

Casos Clínicos: Análise de dataset e seu possível uso para geração

Aquilino Abreu da Silva
aquilinosilva@tecnico.ulisboa.pt

Instituto Superior Técnico, Lisboa, Portugal

Maio 2019

Resumo

A medicina é uma área de estudo e labor em constante desenvolvimento, onde novas técnicas e tratamentos são prosseguidos com o propósito de contribuírem para a prevenção de doenças e para a resolução de casos clínicos. Contributo seminal e relevante para a aquisição de experiência do clínico, a resolução de casos clínicos auxilia a expansão do conhecimento e o contínuo complemento da informação médica. Não obstante, a existência de elevada quantidade de informação não contribui diametralmente para a eficiência na averiguação dos diagnósticos clínicos do pessoal médico. A aprendizagem automática tem vindo a desempenhar um papel notável no campo da medicina, sobretudo no que concerne à análise e tratamento de quantidades consideráveis de informação. No presente relatório é feito um estudo de um *dataset* de casos clínicos com o propósito de se verificar a possibilidade de um modelo, treinado através de redes neuronais, realizar a previsão de casos clínicos que possam vir a ter, ou não, hipertensão como diagnóstico. É também proposto a idealização de um modelo generativo direcionado à área da medicina, dando como exemplo um modelo generativo de dígitos escritos à mão.

Palavra-chave: Aprendizagem Automática, Rede Adversarial Generativa, Diagnósticos, Treino profissional médico, Hipertensão

1. Introdução

A medicina é uma área que se encontra em contínuo desenvolvimento, sendo que, quando em sincronia com os avanços tecnológicos, resulta nos mais variados benefícios, tal como maior longevidade, melhor qualidade de vida e menor tempo de recuperação por parte do paciente [3]. A utilização de tecnologia na medicina, como é o caso dos equipamentos tecnológicos que são utilizados na deteção das possíveis causas de doença, encontra-se cada vez mais presente, tornando assim possível a resolução de um caso clínico de uma forma mais eficiente e precisa. No entanto, a resolução deste tipo de problemas não só depende dos avanços tecnológicos como também da capacidade de um clínico proceder a uma correta análise dos casos clínicos com o qual se depara e consequente diagnóstico.

Os diagnósticos clínicos são um processo de análise efetuado às diversas características e sintomas que um determinado paciente demonstra, sendo este tipo de processo realizado por especialistas clínicos. O processo de aprendizagem pelo qual os alunos de medicina passam durante o período de faculdade não permite experienciar o avultado número de casos clínicos existentes e desenvolver uma vasta rede de conhecimento organizada.

Este tipo de processo é desenvolvido ao longo dos anos, através da prática e resolução de problemas

clínicos, através da prática e resolução de problemas clínicos reais, devido ao seu fator de imprevisibilidade e grau de dificuldade. A ausência de *feedback* por parte do paciente levanta vários constrangimentos, pois impossibilita ao clínico o reconhecimento da certeza do diagnóstico e da consequente eficácia da terapêutica. Pode ainda acarretar a formulação da certeza do diagnóstico e consequentes diagnósticos incorretos, o que pode suceder, além da não resolução da condição exposta, no agravamento dos sintomas e no surgimento de novos problemas de saúde.

Pelo moroso processo de aprendizagem dos estudantes de medicina bem como por fatores subjacentes aos diagnósticos realizados por clínicos menos experientes, torna-se imprescindível o desenvolvimento de um sistema que possibilite a sua aprendizagem de uma forma mais flexível, eficiente, descartando assim a possibilidade de evolução dos riscos associados à tomada de decisões por eles tomados.

Pela indiscutível importância da medicina e da saúde pública e pela crescente utilização da aprendizagem automática em tecnologias do uso quotidiano, o presente documento focar-se-á no desenvolvimento de vários modelos de redes neuronais que sejam capazes de realizar a classificação de diagnósticos, sendo o principal ênfase nos casos de hipertensão, tendo em conta os sintomas e resulta-

dos de testes/exames que um paciente apresenta. Por fim, será proposto um modelo generativo que permita assim a Geração Procedimental de Casos Clínicos, um modelo capaz de gerar casos clínicos "falsos" tendo em conta casos clínicos reais obtidos através da Base de Dados MIMIC-III (*Medical Information Mart for Intensive Care*) [1].

2. Background e Trabalho Relacionado

Nesta Secção são introduzidos alguns conceitos sobre o Diagnóstico Clínico e consequente processo. Posto isso, é abordado a temática da Aprendizagem Automática aplicada aos Diagnósticos Clínicos, Redes Neurais Artificiais e sua estrutura.

2.1. Diagnóstico Clínico

A medicina é uma das inúmeras áreas do conhecimento científico, ou seja, uma ciência que representa todo o conhecimento adquirido através do estudo ou prática, baseado em princípios. Nesta perspectiva, a saúde humana é também objeto de estudo por parte do médico que pesquisa, faz diagnósticos, requer exames, prescreve medicamentos e realiza cirurgias a fim de tratar disfunções e enfermidades. Para que a sua intervenção seja bem-sucedida, deverá estar bem informado e atualizado sobre os novos fármacos, métodos e técnicas capazes de desenvolver grandes avanços na resolução de problemas.

No entanto, nem tudo depende dos avanços, mas sim da capacidade de um médico perceber quais os sintomas relevantes no caso clínico, avaliar as possíveis hipóteses, verificar os sinais e sintomas relevantes no caso clínico para que assim possam apresentar um tratamento adequado quando confrontados com um paciente.

O raciocínio clínico é um processo do qual requer bastante conhecimento e experiência obtida de vários anos de estudo e da resolução de casos clínicos, resultando assim na construção de redes organizadas de conhecimento conhecidos por *scripts* de doença. As *scripts* de doenças são compostas pelas características que se apresentaram com o surgir da doença, descrição do que se encontra a falhar, sinais e sintomas que originaram devido à falha, permitindo assim o desenvolvimento do raciocínio e de reconhecimento de padrões [8].

Um dos passos mais importantes no raciocínio clínico encontra-se na descrição dos sintomas que o paciente faz durante a entrevista médica. Embora os sintomas descritos estejam sujeitos à interpretação do próprio paciente, compete ao profissional recolher informações necessárias ao pleno conhecimento das características dos sintomas relatados, bem como do historial clínico, incluindo exames laboratoriais [2]. A razão pela qual é importante recolher todas as características possíveis durante a observação deve-se ao facto de que caso as mesmas não sejam obtidas na sua totalidade, pode levar a uma

completa mudança do diagnóstico. Por esta mesma razão, torna-se imprescindível uma completa "dissecação" ao caso clínico e seleção dos dados relevantes para a resolução do mesmo.

No entanto, durante o processo de diagnóstico clínico há sempre a possibilidade de o clínico falhar na identificação de informação crucial, levando assim a uma incorreta interpretação e consequente formulação de hipóteses e diagnósticos incorretos. Fatores como a experiência que o clínico possui, ou até mesmo a complexidade do caso clínico, podem contribuir para esta situação.

A fim de contribuir para o ganho de experiência e desenvolvimento das capacidades de diagnóstico por parte do clínico, sem que o mesmo envolva colocar em risco a vida dos pacientes, torna-se imprescindível a aplicação da Aprendizagem Automática nos Diagnósticos Clínicos.

2.2. Aprendizagem Automática aplicada aos Diagnósticos Clínicos

Nas últimas décadas, a aprendizagem computacional tornou-se um dos pilares da tecnologia da informação, gerando um contributo muito importante na resolução de problemas e tomada de decisões por parte do ser humano, devido ao crescente volume de dados adquiridos numa determinada área que necessitam de estudo e aplicação de técnicas para análise de dados. Nesta perspectiva, a ligação entre diagnóstico clínico e aprendizagem computacional proporciona melhor exatidão e maior eficiência nos resultados obtidos.

A Aprendizagem Automática encontra-se em duas grandes áreas, nomeadamente a Inteligência Artificial e a Ciência de Dados. A Aprendizagem Automática aplicada à Inteligência Artificial consiste na produção de um sistema com a finalidade de executar ações que normalmente envolvem a tomada de decisões, resolução de problemas, utilização de agentes inteligentes, como também decisões normalmente tomadas por seres humanos racionais que possuem inteligência [12]. No entanto, a resolução do problema referido no presente documento implica a análise de uma grande quantidade de dados para auxiliar na classificação de diagnósticos dos casos clínicos. Como tal, será discutido sobre a Aprendizagem Automática aplicada à Ciência de Dados.

Arthur Samuel, um pioneiro americano no campo dos jogos de computador e Inteligência Artificial, como: "*the field of study that gives computers the ability to learn without being explicitly programmed*" [11]. Uma definição moderna, mais recente, expressa por Tom Mitchell, um cientista de computação americana, define a Aprendizagem Automática desta forma [9]:

"*A computer program is said to learn from expe-*

rience E with respect to some class of tasks T and performance measure P , if its performance at tasks in T , as measured by P , improves with experience E .”

Dadas estas definições sobre a Aprendizagem Automática, depreende-se que uma máquina não necessita ser explicitamente programada para cada tarefa. Porém, o conhecimento pode ser ensinado ou transmitido para que o sistema possa aprender a resolver problemas, dada nova informação, ou na tomada de decisões.

De acordo com [5], os algoritmos de Aprendizagem Automática foram desenvolvidos, inicialmente, para serem utilizados na análise de dados clínicos a fim de alcançar um potencial diagnóstico, tendo como base a análise de um avultado número de casos clínicos previamente classificados. Contudo, a análise deste volume de informação implica a existência de organização e facilidade de acesso aos casos clínicos de forma a poderem ser introduzidos no sistema que correrá os algoritmos de aprendizagem automática, de forma eficiente.

Com os avanços tecnológicos, modernização dos equipamentos utilizados em hospitais e clínicas, torna-se possível a recolha, armazenamento e partilha de registos médicos eletrónicos com outras entidades e/ou instituições, caso exista a devida autorização para o efeito. Tendo acesso a este tipo de informação, torna-se possível a utilização de programas de computador para execução de algoritmos de Aprendizagem Automática em casos clínicos que tenham sido resolvidos no passado, resultando assim na sua aprendizagem e consequente previsão de resultados de casos clínicos nunca antes resolvidos [4].

2.3. Aprendizagem Automática

Existem vários tipos de Aprendizagem, nomeadamente:

- **Aprendizagem Semi-Supervisionada:** dado um conjunto de dados de treino que possui dados identificados e não-identificados, uma hipótese inicial é formulada tendo em conta os dados identificados, que depois é aperfeiçoada com a utilização de dados não-identificados através de uma determinada estratégia de identificação automática, como por exemplo o algoritmo *Expectation-Maximization* (EM) [6];
- **Aprendizagem Não-Supervisionada:** permite estruturar dados que não se encontram categorizados [10];
- **Aprendizagem Supervisionada:** partindo do suposto que existe uma relação entre os dados de entrada e os dados de saída, dado um conjunto de dados, os algoritmos de aprendizagem supervisionada são capazes de prever o

seu correto output. A aprendizagem supervisionada encontra-se organizada em dois tipos de problemas [10]:

- **Regressão:** dado variáveis de entrada e conhecidas as variáveis de saída, é criado uma função contínua que permite prever o *output*;
 - **Classificação:** dado variáveis de entrada, a previsão das variáveis de saída é representada por 0 ou 1.
- **Aprendizagem por Reforço:** tomada de decisões por parte de um agente de reforço que visa em completar uma determinada tarefa de melhor forma possível através da experiência que adquire ao longo do mesmo;

Para este tipo de problema que se pretende resolver será utilizada a Aprendizagem Supervisionada, onde um modelo será treinado tendo como base um conjunto de casos clínicos resolvidos.

2.3.1 Representação do Modelo

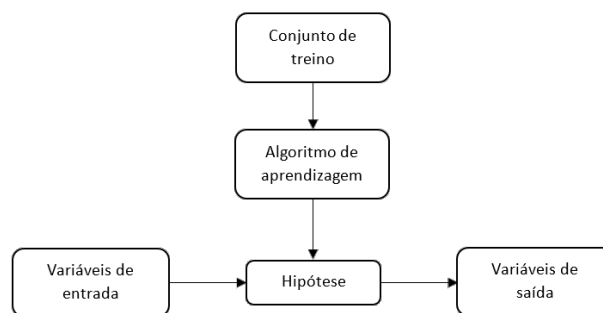


Figura 1: Formulação de hipóteses

A Figura 1 apresenta um esquema de como é representado a formulação de hipóteses. Um conjunto de dados de treino é dado por dois tipos de variáveis, X e Y , no qual X representa diferentes atributos, ou características, e Y denota o *output* originado dado X . Pode acontecer que no conjunto de treinos seja utilizado um elevado número de atributos relativamente ao número de eventos (pacientes) e resulte na formulação de uma hipótese que se ajuste aos dados de treino, mas falhar na generalização de novos exemplos. O mesmo pode acontecer caso seja utilizado um número baixo de atributos. Neste caso pode levar a duas situações: *underfitting* e *overfitting*.

Quando lidamos com quaisquer tipos de algoritmos de aprendizagem, e tendo em conta que os mesmos se adequam aos problemas aos quais são expostos, os mesmos são capazes de funcionar para a maior parte dos casos. No entanto, existe uma

possibilidade de estes mesmos algoritmos ficarem sujeitos a problemas de *performance*, causados pelo *overfitting*, ou *underfitting*, que se encontra relacionado com a capacidade com a qual a função criada possui para generalizar relativamente a novos dados.

Um modelo "*underfit*" é um modelo que não é adequado para fazer a previsão de dados nunca vistos, pois a sua *performance* é muito baixa no conjunto de treino.

Um modelo "*overfit*" é um modelo que tem em conta todos os detalhes e ruído dos dados de treino, levando assim à consequente inabilidade de generalizar para novos dados, ou seja, o modelo considera os ruídos, ou até mesmo pequenas flutuações aleatórias, como sendo algo importante para a aprendizagem. Este tipo de problema é bastante comum em modelos não-lineares, como é o caso das redes neuronais, daí serem utilizadas diversas técnicas e diferentes valores para os mais variados parâmetros que estes modelos possuem a fim de limitar a quantidade de detalhe que o modelo aprende.

Um modelo é considerado como sendo ideal quando encontra o equilíbrio entre o *underfitting* e o *overfitting*. De modo a verificar se o modelo se encontra neste estado, uma das várias opções que temos a nosso dispor para o verificar é o desenho de gráficos que mostre a *performance* do modelo no conjunto de dados de treino e no conjunto de dados de teste.

2.3.2 Validação e Seleção do modelo

Geralmente, um conjunto de dados é dividido em três tipos: treino, validação e teste. O conjunto de dados de treino é utilizado para treinar o classificador, o de validação é utilizado para fazer validações iniciais durante a aprendizagem do modelo, e os de teste tratam-se de um conjunto de dados nunca vistos pelo modelo. A fim de avaliar quão errado um modelo se encontra, são utilizadas as Funções de Custo.

2.3.3 Função de Custo

Na Aprendizagem Automática, as funções de custo têm como principal objetivo ajudar o modelo em questão a corrigir o seu funcionamento de modo a minimizar os erros. São, por sua vez, utilizados métodos de otimização, tais como o *Gradient Descent*, *Stochastic Gradient Descent*, *Adam (Adaptive Moment Estimation)*, que permitem assim minimizar a função de custo.

2.4. Redes Neuronais Artificiais

Existem diversos tipos de Algoritmos de Aprendizagem Automática, tais como *k-vizinhos mais*

próximos, Classificadores *Bayesianos Naïve e Semi-Naïve*, Redes Neuronais Artificiais com *Backpropagation*, Redes Neuronais Convolucionais, Recorrentes, entre outras. No entanto será dado maior ênfase às Redes Neuronais com *Backpropagation* pois trata-se da abordagem utilizada no desenvolvimento deste projeto.

As redes neuronais são compostas por *nodes* ou unidades interligados, no qual a ligação entre duas unidades, i e j , permite a propagação da ativação a_i . Cada ligação possui um determinado peso $w_{i,j}$ associado, a fim de ser definido a força e o sinal dessa mesma ligação. Cada unidade possui o chamado "*dummy input*", $a_0 = 1$, com o respetivo peso $w_{i,j}$ associado. Primeiramente, é calculada a soma ponderada para cada unidade j , que se encontra representada na equação 1. Após o cálculo da soma ponderada da unidade j , é aplicada a função de ativação, g , que se encontra representada na equação 2, a fim de definir o valor de saída [12].

$$input_j = \sum_{i=0}^n w_{i,j} a_i \quad (1)$$

$$activation_j = g(input_j) = g\left(\sum_{i=0}^n w_{i,j} a_i\right) \quad (2)$$

Existem vários tipos de funções de ativação, tais como a Função *Sigmoid*, *Tanh (Hyperbolic Tangent)*, *ReLu*, entre outros. A escolha da Função de Ativação depende, maior parte das vezes, do tipo de problema para o qual se treina um rede e da *performance* que cada uma dessas funções é capaz de entregar, sendo necessário assim ter o conhecimento de como cada uma das funções de ativação funcionam antes de serem utilizadas [13].

3. Implementação

De modo a proceder à resolução do problema exposto na proposta do projeto, foi utilizada uma Base de Dados com imensos dados e informações pertinentes a casos clínicos relativos a pacientes que se encontraram internados nos cuidados intensivos do *Beth Israel Deaconess Medical Center, Boston, Massachusetts*.

A base de dados em questão, conhecida como MIMIC-III (*Medical Information Mart for Intensive Care*) [1] é nada mais nada menos que um conjunto de várias tabelas, ligadas entre si e com algumas dependências, com registos médicos eletrónicos (*EMR - Electronic Medical Record*).

A fim de desenvolver um mecanismo de Aprendizagem Automática que aprenda quais os sintomas, ações e tratamentos que devem ser ensinados, partindo da informação de casos clínicos reais, é necessário perceber como se encontra organizada a base de dados, que tipos de informação esta possui,

como é que as várias tabelas que a compõem se encontram interligadas, e quais os dados que deverão ser tomados em conta de forma a que os vários objetivos delineados sejam alcançados.

Como ponto de partida foi realizada uma pesquisa à base de dados com o propósito de obter os 10 diagnósticos mais comuns, e o respetivo número de pacientes, sendo que a escolha terminou por ser o diagnóstico de Hipertensão.

Tendo em conta os diferentes fatores responsáveis pela Hipertensão Arterial, foram gerados vários *datasets* (conjuntos de dados), sendo que no final foi utilizado o conjunto de dados que possuía as características mais relevantes encontradas nos casos de Hipertensão.

3.1. Construção do *Dataset*

Para a construção do *dataset* a ser utilizado para treinar a rede neuronal, foram obtidos os dados, resultados de exames, análises clínicas, medições do batimento cardíaco, pressão arterial sistólica e diastólica dos pacientes que apresentavam o sintoma de hipertensão, e dos pacientes que não apresentavam o sintoma de hipertensão. Após obtidos os dados dos pacientes que seguiam os critérios previamente mencionados, os mesmos compuseram uma tabela, onde cada linha dessa tabela correspondia a um paciente, e cada coluna correspondia a uma determinada *feature* (i.e. batimento cardíaco, peso, altura, etc). A fim de serem distinguidos quais os pacientes que tinham hipertensão e os que não tinham, foi criada uma nova coluna (posicionada no extremo direito da tabela) identificada como "Class". Para os pacientes com hipertensão eram atribuídos o valor 1 na coluna "Class", e o valor 0 para os pacientes sem hipertensão.

A construção de um modelo preditivo de sucesso depende bastante dos dados selecionados e da forma como os mesmos são utilizados. No entanto, o *dataset* em questão não se encontra preparado para ser introduzido num modelo, pois existem vários fatores que deverão ser tomados em conta antes de ser iniciado o treino e consequente teste do modelo.

3.2. Preparação do *Dataset*

A fim de possibilitar a utilização do *dataset* por parte das Redes Neurais, foi necessário lidar com os valores em falta, converter características categóricas para uma representação numérica e dimensionar as características (normalização).

3.3. Implementação do Modelo de Redes Neurais

Como ponto de partida foi utilizada, para ambos os *datasets*, a abordagem descrita em [15], onde o autor *Huang* apresentou que o número suficiente de neurónios para a primeira e segunda camada intermédia, na pior das hipóteses, são calculadas

através das Equações 3 e 4 respetivamente.

$$\sqrt{(m+2)N} + 2\sqrt{N/(m+2)} \quad (3)$$

$$m\sqrt{N/(m+2)} \quad (4)$$

onde

$m \rightarrow$ número de características;

$N \rightarrow$ número de exemplos;

Como tal foram criadas várias versões do *dataset* escolhido, no qual cada uma das versões continha um número diferente de exemplos (casos clínicos) e de características (*features*), como pode ser observado na Tabela 1. Foram criados vários modelos para essas mesmas versões do *dataset*, utilizando somente duas camadas intermédias, o número de neurónios para cada uma das camadas intermédias como descrito pelas equações 3 e 4 e com os hiperparâmetros *default* do Scikit-Learn [14], uma biblioteca *open-source* de aprendizagem automática utilizada na linguagem de programação *Python*.

| | Versões do dataset | | | | | |
|---------------------|--------------------|----------|----------|----------|----------|----------|
| | 1 | 2 | 3 | 4 | 5 | 6 |
| Nº Caraterísticas | 64 | 70 | 86 | 91 | 91 | 92 |
| Nº Exemplos | 20452 | 12521 | 20452 | 2658 | 7192 | 12521 |
| Camadas Intermédias | (412,82) | (323,64) | (412,82) | (148,29) | (244,48) | (323,64) |

Tabela 1: Informação relativa ao *dataset* com eventos gráficos

Posto isto, foi escolhida a versão do *dataset* que melhores resultados apresentava, com o principal objetivo de melhorar a *performance* do modelo através da alteração do número de neurónios na primeira e segunda camada intermédia, utilização de diferentes métodos que visam à redução do *overfitting*, funções de ativação, métodos de otimização e hiperparâmetros. Na Secção 4 são apresentados os resultados obtidos.

4. Resultados

Na Tabela 2 temos os valores de precisão de treino e teste para as várias versões do *dataset* com eventos gráficos, onde foi seguida a mesma abordagem utilizada para com os *datasets* com eventos de laboratório.

Alguns destes hiperparâmetros *default* são:

- **Função de ativação:** *ReLU*;
- **Alpha:** 0.0001;
- **Método de otimização:** *Adam*;
- **Taxa de aprendizagem inicial:** 0.001;
- **Nº máximo iterações:** 200;

| chartevents dataset | Precisão de treino | Precisão de teste |
|---------------------|--------------------|-------------------|
| 1 | 74% | 60% |
| 2 | 77% | 60% |
| 3 | 80% | 64% |
| 4 | 89% | 64% |
| 5 | 80% | 58% |
| 6 | 77% | 60% |

Tabela 2: Precisão de treino e teste dos *datasets* com eventos gráficos

- **Tolerância:** 0.0001;

Tendo em conta os resultados obtidos, foi escolhida a versão 4 devido aos valores obtidos na precisão durante o treino e teste, e por ter uma média ponderada da precisão e da capacidade para recordar (*f1-score*) superior à das outras versões, como é possível observar na Figura 2.

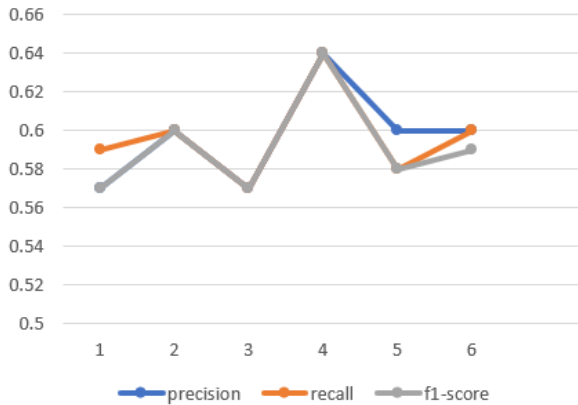
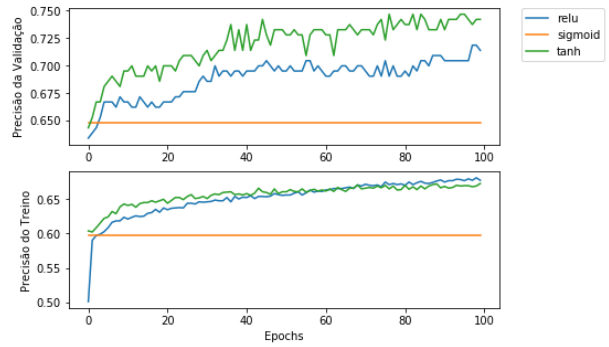
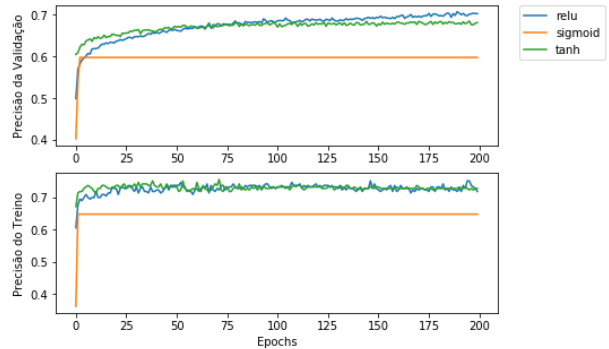


Figura 2: Métricas (*precision*, *recall* e *f1-score*)

Os gráficos representados na Figura 3 ilustram a progressão de vários modelos, com diferentes funções de ativação nas camadas intermédias (*ReLU*, *Sigmoid* e *Hyperbolic Tangent*) e diferentes métodos de otimização (*Adam* e *Stochastic Gradient Descent*). Como é possível observar na Figura 3, são obtidos diferentes resultados para cada uma das diferentes funções de ativação e métodos de otimização. Enquanto que quando utilizado o método de otimização *SGD* é alcançado um nível de *performance* à volta dos 70%, tanto no treino como na validação, a função de ativação *Sigmoid* não passa dos 60%. Relativamente ao modelo no qual é utilizado o método de otimização *Adam*, observa-se uma *performance* superior durante a validação, comparativamente com a do treino. Novamente, a função de ativação *Sigmoid* mantém-se constante nos 65%.



(a) Optimizado com Adam



(b) Optimizado com SGD

Figura 3: Medição da precisão utilizando diferentes funções de ativação, e diferentes métodos de otimização

Após observados os resultados obtidos dos gráficos da Figura 3, obtidos após terem percorrido um valor máximo de 200 *epochs* (1 *epoch* corresponde ao *feedforward* e *backward* de um conjunto de dados, na sua totalidade, de somente uma vez), foi decidido utilizar o método de otimização *Adam* e as funções de ativação *ReLU* e *Tanh*, devido à contínua melhoria observada durante a validação dos modelos. Uma outra razão pelo facto de a função de ativação *ReLU* ter sido escolhida é por ser uma das funções mais utilizadas em redes neuronais artificiais, e o *Tanh* devido aos resultados que apresentou e por este ser uma versão melhorada da função *Sigmoid*.

4.1. Aplicação de soluções para evitar *Overfitting*

O *dataset* com eventos gráficos (*chartevents*) foi construído tendo em conta as características que mais se relacionam com a hipertensão. Como tal o *dataset* inicial continha essas possíveis características e casos clínicos, até ao momento do tratamento dos dados, o que originou às várias versões do *dataset*, onde o número de características e de casos clínicos foi variando. Uma das formas de evitar o *overfitting* é o treino dos modelos com mais dados, no entanto o mesmo não foi possível pois após

uma análise aos resultados dos vários modelos foi escolhido aquele com melhor *performance*.

Ao longo do processo de evitar o *overfitting*, foram utilizados vários métodos:

- Redução do número de neurónios na primeira e segunda camada intermédia, conjuntamente com a alteração das funções de ativação;
- Alteração do valor de α ;
- Alteração da taxa inicial de aprendizagem - controla o tamanho dos 'passos' durante a atualização dos pesos;
- K-Fold Cross-Validation, com $k=5$;
- *Principal Component Analysis* (PCA) - técnica utilizada na redução de características através da extração das mesmas;

Apesar da aplicação dos diversos métodos aqui descritos, não foi possível reduzir o *overfitting* nos modelos gerados. A variação da precisão durante os treinos e testes eram insignificantes, mesmo com a variância do número de neurónios em ambas as camadas intermédias, utilização das funções de ativação *ReLU* e *Tanh* e o método de otimização *Adam*.

Num dos casos foi utilizado o PCA conjuntamente com o *K-Fold Cross-Validation*, onde $K = 5$, em que apesar de a precisão durante o treino ter alcançado os 100%, a precisão durante o teste manteve-se inalterável.

Foram também criados modelos em que foi utilizado o PCA e modificados os hiperparâmetros da taxa inicial da aprendizagem e do valor de α . Notou-se que apesar de os valores de precisão não variarem significativamente com o aumento do valor de α , o valor da precisão variava, consideravelmente, quando o valor da taxa de aprendizagem inicial diminuía.

Foram também criados modelos, que além de ter sido utilizado o PCA foram alterados os hiperparâmetros da taxa inicial da aprendizagem e de α . À medida que o valor de α aumentava, os valores de precisão, tanto de treino como de teste, não variaram muito. No entanto, quando o valor da taxa de aprendizagem inicial era diminuído, os valores de precisão variavam consideravelmente durante o treino.

4.2. Proposta para um modelo generativo

Este projeto tem como finalidade o estudo e análise de um *dataset* de casos clínicos e consequente proposta de um modelo capaz de gerar casos clínicos tendo como base casos clínicos reais. Para tal, é necessária a criação de dois tipos de redes, generativas e discriminativas, conhecidas como Redes Adversariais Generativas (*Generative Adversarial Networks* - GAN).

Os algoritmos discriminativos são algoritmos que realizam a previsão de uma categoria ao ser facultado um conjunto de características, isto é, modelos que aprendem a limitar as diferentes categorias, enquanto que os algoritmos generativos tentam fazer a previsão de características para uma determinada categoria.

As Redes Adversariais Generativas, como o nome dá a entender, consistem em duas redes neuronais: uma chamada "gerador", que gera novos exemplos, e o "discriminador", que avaliará se um determinado exemplo foi gerado ou se pertence a um conjunto de dados real.

Dentro do contexto dos casos clínicos, o papel do gerador é criar novos casos clínicos tendo em conta casos clínicos reais, enquanto que o papel do discriminador é ser capaz de distinguir os casos clínicos reais daqueles que foram gerados. Na Figura 4 encontra-se representado os passos que são tomados pelas GAN.

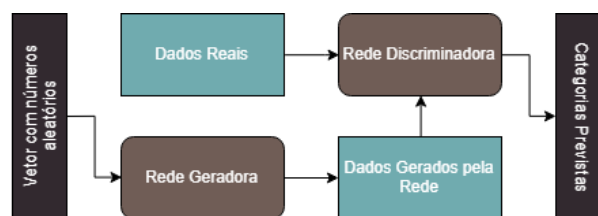


Figura 4: Passos tomados nas Redes Adversariais Generativas

No *GitHub* encontrava-se disponibilizado um repositório que providenciava um algoritmo que permitia a geração de dígitos escritos à mão [7], através do modelo gerador, e a discriminação dos mesmos através do modelo discriminador. Na Figura 5 temos exemplos de alguns dígitos gerados pelas GAN.

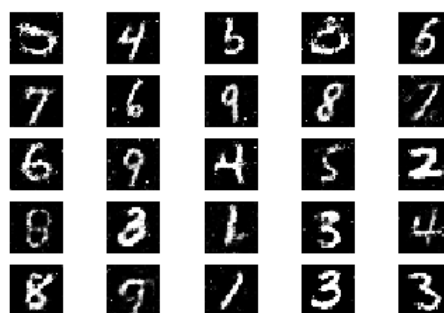


Figura 5: Dígitos gerado pelas GAN

5. Conclusões

O presente projeto tinha como principal objetivo a Geração Procedimental de Casos Clínicos, que

visava a contribuir para o desenvolvimento do raciocínio clínico de estudantes de medicina e clínicos. No entanto, ao longo do projeto foram encontradas dificuldades durante a criação de um modelo *Gold Standard*, um modelo que fosse capaz de classificar casos clínicos cujos diagnósticos fossem de hipertensão ou não. A performance do modelo em questão não alcançou um nível de *performance* aceitável devido à sua incapacidade de generalizar para casos clínicos nunca antes vistos. Apesar das várias técnicas aplicadas, alteração dos hiperparâmetros, do número de camadas intermédias e respetivo número de neurónios das redes neuronais, utilização de várias versões de *datasets* tendo em consideração os aspetos mais tomados em conta durante o diagnóstico de hipertensão, a *performance* nunca atingiu valores satisfatórios e nunca se percebeu uma evolução tendo em conta os hiperparâmetros.

As Redes Adversariais Generativas são o futuro da Aprendizagem Automática pois tratam-se de dois tipos de redes com funções completamente distintas: os modelos generativos, que são responsáveis por "aprender" através de *datasets* reais e gerar dados falsos, e os modelos discriminativos, que são responsáveis por distinguir os dados reais daqueles que foram gerados. Casos como a criação de arte e geração de ambientes paisagísticos através de simples "traços" são algumas das capacidades e potencial que as Redes Adversariais Generativas têm vindo a mostrar. Além da recorrente investigação e desenvolvimento de ferramentas utilizando estas redes, as mesmas demonstram um tremendo potencial na área da medicina, através da geração de casos clínicos e respetivos sintomas que podem vir a ser utilizados por estudantes de medicina e clínicos com o principal objetivos de ajudar na preparação da resolução de casos clínicos e desenvolvimento do próprio raciocínio clínico, evitando assim colocar a vida de um paciente em risco.

Agradecimentos

Primeiramente, quero agradecer à minha família por todo o apoio que me prestaram, aos meus orientadores Pedro Santos e Luís L. Ribeiro por toda a orientação prestada desde o início até ao fim deste projeto e aos meus amigos que fizeram parte deste meu percurso académico.

Referências

- [1] J. AEW, P. TJ, S. L, L. L, F. M, G. M, M. B, S. P, C. LA, and M. RG. Mimic-iii, a freely accessible critical care database. *Scientific Data* (2016). DOI: 10.1038/sdata.2016.35.
- [2] J. L. Bowen. Educational strategies to promote clinical diagnostic reasoning. *New En-*

gland Journal of Medicine, 355(21):2217–2225, 2006.

- [3] D. M. Cutler and M. McClellan. Is technological change in medicine worth it? *Health affairs*, 20(5):11–29, 2001.
- [4] K. R. Foster, R. Koprowski, and J. D. Skufca. Machine learning, medical diagnosis, and biomedical engineering research-commentary. *Biomedical engineering online*, 13(1):94, 2014.
- [5] I. Kononenko. Machine learning for medical diagnosis: history, state of the art and perspective. *Artificial Intelligence in medicine*, 23(1):89–109, 2001.
- [6] M. Li and Z.-H. Zhou. Improve computer-aided diagnosis with machine learning techniques using undiagnosed samples. *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans*, 37(6):1088–1098, 2007.
- [7] E. Linder-Norén. Keras implementations of generative adversarial networks., 2018.
- [8] W. J. McMillan. Teaching for clinical reasoning—helping students make the conceptual links. *Medical Teacher*, 32(10):e436–e442, 2010.
- [9] T. M. Mitchell. *Machine Learning*. McGraw-Hill, Inc., New York, NY, USA, 1 edition, 1997.
- [10] M. Mohri, A. Rostamizadeh, and A. Talwalkar. *Foundations of machine learning*. MIT press, 2012.
- [11] A. Munoz. Machine learning and optimization. URL: https://www.cims.nyu.edu/~munoz/files/ml_optimization.pdf [accessed 2016-03-02][WebCite Cache ID 6fiLfZvnG], 2014.
- [12] S. Russell and P. Norvig. *Artificial Intelligence: A Modern Approach*. Prentice Hall, 3rd edition, 2010., 1994.
- [13] A. Sharma. Understanding activation functions in neural networks. Acedido a: 2018-10-26.
- [14] O. Source. Scikit-learn.
- [15] D. Stathakis. How many hidden layers and nodes? *International Journal of Remote Sensing*, 30(8):2133–2147, 2009.