

Célia Natália Lemos Figueiredo

**Identificação de pacientes com *delirium* em contexto hospitalar através de algoritmos de *machine* *learning***

Dissertação de Mestrado

Mestrado em Engenharia de Sistemas

Trabalho realizado sob a orientação de

Professora Doutora Ana Cristina Silva Braga

DoutorJosé António Briote Mariz

Outubro de 2021

**DIREITOS DE AUTOR E CONDIÇÕES DE UTILIZAÇÃO DO TRABALHO POR TERCEIROS**

Este é um trabalho académico que pode ser utilizado por terceiros desde que respeitadas as regras e boas práticas internacionalmente aceites, no que concerne aos direitos de autor e direitos conexos.

Assim, o presente trabalho pode ser utilizado nos termos previstos na licença abaixo indicada.

Caso o utilizador necessite de permissão para poder fazer um uso do trabalho em condições não previstas no licenciamento indicado, deverá contactar o autor, através do RepositóriUM da Universidade do Minho.

**Licença concedida aos utilizadores deste trabalho**

https://licensebuttons.net/l/by-nc-nd/3.0/88x31.png

**Atribuição-NãoComercial-SemDerivações   
CC BY-NC-ND**

https://creativecommons.org/licenses/by-nc-nd/4.0/

# Agradecimentos

Esta página é opcional para agradecimentos do autor podendo fazer referência a apoio financeiro, se aplicável.

Agradeço profundamente aos meus pais todo o apoio, confiança e pela oportunidade que me proporcionaram de prosseguir os estudos no ensino superior.

**DECLARAÇÃO DE INTEGRIDADE**

Declaro ter atuado com integridade na elaboração do presente trabalho académico e confirmo que não recorri à prática de plágio nem a qualquer forma de utilização indevida ou falsificação de informações ou resultados em nenhuma das etapas conducente à sua elaboração.

Mais declaro que conheço e que respeitei o Código de Conduta Ética da Universidade do Minho.

**Identificação de pacientes com *delirium* em contexto hospitalar através de algoritmos de *machine learning***

# Resumo

Na extensão máxima de uma página.

Palavras-Chave: 3 a 5 palavras, escritas por ordem alfabética

**Identificação de pacientes com delirium em contexto hospitalar através de algoritmos de *machine learning***

# Abstract

Na extensão máxima de uma página.

Keywords: 3 a 5 palavras, escritas por ordem alfabética

Índice

[Agradecimentos iii](#_Toc89964677)

[Resumo v](#_Toc89964678)

[Abstract vi](#_Toc89964679)

[Lista de Abreviaturas e Siglas x](#_Toc89964680)

[Lista de Figuras xii](#_Toc89964681)

[Lista de Tabelas xiii](#_Toc89964682)

[1. Introdução 1](#_Toc89964683)

[1.1 Enquadramento 1](#_Toc89964684)

[1.2 Objetivos 3](#_Toc89964685)

[1.3 Metodologia 4](#_Toc89964686)

[1.4 Estrutura da dissertação 5](#_Toc89964687)

[2. Revisão da literatura 7](#_Toc89964688)

[2.1 *Machine Learning* 7](#_Toc89964689)

[2.1.1 Breve História e Evolução 8](#_Toc89964690)

[2.1.2 Uso de *Machine Learning* na medicina 9](#_Toc89964691)

[2.2 *Delirium* 11](#_Toc89964692)

[2.2.1 Breve introdução histórica 11](#_Toc89964693)

[2.2.2 Definição e características gerais 13](#_Toc89964694)

[3. Machine Learning 16](#_Toc89964695)

[3.1 Aprendizagem supervisionada 16](#_Toc89964696)

[3.1.1 Regressão Logística 17](#_Toc89964697)

[3.1.2 Árvores de decisão 20](#_Toc89964698)

[*3.1.3* *Random Forest* 22](#_Toc89964699)

[3.1.4 Máquina Vetorial de Apoio 22](#_Toc89964700)

[3.1.5 K vizinhos mais próximos 23](#_Toc89964701)

[3.1.6 *Artificial Neural Network* (ANN) 23](#_Toc89964702)

[3.2 Aprendizagem Não-Supervisionada 25](#_Toc89964703)

[*3.2.1* *Clustering* 25](#_Toc89964704)

[*3.2.2* *K-Means* 26](#_Toc89964705)

[3.3 Aprendizagem por Reforço 26](#_Toc89964706)

[3.4 Ferramentas disponíveis em Python para ML 27](#_Toc89964707)

[4. Delirium 28](#_Toc89964708)

[4.1 Prevalência 29](#_Toc89964709)

[4.2 Fatores de risco 30](#_Toc89964710)

[4.2.1 Fatores predisponentes 31](#_Toc89964711)

[4.2.2 Fatores precipitantes 32](#_Toc89964712)

[4.3 Fisiopatologia 34](#_Toc89964713)

[4.4 Instrumentos de rastreio e diagnóstico de *delirium* 38](#_Toc89964714)

[4.4.1 *Richmond Agitation Sedation Scale* (RASS) 40](#_Toc89964715)

[4.4.2 *Confusion Acessment Method* (CAM) 41](#_Toc89964716)

[4.4.3 *Confusion Assessment Method for the Intensive Care Unit* (CAM-ICU) 41](#_Toc89964717)

[5. Preparação e análise exploratória dos dados 42](#_Toc89964718)

[5.1 Preparação e recolha de dados 43](#_Toc89964719)

[5.2 Limpeza dos dados 44](#_Toc89964720)

[5.2.1 Identificação de colunas com dados redundantes 44](#_Toc89964721)

[5.2.2 Identificação de colunas com valores únicos 45](#_Toc89964722)

[5.2.3 Identificação de colunas inócuas 45](#_Toc89964723)

[5.2.4 Exclusão da variável ‘Obito’ 46](#_Toc89964724)

[5.2.5 Identificação de linhas com dados duplicados e omissos 46](#_Toc89964725)

[5.3 Feature Selection 47](#_Toc89964726)

[5.3.1 Identificação de colunas com pouca variância 47](#_Toc89964727)

[5.4 Transformação dos dados 48](#_Toc89964728)

[5.4.1 Agregação dos medicamentos por grupo farmacológico 48](#_Toc89964729)

[5.4.2 Codificação dos dados categóricos 75](#_Toc89964730)

[5.4.3 Normalização dos dados 78](#_Toc89964731)

[6. Modelação, Apresentação e discussão dos resultados 80](#_Toc89964732)

[6.1 Desbalanceamento dos dados 80](#_Toc89964733)

[6.2 Divisão dos dados 82](#_Toc89964734)

[6.3 Treino do modelo Random Forest 83](#_Toc89964735)

[6.4 Previsões 85](#_Toc89964736)

[6.5 Avaliação da performance do modelo 85](#_Toc89964737)

[6.6 Feature importance 87](#_Toc89964738)

[6.7 Tuning dos parâmetros 88](#_Toc89964739)

[6.8 Avaliação da performance do modelo 88](#_Toc89964740)

[6.9 Regressão Logística 89](#_Toc89964741)

[6.9.1 Seleção do melhor modelo 89](#_Toc89964742)

[6.10 Random Forest 90](#_Toc89964743)

[6.11 Seleção do melhor modelo 90](#_Toc89964744)

[6.12 Aplicação *web* 90](#_Toc89964745)

[6.12.1 Página Inicial 90](#_Toc89964746)

[6.12.2 Previsão 90](#_Toc89964747)

[6.12.3 Sobre 91](#_Toc89964748)

[7. Conclusões e Trabalho futuro 92](#_Toc89964749)

[Bibliografia 93](#_Toc89964750)

[Apêndice I – Descrição da Base de dados 111](#_Toc89964751)

[Anexo I – The confusion assessment method instrument 116](#_Toc89964752)

[Anexo II – The Confusion Assessment Method (CAM) Diagnostic Algorithm\* 118](#_Toc89964753)

# Lista de Abreviaturas e Siglas

AD – Árvores de Decisão

ADASYN - *Adaptive Synthetic*

ADN - Ácido Desoxirribonucleico

ADT - Antidepressivos Tricíclicos

IC - *Akaike information criterion*

APACHE--II *– Acute Physiology and Chronic Health Evaluation-II*

AUC *- Area Under the ROC Curve*

BIC -- *Bayesian Information Criterion*

CAM – *Confusion Assessment Method*

CAM-ICU – *Confusion Assessment Method for the Intensive Care Unit*

CART *-- Classification and Regression Tree*

CRISP-DM -- *Cross Industry Standard Process for Data Mining*

DARPA -- *Defense Advanced Research Projects Agency*

DOSS -- *Delirium Observation Screening Scale*

DT – *Decision Tree*

DSM-5 – *Diagnostic and Statistical Manual of Mental Disorders*

GABA -- Ácido Gama-AminoButírico

GBM - *Gradient Boosting Machine*

GNB -- *Gaussian Naïve Bayes*

ICD -- *International Classification of Diseases*

ISRS -- Inibidores seletivos de recaptação da serotonina

ISRSN -- Inibidores seletivos da recaptação da serotonina e da noradrenalina

KNN – *K-Nearest Neighbour*

ML – *Machine Learning*

PRE-DELIRIC – *PREdiction of DELIRium for Intensive Care patients*

POC -- Perturbação Obsessiva-Compulsiva

RF – *Random Forest*

RL – Regressão Logística

ROC -- *Receiver Operating Characteristic*

SIRS – *Systemic Inflammatory Response Syndrome*

SMOTE – *Synthetic Minority Over-sampling Technique*

SNC – Sistema Nervoso Central

SU – Serviço de urgência

SVM – *Support Vector Machine*

TG – Triglicerídeos

UCI – Unidade de Cuidados Intensivos

# Lista de Figuras

[Figura 1 – Diferentes fases implicadas na Metodologia CRISP-DM 4](#_Toc89964984)

[Figura 2 - Comparação gráfica entre o modelo de RL e regressão linear (James, Witten, Hastie, & Tibshirani, 2013) 18](#_Toc89964985)

[Figura 3 Funcionamento do neurónio artificial (Faceli et al., 2011) 24](#_Toc89964986)

[Figura 4 Interação entre um agente e o seu ambiente. Adaptado: (Tom M. Mitchell, 1997) 27](#_Toc89964987)

[Figura 5 – Gráfico de linhas com a variação do limiar em relação ao número de variáveis selecionadas 48](#_Toc89964988)

[Figura 6 - Esquema representativo da divisão dos dados e posterior balanceamento dos dados de treino. A - Contagem de categorias da variável ‘Delirium’ antes da divisão dos dados. B - Contagem de categorias da variável ‘Delirium, para os dados de treino. C - Contagem de categorias da variável ‘Delirium, para os dados de teste. D - Contagem de categorias da variável ‘Delirium, para os dados de treino após o balanceamento dos dados. 81](#_Toc89964989)

[Figura 7 - Esquema de divisão dos dados para treino e teste do modelo gerado 83](#_Toc89964990)

# Lista de Tabelas

[Tabela 1 - Tabela resumo dos fatores considerados precipitantes para o delirium (Sharon K. Inouye et al., 2014; Laurila et al., 2008; Nagari & Suresh Babu, 2019) 33](#_Toc89964998)

[Tabela 2 - Lista das ferramentas para diagnóstico de delirium (De & Wand, 2015) 39](#_Toc89964999)

[Tabela 3 - Escala RASS (Ely et al., 2003) 40](#_Toc89965000)

[Tabela 4 – Resultado da contagem de colunas com valores únicos 45](#_Toc89965001)

[Tabela 5 - Lista de medicamentos considerados para o desenvolvimento de delirium 49](#_Toc89965002)

[Tabela 6 - Opiáceos: indicações e efeitos adversos 51](#_Toc89965003)

[Tabela 7 – Benzodiazepinas: principais indicações e efeitos adversos frequentes 53](#_Toc89965004)

[Tabela 8 - Substâncias ativas dos ADT e respetivas indicações e efeitos secundários 55](#_Toc89965005)

[Tabela 9 - ISRS: indicação terapêutica e efeitos adversos frequentes 56](#_Toc89965006)

[Tabela 10 - Antipsicóticos: indicações e efeitos adversos frequentes 59](#_Toc89965007)

[Tabela 11 – Estatinas: efeitos adversos mais frequentes 64](#_Toc89965008)

[Tabela 12 – Efeitos secundários dos glucocorticoides 73](#_Toc89965009)

[Tabela 13 - Agrupamento das variáveis na categoria 'Outros Med' 74](#_Toc89965010)

[Tabela 14 - Codificação das variáveis categóricas 75](#_Toc89965011)

[Tabela 15 - Codificação das variáveis categóricas com mais do que duas instâncias 76](#_Toc89965012)

[Tabela 16 - Transformação dos valores da 'Local\_SU' numa escala de gravidade e posterior codificação 77](#_Toc89965013)

[Tabela 17 – Variáveis numéricas e os respetivos valores de referência e intervalos de valores ocorridos na base de dados 79](#_Toc89965014)

[Tabela 18 - Parâmetros da função RandomForestClassifier() e respetivo significado 84](#_Toc89965015)

# Introdução

A presente dissertação foi realizada no âmbito do Mestrado em Engenharia de Sistemas. Neste capítulo pretende-se apresentar o contexto do tema da dissertação, os objetivos deste projeto, assim como a metodologia seguida e, por último, será descrita a estrutura do documento.

## Enquadramento

Nas últimas décadas, tem havido um aumento considerável da esperança média de vida, não só em Portugal mas também na Europa (PORDATA, 2020). Este aumento da longevidade na população pode ser explicado por vários fatores, incluindo melhores cuidados de higiene, aumento de conhecimento na área da alimentação e consequente desenvolvimento de boas práticas de agricultura, a descoberta de vacinas, e também o constante desenvolvimento na medicina, que possibilitou a conceção de diagnósticos mais rápidos e certeiros, aliado ao desenvolvimento da farmacologia, esta que auxilia no tratamento de diversas patologias.

No entanto, paralelamente a este envelhecimento da população, tem-se assistido a um acréscimo de doenças crónicas caracterizadas por produzirem elevados graus de incapacidade sendo muitas vezes responsáveis por pressões sobre o sistema de saúde. O envelhecimento biológico a que o indivíduo está sujeito tem uma evolução variável, e passa por uma degradação natural onde podem ocorrer inúmeras alterações no normal funcionamento do organismo, destacando-se aqui a deterioração a nível cognitivo, esta que é uma alteração bastante comum em pessoas idosas. O *delirium* é uma manifestação comum de disfunção neuropsiquiátrica aguda em doentes admitidos em contexto hospitalar. Clinicamente, caracteriza-se pela turvação do estado de consciência, défice de atenção e por um distúrbio cognitivo e comportamental (American Psychiatric Association, 2013; INFARMED, 2010b; Salluh et al., 2015). É uma entidade muito prevalente, sobretudo na população idosa internada (American Psychiatric Association, 2013) e em ambientes de terapia intensiva (J. E. Wilson et al., 2020). Porém, qualquer pessoa, independentemente da idade, que se encontre em contexto hospitalar, tem probabilidade de desenvolver *delirium* (Lee-Archer, von Ungern-Sternberg, Reade, Law, & Long, 2021), sendo por isso fundamental o seu correto diagnóstico para que sejam aplicadas o mais precocemente possível as devidas intervenções clínicas, com vista a minimizar as complicações adjacentes a este problema.

Apesar de grave e potencialmente fatal, o *delirium* é frequentemente subdiagnosticado e negligenciado, acredita-se que devido à sua variabilidade de apresentação clínica principalmente quando o contexto do doente é complexo (Salluh et al., 2015; Van Eijk et al., 2009). Esta síndrome está relacionada não só com maiores taxas de morbilidade e mortalidade, como também com o aumento do tempo de internamento, assim como deterioração do estado físico e mental (Sharon K. Inouye, Westendorp, & Saczynski, 2014; Michaud et al., 2007). Nos últimos anos, tem sido desenvolvidos vários instrumentos de avaliação clínica para o *delirium*, o que representou um importante avanço metodológico no estudo e diagnóstico desta perturbação. Dos instrumentos de triagem do *delirium* existentes, destacam-se o Método de Avaliação da Confusão (CAM) e as suas variações de acordo com o ambiente em que se encontra o doente, a Escala de Sedação de Agitação de Richmond (RAAS) e a Escala de Confusão de Neelon e Champanhe (NEECHAM) (Mariz, Castanho, Teixeira, Sousa, & Santos, 2016). Para além disso, também tem vindo a ser estudados os diferentes fatores de risco, bem como várias perspetivas acerca das vias fisiopatológicas envolvidas nesta perturbação neurocognitiva (MacLullich, Ferguson, Miller, de Rooij, & Cunningham, 2008; Maldonado, 2008, 2017; van der Mast, 1998). No entanto, em ambiente de Serviço de Urgência (SU), o tempo muitas vezes é escasso e há uma necessidade de obter respostas rapidamente, pelo que este tipo de síndrome pode passar despercebida a muitos profissionais de saúde. Daí, ter surgido a necessidade de investigar ferramentas que possam permitir a elaboração de um diagnóstico de forma mais rápida e precisa que o habitual. Uma vez que o doente crítico se encontra exposto a vários fatores de risco, estes que podem precipitar o desenvolvimento do *delirium*, e facilmente passar despercebidos, urge promover a implementação de medidas que permitam o reconhecimento precoce do *delirium*. Nesse sentido, de forma a obter-se informações cada vez mais rigorosas, pensou-se em utilizar as técnicas existentes de *machine learning* (ML) para auxiliar na resolução deste problema. Uma abordagem típica utilizada na identificação das variáveis mais influentes passa pelo uso de modelos preditivos de regressão logística (RL), sendo que, estes modelos permitem igualmente estimar a probabilidade de um determinado indivíduo desenvolver ou não o *delirium,* bem como efetuar uma análise da relação entre as variáveis. Uma outra abordagem passa pelo uso do algoritmo *random forest* (RF), que é um algoritmos comummente utilizado na área de previsão, devido à sua simplicidade e o facto de poder ser utilizado tanto em tarefas de classificação como de regressão (Breiman, 2001). Para que esta tarefa de classificação de indivíduos se realize será necessário, após a seleção do melhor conjunto de variáveis independentes a incluir no modelo, avaliar a significância estatística dos parâmetros estimados. E, para além disso, é necessária a validação do modelo resultante a nível interno, no momento em que se utiliza a amostra original do estudo, e a nível externo, quando são utilizados dados novos, ou seja dados que o modelo nunca teve acesso (A. C. Braga & Carneiro, 2016). Estas validações do modelo desenvolvido podem ser avaliadas através da análise da curva ROC *(Receiver Operating Characteristic)*, na medida em que a área sob a curva ROC permite sumariar a qualidade do modelo produzido (A. Braga, 2000).

A principal motivação para o desenvolvimento deste trabalho, foi a preocupação com a melhoria do bem estar dos indivíduos admitidos em contexto de SU. Neste sentido, na presente dissertação, serão estudados alguns dos contornos que poderão levar ao desenvolvimento do *delirium*, assim como a avaliação do desempenho dos modelos gerados. Assim, face às ferramentas existentes e aos dados de saúde disponíveis eletronicamente no momento da admissão do doente no SU, pretende-se desenvolver uma aplicação, acessível aos profissionais de saúde, que determine o risco de desenvolvimento de *delirium* de um paciente no contexto do SU. Esta ferramenta tem o intuito de facilitar o diagnóstico de *delirium* para os profissionais de saúde e, consequentemente, melhorar a qualidade de vida do paciente.

## Objetivos

Este projeto de dissertação tem como principal objetivo o desenvolvimento de uma aplicação informática que auxiliará os profissionais de saúde no diagnóstico de *delirium* em contexto de SU. Com base nesta premissa, também se tem o objetivo secundário de responder à seguinte questão: “É possível determinar quais as variáveis mais influentes na previsão do desenvolvimento de *delirium*, através do uso dos dados disponíveis aquando da entrada de um doente no SU?”. Para a obtenção de resposta a esta questão, serão construídos modelos de ML que serão aplicados aos dados existentes. Inerentemente, serão estudados algoritmos de ML que melhor se adequem a este tema e selecionado o que produza melhores resultados.

De modo a desenvolver a aplicação de previsão de ocorrência ou não de *delirium*, será utilizada a linguagem *Python*, esta que é uma das linguagens mais populares na computação científica. Graças à sua natureza de alto nível e ao seu ecossistema de bibliotecas científicas, torna-se uma escolha apelativa não só para o desenvolvimento algorítmico, como também para análise exploratória de dados. Ainda dentro desta temática é importante ressaltar em primeiro lugar o módulo *﻿Scikit-learn*, pois integra uma vasta gama de algoritmos de ML (Pedregosa et al., 2012), e em seguida a biblioteca *Streamlit* que facilita a criação e partilha de aplicações web centrada no ML e ciência de dados.

Por fim, é esperado que no final do projeto seja desenvolvida uma aplicação funcional, com recurso a algoritmos estatísticos, capaz de diagnosticar antecipadamente o *delirium* em pacientes admitidos no SU. Desta forma, pretende-se que esta ferramenta facilite o diagnóstico desta doença e consequentemente melhore a qualidade de vida dos pacientes.

## Metodologia

O presente projeto de dissertação seguiu a metodologia CRISP-DM *(Cross Industry Standard Process for Data Mining)*. Esta é uma metodologia recente e surgiu pela necessidade que os profissionais de *Data Mining* (DM) sentiram ao desenvolver projetos relacionados com o processamento e análise de um grande volume de dados (Schröer, Kruse, & Gómez, 2021). O modelo CRISP-DM, define um projeto como um processo cíclico, onde podem ser utilizadas várias iterações para permitir um resultado final sintonizado com os objetivos do projeto. Nesta metodologia estão reunidas as melhores práticas para lidar com projetos relacionados com a mineração de dados. Possibilitando, assim, que se construam modelos da forma mais eficiente possível (Wirth & Hipp, 2000). No ano de 2000, foi concebido um esquema que permite a visualização das fases necessárias para a realização de projetos relacionados com a prospecção de dados. Na Figura 1, é apresentada uma adaptação do esquema original desenvolvido pelos autores, sendo possível identificar as seis etapas pertencentes ao modelo CRISP-DM, sendo elas: definição do problema, compreensão dos dados, tratamento dos dados, modelação, avaliação dos resultados e implementação (Chapman et al., 2000; Wirth & Hipp, 2000).

Diagram

Description automatically generated

Figura 1 – Diferentes fases implicadas na Metodologia CRISP-DM

*(Fonte: adaptado de (Chapman et al., 2000; Wirth & Hipp, 2000)*

Como é possível de observar pela Figura 1, esta metodologia define o ciclo de vida do projeto, dividindo-o em seis etapas, enunciadas anteriormente, que serão explicadas de seguida:

1. **Definição do problema:** A primeira fase da metodologia consiste em identificar os objetivos do projeto, perceber quais os recursos disponíveis (dados, software), determinar o objetivo do uso do ML neste projeto e por fim a construção do plano do projeto.

2. **Compreensão dos dados:** Esta etapa começa com uma recolha inicial de dados e prossegue através de atividades de familiarização com os dados. Assim, será possível não só identificar problemas de qualidade dos dados, como também adquirir conhecimento sobre os dados em estudo.

3. **Preparação dos dados:** A fase de preparação dos dados abrange todas as atividades para a construção do conjunto de dados final. Esta tarefa, pode ser executada várias vezes, e não tem qualquer ordem prescrita. Nesta fase, podem ser incluídas tarefas como a seleção de variáveis, limpeza de dados, construção de novos atributos e transformação de dados para as ferramentas de modelação.

4. **Modelação:** Na etapa de modelação são selecionadas as técnicas de modelação dos dados e verificadas as premissas para a mesma. Além disso, é nesta fase que se desenrola a construção dos modelos de ML. Muitas vezes, é nesta etapa que se constatam problemas nos dados ou, por outro lado, surgem novas ideias para a construção dos modelos.

5. **Avaliação dos resultados:** Nesta fase do projeto, tem-se construído um ou mais modelos de acordo com os objetivos pretendidos. No entanto, antes de proceder à implementação final do modelo, é importante avaliar mais profundamente o modelo, e rever as etapas executadas para a construção o modelo, para garantir que alcança corretamente os objetivos pretendidos no projeto.

6. **Implementação:** A última fase do projeto consiste em implementar os conhecimentos adquiridos nas fases anteriores. E para este projeto específico, será implementada uma aplicação baseada em algoritmos de ML que auxiliará os profissionais de saúde na deteção do *delirium* em doentes hospitalizados.

A utilização desta metodologia no projeto torna-se um processo útil para o planeamento, documentação e implementação, uma vez que permite compreender antecipadamente quais as ações que terão de ser levadas a cabo para a construção dos modelos pretendidos.

## Estrutura da dissertação

A estrutura deste documento consiste na apresentação de sete capítulos, cada um composto por diversas seções. O primeiro capítulo introduz o tema deste projeto, revelando as principais motivações para o seu desenvolvimento, bem como os objetivos a alcançar até ao final do projeto. Termina com uma breve descrição da estrutura desta dissertação.

No segundo capítulo, é apresentada a revisão bibliográfica dos principais temas tratados ao longo deste projeto. Começando por uma breve introdução ao ML, seguindo-se uma resumida abordagem histórica e evolução das técnicas de ML e terminando com um apanhado do estado da arte onde é correlacionado o ML e o *delirium*. De seguida, é apresentada uma história sucinta do *delirium* e, logo após, é descrita a definição de *delirium* bem como são descritas as suas principais características.

No Capítulo xx é apresentada a fundamentação teórica dos conceitos relativos ao ML. Começa-se por fazer uma

o caso de estudo, aqui é feita uma breve descrição do problema, assim como uma análise breve aos dados que se possui.

No Capítulo **Error! Reference source not found.** é apresentado o processo de preparação dos dados, onde são abordados temas como a exploração dos dados recorrendo a gráficos, limpeza de dados e a sua transformação para posterior modelação.

No Capítulo 6 é descrito o processo de modelação dos algoritmos de ML recorrendo à linguagem *Python*. Serão abordados alguns passos importantes na implementação dos algoritmos ao problema apresentado, assim como uma breve explicação do funcionamento da aplicação desenvolvida para a predição de *delirium*. Os algoritmo abordados para a resolução deste problema serão: regressão logística (RL), random forest, ...

No capítulo **Error! Reference source not found.** são apresentados os resultados da predição dos diversos algoritmos de ML, assim como a apresentação dos resultados acerca do funcionamento da aplicação informática desenvolvida.

No último capítulo (Capitulo 7) são apresentadas as principais conclusões deste projeto de dissertação. São ainda feitas sugestões de investigação futura para estudo neste tipo de projetos.

**(TERMINAR NO FINAL)**

# Revisão da literatura

Neste capítulo é apresentada uma revisão das principais áreas de domínio e conceitos utilizados neste projeto. A metodologia utilizada para a pesquisa dos conceitos discutidos neste capítulo, consistiu principalmente na análise de artigos científicos em conferências relevantes, revistas científicas e livros científicos.

## *Machine Learning*

Durante os últimos anos devido aos avanços da tecnologia, o armazenamento de dados tem sido uma prática recorrente (Jamin, Abraham, & Humeau-Heurtier, 2021). O que levou a um crescente interesse na prospeção de dados, ou na utilização de dados históricos para descobrir padrões e melhorar decisões futuras (T. M. Mitchell, 1999). Desde o início da era informática, vários investigadores têm concentrado o seu esforço para implantar a capacidade de aprendizagem em equipamentos informáticos, tendo surgido o conceito de inteligência artificial (IA) (Michalski, Carbonell, & Mitchell, 1983). Este conceito é vasto e tradicionalmente refere-se a criações artificiais que permitem imitar o funcionamento da inteligência humana para resolver problemas do dia-a-dia.

O ML é uma área de investigação da ciência da computação que utiliza conceitos de IA e métodos estatísticos para desenvolver algoritmos que aprendem e fazem previsões sobre os dados. Este campo da IA explora o estudo e a construção de algoritmos, que permitem aprender com dados, identificar padrões em enormes quantidades de dados e tomar decisões. A maior utilidade e impacto do conhecimento extraído a partir de dados e eventos históricos é a previsão de eventos e alterações similares no futuro (Murphy, 2012). Apesar de não ser nova, esta técnica tem vindo a ganhar importância nos últimos anos e é agora utilizada numa grande variedade de aplicações. Com o rápido desenvolvimento da IA, o ML e o reconhecimento inteligente têm sido cada vez mais aplicados às necessidades da vida humana (Xia, Wang, Yan, Dong, & Wang, 2019). Em algumas áreas, tais como, a medicina e cuidados de saúde, tem sido feita uma transição para o uso de instrumentos informáticos dependentes de dados. Este processo foi possibilitado pelos avanços simultâneos tanto no armazenamento de dados como pelo desenvolvimento tecnológico. Um estudo realizado em 2020 por Vellido, afirma que a conjetura atual do desenvolvimento tecnológico desencadeou a ideia que a utilização de ML seria o caminho a seguir para resolver problemas relacionados com a saúde, além de ser uma mais valia para a melhoria da qualidade dos serviços de saúde (Vellido, 2020). Do mesmo modo, Kareemi et al. (2021) destacam o potencial do ML implementado nos cuidados de saúde, ao promover uma melhoria na qualidade da medicina e ao permitir acelerar o ritmo de evolução de técnicas complexas de diagnóstico e terapêuticas. Os autores salientam que as ferramentas de ML utilizam os princípios centrais das abordagens estatísticas tradicionais, ao mesmo tempo que relaxam as limitações sobre o número de variáveis em estudo, variedades de dados de entrada, e os tipos de relações entre as variáveis (Kareemi, Vaillancourt, Rosenberg, Fournier, & Yadav, 2021). Para além disto, foram realizadas investigações que analisaram a utilização de modelos de ML em diversas áreas da saúde, e concluíram que os modelos de ML parecem ter melhor desempenho de diagnóstico e prognóstico em comparação com os cuidados habituais para pacientes admitidos em SU (Serviço de Urgência) (Jauk et al., 2020; Kareemi et al., 2021; Stewart, Sprivulis, & Dwivedi, 2018; Vellido, 2020).

### Breve História e Evolução

Ao longo da linha temporal, foram feitos esforços para mecanizar o pensamento, começando pelos primeiros exemplos mitológicos e literários, seguindo-se os textos filosóficos, fórmulas matemáticas e finalizando com os autómatos e outros dispositivos eletrónicos (McCorduck, 2004). A aventura pela inteligência artificial começou a dar frutos a partir do ano de 1943, quando Warren McCulloch e Walter Pitts escreveram um artigo sobre o funcionamento dos neurónios e desenvolveram um modelo computacional para redes neuronais baseadas em algoritmos de lógica­­­­­ (McCulloch & Pits, 1943). Este estudo foi a rampa de lançamento para o desenvolvimento da área da IA. No ano de 1950, Alan Turing publicou o artigo intitulado “*Computing Machinery and Intelligence*,” no qual propôs o “Teste de Turing” que consistia em o computador ter a capacidade de convencer um humano de que é um humano e não um computador.

Frank Rosenblatt, em 1958, concebeu a primeira rede neural artificial chamada “*The Perceptron*”. Com este estudo descobriu que o sistema desenvolvido era capaz de reconhecer padrões (Rosenblatt, 1958). Em 1959, Arthur Samuel, um pioneiro americano no campo dos jogos de computador, ML e IA estudou procedimentos de ML e verificou que um computador poderia ser programado para que aprendesse a jogar um jogo de damas num curto espaço de tempo, tal como as pessoas. Para tal acontecer, apenas seria necessário programar todas as diretrizes do jogo. E concluiu que este mecanismo de aprendizagem poderia ser aplicado a problemas da vida real (Samuel, 1959).

Já em 1987 Laird, Newell & Rosenbloom apresentaram o “*Soar*” com o objetivo de fornecer a estrutura que permitiria a um sistema executar tarefas cognitivas, aplicar métodos de resolução de problemas, e aprender sobre os aspetos das tarefas e do desempenho (Laird, Newell, & Rosenbloom, 1987). Em 1997, o computador de xadrez chamado de “*Deep Blue*” da IBM venceu o campeão mundial de xadrez (McCorduck, 2004). Desde então, houve muitos avanços no campo de ML, nomeadamente a partir de 2000 sugiram para comercialização os primeiros robots de estimação e brinquedos inteligentes. Em 2003, DARPA *(Defense Advanced Research Projects Agency)* deu inicio a grandes projetos de IA, nomeadamente o “*LifeLog*” (um diário eletrónico permanente da vida das pessoas), porém foi cancelado em 2004 por violar as políticas de privacidade. Em 2015, foi lançada a plataforma *“Amazon Machine Learning Platform”*, esta ferramenta impulsiona os sistemas internos assim como as recomendações de pesquisa, a *“Alexa*”, e serviços como “*Prime Air”* e “*Amazon Go”*.

Atualmente são utilizados variados algoritmos de ML no nosso quotidiano. Muitas empresas na área do comércio eletrónico, hotelaria, restauração e saúde utilizam estes sistemas, não só para proporcionar uma experiência melhor ao consumidor, mas também para retirar as diversas vantagens que deles advêm. A *Amazon* foi a primeira grande empresa a utilizar algoritmos de ML, em 1998, sendo pioneira em recomendações de produtos para milhões de clientes. A plataforma *YouTube*, é mais um exemplo da utilização de ML, neste modelo são selecionados os vídeos de acordo com os interesses do utilizador. Um outro exemplo, são os fornecedores de serviços de *e-mail* que utilizam um modelo de ML que permite detetar e mover automaticamente as mensagens não solicitadas para a pasta de *spam*.

### Uso de *Machine Learning* na medicina

Na literatura podem ser encontrados vários estudos que relacionam o conceito de ML aplicado a diversas áreas da saúde (Alsuliman, Humaidan, & Sliman, 2020; Basu, Faghmous, & Doupe, 2020; Corradi, Thompson, Mather, Waszynski, & Dicks, 2018; Nemati et al., 2018; Taylor & Haimovich, 2021; Xue et al., 2021). Pelo que, esta é uma temática atual e com evidência cientifica em desenvolvimento. Em 2018, foi realizado um estudo que avaliou a predição de *delirium* usando o algoritmo *random forest* (RF). Para tal, previamente foi executada uma recolha de dados, que implicou a realização do rastreio de *delirium* através do CAM *(Confusion Assessment Method)* e também a recolha dos dados de saúde eletrónicos de 64038 pacientes. Estes dados foram divididos aleatoriamente em 80% para treino e 20% para teste e aplicados ao algoritmoRF. Este modelo de previsão produziu uma área sob a curva ROC de 0.909, o que demonstrou que este algoritmo possui um grau elevado de precisão e potencial para fornecer um modelo preditivo útil na prática clínica (Corradi et al., 2018). Já em 2021, foi publicado um estudo de coorte retrospetivo que desenvolveu e validou algoritmos de ML para a deteção do *delirium*. Para a execução deste estudo foram recolhidos dados durante 5 anos e para a realização do rastreio foram utilizados os métodos DOSS (*Delirium Observation Screening Scale*) para os doentes internados e CAM-ICU *(Confusion Assessment Method for the Intensive Care Unit)* para doentes ventilados. Também foram recolhidos dados acerca do histórico médico, medicamentos administrados, medições fisiológicas e resultados laboratoriais. Os algoritmos estudados incluíram regressão logística (RL), Árvore de Decisão (AD), RF, *Gradient Boosting Machine* (GBM), *Gaussian Naïve Bayes* (GNB), *Support Vector Machine* (SVM), e K *Nearest Neighbor* (KNN). Com este estudo, foi possível concluir que os algoritmos de RF, GBM e RL apresentaram a melhor capacidade de previsão com o valor da área sob a curva ROC de 0,85 a 0,86. Tendo demonstrado, que o uso de algoritmos de ML para a identificação de *delirium* pode ser uma boa abordagem na prática clinica, na medida em que podem permitir identificar casos que passariam despercebidos (Lee, Mueller, Nick Street, & M. Carnahan, 2021). Para além deste estudo, também outros estudos salientam que os modelos de previsão baseados em ML permitem auxiliar os profissionais de saúde na identificação de diferentes patologias aquando do internamento hospitalar (Davoudi et al., 2017; Lee et al., 2021; A. Wong et al., 2018).

Perante o que foi descrito anteriormente, destaca-se a importância de investir em métodos que permitam a antecipação do diagnóstico de *delirium.* Neste sentido, é importante frisar que a identificação precoce de doentes com risco de desenvolver *delirium* pode facilitar a prevenção desta perturbação e assim melhorar a qualidade de vida dos pacientes. Neste seguimento, foram desenvolvidos modelos preditivos para a deteção do *delirium*, que demonstraram ser uma vantagem na prática clinica diária (Van Den Boogaard et al., 2012). O PRE-DELIRIC *(PREdiction of DELIRium for Intensive Care patients)*, foi um modelo de previsão do *delirium* criado em 2012 para uso na medicina de cuidados intensivos. Este modelo prevê o desenvolvimento de *delirium* ao longo do internamento, mediante 10 preditores (idade, grupo diagnóstico, coma, admissão urgente, administração de morfina, ureia, infeção, sedação, acidose metabólica, pontuação APACHE-II *(Acute Physiology and Chronic Health Evaluation-II)* avaliáveis 24 horas após a admissão do doente (Liang et al., 2020; Van Den Boogaard et al., 2012). Segundo Liang et al. (2020), o PRE-DELIRIC tem um elevado valor preditivo e é sugerido que este modelo seja adotado nas UCI para a deteção do *delirium* em doentes de alto risco, pois contribui para uma melhor gestão de recursos assim como uma melhoria na vida dos pacientes. Em 2015, foi validado outro modelo para deteção precoce do *delirium* para cuidados intensivos, denominado por E-PRE-DELIRIC *(Early PREdiction of DELIRium for Intensive Care patients)*. Este modelo é constituído por nove preditores: idade, histórico de alterações cognitivas, histórico de abuso de álcool, níveis de ureia no sangue, grupo diagnóstico, admissão urgente, tensão arterial média, administração de corticosteroides e insuficiência respiratória. Este estudo surgiu como necessidade de colmatar a lacuna do modelo anterior ter a limitação de exigir preditores obtidos durante as primeiras 24 h de admissão na UCI. Pelo que, o modelo E-PRE-DELIRIC utiliza os dados disponíveis na admissão à UCI para prever o desenvolvimento do *delirium* durante o tempo de internamento do paciente (Wassenaar et al., 2015).

Com o objetivo de perceber qual dos dois modelos estaria melhor preparado para o uso clinico, foi realizado o estudo “*Delirium prediction in the intensive care unit: comparison of two delirium prediction models*” no ano de 2017. Este estudo concluiu que o modelo PRE-DELIRIC surge como ferramenta mais fiável, no entanto, os médicos da UCI classificaram a ótica do utilizador o E-PRE-DELIRIC como superior ao PRE-DELIRIC. E ainda que, em pacientes de baixo risco, a previsão do *delirium* melhora ainda mais após uma atualização com o modelo PRE-DELIRIC após 24 h (Wassenaar, Van Den Boogaard, Schoonhoven, Donders, & Pickkers, 2017).

Já em 2018, foi desenvolvido um estudo que testou a viabilidade do uso de uma aplicação para telemóvel para monotorização de perturbações a nível de excitação e atenção associadas ao *delirium*. A aplicação para telemóvel denominada por DelApp, mostrou-se promissora como uma ferramenta objetiva para auxiliar na deteção do *delirium*, com potenciais aplicações na prática clínica após desenvolvimento posterior e validação formal (Tang et al., 2018). No entanto, esta aplicação não faz uso de técnicas de ML, apenas usa um conjunto de testes para avaliar o nível de atenção de um paciente e posteriormente atribui uma pontuação, permitindo avaliar a ocorrência ou não de *delirium* (Rutter et al., 2018). Com os resultados destes estudos pode-se concluir que as ferramentas informáticas podem ser uma mais valia no diagnóstico do *delirium.*

## *Delirium*

### Breve introdução histórica

O *delirium* foi uma das primeiras doenças psiquiátricas descritas na literatura médica, há mais de 2500 anos, inicialmente descrita por Hipócrates (Zbigniew J. Lipowski, 1991). As suas obras datam de 460-366 a.c., e embora não tenha sido utilizado o termo “*delirium*” são descritas anomalias mentais causadas pela febre, venenos ou traumatismo craniano, que se pronunciam com sinais de entorpecimento dos sentidos, sonolência, inquietação ou mesmo comportamentos violentos (Isik & Grossberg, 2018). Estes sintomas foram muitas vezes descritos como frenite, que tipicamente envolve insónias e alucinações, e também como letargia que envolve sonolência e uma paralisação injustificada.

O termo *"delirium"*, deriva do latim *"deliro-delirar*e" e traduz-se como "fora do caminho", este termo foi aceite e usado pela primeira vez por Aulus Cornelius Celsus no primeiro século d.c. (Adamis, Treloar, Martin, & Macdonald, 2007; Zbigniew J. Lipowski, 1991). Este enciclopedista ficou conhecido por ser o primeiro a relatar causas não-febris, dando o exemplo do consumo de vinho em excesso, como razão para o desenvolvimento de *delirium*. No segundo século, o escritor médico capadócio, Aretaeus, observou que o *delirium* diferia das doenças crónicas (demência) em termos de duração e foi provavelmente o primeiro a recomendar um espaço calmo e escuro para um paciente delirante e haxixe (papoila cozida) (Isik & Grossberg, 2018).

Entretanto, a documentação mais recente relativa a esta temática aparece no século XVI. No ano 1583, é publicado um livro da autoria de ﻿Barrough, entitulado *“The Method of Physic”.* Neste livro, o autor referiu-se ao *delirium* como *"frenesie"* e observou que envolvia o transtorno de três funções principais: imaginação, cogitação e memória, e realçou que os doentes apresentavam também distúrbios no sono. No decurso do século XVI e XVII foram publicadas várias dissertações sobre o *delirium*, nas quais foram aprofundadas as características mais comuns desta síndrome (Zbigniew J. Lipowski, 1991). Em 1683, Thomas Willis frisou que o *delirium* não se tratava de uma doença, mas sim de um sintoma associado à febre como resultado de envenenamento, hemorragia, falta de sono e/ou embriaguez. Segundo o autor, o *delirium* caracterizava-se principalmente por conceções incongruentes e pensamentos confusos, perceções visuais distorcidas e comportamento perturbado.

Uma descrição mais sofisticada da psicopatologia do *delirium* a aparecer no século XIX foi a de Greiner, no ano de 1817. Este autor referiu-se à síndrome como "insanidade febril", uma desordem que envolvia a turvação da consciência. Esta foi provavelmente a primeira vez que o *delirium* foi ligado explicitamente à consciência desordenada. Esta associação foi estudada por Hughlings Jackson nos anos 1860, que concluiu que o *delirium* era caracterizado por um estado de consciência reduzida que variava de um grau de menor confusão até ao estado de coma. E que esta oscilação acontecia devido a um determinado grau de dissolução da camada dos centros nervosos e consequente libertação da inibição dos centros de camadas inferiores. Desta forma, o *delirium* começou a ser visto como uma manifestação clínica de consciência perturbada (ou turvação da consciência), este diagnóstico foi frequentemente utilizado na segunda metade do século XIX, porém alguns escritores aplicaram-no tanto ao *delirium* como a certas perturbações mentais funcionais, tais como a histeria (Zbigniew J. Lipowski, 1991).

Um outro termo vinculado ao *delirium* no século XIX foi o de "confusão". Este termo foi introduzido por escritores alemães e franceses, e a palavra "confusão" referia-se a uma incapacidade de pensar de forma lógica e coerente, a uma perda de memória e a uma perceção desordenada. Assim, estes dois usos não eram realmente específicos do *delirium*. A partir da palavra "confusão" foram derivados vários outros termos, nomeadamente "estados confusionais agudos" e "insanidade confusional aguda".

Já em 1959, surge uma das contribuições mais importantes para o estudo de *delirium*. Com base em estudos clínicos e experimentais, Engel e Romano concluíram que o *delirium* era um distúrbio no nível de consciência, manifestado por distúrbios cognitivos de atenção. A síndrome era devida a uma redução da taxa metabólica do cérebro, esta indicação foi suportada pelos registos de um eletroencefalograma (EEG), que indicava que quanto maior fosse a perturbação, maior seria a desaceleração (Field & Wall, 2013; Zbigniew J. Lipowski, 1991). Nos anos 60, vários investigadores estudaram o *delirium* induzido experimentalmente em voluntários através da administração de drogas com atividade anticolinérgica. Este trabalho destacou o papel patogénico do bloqueio colinérgico no *delirium* e forneceu uma pista importante para o reconhecimento da suscetibilidade geralmente apresentada pelos doentes com doença de Alzheimer para o desenvolvimento da síndrome do *delirium* (Zbigniew J. Lipowski, 1991).

Apesar de parecer que havia um consenso na nomenclatura desta síndrome, no último século, segundo vários autores, foram empregues múltiplos termos como sinónimos, incluindo: síndrome confusional ou estado confusional agudo, agitação, alteração da consciência, encefalopatia, falência cerebral aguda, síndrome cerebral agudo, psicose, entre outros (Z. J. Lipowski, 1987; Morandi et al., 2008; J. E. Wilson et al., 2020). A falta de uma terminologia consistente afetou negativamente a investigação deste distúrbio (Slooter et al., 2020), além de ter contribuído para uma sub-representação maciça nos dados de alta hospitalar (Casey et al., 2019). Estas questões suscitaram a necessidade de consenso sobre a nomenclatura por parte da comunidade científica. Portanto, em 1980, a *American Psychiatric Association* (APA) publicou o DSM-III *(Diagnostic and Statistical Manual of Mental Disorders)* na tentativa de rever a classificação de síndromes de cérebro. Lipowski, foi um dos responsáveis por essa revisão e propôs a reintrodução dos termos “*delirium*” e “demência” na classificação oficial, conseguindo-se desta forma como resultado uma terminologia uniforme. Para além disso, foram explicitadas as características clínicas essenciais da síndrome e formulados critérios de diagnóstico para a mesma (Zbigniew J. Lipowski, 1991). Nesta edição do DSM-III, o *delirium* foi caracterizado como um estado de consciência turvado, ou seja uma dificuldade em manter a atenção tanto a estímulos externos como internos, uma má perceção sensorial, e uma corrente de pensamento desordenada. Para além disso, foi descrita também a presença de distúrbios no sono-vigília e na atividade psicomotora, e também que o início desta síndrome podia ser relativamente rápido e o curso podia flutuar ao longo do dia (American Psychiatric Association, 1980).

Posteriormente, esta edição foi revista e atualizada com o DSM-IV (1994), que tentou simplificar a definição prévia, classificando o *delirium* segundo a sua etiologia e destacando a alteração de consciência como elemento fundamental para o seu diagnóstico. Em 2013, foi publicado o DSM-5 (American Psychiatric Association, 2013) que classifica o *delirium* no conjunto de perturbações neurocognitivas.

### Definição e características gerais

O *delirium* encontra-se definido e descrito na 5ª edição do DSM (DSM-V), e suportado pela 10ª revisão da *International Classification of Diseases* (ICD-10)*,* como um quadro confusional agudo de evolução flutuante. Apesar das diversas definições descritas para o conceito de *delirium*, uma globalmente aceite é a de “disfunção cognitiva global aguda” (Bourgeois, Hategan, & Losier, 2014). Segundo a definição presente no DSM-V o *delirium* é uma deficiência cognitiva global com capacidade reduzida de concentração e atenção, isto é, capacidade reduzida de atenção, focalização, sustentação e desvio da atenção e orientação reduzida no ambiente. Esta definição salienta ainda que, a perturbação desenvolve-se durante um curto período de tempo (geralmente horas a alguns dias), e tende a flutuar em intensidade no decurso de um dia. É também mencionada uma perturbação adicional na cognição, como por exemplo, défice de memória, desorientação, linguagem, capacidade visio-espacial, ou perceção do ambiente, quando estas ocorrências não são explicadas por outra perturbação preexistente estabelecida ou evolutiva (American Psychiatric Association, 2013).

O *delirium* reflete um estado de falência cerebral aguda, potencialmente reversível se diagnosticado e tratado de forma correta e no tempo certo. De um modo geral, qualquer insulto agudo que comprometa a homeostasia do sistema nervoso central (SNC) pode precipitar um episódio de *delirium*. Por outro lado, a homeostasia do SNC pode apresentar um estado debilitado devido a alterações associadas ao normal envelhecimento ou por alterações neurodegenerativas, como a demência (Saraiva & Cerejeira, 2014). A etiologia é, na maioria dos casos, de natureza multifatorial, e depende da interação entre a vulnerabilidade do doente e a severidade dos fatores precipitantes (Harrison, Cowen, Burns, & Fazel, 2018). E segundo o DSM-V é possível subdividir o *delirium* em cinco categorias etiológicas:

* *Delirium* devido a intoxicação por substância: álcool, cannabis, fenciclidina, opiácios, sedativos, hipnóticos, ansiolíticos, cocaína, anfetaminas, entre outros.
* *Delirium* por abstinência de substância: álcool; opiácios, sedativos, hipnóticos, ansiolíticos ou outras substâncias/medicação desconhecidas.
* *Delirium* induzido por fármaco: medicação opióide, sedativa, hipnótica, ansiolítica, anfetamina ou outras.
* *Delirium* devido a uma condição médica: infeção, desidratação, insuficiência renal, hepática, cardíaca, cirurgia, algaliação, dores, entre outros
* *Delirium* devido a múltiplas etiologias.

Portanto, de um modo geral pode-se afirmar que o *delirium* é definido como uma síndrome neuropsiquiátrica aguda, caracterizada por um transtorno agudo da atenção e cognição, de natureza multifatorial. É uma entidade muito prevalente, sobretudo na população idosa e ocorre em diversos ambientes clínicos. E de acordo com a predominância dos sintomas acompanhantes do doente: a atividade psicomotora e o nível de vigília, o episódio clínico de *delirium* pode ser classificado em três subtipos distintos:

1) *Delirium* Hiperativo

Os pacientes diagnosticados com este subtipo apresentam um quadro de hiperatividade psicomotora e, na maioria das vezes não dormem. Manifestam um aumento de atividade motora e ansiedade e por vezes podem adotar um comportamento agressivo e ameaçador. Podendo também expor um discurso confuso e alucinações (Lee-Archer et al., 2021; Nagari & Suresh Babu, 2019). Esta forma de *delirium* ocorre principalmente no diagnóstico de *delirium tremens* por abstinência alcoólica, síndromes de abstinência de medicamentos e pela ação de drogas anticolinérgicas.

2) *Delirium* Hipoativo

No *delirium* hipoativo os doentes apresentam-se sonolentos, apáticos, movem-se lentamente, falam pouco e podem apresentar um diminuição no apetite assim como diminuição da consciência do ambiente (Nagari & Suresh Babu, 2019). Este diagnóstico pode ser confundido com o quadro de depressão ou falta de motivação. Este subtipo apresenta, portanto, um atraso de diagnóstico considerável e associa-se a doença severa, maior morbi-mortalidade e maior necessidade de cuidados paliativos (Saraiva & Cerejeira, 2014). O subtipo hipoativo pode desenvolver-se por intoxicação de drogas hipnóticas ou sedativas, hipóxia, encefalopatia, sendo este o tipo mais comum do que o subtipo hiperativo nos idosos (Nagari & Suresh Babu, 2019; Saraiva & Cerejeira, 2014).

3) *Delirium* Misto

Neste tipo de *delirium* o doente alterna entre períodos de hiperatividade e de hipoatividade podendo ocorrer durante um dia ou vários (Nagari & Suresh Babu, 2019).

No que diz respeito ao curso da síndrome em ambientes hospitalares, o *delirium* dura geralmente cerca de uma semana, mas alguns sintomas persistem frequentemente mesmo depois de os indivíduos terem alta do hospital (American Psychiatric Association, 2013). Os indivíduos com *delirium* podem alternar rapidamente entre o estado hiperativo e hipoativo, sendo que o estado hiperativo pode ser mais comum ou mais frequentemente reconhecido e está associado a efeitos secundários da medicação e à remoção de medicamentos. Já o estado hipoativo é mais frequente desenvolver-se em adultos mais velhos (American Psychiatric Association, 2013; Field & Wall, 2013).

# Machine Learning

Uma das áreas de aplicação das técnicas de ML que tem uma boa recetividade é a área da saúde. A prova disso é que tem sido utilizadas aplicações de ML que auxiliam os profissionais de saúde na realização de exames clínicos, na monitorização do estado dos pacientes, na dosagem de medicamentos e ainda no diagnóstico médico.

Atualmente, existe um vasto número de algoritmos de ML. E, dependendo da natureza do problema a ser tratado, podem ser aplicadas diferentes abordagens baseadas no tipo e volume dos dados. Sendo possível categorizar os sistemas de ML ao longo de diversas dimensões. Contudo, com base no *feedback* disponível durante a aprendizagem, pode-se distinguir: aprendizagem supervisionada; aprendizagem por reforço e aprendizagem não supervisionada (Basu et al., 2020; Bruha & Berka, 2000; Libbrecht & Noble, 2015). Ao longo desta secção serão discutidas as diferentes abordagens e serão apresentados alguns exemplos e também uma breve explicação do funcionamento de cada uma delas, com especial atenção nos algoritmos RF e RL.

## Aprendizagem supervisionada

A aprendizagem supervisionada é definida pela utilização de conjuntos de dados com informações sobre a classificação dos mesmos (Hurwitz & Kirsch, 2018). Este tipo de aprendizagem usa um conjunto de dados de treino com as entradas (características) emparelhadas com os valores de resposta correspondentes (por exemplo, ‘sim’ ou ‘não’). Durante o treino, o algoritmo procura padrões nos dados que se correlacionem com as saídas desejadas. Após o treino, o algoritmo recebe novos dados e atribui um rótulo, com base nos dados de treino prévio. O objetivo deste modelo de aprendizagem é prever a etiqueta correta para um novo conjunto de dados. Podendo ser usado quando estão disponíveis dados históricos e o objetivo é prever resultados futuros. Desta forma, é possível comparar os valores gerados pelo sistema com os esperados, através dos dados de teste (Friedman, 2002; Libbrecht & Noble, 2015).

Por vezes, os padrões identificados num subconjunto de dados podem não ser detetados num conjunto maior de dados. Quando o modelo está adequado apenas para representar padrões que existam no subconjunto de treino, pode ocorrer o problema chamado *overfitting*. Isto significa que o modelo está precisamente afinado para os dados de treino, mas não pode ser aplicado a grandes conjuntos de dados desconhecidos (Hurwitz & Kirsch, 2018).

### Regressão Logística

A RL é uma técnica estatística que tem como objetivo produzir, a partir de um conjunto de observações, um modelo que permite a predição de valores tomados por uma variável categórica, em função de uma ou mais variáveis independentes contínuas e/ou binárias (Hosmer, Lemeshow, & Sturdivant, 2013). Esta técnica utiliza mais ou menos os mesmos princípios gerais utilizados na regressão linear, a diferença está na variável reposta que na RL é binária (dicotómica) enquanto que no modelo de regressão linear é contínua (Hosmer et al., 2013).

Segundo Stoltzfus (2011), as técnicas de regressão são versáteis no momento da sua aplicação à investigação médica, pois permitem prever resultados e controlar variáveis. Defende ainda que a RL é uma forma eficiente e poderosa de analisar o efeito de um grupo de variáveis independentes sobre um resultado binário, quantificando a contribuição de cada variável (Stoltzfus, 2011).

Tem sido propostas várias funções de distribuição para serem usadas na análise de uma variável resposta dicotómica. Neste caso, é apresentado um modelo baseado na distribuição logística (equação 2.1). Pois, é uma função extremamente flexível, de fácil utilização e os seus parâmetros fornecem a base para a realização de estimativas com precisão (Hosmer et al., 2013).

(2.1)

Foi usada a notação para representar a média condicionada em que Y se refere à variável resultado e x representa um valor especifico da variável independente. E lê-se a probabilidade de ocorrer o valor Y, dado o valor de x. A função logística é conhecida pela produção de uma curva em forma de S, e assim, independentemente do valor de x, permite obter uma previsão mais assertiva quando comparada com o modelo de regressão linear. Na Figura 2 é apresentada a comparação gráfica entre o modelo de regressão linear e logística, que permite verificar os pontos de assertividade dos resultados produzidos pelas respetivas funções. Os pontos amarelos indicam os valores 0 e 1, representado nos dados a tradução para “Não” ou “Sim”, respetivamente. Na imagem da esquerda é estimada a probabilidade utilizando regressão linear e na da direita é estimada utilizando RL. Note-se que a função logística apresentada na figura é muitas vezes abreviada para função sigmoide, cujo valor z pode ser ajustado consoante as necessidades do problema.

Chart, line chart

Description automatically generated

Figura 2 - Comparação gráfica entre o modelo de RL e regressão linear (James, Witten, Hastie, & Tibshirani, 2013)

Após um pouco de manipulação da função apresentada na equação 2.1, chega-se à seguinte expressão:

(2.2)

A expressão p(E)/[1-p(E)] traduz a possibilidade de sucesso do evento E *(odds),* e pode assumir qualquer valor entre 0 e . Resultados próximos de 0 indicam probabilidades muito baixas de falha enquanto que valores próximos de apontam para probabilidades muito altas da não ocorrência do evento. Pela aplicação do logaritmo aos dois lados da equação 2.2 é possível chegar à transformação apresentada na equação 2.3 (James et al., 2013):

(2.3)

Esta transformação é denominada de transformação *logit* da probabilidade. E é importante na medida em que a função *logit* passa a ter muitas das propriedades desejáveis de um modelo de regressão linear. Podendo, ser linear nos parâmetros, ser contínua e variar entre os valores de -a +, dependendo do domínio de x (Hosmer et al., 2013). A função resposta transformada é denominada como função resposta *logit.*

No que diz respeito à distribuição condicional da variável resposta, pode-se expressar o valor da variável resultado através da expressão apresentada na equação 2.4. O símbolo representa o erro e expressa o desvio da observação em relação à média condicional. No caso de uma variável resposta dicotómica, o valor da variável resposta x é dada pela expressão:

(2.4)

em que o valor do erro, , pode assumir dois valores:

* se y = 1 então com a probabilidade ;
* se y = 0 então com a probabilidade de .

Assim, apresenta uma distribuição com média zero e variância igual a ] Ou seja, a distribuição condicional da variável resposta segue uma distribuição binomial com probabilidade dada pela média condicional () (Hosmer et al., 2013).

Para concluir, pode-se afirmar que os pressupostos básicos a ser cumpridos para a avaliação do modelo de RL incluem independência de erros, linearidade no *logit* para variáveis contínuas, ausência de multicolinearidade, e ausência de *outliers* fortemente influentes. A adequação global do modelo de RL resultante é avaliada pelos resultados caracterizados pela menor diferença entre os valores observados e os valores previstos no modelo. Estando comprovado que a RL é uma forma eficiente e poderosa de avaliar as contribuições das variáveis independentes para o resultado binário. Além disso, percebe-se que a obtenção de bons resultados depende em grande parte de uma seleção cuidadosa das variáveis que satisfaçam os pressupostos básicos (Stoltzfus, 2011).

**Critérios de avaliação do modelo de previsão**

Para que o modelo preditivo desenvolvido possa ser utilizado com segurança, será necessário avaliar a qualidade das predições obtidas. Para tal, ao longo do tempo foram desenvolvidas e testadas algumas métricas para esse efeito. E, Na literatura é possível encontrar-se diversas métricas de avaliação de desempenho de modelos de RL. Dentre os critérios para seleção de modelos destacam-se os seguintes:

* Curva ROC

Uma das formas para avaliar o diagnóstico dos modelos estatísticos obtidos passa pela utilização da curva ROC *(Receiver Operating Characteristic).* Esta é uma representação gráfica da fração de verdadeiros positivos, no eixo nas ordenadas, e da fração de falsos positivos, no eixo das abcissas. A análise da curva ROC permite estudar a variação da sensibilidade e especificidade para cada valor de *cut-off.* O objetivo desta análise é identificar ou confirmar a qualidade do ajustamento do modelo. Aquando da observação do gráfico, o cenário ideal implica encontrar uma área sob a curva ROC (AUC) perto de 1, uma vez que, quanto mais próxima estiver a curva do canto superior esquerdo, mais verdadeiros positivos e menos falsos negativos serão obtidos (Stoltzfus, 2011).

* Critério de Informação de Akaike

O *Akaike Information Criterion* (AIC) é um critério amplamente utilizado com a finalidade de selecionar o melhor modelo entre um conjunto de modelos candidatos (Yanagihara, Kamo, Imori, & Satoh, 2012). Cada modelo é caracterizado por um valor de AIC, sendo que, a seleção do melhor modelo é regida por aquele que apresentar o menor valor de AIC. O critério AIC é definido pela seguinte equação (Hosmer et al., 2013):

onde p representa o número de parâmetros independentes e 𝐿 representa o valor obtido da função de máxima verosimilhança do modelo.

* Critério Bayesiano de Schwarz

O *Bayesian Information Criterion* (BIC) é um critério de seleção de modelos de entre um conjunto finito de modelos, aquele que apresentar o menor valor de BIC será o escolhido (James et al., 2013). O BIC baseia-se na função da máxima verosimilhança e está intimamente relacionado com o critério AIC.

### Árvores de decisão

Este algoritmo divide repetidamente o conjunto de dados de acordo com um critério que maximiza a separação dos dados, resultando numa estrutura em forma de árvore. É usada a estratégia de dividir para conquistar com a finalidade de resolver um problema de decisão. Esta é a ideia base dos algoritmos baseados em árvores de decisão (AD). Ao longo do tempo foram desenvolvidos algoritmos baseados no modelo de árvore, pelo que, um dos primeiros algoritmos a ser desenvolvido foi o ID3 *(Iterative Dichotomiser),* em 1979. O algoritmo ID3 foi desenvolvido por J. Ross Quinlan, que criou uma construção simples começando pela raiz e terminando nas folhas *(construção top-down)* (J. Han, Kamber, & Pei, 2012). Porém, este algoritmo apresentava a limitação de só ser capaz de lidar com variáveis nominais e consequentemente apenas poderia ser utilizado em problemas de classificação. Na ânsia de resolver este problema, Quinlan desenvolveu o algoritmo C4.5 que permite variáveis numéricas contínuas, sendo capaz de para lidar com os problemas de regressão (para além dos problemas de classificação). Além disso, outras melhorias surgiram, tais como a possibilidade de poda das árvores. Esta última é considerada uma melhoria importante, uma vez que permite a remoção de ramos que causam "ruído" nos dados e, consequentemente, uma melhoria na precisão da previsão (Nikam, 2015). Este algoritmo é o mais utilizado nos modelos de ML.

Outro algoritmo desenvolvido por um grupo de estatísticos (L. Breiman, J. Friedman, R. Olshen, e C. Stone), em paralelo com o mencionado anteriormente, é o CART. Este é um algoritmo com uma abordagem semelhante à do C4.5, utilizada também para modelos de classificação e regressão (J. Han et al., 2012). As AD são geradas a partir de um conjunto de dados de treino, resultando na criação de uma árvore com a forma da estrutura utilizada para classificar os novos casos. Cada caso é descrito como um conjunto de características ou atributos associados com cada caso dos dados de treino. O resultado é o valor/nome da classe, ou seja, a característica alvo, a que está prevista (Quinlan, 1986).

As medidas de seleção de atributos determinam como as instâncias num determinado nodo devem ser divididas e, devido a isso, são também denominadas regras de divisão (J. Han et al., 2012). Assim, existem três medidas de seleção de atributos, descritas abaixo:

*Information Gain:*

*Information Gain* é uma medida de seleção de atributos baseada na entropia da sua contagem. É a medida utilizada no algoritmo ID3. A entropia é uma medida de incerteza que permite determinar qual deve ter lugar em cada posição de árvore. Está associado a um conjunto de objetos que permite identificação do grau de perturbação dos dados (baixa pureza), calculado para cada um dos atributos sob estudo (J. Han et al., 2012).

*Gain Ratio:*

Esta medida de seleção de atributos tem vindo a resolver algumas das limitações e especificidades do *information gain*. Assim, esta medida é utilizada no algoritmo C4.5 e é considerada uma extensão do ganho de informação. A equação seguinte, mostra a informação potencial que é criada através da partição do conjunto de dados de formação, 𝐷, num total de 𝑣 resultados, resultante do teste no atributo 𝑋 (𝑣 partições).

*Gini index*

Esta medida é aplicada no algoritmo CART. É utilizado para medir a impureza de um conjunto ou subconjunto de dados.

### *Random Forest*

Nos algoritmos RF, cada AD é construída utilizando um subconjunto de dados escolhidos aleatoriamente a partir dos dados de treino (Breiman, 2001). Cada subconjunto irá criar uma AD, e assim, ao testar o modelo com um *input*, cada árvore irá classificá-lo com a respetiva classe (James et al., 2013). Isto significa que no momento da construção do RF, de cada vez que a árvore é dividida não são considerados todos os preditores disponíveis. Esta abordagem permite contornar o problema do algoritmo de *bagging*, em que todas as árvores tinham um aspeto semelhante entre si (James et al., 2013). Este modelo também tem as suas vantagens e limitações. Funciona com sucesso em grandes conjuntos de dados e lida bastante bem com *outliers* e modelos "ruidosos". Tal como as AD, este modelo permite compreender quais as variáveis mais importantes. No entanto, requer mais esforços computacionais mas quando comparado com as AD apresenta bons resultados e previsões de alta precisão (J. Ali, Khan, Ahmad, & Maqsood, 2012).

### Máquina Vetorial de Apoio

O algoritmo máquina vetorial de apoio mais conhecido por SVM (Support Vector Machine) tem recebido uma atenção crescente na comunidade de ML, baseiam-se na teoria de aprendizagem estatística, desenvolvida por Vapnik em 1995 a partir de estudos iniciados por Vladimir N. Vapnik e Alexey Chervonenkis em 1968 (Vapnik, 2000). Esta teoria estabelece uma série de princípios a ser seguidos na obtenção de classificadores com boa capacidade de generalização. Este modelo é constituído por implementações algorítmicas da teoria da aprendizagem estatística que estabelece condições matemáticas que auxiliam na escolha de um classificador a partir de um conjunto de dados de treino (Faceli, Lorena, Gama, & Carvalho, 2011).

Uma característica atrativa deste modelo é a convexidade de otimização formulada durante o treino, que implica sempre a existência de um mínimo global. Além disso, o uso de funções *kernel* na não linearização das SVM torna o algoritmo eficiente, pois permite a construção simples de hiperplanos num espaço de alta dimensão de forma tratável do ponto de vista computacional, resolvendo um problema de otimização quadrática limitada (Burges, Christopher, 1998). Por outro lado, uma das principais limitações centra-se na sensibilidade na escolha de valores dos parâmetros e a dificuldade de interpretação do modelo gerado por esta técnica (Chapelle, Vapnik, Bousquet, & Mukherjee, 2002; Duan, Keerthi, & Poo, 2003).

### K vizinhos mais próximos

O algoritmo dos K vizinhos mais próximos (*K-Nearest Neighbor*, KNN) tem sido utilizado desde a década de 1950 na área de Estatística, e é utilizado para problemas tanto de classificação como de regressão. Está documentado como sendo um algoritmo lento, mas eficiente, e é recomendado para bases de dados que contenham muitas instâncias. Este algoritmo foi introduzido por Fix e Hodges, em 1951, que o classificaram um método não-paramétrico de classificação de padrões. O funcionamento deste algoritmo é relativamente simples, é identificado o valor de *K* vizinhos mais próximos para o ponto de dados desconhecido, sendo identificados com base na medida de distância escolhida, e o ponto desconhecido será classificado com base na classe maioritária entre as classes de pontos de dados mais próximos identificadas. A principal desvantagem do KNN é a complexidade da métrica que calcula a distância dos vizinhos mais próximos para cada amostra (Dreiseitl & Ohno-Machado, 2002).

### *Artificial Neural Network* (ANN)

As redes neuronais poder ser classificadas no quadrante de aprendizagem supervisionada se os dados utilizados possuírem o resultado desejado conhecido. Este tipo de aprendizagem é especialmente útil quando é necessária a aprendizagem de padrões a partir de dados não estruturados. As redes neuronais são concebidas para imitar a forma como o cérebro humano funciona, para que as máquinas possam ser treinadas para lidarem com abstrações e problemas mal definidos. Este tipo de aprendizagem, é frequentemente utilizada em aplicações de reconhecimento de imagem, voz e visão por computador.

O funcionamento da ANN (Artificial Neural Network) é inspirado no modelo físico concebido pela natureza, o cérebro humano (A. de P. Braga, Ludermir, & Carvalho, 2000). No dia-a-dia, são realizadas diversas tarefas que requerem a atenção a diferentes eventos ao mesmo tempo, assim como o processamento de diversas informações, com o intuito de tomar decisões. A complexidade de tais ações, simples para a maioria das pessoas, é evidenciada pela dificuldade encontrada a ensinar robôs. A partir destas motivações, o desenvolvimento de redes neuronais artificiais (RNAs) inspirou-se na estrutura e funcionamento do cérebro humano para a aquisição de conhecimento.

A procura por modelos computacionais ou matemáticos do sistema nervoso teve inicio na mesma época em que foram desenvolvidos os primeiros computadores, na década de 1940. Como já referido McCulloch e Pits (1943), propuseram um modelo matemático de neurónios artificiais, em que os neurónios executavam funções lógicas simples e cada um podia executar uma função diferente (McCulloch & Pits, 1943). Ao longo dos anos, foram realizados diversos estudos para compreender o cérebro e utilizar o conhecimento obtido para desenvolver sistemas de aprendizagem biologicamente aceitáveis. Desta forma, foi possível desenvolver RNAs baseadas em modelos abstratos do funcionamento do cérebro.

As RNAs são sistemas computacionais distribuídos compostos por unidades de processamento simples e densamente interconectadas. Estas unidades são conhecidas como neurónios artificiais que computam funções matemáticas. As unidades são dispostas em uma ou mais camadas e interligadas por um grande número de conexões (Faceli et al., 2011). Na maioria das arquiteturas, essas conexões, que simulam sinapses biológicas, possuem pesos associados, que ponderam a entrada recebida por cada neurónio da rede. Os pesos tem os seus valores ajustados no processo de aprendizagem e codificam o conhecimento adquirido pela rede (A. de P. Braga et al., 2000). O neurónio é a unidade de processamento fundamental de uma RNA (Haykin, 2009). O diagrama apresentado na Figura 3 mostra um modelo do neurónio artificial simplificado. Através da observação da Figura 3 é possível perceber que cada terminal de entrada do neurónio recebe um valor (pesos), os valores recebidos são ponderados e combinados por uma função matemática *fa*. A saída da função é a resposta do neurónio de entrada, e é definida pela aplicação de uma função de ativação. Esta função tem a tarefa de limitar a amplitude da resposta de um neurónio para um valor finito (Haykin, 2009).

Diagram

Description automatically generated

Figura 3 Funcionamento do neurónio artificial (Faceli et al., 2011)

Fonte: (Faceli et al., 2011) página 110

Numa RNA, os neurónios podem estar dispostos numa ou mais camadas. Quando são utilizadas duas ou mais camadas, um neurónio pode receber nos terminais de entrada valores de saída de neurónios da camada anterior e/ou enviar o valor de saída para terminais de entrada de neurónios da camada seguinte. Uma rede com mais de uma camada de neurónios recebe o nome de rede multicamada. Sendo que a camada que gera os valores é denominada de camada de saída, as restantes camadas são designadas por camadas intermediárias, escondidas ou ocultas. Na literatura tem sido propostos vários algoritmos para o ajuste de parâmetros de uma RNA. Este ajuste prende-se principalmente na definição dos valores dos pesos associados às conexões da rede que fazem com que o modelo obtenha o melhor desempenho. Diversos autores propuseram algoritmos de treino para RNA seguindo os paradigmas de aprendizagem supervisionada, não supervisionada e por reforço, pode-se destacar o algoritmo de correção de erro, hebbiano, competitivo e termodinâmico (Faceli et al., 2011).

## Aprendizagem Não-Supervisionada

Enquanto que nos modelos de aprendizagem supervisionada os algoritmos produzem uma previsão com base em informação recolhida previamente, na aprendizagem não-supervisionada são agrupadas amostras de acordo com uma métrica pré-especificada. Neste tipo de aprendizagem, os algoritmos pesquisam por padrões em conjuntos de dados que não possuem etiquetas (Libbrecht & Noble, 2015). A ausência de etiquetas, estas que representam o comportamento desejado para o modelo, significam a ausência de um ponto de referência sólido para avaliar a qualidade do modelo. Uma vez que não há nenhum resultado esperado, a categorização é feita de acordo com a classificação definida pelo algoritmo, este que procura padrões nos conjuntos de dados. O principal objetivo dos algoritmos de aprendizagem não-supervisionada é segmentar os dados em *clusters* ou grupos característicos. Esta segmentação é feita através da identificação padrões que se repetem entre os dados ou semelhanças, que permitem o agrupamento dos dados consoante as características apresentadas (Hurwitz & Kirsch, 2018). De seguida são apresentados dois algoritmos que exemplificam o funcionamento da aprendizagem não-supervisionada.

### *Clustering*

A análise de *clusters* é o estudo formal de algoritmos e métodos para agrupar, ou classificar, objetos. Um objeto é descrito ou por um conjunto de medidas ou por relações entre o objeto e outros objetos. A técnica de *clustering* não utiliza dados previamente etiquetados, o objetivo desta técnica é agrupar os dados de forma que os objetos pertencentes a cada *cluster* compartilhem determinada característica ou propriedade relevante para o domínio do problema em estudo (Jain & Dubes, 1988). Embora a ideia do que constitui um *cluster* seja intuitiva (agrupar objetos semelhantes) não existe uma definição única e precisa para este conceito, existindo uma grande variedade de definições na literatura. Cada definição de *cluster* resulta num critério de agrupamento que, de um modo genérico, é a forma de selecionar um modelo que melhor se adapte a determinado conjunto de dados (Estivill-Castro, 2002).

### *K-Means*

O objectivo-chave de um algoritmo *K-means* é organizar os dados em clusters de modo a que haja elevada semelhança intra-cluster e baixa semelhança inter-cluster. Um item apenas pertencerá a um agrupamento, não a vários, ou seja, gera um número específico e não hierárquico clusters.

O algoritmo K-means utiliza a estratégia de dividir e conquistar, e é um exemplo clássico de um algoritmo de maximização de expectativas. Os algoritmos são constituídos por duas etapas: a primeira etapa é conhecida como expectativa e é utilizada para encontrar o ponto esperado associado a um cluster; e a segunda etapa é conhecida como maximização e é utilizada para melhorar a estimativa do cluster, utilizando o conhecimento da primeira etapa. As duas etapas são processadas repetidamente até se alcançar a convergência. O algoritmo K-means pode ser concebido para a distância Euclidiana (equação 3.1)

|  |  |  |
| --- | --- | --- |
|  |  | (3.1) |

## Aprendizagem por Reforço

A aprendizagem por reforço é um modelo de aprendizagem comportamental. O algoritmo recebe *feedback* da análise dos dados para que o utilizador seja orientado para o melhor resultado. Este tipo de aprendizagem difere de outros tipos de aprendizagem porque o sistema não é treinado com o conjunto de dados da amostra. Em vez disso, o sistema aprende através de tentativa e erro. Portanto, uma sequência de decisões bem sucedidas resultará no "reforço" do processo porque, estas resolvem da melhor forma o problema em questão. Uma das aplicações mais comuns da aprendizagem de reforço é na robótica ou num jogo (Hurwitz & Kirsch, 2018). De um modo geral, o agente deve aprender a escolher ações que alteram o estado do seu ambiente e é utilizada uma função de recompensa cumulativa para definir a qualidade de uma dada sequência de ações (Tom M. Mitchell, 1997). Na Figura 4 é exemplificado o funcionamento da interação entre um agente e o ambiente que o rodeia. Sendo possível inferir que o agente encontra-se inserido num ambiente descrito por um conjunto de estados possíveis *E*, e pode realizar qualquer ação de um conjunto de ações possíveis, *a*. Cada vez que é realizada uma ação *a*, o agente recebe um valor de recompensa *r*, que indica o valor imediato desta transição de estado-ação.

Diagram

Description automatically generated

Figura 4 Interação entre um agente e o seu ambiente. Adaptado: (Tom M. Mitchell, 1997)

Isto produz uma sequência de estados Ei, ações ai, e recompensas imediatas ri, como mostra a Figura 4. A tarefa do agente é aprender com a recompensa indireta retardada, e escolher sequências de ações que produzam a maior recompensa cumulativa (Tom M. Mitchell, 1997). Ou seja, como resultado é esperada a aprendizagem de uma estratégia que determina as ações a serem tomadas na realização de uma tarefa, com uma probabilidade reduzida de ocorrência de falhas.

## Ferramentas disponíveis em Python para ML

A linguagem de programação *Python* é uma das linguagens mais populares para computação científica. Graças à sua natureza interativa de alto nível e ao ecossistema para o desenvolvimento de bibliotecas científicas, é uma escolha interessante para o desenvolvimento algorítmico e a análise exploratória de dados (Pedregosa et al., 2012). Atualmente existem um número elevado de bibliotecas de código-fonte aberto disponíveis para facilitar a construção de modelos de ML. A linguagem *Python* fornece não só, pacotes matemáticos e funcionalidades científicas essenciais para realizar o pré-processamento e transformação de dados, como também, pacotes que fornecem todos os algoritmos e funcionalidades de ML necessários que podem ser aplicados a um conjunto de dados para extrair os padrões.

*Scikit-learn* é uma biblioteca do *Python* que integra uma gama ampla de algoritmos de ML para a resolução de variados problemas. Fornece várias ferramentas para ajuste de modelo, pré-processamento de dados, seleção e avaliação dos modelos (Pedregosa et al., 2012).

A função *train\_test\_split* permite a separação aleatória dos dados em percentagens desejadas para treino e teste.

O *Scikit-learn* também tem implementadas uma grande variedade de métricas de avaliação de performance que se encontram disponíveis através do módulo *metrics*. Pode-se destacar a *accuracy\_score* que calcula o valor de acertos do modelo.

Para além desta, existem pelo menos quatro bibliotecas chave que são amplamente utilizadas para análise de dados: *NumPy*; *SciPy*; *Matplotlib* e *Pandas*. A biblioteca *NumPy* fornece um objeto de matriz multidimensional de alto desempenho, e ferramentas para trabalhar com estas matrizes. Já a biblioteca *Pandas* fornece estruturas de dados rápidas, flexíveis e expressivas, concebidas para trabalhar com dados "relacionais" ou "etiquetados". Foi desenvolvida por Wes McKinney em 2008 quando estava na AQR Capital Management por necessidade de uma ferramenta com alto desempenho e flexível para realizar análises quantitativas sobre dados financeiros. A biblioteca *SciPy* suplementa a biblioteca *NumPy* e tem uma variedade de módulos-chave de ciência e engenharia de alto nível.

# Delirium

O *delirium* é uma síndrome comum e potencialmente evitável em doentes mais idosos hospitalizados, estando associado a elevadas taxas de mortalidade, resultando assim em estadias hospitalares mais longas, elevados custos de saúde, declínio funcional, aumento da probabilidade de ocorrência de quedas, readmissões hospitalares, desenvolvimento de demência ou deficiência cognitiva a longo prazo. O *delirium* também pode causar eventos adversos após a hospitalização, incluindo limitações funcionais duradouras, declínio cognitivo persistente para além da perda de qualidade de vida para o paciente e para os prestadores de cuidados (Rossi Varallo et al., 2021). Embora em alguns casos, o *delirium* possa ser causado por um único medicamento ou doença subjacente, na maioria dos casos esta síndrome surge como o resultado da ação combinada de fatores predisponentes e precipitantes. É, portanto, uma condição multifatorial, que envolve a inter-relação entre a vulnerabilidade do paciente ao *delirium* no momento da admissão hospitalar e fatores precipitantes que podem surgir durante a hospitalização. Nesta perspetiva, os pacientes considerados vulneráveis (por exemplo, aqueles com demência ou uma doença subjacente grave) podem experimentar *delirium* devido à utilização de uma única dose de um fármaco com o intuito de ajudar na recuperação de determinada doença. Em contraste, os pacientes resistentes ao desenvolvimento do *delirium* podem apresentar esta condição após uma série de intervenções médicas, tais como anestesia geral, cirurgia, privação de sono, imobilização e uso de múltiplos fármacos psicoativos (Sharon K Inouye, 2000; Rossi Varallo et al., 2021).

## Prevalência

A população mundial está a envelhecer e todos os países do mundo estão a assistir a um crescimento no número e na proporção de indivíduos com idade igual ou superior a 65 anos (ONU, 2021). No entanto, à medida que a esperança média de vida aumenta, a prevalência de doenças crónicas também cresce, o que implica um incremento das necessidades de cuidados hospitalares quando comparados com os indivíduos mais jovens. O *delirium* é uma síndrome neuropsiquiátrica que se manifesta frequentemente em condições de doença aguda, e com maior prevalência entre os indivíduos mais velhos hospitalizados e pode variar em função das características dos indivíduos, do ambiente dos cuidados e da sensibilidade do método de deteção (American Psychiatric Association, 2013; Pérez-Ros & Martínez-Arnau, 2019).

﻿Em Portugal, não se conhece a prevalência exata de *delirium*, porém presume-se que se assemelhe à dos dados existentes internacionalmente (Prayce, Quaresma, & Neto, 2018). A prevalência de *delirium* pode divergir consoante o ambiente onde se encontra o paciente. Esta síndrome pode ocorrer em residentes em lares, doentes em fim de vida, doentes hospitalizados, doentes pós-cirúrgicos, doentes admitidos na Unidade de Cuidados Intensivos (UCI) ou no Serviço de Urgência (SU). No caso de pacientes em cuidados continuados, a prevalência relatada varia entre 15% e 70%, embora a deteção de *delirium* hiperativo seja mais frequente e os sintomas mais reportados são: desorientação, irritabilidade, agitação psicomotora e alucinações visuais (Hosker & Ward, 2017; Pérez-Ros & Martínez-Arnau, 2019). Como o *delirium* hipoativo é dominado por sintomas de sonolência e inatividade, pode passar despercebido, e o facto de algumas pessoas experimentarem uma mistura de ambos os subtipos pode afetar o diagnóstico (Desforges & Lipowski, 1989).

Segundo os dados apresentados na revisão bibliográfica de Pérez-Ros & Martínez-Arnau (2019), a prevalência do *delirium* em doentes em fim de vida aproxima-se dos 85% em ambientes de cuidados paliativos. Já em doentes hospitalizados mais idosos, o *delirium* é uma complicação bastante frequente, com uma prevalência aproximada de 27%. Em doentes pós-cirúrgicos, a síndrome está condicionada ao tipo de cirurgia, à anestesia e aos procedimentos utilizados. Por exemplo, otorrinolaringologia e cirurgia geral apresentam menor risco, com uma prevalência de 12% e 13%, respetivamente, enquanto a prevalência de *delirium* em pacientes submetidos a cirurgia aórtica, abdominal, ou cardíaca é mais elevada, com valores de 29%, 50%, e 51%, respetivamente. Na UCI, a prevalência varia entre 31% e 35%, sendo mais elevada em doentes submetidos a ventilação mecânica e com comorbilidades podendo chegar até aos 80%. A maioria dos estudos, tanto de idosos hospitalizados como institucionalizados, mostra uma forte associação entre demência e *delirium*, tendo os dados demonstrado uma relação diretamente proporcional ao grau de demência. Havendo a probabilidade de o *delirium* aumentar 45% no caso de pacientes com demências moderadas e 58% em demências graves (Pérez-Ros & Martínez-Arnau, 2019).

A prevalência de *delirium* na comunidade em geral é baixa (1% a 2%) mas aumenta com a idade, subindo para 14% entre os indivíduos com mais de 85 anos de idade. Já nos indivíduos mais idosos, que se encontrem no SU, a prevalência pode variar entre os 10% e 30% (American Psychiatric Association, 2013). O *delirium* em contexto de SU é um problema altamente prevalecente que normalmente passa despercebido pelos prestadores de cuidados de saúde. Além disso, o início de um episódio de *delirium* em SU pode estar associado a um aumento da morbilidade e mortalidade (Pérez-Ros & Martínez-Arnau, 2019). A ocorrência de novos casos de *delirium* durante a hospitalização de pessoas mais velhas varia entre 6-56%; enquanto a prevalência de *delirium* no momento da admissão de pessoas mais velhas varia entre 14-24%. Além disso, com o avanço da idade da população, esta síndrome constitui uma preocupação de saúde pública, uma vez que tende a aumentar no futuro (Sharon K Inouye, 2000; Rossi Varallo et al., 2021).

Por norma, a maioria dos indivíduos com *delirium* tem uma recuperação completa, pelo que o reconhecimento precoce e uma intervenção adequada, geralmente encurta a duração do *delirium*. Esta síndrome pode progredir para o estupor, coma, convulsões ou morte, particularmente se a causa subjacente permanecer sem tratamento (Harrison et al., 2018). A mortalidade entre os indivíduos hospitalizados com *delirium* é elevada, e até 40% dos indivíduos com *delirium*, particularmente os que sofrem de doenças malignas ou tenham outras condições médicas subjacentes significativas, morrem no prazo de um ano após o diagnóstico (American Psychiatric Association, 2013).

## Fatores de risco

Na literatura estão descritos diversos fatores de risco para o desenvolvimento de *delirium*, estando a maioria relacionados com o processo de admissão hospitalar. Além disso, está também documentado que o desenvolvimento de *delirium* em SU ou UCI é um fator preditor do aumento do tempo de internamento (Cano-escalera, Besga, & Graña, 2021; J. H. Han et al., 2011). Segundo vários estudos realizados é possível afirmar que a incidência do *delirium* na UCI pode variar desde valores baixos a muito altos dependendo das diferentes populações de doentes. Pode-se realçar que o *delirium* pós-operatório tem sido a complicação mais comum em pacientes mais velhos que foram submetidos a cirurgia (Ali et al., 2021; Robinson, Raeburn, Tran, Brenner, & Moss, 2011).

Os fatores de risco assumem diversas relevâncias dependendo das características individuais de cada doente, bem como do seu contexto patológico. A causa do *delirium* é quase sempre de origem multifatorial, dependente de determinados fatores (Cano-escalera et al., 2021). Estes fatores podem ser divididos em predisponentes, estando relacionados com o estado basal do doente e respetivas comorbilidades, e precipitantes, referentes ao contexto hospitalar do doente, nomeadamente doença aguda e respetivo tratamento. Alguns dos fatores predisponentes não são modificáveis, como por exemplo, idade, sexo, dependência, deficiência cognitiva pré-existente, doenças cardíacas e pulmonares pré-existentes. Relativamente aos fatores modificáveis, estes podem ter origem nas condições ambientais do local onde se encontra o paciente, tais como ausência de luz do dia visível, a falta de noção temporal, impossibilidade de visitas (M. A. Ali et al., 2021; Cano-escalera et al., 2021). Ou por alteração da doença como por exemplo, sedação, aumento do tempo de internamento, febre, dor, entubação e cateteres (M. A. Ali et al., 2021). Estudos têm demonstrado que a proporção de pacientes que desenvolve *delirium* é diretamente proporcional ao número de fatores de risco presentes à admissão (Mittal et al., 2011). Além disso, foi evidenciada, por diversos estudos, a relação entre *delirium* e a mortalidade durante e após o internamento, em qualquer tipologia (S. K. Inouye et al., 1990; Robinson, Raeburn, Tran, Brenner, & Moss, 2011; Siddiqi, House, & Holmes, 2006; Witlox et al., 2010).

### Fatores predisponentes

O *delirium* partilha muitos dos seus fatores predisponentes com outras síndromes geriátricas, como as quedas e a incontinência. Isto mostra que os fatores predisponentes do *delirium* indicam a vulnerabilidade dos doentes geriátricos (Sharon K. Inouye et al., 2014; Meagher, MacLullich, & Laurila, 2008).

Os principais fatores predisponentes para *delirium*, identificados nas populações de doentes, incluem idade avançada, défice cognitivo, como atraso no desenvolvimento ou demência (Gross et al., 2012), fragilidade, comorbilidades (incluindo doença cardiovascular e renal), depressão ou outra doença psiquiátrica (K. Wilson, Broadhurst, Diver, Jackson, & Mottram, 2005), deficiência visual e auditiva (Sharon K. Inouye et al., 2014; Smith et al., 2017), consumo de álcool e estado nutricional deficiente (Velayati, Vahdat Shariatpanahi, Shahbazi, & Vahdat Shariatpanahi, 2019). Sabe-se que a demência é o fator predisponente mais prevalente no momento da hospitalização, uma vez que é capaz de aumentar em duas a cinco vezes a possibilidade de desenvolver *delirium*, no entanto, qualquer doença crónica pode predispor o aparecimento do *delirium.*

Além disso, estudos de neuroimagem indicam que o risco de *delirium* pode ser superior em indivíduos com atrofia cerebral e/ou doença da substância branca (Nitchingham, Kumar, Shenkin, Ferguson, & Caplan, 2018). Bem como em doentes acima dos 65 anos que estejam em situação de pós-operatório (Hatano et al., 2013).

### Fatores precipitantes

Os fatores precipitantes do *delirium* abrangem uma gama ampla de diferentes tipos de condições médicas em doentes internados (Nagari & Suresh Babu, 2019) incluindo, entre outros, doenças médicas agudas (como sepse, hipoglicemia, acidente vascular cerebral e insuficiência hepática), trauma (como fraturas ou traumatismo craniano), cirurgia, desidratação, problemas psicológicos, stress (Sharon K. Inouye et al., 2014; J. E. Wilson et al., 2020). Para além destes, assinalam-se o uso de medicação psicoativa, imobilização física, algaliação ou alterações eletrolíticas (Sharon K. Inouye & Charpentier, 1996; Sharon K. Inouye et al., 2014; Lawlor et al., 2002). Sendo que, de entre os fatores precipitantes mencionados, o uso de medicação é considerado um fator extremamente habitual durante a hospitalização, originando até 40% dos casos. E, consequentemente, a ocorrência de *delirium* aumenta em proporção direta com o número de medicamentos utilizados, devido à maior probabilidade de ocorrência de efeitos secundários para além das interações medicamentosas (Lôbo, Silva Filho, Lima, Ferriolli, & Moriguti, 2010).

O *delirium* está descrito como uma síndrome heterógenia, como tal, poderá estar presente mais do que um fator precipitante nos pacientes (Cirbus et al., 2019). É importante ressaltar que benzodiazepínicos, diidropiridínicos (bloqueadores dos canais de cálcio do tipo L normalmente usados ​​no tratamento da hipertensão), anti-histamínicos e opióides podem representar um incremento no risco de *delirium* (Clegg & Young, 2011).

Segundo um estudo observacional realizado na Índia e publicado em 2019 verificou que dos 1582 pacientes da UCI a percentagem de doentes que desenvolveram *delirium*  nas primeiras 72 horas de admissão foi de 25,7% (406/1582), 52% dos doentes apresentaram o subtipo hipoativo e 48% deles tinham *delirium* hiperativo. Com este estudo concluíram que o principal fator de risco que contribuiu para o *delirium* foi o consumo de álcool. No que diz respeito aos fatores precipitantes, os mais comuns que resultaram em *delirium* estavam relacionados com os parâmetros metabólicos anormais (Nagari & Suresh Babu, 2019).

Na Tabela 1 são apresentados os fatores precipitantes de *delirium* que foram considerados como as causas mais prováveis desta doença de acordo com vários autores (Sharon K. Inouye et al., 2014; Laurila, Laakkonen, Strandberg, & Tilvis, 2008; Nagari & Suresh Babu, 2019). Os fatores de precipitação do *delirium* foram classificados em toxinas, parâmetros metabólicos anormais, infeções, causas do sistema nervoso central e parâmetros metabólicos. Segundo o estudo realizado por Nagari, em 2019, de entre os parâmetros metabólicos, a uremia, a encefalopatia hepática e a hiponatremia contribuíram para a maioria dos casos de *delirium*. Nos doentes internados, para além dos resultados metabólicos anormais, vale a pena mencionar outros fatores importantes, tais como a polifarmácia (utilização de mais do que cinco drogas), o uso de drogas psicoativas e especialmente as restrições físicas a que os doentes estão sujeitos (Sharon K. Inouye et al., 2014).

Tabela 1 - Tabela resumo dos fatores considerados precipitantes para o delirium (Sharon K. Inouye et al., 2014; Laurila et al., 2008; Nagari & Suresh Babu, 2019)

|  |  |
| --- | --- |
| **Fatores Precipitantes** | |
| Medicamentos | Abstinência de álcool |
| Tramadol |
| Cortisona |
| Medicamentos para Parkinson |
| Medicamentos com propriedades anticolinérgicas |
| Condições Metabólicas | Hiponatremia (redução da concentração plasmática de sódio) |
| Hiperglicemia (nível de glicose no sangue muito elevado) |
| Hipoglicemia (nível de glicose no sangue muito baixo) |
| Hipercarbia (aumento da pressão parcial de CO2 (PCO2) no sangue) |
| Uremia (aumento da ureia no sangue) |
| Encefalopatia hepática (hiperamonemia) |
| Procedimentos médicos/cirurgias | Cirurgia aórtica, abdominal, cardíaca |
| Infeções | Causas infeciosas sistémicas |
| Meningite/Encefalite |
| Urinária |
| Respiratória |
| Causas do SNC | Estados de hipoperfusão |
| Encefalopatia hipertensiva |
| Acidente vascular cerebral (AVC) |
| Lesão de ocupação do espaço intracraniano (ICSOL) |
| Convulsões |
| Doença psiquiátrica |
| Vícios | Consumo de álcool |
| Consumo de drogas |

Segundo Francis (1996) a medicação é responsável em por cerca de 30% dos casos de *delirium*, estando implicado não só o tipo, mas também o número total de fármacos (Francis, 1996). Apesar de, na sua maioria, poderem desencadear um episódio de *delirium*, algumas classes de medicamentos apresentam um maior risco, possivelmente por atuarem ao nível das vias patofisiológicas e dos neuromediadores, como os anticolinérgicos ou os antidopaminérgicos (Nagari & Suresh Babu, 2019). Assim, destacam-se os: antipsicóticos, ansiolíticos, antidepressivos, opióides, corticosteroides, anticonvulsivantes e anti-histamínicos (Clegg & Young, 2011; Gaudreau, Gagnon, Harel, Roy, & Tremblay, 2005; Sharon K. Inouye et al., 2014).

## Fisiopatologia

O mecanismo fisiopatológico do *delirium* ainda não é bem conhecido, no entanto, pode ser explicado pelos mecanismos que comprometem a neurobiologia da atenção, a função cortical e subcortical e a neurotransmissão (Flacker & Lipsitz, 1999; Sharon K. Inouye, 2006; Maldonado, 2017; Mittal et al., 2011). A existência de um amplo espetro de problemas clínicos quer em idosos ou não pode explicar o desenvolvimento de *delirium* nesta faixa etária (van der Mast, 1998). A disfunção neuronal, tem na sua origem fatores de risco, tais como: idade, nível de função cognitiva, doença grave e mau estado funcional, distúrbios metabólicos, e deficiências visuais e auditivas. Sabe-se que o processo de envelhecimento tem associadas alterações cerebrais, nomeadamente, alteração na proporção de neurotransmissores que regulam o stress, redução do fluxo sanguíneo cerebral, diminuição da densidade vascular, perda de neurónios e alteração ao nível dos sistemas de transdução do sinal intracelular. Por estes motivos, as pessoas idosas são mais suscetíveis a desenvolverem *delirium* do que indivíduos jovens (Sharon K. Inouye & Charpentier, 1996; Martins & Fernandes, 2012).

De entre as principais hipóteses que visam explicar os mecanismos envolvidos na fisiopatologia de *delirium*, destacam-se, as anomalias na síntese, libertação e inativação de neurotransmissores e a hipótese inflamatória.

Começando pela hipótese dos neurotransmissores, antes de mais é importante realçar que o correto funcionamento do sistema nervoso depende de uma comunicação adequada entre os neurónios. De um modo geral, este tipo de comunicação envolve sinais químicos que passam dos axónios para as dendrites que são transformados em sinais eléctricos. Nos locais da célula responsáveis por receber os sinais, as dendrites estabelecem contatos com os axónios de outras células, cada um destes contatos está separado pela a fenda sináptica, para onde são libertados os neurotransmissores. Estes transmissores químicos são armazenados nos terminais dos axónios, em pequenos sacos esféricos designados vesículas sinápticas que serão libertados para as junções sinápticas quando induzidas pela chegada de um potencial de ação (R. Webster, 2001). Isto significa que os neurotransmissores atuam como mediadores químicos na comunicação intercelular através da ativação de recetores específicos e mensageiros secundários nas células pós-sinápticas. E, embora exista uma da grande variedade de neurotransmissores, estes podem classificar-se em monoaminas (como a acetilcolina e serotonina); em catecolaminas (como a dopamina, adrenalina e noradrenalina); e em aminoácidos (como o GABA).

Como já referido anteriormente, o processo de envelhecimento está associado não só a uma diminuição do volume de produção de acetilcolina, como também a perturbações na atividade do sistema colinérgico de vários neurotransmissores, pois este sistema torna-se progressivamente deficitário, especialmente no caso da dopamina, da noradrenalina, da serotonina, da acetilcolina, e do GABA (Maldonado, 2017). Já há algum tempo que foi reconhecido que um défice no sistema colinérgico poderia estar envolvido na fisiopatologia do *delirium* (Zbigniew J. Lipowski, 1991). Na verdade, a acetilcolina participa em vários processos neurofisiológicos cruciais a nível da consciência, atenção e função cognitiva. Este facto, implica que qualquer agressão que perturbe a cadeia oxidativa, como inflamação ou hipoxia, possa afetar a quantidade de acetilcolina disponível no cérebro e assim prejudicar os processos cognitivos (Z. J. Lipowski, 1987; Zbigniew J. Lipowski, 1991; Meagher et al., 2008). Sabe-se que a síntese de acetilcolina envolve vários precursores, enzimas e recetores e qualquer disfunção nestes componentes leva a excogitar que a deficiência colinérgica poderá estar diretamente relacionada com o défice cognitivo observado nos pacientes com *delirium* (Hshieh, Fong, Marcantonio, & Inouye, 2008)*.* Já vários estudos demonstraram que quando são documentados níveis baixos de acetilcolina no plasma e líquido cefalorraquidiano a probabilidade de pertencer a um doente delirante aumenta. Por outro lado, níveis elevados de atividade anticolinérgica sérica têm sido associados a uma maior probabilidade de *delirium* entre doentes a receber medicação, bem como no *delirium* pós-operatório. Estes dados coincidem com a relação clínica observada entre o potencial anticolinérgico de um medicamento e os seus efeitos deliriogénicos (Maldonado, 2017).

Um outro neurotransmissor relacionado com a fisiopatologia do *delirium* é a dopamina. Este neurotransmissor desempenha um papel importante na atividade motora e funções cognitivas como a atenção, o pensamento e a perceção, sendo também muito relevante no desenvolvimento de sintomas psicóticos (van der Mast, 1998). A atividade dopaminérgica em excesso tem sido apontada como fator contribuinte para a síndrome do *delirium*. Postula-se que um aumento da dopamina extracelular possa fomentar a morte celular e causar sintomas psicóticos do *delirium*, talvez devido à sua influência reguladora na libertação de acetilcolina, que resulta geralmente numa diminuição dos níveis de acetilcolina. Este aumento da dopamina causa geralmente o aparecimento dos sintomas tais como aumento da atividade psicomotora, irritabilidade, agitação, perturbação, agressividade e psicose (Maldonado, 2008). De modo a mitigar os efeitos causados pela desregulação dos níveis de dopamina no organismo, utilizam-se bloqueadores da dopamina para o tratamento do *delirium* até que as causas subjacentes sejam melhoradas, uma vez que ajudam ao equilíbrio temporário das atividades colinérgicas e dopaminérgicas (Mittal et al., 2011). Os antagonistas da dopamina (por exemplo, agentes antipsicóticos) diminuem eficazmente os sintomas do *delirium* anticolinérgico, porém os medicamentos dopaminérgicos (por exemplo, levodopa) são precipitantes reconhecidos do *delirium* (Trzepacz, 1999).

Ainda na temática da hipótese dos neurotransmissores, a serotonina também parece estar envolvida no decurso de comportamentos característicos do *delirium*. A serotonina é um neurotransmissor muito abundante no cérebro, e desempenha um papel importante em várias funções cerebrais, tais como a cognição, sono e atividade motora. A síntese e libertação normal de serotonina depende, entre outros, da disponibilidade do triptofano (Maldonado, 2008), um aminoácido essencial obtido através da alimentação. A ocorrência de desequilíbrios nos níveis de serotonina têm sido associados a uma variedade de distúrbios psiquiátricos, incluindo depressão, ansiedade, distúrbios no humor, psicoses. Na verdade, tanto o aumento como a diminuição da atividade serotonérgica têm sido associados ao *delirium*. Mais especificamente, o aumento dos níveis de serotonina tem sido associados a indivíduos com encefalopatia hepática e em doentes que sofrem de *delirium* hipoativo. Já a medição de níveis baixos de serotonina, como ocorre na hipoxia, podem estar associados ao *delirium* hiperativo (Maldonado, 2008).

O sistema colinérgico é equilibrado pela atividade das monoaminas como a dopamina, noradrenalina e serotonina, estas que desempenham papéis importantes a nível motor e no ciclo sono-vigília, mediando as respostas fisiológicas aos estímulos. Estas respostas são moduladas pela via colinérgica, o que leva a concluir que o desenvolvimento do *delirium* muito provavelmente estará relacionado com a interação entre a via colinérgica e estas monoaminas (Hshieh et al., 2008). Assim, é de esperar que fármacos anticolinérgicos constituam fatores de risco, bem como outros fármacos que tenham também capacidade de ligação ao recetores muscarínicos (MacLullich et al., 2008). ﻿

A inflamação sistémica é uma característica associada a determinadas condições médicas e cirúrgicas que por vezes está também relacionada com o *delirium*, particularmente quando está envolvida a destruição e/ou infeção de tecidos (Cerejeira, Firmino, Vaz-Serra, & Mukaetova-Ladinska, 2010). Esta condição está muitas vezes ligada a ocorrências conhecidas por precipitar disfunções cognitivas, principalmente na população idosa, que inclui infeções, cancro, e cirurgia. O *delirium* é um exemplo de declínio agudo na cognição, igualmente associado a estados inflamatórios agudos (Simone & Tan, 2011). Têm sido associadas várias condições relacionadas com o *delirium* que se caracterizam pela resposta inflamatória com libertação aguda de mediadores inflamatórios na corrente sanguínea (Cerejeira et al., 2010). As células do sistema imunitário, como os neutrófilos e os macrófagos, comunicam através de citocinas, que são proteínas semelhantes a hormonas produzidas por células especializadas do sistema imunitário. As citocinas são importantes mediadores da resposta imunitária, que iniciam, perpetuam, ou desregulam a resposta, e podem ser classificadas em interleucinas (IL), fatores de necrose tumoral, e fatores de crescimento transformadores. Algumas citocinas são consideradas pró-inflamatórias como a IL-1, IL-6, IL-8, TNF-α e PCR, e outras são anti-inflamatórias, como é o caso da IL-4, IL-10, e IL-13 (Simone & Tan, 2011).

Há provas convincentes de que a estimulação inflamatória induz a ativação de células parenquimatosas do cérebro, expressão de citocinas pró-inflamatórias e mediadores inflamatórios no sistema nervoso central. Estas alterações neuroinflamatórias induzem disfunções neuronais e sinápticas e subsequentes sintomas neurocomportamentais e cognitivos (Cerejeira et al., 2010; Simone & Tan, 2011). Acredita-se que a desregulação das citocinas seja o principal incitador da neurodegeneração e do subsequente comprometimento cognitivo no *delirium* que resulta da ativação de uma resposta inflamatória sistémica (Simone & Tan, 2011). De modo a compreender melhor o papel que a inflamação desempenha no *delirium*, os investigadores estudaram e mediram os níveis de marcadores inflamatórios em pacientes doentes. E, em 2007 foi publicado um estudo que visava examinar os padrões de expressão das citocinas pró e anti-inflamatórias em doentes idosos com e sem *delirium*, admitidos no hospital. E verificaram que em doentes com *delirium*, os níveis de IL-6 (53% vs. 31%) e IL-8 (45% vs. 22%) estavam acima do limite de deteção, em comparação com os doentes sem *delirium*. O que permitiu concluir que as citocinas pró-inflamatórias poderiam contribuir para a patogénese do *delirium* em doentes idosos hospitalizados (de Rooij, van Munster, Korevaar, & Levi, 2007). Curiosamente, também um outro estudo prospectivo de doentes idosos admitidos para a reparação cirúrgica aguda da fratura da anca, realizado em 2008, concluiu que os níveis periféricos de IL-6 e IL-8 foram igualmente elevados em doentes que desenvolveram *delirium*. Para além disso, também foi verificado que a IL-6 foi especificamente associada à forma hiperativa de *delirium*, enquanto a IL-8 foi mais elevada nos dias que antecederam o início de *delirium* (Van Munster et al., 2008). Para além das citocinas, também a proteína C-reativa (PCR), que tem sido comumente utilizada como marcador de infeção, inflamação e lesão tecidual, pode estar implicada no desenvolvimento do *delirium* (De Morton et al., 2007). E para comprovar isso mesmo, em 2014 foi desenvolvido um estudo onde foi avaliada a relação entre a PCR e o *delirium*. Que, por conseguinte, concluiu a existência de uma associação entre a PCR elevada e o *delirium,* principalmente em doentes admitidos com doença músculo-esquelética mas não noutros grupos de diagnóstico. O que leva a excogitar que a PCR está envolvida na génese do *delirium* no caso de doença músculo-esquelética, mas que outros fatores ou processos podem ser mais importantes nos doentes com patologias noutros grupos de diagnóstico (Ritchie, Newman, Leurent, & Sampson, 2014).

## Instrumentos de rastreio e diagnóstico de *delirium*

A capacidade para avaliar o *delirium* é uma componente essencial na estratégia de avaliação do doente de modo a prevenir ou tratar a ocorrência desta perturbação. Estudos apontam que o diagnóstico precoce e uma abordagem adequada, estão associados a uma redução das taxas de morbilidade e mortalidade associadas ao *delirium* (Sharon K. Inouye et al., 2014; Mittal et al., 2011). Trata-se de um diagnóstico clínico realizado através de um histórico clínico e exames, complementados com uma avaliação cognitiva formal perante a suspeita de alteração cognitiva, e em caso positivo, a confirmação do diagnóstico de *delirium* através de um instrumento de diagnóstico validado (Bourgeois et al., 2014). Como o *delirium* pode passar facilmente despercebido aos profissionais de saúde, especialmente em doentes internados em UCI e SU, torna-se importante o uso de ferramentas de rastreio de forma a detetar mais precoce e facilmente este distúrbio. Não só, pela possível melhoria na qualidade de vida dos pacientes, mas também pela possível contenção de custos relacionados com o tratamento dos doentes. Face a esta necessidade, têm vindo a ser desenvolvidas e validadas ferramentas de rastreio para esta perturbação, com o intuito de serem usadas na prática clínica diária (De & Wand, 2015). Atualmente existem vários instrumentos validados de apoio ao diagnóstico do *delirium*, que foram adequados consoantes a tipologia de doentes envolvidos (Leonard et al., 2014). Assim, já existem mais de 30 instrumentos desenvolvidos e testados para a avaliação do *delirium* (Adamis, Sharma, Whelan, & MacDonald, 2010; C. L. Wong, Holroyd-Leduc, Simel, & Straus, 2010). Na Tabela 2 estão reunidos alguns dos instrumentos usados mundialmente no rastreio e avaliação de *delirium* (De & Wand, 2015).

Tabela 2 - Lista das ferramentas para diagnóstico de delirium (De & Wand, 2015)

|  |
| --- |
| **Ferramentas para deteção do *delirium*** |
| Confusion Assessment Method (CAM) |
| Memorial Delirium Assessment Scale (MDAS) |
| Confusion Assessment Method for the Intensive Care Unit (CAM-ICU) |
| Delirium Rating Scale (DRS) |
| Delirium Rating Scale, Revised (DRS-R-98) |
| Nursing Delirium Screening Checklist (NuDESC) |
| Delirium Detection Score (DDS) |
| Delirium Observation Screening Scale (DOSS) |
| Digit Span Test (DST) |
| Single Question in Delirium (SQiD) |
| Delirium Symptom Interview (DSI) |
| NEECHAM Cofusion Scale |
| Brief CAM (bCAM) |
| Clinical Assessment of Confusion (CAC) |
| Delirium Diagnostic Tool-provisional (DDT-Pro) |
| *Richmond Agitation Sedation Scale* (RASS) |
| Delirium triage screen (DTS) |
| Intensive Care Delirium Screening Checklist (ICDSC) |
| Modified Richmond Agitation Sedation Scale (mRASS) |
| Simple Question for Easy Evaluation of Consciousness (SQUEEC) |
| Short Portable Mental Status Questionnaire (SPMSQ) |
| The 4As Test (4AT) |
| Vigilance A Test |

De seguida, serão explicadas as duas ferramentas de diagnóstico mais relevantes para este projeto. A RASS, pois foi a escala utilizada para auxiliar no diagnóstico de *delirium* dos pacientes que se encontram no conjunto de dados deste projeto. E o CAM, pois segundo uma revisão sistemática publicada em 2016, dos instrumentos de diagnóstico identificados em contexto de SU, o método de avaliação de confusão para a unidade de cuidados intensivos (CAM-ICU) foi o mais utilizado (Mariz et al., 2016).

### *Richmond Agitation Sedation Scale* (RASS)

A escala RASS *(Richmond Agitation Sedation Scale)* é uma das escalas de avaliação mais completas, pois inclui medidas de avaliação tanto para agitação como para sedação. Tal como descrito na Tabela 3, é uma escala que varia de +4 a -5, na qual, a pontuação de RASS de 0 indica um doente calmo e desperto, as pontuações positivas (0 a +4) correspondem aos níveis de agitação enquanto as negativas (-1 a -5) representam os níveis de sedação. No caso de os pacientes não estarem espontaneamente em alerta, serão então chamados pelo nome, para estabelecerem contacto visual com o avaliador. A duração do contacto visual é medida e será atribuída uma classificação entre -1 e -3 consoante a resposta do doente. Se os pacientes não respondessem à estimulação verbal, podem então ser estimulados fisicamente (ou seja, toque do ombro e/ou fricção esternal) e pontuados de acordo com a sua resposta como -4 ou -5 (Ely et al., 2003).

Tabela 3 - Escala RASS (Ely et al., 2003)

|  |  |  |
| --- | --- | --- |
| **Pontuação** | **Classificação** | **Descrição** |
| +4 | Combativo | Combativo, violento, risco para a equipa |
| +3 | Muito agitado | Conduta agressiva, puxa ou remove tubos ou cateteres, agressivo verbalmente |
| +2 | Agitado | Movimentos despropositados frequentes, vigorosos ou agressivos |
| +1 | Inquieto | Intranquilo, ansioso, sem movimentos vigorosos ou agressivos |
| 0 | Alerta e calmo | Alerta, calmo |
| -1 | Sonolento | Adormecido, facilmente despertável, mantém contacto visual por mais de 10 segundos |
| -2 | Sedação Leve | Despertar precoce ao estimulo verbal, mantem contacto visual por menos de 10 segundos |
| -3 | Sedação Moderada | Movimentos e abertura ocular ao estímulo verbal, mas sem contacto visual. |
| -4 | Sedação Intensa | Sem resposta ao estímulo verbal, mas apresenta movimentos ou abertura ocular ao toque (estimulo físico) |
| -5 | Não desperta | Sem resposta a estímulo verbal ou físico |

### *Confusion Acessment Method* (CAM)

O Método de Avaliação da Confusão (CAM) constitui uma ferramenta de diagnóstico de *delirium* publicada em 1990, originalmente desenvolvida a partir de uma revisão da literatura com o consenso de especialistas e validada com base nos critérios da 3ª edição do DSM (S. K. Inouye et al., 1990). Este tem sido o instrumento mais amplamente utilizado para identificar *delirium*, no entanto, é necessária formação específica para assegurar bons resultados (De & Wand, 2015).

Esta ferramenta foi desenhada com o objetivo de melhorar a identificação de *delirium*, permitindo a realização de um diagnóstico mais rápido e eficaz através de uma breve avaliação cognitiva. Desta forma, permitiu a redução da morbilidade e mortalidade desta condição devastadora, particularmente em pacientes idosos hospitalizados de alto risco tratados em serviços médicos e cirúrgicos gerais. Além disso, a CAM apresentou-se não só como uma ferramenta de diagnóstico padronizada como também como um meio de sistematização e registo de observações clínicas (S. K. Inouye et al., 1990).

O questionário resultante (ver Anexo I – The confusion assessment method instrument) baseou-se em observações específicas relevantes para cada uma das nove características do *delirium*. As características clínicas identificadas para o desenvolvimento do CAM, estão incluídas nos critérios de diagnóstico DSM-III, e foram as seguintes: início agudo e curso flutuante, desatenção, pensamento desorganizado, alteração do nível de consciência, desorientação, perturbações da memória, perturbações de perceção, aumento ou diminuição da atividade psicomotora, e perturbação do ciclo sono-vigília (S. K. Inouye et al., 1990). O CAM já foi utilizado em mais de 4000 estudos publicados e traduzido em pelo menos 12 línguas, estando também adaptado para UCI (CAM-UCI) e departamentos de emergência (Sharon K. Inouye et al., 2014). E , apesar da possibilidade do CAM de identificar os presença ou ausência de *delirium* de uma forma rápida e fácil, não avalia a gravidade da condição (Mariz et al., 2016).

### *Confusion Assessment Method for the Intensive Care Unit* (CAM-ICU)

O CAM-ICU foi adaptado do CAM para avaliar doentes adultos críticos para o *delirium* (Ely et al., 2001). Embora o CAM-ICU seja um algoritmo que se baseia na presença de quatro elementos característicos do *delirium*: início súbito, flutuação dos sintomas, inatenção e pensamento desorganizado ou alteração da consciência (S. K. Inouye et al., 1990) (ver Anexo II – The Confusion Assessment Method (CAM) Diagnostic Algorithm\*), foi validado utilizando os critérios da quarta edição do DSM.

O CAM-­ICU permite identificar o *delirium* em doentes críticos, principalmente doentes em ventilação mecânica. Utiliza métodos de avaliação não-­verbal para avaliar as características importantes de *delirium*. A maioria dos questionários CAM-ICU são rapidamente realizados, não demorando geralmente mais do que alguns minutos. Este é o instrumento de diagnóstico melhor estudado e mais amplamente utilizado (Ely et al., 2001; S. K. Inouye et al., 1990; Nagari & Suresh Babu, 2019) Estudos indicam que o CAM e CAM-ICU são os dois melhores instrumentos diagnósticos de *delirium* atualmente disponíveis (Mittal et al., 2011).

# Preparação e análise exploratória dos dados

A análise das características presentes em conjuntos de dados permite a descoberta, não só de padrões e tendências que podem fornecer informações valiosas, como também podem facilitar a compreensão do processo que gerou determinado resultado. O propósito desta análise é extrair, contextualizar e organizar toda a informação possível, para assim gerar novas hipóteses ou modelos que permitam a resolução do problema (Kuhn & Johnson, 2020). Para além disso, este estudo possibilita a definição das possíveis técnicas a serem utilizadas e os procedimentos a serem adotados. Neste sentido, no presente capítulo serão descritas as principais características de cada variável, assim como a sua relação, previamente documentada em artigos e livros científicos, com a síndrome do *delirium*. Para além disso, esta informação será completada com o apoio de tabelas, para ajudar a organizar a informação, e representações gráficas com vista a facilitar a perceção da variabilidade no conjunto de dados observados.

Uma das etapas mais importantes, para o desenvolvimento de modelos preditivos é o tratamento da base de dados. Este processo é importante na medida em que permite que não sejam incluídos valores errados, valores omissos e variáveis inócuas, valores estes que podem comprometer a precisão das previsões (Rahm & Do, 2000).

A base de dados inicial para este projeto, apresentava-se com informação de 511 indivíduos e 124 variáveis, com informação relativa a indivíduos que passaram pelas Urgências de um Hospital português entre 2014 e 2016. Neste conjunto de dados encontravam-se informações como: idade, tempo de permanência na unidade de urgência, análises clinicas (glicose, ureia, creatinina, PCR), gasometria (pH sanguíneo, pressão parcial de oxigénio (PO2), pressão parcial de dióxido de carbono (PCO2), e ião bicarbonato (HCO3)), a medicação habitual dos pacientes, esta selecionada de acordo com os resultados da evidência científica para o desenvolvimento do *delirium*, entre outras variáveis.

## Preparação e recolha de dados

Num projeto de modelação preditiva, os algoritmos de ML aprendem um mapeamento que permite prever, através das variáveis de entrada, uma resposta. Neste tipo de modelação estão envolvidos os chamados dados estruturados, ou seja, dados tal como aparecem numa folha de cálculo ou numa matriz, com linhas de exemplos e colunas de características para cada exemplo. É também bastante comum, que nestes conjuntos de dados, as variáveis apresentem diferentes tipos de dados. Algumas variáveis podem ser numéricas, tais como números inteiros, decimais, posições ou percentagens. Já outras podem ser categóricas como nomes, categorias ou siglas, ou ainda binárias, representadas com 0 e 1 ou Verdadeiro e Falso. O problema é que os algoritmos de ML no seu núcleo funcionam com base em dados numéricos. Por esse motivo, é necessária a transformação dos dados de modo a satisfazer este requisito dos algoritmos, tarefa que é vulgarmente conhecida como preparação ou pré-processamento de dados. As técnicas de pré-processamento de dados referem-se geralmente à adição, eliminação, ou transformação conjunto de dados de treino.

O primeiro passo efetuado neste conjunto de dados consistiu numa análise gráfica que permitiu não só verificar a tipologia dos dados, como também avaliar a quantidade de valores omissos presentes nos dados. De seguida, foram estudadas as informações contidas em cada coluna e, após este estudo inicial verificou-se que a quantidade de valores omissos era superior à dos valores presentes pelo que se resolveu proceder a uma recolha de dados manual diretamente do sistema informático do hospital. Um outro problema encontrado, situou-se em algumas entradas na coluna da informação relativa à PCR, pois verificou-se que o valor apresentado na coluna da glicose e da PCR eram iguais, o que causou desconfiança da veracidade dos dados, pelo que também foi necessário retificar estes valores. Durante o período de recolha, constatou-se que alguns valores seriam impossíveis de recolher, pois não constavam no sistema. Devido a esta situação, decidiu-se eliminar as linhas cuja informação era impossível de obter, para não se deduzirem valores que poderiam não coincidir com a realidade do paciente em questão, e de alguma maneira induzir em erro o modelo de previsão. Ainda durante a fase de recolha de dados, constatou-se que poderia ser relevante adicionar dados relativos ao consumo de álcool, e também informação relativa a quatro medicamentos, nomeadamente, sertralina, tramadol, clonidina, mexazolam e lorazepam. No final desta recolha, conseguiu-se uma a base de dados sem valores omissos.

## Limpeza dos dados

É bastante comum que as bases de dados contenham dados em falta ou até dados incorretos. A identificação destes problemas é um ponto crucial para a modelação de um sistema preditivo. A limpeza dos dados refere-se à identificação e correção de erros no conjunto de dados (Rahm & Do, 2000). Como tal, nos pontos seguintes serão descritos os passos efetuados para corrigir este problema.

### Identificação de colunas com dados redundantes

Numa base de dados é bastante comum a existência de informação repetida. E, neste caso, o armazenamento de informação repetida também se verificou. Neste sentido, foram estudadas as diversas colunas e excluídas aquelas que apresentaram dados redundantes. Começou-se por eliminar as colunas geradas a partir da coluna ‘Proveniência’, sendo elas a variável ‘casa’, ‘lar’, ‘Intra-Hospitalar’, ‘Inter-Hospitalar’). A coluna referente ao local de proveniência do SU (‘Local\_SU’) apresentou o mesmo problema da anterior, pelo que as variáveis ‘UDC1’, ‘UDC2’ Unidade Cuidados Intermédios (UCISU) e Ambulatório (‘AMBUL’) também foram descartadas. Do mesmo modo se procedeu com a informação relativa à variável do grupo de diagnóstico (‘Grupo\_Diagn’) para as variáveis ‘Neurológico’, ‘Cardiovascular’, ‘Gastrointestinal’, ‘Respiratório’, ‘Geniturinário’, ‘Musculosquelético’, ‘Toxicidade de Drogas’, ‘Outro’ e ‘Hemato-Oncológico’. Foram excluídas as colunas referentes à informação que levou ao resultado para os critérios SIRS (Freq\_Resp’, 'Freq\_Card', 'Temp\_Corporal', 'Cont\_Leucócitos'), bem como a coluna SIRS \_2’ que continha informação já presente na coluna SIRS. Da mesma maneira, também as colunas ‘AC0’, ‘AC1’, ‘AC2\_3’, ‘AC\_mais\_q\_3’, ‘AC\_1\_todos’ e ‘AC\_2\_3\_todos’ foram descartadas, pois para além de não se ter entendido a razão da sua existência continham 421 valores omissos. Para além destas, o mesmo sucedeu às colunas 'Falta\_atenção\_auditiva', 'Falta\_atenção\_visual', 'RASS\_2', 'RASS\_3', pois foram dados auxiliares para se chegar ao diagnóstico do *delirium*. Ficando apenas a coluna ‘Delirium’ que expressa os resultados relativos à presença ou não de *delirium* no individuo. Assim como, a coluna ‘resulta\_hipo\_hiper’, foi também descartada, pois a informação contida é a mesma da coluna ‘Delirium’, com a diferença que especificava o tipo de *delirium* ocorrido em cada individuo.

Ainda dentro desta temática, procedeu-se ao arranjo da coluna ‘Interna\_Dias’, devido ao facto de esta coluna e a ‘Interna\_Horas’representarem o mesmo tipo de informação, mas em unidades de medida diferentes. Então, optou-se por colocar a informação contida nas duas colunas apenas numa e fez-se a conversão das horas para dias (1 hora corresponde a 0.0417 dias).

### Identificação de colunas com valores únicos

As colunas que apresentam um único valor para todas as linhas não contêm qualquer informação relevante para o processo de modelação e podem também causar erros ou resultados inesperados. Estas colunas ou preditores referem-se a preditores de variação zero, porque não há variação de valores na variável. É possível detetar linhas com esta propriedade usando a função *unique()* da biblioteca *NumPy* que devolve o número de valores únicos de cada coluna. A **Error! Reference source not found.** exibe o conjunto de variáveis que possuem apenas um valor. Sendo possível verificar que as variáveis exibidas apresentam o valor de 1, significando que as variáveis contém o mesmo valor para todas as entradas. E, por este motivo, estas quatro variáveis foram eliminadas do conjunto de dados.

Tabela 4 – Resultado da contagem de colunas com valores únicos

|  |  |
| --- | --- |
| **Coluna** | **Resultado** |
| ‘Anticolinerg\_Central’ | 1 |
| ‘Relaxante\_Musc’ | 1 |
| ‘Alverine’ | 1 |
| ‘Codeine’ | 1 |

Após a verificação das colunas que não apresentavam variação nos valores, constatou-se que as colunas: 'Antihistamínico', 'Antidepressivo', 'Antipsicótico', 'Antiespasmódicos\_GInt', 'Antiespasmódico\_GUrin', 'Antihemético', 'Analgésico' e 'Antiepiléptico', não continham a informação correta relativa aos medicamentos pertencentes a cada uma desta categorias, pelo que se decidiu eliminar estas colunas e manter as colunas dos diferentes medicamentos.

### Identificação de colunas inócuas

Por vezes torna-se útil a utilização de identificadores para o reconhecimento de determinada entrada no conjunto de dados. Porém, para este modelo de previsão, a informação contida nas variáveis de identificação, não se demonstra relevante, uma vez que o resultado não depende destes valores. Por este motivo, as colunas com informações relativas à identificação do individuo no sistema informático, como o caso das colunas ‘Id\_App’ e ‘N.ºProc’, foram excluídas.

### Exclusão da variável ‘Obito’

A variável ‘Obito’ contém informação relativa à data de óbito de determinado indivíduo, ocorrida durante o período de recolha dos dados. A informação desta coluna encontra-se apresentada em formato de data (dd/mm/aa), para o caso de ter ocorrido o óbito, e o valor ‘NULL’ para o evento oposto. De modo a retirar os valores nulos apresentados por esta variável, esta informação foi codificada com o valor 1 para o caso de ter uma data de óbito e com o valor 0, caso contrário.

A data da morte do indivíduo é uma informação relevante para a avaliar a taxa de mortalidade no *delirium*. No entanto, os dados relativos à data da morte apenas indicam se o individuo sucumbiu, havendo uma lacuna na relação entre a ocorrência da morte e o desenvolvimento do *delirium*. Por este motivo, esta variável não poderá constar no conjunto de variáveis para a modelação, uma vez que pode induzir em erro a previsão, para além disso, não faz sentido prever o *delirium* num individuo já falecido.

### Identificação de linhas com dados duplicados e omissos

As linhas com dados duplicados podem ser enganadoras durante a avaliação do modelo. Por isso, a remoção de dados duplicados será um passo importante para garantir que os dados possam ser utilizados com precisão. Uma linha duplicada é uma linha que apresenta exatamente os mesmos valores numa outra linha pela mesma ordem. A função *duplicated()* da biblioteca *Pandas* reporta se determinada linha está ou não duplicada. Marcando como Falso as linhas que não estão duplicadas ou como Verdadeiro as linhas que se apresentem duplicadas. Após a execução da função nesta base de dados, não se verificou a existência de linhas com informação duplicada.

Os dados podem ter valores em falta por uma série de razões tais como observações que não foram registadas ou dados perdidos. Por este motivo, o tratamento de dados em falta é importante, visto que muitos algoritmos de ML não suportam dados com valores em falta. Numa primeira fase foi necessário proceder à identificação das características identificadoras dos dados omissos. Após uma breve análise percebeu-se que os dados em falta estavam identificados com ‘NA’ e que os valores ‘0’ correspondiam a valores binários, sendo portanto o valor ‘0’ importante para a modelação. Portanto, após a eliminação das colunas mencionadas nos pontos anteriores e a recolha extra dos dados verificou-se com a função *isnull()* da biblioteca Pandas a não existência de linhas com dados omissos.

Portanto, após este arranjo inicial, a base de dados resultante inclui informação relativa a 434 indivíduos e 57 variáveis. E, no Apêndice I – Descrição da Base de dados, pode ser consultado o resultado deste arranjo da base de dados e também a descrição e respetiva codificação, quando aplicável, de cada uma das variáveis.

## Feature Selection

### Identificação de colunas com pouca variância

Uma outra abordagem ao problema da limpeza dos dados é a identificação de colunas com baixa variância. A variância é um conceito estatístico que mede a variabilidade de um conjunto de valores. Sendo útil para determinar o afastamento da média que os dados de um conjunto apresentam, através da determinação do valor médio das diferenças quadradas da média. Esta medida de dispersão pode ser utilizada como filtro para identificar colunas a serem removidas do conjunto de dados. Com esta abordagem assume-se que as características com uma variância superior podem conter mais informação útil, mas note-se que não se tem em conta a relação entre variáveis de característica ou características e variáveis-alvo, o que é uma das desvantagens deste método de filtragem.

A classe *‘VarianceThreshold’* da biblioteca *scikit-learn* permite selecionar as variáveis com menos variação e descartá-las. Por defeito, remove todas as variáveis com variação zero, ou seja, características que têm o mesmo valor em todas as amostras. E, através da função *fit\_transform()*, é possível criar uma versão transformada do conjunto de dados, cujas colunas que apresentem uma variância inferior ao valor limite definido são removidas. Com a execução deste filtro, como é possível verificar pelo gráfico apresentado na **Error! Reference source not found.**, com a variação da variância entre 0 e 0,55 com incremento de 0,01, verificou-se que o número de características no conjunto de dados cai rapidamente de 57 para 28, com um limiar de 0,03. Mais tarde, o número de características selecionas cai para 15 (com um total de 42 colunas suprimidas) com um limiar de 0,25. No entanto, estas colunas não serão já descartadas, pois ainda é possível agrupar a informação relativa a cada substância ativa presente nos medicamentos.

Chart

Description automatically generated

Figura 5 – Gráfico de linhas com a variação do limiar em relação ao número de variáveis selecionadas

## Transformação dos dados

### Agregação dos medicamentos por grupo farmacológico

A agregação das variáveis relativas aos medicamentos, vai permitir uma melhor organização dos dados, bem como uma diminuição das variáveis sem se perder qualquer informação. Neste sentido, foi realizada uma pesquisa para que se conseguisse efetuar a agregação dos medicamentos pelo respetivo grupo farmacológico pertencente. Nos pontos seguintes serão apresentados os diversos grupos farmacológicos bem como será dada uma contextualização do grupo farmacológico e o papel de cada substância ativa no *delirium*.

O uso de drogas têm sido associado ao desenvolvimento de *delirium* na população geriátrica e é também considerado um dos principais fatores de desencadeamento desta síndrome mais facilmente reversível (Alagiakrishnan & Wiens, 2004). O *delirium* induzido por drogas é comumente visto na prática médica, especialmente em ambientes hospitalares. A polifarmácia, a prescrição de fármacos delirogénicos, anticolinérgicos e potencialmente inadequados são fatores que contribuem para a ocorrência do distúrbio (Alagiakrishnan & Wiens, 2004; Rossi Varallo et al., 2021). Tem sido associados muitos medicamentos ao desenvolvimento do *delirium*, porém determinadas classes de medicamentos, expressas na **Error! Reference source not found.**, são mais comummente vistos como agentes causadores do *delirium*. Esta tabela classificativa foi construída de acordo com a classificação das substâncias ativas presentes no prontuário terapêutico disponibilizado pelo INFARMED.

As drogas com efeitos delirantes mais comuns incluem narcóticos em doses elevadas, benzodiazepinas, e medicamentos anticolinérgicos. A atividade anticolinérgica está também associada à ocorrência e gravidade do *delirium*. Vários estudos demonstraram que o uso de medicamentos anticolinérgicos é um fator de risco precipitante comum. E, embora o *delirium* seja uma síndrome de origem multifatorial, estima-se que os medicamentos por si só possam representar 12%-39% de todos os casos de *delirium* (Rudberg, Pompei, Foreman, Ross, & Cassel, 1997).

Tabela 5 - Lista de medicamentos considerados para o desenvolvimento de delirium

|  |  |  |  |
| --- | --- | --- | --- |
| **Grupo do fármaco** | | **Classe do fármaco** | **Substância ativa** |
| Analgésicos estupefacientes | | Analgésicos opiáceos | Morfina  Tramadol |
| Psicofármacos | Ansiolíticos, sedativos e hipnóticos | Benzodiazepinas | Alprazolam  Diazepam  Lorazepam  Mexazolam |
| Antipsicóticos | Típicos (Bloqueiam os recetores dopaminérgicos D2) e Atípicos (Bloqueiam os recetores dopaminérgicos D2 e os recetores serotoninérgicos 5HT2A) | Haloperidol  Quetiapina  Risperidona  Paliperidona  Iloperidona |
| Antidepressores | Inibidores seletivos de recaptação da serotonina (ISRS) | Fluvoxamina  Paroxetina  Sertralina |
| Inibidores seletivos da recaptação da serotonina e da noradrenalina (ISRSN) | Venlafaxina |
| Tricíclicos e afins | Trazodona  Amitriptilina |
| Antiácidos e antiulcerosos | | Antagonistas dos recetores H2 | Ranitidina |
| Anticoagulantes | | Antivitaminicos K | Varfarina |
| Antiagregantes plaquetários | Dipiridamol |
| Antidislipidémicos | | Estatinas | Rosuvastatina  Atorvastatina  Pravastatina  Sinvastatina  Fluvastatina |
| Antiespasmódicos | | | Escopolamina |
| Antiespasmódico musculotrópico | | | Alverina |
| Anti-hipertensores | | Bloqueadores da entrada do cálcio | Nifedipina |
| Inibidores da enzima de conversão da angiotensina (IECAs) | Captopril |
| Depressores da atividade adrenérgica (Agonistas alfa 2 centrais) | Clonidina |
| Diuréticos da ansa | Furosemida |
| Anti-histamínicos | | Anti-histamínicos H1 sedativos | Hidroxizina |
| Anti-histamínicos H1 não sedativos | Desloratadina |
| Antiparkinsónicos | | Anticolinérgicos | Tri-hexifenidilo |
| Antitússicos e expetorantes | | Antitússicos | Codeína |
| Cardiotónicos | | Digitálicos | Digoxina |
| Corticosteroides | | Glucocorticoides | Hidrocortisona  Prednisolona |
| Disfunções geniturinárias | | Incontinência urinária | Cloreto de tróspio |

O conjunto de dados que servirá como objeto de estudo é constituído pela informação acerca do uso ou não de determinado medicamento, sendo que a classificação de ‘Ausente’ significa que o paciente naquele momento não estava a efetuar a terapia farmacologia daquela substância ativa, já o valor ‘Presente’ indica que estava a tomar o medicamento. Nos subcapítulos que se seguem serão abordadas as principais características de cada grupo farmacológico, assim como serão avaliados os principais efeitos secundários de cada substância ativa.

**Analgésicos estupefacientes**

Os analgésicos opiáceos, compostos naturais ou sintéticos, são a base no tratamento de dores agudas e certos tipos de dores crónicas. Clinicamente, esta família de medicamentos é reconhecida por proporcionar um alívio da dor (DiFazio, 1989).

De acordo com a informação disponibilizada pelo INFARMED, das reações adversas que os opiáceos apresentam destacam-se as tonturas, sedação, náuseas e vómitos e sudação. Para além destes, podem acontecer também euforia, disforia, estados confusionais, insónia, agitação, medo, alucinações, sonolência, incoordenação dos movimentos, alteração do humor, cefaleias, alterações da visão, miose, tremor, convulsões, aumento da pressão intracraniana (INFARMED, 2010b). Daqui, pode-se constatar que os opiáceos podem efetivamente desencadear sintomas característicos do *delirium*, em resultado das propriedades farmacocinéticas e farmacodinâmicas específicas. Por exemplo, o tramadol têm metabolitos com elevadas propriedades anticolinérgicas.

Na literatura, os estudos existentes que relacionam o uso de opiáceos e o desenvolvimento de *delirium* indiciam que ainda não existem dados suficientes que permitam retirar conclusões assertivas relativas ao uso deste tipo de drogas. Mesmo assim, foi realizada uma revisão sistemática que analisou a possibilidade de ocorrência de *delirium* devido a diferentes opiáceos. Esta revisão contrasta os resultados de seis estudos observacionais que foram classificados como estudos de qualidade média-baixa, ainda assim, com estes estudos foi possível concluir que o uso de tramadol estava associado a um risco acrescido de *delirium*, enquanto que o uso de morfina, fentanil, oxicodona e codeína não. No entanto, os autores desta revisão, referem que não foram registados vários fatores importantes para a avaliação correta da influência dos opiáceos na síndrome de *delirium*, incluindo o nível de dor, a deficiência cognitiva e as comorbilidades dos pacientes. Defendem ainda que a dor neuropática pode ativar células de microglia e desta forma levar a um estado inflamatório. E como este estado está muitas vezes associado ao desenvolvimento de *delirium,* esta condição pode ter surgido devido à dor e não como resultado do uso de opiáceos (Swart, van der Zanden, Spies, de Rooij, & van Munster, 2017). Na **Error! Reference source not found.** estão apresentas as principais indicações terapêuticas e efeitos secundários adjacentes às substâncias ativas estudadas neste grupo de fármacos.

Tabela 6 - Opiáceos: indicações e efeitos adversos

|  |  |  |  |
| --- | --- | --- | --- |
| **Substância Ativa** | **Indicação** | **Efeitos adversos frequentes** | **Número de registos** |
| Tramadol | Tratamento da dor moderada a  grave de carácter agudo e crónico | Vertigens; cefaleias; suores; visão turva; sonolência; agitação; ansiedade; nervosismo; insónia; tremor; (INFARMED, 2018f) | 11 |
| Morfina | Alívio da dor grave e intratável; Alívio da dor pós-operatória | Confusão; insónias; alterações do pensamento; agitação; cefaleias; sonolência; (INFARMED, 2020g) | 4 |

Como é possível verificar pelos resultados obtidos através da contagem de linhas para cada substância ativa, expressa na Tabela 6, verifica-se que esta classe apresenta um total de 15 registos. No entanto, numa amostra de 434, este valor pode não apresentar significância no modelo.

**Ansiolíticos, sedativos e hipnóticos**

Ansiolíticos, sedativos e hipnóticos são um grupo de fármacos que têm como principal indicação o tratamento das síndromes de ansiedade, e/ou a indução ou manutenção do sono (INFARMED, 2010b). Neste grupo estão incluídas as benzodiazepinas que são psicofármacos com efeitos depressores a nível do SNC. O seu consumo provoca efeitos ansiolíticos, relaxantes, anticonvulsivantes e hipnóticos, tendo como principais finalidades terapêuticas o tratamento da ansiedade e de insónias e epilepsia. Esta classe farmacológica apresentam um amplo espectro de atividade e trata-se de uma das mais prescritas em todo o mundo, sendo que Portugal apresenta um dos maiores níveis de utilização a nível europeu (INFARMED - Direção de informação e Planeamento Estratégico, 2017).

Relativamente ao mecanismo de ação, as benzodiazepinas, facilitam a ação do ácido gamma-aminobutírico (GABA), um neurotransmissor inibidor do SNC, sobre os seus recetores. Mais especificamente, os neurónios inibidores medulares e cerebrais utilizam principalmente o ácido gamma-aminobutírico (GABA) como substância transmissora, esta que é responsável pela diminuição da excitabilidade das células-alvo através dos recetores GABA. O efeito desencadeado pela ligação do GABA aos recetores GABAA promove a abertura do canal de cloro, o que promove a hiperpolarização neuronal e, desta forma, diminui a excitabilidade da célula-alvo. O recetor, constituído de cinco subunidades, tem um local de ligação com alta afinidade pelas benzodiazepinas, para além do local de ligação do GABA. Assim, a ligação da benzodiazepina aumenta alostericamente a ligação e a ação do GABA (Lüllmann, Mohr, & Hein, 2018).

As benzodiazepinas têm todas um mecanismo de ação e efeitos adversos similares, no entanto, diferem nas suas características farmacocinéticas. E são estas diferenças que lhes conferem características que definem a escolha de um fármaco em detrimento de outro. As reações adversas causadas pelas benzodiazepinas que se observam com maior frequência são sonolência, incoordenação motora, alteração da memória a curto prazo, depressão, vertigem, confusão, alterações gastrintestinais, alterações visuais e irregularidades cardiovasculares (INFARMED, 2010b). Na **Error! Reference source not found.** estão reunidas as principais indicações e efeitos adversos mais frequentes das benzodiazepinas estudadas neste projeto, sendo elas as substâncias ativas batizadas como alprazolam, lorazepam diazepam e mexazolam. De um modo geral, as benzodiazepinas tem um uso muito difundido como sonífero, especialmente o alprazolam e o lorazepam. Já, os fármacos cuja indicação principal se prende com perturbações da ansiedade são alprazolam, lorazepam, mexazolam e diazepam.

Tabela 7 – Benzodiazepinas: principais indicações e efeitos adversos frequentes

|  |  |  |  |
| --- | --- | --- | --- |
| **Substância Ativa** | **Indicação** | **Efeitos adversos** | **Número de registos** |
| Alprazolam | Ansiedade; perturbações relacionadas com o pânico e fobias | Depressão; estado confusional; desorientação; ansiedade; nervosismo; sedação; sonolência; compromisso da memória; cansaço; irritabilidade; ataques de raiva; pesadelos; alucinações (INFARMED, 2021a) | 23 |
| Lorazepam | Ansiedade; insónia devida à ansiedade | Inquietação; agitação; irritabilidade; agressividade; ilusão; ataques de raiva; pesadelos; alucinações; fraqueza muscular; sedação; confusão; depressão; (INFARMED, 2010a) | 42 |
| Diazepam | Ansiedade; ansiedade associada a desordens psiquiátricas; | Sonolência; fala indistinta; cefaleias; tremor; tonturas; inquietação; agitação; irritabilidade; agressividade; ataques de raiva; pesadelos; alucinações; (INFARMED, 2014) | 19 |
| Mexazolam | Ansiedade associada ou não a estados psiconeuróticos | Redução do estado de alerta; confusão; fadiga; cefaleias; inquietação; agitação; irritabilidade; agressividade; ilusões; fúria; pesadelos; alucinações (MediBIAL – Produtos Médicos e Farmacêuticos, n.d.) | 6 |

Nesta categoria de fármacos verifica-se que no total estão presentes 90 entradas. No entanto destas, verificou-se que algumas poderiam pertencer ao mesmo individuo, pelo que se ignoraram as linhas que apresentavam mais do que um medicamento tendo-se verificado um total de 86 entradas para esta categoria.

**Antidepressivos**

A depressão constitui um problema grave de saúde, podendo afetar os indivíduos a nível social, pessoal e profissional prejudicando a qualidade de vida e a produtividade. Esta patologia está associada a episódios de grande duração, altas taxas de cronicidade, recaídas e a elevada morbilidade e mortalidade, uma vez que os pacientes com formas graves apresentam um risco de morte por suicídio (INFARMED & Ministério da Saúde, 2012). O tratamento de primeira linha, para a maioria dos doentes com depressão, consiste em medicação antidepressiva, psicoterapia ou uma combinação das duas (INFARMED, 2002). Os medicamentos antidepressivos têm demonstrado eficácia no alívio dos sintomas da perturbação depressiva do humor, quer se trate de depressão *major* ou de alteração distímica (INFARMED, 2010b). O seu desenvolvimento a partir da década de 1950 marcou uma das revoluções terapêuticas na área da saúde mental. De um modo geral, os antidepressivos interferem com a recaptação sináptica de um ou mais neurotransmissores, sendo a noradrenalina e a serotonina os mais relevantes no processo (INFARMED, 2002). Os antidepressivos são divididos em vários grupos, nomeadamente os inibidores seletivos da recaptação da serotonina (ISRS), os inibidores seletivos da recaptação da serotonina-noradrenalina (ISRSN), os antidepressivos tricíclicos (ADT), os inibidores da monoamina oxidase (IMAOs) e outros antidepressivos.

Antidepressivos Tricíclicos

Os ADT são assim denominados devido à estrutura química ser constituída por três anéis. São considerados antidepressivos de primeira geração e exibem vários graus de seletividade pelas bombas de recaptação de noradrenalina e serotonina. A ação antidepressiva dos ADT parece estar relacionada com o bloqueio da recaptação tanto da noradrenalina como da serotonina ao nível dos respetivos terminais nervosos, aumentando deste modo a concentração destes neurotransmissores na vizinhança dos recetores (INFARMED, 2019a). Verifica-se ainda o bloqueio de outros recetores que de certa forma podem explicar os efeitos adversos destes medicamentos. O bloqueio dos recetores H1 histaminérgicos está na origem da sedação, da mesma forma que o bloqueio dos recetores α1-adrenérgicos justificam a hipotensão ortostática. É ainda de referir que devido à inibição colinérgica, os ADT podem causar efeitos cognitivos, principalmente em doentes idosos, e em doses elevados podem causar *delirium*. Os efeitos secundários mais comuns incluem: hipotensão ortostática, convulsão, visão turva, boca seca, tremores, taquicardia e hipotensão ortostática (INFARMED, 2019a; INFARMED & Ministério da Saúde, 2012). De modo a facilitar a interpretação dos efeitos secundários mais comuns dos medicamentos desta categoria, foi construída a **Error! Reference source not found.**. Nesta tabela são descritas as principais indicações terapêuticas e os respetivos efeitos secundários mais frequentes para as substâncias ativas em estudo, a amitriptilina e a trazodona.

Tabela 8 - Substâncias ativas dos ADT e respetivas indicações e efeitos secundários

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Substância Ativa** | **Indicação** | **Efeitos adversos** | **Número de registos** | |
| Amitriptilina | Estado depressivo; depressão ansiosa | Confusão; insónia; *delirium;* perturbações da concentração; desorientação; ilusões; alucinações; excitação; ansiedade; agitação; pesadelos; (INFARMED, 2019a) | | 7 |
| Trazodona | Tratamento da depressão | Ideação suicida ou comportamento suicida; estado de confusão; insónia; desorientação; mania; ansiedade; nervosismo; agitação; ilusão; agressividade; alucinações; pesadelos; inquietação; atenção diminuída; perturbações da memória; (INFARMED, 2020i) | | 11 |

Inibidores seletivos de recaptação da serotonina (ISRS)

Os ISRS são inibidores específicos que inibem a recaptação da serotonina pelo neurónio pré-sináptico, aumentando assim a neurotransmissão seratonérgica. Esta classe de antidepressores potenciam seletivamente a atividade da serotonina [5-hidroxitriptamina (5-HT)], através da inibição da bomba de recaptação da serotonina (recetores 5-HT) (Telles-Correia, Guerreiro, Oliveira, & Figueira, 2007). O primeiro fármaco da classe surgiu na década de 1980, a fluoxetina, seguida pela paroxetina, fluvoxamina, sertralina, escitalopram e citalopram (Souza, Wildner, Gazdzichi, & Nink, 2020). Todos estes fármacos, pertencentes a famílias químicas diferentes, têm em comum a capacidade de inibir seletivamente a recaptação de serotonina (Telles-Correia et al., 2007). Normalmente são prescritos em casos de depressão leve, moderada ou grave, ansiedade, fibromialgia, diabetes e outras doenças neuropáticas. Os efeitos adversos mais comuns dos ISRS são gastrintestinais, no entanto, também são relatados casos de ansiedade, confusão, perturbações da atenção, insónias, alucinações, sonolência, amnésia, entre outros. Na **Error! Reference source not found.**, estão sintetizadas, para cada substância ativa, as principais indicações terapêuticas assim como os respetivos efeitos adversos mais relevantes para este estudo e ainda, o número registado na contagem de cada entrada nos dados relativamente a cada substância ativa.

Tabela 9 - ISRS: indicação terapêutica e efeitos adversos frequentes

|  |  |  |  |
| --- | --- | --- | --- |
| **Substância Ativa** | **Indicação** | **Efeitos adversos** | **Número de registos** |
| Fluvoxamina | Episódio depressivo *major*;  POC | Alucinação; estado confusional; agressividade; agitação; nervosismo; ansiedade; insónia; (INFARMED, 2021c) | 4 |
| Paroxetina | Episódio depressivo *major;* POC; Pânico; Ansiedade; Stress pós-traumático. | Sonolência; insónia; agitação; sonhos estranhos (incluindo pesadelos); confusão; alucinações; dificuldade de concentração; tonturas; visão turva; distúrbios sensoriais; perturbações de sono; ansiedade (INFARMED, 2021e) | 3 |
| Sertralina | Episódio depressivo *major;* Pânico; POC; Ansiedade social; Stress pós-traumático. | Depressão; pesadelos; ansiedade; agitação; nervosismo; alucinações; euforia; apatia; pensamentos anómalos; insónia; sonolência; perturbação na atenção; (INFARMED, 2020k) | 13 |

Inibidores seletivos de recaptação da serotonina-noradrenalina (ISRSN)

Os ISRSN aliviam a depressão ao afetar os mensageiros químicos utilizados para comunicar entre as células cerebrais. Tal como a maioria dos antidepressivos, os ISRSN bloqueiam a reabsorção da serotonina e noradrenalina no cérebro. Acredita-se que os ISRSN ajudam a tratar a depressão mantendo os níveis destes dois mensageiros químicos no cérebro. Fazem-no impedindo que a serotonina e a noradrenalina regressem às células que os libertaram (Stahl, Grady, Moret, & Briley, 2005).

O mecanismo da atividade antidepressiva da venlafaxina em seres humanos está relacionado com a potenciação da atividade neurotransmissora no SNC. Estudos pré-clínicos demonstraram que a venlafaxina e o seu metabolito principal, a O-desmetilvenlafaxina (ODV), são inibidores da recaptação neuronal da serotonina e da noradrenalina, sendo que a venlafaxina é também um inibidor fraco da recaptação da dopamina. A venlafaxina e o seu metabolito ativo reduzem a resposta b-adrenérgica, quer após a administração de dose única, quer crónica (INFARMED, 2021h). A venlafaxina tem como principais indicações terapêuticas a depressão e ansiedade generalizada. Os efeitos secundários mais comuns desta substância ativa são insónias, sonolência, tonturas, sedação, estado confusional, agitação, alucinações, apatia, nervosismo, hipertensão e irritabilidade (INFARMED, 2021h; INFARMED & Ministério da Saúde, 2012). Para esta substância existem apenas 5 registos nos dados.

Após análise da quantidade de registos de substâncias ativas antidepressivas, verificou-se que existem 43 linhas com os diferentes medicamentos e, que destas, 39 correspondem a valores não repetidos, ou seja, foram excluídas as entradas de indivíduos que utilizam mais do que um medicamento desta classe.

**Antipsicóticos**

Atualmente, Portugal apresenta uma elevada prevalência de doenças mentais, onde se inclui a psicose, sendo esta uma causa importante de morbilidade com elevado impacto na sociedade. Com efeito, estas doenças do foro mental trazem repercussões na qualidade de vida dos doentes, uma vez que são doenças crónicas, de evolução variável e fortemente incapacitantes (Conselho Nacional da Saúde, 2019). A psicose é um sintoma de doença mental caracterizado como um distúrbio da perceção da realidade. Os transtornos psicóticos têm diversas etiologias, e cada uma delas exige uma abordagem de tratamento exclusiva. Os mais comuns incluem os transtornos de humor, psicose induzida por substâncias, demência, transtorno psicótico breve, transtorno delirante, transtorno esquizoafetivo e esquizofrenia (Brunton, Chabner, & Knollmann, 2011).

De forma a contornar a problemática da doença psicótica, tem vindo a ser estudados componentes farmacológicos com o intuito de atenuar ou controlar a sintomatologia. O desenvolvimento dos antipsicóticos tem um história longa e também marcada por descobertas ocasionais. Em 1891, Paul Ehrlich descobriu que o azul-de-metileno, um derivado da fenotiazina, apresentava efeitos anti-maláricos. Mais tarde, em 1951, Laborit e Huguenard administraram fenotiazina alifática, a clorpromazina, aos pacientes pelos seus potenciais efeitos anestésicos durante a cirurgia. Pouco tempo depois, Hamon e Delay alargaram a utilização deste tratamento em doentes psiquiátricos e descobriram serendipitadamente a sua atividade antipsicótica (Shen, 1999). Esta descoberta marcou o início de uma era de novidade no tratamento das doenças mentais. A clorpromazina e os antipsicóticos que lhe seguiram durante a década de 1950, como a reserpina e o haloperidol, abriram caminho a uma espiral de conhecimento e inovação que se materializou numa nova conceção de saúde mental. A hipótese da hiperatividade da dopamina levou ao desenvolvimento desta classe terapêutica, atualmente conhecida como antipsicóticos típicos ou de primeira geração. Por sua vez, mais tarde com a contínua investigação, a estrutura base de alguns compostos foi alterada deixando de ser usada a fenotiazina, e passando a ser utilizados derivados de tioxantenos e butirofenonas, o que conduziu ao desenvolvimento de outros antipsicóticos conhecidos por atípicos ou de segunda geração, como a clozapina, risperidona, olanzapina, a quetiapina (Shen, 1999) ou o aripiprazol.

O mecanismo de ação exato é ainda desconhecido, mas pensa-se que este se deva ao antagonismo dos recetores dopaminérgicos, mais propriamente ao recetor D2, impedindo a libertação da dopamina, neurotransmissor associado à génese da psicose e de certos serotonérgicos, nomeadamente o 5-HT2A, recetor que se localiza também nos núcleos da base e está associado a sintomas relacionados com a esquizofrenia (INFARMED & Ministério da Saúde, 2012; Meltzer, 2013). Os antipsicóticos podem ser classificados em dois grandes grupos: os típicos ou de primeira geração e os atípicos ou de segunda geração. Esta distinção tem por base essencialmente diferenças a nível dos efeitos adversos, no mecanismo de ação e a na afinidade aos recetores (Meltzer, 2013). No entanto, esta classificação nem sempre é muito clara, pois de uma forma generalizada, é tida em consideração a afinidade para os recetores D2 e consequente risco de indução de efeitos extrapiramidais. Ora, os antipsicóticos típicos têm elevada afinidade para os recetores D2 e produzem efeitos extrapiramidais graves com maior frequência. Já os antipsicóticos atípicos são os que têm menor afinidade para os recetores D2, tendo menor probabilidade de causar efeitos extrapiramidais. Porém, recentemente foi desenvolvido um subgrupo de antipsicóticos que são simultaneamente antagonistas dos recetores da dopamina e dos da serotonina, pelo que a classificação suscita algumas dúvidas. Neste subgrupo incluem-se a risperidona, olanzapina, sertindol e clozapina (INFARMED & Ministério da Saúde, 2012).

De um modo geral, a lista das reações adversas desta classe de medicamentos inclui sintomas e sinais extrapiramidais e podem produzir, em graus variáveis, sedação e efeitos anticolinérgicos, hipotensão ortostática e arritmias (ver **Error! Reference source not found.**). Foram também registados casos de náuseas, vómitos, irritação gástrica, crises convulsivas, alterações endócrinas, alterações hematológicas, erupções cutâneas, entre outros (INFARMED & Ministério da Saúde, 2012). Dado que nesta classe de fármacos estão incluídas várias substâncias ativas, a **Error! Reference source not found.**, expõe não só as indicações mais comuns de cada substância ativa como também as reações adversas mais relevantes.

Tabela 10 - Antipsicóticos: indicações e efeitos adversos frequentes

|  |  |  |  |
| --- | --- | --- | --- |
| **Substância Ativa** | **Indicação** | **Efeitos adversos** | **Número de registos** |
| Haloperidol | Esquizofrenia; Tratamento de *delirium*; mania, tiques e síndrome de *Gilles de la Tourette*. | Agitação; insónia; perturbações psicóticas; depressão; estado de confusão; irrequietude; tonturas; sonolência; tremor; (INFARMED, 2020d) | 6 |
| Quetiapina | Esquizofrenia; tratamento da perturbação bipolar; | Sonhos anormais e pesadelos; ideação suicida e comportamento suicida; tonturas; sonolência; (INFARMED, 2018c). | 11 |
| Risperidona | Esquizofrenia; episódios de mania; | Parkinsonismo; agitação; irritabilidade; ansiedade; distúrbios do sono; depressão; medo; sedação; falta de atenção; (INFARMED, 2021f) | 3 |

Dentro desta temática, é importante realçar que as psicoses relacionadas com o *delirium* e demência, particularmente a do tipo Alzheimer, partilham uma etiologia comum relacionada com uma deficiência na neurotransmissão colinérgica, e pode ser devida a propriedades anticolinérgicas dos medicamentos, perda neuronial relacionada com a idade ou com doenças ou ambos. Entre os pacientes idosos hospitalizados, o aumento das concentrações plasmáticas de medicamentos anticolinérgicos estão diretamente associados ao aumento do risco de *delirium*. No entanto, ao contrário de pacientes com demência de Alzheimer, cujos sintomas psicóticos estão diretamente relacionados com a perda de neurónios colinérgicos e podem responder ao tratamento com acetilcolinesterase, o *delirium* pode ter inúmeros fatores precipitantes para além das propriedades anticolinérgicas associadas a medicamentos, sendo que será necessário um tratamento específico para além da remoção dos medicamentos anticolinérgicos causadores da síndrome (Brunton et al., 2011). Todavia, os resultados de um estudo, publicado em 2013, indicam que os antipsicóticos atípicos podem ser eficazes e seguros no tratamento do *delirium*. Em particular, estudos comparativos com haloperidol mostraram que a eficácia dos antipsicóticos atípicos era semelhante à do haloperidol (Wang, Woo, & Bahk, 2013).

Para concluir, os medicamentos antipsicóticos atenuam alucinações e *delirium* em doentes com perturbações neuropsiquiátricas, particularmente esquizofrenia e doença bipolar, mas variam em eficácia e efeitos secundários. Os antipsicóticos são utilizados para o tratamento de muitas perturbações, porém com mais frequência para casos de depressão resistente ao tratamento, demência, distúrbios do espectro do autismo, distúrbios do desenvolvimento generalizado e distúrbios do sono, de salientar que os distúrbios psicóticos são as suas principais indicações (Meltzer, 2013).

Após a contabilização das linhas com informação relativa aos medicamentos pertencentes à classe dos antipsicóticos, exibida na Tabela 10 verificou-se que das 20 entradas registadas todas correspondem a valores únicos de cada medicamento.

**Antiácidos e antiulcerosos**

Neste grupo agrupam-se fármacos muito heterogéneos que partilham utilização terapêutica comum como a neutralização da acidez gástrica e/ou tratamento da úlcera péptica. A Ranitidina está indicada para o tratamento de úlcera péptica, esofagite de refluxo ou síndrome de Zollinger-Ellison, sendo que está também indicado na profilaxia de úlcera de stress em situações graves; hemorragia recorrente em doentes com úlcera péptica hemorrágica ou síndrome de Mendelson. A ranitidina é um fármaco antagonista dos recetores H2 da histamina, inibindo assim a produção de ácido pelo estômago, induzida pela histamina e gastrina. Isto favorece a cicatrização da gastrite, de úlceras pépticas do estômago e do duodeno, e previne o surgimento do mal estar e azia e de outras complicações, uma vez que reduz tanto o volume secretado como o seu conteúdo em ácido e pepsina (INFARMED & Ministério da Saúde, 2012). Segundo o INFARMED, com a utilização de ranitidina não se têm observado estados de confusão, agitação ou alucinações, ao contrário da cimetidina quando usada em insuficientes renais e em doentes idosos. No entanto, na bula deste fármaco estão relatados efeitos secundários do foro psiquiátrico classificados como muito raros designadamente confusão mental reversível, depressão e alucinações, predominantemente em indivíduos gravemente doentes, em doentes idosos e em doentes nefropáticos (INFARMED, 2018d). De realçar, que no espaço de tempo da recolha dos dados, este medicamento era comumente usado. No entanto, só no início do ano de 2021 é que o INFARMED recomendou a recolha e suspensão da comercialização dos lotes de medicamentos com ranitidina, como precaução devido à presença de níveis baixos de uma impureza chamada N-nitrosodimetilamina (NDMA) (INFARMED, 2021i).

Neste grupo de medicamentos, apenas se enquadrou a ranitidina, e verificou-se que esta substância ativa apenas apresentou duas entradas, pelo que se decidiu agrupar esta categoria numa outra intitulada de ‘Outros Med’.

**Anticoagulantes**

Os sistemas fisiológicos que controlam a fluidez do sangue são bastante complexos, pois, o sangue deve permanecer líquido no interior dos vasos sanguíneos e ao mesmo tempo coagular rapidamente quando exposto. Em circunstâncias normais, ocorre um equilíbrio entre a coagulação e a fibrinólise, para impedir não só a trombose como também as hemorragias. Qualquer alteração que ocorra neste equilíbrio da coagulação poderá levar à trombose (Brunton et al., 2011).

Os trombos, também chamados de coágulos são compostos por agregados de plaquetas, fibrina e hemácias, que se podem formar nas artérias ou veias. Devido à predominância de plaquetas e fibrina nos trombos, os fármacos antitrombóticos usados para tratar a trombose incluem fármacos antiplaquetários, que inibem a agregação plaquetária, anticoagulantes, que atenuam a formação da fibrina e agentes fibrinolíticos, que degradam a fibrina (Brunton et al., 2011). Estes medicamentos são utilizados na profilaxia das patologias tromboembólicas e podem ser classificados em anticoagulantes diretos como as heparinas e em anticoagulantes indiretos como os derivados cumarínicos.

Antivitamínicos K

O potencial da varfarina como anticoagulante foi reconhecido por volta do ano de 1951. Foram sintetizados inúmeros anticoagulantes como derivados da 4-hidroxicumarina e do composto indan-1,3-diona, porém apenas os derivados da cumarina são os mais amplamente utilizados (Brunton et al., 2011). Os anticoagulantes orais derivados da 4-hidroxicumarina (acenocumarol e varfarina sódica) são antagonistas da vitamina K, uma vitamina lipossolúvel essencial para a biossíntese normal de vários fatores da coagulação. São reconhecidos como anticoagulantes de ação indireta que reduzem a síntese hepática dos fatores II, VII, IX e X da coagulação, bem como as proteínas anticoagulantes C e S (INFARMED & Ministério da Saúde, 2012). A varfarina está indicada na terapêutica e profilaxia de tromboses das veias profundas e de tromboembolismo pulmonar e também na prevenção do tromboembolismo em doentes com fibrilação auricular ou submetidos a plastias valvulares. O maior risco da terapêutica deste fármaco é a hemorragia, que pode afetar qualquer órgão, com a consequente formação de hematomas ou desenvolvimento de anemia (INFARMED, 2016b). No entanto, no documento com as características do medicamento, que foram aprovados pelo INFARMED, não estão descritos efeitos secundários afetos ao foro psicológico. Porém, segundo os resultados de um estudo realizado na Índia, em que foi avaliada a prevalência de *delirium* em doentes admitidos numa UCI com emergências cardíacas, verificou-se que de um conjunto de 27 doentes que estavam a receber a varfarina como terapêutica farmacológica 16 desenvolveram *delirium* (Lahariya, Grover, Bagga, & Sharma, 2014)*.*

No que diz respeito à contabilização do número de linhas com informação relativa à utilização da varfarina apurou-se um total de 57 indivíduos utilizadores deste fármaco.

Antiagregantes plaquetários

As plaquetas são células sanguíneas que possuem um papel fundamental no processo de coagulação, pois produzem um tampão hemostático inicial nos locais de lesão vascular (Brunton et al., 2011). Os antiagregantes plaquetários previnem as oclusões tromboembólicas agudas da circulação arterial, por reduzirem a agregação das plaquetas e inibirem a formação de trombos na sequência de uma disfunção do endotélio. O dipiridamol é um vasodilatador que interfere com a função plaquetária através do aumento da concentração celular de adenosina monofosfato cíclico (AMPc), este efeito é mediado pela inibição da fosfodiesterase e/ou pelo bloqueio da captação da adenosina. Não tem qualquer benefício por si mesmo mas, associado à varfarina, usa-se na profilaxia da doença tromboembólica em doentes com próteses valvulares e associado ao ácido acetilsalicílico na prevenção secundária de AVC isquémico e de acidentes isquémicos transitórios (INFARMED & Ministério da Saúde, 2012). O dipiridamol é um medicamento antiagregante plaquetário, utilizado como adjuvante dos anticoagulantes orais na profilaxia do tromboembolismo associado a próteses valvulares cardíacas. As reações adversas com doses terapêuticas são geralmente ligeiras e transitórias, tendo sido relatados casos de vómitos, diarreia e sintomas tais como tonturas, náuseas, dores de cabeça e dores musculares (INFARMED, 2020b).

Este medicamento apresenta apenas um registo de utilização na base de dados. No entanto, o somatório destes dois medicamentos perfazem um total de 58 entradas, sendo que são todas correspondentes a valores únicos. Portanto a variável ‘Anticoagulantes’, irá apresentar 58 registos para o valor ‘Presente’ e 376 para o valor ‘Ausente’.

**Antidislipidémicos**

A dislipidemia é considerada um fator de risco clínico para o desenvolvimento de doenças cardiovasculares, visto que está implicada na génese da aterosclerose, condição na qual ocorre uma acumulação anormal de gordura nas paredes das artérias dando origem à formação de placas compostas por gordura e tecido fibroso (Associação Portuguesa de Nutrição, 2018; Brunton et al., 2011; INFARMED & Ministério da Saúde, 2012). Por definição, a dislipidemia é um termo usado para designar um conjunto de anomalias quantitativas ou qualitativas dos lípidos no sangue. As dislipidemias classificam-se em vários tipos e podem manifestar-se através da hipercolesterolemia, que se caracteriza por um aumento do colesterol total (CT) e/ou do colesterol LDL *(low-density lipoprotein)*; da hipertrigliceridemia, definida por um aumento dos triglicerídeos (TG); da dislipidemia mista que combina os dois fatores anteriores; e da hipolipidemia definida por uma redução do níveis de colesterol HDL *(high-density lipoprotein)* (Associação Portuguesa de Nutrição, 2018). Esta condição tem sido observada globalmente e pode ser provocada não só por problemas relacionados com a genética, mas também devido à adoção de um estilo de vida associado a comportamentos sedentários, tabagismo e dietas ricas em calorias, gordura saturada e colesterol. Para muitos indivíduos, a correção destes fatores de risco pode ser fundamental para uma prevenção primária de doenças cardiovasculares resultando num aumento da qualidade de vida e diminuição de custos com terapia medicamentosa (Brunton et al., 2011).

Ao longo de quase um século de investigação, os cientistas desenvolveram várias linhas de evidência que estabelecem a relação causal entre colesterol sanguíneo, aterosclerose, e doença coronária. Com base nesse conhecimento, os cientistas e a indústria farmacêutica desenvolveram com sucesso uma classe eficaz de medicamentos, apelidada de estatinas. Estas reduzem os níveis de colesterol no sangue e reduzem a frequência dos ataques cardíacos. As estatinas foram isoladas do fungo *Penicillium citrinum*, e identificadas como inibidoras da biossíntese de colesterol em 1976. Estudos subsequentes estabeleceram que as estatinas atuam mediante a inibição da HMG-CoA (3-hidroxi-3-metil-glutaril coenzima A) redutase. A primeira estatina estudada em seres humanos foi a *compactina*, renomeada *mevastatina*, que demonstrou o potencial terapêutico desta classe de fármacos. No entanto, a primeira estatina aprovada para uso em humanos foi a lovastatina (também conhecida como mevinolina), esta que foi isolada do fungo *Aspergillus terreus.* Atualmente, existem mais seis estatinas, sendo que a pravastatina e a sinvastatina são metabólitos fúngicos, e cada uma contém um anel hexaidronaftaleno. A fluvastatina, a atorvastatina, rosuvastatina e a pitavastatina são compostos totalmente sintéticos, que contém uma cadeia lateral de ácido heptanoico, que forma um análogo estrutural do intermediário da HMG-CoA (Brunton et al., 2011).

O mecanismo de ação das estatinas prende-se com a redução dos níveis de LDL, devido a uma porção semelhante ao ácido mevalónico, que inibe competitivamente a HMG-CoA redutase. Ao reduzir a conversão da HMG-CoA em mevalonato, as estatinas inibem uma etapa inicial e limitante de velocidade na biossíntese do colesterol. Assim, as estatinas afetam os níveis sanguíneos de colesterol ao inibir a síntese hepática de colesterol, resultando num aumento da expressão do gene recetor de LDL. O tratamento das dislipidemias e em particular da hipercolesterolemia, assumiu uma importância renovada pela introdução na prática clinica dos inibidores da redutase da HMG-CoA (Brunton et al., 2011).

Os estudos e vigilância iniciais após a comercialização das estatinas revelaram uma elevação dos níveis de transaminases hepáticas para valores acima do triplo do limite superior da normalidade, que parece estar relacionado com a dose (Brunton et al., 2011). Em 2012, a Food and Drug Administration (FDA), a agência de vigilância sanitária dos Estados Unidos, obrigou a que fossem introduzidos rótulos de segurança nas estatinas para advertir médicos e pacientes sobre os potenciais efeitos colaterais relacionados com cérebro, como perda de memória e confusão mental. Para além disso, segundo o resumo das características do medicamento, foram notificados na pós-comercialização casos de insuficiência cognitiva associados à utilização de estatinas (INFARMED, 2018e). Das reações adversas mais frequentes que podem provocar são de referir: insónias, pesadelos, dores abdominais, náuseas, anorexia, astenia e perturbações psíquicas. Estes são alguns dos efeitos adversos já conhecidos relacionados com as estatinas, e estão descritos na **Error! Reference source not found.** de acordo com a substância ativa em estudo.

Tabela 11 – Estatinas: efeitos adversos mais frequentes

|  |  |  |  |
| --- | --- | --- | --- |
| **Substância Ativa** | **Indicação** | **Efeitos adversos relevantes** | **Número de registos** |
| Atorvastatina | Hipercolesterolemia; Prevenção da doença cardiovascular | Pesadelos; insónias; cefaleias; epistaxis; dispepsia; náuseas; mialgia; (INFARMED, 2021b) | 39 |
| Fluvastatina | Hipercolesterolemia; Dislipidemia mista; Prevenção da doença cardíaca coronária | Insónias; cefaleias; náuseas; dispepsia; dor abdominal; creatina fosfoquinase aumentada; aumento de transaminases no sangue (INFARMED, 2020c) | 2 |
| Pravastatina | Hipercolesterolemia; Dislipidemia mista; hiperlipidemia pós-transplante | Pesadelos; perda de memória; depressão; tonturas; cefaleias (INFARMED, 2018b) | 15 |
| Rosuvastatina | Hipercolesterolemia; Dislipidemia mista; Prevenção cardiovascular | Cefaleias; tonturas; diabetes *mellitus*; depressão; alterações do sono; perda de memória; (INFARMED, 2019c) | 11 |
| Sinvastatina | ﻿Hipercolesterolemia; Prevenção cardiovascular; Dislipidemia mista; | Perda de memória; esquecimento; amnésia; defeito de memória; confusão; distúrbios do sono; tonturas; (INFARMED, 2018e) | 74 |
| **Total:** |  |  | 141 |

Do total de 141 registos, tal como é possível deduzir da Tabela 11, verificou-se que 139 correspondem a entradas únicas, isto é, não foram contabilizados as linhas de medicamentos que pertencem ao mesmo individuo. Por exemplo, no caso de um individuo possuir três medicamentos desta categoria, apenas se contabilizou um.

**Antiespasmódicos**

A dor abdominal crónica é um sintoma gastrointestinal comum que se caracteriza pelos distúrbios gastrointestinais funcionais da interação entre o intestino e o cérebro, incluindo síndrome do cólon irritável e dispepsia funcional. Os sintomas da dor abdominal são frequentemente tratados com agentes antiespasmódicos. Estes, são um grupo de substâncias que previnem ou interrompem a contração dolorosa e involuntária (espasmo) do músculo liso intestinal, um dos mecanismos referidos na génese da dor nas patologias gastrointestinais. O tratamento inclui uma gama de classes terapêuticas que são classificadas de acordo com o mecanismo de ação, dos quais se podem destacar os agentes anticolinérgicos/antimuscarínicos, que inibem a contração muscular lisa; inibidores dos canais de cálcio que inibem o transporte de cálcio para o músculo liso; e relaxantes musculares lisos diretos que inibem o transporte de sódio e cálcio (Brenner & Lacy, 2021).

Nesta classe de fármacos, destaca-se o butilbrometo de escopolamina, um alcalóide encontrado em plantas da família das solanáceas, que atua bloqueando os recetores muscarínicos da acetilcolina do músculo liso, e desta forma, impede a sua contração, o que diminui a dor e o desconforto gástrico. Segundo a bula desta substância ativa, não ocorrem efeitos adversos anticolinérgicos no SNC devido ao facto de possuir um derivado de amónio quaternário, o brometo de N-butil-hioscina, este que não passa para o SNC. Os efeitos secundários desta classe de medicamentos são raros e podem incluir reações cutâneas ou cardiopatias. É referido também que a ação anticolinérgica periférica resulta da ação bloqueadora dos gânglios na parede visceral, assim como da atividade antimuscarínica. E ainda que, o efeito anticolinérgico de antidepressivos tricíclicos e tetracíclicos, antipsicóticos ou anti-histamínicos pode ser potenciado por este medicamento (INFARMED, 2020j). No entanto, um estudo realizado em animais que teve como objetivo a avaliação dos efeitos da escopolamina no desenvolvimento da resposta neuroinflamatória do *delirium* demonstrou que nos modelos animais a escopolamina provocou alterações neuropsiquiátricas semelhantes ao *delirium* (Cheon et al., 2021).

Relativamente a esta substância ativa, apenas se registaram três linhas, pelo que também esta categoria de medicamentos será agrupada na ‘Outros Med’.

**Anti-hipertensores**

A hipertensão arterial é um importante fator de risco responsável pelo aparecimento de várias complicações cardiovasculares, tais como, acidente vascular cerebral (AVC), insuficiência cardíaca, enfarte do miocárdio, insuficiência renal, morte súbita e doença arterial periférica (INFARMED & Ministério da Saúde, 2012). Esta patologia é considerada um problema importante de saúde pública devido à prevalência da hipertensão arterial aumentar com a longevidade e com a adoção de estilos de vida desadequados da população. Porém, nem sempre a mudança de hábitos de vida é suficiente, pelo que o recurso aos anti-hipertensores é frequentemente necessário. Os anti-hipertensores podem ser classificados de acordo com seu mecanismo de ação e os principais grupos desta classe de fármacos são os inibidores do enzima de conversão da angiotensina, os antagonistas dos receptores da angiotensina, os bloqueadores dos canais de cálcio, os vasodilatadores, os diuréticos e os bloqueadores β. As substancias ativas relevantes para este estudo, de cada classe destes fármacos, serão abordadas de seguida.

Bloqueadores dos canais de cálcio

A nifedipina está no conjunto de substâncias que bloqueiam os canais de cálcio. Os antagonistas do cálcio reduzem o fluxo transmembranar dos iões cálcio para a célula, através dos canais lentos de cálcio. A nifedipina dilata as artérias coronárias, reduz o tónus vascular do músculo liso nas artérias coronárias e evita o vasospasmo. O resultado final traduz-se por um aumento do fluxo sanguíneo pós-estenótico e por um maior aporte de oxigénio. Com os bloqueadores dos canais de cálcio podem ocorrer diversos efeitos laterais tais como cefaleias, tonturas e náuseas. A nifedipina e pode ainda provocar perturbações do sono, ansiedade e sonolência (INFARMED, 2021d).

No que diz respeito à contabilização do número de entradas deste medicamento, foram apurados três indivíduos a usar a nifedipina.

Inibidores da enzima conversora de angiotensina (IECAs)

O captopril e os outros fármacos incluídos neste grupo atuam sobre o sistema renina-angiotensina-aldosterona inibindo a enzima que converte a angiotensina-I em angiotensina-II. Esta última é ﻿uma vasoconstritora potente, responsável pela vasoconstrição arterial, pelo aumento da pressão arterial, e pela estimulação da glândula suprarrenal para produzir aldosterona. Através do uso do captopril ou similares ocorre uma inibição da enzima conversora de angiotensina, o que resulta na diminuição plasmática da angiotensina-II, o que leva a uma atividade vasopressora diminuída e a uma produção de aldosterona reduzida (INFARMED, 2018a).

Os IECAs são anti-hipertensores de 1ª linha, capazes de modificarem também certos parâmetros, tais como, a resistência à insulina e a ﻿hipertrofia ventricular esquerda. São utilizados no tratamento da hipertensão e têm sido utilizados também com sucesso no tratamento da insuficiência cardíaca, no tratamento do enfarte do miocárdio e na prevenção da nefropatia diabética do tipo I (INFARMED, 2018a). Os IECAs podem originar insónias, confusão, depressão, tonturas, cefaleias, sonolência e muito raramente causar AVC ou ﻿insuficiência vascular cerebral. Outros efeitos adversos comuns dos IECA incluem náuseas, vómitos, boca seca, tosse seca ou dispneia (INFARMED, 2018a).

Para este medicamento verificou-se a existência de sete linhas que registaram a utilização do fármaco.

Diuréticos da ansa

Os diuréticos da ansa inibem a reabsorção de sódio no ramo ascendente da ansa de Henle, e neste grupo de fármacos está incluída a furosemida. Estes diuréticos exercem também efeitos diretos sobre o fluxo sanguíneo, originando vasodilatação e redução da resistência vascular renal. Do ponto de vista farmacológico, a furosemida inibe o sistema de reabsorção dos eletrólitos sódio, potássio e cloro, localizado ao nível da membrana da célula luminal do ramo ascendente da ansa de Henle. Consequentemente, a eficácia da ação salurética da furosemida, depende do alcance do medicamento no lúmen tubular através do mecanismo de transporte aniónico. A ação diurética resulta da inibição da reabsorção de cloreto de sódio neste segmento da ansa de Henle. Como resultado, tem-se um incremento na taxa de excreção do sódio que implica um aumento da excreção urinária (INFARMED, 2016a).

A furosemida está indicada para o tratamento de hipertensão arterial e ainda para o tratamento de edemas de origem cardíaca, hepática ou renal. As reações adversas mais comuns dos diuréticos da ansa resultam em hemoconcentração, desequilíbrios eletrolíticos, desidratação, aumento da creatinina, colesterol e ácido úrico e/ou encefalopatia hepática. Embora menos frequentes, também foram relatados casos de cefaleias, perda auditiva e vómitos (INFARMED, 2016a).

Para a furosemida, foram registados 92 indivíduos que apresentaram este medicamento na sua medicação habitual.

Depressores da atividade adrenérgica

Os depressores da atividade adrenérgica são fármacos que atuam em locais distintos do sistema nervoso simpático e tem a função de impedir que a neurotransmissão adrenérgica atinja a célula efetora localizada nos vasos, coração e/ou rins, impedindo a elevação da pressão arterial. São classificados em bloqueadores alfa, bloqueadores beta e agonistas alfa-2-centrais. Os bloqueadores alfa obtêm o seu efeito anti-hipertensor através do bloqueio dos recetores alfa-adrenérgicos tendo como resultado a inibição do efeito vasoconstritor. Os bloqueadores beta antagonizam a ação das catecolaminas nos recetores beta adrenérgicos no sistema nervoso central e periférico e assim reduzem o débito cardíaco. Os agonistas alfa-2-centrais inibem a atividade simpática nos centros vasomotores cerebrais, principalmente por impedirem a libertação de noradrenalina nas sinapses nervosas e, deste modo, inibem a vasoconstrição por ela provocada (WHO & EMRO, 2005). A substância ativa que se está a estudar pertence a este último e é denominada de clonidina. Este fármaco, foi inicialmente desenvolvido como um descongestionante nasal vasoconstritor atenuante nos recetores alfa-2 periféricos. Durante os ensaios clínicos foi observado que a clonidina causava hipotensão, sedação e bradicardia (Brunton et al., 2011). Hoje, sabe-se que a clonidina trata-se de um ﻿estimulante dos receptores alfa-adrenérgicos. No SNC, a ativação dos alfa adrenorecetores em estruturas relacionadas com o centro vasomotor induz uma inibição prolongada da atividade simpática com predomínio vagal, o que provoca a hipotensão. Assim sendo, a clonidina é um medicamento usado para tratar a hipertensão arterial, transtorno de atenção e hiperatividade, abstinência de drogas, espasticidade e certas condições de dor. Está também indicada para o tratamento de casos de glaucoma nas suas várias formas, na prevenção da hipertensão ocular no pós-operatório e na redução da pressão ocular em situações refratárias a outros tratamentos (INFARMED, 2019b).

A utilização de clonidina pode causar xerostomia, astenia, hipotensão moderada, alterações do paladar, desconforto e prurido ocular, cefaleias ou sonolência (INFARMED, 2019b). A relação entre o uso de clonidina e a potenciação de *delirium* ainda é pouco estudada e é um efeito secundário muito pouco reportado. E, segundo um estudo publicado em 2006, a clonidina pode induzir uma variedade de efeitos secundários psicológicos que vão desde a depressão, à alucinação aguda e ao *delirium* (Delaney, Spevack, Doddamani, & Ostfeld, 2006). No entanto, um outro estudo piloto publicado em 2010, concluiu que a administração de clonidina em pacientes submetidos a cirurgia cardiotorácica mostrou resultados promissores relativamente à gravidade da síndrome de *delirium* (Rubino et al., 2010).

Já este fármaco apresentou apenas dois registos na base de dados. No entanto, a soma dos registos das substâncias ativas pertencentes ao grupo dos anti-hipertensores perfazem um total de 104, sendo que 98 registos são únicos, e serão estes os utilizados para a modelação.

**Anti-histamínicos**

No inicio do século XX, foram realizados estudos farmacológicos extensivos a um produto de putrefação da histidina, e descobriu-se que este estimulava vários músculos lisos e tinha ação vasodepressora intensa. Em 1927, este produto foi isolado de amostras frescas de pulmão e fígado e, deste modo, demonstrou-se que esta amina era um constituinte natural dos tecidos dos mamíferos, resultando na adoção do termo histamina, que se origina da palavra grega *histos* (tecido) (Brunton et al., 2011; Church & Maurer, 2014). Esta descoberta incitou a que se iniciasse uma investigação na procura de uma substância que exercesse um antagonismo especifico em relação à histamina. Consequentemente, no ano de 1937 surge a primeira substância com propriedades anti-histamínicas, a timoxietildietilamina, tendo sido a rampa de lançamento para o aparecimento de outras substâncias classificadas como anti-histamínicos de primeira geração (ou anti-histamínicos H1 sedativos) (Church & Maurer, 2014). Por volta da década de 1980, ocorreu um grande avanço no desenvolvimento de anti-histamínicos com a introdução de anti-histamínicos H1 de segunda geração, que são minimamente ou não sedativos devido à sua penetração limitada da barreira hemato-encefálica. Além disso, estes medicamentos são altamente seletivos para o receptor de histamina H1 e não têm efeitos anticolinérgicos (Church & Maurer, 2014)

Hoje sabe-se que a histamina endógena participa da resposta alérgica imediata e é um importante regulador da secreção gástrica. Para além disso, foi demonstrado que a histamina é responsável pela modulação de processos fisiológicos, atuando tanto no SNC como na periferia, através dos seus recetores, designados H1, H2, H3 e H4 (Brunton et al., 2011).

De um modo geral, os anti-histamínicos são fármacos que inibem a ação da histamina, bloqueando a sua ligação aos recetores H1. Sendo normalmente utilizados para alivio de alergias no tratamento de reações de hipersensibilidade. No caso especifico da hidroxizina, um fármaco classificado como anti-histamínico H1 sedativo, é indicado para o tratamento sintomático da ansiedade e de prurido. Os efeitos indesejáveis associados a este fármaco são principalmente relacionados com efeitos depressores do SNC ou com a atividade anticolinérgica ou com reações de hipersensibilidade. De entre os efeitos secundários podem destacar-se sonolência, cefaleias, fadiga, xerostomia, agitação, confusão, desorientação, alucinações, sedação, insónias, tremores e náuseas (INFARMED, 2020f). Já a substância desloratadina, classificada como anti-histamínico H1 não sedativo, é indicada para o tratamento de rinite alérgica e urticária. Dos efeitos colaterais apresentados, salienta-se as alucinações, cefaleias, tonturas, sonolência, insónias, hiperatividade psicomotora, convulsões xerostomia e fadiga (INFARMED, 2020a).

No que diz respeito à contagem de entradas deste fármaco, observou-se que apenas se efetuaram seis registos para a hidroxizina e apenas um registo para a desloratadina. Por esse motivo, decidiu-se agrupar também esta classe de medicamentos na variável ‘Outros Med’.

**Antiparkinsónicos**

A doença de Parkinson é a segunda doença degenerativa do SNC mais prevalente nas sociedades (Cacabelos, 2017). É um distúrbio degenerativo do SNC que se caracteriza pela lentidão e falta de movimentos, rigidez muscular, tremores em repouso e desequilíbrio postural (Brunton et al., 2011). De uma forma simplista na fisiopatologia da doença de Parkinson, os medicamentos utilizados no seu tratamento dividem-se em anticolinérgicos e dopaminomiméticos. Segundo a história, os fármacos antiparkinsónicos anticolinérgicos foram os primeiros fármacos a mostrarem-se eficazes no controlo de sintomas do parkinsonismo (INFARMED & Ministério da Saúde, 2012). No caso especifico do tri-hexifenidilo, um medicamento pertence à classe mencionada anteriormente, é indicado para o tratamento da doença de Parkinson e outras formas de parkinsonismo, nomeadamente o induzido por drogas. É também eficaz na diminuição da rigidez muscular, do tremor das extremidades e da bradicinésia associados à Doença de Parkinson, e também, dos efeitos extrapiramidais causados por medicamentos antipsicóticos. Tem sido reportados efeitos adversos relacionados com esquecimento, agitação psico-motora, confusão mental, alucinações, paranóia, psicoses, alterações do sono, sonolência (INFARMED, 2009). Embora seja mencionando que estes efeitos colaterais ocorrem raramente, são relevantes o estudo da predição da síndrome de *delirium*.

Na base de dados, este fármaco apresentou apenas duas entradas nos registos. Por esse mesmo motivo também se acrescentou esta variável ao conjunto de variáveis a agrupar na variável ‘Outros Med’.

**Digitálicos**

A insuficiência cardíaca é uma síndrome causada por uma anomalia da função cardíaca, causando um débito sanguíneo inadequado às necessidades metabólicas do organismo quer em repouso quer em esforço. A prevalência desta doença aumenta não só com o avanço da idade, como também pela ocorrência de hipertensão arterial e de enfarte do miocárdio em pessoas a partir dos 65 anos (Fonseca et al., 2017). De modo a melhorar a qualidade de vida dos indivíduos com esta patologia, ao longo dos anos foram estudados diversos fármacos para o tratamento desta doença, alguns deles já referidos anteriormente e que não pertencem a esta classe. Dentro desta classe de medicamentos, frisa-se a digoxina. Este medicamento é usado para o tratamento da a insuficiência cardíaca crónica e para o tratamento de algumas arritmias supraventriculares (INFARMED, 2011).

A digoxina tem dois mecanismos de ação principais que são utilizados seletivamente, um deles é o efeito inotrópico positivo em que aumenta a contratilidade do miocárdio por ação direta. Este efeito faz com que a força de contração do coração seja aumentada através da inibição específica da adenosina trifosfatase alterando assim a atividade das trocas sódio-potássio. Esta alteração iónica distribuída pela membrana, resulta num aumento do influxo do ião cálcio no coração causando um aumento da contractilidade (Brunton et al., 2011; INFARMED, 2011). Por outro lado, a digoxina exerce o mesmo efeito fundamental de inibição do mecanismo de troca sódio-potássio nas células do sistema nervoso autónomo, estimulando-as a exercer atividade cardíaca indireta. Este mecanismo retarda a condução elétrica no nódulo atrioventricular diminuindo assim a frequência cardíaca (INFARMED, 2011).

Este medicamento apresenta algumas reações adversas ocorridas devido ao uso da substância destacando-se perturbações do foro psiquiátrico que envolvem depressão, psicose, apatia e estados de confusão (INFARMED, 2011).

Relativamente à contagem de linhas da categoria deste fármaco, foram contabilizadas 28 entradas.

**Glucocorticóides**

O sistema endócrino é formado por um conjunto de glândulas produtoras de hormonas, muitas das quais reguladas por hormonas estimuladoras segregadas pela hipófise, a qual estabelece a ligação entre o sistema endócrino e o hipotálamo. É uma rede complexa de sinais e mensagens químicas que controla as funções e reações corporais interagindo diretamente com o sistema nervoso (Nussey & Whitehead, 2001). Existe uma grande diversidade de glândulas e hormonas que desempenham um papel fundamental no crescimento e desenvolvimento, na reprodução e na diferenciação sexual e ainda na formação do sistema nervoso e imunológico (Chrousos, 1993; Jaszczyk & Juszczak, 2021).

As hormonas estão presentes na natureza, tanto em espécies animais como em espécies vegetais, e embora existam diferenças de estrutura e função, o mecanismo de ação é semelhante. Esta é a razão pela qual, qualquer substância que interfira no mecanismo de ação hormonal pode alterar o desenvolvimento, a reprodução e outras funções dos seres vivos das diversas espécies (Brunton et al., 2011). As hormonas podem ser classificadas em diversas categorias, de acordo com o critério de classificação. Mas aqui, apenas se vai referir as hormonas esteróides, estas que possuem a capacidade de entrar no interior da célula e interagir ao nível do ADN. A síntese das hormonas esteróides ocorre a partir do colesterol em diversos órgãos como o fígado, cérebro, órgãos reprodutores femininos e masculinos. A classe das hormonas esteróides liga-se a recetores intracelulares que possuem características funcionais idênticas. Assim, esta classe pode ser subdividida em cinco grupos de acordo com o recetor ao qual se liga: glucocorticóides, mineralcorticóides, androgénios, estrogénios e progestrogénios (Nussey & Whitehead, 2001). Destes grupos, apenas se dará realce aos glucocorticóides que se caracterizam pela sua ligação ao cortisol.

De um modo geral, os glucocorticóides são hormonas esteróides produzidas e segregadas pelo córtex da supra-renal e desempenham um papel fundamental na regulação fisiológica e na manutenção da homeostasia de vários sistemas, como o cardiovascular, imune, metabólico ou nervoso central (Chrousos, 1993). Esta classe de medicamentos é comumente utilizada para o tratamento de inflamações, doenças autoimunes e cancro. A descoberta deste fármaco iniciou-se por volta do século XX, quando Edward Calvin Kendall isolou quatro compostos esteroidais de adrenalina. No entanto, o potencial terapêutico foi descoberto pelo reumatologista Philip Hench num paciente que sofria de artrite reumatoide. A partir da descoberta do potencial anti-inflamatório, estes medicamentos foram amplamente utilizados e estudados (Timmermans, Souffriau, & Libert, 2019). Neste grupo de fármacos, estão incluídas a hidrocortisona e a prednisolona, que são substâncias a ser avaliadas na interferência do *delirium,* em que as principais áreas de indicação e efeitos secundários estão descritos na **Error! Reference source not found.**.

Tabela 12 – Efeitos secundários dos glucocorticoides

|  |  |  |  |
| --- | --- | --- | --- |
| **Substância Ativa** | **Indicação** | **Efeitos adversos relevantes** | **Número de registos** |
| Hidrocortisona | Insuficiência supra-renal aguda; anafilaxia aguda; asma, edema cerebral, encefalites | Inibição secundária da reatividade hipotálamo-hipófise-suprarrenal; diminuição da tolerância aos hidratos de carbono; (INFARMED, 2020e) | 6 |
| Prednisolona | Doenças reumáticas, respiratórias, neoplásicas, hematológicas; alergias | Euforia; insónias; depressão; vertigens; cefaleias; hipercinesia; supressão do eixo hipotálamo-hipófise-suprarrenal; retenção de sódio; (INFARMED, 2020h) | 22 |
| **Total:** |  |  | 28 |

Tal como é possível verificar pela Tabela 12, as duas substâncias ativas são responsáveis pela informação contida em 28 linhas, no entanto, apenas 25 correspondem a indivíduos utilizadores desta classe de medicamentos, pois três indivíduos apresentaram a utilização tanto da prednisolona como da hidrocortisona. Por este motivo, a variável ‘Corticosteroides’ apresenta apenas 25 registos com o valor ‘Presente’ e 409 para o valor ‘Ausente.

**Medicamentos usados na incontinência urinária**

A incontinência urinária é definida como uma condição em que há perda involuntária de urina. Estas perdas podem ser causadas por anomalias da bexiga, locais ou neurogénicas e/ou por anomalias do esfíncter que conduzem a uma perda na sua eficácia. A capacidade de armazenar e eliminar a urina exige uma série complexa de mecanismos inter-relacionados que envolvem o cérebro, a bexiga, a uretra, os músculos e os nervos do pavimento pélvico. À medida que a urina é produzida e armazenada na bexiga, o músculo detrusor relaxa e distende-se para a poder acomodar. A falha em algum destes mecanismos é o que pode provocar os vários tipos de incontinência urinária. E é para a atenuação destes sintomas que se utilizam os fármacos atenuadores dos sintomas de incontinência urinária. Um dos fármacos utilizados para este fim é o cloreto de tróspio que é um derivado pertencente à classe dos fármacos parassimpaticolíticos ou anticolinérgicos, pois compete com a acetilocolina, de modo dependente da concentração, para os recetores de ligação parassimpáticos pós-sinápticos do transmissor endogéno. O cloreto de tróspio liga-se com elevada afinidade aos recetores muscarínicos dos subtipos M1, M2 e M3 e demonstra afinidade para os recetores nicotínicos. Consequentemente, o efeito anticolinérgico do cloreto de tróspio exerce uma ação relaxante do tecido muscular liso e sobre as funções orgânicas mediadas pelos recetores muscarínicos. O cloreto de tróspio diminui o tónus do músculo liso dos aparelhos gastrointestinal e geniturinário. Além disto, pode inibir a secreção do muco brônquico, saliva, suor e a acomodação ocular. Este fármaco é utilizado para o tratamento sintomático da incontinência de urgência e/ou da frequência e urgência urinárias aumentadas que podem ocorrer em doentes com bexiga hiperativa. Foram relatados casos de cefaleias, alucinações, confusão e estados de agitação (INFARMED, 2021g).

Relativamente ao aparecimento do cloreto de tróspio sabe-se que apresenta quatro entradas, pelo que também será uma variável constituinte do novo grupo ‘Outros Med’.

**Categoria outros medicamentos**

Como referido ao longo deste subcapítulo, foi criada uma nova categoria, a ‘Outros Med’, para acolher os registos dos medicamentos com menos representação nos dados, nomeadamente o grupo farmacológico que apresentasse 10 ou menos entradas. De modo a resumir quais variáveis consideradas, foi construída a Tabela 13, onde se verifica a transformação da informação contida em seis variáveis, transferida apenas para uma. Tendo resultado uma variável com 18 entradas com o valor ‘Presente’.

Tabela 13 - Agrupamento das variáveis na categoria 'Outros Med'

|  |  |  |  |
| --- | --- | --- | --- |
| **Substância ativa** | **Número de registos** | **Grupo farmacológico** | **Nova categoria** |
| Ranitidina | 2 | Antiácidos | Outros Med |
| Butilbrometro de escopolamina | 3 | Antiespasmódicos |
| Desloratadina  Hidroxizina | 1  6 | Anti-histamínico |
| Tri-hexifenidilo | 2 | Antiparkinsónicos |
| Cloreto de tróspio | 4 | Geniturinário |
| **Total:** | 18 |  |  |

### Codificação dos dados categóricos

Os modelos de ML requerem que as variáveis de entrada e saída sejam numéricas (Brownlee, 2017). Isto significa que será necessário proceder a uma codificação das variáveis categóricas presentes nesta base de dados. Para lidar com este problema o Python tem disponível na biblioteca Pandas a função *get\_dummies()*, que foi aplicada às variáveis apresentadas na Tabela 14, e onde é possível ver a codificação gerada. Esta função, conta o número de instâncias (n) e vai criar uma nova coluna para cada instancia e atribui valores numéricos de 1 para o caso de se verificar a característica e 0 caso não se verifique. Como é de esperar, no caso desta colunas que apresentam apenas dois valores, a informação vai aparecer duplicada, pelo que se torna útil a passagem do parâmetro *drop\_first=True,* que permite eliminar a primeira coluna gerada. Exemplificando, para a coluna ‘Genero’, foram geradas duas colunas uma ‘Genero\_Feminino’ e outra ‘Genero\_Masculino’. E cada uma delas foi codificada com ‘1’ para o caso de ser a característica daquela coluna e com ‘0’ caso contrário. Como a informação final nas duas colunas seria a mesma, pois uma era o oposto da outra, optou-se por excluir a coluna que foi gerada em primeiro lugar, neste caso a ‘Genero\_Feminino’.

Tabela 14 - Codificação das variáveis categóricas

|  |  |  |
| --- | --- | --- |
| **Variável** | **Codificação** | **Contagem** |
| ‘Genero’ | 1 – Masculino; 0 - Feminino | Masculino - 257; Feminino - 177 |
| 'Antidislipidemicos' | 1 – Presente ; 0 – Ausente | Presente - 139; Ausente - 294 |
| 'Antipsicóticos' | 1 – Presente ; 0 – Ausente | Presente - 20; Ausente - 414 |
| 'Antidepressores' | 1 – Presente ; 0 – Ausente | Presente - 39; Ausente- 395 |
| 'Anti-hipertensores' | 1 – Presente ; 0 – Ausente | Presente - 98; Ausente - 336 |
| 'Ansioliticos' | 1 – Presente ; 0 – Ausente | Presente - 86; Ausente - 348 |
| 'Analgésicos ' | 1 – Presente ; 0 – Ausente | Presente - 14; Ausente - 420 |
| 'Anticoagulantes ' | 1 – Presente ; 0 – Ausente | Presente - 58; Ausente - 376 |
| 'Corticosteroides' | 1 – Presente ; 0 – Ausente | Presente - 25; Ausente - 409 |
| 'Digitálicos' | 1 – Presente ; 0 – Ausente | Presente - 28; Ausente - 406 |
| ‘Outros Med’ | 1 – Presente ; 0 – Ausente | Presente - 18; Ausente - 416 |
| 'Alcoolico' | 1 – Sim ; 0 – Não | Sim - 25 ; Não - 409 |
| 'Delirium' | 1 – Sim ; 0 – Não | Sim - 98 ; Não - 338 |

Já para as variáveis que apresentavam mais do que duas instâncias, como o caso da ‘Proveniência’ e ‘GrupoDiagnostico’, o procedimento foi muito semelhante ao anterior, com a diferença que foi gerada uma coluna para cada instancia, não se eliminado nenhuma.

A variável ‘Proveniencia’ apresenta informação relativa à proveniência do doente aquando da sua entrada no SU. Para esta variável foram mencionados quatro locais de proveniência, sendo eles casa, lar, outro hospital (inter-hospitalar) ou doente vindo do próprio hospital (intra-hospitalar).

O grupo de diagnóstico, representado pela variável ‘GrupoDiagnostico’, representa os serviços oferecidos pelo hospital, estes que definem a organização hospitalar. Sabendo que a medicina é um campo vasto, que se propõe a estudar a anatomia, funcionamento e interações dos órgãos, sistemas e tecidos que compõe o corpo humano, à medida que se foram aprofundando os conhecimentos nestes campos foram sendo criadas novas especialidades. Esta divisão de conhecimentos por especialidade permitiu dar continuidade às descobertas e aprofundar conhecimentos, para assim prevenir e tratar doenças de forma mais acertada. Neste sentido, as especialidades que são abordadas neste projeto estão relacionadas com a causa que levou determinado doente a recorrer ao SU. Nesta variável, estão incluídas as seguintes especialidades: cardiovascular, gastrointestinal, neurológico, respiratório, geniturinário, musculosquelético, toxicidade e drogas, outros e hemato-oncologico. Antes da realização desta codificação verificou-se que a variável ‘Toxicidade de Drogas’ apresentava apenas 9 entradas, por isso decidiu-se adicionar esta informação, na variável ‘Outros’, excluindo-se a variável referida em primeiro lugar.

De modo a compilar a informação descrita anteriormente, foi construída a Tabela 15, onde se apresentam as variáveis iniciais e as respetivas variáveis após a codificação. E, como é possível observar pela Tabela 15, cada individuo só pertence a uma categoria após a codificação, perfazendo sempre um total de 434 entradas para o conjunto de variáveis codificadas de cada categoria.

Tabela 15 - Codificação das variáveis categóricas com mais do que duas instâncias

|  |  |  |  |
| --- | --- | --- | --- |
| **Variável inicial** | **Variável após codificação** | **Numero de entradas** | **Total** |
| ‘Proveniencia’ | ‘Proveniencia\_Casa’ | 281 | 434 |
| ‘Proveniencia\_InterHosp’ | 39 |
| ‘Proveniencia\_IntraHosp’ | 102 |
| ‘Proveniencia\_Lar’ | 12 |
| ‘Local\_SU’ | ‘Local\_SU\_AMBUL’ | 4 | 434 |
| ‘Local\_SU\_UCISU’ | 131 |
| ‘Local\_SU\_UDC1’ | 128 |
| ‘Local\_SU\_UDC2’ | 171 |
| ‘GrupoDiag’ | ‘GrupoDiagn\_Neurológico’ | 47 | 434 |
| ‘GrupoDiagn\_Cardiovascular’ | 67 |
| ‘GrupoDiagn\_Gastrointestinal’ | 75 |
| ‘GrupoDiagn\_Respiratório’ | 68 |
| ‘GrupoDiagn\_Geniturinário’ | 69 |
| ‘GrupoDiagn\_Musculosquelético’ | 26 |
| ‘GrupoDiagn\_Outro’ | 70 |
| ‘GrupoDiagn\_Hemato-Oncológico’ | 12 |

Optou-se por este tipo de codificação, pois como para estas variáveis categóricas não existe uma relação ordinal, a codificação inteira podia não ser suficiente ou mesmo enganadora para o modelo. Assim, desta forma, permitir-se-ia que o modelo assumisse uma ordenação natural entre categorias o que não é correto.

Já para a variável ‘Local\_SU’, optou-se pela codificação ordinal. Uma vez que a variável representa a gravidade do estado do doente aquando da sua chegada ao SU. Sendo que, para esta variável são apresentados quatros estados possíveis: ambulatório (AMBUL), para casos de baixa complexidade e que não põe em risco imediato a vida do paciente; Unidade de Cuidados Intermédios do Serviço de Urgência (UCISU), para o caso de doentes que necessitam de cuidados mais diferenciados e de maior vigilância; ‘UDC1’ representa a triagem de Manchester para os indivíduos que foi atribuída a pulseira amarela, significando que são doentes urgentes; e por fim a ‘UCD2’, representando os doentes com a atribuição de pulseira laranja, ou seja doentes muito urgentes.

Como estes valores aparentam uma relação entre si, pensou-se em transforma-los numa escala de gravidade, traduzida na Tabela 16. Para a realização desta codificação usou-se a função *OrdinalEncoder()* que atribui valores numéricos a cada valor da variável ‘Local\_SU’.

Tabela 16 - Transformação dos valores da 'Local\_SU' numa escala de gravidade e posterior codificação

|  |  |  |
| --- | --- | --- |
| **Valores da variável ‘Local\_SU’** | **Escala de gravidade** | **Após codificação** |
| AMBUL | Sem gravidade | 0 |
| UCISU | Pouco grave | 1 |
| UDC1 | Grave | 2 |
| UDC2 | Muito Grave | 3 |

### Normalização dos dados

A base de dados é composta por 13 variáveis numéricas que incluem diferentes escalas de medida, sendo por este motivo difíceis de comparar. Neste sentido, será feita uma breve análise ao conteúdo destas variáveis numéricas e uma posterior transformação para que possam ser comparadas.

Começando pela variável ‘Idade’, esta apresenta-se medida em anos completos, até à data de recolha dos dados, exibindo valores que variam entre os 18 e 100 anos. A idade é um fator muito importante no diagnóstico de *delirium*, pois a maioria das doenças do foro psicológico tendem a acontecer em idades mais avançadas. E, embora o *delirium* possa ocorrer em qualquer idade, ocorre mais frequentemente em doentes idosos e com um estado mental comprometido (Maldonado, 2017).

Em seguida tem-se a variável 'Interna\_Dias', que se refere ao tempo de permanência no SU, e tal como já referido, esta coluna sofreu um arranjo na unidade de medida temporal. Tendo-se optado por escolher o tempo de permanência na urgência codificado em dias, verificando-se que os valores desta variável, após conversão de horas para dias, variam desde 0.083 a 12 dias.

Posteriormente, segue-se a variável ‘SIRS’ que contém informação relativa à quantidade de critérios SIRS (*Systemic Inflammatory Response Syndrome*) presentes. Clinicamente, a SIRS é identificada pela ocorrência de pelo menos dois dos seguintes critérios: febre > 38,0°C ou hipotermia <36,0°C; taquicardia > 90 batimentos/minuto; taquipneia > 20 respirações/minuto; alteração na contagem de leucócitos sanguíneos (Bone et al., 1992). Os valores possíveis desta variável variam desde 0 até 4, significando que 0 não apresentou qualquer critério e 4 apresentou todos os critérios .

Por fim, são apresentadas as variáveis relativas às análises clinicas recolhidas na data de entrada do individuo no SU. Como já referido, os dados laboratoriais existentes foram atualizados e corrigidos através de uma consulta manual executada no sistema informático do hospital. Com posterior registo dos valores na folha de *Excel* existente. A recolha da informação relativa aos dados metabólicos pareceu importante, na medida em que, já outros estudos foram realizados e comprovaram a importância destas variáveis. Tal como referido no subcapítulo 4.2.2, onde são introduzidos alguns fatores precipitantes, é mencionado, segundo alguns estudos, que as alterações metabólicas parecem interferir no desenvolvimento de *delirium*. Assim, no conjunto de dados recolhidos encontram-se valores relativos aos níveis sanguíneos de glicose, sódio, creatinina, ureia e PCR. Bem como também são apresentados valores relativos à gasometria nomeadamente o pH sanguíneo, cálcio ionizado, pressão parcial de oxigénio (PO2), pressão parcial de dióxido de carbono (PCO2) e o ião bicarbonato (HCO3). A variação dos valores destas variáveis, bem como as unidades de medida podem ser consultadas na Tabela 17.

Pelos motivos referidos supra, procedeu-se ao método de normalização destas variáveis. A normalização é um redimensionamento dos dados a partir dos dados originais para que todos os valores se situem num intervalo fixo, geralmente de 0 a 1. Este processo facilita a comparação de valores que se encontram medidos em diferentes escalas. Como é possível verificar através da observação da Tabela 17, os dados apresentam diferentes escalas de medidas. E, portanto, para que estas diferenças não sejam tão notórias recorreu-se à normalização dos dados. Um valor pode ser normalizado através da seguinte expressão:

Onde os valores mínimo e máximo dizem respeito ao valor x a ser normalizado. É possível normalizar os dados através da função *MinMaxScaler*, da biblioteca *scikit-learn*, que por defeito transforma as variáveis numa escala de [0,1].

Tabela 17 – Variáveis numéricas e os respetivos valores de referência e intervalos de valores ocorridos na base de dados

|  |  |  |
| --- | --- | --- |
| **Variável** | **Valores de referência** | **Intervalo de valores registados** |
| ‘Idade’ | - | 18-100 anos |
| 'Interna\_Dias' | - | 0.083-12 dias |
| 'SIRS' | - | 0-4 |
| 'Glicose' | 90 -130 mg/dL | 41-1000 mg/dL |
| 'Sodio' | 135-146 mEq/L | 42-151 mEq/L |
| 'Ureia' | 19-49 mEq/L | 4-275 mEq/L |
| 'Creatinina' | 0.6-1.2 mg/dL | 0.1-19.5 mg/dL |
| 'PCR' | <5 mg/L | 2.3-499 mg/L |
| 'pH' | 7.35-7.45 | 7.026-7.625 |
| 'Ca\_ionizado' | 1.15-1.35 mmol/L | 0.84-1.37 mmol/L |
| 'pCO2' | 33-45 mm Hg | 13.2-121.3 mm Hg |
| 'pO2' | 75-105 mm Hg | 34.1-178.1 mm Hg |
| 'HCO3' | 22-28 mEq/L | 7.40-39.1 mEq/L |

# Modelação, Apresentação e discussão dos resultados

Neste capítulo estão descritos os passos efetuados para a modelação do problema. Aqui, são apresentados os passos efetuados na construção do algoritmo de RL e RF para a predição do *delirium*. O objetivo é modelar os dois modelos e no final escolher aquele em que se obtiverem melhores resultados. Esta modelação terá como finalidade a utilização por parte de profissionais de saúde, para que seja possível prever com antecedência o desenvolvimento de *delirium* nos pacientes admitidos em SU.

## Desbalanceamento dos dados

Um dos primeiros problemas encontrados na fase de modelação centrou-se na obtenção de resultados de previsão com apenas uma categoria, o ‘Não’. E, embora até se tenha conseguido resultados de cerca de 80% nas medidas de sensibilidade, especificidade e taxa de acertos, verificou-se que os dados de treino estavam enviesados, e a sua utilização estava a originar interpretações erradas. Pois, como nos dados de treino, existia uma maior quantidade de dados para a categoria ‘Não’, o modelo interpretou que se classificasse todas as previsões com a classe maioritária, conseguiria boas taxas de acerto, concluindo-se desta forma que o modelo não estaria apto para detetar casos de *delirium*. Uma forma de contornar este problema, passou pelo desbalanceamento dos dados.

Os dados desbalanceados definem-se pela pequena incidência de uma categoria num conjunto de dados (classe minoritária) em comparação com as restantes categorias (Hoang, Bouzerdoum, & Lam, 2009). O desbalanceamento nos dados mostra-se presente em diversos setores e campos do conhecimento, não sendo incomum serem encontrados regularmente em contextos variados. Como, por exemplo, nos dados referentes ao diagnóstico de cancro, em que a maioria das pessoas que realizam a experiência apresentam um resultado negativo (Vluymans, 2019). O mesmo sucedeu com os dados relativos à ocorrência ou não de *delirium*, nos quais, a categoria ‘Sim’ apresentou uma quantidade bastante mais baixa que a categoria oposta. A ocorrência deste desbalanceamento nos dados pode causar problemas não só, na construção do modelos de ML, como também, na geração de previsões (Vluymans, 2019).Exemplificando, tal como é possível verificar pela Figura 6, no gráfico A, a variável resposta (‘*Delirium’*), possui mais entradas para o valor ‘Não’, correspondendo a cerca de 77%. E, supondo que o modelo apenas efetua classificações com o valor ‘Não’ para todas as previsões, obter-se-á uma taxa de acertos de cerca de 77%. O que causaria prejuízo, pois, se o modelo não diferenciar com sucesso os diagnósticos positivos de *delirium*, estar-se-ia a considerar que nenhum paciente teria este distúrbio.

Chart, waterfall chart

Description automatically generated

Figura 6 - Esquema representativo da divisão dos dados e posterior balanceamento dos dados de treino. A - Contagem de categorias da variável ‘Delirium’ antes da divisão dos dados. B - Contagem de categorias da variável ‘Delirium, para os dados de treino. C - Contagem de categorias da variável ‘Delirium, para os dados de teste. D - Contagem de categorias da variável ‘Delirium, para os dados de treino após o balanceamento dos dados.

Uma forma de contornar este problema de diferença na proporção entre as categorias, consistiu em manipular a quantidade de dados utilizados pelo modelo de ML, tentando-se igualar o número de observações entre as classes. De salientar, tal como é possível verificar pelos gráficos B e C da Figura 6, que esta operação apenas se realizará no conjunto de dados destinados ao treino do modelo. Pois, nos dados de teste, não deverá constar nenhuma das informações contidas nos dados de treino, garantindo-se, desta forma, que nenhuma das observações criadas sinteticamente, consta nos dados de teste. Para a execução da tarefa de criação de observações sintéticas, utilizou-se a função *ADASYN()* disponível na biblioteca *imblearn.* A ideia principal do algoritmo ADASYN *(﻿Adaptive Synthetic Sampling)* consiste em utilizar uma distribuição de densidade como critério para decidir o número de amostras sintéticas a ser geradas para cada exemplo minoritário, alterando adaptativamente os pesos dos diferentes exemplos minoritários (Vluymans, 2019). Portanto, com esta abordagem são criadas sinteticamente novas observações da classe minoritária, com o objetivo de igualar a proporção das categorias. Tal como é possível verificar pelo gráfico D da Figura 6, foram criadas sinteticamente 116 novas linhas para a categoria ‘Sim’, no conjunto de dados destinados ao treino do modelo

A escolha desta estratégia em detrimento de outras existentes centrou-se no método utilizado por cada método, no momento da criação de novas linhas. Uma vez que os dados apresentam poucas linhas, a estratégia de *UnderSampling*, foi descartada, devido ao facto de serem excluídas linhas da classe maioritária até perfazer o número de linhas da categoria minoritária. Por outro lado, na estratégia de *OverSampling* foram encontradas várias alternativas, estas diferenciadas no método de criação de novas linhas. A técnica *RandomOverSampler* consiste na replica integral de linhas da classe minoritária até se atingir o número de linhas da classe maioritária. E, embora aumente o número de dados, não dá qualquer nova informação ou variação ao modelo de ML. Já a técnica SMOTE (*Synthetic Minority Oversampling Technique*), utiliza o algoritmo de KNN para gerar amostras sintéticas baseado nas similaridades entre amostras no espaço n-dimensional de variáveis (Chawla, Bowyer, Hall, & Kegelmeyer, 2002). Uma desvantagem desta abordagem é que os exemplos sintéticos são criados sem considerar a classe maioritária, resultando possivelmente em exemplos ambíguos caso haja uma sobreposição nas classes. A partir desta estratégia, foram desenvolvidas outras variações, e uma delas é a ADASYN, que após encontrar os k- vizinhos mais próximos, desenhar a linha entre os vizinhos e gerar pontos aleatórios, é adicionado um pequeno valor aleatório aos pontos. Ou seja, em vez da amostra estar linearmente correlacionada com o pai, têm um pouco mais de variação.

## Divisão dos dados

A aprendizagem supervisionada consiste em criar modelos que mapeiem, especificamente, para as entradas dadas (variáveis independentes, ou preditores) as respetivas saídas (variáveis dependentes, ou respostas). E, para tal ser possível de realizar, é necessário dividir os dados de forma a que uma parte possa ser utilizada no treino, e outra para teste do modelo. Para a realização desta tarefa, utilizou-se a função *train\_test\_split()*, que permite dividir o conjunto de dados antes de o utilizar. Esta divisão efetuada pela função, encontra-se esquematizada na Figura 7, através da qual é possível perceber que as células coloridas a verde correspondem às variáveis ‘X’ e ‘y’ utilizadas para o treino do modelo, e as restantes sem coloração serão utilizadas para a realização de avaliações ao modelo. Pois, para que se consiga efetuar uma avaliação imparcial do desempenho preditivo do modelo, é necessário utilizar dados diferentes para o treino e para o teste do modelo. Isto significa que não se pode avaliar o desempenho do modelo com os mesmos dados que se utilizaram para treino, pois desta forma o modelo terá facilidade em prever o resultado, uma vez que os dados já são do conhecimento do modelo.

Nesta função, também é possível definir o argumento ‘*test\_size*’ que determina o tamanho do conjunto de teste. A percentagem padrão estipulada para este argumento é de 25%, no entanto, este valor poder ser otimizado consoante a tipologia dos dados.

Table

Description automatically generated

Figura 7 - Esquema de divisão dos dados para treino e teste do modelo

## Feature Selection / Seleção de variáveis do modelo

Para a seleção de variáveis do modelo recorreu-se à Feature Selection

A Feature Selection deve ser feita depois da etapa de pré-processamento dos dados. O objetivo desta etapa consiste em selecionar as melhores variáveis como possíveis variáveis preditoras. A etapa de feature selection pode ajudar a reduzir o overfitting, aumentar a acurácia do modelo e uma redução no tempo de treino.

### Tipos de algoritmos e métodos

Filter Methods : Métodos de seleção que utiliza medidas estatísticas para atribuir um score para cada feature. As features são classificadas pelo score para serem mantidas ou removidas do modelo. Normalmente são usados testes univariados que consideram a independência da variável com a variável alvo. Exemplo: chi squared, scores com coeficiente de correlação.

Wrapper Methods : Métodos de seleção que selecionam um conjunto de variáveis, onde são preparadas diferentes combinações e de seguida são avaliadas e comparadas. Um modelo preditivo é usado para avaliar a combinação de features a atribuir um score baseado em uma acurácia de modelo. Exemplo: algoritmo RFE

Embedded Methods : Métodos Embedded aprendem quais as variáveis que melhor contribuem para a acurácia do modelo no momento de construção do modelo. Exemplo: métodos de penalização, algoritmos Lasso, Elastic NEt e Ridge Regression.

Antes de se proceder à construção do modelo, foi necessário escolher o melhor conjunto de varáveis independentes a introduzir no modelo preditivo, eliminando aquelas que não apresentam variabilidade. Para isso, foram utilizadas técnicas estatísticas e exploratórias que permitiram eliminar variáveis que à priori não iriam contribuir matematicamente para o modelo.

O primeiro passo, passou por realizar um procedimento univariado variável a variável, tendo-se eliminado aquelas que não eram significativas, ou seja, que apresentavam um p-value acima de 0,25. Para tal, utilizou-se o *SelectKBest*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Variável | Valor | Erro |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |

Após esta seleção das variáveis candidatas a serem implementadas no modelo de regressão logística, procedeu-se à construção de possíveis modelos multivariados.

## Treino do modelo Random Forest

O algoritmo RF, tal como o seu nome implica, consiste num grande número de árvores de decisão individuais que funcionam como um conjunto. Cada árvore individual no RF emite uma previsão de classe e a classe com mais votos torna-se a previsão do modelo, por isso o facto de aumentar o número de árvores pode aumentar a precisão do resultado. As AD são muito sensíveis aos dados sobre os quais são treinadas, pois pequenas alterações ao conjunto de treino podem resultar em estruturas de árvores significativamente diferentes. E, o RF tira partido disto, permitindo que cada árvore individual recolha amostras aleatórias do conjunto de dados com substituição, resultando em árvores diferentes. Este processo é conhecido como *bagging* (Bootstrap agregation).

A criação e treino do modelo é um processo simples, bastando utilizar a biblioteca Scikit-learn, na qual é possível utilizar a função *RandomForestClassifier()*. Nesta função é possível alterar os parâmetros, e de acordo com os resultados gerados, selecionar aqueles em que se obtenham melhores resultados. De modo a compreender o significado de cada parâmetro construiu-se a Tabela 18, onde estão descritos os parâmetros e o seu respetivo significado.

Tabela 18 - Parâmetros da função RandomForestClassifier() e respetivo significado

|  |  |  |
| --- | --- | --- |
| **Parâmetro** | **Valor padrão** | **Significado** |
| n\_estimators | 100 | Número de árvores |
| criterion | gini | Mede a qualidade de uma divisão. Os critérios suportados são "gini" e “entropy” |
| max\_depth | None | A profundidade máxima da árvore. Se não for definido nenhum, então os nós são expandidos até todas as folhas serem puras ou até todas as folhas conterem menos do que amostras de “min\_samples\_split”. |
| min\_samples\_split | 2 | O número mínimo de amostras necessárias para dividir um nodo interno. |
| min\_samples\_leaf | 1 | O número mínimo de amostras necessárias a estar num nodo. |
| max\_features | “auto” | O número de características a considerar quando se procura a melhor divisão. Valores possíveis: “auto”, “sqrt”, “log2” |
| bootstrap | True | São utilizadas amostras na construção de árvores. Se *False*, é utilizado o conjunto de dados total. |
| oob\_score | False | Se usar amostras fora do saco para estimar a pontuação de generalização. Apenas disponível se “bootstrap=True”. |
| random\_state | None | Controla tanto a aleatoriedade do bootstrapping das amostras utilizadas na construção de árvores como a amostragem das características a considerar quando se procura a melhor divisão em cada nó. |
| n\_jobs | None | O número de trabalhos a executar em paralelo.  -1 significa usar todas as CPU |

A conjugação destes valores serão estudados no momento da realização do tuning do modelo na secção X.

A entropia é uma métrica para calcular a incerteza. O ganho de informação é uma medida de como a incerteza na variável alvo é reduzida, dado um conjunto de variáveis independentes.

## Previsões

Após o treino do modelo, este está pronto para alguma análise. Nesta etapa, os dados que foram reservados para testar o modelo são utilizados para executar as previsões. Os dados de teste (X\_test) são submetidos ao modelo de previsão e os resultados produzidos são avaliados. Para a realização desta avaliação são comparados os dados produzidos pelo modelo com os dados reais (y\_test) que se reservaram no início da modelação.

## Avaliação da performance do modelo

Os modelos de classificação devem ser avaliados antes de serem adotados em contexto real, pois, caso o classificador esteja mal calibrado pode induzir em erro os profissionais de saúde e causar danos no paciente alvo do estudo. Assim, de forma a minimizar a ocorrência desta situação, o modelo é submetido a métricas que permitem avaliar o desempenho do modelo. Esta seleção das métricas, pode influenciar a forma como o desempenho dos algoritmos de ML são medidos e comparados. Pois, influenciam a forma como se pondera a importância das diferentes características nos resultados e também interferem na escolha final sobre o algoritmo a selecionar.

Os resultados da previsão adquiridos na etapa anterior são comparados, através dos resultados reais que deveriam ter sido obtidos. Para a avaliação do desempenho do modelo podem ser calculadas várias métricas de avaliação. Além disso, em contexto médico, medidas como a sensibilidade (a capacidade de prever *delirium* quando a condição está presente), a especificidade (a capacidade de prever a ausência da *delirium* quando a condição não está presente), e também a área sob a curva ROC, são características chave que devem ser consideradas quando se pretende comparar classificadores (Maroco et al., 2011). De modo a calcular estas métricas é necessário, em primeiro lugar, representar os diferentes tipos de acertos ou erros cometidos. E, a representação desses diferentes tipos é apresentada de seguida:

* Verdadeiro Positivo (VP): modelo classifica corretamente a classe positiva ‘1’;
* Falso Positivo (FP): modelo classifica incorretamente a classe positiva ‘1’;
* Verdadeiro Negativo (VN): modelo classifica corretamente a classe negativa ‘0’;
* Falso Negativo (FN): modelo classifica incorretamente a classe negativa ‘0’.

Começando pela taxa de acertos *(accuracy)* do modelo, esta é caracterizada pelo número de previsões corretas feitas como um rácio de todas as previsões efetuadas. Esta medida deve ser usada em dados com a mesma proporção de exemplos para cada classe.

A sensibilidade (*recall*) mede uma proporção dos valores que são de facto positivos e que foram preditos corretamente. É uma proporção entre os verdadeiros positivos (VP) e os falsos negativos (FN), conforme apresentado na fórmula.

Já a especificidade (*precision*) mede a proporção de predições positivas que estão corretas, ou seja, quão bem o modelo prevê os valores positivos. É uma proporção entre os verdadeiros positivos (VP) e os falso positivos (FP), conforme apresentado na fórmula. A sensibilidade e a especificidade, estão entre os principais indicadores para avaliação de modelos, isso acontece porque ambos comparam VP com os erros cometidos. E conforme vimos anteriormente, dependendo do problema, o objetivo será sempre a minimização ou de FP ou de FN.

A métrica F-1 é uma média harmónica entre a especificidade e a sensibilidade. Portanto, esta métrica tem em consideração tanto os FP como os FN, sendo normalmente mais útil do que a taxa de acertos, no caso de os dados apresentarem uma distribuição da classe desbalanceada. Quando se obtém um F1-Score baixo, indica que ou a especificidade ou a sensibilidade está baixa.

|  |  |  |  |
| --- | --- | --- | --- |
| **Taxa de acertos** | **Especificidade** | **Sensibilidade** | **F1-score** |
| 0,80 | 0,80 | 0,80 | 0,80 |
| 0,90 | 0,90 | 0,90 | 0,90 |
| 0,92 | 0,93 | 0,92 | 0,92 |

Area Under Roc Curve (AUROC)

Outra métrica para avaliar a performance de modelos é a AUROC, geralmente tratada como AUC (Area Under The Curve), que é simplesmente a área sob a curva ROC, como indicado ao lado.

Podemos compreender esse indicador como a capacidade do modelo de categorizar corretamente os dados, com valor máximo em 1 e mínimo em 0.

A matriz de confusão nos permite analisar o desempenho do modelo de forma mais cuidadosa, especialmente no que concerne ao problema da acurácia discutido anteriormente.

## Feature importance

A seleção de características é o processo de reduzir a variável de entrada do modelo, utilizando apenas dados relevantes e livrando-se do ruído nos dados.

Filter method – Coeficiente de Pearson, Chi-quadrado, ANOVA

Wrapper method – RFE, algoritmos genéticos

Intrinsic method – Lasso regularization (regreesion models), decision trees ()

## Tuning dos parâmetros

## Avaliação da performance do modelo

Os resultados da previsão adquiridos na etapa anterior são comparados utilizando os resultados reais que deveriam ter sido obtidos. Várias métricas de avaliação são geradas para calcular o desempenho do modelo.

A matriz de confusão nos permite analisar o desempenho do nosso modelo de forma mais cuidadosa, especialmente no que concerne ao problema da acurácia discutido anteriormente.

Cada quadrante da matriz de confusão representa uma contagem referente aos tipos de acertos ou erros cometidos.

True Positives (TP): modelo prediz corretamente a classe positiva ‘1’;

False Positives (FP): modelo prediz incorretamente a classe positiva ‘1’;

True Negatives (TN): modelo prediz corretamente a classe negativa ‘0’;

False Negatives (FN): modelo prediz incorretamente a classe negativa ‘0’;

Recall, True Positive Rate (TPR) ou Sensitividade

Mede uma proporção dos valores que são de facto positivos e que foram preditos corretamente.

É uma proporção entre os positivos verdadeiros (TP) e os falsos negativos (FN), conforme apresentado na fórmula.

Precision

Mede a proporção de predições positivas que estão corretas, ou seja, quão bem o modelo predisse os valores positivos. É uma proporção entre os positivos verdadeiros (TP) e os falso positivos (FP), conforme apresentado na fórmula.

Precision e Recall estão entre os principais indicadores para avaliação de modelos, isso acontece porque ambos comparam TP com os erros cometidos. E conforme vimos anteriormente, dependendo do problema desejamos minimizar apenas um desses erros, FP ou FN.

Area Under Roc Curve (AUROC)

Outra métrica para avaliar a performance de modelos é a AUROC, geralmente tratada como AUC (Area Under The Curve), que é simplesmente a área sob a curva ROC, como indicado ao lado.

Podemos compreender esse indicador como a habilidade do modelo de categorizar corretamente os dados, com valor máximo em 1 e mínimo em 0.

## Regressão Logística

Um dos primeiros passos a efetuar, nos modelos de RL é a seleção de variáveis independentes a introduzir no modelo preditivo. Para tal, existem técnicas estatísticas que permitem a eliminação de variáveis menos contributivas para o modelo. Uma variável dependente é uma medida que dependerá do valor de outra variável ou de um conjunto de variáveis. Neste sentido, a variável dependente é dicotómica, atribuindo-se o valor 1 ao acontecimento de ter *delirium* e 0 ao acontecimento complementar, ou seja, esta variável define se o paciente desenvolveu ou não a síndrome

### Seleção do melhor modelo

Após a seleção das variáveis candidatas a serem implementadas no modelo de regressão logística, procedeu-se à construção de possíveis modelos multivariados. Ao longo desta secção será explicado o motivo da escolha do conjunto de variáveis independentes, integrantes de cada modelo . E também será feita uma análise dos resultados relativos à qualidade de ajuste e à capacidade de desempenho, de forma a escolher o melhor modelo a ser utilizado como classificador.

## Random Forest

Nesta secção será feita uma avaliação dos resultados obtidos pelos modelos para a predição do *delirium*, e será elaborada uma comparação às métricas dos dois modelos.

## Aplicação *web*

Como definido no subcapítulo anterior, o modelo de previsão que apresenta uma melhor capacidade preditiva é o que utiliza o algoritmo de RF. Como tal, para facilitar a utilização deste modelo preditivo contrui-se uma aplicação web. Este tipo formato permite facilitar a predição em tempo real, de forma rápida e bastante assertiva a deteção de delirium em doentes admitidos em SU. Esta aplicação pode ser consultada através do seguinte link:

### Página Inicial

### Previsão

### Sobre

# Conclusões e Trabalho futuro

O *delirium* é uma síndrome grave, muito prevalente no ambiente hospitalar, tendo como causa uma patologia ou condição clínica potencialmente reversível, mas altamente subdiagnosticada e negligenciada.

Os idosos são o grupo etário mais afetado, devendo ser alvo de rastreio do risco de *delirium* para permitir ações preventivas. Fatores de risco importantes incluem, sobretudo, a demência, mas também pluripatologia e outros fatores incluídos em modelos preditivos.

Para finalizar, sendo que o *delirium* em ambiente hospitalar é frequentemente de natureza multifatorial, deve ser seguida uma abordagem preventiva abrangente, no sentido de se tentar minimizar os vários fatores de risco. A redução da sua incidência em ambiente hospitalar deve ser considerada como um indicador de qualidade na prestação de serviços de saúde e também um marcador na melhoria da qualidade de vida dos pacientes.

# Bibliografia

Adamis, D., Sharma, N., Whelan, P. J. P., & MacDonald, A. J. D. (2010). Delirium scales: A review of current evidence. *Aging and Mental Health*, *14*(5), 543–555. https://doi.org/10.1080/13607860903421011

Adamis, D., Treloar, A., Martin, F. C., & Macdonald, A. J. D. (2007). A brief review of the history of delirium as a mental disorder. *History of Psychiatry*, *18*(4), 459–469. https://doi.org/10.1177/0957154X07076467

Alagiakrishnan, K., & Wiens, C. A. (2004). An approach to drug induced delirium in the elderly. *Postgraduate Medical Journal*, *80*(945), 388–393. https://doi.org/10.1136/pgmj.2003.017236

Ali, J., Khan, R., Ahmad, N., & Maqsood, I. (2012). Random Forests and Decision Trees. *International Journal of Computer Science Issues*, *9*(5), 272–278.

Ali, M. A., Hashmi, M., Ahmed, W., Raza, S. A., Khan, M. F., & Salim, B. (2021). Incidence and risk factors of delirium in surgical intensive care unit. *Trauma Surgery and Acute Care Open*, *6*(1). https://doi.org/10.1136/tsaco-2020-000564

Alsuliman, T., Humaidan, D., & Sliman, L. (2020). Machine learning and artificial intelligence in the service of medicine: Necessity or potentiality? *Current Research in Translational Medicine*, *68*(4), 245–251. https://doi.org/10.1016/j.retram.2020.01.002

American Psychiatric Association. (1980). *Diagnostic and Statistical Manual of Mental Disorders - III*. Washington.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders - DSM-5*. *Pediatria Integral* (fifth, Vol. 17).

Associação Portuguesa de Nutrição. (2018). *Dislipidemias: Caracterização e Tratamento nutricional*. Associação Portuguesa de Nutrição. Retrieved from http://www.jstage.jst.go.jp/article/pjab/86/5/86\_5\_484/\_article

Basu, S., Faghmous, J. H., & Doupe, P. (2020). Machine learning methods for precision medicine research designed to reduce health disparities: A structured tutorial. *Ethnicity and Disease*, *30*, 217–228. https://doi.org/10.18865/ed.30.S1.217

Bone, R. C., Balk, R. A., Cerra, F. B., Dellinger, R. P., Fein, A. M., Knaus, W. A., … Sibbald, W. J. (1992). Definitions for Sepsis and Organ Failure and Guidelines for the Use of Innovative Therapies in Sepsis. *Chest*, *101*(6), 1644–1655. https://doi.org/10.1378/chest.101.6.1644

Bourgeois, J. A., Hategan, A., & Losier, B. (2014). Delirium in the hospital: Emphasis on the management of geriatric patients. *Current Psychiatry*, *13*(8), 29–42.

Braga, A. (2000). *Curvas ROC: aspectos funcionais e aplicações*. *Tese*. University of Minho. Retrieved from http://repositorium.sdum.uminho.pt/bitstream/1822/195/1/tese\_doutACB.pdf

Braga, A. C., & Carneiro, P. (2016). Development and Validation of a Logistic Regression Model to Estimate the Risk of WMSDs in Portuguese Home Care Nurses. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* (Vol. 9786, pp. 97–109). https://doi.org/10.1007/978-3-319-42085-1\_8

Braga, A. de P., Ludermir, T. B., & Carvalho, A. C. P. de L. F. (2000). *Redes Neurais Artificiais: Teoria e Aplicações*. Rio de Janeiro: LTC - Livros técnicos e científicos editora S.A.

Breiman, L. (2001). Random forests. *Machine Learning*, *45*(1), 5–32.

Brenner, D. M., & Lacy, B. E. (2021). Antispasmodics for Chronic Abdominal Pain. *American Journal of Gastroenterology*, *Publish Ah*, 1587–1600. https://doi.org/10.14309/ajg.0000000000001266

Brownlee, J. (2017). How to One Hot Encode Sequence Data in Python. Retrieved November 24, 2021, from https://machinelearningmastery.com/how-to-one-hot-encode-sequence-data-in-python/

Bruha, I., & Berka, P. (2000). Discretization and Fuzzification of Numerical Attributes in Attribute-Based Learning. In *Fuzzy Systems in Medicine* (Vol. 41, pp. 112–138). https://doi.org/10.1007/978-3-7908-1859-8\_6

Brunton, L. L., Chabner, B. A., & Knollmann, B. C. (2011). *Goodman & Gilman’s pharmacological basis of therapeutics* (twelfth). New York, NY: McGraw-Hill Medical.

Burges, Christopher, J. C. (1998). A Tutorial on Support Vector Machines for Pattern Recognition. *Data Mining and Knowledge Discovery*, *2*, 121–167. https://doi.org/https://doi.org/10.1023/A:1009715923555

Cacabelos, R. (2017). Parkinson’s Disease: From Pathogenesis to Pharmacogenomics. *International Journal of Molecular Sciences*, *18*(3), 551. https://doi.org/10.3390/ijms18030551

Cano-escalera, G., Besga, A., & Graña, M. (2021). Risk factors for prediction of delirium at hospital admittance. *Expert Systems*, *e12698*(December 2020), 1–10. https://doi.org/10.1111/exsy.12698

Casey, P., Cross, W., Mart, M. W. S., Baldwin, C., Riddell, K., & Dārziņš, P. (2019). Hospital discharge data under-reports delirium occurrence: results from a point prevalence survey of delirium in a major Australian health service. *Internal Medicine Journal*, *49*(3), 338–344. https://doi.org/10.1111/imj.14066

Cerejeira, J., Firmino, H., Vaz-Serra, A., & Mukaetova-Ladinska, E. B. (2010). The neuroinflammatory hypothesis of delirium. *Acta Neuropathologica*, *119*(6), 737–754. https://doi.org/10.1007/s00401-010-0674-1

Chapelle, O., Vapnik, V., Bousquet, O., & Mukherjee, S. (2002). Choosing Multiple Parameters for Support Vector Machines. *Machine Learning*, *46*, 131–159. https://doi.org/https://doi.org/10.1023/A:1012450327387

Chapman, P., Clinton, J., Kerber, R., Khabaza, T., Reinartz, T., Shearer, C., & Wirth, R. (2000). *CRISP-DM 1.0 Step-by-step data mining guide*. *SPSS inc*. Retrieved from http://www.crisp-dm.org/CRISPWP-0800.pdf

Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE: Synthetic Minority Over-sampling Technique. *Journal of Artificial Intelligence Research*, *16*(2), 321–357. https://doi.org/10.1613/jair.953

Cheon, S. Y., Koo, B.-N., Kim, S. Y., Kam, E. H., Nam, J., & Kim, E. J. (2021). Scopolamine promotes neuroinflammation and delirium-like neuropsychiatric disorder in mice. *Scientific Reports*, *11*(1), 8376. https://doi.org/10.1038/s41598-021-87790-y

Chrousos, G. P. (1993). Syndromes of Glucocorticoid Resistance. *Annals of Internal Medicine*, *119*(11), 1113. https://doi.org/10.7326/0003-4819-119-11-199312010-00009

Church, M. K., & Maurer, M. (2014). Antihistamines. In *History of Allergy* (Vol. 100, pp. 302–310). https://doi.org/10.1159/000359963

Cirbus, J., MacLullich, A. M. J., Noel, C., Ely, E. W., Chandrasekhar, R., & Han, J. H. (2019). Delirium etiology subtypes and their effect on six-month function and cognition in older emergency department patients. *International Psychogeriatrics*, *31*(2), 267–276. https://doi.org/10.1017/S1041610218000777

Clegg, A., & Young, J. B. (2011). Which medications to avoid in people at risk of delirium: A systematic review. *Age and Ageing*, *40*(1), 23–29. https://doi.org/10.1093/ageing/afq140

Conselho Nacional da Saúde. (2019). *Sem mais tempo a perder - Saúde mental em Portugal: um desafio para a próxima década*. *Conselho Nacional de Saúde*. Lisboa: CNS.

Corradi, J. P., Thompson, S., Mather, J. F., Waszynski, C. M., & Dicks, R. S. (2018). Prediction of Incident Delirium Using a Random Forest classifier. *Journal of Medical Systems*, *42*(12). https://doi.org/10.1007/s10916-018-1109-0

Davoudi, A., Ebadi, A., Rashidi, P., Ozrazgat-Baslanti, T., Bihorac, A., & Bursian, A. C. (2017). Delirium Prediction using Machine Learning Models on Predictive Electronic Health Records Data. *Proceedings - IEEE 17th International Symposium on Bioinformatics and Bioengineering, BIBE 2017*, 568–573. https://doi.org/10.1109/BIBE.2017.00014

De, J., & Wand, A. P. F. (2015). Delirium screening: A systematic review of delirium screening tools in hospitalized patients. *Gerontologist*, *55*(6), 1079–1099. https://doi.org/10.1093/geront/gnv100

De Morton, N. A., Jones, C. T., Keating, J. L., Berlowitz, D. J., MacGregor, L., Lim, W. K., … Brand, C. A. (2007). The effect of exercise on outcomes for hospitalised older acute medical patients: an individual patient data meta-analysis. *Age and Ageing*, *36*(2), 219–222. https://doi.org/10.1093/ageing/afl118

de Rooij, S. E., van Munster, B. C., Korevaar, J. C., & Levi, M. (2007). Cytokines and acute phase response in delirium. *Journal of Psychosomatic Research*, *62*(5), 521–525. https://doi.org/10.1016/j.jpsychores.2006.11.013

Delaney, J., Spevack, D., Doddamani, S., & Ostfeld, R. (2006). Clonidine-induced delirium. *International Journal of Cardiology*, *113*(2), 276–278. https://doi.org/10.1016/j.ijcard.2005.09.032

Desforges, J. F., & Lipowski, Z. J. (1989). Delirium in the Elderly Patient. *New England Journal of Medicine*, *320*(9), 578–582. https://doi.org/10.1056/NEJM198903023200907

DiFazio, C. A. (1989). Pharmacology of Narcotic Analgesics. *The Clinical Journal of Pain*, *5*, S5–S7. https://doi.org/10.1097/00002508-198903001-00004

Dreiseitl, S., & Ohno-Machado, L. (2002). Logistic regression and artificial neural network classification models: A methodology review. *Journal of Biomedical Informatics*, *35*(5–6), 352–359. https://doi.org/10.1016/S1532-0464(03)00034-0

Duan, K., Keerthi, S., & Poo, A. (2003). Evaluation of simple performance measures for tuning SVM hyper parameters. Technical report. *Neurocomputing*, *51*, 41–59. https://doi.org/https://doi.org/10.1016/S0925-2312(02)00601-X

Ely, E. W., Margolin, R., Francis, J., May, L., Truman, B., Dittus, R., … Inouye, S. K. (2001). Evaluation of delirium in critically ill patients: Validation of the Confusion Assessment Method for the intensive care unit (CAM-ICU). *Critical Care Medicine*, *29*(7), 1370–1379. https://doi.org/10.1097/00003246-200107000-00012

Ely, E. W., Truman, B., Shintani, A., Thomason, J. W. W., Wheeler, A. P., Gordon, S., … Bernard, G. R. (2003). Monitoring Sedation Status Over Time in ICU Patients. *JAMA*, *289*(22), 2983. https://doi.org/10.1001/jama.289.22.2983

Estivill-Castro, V. (2002). Why so many clustering algorithms - a position paper. *ACM SIGKDD Explorations Newsletter*, *4*(1), 65–75. https://doi.org/10.1145/568574.568575

Faceli, K., Lorena, A. C., Gama, J., & Carvalho, A. C. P. L. . de C. (2011). *Inteligência Artificial. Uma Abordagem de Aprendizado de Máquina (Em Portugues do Brasil)*. GEN: LTC. Retrieved from http://amazon.com/o/ASIN/8521618808/

Field, R. R., & Wall, M. H. (2013). Delirium: Past, present, and future. *Seminars in Cardiothoracic and Vascular Anesthesia*, *17*(3), 170–179. https://doi.org/10.1177/1089253213476957

Flacker, J. M., & Lipsitz, L. A. (1999). Neural Mechanisms of Delirium: Current Hypotheses and Evolving Concepts. *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences*, *54*(6), B239–B246. https://doi.org/10.1093/gerona/54.6.B239

Fonseca, C., Brito, D., Cernadas, R., Ferreira, J., Franco, F., Rodrigues, T., … Silva Cardoso, J. (2017). Pela melhoria do tratamento da insuficiência cardíaca em Portugal – documento de consenso. *Revista Portuguesa de Cardiologia*, *36*(1), 1–8. https://doi.org/10.1016/j.repc.2016.10.006

Francis, J. (1996). Drug-Induced Delirium : Diagnosis and treatment. *CNS Drugs*, *5*(2), 103–114.

Friedman, J. H. (2002). Stochastic gradient boosting. *Computational Statistics and Data Analysis*, *38*(4), 367–378. https://doi.org/10.1016/S0167-9473(01)00065-2

Gaudreau, J. D., Gagnon, P., Harel, F., Roy, M. A., & Tremblay, A. (2005). Psychoactive medications and risk of delirium in hospitalized cancer patients. *Journal of Clinical Oncology*, *23*(27), 6712–6718. https://doi.org/10.1200/JCO.2005.05.140

Gross, A. L., Jones, R. N., Habtemariam, D. A., Fong, T. G., Tommet, D., Quach, L., … Inouye, S. K. (2012). Delirium and long-term cognitive trajectory among persons with dementia. *Archives of Internal Medicine*, *172*(17), 1324–1331. https://doi.org/10.1001/archinternmed.2012.3203

Han, J. H., Eden, S., Shintani, A., Morandi, A., Schnelle, J., Dittus, R. S., … Ely, E. W. (2011). Delirium in older emergency department patients is an independent predictor of hospital length of stay. *Academic Emergency Medicine*, *18*(5), 451–457. https://doi.org/10.1111/j.1553-2712.2011.01065.x

Han, J., Kamber, M., & Pei, J. (2012). *Data mining: Concepts and Techniques*. *Elsevier*. Morgan Kaufmann.

Harrison, P., Cowen, P., Burns, T., & Fazel, M. (2018). *Shorter Oxford Textbook of Psychiatry* (seventh). Oxford: Oxford University Press.

Hatano, Y., Narumoto, J., Shibata, K., Matsuoka, T., Taniguchi, S., Hata, Y., … Fukui, K. (2013). White-matter hyperintensities predict delirium after cardiac surgery. *American Journal of Geriatric Psychiatry*, *21*(10), 938–945. https://doi.org/10.1016/j.jagp.2013.01.061

Haykin, S. (2009). *Neural Networks and Learning Machines* (3rd ed., Vol. 1–3). New Jersey: Pearson.

Hoang, G., Bouzerdoum, A., & Lam, S. (2009). Learning Pattern Classification Tasks with Imbalanced Data Sets. In *Pattern Recognition* (pp. 193–208). InTech. https://doi.org/10.5772/7544

Hosker, C., & Ward, D. (2017). Hypoactive delirium. *BMJ*, *357*. https://doi.org/10.1136/bmj.j2047

Hosmer, D. W., Lemeshow, S., & Sturdivant, R. X. (2013). *Applied Logistic Regression.* *Wiley Series in Probability and Statistics*.

Hshieh, T. T., Fong, T. G., Marcantonio, E. R., & Inouye, S. K. (2008). Cholinergic Deficiency Hypothesis in Delirium: A Synthesis of Current Evidence. *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences*, *63*(7), 764–772. https://doi.org/10.1093/gerona/63.7.764

Hurwitz, J., & Kirsch, D. (2018). *Machine learning*. John Wiley & Sons Inc.

INFARMED - Direção de informação e Planeamento Estratégico. (2017). *Utlilização de Benzodiazepinas e análogos*. Retrieved from https://www.infarmed.pt/documents/15786/2219894/Utlilização+de+Benzodiazepinas+e+análogos/adb100fa-4a77-4eb7-9e67-99229e13154f

INFARMED, A. N. do M. e P. de S. (2002). *Evolução do consumo de antidepressivos em Portugal continental de 1995 a 2001: impacto das medidas reguladoras*. *Observatório dos Medicamentos e Produtos de Saúde - Infarmed*. Retrieved from http://www.infarmed.pt/web/infarmed/profissionais-de-saude/utilizacao-e-despesa/estudos-medicamentos-por-grupo-terapeutico

INFARMED, A. N. do M. e P. de S. (2009). Infomed – base de dados de medicamentos do Infarmed - Tri-hexifenidilo. Retrieved July 12, 2021, from extranet.infarmed.pt/INFOMED-fo/index.xhtml

INFARMED, A. N. do M. e P. de S. (2010a). Infomed – base de dados de medicamentos do Infarmed - Lorazepam. Retrieved July 21, 2021, from extranet.infarmed.pt/INFOMED-fo/index.xhtml

INFARMED, A. N. do M. e P. de S. (2010b). *Prontuário Terapêutico - 9*. (I. de S. INFARMED - Autoridade Nacional do Medicamento e Produtos de Saúde, Ed.).

INFARMED, A. N. do M. e P. de S. (2011). Infomed – base de dados de medicamentos do Infarmed - Digoxina. Retrieved June 16, 2021, from http://www.infarmed.pt/infomed/download\_ficheiro.php?med\_id=4892&tipo\_doc=rcm

INFARMED, A. N. do M. e P. de S. (2014). Infomed – base de dados de medicamentos do Infarmed - Diazepam. Retrieved July 21, 2021, from extranet.infarmed.pt/INFOMED-fo/index.xhtml

INFARMED, A. N. do M. e P. de S. (2016a). Infomed – base de dados de medicamentos do Infarmed - Furosemida. Retrieved June 13, 2021, from extranet.infarmed.pt/INFOMED-fo/index.xhtml

INFARMED, A. N. do M. e P. de S. (2016b). Infomed – base de dados de medicamentos do Infarmed - Varfarina. Retrieved July 16, 2021, from extranet.infarmed.pt/INFOMED-fo/index.xhtml

INFARMED, A. N. do M. e P. de S. (2018a). Infomed – base de dados de medicamentos do Infarmed - Captopril. Retrieved September 13, 2021, from extranet.infarmed.pt/INFOMED-fo/index.xhtml

INFARMED, A. N. do M. e P. de S. (2018b). Infomed – base de dados de medicamentos do Infarmed - Pravastatina. Retrieved July 10, 2021, from https://extranet.infarmed.pt/INFOMED-fo/pesquisa-avancada.xhtml

INFARMED, A. N. do M. e P. de S. (2018c). Infomed – base de dados de medicamentos do Infarmed - Quetiapina. Retrieved July 13, 2021, from extranet.infarmed.pt/INFOMED-fo/index.xhtml

INFARMED, A. N. do M. e P. de S. (2018d). Infomed – base de dados de medicamentos do Infarmed - Ranitidina. Retrieved July 1, 2021, from extranet.infarmed.pt/INFOMED-fo/index.xhtml

INFARMED, A. N. do M. e P. de S. (2018e). Infomed – base de dados de medicamentos do Infarmed - Sinvastatina. Retrieved September 23, 2021, from extranet.infarmed.pt/INFOMED-fo/index.xhtml

INFARMED, A. N. do M. e P. de S. (2018f). Infomed – base de dados de medicamentos do Infarmed - Tramadol. Retrieved September 2, 2021, from https://extranet.infarmed.pt/INFOMED-fo/pesquisa-avancada.xhtml

INFARMED, A. N. do M. e P. de S. (2019a). Infomed – base de dados de medicamentos do Infarmed - Amitriptilina. Retrieved August 13, 2021, from extranet.infarmed.pt/INFOMED-fo/index.xhtml

INFARMED, A. N. do M. e P. de S. (2019b). Infomed – base de dados de medicamentos do Infarmed - Clonidina. Retrieved June 17, 2021, from https://extranet.infarmed.pt/INFOMED-fo/pesquisa-avancada.xhtml

INFARMED, A. N. do M. e P. de S. (2019c). Infomed – base de dados de medicamentos do Infarmed - Rosuvastatina. Retrieved August 28, 2021, from https://extranet.infarmed.pt/INFOMED-fo/pesquisa-avancada.xhtml

INFARMED, A. N. do M. e P. de S. (2020a). Infomed – base de dados de medicamentos do Infarmed - Desloratadina. Retrieved August 17, 2021, from https://extranet.infarmed.pt/INFOMED-fo/pesquisa-avancada.xhtml

INFARMED, A. N. do M. e P. de S. (2020b). Infomed – base de dados de medicamentos do Infarmed - Dipiridamol. Retrieved June 12, 2021, from https://extranet.infarmed.pt/INFOMED-fo/pesquisa-avancada.xhtml

INFARMED, A. N. do M. e P. de S. (2020c). Infomed – base de dados de medicamentos do Infarmed - Fluvastatina. Retrieved June 13, 2021, from extranet.infarmed.pt/INFOMED-fo/index.xhtml

INFARMED, A. N. do M. e P. de S. (2020d). Infomed – base de dados de medicamentos do Infarmed - Haloperidol. Retrieved June 24, 2021, from extranet.infarmed.pt/INFOMED-fo/index.xhtml

INFARMED, A. N. do M. e P. de S. (2020e). Infomed – base de dados de medicamentos do Infarmed - Hidrocortisona. Retrieved July 14, 2021, from extranet.infarmed.pt/INFOMED-fo/index.xhtml

INFARMED, A. N. do M. e P. de S. (2020f). Infomed – base de dados de medicamentos do Infarmed - Hidroxizina. Retrieved August 13, 2021, from extranet.infarmed.pt/INFOMED-fo/index.xhtml

INFARMED, A. N. do M. e P. de S. (2020g). Infomed – base de dados de medicamentos do Infarmed - Morfina. Retrieved September 20, 2021, from extranet.infarmed.pt/INFOMED-fo/index.xhtml

INFARMED, A. N. do M. e P. de S. (2020h). Infomed – base de dados de medicamentos do Infarmed - Prednisolona. Retrieved June 15, 2021, from https://extranet.infarmed.pt/INFOMED-fo/pesquisa-avancada.xhtml

INFARMED, A. N. do M. e P. de S. (2020i). Infomed – base de dados de medicamentos do Infarmed - Trazodona. Retrieved September 16, 2021, from extranet.infarmed.pt/INFOMED-fo/index.xhtml

INFARMED, A. N. do M. e P. de S. (2020j). Infomed – base de dados de medicamentos do Infarmed Brometo de Butilescopolamina. Retrieved September 8, 2021, from https://extranet.infarmed.pt/INFOMED-fo/pesquisa-avancada.xhtml

INFARMED, A. N. do M. e P. de S. (2020k). Infomed – base de dados de medicamentos do Infarmed Sertralina. Retrieved August 18, 2021, from https://extranet.infarmed.pt/INFOMED-fo/pesquisa-avancada.xhtml

INFARMED, A. N. do M. e P. de S. (2021a). Infomed – base de dados de medicamentos do Infarmed - Alprazolam. Retrieved July 21, 2021, from extranet.infarmed.pt/INFOMED-fo/index.xhtml

INFARMED, A. N. do M. e P. de S. (2021b). Infomed – base de dados de medicamentos do Infarmed - Atorvastatina. Retrieved June 15, 2021, from https://extranet.infarmed.pt/INFOMED-fo/pesquisa-avancada.xhtml

INFARMED, A. N. do M. e P. de S. (2021c). Infomed – base de dados de medicamentos do Infarmed - Fluvoxamina. Retrieved August 26, 2021, from https://extranet.infarmed.pt/INFOMED-fo/pesquisa-avancada.xhtml

INFARMED, A. N. do M. e P. de S. (2021d). Infomed – base de dados de medicamentos do Infarmed - Nifedipina. Retrieved August 28, 2021, from extranet.infarmed.pt/INFOMED-fo/index.xhtml

INFARMED, A. N. do M. e P. de S. (2021e). Infomed – base de dados de medicamentos do Infarmed - Paroxetina. Retrieved June 13, 2021, from extranet.infarmed.pt/INFOMED-fo/index.xhtml

INFARMED, A. N. do M. e P. de S. (2021f). Infomed – base de dados de medicamentos do Infarmed - Risperidona. Retrieved July 15, 2021, from extranet.infarmed.pt/INFOMED-fo/index.xhtml

INFARMED, A. N. do M. e P. de S. (2021g). Infomed – base de dados de medicamentos do Infarmed - Tróspio. Retrieved August 13, 2021, from https://extranet.infarmed.pt/INFOMED-fo/pesquisa-avancada.xhtml

INFARMED, A. N. do M. e P. de S. (2021h). Infomed – base de dados de medicamentos do Infarmed - Venlafaxina. Retrieved August 13, 2021, from extranet.infarmed.pt/INFOMED-fo/index.xhtml

INFARMED, A. N. do M. e P. de S. (2021i). Suspensão da AIM de medicamentos contendo ranitidina. Retrieved October 27, 2021, from https://www.infarmed.pt/documents/15786/4183412/Suspensão+da+AIM+de+medicamentos+contendo+ranitidina/0776856c-4480-d062-dd58-a511a62c3786

INFARMED, A. N. do M. e P. de S., & Ministério da Saúde, I. (2012). *Prontuário Terapêutico - 11*. (I. P. / M. de S. INFARMED – Autoridade Nacional do Medicamento e Produtos de Saúde, Ed.) (Caramona,).

Inouye, S. K., Van Dyck, C. H., Alessi, C. A., Balkin, S., Siegal, A. P., & Horwitz, R. I. (1990). Clarifying confusion: The confusion assessment method: A new method for detection of delirium. *Annals of Internal Medicine*, *113*(12), 941–948. https://doi.org/10.7326/0003-4819-113-12-941

Inouye, Sharon K. (2006). Delirium in Older Persons. *The New England Journal of Medicine*, *354*(11), 1157–1165.

Inouye, Sharon K., & Charpentier, P. A. (1996). Precipitating factors for delirium in hospitalized elderly persons: Predictive model and interrelationship with baseline vulnerability. *Journal of the American Medical Association*, *275*(11), 852–857. https://doi.org/10.1001/jama.275.11.852

Inouye, Sharon K., Westendorp, R. G. J., & Saczynski, J. S. (2014). Delirium in elderly people. *The Lancet*, *383*(9920), 911–922. https://doi.org/10.1016/S0140-6736(13)60688-1

Inouye, Sharon K. (2000). Prevention of delirium in hospitalized older patients: risk factors and targeted intervention strategies. *Annals of Medicine*, *32*(4), 257–263. https://doi.org/10.3109/07853890009011770

Isik, A. T., & Grossberg, G. T. (2018). *Delirium in Elderly Patients*. (A. T. Isik & G. T. Grossberg, Eds.), *Delirium in Elderly Patients*. Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-65239-9

Jain, A. K., & Dubes, R. C. (1988). *Algorithms for Clustering Data*. Prentice Hall.

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). *An Introduction to Statistical Learning - with Applications in R*. Springer Texts in Statistics.

Jamin, A., Abraham, P., & Humeau-Heurtier, A. (2021). Machine learning for predictive data analytics in medicine: A review illustrated by cardiovascular and nuclear medicine examples. *Clinical Physiology and Functional Imaging*, *41*(2), 113–127. https://doi.org/10.1111/cpf.12686

Jaszczyk, A., & Juszczak, G. R. (2021). Glucocorticoids, metabolism and brain activity. *Neuroscience & Biobehavioral Reviews*, *126*(March 2020), 113–145. https://doi.org/10.1016/j.neubiorev.2021.03.007

Jauk, S., Kramer, D., Großauer, B., Rienmüller, S., Avian, A., Berghold, A., … Schulz, S. (2020). Risk prediction of delirium in hospitalized patients using machine learning: An implementation and prospective evaluation study. *Journal of the American Medical Informatics Association*, *27*(9), 1383–1392. https://doi.org/10.1093/jamia/ocaa113

Kareemi, H., Vaillancourt, C., Rosenberg, H., Fournier, K., & Yadav, K. (2021). Machine Learning Versus Usual Care for Diagnostic and Prognostic Prediction in the Emergency Department: A Systematic Review. *Academic Emergency Medicine*, *28*(2), 184–196. https://doi.org/10.1111/acem.14190

Kuhn, M., & Johnson, K. (2020). *Feature Engineering and Selection: A Practical Approach for Predictive Models*. CRC Press.

Lahariya, S., Grover, S., Bagga, S., & Sharma, A. (2014). Delirium in patients admitted to a cardiac intensive care unit with cardiac emergencies in a developing country: Incidence, prevalence, risk factor and outcome. *General Hospital Psychiatry*, *36*(2), 156–164. https://doi.org/10.1016/j.genhosppsych.2013.10.010

Laird, J. E., Newell, A., & Rosenbloom, P. S. (1987). An integrative architecture for general intelligence and. *Artificial Intelligence*, *33*(1987), 1–64.

Laurila, J. V., Laakkonen, M. L., Strandberg, T. E., & Tilvis, R. S. (2008). Predisposing and precipitating factors for delirium in a frail geriatric population. *Journal of Psychosomatic Research*, *65*(3), 249–254. https://doi.org/10.1016/j.jpsychores.2008.05.026

Lawlor, P. G., Gagnon, B., Mancini, I. L., Pereira, J. L., Hanson, J., Suarez-Almazor, M. E., & Bruera, E. D. (2002). Occurrence, causes, and outcome of delirium in patients with advanced cancer: A prospective study. *Archives of Internal Medicine*, *160*(6), 786–794. https://doi.org/10.1001/archinte.160.6.786

Lee-Archer, P. F., von Ungern-Sternberg, B. S., Reade, M. C., Law, K. C., & Long, D. (2021). An observational study of hypoactive delirium in the post-anesthesia recovery unit of a pediatric hospital. *Paediatric Anaesthesia*, (September 2020), 429–435. https://doi.org/10.1111/pan.14122

Lee, S., Mueller, B., Nick Street, W., & M. Carnahan, R. (2021). Machine learning algorithm to predict delirium from emergency department data. *BMJ*. https://doi.org/https://doi.org/10.1101/2021.02.19.21251956

Leonard, M. M., Nekolaichuk, C., Meagher, D. J., Barnes, C., Gaudreau, J. D., Watanabe, S., … Lawlor, P. G. (2014). Practical assessment of delirium in palliative care. *Journal of Pain and Symptom Management*, *48*(2), 176–190. https://doi.org/10.1016/j.jpainsymman.2013.10.024

Liang, S., Chau, J. P. C., Lo, S. H. S., Bai, L., Yao, L., & Choi, K. C. (2020). Validation of PREdiction of DELIRium in ICu patients (PRE-DELIRIC) among patients in intensive care units: A retrospective cohort study. *Nursing in Critical Care*, (August), 1–7. https://doi.org/10.1111/nicc.12550

Libbrecht, M. W., & Noble, W. S. (2015). Machine learning applications in genetics and genomics. *Nature Reviews Genetics*, *16*(6), 321–332. https://doi.org/10.1038/nrg3920

Lipowski, Z. J. (1987). Delirium (Acute Confusional States). *JAMA: The Journal of the American Medical Association*, *258*(13), 1789. https://doi.org/10.1001/jama.1987.03400130103041

Lipowski, Zbigniew J. (1991). Delirium: How Its Concept Has Developed. *International Psychogeriatrics*, *3*(2), 115–120. https://doi.org/10.1017/S1041610291000595

Lôbo, R. R., Silva Filho, S. R. B., Lima, N. K. C., Ferriolli, E., & Moriguti, J. C. (2010). Delirium. *Medicina (Ribeirao Preto. Online)*, *43*(3), 249–257. https://doi.org/10.11606/issn.2176-7262.v43i3p249-257

Lüllmann, H., Mohr, K., & Hein, L. (2018). *Color Atlas of Pharmacology* (Fifth). Thieme.

MacLullich, A. M. J., Ferguson, K. J., Miller, T., de Rooij, S. E. J. A., & Cunningham, C. (2008). Unravelling the pathophysiology of delirium: A focus on the role of aberrant stress responses. *Journal of Psychosomatic Research*, *65*(3), 229–238. https://doi.org/10.1016/j.jpsychores.2008.05.019

Maldonado, J. R. (2008). Pathoetiological Model of Delirium: a Comprehensive Understanding of the Neurobiology of Delirium and an Evidence-Based Approach to Prevention and Treatment. *Critical Care Clinics*, *24*(4), 789–856. https://doi.org/10.1016/j.ccc.2008.06.004

Maldonado, J. R. (2017). Delirium pathophysiology: An updated hypothesis of the etiology of acute brain failure. *International Journal of GeriSatric Psychiatry*, *33*(11), 1428–1457. https://doi.org/10.1002/gps.4823

Mariz, J., Castanho, T. C., Teixeira, J., Sousa, N., & Santos, N. C. (2016). Delirium diagnostic and screening instruments in the emergency department: An up-to-date systematic review. *Geriatrics (Switzerland)*, *1*(3), 1–14. https://doi.org/10.3390/geriatrics1030022

Maroco, J., Silva, D., Rodrigues, A., Guerreiro, M., Santana, I., & de Mendonça, A. (2011). Data mining methods in the prediction of Dementia: A real-data comparison of the accuracy, sensitivity and specificity of linear discriminant analysis, logistic regression, neural networks, support vector machines, classification trees and random forests. *BMC Research Notes*, *4*(1), 299. https://doi.org/10.1186/1756-0500-4-299

Martins, S., & Fernandes, L. (2012). Delirium in elderly people: A review. *Frontiers in Neurology*, *JUN*(June), 1–12. https://doi.org/10.3389/fneur.2012.00101

McCorduck, P. (2004). *Machines Who Think*. *A K Peters, Ltd.* Natick, MA: A K Peters, Ltd. https://doi.org/10.1126/science.254.5036.1291-a

McCulloch, W. S., & Pits, W. (1943). A LOGICAL CALCULUS OF THE IDEAS IMMANENT IN NERVOUS ACTIVITY. *Bullentin of Mathematical*, *5*, 115–133.

Meagher, D. J., MacLullich, A. M. J., & Laurila, J. V. (2008). Defining delirium for the International Classification of Diseases, 11th Revision. *Journal of Psychosomatic Research*, *65*(3), 207–214. https://doi.org/10.1016/j.jpsychores.2008.05.015

MediBIAL – Produtos Médicos e Farmacêuticos. (n.d.). Mexazolam. Retrieved July 21, 2021, from https://www.bial.com/media/2780/sedoxil.pdf

Meltzer, H. Y. (2013). Update on Typical and Atypical Antipsychotic Drugs. *Annual Review of Medicine*, *64*(1), 393–406. https://doi.org/10.1146/annurev-med-050911-161504

Michalski, R. S., Carbonell, J. G., & Mitchell, T. M. (1983). *Machine Learning*. Berlin, Heidelberg: Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-662-12405-5

Michaud, L., Büla, C., Berney, A., Camus, V., Voellinger, R., Stiefel, F., & Burnand, B. (2007). Delirium: Guidelines for general hospitals. *Journal of Psychosomatic Research*, *62*(3), 371–383. https://doi.org/10.1016/j.jpsychores.2006.10.004

Mitchell, T. M. (1999). Machine learning and data mining. *Communications of the ACM*, *42*(11), 30–46. https://doi.org/10.1145/319382.319388

Mitchell, Tom M. (1997). *Machine Learning*. McGraw-Hill Science/Engineering/Math.

Mittal, V., Muralee, S., Williamson, D., McEnerney, N., Thomas, J., Cash, M., & Tampi, R. R. (2011). Delirium in the elderly: A comprehensive review. *American Journal of Alzheimer’s Disease and Other Dementias*, *26*(2), 97–109. https://doi.org/10.1177/1533317510397331

Morandi, A., Pandharipande, P., Trabucchi, M., Rozzini, R., Mistraletti, G., Trompeo, A. C., … Ely, E. W. (2008). Understanding international differences in terminology for delirium and other types of acute brain dysfunction in critically ill patients. *Intensive Care Medicine*, *34*(10), 1907–1915. https://doi.org/10.1007/s00134-008-1177-6

Murphy, K. P. (2012). *Machine learning: A Probabilistic Perspective*. *Expert Systems* (Vol. 5). The MIT Press. https://doi.org/10.1111/j.1468-0394.1988.tb00341.x

Nagari, N., & Suresh Babu, M. (2019). Assessment of risk factors and precipitating factors of delirium in patients admitted to intensive care unit of a tertiary care hospital. *British Journal of Medical Practitioners*, *12*(2).

Nemati, S., Holder, A., Razmi, F., Stanley, M. D., Clifford, G. D., & Buchman, T. G. (2018). An Interpretable Machine Learning Model for Accurate Prediction of Sepsis in the ICU. *Critical Care Medicine*, *46*(4), 547–553. https://doi.org/10.1097/CCM.0000000000002936

Nikam, S. S. (2015). A Comparative Study of Classification Techniques in Data Mining Algorithms. *ORIENTAL JOURNAL OF COMPUTER SCIENCE & TECHNOLOGY*, *8*(1), 13–19.

Nitchingham, A., Kumar, V., Shenkin, S., Ferguson, K. J., & Caplan, G. A. (2018). A systematic review of neuroimaging in delirium: predictors, correlates and consequences. *International Journal of Geriatric Psychiatry*, *33*(11), 1458–1478. https://doi.org/10.1002/gps.4724

Nussey, S., & Whitehead, S. (2001). *Endocrinology: An Integrated Approach*. Informa HealthCare 2001. Retrieved from https://www.ncbi.nlm.nih.gov/books/NBK20/#A10

ONU, P. (2021). Envelhecimento. Retrieved May 31, 2021, from https://unric.org/pt/envelhecimento/

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., … Duchesnay, É. (2012). Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*, *12*(May 2014), 2825–2830. Retrieved from http://arxiv.org/abs/1201.0490

Pérez-Ros, P., & Martínez-Arnau, F. (2019). Delirium Assessment in Older People in Emergency Departments. A Literature Review. *Diseases*, *7*(1), 14. https://doi.org/10.3390/diseases7010014

PORDATA. (2020). Esperança de vida à nascença: total e por sexo. Retrieved October 20, 2021, from https://www.pordata.pt/Europa/Esperança+de+vida+à+nascença+total+e+por+sexo-1260

Prayce, R., Quaresma, F., & Neto, I. G. (2018). Delirium: O 7o Parâmetro Vital? *Acta Médica Portuguesa*, *31*(1), 51. https://doi.org/10.20344/amp.9670

Quinlan, J. R. (1986). Induction of decision trees. *Machine Learning*, *1*(1), 81–106. https://doi.org/10.1007/bf00116251

Rahm, E., & Do, H. (2000). Data cleaning: Problems and current approaches. *IEEE Data Eng. Bull.*, *23*(4), 3–13. Retrieved from http://wwwiti.cs.uni-magdeburg.de/iti\_db/lehre/dw/paper/data\_cleaning.pdf%5Cnpapers2://publication/uuid/17B58056-3A7F-4184-8E8B-0E4D82EFEA1A%5Cnhttp://dc-pubs.dbs.uni-leipzig.de/files/Rahm2000DataCleaningProblemsand.pdf

Ritchie, C. W., Newman, T. H., Leurent, B., & Sampson, E. L. (2014). The association between C-reactive protein and delirium in 710 acute elderly hospital admissions. *International Psychogeriatrics*, *26*(5), 717–724. https://doi.org/10.1017/S1041610213002433

Robinson, T. N., Raeburn, C. D., Tran, Z. V., Brenner, L. A., & Moss, M. (2011). Motor subtypes of postoperative delirium in older adults. *Archives of Surgery*, *146*(3), 295–300. https://doi.org/10.1001/archsurg.2011.14

Rosenblatt, F. (1958). The perceptron: A probabilistic model for information storage and organization in the brain. *Psychological Review*, *65*(6), 386–408. https://doi.org/10.1037/h0042519

Rossi Varallo, F., Maicon de Oliveira, A., Cristina Barboza Zanetti, A., Carneiro Capucho, H., Régis Leira Pereira, L., Borges Pereira, L., … Detoni Lopes, V. (2021). Drug-Induced Delirium among Older People. In *New Insights into the Future of Pharmacoepidemiology and Drug Safety [Working Title]* (Vol. 32, pp. 137–144). IntechOpen. https://doi.org/10.5772/intechopen.95470

Rubino, A. S., Onorati, F., Caroleo, S., Galato, E., Nucera, S., Amantea, B., … Renzulli, A. (2010). Impact of clonidine administration on delirium and related respiratory weaning after surgical correction of acute type-A aortic dissection: results of a pilot study. *Interactive CardioVascular and Thoracic Surgery*, *10*(1), 58–62. https://doi.org/10.1510/icvts.2009.217562

Rudberg, M. A., Pompei, P., Foreman, M. D., Ross, R. E., & Cassel, C. K. (1997). The natural history of delirium in older hospitalized patients: A syndrome of heterogeneity. *Age and Ageing*, *26*(3), 169–174. https://doi.org/10.1093/ageing/26.3.169

Rutter, L.-M., Nouzova, E., Stott, D. J., Weir, C. J., Assi, V., Barnett, J. H., … Tieges, Z. (2018). Diagnostic test accuracy of a novel smartphone application for the assessment of attention deficits in delirium in older hospitalised patients: a prospective cohort study protocol. *BMC Geriatrics*, *18*(1), 217. https://doi.org/10.1186/s12877-018-0901-5

Salluh, J. I. F., Wang, H., Schneider, E. B., Nagaraja, N., Yenokyan, G., Damluji, A., … Stevens, R. D. (2015). Outcome of delirium in critically ill patients: Systematic review and meta-analysis. *BMJ (Online)*, *350*, 1–10. https://doi.org/10.1136/bmj.h2538

Samuel, A. L. (1959). Some Studies in Machine Learning. *IBM Journal of Research and Development*, *3*(3), 210–229. Retrieved from https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=5392560

Saraiva, C. B., & Cerejeira, J. (2014). *Psiquiatria fundamental*. Lisboa: Lidel. Retrieved from http://id.bnportugal.gov.pt/bib/bibnacional/1938223

Schröer, C., Kruse, F., & Gómez, J. M. (2021). A Systematic Literature Review on Applying CRISP-DM Process Model. *Procedia Computer Science*, *181*(2019), 526–534. https://doi.org/10.1016/j.procs.2021.01.199

Shen, W. W. (1999). A history of antipsychotic drug development. *Comprehensive Psychiatry*, *40*(6), 407–414. https://doi.org/10.1016/S0010-440X(99)90082-2

Siddiqi, N., House, A. O., & Holmes, J. D. (2006). Occurrence and outcome of delirium in medical in-patients: A systematic literature review. *Age and Ageing*, *35*(4), 350–364. https://doi.org/10.1093/ageing/afl005

Simone, M. J., & Tan, Z. S. (2011). The Role of Inflammation in the Pathogenesis of Delirium and Dementia in Older Adults: A Review. *CNS Neuroscience & Therapeutics*, *17*(5), 506–513. https://doi.org/10.1111/j.1755-5949.2010.00173.x

Slooter, A. J. C., Otte, W. M., Devlin, J. W., Arora, R. C., Bleck, T. P., Claassen, J., … Stevens, R. D. (2020). Updated nomenclature of delirium and acute encephalopathy: statement of ten Societies. *Intensive Care Medicine*, *46*(5), 1020–1022. https://doi.org/10.1007/s00134-019-05907-4

Smith, T. O., Cooper, A., Peryer, G., Griffiths, R., Fox, C., & Cross, J. (2017). Factors predicting incidence of post-operative delirium in older people following hip fracture surgery: a systematic review and meta-analysis. *International Journal of Geriatric Psychiatry*, *32*(4), 386–396. https://doi.org/10.1002/gps.4655

Souza, I. T., Wildner, D. P. da S., Gazdzichi, A. K., & Nink, R. F. O. (2020). THE EVOLUTION OF PSYCHOPHARMACES IN THE TREATMENT OF DEPRESSION. *Brazilian Journal of Surgery and Clinical Research – BJSCR*, *33*(2), 109–114. Retrieved from http://www.mastereditora.com.br/bjscr

Stahl, S. M., Grady, M. M., Moret, C., & Briley, M. (2005). SNRIs: The Pharmacology, Clinical Efficacy, and Tolerability in Comparison with Other Classes of Antidepressants. *CNS Spectrums*, *10*(9), 732–747. https://doi.org/10.1017/S1092852900019726

Stewart, J., Sprivulis, P., & Dwivedi, G. (2018). Artificial intelligence and machine learning in emergency medicine. *EMA - Emergency Medicine Australasia*, *30*(6), 870–874. https://doi.org/10.1111/1742-6723.13145

Stoltzfus, J. C. (2011). Logistic regression: A brief primer. *Academic Emergency Medicine*, *18*(10), 1099–1104. https://doi.org/10.1111/j.1553-2712.2011.01185.x

Swart, L. M., van der Zanden, V., Spies, P. E., de Rooij, S. E., & van Munster, B. C. (2017). The Comparative Risk of Delirium with Different Opioids: A Systematic Review. *Drugs & Aging*, *34*(6), 437–443. https://doi.org/10.1007/s40266-017-0455-9

Tang, E., Laverty, M., Weir, A., Wilson, E. S., Walsh, T. S., Allerhand, M., … Tieges, Z. (2018). Development and feasibility of a smartphone-based test for the objective detection and monitoring of attention impairments in delirium in the ICU. *Journal of Critical Care*, *48*, 104–111. https://doi.org/10.1016/j.jcrc.2018.08.019

Taylor, R. A., & Haimovich, A. D. (2021). Machine Learning in Emergency Medicine: Keys to Future Success. *Academic Emergency Medicine*, *28*(2), 263–267. https://doi.org/10.1111/acem.14189

Telles-Correia, D., Guerreiro, D. F., Oliveira, S., & Figueira, M. L. (2007). Diferenças farmacodinâmicas e farmacocinéticas entre os SSRI: Implicações na prática clínica. *Acta Medica Portuguesa*, *20*(2), 167–174.

Timmermans, S., Souffriau, J., & Libert, C. (2019). A General Introduction to Glucocorticoid Biology. *Frontiers in Immunology*, *10*(JULY). https://doi.org/10.3389/fimmu.2019.01545

Trzepacz, P. T. (1999). Update on the Neuropathogenesis of Delirium. *Dementia and Geriatric Cognitive Disorders*, *10*(5), 330–334. https://doi.org/10.1159/000017164

Van Den Boogaard, M., Pickkers, P., Slooter, A. J. C., Kuiper, M. A., Spronk, P. E., Van Der Voort, P. H. J., … Schoonhoven, L. (2012). Development and validation of PRE-DELIRIC (PREdiction of DELIRium in ICu patients) delirium prediction model for intensive care patients: Observational multicentre study. *BMJ (Online)*, *344*(7845), 17. https://doi.org/10.1136/bmj.e420

van der Mast, R. C. (1998). Pathophysiology of Delirium. *Journal of Geriatric Psychiatry and Neurology*, *11*(3), 138–145. https://doi.org/10.1177/089198879801100304

Van Eijk, M. M. J., Van Marum, R. J., Klijn, I. A. M., De Wit, N., Kesecioglu, J., & Slooter, A. J. C. (2009). Comparison of delirium assessment tools in a mixed intensive care unit. *Critical Care Medicine*, *37*(6), 1881–1885. https://doi.org/10.1097/CCM.0b013e3181a00118

Van Munster, B. C., Korevaar, J. C., Zwinderman, A. H., Levi, M., Wiersinga, W. J., & De Rooij, S. E. (2008). Time-course of cytokines during delirium in elderly patients with hip fractures. *Journal of the American Geriatrics Society*, *56*(9), 1704–1709. https://doi.org/10.1111/j.1532-5415.2008.01851.x

Vapnik, V. N. (2000). *The Nature of Statistical Learning Theory*. *Statistics for Engineering and Information Science* (2nd ed.). Springer New York. Retrieved from https://ci.nii.ac.jp/naid/10020951890

Velayati, A., Vahdat Shariatpanahi, M., Shahbazi, E., & Vahdat Shariatpanahi, Z. (2019). Association between preoperative nutritional status and postoperative delirium in individuals with coronary artery bypass graft surgery: A prospective cohort study. *Nutrition*, *66*, 227–232. https://doi.org/10.1016/j.nut.2019.06.006

Vellido, A. (2020). The importance of interpretability and visualization in machine learning for applications in medicine and health care. *Neural Computing and Applications*, *32*(24), 18069–18083. https://doi.org/10.1007/s00521-019-04051-w

Vluymans, S. (2019). Learning from imbalanced data. *Studies in Computational Intelligence*, *807*(9), 81–110. https://doi.org/10.1007/978-3-030-04663-7\_4

Wang, H. R., Woo, Y. S., & Bahk, W.-M. (2013). Atypical antipsychotics in the treatment of delirium. *Psychiatry and Clinical Neurosciences*, *67*(5), 323–331. https://doi.org/10.1111/pcn.12066

Wassenaar, A., Van Den Boogaard, M., Schoonhoven, L., Donders, R., & Pickkers, P. (2017). Delirium prediction in the intensive care unit: Head to head comparison of two delirium prediction models. *Intensive Care Medicine Experimental*, *5*(2), 1–9. Retrieved from http://www.embase.com/search/results?subaction=viewrecord&from=export&id=L619043579%0Ahttp://dx.doi.org/10.1186/s40635-017-0151-4%0Ahttp://sfx.library.uu.nl/utrecht?sid=EMBASE&issn=2197425X&id=doi:10.1186%2Fs40635-017-0151-4&atitle=Delirium+prediction+in+

Wassenaar, A., van den Boogaard, M., van Achterberg, T., Slooter, A. J. C., Kuiper, M. A., Hoogendoorn, M. E., … Pickkers, P. (2015). Multinational development and validation of an early prediction model for delirium in ICU patients. *Intensive Care Medicine*, *41*(6), 1048–1056. https://doi.org/10.1007/s00134-015-3777-2

Webster, R. (2001). Neurotransmitter Systems and Function: Overview. In R. A. Webster (Ed.), *Neurotransmitters, Drugs and Brain Function* (pp. 1–32). Chichester, UK: John Wiley & Sons, Ltd. https://doi.org/10.1002/0470846577.ch1

WHO, & EMRO. (2005). *Clinical guidelines for the management od hypertension*. Retrieved from https://apps.who.int/iris/handle/10665/119738

Wilson, J. E., Mart, M. F., Cunningham, C., Shehabi, Y., Girard, T. D., MacLullich, A. M. J., … Ely, E. W. (2020). Delirium. *Nature Reviews Disease Primers*, *6*(1), 90. https://doi.org/10.1038/s41572-020-00223-4

Wilson, K., Broadhurst, C., Diver, M., Jackson, M., & Mottram, P. (2005). Plasma insulin growth factor - 1 and incident delirium in older people. *International Journal of Geriatric Psychiatry*, *20*(2), 154–159. https://doi.org/10.1002/gps.1265

Wirth, R., & Hipp, J. (2000). CRISP-DM : Towards a Standard Process Model for Data Mining. *Proceedings of the Fourth International Conference on the Practical Application of Knowledge Discovery and Data Mining*, (24959), 29–39.

Witlox, J., Eurelings, L. S. M., De Jonghe, J. F. M., Kalisvaart, K. J., Eikelenboom, P., & Van Gool, W. A. (2010). Delirium in elderly patients and the risk of postdischarge mortality, institutionalization, and dementia: A meta-analysis. *JAMA - Journal of the American Medical Association*, *304*(4), 443–451. https://doi.org/10.1001/jama.2010.1013

Wong, A., Young, A. T., Liang, A. S., Gonzales, R., Douglas, V. C., & Hadley, D. (2018). Development and Validation of an Electronic Health Record-Based Machine Learning Model to Estimate Delirium Risk in Newly Hospitalized Patients Without Known Cognitive Impairment. *JAMA Network Open*, *1*(4), e181018. https://doi.org/10.1001/jamanetworkopen.2018.1018

Wong, C. L., Holroyd-Leduc, J., Simel, D. L., & Straus, S. E. (2010). Does this patient have delirium?: value of bedside instruments. *Jama*, *304*(7), 779–786.

Xia, H., Wang, C., Yan, L., Dong, X., & Wang, Y. (2019). Machine Learning Based Medicine Distribution System. In *Proceedings of the 2019 10th IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications, IDAACS 2019* (Vol. 2, pp. 912–915). IEEE. https://doi.org/10.1109/IDAACS.2019.8924236

Xue, B., Li, D., Lu, C., King, C. R., Wildes, T., Avidan, M. S., … Abraham, J. (2021). Use of Machine Learning to Develop and Evaluate Models Using Preoperative and Intraoperative Data to Identify Risks of Postoperative Complications. *JAMA Network Open*, *4*(3), e212240. https://doi.org/10.1001/jamanetworkopen.2021.2240

Yanagihara, H., Kamo, K. I., Imori, S., & Satoh, K. (2012). Bias-corrected AIC for selecting variables in multinomial logistic regression models. *Linear Algebra and Its Applications*, *436*(11), 4329–4341. https://doi.org/10.1016/j.laa.2012.01.018

# Apêndice I – Descrição da Base de dados

|  |  |  |
| --- | --- | --- |
| **Variável** | **Descrição** | **Codificação** |
| **Proveniência** | Local de proveniência do doente | 0 - Casa  1 - Inter-Hopitlar  2 - Intra-Hospitalar  3 - Lar |
| **Local\_SU** | Local de proveniência do doente dentro do SU | 0 – AMBUL  1 – UCISU  2 – UDC1  3 – UDC2 |
| **Idade** | Idade, em anos, até à data da recolha dos dados | - - - - |
| **Genero** | Variável que indicativa do género do individuo que entrou no SU | 1 – Masculino  0 – Feminino |
| **Interna\_Dias** | Tempo, em dias, que o individuo passou no SU | - - - - |
| **Grupo\_Diagn** | Indica a especialidade hospitalar que levou o individuo a recorrer ao SU | 0 – Cardiovascular  1 – Gastrointestinal  2 – Genitourinário  3 -- Hemato-Oncológico  4 – Musculoesquelético  5 – Neurológico  6 – Outro  7 – Respiratório  8 -- Toxicidade de Drogas |
| **SIRS** | Quantidade de critérios SIRS presentes.  Critérios: temperatura corporal, frequência respiratória, frequência cardíaca, número de leucócitos.  Valores possíveis: 0-4 | - - - - |
| **Glicose** | Valor dos níveis de glicose no sangue  Valores de referência: 90 a 130 mg/dL | - - - - |
| **Sodio** | Valor dos níveis de sódio no sangue  Valores de referência: 135 a 146 mEq/L | - - - - |
| **Ureia** | Valor dos níveis de ureia no sangue  Valores de referência: 19 a 49 mg/dL | - - - - |
| **Creatinina** | Valor dos níveis de creatinina no sangue  Valores de referência: 0.6 a 1.2 mg/dL | - - - - |
| **PCR** | Valor relativo à PCR  Valores de referência: < 5mg/L | - - - - |
| **pH** | Valor do pH sanguíneo  Valores de referência: 7.35 a 7.45 | - - - - |
| **Ca\_ionizado** | Valor cálcio ionizado  Valores de referência: 1.15 a 1.35 mmol/L | - - - - |
| **pCO2** | Pressão parcial de dióxido de carbono (PCO2)  Valores de referência: 33 a 45 mm Hg | - - - - |
| **pO2** | Pressão parcial de oxigénio (PO2)  Valores de referência: 75 a 105 mm Hg | - - - - |
| **HC03** | Valor relativo ao ião bicarbonato  Valores de referência: 22 a 28 mEq/L | - - - - |
| **Rosuvastatina** | Medicamento antidislipidémico | 0 – Ausente  1 – Presente |
| **Atorvastatina** | Medicamento antidislipidémico | 0 – Ausente  1 – Presente |
| **Pravastatina** | Medicamento antidislipidémico | 0 – Ausente  1 – Presente |
| **Sinvastatina** | Medicamento antidislipidémico | 0 – Ausente  1 – Presente |
| **Fluvastatina** | Medicamento antidislipidémico | 0 – Ausente  1 – Presente |
| **Alprazolam** | Psicofármaco pertencente à classe dos ansiolíticos, sedativos e hipnóticos | 0 – Ausente  1 – Presente |
| **Captopril** | Anti-hipertensor pertencente à categoria dos IECA | 0 – Ausente  1 – Presente |
| **Desloratadine** | Anti-histamínico não sedativo | 0 – Ausente  1 – Presente |
| **Diazepam** | Psicofármaco pertencente à classe dos ansiolíticos, sedativos e hipnóticos | 0 – Ausente  1 – Presente |
| **Lorazepam** | Psicofármaco pertencente à classe dos ansiolíticos, sedativos e hipnóticos | 0 – Ausente  1 – Presente |
| **Digoxin** | Medicamento cardiotónico digitálico | 0 – Ausente  1 – Presente |
| **Dipyridamole** | Fármaco da categoria dos anticoagulantes | 0 – Ausente  1 – Presente |
| **Furosemide** | Anti-hipertensor pertencente à categoria dos diuréticos de ansa | 0 – Ausente  1 – Presente |
| **Fluvoxamine** | Psicofármaco pertencente à classe dos antidepressivos | 0 – Ausente  1 – Presente |
| **Haloperidol** | Fármaco pertencente à categoria dos antipsicóticos | 0 – Ausente  1 – Presente |
| **Hydrocortisone** | Medicamento glucocorticoide | 0 – Ausente  1 – Presente |
| **Iloperidone** | Fármaco pertencente à categoria dos antipsicóticos | 0 – Ausente  1 – Presente |
| **Morphine** | Medicamento pertencente ao grupo dos analgésicos estupefacientes | 0 – Ausente  1 – Presente |
| **Nifedipine** | Anti-hipertensor pertencente à categoria dos bloqueadores de Ca2+ | 0 – Ausente  1 – Presente |
| **Paliperidone** | Fármaco pertencente à categoria dos antipsicóticos | 0 – Ausente  1 – Presente |
| **Prednisone** | Medicamento glucocorticoide | 0 – Ausente  1 – Presente |
| **Ranitidine** | Medicamento da categoria dos antiácidos e antiulcerosos | 0 – Ausente  1 – Presente |
| **Risperidone** | Fármaco pertencente à categoria dos antipsicóticos | 0 – Ausente  1 – Presente |
| **Trazodone** | Psicofármaco antidepressor pertencente à classe dos tricíclicos | 0 – Ausente  1 – Presente |
| **Venlafaxine** | Psicofármaco pertencente à classe dos ISRSN | 0 – Ausente  1 – Presente |
| **Warfarin** | Fármaco da categoria dos anticoagulantes | 0 – Ausente  1 – Presente |
| **Amitriptyline** | Psicofármaco antidepressor pertencente à classe dos tricíclicos | 0 – Ausente  1 – Presente |
| **Hydroxyzine** | Anti-histamínico sedativo | 0 – Ausente  1 – Presente |
| **Paroxetine** | Psicofármaco pertencente à classe dos antidepressivos | 0 – Ausente  1 – Presente |
| **Quetiapine** | Fármaco pertencente à categoria dos antipsicóticos | 0 – Ausente  1 – Presente |
| **Scopolamine** | Antiespasmódico | 0 – Ausente  1 – Presente |
| **Trihexyphenidyl** | Medicamento antiparkinsónico | 0 – Ausente  1 – Presente |
| **Clonidine** | Anti-hipertensor pertencente à categoria dos alfa2 | 0 – Ausente  1 – Presente |
| **Sertralina** | Psicofármaco pertencente à classe dos antidepressivos | 0 – Ausente  1 – Presente |
| **Tramadol** | Medicamento pertencente ao grupo dos analgésicos estupefacientes | 0 – Ausente  1 – Presente |
| **Mexazolam** | Psicofármaco pertencente à classe dos ansiolíticos, sedativos e hipnóticos | 0 – Ausente  1 – Presente |
| **Trospium** | Medicamento utilizado para disfunções geniturinárias | 0 – Ausente  1 – Presente |
| **Obito** | Indica se o individuo faleceu ou não até ao mês Outubro do ano de 2021 | 0 – Não  1 – Sim |
| **Alcoolico** | Variável indicativa acerca do individuo apresentar ou não o vício do álcool | 0 – Não  1 – Sim |
| **Delirium** | Variável que indica se o individuo desenvolveu *delirium* ou não | 0 – Não  1 – Sim |

# Anexo I – The confusion assessment method instrument

**Acute onset**

1. Is there evidence of an acute change in mental status from the patient's baseline?

**Inattention\***

2. A. Did the patient have difficulty focusing attention, for example, being easily distractible, or having difficulty keeping track of what was being said?

Not present at any time during interview.

Present at some time during interview, but in mild form.

Present at some time during interview, in marked form.

Uncertain.

B. (If present or abnormal) Did this behavior fluctuate during the interview, that is, tend to come and go or increase and decrease in severity?

Yes.

No.

Uncertain

Not applicable.

C. (If present or abnormal) Please describe this behavior:

**Disorganized thinking**

3. Was the patient's thinking disorganized or incoherent, such as rambling or irrelevant conversation, unclear or illogical flow of ideas, or unpredictable switching from subject to subject?

**Altered level of consciousness**

4. Overall, how would you rate this patient's level of consciousness?

Alert (normal).

Vigilant (hyperalert, overly sensitive to environmental stimuli, startled very easily).

Lethargic (drowsy, easily aroused).

Stupor (difficult to arouse).

Coma (unarousable).

Uncertain.

**Disorientation**

5. Was the patient disoriented at any time during the interview, such as thinking that he or she was somewhere other than the hospital, using the wrong bed, or misjudging the time of day?

**Memory impairment**

6. Did the patient demonstrate any memory problems during the interview, such as inability to remember events in the hospital or difficulty remembering instructions?

**Perceptual disturbances**

7. Did the patient have any evidence of perceptual disturbances, for example, hallucinations, illusions, or misinterpretations (such as thinking something was moving when it was not)?

**Psychomotor agitation**

8. Part 1. At any time during the interview, did the patient have an unusually increased level of motor activity, such as restlessness, picking at bedclothes, tapping fingers, or making frequent sudden changes of position?

**Psychomotor retardation**

8. Part 2. At any time during the interview, did the patient have an unusually decreased level of motor activity, such as sluggishness, staring into space, staying in one position for a long time, or moving very slowly?

**Altered sleep-wake cycle**

9. Did the patient have evidence of disturbance of the sleep-wake cycle, such as excessive daytime sleepiness with insomnia at night?

\* The questions listed under this topic were repeated for each topic where applicable.

# Anexo II – The Confusion Assessment Method (CAM) Diagnostic Algorithm\*

Feature 1. Acute Onset and Fluctuating Course

This feature is usually obtained from a family member or nurse and is shown by positive responses to the following questions: Is there evidence of an acute change in mental status from the patient's baseline? Did the (abnormal) behavior fluctuate during the day, that is, tend to come and go, or increase and decrease in severity?

Feature 2. Inattention

This feature is shown by a positive response to the following question: Did the patient have difficulty focusing attention, for example, being easily distractible, or having difficulty keeping track of what was being said?

Feature 3. Disorganized Thinking

This feature is shown by a positive response to the following question: Was the patient's thinking disorganized or incoherent, such as rambling or irrelevant conversation, unclear or illogical flow of ideas, or unpredictable switching from subject to subject?

Feature 4. Altered Level of Consciousness

This feature is shown by any answer other than "alert" to the following question: Overall, how would you rate this patient's level of consciousness? (alert [normal], vigilant [hyperalert], lethargic [drowsy, easily aroused], stupor [difficult to arouse], or coma [unarousable])

\* The diagnosis of delirium by CAM requires the presence of features 1 and 2 and either 3 or 4.