

Renato Ferreira Magalhães

Validação de algoritmos de *text mining* em contexto de
oncologia

Dissertação realizada no âmbito do Mestrado em Ciência da Informação,
orientada pelo Professor Doutor Mário Amorim Lopes e coorientada pelo
Professor Doutor Lúcio Lara Santos

Faculdade de Engenharia e Faculdade de Letras
Universidade do Porto

Setembro de 2020

Validação de algoritmos de *text mining* em contexto de oncologia

Renato Ferreira Magalhães

Dissertação realizada no âmbito do Mestrado em Ciência da Informação, orientada pelo Professor Doutor Mário Amorim Lopes e coorientada pelo Professor Doutor Lúcio Lara Santos

Membros do Júri

Professor Doutor António Lucas Soares

Faculdade de Engenharia - Universidade do Porto

Professor Doutor Francisco Moreira Couto

Faculdade de Ciências - Universidade de Lisboa

Professor Doutor Mário Amorim Lopes

Faculdade de Engenharia - Universidade do Porto

*Dedico este presente trabalhos aos clínicos do SNS e
em especial aos doentes oncológicos pelos desafios
que enfrentam todos os dias.*

*Nada é mais difícil e, portanto, tão precioso, do que
ser capaz de decidir. (Napoleão Bonaparte)*

Agradecimentos

Ao **Professor Doutor Mário Amorim Lopes**, pelo suporte ao longo desta jornada pela ciência da informação fazendo-me sentir que não estou sozinho no desafio de comprovar que a “máquina” e o “homem” não competirão pelo mesmo lugar. Cada um tem, e continuará a ter, o seu lugar muito especial e diferenciador.

Ao **Professor Doutor Lúcio Lara Santos**, a quem não posso agradecer o suficiente. Sem a sua orientação, disponibilidade e suporte, com certeza não teria chegado aqui. Obrigado pelo que tem feito pelos nossos doentes e pelo Serviço Nacional de Saúde.

Ao **Professor Doutor Alípio Jorge** e ao **Arian Pasquali**, pelas conversas desafiantes e extremamente úteis.

Às anotadoras **Clara Daher** e **Isabel Alves**, pelo tempo e dedicação a um trabalho que para além de absolutamente necessário, poderá decerto contribuir para um bem maior.

À **Rita**, minha mulher, pela motivação, compreensão e ajuda e por acreditar que este trabalho pretende ser parte de uma solução e não de um problema para todos os Clínicos que diariamente lutam para melhorar a qualidade de vida dos seus doentes.

Resumo

O caso de estudo do presente trabalho teve lugar num hospital oncológico que implementou em 2012 um processo clínico eletrónico (EPR). Todos os registos clínicos efetuados no contexto da consulta são efetuados em “texto livre”, isto é, sem anotação semântica, clínica ou não, dos episódios de consultas e internamentos no diário clínico do EPR. Outros documentos fazem parte do EPR, como por exemplo notas de alta, resultados da patologia clínica, resultados de anatomia patológica, ficando agregados a um visualizador de exames e resultados.

Estes registos são utilizados pelos clínicos no apoio à decisão clínica. No entanto, a complexidade e a extensão desta informação tornam por vezes moroso o trabalho manual de caracterização de doente e do seu historial, isto porque não existe uma forma estruturada de realizar os registos nem existe um padrão de escrita por parte dos clínicos, o que não facilita a análise dos dados.

A maior barreira/dificuldade à implementação da metodologia proposta é a identificação (semi-manual) dos termos nos diários clínicos com o objetivo de criar um dicionário que permitisse posteriormente, com auxílio a ferramentas de *text mining* (TM), identificar conteúdo relevante, agregá-lo e apresentá-lo aos clínicos de forma organizada, sendo então possível estudar padrões, sua evolução, respostas a tratamentos, complicações, sobrevivência e com base nesses dados desenvolver aplicações de suporte à tomada de decisão.

Ao nível dos processos clínicos eletrónicos foram encontrados muitos textos desorganizados, não respeitando regras de pontuação, erros ortográficos e recurso frequente a abreviaturas não oficiais ou não reconhecidas.

Para análise e compreensão dos textos do EPR recorreu-se ao processamento de linguagem natural (PLN).

Os resultados quantitativos dos algoritmos treinados com anotações realizadas por especialistas tiveram uma precisão superior a 60%. Não foi calculada a acurácia porque os números de exemplos para cada classe não eram iguais. Realizou-se também uma análise qualitativa aos modelos, da qual se concluiu que a metodologia utilizada permitirá cumprir os objetivos, mas será ainda necessário reforçar a formação dos anotadores, corrigir as anotações e realizar um novo estudo piloto com mais casos e consequente validação clínica.

Palavras-chave: Inteligência artificial, *Machine learning*, *Text mining*, Suporte de apoio à decisão clínica, Processamento de linguagem natural.

Abstract

The case of the present study took place in an oncology hospital that implemented an electronic patient record (EPR) in 2012. All clinical records that occurred in the context of medical consultation are made in "free text", therefore, without semantic annotation, clinical or not, of the episodes of consultations and hospitalizations in the daily clinical notes of the EPR. Other documents are part of the EPR, such as discharge notes, clinical pathology results, pathological anatomy results, these are aggregated to a viewer of exams and results.

These records are used by clinicians to support clinical decision-making. However, the complexity and extent of this information sometimes makes manual patient characterization and history work very time consuming, this happens because there is no structured way to record and no pattern of writing by the clinicians, which does not facilitate data analysis.

The biggest barrier/difficulty of the proposed methodology implementation is the identification (semi-manual) of the terms in clinical diaries with the objective of creating a dictionary that, with the aid of text mining tools to identify relevant content, would later allow to aggregate it and present it to clinicians in an organized way, and then it would be possible to study patterns, their evolution, responses to treatments, complications, survival and based on these data develop applications that support decision-making.

It was found on the electronic clinical processes, disorganized texts, unrespecting scoring rules, spelling errors and frequent use of unofficial or unrecognized abbreviations.

Natural language Processing was used for the analysis and understanding of the texts of the EPR.

The quantitative results of the algorithms trained with annotations performed by specialists had a precision greater than 60%. Accuracy was not calculated because the number of examples for each class were not equals. A qualitative analysis was also performed to the models, from which it was concluded that the methodology used will allow to meet the objectives, but it will be necessary to reinforce the training of the annotators, correct the annotations and perform a new pilot study with more cases and consequent clinical validation.

Keywords: Artificial intelligence, Machine learning, Text mining, Clinical decision support system, Natural processing language.

Lista de ilustrações

Figura 1 - Componentes da decisão clínica (DC)	23
Figura 2 - Inteligência artificial e Machine learning.....	27
Figura 3 - Alguns tipos de redes neurais.....	32
Figura 4 - Processo de descoberta de conhecimento, adaptada Fayyad et al. (1996).....	35
Figura 5 - Exemplo do bag of words.....	39
Figura 6 - Recursos UMLS	46
Figura 7 - Segmentação de frases.....	47
Figura 8 - Exemplo da separação por palavras (tokens).....	48
Figura 9 - Exemplo de anotação para a tarefa NER.....	50
Figura 10 - Princípio de funcionamento do treino com o spaCy	51
Figura 11 - Pipeline do spaCy	52
Figura 12 - Exemplo de NER.....	52
Figura 13 - Exemplo de um conjunto de dados de treino	53
Figura 14 - Passos necessário para a desambiguação de uma entidade	54
Figura 15 - Exemplo de um episódio do EPR.....	63
Figura 16 - Exemplo do diário clínico	64
Figura 17 - Metodologia de validação dos modelos	66
Figura 18 - Modelo baseado em regras	66
Figura 19 - Modelo semi-supervisionado.....	67
Figura 20 - Modelo supervisionado	67
Figura 21 - Exemplo de anotação de texto como sendo "drugs"	68
Figura 22 - Arquitetura técnica conceptual	72

Lista de tabelas

Tabela 1 - Exemplo de matriz de confusão.....	33
Tabela 2 - Quantidade de conceitos suportados pelo UMLS	65
Tabela 3 – Avaliação global do modelo semi-supervisionado	73
Tabela 4 - Resultados do modelo semi-supervisionado por classe.....	73
Tabela 5 – Avaliação global do modelo supervisionado com as anotações do anotador A	74
Tabela 6 - Resultados do modelo supervisionado com as anotações do anotador A por classe	74
Tabela 7 - Avaliação global do modelo supervisionado com as anotações do anotador B	74
Tabela 8 - Resultados do modelo supervisionado com as anotações do anotador B por classe	74
Tabela 9 – Exemplos da avaliação realizada pelo especialista C a algumas das frases que tiveram a mesma anotação por parte dos anotadores A e B	75
Tabela 10 – Exemplos da avaliação realizada pelo especialista C ao modelo semi-supervisionado	76
Tabela 11 - Exemplos da avaliação realizada pelo especialista C aos modelos dos anotadores A e B.....	78
Tabela 12 – Exemplos da avaliação realizada pelo especialista C aos modelos dos anotadores A e B, em frases em que apenas um anotou	80
Tabela 13 – Exemplos de diferenças entre o modelo semi-supervisionado e os modelos dos anotadores A e B.....	81

Lista de abreviaturas e siglas

Sigla	Descrição
BOW	Bag of words
CRF	Conditional random field
cTAKES	Clinical text analysis and knowledge extraction system
DM	Data mining
DC	Decisão clínica
DO	Disease ontology
DST	Doente sentado no cadeirão
EL	Entity linking
EPR	Processo clínico eletrônico
FMA	Foundation model of anatomy
FN	Falsos negativos
FP	Falsos positivos
GATE	General architecture for text engineering
GB	Gradient boosting
GO	Gene ontology
I2B2	Informatics for integrating biology and the bedside
IA	Inteligência artificial
ICD	International classification of diseases
ICD-10	International classification of diseases, 10 ^a revisão
ICPC	The international classification of primary care
IDF	Inverse document frequency
KDD	Knowledge discovery ou knowledge discovery in databases
KDT	Knowledge discovery from textual databases
k-NN	K-nearest neighbors
LDA	Latent dirichlet allocation
LN	Linguagem natural
LOC	Localizações
MCDDT's	Meios complementares de diagnóstico
MedLEE	Medical language extraction and encoding system
MeSH	Medical subject headings
MISC	Diversos
ML	Machine learning
NCI	National cancer institute
NED	Named entity disambiguation
NEL	Named entity linking
NEN	Named entity normalization
NER	Reconhecimento de nomes de entidades

Sigla	Descrição
NERD	Named entity recognition and disambiguation
NLTK	Natural language toolkit
OAE	Ontology of adverse events
OMIM	Online mendelian inheritance in man
ORG	Organizações
PDF	Portable document format
PLN	Processamento de linguagem natural
pLSA	Probabilistic latent semantic analysis
POS	Pessoas
QALYs	Quality-adjusted life years
RadLex	Radiology lexicon
RF	Radom forest
RON	Registo oncológico nacional
RSE	Registo de saúde eletrónico
SADC	Sistemas de apoio à decisão clínica
SIGIC	Sistema integrado de gestão de inscritos para cirurgia
SNOMED	Systematized nomenclature of medicine
SNOMED CT	Systematized nomenclature of medicine clinical terms
SOAP	Metodologia subjetivo, objetivo, avaliação e plano
spaCy	Open-source library for natural language processing in python
SPAM	Sending and posting advertisement in mass
SVM	Support-vector machines
TDM	Text data mining
TF	Term frequency
TM	Text mining
UMLS	Unified medical language system
VN	Verdadeiros negativos
VP	Verdadeiros positivos
W3C	The world wide web consortium

Índice de conteúdo

1.	Introdução	14
2.	Revisão da literatura	17
2.1.	Tomada de decisão clínica.....	17
2.1.1.	Tomada de decisão	17
2.1.2.	Tomada de decisão clínica	20
2.2.	Inteligência Artificial (IA).....	25
2.2.1.	Machine Learning (ML).....	26
2.3.	Text Mining (TM)	35
2.3.1.	Métodos de representação de dados	38
2.3.2.	Processamento de Linguagem Natural (PLN).....	40
2.4.	TM aplicado à saúde.....	43
2.4.1.	Ontologias	44
2.4.2.	Extração de Informação.....	47
2.4.3.	Reconhecimento de nomes de entidades (NER)	49
2.5.	Sistemas PLN no domínio clínico	54
2.5.1.	Clinical History Extractor for Syndromic Surveillance (CHESS).....	57
2.5.2.	Medical Language Extraction and Encoding System (MedLEE).....	58
2.5.3.	clinical Text Analysis and Knowledge Extraction System (cTAKES).....	59
2.5.4.	Extração de entidade em notas clínicas.....	60
2.6.	Sistemas de apoio à decisão clínica (SADC).....	60
3.	Metodologia	63
3.1.	Trabalho inicial.....	63
3.2.	Preparação dos documentos.....	63
3.2.1.	Extração dos PDF para texto	63
3.2.2.	Normalização de frases	64
3.3.	Construção dos modelos.....	65
3.3.1.	Validação dos modelos	65
3.4.	Tarefas adicionais de pós-processamento.....	69
3.4.1.	Deteção de presença ou ausência (Assertion Status Detection).....	69
3.4.2.	Entity Linking (EL)	69

3.5. Tarefa de anotação do SOAP, de frases do EPR, para organização dos resumos num <i>dashboard</i>	69
4. Caso de estudo	70
4.1. Informação sobre o processo clínico.....	70
4.2. Objetivo da implementação do TM	71
4.3. Estrutura conceptual da solução técnica futura	72
4.4. Resultados do trabalho da validação dos algoritmos de TM	73
4.5. Avaliação dos resultados	75
4.5.1. Avaliação do modelo semi-supervisionado com pré-anotação do UMLS	76
4.5.2. Avaliação do modelo supervisionado com base nas anotações dos anotadores A e B	77
4.5.2.1. Comparativo de frases em que ambos os anotadores A e B registaram	78
4.5.2.2. Comparativo de frases em que apenas um dos anotadores registou	80
4.5.3. Comparativo das anotações dos modelos semi-supervisionado, A e B....	81
5. Conclusões e perspetivas de desenvolvimento.....	83
6. Referências bibliográficas.....	84
Anexos.....	92
Anexo 1 – Exemplo de um diário clínico de um doente	93
Anexo 2 – Exemplos do levantamento de siglas que constam nos diários clínicos e respetivo significado	94
Anexo 3 – Exemplos de dificuldades na segmentação das frases, devido a pouca pontuação, pontuação em abreviaturas e também falta de pontuação no final de frases	96
Anexo 4 – Anotações dos conceitos mais importantes	97
Anexo 5 – Documento base para realização da anotação manual por parte de cada anotador especialista	98
Anexo 6 – Análise de registos clínicos.....	99
Anexo 7 - Anotação de frases segundo o SOAP	101
Anexo 8 – Orientações para a anotação SOAP	103
Anexo 9 - Resultados da avaliação realizada pelo especialista C a algumas das frases (83) que tiveram a mesma anotação por parte dos anotadores A e B	105
Anexo 10 - Resultados da avaliação do modelo semi-supervisionado com pré-anotação do UMLS.....	110
Anexo 11 - Resultados do comparativo de frases em que ambos os anotadores A e B realizaram anotação	113

Anexo 12 - Resultados do comparativo de frases em que apenas um dos anotadores realizou anotação.....	118
Anexo 13 – Exemplo de um <i>dashboard</i> para auxiliar o clínico na tomada de decisão	120

1. Introdução

As atividades no sector da saúde baseiam-se em informações adquiridas de forma sistemática e constante e que auxiliem na organização do conhecimento formado por essas. Essas informações contribuem para o processo de tomada de decisão sobre o possível diagnóstico e seu plano de tratamento.

Vários autores ressaltam a incipiente utilização de sistemas de informação e apoio à decisão clínica pelos profissionais de saúde (Coiera, 1998; Regueira, 2018; Shortliffe, 1990). Apesar disso, as áreas das tecnologias continuam a desenvolver sistemas de organização e divulgação de dados e informações clínicas cada vez mais complexos e integrados.

Os hospitais acumulam grandes volumes de dados, os quais contêm o estado de saúde dos doentes, o diagnóstico e a terapêutica, assim como os resultados de meios complementares de diagnóstico e a resposta aos tratamentos prescritos.

Atualmente, os profissionais têm à sua disposição uma panóplia de sistemas de informação, nem sempre conectados e nem sempre de acesso direto ou simples. Nas últimas décadas, a introdução de registos de saúde eletrónicos aumentou diariamente a quantidade de informação registada, tornando-se um padrão na prática clínica, nos quais se passaram a incluir documentos escritos por médicos, enfermeiros e outros profissionais em contacto com o doente (Feldman et al., 2016). Assim, os tempos reduzidos para a realização das consultas, a repetição de tarefas e a forma não organizada da informação do processo clínico dificulta, acreditamos, a análise da informação anterior e o seu correto registo, o que terá influência na tomada de decisão dos clínicos (Ordem dos Médicos, 2019). Esta falta de tempo para o registo e acesso à informação aumenta ainda o risco de burnout destes profissionais (R. L. Gardner et al., 2019) levando a que estejam menos disponíveis para uma maior adesão aos registos de forma eletrónica.

É frequente ouvirmos no nosso dia-a-dia que certa informação não se regista ou não se consulta porque o sistema de informação solicita demasiados “clicks” (Levinson et al., 2017) ou porque é “lento”, mas na verdade acreditamos que, chegando às variáveis preditoras necessárias e que poderão ser importantes na tomada de decisão de um médico poderemos criar um sistema de apoio à tomada de decisão que cumpra efetivamente o seu verdadeiro propósito de simplificar, apoiar ou mesmo interferir na sua tomada de decisão clínica (Walsh et al., 2019).

Como já referido, são, portanto, produzidas quantidades de dados elevados, registadas em texto livre sem obedecer a uma organização específica. É neste sentido que se tem vindo a desenvolver na área da saúde, e muito concretamente em meios hospitalares, sistemas que

permitem organizar a informação clínica produzida.

Um dos objetivos deste trabalho é tentar demonstrar que a utilização de técnicas como o *text mining* (TM) e o processamento de linguagem natural (PLN), através da sua classificação, quantificação e organização, passam a disponibilizar a informação de modo categorizado e padronizado, possibilitando assim a construção de um resumo estruturado para auxiliar o clínico na tomada de decisão. Para o efeito serão utilizados ficheiros PDF com os processos clínicos em texto livre (anexo 1) com aplicação de modelos de *Machine Learning* (ML) supervisionados para a identificação de procedimentos, medicamentos e doenças. Para a construção do modelo serão utilizados dados anotados por especialistas. Depois de treinado, será validado e avaliada a acurácia do modelo.

Este trabalho assenta assim na premissa de que a informação estruturada apresentada no momento da tomada de decisão influencia e melhora a tomada de decisão clínica.

Assim, a pertinência do presente trabalho centra-se em premissas, tais como: o médico toma diariamente decisões complexas e baseada em evidências (documentos fonte e avaliação clínica do doente) no contexto da consulta; as ferramentas informáticas disponíveis para o médico, nomeadamente a forma como permitem aceder à informação, influenciam a tomada de decisão clínica; o desconhecimento de como utilizar o sistema de informação e a dificuldade de preenchimento dos dados e o formato de retorno das informações podem contribuir para a resistência do profissional para o seu uso e por outro lado levar a tomadas de decisão menos eficazes e ao *burnout*; um sistema de informação que permita ao médico aceder de forma mais direta e rápida aos dados identificados como sendo cruciais no seu processo de tomada de decisão clínica aumenta a motivação do médico, proporcionado a sua utilização, e aumenta a sua produtividade e eficácia; perante a quantidade de dados que estão inseridos nos sistemas de informação clínicos, o uso de ferramentas de inteligência artificial (IA), através de algoritmos definidos por especialistas, poderá revelar-se uma forma capaz de analisar os dados e tomar decisões, já que tem na sua base, a capacidade de as máquinas simularem características humanas (Lobo, 2018).

Pretende-se com este trabalho fazer a ponte entre o clínico, que tem os conhecimentos específicos da área da saúde, e o programador, que domina as tecnologias, disponibilizando informação organizada, que se encontrava registada em texto livre não estruturado, permitindo um acesso mais rápido, auxiliando no processo de tomada de decisão.

Para uma melhor compreensão, este trabalho está organizado pelos seguintes capítulos; no primeiro é realizada uma introdução, onde se efetua o enquadramento do âmbito e tema da dissertação, assim como a descrição dos objetivos propostos, seguidamente no segundo

capítulo, é realizada a revisão da literatura para as temáticas da tomada de decisão clínica, inteligência artificial, *text mining*, sistemas de apoio à decisão clínica e sistemas de processamento de linguagem natural no domínio clínico. O terceiro capítulo é descrita a metodologia de validação clínica de algoritmos de TM em contexto de oncologia. O quarto capítulo é apresentado o caso de estudo, com o enquadramento da Unidade de Saúde e os dados utilizados. No quinto capítulo são apresentadas as conclusões, perspectivas de desenvolvimento e ainda algumas sugestões. Por fim, será referenciada a bibliografia que serviu de suporte à elaboração deste trabalho, bem como os anexos, que complementarão as atividades descritas, permitindo uma melhor compreensão das mesmas.

2. Revisão da literatura

2.1. Tomada de decisão clínica

2.1.1. Tomada de decisão

A tomada de decisão é uma área clássica da Psicologia Cognitiva, tendo uma grande relevância internacional e sendo uma disciplina de interesse para as mais diversas áreas do conhecimento (Liberali, 2012).

O raciocínio clínico (*clinical reasoning*) constitui o núcleo do ato médico, sendo a unidade estrutural base que define o clínico como perito. A visão tão prevalente – nomeadamente nos gestores e políticos da saúde – de que a prática médica é primeiramente uma atividade científica, e consequentemente as decisões clínicas devem ser exclusivamente baseadas na evidência científica, é redutora e não tem em conta a insubstituível necessidade do julgamento, da tomada de decisão (Vaz Carneiro, 2003)

No paradigma da tomada de decisão, o diagnóstico é visto como uma atualização de uma opinião baseada em informação imperfeita e incompleta; a regra normativa aplicável a este processo é o teorema de Bayes que mostra a relação entre uma probabilidade condicional (probabilidade que um evento aconteça, dado que outro aconteceu) e a sua inversa; por exemplo, a probabilidade de uma hipótese dada a observação de uma evidência e a probabilidade da evidência dada pela hipótese.

Assim, como noções básicas apontadas pelos autores Eysenck e Keane (2015) e que são consenso entre os diversos modelos:

1. a mente é formada por processos cognitivos interrelacionados;
2. o principal responsável pela vida mental é a organização do conhecimento;
3. processos cognitivos que sustentam eventos mentais devem ocorrer dentro de uma área específica, pelo menos em algumas situações;
4. já que eventos mentais são abstratos, serão mais facilmente compreendidos utilizando uma análise abstrata e, apesar de depender de substrato neurológico, não se restringem a ele;
5. o ser humano é autónomo e interage com o mundo externo intencionalmente;
6. a interação dá-se por meio da mente, que é um processador de símbolos e significados e que terão relação com as coisas do mundo externo (Neufeld et al., 2011).

À luz da psicologia cognitiva, os fatores envolvidos na tomada de decisão dependem da importância da decisão. Normalmente tendemos a avaliar a qualidade das nossas decisões em

termos das consequências. Contudo, por vezes esta avaliação tem as suas limitações, como se comprova, no caso, por exemplo, de um cirurgião que diz “A operação foi um sucesso, mas infelizmente o doente faleceu”. De facto, a decisão pode ser boa com base na informação disponível quando se a toma, mesmo que seja pobre em termos das suas consequências (Eysenck & Keane, 2015). Se a informação disponível for insuficiente, então poderá ser necessário o julgamento (uma avaliação da probabilidade de um determinado evento ocorrer baseado em informação incompleta).

No nosso dia-a-dia a força das nossas crenças é frequentemente aumentada ou diminuída por novas informações. Então como disciplinar a intuição? Foi a Thomas Bayes¹ a quem se atribuiu a primeira contribuição para um grande problema: a lógica de como as pessoas deverão mudar a sua opinião à luz da evidência (Eysenck & Keane, 2015; Kahneman, 2012). A regra de Bayes especifica como as crenças anteriores deverão ser combinadas com a diagnosticidade da evidência, o grau a que é favorecida a hipótese sobre a alternativa. Na sua regra há duas ideias a reter: as taxas básicas importam, mesmo em presença de evidência acerca do caso em questão; as impressões intuitivas da diagnosticidade da evidência são muitas vezes exageradas. De acordo com este Teorema precisamos avaliar as probabilidades das duas hipóteses antes dos dados serem obtidos.

Frequentemente utilizamos as taxas básicas no nosso dia-a-dia quando possuímos conhecimento causal relevante. Também usamos essa informação quando essa é vantajosa para nós, mas ignoramo-la se a mesma não o é (Eysenck & Keane, 2015).

Muitas investigações envolveram o efeito enquadramento no qual as decisões são influenciadas por aspetos irrelevantes da situação.

O efeito enquadramento acontece quando a decisão do indivíduo é dependente da forma como as opções são apresentadas. Especificamente, as escolhas das pessoas, quando confrontados com problemas de decisão idênticos, são influenciadas quando são enquadradas numa visão positiva das coisas (perspetiva de ganhos) versus uma visão negativa (perspetiva de perdas), são muitas vezes contraditórias.

O problema da doença asiática descrita por Tversky e Kahneman (1981) é um exemplo clássico do efeito enquadramento. Os participantes foram convidados a escolher entre uma opção absoluta (ou seja, com certeza) ou uma probabilística (ou seja, envolvendo risco) para salvar

¹ Thomas Bayes (ca. 1701 — 7 de abril de 1761) matemático inglês, conhecido por ter formulado o caso especial do teorema de Bayes.

vidas (enquadramento positivo) ou minimizar mortes (enquadramento negativo) de uma doença incomum. Deste estudo concluíram que as pessoas parecem, de maneira geral, apresentar uma propensão ao risco quando confrontados com problemas enquadrados de forma negativa e avessos ao risco quando apresentados aos mesmos problemas, mas enquadrados de forma positiva (Tversky & Kahneman, 1981).

Muitos estudos suportaram o que previa a Teoria da Prospeção, na qual as pessoas sobrevalorizam a probabilidade de eventos raros na tomada de decisão. Contudo, aos participantes era geralmente fornecida uma descrição da possibilidade de resultados e a sua probabilidade associada. Na vida real as pessoas tomam muitas das suas decisões com base na sua limitada experiência pessoal e os médicos não serão exceção.

Os estudos nesta área apontam para que o peso dos eventos raros na tomada de decisão seja influenciado pela informação precisa transmitida aos participantes. Quando as decisões são baseadas em descrições, os participantes sobrevalorizam os eventos raros. Quando as decisões são baseadas na experiência, os participantes subvalorizam os eventos raros (Eysenck & Keane, 2015).

Para ultrapassar esta questão deveríamos ter acesso a informação detalhada e de forma rápida no momento da tomada de decisão, mais ainda quando nos encontramos com decisões importantes e complexas, como é o caso de algumas decisões dos médicos que envolvem muitas vezes a escolha de um diagnóstico que pode matar ou salvar pessoas, sendo necessária a capacidade dos médicos chegarem a diagnósticos corretos e precisos rapidamente. Aliás, é precisamente uma das estratégias que distingue um médico iniciante de um médico experiente a utilização por estes últimos de processos mais rápidos, quase automáticos e de forma mais eficaz confirmada, por exemplo, pelo estudo dos movimentos oculares durante a leitura de imagens radiológicas (Kundel et al., 2007). Claro que estes estudos em laboratório apresentam algumas limitações, podendo não representar o mesmo processo de tomada de decisão em contexto real.

Uma tomada de decisão complexa envolve os seguintes pontos (Wright, 1984):

1. Identificar atributos relevantes para a decisão;
2. Decidir o peso de cada atributo;
3. Listar todas as opções a considerar;
4. Relacionar cada opção com cada atributo;
5. Obter a utilidade total e seleccionar aquele com o maior peso total.

Na vida real, muitas vezes não seguimos esta sequência e optamos por reduzir o número de opções, eliminando alguns aspetos o que poderá conduzir à escolha de uma opção que poderá não ser a melhor opção. Contudo, os mais experientes frequentemente consideram apenas uma opção e tomam decisões rápidas e intuitivas (Eysenck & Keane, 2015; Kahneman, 2012). Tomar uma decisão pode levar os decisores a esquecer informação factual relevante e a exponenciar o que suportará a sua decisão. O pensamento inconsciente é mais útil do que o pensamento consciente na tomada de decisões complexas. Contudo, a tomada de decisão é por vezes melhor se aliarmos ambos os pensamentos.

A eliminação de alguns aspetos na tomada de decisão provou ter sucesso quando os sujeitos têm de escolher entre muitas opções. Contudo, tal não acontece quando existem poucas opções.

Ora, acreditamos que a tomada de decisão clínica é um evidente exemplo de uma tomada de decisão complexa.

2.1.2. Tomada de decisão clínica

O estudo da tomada de decisão levanta questões de racionalidade e lógica da decisão, bem como de crenças e preferências das pessoas.

As análises da tomada de decisão distinguem, em geral, escolhas com e sem risco (Kahneman, 2012). Este autor defende ainda as duas formas de pensar: o pensamento rápido, intuitivo e emocional, e o devagar, lógico e ponderado. Kahneman (2012) mostra-nos a capacidade do pensamento rápido, a sua influência persuasiva nas nossas decisões e até onde podemos ou não confiar nele. O entendimento do funcionamento dessas duas formas de pensar pode ajudar nas nossas decisões pessoais e profissionais.

Recordemos a história da Dra. Apgar². Um dia um interno perguntou como é que a Dra. Apgar faria uma avaliação sistemática de um recém-nascido, ao que prontamente respondeu com 5 variáveis e 3 pontuações e percebeu que tinha encontrado uma forma consistente de rapidamente tomar a decisão de intervir de modo imediato nos recém-nascidos mesmo antes dos primeiros sintomas se manifestassem. Até à Dra. Apgar não havia um procedimento

² Virginia Apgar (Westfield, 7 de junho de 1909 — Nova Iorque, 7 de agosto de 1974) foi uma médica norte-americana que se especializou em obstetrícia e anestesiologia.

pradonizado e os sinais de perigo eram muitas vezes ignorados e muitos recém-nascidos morriam. Os testes de Apgar ainda hoje são utilizados todos os dias em todas as salas de parto.

O recente livro de Atul Gawande (2015), o *Efeito Checklist*, fornece muitos outros exemplos das vantagens da informação sistematizada e acessível e utilização das listas de verificação e das regras simples.

O autor desenvolve um ensaio sobre um mundo cada vez mais dependente da tecnologia e onde, paradoxalmente, o erro humano pode ser evitado de uma forma tão simples. Defende ainda que a forma esquematizada como a informação aparece e/ou é registada, nomeadamente em forma de *checklist*, parece defender qualquer pessoa contra o insucesso, mesmo a mais experiente, em muito mais tarefas do que aquelas que pensávamos. Detetam as falhas mentais de todos nós e podem ser instrumentos poderosos no apoio à decisão clínica mesmo em situações complexas. E, no caso dos médicos, os doentes são distintos, são complexos. Não há dois doentes com pneumonia iguais. Mesmo com a mesma bactéria, a mesma tosse, com o mesmo antibiótico, um doente pode melhorar e o outro não. A medicina integra assim todo o tipo de problemas, os simples, os complicados e os complexos (p. ex. problemas que nunca aconteceram antes ou não possuem soluções conhecidas).

O raciocínio clínico é uma função essencial da atividade médica e a eficiência da prestação de cuidados é altamente dependente da análise e síntese adequadas dos dados clínicos e da qualidade das decisões envolvendo risco e benefícios dos testes diagnósticos e do tratamento (Réa-Neto, 1998).

Vários estudos têm mostrado que quando um médico está na presença de um doente que apresenta um problema, o médico utiliza um método cognitivo para resolver problemas muito semelhante ao método hipotético-dedutivo e que se separa em duas grandes componentes: o conteúdo (corresponde à base de conhecimento rica e extensa que reside na memória do médico) e o processo (corresponde ao método de aplicação do conhecimento utilizado pelo médico na procura de uma solução para o problema do doente).

Para Réa-Neto (1998), o que os médicos fazem é utilizar hipóteses sugeridas pelo problema para, a partir delas, determinar exatamente quais as informações que são necessárias para deduzir qual das hipóteses é a mais correta.

Partindo do pressuposto de que o sucesso na resolução do problema do doente é o objetivo de todo o processo, podemos inferir que o sucesso estará na obtenção de um diagnóstico correto e de um tratamento efetivo.

Desta forma, a resolução desse problema poderá ser caracterizada por duas grandes fases de tomada de decisão: a designação de um diagnóstico e a seleção de um tratamento.

Este processo de tomada de decisão envolve ainda processos de seleção de perguntas a fazer, decidir que respostas são confiáveis, interpretar um sinal físico, selecionar um ou mais meios complementares de diagnóstico e escolher uma ou mais formas de tratamento da doença (Réa-Neto, 1998).

Podemos dizer que o primeiro elemento na tentativa de solucionar um problema clínico é obter informações relacionadas com o problema do doente. Consequentemente, surgem várias hipóteses sobre a questão. Neste processo, o médico pode basear estas suas hipóteses em dados estatísticos de prevalência das possíveis explicações para cada dado ou conjunto de dados obtidos.

Mas mais comum é os médicos recorrerem a heurísticas (associações rápidas que os médicos fazem entre dados e explicações potenciais) baseadas em experiências passadas, utilizando os conhecimentos que possuem na memória a longo prazo.

Perante os dados, o médico é capaz de tomar decisões, de estabelecer um diagnóstico. A partir daí começa a desenhar o seu plano de tratamentos. Assim se conclui que a fase da análise dos dados, da informação disponível, é crucial para a correta determinação da estratégia de cura do doente.

O bom suporte à decisão clínica (DC), para além do *know-how* do clínico, depende de dados disponíveis sobre o doente e do conhecimento dos procedimentos terapêuticos eficazes de resolução de problemas. Nesse cenário, as decisões em saúde sofrem ainda com a falta de padronização na terminologia médica, com a diversidade de formato de dados e rotinas e com a velocidade com que as inovações são divulgadas e devem ser incorporadas pelos médicos.

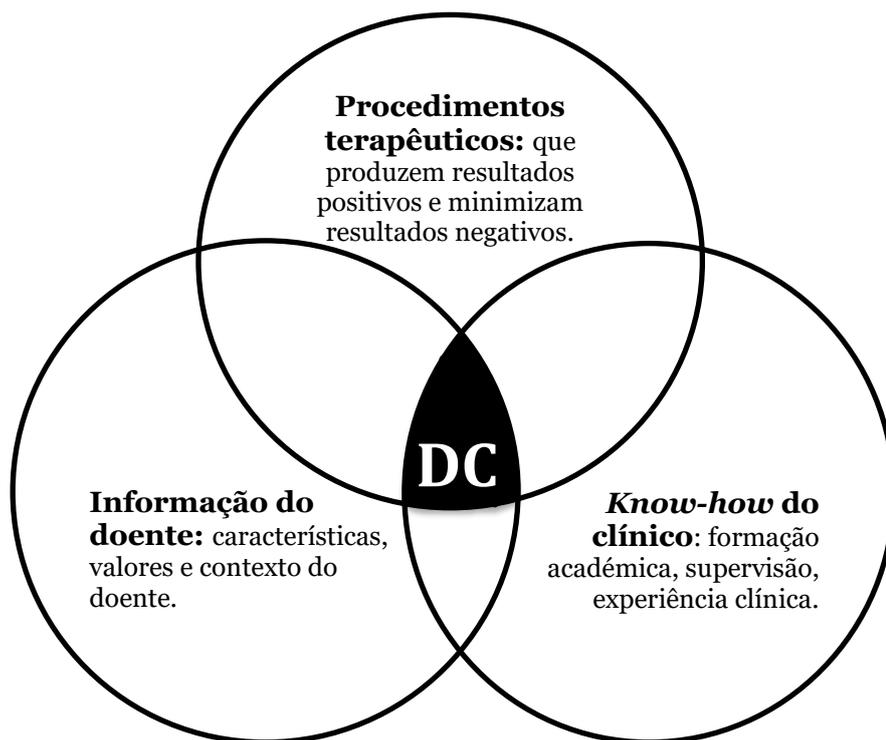


Figura 1 - Componentes da decisão clínica (DC)

Ledley e Lusted (1959) verificaram que a maior parte dos erros médicos ocorrem por omissão, ou seja, por falta da análise de uma componente que iria influenciar a decisão do Médico.

Segundo Valusek (2002), a principal questão de fiabilidade e segurança do doente em relação ao médico é a desintegração da informação – desde o momento em que o profissional a recebe ao momento em que a processa -, capaz de causar o erro por mecanismos como a distração ou o desvio.

Não só a fase do diagnóstico, mas também a fase da definição do tratamento, envolve a tomada de decisão dos médicos e que, fazendo a analogia ao modelo de Simon (1997), corresponderá à fase da escolha que visa a seleção de uma das alternativas identificadas na fase anterior.

São vários os dados importantes para o processo de atendimento de um doente, tais como: alguns dados demográficos do doente, antecedentes familiares ou hereditários, dados vitais, consultas anteriores e exames complementares (Werner et al., 2009). Em paralelo, é necessário conhecimento clínico para análise dos exames, como por exemplo valores de referência de resultados de exames e doses de fármacos. É com base nestes dados que o

processo de decisão clínica ocorre. Os dados clínicos são o conjunto das observações feitas pelo Médico nos processos de anamnese³, exame físico e exames complementares.

É comum os médicos utilizarem a metodologia *SOAP*: *S* – *subjetivo, a razão da visita*; *O* – *objetivo, as descobertas ao observar*; *A* – *avaliação, análise* e *P* – *plano, tratamento a realizar*. O SOAP é um sistema racional de resolução de problemas: primeiro porque é a tradução escrita resumida sem informações supérfluas de um episódio clínico e, segundo, porque as suas componentes “S”, “O”, “A”, “P” se coadunam com os passos de uma consulta médica.

A informação sistematizada permite poupar muito tempo na consulta e no registo de dados dos doentes (Enelow et al., 1999). Para Queiroz (2009), o registo é sobretudo um suporte informativo, refletindo o estado de saúde do utente, e um meio de comunicação, permitindo comunicar factos relevantes. Contribuí para as cinco áreas no âmbito da atividade médica: 1) Auxiliar de memória; 2) Transferência de informação; 3) Avaliação de qualidade; 4) Formação; 5) Investigação.

A possível complexidade de cada doente, quando alargada a muitos, obriga o médico a auxiliar-se de uma memória escrita, extensível, dinâmica e eficaz. Os dados armazenados servem o médico assistente ou outro que, transitória ou definitivamente, o substitua, na medida em que o processo clínico acompanha a transferência do doente, no Serviço Nacional de Saúde. É também do processo clínico que são transferidos dados para outras bases de dados (Registo de Saúde Eletrónico-RSE, Registo Oncológico Nacional-RON, Sistema DiabCare, entre outros) ou elaborados relatórios para referência, para outras finalidades ou a pedido do doente.

Esta quantidade e complexidade do formato dos dados de saúde atraiu muitas disciplinas ligadas ao estudo e organização de informações. Para Wong (2000), o sector da saúde é considerado essencialmente um ‘*negócio de informação*’. Daí a necessidade emergente de dotar as Instituições de Saúde em geral, e os médicos em particular, de sistemas de apoio à decisão clínica mais fiáveis e eficazes.

³ Anamnese é uma entrevista realizada pelo médico ao doente, que tem a intenção de ser um ponto inicial no diagnóstico de uma doença.

2.2. Inteligência Artificial (IA)

No Verão de 1956, um grupo de jovens académicos da área da matemática e da lógica que se interessava pela potencialidade dos computadores ao nível da resolução de problemas reuniu-se na Universidade de Dartmouth para discutir interesses comuns. Entre os presentes estavam os académicos que trabalhavam no que mais tarde se viria a designar de IA, incluindo os quatro investigadores designados como os seus pais fundadores: John McCarthy, Marvin Minsky, Allen Newell e Herbert Simon (H. Gardner, 2002).

Inicialmente a IA tinha duas abordagens, a “baseada em regras” e a das “redes neuronais”. Na primeira a aplicação tinha por base uma série de regras lógicas, se X, então Y. Já nas “redes neuronais” o importante eram os dados e a capacidade de computação. Os dados serviam para treinar o programa a reconhecer padrões e a capacidade de computação para analisar com rapidez as situações apresentadas. A título de exemplo, perante um problema de identificação de uma laranja, na abordagem “baseada por regras” uma forma de validação seria através do teste de existência de uma forma circular. Na abordagem “redes neuronais”, seriam carregadas previamente inúmeras imagens com exemplos de presença ou ausência de uma laranja. Quando apresentada a nova imagem com um fruto, o programa tem a capacidade de dizer se é ou não uma laranja.

A IA tem vindo a ganhar domínio expressivo em diversas áreas do conhecimento humano, não se aplicando apenas no âmbito das ciências computacionais. Estes avanços tecnológicos permitem aplicar tecnologias de génese racional em áreas mais humanas e empíricas, como é o caso da medicina (Andrade, 2018). Importa ainda acrescentar que a IA não é em si uma tecnologia, mas sim uma coleção de tecnologias.

Kaplan & Haenlein (2018) definem a IA como *“uma capacidade do sistema para interpretar corretamente dados externos, aprender a partir desses dados e utilizar essas aprendizagens para atingir objetivos e tarefas específicas através de adaptação flexível”*.

A IA médica inicialmente preocupou-se com a construção de programas com o objetivo de realizar diagnóstico e fazer recomendações terapêuticas (Clancey & Shortliffe, 1984).

As máquinas através de processos de IA são capazes de entender características humanas, o que potencia a análise de uma panóplia de variáveis de um modo ilimitado e organizado. A IA apoia assim a informação clínica e ajuda a estipular métodos e diagnósticos mais eficazes, menos invasivos e menos dispendiosos. A IA tem apresentado potencial para ser mais precisa do que os clínicos em diagnósticos de especialidades como radiologia, dermatologia e também na geração de modelos prognósticos e na execução de intervenções cirúrgicas (Goldhahn et al.,

2018). A IA aplicada à medicina permite que os médicos tomem decisões mais céleres, assentes numa maior probabilidade de sucesso ao nível da prescrição terapêutica (Andrade, 2018).

A IA tem ao seu dispor “*toda a informação disponível sobre cada entidade clínica, toda a informação epidemiológica e estatística, todas as bases de dados relativas a casos semelhantes e poderá decidir em frações de segundos o rumo a seguir. Os algoritmos ou árvores de decisão (...) poderão ser seguidos pela IA, com maior rapidez, maior número de variáveis analisadas e, provavelmente, maior acuidade do que qualquer médico seria capaz*” (Andrade, 2018, p. 54).

São vários os exemplos de utilização de IA na área da saúde. A título de exemplo, a *Food and Drug Administration*, a agência de aprovação de tecnologias da saúde dos EUA, aprovou um programa informático de IA, designado como *Imagen OsteoDetect*, para diagnóstico das fraturas do pulso em adultos, através de radiografias bidimensionais e mais recentemente aprovou um software, *Caption Guidance*, que usa IA para orientar profissionais médicos na aquisição de imagens de ultrassom cardíaco ou ecocardiografia para exame ao coração (U.S. Food and Drug Administration, 2018, 2020).

Com o aumento da capacidade de processamento, a IA tem vindo a ser gradualmente aplicada em oncologia, tanto em ferramentas de diagnóstico, medicina de precisão e também redução dos custos dos ensaios clínicos. Alguns especialistas acreditam que, com o acesso à IA, os médicos terão mais tempo para interagir com os doentes, sendo certo que esta mudança só ocorrerá se a duração das consultas permanecer a mesma e a IA não for usada para cobrir a falta de pessoal (“What to expect from AI in oncology,” 2019).

2.2.1. Machine Learning (ML)

ML é uma tecnologia de IA que possibilita aos sistemas a capacidade de aprenderem automaticamente com a experiência (figura 2), ou seja, através de treino, sem a necessidade de programação explícita.

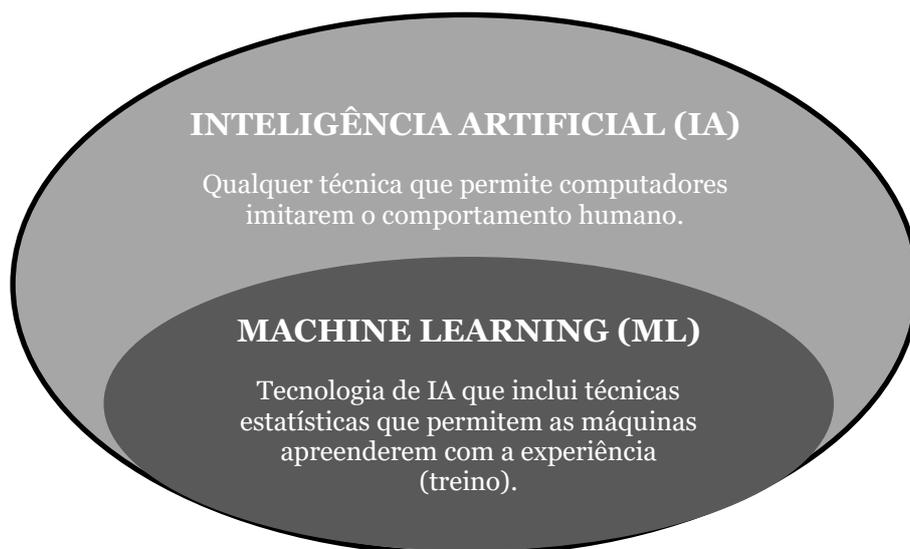


Figura 2 - Inteligência artificial e Machine learning

ML tem sido o termo utilizado para explicar o processo através do qual é possível que os computadores lidem com grandes quantidades de dados e informações, aprendendo com os dados e com a experiência (Kavakiotis et al., 2017; Mjolsness & Decoste, 2001).

Diferentes autores consideram os conceitos de ML e IA semelhantes, pois ambos permitem a aprendizagem, sendo esta competência a principal característica de uma entidade considerada inteligente (Andrade, 2018; Kavakiotis et al., 2017; Kourou et al., 2015). Este ramo da IA estuda as ferramentas e os métodos para identificar padrões nos dados.

A criação de algoritmos, fórmulas, bem como a identificação de padrões de dados e a criação de conexões entre os dados, permite que os computadores consigam executar tarefas sem a ajuda do homem e adaptar-se a novas circunstâncias, de modo independente. Os algoritmos, baseados em análises estatísticas, permitem prever respostas e sugerem o melhor resultado preditivo, com menor probabilidade de erro (Kourou et al., 2015). Deste modo, as técnicas de ML visam construir sistemas computacionais capazes de se adaptar e aprender a partir da experiência (Kavakiotis et al., 2017).

Uma das vantagens que tem vindo a ser associada às técnicas de ML é a capacidade de processar quantidades ilimitadas de informação, provenientes de fontes diversas, em tempo real. Paralelamente, os modelos criados aprendem com resultados anteriores e podem melhorar continuamente as suas previsões com base em novas informações, ajustando-se para dar uma resposta de acordo com os dados disponíveis para análise (Wiens & Shenoy, 2018).

Em oncologia, os métodos de ML têm sido aplicados para criar modelos preditivos da suscetibilidade ao cancro, reincidência e taxa de sobrevivência, com o objetivo de permitir uma tomada de decisão precisa (Kourou et al., 2015). Em processamento de imagens, os métodos de ML podem ser usados como uma ferramenta de suporte para, por exemplo, reduzir a complexidade dos protocolos de imagem de cancro e aumentar a reprodutibilidade e padronização na imagem clínica.

O ML pode ser aplicado com recurso a aprendizagem supervisionada, aprendizagem não supervisionada, semi-supervisionada e aprendizagem por reforço.

Na **aprendizagem supervisionada** são utilizados modelos preditivos, com utilização de dados pré-categorizados ou numéricos, onde são apresentados à máquina exemplos de entradas e saídas desejadas com o objetivo de aprender uma regra geral para mapear as entradas para as saídas. Para dar conhecimento à máquina são definidos manualmente dados e com estes é realizado o treino do modelo, processo que pode ser efetuado por:

- Classificação, que consiste em criar um mapeamento (classificador) de cada objeto de uma coleção (*dataset*) num conjunto de categorias ou classes (onde se quer aprender). Para ajudar a prever uma classe são utilizadas variáveis, denominadas atributos (*feature*), cujos valores podem ser determinados;
- Regressão, em que a saída é um valor e não uma categoria, sendo a principal dificuldade encontrar a função que melhor represente o problema.

Neste tipo de aprendizagem, através da classificação ou da regressão, o objetivo é ter uma função que transforme atributos (variáveis) num rótulo. Fazendo uma analogia com as funções seria $y=f(x)$, onde x é uma matriz, em que cada linha representa a amostras dos dados (p. ex. informação sobre uma pessoa) e cada coluna é um atributo (p. ex. idade, peso, género). A saída y será um vetor que contém rótulos no caso da classificação ou valores na regressão.

A classificação depende de dados intensos e corretamente anotados, isto porque o sistema aprende com os dados. Se forem utilizados dados maus na aprendizagem, a máquina irá replicar.

Na **aprendizagem não supervisionada** são utilizados modelos descritivos, com dados não rotulados, onde a máquina não tem resposta para os casos, deixando-a sozinha para encontrar estrutura nas entradas fornecidas, ou seja, descobrir padrões nos dados. Para tal pode ser aplicada a *clusterização*, agrupar os casos em grupos que pareçam ser mais semelhantes (clusters), com base nas características que eles apresentam, ou a redução de

dimensionalidade, diminuir a complexidade dos dados escolhendo uma melhor forma de representá-los.

A **aprendizagem semi-supervisionada** situa-se entre a aprendizagem supervisionada e a não supervisionada, sendo utilizada principalmente em situações em que o custo para etiquetar é bastante alto, uma vez que requer especialistas humanos qualificados para fazer isso.

Na **aprendizagem por reforço** a máquina tenta aprender qual é a melhor ação a ser tomada, dependendo do ambiente em que essa ação será executada. A cada ação está associada uma recompensa ou punição, pretendendo-se evitar ações que geram punição ou recompensa menor. Como exemplo de aplicação são os automóveis autônomos ou jogos eletrônicos.

Existem muitos algoritmos baseados em ML, que foram surgindo ao longo dos anos, tais como Support-Vector Machines (SVMs), Decision Trees, Random Forest (RF), Gradient Boosting (GB), Redes Neurais e k-Nearest Neighbors (k-NN). A escolha de um algoritmo dependerá do tipo de problema. De acordo com o Teorema “No Free Lunch” nenhum algoritmo tem um bom desempenho em todos os dados (Wolpert & Macready, 1997), ou seja, um único algoritmo de ML não dará o melhor desempenho em todos os problemas (Kotthoff, 2016). Perante um problema e dada a existência de vários algoritmos de ML é necessário ter em consideração a precisão versus interpretabilidade, em que a precisão consiste no grau de acerto e a interpretabilidade é a capacidade de indicar aos utilizadores como uma determinada decisão ou resposta é obtida (Lee & Shin, 2020).

Na aprendizagem **supervisionada** os algoritmos funcionam como um mapeador, ou seja, mapeia uma variável de entrada x numa variável de saída y . Podem ser usados vários modelos, tais como:

- Naive Bayes: modelo probabilístico baseado no teorema de Bayes, sustentado na probabilidade de cada evento ocorrer, desconsiderando a correlação entre variáveis. Este classificador pressupõe uma independência entre os atributos dos dados. Quando aplicado a dados de texto tem em consideração a frequência das palavras usadas, não considerando a ordem das palavras no documento, mas sim as repetições de palavras. As várias palavras que compõe o documento aparecem independentemente uma da outra, dada a classe do documento (p. ex. notícias, saúde).
- Decision Tree (árvores de decisão): tabela de decisão em forma de árvore, com nós, ramos e folhas, que abrange processos de classificação (sim/não) e regressão (dados

contínuos). É usado como um modelo para tomar decisões e as possíveis consequências, onde os dados são classificados de forma hierarquizada. Partem de um conjunto de testes que são posteriormente subdivididos em nó ou ramos. Para a sua representação é utilizado um gráfico ou modelo de decisão e consiste numa série de declarações condicionais, que têm um seguimento em função do resultado da condição, verdadeiro ou falso. Após o período de classificação estes algoritmos têm a capacidade de aprendizagem.

- Random Forest (RF): consiste num grande número de árvores de decisão individuais (*decision tree*), combinando-as para obter uma predição com maior exatidão. As árvores são criadas uma de cada vez, onde cada nova árvore ajuda a corrigir erros cometidos por árvores previamente treinadas. É usado para tarefas de classificação e também de regressão.
- Gradient Boosting (GB): é outra técnica para realizar tarefas supervisionadas, como classificação e regressão, cujo objetivo é um modelo mais poderoso e robusto através da criação de vários modelos de árvore de decisão considerados mais simples (ou fracos), por combinação dos resultados desses vários modelos fracos. As árvores são treinadas de forma independente, usando uma amostra aleatória dos dados.
- Regressão Linear de Mínimos Quadrados: A técnica de regressão tem por objetivo prever algo contínuo e numérico, para que se possa identificar oscilações entre os valores esperados. Os mínimos quadrados consistem no tipo de métrica de erro que está a ser minimizado. É traçada uma linha reta e cada dado é representado por um ponto ao longo da linha, ficando acima ou abaixo dela, somando-se as distâncias verticais entre a linha e os pontos. A linha ajustada seria aquela em que esta soma de distâncias é a menor possível.
- Support Vector Machines (SVM): este algoritmo de classificação baseia-se em características/combinções lineares. Tende a encontrar “um bom separador linear entre” classes distintas. Partindo de um conjunto de dados de treino, representados num determinado espaço vetorial, encontra o melhor hiperplano (que em duas dimensões é simplesmente uma linha), sendo este o limite da decisão, que divide as duas categorias.
- Conditional Random Field (CRF): este algoritmo utiliza uma estrutura probabilística para rotular e segmentar dados estruturados, como por exemplo prever sequências de palavras tendo em consideração o contexto. Não pressupõe apenas que os dados são dependentes uns dos outros, mas também considera os dados seguintes ao aprender um padrão.

- K-Nearest Neighbor (k-NN): é usado para resolver problemas de classificação (sim/não, 1/0, spam/não spam, verdadeiro/falso, grávida/não grávida) e regressão (dados contínuos, ex. temperatura), no qual usa medidas de distância para classificar informação. Documentos pertencentes à mesma classe estão “mais próximos” ou são mais similares. Para a entrada são dados k exemplos de treino mais próximos no espaço de características e a saída de classificação é uma associação de classe, caso seja regressão é a média dos valores dos k vizinhos mais próximos.

Os métodos de classificação têm sido utilizados para diversos objetivos, tais como, filtros de spam, reconhecimento de caracteres, análise de sentimentos ou ainda pesquisa de documentos semelhantes, enquanto os métodos de regressão têm sido aplicados em situações como diagnóstico médico, preço de um imóvel ou de um carro baseado na quilometragem.

Na aprendizagem **não supervisionada** um dos algoritmos mais populares é o *clustering* (agrupamento), que tem como função agrupar em *clusters* (grupos) a informação existente nos documentos por grau de similaridade (Dalianis, 2018). O *clustering* consiste numa técnica que agrupa pontos de dados semelhantes, de forma a que os pontos no mesmo grupo sejam mais semelhantes entre si do que os pontos nos outros grupos. O grupo de pontos de dados semelhantes é chamado de *cluster*. Existem ainda os algoritmos de redução da dimensão (generalização), em que consiste na redução do número de variáveis aleatórias em consideração mediante a obtenção de um conjunto de variáveis principais, e os algoritmos de associação, que consistem em descobrir regras que descrevam grandes porções dos dados, como “pessoas que tiveram A tendem a ter B”.

Segundo Allahyari et al. (2017), a aplicação de *clustering* pode ser efetuada pelos algoritmos Hierarchical Clustering, k-means Clustering e Topic modeling, que são alguns dos modelos utilizados neste tipo de aprendizagem:

- Hierarchical Clustering: é dividido em 2 tipos, o *Agglomerative*, onde inicialmente cada ponto de dados é considerado como um agrupamento individual e, a cada iteração, os *clusters* semelhantes são anexados a outros clusters; e o *Divisive*, ou seja, todos os pontos de dados são inicialmente um único cluster e, em cada iteração, são separados os pontos de dados do cluster que não são semelhantes. Cada ponto de dados separado é considerado como um *cluster* individual.
- k-means clustering: consiste em agrupar pontos de dados semelhantes e descobrir padrões subjacentes.

- Topic modeling: consiste em descobrir os “tópicos” abstratos que ocorrem numa coleção de documentos. Os dois modelos mais utilizados são o “*Probabilistic Latent Semantic Analysis (pLSA)* e *Latent Dirichlet Allocation (LDA)*. (Allahyari et al., 2017, p. 7).

A título de exemplo, métodos de agrupamento podem ser aplicados para detetar segmentos de mercado (p. ex. tipo de clientes) ou detetar um comportamento anormal, enquanto os métodos de redução de dimensionalidade são utilizados para criar sistemas de recomendação (p. ex. recomendar um filme) ou análise de imagens falsas. Os métodos de associação são utilizados em plataforma de compras online ou ainda de posicionamento de compras num supermercado.

Existem ainda algoritmos que funcionam tanto na aprendizagem supervisionada como aprendizagem não supervisionada, como por exemplo:

- Rede neuronal artificial: Estas redes são inspiradas nas redes neuronais biológicas, são um conjunto de modelos matemáticos, que têm por base uma variável de entrada que corresponde aos atributos ou características do problema e que liga a uma variável de saída, que corresponde ao que pretendemos como resultado. Existem várias variantes deste tipo de rede, figura 3, como por exemplo a rede neuronal recorrente, utilizada para analisar dados sequenciais (p. ex. texto ou fala) dado ter um estado de memória interna que pode armazenar informações sobre pontos de dados anteriores (Sahiner et al., 2019).

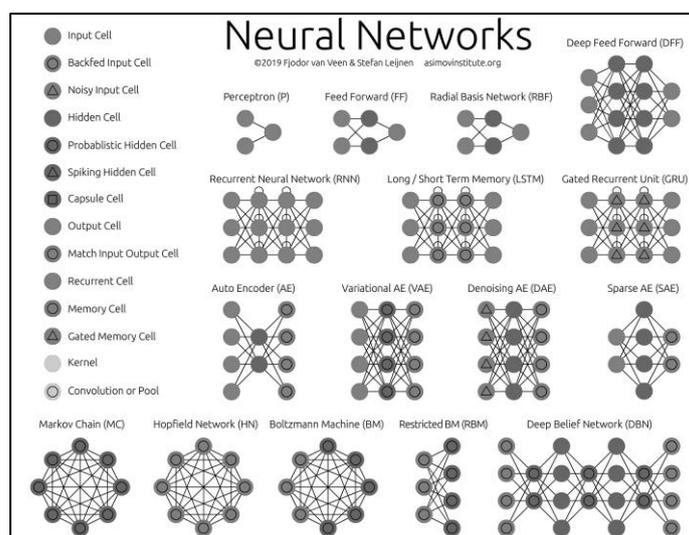


Figura 3 - Alguns tipos de redes neuronais⁴

⁴ Imagem parcial do site <https://www.asimovinstitute.org/neural-network-zoo>, acedido em 19/09/2020.

(Sarkiss & Germano, 2019) na revisão sistemática da literatura realizada sobre as aplicações de ML em neuro-oncologia confirmaram que o método SVM foi o mais usado. Também (Lao et al., 2017; Li et al., 2017) usaram com sucesso o método *radiomics*, que consiste em extrair um grande número de características de imagens médicas radiográficas usando algoritmos de caracterização de dados, para realizar o agrupamento de doentes e prever a sobrevivência a Glioblastoma Multiforme.

De uma forma simplista, a avaliação de um classificador pode ser realizada pelo número de acertos. No entanto, esta simplicidade apenas permitiria identificar pontos fortes e fracos do classificador. Serão, portanto, necessárias outras métricas para realizar a avaliação. Para diferentes tipos de algoritmos, as métricas e a avaliação também são diferentes. Para modelos de classificação existem métricas definidas, tais como:

- A **acurácia** é um dos critérios mais importantes e amplamente utilizado no processo de classificação (Halkidi & Vazirgiannis, 2009), que consiste na percentagem de acerto do classificador. A título de exemplo, pretendemos saber se uma mulher está grávida. Supondo que temos 101 amostras dos resultados de testes de gravidez que o classificador acertou em 91, a acurácia seria $(91/101) \times 100 = 90,10\%$

Apesar da elevada percentagem não é possível afirmar de imediato que o classificador tem um bom desempenho. O erro estaria em situações em que estar grávida, mas o resultado é contrário e vice-versa. Nesta situação seria possível afirmar que existia um erro de classificação. Para facilitar a análise do desempenho é construída uma tabela com o nome matriz de confusão (confusion matrix), em que cada linha representa a classe das amostras que o classificador retornou e cada coluna representa a classe real.

	Sim	Não	Total
Sim	VP= 60	FN=3	63
Não	FP=7	VN= 31	38
Total	67	34	101

Tabela 1 - Exemplo de matriz de confusão

Neste exemplo o classificador indica 3 casos que não estava grávida, mas estava e 7 casos em que indicava que estava grávida, mas não estava.

A nomenclatura aplicada aos valores da tabela com fundo a cinzento é de VP (Verdadeiros Positivos), VN (Verdadeiros Negativos), FP (Falsos Positivos) e FN (Falsos Negativos), onde apenas as duas primeiras classificações estão corretas, ou seja, “estar grávida” ou “não estar grávida”. Para as duas últimas classificações um falso positivo seria afirmar que um homem estava grávido e um falso negativo seria afirmar que uma mulher não está grávida quando existia evidência de gravidez.

Seguindo esta nomenclatura, a acurácia (A) também pode ser calculada da seguinte forma:

$$A = \frac{VP + VN}{VP + VN + FP + FN}$$

- A **precisão** (P) do classificador é a percentagem de predições positivas que estavam corretas. O valor é calculado pelo quociente entre número de VP pela soma dos VP e FP. Um baixo valor P significa grande número de falsos positivos.

$$P = \frac{VP}{VP + FP}$$

- O cálculo da métrica **recall** (R) indica a percentagem de resultados relevantes classificados corretamente, é efetuado através do quociente entre número de VP pela soma dos VP e FN. Um alto valor R significa que os verdadeiros positivos são identificados.

$$R = \frac{VP}{VP + FN}$$

- A eficácia do classificador em reconhecer amostras negativas é efetuada através da métrica **especificidade** (E).

$$E = \frac{VN}{VN + FP}$$

- A eficácia do classificador em reconhecer amostras positivas é efetuada através da métrica **sensibilidade** (S).

$$S = \frac{VP}{VN + FP}$$

- A métrica **f1 score** (F1) é definida como sendo duas vezes a média harmônica entre R e P. O modelo terá melhor desempenho quanto o F1 é mais próximo de 1.

$$F1 = 2 \times \frac{P \times R}{P + R}$$

2.3. Text Mining (TM)

O aumento de dados textuais não estruturados cria um desafio às organizações, dada a necessidade de obtenção de informações para análise. A pesquisa de informação em enormes quantidades de texto torna-se impraticável sem o auxílio de ferramentas informáticas adequadas que permitam a análise, extração de padrões e relações a partir dos dados de texto.

O TM é o processo de explorar e analisar grandes quantidades de dados de texto, não estruturados ou semi-estruturados. TM é um conjunto de técnicas que tem como intuito a identificação de conhecimento em documentos de texto (Silva, 2016). Embora para dados estruturados também possa ser utilizado o TM, são normalmente tratados por *data mining* (DM).

O DM consiste na aplicação de algoritmos específicos para extrair padrões de dados, sendo o processo de compreensão e descoberta de padrões em enormes volumes de dados para recuperar conhecimentos úteis (Fayyad et al., 1996), conforme ilustrado na figura seguinte:

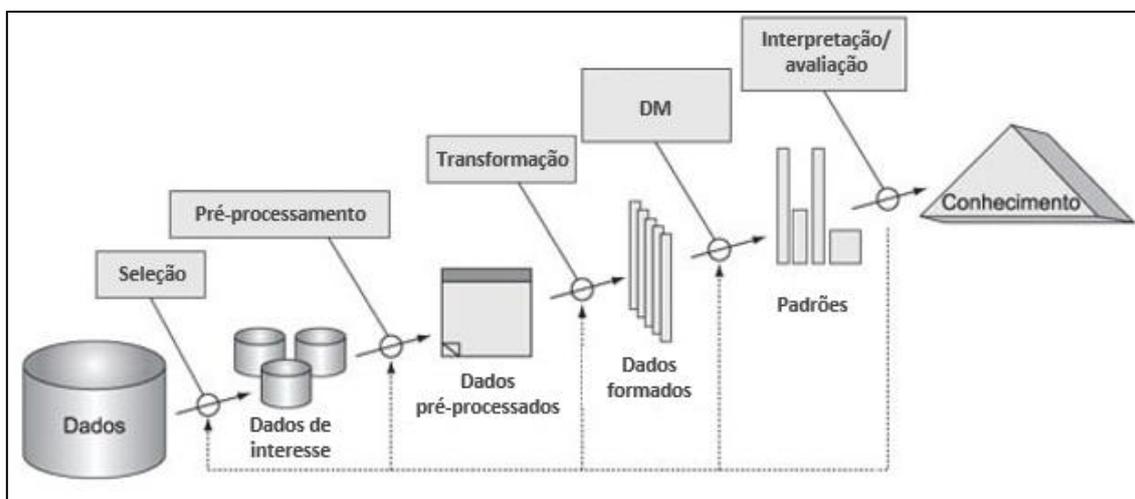


Figura 4 - Processo de descoberta de conhecimento, adaptada Fayyad et al. (1996)

Para Tan (1999), TM é o mesmo que Text Data Mining (TDM) ou ainda Knowledge Discovery from textual databases (KDT), sendo o processo de extrair padrões ou conhecimentos interessantes e não triviais de documentos de texto.

Segundo Hotho et al. (2005), é possível definir TM por três perspectivas diferentes, a saber:

1. Extração de informação, extração automática de informação estruturada a partir de documentos lidos por máquinas, não estruturados ou semi-estruturados;
2. TDM, semelhante ao DM, utilizado para descobrir padrões úteis;
3. Processo KDD - Knowledge Discovery ou Knowledge Discovery in Databases, processo de descoberta de conhecimento através da extração de informação ainda não descoberta em grandes coleções de textos.

O KDD visa descobrir informação a partir de base de dados, nomeadamente as associações e padrões que estão ocultos nos dados e que são descobertos através de análise estatística.

Por sua vez, o processo de descoberta em texto (Knowledge Discovery in Text- KDT), objetiva obter informação a partir de uma panóplia de documentos (através da identificação de palavras importantes do texto - palavras-chave/atributos) e fazer a sua associação/relação.

O processo de extração de informação deve seguir alguns passos, tais como: seleção de um conjunto de dados e identificação de objetivos no processo de descoberta de informação; pré-processamento dos dados (converter dados de texto em dados com interesse - num formato que permita a utilização de algoritmos); transformação dos dados com interesse (início da extração de dados e criação de modelos/padrões); e avaliação e interpretação (Allahyari et al., 2017; Ferreira, 2015).

Para Blake (2011), o TM contempla os passos de seleção, pré-processamento, transformação, DM, interpretação e avaliação. No contexto TM, DM é um passo dentro do processo de identificação de conhecimento

Existem diferentes abordagens de TM e várias etapas de **pré-processamento** que podem ser aplicados à classificação de textos complexos. Aplicar o TM torna-se um procedimento complexo, na medida em que são analisados um vasto número de documentos. É então importante ter uma base estruturada de dados que facilite a análise posterior dos documentos (Allahyari et al., 2017).

Para ShrihariR e Desai (2015), no TM são utilizadas técnicas de extração de informações, *clustering*, sumarização, visualização e categorização:

- A extração de informação é a principal etapa para analisar texto não estruturado e a sua relação. Esse processo é realizado pela correspondência de padrões e é usada para procurar uma sequência predefinida de texto;
- O *clustering* ou agrupamento dos dados é um método não supervisionado, que consiste em reunir dados de acordo com o grau de semelhança;
- A sumarização que consiste em realizar o resumo de vários documentos num só sem mudar o sentido do texto;
- A visualização, no TM, reduz a dificuldade de descobrir informação. Este método fornece informação compreensível e em quantidade, o que ajuda a descobrir o padrão da coleção de documentos. Para o efeito são utilizadas, por exemplo, cores diferentes e distâncias relacionais.
- A categorização é uma técnica supervisionada que parte de uma base de conhecimento (exemplos) e *input* e *output* com o objetivo de treinar o classificador. O classificador de texto é usado para categorizar o documento de texto e para definir previamente a classe. Um exemplo típico de categorização de texto consiste em pré-processamento, indexação, reduções de dimensões e classificação.

As diferentes abordagens ao TM possibilitam obter informação útil, como a identificação de conceitos importantes num texto (Rodrigues, 2016). Os algoritmos de TM tratam a informação desestruturada existente nos textos, ou seja, tratam a informação como um conjunto de palavras sem relação entre si. O objetivo da aplicação dos métodos é descobrir informação, que inicialmente era desconhecida, de um modo automático. Para tal recorre à extração através de um grande número de textos que se encontram desorganizados estruturalmente (Rodrigues, 2016).

Os processos de extração, seleção e algoritmos de classificação são influenciados pelo pré-processamento de texto ou normalização do texto. Como tal, para o pré-processamento inicialmente é aplicado um processo de limpeza (p. ex. verificação ortográfica, substituição de siglas, remoção de caracteres duplicados e aplicação de regras gramaticais), seguido de tarefas fundamentais que devem ser tidas em conta, como a tokenização (*tokenization*), a filtragem (*filtering*), a lematização (*lemmatization*), o *stemming*, a marcação de partes do discurso (*part-of-speech-tagging*):

- A *tokenização* é uma tarefa que divide sequências de texto mais longas em pedaços menores (palavra/frases) ou *tokens*, podendo ao mesmo tempo retirar alguns caracteres, como a pontuação.

- A filtragem consiste em retirar do texto palavras irrelevantes (stop-words). Esta eliminação é efetuada para palavras que não contenham muita informação (p. ex. determinantes, preposições e conjunções) ou não agregam valor à interpretação do texto;
- A lematização ou análise morfológica das palavras é o processo de agrupar palavras com o mesmo significado para que possam ser analisados como um item, ou seja, retornar a forma de base ou dicionário de uma palavra, que é conhecido como o lema, com por exemplo: tiver, tenho, tinha e tem são do mesmo lema ter; bom, melhor e ótimo são lexemas do lema bom.
- O *stemming* é uma versão mais simples de lematização que consiste principalmente em retirar os sufixos do final da palavra, ou seja, é o processo de produção de variantes morfológicas de uma palavra raiz/base, por exemplo “recuperação”, "recuperado", "recuperados" e reduz para o radical "recupera".
- A marcação de parte do discurso é o processo responsável pela categorização morfológica de cada palavra no texto (substantivo, verbo, adjetivo, etc.) e das palavras vizinhas. A fase de rotulagem (*labeling phase*) é realmente complexa devido à ambiguidade, um dos principais problemas da interpretação da linguagem natural.

2.3.1. Métodos de representação de dados

Num enorme número de documentos, a aplicação do TM é um processo complexo, sendo necessário um modelo de representação adequado que permita uma análise mais eficiente em fases seguintes.

Existem vários métodos de representação das palavras ou frases de forma vetorial, como o “bag of words” (BOW), em que os mapeamentos são realizados para vetores muito grandes, e o Word Embeddings em que os mapeamentos são realizados para vetores distribuídos, densos, contínuos, de tamanhos fixos, capazes de armazenar informações sobre o contexto e significado dos documentos.

A título de exemplo, no modelo BOW, para a representação é considerado o número de ocorrências de cada termo, sem qualquer ordem específica, com se pode verificar na figura seguinte.

Texto	Bag of words	
“sem náuseas e sem vômitos”	Termo	Frequência
	sem	2
	náuseas	1
	e	1
	vômitos	1

Figura 5 - Exemplo do bag of words

Este processo leva a um vetor de representação que poderá ser analisado através de algoritmos de resolução de dimensão a partir de ML e estatísticas (Allahyari et al., 2017), como o k-means clustering.

Existem outros métodos, como o TF-IDF (*term frequency-inverse document frequency*), que significa frequência do termo-inverso da frequência nos documentos, é uma medida estatística que tem como objetivo indicar a importância de uma palavra de um documento em relação a uma coleção de documentos, em que:

$$TF(t) = \frac{\text{número de vezes que o termo } t \text{ aparece num documento}}{\text{número de termos no documento}}$$

$$IDF(t) = \log \left(\frac{\text{número de documentos}}{\text{número de documentos em que o termo } t \text{ apareceu}} \right)$$

$$TF - IDF(t) = TF(t) \times IDF(t)$$

Utilizando este método, se considerarmos um documento com 100 palavras no qual a palavra “insônia” aparece 4 vezes, o TF será $(4/100) = 0,04$. Num universo de 10 milhões de documentos se a palavra aparece em 1.000, o IDF é o $\log(10.000.000 / 1.000) = 4$. Neste caso o peso TF-IDF é o produto de 0,04 por 4, ou seja, 0,16.

A técnica TF-IDF é normalmente usada para pontuar a importância de uma palavra num documento com base na frequência com que a mesma aparece nesse documento. O objetivo é destacar palavras distintas, ou seja, atribuir mais peso a essas palavras. Palavras com muita frequência têm uma pontuação baixa e para palavras com pouca frequência têm pontuação mais alta.

A fase seguinte é a descoberta, onde informações úteis, inesperadas e desconhecidas são extraídas de recolha de dados textuais através da aplicação de certos métodos e técnicas, por exemplo, classificação e o *clustering*.

Em suma, o TM aplica uma variedade de metodologias para processar o texto, sendo uma das mais importantes o PLN.

2.3.2. Processamento de Linguagem Natural (PLN)

Linguagem natural (LN), como o nome indica, é a linguagem usada na comunicação humana.

A LN acarreta significados e ambiguidades que dependem do contexto em que se inserem, neste sentido é necessário desenvolver sistemas que a contemplem e a compreendam, tendo por base os parâmetros da linguagem formal (Allahyari et al., 2017).

O PLN visa conhecer a LN (não formal, linguagem utilizada pelos humanos) através dos computadores (Allahyari et al., 2017), ou seja, é utilizado para transformar automaticamente texto (neste trabalho texto clínico) em dados clínicos estruturados que podem ser processados por algoritmos de ML.

As abordagens clássicas do PLN, tal como já indicado no TM, podem ser vistas como sendo uma série de transformações que incluem *tokenização*, deteção de limites, deteção de limite da sentença, lematização e/ou *stemming*, rotulagem dos componentes do discurso, deteção de idioma e identificação de relações semânticas.

O PLN requer análise de alguns desafios, tais como:

- **Ambiguidade:** na linguagem natural, as palavras são únicas, mas podem ter significados diferentes dependendo do contexto;
- **Sinónimos:** palavras diferentes podem expressar a mesma ideia;
- **Sintaxe:** apesar das regras linguísticas de construção de frases verificam-se muitas vezes irregularidades sintáticas;
- **Correlação:** quando os conceitos são omitidos por referências porque foram mencionados anteriormente, tais como “A doente B fez a medicação. Ela vive no Porto”, onde “Ela” é referente à doente “B”.

Embora não seja uma ciência nova, o PLN tem avançando rapidamente devido à comunicação entre o humano e a máquina, associada ao aumento da capacidade de computação. O PLN envolve o uso de conhecimentos léxicos, sintáticos, e semânticos do idioma, entre outras

informações necessárias. São utilizadas técnicas diversas, desde métodos estatísticos e de ML, abordagens algorítmicas e também baseadas em regras, para interpretar a linguagem humana.

Diariamente, nos hospitais, são geradas quantidades significativas de informação, em registros clínicos, em muitos dos casos em “texto livre”, armazenadas eletronicamente, com um crescimento exponencial devido à implementação de processos clínicos eletrônicos. Face à enorme quantidade de texto torna-se necessário realizar uma análise adequada de forma a transformar dados em valor, podendo para o efeito serem utilizados sistemas de análise de texto e PLN.

Os sistemas de PLN na área da medicina foram desenvolvidos para identificar, extrair e codificar informações no texto narrativo clínico (Al-Haddad et al., 2010). Na fase de pré-processamento são utilizadas diferentes técnicas de PLN para normalizar e limpar dados de entrada textual e não estruturados.

O trabalho com texto clínico tem particularidades que um jornal ou uma rede social não têm. O texto clínico varia ao nível de estrutura e de estilo de escrita de profissional para profissional, utiliza linguagem comum juntamente com termos clínicos. Por exemplo, a palavra “alta”, que pode ser uma “pessoa alta”, mas no domínio clínico uma “nota de alta”, que tem um significado distinto. Neste caso, a polissemia é um dos obstáculos no PLN. O mesmo se aplica aos sinónimos, onde “consulta” e “ato médico” podem representar o mesmo conceito.

O PLN no contexto clínico, em documentos tais como diários clínicos, relatórios e notas de alta, para posteriormente serem utilizados no ML, requer a utilização de sistemas de organização de conhecimento, ontologias e tesouros, e ainda taxonomias (p. ex. UMLS, SNOMED e ICD), por forma a permitir realizar a correspondência entre diferentes termos em função do contexto. É ainda importante na análise do texto a capacidade de medir alguns dados, tais como intervalos de datas, valores numéricos (p. ex. idades), percentagens (p. ex. doses), duração, altura e peso.

O PLN é utilizado em conjunto com a análise de texto. Os textos clínicos podem ser complexos, pois são compostos por texto corrente, dados estruturados e não estruturados combinados, vocabulário complexo, alta ambiguidade, frases curtas, termos técnicos, abreviaturas, uso excessivo de siglas (p. ex. DSC – doente sentado no cadeirão), escalas numéricas, erros gramaticais, dados redundantes, entre outros. Na análise de textos clínicos é possível descobrir doentes com algumas condições e fazer algumas equivalências semânticas, como por exemplo, temperatura 39° é igual a “febre”. É possível também verificar a dependência entre as palavras, como por exemplo “Instituto Português”, em que a palavra seguinte estatisticamente mais provável será “Oncologia”.

Estatisticamente é aquilo o que surge pela probabilidade. As palavras são dependentes. No entanto, tendo por base a estatística numa análise de texto é possível antecipar as palavras com base nas anteriores ou nas seguintes. Com base nas 3 sílabas anteriores as seguintes são 2 ou 3 possíveis. Assim, o que foi analisado anteriormente servirá como guia para o que vem a seguir.

Não menos importante é o processo de desambiguação, que consiste em perceber o sentido da palavra, como por exemplo para a palavra “análise” é necessário perceber se o doente está em análise ou se foi solicitada uma análise. A resolução da correferência permite fazer a relação, por exemplo “A Sofia está doente. Ela ...”, em que “Ela” é entendido pelo sistema como sendo a “Sofia”.

A descoberta de conhecimento em textos clínicos não estruturados em informação é essencial para fornecer apoio aos clínicos no processo de tomada de decisão.

Normalmente um sistema real de PLN que utiliza reconhecimento de nomes de entidades (NER), o qual consiste no “*processo de identificação de tipos específicos de nomes ou símbolos em coleções de documentos*” e que normalmente passa por identificar o limite e a classe das identidades, combina vários métodos, sendo os mais utilizados os seguintes (Sun et al., 2018):

1. Modelo simples de regras: baseia-se na construção de cadeias de caracteres que se chamam expressões regulares. Um programador humano escreve regras para imitar o comportamento exigido de um programa, indicando as entradas e o que pretende na saída. Consistem em identificar regras da entidade a partir de textos médicos em conjuntos específicos de dados. Tanto este ponto (1) como o ponto seguinte (2) precisam de um especialista médico que construa modelos de regras e dicionários de modo manual.
2. Aprendizagem automática (ML): Este modelo transforma cada uma das tarefas anteriores em problemas de classificação supervisionadas. Tem uma boa performance ao reconhecer entidades em texto clínico. Este algoritmo é usado para reconhecer as entidades, usando características estatísticas e parâmetros dos dados de amostra. Requer um conjunto de dados treinados a partir de anotações padrão.
3. Aprendizagem simbólica, abordagem dicionário ou abordagem gramatical (PLN clássico): nesta abordagem tratam-se cada uma das palavras como símbolos que representem coisas, como por exemplo mapear os códigos médicos de forma a desambiguar. No entanto podem surgir alguns problemas, dado que a linguagem dos médicos é geralmente sucinta por um lado e por outro existem muitas variantes, o que torna este processo muito complicado para um dicionário cobrir tudo. É utilizado para

pesquisas mais precisas, como por exemplo na anotação de textos clínicos em grande escala. Reconhecem um termo através da pesquisa do termo mais semelhante no dicionário.

2.4. TM aplicado à saúde

Quando um médico descreve um doente num processo clínico eletrónico faz esse registo com linguagem natural, ou seja, de forma não estruturada. Normalmente nos textos clínicos incluem notas do clínico, observações, cartas de referência, relatórios de especialistas, notas de alta e comunicações entre médicos e doentes. Estes registos são realizados não só como um suporte de memória, mas também para serem utilizados por outros profissionais envolvidos no tratamento do doente.

As narrativas clínicas vão para além de uma classificação binária (se tem ou não tem doença), pois para além do registo da doença é efetuado o registo de como é que esta evolui ao longo do tempo e como se estabelece, o grau de gravidade associada e os fatores quer causais quer circunstâncias que se correlacionam. Verifica-se também que num processo clínico de um doente são realizados registos por diferentes profissionais, tais como médicos, enfermeiros, fisioterapeutas, nutricionistas, psicólogos, entre outros, com formas de escrita e estruturas diferentes. Todos estes dados estão presentes nas narrativas, sendo muito difícil de captar numa classificação de um diagnóstico.

Perante todos os conceitos na área clínica torna-se necessária a respetiva representação dentro de um domínio (p. ex. oncologia) e os relacionamentos entre este. Neste contexto, podem ser utilizadas ontologias, uma vez que permitem representar vocabulários formais que descrevem as premissas básicas de um determinado domínio (Freitas & Schulz, 2009).

A utilização do ML na saúde permite que a máquina entenda padrões de uma doença e seja capaz de realizar algumas previsões. Um modelo treinado com informações detalhadas, nomeadamente doenças, diagnósticos e atividade realizada, pode ajudar a que a máquina identifique e avance com propostas adequadas em função dos sintomas que um doente apresente. É certo que na medicina existem várias variáveis a ter em consideração na análise de um caso. No entanto, a máquina pela análise dos dados pode de imediato descartar algumas possibilidades, o que só por si já é uma ajuda na tomada de decisão, e permitindo ao doente eliminar a realização de MCDT's desnecessários.

O auxílio na tomada de decisão permite ao clínico orientar a análise da doença, evitando perda de tempo em hipóteses desnecessárias, levando-o a uma conclusão mais célere.

A aplicação de ML na área de saúde para a tomada de decisão pode ser tanto ao nível dos diários clínicos como dos MCDT's, por exemplo na imagiologia, podendo esta última facilitar o trabalho do clínico, através da análise da imagem e comparação no modelo, na identificação de tumor como sendo benigno ou maligno.

Os doentes também podem tirar partido de ferramentas de ML. Numa abordagem clássica um doente que necessite de saber em que data será a próxima consulta teria de consulta a APP e aceder à opção de agendamento. Com uma ferramenta de ML, como por exemplo um *chatbot*, é possível ao doente introduzir o pedido “qual a minha próxima consulta?” e a ferramenta apresentar de imediato o ato médico e a respetiva data.

2.4.1. Ontologias

Os sistemas de saúde deparam-se com muitos desafios, entre eles a gestão do aumento da quantidade de dados que precisam ser adquiridos, armazenados, processados e apresentados quando necessários e de forma acessível (Focsa, 2016).

Um sistema de informação na área da saúde tem como propósito armazenar os registos clínicos. Visto que estes sistemas não são capazes de definir/compreender todos os dados disponíveis, desde logo porque grande parte dos dados estão em “texto livre”, a construção de ontologias surge com o intuito de colmatar estas dificuldades. Para tal, deverão ser criadas ontologias específicas com o propósito de uniformizar informação (Sezer et al., 2019). Para este autor, “*As ontologias são usadas para descrever as entidades do mundo real e os relacionamentos entre essas entidades em um domínio específico*” (Sezer et al., 2019, p. 148).

As ontologias são classificações que procuram proporcionar compreensão comum num determinado domínio através do relacionamento entre palavras e expressões indicativas que representem um contexto (Studer et al., 1998).

Segundo Maculan (2020), em ciências da informação, as ontologias são um tipo de sistema de organização de conhecimento, desenvolvidas como um modelo de dados dentro de um domínio, para serem interpretáveis pela máquina. Quem constrói ontologias deverá definir processos de avaliação de construção e de refinamento da informação (Silva, 2016).

De acordo com o *The World Wide Web Consortium (W3C)*, “*An ontology formally defines a common set of terms that are used to describe and represent a domain. An ontology defines the terms used to describe and represent an area of knowledge.*” (Allahyari et al., 2017, p. 9).

Segundo Gruber (1995), uma ontologia é uma descrição formal de conceitos e as relações entre conceitos. Esta define os termos básicos e seus relacionamentos a partir do vocabulário de um domínio e as definições de classes, relações, funções e outros objetos. As ontologias geralmente descrevem – classes, que são conjuntos; coleções ou tipo de objetos – atributos, que são propriedades, características ou parâmetros que os objetos podem ter e compartilhar – relacionamentos, que são as formas como os objetos podem se relacionar com outros objetos. Em termos de filosofia e teorias da IA, a ontologia é uma descrição sistemática da existência onde tudo o que existe pode ser representado (Gruber, 1995).

Ao longo dos tempos tem se vindo a desenvolver modelos de dados padronizados na saúde. Estes modelos visam partilhar dados estruturados (com informação importante) entre diferentes sistemas, a chamada interoperabilidade, que consiste na capacidade de interação e comunicação entre sistemas. Como tal, a padronização da informação e definição de regras de comunicação é essencial para que a mesma seja transmitida entre os diferentes sistemas, sem que ocorra a alteração do significado da informação (Sezer et al., 2019).

As ontologias na saúde permitem entender o significado real do texto, independentemente dos termos que são utilizados, como por exemplo, Panadol vs. Paracetamol.

Para uma tomada de decisão clínica adequada é necessária a recuperação das informações de forma precisa.

Os sistemas de extração de informação de texto ainda dependem muito do domínio, neste sentido, existem ontologias diversificadas que visam fornecer o conhecimento semântico (sobre o domínio), através das quais se extrai a informação relevante.

Também o uso de tesouros é importante, pois consiste em vocabulário controlado e estruturado que contém termos, relacionamentos, hierarquias, sinónimos e outros níveis de dependências que representam conhecimento num determinado campo.

Uma coleção de informação muito usada em biomédica é o Unified Medical Language System (UMLS) (Bodenreider, 2004), que abarca mais de 100 dicionários distintos, terminologias e ontologias. O UMLS usa mecanismos que integram um vasto vocabulário biomédico, como: Systematized Nomenclature of Medicine Clinical Terms (SNOMED CT), Gene Ontology (GO), etc (Allahyari et al., 2017). Esta utilização facilita o desenvolvimento de sistemas de informação que necessitam de processar textos nos domínios de biomedicina e da saúde. O UMLS contempla o Metathesaurus, em vários idiomas, com mais de um milhão de conceitos biomédicos de mais de 100 dicionários distintos, que define 133 categorias amplas e 54 relações entre categorias para etiquetar o domínio biomédico e SPECIALIST Lexicon & Lexical Tools

(figura 6), que compreende uma coleção de ferramentas. Para cada termo é atribuído um identificador único que é mapeado a um identificador de conceito. Um conceito representa uma definição semântica e possui também um identificador único do conceito.

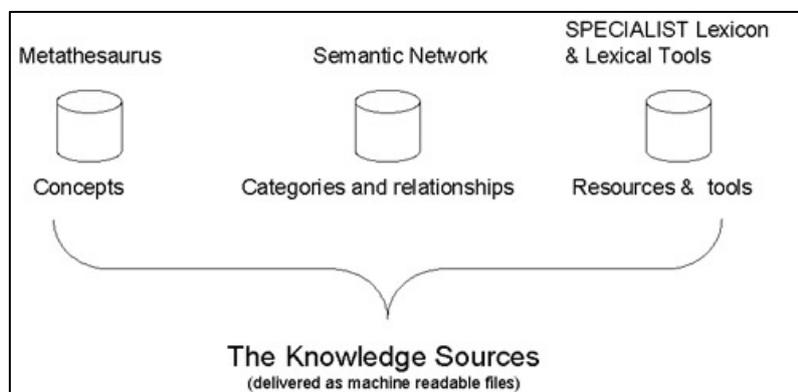


Figura 6 - Recursos UMLS

A FMA (Foundation Model of Anatomy), RadLex (Radiology Lexicon), ICD (International Classification of Diseases) são outros exemplos de ontologias usadas em medicina (Sezer et al., 2019). Existem ainda outras ontologias e terminologias utilizadas para vincular informações textuais a fontes existentes de conhecimento estruturado, tais como:

- *The Medical Subject Headings (MeSH) vocabulary e The SNOMED Clinical Terms (SNOMED CT) collection of medical terms;*
- *The International Classification of Primary Care (ICPC) for classifying a patient's reason for encounter;*
- *Disease Ontology (DO)* - é uma ontologia de doenças humanas hereditárias, de desenvolvimento e adquiridas, que integra os recursos terminológicos, MeSH, SNOMED, OMIM e o tesouro do NCI;
- *GALEN* – é uma ontologia para a representação de diferentes terminologias médicas para diversos usos clínicos, que inclui conceitos anatômicos, doenças, sintomas, medicamentos e procedimentos.
- *Ontology of Adverse Events (OAE)* - é uma ontologia de definição e classificação de eventos adversos que ocorrem após um ato médico.

Devido à complexidade de construção ontológica, é necessário por vezes optar pela construção através de níveis (quando são utilizadas técnicas de ML). Neste âmbito considera-se o nível lexical, de vocabulário ou camada de dados; de hierarquia ou taxonomia; de outras relações semânticas; do contexto; sintático e de estrutura, arquitetura e design (Silva, 2016, p. 35).

Existem várias maneiras de construção de uma ontologia. Em Sezer, et al.,(2019) são apresentados os passos necessários para a construção de ontologias. Os autores partem da informação disponível nos sistemas de informação (em saúde) e definem o seguinte processo:

- i. determinar o domínio e alcance da ontologia;
- ii. dividir em “subdomínios” (testar se o alcance da ontologia é amplo, para tal incluem termos do SNOMED CT);
- iii. reutilizar as ontologias existentes e inserir informação existente diversa;
- iv. identificar os termos importantes (deve ser feito pelos profissionais de saúde);
- v. depois de listar os termos deve-se definir conceitos de forma hierarquizada (estrutura é importante para fazer raciocínio e inferência/dedução. Note-se que neste ponto é difícil definir se um conceito é uma classe ou uma instância individual. É também importante manter o sistema atualizado);
- vi. definir as propriedades do objeto entre conceitos;
- vii. definir as propriedades dos dados (neste passo definem-se os atributos dos conceitos individuais e o seu alcance).

2.4.2. Extração de Informação

A extração de informação consiste na análise de documentos não estruturados extraíndo contexto relevante, tais como entidades (p. ex. pessoa, local, organização), eventos e relações existentes entre elementos num texto ou grupos de textos. A extração de relações permite identificar as relações semânticas entre duas entidades, por exemplo, relacionamento conjugal entre duas pessoas, enquanto na extração de eventos é analisado o texto para obter: “quem fez o quê, para quem, onde e quando”.

O processo de extração de informação envolve os seguintes procedimentos:

- Segmentação de frases: O texto é dividido em frases (p. ex. figura 7). Para esta divisão são utilizados os caracteres definidos para identificar o fim das frases, tais como “.”, “!” e “?”. No entanto, será de ter em atenção que nem sempre um “.” indica o final da frase, mas sim uma abreviação, ou ambas em simultâneo.

Frase 1	Frase 2
Conectada bomba infusora ao Cvc.	Retira bomba no domicílio.

Figura 7 - Segmentação de frases

- Separação por palavras (*tokens*): Cada frase é dividida em palavras, conforme exemplo figura 8, sendo utilizada espaços em branco ou caracteres especiais tais como "-" e "&".

Conectada	bomba	infusora	ao	Cvc	.	Retira	bomba	no	domicílio	.
1	2	3	4	5	6	7	8	9	10	11

Figura 8 - Exemplo da separação por palavras (tokens)

- Lematização: análise morfológica das palavras.
- Marcação de partes do discurso: são aplicadas etiquetas aos *tokens* definindo para cada um se é um verbo, sujeito, preposição, etc.
- NER: Identificação de possíveis entidades tais como pessoas, locais, tempo, etc.
- Reconhecimento de relações: definição de prováveis relações entre entidades encontradas no texto.

O texto não estruturado das notas clínicas dos EPR é uma fonte rica de informações de genótipo (constituição genética do indivíduo) e fenótipo (características apresentadas por um indivíduo). O TM pode tornar os textos acessíveis computacionalmente, oferecer suporte à extração de informações e à geração de hipóteses, promovendo maior valor para doentes e sistemas de saúde (Simmons et al., 2016).

A medicina tem vindo a utilizar técnicas de TM, como extração de informação de dados clínicos, sem as quais a capacidade de descoberta de conhecimento seria muito limitada. Contudo, este procedimento contém alguns desafios, nomeadamente o NER ou referências de tempo (por ex. "1/dia", "duas vezes por dia" ou "semana passada") de eventos numa linha de tempo (Tissot et al., 2015).

O nome de uma entidade pode ser definida como uma palavra ou conjunto de palavras que identifica uma pessoa, uma organização, um local, uma data, uma hora específica, uma percentagem ou quantidade (Luque et al., 2019).

As regulares descobertas na área da saúde promovem o aumento significativo de novos dados, o que muitas vezes se constitui num problema para os sistemas NER, uma vez que estes dependem do uso de dicionários.

Por forma a obter a informação de qualidade é sempre necessário contextualizar termos e conceitos, visto que os mesmos conceitos podem ter nomes diferentes. De considerar que em

medicina são usados muitos termos e abreviações, o que dificulta o processo de identificação. Devido a estas dificuldades é extremamente necessário que os sistemas NER analisem os textos de um modo preciso e com alta qualidade. Para tal, existem as seguintes abordagens (Allahyari et al., 2017):

- baseadas em dicionário, parte-se de um dicionário de termos biomédicos para identificar ocorrências palavras ou frases no texto;
- baseada em regras, define regras que especificam padrões de entidades biomédicas. Dependendo da regra que selecionar a palavra (p. ex. tem a primeira letra maiúscula e tem mais de 3 caracteres então é um pessoa), esta será definida como uma entidade de uma determinada classe (Nadeau & Sekine, 2007);
- abordagens estatísticas / ML, utilizam algoritmos que podem ser supervisionados ou não, e que normalmente são categorizados em abordagens baseadas em classificação (converter tarefas NER em classificação de problemas), métodos baseados em sequências, (usar sequências de palavras onde se tenta prever rótulos para um conjunto de palavras) e métodos híbridos (junta diferentes abordagens - técnicas de dicionário e estatísticas). O modelo de Naive Bayes é a abordagem mais frequente para pesquisar e classificar registos do EPR relacionados com o cancro e o modelo Random Forest para classificar tipos de cancro (Sheikhalishahi et al., 2019).

2.4.3. Reconhecimento de nomes de entidades (NER)

A linguagem natural é complexa. Uma frase pode conter informações contextuais, como por exemplo “A doente foi observada no centro de saúde da residência”, informação sintática onde o significado depende da estrutura da frase, como por exemplo “Doente tomou medicação de acordo com a orientação”, ou ainda informações semânticas, em que o significado depende de uma palavra individual, como por exemplo “Pessoa alta; nota de alta; febre alta”.

A extração de informações é o processo de aquisição de conhecimento num texto e da procura de ocorrências de uma classe particular de objetos, bem como relações entre objetos (Russel & Norvig, 2012).

O NER é importante no PLN, sendo uma sub tarefa da extração de informação que consiste em obtenção de informações que identificam automaticamente entidades num texto e as classifica em categorias predefinidas. Segundo Culotta et al. (2006), NER pode ser definido como “*Um documento D é uma sequência de palavras tokens $x = \langle x_1 \dots x_n \rangle$. O objetivo do NER é extrair de D um conjunto de campos $F = \{F_1 \dots F_k\}$, onde cada campo é um par atributo-valor,*

$F_i = \langle a, v \rangle$ (por exemplo $F_i = \langle Cidade, Porto \rangle$). O valor do campo pode ser repetido para múltiplos tokens”. A título de exemplo da tarefa de NER apresentamos o seguinte: ”Maria nascida no Porto.”, onde são encontradas duas entidades, $F_1 = \langle Nome, Maria \rangle$, $F_2 = \langle Cidade, Porto \rangle$.

As entidades podem ser nomes de pessoas, organizações, locais, tempo, quantidades, percentagens, entre outros. Quanto o modelo não conhece um tipo ou classe de uma entidade é necessário treiná-lo, sendo necessário realizar previamente a anotação (figura 9).

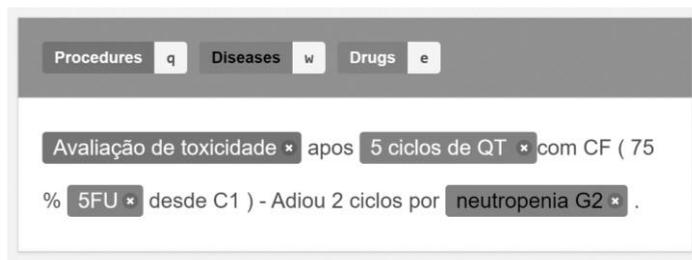


Figura 9 - Exemplo de anotação para a tarefa NER

A leitura e interpretação da frase apresentada na figura anterior para um clínico é simples. No entanto, para uma máquina nem sempre é objetivo sendo necessário indicar as entidades, neste caso os textos, expressões ou palavras que identificam procedimentos, doenças ou medicamentos.

O NER, no âmbito do processamento de linguagem natural (PLN), é útil porque o significado de algo como por exemplo “Apple” é finalmente fundamentado fora da rede linguística. No contexto tecnológico, "Apple" significa a empresa Apple. Outro exemplo, para uma entidade monetária como "1 bilhão de €", é possível anexar um valor numérico a indicar que o significado de "1 bilhão de €" é “1.000.000.000 €”. Ou seja, fornece a base para o resto das tarefas de semântica, o que torna o NER uma tarefa tão útil.

O funcionamento de um modelo NER está assente num processo constituído por duas etapas:

- Detetar uma entidade nomeada: consiste na deteção de uma palavra ou sequência de palavras que forma uma entidade. Cada palavra representa um *token*, por exemplo “avaliação de toxicidade”, é composta por três *tokens* e que representa a entidade “Procedures”;
- Categorizar a entidade: consiste na criação de categorias de entidades, como por exemplo “Procedures”, “Disorders” e “Drugs”.

Existem várias bibliotecas que podem ser utilizadas para PLN, que suportam a maioria das tarefas, nomeadamente tokenização, segmentação de frases, marcação de partes do discurso,

parsing, reconhecimento de idiomas (algumas não possuem algoritmos para Português) e NER, das quais destacamos o Apache OpenNLP⁵ é uma biblioteca Java, o Natural Language Toolkit (NLTK⁶) utilizada na construção de programas na linguagem de programação Python, a general architecture for text engineering (GATE⁷) constituída por um conjunto de ferramentas JAVA e a spaCy⁸. Esta última é uma biblioteca open-source para PLN, em Python, que utiliza ML e modelos de redes neurais, disponibiliza modelos pré-treinados em 33 idiomas, inclusive para português.

Os modelos da biblioteca spaCy são estatísticos e cada decisão que é tomada é considerada uma previsão baseada nos exemplos que o modelo viu durante o treino. Como é conhecida a resposta correta é dado feedback ao modelo na forma de gradiente de erro, que calcula a diferença entre o exemplo de treino e a saída esperada (figura 10).

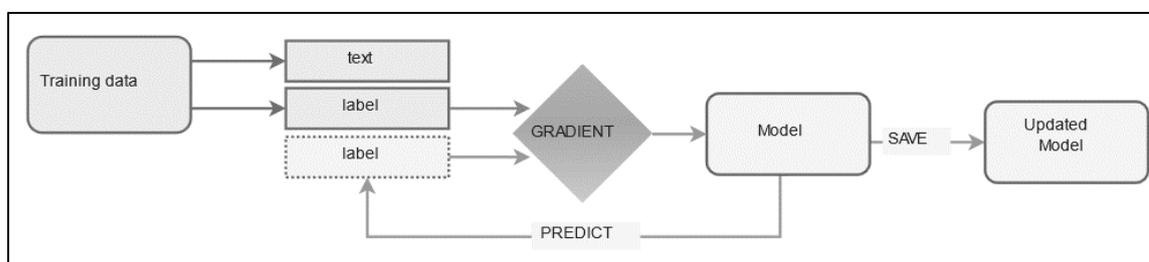


Figura 10 - Princípio de funcionamento do treino com o spaCy

O treino no spaCy, representado na figura anterior, é realizado da seguinte forma:

- Dados de treino (training data): os dados anotados contêm texto e os respectivos rótulos;
- Texto (text): É o texto de entrada para o qual o modelo deve prever um rótulo;
- Rótulo (label): É o rótulo que o modelo deve prever;
- Gradiente (Gradient): Consiste num algoritmo de otimização que calcula como alterar os pesos para melhorar as previsões, para auxiliar o treino da rede neuronal;
- Updated model: consiste em guardar o modelo.

⁵ <https://github.com/apache/opennlp>, acessido em 9/09/2020.

⁶ <https://www.nltk.org>, acessido em 9/09/2020.

⁷ <https://gate.ac.uk>, acessido em 9/09/2020.

⁸ <https://spacy.io>, acessido em 9/09/2020.

Para a realização do PLN, após a passagem do texto é realizada a tokenização e de seguida são executadas várias etapas, conhecidas como pipeline de processamento (figura 11), tais como a atribuição classes gramaticais (*tagger*), rótulos de dependência (*parser*), NER e atribuição de rótulos ao documento.

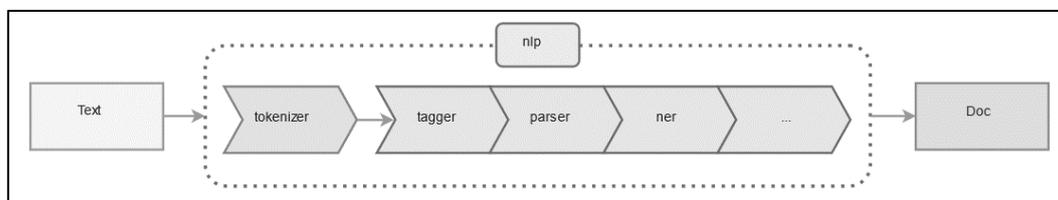


Figura 11 - Pipeline do spaCy

A tarefa de NER é capaz de identificar vários tipos de conceitos, desde nomes de pessoas, países, locais, veículos, alimentos, títulos de livros, datas e quantidades numéricas. Alguns modelos estão disponíveis para download os quais já contêm informações sobre linguagens, vocabulários, vetores treinados, sintaxes e entidades.

O objetivo do spaCy consiste em utilizar textos completos e retornar um objeto (Doc) com as diversas informações, através de um *parser* sintático rápido e com um elevado grau de acurácia. O NER base já tem etiquetas pré-treinadas que permitem a classificação de Pessoas (PER), Locais (LOC), Organizações (ORG), Diversos (MISC), como por exemplo na figura seguinte:

António Costa PER é político, mora em Lisboa LOC desde 01/01/2000 e realizou uma inauguração.

Figura 12 - Exemplo de NER

Além das entidades base que a biblioteca spaCy disponibiliza é possível adicionar outras classes ao modelo NER por treinar, para atualizá-lo, com exemplos mais recentes. Ainda no exemplo, para que fosse possível passar a identificar a data teria de ser criado um conjunto de dados de treino (figura 13):

```
train_data = [  
    ("Em 15/09/1979 foi criado o Serviço Nacional de Saúde", [(3, 13, 'DATE')]),  
    ("Dia 10/06/2020 foi feriado", [(4, 14, 'DATE')])]
```

Figura 13 - Exemplo de um conjunto de dados de treino

No seguimento do exemplo anterior, para cada entidade é necessária a sua identificação. Para melhor entender este procedimento, vejamos o exemplo acima apresentado na figura 12, no qual o NER identificou “António Costa” como uma pessoa. No entanto, ainda não sabemos exatamente a que “António Costa” o texto se refere. O Entity Linking (EL) também conhecido por Named Entity Linking (NEL), Named Entity Disambiguation (NED), Named Entity Recognition and Disambiguation (NERD) ou Named Entity Normalization (NEN), ou seja, entidade nomeada, desambiguação ou normalização, ajuda na obtenção desta resposta, atribuindo uma identidade única às entidades mencionadas no texto (Shen et al., 2015), tendo por base uma base de dados de conhecimento, como por exemplo o ULMS ou a DBpedia (base de conhecimento estruturada extraída da Wikipedia).

Apesar do modelo ter identificado “António Costa” como uma pessoa não é possível saber a que “António Costa” se está a referir. Existem muitas pessoas com o nome “António Costa” que são políticos e que vivem em Lisboa.

Uma forma inicial de reduzir o número de pessoas seria pesquisar no base de conhecimento, como a wikidata, que contém vários *aliases* e sinónimos e teríamos pelo menos 3 candidatos, a saber:

- António Costa (Q610788), Primeiro Ministro
- Marco António Costa (Q11767932), Político
- António da Costa Leal (Q16491405), Político

Agora teríamos de encontrar a pessoa, reduzindo a lista a uma pessoa que represente o pretendido. Seria necessário treinar o modelo com várias frases que contivessem o nome “António Costa”.

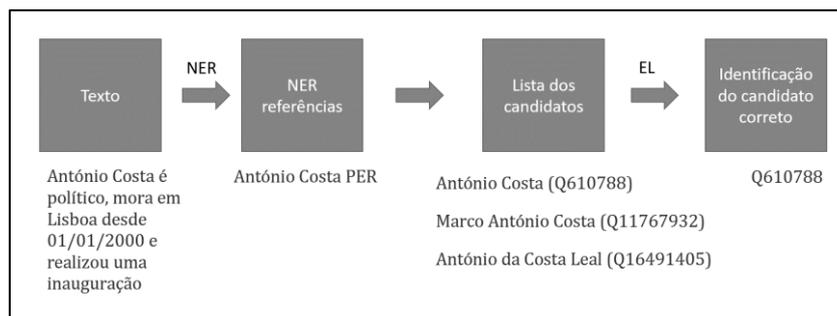


Figura 14 - Passos necessário para a desambiguação de uma entidade

2.5. Sistemas PLN no domínio clínico

Para análise de metodologias, numa análise de alguns autores concluiu-se que os métodos a utilizar dependem dos dados que se pretende analisar, dependendo se esses dados se apresentam de um modo estruturado ou não estruturado (Delespierre et al., 2017; Kocbek et al., 2016; Sun et al., 2018; van Laar et al., 2020). De um modo geral, o processamento de extração de dados dos registos médicos eletrónicos engloba a recolha, o pré-processamento, a extração, a evolução e a aplicação do conhecimento (Sun et al., 2018).

Para a recolha de dados parte-se de registos eletrónicos e procede-se à extração de dados estruturados e não estruturados. No que concerne aos dados não estruturados, a partir de notas clínicas, a extração é feita através da pesquisa de palavra a palavra havendo posteriormente uma validação humana. Por sua vez, para dados estruturados (medicamentos, dados laboratoriais, etc.) a extração é feita automaticamente através de consulta de dados estruturados (van Laar et al., 2020).

O primeiro passo, limpeza de dados inconsistentes, centra-se, portanto, em retirar a informação não útil, em anonimizar a informação, isto é, retirar meta-dados demográficos relacionados com o hospital (códigos, data, hora, etc). (Kocbek et al., 2016) extraíram a forma textual de todos os relatórios da plataforma REASON (*data warehouse* com múltiplas fontes de dados dentro do sistema) (Bain & Manus, 2014) e seguiram os seguintes passos: encurtaram caracteres, removeram a acentuação, datas, etc.; removeram sobreposições de narrativas de enfermagem; excluíram palavras ou expressões sem sentido utilizando o ORACLE® com a função SQL LIKE e *wildcards* para realizar correspondência de padrões; utilizaram as mesmas técnicas na construção das variáveis de fisioterapia para detetar palavras, descrever os cuidados de fisioterapia e contá-los (encontrar as palavras úteis). Através deste processo foram capazes de identificar as palavras com maior ocorrência e aquelas que pertencem ao texto livre.

Para tudo aquilo que não foi classificado, recorreram à linguagem R, onde foram usadas técnicas de *stemming* e lematização (que é a combinação de palavras da mesma família no mesmo grupo) (Kocbek et al., 2016).

Para cada registo estava associado um conjunto de códigos ICD-10, que eram anotados (construção de um *dataset*). Os relatórios de radiologia e patologia clínica foram processados com recurso à *metamap tool*⁹, ferramenta que identifica e normaliza a terminologia biomédica da UMLS. Para avaliar a precisão dos classificadores, os códigos inferidos pelos classificadores são comparados com os verdadeiros códigos ICD (que foram removidos dos conjuntos de dados durante a avaliação experimental). Após isto, seleccionaram 7 doenças. Para cada doença foi construído um classificador utilizando a ferramenta Weka Toolkit (Hall et al., 2009) que recorreu ao algoritmo de SVM, usando *polinomial kernel e parâmetro complexo*. Foi ainda utilizado o método Naive Bayes para fazer um treino mais rápido (Kocbek et al., 2016).

Já para Sun et al. (2018), as etapas de TM normalmente passam por: *information retrieval* (identificar texto desejado) – *data collection* (recolha dos dados), *information extraction* (extrair informação pré-definida), *knowledge discovery* (extrair conhecimento novo do texto) e *knowledge application* (factos desconhecidos dos textos para a prática). TM em medicina consiste em obter dados a partir de textos semiestruturados ou não estruturados, e convertê-los numa base de dados, a partir de técnicas de extração de informação de PLN – NER e a relação de entidades (Sun et al., 2018). Obstáculos à aplicação do NER são os erros gramaticais, as diferentes formas de escrever e as abreviaturas (Sun et al., 2018).

A extração de informação deve ser validada através de parâmetros de pontuação e comparada a sua utilidade com os resultados recuperados manualmente (van Laar et al., 2020). Deste modo, o TM e NLP tem de seguir 3 métricas de avaliação, Precisão (P), *Recall* (R), *F1 score* (pontuação) (Delespierre et al., 2017; Kocbek et al., 2016; Sun et al., 2018; van Laar et al., 2020), sendo que quanto maior for a pontuação (*F1 score*), melhor serão os resultados.

Para a construção de classificadores os vários autores usaram diferentes formas de recolha e extração de dados. O TM e o DM constituem-se ferramentas importantes para a melhoria contínua das plataformas médicas eletrónicas, o que facilitará, por sua vez, a melhoria contínua na prestação de cuidados de saúde aos doentes.

⁹ https://www.nlm.nih.gov/research/umls/implementation_resources/metamap.html, acedido em 10/09/2020.

Avaliar sistemas PLN no domínio clínico não é tarefa fácil. Friedman (1998) sugere que uma avaliação de um sistema PLN em medicina deve contemplar 20 tópicos, organizados em 5 grupos de acordo com seu propósito.

Minimizar o *Bias*:

1. O programador não deve ver o conjunto de documentos de teste.
2. Se forem escolhidos especialistas de domínio para estabelecer o padrão de referência, então eles não podem ser os programadores do sistema, nem participar no projeto.
3. O programador não deve realizar a avaliação.
4. O sistema PLN deve estar fechado antes da fase de testes.
5. Se a capacidade de generalização do sistema está a ser testada o programador não deve saber antecipadamente os detalhes do estudo.
6. Idealmente, a pessoa que projeta o estudo de avaliação não deve ser o programador do sistema.

Para estabelecer um padrão de referência:

7. Se forem escolhidos especialistas para determinar o padrão de referência, deve haver um número suficiente deles para avaliar a variabilidade do padrão de referência.
8. O conjunto de testes deve ser grande o suficiente para que seja possível distinguir os níveis de desempenho.
9. A escolha do padrão de referência deve ser baseada nos objetivos do estudo (por exemplo, capacidade de extração vs. desempenho do aplicativo).
10. E se são utilizados especialistas do domínio para determinar a referência padrão, o tipo de especialista deve ser apropriado (p. ex. radiologista vs. internista).

Métodos de avaliação:

11. O método usado para determinar o padrão de referência deve ser claramente descrito, especialmente se houver especialistas de domínio.
12. A forma de escolha dos documentos de teste deve estar descrita.
13. Os métodos usados para calcular o desempenho devem ser claramente apresentados, principalmente na utilização de métodos não padronizados.

Apresentação dos resultados:

14. Os resultados de desempenho devem estar relacionados ao conjunto inteiro de teste.

15. Se forem utilizados especialistas humanos deve haver troca de pareceres entre eles.
16. Intervalos de confiança devem ser fornecidos para todas as medidas.

Discussão das conclusões:

17. Limitações do estudo devem ser discutidas.
18. Os resultados devem ser apresentados segundo os requisitos da aplicação.
19. A generalização dos resultados deve ser evitada.
20. Uma análise das falhas do sistema deve ser fornecida juntamente com a discussão das dificuldades em se obter as referidas correções.

Existem vários sistemas de PLN disponíveis para extração de informações de narrativas clínicas, em inglês. Os sistemas seguintes utilizam o PLN com dicionários clínicos integrados, terminologias e classificações, desenvolvidos para a área clínica com diversas finalidades e das analisadas, cTAKES (Savova et al., 2010), MedLEE (Carol Friedman et al., 2004), CHES (Hardjojo et al., 2018), concluiu-se que o UMLS é utilizado para validação das extrações realizadas.

2.5.1. Clinical History Extractor for Syndromic Surveillance (CHES)

O CHES é um algoritmo de processamento de linguagem natural baseado em regras para extrair informações clínicas de registros em texto livre nos cuidados primários do National Healthcare Group Polyclinics, de Singapura, com capacidade de extrair 48 sinais e sintomas sugerindo infecções respiratórias, infecções gastrointestinais, constitucionais, bem como outros sinais e sintomas potencialmente associados a doenças infecciosas. Para além da utilização de um conjunto de palavras-chave foi também utilizada o metathesaurus da UMLS.

O sistema é constituído por uma série de etapas, iniciando-se pela segmentação das frases, para facilitar a categorização de acordo com um conjunto de regras. Estes *tokens* são categorizados como sintomas (presentes ou ausentes) e medidas (p. ex. quantidades ou tempo). Para cada sintoma que é identificado pelo dicionário tem uma relação com uma base de dados relacional para incorporar erros comuns de ortografia, abreviações e sinônimos, e remoção de palavras não essenciais (p. ex. “em”).

Foi utilizada uma ferramenta para realizar a classificação manual dos sintomas em 1680 registros, com o objetivo de treinar e validar o algoritmo do CHES. Este trabalho foi realizado por dois especialistas, profissionais de saúde com experiência em pesquisa clínica e

interpretação de notas clínicas. Foi ainda utilizado um terceiro especialista que serviu como juiz para os casos em que não houve acordo.

O treino, com várias interações, foi realizado em 50% dos registos anotados, tendo sido identificados novos termos e redução de erros ortográficos e serviu também para eliminar os falsos negativos.

A validação do desempenho foi realizada nos restantes 50% dos registos, permitindo assim testar a robustez do algoritmo na identificação correta de sinais e sintomas, sendo a avaliação realizada através da precisão, *recall* e *f1 score*.

Foi concluído que o CHES consegue identificar com bom desempenho, determinando a existência ou não, ou ainda a suspeita, de sinais e sintomas, com uma precisão muito alta (97,2% para o conjunto de validação).

2.5.2. Medical Language Extraction and Encoding System (MedLEE)

Este sistema baseado em PLN, com desenvolvimento iniciado em 1991, inicialmente projetado para tratamento de relatórios de radiologia, mas posteriormente alargado a outros domínios, tem sido utilizado no Columbia University Medical Center e no New York Presbyterian Hospital, para processar relatórios de radiologia, cardiologia, patologia, notas de admissão, notas de alta e também para apoio à decisão clínica (Mendonça et al., 2005). Tem como objetivo, o mapeamento automático de narrativas clínicas para um formato estruturado usando um vocabulário controlado e faz também a respetiva avaliação quantitativa.

O sistema é constituído por uma série de fases (Carol Friedman et al., 2004; Mendonça et al., 2005), a saber:

- **Pré-processamento:** é realizado o reconhecimento dos limites da frase e são também realizadas pesquisas lexicais para reconhecer e categorizar palavras, frases e abreviações, bem como a desambiguação;
- **Parser** ou análise gramatical: as definições gramaticais e lexicais identificam e interpretam a estrutura da frase, e servem para elaborar uma estrutura intermediária fundamentada nas especificações gramaticais.
- **Composição / Regulação de frases:** é necessária para construir e normalizar frases com várias palavras para facilitar a recuperação nos casos em que as palavras individuais da frase foram separadas umas das outras nos textos de entrada.

- Codificador: são realizados os mapeamentos dos termos encontrados para um vocabulário clínico padrão (p. ex. realizar a codificação UMLS, C0006826 – cancro).

Para a avaliação da codificação do UMLS utilizaram um universo de 818.000 frases, através de dois estudos separados, ambos com 150 frases selecionadas aleatoriamente, um para medir o *recall* e o outro a precisão.

Foi então utilizado o MedLEE foi usado para processar as frases, extrair informações clínicas relevantes e realizar a codificação UMLS. Foram utilizados seis especialistas, com instruções por escrito e treinados previamente, para analisar dos resultados, em que para cada analisou 75 frases, em que uma frase foi analisada pelo menos por dois especialistas, através de uma ferramenta informática, efetuado o registo: correto, incorreto ou talvez correto. Foi realizada a mesma avaliação manual com grupos de especialistas, através de anotação manual junto das frases.

Os resultados da avaliação do método mostraram que o MedLEE teve um desempenho similar ou melhor do que os especialistas tanto em *recall*, quanto em precisão.

2.5.3. clinical Text Analysis and Knowledge Extraction System (cTAKES)

O cTAKES é um sistema para extração de informações de texto livre do processo clínico, que tem por base tecnologias *open source* e NLP, combinando técnicas baseadas em regras e de ML com objetivo de realizar extração de informações de textos clínicos, sendo utilizado pela Mayo Clinic. Métricas de avaliação standard (*recall*, precisão, *f1 score* e acurácia) são usadas para medir a qualidade dos padrões associados aos rótulos linguísticos e conceitos clínicos, estando disponíveis na versão *open source* os seguintes componentes:

- identificação de frases, que prevê se um ponto final, ponto de interrogação ou ponto de exclamação é o final de uma frase;
- *tokenizador*, faz a divisão da frase segundo regras, como por exemplo pontuação;
- normalizador, faz uma representação para cada palavra no texto de entrada para uma série de propriedades lexicais, incluindo variantes de escrita e pontuação.
- marcação de partes do discurso;
- NER, em que cada entidade descoberta pertence a um dos tipos semânticos do dicionário e tem atributos para o intervalo de texto associado à respetiva entidade.

O cTAKES está acessível via API requerendo como *endpoint* uma nota clínica, sendo avaliada pelo UMLS standard, podendo ser criados dicionários personalizados e como output disponibiliza condições, observações, medicamentos e procedimentos.

2.5.4. Extração de entidade em notas clínicas

Kumar (2015) aplicou o sistema de extração de entidades com recurso ao CRF na área clínica, utilizando o *dataset* da I2B2¹⁰, com 2100 frases, com o objetivo de abranger várias entidades, utilizando 70% dos frases para treino e 30% para testes.

A decisão de utilizar o CRF teve por base a análise da literatura na qual concluíram que estes modelos têm melhor desempenho na aplicação de sequências, com utilização de uma biblioteca em JAVA, a Stanford NER¹¹, que integra com ferramentas PLN. Este NER tem a capacidade de rotular sequências de palavras de um texto, como nomes de coisas, pessoas, empresas, ou ainda nomes de genes e proteínas.

Neste trabalho extraíram um dado número de sequências para entender o contexto das palavras, permitindo assim entender algumas palavras como sendo temperatura, dosagem ou frequência, usando também características morfológicas como prefixo e sufixo, bem como o conhecimento semântico das siglas e sinónimos.

2.6. Sistemas de apoio à decisão clínica (SADC)

Na atividade clínica, a informática é utilizada em diversas fases do processo de interação entre o doente e o médico: na recolha de dados, na realização de exames, na monitorização e no tratamento.

A Informática Médica comporta vários sistemas, entre os quais sistemas de informação clínicos, para registar os dados de saúde e doença de utentes e SADC, com o objetivo de fornecer informação aos profissionais de saúde.

Um SADC é um meio de suporte à decisão através de conhecimentos na área de saúde. Estes

¹⁰ <https://www.i2b2.org/>, acedido em 9/09/2020.

¹¹ <https://nlp.stanford.edu/software/CRF-NER.html>, acedido em 9/09/2020.

sistemas são maioritariamente utilizados na descoberta do diagnóstico e são bastante procurados porque ajudam a reduzir os erros médicos e melhoram a eficiência e qualidade dos cuidados prestados (Sheng, 2000).

O resultado de um diagnóstico deve ser objetivo, claro, estruturado e sistematizado de forma a descrever o problema de saúde do doente. De forma que o principal objetivo de um SADC é a diminuição da incerteza durante o processo de decisão, devendo fornecer sugestões e respostas, e ajudar a traçar caminhos para obter um diagnóstico.

Assim, poderiam ser denominados SADC todos os sistemas que utilizam dados, informações e conhecimento; que participem do processo de atenção ao doente e que forneçam *“apoio aos médicos no processo diário de tomada de decisão, (...) baseado em conhecimento especializado”* (Sigulem, et al., 1995, p.79).

Os SADC oferecem dois tipos de assistência no processo de tomada de decisão (Vasconcelos et al., 2004):

- i. Assistência direta: através da aplicação de conhecimento científico (ou através de respostas de medicina baseada em evidências), os SADC fornecem respostas, sugestões ou alternativas a um determinado problema, por forma a reduzir significativamente a incerteza na tomada de decisão.
- ii. Assistência indireta: estes sistemas devem ser *“user-friendly”* para quem os usa, apresentar a informação numa linguagem clara para os utilizadores e possibilitar a seleção da quantidade de informação que um utilizador deseja ver para facilitar o processo da tomada de decisão, não causando sobrecarga cognitiva.

Os sistemas de apoio à tomada de decisão clínica são concebidos para auxiliar os profissionais de saúde a tomar decisões clínicas. Estes sistemas são utilizados na elaboração do diagnóstico médico e durante a tomada de decisão de uma terapia. O SADC utiliza informações médicas e dados dos doentes e é baseado em conhecimentos médicos.

O processo de tomada de decisão clínica depende da forma como o registo eletrónico do doente se encontra estruturado e atualizado. O registo clínico do doente é de crucial importância nas diferentes etapas da tomada de decisão clínica.

Não é objetivo do SADC substituir o médico no ato clínico. O seu objetivo é auxiliá-lo a simplificar e otimizar o seu desempenho durante a tomada de decisão. A disponibilização de SADC com IA visa ampliar as capacidades do médico e não a substituí-las. Neste momento, concentramo-nos em aplicar a inteligência artificial como forma de complementar e ampliar as capacidades humanas.

O SADC assume-se como um complemento entre o conhecimento do médico e a informação do próprio sistema. Estes sistemas têm vantagens compreensíveis, tais como: os sistemas informáticos não são afetados por falhas normalmente humanas, como o cansaço, distração, entre outras; existe possibilidade de redução do tempo de atendimento ao doente: e redução do tempo de procura da informação do doente, sobrando mais tempo para atendimento ao doente.

Num SADC deve ser possível visualizar todos os dados existentes sobre o doente e o seu historial clínico, pesquisar informações médicas que suportem a tomada de decisão e ter acesso a relatórios de exames, resultados de análises, e relatórios de consultas.

Com a utilização do PLN é possível utilizar informações que estão em texto livre, aproveitando todo o conteúdo da narrativa clínica, para auxiliar o clínico no suporte à decisão.

Pretende-se com este trabalho a validação de vários modelos com o objetivo de extrair os dados necessários para permitir ao clínico:

1. Aceder rapidamente aos dados básicos do doente;
2. Disponibilizar dados contínuos sobre os problemas de saúde;
3. Realizar o acompanhamento das consultas do doente;
4. Disponibilizar dados para organizar medidas preventivas ou terapêuticas;
5. Contribuir para a educação contínua.

3. Metodologia

Pretende-se a validação clínica de algoritmos de TM em contexto de oncologia, tendo sido utilizados como exemplo 2 processos clínicos completos e anonimizados. Para o efeito foram realizadas as atividades de preparação dos documentos e construção de modelos.

3.1. Trabalho inicial

Os textos clínicos são compostos por texto corrente, dados estruturados e não estruturados combinados, vocabulário complexo, alta ambiguidade, termos técnicos, abreviaturas, uso excessivo de siglas (figura 15). Com a colaboração de especialistas de domínio foi realizado um levantamento do léxico oncológico e não oncológico (figura 16), incluindo abreviaturas e siglas (p. ex. Hx, ADC, QT, RT), que constam dos EPR, tendo sido utilizados 5 processo clínicos, devidamente anonimizados. No anexo 2 é possível ver exemplos de abreviaturas e siglas.

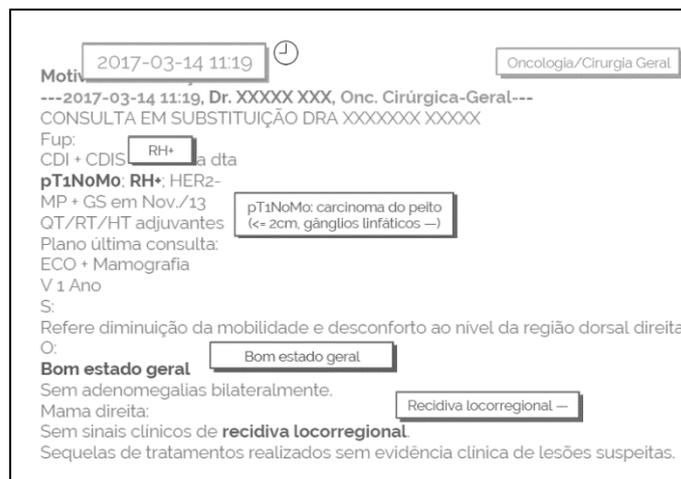


Figura 15 - Exemplo de um episódio do EPR

3.2. Preparação dos documentos

3.2.1. Extração dos PDF para texto

Neste processo foram utilizados ficheiros PDF, em que cada documento tem o diário clínico completo de cada doente, organizado por episódios, conforme apresentado na figura seguinte. Para utilizar estes documentos foi necessário realizar a conversão para formato texto para de seguida realizar a normalização.

<p>CE em especialidade, aaaa-mm-dd, Nome do clínico</p> <p>Motivo/ Observação</p> <p>---aaaa-mm-dd hh:mm, Nome do clínico, especialidade---</p> <p>Motivo: seguimento</p> <p>Doente do sexo feminino, xx anos, profissão, estado civil, reside em xxxxx, com Hx obesidade.</p> <p># 1. ADC do recto/transição rectosigmoideia diagnosticado por colonoscopia a dd.mm.aaaa na sequência de rastreio com PSOF +. Estádio cT2-4aN+Mo.</p> <p>.</p> <p>..</p> <p>...</p> <p>CE em especialidade, aaaa-mm-dd, Nome do clínico</p> <p>Motivo/ Observação</p> <p>---aaaa-mm-dd hh:mm, Nome do clínico, especialidade---</p> <p>Faz QT + RT</p> <p>.</p>
--

Figura 16 - Exemplo do diário clínico

3.2.2. Normalização de frases

Estando o documento em formato texto foi realizada a segmentação do texto em frases. Sendo uma tarefa aparentemente simples, verificamos algumas dificuldades como pouca pontuação, pontuação em abreviaturas e também falta de pontuação no final de frases (anexo 3). Em alguns processos o clínico utiliza a quebra de linha para delimitar o final da frase, tendo-se concluído que é bastante comum nos textos clínicos que foram analisados.

De seguida é aplicada a normalização com vista a traduzir sempre que possível o termo abreviado para o termo escrito por extenso. Esta atividade é importante pois ajuda na padronização do texto e também a minimizar erros na segmentação das frases. Essa tarefa é realizada com a combinação de um algoritmo baseado em regras, dicionário e também por ML.

3.3. Construção dos modelos

Foram utilizados modelos sem supervisão, semi-supervisionado e supervisionados. Para a construção do modelo semi-supervisionado, foram definidas apenas três categorias do UMLS, “*Procedures*”, “*Disorders*” e “*Drugs*”. Na tabela seguinte é possível verificar para as categorias selecionadas a quantidade de conceitos suportados, relevantes para este trabalho.

Categorias	Subcategorias	Conceitos encontrados
Chemicals & Drugs	26	356,211
Disorders	12	136,389
Procedures	7	81,847
Total	45	574,447

Tabela 2 - Quantidade de conceitos suportados pelo UMLS¹²

3.3.1. Validação dos modelos

Para o *dataset* foram utilizadas 2055 frases, cerca de 80% de todas as frases dos 2 processos utilizados para este trabalho. Como metodologia de validação dos modelos (figura 17), foram realizadas 2 anotações por anotadores humanos (A e B) e outra anotação com recurso ao UMLS. Com estes rótulos foram aplicados 3 modelos e por fim realizada uma análise qualitativa dos resultados por um médico sénior (especialista C).

¹² Dados consultados em <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC43>, em 9/09/2020.

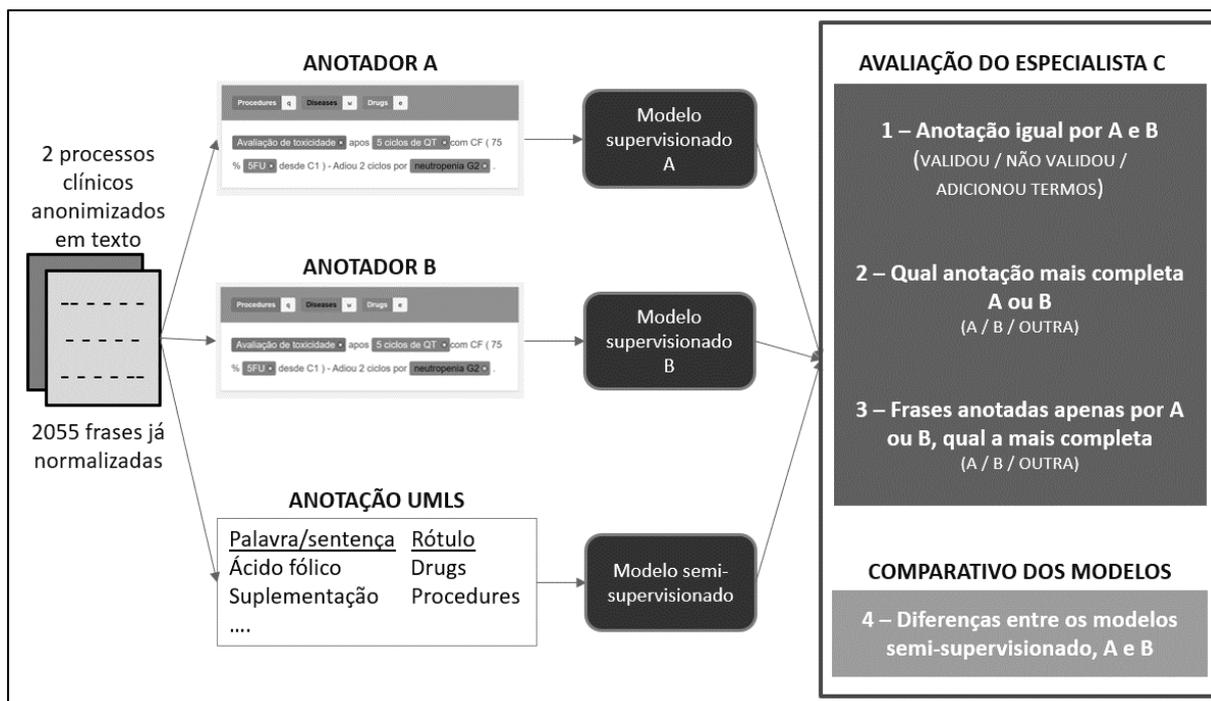


Figura 17 - Metodologia de validação dos modelos

Para a validação foram utilizados os seguintes modelos:

1. **Modelo baseado em regras**, ou seja, sem qualquer treino, com dados extraídos usando o UMLS. Este modelo serviu apenas para realizar a descoberta de termos para aplicar no modelo semi-supervisionado.

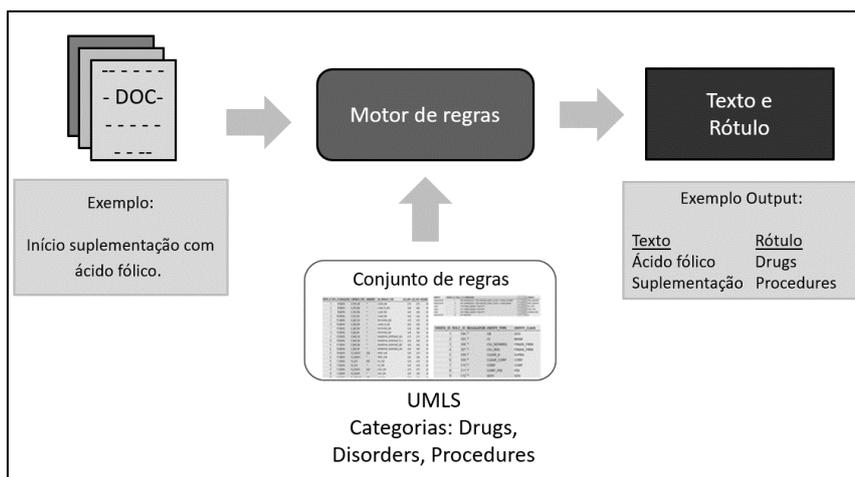


Figura 18 - Modelo baseado em regras

2. Modelo semi-supervisionado com pré-anotação automática do UMLS.

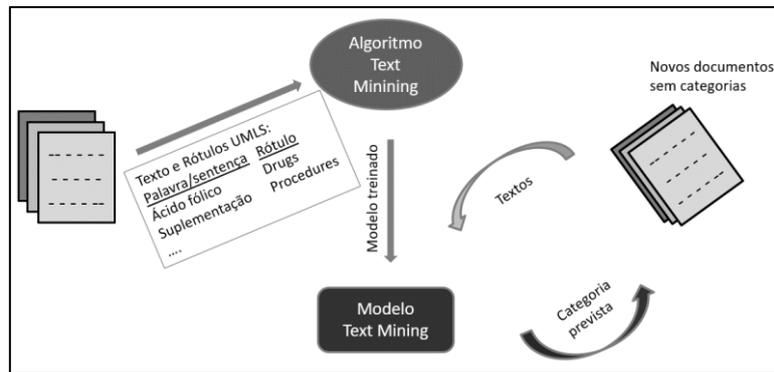


Figura 19 - Modelo semi-supervisionado

3. Modelo supervisionado, com anotações de especialistas.

Para este modelo foram utilizados dois anotadores especialistas (A e B), um sem prática (A), mas da área da saúde e outro com prática de anotação (B), mas não é da área da saúde, para realizarem anotação de entidades (conforme anexo 4) às frases (cerca de 80% do total das frases de 2 processos clínicos) extraídas dos processos que contêm toda a história clínica, seguindo as orientações identificadas no anexo 5.

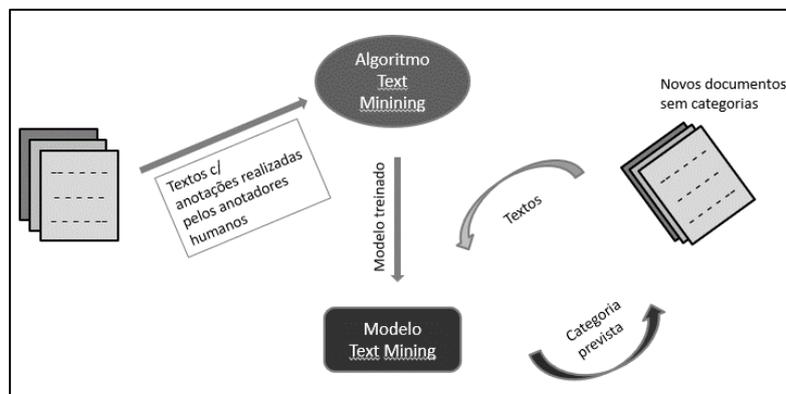


Figura 20 - Modelo supervisionado

- i. **Anotação do dataset para treino:** para a anotação foi utilizada a ferramenta Doccano¹³ que oferece suporte a recursos de anotação para classificação de texto, tendo sido criados 2 projetos iguais, um para cada anotador. Estes dados anotados servirão

¹³ Disponível em <https://github.com/doccano/doccano>

posteriormente para alimentar um sistema de ML capaz de aprender a identificar esses elementos de maneira automatizada.

Para a anotação foram criados os “*procedures*”, “*diseases*” e “*drugs*”, cada um com uma cor diferentes para simplificar o processo.

O processo de anotação consiste em selecionar o documento sem nenhuma anotação, destacar o texto a anotar e pressionar no rótulo adequado ou a letra associada ao rótulo, conforme apresentado na figura seguinte.



Figura 21 - Exemplo de anotação de texto como sendo "drugs"

Após a anotação dos dados pelos anotadores especialistas do domínio, foi possível realizar o download do projeto com rótulos no formato JSON / JSONL para treinar o modelo de ML.

- ii. **Treino do modelo:** Para esta tarefa foi utilizada a biblioteca spaCy descrita anteriormente, tendo sido utilizados os dados anotados pelos anotadores especialistas e o treino foi realizado com todas as frases dos 2 processos clínicos, com o objetivo de realizar uma aprendizagem de todo o contexto e que passe a utilizar noutros casos, pois caso fosse treinado apenas com os dados iniciais o modelo deixaria de ter a capacidade de generalizar.

3.4. Tarefas adicionais de pós-processamento

As tarefas seguintes são importantes, mas não serão avaliadas neste trabalho. São descritas como trabalhos necessários no âmbito do PLN e também na organização dos dados.

3.4.1. Detecção de presença ou ausência (Assertion Status Detection)

Nos diários clínicos o médico também menciona sintomas ausentes no doente, tais como “nega febre”. Numa leitura simples do texto seria de indicar a presença de “febre”. É, portanto, necessário detetar se o sintoma está presente ou não. O significado das entidades clínicas é fortemente afetado pela negação, o incerto, o hipotético, entre outros (Chen, 2019).

Uma atribuição incorreta poderia causar um diagnóstico errado face à condição do doente, influenciando as etapas de tratamento seguintes.

3.4.2. Entity Linking (EL)

O NER indica que palavras são entidades e de que tipo são, pelo que não é suficiente. Para além de detetar conceitos é preciso normalizar e atribuir o seu respetivo código na base de dados da Ontologia UMLS. Para isso, efetuamos para cada elemento detetado no texto uma consulta na base de dados e retornamos o identificador único do elemento. Em termos práticos, é detetada uma expressão (dor de garganta) mas é necessário saber qual é o código UMLS. Com esta tarefa é atribuído o identificador internacional para a categoria (p. ex. “*Procedures*” ou “*Disorders*”), que em conjunto com o ponto anterior, dá origem à tabela de presença ou ausência do termo, conforme exemplo apresentado no anexo 6.

3.5. Tarefa de anotação do SOAP, de frases do EPR, para organização dos resumos num *dashboard*

Seguindo a mesma forma de anotação, utilizando os mesmos processos, foram realizadas anotações (anexo 7) para as entidades do SOAP, tendo por base as orientações que constam do anexo 8, o que permitiu identificar a seguinte organização:

- Dados subjetivos, tais como queixas do doente;
- Dados objetivos, observados pelos clínicos, incluindo resultados de exames;
- Elementos de avaliação, incluindo diagnósticos inferidos a partir dos dados;
- Ações a realizar em etapas seguintes (plano).

4. Caso de estudo

A Instituição onde foi realizado este trabalho é uma Unidade de Saúde de referência nacional e internacional no domínio do tratamento, da investigação e do ensino no cancro. Posiciona-se na rede pública de cuidados hospitalares como instituição altamente diferenciada, de cuidados especializados e de natureza muito específica, centrados no doente. Uma das áreas de intervenção é a clínica na qual são realizadas as ações de prevenção primária, rastreio e diagnóstico precoce, tratamento multidisciplinar, seguimento pós tratamento, prestação de cuidados paliativos e aconselhamento genético.

Apesar da Instituição estar dotada de muitos aplicativos para suporte à atividade assistencial, alguns preparados para a introdução estruturada de dados, facilitando assim o seu uso em modelos analíticos, os registos de cada doente que são efetuados no diário clínico são introduzido sob a forma de “texto livre”, não sendo possível extrair a semântica do texto de uma forma automatizada.

Pretende-se com este caso de uso a utilização de ferramentas de TM, com o objetivo estruturar “texto livre”, com a validação clínica dos algoritmos, para que a informação possa ser utilizada no suporte à decisão clínica. Para o efeito foram utilizados 2 processos clínicos completos anonimizados que estavam em PDF e com a colaboração de anotadores especialistas que realizaram a anotação das classes “*Procedures*”, “*Disorders*” e “*Drugs*” em 2055 frases, tendo posteriormente sido validado por um terceiro especialista, um médico sénior.

4.1. Informação sobre o processo clínico

O processo clínico eletrónico desta Instituição de saúde permite ao médico o registo de atos médicos programados ou não, registo do ato médico em texto livre, o registo de diagnósticos codificados com a possibilidade de configurar diagnósticos por especialidade, registo de alertas e problemas, indicando informação relevante que necessite de alertar o clínico que acede ao processo do doente, a requisição de meios complementares de diagnóstico e terapêutica, marcação de consultas para o próprio ou outros médicos, pedido de consulta para outra especialidade sendo possível indicar a necessidade de efetuar a transferência de responsabilidade, elaboração da proposta cirúrgica segundo as regras definidas pelo SIGIC - sistema integrado de gestão de inscritos para cirurgia, elaboração de pedido de internamento racionalizando o número de pedidos de internamento possíveis para um dia/especialidade, possibilidade de tipificar a alta da instituição: alta médica ou alta por óbito, elaboração de relatórios do episódio com a possibilidade de inclusão de *templates* com carregamento

automático de informação pré-definida, disponibilização do acesso aos resultados dos exames.

Apesar do registo do texto no processo clínico eletrónico ser livre, alguns médicos fazem o registo utilizando a ordem do SOAP, registando o Subjetivo, Objetivo, Avaliação e Plano. Este registo contém informação de diagnóstico, resultados dos meios complementares de diagnóstico, estado de saúde em geral e ainda características do doente. Com todos estes dados é possível analisar a resposta de determinados case mixes de tratamentos oncológicos, taxas de sobrevivência, estimar QALYs (Quality-adjusted Life Years), confirmar valores de custo-efetividade reportados pelas farmacêuticas e auxiliar na tomada de decisão clínica.

Diariamente, nesta unidade de saúde são realizados mais de 2000 atos médicos em ambulatório e internamento, com vários registos no diário clínico.

4.2. Objetivo da implementação do TM

Pretende-se obter um resumo estruturado através dos diários clínicos escritos em “texto livre”, com recurso ao TM. Para o desenvolvimento do algoritmo de TM, adaptado ao contexto oncológico, foi realizada a extração de registos clínicos, em particular informação referente ao diagnóstico, análises, estado de saúde, características do doente e planos SOAP.

Esta atividade compreende tarefas, tais como o levantamento do léxico de suporte, a integração com o processo clínico eletrónico, o treino do algoritmo recorrendo a registos clínicos passados e a própria extração de informação estruturada.

Recorrendo a todo o histórico de registos clínicos eletrónicos, acreditamos que é possível criar e disponibilizar uma ferramenta de apoio à decisão clínica, que permitirá a adesão a protocolos e *guidelines* de tratamento, automatizar as dosagens de medicação no bloco operatório, usar indicadores indiretos para efetuar *screening* de doenças, prever a probabilidade de readmissão hospitalar, criar um índice de expectativa de vida para doentes idosos hospitalizados, determinar indicadores de deteção precoce de deterioração do estado de saúde dos doentes. Este sistema de apoio à decisão permitirá auxiliar os clínicos na tomada de decisão, nomeadamente qual o tratamento a seguir e alertar para condições que tornam o tratamento ineficaz ou ainda informação sobre o custo-efetividade.

4.3. Estrutura conceptual da solução técnica futura

A solução técnica futura será constituída por duas componentes-chave, a extração e consolidação de dados por TM e o SADC. A primeira consiste na obtenção de informação a partir dos registos clínicos existentes e a segunda, como objetivo final, é pretendido o desenvolvimento de um sistema de suporte à decisão clínica (MLBA – *Clinical decision support system*) que auxilie os clínicos na tomada de decisões clínicas.

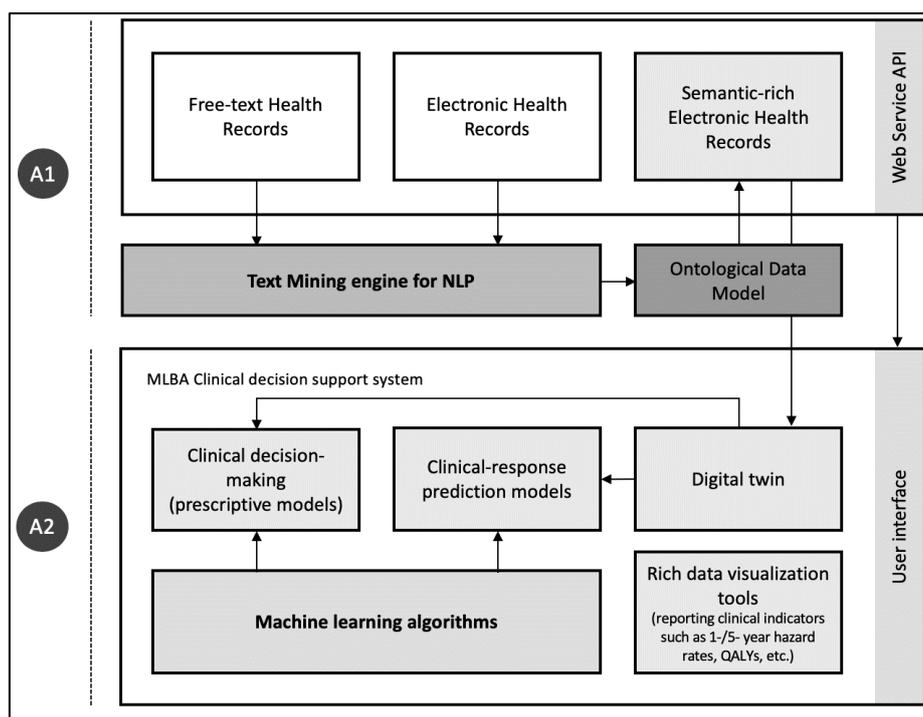


Figura 22 - Arquitetura técnica conceptual

A arquitetura técnica conceptual está dividida em duas atividades distintas (figura 22), mas interrelacionadas, que englobam várias tarefas objetivas da seguinte forma:

- Atividade 1 (A1) – Consiste no desenvolvimento de um algoritmo de TM para extrair dados estruturados a partir de registos clínicos/oncológicos, nomeadamente dados referentes a diagnósticos, dados da patologia clínica, condição de saúde, características do doente e planos SOAP. Inclui ainda outras subactividades, como a recolha do léxico de suporte, a adaptação das ontologias oncológicas existentes (como o *National Cancer Institute Thesaurus*), o treino do algoritmo utilizando registos clínicos armazenados e a extração da informação estruturada num modelo de dados semânticos.

- Atividade 2 (A2) – Serão desenvolvidos métodos capazes de fornecer assistência clínica, reportar dados estatísticos relevantes e aplicar modelos de ML e IA que permitam estimar o possível impacto de um determinado tratamento no doente. Para o efeito serão utilizados os dados recolhidos na A1. Será construído um repositório de dados estruturado com conhecimento extraído dos registos clínicos que servirá para melhorar os modelos de ML, permitindo assim transformar dados clínicos em informações úteis para melhorar os cuidados de saúde dos doentes. Por último, a tarefa principal desta atividade diz respeito ao desenvolvimento de um SADC.

4.4. Resultados do trabalho da validação dos algoritmos de TM

Para cada modelo foram realizados os cálculos para precisão, *recall* e *f1 score* para as classes “*Procedures*”, “*Disorders*”, e “*Drugs*”, cujos dados são apresentados nos quadros seguintes:

i. Modelo semi-supervisionado com pré-anotação do UMLS

	Avaliação Global
Precisão	86.609%
Recall	83.716%
F1	85.138%

Tabela 3 – Avaliação global do modelo semi-supervisionado

	Disorders	Procedures	Drugs
Precisão	80.213%	90.909%	90.990%
Recall	75.376%	87.719%	92.660%
F1	77.720%	89.285%	91.818%

Tabela 4 - Resultados do modelo semi-supervisionado por classe

ii. Modelo supervisionado com base nas anotações do especialista

a. Anotador A

	Avaliação Global
Precisão	63.841%
Recall	54.854%
F1	50.007%

Tabela 5 – Avaliação global do modelo supervisionado com as anotações do anotador A

	Disorders	Drugs	Procedures
Precisão	64.062%	72.549%	56.451%
Recall	52.564%	55.223%	57.377%
F1	57.746%	62.711%	56.910%

Tabela 6 - Resultados do modelo supervisionado com as anotações do anotador A por classe

b. Anotador B

	Avaliação Global
Precisão	65.671%
Recall	63.768%
F1	64.705%

Tabela 7 - Avaliação global do modelo supervisionado com as anotações do anotador B

	Disorders	Drugs	Procedures
Precisão	62.5%	78.571%	61.538%
Recall	71.428%	68.75%	44.444%
F1	66.666%	73.333%	51.612%

Tabela 8 - Resultados do modelo supervisionado com as anotações do anotador B por classe

4.5. Avaliação dos resultados

A avaliação dos modelos não poderá ser realizada através do cálculo da acurácia porque o número de exemplos para cada classe não é igual. Assim, nas tabelas seguintes realizou-se uma análise qualitativa aos vários modelos. O número de frases anotadas foram 2055, no entanto, foram detetadas 44 frases iguais.

Para os modelos anotados pelos anotadores A e B, e para algumas das frases que tiveram a mesma anotação foram analisadas por um especialista C, um médico sénior, 83 frases e respetivas anotações, demonstrado no anexo 9, com os respetivos comentários, dos quais são apresentados alguns exemplos na tabela 9, com os seguintes resultados:

- concordou totalmente com 60;
- adicionou 1 termo/entidade a 18 frases;
- adicionou 2 termos/entidades a 5 frases.

Frases	Anotação dos anotadores A e B	Validação do especialista C (VALIDOU/ NÃO VALIDOU/ ADICIONOU TERMOS)
Realizada endoscopia digestiva alta , sem complicações imediatas (ver relatório no Siima) .	endoscopia digestiva alta (Procedures)	Validou.
Protocolo analgesia : avaliar às 14h45 .	Protocolo analgesia (Procedures)	Validou.
Volta em março de 2015 com TAC , pois passa a vigilância .	TAC (Procedures)	Adicionou 1 termo: 'Vigilância' (Procedures)
Diagnóstico : Adenocarcinoma do reto , encontra.-se horas após RAR .	Adenocarcinoma do reto (Disorders), RAR (Procedures)	Adicionou 1 termo: 'Diagnóstico' (Procedures)
Início suplementação com ácido fólico .	ácido fólico (Drugs)	Adicionou 1 termo: 'suplementação' (Procedures)
Documentado défice de ác . fólico , iniciou suplementação .	ác . fólico (Drugs)	Adicionou 2 termos: 'Suplementação' (Procedures), 'défice' (Disorders)

Tabela 9 – Exemplos da avaliação realizada pelo especialista C a algumas das frases que tiveram a mesma anotação por parte dos anotadores A e B

4.5.1. Avaliação do modelo semi-supervisionado com pré-anotação do UMLS

Para este modelo concluiu-se que tem o modelo uma precisão de 86,609%, ou seja, valor relativo aos positivos previstos, enquanto o *recall* é de 83,713%, ou seja, o número de positivos reais que o modelo encontrou. Neste caso o equilíbrio dado pelo *f1 score* é de 85,138%.

É relevante referir que o modelo semi-supervisionado não realizou qualquer anotação em 795 frases, quando existam termos para anotar.

Pelos valores quantitativos os resultados parecem animadores relativamente à aprendizagem realizada, mas ao serem analisados qualitativamente por um especialista C, um médico sénior, foram detetadas algumas falhas demonstradas no anexo 10, com os respetivos comentários, dos quais são apresentados no quadro seguinte alguns exemplos:

Frases	Entidades detetadas	Avaliação do especialista C sobre a anotação: SIM/NÃO/PARCIAL
Colhido hemograma + bioquímica .		NÃO. Deveria ter detetado 'bioquímica'
Sem sinais de alterações de mobilidade e/ou sensibilidade dos membros inferiores .	sensibilidade (Disorders)	PARCIAL. Deveria ter classificado 'mobilidade' (Disorders)
Peso : 58 kg (perda de 2 Kg dese a última consulta) Orofaringe sem lesões , sem exsudados .	Peso (Procedures)	PARCIAL. Deveria ter classificado 'orofaringe sem lesões' (Disorders)
Circulante Ryan Amanda Geada de Michael Trigo Lourenco (1214) - Enf .	Circulante (Drugs)	NÃO. Circulante é uma profissão
Faz tratamento com suporte de G CSF .	tratamento (Procedures)	PARCIAL. Deveria ter classificado 'G CSF' (Drugs)
Antropometria : P atual : 52kg P hab : 62 Kg (Abr2018) E : 1.53 m IMC : Magreza Ingestão alimentar hab : Via : oral Consistência : geral Levanta se pelas 7h30 PA (7:30h) - 1 chav de leite MG c / cevada s/açúcar c / Nestum ou 3-4 bolachas Cynthia MM (10h30) - 1 iog sól ou banana A (12:30h) - Prato : (50 % hab , melhor tolerância para o peixe) MT (16h30) - = PA 18h - canja ou sopa legumes s/HC , c/leguminosas , as x c / FP .	Magreza (Disorders), leite (Disorders), Antropometria (Procedures), açúcar (Drugs)	NÃO.

Tabela 10 – Exemplos da avaliação realizada pelo especialista C ao modelo semi-supervisionado

Concluiu-se, portanto, que neste caso a análise quantitativa dos resultados não é a mais adequada, pois na análise qualitativa dos resultados do modelo semi-supervisionado com UMLS verificou-se que ainda tem falhas e, portanto, necessita de ser melhorado, por exemplo através da adição de regras.

4.5.2. Avaliação do modelo supervisionado com base nas anotações dos anotadores A e B

Analisados os resultados dos modelos treinados com as anotações individuais de cada anotador concluiu-se que as diferenças não são pontuais, mas sim representativas. Apesar dos resultados quantitativos do anotador B serem melhores do que os do anotador A, para avaliar as anotações efetuadas, foi realizada uma terceira validação por um especialista C, um médico sénior, utilizando uma amostra de frases, apresentando-se nos quadros seguintes os comparativos relativos a frases em que ambos os anotadores registaram e a informação sobre as frases anotadas, frases em que apenas um dos anotadores registou, e anotações do modelo semi-supervisionada e os modelos dos anotadores A e B.

4.5.2.1. Comparativo de frases em que ambos os anotadores A e B registaram

Os resultados são apresentados no anexo 11, com os respectivos comentários, dos quais são apresentados alguns exemplos:

Frases	Modelo anotador A	Modelo anotador B	Informação do especialista C sobre a anotação mais completa: A/B/OUTRA
Pedido de TAC Toracico de estadiamento e análises .	TAC Torácico (Procedures), análises (Procedures)	TAC (Procedures), análises (Procedures)	A está correta.
PCEA a 5ml/h que perfunde sem intercorrências .	PCEA (Drugs)	PCEA (Procedures)	A e B corretas. A frase não deveria ter a medicação como PCEA.
ECOG o Pele e mucosas coradas e hidratadas Altura - 152 cm Peso 64 kg SC - 1.61 m2 SPP resolvido Análises de hoje : neutropenia grau 2 (1.290/u L) .	Análises (Procedures), neutropenia grau 2 (Disorders)	Análises (Procedures), neutropenia (Disorders)	A mais precisa. Falta classificação 'ECOG' (Disorders), 'Pele e mucosas coradas e hidratadas' (Disorders)
ECOG o Pele e mucosas coradas e hidratadas Altura - 152 cm Peso 69 kg (aumento de 5 kg) ACP - sem alterações de relevo Abdomen - sem alterações de relevo Sem edemas periféricos Análises de hoje : anemia resolvida ; neutropenia em recuperação , neste momento grau 1 .	Análises (Procedures), anemia (Disorders), neutropenia (Disorders)	Análises (Procedures), anemia (Disorders), edemas periféricos (Disorders), neutropenia (Disorders)	B. Falta classificação 'ECOG' (Disorders), 'Pele e mucosas coradas e hidratadas' (Disorders)
ECOG o Pele e mucosas coradas e hidratadas Altura - 152 cm Peso 64 kg SC - 1.61 m2 Sem alterações de relevo de novo E.O . Análises de hoje : mantém neutropenia grau 2 (1.370/u L) .	Análises (Procedures), neutropenia grau 2 (Disorders)	Análises (Procedures), neutropenia (Disorders)	A mais precisa. Falta classificação 'Pele e mucosas coradas e hidratadas' (Disorders)
Colocada bomba de 5FU.Saíu bem .	bomba de 5FU (Drugs)	5FU (Drugs)	B.

Tabela 11 - Exemplos da avaliação realizada pelo especialista C aos modelos dos anotadores A e B

Desta avaliação concluiu-se que o anotador A é mais preciso nas anotações, mas o anotador B tem mais anotações corretas.

Apresentam-se ainda mais alguns dados referentes às anotações dos anotadores A e B:

- Não realizaram qualquer anotação em 623 frases;
- Realizaram as mesmas anotações em 299 frases e nestas:
 - Anotaram a entidade “*Procedures*” em 117 frases;
 - Anotaram a entidade “*Disorders*” em 115 frases;
 - Anotaram a entidade “*Drugs*” em 112 frases.

4.5.2.2. Comparativo de frases em que apenas um dos anotadores registou

Os resultados são apresentados no anexo 12, com os respetivos comentários, dos quais são apresentados no quadro seguinte alguns exemplos:

Frases	Modelo anotador A	Modelo anotador B	Informação do especialista C sobre a anotação mais completa: A/B/OUTRA
Dadas orientações sobre o procedimento e sua colaboração durante o mesmo .		Orientações (Procedures)	B.
Doente com clearance de creatinina de 86.89 mg/dl .	clearance de creatinina (Procedures)		A.
Estadio patológico p T1b N1 Mo Ro FU de 8 meses .	Estadio patológico p T1b N1 Mo Ro FU (Disorders)		A.
Peso : 58 kg (perda de 2 Kg dese a última consulta) Orofaringe sem lesões , sem exsudados .		lesões (Disorders)	B. Falta 'exsudados' (Disorders)
Sem icterícia Sem adenomegalias ABD : Mole e depressível .		adenomegalias (Disorders), icterícia (Disorders)	B.
Faz tratamento com suporte de G CSF .	G CSF (Drugs)		A.

Tabela 12 – Exemplos da avaliação realizada pelo especialista C aos modelos dos anotadores A e B, em frases em que apenas um anotou

Desta análise concluiu-se que para todas as frases existiam termos para anotar e que tanto o anotador A como o anotador B, não anotaram. Ainda assim, para as frases anotadas algumas anotações não estavam completas. De referir ainda que para as 2055 frases, o anotador A não anotou 906 e o anotador B não anotou 764.

4.5.3. Comparativo das anotações dos modelos semi-supervisionado, A e B.

Na tabela seguinte é possível verificar várias diferenças entre os 3 modelos que foram treinados, para as classes “*Procedures*”, “*Disorders*”, “*Drugs*”.

Frase	Modelo semi-supervisionado	Modelo do anotador A	Modelo do anotador B
Fica internada por anemia sintomática	'anemia' (Disorders)	'anemia sintomática' (Disorders)	'anemia' (Disorders)
Internamento por anemia sintomática por perda tumoral gástrica , em 05/2019 .	'anemia sintomática' (Disorders)	'anemia sintomática' (Disorders)	'anemia' (Disorders)
	'perda tumoral gástrica' (Disorders)		'perda tumoral gástrica' (Disorders)
Frequência : 12/12 h Nome Genérico : Sucralfato 1000 mg/5 ml Susp oral Saq Via Administração : ORAL ;	'Sucralfato 1000' (Drugs)	'Sucralfato 1000 mg/5 ml' (Drugs)	'Sucralfato' (Drugs)
Doente com diagnóstico de Ca gástrico metastizado .	'Ca gástrico' (Disorders)	'Ca gástrico metastizado' (Disorders)	'Ca gástrico metastizado' (Disorders)
Antenia marcada e sensação de desmaio .		'sensação de desmaio' (Disorders)	'sensação de desmaio' (Disorders)
Desconforto abdominal agora melhroado.		'Desconforto abdominal' (Disorder)	'Desconforto abdominal' (Disorder)
Foi realizada análise imunohistoquímica de MLH1/MSH2/MSH6/PMS2 , tendo sido normal .		'análise imunohistoquímica' (Procedures)	'análise imunohistoquímica' (Procedures)
Efetuada endoscopia digestiva baixa - ver relatório Sii MA .	'endoscopia digestiva' (Procedures)	'endoscopia digestiva baixa' (Procedures)	'endoscopia digestiva baixa' (Procedures)
Pedido anterior recusado por falta de TAC e resultado de estudo histológico das biopsias.	'TAC' (Procedures)	'TAC' (Procedures)	'TAC', (Procedures)
	'estudo histológico' (Procedures)	'estudo histológico' (Procedures)	'estudo histológico' (Procedures)
			'biopsias' (Procedures)

Tabela 13 – Exemplos de diferenças entre o modelo semi-supervisionado e os modelos dos anotadores A e B

Pelas diferenças encontradas é possível concluir a existência de acertos parciais de termos. Algumas anotações não foram anotadas da mesma forma, principalmente quando a frase contém valores e unidades de medida. Será, portanto, necessário melhorar a anotação para aumentar o grau de acerto.

5. Conclusões e perspectivas de desenvolvimento

Os registos clínicos em “texto livre” permitem liberdade da escrita, mas não facilitam à análise dos dados. Estando os registos codificados é possível estudar padrões, a sua evolução, as respostas a tratamentos, as complicações, a sobrevivência e com base nesses dados desenvolver aplicativos de suporte à tomada de decisão, podendo também ser utilizadas para estudos.

Do trabalho realizado pelos anotadores foi possível obter resultados com uma precisão superior a 60%. Concluiu-se ainda que a metodologia permitirá cumprir os objetivos, mas será necessário reforçar a formação dos anotadores, corrigir as anotações e realizar novo estudo piloto com mais casos e consequente validação clínica.

Foram encontradas algumas limitações na realização deste trabalho, das quais destacamos a pouca informação disponível sobre TM e PLN aplicado à área da medicina. Ao nível dos processos clínicos foram encontrados muitos textos desorganizados, não respeitando regras de pontuação, erros ortográficos e recurso frequente a abreviaturas não oficiais ou não reconhecidas. No decurso deste trabalho realizado em âmbito hospitalar ocorreu a pandemia o que dificultou todas as ações em curso, nomeadamente a alocação de profissionais de saúde às tarefas de anotação.

Como perspectivas de desenvolvimento, para melhorar o trabalho e a acurácia, seria aconselhado aumentar o número de anotadores e de textos a serem analisados para que o modelo aprenda de forma mais rápida e proficiente.

Sugere-se que como passo seguinte a implementação de um *dashboard*, semelhante ao do anexo 13, que incorpore vários elementos, como é o caso de uma linha temporal que indique o fluxo de cuidados, de diagnóstico e de decisões tomadas, possibilitando perceber o momento exato em que o doente se encontra e as respetivas decisões clínicas até então tomadas. Esta deverá ser de utilização simples e intuitiva, que apoie e facilite o dia a dia de cada médico, dando assim uma nova vitalidade no apoio à tomada de decisão médica, que se traduzirá num reforço contínuo de melhoria nos serviços prestados aos doentes da instituição.

6. Referências bibliográficas

- Al-Haddad, M. A., Friedlin, J., Kesterson, J., Waters, J. A., Aguilar-Saavedra, J. R., & Schmidt, C. M. (2010). Natural language processing for the development of a clinical registry: A validation study in intraductal papillary mucinous neoplasms. *HPB*.
<https://doi.org/10.1111/j.1477-2574.2010.00235.x>
- Allahyari, M., Pouriye, S. A., Assefi, M., Safaei, S., Trippe, E. D., Gutierrez, J. B., & Kochut, K. (2017). A Brief Survey of Text Mining: Classification, Clustering and Extraction Techniques. *CoRR*, *abs/1707.0*.
- Andrade, L. (2018). Inteligência Artificial e Medicina: Haverá Lugar para o Fator Humano?... *Revista Portuguesa de Farmacoterapia*, *10*(1), 53–55.
<https://doi.org/10.25756/rpf.v10i1.189>
- Bain, C., & Manus, C. Mac. (2014). Advancing data management and usage in a major Australian health service: The REASON discovery platform™. *Proceedings - 2014 International Conference on Data Science and Engineering, ICDSE 2014*.
<https://doi.org/10.1109/ICDSE.2014.6974609>
- Blake, C. (2011). Text mining. *Annual Review of Information Science and Technology*, *45*(1), 121–155. <https://doi.org/10.1002/aris.2011.1440450110>
- Bodenreider, O. (2004). The Unified Medical Language System (UMLS): Integrating biomedical terminology. *Nucleic Acids Research*. <https://doi.org/10.1093/nar/gkh061>
- Chen, L. (2019). Assertion Detection in Clinical Natural Language Processing: A Knowledge-Poor Machine Learning Approach. *2019 IEEE 2nd International Conference on Information and Computer Technologies (ICICT)*, 37–40.
<https://doi.org/10.1109/INFOCT.2019.8710921>
- Clancey, W., & Shortliffe, E. (1984). *Readings in Medical Artificial Intelligence : The First Decade*. Addison Wesley.
- Coiera, E. W. (1998). *Inteligência Artificial em Medicina*. *Revista de Informática Médica*. 1 (4).
- Culotta, A., Kristjansson, T., McCallum, A., & Viola, P. (2006). Corrective feedback and persistent learning for information extraction. *Artificial Intelligence*.
<https://doi.org/10.1016/j.artint.2006.08.001>
- Dalianis, H. (2018). Clinical text mining: Secondary use of electronic patient records. In

- Clinical Text Mining: Secondary Use of Electronic Patient Records*. Springer International Publishing. <https://doi.org/10.1007/978-3-319-78503-5>
- Delespierre, T., Denormandie, P., Bar-Hen, A., & Josseran, L. (2017). Empirical advances with text mining of electronic health records. *BMC Medical Informatics and Decision Making*. <https://doi.org/10.1186/s12911-017-0519-0>
- Enelow, A. J., Forde, D. L., & Brummel-Smith, K. (1999). *Entrevista clínica e cuidados ao paciente*. Climepsi Editores.
- Eysenck, M. W., & Keane, M. T. (2015). Cognitive psychology: A student's handbook, 7th ed. In *Cognitive psychology: A student's handbook, 7th ed.* Psychology Press. <https://doi.org/10.4324/9781315778006>
- Fayyad, U., Piatetsky-Shapiro, G., & Smyth, P. (1996). From data mining to knowledge discovery in databases. *AI Magazine*.
- Feldman, K., Hazekamp, N., & Chawla, N. V. (2016). Mining the Clinical Narrative: All Text are Not Equal. *Proceedings - 2016 IEEE International Conference on Healthcare Informatics, ICHI 2016*. <https://doi.org/10.1109/ICHI.2016.37>
- Ferreira, E. A. L. (2015). *Mineração de Dados Aplicada à Dados Médicos*. Universidade Federal de Minas Gerais.
- Focsa, M. A. (2016). An architectural approach for building medical ontologies. *Applied Medical Informatics*, 38(2), 66–72.
- Freitas, F., & Schulz, S. (2009). Ontologies, semantic Web and health. *RECIIS*. <https://doi.org/10.3395/reciis.v3i1.238en>
- Friedman, C., & Hripcsak, G. (1998). Evaluating Natural Language Processors in the Clinical Domain. *Methods of Information in Medicine*, 37(04/05), 334–344. <https://doi.org/10.1055/s-0038-1634566>
- Friedman, Carol, Shagina, L., Lussier, Y., & Hripcsak, G. (2004). Automated encoding of clinical documents based on natural language processing. *Journal of the American Medical Informatics Association*. <https://doi.org/10.1197/jamia.M1552>
- Gardner, H. (2002). *A nova ciência da mente*. Relógio D'Água.
- Gardner, R. L., Cooper, E., Haskell, J., Harris, D. A., Poplau, S., Kroth, P. J., & Linzer, M. (2019). Physician stress and burnout: the impact of health information technology. *Journal of the American Medical Informatics Association*, 26(2), 106–114. <https://doi.org/10.1093/jamia/ocy145>

- Gawande, A. (2015). *O Efeito Checklist*. Lua de Papel.
- Goldhahn, J., Rampton, V., & Spinass, G. A. (2018). Could artificial intelligence make doctors obsolete? *BMJ*, k4563. <https://doi.org/10.1136/bmj.k4563>
- Gruber, T. R. (1995). Toward principles for the design of ontologies used for knowledge sharing? *International Journal of Human-Computer Studies*, 43(5), 907–928. <https://doi.org/https://doi.org/10.1006/ijhc.1995.1081>
- Halkidi, M., & Vazirgiannis, M. (2009). Quality Assessment Approaches in Data Mining. In *Data Mining and Knowledge Discovery Handbook*. https://doi.org/10.1007/978-0-387-09823-4_31
- Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P., & Witten, I. H. (2009). The WEKA data mining software: An update. *ACM SIGKDD Explorations Newsletter*.
- Hardjojo, A., Gunachandran, A., Pang, L., Abdullah, M. R. Bin, Wah, W., Chong, J. W. C., Goh, E. H., Teo, S. H., Lim, G., Lee, M. L., Hsu, W., Lee, V., Chen, M. I. C., Wong, F., & Phang, J. S. K. (2018). Validation of a natural language processing algorithm for detecting infectious disease symptoms in primary care electronic medical records in Singapore. *Journal of Medical Internet Research*. <https://doi.org/10.2196/medinform.8204>
- Hotho, A., Nürnberger, A., & Paass, G. (2005). A Brief Survey of Text Mining. *LDV Forum - GLDV Journal for Computational Linguistics and Language Technology*, 20, 19–62.
- Kahneman, D. (2012). *Pensar, Depressa e Devagar*. Temas & Debates.
- Kaplan, A., & Haenlein, M. (2018). Siri, Siri, in my hand: Who's the fairest in the land? On the interpretations, illustrations, and implications of artificial intelligence. *Business Horizons*, 62. <https://doi.org/10.1016/j.bushor.2018.08.004>
- Kavakiotis, I., Tsave, O., Salifoglou, A., Maglaveras, N., Vlahavas, I., & Chouvarda, I. (2017). Machine Learning and Data Mining Methods in Diabetes Research. *Computational and Structural Biotechnology Journal*, 15, 104–116. <https://doi.org/10.1016/j.csbj.2016.12.005>
- Kocbek, S., Cavedon, L., Martinez, D., Bain, C., Manus, C. Mac, Haffari, G., Zukerman, I., & Verspoor, K. (2016). Text mining electronic hospital records to automatically classify admissions against disease: Measuring the impact of linking data sources. *Journal of Biomedical Informatics*. <https://doi.org/10.1016/j.jbi.2016.10.008>
- Kotthoff, L. (2016). Algorithm selection for combinatorial search problems: A survey. In

Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). https://doi.org/10.1007/978-3-319-50137-6_7

- Kourou, K., Exarchos, T. P., Exarchos, K. P., Karamouzis, M. V., & Fotiadis, D. I. (2015). Machine learning applications in cancer prognosis and prediction. *Computational and Structural Biotechnology Journal*, 13, 8–17. <https://doi.org/10.1016/j.csbj.2014.11.005>
- Kumar, A., Alam, H., Kumar, R., & Sheel, S. (2015). Understanding Medical Named Entity Extraction in Clinical Notes. In *Int'l Conf. Health Informatics and Medical Systems*.
- Kundel, H. L., Nodine, C. F., Conant, E. F., & Weinstein, S. P. (2007). Holistic Component of Image Perception in Mammogram Interpretation: Gaze-tracking Study. *Radiology*, 242(2), 396–402. <https://doi.org/10.1148/radiol.2422051997>
- Lao, J., Chen, Y., Li, Z. C., Li, Q., Zhang, J., Liu, J., & Zhai, G. (2017). A Deep Learning-Based Radiomics Model for Prediction of Survival in Glioblastoma Multiforme. *Scientific Reports*. <https://doi.org/10.1038/s41598-017-10649-8>
- Ledley, R. S., & Lusted, L. B. (1959). Reasoning foundations of medical diagnosis. *Science*. <https://doi.org/10.1126/science.130.3366.9>
- Lee, I., & Shin, Y. J. (2020). Machine learning for enterprises: Applications, algorithm selection, and challenges. *Business Horizons*. <https://doi.org/10.1016/j.bushor.2019.10.005>
- Levinson, J., Price, B. H., & Saini, V. (2017). *Death By A Thousand Clicks: Leading Boston Doctors Decry Electronic Medical Records*. Wbur. <https://www.wbur.org/commonhealth/2017/05/12/boston-electronic-medical-records>
- Li, Q., Bai, H., Chen, Y., Sun, Q., Liu, L., Zhou, S., Wang, G., Liang, C., & Li, Z. C. (2017). A Fully-Automatic Multiparametric Radiomics Model: Towards Reproducible and Prognostic Imaging Signature for Prediction of Overall Survival in Glioblastoma Multiforme. *Scientific Reports*. <https://doi.org/10.1038/s41598-017-14753-7>
- Liberali, J. F. de M. (2012). *O papel da memória no julgamento e tomada de decisão*. Universidade Católica do Rio Grande do Sul.
- Lobo, L. C. (2018). Inteligência artificial, o Futuro da Medicina e a Educação. *Revista Brasileira de Educação*, 42, 3–8.
- Luque, C., Luna, J. M., Luque, M., & Ventura, S. (2019). An advanced review on text mining in medicine. In *Wiley Interdisciplinary Reviews: Data Mining and Knowledge*

- Discovery*. <https://doi.org/10.1002/widm.1302>
- Maculan, B., & Aganette, E. (2020). Desambiguação de relações em tesouros e o seu reuso em ontologias. *Ciência Da Informação*, 46, 119.
- Mendonça, E. A., Haas, J., Shagina, L., Larson, E., & Friedman, C. (2005). Extracting information on pneumonia in infants using natural language processing of radiology reports. *Journal of Biomedical Informatics*, 38(4), 314–321.
<https://doi.org/10.1016/j.jbi.2005.02.003>
- Mjolsness, E., & Decoste, D. (2001). Machine Learning for Science: State of the Art and Future Prospects. *Science*, 293(5537), 2051–2055.
<https://doi.org/10.1126/science.293.5537.2051>
- Nadeau, D., & Sekine, S. (2007). A survey of named entity recognition and classification. *Linguisticae Investigationes*, 30(1), 3–26. <https://doi.org/10.1075/li.30.1.03nad>
- Neufeld, C. B., Brust, P. G., & Stein, L. M. (2011). Bases epistemológicas da psicologia cognitiva experimental. *Psicologia: Teoria e Pesquisa*. <https://doi.org/10.1590/S0102-37722011000100013>
- Ordem dos Médicos. (2019). *Tempos padrão para consultas – discussão pública*.
<https://ordemosmedicos.pt/tempos-padrao-para-consultas-discussao-publica/>
- Queiroz, M. J. (2009). SOAP revisitado. *Revista Portuguesa de Clínica Geral*.
<https://doi.org/10.32385/rpmgf.v25i2.10610>
- Réa-Neto, A. (1998). Raciocínio clínico - o processo de decisão diagnóstica e terapêutica. *Revista Da Associação Médica Brasileira*, 44, 301–311.
- Regueira, Ca. (2018). *A Aceitação e Uso dos Sistemas de Informação pelos Técnicos Superiores de Diagnóstico e Terapêutica. Um Estudo de Caso no Hospital de Santarém*. Instituto Politécnico de Tomar.
- Rodrigues, H. J. F. (2016). *Ferramenta para Text Mining em Textos completos*. Universidade do Porto.
- Russel, S., & Norvig, P. (2012). Artificial intelligence—a modern approach 3rd Edition. In *The Knowledge Engineering Review*. <https://doi.org/10.1017/S0269888900007724>
- Sahiner, B., Pezeshk, A., Hadjiiski, L. M., Wang, X., Drukker, K., Cha, K. H., Summers, R. M., & Giger, M. L. (2019). Deep learning in medical imaging and radiation therapy. In *Medical Physics*. <https://doi.org/10.1002/mp.13264>

- Sarkiss, C. A., & Germano, I. M. (2019). Machine Learning in Neuro-Oncology: Can Data Analysis From 5346 Patients Change Decision-Making Paradigms? In *World Neurosurgery*. <https://doi.org/10.1016/j.wneu.2019.01.046>
- Savova, G. K., Masanz, J. J., Ogren, P. V., Zheng, J., Sohn, S., Kipper-Schuler, K. C., & Chute, C. G. (2010). Mayo clinical Text Analysis and Knowledge Extraction System (cTAKES): Architecture, component evaluation and applications. *Journal of the American Medical Informatics Association*. <https://doi.org/10.1136/jamia.2009.001560>
- Sezer, E., Unalir, M. O., Can, O., & Bursa, O. (2019). A Methodology for Defining Working Healthcare Ontologies. *International Journal of Scientific and Technological Research*, 5, 146–156. <https://doi.org/10.7176/jstr/5-4-15>
- Sheikhalishahi, S., Miotto, R., Dudley, J. T., Lavelli, A., Rinaldi, F., & Osmani, V. (2019). Natural language processing of clinical notes on chronic diseases: Systematic review. In *Journal of Medical Internet Research*. <https://doi.org/10.2196/12239>
- Shen, W., Wang, J., & Han, J. (2015). Entity linking with a knowledge base: Issues, techniques, and solutions. *IEEE Transactions on Knowledge and Data Engineering*. <https://doi.org/10.1109/TKDE.2014.2327028>
- Shortliffe, E. H. (1990). *Medical Informatics – Computer Applications in Medical Care*. Addison-Wesley.
- ShrihariR, C., & Desai, A. (2015). A Review on Knowledge Discovery using Text Classification Techniques in Text Mining. *International Journal of Computer Applications*. <https://doi.org/10.5120/19542-0784>
- Silva, A. J. O. (2016). *Text Mining e processamento de linguagem natural para interpretação de notas clínicas*. Universidade do Minho.
- Simmons, M., Singhal, A., & Lu, Z. (2016). Text Mining for Precision Medicine: Bringing Structure to EHRs and Biomedical Literature to Understand Genes and Health. *Advances in Experimental Medicine and Biology*, 939, 139–166. https://doi.org/10.1007/978-981-10-1503-8_7
- Simon, H. A. (1997). *Administrative Behavior, 4th Edition*. Free Press; 0004- edition (March 1, 1997).
- Studer, R., Benjamins, V. R., & Fensel, D. (1998). Knowledge engineering: Principles and methods. *Data & Knowledge Engineering*, 25(1–2), 161–197. [https://doi.org/10.1016/S0169-023X\(97\)00056-6](https://doi.org/10.1016/S0169-023X(97)00056-6)

- Sun, W., Cai, Z., Li, Y., Liu, F., Fang, S., & Wang, G. (2018). Data processing and text mining technologies on electronic medical records: A review. In *Journal of Healthcare Engineering*. <https://doi.org/10.1155/2018/4302425>
- Tan, A.-H. (1999). Text Mining: The state of the art and the challenges. *Proceedings of the PAKDD 1999 Workshop on Knowledge Discovery from Advanced Databases*. <https://doi.org/10.1.1.38.7672>
- Tissot, H., Roberts, A., Derczynski, L., Gorrell, G., & Del Fabro, M. D. (2015). Analysis of temporal expressions annotated in clinical notes. *Proceedings 11th Joint ACL-ISO Workshop on Interoperable Semantic Annotation (Isa-11)*.
- Tversky, A., & Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science*, 211(4481), 453–458. <https://doi.org/10.1126/science.7455683>
- U.S. Food and Drug Administration. (2018). *FDA permits marketing of artificial intelligence algorithm for aiding providers in detecting wrist fractures*. <https://www.fda.gov/news-events/press-announcements/fda-permits-marketing-artificial-intelligence-algorithm-aiding-providers-detecting-wrist-fractures>
- U.S. Food and Drug Administration. (2020). *FDA Authorizes Marketing of First Cardiac Ultrasound Software That Uses Artificial Intelligence to Guide User*. <https://www.fda.gov/news-events/press-announcements/fda-authorizes-marketing-first-cardiac-ultrasound-software-uses-artificial-intelligence-guide-user>
- Valusek, J. R. (2002). Decision support: a paradigm addition for patient safety. *Journal of Healthcare Information Management : JHIM*.
- van Laar, S. A., Gombert-Handoko, K. B., Guchelaar, H. J., & Zwaveling, J. (2020). An Electronic Health Record Text Mining Tool to Collect Real-World Drug Treatment Outcomes: A Validation Study in Patients With Metastatic Renal Cell Carcinoma. *Clinical Pharmacology and Therapeutics*. <https://doi.org/10.1002/cpt.1966>
- Vasconcelos, J., Rocha, Á., & Gomes, R. (2004). Sistemas de Informação de Apoio à Decisão Clínica: Estudo de um caso de uma Instituição de Saúde. *5a Conferência Da Associação Portuguesa de Sistemas de Informação*. <https://doi.org/10.3389/fpsyg.2015.00578>
- Vaz Carneiro, A. (2003). O raciocínio clínico. Qual a sua natureza? Pode ensinar-se? In *Revista Portuguesa de Cardiologia* (pp. 22:433–443).
- Walsh, S., de Jong, E. E. C., van Timmeren, J. E., Ibrahim, A., Compter, I., Peerlings, J., Sanduleanu, S., Refaee, T., Keek, S., Larue, R. T. H. M., van Wijk, Y., Even, A. J. G.,

- Jochems, A., Barakat, M. S., Leijenaar, R. T. H., & Lambin, P. (2019). Decision Support Systems in Oncology. *JCO Clinical Cancer Informatics*.
<https://doi.org/10.1200/cci.18.00001>
- Werner, D., Cliff, J., Mariano, A., & Munguambe, K. (2009). *Onde Nao Ha Medico* (Hesperian Foundation (ed.); pp. 60–85). Health Books International/TALC.
- What to expect from AI in oncology. (2019). *Nature Reviews Clinical Oncology*, 16(11), 655–655. <https://doi.org/10.1038/s41571-019-0283-4>
- Wiens, J., & Shenoy, E. S. (2018). Machine Learning for Healthcare: On the Verge of a Major Shift in Healthcare Epidemiology. *Clinical Infectious Diseases*, 66(1), 149–153.
<https://doi.org/10.1093/cid/cix731>
- Wolpert, D. H., & Macready, W. G. (1997). No free lunch theorems for optimization. *IEEE Transactions on Evolutionary Computation*. <https://doi.org/10.1109/4235.585893>
- Wong, H. J., Legnini, M. W., & Whitmore, H. H. (2000). The diffusion of decision support systems in healthcare: Are we there yet? *Journal of Healthcare Management*.
<https://doi.org/10.1097/00115514-200007000-00009>
- Wright, G. (1984). *Behavioural Decision Theory: An Introduction*. Penguin Education.

Anexos

Anexo 1 – Exemplo de um diário clínico de um doente

Peso: ___ kg

Doente de ___ anos

Motivo de referenciação: Neoplasia do sigmoide

HDA: Doente realiza tds os anos PSOF, que foi positiva. O que motivou a realização de colonoscopia. Doente assintomática, sem perdas hemáticas, sem dor associada. sem alterações do TGI, sem emagrecimento.

Antecedentes pessoais: Alterações do metabolismo lipídico; Perturbações depressivas; Hipertensão não complicada

Antecedentes cirurgicos: Pulso direito, varizes, raspagem uterina

Medicação crónica: Losartan

Alergias: Nega

Historia familiar: Desconhece antecedentes oncologicos

E.O.

Peso: ___ kg

Ecog. o

Sem adenopatias supraclaviculares

Sem organomegalias ou tumefações palpaveis.

TR não tóco lesões, sem sangue na luva

Exames complementares:

Colonoscopia total - Colonoscopia total com identificação do orifício apendicular e da válvula ileo-cecal. Válvula ileo-cecal sem alterações. Má preparação intestinal com lagos de fezes e resíduos fecais sólidos não aspiráveis na totalidade e limitando a observação da mucosa, em particular do colon direito.

Observados divertículos ao longo de todos os segmentos cólicos explorados.

Entre os 18-23cm da margem anal, observada neoplasia ulcerada, friável, ocupando cerca de 3/4 da circunferência luminal _ realizadas biópsias e tatuagem 3cm distalmente em 2 pontos.

Restante mucosa do recto (observada em retroflexão) e colon com padrão vascular preservado, não friável. Não foram observadas outras lesões da mucosa explorada.

Histologia: Adenocarcinoma SOE de baixo grau no colon sigmoide.

TAC estadiamento: Diverticulose sigmóide exuberante, todo o cólon descendente, transverso e menos do ascendente. Relativa estenose do reto, transição sigmóide, espessamento difuso, ao longo de 8 cm, lesão proliferativa e pequenas formações ganglionares secundárias adjacentes suspeitas, as maiores com 7 mm à dir e esq, a estudar por RMN.

P- Peça RM pelvica pela duvida se estenose do recto e extensão de 8 cm- distal?, apesar da lesão estar localizada acerca de 18-23 cm . Peça estudo pre-op.

Consulta de grupo multidisciplinar com doente, cirurgia primaria? QT/RT neoadjuvante.

Anexo 2 – Exemplos do levantamento de siglas que constam nos diários clínicos e respetivo significado

Sigla	Significado
ADC	Adenocarcinoma
AIT	Ataque Isquémico Transitório
APS	Atenção Primária à Saúde
AVC	Acidente Vascular Cerebral
BFAE	Bloqueio Fascicular Anterior Esquerdo
BIRD	Bloqueio incompleto do ramo direito
BMN	Bócio multinodular
Ca	Cancro
CVP	Cateter Venoso Periférico
Cx	Cirurgia
DM	Diabetes mellitus
DMID	Diabetes Mellitus insulino-dependente
DMNID	Diabetes Mellitus não insulino-dependente
DPC	Duodenopancreatectomia cefálica
DPOC	Doença pulmonar obstrutiva crónica
DRC	Doença Renal Crónica
DRC	Doença renal crónica
EA	Estenose aórtica
EAM	Enfarte Agudo Miocárdio
FA	Fibrilação auricular
FAV	Fístula arteriovenosa
HBP	Hiperplasia benigna da próstata

Sigla	Significado
HDA	Hemorragia digestiva alta
HT	Hormonoterapia
HTA	Hipertensão arterial
Hx	História
IAP	Insuficiência Arterial Periférica
IC	Insuficiência cardíaca
ICC	Insuficiência cardíaca congestiva
IRC	Insuficiência Renal Aguda
ITU	Infeção do trato urinário
IVP	Insuficiência Vascular Periférica
MID/E	Membro inferior direito/esquerdo
MM	Mieloma Múltiplo
MSD	Membro Superior Direito
MSD/E	Membro superior direito/esquerdo
Mtx	Metástases
QT	Quimioterapia
RT	Radioterapia
SAOS	Síndrome da Apneia Obstrutiva do Sono
SR	Supra-renal
TAC	Tomografia Axial Computorizada
TC	Tomografia Computorizada
TEP ou TP	Tromboembolismo Pulmonar
TVP	Trombose Venosa Profunda

Anexo 3 – Exemplos de dificuldades na segmentação das frases, devido a pouca pontuação, pontuação em abreviaturas e também falta de pontuação no final de frases

Frase
<p>Abertura da boca : sem alterações relevantes ALERGIAS:o EXAMES COMPLEMENTARES DE CURTIS ÓSTICO A . Hemograma : Hgb 14.2 gr/dl Plaquetas:232.000 B . Bioquímica:N C . Hemostase:N D . E.C.G:N E . Ecocardiograma : F . Rx Tórax:ok G . P.F.R : H . Outros MEDICACAO HABITUAL : o SINAIS VITAIS Pulso : 92 / min TA : 119/83 / mm Hg SAT O2 : 100 % ASA:II MEDICACAO PRE ANESTESICA : PROPOSTA ANESTESICA : Anestesia geral CONSENTIMENTO (S/N) : S NECESSIDADES ESPECIAIS:S UC .</p>
<p>Benjamin Horton Thomas , ONCOLOGIA CIRURGICA - AREA A - - - < < # # date:13832 # # > > : Angiografia hepática e embolização simples de adenoma hepatico Sem queixas Ligeira equimose antebraço esquerdo Abdomen indolor , sem massas palpaveis Tem ja novo TC e nova embolização marcada para Janeiro 2018 Peça Benjaminlises e consulta para fev 2018 Radiologia Intervenção p / Rx em Radiologia de Intervenção Geral (Rad Int) , < < # # date:13832 # # > > , Dra .</p>
<p>Abertura da boca : sem alterações relevantes ALERGIAS:o EXAMES COMPLEMENTARES DE DIAGNÓSTICO A . Hemograma : Hgb 14.2 gr/dl Plaquetas:232.000 B . Bioquímica:N C . Hemostase:N D . E.C.G:N E . Ecocardiograma : F . Rx Tórax:ok G . P.F.R : H . Outros MEDICACAO HABITUAL : o SINAIS VITAIS Pulso : 92 / min TA : 119/83 / mm Hg SAT O2 : 100 % ASA:II MEDICACAO PRE ANESTESICA : PROPOSTA ANESTESICA : Anestesia geral CONSENTIMENTO (S/N) : S NECESSIDADES ESPECIAIS:S UC .</p>
<p>Pele e mucosas descoradas Ligeiras queixas algicas no local operado protocolo de Meganlgesia eficaz CONV4 Boas sat O2 com C/N 2 l/mim Apresenta edemas dos MS 's e MI 's .</p>
<p>- vesical urina carregada Soroterapia em cvp no dorso da mão esqd^a .</p>

Anexo 4 – Anotações dos conceitos mais importantes

About 1675 results (page 11 of 335)

Sort by

- Stephanie Gould
Curtisósticos ; Neoplasia Maligna do Fígado ...
- Enferm.-Anestesia , << # # date:14009 # # >> , BIOMETRIAS ...
- Sinusite
- Shaun Castro , Anestesiologia Geral - - - AVALIAÇÃO PRÉ ANES...
- (HS Sebastião) 2 - 2 cesare Benjamins há 7 e 3 anos respec...

Chemicals & Drugs c Disorders d Procedures S-p

Shaun Castro , Anestesiologia Geral - - -

AVALIAÇÃO PRÉ ANESTÉSICA x IDADE:38

PESO x : 57 Kg Curtis ÓSTICO : Adenoma x

hepatico OPERAÇÃO x PROPOSTA :

hepatectomia x Lateral esquerda PATOLOGIA x

ASSOCIADA : Respiratório : Sinusite x

Gastrointestinal : Gastrite x Metabólico :

54 / 1675

Anexo 5 – Documento base para realização da anotação manual por parte de cada anotador especialista

Disorders

- Anormalidade Adquirida
- Anormalidade Anatômica
- Disfunção celular ou molecular
- Anormalidade congênita
- Doença ou Síndrome
- Modelo Experimental de Doença
- Encontrando
- Lesão ou envenenamento
- Disfunção mental ou comportamental
- Processo Neoplásico
- Função Patológica
- Sinal ou sintoma

Procedures

- Procedimento de diagnóstico
- Atividade Educativa
- Atividade de Saúde
- Procedimento Laboratorial
- Técnica de Pesquisa em Biologia Molecular
- Atividade de Pesquisa
- Procedimento terapêutico ou preventivo

Chemicals & Drugs

- Aminoácido, peptídeo ou proteína
- Antibiótico
- Substância biologicamente ativa
- Material biomédico ou dentário
- Químico
- Produto químico visto funcionalmente e Produto químico visto estruturalmente
- Medicamento Clínico
- Elemento, Íon ou Isótopo
- Enzima
- Substância perigosa ou venenosa
- Hormônio
- Fator imunológico
- Indicador, reagente ou auxílio diagnóstico
- Produto químico inorgânico
- Ácido Nucleico, Nucleósido ou Nucleotídeo
- Produto químico orgânico
- Substância Farmacológica
- Recetor
- Vitamina

Anexo 6 – Análise de registos clínicos

>

☰

Mine4Health IPO: Componente NLP

Componente de processamento de linguagem natural aplicado à registros clínicos.

Essa aplicação tem como objetivo demonstrar o funcionamento do componente de processamento de linguagem que vai analisar os registros clínicos dos pacientes.

Análise de registo clínico

Exemplos

Exemplo 1

Insira aqui o texto do registo clínico

Peso: 68 kg
Doente de 64 anos
Motivo de referenciação: Neoplasia do sigmoide
HDA: Doente realiza tds os anos PSOF, que foi positiva. O que motivou a realização de colonoscopia. Doente assintomática, sem perdas hemáticas, sem dor associada. sem alterações do TGI, sem emagrecimento.
Antecedentes pessoais: Alterações do metabolismo lipídico; Perturbações depressivas; Hipertensão não complicada
Antecedentes cirúrgicos: Pulso direito, varizes, raspagem uterina
Medicação crónica: Losartan
Alergias: Nega
História familiar: Doenças antecedentes: nenhuma

Analisar

	Conceito	Categoria	Presente
0	Peso	Procedures	Sim
1	Motivo	Physiology	Sim
2	Neoplasia	Disorders	Sim
3	sigmoide	Anatomy	Sim
4	realização	Disorders	Sim
5	colonoscopia	Procedures	Sim
6	assintomática	Disorders	Sim
7	emagrecimento	Disorders	Não
8	Hipertensão	Disorders	Sim
9	Pulso	Procedures	Sim
10	varizes	Disorders	Sim
11	uterina	Disorders	Sim
12	Medicação	Procedures	Sim
13	Losartan	Disorders	Sim
14	Alergias	Disorders	Sim
15	Peso	Procedures	Sim
16	adenopatias	Disorders	Não
17	Exames complementares	Procedures	Sim
18	Colonoscopia	Procedures	Sim
19	Colonoscopia	Procedures	Sim
20	preparação	Disorders	Sim
21	intestinal	Disorders	Sim
22	totalidade	Disorders	Sim
23	particular	Disorders	Sim
24	cólicos	Procedures	Sim
25	neoplasia	Disorders	Sim
26	tatuagem	Procedures	Sim
27	padrão vascular	Disorders	Sim
28	Histologia	Procedures	Sim
29	Adenocarcinoma	Disorders	Sim
30	sigmoide	Anatomy	Sim
31	Diverticulose	Disorders	Sim
32	estenose	Disorders	Sim
33	estenose	Disorders	Sim
34	Consulta	Procedures	Sim
35	cirurgia	Procedures	Sim

Anexo 7 - Anotação de frases segundo o SOAP

The screenshot shows the doccano interface with a search bar and a list of results. The selected result is 'Avaliação' with a count of 1. The main content area displays the text: 'Avaliação comparativa com estudo prévio realizado a < ## date:14212 ## > , registando se aumento do número e dimensões das metástases hepáticas , de distribuição bilobar , a maior no segmento IV com 3 cm e segmento VII com 4,2 cm .'. The interface includes a search bar, a 'Sort by' dropdown, and a list of filters on the left side.

doccano

Search document

About 205 results (page 37 of 41)

Sort by

- Paredes ventriculares com espessura e ecogenicidade normais ...
- Avaliação comparativa com estudo prévio realizado a < ## ...
- Plano : Convoco a doente dentro 1 semana
- P / tem consulta em Março , mantém plano
- Plano Operatorio para Onc

205/205

Avaliação 1 Objetivo o Plano p Subjetivo s

Avaliação x

Avaliação comparativa com estudo prévio realizado a < ## date:14212 ## > , registando se aumento do número e dimensões das metástases hepáticas , de distribuição bilobar , a maior no segmento IV com 3 cm e segmento VII com 4,2 cm .

The screenshot shows the doccano interface with a search bar and a list of results. The selected result is 'Objetivo' with a count of 1. The main content area displays the text: 'TAC - Foi realizado estudo TC do abdómen superior aps administração de produto de contraste iodado endovenoso .'. The interface includes a search bar, a 'Sort by' dropdown, and a list of filters on the left side.

doccano

Search document

About 205 results (page 1 of 41)

Sort by

- Sem queixas de foro digestivo
- Mantém plano de cuidados
- Nega dor abdominal , febre , anorexia , emagrecimento , perd...
- ECOG 0 coradas , hidratadase anictérica
- TAC - Foi realizado estudo TC do abdómen superior aps admini...

205/205

Avaliação 1 Objetivo o Plano p Subjetivo s

Objetivo x

TAC - Foi realizado estudo TC do abdómen superior aps administração de produto de contraste iodado endovenoso .

5 / 205

doccano Edit Data

Search document

About 205 results (page 1 of 41)

Sort by

- Sem queixas de foro digestivo
- Mantém plano de cuidados
- Nega dor abdominal , febre , anorexia , emagrecimento , perd...
- ECOG 0 coradas , hidratadase anictérica
- TAC - Foi realizado estudo TC do abdómen superior aps admini...

205/205

Avaliação 1 Objetivo o Plano p **Subjetivo s**

Subjetivo x

Nega dor abdominal , febre , anorexia , emagrecimento , perdas hemáticas evidentes e tosse seca .

3 / 205

doccano Edit Data

Search document

About 205 results (page 1 of 41)

Sort by

- Sem queixas de foro digestivo
- Mantém plano de cuidados
- Nega dor abdominal , febre , anorexia , emagrecimento , perd...
- ECOG 0 coradas , hidratadase anictérica
- TAC - Foi realizado estudo TC do abdómen superior aps admini...

205/205

Avaliação 1 Objetivo o **Plano p** Subjetivo s

Plano x

Mantém plano de cuidados .

2 / 205

Anexo 8 – Orientações para a anotação SOAP

Componente subjetivo (S)

- Queixa principal do paciente
- Declaração muito breve do paciente
- Histórico da doença atual
- Condição atual do paciente em forma narrativa
- História e o estado dos sintomas experimentados
- Histórico médico pertinente, histórico cirúrgico
- Histórico familiar e histórico social
- Medicamentos atuais (nome, dose, rota e com que frequência) e alergias

Componente objetivo (O)

- Inclui informações que o profissional de saúde observa ou mede a partir da apresentação atual do paciente, como: Os sinais vitais e medidas, como peso.
- Resultados de exames físicos, incluindo sistemas básicos cardíacos e respiratórios, os sistemas afetados, possível envolvimento de outros sistemas, achados normais pertinentes e anormalidades. As seguintes áreas devem ser incluídas:
 - Apresentação física
 - Caracterização de desconforto ou dor
 - Status psicológico
- Resultados de laboratório e outros testes de diagnóstico já concluídos.

Componente de avaliação (A)

- Diagnóstico médico para a finalidade da visita médica
- Resumo rápido do paciente com os principais sintomas / diagnóstico
- Incluirá possíveis e prováveis etiologias do problema do paciente
- É o progresso do paciente desde a última visita e o progresso geral em direção à meta do paciente da perspectiva do médico.
- Incluirá etiologia e fatores de risco, avaliações da necessidade de terapia, terapia atual e opções de terapia.

Componente plano (P)

- É o que o profissional de saúde fará para tratar as preocupações do paciente, como solicitar novos laboratórios, exames radiológicos, encaminhamentos, procedimentos realizados, medicamentos e educação fornecidos.
- Incluirá objetivos de terapia e parâmetros específicos de monitoramento de medicamentos e doenças
- O plano em si inclui vários componentes:
 - Componente de diagnóstico: continue a monitorar laboratórios
 - Componente terapêutico: dieta avançada
 - Encaminhamentos: acompanhamento com cardiologia dentro de três dias após a alta para teste de estresse como paciente externo.
 - Componente de educação do paciente: que está progredindo bem
 - Componente de disposição: alta para casa pela manhã

EXEMPLO:

Serviço de Cirurgia, Dr. Jones

S: Não há mais dor no peito ou falta de ar. "Sentindo-se melhor hoje." Paciente relata dor de cabeça.

O: Afebril, P 84, R 16, BP 130/82. Sem sofrimento agudo.

Pescoço sem JVD, pulmões limpos

Cor RRR

Abd Bowel parece presente, leve sensibilidade ao RLQ, menos do que ontem. As feridas parecem limpas.

Ext sem edema

A: Paciente masculino, 37 anos, no dia 2 do pós-operatório de apendicectomia laparoscópica. Recuperando bem.

P: Dieta avançada. Continue a monitorar laboratórios. Acompanhamento com Cardiologia dentro de três dias após a alta para teste de estresse como paciente externo. Prepare-se para a alta em casa amanhã de manhã.

Anexo 9 - Resultados da avaliação realizada pelo especialista C a algumas das frases (83) que tiveram a mesma anotação por parte dos anotadores A e B

Frases	Anotação dos anotadores A e B	Validação do especialista C (VALIDOU/ NÃO VALIDOU/ ADICIONOU TERMOS)
Realizada endoscopia digestiva alta , sem complicações imediatas (ver relatório no Siima) .	endoscopia digestiva alta (Procedures)	Validou.
Protocolo analgesia : avaliar às 14h45 .	Protocolo analgesia (Procedures)	Validou.
Volta em março de 2015 com TAC , pois passa a vigilância .	TAC (Procedures)	Adicionou 1 termo: 'Vigilância' (procedures)
Entre D2-4 - alguma astenia mas a melhorar .	astenia (Disorders)	Validou.
Sai sem queixas , volta a 01/12 com analises bloco2 .	analises (Procedures)	Validou.
Fez RX do tórax foi observado pelo médico .	RX do tórax (Procedures)	Validou.
Mantem queixas de enfartamento e regurgitações .	enfartamento (Disorders), regurgitações (Disorders)	Validou.
Protocolo analgesia : avaliar às 20:45H .	Protocolo analgesia (Procedures)	Validou.
Fez descida para o bloco Cefazolina protocolo .	Cefazolina (Drugs)	Validou.
Diagnóstico : Adenocarcinoma do reto , encontra.-se horas após RAR .	Adenocarcinoma do reto (Disorders), RAR (Procedures)	Adicionou 1 termo: 'Diagnóstico' (procedures)
Proposta para realizar amanhã RAR laparoscópica .	RAR laparoscópica (Procedures)	Validou.
Início suplementação com ácido fólico .	ácido fólico (Drugs)	Adicionou 1 termo: 'suplementação' (Procedures)
Iniciou preparação cólica cerca das 16h que foi eficaz .	preparação cólica (Procedures)	Validou.
Início : 10 Out 2018 Perturbações depressivas ;	Perturbações depressivas (Disorders)	Validou.
Documentado défice de ác . fólico , iniciou suplementação .	ác . fólico (Drugs)	Adicionou 2 termos: 'Suplementação' (Procedures), 'défice' (Disorders)
Doente refere apenas `` secura da boca `` aconselhada reforço hidrico que a doente refere que já está a cumprir .	secura da boca (Disorders)	Adicionou 1 termo: 'reforço hidrico' (Procedures)

Frases	Anotação dos anotadores A e B	Validação do especialista C (VALIDOU/ NÃO VALIDOU/ ADICIONOU TERMOS)
Divertículos em todos os segmentos cólicos .	Divertículos (Disorders)	Validou.
Colhido hemograma + bioquímica .	hemograma (Procedures), bioquímica (Procedures)	Validou.
Artroscopia convertida para < IDADE > anos .	Artroscopia (Procedures)	Validou.
Protocolo de analgesia SUF 2 eficaz .	Protocolo de analgesia (Procedures)	Adicionou 1 termo: 'SUF 2' (Drugs)
Victan em SOS Sonolento mas facilmente despertável .	Victan (Drugs)	Adicionou 1 termo: 'Sonolento' (Disorders)
Reiniciado sufentanil a 1ml/h .	sufentanil (Drugs)	Validou.
Observam se imagens de invasão linfovascular .	invasão linfovascular (Disorders)	Validou.
Após reavaliação por oncologia médica decidido internamento por alterações hepáticas de novo .	alterações hepáticas (Disorders)	Validou.
Macrogol 1 a 2 saquetas por dia (se tolerar) .	Macrogol (Drugs)	Validou.
Discreta ectasia do wirsung a a nível corpo sensivelmente sobreponível .	ectasia do wirsung (Disorders)	Validou.
Leva bomba perfusora de 5fu para ser removida às 46h no c . saúde .	5fu (Drugs)	Validou.
Ictérica Razoavelmente hidratada TA 125/72 mm Hg ;	Ictérica (Disorders)	Validou.
Vigilância de Taur - às 3:30h por Taur de 38.4°C foi administrado paracetamol .	paracetamol (Drugs)	Adicionou 1 termo: 'Vigilância' (Procedures)
Espessamento das paredes do antro gástrico , com uma massa exofítica relação com neoplasia gástrica , actualmente atinge 36 mm de diâmetro , francamente menor em confronto com exame anterior .	neoplasia gástrica (Disorders)	Validou.
Leva bomba de 5 FU que remove no Centro de Saúde .	5 FU (Drugs)	Adicionou 1 termo: 'remove' (Procedures)
Tem pedido de rx contrastado para amanhã , e só após observação deste terá indicação de iniciar dieta segundo nutrição .	rx contrastado (Procedures)	Adicionou 1 termo: 'iniciar dieta' (Procedures)
Estabilidade aparente das restantes Fígado de contornos regulares , morfologia preservada com dimensões dentro da normalidade , identificando se algumas imagens hipodensas sugestivas de quistos hepatobiliares , já presentes em estudos anteriores .	quistos hepatobiliares (Disorders)	Validou.

Frases	Anotação dos anotadores A e B	Validação do especialista C (VALIDOU/ NÃO VALIDOU/ ADICIONOU TERMOS)
Sem outras queixas O : ECOG2 COC Melhoria sustentada da Ictérica AC P sem alterações relevantes ABD inocente	Ictérica (Disorders)	Adicionou 1 termo: 'ECOG2' (Disorders)
Colheu sangue para Hemograma .	Hemograma (Procedures)	Validou.
Refere há 1 semana episódio isolado de febre , sem outros sintomas .	febre (Disorders)	Validou.
Com o diagnóstico de Ca gástrico metastizado .	Ca gástrico metastizado (Disorders)	Validou.
Fez RX do tórax foi observado pelo médico .	RX do tórax (Procedures)	Validou.
Refere dor generalizada na data de internamento hoje também melhorada .	dor generalizada (Disorders)	Validou.
Diagnóstico : Adenocarcinoma Gastrico mtx , proposta para COLPALO2 (FOLFIRI / substituído o esquema) .	Adenocarcinoma Gastrico (Disorders), FOLFIRI (Drugs)	Validou.
Crescimento claro da lesão gástrica .	lesão gástrica (Disorders)	Validou.
Internada no HDia onde vinha iniciar FOLFIRI que não realizou AP : Hm hoje Inicia água e chá , a evoluir ao almoço .	FOLFIRI (Drugs)	Adicionou 1 termo: 'Internada' (Procedures)
Plano : Peço hemograma para hoje .	hemograma (Procedures)	Validou.
Ictérica de instalação progressiva .	Ictérica (Disorders)	Validou.
Colheu sangue para hemograma .	hemograma (Procedures)	Validou.
Foi efectur RX torax e após observação teve autorização médica para prosseguir tratamento com FOLFIRI que decorreu sem incidentes .	RX torax (Procedures), FOLFIRI (Drugs)	Adicionou 1 termo: 'tratamento' (Procedures)
Colhe estudo analítico a 30/12 e mantém dieta instituída .	estudo analítico (Procedures)	Adicionou 1 termo: 'mantém dieta' (Procedures)
Colheu sangue para Hemograma e Bioquímica .	Hemograma (Procedures), Bioquímica (Procedures)	Validou.
Internada no HDia onde vinha iniciar FOLFIRI que não realizou AP : Hm e BQ hoje Evolui para dieta cremosa na tentativa de aumentar eficácia gravítica Hm amanhã .	FOLFIRI (Drugs)	Adicionou 2 termos: 'Internada' (Procedures), 'dieta cremosa' (Procedures)
Colheu hemoculturas pelo implantofix e periférico , bem como provas .	hemoculturas (Procedures)	Adicionou 1 termo: 'provas' (Procedures)
Suspeita de quadro obstrutivo alto .	quadro obstrutivo alto (Disorders)	Validou.

Frases	Anotação dos anotadores A e B	Validação do especialista C (VALIDOU/ NÃO VALIDOU/ ADICIONOU TERMOS)
Mantém pausa alimentar até nova avaliação médica pelo que não cumpriu laxantes PO .	laxantes (Drugs)	Adicionou 1 termo: 'pausa alimentar' (Procedures)
Empastamento epigástrico , sem defesa ou SIP .	Empastamento epigástrico (Disorders)	Adicionou 2 termos: 'sem defesa' (Disorders), 'SIP' (Disorders)
Refere ainda que mantém edema dos pés , mau estar generalizado .	edema dos pés (Disorders), mau estar generalizado (Disorders)	Validou.
Efectuada colheita de sangue para hemograma e bioquímica e enviado para o laboratório .	hemograma (Procedures), bioquímica (Procedures)	Validou.
Penso no local de colocação de endoprótese limpo e seco externamente , a reavaliar a 26/12 .	colocação de endoprótese (Procedures)	Adicionou 2 termos: 'penso' (Procedures), 'reavaliar' (Procedures)
Protocolo de analgesia eficaz .	Protocolo de analgesia (Procedures)	Validou.
12:45 Horas por dor generalizada , mas mais acentuada no flanco á direita fez morfina 2,5mg SC , que foi eficaz .	dor generalizada (Disorders), morfina (Drugs)	Validou.
Efectuou Rx contrastado e aguarda decisão para início de dieta .	Rx contrastado (Procedures)	Adicionou 1 termo: 'dieta' (Procedures)
Contacta intimamente o lobo esquerdo , o pâncreas , a vesicula biliar e o colon < IDADE > anos adiposos envolventes , observando se também adenopatias peri gástricas , no hilo hepático e no tronco celíaco .	adenopatias peri gástricas (Disorders)	Validou.
Doente com diagnóstico de Ca gástrico metastizado .	Ca gástrico metastizado (Disorders)	Validou.
Antenia marcada e sensação de desmaio .	sensação de desmaio (Disorders)	Validou.
Efetua amanhã 4ª de 5 fracções de RT Hemostatica .	RT Hemostatica (Procedures)	Validou.
Colhido sangue para hemograma as 7h .	hemograma (Procedures)	Validou.
Por hipocaliémia iniciou reposição .	hipocaliémia (Disorders)	Validou.
Desconforto abdominal agora melhroado .	Desconforto abdominal (Disorders)	Validou.
Hipotensao 95/63 mmhg p=96 sat O2 96 % .	Hipotensao (Disorders)	Validou.

Frases	Anotação dos anotadores A e B	Validação do especialista C (VALIDOU/ NÃO VALIDOU/ ADICIONOU TERMOS)
16h35 por dor abdominal fez SOS Morfina , 5mg , SC , que surtiu efeito .	dor abdominal (Disorders), Morfina (Drugs)	Validou.
Abdómen com massa epigástrica .	massa epigástrica (Disorders)	Validou.
Realizado penso no local de colocação de endoprótese limpo e seco externamente .	colocação de endoprótese (Procedures)	Validou.
Aguarda colocação de prótese .	colocação de prótese (Procedures)	Validou.
Consciente , orientada , debilitada mas colaborante .	debilitada (Disorders)	Adicionou 2 termos: 'Consciente' (Disorders), 'orientada' (Disorders)
Explicado sinais de alerta , por lesão estenosante .	lesão estenosante (Disorders)	Validou.
Colhido sangue para Hemograma e Bioquímica .	Hemograma (Procedures), Bioquímica (Procedures)	Validou.
Apresenta : Sinais vitais estáveis , eupneico ao ar ambiente Pensos abdominais limpos e secos externamente Drenagem vesical que se faz de urina clara Protocolo de analgesia .	Protocolo de analgesia (Procedures)	Adicionou 1 termo: 'Drenagem vesical' (Procedures)
Leva DIB de 5 FU para 46 h que remove no Hospital Boone PLC Chang , Fisher and Green , porque vem a uma consulta .	5 FU (Drugs)	Validou.
Foi realizada análise imunohistoquímica de MLH1/MSH2/MSH6/PMS2 , tendo sido normal .	análise imunohistoquímica (Procedures)	Validou.
Pedida nova marcação de HDia , com análises bloco 1 , para 25.9 mas apenas entrou para 27.9.19	análises (Procedures)	Validou.
Adenocarcinoma do colon sigmoide .	Adenocarcinoma do colon sigmoide (Disorders)	Validou.
Leva DIB de 5 FU para 46 h que remove no Centro de Saúde .	5 FU (Drugs)	Adicionou 1 termo: 'DIB' (procedures)
Fez RX do tórax foi observado pelo médico .	RX do tórax (Procedures)	Validou.
Fica 1 CD de TAC para digitalizar .	TAC (Procedures)	Validou.
Pedida nova marcação de HDia , com análises bloco 1 , para 24.7.19 (Tem COM a 26.7 , não quis marcação de QT para o mesmo dia para a retirada de DIB não ser ao fim de semana) .	análises (Procedures), QT (Drugs)	Validou.

Anexo 10 - Resultados da avaliação do modelo semi-supervisionado com pré-anotação do UMLS

Frases	Entidades detetadas	Avaliação do especialista C sobre a anotação: SIM/NÃO/PARCIAL
Colhido hemograma + bioquímica .		NÃO. Deveria ter detetado 'bioquímica'
Sem sinais de alterações de mobilidade e/ou sensibilidade dos membros inferiores .	sensibilidade (Disorders)	PARCIAL. Deveria ter classificado 'mobilidade' (Disorders)
Peso : 58 kg (perda de 2 Kg dese a última consulta) Orofaringe sem lesões , sem exsudados .	Peso (Procedures)	PARCIAL. Deveria ter classificado 'orofaringe sem lesões' (disorders)
Circulante Ryan Amanda Geada de Michael Trigo Lourenco (1214) - Enf .	Circulante (Drugs)	NÃO. 'Circulante' é uma profissão
Faz tratamento com suporte de G CSF .	tratamento (Procedures)	PARCIAL. Deveria ter classificado 'G CSF' (drugs)
Antecedentes : Prótese Total Anca direita ;	Anca (Drugs)	NÃO. Falta 'Próteses' (drugs)
Peço análises para essa consulta .		NÃO.
Administrada Enoxaparina 40mg SC às 21h , não tendo sido possível validar por alteração de horário .	Enoxaparina (Drugs), tendo (Disorders)	PARCIAL. Não deveria ter classificado 'tendo' (Disorders).
Frequência : 8/8 h Descrição : AVALIAÇÃO DE SINAIS VITAIS Frequência : Sem Frequência Descrição : MEIAS ELÁSTICAS DE COMPRESSÃO GRADUAL (MECG) Frequência : Sem Frequência Descrição : BALANÇO HÍDRICO Frequência : Sem Frequência Descrição : O2 POR CATETER NASAL 3 L/MIN Frequência : Sem Frequência Descrição : 1º LEVANTE Frequência : Sem Frequência Descrição : REMOVER ALGÁLIA Frequência : Sem Frequência Descrição : RETIRAR DRENO Frequência : Sem Frequência Descrição : RETIRAR CATÉTER VENOSO CENTRAL Frequência : Sem Frequência	AVALIAÇÃO DE SINAIS VITAIS (Procedures), Descrição (Procedures)	NÃO. Deveria ter classificados muitos itens.
Mantém : - Soro composto por CVC para substituir às 10h .	composto (Drugs)	NÃO. Deveria ter classificado 'soro' (Drugs)
ECG : RS FC 95 bpm , sem alterações BNM : sem alterações Combur negativo Rx Torax : sem infiltrados ou condensações .	ECG (Procedures), negativo (Disorders)	PARCIAL. Não deveria ter classificado 'negativo' (Disorders) e deveria ter classificado mais itens.

Frases	Entidades detetadas	Avaliação do especialista C sobre a anotação: SIM/NÃO/PARCIAL
Por referir abdominalgia e dor generalizada , fez terapêutica de resgate às 9:15h e às 13:45h , que surtiram efeito reduzido .	reduzido (Disorders), terapêutica (Procedures)	PARCIAL. Não deveria ter classificado 'reduzido' (Disorders) e deveria ter classificado mais itens.
Colhe Hemograma + Bioquímica + estudo da coagulação .	Hemograma (Procedures), Bioquímica + (Procedures), coagulação (Procedures)	SIM.
Ausentou se para realização de TAC .	TAC (Procedures), realização (Disorders)	PARCIAL. Não deveria ter classificado 'realização' (Disorders).
Avaliada por gastro , segundo mural M , não é previsível qualquer benefício da realização de EDA , pelo que cancela a mesma .	gastro (Disorders), realização (Disorders)	NÃO.
Observa se marcado espessamento parietal da região do antro gástrico com formação vegetante da parede anterior com 4.7cm , associando se adenopatias / implantes satélites , a maior anteriormente aos grandes vasos abdominais , com 2.6cm . Associa se pequena quantidade de líquido livre intraperitoneal , em particular no fundo do saco posterior pélvico , achados a correlacionar com dados endoscópicos e a melhor caracterizar por TC abdominopélvico .	livre intraperitoneal (Disorders), anterior (Disorders), adenopatias (Disorders), implantes (Procedures), gástrico (Procedures), formação (Disorders)	NÃO.
Mantém : - CVC com soro em perfusão ;	perfusão (Procedures)	PARCIAL. Deveria ter classificado 'soro' (Drugs)
Admite se ingrowth tumoral , a correlacionar com TC e endoscopia .	e endoscopia (Procedures)	PARCIAL. Deveria ter classificado 'TC' (Procedures)
Altero medicação : Hidromorfona 8mg - 1 cp ao almoço Esomeprazol 40 mg - 1 cp jejum e antes jantar Prednisolona 20 mg - 1 cp ao PA Metoclopramida 10 mg - 1 cp antes refeições Macrogol 10000mg - 1 saqueta ao almoço Ácido Aminocapróico 3 g - 1 saqueta 8/8hh SOS Sevredol 10 mg - meio cp em SOS , se dor , máx 1/1 h (até 5/dia) Paracetamol 1000mg - SOS , se febre , máx 8/8h Macrogol 10000mg - SOS , se obstipação , máx 3/dia Ondansetrom 4 mg - SOS , se náuseas ou vômitos , máx 12/12h Lorazepam 1 mg - meio cp em SOS , se insónia , até 2 tomas/dia Reavaliação telefónica na próxima 6ª feira .	Metoclopramida (Drugs), Hidromorfona (Drugs), Prednisolona (Drugs), Esomeprazol (Drugs), Paracetamol (Drugs), Lorazepam (Drugs), jejum (Disorders), Ondansetrom (Drugs), antes (Drugs), antes (Drugs), Ácido Aminocapróico (Drugs), obstipação (Disorders), medicação (Procedures)	PARCIAL. Não deveria ter classificado 'antes' (Drugs).
Tem prescrição de Ondansetrom em SOS prévio .	prescrição (Procedures), SOS prévio (Disorders)	PARCIAL. Deveria ter classificado 'ondansetrom' (Drugs)

Frases	Entidades detetadas	Avaliação do especialista C sobre a anotação: SIM/NÃO/PARCIAL
Segundo nota de alta : `` Tem alta para prosseguir plano terapêutico de quimioterapia ". hemograma - hb de hoje : 11 P) pode fazer tratamento programado .	quimioterapia (Procedures), tratamento (Procedures), hemograma (Procedures), terapêutico (Procedures)	PARCIAL. Deveria ter classificado 'quimioterapia' (drugs)
Hoje por hemoglobina 6,1 transfunde 2UCE , foi pedido cadeirão e marcado tipagem central colheitas .	hemoglobina (Drugs)	NÃO.
Frequência : 8/8 h Nome Genérico : Glucose 55 mg/ml + Cloreto de sódio 9 mg/ml Sol inj Fr 500 ml IV Via Administração : ENDOVENOSO ;	Glucose (Drugs), Cloreto (Procedures), Administração (Procedures), Genérico (Drugs), sódio (Drugs)	NÃO.
Revejo medicação : fica com sucralfato , epsicaprom , e em pausa alimentar .	sucralfato (Drugs), pausa alimentar (Disorders), medicação (Procedures)	PARCIAL. Deveria ter classificado 'epsicaprom' (Drugs) e 'pausa alimentar' (Procedures)
Antropometria : P atual : 52kg P hab : 62 Kg (Abr2018) E : 1.53 m IMC : Magreza Ingestão alimentar hab : Via : oral Consistência : geral Levanta se pelas 7h30 PA (7:30h) - 1 chav de leite MG c / cevada s/açúcar c / Nestum ou 3-4 bolachas Cynthia MM (10h30) - 1 iog sól ou banana A (12:30h) - Prato : (50 % hab , melhor tolerância para o peixe) MT (16h30) - = PA 18h - canja ou sopa legumes s/HC , c/leguminosas , as x c / FP .	Magreza (Disorders), leite (Disorders), Antropometria (Procedures), açúcar (Drugs)	NÃO.
Nos rins refere se um quisto denso de 11 mm no terço médio do rim esquerdo , e um quisto cortical simples de 2 , Scm no polo inferior do mesmo rim .	e um quisto cortical (Disorders), inferior (Disorders), terço (Disorders), médio (Drugs)	PARCIAL. Não deveria ter classificado 'médio' (Drugs).

Anexo 11 - Resultados do comparativo de frases em que ambos os anotadores A e B realizaram anotação

Frases	Modelo anotador A	Modelo anotador B	Informação do especialista C sobre a anotação mais completa: A/B/OUTRA
Pedido de TAC Toracico de estadiamento e análises .	TAC Torácico (Procedures), análises (Procedures)	TAC (Procedures), análises (Procedures)	A está correta.
PCEA a 5ml/h que perfunde sem intercorrências .	PCEA (Drugs)	PCEA (Procedures)	A e B corretas. A frase não deveria ter a medicação como PCEA
ECOG o Pele e mucosas coradas e hidratadas Altura - 152 cm Peso 64 kg SC - 1.61 m2 SPP resolvido Análises de hoje : neutropenia grau 2 (1.290/u L) .	Análises (Procedures), neutropenia grau 2 (Disorders)	Análises (Procedures), neutropenia (Disorders)	A mais precisa. Falta classificação 'ECOG' (Disorders), 'Pele e mucosas coradas e hidratadas' (Disorders)
ECOG o Pele e mucosas coradas e hidratadas Altura - 152 cm Peso 69 kg (aumento de 5 kg) ACP - sem alterações de relevo Abdomen - sem alterações de relevo Sem edemas periféricos Análises de hoje : anemia resolvida ; neutropenia em recuperação , neste momento grau 1 .	Análises (Procedures), anemia (Disorders), neutropenia (Disorders)	Análises (Procedures), anemia (Disorders), edemas periféricos (Disorders), neutropenia (Disorders)	B. Falta classificação 'ECOG' (Disorders), 'Pele e mucosas coradas e hidratadas' (Disorders)
ECOG o Pele e mucosas coradas e hidratadas Altura - 152 cm Peso 64 kg SC - 1.61 m2 Sem alterações de relevo de novo E.O . Análises de hoje : mantém neutropenia grau 2 (1.370/u L) .	Análises (Procedures), neutropenia grau 2 (Disorders)	Análises (Procedures), neutropenia (Disorders)	A mais precisa. Falta classificação 'Pele e mucosas coradas e hidratadas' (Disorders)
Colocada bomba de 5FU.Saíu bem .	bomba de 5FU (Drugs)	5FU (Drugs)	B.
Doente calma e colaborante dentro das possibilidades , palidez pele e mucosas , deambulou pelo serviço .	palidez pele (Disorders)	palidez (Disorders)	B. Falta classificação 'calma e colaborante' (Disorders)
Sem alterações aparentes da sensibilidade e/ou mobilidade dos MI 's . ~ Em vigilância de TA , por hipotensão no recobro , que não apresentou durante o turno da tarde .	hipotensão (Disorders), vigilância de TA (Procedures)	hipotensão (Disorders), vigilância (Procedures)	A mais precisa. Falta 'sensibilidade' (Disorders)e 'mobilidade' (Disorders)

Frases	Modelo anotador A	Modelo anotador B	Informação do especialista C sobre a anotação mais completa: A/B/OUTRA
Sem alterações aparentes da sensibilidade e/ou mobilidade dos MI 's . ~ Em vigilância de TA , por hipotensão no recobro .	vigilância de TA (Procedures), hipotensão (Disorders)	vigilância (Procedures), hipotensão (Disorders)	A mais precisa. Falta 'sensibilidade' (Disorders)e 'mobilidade' (Disorders)
Foi feito esclarecimento de dúvidas alimentares .	esclarecimento de dúvidas alimentares (Procedures)	esclarecimento de dúvidas (Procedures)	A mais precisa.
Remove bomba infusora no domicilio (nora Enfermeira) e forneceu carta+Fibrilin .	Remove bomba infusora (Procedures), Fibrilin (Drugs)	Fibrilin (Drugs)	A.
Colocada bomba de 5FU.Saíu bem .	bomba de 5FU (Drugs)	5FU (Drugs)	B. Falta 'Colocar bomba' (Procedures)
Administrado ondasetron 4mg EV pelas 12h por indicação da Dra Dominique Dalton , que foi eficaz .	ondasetron 4mg (Drugs)	ondasetron (Drugs)	B.
Dor controlada com analgesia epidural .	analgesia epidural (Drugs)	analgesia epidural (Procedures)	B.
Tem previsto QT adjuvante p / breve .	QT adjuvante (Drugs)	QT (Drugs)	A mais precisa.
Avaliação de toxicidade apos 3 ciclos de QT com CF - Adiou 2 ciclos por neutropenia G2 .	Avaliação de toxicidade (Procedures), QT (Drugs), CF (Drugs), neutropenia G2 (Disorders)	Avaliação de toxicidade (Procedures), CF (Drugs), neutropenia (Disorders), QT (Drugs)	A mais precisa.
Incentivada a executar exercicios respiratorios , tosse eficaz , requer incentivo .	executar exercicios respiratorios (Procedures)	exercicios respiratorios (Procedures)	B.
Medicação Habitual : Losartan/HCTZ 100/25mg 1/2 comp de manhã Alergias : desconhecidas Boa tolerância clinica a 1º ciclo de capecitabina .	Losartan/HCTZ 100/25mg (Drugs), capecitabina (Drugs)	capecitabina (Drugs), Losartan (Drugs), HCTZ (Drugs)	B.

Frases	Modelo anotador A	Modelo anotador B	Informação do especialista C sobre a anotação mais completa: A/B/OUTRA
Encontra se no 6º dia após RAR com colostomia derivativa do cólon transversos .	RAR (Procedures), colostomia derivativa (Procedures)	RAR (Procedures), colostomia (Procedures)	B.
Feitos ensinamentos sobre a dieta pobre em resíduos com início a 12/03 .	ensinamentos sobre a dieta (Procedures)	ensinamentos (Procedures)	A mais precisa.
Frequência : 1 id Nome Genérico : Paracetamol 500 mg Comp Via Administração : ORAL ;	Paracetamol 500 mg (Drugs)	Paracetamol (Drugs)	B.
Medicação Habitual : Losartan/HCTZ 100/25mg 1/2 comp de manhã Alergias : desconhecidas Boa recuperação de pós operatório .	Losartan/HCTZ 100/25mg (Drugs)	Losartan (Drugs), HCTZ (Drugs)	B.
Administrada Enoxaparina 40mg SC às 21h , não tendo sido possível validar por alteração de horário .	Enoxaparina 40mg (Drugs)	Enoxaparina (Drugs)	B.
TVP do MID (eixos popliteu femoral e tronco tibio peroneal direitos) , diagnóstico em fev 2018 ;	TVP do MID (Disorders)	TVP (Disorders)	A mais precisa.
19h T (aur) = 37.9°C . 19h30 T (aur) = 38.1°C , contactado médico de permanência Dr . Benjamin que deu indicação para iniciar transfusão .	transfusão (Drugs)	transfusão (Procedures)	A.
condicionados pela prótese metálica gastroduodenal .	prótese metálica gastroduodenal (Procedures)	prótese metálica gastroduodenal (Disorders)	OUTRA: 'protese metálica' (Drugs)
Está sob RT hemostática - amanhã tem sessão .	RT hemostática (Procedures)	RT hemostática (Drugs)	B.
Doente a aguardar SCP , que vinha hoje para reavaliação de hemograma após transfusão e RT .	hemograma (Procedures), transfusão (Drugs), RT (Procedures)	hemograma (Procedures), transfusão (Procedures), RT (Drugs)	B. Exceto 'transfusão' (Drugs)
Protese gastrica agendada para 11h30 .	Protese gastrica (Procedures)	Protese gastrica (Disorders)	OUTRA: 'Protese gástrica' (Drugs)

Frases	Modelo anotador A	Modelo anotador B	Informação do especialista C sobre a anotação mais completa: A/B/OUTRA
Ausentou se para realizar 3 ^o fx (5fx) de RT hemostática .	RT hemostática (Procedures)	RT hemostática (Drugs)	B.
Explico situação atual e prognóstico reservado no caso de se manter hemorragia ativa .	hemorragia ativa (Disorders)	hemorragia (Disorders)	B.
Trânsito gastro duodenal , segundo registo clínico : ÷ Identifica se protese metálica projectada no antro gástrico/primeira porção do duodeno , sem evidente deslocamento em comparação com estudos prévios , a protese encontra se globalmente bem expandida .	protese metálica (Procedures)	protese metálica (Disorders)	OUTRA: 'prótese metálica' (Drugs)
CVC funcionante por onde efetuou HERCEPTIN EV sem intercorrências .	HERCEPTIN EV (Drugs)	HERCEPTIN (Drugs)	B.
Efetuo 1U de CE que perfundi sem reação anormal aparente - repuncionada para efetuar a transfusão por mau funcionamento do anterior CVP .	CE (Drugs)	transfusão (Procedures)	A.
Através de fio guia foi colocada endoprótese metálica Wallstent 94x10mm .	endoprótese metálica (Procedures)	endoprótese metálica (Disorders)	OUTRA: 'prótese metálica' (Drugs)
Tem pedido de H para Domingo e de H + hemostase para 2 ^a feira .	hemostase (Drugs), H (Drugs)	hemostase (Procedures)	OUTRA: 'hemostase' (Disorders)
Falo com doente e familiar , proponho internamento que aceitam , decido iniciar antibioterapia de largo espectro , por febre sem foco associada a cuidados de saude (inf do CVC ?) e vomitos nao controlados G3 .	antibioterapia (Drugs), vomitos nao controlados G3 (Disorders)	antibioterapia (Procedures), febre (Disorders), vomitos (Disorders)	A.
Marca se consulta de anestesia para 27/2/19 , entregue folheto informativo .	consulta de anestesia (Procedures)	anestesia (Procedures)	A.

Frases	Modelo anotador A	Modelo anotador B	Informação do especialista C sobre a anotação mais completa: A/B/OUTRA
Efetuado ensino sobre cuidados pré e pós operatórios , que compreendeu .	ensino sobre cuidados pré e pós operatórios (Procedures)	ensino (Procedures)	A mais precisa.
Com este resultado , num doente com critérios de Bethesda , não há indicação para pesquisa de mutações germinativas associadas com Síndrome de Lynch .	Síndrome de Lynch (Disorders), pesquisa de mutações germinativas (Procedures)	pesquisa de mutações germinativas (Procedures)	A.

Anexo 12 - Resultados do comparativo de frases em que apenas um dos anotadores realizou anotação

Frases	Modelo anotador A	Modelo anotador B	Informação do especialista C sobre a anotação mais completa: A/B/OUTRA
Dadas orientações sobre o procedimento e sua colaboração durante o mesmo .		Orientações (Procedures)	B.
Doente com clearance de creatinina de 86.89 mg/dl .	clearance de creatinina (Procedures)		A.
Estadio patológico p T1b N1 Mo Ro FU de 8 meses .	Estadio patológico p T1b N1 Mo Ro FU (Disorders)		A.
Peso : 58 kg (perda de 2 Kg dese a última consulta) Orofaringe sem lesões , sem exsudados .		lesões (Disorders)	B. Falta 'exsudados' (Disorders)
Sem icterícia Sem adenomegalias ABD : Mole e depressível .		adenomegalias (Disorders), icterícia (Disorders)	B.
Faz tratamento com suporte de G CSF .	G CSF (Drugs)		A.
Intestino funcionando em pequena quantidade de fezes duras segundo a doente nas 24h .	fezes duras (Disorders)		A.
Posologia : 1 comp 6/6h se dor ou febre ;		dor (Disorders), febre (Disorders)	B.
Assim , sem noção quanto a nodulos no apex do LDireito .		nodulos (Disorders)	B.
Volta em agosto com BQ e set . Peço protese capilar .	BQ (Procedures)		A.
Sem alterações neurovasculares aparentes das extremidades .		alterações neurovasculares (Disorders)	B.
Removida vareta conforme indicação médica .	Removida vareta (Procedures)		A.
Tem indicação de remover vareta a 22/03 .	remover vareta (Procedures)		A.
Peso atual - 47.5kg . Ajusto dose de quimioterapia ao peso atual .		quimioterapia (Drugs)	B.
Instruidos para os exames e consultas seguintes .		exames (Procedures)	B.

Frases	Modelo anotador A	Modelo anotador B	Informação do especialista C sobre a anotação mais completa: A/B/OUTRA
Por 7.2 g/dl de Hg , foi pedido 1UGR , que aguarda .	1UGR (Drugs)		A.
Mantem para já ATB em curso até resultado definitivo de HCs .	ATB (Drugs)		A.
Nega dispneia ou cansaço para pequenos esforços , tem tolerado caminhadas pelo quarto e serviço .		dispneia (Disorders), cansaço (Disorders)	B.
Deverá ir no final das transfusões para Radioterapia .		Radioterapia (Drugs)	B. Falta 'transfusões' (Drugs)
Fígado de dimensões normais , evidenciando algumas imagens hipodensas sugestivas de quistos .		quistos (Disorders)	B.
Hoje por hemoglobina 6,1 transfunde 2UCE , foi pedido cadeirão e marcado tipagem central colheitas .	2UCE (Drugs), tipagem (Procedures)		A.
Tem pedido de colaboração por nutrição em anexo .	pedido de colaboração por nutrição (Procedures)		A.

Anexo 13 – Exemplo de um *dashboard* para auxiliar o clínico na tomada de decisão

