UCI Machine Learning*. Para cada uma das equações foram gerados valores aleatórios para as variáveis que faziam parte de cada equação e variáveis que não tinham qualquer influência sobre a mesma. Desse modo para o algoritmo ser considerado satisfatório este deve atribuir valores de importância relativa maiores para as variáveis que fazem parte da equação. Foi observado que os algoritmos original e modificado de Garson obtinham resultados idênticos, e que a análise de sensibilidades fazia uma melhor distinção entre as variáveis contidas na equação e as não contidas. Com as bases de dados a análise de sensibilidades apresentou melhor poda nos atributos.

Como a análise de sensibilidades apresentou melhores resultados de seleção, este algoritmo foi escolhido para se determinar a contribuição de cada fator no favorecimento de lesões em atletas.

Dada a complexidade do problema, o ideal seria ter uma base de dados de tamanho considerável. Porém há uma grande dificuldade em se encontrar dados sobre lesões em atletas que tenham grande quantidade de registros e que estejam disponíveis para estudo. Assim, com a base disponibilizada para este trabalho, o melhor resultado encontrado foi de 61% de taxa de acerto.

Aplicada a análise de sensibilidades sobre o resultado do treinamento, alguns resultados podem ser discutidos. Com os resultados obtidos podemos fazer algumas observações:

- ❖ Ao fator Freqüência Ocorrida não foi atribuído qualquer valor de importância, e o fator Modalidade também foi considerado ter pouca contribuição no surgimento de lesões, este fato pode ser explicado através dados da tabela 11, em que a quantidade de atletas lesionados em cada modalidade são exibidos.

Tabela 11. Distribuição da quantidade de atletas lesionados por modalidade.

Modalidade	Atletas na modalidade	Atletas lesionados na modalidade	Percentual de lesões por modalidade
Velocistas	50	30	60%
Fundistas	37	17	46,4%
Nado Borboleta	43	23	53,48%
Nado Costas	39	24	61,54%
Nado Peito	52	27	51,9%

Observa-se que a quantidade de atletas lesionados está bem distribuída entre as modalidades, o que nos leva a concluir que para os dados apresentados as diferentes modalidades não possuem uma grande influência sobre lesionamento do atleta.

- ❖ O atributo de maior importância foi o tempo em que o atleta vem praticando o esporte, seguido do tempo em horas de treino diário. Tais resultados fazem sentido, pois segundo Hootman [25] e Hino *et al.* [26] à medida que se aumenta o tempo de treino, o risco de lesões tende a crescer.

- ❖ Os fatores Idade e Sexo foram considerados de maior importância que os fatores Peso e IMC. Este fato pode ser explicado pelo fato de que por serem atletas de alta performance, menos de 8% deles apresentou sobrepeso, tornando estes atributos irrelevantes.
- ❖ O atributo Estatura teve maior importância que Peso e IMC, porém menor importância que Sexo e Idade.

Diante do observado, os resultados de seleção dos fatores sobre o lesionamento de atletas foi considerado satisfatório. Isto mostra que a seleção de variáveis utilizando o algoritmo de análise de sensibilidades é uma boa alternativa para problemas deste tipo.

A maior dificuldade encontrada para dar início à realização deste trabalho foi a obtenção dos dados dos atletas. E após a obtenção dos dados, durante os experimentos, a atividade que mais demandou tempo foi a de implementação, devido a correções que foram necessárias. Porém com todas as dificuldades, o aprendizado foi bastante proveitoso.

Durante os testes realizados, vários parâmetros da RNA MLP, como taxa de aprendizagem, número de ciclos e quantidade de neurônios escondidos tiveram de ser ajustados a fim de se encontrar o melhor resultado. A nível de pesquisa isto é algo aceitável, porém no caso de se construir uma ferramenta para usuário final, sendo este da área de saúde, por exemplo, a tarefa de ajustar tais parâmetros torna-se um problema. Portanto, como trabalho futuro, seria interessante a busca da automação da arquitetura da rede, de forma que seus parâmetros não precisem sofrer interferência humana. Há também a possibilidade de se comparar as técnicas de seleção aqui apresentadas com outras técnicas tradicionais da literatura, as quais não trabalham com extração de conhecimento dos pesos da rede.

Referências

- [1] SOARES, S. T. M. Relação entre técnicas e lesões em praticantes de judô. Rev Bras. Med. Esporte, v.12,n.1, Niterói jan./fev. 2006.
- [2] Rose, G.; Tadiello F. F.; Rose Jr, D. Lesões esportivas: um estudo com atletas do basquetebol brasileiro. Efdeports. Ano 10, n.94, Buenos Aires, mar. 2006.
- [3] Pastre CM, Filho GC, Monteiro HL, Júnior JN, Padovani CR. Lesões desportivas na elite do atletismo brasileiro: estudo a partir de morbidade referida. Rev Bras Med Esporte. 2005;11:43-7.
- [4] Leite, C. B. S.; Neto, F. F. C. Incidência de lesões traumato-ortopédicas no futebol de campo feminino e sua relação com alterações posturais. Efdeports. Ano 9, n.61, Buenos Aires, jun. 2003.
- [5] Kettunen J.A., Kujala U.M., Kaprio J., Koskenvuo M., Sarna S. Lower-limb function among former elite male athletes. Am J Sports Med 2001;29:2-8.
- [6] Flegel MJ, Sport first aid. Updated ed. Champaign: Human Kinetics, 1997.
- [7] http://www.saudenarede.com.br/?p=av&id=ENTENDENDO_AS_LESOES
- [8] Weineck J. Biologia do esporte. 3a ed. São Paulo: Manole, 1991.
- [9] Hernandez Jr, BDO. Treinamento desportivo. Rio de Janeiro: Sprint, 2002
- [10] Sharkey BJ. Condicionamento físico e saúde. Porto Alegre: Artmed, 1998.
- [11] Horta L. Prevenção de lesões no desporto. Lisboa: Caminho, 1995.
- [12] Hollmann W, Hettinger T. Medicina do esporte. São Paulo: Manole, 1992.
- [13] Watson AWS. Sports injuries, incidence, causes, prevention. Rev Fisiot USP 1997;4:16-7.

- [14] Brunner e Sudarth. Tratado de Enfermagem Médico-Cirúrgica, volume 3. Rio de Janeiro: Guanabara Koogan, 2009
- [15] Mcculloch,W. S.; Pitts,W. A logical calculus of the ideas immanent in nervous activity. Bulletin os Mathematical Biophysics, p. 115 - 133, 1943.
- [16] Garson, G.D., 1991. Interpreting neural network connection weights. Artificial Intelligence Expert 6, 47- 51.
- [17] GOH, A.T.C. (1995). Back-propagation neural networks for modeling complex systems. Artif. Intell. Eng. 9, 143–151.
- [18] MASTRORILLO, S.; LEK, S.; DAUBA, F.; BELAND, A.(1997). The use of artificial neural networks to predict the presence of small-bodied fish in a river. Fresh. Biol. 38, 237–246.
- [19] M.J.S. Valença, T. B. Ludemir. Explicando a relação entre as variáveis de uma rede neural – Iluminando a “Caixa Preta”. XVII Simpósio Brasileiro de Recursos Hídricos, São Paulo, Novembro, 2007.
- [20] VALENÇA, Mêuser. Fundamentos das Redes Neurais: exemplos em Java. 2. ed. ver. ampl. Olinda: Livro Rápido, 2009. 384 p.
- [21] <http://archive.ics.uci.edu/ml/datasets/>
- [22] <http://www.cs.waikato.ac.nz/ml/weka/>
- [23] Pastre et al. Lesões desportivas na elite do atletismo brasileiro: estudo a partir de morbidade referida. Ver Brás Med Esporte 2005;11: 43-47.
- [24] <http://www.who.int/topics/obesity/es/>
- [25] Hootman JM, Macera CA, Ainsworth BE, Martin M, Addy CL, Blair SN. Association among physical activity level, cardiorespiratory fitness, and risk of musculoskeletal injury. Am J Epidemiol. 2001;154(3):251-8

- [26] Hino AAF, Reis RS, Rodriguez-Añez CR,R, Fermino RC. Prevalência de Lesões em corredores de rua e Fatores Associados. Rev Bras Med Esporte – Vol. 15, No 1 – Jan/Fev, 2009.