

ProSmartHealth: Uma ferramenta de decisão

Inês Prudêncio Sena

Dissertação realizada pela orientação da

Prof^ª. Ana Isabel Pereira

Mestrado em Tecnologia Biomédica - Instrumentação e Sinais Médicos

Outubro 2020



ProSmartHealth: Uma ferramenta de decisão

Relatório Final da
Dissertação do Mestrado em Tecnologia Biomédica
Escola Superior de Tecnologia e Gestão

Inês Prudêncio Sena

Outubro 2020

Agradecimentos

Um trabalho de mestrado é uma longa viagem, que inclui uma trajetória permeada por inúmeros desafios, tristezas, incertezas, alegrias e muitos percalços pelo caminho, mas apesar do processo solitário reúne contributos de várias pessoas, indispensáveis para encontrar o melhor rumo em cada momento da caminhada. Trilhar este caminho só foi possível com o apoio, energia e força de várias pessoas, a quem dedico especialmente este projeto.

Agradeço à Professora Ana Isabel Pereira, minha orientadora, pela disponibilidade, ensinamentos, incentivos, ajuda e dedicação incondicional, pela partilha do saber e valioso contributo na elaboração desta dissertação.

À Escola Superior de Tecnologia e Gestão de Bragança pela disponibilidade, abertura e apoios proporcionados, que possibilitou ao longo de todo o meu percurso escolar.

Ao Rui, meu companheiro, amigo e maior apoio, por ter caminhado ao meu lado, pela sua paciência, carinho, compreensão e ajuda prestada durante a minha vida académica.

Aos meus amigos, em especial ao João Mendes que esteve ao meu lado durante todo este percurso. Agradeço a todos pelo companheirismo, paciência, atenção, força e apoio prestado em momentos difíceis.

Por último, um agradecimento especial aos meus pais, por serem modelos de coragem, pelo seu apoio incondicional, incentivo, amizade, força, carinho e paciência demonstrados e total ajuda na superação dos obstáculos que ao longo desta caminhada foram surgindo. Sem esquecer, a minha família pela ajuda, apoio, incentivo e encorajamento.

Resumo

A segurança do paciente constitui um dos grandes desafios dos cuidados de saúde do século XXI. Deste modo um fator importante para estabelecer diagnósticos clínicos e prevenção de erros é a comunicação médico-doente.

Diariamente os profissionais de saúde são procurados por pacientes, motivados pela vontade de permanecer saudável, realizando inúmeros diagnósticos que podem estar errados por diversos motivos.

O objetivo principal desta dissertação é desenvolver uma aplicação com o intuito de evitar erros no diagnóstico médico, auxiliando assim os profissionais de saúde nos procedimentos de diagnóstico. A aplicação ProSmartHealth é baseada em treinar e testar um procedimento de aprendizagem com o conjunto de dados de saúde para identificar clusters e padrões nos sintomas humanos. No início terá um questionário integrado, onde os profissionais de saúde preenchem perguntas chave com os resultados dos pacientes e, no final, a aplicação indicará uma sugestão de diagnóstico. Até agora, a ProSmartHealth considera o diagnóstico associado a doença cardíaca, cancro da mama e demência.

A ProSmartHealth utiliza o algoritmo de classificação Support Vector Machine, do Supervised Learning, para treinar e testar as sugestões de diagnóstico.

Neste estudo obteve-se uma precisão média de 84% na identificação das doenças estudadas.

Palavras Chaves: Inteligência Artificial, Machine Learning, Algoritmo de Classificação, Support Vector Machine

Abstract

Patient safety is one of the major health care challenges of the 21st century. Thus, an important factor for establishing clinical diagnoses and preventing errors is doctor-patient communication.

Daily, health professionals are sought out by patients, motivated by the will to stay healthy, making numerous diagnoses that can be wrong for several reasons.

The main objective of this work is to develop an application in order to avoid errors in medical diagnosis, thus assisting health professionals in diagnostic procedures. The ProSmartHealth application is based on intelligent algorithms to identify clusters and patterns in human symptoms. In the beginning, there will be an integrated questionnaire, where health professionals fill in key questions with the results of the patients and, in the end, the application will indicate a suggestion of diagnosis. So far, ProSmartHealth considers the diagnosis associated with heart disease, breast cancer, and dementia.

ProSmartHealth uses Supervised Learning's Support Vector Machine rating algorithm to train and test diagnostic suggestions.

In this study, an average accuracy of 84% was obtained in identifying the diseases studied.

Keywords: Artificial Intelligence, Machine Learning, Classification Algorithm, Support Vector Machine

Conteúdo

1	Introdução	1
1.1	Estado da Arte	1
1.2	Objetivos	3
1.3	Apresentação Sumária dos Capítulos	4
2	Introdução à Inteligência Artificial	7
2.1	<i>Machine Learning</i>	8
2.1.1	<i>Naive Bayes Classification</i>	10
2.1.2	<i>Decision Tree</i>	11
2.1.3	<i>K-Nearest Neighbor</i>	11
2.1.4	<i>Fuzzy Logic</i>	12
2.1.5	Redes Neurais Artificiais	12
2.1.6	<i>Support Vector Machine</i>	13
2.2	<i>Deep Learning</i>	14
2.3	Aplicações na área da saúde	15
2.3.1	Inteligência Artificial	15
2.3.2	<i>Machine Learning</i>	16
2.3.3	<i>Support Vector Machine</i>	17
2.3.4	<i>Deep Learning</i>	17
3	Metodologia	19
3.1	Teoria de Aprendizagem	20

3.2	Otimização Numérica	24
3.3	Classificação Binária	26
3.3.1	Classificação Binária Linear	26
3.3.2	Classificação Binária Não Linear	32
3.4	Classificação Multi-classe	35
3.4.1	Um contra Um	36
3.4.2	Um contra todos	37
4	Caracterização da Base de Dados	39
4.1	Cancro da Mama	39
4.2	Demência	41
4.3	Doença Cardíaca	43
5	Análise e Discussão dos Resultados	47
5.1	Cancro da Mama	48
5.2	Demência	52
5.3	Doença Cardíaca	56
6	Aplicação ProSmartHealth	61
7	Conclusões e Trabalhos Futuros	65
A	Anexos	A73

Lista de Tabelas

3.1	Funções de <i>Kernel</i>	33
5.1	Cancro da Mama - Amostra da base de dados utilizada.	48
5.2	Cancro da Mama - Amostra dos resultados de previsão obtidos.	50
5.3	Cancro da Mama - Relação de diagnósticos.	50
5.4	Demência - Amostra da base de dados utilizada.	52
5.5	Demência - Amostra dos resultados obtidos.	54
5.6	Demência - Relação entre resultados.	54
5.7	Doença Cardíaca - Amostra da base de dados utilizada.	56
5.8	Doença Cardíaca - Amostra dos resultados de sugestão de diagnóstico obtidos.	58
5.9	Doença Cardíaca - Matriz de confusão dos resultados obtidos.	58

Lista de Figuras

2.1	Modelo representativo da ligação entre a IA e as ciências de computação.	7
2.2	Esquema representativo do processo de <i>Machine Learning</i> [4].	8
2.3	Esquema representativo dos diferentes tipos de <i>Machine Learning</i>	10
2.4	Hiperplano ideal que separa duas classes [14].	13
2.5	Representação do grau de utilização dos algoritmos pertencentes ao <i>Machine Learning</i> [2].	14
2.6	Esquema representativo do processo de <i>Deep Learning</i> [4].	14
3.1	O limite de risco é a soma do risco empírico com o intervalo de confiança [6].	24
3.2	Representação da linha de separação e das linhas de suporte [28].	28
3.3	Representação de um hiperplano ideal [28].	28
3.4	Exemplo de classificação binária não linear [28].	32
4.1	Biopsia aspirativa ou PAAF (punção aspirativa por agulha fina) [30]	40
5.1	Cancro da Mama - Gráfico representativo da relação entre o parâmetro 1 (Raio) e o parâmetro 2 (Textura).	49
5.2	Cancro da Mama - Gráfico com a comparação dos diagnóstico reais com os previstos do parâmetro 1 (raio) e o parâmetro 2 (textura).	51
5.3	Demência - Gráfico representativo da relação entre o parâmetro 1 (Idade) e o parâmetro 2 (CDR).	53
5.4	Demência - Gráfico com a comparação dos diagnósticos reais com os previstos do parâmetro 1 (Idade) e o parâmetro 2 (CDR).	55

5.5	Doença Cardíaca - Gráfico representativo da relação entre o parâmetro 4 (Pressão Arterial) e o parâmetro 5 (Colesterol)	57
5.6	Doença Cardíaca - Gráfico com a comparação dos diagnósticos reais com os previstos do parâmetro 4 (Pressão Arterial) e o parâmetro 5 (Colesterol).	59
6.1	Página inicial da aplicação ProSmartHealth.	61
6.2	Questionário do cancro da mama preenchido.	62
6.3	Questionário da Demência preenchido.	63
6.4	Questionário da Doença Cardíaca preenchido.	64

Lista de Acrónimos

IA Inteligência Artificial

ML Machine Learning

DL Deep Learning

SVM Support Vector Machine

DT Decision Tree

CART Classification And Regression Tree

RNA Redes Neurais Artificiais

DR Rerinopatia Diabética

MA Micro Aneurisma

FallDS-IoT Sistema de detecção de quedas baseada na internet das coisas

ECG Eletrocardiograma

TAE Teoria de Aprendizagem Estatística

ERM Minimização do Risco Empírico

VC Vapnik-Chervonenkis

SRM Minimização do Risco Estrutural

SMO Otimização Sequencial Mínima

ECOC Error Correctiong Output Code

PAAF Punção Aspirativa por Agulha Fina

CDR Clinical Dementia Rating

MMSE Mini Mental State Examination

n-WBV Normalized Whole Brain Volume

Lista de Símbolos

f função de decisão

g função real

l elementos de treino

x valores de entrada

y valores de saída

T conjunto de treino

$F(z)$ medida de probabilidade

$Q(z, \beta)$ função de perda

Q conjunto de funções

β parâmetro de aprendizagem

z par (x, y)

$R(\beta)$ função de risco

$R_{emp}(\beta)$ função de risco empírico

h dimensão VC

L função de Lagrange

α multiplicador de Lagrange

w hiperplano

b ponto de interseção da reta

Y conjunto de classes

sgn função de sinal

K função de Kernel

Φ mapa

H espaço Hilbert

ξ variáveis de folga

C parâmetro de penalidade

M regiões

m comprimento da palavra de código

Capítulo 1

Introdução

Este trabalho foca-se na utilização do algoritmo de classificação *Support Vector Machine* do *Supervised Learning* para desenvolver uma aplicação que apoie os profissionais de saúde no diagnóstico, com o propósito de diminuir os erros no diagnóstico médico, até agora, de três doenças: demência, doença cardíaca e cancro da mama. A aplicação denominada ProSmartHealth, foi elaborada através da junção de três códigos modelos que identificam as três doenças individualmente no *software Matlab*.

1.1 Estado da Arte

Atualmente, existe uma maior disponibilidade de dados, uma maior variedade de técnicas de análise de dados e um aumento na capacidade de processamento de informação dos computadores, o que permite que a Inteligência Artificial, e as suas ciências de computação, *Machine Learning* e *Deep Learning* promovam uma mudança de paradigma em diversas áreas, principalmente na saúde, contribuindo na indústria médica, para apoiar os médicos em diferentes domínios e aplicações, como o estudo da transmissão e identificação de risco de doença, entre outros [1], [2].

A Inteligência Artificial estuda os procedimentos inteligentes em máquinas, que envolve os conceitos de perceber, aprender, comunicar e agir em ambientes complexos [3]. A IA tem diversos objetivos, um deles é desenvolver uma máquina que consiga fazer tudo isto

de igual forma ou melhor que os humanos; outro objetivo é conseguir compreender esse tipo de comportamento, seja em máquinas ou humanos [3].

Atualmente, consegue-se capturar, processar, criar e armazenar um elevado número de dados através de procedimentos computacionais e dos sistemas de comunicação. Cada vez mais existem aplicações que guardam vários tipos de dados, promovendo a disponibilidade de base de dados, cada vez maiores, e com elevada variedade de informação. As evoluções dos sistemas de base de dados permitem o desenvolvimento de procedimentos que identificam padrões, tendências, *cluster*, entre outros [3].

Nesse sentido, a Inteligência Artificial tem sido aplicada em diversos setores, como na eletrônica, mecânica, computação, saúde, entre outras. Para tentar apoiar a área da saúde utilizam-se métodos computacionais, como o *Deep Learning* e o *Machine Learning* [4].

O ML, surgiu na década de 80 como a primeira forma de colocar os conceitos de Inteligência Artificial em prática, baseia-se na ideia de que os sistemas podem aprender com dados a identificar padrões e a tomar decisões com o mínimo de intervenção humana [1]. O DL, surgiu em 2000 através do aparecimento de computadores poderosos e da disponibilidade de um maior número de dados, tornando-se uma ferramenta de computação mais sofisticada que o *Machine Learning*. É constituída por redes neuronais, que através de um grande base de dados, conseguem treinar modelos computacionais para identificarem o que um humano já sabe [1].

O *Machine Learning* utiliza diversos métodos que podem ser utilizados em diferentes tipos de ações, como melhorar aplicações de assistência médica, as capacidades de diagnóstico, entre outros. O ML é constituído por quatro técnicas diferentes, *Unsupervised Learning*, *Supervised Learning*, *Reinforcement Learning* e *Semi-Supervised Learning* [5].

O *Supervised Learning* é o primeiro conceito a ser estudado quando se começa a estudar o *Machine Learning*, que é muito semelhante ao processo do ser humano aprender [4]. As duas técnicas possíveis são a classificação e a regressão. A classificação é a mais prevalente, encontra as classes às quais os dados pertencem, já a regressão não determina as classes, estima sim um valor [4].

Estas duas técnicas de aplicação do *Supervised Learning* têm diferentes algoritmos para diferentes tipos de problemas, para este estudo baseamo-nos na técnica de classificação com o algoritmo de *Support Vector Machine*. Este algoritmo tem sido muito usado recentemente, no entanto o algoritmo foi desenvolvido entre a década de 60 e 70, com os conceitos de funções *Kernel* e hiperplano ideal, descobertos em 1964 [6].

Em 1964, Vladimir Vapnik e Alexey Chervonenkis desenvolveram o algoritmo de SVM, mas durante trinta anos a investigação ficou parada devido a existência de dois problemas no algoritmo original [6]. O primeiro problema é que o algoritmo original era limitado a problemas de aprendizagem linearmente separáveis e o segundo era que não apresentava uma forma de tratar os dados isolados ou muito distantes do padrão original. Assim, através das primeiras investigações desenvolvidas para a criação deste algoritmo conseguiu-se resolver esses problemas, em que o primeiro foi solucionado com o uso de funções *Kernel* e o segundo foi resolvido através da introdução de variáveis de folga. Ao juntar esses dois conceitos formulou-se o SVM que foi apresentado oficialmente por Vapnik e seus colaboradores em 1992 [6].

No decorrer de uma revisão bibliográfica do que existe relacionado com este a área, constatou-se que os algoritmos de *Machine Learning* e *Deep Learning* são muito utilizados para identificar doenças antes de elas se manifestarem, por exemplo, para identificar se o paciente tem, ou poderá vir a contrair, diabetes, alzheimer, algum tipo de cancro, entre outros. O *Support Vector Machine* e Redes Neurais do *Supervised Learning* são os algoritmos com uma elevada popularidade de utilização em aplicações médicas para detecção e diagnóstico precoces, tratamento, previsão de resultados e avaliação de prognóstico [2].

1.2 Objetivos

O presente estudo tem como objetivo principal desenvolver e aplicar diversas estratégias, baseadas em processos de otimização, para desenvolver uma aplicação para apoiar os profissionais de saúde na tomada de decisão do diagnóstico final.

Com o intuito de alcançar os objetivos elencados, o trabalho apresentado neste relatório

dividiu-se nas seguintes etapas:

- Estudo do problema e estado de arte;
- Modelação do problema;
- Estabelecer comparações com as diversas modelações matemáticas já existentes;
- Concluir e discutir os resultados obtidos.

1.3 Apresentação Sumária dos Capítulos

A dissertação encontra-se organizada em sete capítulos, de forma a facilitar a sua compreensão. O conteúdo destes pode ser sintetizado da seguinte forma:

- O **Capítulo 1 'Introdução'** tem a finalidade de apresentar um enquadramento do trabalho, fazendo referência aos principais objetivos da dissertação e uma breve descrição da sua estrutura.
- O **Capítulo 2 "Introdução à Inteligência Artificial"** visa apresentar alguns fundamentos teóricos considerados relevantes para a correta compreensão do estudo apresentado nesta dissertação. No início, apresenta-se uma descrição da Inteligência Artificial e as ciências de computação que a constituem, dando mais importância ao *Machine Learning*. Seguindo-se de uma breve descrição do *software* utilizado para a realização deste estudo.
- O **Capítulo 3 "Metodologia"** visa a apresentar teoricamente o algoritmo de *Support Vector Machine* considerando a sua teoria de aprendizagem, o problema de otimização utilizado e os tipos de classificação existentes.
- O **Capítulo 4 "Caracterização da Base de Dados"** pretende mostrar a base de dados utilizada para o presente estudo.

- O **Capítulo 5 “Análise e Discussão dos Resultados”** apresenta o procedimento utilizado, os resultados obtidos ao longo do estudo e analisa-os efetuando breves considerações dos mesmos.
- O **Capítulo 6 "Aplicação ProSmartHealth"** apresenta a aplicação desenvolvida, com uma breve explicação do seu funcionamento.
- O **Capítulo 7 "Conclusões e Trabalhos Futuros"** são apresentadas as conclusões gerais do presente trabalho e apresentam-se também algumas propostas para trabalhos futuros.

Capítulo 2

Introdução à Inteligência Artificial

A Inteligência Artificial é a ciência que estuda e procura compreender o fenómeno da inteligência. Dá apoio aos humanos, apoiando-os na realização de tarefas rotineiras [7].

Um sistema IA, além de ser capaz de armazenar e manipular dados, consegue também adquirir, representar e manipular conhecimento. Atualmente pode ser implementada em diversas áreas, nomeadamente, no carro autónomo, no chão da fábrica, no sistema de atendimento dos hospitais, nas redes sociais, no telemóvel, entre outros [7].

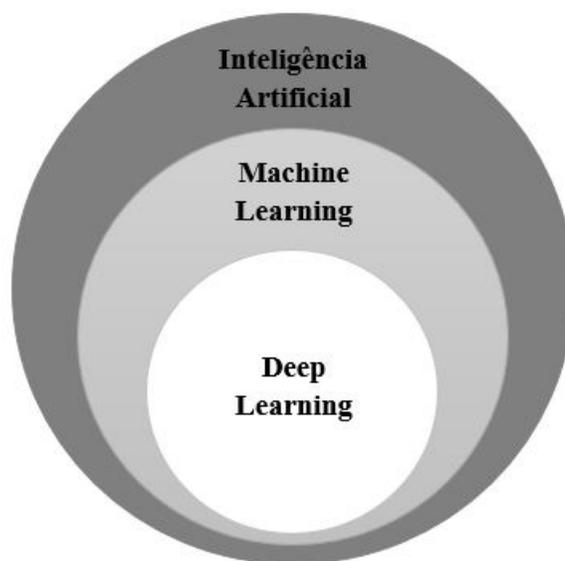


Figura 2.1: Modelo representativo da ligação entre a IA e as ciências de computação.

Como se pode observar na Figura 2.1, o *Machine Learning* e o *Deep Learning* são duas estratégias da Inteligência Artificial.

2.1 *Machine Learning*

O *Machine Learning* é uma técnica de análise e modelação de dados, que tem como objetivo descobrir o modelo dos dados, como a previsão de situações.

O ML analisa dados e encontra o modelo por si só, sem ajuda de um humano, designa-se isso por aprendizagem porque o processo assemelha-se a ser treinado com os dados para resolver o problema de encontrar um modelo [4].

Na Figura 2.2 pode-se observar o procedimento do ML, que através de um algoritmo de *Machine Learning* usa os dados de treino (*training data*), para criar o modelo. Assim, ao fornecer os dados de teste (*input data*) gera-se os dados de saída (*output*) [4].

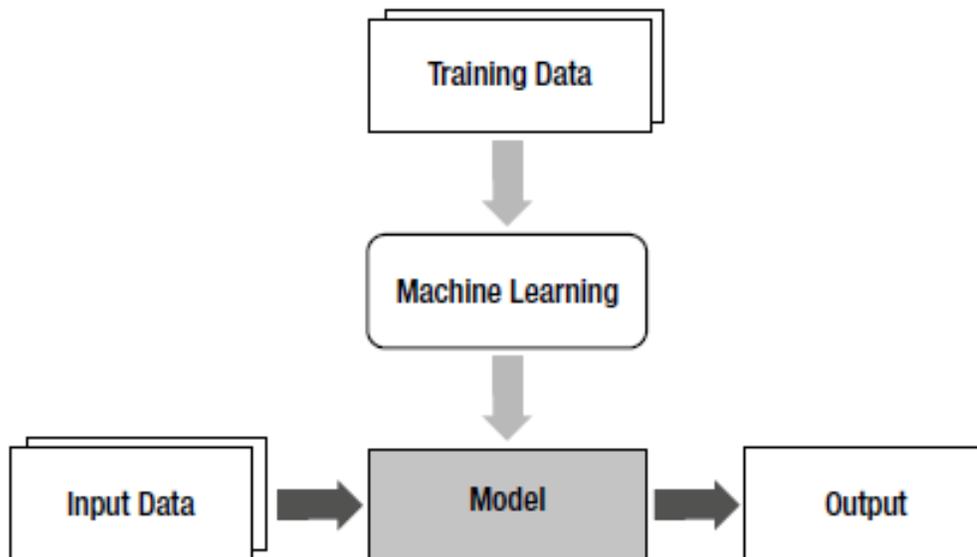


Figura 2.2: Esquema representativo do processo de *Machine Learning* [4].

Existem diferentes tipos de *Machine Learning* que servem para resolver problemas em vários campos, representados na Figura 2.3.

- ***Supervised Learning*** - é uma técnica muito semelhante ao processo de aprendizagem do ser humano [4]. É realizado sob treino, em que o algoritmo é constituído por entradas e pares de saída corretos [5]. Divide-se em diferentes técnicas com diferentes algoritmos:
 - **Regressão:** Regressão Linear, Regressão Lógica e *Decision Tree*.
 - **Classificação:** *Support Vector Machine*, *Naive Bayes Classification*, *Decision Tree*, *K-Nearest Neighbor*, Redes Neurais Artificiais e *Fuzzy Logic*.

- ***Unsupervised Learning*** - é utilizado, geralmente, para investigar as características e pré-processamento dos dados [4]. É um método semelhante a um aluno que apenas resolve os problemas por construção e atributo, mas que não aprende a resolvê-los porque não têm saídas corretas conhecidas, assim os seus dados para treinar o modelo contêm apenas dados de entrada [4]. É constituído por diferentes algoritmos:
 - **Clustering:** *K-means*, *Hierarchical model* e *Mixture model*.

- ***Reinforcement Learning*** - é usado, geralmente, quando é necessária uma interação ideal, como controlar as jogadas em jogos, um exemplo é o jogo de xadrez [4]. Divide-se em diferentes algoritmos:
 - **Classificação e Clustering:** *Low density Separation*, *Graph Based method* e *Heuristic Approach*.

- ***Semi-Supervised Learning*** - é uma técnica que identifica o melhor classificador de cada informação não rotulada e rotulada [5]. Utilizando informações não rotuladas, que consiste em ocorrências diferentes dos próprios sinais de entrada, consegue-se obter alto desempenho de classificação [5]. Divide-se em diferentes algoritmos:
 - **Classificação e Controlo:** *Monte Carlo Method*, *Temporal Difference method* e *Direct policy Search*.

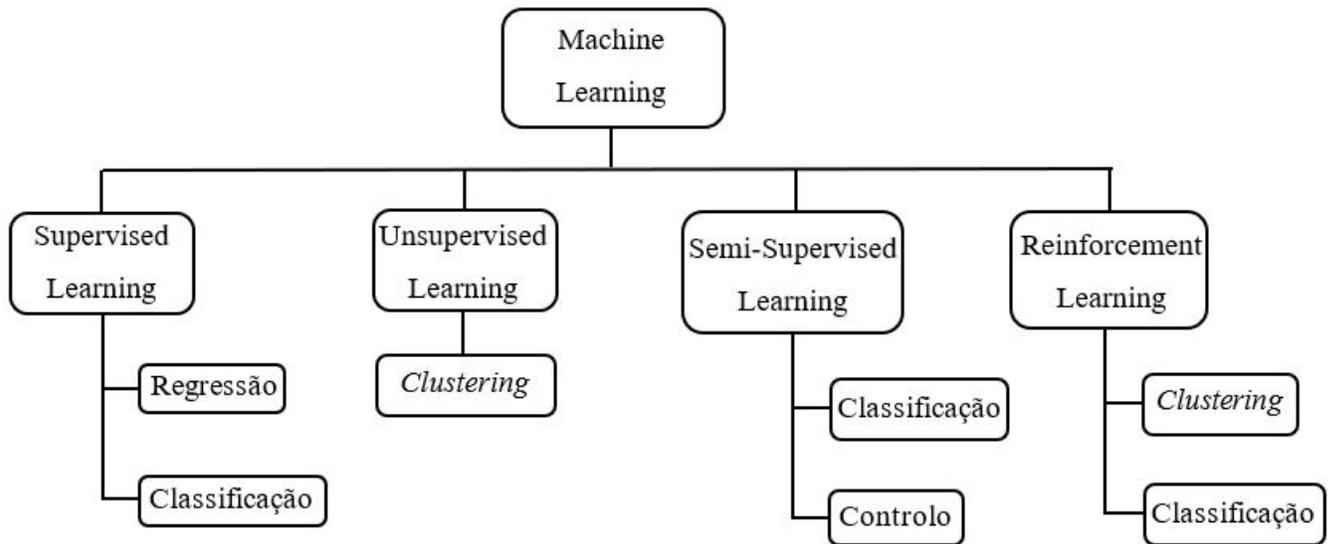


Figura 2.3: Esquema representativo dos diferentes tipos de *Machine Learning*.

O *Supervised Learning* é a técnica de ML escolhida para realizar este estudo, como referido anteriormente, é dividido em técnicas de classificação e regressão. Enquanto que a classificação prevê um rótulo de classe discreto a regressão prevê uma quantidade contínua.

As subsecções seguintes descrevem brevemente alguns algoritmos constituintes da técnica de classificação do *Supervised Learning*, uma vez que é o método utilizado ao longo do estudo.

2.1.1 *Naive Bayes Classification*

O método de *Naive Bayes*, é um exemplo de classificador probabilístico que consegue prever as probabilidades de associação de classe baseadas no rótulo de determinada classe. Executa uma digitalização de dados e, portanto, a classificação é fácil.

O *Naive Bayes Classification* pode ser utilizado, por exemplo, para desenvolver um sistema probabilístico de identificação do cancro da mama. Este classificador pode ser utilizado para ajudar um especialista numa decisão com uma alta precisão [8].

2.1.2 *Decision Tree*

A *Decision Tree* é uma ilustração gráfica de uma decisão exata, que pode ser utilizado para regressão ou classificação. Os nós, as raízes e as decisões de ramificação são os principais componentes de uma DT.

A DT pode ser utilizada, quando existem problemas médicos com um elevado número de parâmetros de classificação de um conjunto de dados. As duas principais desvantagens deste algoritmo são o *overfitting*, que acontece quando existe um elevado número de dados e não existem nós suficientes para espalhá-los, e o ser ganancioso que leva-nos a ter uma árvore menos ideal, mas esta desvantagem é resolvida através da *J48 Decision Tree* [9].

Existem diferentes abordagens que podem ser utilizadas na *Decision Tree*, como o ID3, CYT, C5.0, J48 e o CART. O CART, é a variação do algoritmo *Decision Tree* mais utilizada e está estruturado como uma sequência de perguntas, cujas respostas determinam qual será a próxima pergunta, se existir [10]. Os resultados a essas perguntas aparecem em estrutura de árvore onde as extremidades são os nós terminais.

2.1.3 *K-Nearest Neighbor*

A técnica de *K-Nearest Neighbor* é frequentemente usada na classificação das amostras, ao usar esta técnica a distância é calculada para treinar o conjunto de dados para cada K pontos de dados mais próximos a partir do número N de amostras de treino. O princípio de funcionamento do *K-Nearest Neighbor* é baseado numa atribuição de peso para cada ponto de dados conhecido como vizinho [9].

Este algoritmo é dividido em duas etapas:

- Treinar a partir dos dados;
- Testar em novas instâncias;

É um algoritmo supervisionado e muito lento, visto que demora mais tempo para conseguir obter classificação treinadas, mas é especialmente eficaz para grandes conjuntos de dados.

2.1.4 *Fuzzy Logic*

As bases lógicas do raciocínio podem ser distinguidas essencialmente por três itens neutros: valores de verdade, vocabulário (operadores) e procedimento de raciocínio [11].

Este algoritmo tem uma forma lógica com muitos valores, na qual os valores verdadeiros das variáveis podem ser qualquer número real entre $[0, 1]$ [10]. A lógica difusa, é aplicável em muitos campos da teoria do controle à IA e é utilizada principalmente para lidar com o conceito de verdade parcial, onde o valor da verdade pode variar entre completamente verdadeiro e falso [5].

2.1.5 Redes Neurais Artificiais

As Redes Neurais Artificiais são utilizadas na ciência da computação como modelos computacionais inspirados pelo sistema nervoso central. Todos os organismos multi celulares possuem algum tipo de sistema nervoso, cuja sua complexidade e organização varia de acordo com o tipo de animal [12].

Este sistema é responsável por dotar o organismo, através de entradas sensoriais, de informações sobre o estado do ambiente no qual vive e se move. A informação de entrada é processada, comparada com as experiências passadas, e transformada em ações apropriadas ou absorvidas sob a forma de conhecimento. O sistema nervoso pode ser organizado em diferentes níveis: moléculas, sinapses, neurónios, camadas, mapas e sistemas [12].

A representação de conhecimentos nas redes neurais, como diz o próprio nome, é fortemente ligada a noção de conexão entre neurónios (elementos processadores de informação) que interagem uns com os outros através das ligações. O conhecimento de uma rede neural artificial está codificado na estrutura da rede, onde se destacam as conexões (sinapses) entre as unidades (neurónios) que a compõe.

Assim, as redes neuronais artificiais podem ser definidas como estruturas de processamento (rede) compostas por um número de unidades interconectadas (neurónios artificiais), sendo que cada unidade apresenta um comportamento específico de entrada/saída determinado pela sua função de transferência [12].

2.1.6 *Support Vector Machine*

O *Support Vector Machine* é um conceito na ciência da computação projetado para obter suporte às tarefas de *Machine Learning* [13]. É uma ferramenta simples e com um processo de destaque. Originalmente projetada para classificação binária, podendo ser estendida para classificação multi-classe.

O modelo SVM representa pontos no espaço bidimensional, com o intuito de encontrar um hiperplano ideal que consiga separar os pontos, em duas classes, caso se trate de uma classificação binária, como se pode observar na Figura 2.4.

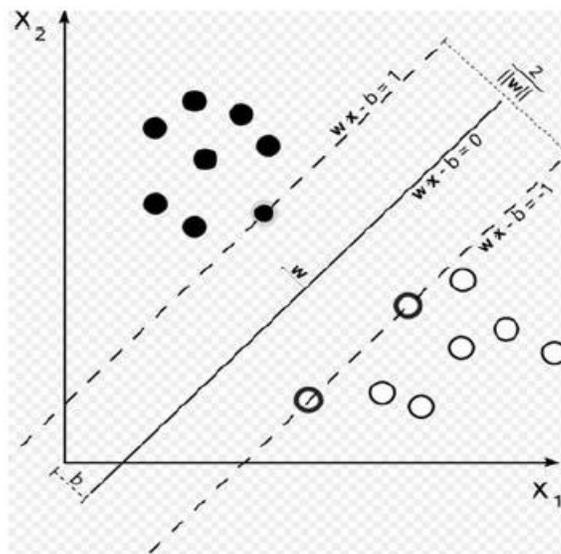


Figura 2.4: Hiperplano ideal que separa duas classes [14].

Com o hiperplano pretende-se minimizar a distância entre os pontos mais próximos em relação a cada uma das classes. Os pontos que se encontrarem mais próximos do hiperplano designam-se por vetores de suporte.

A Figura 2.5 faz referência ao grau de utilização de cada algoritmo demonstrando que o *Support Vector Machine* e as Redes Neurais Artificiais são os mais utilizados.

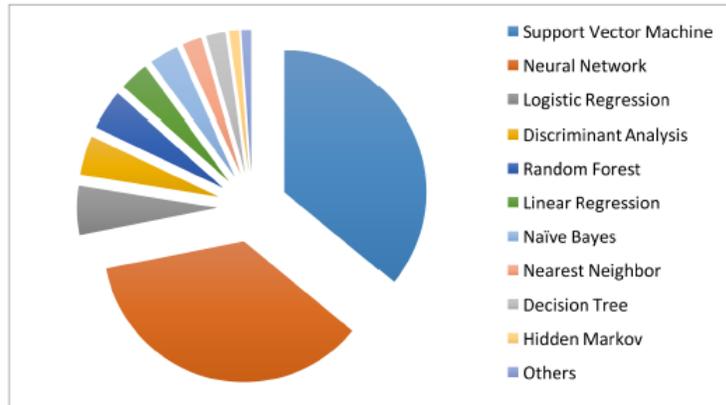


Figura 2.5: Representação do grau de utilização dos algoritmos pertencentes ao *Machine Learning* [2].

2.2 *Deep Learning*

O *Deep Learning* (DL) é um método que utiliza rede neural profunda, a multi camada rede neural que contém duas ou mais camadas ocultas [4].

Na Figura 2.6 pode-se observar como se processa o *Deep Learning*. Os dados de treino (*training data*) são usados através da *Deep Neural Network* para identificar a *Learning Rule*. Assim que a rede esteja treinada são fornecidos dados de entrada (*input data*) e é gerada uma saída (*output*) [4].

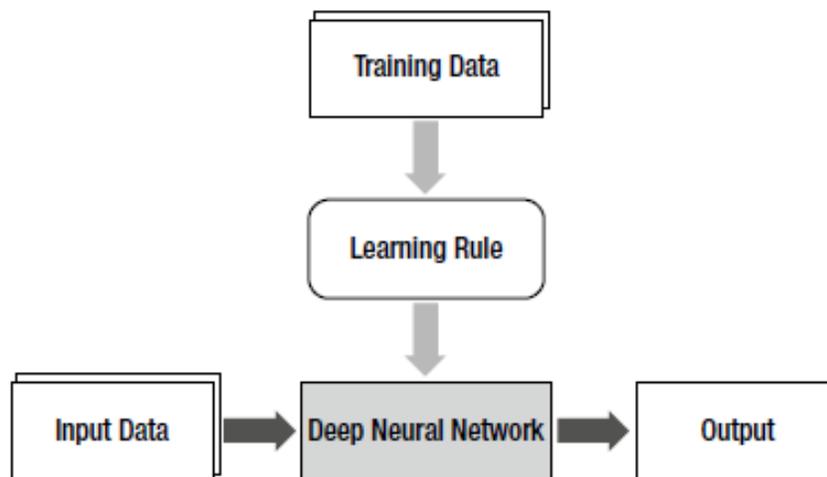


Figura 2.6: Esquema representativo do processo de *Deep Learning* [4].

Através do *Deep Learning* consegue-se treinar procedimentos para que se consiga identificar objetos que naturalmente os seres humanos conseguem identificar, por exemplo, desde tenra idade que somos ensinados a distinguir os números, em que sabemos que isto é 0 (zero) e isto é 1 (um), ou seja, pode-se dizer que fomos treinados para distingui-los, assim também se poderá treinar os computadores para distingui-los [4].

Esta técnica é muito importante atualmente devido ao seu valor de precisão, mas para esse mesmo valor necessita de uma quantidade elevada de resultados reais, por exemplo, hoje em dia existem automóveis que se conseguem mover sem precisar de um motorista, isto deve-se a evolução do *Deep Learning* e ao facto do modelo ter sido treinado com milhões de imagens e milhares de horas de vídeo [4].

2.3 Aplicações na área da saúde

Nesta secção serão apresentadas diversas aplicações da Inteligência Artificial, *Machine Learning*, *Support Vector Machine* e do *Deep Learning* na área da saúde.

2.3.1 Inteligência Artificial

A IA através das suas técnicas de computação pode apoiar médicos, criando modelos que detetam e apoiam o diagnóstico, tratam resultados e sugerem previsões, por exemplo, de doenças. Além disso, com um elevado número de informações consegue-se criar *softwares* que alertam para o risco, ou que conseguem diagnosticar e identificar o aparecimento de doenças [15].

A *Google DeepMind* e a *Enlitic* são duas empresas que estudam e aplicam a Inteligência Artificial na saúde. A *DeepMind* é uma empresa que pesquisa e constrói sistemas de Inteligência Artificial, um dos programas que estão a desenvolver é para diagnosticar doenças oculares com a mesma eficácia que os médicos, em que a partir de 2014 uniram forças com a Google para acelerar o processo [16]. A *Enlitic* usa dados para avançar no diagnóstico médico, com o intuito de utilizar a Inteligência Artificial para criar um

software que permita aos médicos diagnosticar mais rapidamente e com uma precisão reconhecida [17].

A IA tem sido muito aplicada em vários setores da saúde como a neurologia e a cardiologia. Um exemplo de utilização é no diagnóstico e previsão de acidentes vasculares cerebrais [2].

2.3.2 *Machine Learning*

O ML evoluiu rapidamente na indústria médica, podendo ser aplicado no diagnóstico, na visualização, no estudo da transmissão de doenças, na segmentação de imagens médicas, registros de imagens, fusão de imagens, diagnóstico assistido por computador, terapia guiada por imagem, anotação de imagem e recuperação do banco de dados de imagem [2], [18].

Um exemplo de utilização é o desenvolvimento de um modelo para uma classificação hospitalar baseada em diagnóstico de internamentos de emergência ou não, porque cada vez mais existe um aumento nos serviços de emergência nos hospitais o que começa a preocupar muitos sistemas. Os internamentos hospitalares são caracterizadas em dois tipos, o primeiro tipo é o internamento de não emergência, em que o paciente pode ter sido encaminhado por um médico de cuidados primários, e o segundo tipo é internamentos não programados, ou seja, emergência, que inclui paciente com urgência ou risco de vida que requerem assistência médica imediata. Uma maneira de utilizar esta técnica é num estudo para desenvolver um modelo para uma classificação hospitalar baseada em diagnóstico de internamentos com atendimento de emergência ou não, com isto, pretende-se prever a urgência das admissões com um valor numérico que reflete o grau de planificação disponível para um hospital. Para este fim foram utilizadas técnicas de *Machine Learning* para desenvolver um modelo de classificação, e assim treiná-lo com inúmeros dados julgados por especialistas [19].

Os cuidados paliativos são associados a pacientes que sofrem com uma doença terminal. O ML foi utilizado para avaliar o risco de mortalidade de pacientes nos cuidados paliativos

num ano, através da previsão de saída de pacientes no próximo ano usando dados coletados na admissão hospitalar [20].

2.3.3 *Support Vector Machine*

O *Support Vector Machine* é um algoritmo que pretende melhorar os cuidados de saúde através do desenvolvimento de modelos preditivos para aumentar a precisão de diagnóstico. De seguida serão apresentados alguns exemplos onde o SVM é utilizado.

Um exemplo de aplicação do SVM é a classificação arritmica do batimento do coração. Neste caso a técnica é usada principalmente na deteção de anormalidade do eletrocardiograma para identificar problemas relacionados com o coração. A classificação de batimentos arrítmicos baseada em treino por máquina é realizada para categorizar em indivíduos normais e anormais. Utilizou-se o algoritmo de SVM com o intuito de conseguir classificar a batida num sinal extraído do recurso de remoção de ruído [21].

O SVM pode ser utilizado para diagnosticar o cancro da mama, através de uma base de dados com o histórico médico dos pacientes. Foram efetuados estudos com diferentes números de classes de forma a prever o diagnóstico com alta precisão. A base de dados baseia-se no diagnóstico desta doença, ou seja, possuiu parâmetros de entrada e respetivos diagnósticos [22].

O SVM também pode ser usado para prever a doença de Alzheimer, através do uso de uma base de dados foi realizado o treino do processo para obter um diagnóstico final fornecido com os valores correspondentes, esse treino é efetuado através de um método de classificação binária. Logo, este algoritmo consegue prever a demência e validar seu desempenho [23].

2.3.4 *Deep Learning*

Nesta era digital, as instituições de saúde dispõem de um elevado número de dados biomédicos que podem ser usados no auxílio de tomada de decisões.

Alguns exemplos de como o *Deep Learning* pode ser utilizado:

- No processamento de imagens clínicas, especialmente na análise de imagens de ressonância magnética do cérebro para prever a doença de Alzheimer e suas variações [15];
- Na biologia de alto rendimento, a tecnologia é usada para capturar a estrutura interna de conjuntos de dados cada vez maiores e de alta dimensão, como por exemplo, sequências de DNA e medições de RNA [15];
- Os smartphones e tablets equipados com sensores estão a transformar uma variedade de aplicações móveis, incluindo a monitorização da saúde. Estes dispositivos podem dar aos pacientes acesso direto a análises pessoais, facilitando os cuidados preventivos e auxiliar pacientes que já tenham doenças desenvolvidas [15].

Assim, o DL pode abrir caminho para uma próxima geração de sistemas de assistência médica que podem incluir muitos milhões a biliões de registos de pacientes e usá-los para apoiar efetivamente os médicos em diferentes domínios e aplicações, como na previsão de risco de doença, prescrições personalizadas, recomendações de tratamento, recrutamento de ensaios clínicos, pesquisa e análise de dados, entre outros [15].

Capítulo 3

Metodologia

Este capítulo visa a introduzir conceitos sobre o *Support Vector Machine*, ou seja, a sua teoria de aprendizagem, as estratégias de otimização e os tipos de classificação existentes.

Como já foi referido anteriormente, o *Support Vector Machine* é um algoritmo de classificação que pertence ao *Supervised Learning*, isto porque se trata de um gerador que recebe um conjunto de dados em que cada dado é formado por um conjunto de parâmetros (características) de entrada e um conjunto de classes, que representam o fenómeno de interesse sobre o qual se pretende fazer previsões.

No *Supervised Learning* existem duas técnicas possíveis de se utilizar, a regressão e a classificação. O que determina qual técnica aplicar é o tipo de classes a utilizar, caso as classes contenham valores contínuos, tem-se uma regressão, caso assumam valores discretos tem-se um problema de classificação.

O SVM inicialmente foi projetado para classificação binária, mas pode ser estendido para classificação multi-classe, o que determina o tipo de classificação a utilizar é o valor discreto da classe, pois se for igual a dois é denominado classificação binária e se for superior a dois considera-se uma classificação multi-classe.

Este algoritmo constitui um método de aprendizagem que tem vindo a receber grande atenção nos últimos anos, devido a ter uma boa capacidade de generalização e robustez na

presença de dados de grande dimensão [24]. Isto deve-se a Teoria de Aprendizagem Estatística (TAE), que apresenta diversos limites para controlar a capacidade de generalização de um classificador linear [6].

A generalização de um modelo define a capacidade de prever corretamente a classe de novos dados, que se divide em *overfitting* ou *underfitting*. O *overfitting* acontece quando o modelo se especializa nos dados utilizados no seu treino e obtém uma taxa de acerto baixa quando confrontado com os novos dados. Também é possível induzir hipóteses que apresentem uma baixa taxa de acerto mesmo no subconjunto de treino, acontecendo o *underfitting*.

3.1 Teoria de Aprendizagem

Esta secção tem como objetivo descrever o conceito do modelo para os processos de aprendizagem que são usados na construção de algoritmos do *Machine Learning*, em especial, o SVM.

Para construir qualquer teoria é necessário o uso de alguns conceitos em torno dos quais a teoria é desenvolvida. É importante utilizar conceitos que descrevam condições necessárias e suficientes para a capacidade de um método de aprendizagem convergir para uma solução ótima.

O modelo geral de aprendizagem é constituído por três componentes:

- Um gerador (G) de vetores aleatórios, $x \in \mathbb{R}^n$, selecionados de forma independente a partir de uma função de distribuição de probabilidade cumulativa $F(x)$ fixa, porém desconhecida [6];
- Um supervisor (S) que retorna um valor de saída desejada y , para cada vetor de entrada x , de acordo com uma função de distribuição condicional $F(y|x)$, também fixa e desconhecida [6];
- Um algoritmo de Machine Learning capaz de implementar um conjunto de funções $f(x, \beta)$, $\beta \in \Lambda$, em que Λ é um conjunto de parâmetros [6].

Assim, o problema de aprendizagem consiste em escolher o melhor conjunto de funções, $f(x, \beta), \beta \in \Lambda$, que melhor se aproxima a resposta do supervisor, ou seja, um valor de saída desejado.

A função desejada é baseada em um conjunto de treino, T , com l observações independentes e identicamente distribuídas (i.i.d) desenhadas de acordo com $F(x, y) = F(x)F(y|x)$:

$$T : \{(x_1, y_1), \dots, (x_l, y_l)\}.$$

Para alcançar a melhor aproximação entre a saída produzida pelo SVM e a saída desejada, é necessário trabalhar com uma medida designada por função de risco (erro).

Com o propósito de escolher o melhor conjunto de funções para a resposta do supervisor, mede-se a perda ou discrepância $Q(y, f(x, \beta))$ entre a resposta y do supervisor a uma determinada entrada x e a resposta $f(x, \beta)$ fornecida pelo *Support Vector Machine*. Através de um problema de aprendizagem mais geral, considera-se a medida de probabilidade $F(z)$, definida no espaço Z e uma função de perda específica de cada problema $Q(z, \beta), \beta \in \Lambda$, onde z descreve um par (x, y) . O valor esperado da perda é definido pela seguinte fórmula, designada por Função de Risco, $R(\beta)$:

$$R(\beta) = \int Q(z, \beta) dF(z), \quad \beta \in \Lambda,$$

na qual a medida de probabilidade $F(z)$ é desconhecida, porém z representa um par (x, y) e compreende dados entre z_1, \dots, z_l . Para aplicar essas definições a um problema particular, deve-se substituir a função de perda correspondente na fórmula.

A função de risco determina, para essa função de perda, a probabilidade de erro de classificação, deste modo tem como objetivo encontrar a função $Q(z, \beta_0)$ que minimiza o risco. Contudo a função de distribuição de probabilidade conjunta $F(z)$ é desconhecida e as únicas informações disponíveis encontram-se no conjunto de treino, assim temos um problema, que é encontrar uma função que minimize a probabilidade de erro de classificação quando a medida de probabilidade $F(z)$ é desconhecida, mas os dados são fornecidos.

Para minimizar a Função de Risco $R(\beta)$ pode-se aplicar o princípio indutivo de Minimização do Risco Empírico [6]. A Minimização do Risco Empírico é um princípio na Teoria de Aprendizagem Estatística que define uma família de algoritmos de aprendizagem e é usado para fornecer limites teóricos sobre seu desempenho.

Este princípio depende inteiramente do conjunto de dados de treino. O risco empírico é a medida da taxa de erro no conjunto de treino para um número de observações finitas e fixas. Esse princípio indutivo pode ser descrito como:

1. Substituir a função de Risco $R(\beta)$ pela função de Risco Empírico $R_{emp}(\beta)$:

$$R_{emp}(\beta) = \frac{1}{l} \sum_{i=1}^l Q(z_i, \beta),$$

construída a partir do conjunto de treino z_1, \dots, z_l .

2. Aproximar a função $Q(z, \beta_0)$ que minimiza a função de Risco pela função $Q(z, \beta_l)$ que minimiza o risco empírico.

A função de Risco Empírico difere da função de Risco por dois motivos principais: não depende, de forma explícita, da função de distribuição desconhecida, e pode ser minimizada em relação aos parâmetros de aprendizagem, β [6].

O erro de generalização não se consegue obter apenas pelo cálculo ERM, por isso, recorre-se a Teoria Vapnik-Chervonenkis que demonstra que é imperativo restringir a classe de funções a partir da qual f é escolhida para que a capacidade seja adequada à quantidade de dados de treino disponíveis [25]. O conceito de capacidade mais conhecido da Teoria VC é a dimensão Vapnik-Chervonenkis, que é a componente importante para o limite de generalização de SVM.

A dimensão VC, h , do conjunto de funções Q , é a responsável pelo limite de generalização de SVM. Quanto maior o seu valor, mais complexas são as funções de classificação que podem ser induzidas a partir de Q .

A dimensão Vapnik-Chervonenkis de um conjunto de funções indicadores $Q(z, \beta)$, $\beta \in \Lambda$, é o número máximo h de vetores z_1, \dots, z_h (relembrando que z descreve um par (x, y))

que podem ser separados em duas classes em todas as 2^h maneiras possíveis usando este conjuntos de funções.

Como foi referido anteriormente o ERM não determina um erro de teste pequeno. Para superar essa limitação, o SVM realiza a Minimização do Risco Estrutural. Este princípio demonstra o erro de generalização limitado pela soma do erro do conjunto de treino (risco empírico) e um termo de confiança que depende da dimensão VC do classificador.

Assim, o SRM é baseado num limite importante fornecido pela TAE que relaciona o risco esperado de uma função ao seu risco empírico e a um termo de confiança. Esse limite, apresentado na Equação (3.1), é garantido com a probabilidade $1 - \eta$, em que $\eta \in [0, 1]$.

$$R(\beta) \leq R_{emp}(\beta) + \sqrt{\frac{h(\log(\frac{2l}{h}) + 1) - \log(\frac{\eta}{4})}{l}}, \quad (3.1)$$

onde l é o tamanho do conjunto de treino, h a dimensão Vapnik-Chervonenkis (VC) da classe de funções Q à qual z pertence, e a parcela de raiz na soma é referenciada como termo de confiança.

O primeiro termo da inequação é designado por Limite de Risco, pois é o limite para o erro de generalização, e o segundo termo por Confiança VC [25]

Segundo Haykin o erro de treino, para um número fixo de exemplos, diminui quando a dimensão VC e o intervalo de confiança aumentam, assim, o Limite de Risco e a Confiança VC convergem para o mínimo [26].

Então, dado o conjunto S de funções $Q(z, \beta), \beta \in \Lambda$, é produzida uma estrutura que consiste na divisão da classe de funções de entrada em subconjuntos aninhados de funções $S_k = \{Q(z, \beta), \beta \in \Lambda_k\}$, tal que $S_1 \subset S_2 \subset \dots \subset S_l$.

Dado um conjunto de observações z_1, \dots, z_l , o princípio SRM escolhe a função $Q(z, \beta_l^k)$, minimizando o risco empírico do subconjunto S_k , que apresente o menor Limite de Risco, procurando assim um compromisso entre a qualidade de erro de treino e a complexidade do intervalo de confiança.

O risco empírico diminui com o aumento no índice do elemento da estrutura, enquanto o intervalo de confiança aumenta. O limite mínimo do risco é alcançado para algum

elemento apropriada da estrutura (S^* da Figura 3.1).

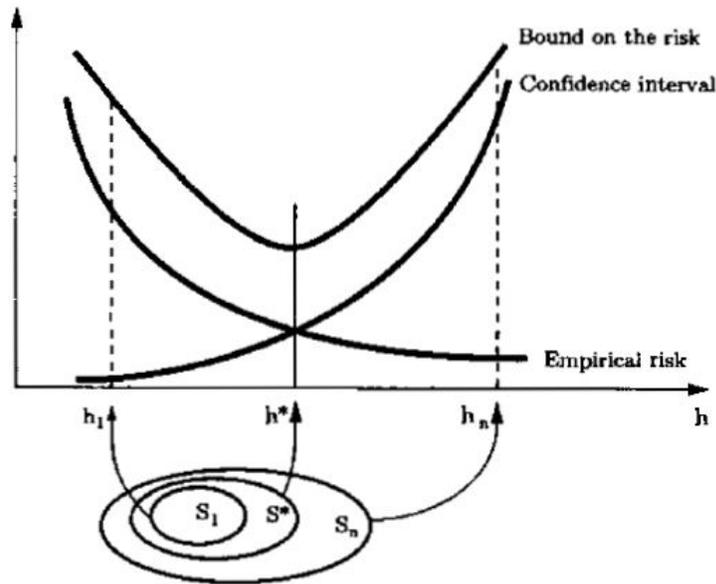


Figura 3.1: O limite de risco é a soma do risco empírico com o intervalo de confiança [6].

A Figura 3.1 representa o que se pretende para obter o *Support Vector Machine*. Observando a figura pode-se esclarecer que o *underfitting* acontece entre h_1 e h^* , e o *overfitting* entre h^* e h_n . Assim, o SVM implementa esse princípio estatístico para controlar a capacidade e prevenir o *overfitting*.

Na secção seguinte serão apresentadas definições de Otimização Matemática, porque os tipos de classificação do SVM resolvem-se através da resolução de problemas quadráticos convexos, e do algoritmo de Otimização Mínima Sequencial por ser o solvente utilizado pelo *software Matlab* para resolver problemas de classificação.

3.2 Otimização Numérica

A otimização é um ramo de matemática que se refere ao estudo de problemas de minimizar ou maximizar funções de qualquer natureza. No contexto SVM a otimização envolve a caracterização de soluções para classes de problemas compostos por funções que devem ser escolhidas para minimizar ou maximizar uma certa função objetivo.

Existem diferentes tipos de otimização, dependendo do tipo de restrições:

- **Otimização sem restrições:** abrange os problemas em que as variáveis podem assumir qualquer valor sem uma limitação de regras (restrições);
- **Otimização com restrições:** abrange a classe de problemas em que as variáveis podem assumir valores condicionados, ou seja, sujeitos a restrições. Essas restrições podem ser de igualdade e desigualdade. Uma restrição de desigualdade $g_i(w) \geq 0$ é ativa se $g_i(w) = 0$ e inativa se $g_i(w) > 0$. Em algumas situações são introduzidas variáveis de folga, ξ , para transformar uma restrição de desigualdade numa restrição de igualdade, como se poderá observar mais a frente no capítulo.

Pode-se identificar diversos tipos de otimização:

- **Programação linear:** a função objetivo e todas as funções de restrição são lineares;
- **Programação não-linear:** pelo menos a função objetivo ou uma das funções de restrição são não lineares;
- **Programação Quadrática:** a função objetivo é quadrática e as restrições são lineares;
- **Programação Convexa:** um conjunto de pontos é dito convexo se, dados quaisquer dois pontos no conjunto, o segmento de reta que os conecta está também no conjunto, ou seja, qualquer mínimo local é um mínimo global;
- **Programação Quadrática Convexa:** a função objetivo é convexa e quadrática, e considera problemas em que as restrições são lineares.

No contexto SVM trata-se de problemas quadráticos convexos e esses problemas serão solucionados através da função *Lagrange*. Todos os tipos de problemas de classificação baseiam-se no mesmo problema de otimização quadrático convexo.

O algoritmo de otimização utilizado para o modelo SVM é o algoritmo de Otimização Mínima Sequencial (SMO), que foi criado especificamente para resolver problemas de otimização SVM, ajudando-o a resolver o treino da classificação binária linear.

A grande vantagem desta abordagem é que não precisamos de um programa para resolver os problemas quadráticos com dois multiplicadores de Lagrange, pois conseguimos resolver o problema analiticamente. Como consequência, não é necessário armazenar uma matriz enorme, o que pode causar problemas na memória da máquina.

Além disso, o SMO usa várias heurísticas para acelerar o cálculo [27]. O algoritmo SMO é composto de três partes:

- Uma heurística para escolher o primeiro multiplicador de Lagrange;
- Uma heurística para escolher o segundo multiplicador de Lagrange;
- O código para resolver o problema de otimização analiticamente para os dois multiplicadores escolhidos.

Mais detalhes sobre a otimização no contexto SVM podem ser consultados em [27].

3.3 Classificação Binária

Como já foi referido inúmeras vezes ao longo deste estudo, o *Support Vector Machine* foi inicialmente projetado para classificação binária. Assim, esta secção visa apresentar os diferentes tipos de classificação binária existentes.

3.3.1 Classificação Binária Linear

A classificação binária linear surgiu pela aplicação direta dos resultados fornecidos pela TAE. O classificador binário consiste em dividir um conjunto de dados em duas classes distintas, por isso, geralmente pode-se considerar um problema de classificação que contém n características, ou seja $x \in \mathbb{R}^n$ e l elementos de treino [28]. Assim, teremos o seguinte conjunto de treino, T :

$$T = \{(x_1, y_1), \dots, (x_l, y_l)\},$$

onde $x_i \in \mathbb{R}^n$, $y_i \in Y = \{-1, 1\}$, $i = 1, \dots, l$. É preciso encontrar uma função real $g(x)$ em \mathbb{R}^n , que descubra o valor de y para qualquer x pela função de decisão:

$$f(x) = \text{sgn}(g(x)),$$

onde $\text{sgn}(\cdot)$ é a função de sinal definida por:

$$\text{sgn}(a) = \begin{cases} 1, & a \geq 0; \\ -1, & a < 0. \end{cases} .$$

Assim, pode-se observar que para resolver um problema de classificação é necessário encontrar um critério para separar o espaço \mathbb{R}^n em duas regiões, de acordo com o conjunto de treino T [28].

Através de conhecimentos matemáticos já adquiridos sabemos que qualquer linha reta, $w \cdot x + b = 0$ (onde $(w \cdot x)$ é o produto escalar entre $w = (w_1, w_2)^T$ e $x = (x_1, x_2)^T$) consegue dividir um plano em duas regiões: $(w \cdot x) + b \geq 0$ e $(w \cdot x) + b < 0$ [28]. Ou seja, assim podemos determinar o valor de y correspondente a qualquer ponto de x através da função de decisão:

$$y = \text{sgn}((w \cdot x) + b)$$

A classificação binária linear pode ser estendida para problemas linearmente separáveis ou não separáveis.

Problemas Linearmente Separáveis

Um problema linearmente separável é um problema em que um conjunto de treino pode ser separado por um hiperplano, $y = f(x) = \text{sgn}((w \cdot x) + b)$ [28].

O método de margem rígida ou máxima pretende encontrar a melhor direção de w , mas para isso necessita de três linhas, a linha l_1 que é designada por hiperplano ou linha de separação, e as linhas l_2 e l_3 que são as linhas de suporte, como se pode observar na Figura 3.2.

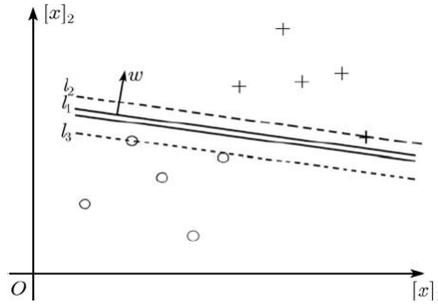


Figura 3.2: Representação da linha de separação e das linhas de suporte [28].

As linhas de suporte encontram-se em paralelo, e a distância entre elas designa-se por margem, assim quando a margem for máxima, $\frac{2}{\|w\|}$, o hiperplano é ideal, como se pode observar na Figura 3.3.

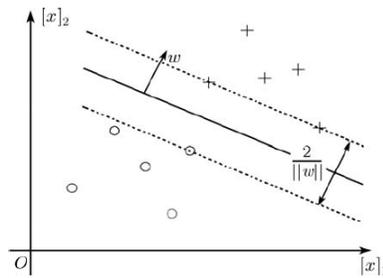


Figura 3.3: Representação de um hiperplano ideal [28].

Através do método de margem rígida obtém-se o seguinte problema de otimização para w e b , que corresponde ao cálculo da minimização da margem entre as duas linhas de suporte no espaço bidimensional:

$$\begin{aligned} & \min_{w,b} \frac{1}{2} \|w\|^2 \\ \text{s.t. } & y_i(w \cdot x_i) + b \geq 1, \quad i = 1, \dots, l \end{aligned} \tag{3.2}$$

O problema de otimização obtido é quadrático, cuja solução possui uma ampla e estabelecida teoria matemática [28]. Como a função objetivo é minimizada e os pontos que satisfazem as restrições formam um conjunto convexo, esse problema possui um único mínimo global. Problemas desse tipo podem ser solucionados com a introdução de uma função de *Lagrange*, que engloba as restrições à função objetivo, associadas a parâmetros

denominados multiplicadores de *Lagrange*, α .

$$L(w, b, \alpha) = \frac{1}{2} \|w\|^2 - \sum_{i=1}^l \alpha_i [y_i (w \cdot x_i) + b - 1],$$

onde $\alpha = (\alpha_1, \dots, \alpha_l)^T$ é o vetor multiplicador *Lagrange*.

A função *Lagrange* $L(w, b, \alpha)$ é uma função quadrática estritamente convexa de w , que alcança seu valor mínimo se w satisfizer:

$$\begin{aligned} \nabla_w L(w, b, \alpha) &= w - \sum_{i=1}^l \alpha_i y_i x_i = 0 \\ w &= \sum_{i=1}^l \alpha_i y_i x_i \end{aligned} \tag{3.3}$$

Assim, ao encontrar o valor de w através da função de Lagrange obtém-se o seguinte algoritmo para calcular o hiperplano:

Algorithm 1 Problemas Linearmente Separáveis

- 1: Obtenha o conjunto de treino $T = \{(x_1, y_1), \dots, (x_l, y_l)\}$, onde $x_i \in \mathbb{R}^n$, $y_i \in Y = \{-1, 1\}$, $i = 1, \dots, l$;
- 2: Construa e resolva o problema quadrático convexo:

$$\begin{aligned} \min_{\alpha} \quad & \frac{1}{2} \sum_{j=1}^l \sum_{i=1}^l \alpha_i \alpha_j y_i y_j (x_i \cdot x_j) - \sum_{j=1}^l \alpha_j \\ \text{s.t.} \quad & \sum_{i=1}^l \alpha_i y_i = 0, \\ & \alpha_i \geq 0, i = 1, \dots, l \end{aligned}$$

obtendo a solução $\alpha^* = (\alpha_1^*, \dots, \alpha_l^*)$;

- 3: Calcule $w^* = \sum_{i=1}^l \alpha_i^* y_i x_i$ e escolha um componente positivo de α^*, α_j^* e depois calcule b^* :

$$b^* = y_j - \sum_{i=1}^l \alpha_i^* y_i (x_i \cdot x_j);$$

- 4: Construa o hiperplano ideal $(w^* \cdot x) + b^* = 0$, e sua função de decisão $f(x) = \text{sgn}(g(x))$, onde:

$$g(x) = (w^* \cdot x) + b^* = \sum_{i=1}^l (y_i \alpha_i^* x_i \cdot x) + b^*.$$

Mais detalhes sobre o algoritmo SVM para problemas lineares separáveis podem ser consultados em [28].

Problema Linearmente Não Separável

Para um problema de classificação geral é possível que qualquer hiperplano não consiga separar corretamente todas as entradas positivas e negativas, mas caso se queira utilizar o hiperplano como separador é necessário ter em conta algumas estratégias a adotar.

Para este caso o problema linear separável de margem máxima estende-se para lidar com conjuntos de treino mais gerais, passando a designar-se por método de margem suave. Este método permite que alguns dados possam violar a restrição, $y_i((w \cdot x_i) + b) \geq 1$, através da introdução de variáveis de folga, ξ [28]. Essas variáveis relaxam as restrições impostas ao problema de otimização principal (3.2) que se torna:

$$y_i((w \cdot x_i) + b) \geq 1 - \xi_i, \quad i = 1, \dots, l.$$

Outra estratégia é violar as restrições o menos possível, ou seja, evitar aumentar muito ξ_i , impondo-lhes uma penalidade na função objetivo (3.2), por exemplo, pode-se adicionar um termo $\sum_i \xi_i$ à função objetivo (3.2), resultando na alteração do problema principal [28]:

$$\begin{aligned} \min_{w,b,\xi} \quad & \frac{1}{2} \|w\|^2 + C \sum_{i=1}^l \xi_i, \\ \text{s.t.} \quad & y_i(w \cdot x_i) + b \geq 1 - \xi_i, \quad i = 1, \dots, l, \\ & \xi_i \geq 0, \quad i = 1, \dots, l, \end{aligned} \tag{3.4}$$

onde $\xi_i = (\xi_1, \dots, \xi_l)^T$, e $C > 0$ é um parâmetro de penalidade.

Problemas desse tipo podem ser solucionados com a introdução de uma função de Lagrange, que engloba as restrições à função objetivo, associadas a parâmetros denominados multiplicadores de *Lagrange*, α_i [28]:

$$L(w, b, \alpha) = \frac{1}{2} \|w\|^2 + C \left[\sum_{i=1}^l \xi_i - \sum_{i=1}^l \alpha_i (y_i(w \cdot x_i) + b - 1 + \xi_i) - \sum_{i=1}^l \xi_i \right],$$

onde $\alpha = (\alpha_1, \dots, \alpha_l)^T$ é um multiplicador de *Lagrange*.

Assim, ao encontrar o valor de w através da função de *Lagrange* (3.3) obtém-se o algoritmo para se obter o hiperplano ideal:

Algorithm 2 Problemas Linearmente Não Separáveis

- 1: Obtenha o conjunto de treino $T = \{(x_1, y_1), \dots, (x_l, y_l)\}$, onde $x_i \in \mathbb{R}^n$, $y_i \in Y = \{-1, 1\}$, $i = 1, \dots, l$;
- 2: Escolha um parâmetro de penalidade apropriado $C > 0$;
- 3: Construa e resolva o problema quadrático convexo:

$$\begin{aligned} \min_{\alpha} \quad & \frac{1}{2} \sum_{i=1}^l \sum_{j=1}^l y_i y_j (x_i \cdot x_j) \alpha_i \alpha_j - \sum_{j=1}^l \alpha_j, \\ \text{s.t.} \quad & \sum_{i=1}^l \alpha_i y_i = 0, \\ & 0 \leq \alpha_i \leq C, \quad i = 1, \dots, l \end{aligned}$$

obtendo a solução $\alpha^* = (\alpha_1^*, \dots, \alpha_l^*)$;

- 4: Calcule b^* e escolha um componente positivo de $\alpha^*, \alpha_j^* \in (0, C)$, depois calcule:

$$b^* = y_j - \sum_{i=1}^l \alpha_i^* y_i (x_i \cdot x_j);$$

- 5: Construa a função de decisão $f(x) = \text{sgn}(g(x))$, onde:

$$g(x) = \sum_{i=1}^l (y_i \alpha_i^* x_i \cdot x) + b^*.$$

Mais detalhes sobre o algoritmo de SVM para problemas lineares não separáveis podem ser consultados em [28].

Como já se sabe encontrar os hiperplanos para resolver problemas lineares, é necessário compreender que existiram pontos sobrepostos a esse hiperplano, designando-os como vetores de suporte:

- Se for vetor de suporte x_i satisfaz $y_i g(x_i) = y_i((w^* \cdot x_i) + b^*) = 1$ ou $y_i g(x_i) = y_i((w^* \cdot x_i) + b^*) = -1$, ou seja, todos os vetores de suporte estão nos dois hiperplanos de suporte [28];
- Se não for um vetor suporte x_i satisfaz $y_i g(x_i) = y_i((w^* \cdot x_i) + b^*) \geq 1$ [28].

3.3.2 Classificação Binária Não Linear

Existem alguns tipos de problemas que a classificação linear não consegue resolver, como por exemplo o problema apresentado na Figura 3.4.

Na Figura 3.4 observa-se o hiperplano apropriado para a resolução de problemas não lineares. O hiperplano ideal parece-se com uma elipse centrada na origem no plano $([x]_1, O[x]_2)$ [28].

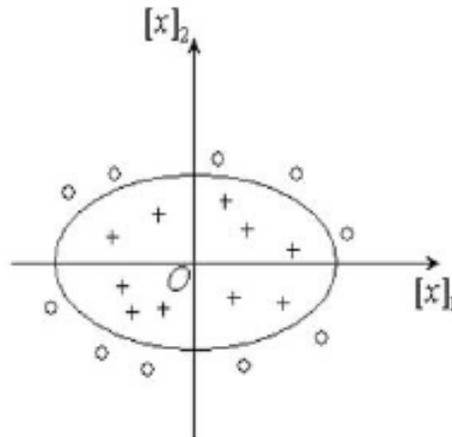


Figura 3.4: Exemplo de classificação binária não linear [28].

Para transformar um elipse em linha reta é necessário considerar um mapa $x = \Phi(x)$ dos pontos no plano $([x]_1, O[x]_2)$, para os pontos $([x]_1, [x]_2)$, no plano $([x]_1, O[x]_2)$ [28]:

$$\Phi : \begin{cases} [x]_1 = [x]_1^2, \\ [x]_2 = [x]_2^2, \end{cases} . \quad (3.5)$$

A Equação (3.5) mostra que para modificar uma classificação com separação não linear para linear, é somente necessário obter um mapa apropriado.

Esse mapa é calculado através da função *Kernel*, que tem como tarefa descobrir uma função de decisão e deduzir a saída correspondente y de qualquer nova entrada x . O ponto de partida para resolver esse problema é que entradas semelhantes devem ter as mesmas saídas. Portanto, é necessário medir a nova entrada e ver se x é mais parecido às entradas

positivas ou negativas; isto envolve o conceito de “similaridade” [28]:

$$K(x_i, x_j) = \Phi(x_i) \cdot \Phi(x_j) \quad (3.6)$$

Essa semelhança entre as duas entradas é medida pela distância, quanto menor a distância mais semelhante, são as entradas. A distância é decidida pelo *Kernel* selecionado através da Equação (3.6).

Para uma melhor precisão de resultados, a distância entre duas entradas x_i e x_j é definida pela distância de duas normas entre os dois vetores, $\|x_i - x_j\|^2$ [28]. Na Tabela 3.1 pode-se observar as funções *Kernel* mais comuns.

Tabela 3.1: Funções de *Kernel*.

	Tipo de Kernel	Função $K(x_i, x_j)$
	Linear	$K(x_i, x_j) = x_i \cdot x_j$
Polinomial	Homogénea	$K(x_i, x_j) = (x_i \cdot x_j)^d$
	Não Homogénea	$K(x_i, x_j) = ((x_i \cdot x_j) + 1)^d$
	Gaussiano	$K(x_i, x_j) = \frac{e^{-\ x_i - x_j\ ^2}}{\sigma^2} (x_i \cdot x_j)^d$

Para um problema n -dimensional geral é permitido que um mapa transforme um vetor n -dimensional x em outro vetor dimensional infinito x no espaço *Hilbert*, H . O espaço de *Hilbert* é uma generalização de um espaço vetorial real de dimensão finita distribuída por um produto interno, que não precisa estar restrita a um número finito de dimensões. Assim, o mapa pode ser expresso como [28]:

$$\Phi : \begin{cases} \mathbb{R}^n \rightarrow H, \\ x = ([x]_1, \dots, [x]_n)^T \rightarrow x = ([x]_1, [x]_2, \dots)^T = \Phi_x, \end{cases} .$$

A diferença entre o conjunto de treino para resolver problemas lineares e o conjunto de treino utilizado para resolver problemas não lineares, é a adição do mapa:

$$T_\Phi = \{x_1, y_1), \dots, (x_l, y_l)\}, \quad (3.7)$$

onde $x_i = \Phi(x_i) \in H, y_i \in Y \{-1, 1\}, i = 1, \dots, l$.

Para este tipo de problemas também se utiliza o método de margem suave, utilizando como base a função objetivo na Equação (3.4):

$$\begin{aligned} & \min_{w,b} \frac{1}{2} \|w\|^2 + C \sum_{i=1}^l \xi_i, \\ \text{s.t. } & y_i(w \cdot \Phi(x_i)) + b \geq 1 - \xi_i, \quad i = 1, \dots, l, \\ & \xi_i \geq 0, \quad i = 1, \dots, l, \end{aligned}$$

Como referido anteriormente trata-se de um problema quadrático convexo, logo introduzimos a função de *Lagrange*:

$$L(w, b, \xi, \alpha) = \frac{1}{2} \|w\|^2 + C \left[\sum_{i=1}^l \xi_i - \sum_{i=1}^l \alpha_i (y_i(w \cdot \Phi(x_i)) + b - 1 + \xi_i) - \sum_{i=1}^l \xi_i \right],$$

onde $\alpha = (\alpha_1, \dots, \alpha_l)^T$ é um multiplicador *Lagrange*.

Baseando-se na função de *Lagrange* (3.3), com variáveis de folga e funções *Kernel*, obtém-se $w = \sum_{i=1}^l \alpha_i y_i \Phi(x_i)$. O Algoritmo 3 apresenta o procedimento para encontrar o hiperplano ideal.

Para se obter os vetores de suporte do classificador binário não linear com *Kernel*, é necessário que a entrada x_i , esteja associada a um ponto de treino (x_i, y_i) , e que os componentes correspondentes, α_i^* e α^* , sejam diferentes de zero. Se $g(x)$ é definido pela equação (3.9):

- Se o vetor de suporte x_i correspondente a $\alpha_i \in (0, C)$ então satisfaz $y_i g(x_i) = 1$;
- Se o vetor de suporte x_i correspondente a $\alpha_i = C$ então satisfaz $y_i g(x_i) = -1$;
- Se o vetor x_i for não suportado satisfaz $y_i g(x_i) > 1$.

Mais detalhes sobre o processo podem ser consultados em [28]. A secção seguinte visa a apresentar o outro tipo de classificação existente.

Algorithm 3 Problemas Não Lineares

- 1: Obtenha o conjunto de treino original $T = \{(x_1, y_1), \dots, (x_l, y_l)\}$, onde $x_i \in \mathbb{R}^n$, $y_i \in Y = \{-1, 1\}$, $i = 1, \dots, l$;
- 2: Escolha um *Kernel* do espaço \mathbb{R}^n ao espaço *Hilbert*, e um parâmetro de penalidade apropriado $C > 0$, para se obter o conjunto de treino em 3.7;
- 3: Construa e resolva o problema quadrático convexo:

$$\begin{aligned} \min_{\alpha} \quad & \frac{1}{2} \sum_{i=1}^l \sum_{j=1}^l y_i y_j \alpha_i \alpha_j K(x_i \cdot x_j) - \sum_{j=1}^l \alpha_j, \\ \text{s.t.} \quad & \sum_{i=1}^l \alpha_i y_i = 0, 0 \leq \alpha_i \leq C, \quad i = 1, \dots, l \end{aligned}$$

obtendo a solução $\alpha^* = (\alpha_1^*, \dots, \alpha_l^*)$;

- 4: Calcule b^* e escolha um componente positivo de $\alpha^*, \alpha_j^* \in (0, C)$ e depois calcule:

$$b^* = y_j - \sum_{i=1}^l \sum_{j=1}^l y_i \alpha_i^* (\Phi(x_i) \cdot \Phi(x_j)); \quad (3.8)$$

- 5: Construa a função de decisão $f(x) = \text{sgn}(g(x))$, onde:

$$g(x) = \sum_{i=1}^l (y_i \alpha_i^* (\Phi(x_i) \cdot \Phi(x))) + b^*. \quad (3.9)$$

3.4 Classificação Multi-classe

Existem várias extensões do problema de classificação binária. A extensão mais utilizada para resolver problemas com mais do que duas classes é a de classificação de multi-classe que é formulada matematicamente da seguinte maneira:

$$T = \{(x_1, y_1), \dots, (x_l, y_l)\},$$

onde $x_i \in \mathbb{R}^n$, $y_i \in Y = \{1, 2, \dots, M\}$, $i = 1, \dots, l$.

É necessário encontrar uma função de decisão $f(x)$ em \mathbb{R}^n de modo a que o número de classe y para qualquer valor de x possa ser previsto pela função de decisão, $y = f(x)$. A solução é encontrar um critério para separar o espaço \mathbb{R}^n em regiões M de acordo com o conjunto de treino [28].

É necessário entender que o parâmetro M é escolhido por validação cruzada e interage fortemente com os parâmetros de *Kernel* [13].

Para resolver problemas de classificação multi-classe utiliza-se um método que pode ser usado através de vários métodos de esquema de votação baseados na combinação de muitas funções de decisão de classificação binária.

O ECOC (*Error-Correction-Output-Code*) representa uma estrutura bem-sucedida para lidar com problemas de categorização de várias classes com base na combinação de classificação binária [29].

Este método é dividido em dois estágios: codificação e decodificação [29]. O estágio de codificação cria uma palavra-código para cada classe, de um conjunto de classes, com base em diferentes problemas binários, que é obtida através de um conjunto de M classes a serem aprendidas, m diferentes bipartições (grupos de classe) são formadas e m problemas binários são treinados. A palavra-código é uma sequência de bits de um código que representa cada classe, onde cada bit identifica a associação da classe para um determinado classificador binário. O estágio de decodificação toma uma decisão de classificação para uma determinada amostra de teste com base no valor do código de saída.

As duas sub-subsecções seguintes apresentam dois métodos do ECOC utilizados para treinar o modelo, o um contra um ("*one versus one*"), e um contra todos ("*one versus the rest*"). O método mais usual e que é automaticamente utilizado através do *software Matlab* é o um contra um.

3.4.1 Um contra Um

É uma abordagem que tenta distinguir uma classe de outra. Considerando o conjunto de treino anterior, para cada par, $(i, j) \in \{(i, j) | i < j, i, j = 1, \dots, M\}$, constrói-se um problema de classificação binária para separar a i -ésima classe da j -ésima classe, resultando

em uma função $g^{i-j}(x)$ e na sua função de decisão correspondente [28]:

$$f^{i-j}(x) =: \begin{cases} i, & g^{i-j}(x) > 0, \\ j, & g^{i-j}(x) < 0, \end{cases} .$$

Para prever o rótulo da classe y para qualquer entrada x , utiliza-se a função de decisão do tipo $\frac{M(M-1)}{2}$, que obtém o rótulo através do qual a classe obtém um maior número de votos.

Um voto para uma determinada classe é definido como uma função de decisão que coloca a entrada x na classe [28].

3.4.2 Um contra todos

O método um contra todos, consiste em M problemas binários, em que o j -ésimo separa a j -ésima classe do resto, produzindo a função de decisão $f^j(x) = \text{sgn}(g^j(x))$, $j = 1, \dots, M$.

O próximo passo é fazer a classificação multi-classe de acordo com $g^1(x), \dots, g^M(x)$. Obviamente, se essas funções de decisão estiverem exatamente corretas, para qualquer entrada x existe um valor único de $g^j(x) > 0$ entre os valores de $g^1(x), \dots, g^M(x)$ e a entrada x pode ser prevista na classe J [28].

Uma dificuldade que pode aparecer através do algoritmo acima é que os conjuntos de dados nos problemas de classificação binária estão desequilibrados. A característica do desequilíbrio de classe geralmente leva a resultados insatisfatórios.

Uma solução para resolver este problema é selecionar parâmetros de penalidade para diferentes classes, penalidade maior para a classe minoritária e penalidade menor para a classe majoritária [28]. Por exemplo, para o reconhecimento de dez classes, precisamos resolver problemas de classificação binária que separem um dígito (classe positiva) dos restantes (classe negativa), portanto, a classe positiva seria a classe minoritária e a classe negativa seria a majoritária, para solucionar este desequilíbrio criávamos parâmetros de penalidade para cada uma delas. Assim o problema é resolvido com base no algoritmo de

classificação binária não linear com *Kernel*, equações (3.8) e (3.9).

O Algoritmo 4 apresenta o procedimento para encontrar o hiperplano ideal:

Algorithm 4 Um contra Todos

1: Insira o conjunto de treino:

$$T = \{(x_1, y_1), \dots, (x_l, y_l)\},$$

onde $x_i \in \mathbb{R}^n$, $y_i \in Y = \{1, 2, \dots, M\}$, $i = 1, \dots, l$.

2: Para $j = 1, \dots, M$, construa um conjunto de treino do j -ésimo problema binário com o conjunto de treino:

$$T^j = \{(x_1, y_1^j), \dots, (x_l, y_l^j)\},$$
$$y_i^j(x) =: \begin{cases} 1, & y^j = 1, \\ 0, & y^j \neq 1. \end{cases}$$

3: Construa a função de decisão:

$$f(x) = \arg \max_{j=1, \dots, M} ((w_j \cdot \Phi(x)) + b_j)$$

Capítulo 4

Caracterização da Base de Dados

O presente capítulo pretende apresentar a base de dados utilizada. De maneira a cumprir o objetivo principal deste estudo, o desenvolvimento de uma aplicação para apoiar os profissionais de saúde nos diagnósticos médicos, implementaram-se estratégias para treinar e testar um modelo que consiga identificar individualmente as doenças.

Com o objetivo a obter um sistema inteligente capaz de identificar doenças é necessário uma base de dados com parâmetros que caracterizem a doença, porque o desempenho depende em grande parte do número de parâmetros de entrada levados em consideração. Considerando o método SVM para treinar o modelo é fundamental que nos dados esteja identificado o diagnóstico final correspondente a esses resultados.

Neste trabalho foram consideradas três bases de dados referentes ao cancro da mama [30], a diferentes tipos de demência [23] e doença cardíaca [31].

4.1 Cancro da Mama

O cancro da mama tem origem na glândula mamária e é o tumor maligno mais frequente na mulher. A glândula mamária é formada por lóbulos, onde é produzido o leite, e por ductos, que o transportam até ao mamilo. Quando as células dos ductos ou dos lóbulos mamários se transformam em células malignas, surge um carcinoma ductal ou um carcinoma lobular [32].

Para este estudo utilizou-se uma base de dados que foi recolhida pela universidade de *Wisconsin*, que possui 569 amostras e 10 parâmetros relevantes que foram calculados a partir de uma biopsia aspirativa ou PAAF (punção aspirativa por agulha fina) [30].

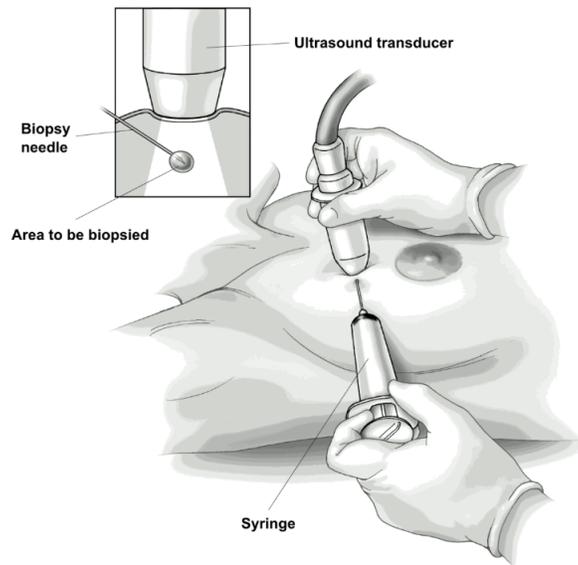


Figura 4.1: Biopsia aspirativa ou PAAF (punção aspirativa por agulha fina) [30]

A biópsia aspirativa é um procedimento de diagnóstico que consiste na remoção de uma amostra de células do tecido mamário suspeito para exame, é realizado através de uma agulha fina que é guiada pelo ultrassom, que permite recolher pequenas quantidades de células ou líquido para o estudo citológico, como se pode observar na Figura 4.1 [32].

O conjunto de dados é composto por 357 pacientes que foram diagnosticadas com nódulo benigno e 212 pacientes foram diagnosticadas com nódulo maligno. Os nove parâmetros de entrada utilizados neste estudo foram:

- **Raio:** Calculado através da média das distância do centro aos pontos da fronteira. Este parâmetro compreende valores entre 6.98 *mm* e 28.11 *mm*;
- **Textura:** Calculado através do desvio padrão dos valores da escala de cinza. Os valores variam entre 9.71 e 39.3;
- **Perímetro:** Calculado através da distância do centro aos pontos da fronteira, $2\pi \times r$. A variação de valores acontece entre 43.79 *mm* e 188.5 *mm*;

- **Área:** Calculado através de $\pi \times r^2$. Este parâmetro compreende valores entre 143.5 mm^2 e 2501 mm^2 ;
- **Suavidade:** Variação local nos comprimentos do raio. A variação dos valores deste parâmetro está compreendida em 0.0526 e 1158;
- **Compactação:** É calculado através de $\frac{\text{perímetro}^2}{\text{área}} - 1.0$. Os valores variam entre 0.0107 e 0.3;
- **Concavidade:** Severidade das porções côncavas do contorno. A variação de valores deste parâmetro está compreendida entre 0 e 0.5;
- **Côncava:** Medida da superfície mais profunda no centro do que na extremidade. Este parâmetro compreende valores entre 0 e 0.27;
- **Simetria:** Simetria dos orifícios dos ductos. Este parâmetro compreende valores entre 0 e 0.3;
- **Dimensão Fractal:** Valor da "aproximação da costa". A variação de valores deste parâmetro está compreendida entre 0 e 1.2.

Na identificação dos nódulos do cancro da mama pode-se obter dois diagnósticos, que são representados por 0 e 1, nódulo benigno e nódulo maligno, respetivamente.

4.2 Demência

Uma das maiores causas de mortalidade entre idosos é a demência [23]. A demência é uma síndrome caracterizada pelo declínio progressivo e global das funções cognitivas, na ausência de um comportamento agudo do estado de consciência, o que interfere nas atividades sociais e ocupacionais do indivíduo [23].

É o termo utilizado para descrever os sintomas de um grupo alargado de doenças que causam um declínio progressivo no funcionamento da pessoa. Assim, existem diversas

formas de doenças pertencentes a demência, mas a doença de Alzheimer é a causa mais comum de demência no idoso.

Esta base de dados consiste em uma coleção longitudinal de 150 indivíduos com cinco parâmetros diferentes recolhidos pelo Open Access Series of Imaging Studies (OASIS) [23]. Neste estudo 77 dos sujeitos representam o diagnóstico não demente, 58 dos indivíduos o demente, e os restantes 15 o convertidos, ou seja, não dementes no momento de sua visita inicial e subsequentemente caracterizados como demente em uma visita posterior. Os cinco parâmetros de entrada levados em consideração foram os seguintes:

- **Idade:** É o principal fator de risco para as doenças do nível de demência. Este parâmetro compreende idades dos 60 aos 93 anos;
- **Avaliação Clínica da Demência (Clinical Dementia Rating - CDR):** É um instrumento de estudo da demência que classifica os indivíduos por deficiência em cada um dos sete domínios: memória, orientação, julgamento e solução de problemas, função nos assuntos da comunidade, casa, hobbies e cuidados pessoais. Com base na fonte colateral e na entrevista ao paciente obtém-se uma pontuação, se CDR = 0 sem Alzheimer, se CDR = 0.5 Alzheimer muito leve, CDR = 1 Alzheimer leve, CDR = 2 Alzheimer moderado;
- **Exame do Estado Mental (Mini-Mental State Examination - MMSE):** É um teste neuropsicológico que avalia, de 0 a 30, diversas capacidades, como leitura, escrita, orientação e memória a curto prazo. Assim, uma pontuação maior que 24 pontos indica desempenho cognitivo saudável, entre 19 e 23 pontos corresponde a uma demência leve, entre 10 e 18 pontos representa uma demência moderada e inferior a 9 pontos representa demência grave;
- **MR Delay:** É o intervalo entre cada imagem retirada; Este parâmetro compreende valores que variam entre 0 e 2639;
- **Normalized Whole-Brain Volume (*n*-WBV):** é calculado através de uma imagem que é inicialmente segmentada para classificar o tecido cerebral como líquido

espinhal cerebral, substância cinzenta ou branca. O procedimento de segmentação atribuiu iterativamente *voxels* (matriz de elementos de volume que constitui um espaço tridimensional) a classes de tecidos. O *n*-WBV é então calculado como a quantidade de todos os *voxels* dentro de cada classe de tecido. Este parâmetro compreende valores que variam entre 0.6 e 0.8.

Pode-se obter três tipos de diagnóstico, em que o 0 representa não demente, o 1 de convertido e o 2 o diagnóstico como demente.

4.3 Doença Cardíaca

As doenças cardiovasculares são um conjunto de problemas que atingem o coração e os vasos sanguíneos, afetando, geralmente, mais homens do que mulheres com idades acima dos 50 anos. As doenças cardiovasculares estão associadas a um conjunto de fatores, que se designam habitualmente por fatores de risco. As doenças cardíacas são uma das maiores causas de mortalidade entre a população do mundo [31].

O conjunto de dados utilizado é do banco de dados de Cleveland, que consiste em uma coleção longitudinal de 303 indivíduos com treze parâmetros diferentes [31]. Neste conjunto de dados existem dois diagnósticos possíveis, onde 138 dos sujeitos tem ausência de doença cardíaca, e 165 dos indivíduos têm presença de doença cardíaca. Os treze parâmetros de entrada utilizados neste estudo foram:

- **Idade:** O fator de risco mais importante no desenvolvimento de doenças cardíacas, porque a pressão arterial tende a aumentar com a idade, para tal contribui o fato dos vasos sanguíneos perderem a elasticidade. Este parâmetro compreende idades dos 29 aos 77 anos;
- **Gênero:** Os homens têm maior risco de doença cardíaca do que as mulheres na pré-menopausa, mas depois da menopausa o risco é semelhante ao de um homem. Na base de dados, o 0 corresponde a mulher e o 1 ao homem.

- **Angina (dor no peito):** A angina é uma dor no peito ou um desconforto causado quando o músculo cardíaco não recebe sangue rico em oxigênio suficiente. O tipo de dor no peito que o paciente sente é exibido por: 1 (angina típica), 2 (angina atípica), 3 (dor não angina) e 4 (assintomático).
- **Pressão Arterial em repouso:** Um dos fatores de risco, porque a pressão alta pode danificar as artérias que alimentam o coração, e o aumento da pressão do sangue dentro das artérias também leva a que o coração tenha de fazer um maior esforço para bombear o sangue. O valor é exibido em mmHg, em que os valores para a sistólica (máxima) variam entre 120 mmHg e 180 mmHg e para a diastólica (mínima) variam entre 80 mmHg e 110 mmHg;
- **Colesterol:** Um alto nível do colesterol de lipoproteína de baixa densidade (LDL) é um fator de risco para doenças cardiovasculares. O valor é representado em mg/dl, que varia entre 130 mg/dl e 50 mg/dl, significando risco cardiovascular baixo e risco cardiovascular muito alto, respetivamente.
- **Açúcar no sangue em jejum:** Não produzir hormônio suficiente secretado pelo pâncreas (insulina) ou não responder adequadamente à insulina faz com que os níveis de açúcar no sangue do seu corpo aumentem, aumentando o risco de um ataque cardíaco. Compara-se o valor de açúcar no sangue em jejum de um indivíduo com 120 mg/dl, se o valor for maior que 120 mg/dl, então possui valor 1 (verdadeiro) caso contrário valor 0 (falso).
- **ECG repouso:** É um exame que deteta a atividade elétrica do coração, sendo o mais utilizado para avaliar arritmias cardíacas. Exibe resultados como: 0 (normal), 1 (com anormalidade da onda ST-T) e 2 (hipertrofia ventricular esquerda).
- **Frequência cardíaca máxima atingida:** O aumento do risco cardiovascular, está associado à aceleração de frequência cardíaca. Os valores variam entre 60 bpm e 100 bpm, que representam, frequência cardíaca baixa e frequência cardíaca elevada, respetivamente.

- **Angina induzida pelo exercício:** A dor ou desconforto associado à angina pode variar de leve a grave, com valor em caso afirmativo de 1 e em caso negativo 0.
- **Pico do segmento ST do exercício:** Um teste de esforço no ECG é considerado anormal quando há uma depressão do segmento ST horizontal ou inclinada para baixo ≥ 1 mm a 60-80 ms após o ponto J. O parâmetro compreende valores que variam entre 0 e 6.2 mm.
- **Segmento ST de pico de exercício:** A duração da depressão do segmento ST também é importante, pois a recuperação prolongada após o pico de stress é consistente com um teste de stress positivo no ECG, pode ser representado em caso de subida por 1, em caso de estabilidade/plano por 2 e em caso de descendente por 3.
- **Número de vasos principais:** Que vão de 0 até 4 que são coloridos por fluoroscopia, é exibido por um valor inteiro.
- **Talessemia:** É um grupo de doenças hereditárias resultantes de um desequilíbrio na produção de uma das quatro cadeias de aminoácidos que compõem a hemoglobina, em 3 (normal), 6 (defeito corrigido) e 7 (defeito reversível).

Para a doença cardíaca pode-se indicar dois tipos de diagnósticos diferentes, representados por 0 ou 1, ausência ou presença da doença cardíaca, respetivamente.

Capítulo 5

Análise e Discussão dos Resultados

Este capítulo pretende demonstrar como foi implementado o algoritmo de *Support Vector Machine* neste estudo e os resultados obtidos para cada conjunto de dados utilizado.

Primeiramente começa-se por recolher o conjunto de dados necessário para o estudo, e de seguida, divide-se o conjunto de dados em dois conjuntos de dados: treino e teste. O conjunto de dados de treino tem 85% dos dados numéricos e o conjunto de dados de teste tem 15% dos dados numéricos.

Para uma melhor compreensão da base de dados foram realizados gráficos que ilustram o comportamento dos dados em relação a dois parâmetros e apresentam o hiperplano calculado através do SVM.

Os modelos desenvolvidos para este estudo foram efetuados no *software Matlab* através de funções predefinidas. O SVM têm duas etapas necessárias, treinar e classificar:

- O treino é realizado, em 100 iterações através da função *fitcsvm* para o conjunto de dados do cancro da mama e doença cardíaca, onde é utilizado a classificação binária linear de problemas linearmente separáveis, com o algoritmo SMO e uma função *Kernel* linear, já para a demência utilizou-se a função *fitcecoc* com o método "um contra um" e uma função *Kernel* linear.
- A classificação pretende classificar os novos dados usando a previsão, para o modelo do cancro da mama e doença cardíaca utiliza-se a função *fitPosterior* e a *predict*,

para a demência é só necessário a função *predict*.

Os resultados são apresentados numa tabela que compara o diagnóstico real com os diagnósticos previstos, indicando a probabilidade de pertencer a cada classe. Para perceber se os modelos alcançaram bons resultados, calculou-se a precisão dos mesmos e a percentagem de falsos negativos e positivos.

Através da utilização da aplicação *Classification Learner* do *software Matlab* e da análise estrutural dos dados obteve-se que este procedimento alcançou a melhor precisão de resultados para as três bases de dados.

5.1 Cancro da Mama

Na Tabela 5.1 é possível observar uma amostra da base de dados usada para o treino e teste do modelo onde a primeira coluna representa o índice das pacientes (ID) e as colunas seguintes possuem informação relativa ao raio (R), textura (Te), perímetro (P), área (A), suavidade (Su), compactação (Cp), concavidade (Ca), côncava (Co), simetria (Si), dimensão fractal (DF) e o diagnóstico. Os restantes valores da base de dados podem ser observados no Anexo 1A.

Tabela 5.1: Cancro da Mama - Amostra da base de dados utilizada.

ID	R	Te	P	A	Su	Cp	Ca	Co	Si	DF	Diag
1	13.54	14.36	87.46	566.3	0.0978	0.0813	0.0666	0.0478	0.1885	0.0577	0
2	13.08	15.71	85.63	520.0	0.1075	0.1270	0.0457	0.0311	0.1967	0.0681	0
3	9.50	12.44	60.34	273.9	0.1024	0.0649	0.0296	0.0208	0.1815	0.0691	0
...
357	7.76	24.54	47.92	181.0	0.0526	0.0436	0.0000	0.0000	0.1587	0.0588	0
358	17.99	10.38	122.8	1001.0	0.1184	0.2776	0.3001	0.1471	0.2419	0.0787	1
...
567	20.13	28.25	131.2	1261.0	0.0978	0.1034	0.1440	0.0979	0.1752	0.0553	1
568	16.60	28.08	108.3	858.1	0.0846	0.1023	0.0925	0.0530	0.1590	0.0565	1
569	20.60	29.33	140.1	1265.0	0.1178	0.2770	0.3514	0.1520	0.2397	0.0702	1

Posteriormente à divisão dos dados, na qual, 484 pertencem ao conjunto de treino e 85 ao conjunto de teste, ilustrou-se o padrão da base de dados de treino numa relação

entre dois parâmetros e o respectivo hiperplano (calculado através do SVM) para que se consiga observar as regiões de caracterização dos parâmetros.

Afirma-se que para esta doença a representação dos dados graficamente torna-se um pouco difícil devido a existirem relações de parâmetros onde os dados tendem a formar um polinómio, e não uma linha reta (hiperplano), o que torna árduo a separação desses dados. Na figura seguinte pode-se observar um dos gráficos que apresenta o hiperplano ideal, o parâmetro 1 (raio) versus parâmetro 2 (textura), os restantes gráficos poderão ser observados no Anexo 1B.

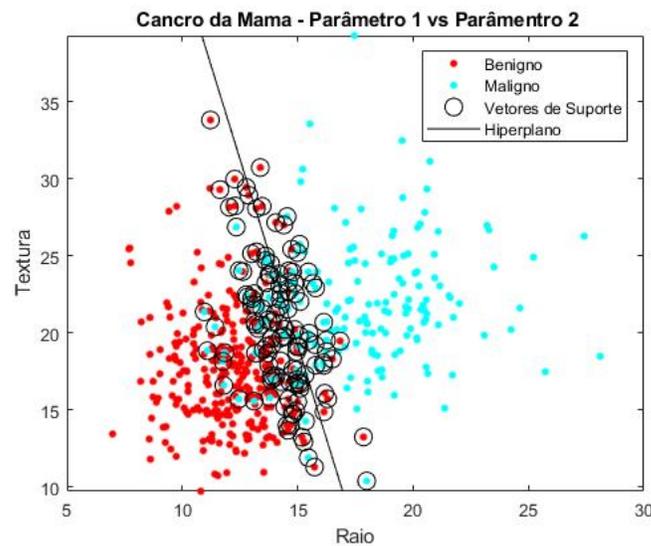


Figura 5.1: Cancro da Mama - Gráfico representativo da relação entre o parâmetro 1 (Raio) e o parâmetro 2 (Textura).

Analisando a Figura 5.1, observa-se que há uma separação dos dados, existindo bastantes pontos sobrepostos ao hiperplano, e alguns pontos representativos da classe maligna que o ultrapassam, encontrando-se na região do nódulo benigno, isto poderá significar alguma dificuldade de previsão no diagnóstico.

Sucessivamente, o modelo foi testado com o conjunto de dados de teste, alcançando-se a Tabela 5.2 que apresenta, alguns resultados da comparação entre os diagnósticos reais e os previstos, indicando a probabilidade daquele dado corresponder a cada classe. Os restantes poderão ser observados na Tabela 4 do Anexo 1C.

Tabela 5.2: Cancro da Mama - Amostra dos resultados de previsão obtidos.

ID Paciente	Diag Real	Diag Previsto	Probabilidade (%)	
			Benigno (0)	Maligno (1)
7	0	0	86.32	13.68
21	0	0	99.90	0.10
29	0	1	29.65	70.35
43	0	0	99.97	0.03
...
388	1	1	0.02	99.98
393	1	0	66.43	33.57
404	1	1	3.51	96.49
406	1	1	25.84	74.16

Considerando todos os resultados, e calculando a taxa de acerto, 90.57% e 75.00%, do nódulo benigno e nódulo maligno, respetivamente, verifica-se que o modelo tem alguma facilidade em identificar a classe benigna. Isto ocorre porque a maior parte dos pontos representantes dessa classe encontram-se bem divididos e são poucos os pontos que ultrapassam o hiperplano, facilitando a sua previsão.

De seguida, na Tabela 5.3 é apresentada uma tabela que relaciona os diagnósticos reais com os obtidos pelo modelo, com o objetivo de mostrar o desempenho do modelo ao analisar a quantidade de acertos e de erros para cada classe.

Tabela 5.3: Cancro da Mama - Relação de diagnósticos.

		Diagnóstico previsto	
		Benigno (0)	Maligno (1)
Diagnóstico real	Benigno (0)	48	5
	Maligno (1)	8	24

Averiguando a Tabela 5.3 pode-se indicar que o modelo obtém as seguintes taxas: 56.47% verdadeiros positivos, 28.24% verdadeiros negativos, 9.41% falsos negativos e 5.88% falsos positivos. Verifica-se que este modelo obtém resultados que requerem alguma atenção, porque apresenta uma maior percentagem de diagnóstico negativo com presença de doença.

A precisão é calculada através da seguinte Equação:

$$Precisão = \frac{\sum Verdadeiros\ Positivos + \sum Verdadeiros\ Negativos}{\sum Diagnósticos\ Previstos} \times 100 \quad (5.1)$$

Através da Tabela 5.3 observou-se que existiram 48 verdadeiros positivos e 24 verdadeiros negativos, substituindo na Equação 5.1 e dividir por 85 obtém-se uma precisão total de 84.71%.

Para uma melhor visualização dos resultados da Tabela 5.2, foram realizados dois gráficos, onde, o gráfico dos diagnósticos reais representa o comportamento dos dados da coluna "Diagnóstico Real", e o gráfico dos diagnósticos previstos representa o comportamento dos dados da coluna "Diagnóstico Previsto".

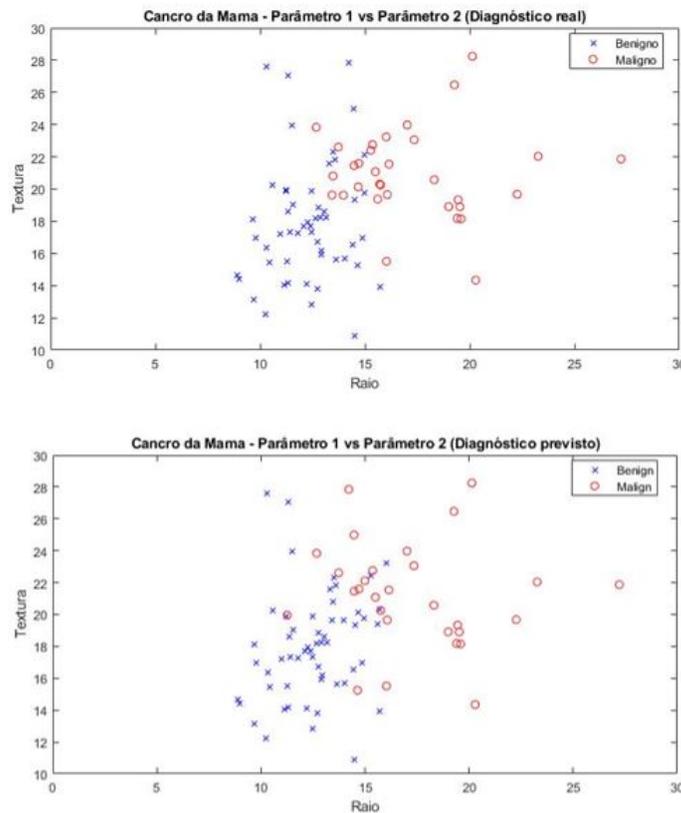


Figura 5.2: Cancro da Mama - Gráfico com a comparação dos diagnóstico reais com os previstos do parâmetro 1 (raio) e o parâmetro 2 (textura).

Na Figura 5.2, pode-se observar que existem pontos que num gráfico têm um diagnóstico e no outro gráfico um diagnóstico diferentes, um exemplo é o ponto (14.64, 15.24) referente ao paciente n°29.

5.2 Demência

Na Tabela 5.4 é possível observar uma amostra da base de dados usada para o treino e teste do modelo onde a primeira coluna representa o índice do paciente e as colunas seguintes possuem informação relativa a idade, CDR, MMSE, MR Delay, n -WBV e diagnóstico. Os restantes valores da base de dados podem ser observados no Anexo 2A.

Tabela 5.4: Demência - Amostra da base de dados utilizada.

ID Paciente	Idade	CDR	MMSE	MR Delay	n -WBV	Diagnóstico
1	87	0.0	27	0	0.7	0
2	88	0.0	28	0	0.7	0
3	90	0.0	27	538	0.7	0
...
78	87	0.0	30	0	0.7	1
79	88	0.0	29	489	0.7	1
80	92	0.5	27	1933	0.7	1
...
148	74	0.5	26	0	0.7	2
149	73	0.5	23	0	0.7	2
150	86	0.5	26	2297	0.7	2

Posteriormente à divisão dos dados, na qual 128 pertencem ao conjunto de treino e 22 ao conjunto de teste, ilustrou-se o padrão da base de dados de treino numa relação entre dois parâmetros e o respectivo hiperplano (calculado através do SVM) para que se consiga observar as regiões de caracterização dos parâmetros e os vetores de suporte.

Como foi referido no Capítulo 3, o hiperplano para classificações de multi-classe pode ser calculado pelo método um contra um, o que irá resultar num hiperplano para cada relação entre classes, ou seja, um hiperplano que separe a classe não demente com a classe convertida, a não demente com a demente, e por último a convertida com a demente.

Após a divisão dos dados, na qual, 128 pertencem ao conjunto de treino e 22 ao

conjunto de teste, representou-se graficamente o conjunto de treino para que se consiga observar qual o comportamento dos mesmo separando-os em classes através de um hiperplano com o intuito de verificar se existe uma separação ideal entre elas. Os dados irão ser visualizados através do comportamento em cada parâmetro, ou seja, fez-se parâmetro contra parâmetro para encontrar os que apresentam um hiperplano ideal, como se pode verificar na legenda da Figura 5.3.

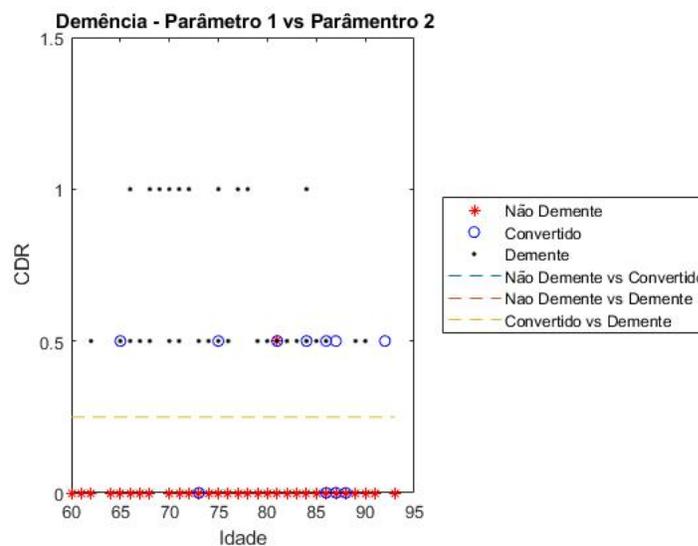


Figura 5.3: Demência - Gráfico representativo da relação entre o parâmetro 1 (Idade) e o parâmetro 2 (CDR).

Na Figura 5.3 pode-se observar que as classes não demente e demente são dois clusters diferentes, agrupados no gráfico, mas a classe convertida não, porque tem pontos entre a classe não demente e demente, logo não se consegue agrupar sozinha, assim através da análise do gráfico pode-se referir que não existirá previsões de diagnósticos da classe convertida corretos.

No gráfico deveriam de estar representados três hiperplanos, mas apenas o hiperplano "Convertido vs Demente" é observado, isto deve-se ao fato de os restantes encontrarem-se sobrepostos. Porém, o que fazia mais sentido era o hiperplano "Não Demente vs Demente" porque são as únicas classes agrupadas e que se conseguem separar. Os restantes gráficos poderão ser observados no Anexo 2B.

De seguida, o modelo foi testado com o conjunto de dados de teste, obtendo-se a Tabela 5.5 que apresenta, alguns resultados da comparação entre os diagnósticos reais e os previstos, indicando a probabilidade daquele dado corresponder a cada classe. Os restantes poderão ser observados na Tabela 7 do Anexo 2C.

Tabela 5.5: Demência - Amostra dos resultados obtidos.

ID	Diag. Real	Diag. Previsto	Probabilidade (%)		
			Não Demente (0)	Convertido (1)	Demente (2)
20	0	0	92.24	6.68	1.08
23	0	0	91.80	7.72	1.28
24	0	0	92.27	6.61	1.12
...
83	1	0	92.61	5.75	1.64
87	1	0	92.77	5.34	1.89
91	1	0	92.94	4.90	2.16
...
128	2	2	0.92	3.61	95.47
130	2	2	1.39	6.08	92.53
133	2	2	3.02	16.62	80.36

Analisando todos os resultados, verifica-se que não existe diagnósticos previstos corretamente relativamente a classe convertida, isto porque no conjunto de dados existem poucos dados representativos da mesma, dificultando assim a sua previsão.

De seguida, na Tabela 5.6 é apresentada uma tabela que relaciona os diagnósticos reais com os obtidos pelo modelo, demonstrando o desempenho do modelo ao analisar a quantidade de acertos e de erros para cada classe.

Tabela 5.6: Demência - Relação entre resultados.

		Diagnóstico previsto		
		Não Demente (0)	Convertido (1)	Demente (2)
Diagnóstico real	Não Demente (0)	15	0	0
	Convertido (1)	3	0	0
	Demente (2)	0	0	4

Averiguando a Tabela 5.6 pode-se indicar que o modelo obtém as seguintes taxas de: 68.18% verdadeiros positivos, 18.18% verdadeiros negativos, 6.82% falsos negativos e 6.82 falsos positivos. Verifica-se que este modelo tem uma maior facilidade a prever a classe não

demente, porém os resultados são obtidos de maneira diferentes que os modelos passados. Neste caso, o modelo requer muita atenção por parte dos profissionais de saúde visto que a porcentagem de obter um diagnóstico negativo com presença de doença ou um diagnóstico positivo e ausência de doença é igual.

Para uma melhor visualização dos resultados da Tabela 5.6, foram realizados dois gráficos, onde, o gráfico dos diagnósticos reais representa o comportamento dos dados da coluna "Diagnóstico Real", e o gráfico dos diagnósticos previstos representa o comportamento dos dados da coluna "Diagnóstico Previsto".

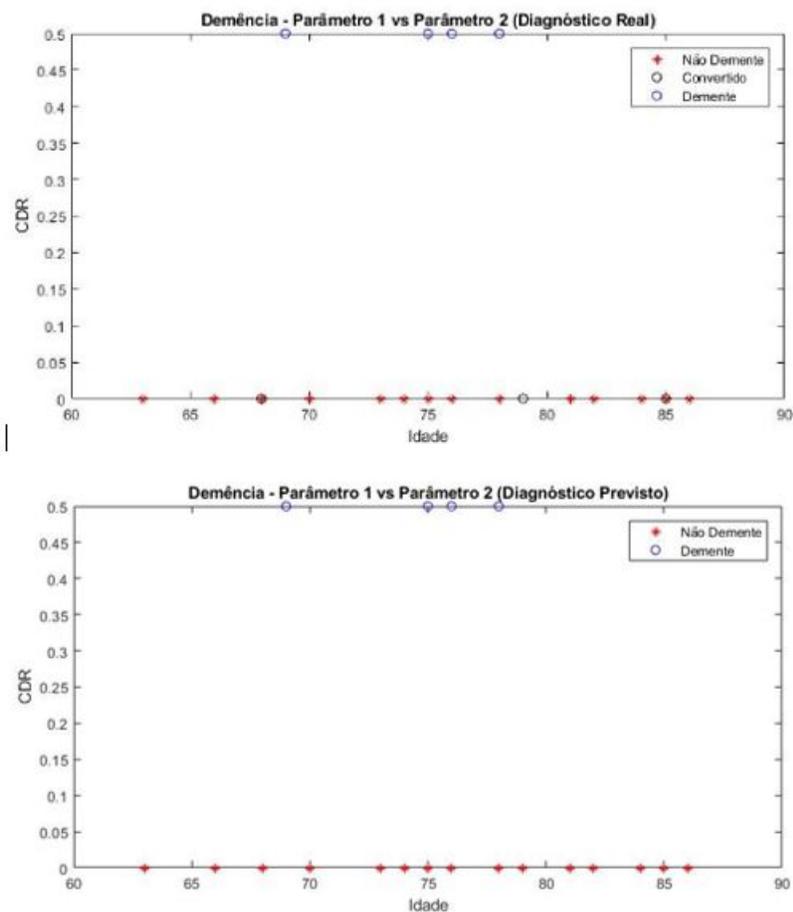


Figura 5.4: Demência - Gráfico com a comparação dos diagnósticos reais com os previstos do parâmetro 1 (Idade) e o parâmetro 2 (CDR).

Na Figura 5.4, pode-se concluir que existem casos que não foram previstos corretamente, como por exemplo no gráfico do diagnóstico real é notável que existem pontos

representativos da classe convertida e no de diagnóstico previsto isso não acontece, um exemplo é o ponto (79, 0), referente ao paciente nº87. Proveniente disso, existem mais pontos da classe demente no gráfico de diagnóstico previsto.

Por fim é efetuado o cálculo de precisão do modelo através da Equação (5.1) substituindo o verdadeiro positivo por 15, o verdadeiro negativo por 4 e o número de diagnósticos previstos por 22, e obtém-se uma precisão total de 86.36%.

5.3 Doença Cardíaca

Na Tabela 5.7 é possível observar uma amostra da base de dados usada para o treino e teste do modelo onde a primeira coluna representa o índice do paciente (ID) e as colunas seguintes possuem informação relativa a idade (I), género (G), angina (A), pressão arterial (PA), colesterol (Co), açúcar em jejum (Açu), ECG em repouso (ECG), frequência cardíaca (FC), angina induzida (AI), pico do segmento ST do exercício (P. ST), segmento ST de pico de exercício (ST), número de vasos principais (Va), talassemia (Tal) e diagnóstico (Diag). Os restantes valores da base de dados podem ser observados no Anexo 3A.

Tabela 5.7: Doença Cardíaca - Amostra da base de dados utilizada.

ID	I	G	Ang	PA	Co	Açu	ECG	FC	AI	P. ST	ST	Va	Tal	Diag
1	67	1	0	160	286	0	0	108	1	1.5	1	3	2	0
2	67	1	0	120	229	0	0	129	1	2.6	1	2	3	0
3	62	0	0	140	268	0	0	160	0	3.6	0	2	2	0
...
138	57	0	1	130	236	0	0	174	0	0.0	1	1	2	0
139	63	1	3	145	233	1	0	150	0	2.3	0	0	1	1
...
301	41	1	1	120	157	0	1	182	0	0.0	2	0	2	1
302	38	1	2	138	175	0	1	173	0	0.0	2	4	2	1
303	38	1	2	138	175	0	1	173	0	0.0	2	4	2	1

Posteriormente à divisão dos dados, na qual, 258 pertencem ao conjunto de treino e 45 ao conjunto de teste, ilustrou-se o padrão da base de dados de treino numa relação entre dois parâmetros e o respectivo hiperplano (calculado através do SVM) para que se consiga observar as regiões de caracterização dos parâmetros.

De todos os gráficos obtidos, nove apresentam um hiperplano ideal, sendo apresentado na Figura 5.5 o gráfico que relaciona os dois principais fatores de risco para doenças cardíacas. Os restantes gráficos podem ser observados no anexo 3B.

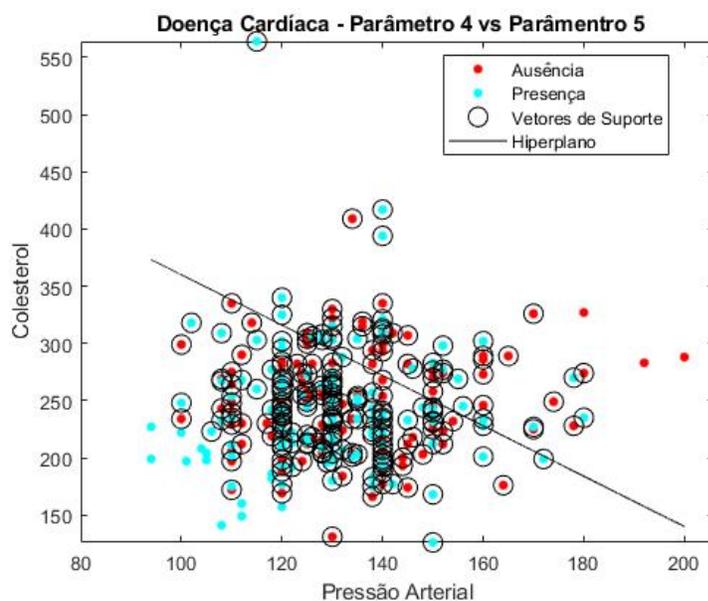


Figura 5.5: Doença Cardíaca - Gráfico representativo da relação entre o parâmetro 4 (Pressão Arterial) e o parâmetro 5 (Colesterol)

Visualizando a Figura 5.5 pode-se constatar que devido ao fato de os valores serem bastante lineares e próximos faz com que não esteja nítido a presença de dois clusters, porém consegue-se obter um hiperplano ideal que consegue separar as classes. Observa-se, também, que a maior parte dos dados de ausência estão na zona de presença, assim, expõe-se que existirá uma dificuldade acrescida na previsão de diagnóstico, uma vez que os valores são difíceis de identificar a qual classe corresponde.

Posteriormente, o modelo foi testado com o conjunto de dados de teste, obtendo-se a Tabela 5.8 que apresenta, uma amostra da comparação do diagnóstico real com o previsto, indicando a probabilidade daquele dado corresponder a cada classe. Os restantes poderão ser observados no Anexo 3C.

Analisando todos os resultados, Tabela 2 do Anexo 1C, é possível verificar através do cálculo da taxa de acerto, ausência de 68.75% e para a presença de 89.66%, que o modelo

Tabela 5.8: Doença Cardíaca - Amostra dos resultados de sugestão de diagnóstico obtidos.

Paciente	Diagnóstico real	Diagnóstico previsto	Probabilidade (%)	
			Ausência (0)	Presença (1)
4	0	0	91.89	8.12
6	0	1	26.04	73.96
11	0	0	99.20	0.80
12	0	0	96.17	3.83
...
223	1	1	27.27	72.73
234	1	0	90.92	9.08
242	1	1	23.32	76.68
246	1	1	48.32	51.68

tem uma maior dificuldade a identificar a ausência da doença cardíaca. Justifica-se a dificuldade do modelo para identificar a ausência da doença, através da grande parte dos pontos representativos da mesma se encontrarem na zona da classe de presença.

De seguida, é apresentada a matriz de confusão com o intuito de avaliar o desempenho do modelo, agrupando os resultados da Tabela 5.8 consoante os seus acertos ou erros para ambas as classes.

Tabela 5.9: Doença Cardíaca - Matriz de confusão dos resultados obtidos.

		Diagnóstico previsto	
		Ausência (0)	Presença (1)
Diagnóstico real	Ausência (0)	11	5
	Presença (1)	3	26

Explorando a Tabela 5.9 pode-se indicar que o modelo obtém as seguintes taxas: 57.78% verdadeiros negativos, 24.44% verdadeiros positivos, 11.11% falsos positivos e 6.67% falsos negativos. Verificando que o modelo identifica melhor a presença de doença cardíaca, e mais resultados com diagnóstico positivo e ausência de doença do que diagnóstico negativo com doença presente.

Para uma melhor visualização dos resultados da Tabela 5.8, foram realizados dois gráficos, onde, o gráfico dos diagnósticos reais representa o comportamento dos dados da coluna "Diagnóstico Real", e o gráfico dos diagnósticos previstos representa o comportamento dos dados da coluna "Diagnóstico Previsto".

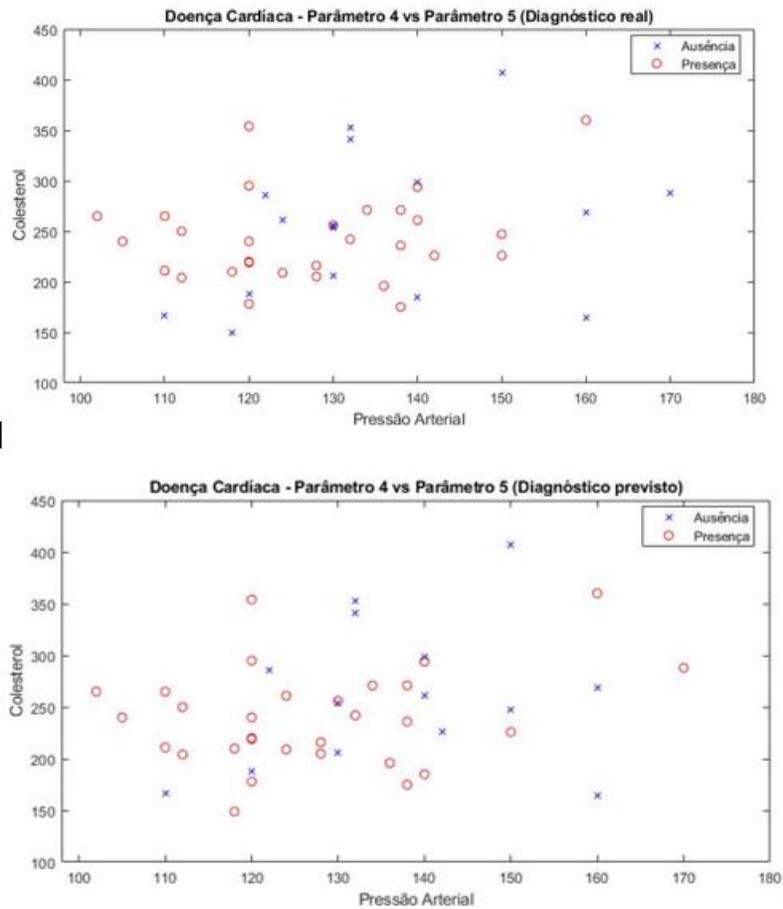


Figura 5.6: Doença Cardíaca - Gráfico com a comparação dos diagnósticos reais com os previstos do parâmetro 4 (Pressão Arterial) e o parâmetro 5 (Colesterol).

Através da Figura 5.6, pode-se observar que existem pontos que num gráfico têm um diagnóstico e no outro gráfico um diagnóstico diferentes, um exemplo é o ponto (170, 288) referente ao sujeito n°64.

Por fim, é efetuado o cálculo de precisão do modelo através da Equação (5.1) substituindo por 11 verdadeiros positivos, 26 verdadeiros negativos e 45 pelo número total diagnósticos previstos, obtém-se uma precisão total de 82.22%.

Capítulo 6

Aplicação ProSmartHealth

Este capítulo visa apresentar a aplicação ProSmartHealth. Posteriormente à previsão do diagnóstico procedeu-se ao desenvolvimento da aplicação ProSmartHealth, no *App-Designer*, onde se modificou os modelos para treinar todo o conjunto de dados e testar apenas os dados inseridos no questionário, com o propósito de apoiar os profissionais de saúde no diagnóstico.

Ao iniciar a ProSmartHealth aparecerá uma página inicial com as doenças possíveis de obter uma sugestão de diagnóstico, como demonstra a Figura 6.1.

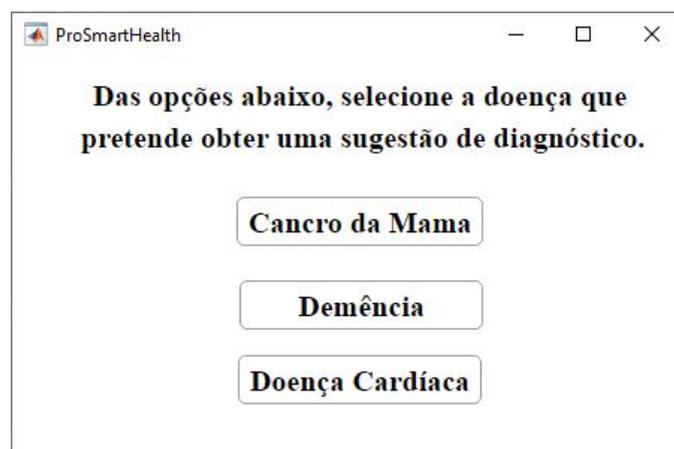
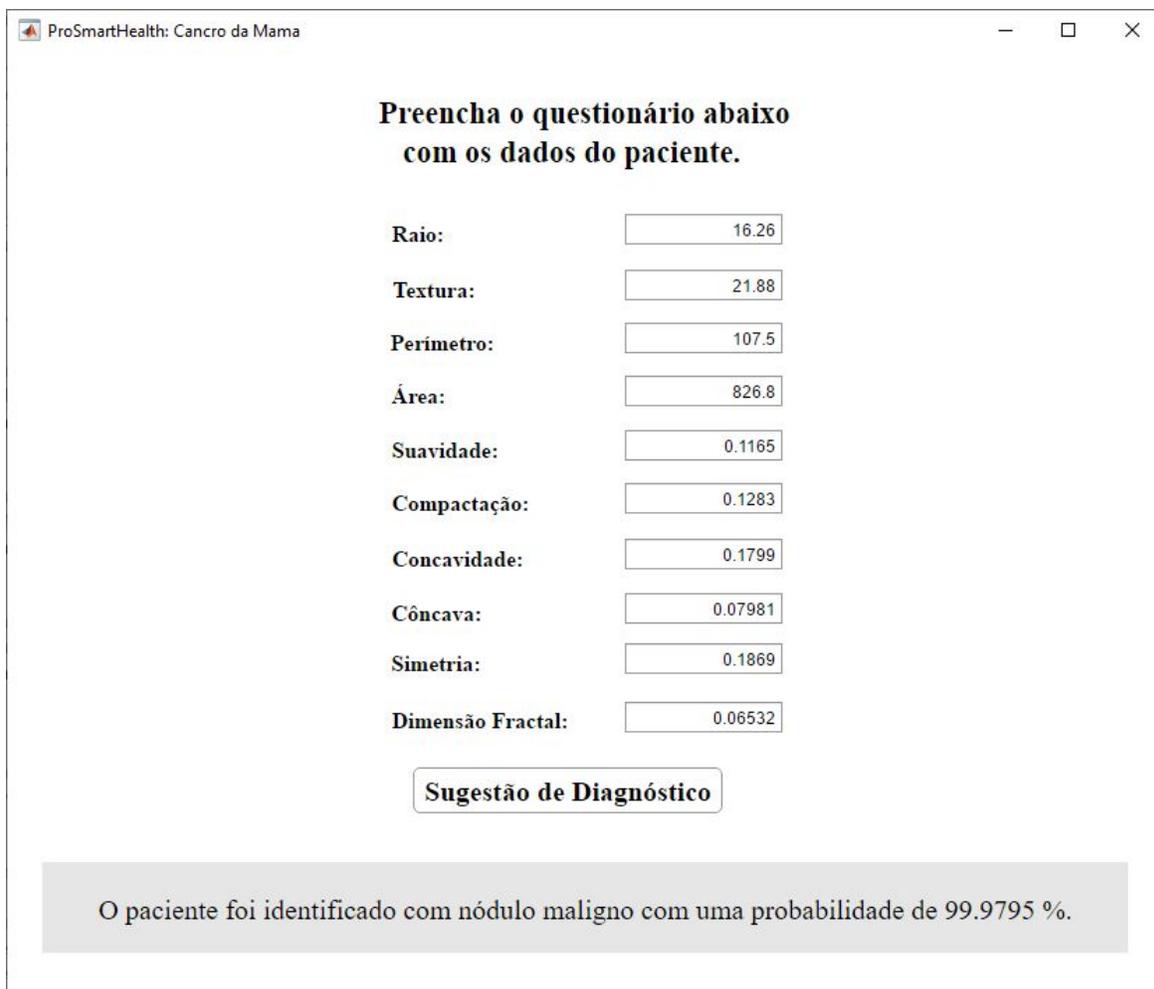


Figura 6.1: Página inicial da aplicação ProSmartHealth.

Na página inicial, o profissional de saúde terá de selecionar a doença que pretende obter uma sugestão de diagnóstico.

Na circunstância do profissional de saúde selecionar o cancro da mama, aparecerá uma janela com um questionário composto por dez questões, correspondentes aos parâmetros utilizados, para serem respondidas com os resultados médicos da biopsia aspirativa efetuada pela paciente, futuramente estes resultados poderão migrar automaticamente. Assim que os parâmetros estejam todos preenchidos, a aplicação sugere um diagnóstico e aparecerá uma mensagem que nos indica uma sugestão de diagnóstico final do paciente, como se pode observar na Figura 6.2 referente ao paciente n°331.



Raio:	16.26
Textura:	21.88
Perimetro:	107.5
Área:	826.8
Suavidade:	0.1165
Compactação:	0.1283
Concavidade:	0.1799
Côncava:	0.07981
Simetria:	0.1869
Dimensão Fractal:	0.06532

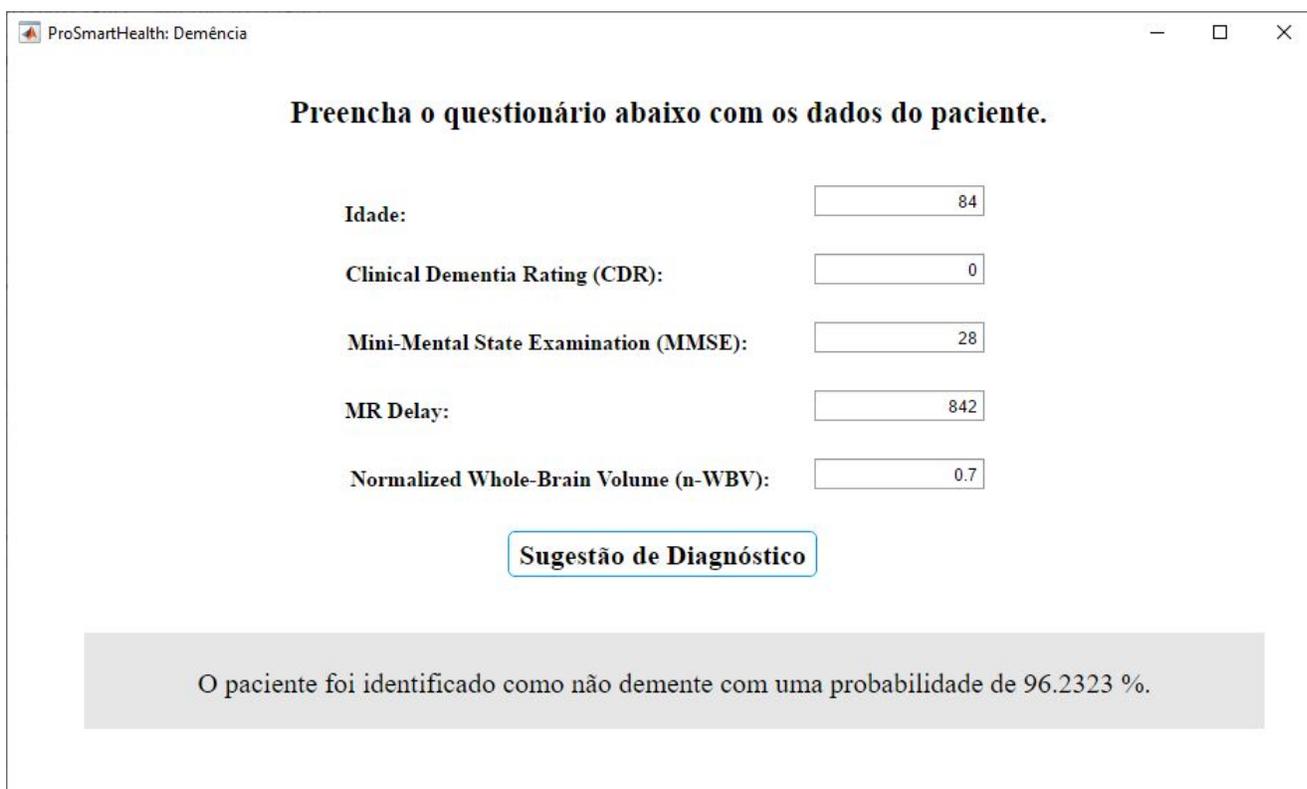
Sugestão de Diagnóstico

O paciente foi identificado com nódulo maligno com uma probabilidade de 99.9795 %.

Figura 6.2: Questionário do cancro da mama preenchido.

Na Figura 6.2 pode-se observar a sugestão de diagnóstico correspondente a paciente N°331, e ao comparar com o resultado dessa paciente na Tabela 2 do Anexo 1C conclui-se que a aplicação faz uma sugestão de diagnóstico corretamente.

Caso, o profissional de saúde opte pela demência aparece um questionário com cinco questões, correspondente aos parâmetros utilizados, que o profissional de saúde terá de preencher com os resultados do paciente. Assim que o profissional de saúde terminar de preencher, o sistema sugere um diagnóstico, onde lhe aparece uma mensagem com uma probabilidade de sugestão de diagnóstico correspondente aos dados preenchidos. Na Figura 6.3 pode-se observar o questionário preenchido com os dados do paciente nº40.



The screenshot shows a window titled "ProSmartHealth: Demência". The main heading is "Preencha o questionário abaixo com os dados do paciente." Below this, there are five input fields with their respective labels and values:

Idade:	84
Clinical Dementia Rating (CDR):	0
Mini-Mental State Examination (MMSE):	28
MR Delay:	842
Normalized Whole-Brain Volume (n-WBV):	0.7

Below the form is a button labeled "Sugestão de Diagnóstico". At the bottom, a grey box contains the message: "O paciente foi identificado como não demente com uma probabilidade de 96.2323 %."

Figura 6.3: Questionário da Demência preenchido.

Na Figura 6.3 pode-se observar a sugestão de diagnóstico calculada pela aplicação para o sujeito nº40, que ao comparar com o resultado desse mesmo paciente na Tabela 4 do Anexo 2C verifica-se a mesma sugestão de diagnóstico.

Por fim, o profissional de saúde ao optar pela doença cardíaca, aparecerá uma janela com um questionário composto por treze questões, referentes aos parâmetros utilizados, para serem respondidas com os resultados do paciente. Assim que os parâmetros estejam todos preenchidos o sistema sugere um diagnóstico e aparece uma mensagem que nos

indica a sugestão de diagnóstico final do sujeito, como se pode observar na Figura 6.4 referente ao paciente n°242.

Item	Valor	Explicação
Idade:	63	
Gênero:	0	0 para mulher e 1 para homem
Angina (Dor no peito):	1	1 (angina típica), 2 (angina atípica), 3 (dor não angina) e 4 (assintótico)
Pressão Arterial em Repouso:	140	
Colesterol:	195	
Açúcar em jejum:	0	se o valor for maior que 120 mg/dl, então 1 (verdadeiro) e 0 (falso)
ECG em repouso:	1	0 (normal), 1 (com anormalidade da onda ST-T) e 2 (hipertrofia ventricular esquerda)
Frequência Cardíaca máxima atingida:	179	
Angina induzida por exercício:	0	0 para sim teve dor no peito e 1 para não teve dor no peito
Pico do segmento ST do exercício:	0	
Segmento ST de pico de exercício:	2	0 (subida), 1 (plano) e 2 (descendente).
Número de vasos principais:	2	
Talassemia:	2	1 (normal), 2 (defeito corrigido) e 3 (defeito reversível)

Sugestão de Diagnóstico

O paciente foi identificado com presença de Doença Cardíaca com uma probabilidade de 88.8856 %.

Figura 6.4: Questionário da Doença Cardíaca preenchido.

Na Figura 6.4 pode-se observar o resultado para o sujeito n°242, que ao comparar com a Tabela 5.8 observa-se que a sugestão de diagnóstico está correta.

Capítulo 7

Conclusões e Trabalhos Futuros

A aplicação do algoritmo *Support Vector Machine* comprova que é possível identificar a presença ou não de doenças, apenas necessita de uma base de dados com dados rotulados (classes) e com parâmetros correspondentes à doença.

O *Support Vector Machine* é uma boa solução para a indústria médica apoiando em inúmeros setores da área, e neste caso, ao diagnosticar precocemente os pacientes, é possível apoiar os profissionais de saúde a reduzir erros no diagnóstico, como uma segunda opinião.

Das três doenças possíveis de identificar pode-se concluir que o modelo para a doença cardíaca é o que alcança melhores resultados a nível de obter menos falsos negativos (diagnóstico negativo mas presença de doença), porém é o que tem uma menor precisão nos resultados devido a existirem pontos muito próximos ou idênticos de ambas as classes. Já o cancro da mama é o que obtém os melhores resultados em relação a percentagem de obtenção de falsos positivos (diagnóstico positivo mas ausência da doença), todavia, tem alguma facilidade em conseguir identificar o nódulo benigno na paciente.

Para a demência pode-se concluir que mesmo sendo o modelo que alcança uma melhor precisão nos resultados é onde pode surgir mais receio dos mesmos, devido a existir uma percentagem um pouco elevada de falsos positivos e negativos, isto porque o modelo não consegue identificar a classe convertida, por dois motivos, existirem poucos pontos representativos e porque essa classe pretende saber se o paciente numa primeira consulta foi

identificado com não demente e numa segunda como demente, ou ao contrário, o que pode derivar alguma dificuldade porque os valores são idênticos aos que representam as outras classes. Porém, se apenas fossem possíveis dois diagnósticos, demente e não demente, atingia-se melhores resultados de precisão e de poucos falsos negativos e positivos, porque os restantes diagnósticos o modelo consegue identificar com facilidade.

Em relação a aplicação ProSmartHealth pode-se concluir que atingiu uma precisão média de 84%, alcançando em média, 49.70% de verdadeiros positivos, 34.73% de verdadeiros negativos, 7.63% de falsos negativos e 7.94% de falsos positivos. Assim, é possível verificar que ainda existe alguma probabilidade de obter falsos diagnósticos, mas tendo em conta que para a aplicação é sempre necessário existir uma segunda opinião, a de um profissional de saúde, porque esta aplicação não tem o objetivo de substituir os profissionais de saúde mas sim, apoia-á-los na redução de erros de diagnóstico e em situações de maior aperto, porque o diagnóstico final é sempre do profissional de saúde.

Ao longo da realização deste estudo aconteceram alguns percalços, como a dificuldade de encontrar base de dados de encontro com o pretendido, ter os resultados de cada parâmetro com o diagnóstico final.

Para trabalho futuros será muito importante melhorar a aplicação, trabalhando na análise de sensibilidade, reduzir a percentagem de falsos negativos e positivos, e abranger um maior número de doenças.

Futuramente, o aplicativo ProSmartHealth poderá ser aplicado a um conjunto de dados que envolva vários sintomas de variadas doença, para que seja possível abranger um maior número de doenças, para que se obtenha uma sugestão de diagnóstico com o percentual de contrair cada uma dessas doenças que podem ser observados com os dados inseridos no questionário.

Outras técnicas de *Machine Learning* poderão ser exploradas para melhorar os resultados.

Bibliografia

- [1] D. S. Academy, *Deep Learning Book*, <http://deeplearningbook.com.br/uma-breve-historia-das-redes-neurais-artificiais/>, Online; accessed: 01 April 2020.
- [2] F. Jiang, Y. Jiang, H. Zhi, Y. Dong, H. Li, S. Ma, Y. Wang, Q. Dong, H. Shen and Y. Wang, “Artificial Intelligence in healthcare: past, present and future”, *Stroke and Vascular Neurology*, vol. 27, n.º 2, pp. 97–111, 2017.
- [3] N. J. Nilson, *Inteligencia Artificial – Una nueva síntesis*. Mc Graw Hill, 1987.
- [4] P. Kim, *Matlab, Deep Learning – With Machine Learning, Neural Networks and Artificial Intelligence*. Apress, 2017.
- [5] K. Shailaja, B. Seetharamulu and M. A. Jabbar, “Machine Learning in Healthcare: A Review”, *Proceedings of the 2nd International conference on Electronics, Communication and Aerospace Technology*, pp. 910–914, 2018.
- [6] V. N. Vapnik, *The Nature of Statistical Learning Theory*. Springer, 1995.
- [7] S. Russel and P. Norvig, *Inteligência Artificial: Un Enfoque Moderno*. Pearson Educación, S.A, 2004.
- [8] A. Hazra, S. K. Mandal and A. Gupta, “Study and Analysis of Breast Cancer Cell Detection using Naive Bayes, SVM and Ensemble Algorithms”, *International Journal of Computer Applications*, vol. 145, n.º 2, pp. 0975–8887, 2016.

- [9] P. Sharma and R. Bhartiya, “Implementation of Decision Tree Algorithm to Analysis the Performance”, *International Journal of Advanced Research in Computer and Communication Engineering*, vol. 1, n.º 10, pp. 861–864, 2012.
- [10] A. Hazra, S. K. Mandal, A. Gupta and A. Mukherjee, “Heart Disease Diagnosis and Prediction Using Machine Learning and Data Mining Techniques: A Review”, *Advances in Computational Sciences and Technology*, vol. 10, n.º 7, pp. 2137–2159, 2017.
- [11] H. Zimmermann, *Fuzzy set theory and its applications*. Springer Science+Business Media, LLC, 2001.
- [12] F. Osório e J. Bittencourt, “Sistemas Inteligentes baseados em Redes Neurais Artificiais aplicados ao Processamento de Imagens”, tese de mestrado, Universidade do Vale do Rio dos Sinos, São Leopoldo, Brasil, 2000.
- [13] K. P. Murphy, *Machine Learning: A Probabilistic Perspective*. The MIT Press, 2012.
- [14] K. Xie, *Support Vector Machine - Concept and matlab build*. ECE 480 - Team 4, 2012.
- [15] R. Miotto, F. Wang, S. Wang, X. Jinag and J. T. Dudley, “Deep Learning for healthcare: Review, Opportunities and Challenges”, *Briefings in Bioinformatics*, vol. 6, n.º 19, pp. 1236–1246, 2017.
- [16] Deep Mind, *Using AI to predict retinal disease progression*, <https://deepmind.com/>, Online; aceso: 30 de janeiro de 2020.
- [17] Enlitic, *Bridging human and artificial intelligence to advance medical diagnostics*, <https://www.enlitic.com>, Online; aceso: 30 de janeiro de 2020.
- [18] A. Khare, M. Jeon, I. K. Sethi and B. Xu, “Machine Learning Theory and Applications for Healthcare”, *Journal of Healthcare Engineering*, p. 2, 2017.
- [19] J. Krämer, J. Schreyögg and R. Busse, “Classification of hospital admissions into emergency and elective care: a machine learning approach”, *Health Care Manag Sci*, n.º 22, pp. 85–105, 2019.

- [20] V. B. Selva, V. R. Garcia, S. Tortajada, J. M. Benidi, B. Valdivieso and J. M. G. Gomez, “Design of one-year mortality forecast at hospital admission based: a Machine Learning Approach”, 2019.
- [21] C. Venkatesan, P. Karthigaikumar, A. Paul, S. Satheeskumaran and R. Kumar, “ECG Signal Preprocessing and SVM Classifier-Based Abnormality Detection in Remote Healthcare Applications”, *Special Section on Sopt Computing Techniques for Image Analysis in the Medical Industry Current Trends, Challenges and Solutions*, vol. XX, pp. 1–8, 2017.
- [22] H. Wang, B. Zheng, S. W. Yoon and H. S. Ko, “A support vector machine-based ensemble algorithm for breast cancer diagnosis”, *European Journal of Operational Research*, n.º 267, pp. 687–699, 2018.
- [23] G. Battineni, N. Chintalapudi and F. Amenta, “Machine Learning in medicine: Performance calculation of dementia prediction by support vector machine (SVM)”, *Informatics in Medicine Unlocked*, n.º 16, pp. 100–200, 2019.
- [24] M. A. Hearst, “Support Vector Machine”, *IEEE Intelligent Systems*, pp. 18–28, 1998.
- [25] C. J. C. Burges, “A tutorial on support vector machines for pattern recognition”, *Data Mining and Knowledge Discovery*, n.º 2, pp. 1–43, 1998.
- [26] S. Haykin, *Neural Networks a Comprehensive Foundation*. PrenticeHall, 1999.
- [27] A. Kowalczyk, *Support Vector Machines Succinctly*. Synctfusion. Inc., 2017.
- [28] N. Deng, Y. Tian and C. Zhang, *Support Vector Machine - Optimization Based Theory, Algorithms, and Extensions*. Chapman Hall Book/CRC Press, 2013.
- [29] S. Escalera, O. Pujol and P. Radeva, “Separability of ternary codes for sparse designs od error-correcting output codes”, *Pattern Recognition Letters*, n.º 30, pp. 285–297, 2009.
- [30] A. Qassim, *Breast Cancer Cell Type Classifier*, <https://towardsdatascience.com/breast-cancer-cell-type-classifier-ace4e82f9a79>, Online; acedido: 03 de maio de 2020.

- [31] C. B. C. Latha and S. C. Jeeva, “Improving the accuracy of prediction of heart disease risk based on ensemble classification techniques”, *Informatics in Medicine Unlocked*, n.º 16, pp. 1–9, 2019.
- [32] S. B. Estar, *Biópsia da mama*, <https://www.saudebemestar.pt/pt/exame/imagiologia/biopsia-da-mama/>, Online; acedido: 30 de abril de 2020.

Apêndice A

Anexos

Anexo 1 – Cancro da Mama

Anexo 1A

Tabela 1: Conjunto de dados utilizado.

ID Paciente	Raio	Textura	Perímetro	Área	Suavidade	Compactação	Concavidade	Côncava	Simetria	Dimensão Fractal	Diagnóstico
1	13.54	14.36	87.46	566.3	0.09779	0.08129	0.06664	0.04781	0.1885	0.05766	0
2	13.08	15.71	85.63	520	0.1075	0.127	0.04568	0.0311	0.1967	0.06811	0
3	9.504	12.44	60.34	273.9	0.1024	0.06492	0.02956	0.02076	0.1815	0.06905	0
4	13.03	18.42	82.61	523.8	0.08983	0.03766	0.02562	0.02923	0.1467	0.05863	0
5	8.196	16.84	51.71	201.9	0.086	0.05943	0.01588	0.005917	0.1769	0.06503	0
6	12.05	14.63	78.04	449.3	0.1031	0.09092	0.06592	0.02749	0.1675	0.06043	0
7	13.49	22.3	86.91	561	0.08752	0.07698	0.04751	0.03384	0.1809	0.05718	0
8	11.76	21.6	74.72	427.9	0.08637	0.04966	0.01657	0.01115	0.1495	0.05888	0
9	13.64	16.34	87.21	571.8	0.07685	0.06059	0.01857	0.01723	0.1353	0.05953	0
10	11.94	18.24	75.71	437.6	0.08261	0.04751	0.01972	0.01349	0.1868	0.0611	0
11	11.52	18.75	73.34	409	0.09524	0.05473	0.03036	0.02278	0.192	0.05907	0
12	13.05	19.31	82.61	527.2	0.0806	0.03789	0.000692	0.004167	0.1819	0.05501	0
13	8.618	11.79	54.34	224.5	0.09752	0.05272	0.02061	0.007799	0.1683	0.07187	0
14	10.17	14.88	64.55	311.9	0.1134	0.08061	0.01084	0.0129	0.2743	0.0696	0
15	8.598	20.98	54.66	221.8	0.1243	0.08963	0.03	0.009259	0.1828	0.06757	0
16	9.173	13.86	59.2	260.9	0.07721	0.08751	0.05988	0.0218	0.2341	0.06963	0
17	9.465	21.01	60.11	269.4	0.1044	0.07773	0.02172	0.01504	0.1717	0.06899	0
18	11.31	19.04	71.8	394.1	0.08139	0.04701	0.03709	0.0223	0.1516	0.05667	0
19	9.029	17.33	58.79	250.5	0.1066	0.1413	0.313	0.04375	0.2111	0.08046	0
20	12.78	16.49	81.37	502.5	0.09831	0.05234	0.03653	0.02864	0.159	0.05653	0

21	8.888	14.64	58.79	244	0.09783	0.1531	0.08606	0.02872	0.1902	0.0898	0
22	12.31	16.52	79.19	470.9	0.09172	0.06829	0.03372	0.02272	0.172	0.05914	0
23	13.53	10.94	87.91	559.2	0.1291	0.1047	0.06877	0.06556	0.2403	0.06641	0
24	12.86	18	83.19	506.3	0.09934	0.09546	0.3883	0.02315	0.1718	0.05997	0
25	11.45	20.97	73.81	401.5	0.1102	0.09362	0.04591	0.02233	0.1842	0.07005	0
26	13.34	15.86	86.49	520	0.1078	0.1535	0.1169	0.06987	0.1942	0.06902	0
27	12	15.65	76.95	443.3	0.09723	0.07165	0.04151	0.01863	0.2079	0.05968	0
28	12.36	21.8	79.78	466.1	0.08772	0.09445	0.06015	0.03745	0.193	0.06404	0
29	14.64	15.24	95.77	651.9	0.1132	0.1339	0.09966	0.07064	0.2116	0.06346	0
30	14.62	24.02	94.57	662.7	0.08974	0.08606	0.03102	0.02957	0.1685	0.05866	0
31	13.27	14.76	84.74	551.7	0.07355	0.05055	0.03261	0.02648	0.1386	0.05318	0
32	13.45	18.3	86.6	555.1	0.1022	0.08165	0.03974	0.0278	0.1638	0.0571	0
33	12.18	17.84	77.79	451.1	0.1045	0.07057	0.0249	0.02491	0.19	0.06635	0
34	9.787	19.94	62.11	294.5	0.1024	0.05301	0.006829	0.007937	0.135	0.0689	0
35	11.6	12.84	74.34	412.6	0.08983	0.07525	0.04196	0.0335	0.162	0.06582	0
36	6.981	13.43	43.79	143.5	0.117	0.07568	0	0	0.193	0.07818	0
37	12.18	20.52	77.22	458.7	0.08013	0.04038	0.02383	0.0177	0.1739	0.05677	0
38	9.876	19.4	63.95	298.3	0.1005	0.09697	0.06154	0.03029	0.10945	0.06322	0
39	10.49	19.29	67.41	336.1	0.09989	0.08578	0.02995	0.01201	0.2217	0.06481	0
40	11.64	18.33	75.17	412.5	0.1142	0.1017	0.0707	0.03485	0.1801	0.0652	0
41	12.36	18.54	79.01	466.7	0.08477	0.06815	0.02643	0.01921	0.1602	0.06066	0
42	11.34	21.26	72.48	396.5	0.08759	0.06575	0.05133	0.01899	0.1487	0.06529	0
43	9.777	16.99	62.5	290.2	0.1037	0.08404	0.04334	0.01778	0.1584	0.07065	0
44	12.63	20.76	82.15	480.4	0.09933	0.1209	0.1065	0.06021	0.1735	0.0707	0
45	14.26	19.65	97.83	629.9	0.07837	0.2233	0.3003	0.07798	0.1704	0.07769	0
46	10.51	20.19	68.64	334.2	0.1122	0.1303	0.06476	0.03068	0.1922	0.07782	0
47	8.726	15.83	55.84	230.9	0.115	0.08201	0.04132	0.01924	0.1649	0.07633	0

48	11.93	21.53	76.53	438.6	0.09768	0.07849	0.03328	0.02008	0.1688	0.06194	0
49	8.95	15.76	58.74	245.2	0.09462	0.1243	0.09263	0.02308	0.1305	0.07163	0
50	11.41	10.82	73.34	403.3	0.09373	0.06685	0.03512	0.02623	0.1667	0.06113	0
51	14.5	10.89	94.28	640.7	0.1101	0.1099	0.08842	0.05778	0.1856	0.06402	0
52	13.37	16.39	86.1	553.3	0.07115	0.07325	0.08092	0.028	0.1422	0.05823	0
53	13.85	17.21	88.44	588.7	0.08785	0.06136	0.0142	0.01141	0.1614	0.0589	0
54	15.1	16.39	99.58	674.5	0.115	0.1807	0.1138	0.08534	0.2001	0.06467	0
55	12.19	13.29	79.08	455.8	0.1066	0.09509	0.02855	0.02882	0.188	0.06471	0
56	15.71	13.93	102	761.7	0.09462	0.09462	0.07135	0.05933	0.1816	0.05723	0
57	11.71	16.67	74.72	423.6	0.1051	0.06095	0.03592	0.026	0.1339	0.05945	0
58	11.43	15.39	73.06	399.8	0.09639	0.06889	0.03503	0.02875	0.1734	0.05865	0
59	11.28	13.39	73	384.8	0.1164	0.1136	0.04635	0.04796	0.1771	0.06072	0
60	9.738	11.97	61.24	288.5	0.0925	0.04102	0	0	0.1903	0.06422	0
61	11.43	17.31	73.66	398	0.1092	0.09486	0.02031	0.01861	0.1645	0.06562	0
62	12.9	15.92	83.74	512.2	0.08677	0.09509	0.04894	0.03088	0.1778	0.06235	0
63	10.75	14.97	68.26	355.3	0.07793	0.05139	0.02251	0.007875	0.1399	0.05688	0
64	11.9	14.65	78.11	432.8	0.1152	0.1296	0.1152	0.1296	0.0371	0.03003	0
65	14.95	18.77	97.84	689.5	0.08138	0.1167	0.0905	0.03562	0.1744	0.06493	0
66	14.44	15.18	93.97	640.1	0.0997	0.1021	0.08487	0.05532	0.1724	0.06081	0
67	13.74	17.91	88.12	585	0.07944	0.06376	0.02881	0.01329	0.1473	0.0558	0
68	13	20.78	83.51	519.4	0.1135	0.07589	0.03136	0.02645	0.254	0.06087	0
69	8.219	20.7	53.27	203.9	0.09405	0.1305	0.1321	0.02168	0.2222	0.08261	0
70	9.731	15.34	63.78	300.2	0.1072	0.1599	0.4108	0.07857	0.2548	0.09296	0
71	11.15	13.08	70.87	381.9	0.09754	0.05113	0.01982	0.01786	0.183	0.06105	0
72	13.15	15.34	85.31	538.9	0.09384	0.08498	0.09293	0.03483	0.1822	0.06207	0
73	12.25	17.94	78.27	460.3	0.08654	0.06679	0.3885	0.02331	0.197	0.06228	0
74	16.84	19.46	108.4	880.2	0.07445	0.07223	0.0515	0.02771	0.1844	0.05268	0

75	12.06	12.74	76.84	448.6	0.09311	0.05241	0.01972	0.01963	0.159	0.05907	0
76	10.9	12.96	68.69	366.8	0.07515	0.03718	0.00309	0.006588	0.1442	0.05743	0
77	11.75	20.18	76.1	419.8	0.1089	0.1141	0.06843	0.03738	0.1993	0.06453	0
78	12.34	22.22	79.85	464.5	0.1012	0.1015	0.0537	0.02822	0.1551	0.06761	0
79	14.97	19.76	95.5	690.2	0.08421	0.05352	0.01947	0.01939	0.1515	0.05266	0
80	10.8	9.71	68.77	357.6	0.09594	0.05736	0.02531	0.01698	0.1381	0.064	0
81	14.97	16.95	96.22	685.9	0.09855	0.07885	0.02602	0.03781	0.178	0.9565	0
82	12.32	12.39	78.85	464.1	0.1028	0.06981	0.03987	0.037	0.1959	0.05955	0
83	11.08	14.71	70.21	372.7	0.1006	0.05743	0.02363	0.02583	0.1566	0.06669	0
84	10.66	15.15	67.49	349.6	0.08792	0.04302	0	0	0.1928	0.05975	0
85	8.671	14.45	54.42	277.2	0.09138	0.04276	0	0	0.1722	0.06724	0
86	9.904	18.06	64.6	302.4	0.09699	0.1294	0.1307	0.03176	0.1669	0.08116	0
87	13.01	22.22	82.01	526.4	0.06251	0.01938	0.001595	0.001852	0.1395	0.05234	0
88	12.81	13.06	81.29	508.8	0.08739	0.03774	0.009193	0.0133	0.1466	0.06133	0
89	11.41	14.92	73.53	402	0.09057	0.08155	0.06181	0.02361	0.1167	0.06217	0
90	10.08	15.11	63.76	317.5	0.09267	0.04695	0.001597	0.002404	0.1703	0.06048	0
91	11.71	17.19	74.68	420.3	0.09774	0.06141	0.03809	0.03239	0.1516	0.06095	0
92	11.81	17.39	75.27	428.9	0.1007	0.05562	0.02353	0.01553	0.1718	0.0578	0
93	12.3	15.9	78.83	463.7	0.0808	0.07253	0.03844	0.01654	0.1667	0.05474	0
94	12.77	21.41	82.02	507.4	0.08749	0.06601	0.03112	0.02864	0.1694	0.06287	0
95	9.72	18.22	60.73	288.1	0.0695	0.02344	0	0	0.1653	0.06447	0
96	12.91	16.33	82.53	516.4	0.07941	0.05366	0.03873	0.02377	0.1829	0.05667	0
97	12.23	19.56	78.54	461	0.09586	0.08087	0.04187	0.04107	0.1979	0.06013	0
98	12.47	18.6	81.09	481.9	0.09965	0.1058	0.08005	0.03821	0.1925	0.06373	0
99	9.876	17.27	62.92	295.4	0.1089	0.07232	0.01756	0.01952	0.1934	0.05986	0
100	13.11	22.54	87.02	529.4	0.1002	0.1483	0.08705	0.05102	0.185	0.0731	0
101	15.27	12.91	98.17	725.5	0.08182	0.0623	0.05892	0.03157	0.1359	0.05526	0

102	11.84	18.94	75.51	428	0.08871	0.069	0.02669	0.01393	0.1533	0.06057	0
103	11.89	18.35	77.32	432.2	0.09363	0.1154	0.06636	0.03142	0.1967	0.06314	0
104	10.2	17.48	65.05	321.2	0.08054	0.05097	0.05774	0.01071	0.1964	0.06315	0
105	13.65	13.16	87.88	568.9	0.09646	0.08711	0.03888	0.02563	0.136	0.06344	0
106	13.56	13.9	88.59	561.3	0.1051	0.1192	0.0786	0.04451	0.1962	0.06303	0
107	10.18	17.53	65.12	313.1	0.1061	0.08502	0.01768	0.01915	0.191	0.06908	0
108	13.27	17.02	84.55	546.4	0.08445	0.04994	0.03554	0.02456	0.1496	0.05674	0
109	14.34	13.47	92.51	641.2	0.09906	0.07624	0.05724	0.04603	0.2075	0.05448	0
110	10.44	15.46	66.62	329.6	0.1053	0.07722	0.006643	0.01216	0.1788	0.0645	0
111	15	15.51	97.45	684.5	0.08371	0.1096	0.06505	0.0378	0.1881	0.05907	0
112	12.62	23.97	81.35	496.4	0.07903	0.07529	0.05438	0.02036	0.1514	0.06019	0
113	11.32	27.08	71.76	395.7	0.06883	0.03813	0.01633	0.003125	0.1869	0.05628	0
114	11.22	33.81	70.79	386.8	0.0778	0.03574	0.004967	0.006434	0.1845	0.05828	0
115	9.567	15.91	60.21	279.6	0.08464	0.04087	0.01652	0.01667	0.1551	0.06403	0
116	14.03	21.25	89.79	603.4	0.0907	0.06945	0.01462	0.01896	0.1517	0.05835	0
117	14.22	27.85	92.55	623.9	0.08223	0.1039	0.1103	0.04408	0.1342	0.06129	0
118	13.64	15.6	87.38	575.3	0.09423	0.0663	0.04705	0.03731	0.1717	0.0566	0
119	12.42	15.04	78.61	476.5	0.07926	0.03393	0.01053	0.01108	0.1546	0.05754	0
120	11.3	18.19	73.93	389.4	0.09592	0.1325	0.1548	0.02854	0.2054	0.07669	0
121	13.75	23.77	88.54	590	0.08043	0.06807	0.04697	0.02344	0.1773	0.05429	0
122	10.48	19.86	66.72	337.7	0.107	0.05971	0.04336	0.01105	0.1487	0.05635	0
123	13.2	17.43	84.13	541.6	0.07215	0.04524	0.04336	0.01105	0.1487	0.05635	0
124	12.89	14.11	84.95	512.2	0.0876	0.1346	0.1374	0.0398	0.1596	0.06409	0
125	10.65	25.22	68.01	347	0.09657	0.07234	0.02379	0.01615	0.1897	0.06329	0
126	11.52	14.93	73.87	406.3	0.1013	0.07808	0.04328	0.02929	0.1883	0.06168	0
127	11.5	18.45	73.28	407.4	0.09345	0.05991	0.02638	0.02069	0.1834	0.05934	0
128	10.6	18.95	69.28	346.4	0.09688	0.1147	0.06387	0.02642	0.1922	0.06491	0

129	13.59	21.84	87.16	561	0.07956	0.08259	0.04072	0.02142	0.1635	0.05859	0
130	12.87	16.21	82.38	512.2	0.09425	0.06219	0.039	0.01615	0.201	0.05769	0
131	10.71	20.39	69.5	344.9	0.1082	0.1289	0.08448	0.02867	0.1668	0.06862	0
132	14.29	16.82	90.3	632.6	0.06429	0.02675	0.00725	0.00625	0.1508	0.05376	0
133	11.29	13.04	72.23	388	0.09834	0.07608	0.03265	0.02755	0.1769	0.0627	0
134	9.742	15.67	61.5	289.9	0.09037	0.04689	0.01103	0.01407	0.2081	0.06312	0
135	11.89	17.36	76.2	435.6	0.1225	0.0721	0.05929	0.07404	0.2015	0.05875	0
136	11.33	14.16	71.79	396.6	0.09379	0.03872	0.001487	0.003333	0.1954	0.05821	0
137	13.59	17.84	86.24	572.3	0.07948	0.04052	0.01997	0.01238	0.1573	0.0552	0
138	13.85	15.18	88.99	587.4	0.09516	0.07688	0.04479	0.03711	0.211	0.05853	0
139	11.74	14.02	74.24	427.3	0.07813	0.0434	0.02245	0.02763	0.2101	0.06113	0
140	12.89	15.7	84.08	516.6	0.07818	0.0958	0.1115	0.0339	0.1432	0.05935	0
141	12.58	18.4	79.83	489	0.08393	0.04216	0.00186	0.002924	0.1697	0.05855	0
142	11.94	20.76	77.87	441	0.08605	0.1011	0.06574	0.03791	0.1588	0.06766	0
143	12.89	13.12	81.89	515.9	0.06955	0.03729	0.0226	0.01171	0.1337	0.05581	0
144	11.26	19.96	73.72	394.1	0.0802	0.1181	0.09274	0.05588	0.2595	0.06233	0
145	11.37	18.89	72.17	396	0.08713	0.05008	0.02399	0.02173	0.2013	0.05955	0
146	14.41	19.73	96.03	651	0.08757	0.1676	0.1362	0.06602	0.1714	0.07192	0
147	14.96	19.1	97.03	687.3	0.08992	0.09823	0.0594	0.04819	0.1879	0.05852	0
148	12.95	16.02	83.14	513.7	0.1005	0.07943	0.06155	0.0337	0.173	0.0647	0
149	11.85	17.46	75.54	432.7	0.08372	0.05642	0.02688	0.0228	0.1875	0.05715	0
150	12.72	13.78	81.78	492.1	0.09667	0.08393	0.01288	0.01924	0.1638	0.061	0
151	13.77	13.27	88.06	582.7	0.09198	0.06221	0.01063	0.01917	0.1592	0.05912	0
152	10.91	12.35	69.14	363.7	0.08518	0.04721	0.01236	0.01369	0.1449	0.06031	0
153	14.26	18.17	91.22	633.1	0.06576	0.0522	0.02475	0.01374	0.1635	0.05586	0
154	10.51	23.29	66.85	334.2	0.1015	0.06797	0.02495	0.01875	0.1695	0.06556	0
155	12.46	19.89	80.43	471.3	0.08451	0.1014	0.0683	0.03099	0.1781	0.06249	0

156	10.49	18.61	66.86	334.3	0.1068	0.06678	0.02297	0.0178	0.1482	0.066	0
157	11.46	18.16	73.59	403.1	0.08853	0.07694	0.03344	0.01502	0.1411	0.06243	0
158	11.6	24.49	74.23	417.2	0.07474	0.05688	0.01974	0.01313	0.1935	0.05878	0
159	13.2	15.82	84.07	537.3	0.08511	0.05251	0.001461	0.003261	0.1632	0.05894	0
160	9	14.4	56.36	246.3	0.07005	0.03116	0.003681	0.003472	0.1788	0.06833	0
161	13.5	12.71	85.69	566.2	0.07376	0.03614	0.002758	0.004419	0.1365	0.05335	0
162	13.05	13.84	82.71	530.6	0.08352	0.03735	0.004559	0.008829	0.1453	0.05518	0
163	11.7	19.11	74.33	418.7	0.08814	0.05253	0.01583	0.01148	0.1936	0.06128	0
164	14.61	15.69	92.68	664.9	0.07618	0.03515	0.01447	0.01877	0.1632	0.05255	0
165	12.76	13.37	82.29	504.1	0.08794	0.07948	0.04052	0.02548	0.1601	0.0614	0
166	11.54	10.72	73.73	409.1	0.08597	0.05969	0.01367	0.008907	0.1833	0.061	0
167	8.597	18.6	54.09	221.2	0.1074	0.05847	0.05847	0	0	0.2163	0
168	12.49	16.85	79.19	481.6	0.08511	0.03834	0.004473	0.006423	0.1215	0.05673	0
169	12.18	14.08	77.25	461.4	0.07734	0.03212	0.01123	0.005051	0.1673	0.05649	0
170	9.042	18.9	60.07	244.5	0.09968	0.1972	0.1975	0.04908	0.233	0.08743	0
171	12.43	17.78	78.6	477.3	0.07557	0.03454	0.01342	0.01699	0.1472	0.05561	0
172	10.25	16.18	66.52	324.2	0.1061	0.1111	0.06726	0.03965	0.1743	0.07279	0
173	12.86	13.32	82.82	504.8	0.1134	0.08834	0.038	0.034	0.1543	0.07726	0
174	12.2	15.21	78.01	457.9	0.08673	0.06545	0.01994	0.01692	0.1638	0.06129	0
175	12.67	17.3	81.25	489.9	0.1028	0.07664	0.03193	0.02107	0.1707	0.05984	0
176	14.11	12.88	90.03	616.5	0.09309	0.05306	0.01765	0.02733	0.1373	0.057	0
177	12.03	17.93	76.09	446	0.07683	0.03892	0.001546	0.005592	0.1382	0.0607	0
178	12.98	19.35	84.52	514	0.9579	0.1125	0.07107	0.0295	0.1761	0.0654	0
179	11.22	19.86	71.94	387.3	0.1054	0.06779	0.005006	0.007583	0.194	0.06028	0
180	11.25	14.78	71.38	390	0.08306	0.04458	0.0009737	0.002941	0.1773	0.06081	0
181	12.3	19.02	77.88	464.4	0.08313	0.04202	0.007756	0.008535	0.1539	0.05945	0
182	12.99	14.23	84.08	514.3	0.09462	0.09965	0.03738	0.02098	0.1652	0.07238	0

183	10.05	17.53	64.41	310.8	0.1007	0.07326	0.02511	0.01775	0.189	0.06331	0
184	14.42	16.54	94.15	641.2	0.09751	0.1139	0.08007	0.04223	0.1912	0.06412	0
185	9.606	16.84	61.64	280.5	0.08481	0.09228	0.08422	0.02292	0.2036	0.07125	0
186	11.06	14.96	71.49	373.9	0.1033	0.09097	0.05397	0.03341	0.1776	0.06907	0
187	11.71	15.45	75.03	420.3	0.115	0.07281	0.04006	0.0325	0.2009	0.06506	0
188	10.26	14.71	66.2	321.6	0.09882	0.09159	0.03581	0.02037	0.1633	0.07005	0
189	12.06	18.9	76.66	445.3	0.08386	0.05794	0.00751	0.008488	0.1555	0.06048	0
190	14.76	14.74	94.87	668.7	0.08875	0.0778	0.04608	0.03528	0.1521	0.05912	0
191	11.47	16.03	73.02	402.7	0.09076	0.05886	0.02587	0.02322	0.1634	0.06372	0
192	11.95	14.96	77.23	426.7	0.1158	0.1206	0.01171	0.01787	0.2459	0.06581	0
193	11.66	17.07	73.7	421	0.07561	0.0363	0.008306	0.01162	0.1671	0.05731	0
194	11.14	14.07	71.24	384.6	0.07276	0.06064	0.04505	0.01471	0.169	0.06083	0
195	12.56	19.07	81.92	485.8	0.0876	0.1038	0.103	0.04391	0.1533	0.06184	0
196	13.05	18.59	85.09	512	0.1082	0.1304	0.09603	0.05603	0.2035	0.06501	0
197	13.87	16.21	88.52	593.7	0.08743	0.05492	0.01502	0.02088	0.1424	0.05883	0
198	8.878	15.49	56.74	241	0.08293	0.07698	0.04721	0.02381	0.193	0.06621	0
199	9.436	18.32	59.82	278.6	0.1009	0.05956	0.0271	0.01406	0.1506	0.06959	0
200	12.54	18.07	79.42	491.9	0.07436	0.0265	0.001194	0.005449	0.1528	0.06959	0
201	13.3	21.57	85.24	546.1	0.08582	0.06373	0.03344	0.02424	0.1815	0.05696	0
202	12.76	18.84	81.87	496.6	0.09676	0.07952	0.02688	0.01781	0.1759	0.06183	0
203	16.5	18.29	106.6	838.1	0.09686	0.08468	0.05862	0.04835	0.1495	0.05593	0
204	13.4	16.95	85.48	552.4	0.07937	0.05696	0.02181	0.01473	0.165	0.05701	0
205	12.21	18.02	78.31	458.4	0.09231	0.07175	0.04392	0.02027	0.1695	0.05916	0
206	15.19	13.21	97.65	711.8	0.07963	0.06934	0.03393	0.02657	0.1721	0.05544	0
207	13.69	16.07	87.84	579.1	0.08302	0.06374	0.02556	0.02031	0.1872	0.05669	0
208	16.17	16.07	106.3	788.5	0.0988	0.1438	0.06651	0.05397	0.199	0.06572	0
209	10.57	20.22	70.15	338.3	0.09073	0.166	0.228	0.05941	0.2188	0.0845	0

210	13.46	28.21	85.89	562.1	0.07517	0.04726	0.01271	0.01117	0.1421	0.05763	0
211	13.66	15.15	88.27	580.6	0.08268	0.07548	0.04249	0.02471	0.1792	0.05897	0
212	11.27	12.96	73.16	386.3	0.1237	0.1111	0.079	0.0555	0.2018	0.06914	0
213	11.04	14.93	70.67	372.7	0.07987	0.07079	0.03546	0.02074	0.2003	0.06246	0
214	12.05	22.72	78.75	447.8	0.06935	0.1073	0.07943	0.02978	0.1203	0.06659	0
215	12.39	17.48	80.64	462.9	0.1042	0.1297	0.05892	0.0288	0.1779	0.06588	0
216	13.28	13.72	85.79	541.8	0.08363	0.08575	0.05077	0.02864	0.1617	0.05594	0
217	12.21	14.09	78.78	462	0.08108	0.07823	0.06839	0.02534	0.1646	0.06154	0
218	13.88	16.16	88.37	596.6	0.07026	0.04831	0.02045	0.008507	0.1607	0.05474	0
219	11.27	15.5	73.38	392	0.08365	0.1114	0.1007	0.02757	0.181	0.07252	0
220	10.26	12.22	65.75	321.6	0.09996	0.07542	0.01923	0.01968	0.18	0.06569	0
221	8.734	16.84	55.27	234.3	0.1039	0.07428	0	0	0.1985	0.07098	0
222	12.1	17.72	78.07	446.2	0.1029	0.09758	0.04783	0.03326	0.1937	0.06161	0
223	14.06	17.18	89.75	609.1	0.08045	0.05361	0.02681	0.03251	0.1641	0.05764	0
224	13.51	18.89	88.1	558.1	0.1059	0.1147	0.0858	0.05381	0.1806	0.06079	0
225	12.8	17.46	83.05	508.3	0.08044	0.08895	0.0739	0.04083	0.1574	0.0575	0
226	11.06	14.83	70.31	378.2	0.07741	0.04768	0.02712	0.007246	0.1535	0.06214	0
227	11.8	17.26	75.26	431.9	0.09087	0.06232	0.02853	0.01638	0.1847	0.06019	0
228	11.93	10.91	76.14	442.7	0.08872	0.05242	0.02606	0.01796	0.1601	0.05541	0
229	12.96	18.29	84.18	525.2	0.07351	0.07899	0.04057	0.01883	0.1874	0.05899	0
230	12.94	16.17	83.18	507.6	0.09879	0.08836	0.03296	0.0239	0.1735	0.062	0
231	12.34	14.95	78.29	469.1	0.08682	0.04571	0.02109	0.02054	0.1571	0.05708	0
232	10.94	18.59	70.39	370	0.1004	0.0746	0.04944	0.02932	0.1486	0.06615	0
233	16.14	14.86	104.3	800	0.09495	0.08501	0.055	0.04528	0.1735	0.05875	0
234	12.85	21.37	82.63	514.5	0.07551	0.08316	0.06126	0.01867	0.158	0.06114	0
235	12.27	17.92	78.41	466.1	0.08685	0.06526	0.03211	0.02653	0.1966	0.05597	0
236	11.36	17.57	72.49	399.8	0.08858	0.05313	0.02783	0.021	0.1601	0.05913	0

237	11.04	16.83	70.92	373.2	0.1077	0.07804	0.03046	0.0248	0.1714	0.0634	0
238	9.397	21.68	59.75	268.8	0.07969	0.06253	0.03735	0.005128	0.1274	0.06724	0
239	14.99	22.11	97.53	693.7	0.08515	0.1025	0.06859	0.03876	0.1944	0.05913	0
240	11.89	21.17	76.39	433.8	0.09773	0.0812	0.02555	0.02179	0.2019	0.0629	0
241	9.405	21.7	59.6	271.2	0.1044	0.06159	0.02047	0.01257	0.2025	0.06601	0
242	12.7	12.17	80.88	495	0.08785	0.05794	0.0236	0.02402	0.1583	0.06275	0
243	11.16	21.41	70.95	380.3	0.1018	0.05978	0.008955	0.01076	0.1615	0.06144	0
244	11.57	19.04	74.2	409.7	0.08546	0.07722	0.05485	0.01428	0.2031	0.06267	0
245	14.69	13.98	98.22	656.1	0.1031	0.1836	0.145	0.063	0.2086	0.07406	0
246	11.61	16.02	75.46	408.2	0.1088	0.1168	0.07097	0.04497	0.1886	0.0632	0
247	13.66	19.13	89.46	575.3	0.09057	0.1147	0.09657	0.04812	0.1848	0.06181	0
248	9.742	19.12	61.93	289.7	0.1075	0.08333	0.008934	0.01967	0.2538	0.07029	0
249	10.03	21.28	63.19	307.3	0.08117	0.03912	0.00247	0.005159	0.163	0.06439	0
250	10.48	14.98	67.49	333.6	0.09816	0.1013	0.06335	0.02218	0.1925	0.06915	0
251	10.8	21.98	68.79	359.9	0.08801	0.05743	0.03614	0.01404	0.2016	0.05977	0
252	11.13	16.62	70.47	381.1	0.08151	0.03834	0.01369	0.0137	0.1511	0.06148	0
253	12.72	17.67	80.98	501.3	0.07896	0.04522	0.01402	0.01835	0.1459	0.05544	0
254	12.4	17.68	81.47	467.8	0.1054	0.1316	0.07741	0.02799	0.1811	0.07102	0
255	14.86	16.94	94.89	673.7	0.08924	0.07074	0.03346	0.02877	0.1573	0.05703	0
256	12.87	19.54	82.67	509.2	0.09136	0.07883	0.01797	0.0209	0.1861	0.06347	0
257	14.04	15.98	89.78	611.2	0.08458	0.05895	0.03534	0.02944	0.1714	0.05898	0
258	13.85	19.6	88.68	592.6	0.08684	0.0633	0.01342	0.02293	0.1555	0.05673	0
259	14.02	15.66	89.59	606.5	0.07966	0.05581	0.02087	0.02652	0.1589	0.05586	0
260	10.97	17.2	71.73	371.5	0.08915	0.1113	0.09457	0.03613	0.1489	0.0664	0
261	13.78	15.79	88.37	585.9	0.08817	0.06718	0.01055	0.009937	0.1405	0.05848	0
262	10.57	18.32	66.82	340.9	0.08142	0.04462	0.01993	0.01111	0.2372	0.0578	0
263	11.99	24.89	77.61	441.3	0.103	0.09218	0.05441	0.04274	0.182	0.0685	0

264	14.8	17.66	95.88	674.8	0.09179	0.0889	0.04069	0.0226	0.1893	0.05886	0
265	14.53	19.34	94.25	659.7	0.08388	0.078	0.08817	0.02925	0.1473	0.05746	0
266	11.87	21.54	76.83	432	0.06613	0.1064	0.08777	0.02386	0.1349	0.06612	0
267	12.28	28.23	76.77	442.5	0.08437	0.0645	0.04055	0.01945	0.1615	0.06104	0
268	14.53	13.98	93.86	644.2	0.1099	0.09242	0.06895	0.06495	0.165	0.06121	0
269	12.62	17.15	80.62	492.9	0.08583	0.0543	0.02966	0.02272	0.1799	0.05826	0
270	13.38	30.72	86.34	557.2	0.09245	0.07426	0.02819	0.03264	0.1375	0.06016	0
271	11.63	29.29	74.87	415.1	0.09357	0.08574	0.0716	0.02017	0.1799	0.06166	0
272	13.21	25.25	84.1	537.9	0.08791	0.05205	0.02772	0.02068	0.1619	0.05584	0
273	13	25.13	82.61	520.2	0.08369	0.05073	0.01206	0.01762	0.1667	0.05449	0
274	9.755	28.2	61.68	290.9	0.07984	0.04626	0.01541	0.01043	0.1621	0.05952	0
275	14.4	26.99	92.25	646.1	0.06995	0.05223	0.03476	0.01737	0.1707	0.05433	0
276	11.6	18.36	73.88	412.7	0.08508	0.05855	0.03367	0.01777	0.1516	0.05859	0
277	13.17	18.22	84.28	537.3	0.07466	0.05994	0.04859	0.0287	0.1454	0.05549	0
278	13.24	20.13	86.87	542.9	0.08284	0.1223	0.101	0.02833	0.1601	0.06432	0
279	13.14	20.74	85.98	536.9	0.08675	0.1089	0.1085	0.0351	0.1562	0.0602	0
280	9.668	18.1	61.06	286.3	0.08311	0.05428	0.01479	0.005769	0.168	0.06412	0
281	11.62	18.18	76.38	408.8	0.1175	0.1483	0.102	0.05564	0.1957	0.07255	0
282	9.667	18.49	61.49	289.1	0.08946	0.06258	0.02948	0.01514	0.2238	0.06413	0
283	12.04	28.14	76.85	449.9	0.08752	0.06	0.02367	0.02377	0.1854	0.05698	0
284	14.92	14.93	96.45	686.9	0.08098	0.08549	0.05539	0.03221	0.1687	0.05669	0
285	12.27	29.97	77.42	465.4	0.07699	0.03398	0	0	0.1701	0.0596	0
286	10.88	15.62	70.41	358.9	1007	0.1069	0.05115	0.01571	0.1861	0.06837	0
287	12.83	15.73	82.89	506.9	0.0904	0.08269	0.05835	0.03078	0.1705	0.05913	0
288	14.2	20.53	92.41	618.4	0.08931	0.1108	0.05063	0.03058	0.1506	0.06009	0
289	13.9	16.62	88.97	599.4	0.06828	0.05319	0.02224	0.01339	0.1813	0.05536	0
290	11.49	14.59	73.99	404.9	0.1046	0.08228	0.05308	0.01969	0.1779	0.06574	0

291	12.16	18.03	78.29	455.3	0.09087	0.07838	0.02916	0.01527	0.1464	0.06284	0
292	13.9	19.24	88.73	602.9	0.07991	0.05326	0.02995	0.0207	0.1579	0.05594	0
293	13.47	14.06	87.32	546.3	0.1071	0.1155	0.05786	0.05266	0.1779	0.06639	0
294	13.7	17.64	87.76	571.1	0.0995	0.07957	0.04548	0.0316	0.1732	0.06088	0
295	15.73	11.28	102.8	747.2	0.1043	0.1299	0.1191	0.06211	0.1784	0.06259	0
296	12.45	16.41	82.85	476.7	0.09514	0.1511	0.1544	0.04846	0.2082	0.07325	0
297	14.64	16.85	94.21	666	0.08641	0.06698	0.05192	0.02791	0.1409	0.05355	0
298	11.68	16.17	75.49	420.5	0.1128	0.09263	0.04279	0.03132	0.1853	0.06401	0
299	12.25	22.44	78.18	466.5	0.08192	0.052	0.01714	0.01261	0.1544	0.05976	0
300	17.85	13.23	114.6	992.1	0.07838	0.06217	0.04445	0.04178	0.122	0.05243	0
301	12.46	12.83	78.83	477.3	0.07372	0.04043	0.007173	0.01149	0.1613	0.06013	0
302	13.16	20.54	84.06	538.7	0.07335	0.05275	0.018	0.01256	0.1713	0.05888	0
303	14.87	20.21	96.12	680.9	0.09587	0.08345	0.06824	0.04951	0.1487	0.05748	0
304	12.65	18.17	82.69	485.6	0.1076	0.1334	0.08017	0.05074	0.1641	0.06854	0
305	12.47	17.31	80.45	480.1	0.08928	0.0763	0.03609	0.02369	0.1526	0.06046	0
306	15.04	16.74	98.73	689.4	0.09883	0.1364	0.07721	0.06142	0.1668	0.06869	0
307	12.54	16.32	81.25	476.3	0.1158	0.1085	0.05928	0.03279	0.1943	0.06612	0
308	9.268	12.87	61.49	248.7	0.1634	0.2239	0.0973	0.05252	0.2378	0.09502	0
309	9.676	13.14	64.12	272.5	0.1255	0.2204	0.1188	0.07038	0.2057	0.09575	0
310	12.22	20.04	79.47	453.1	0.1096	0.1152	0.08175	0.02166	0.2124	0.06894	0
311	11.06	17.12	71.25	366.5	0.1194	0.1071	0.04063	0.04268	0.1954	0.07976	0
312	16.3	15.7	104.7	819.8	0.09427	0.06712	0.05526	0.04563	0.1711	0.05657	0
313	11.74	14.69	76.31	426	0.08099	0.09661	0.06726	0.02639	0.1499	0.06758	0
314	14.81	14.7	94.66	680.7	0.08472	0.05016	0.03416	0.02541	0.1659	0.05348	0
315	14.58	13.66	94.29	658.8	0.09832	0.08918	0.08222	0.04349	0.1739	0.0564	0
316	11.34	18.61	72.76	391.2	0.1049	0.08499	0.04302	0.02594	0.1927	0.06211	0
317	12.88	18.22	84.45	493.1	0.1218	0.1661	0.04825	0.05303	0.1709	0.07253	0

318	12.75	16.7	82.51	493.8	0.1125	0.1117	0.0388	0.02995	0.212	0.06623	0
319	9.295	13.9	59.96	257.8	0.1371	0.1225	0.03332	0.02421	0.2197	0.07696	0
320	11.26	19.83	71.3	388.1	0.08511	0.04413	0.005067	0.005664	0.1637	0.06343	0
321	13.71	18.68	88.73	571	0.09916	0.107	0.05385	0.03783	0.1714	0.06843	0
322	9.847	15.68	63	293.2	0.09492	0.08419	0.0233	0.02416	0.1387	0.06891	0
323	8.571	13.1	54.53	221.3	0.1036	0.07632	0.02565	0.0151	0.1678	0.07126	0
324	13.46	18.75	87.44	551.1	0.1075	0.1138	0.04201	0.03152	0.1723	0.06317	0
325	12.34	12.27	78.94	468.5	0.09003	0.06307	0.02958	0.02647	0.1689	0.05808	0
326	13.94	13.17	90.31	594.2	0.1248	0.09755	0.101	0.06615	0.1976	0.06457	0
327	12.07	13.44	77.83	445.2	0.11	0.09009	0.03781	0.02798	0.1657	0.06608	0
328	11.75	17.56	75.89	422.9	0.1073	0.09713	0.05282	0.0444	0.1598	0.06677	0
329	11.67	20.02	75.21	416.2	0.1016	0.09453	0.042	0.02157	0.1859	0.06461	0
330	13.68	16.33	87.76	575.5	0.09277	0.07255	0.01752	0.0188	0.1631	0.06155	0
331	10.96	17.62	70.79	365.6	0.09687	0.09752	0.05263	0.02788	0.1619	0.06408	0
332	11.69	24.44	76.37	406.4	0.1236	0.1552	0.04515	0.04531	0.2131	0.07405	0
333	7.729	25.49	47.98	178.8	0.08098	0.04878	0	0	0.187	0.07285	0
334	7.691	25.44	48.34	170.4	0.08668	0.1199	0.09252	0.01364	0.2037	0.07751	0
335	11.54	14.44	74.65	402.9	0.09984	0.112	0.06737	0.02594	0.1818	0.06782	0
336	14.47	24.99	95.81	656.4	0.08837	0.123	0.1009	0.0389	0.1872	0.06341	0
337	14.74	25.42	94.7	668.6	0.08275	0.07214	0.04105	0.03027	0.184	0.0568	0
338	13.21	28.06	84.88	538.4	0.08671	0.06877	0.02987	0.03275	0.1628	0.0568	0
339	13.87	20.7	89.77	584.8	0.09578	0.1018	0.03688	0.02369	0.162	0.06688	0
340	13.62	23.23	87.19	573.2	0.09246	0.06747	0.02974	0.02443	0.1664	0.05781	0
341	10.32	16.35	65.31	324.9	0.09434	0.04994	0.01012	0.005495	0.1885	0.06201	0
342	10.26	16.58	65.85	320.8	0.08877	0.08066	0.04358	0.02438	0.1669	0.06714	0
343	9.683	19.34	61.05	285.7	0.08491	0.0503	0.02337	0.009615	0.158	0.06235	0
344	10.82	24.21	68.89	361.6	0.08192	0.06602	0.01548	0.00816	0.1976	0.06328	0

345	10.86	21.48	68.51	360.5	0.07431	0.04227	0	0	0.1661	0.05948	0
346	11.13	22.44	71.49	378.4	0.09566	0.08194	0.04824	0.02257	0.203	0.06552	0
347	12.77	29.43	81.35	507.9	0.08276	0.04234	0.01997	0.01499	0.1539	0.05637	0
348	9.333	21.94	59.01	264	0.0924	0.05605	0.03996	0.01282	0.1692	0.06576	0
349	12.88	28.92	82.5	514.3	0.08123	0.05824	0.06195	0.02343	0.1566	0.05708	0
350	10.29	27.61	65.67	321.4	0.0903	0.04658	0.05999	0.02738	0.1593	0.06127	0
351	10.16	19.59	64.73	311.7	0.1003	0.07504	0.005025	0.01116	0.1791	0.06331	0
352	9.423	27.88	59.26	271.3	0.08123	0.04971	0	0	0.1742	0.06059	0
353	14.59	22.68	96.39	657.1	0.08473	0.133	0.1029	0.03736	0.1454	0.06147	0
354	11.51	23.93	74.52	403.5	0.09261	0.1021	0.1112	0.04105	0.1388	0.0657	0
355	14.05	27.15	91.38	600.4	0.09929	0.1126	0.04462	0.04304	0.1537	0.06171	0
356	11.2	29.37	70.67	386	0.07449	0.03558	0	0	0.106	0.05502	0
357	7.76	24.54	47.92	181	0.05263	0.04362	0	0	0.1587	0.05884	0
358	17.99	10.38	122.8	1001	0.1184	0.2776	0.3001	0.1471	0.2419	0.07871	1
359	20.57	17.77	132.9	1326	0.08474	0.07864	0.0869	0.0701	0.1812	0.05667	1
360	19.69	21.25	130	1203	0.1096	0.1599	0.1974	0.1279	0.2069	0.05999	1
361	11.42	20.38	77.58	386.1	0.1425	0.2839	0.2414	0.1052	0.2597	0.09744	1
362	20.29	14.34	135.1	1297	0.1003	0.1328	0.198	0.1043	0.1809	0.05883	1
363	12.45	15.7	82.57	477.1	0.1278	0.17	0.1578	0.08089	0.2087	0.07613	1
364	18.25	19.98	119.6	1040	0.09463	0.109	0.1127	0.074	0.1794	0.05742	1
365	13.71	20.83	90.2	577.9	0.1189	0.1645	0.09366	0.05985	0.2196	0.07451	1
366	13	21.82	87.5	519.8	0.1273	0.1932	0.1859	0.09353	0.235	0.07389	1
367	12.46	24.04	83.97	475.9	0.1186	0.2396	0.2273	0.08543	0.203	0.08243	1
368	16.02	23.24	102.7	797.8	0.08206	0.06669	0.03299	0.03323	0.1528	0.05697	1
369	15.78	17.89	103.6	781	0.0971	0.1292	0.09954	0.06606	0.1842	0.06082	1
370	19.17	24.8	132.4	1123	0.0974	0.2458	0.2065	0.1118	0.2397	0.078	1
371	15.85	23.95	103.7	782.7	0.08401	0.1002	0.09938	0.05364	0.1847	0.05338	1

372	13.73	22.61	93.6	578.3	0.1131	0.2293	0.2128	0.08025	0.2069	0.07682	1
373	14.54	27.54	96.73	658.8	0.1139	0.1595	0.1639	0.07364	0.2303	0.07077	1
374	14.68	20.13	94.74	684.5	0.09867	0.072	0.07395	0.05259	0.1586	0.05922	1
375	16.13	20.68	108.1	798.8	0.117	0.2022	0.1722	0.1028	0.2164	0.07356	1
376	19.81	22.15	130	1260	0.09831	0.1027	0.1479	0.09498	0.1582	0.05395	1
377	15.34	14.26	102.5	704.4	0.1073	0.2135	0.2077	0.09756	0.2521	0.07032	1
378	21.16	23.04	137.2	1404	0.09428	0.1022	0.1097	0.08632	0.1769	0.05278	1
379	16.65	21.38	110	904.6	0.1121	0.1457	0.1525	0.0917	0.1995	0.0633	1
380	17.14	16.4	116	912.7	0.1186	0.2276	0.2229	0.1401	0.304	0.07413	1
381	14.58	21.53	97.41	644.8	0.1054	0.1868	0.1425	0.08783	0.2252	0.06924	1
382	18.61	20.25	122.1	1094	0.0944	0.1066	0.149	0.07731	0.1697	0.05699	1
383	15.3	25.27	102.4	732.4	0.1082	0.1697	0.1683	0.08751	0.1926	0.0654	1
384	17.57	15.05	115	955.1	0.09847	0.1157	0.09875	0.07953	0.1739	0.06149	1
385	18.63	25.11	124.8	1088	0.1064	0.1887	0.2319	0.1244	0.2183	0.06197	1
386	11.84	18.7	77.93	440.6	0.1109	0.1516	0.1218	0.05182	0.2301	0.07799	1
387	17.02	23.98	112.8	899.3	0.1197	0.1496	0.2417	0.1203	0.2248	0.06382	1
388	19.27	26.47	127.9	1162	0.09401	0.1719	0.1657	0.07593	0.1853	0.06261	1
389	16.13	17.88	107	807.2	0.104	0.1559	0.1354	0.07752	0.1998	0.06515	1
390	16.74	21.59	110.1	869.5	0.0961	0.1336	0.1348	0.06018	0.1896	0.05656	1
391	14.25	21.72	93.63	633	0.09823	0.1098	0.1319	0.05598	0.1885	0.061125	1
392	14.99	25.2	95.54	698.8	0.09387	0.05131	0.02398	0.1565	0.05504	1.214	1
393	13.48	20.82	88.4	559.2	0.1016	0.1255	0.1063	0.05439	0.172	0.06419	1
394	13.44	21.58	86.18	563	0.08162	0.06031	0.0311	0.02031	0.1784	0.05587	1
395	10.95	21.35	71.9	371.1	0.1227	0.1218	0.1044	0.05669	0.1895	0.0687	1
396	19.07	24.81	128.3	1104	0.09081	0.219	0.2107	0.09961	0.231	0.06343	1
397	13.28	20.28	87.32	545.2	0.1041	0.1436	0.09847	0.06158	0.1974	0.06782	1
398	13.17	21.81	85.42	531.5	0.09714	0.1047	0.08259	0.05252	0.1746	0.06177	1

399	18.65	17.6	123.7	1076	0.1099	0.1686	0.1974	0.1009	0.1907	0.06049	1
400	13.17	18.66	85.98	534.6	1158	0.1231	0.1226	0.0734	0.2128	0.06777	1
401	18.22	18.7	120.3	1033	0.1148	0.1485	0.1772	0.106	0.2092	0.0631	1
402	15.1	22.02	97.26	712.8	0.09056	0.07081	0.05253	0.03334	0.1616	0.05684	1
403	19.21	18.57	125.5	1152	0.1053	0.1267	0.1323	0.08994	0.1917	0.05961	1
404	14.71	21.59	95.55	656.9	0.1137	0.1365	0.1293	0.08123	0.2027	0.06758	1
405	14.25	22.15	96.42	645.7	0.1049	0.2008	0.2135	0.08653	0.1949	0.07292	1
406	12.68	23.84	82.69	499	0.1122	0.1262	0.1128	0.06873	0.1905	0.0659	1
407	14.78	23.94	97.4	668.3	0.1172	0.1479	0.1267	0.09029	0.1953	0.06654	1
408	18.94	21.31	123.6	1130	0.09009	0.1029	0.108	0.07951	0.1582	0.05461	1
409	17.2	24.52	114.2	929.4	0.1071	0.183	0.1692	0.07944	0.1927	0.06487	1
410	13.8	15.79	90.43	584.1	0.1007	0.128	0.07789	0.05069	0.1662	0.06566	1
411	16.07	19.65	104.1	817.7	0.09168	0.08424	0.09769	0.06638	0.1798	0.05391	1
412	18.05	16.15	120.2	1006	0.1065	0.2146	0.1684	0.108	0.2152	0.06673	1
413	20.18	23.97	143.7	1245	0.1286	0.3454	0.3754	0.1604	0.2906	0.08142	1
414	25.22	24.91	171.5	1878	0.1063	0.2665	0.3339	0.1845	0.1829	0.06782	1
415	19.1	26.29	129.1	1132	0.1215	0.1791	0.1937	0.1469	0.1634	0.07224	1
416	18.46	18.52	121.1	1075	0.09874	0.1053	0.1335	0.08795	0.2132	0.06022	1
417	14.48	21.46	94.25	648.2	0.09444	0.09947	0.1204	0.04938	0.2075	0.05636	1
418	19.02	24.59	122	1076	0.09029	0.1206	0.1468	0.08271	0.1953	0.05629	1
419	15.37	22.76	100.2	728.2	0.092	0.1036	0.1122	0.07483	0.1717	0.06097	1
420	15.06	19.83	100.3	705.6	0.1039	0.1553	0.17	0.08815	0.1855	0.06284	1
421	20.26	23.03	132.4	1264	0.09078	0.1313	0.1465	0.08683	0.2095	0.05649	1
422	14.42	19.77	94.48	642.5	0.09752	0.1141	0.09388	0.05839	0.1879	0.0639	1
423	13.61	24.98	88.05	582.7	0.09488	0.08511	0.08625	0.04489	0.1609	0.05871	1
424	13.11	15.56	87.21	530.2	0.1398	0.1765	0.2071	0.09601	0.1925	0.07692	1
425	22.27	19.67	152.8	1509	0.1326	0.2768	0.4264	0.1823	0.2556	0.07039	1

426	14.87	16.67	98.64	682.5	0.1162	0.1649	0.169	0.08923	0.2157	0.06768	1
427	15.78	22.91	105.7	782.6	0.1155	0.1752	0.2133	0.09479	0.2096	0.07331	1
428	17.95	20.01	114.2	982	0.08402	0.06722	0.07293	0.05596	0.2129	0.05025	1
429	18.66	17.12	121.4	1077	0.1054	0.11	0.1457	0.08665	0.1966	0.06213	1
430	24.25	20.2	166.2	1761	0.1447	0.2867	0.4268	0.2012	0.2655	0.06877	1
431	13.61	24.69	87.76	572.6	0.09258	0.07862	0.05285	0.03085	0.1761	0.0613	1
432	19	18.91	123.4	1138	0.08217	0.08028	0.09271	0.05627	0.1946	0.05044	1
433	19.79	25.12	130.4	1192	0.1015	0.1589	0.2545	0.1149	0.2202	0.06113	1
434	15.46	19.48	101.7	748.9	0.1092	0.1223	0.1466	0.08087	0.1931	0.05796	1
435	16.15	21.54	106.2	809.8	0.1008	0.1284	0.1043	0.05613	0.216	0.05891	1
436	18.45	21.91	120.2	1075	0.0943	0.09709	0.1153	0.06847	0.1692	0.05727	1
437	12.77	22.47	81.72	506.3	0.09055	0.05761	0.04711	0.2704	0.1585	0.06065	1
438	14.95	17.57	96.85	678.1	0.1167	0.1136	0.04635	0.04796	0.1771	0.06072	1
439	16.11	18.05	105.1	813	0.09721	0.11137	0.09447	0.05943	0.1861	0.06248	1
440	11.8	16.58	78.99	432	0.1091	0.17	0.1659	0.07415	0.2678	0.07371	1
441	17.68	20.74	117.4	963.7	0.1115	0.1665	0.1855	0.1054	0.1971	0.06166	1
442	19.19	15.94	126.3	1157	0.08694	0.1185	0.1193	0.09667	0.1741	0.05176	1
443	19.59	18.15	130.7	1214	0.112	0.1666	0.2508	0.1286	0.2027	0.06082	1
444	23.27	22.04	152.1	1686	0.08439	0.1145	0.1324	0.09702	0.1801	0.05553	1
445	16.78	18.8	109.3	886.3	0.08865	0.09182	0.08422	0.06576	0.1893	0.05534	1
446	17.47	24.68	116.1	984.6	0.1049	0.1603	0.2159	0.1043	0.1538	0.06365	1
447	13.43	19.63	85.84	565.4	0.09048	0.06288	0.05858	0.03438	0.1598	0.05671	1
448	15.46	11.89	102.5	736.9	0.1257	0.1555	0.2032	0.1097	0.1966	0.07069	1
449	16.46	20.11	109.3	832.9	0.09831	0.1556	0.1793	0.08866	0.1794	0.06323	1
450	27.22	21.87	182.1	2250	0.1094	0.1914	0.2871	0.1878	0.18	0.0577	1
451	21.09	26.57	142.7	1311	0.1141	0.2832	0.2487	0.1496	0.2395	0.07398	1
452	15.7	20.31	101.2	766.6	0.09597	0.08799	0.06593	0.05189	0.1618	0.05549	1

453	15.28	22.41	98.92	710.6	0.09057	0.1052	0.05375	0.03263	0.1727	0.06317	1
454	18.31	18.58	118.6	1041	0.08588	0.08468	0.08169	0.05814	0.1621	0.05425	1
455	14.22	23.12	94.37	609.9	0.1075	0.2413	0.1981	0.06618	0.2384	0.07542	1
456	12.34	26.86	81.15	477.4	0.1034	0.1353	0.1085	0.04562	0.1943	0.06937	1
457	14.86	23.21	100.4	671.4	0.1044	0.198	0.1697	0.08878	0.1737	0.06672	1
458	13.77	22.29	90.63	588.9	0.12	0.1267	0.1385	0.06526	0.1834	0.06877	1
459	18.08	21.84	117.4	1024	0.07371	0.08642	0.1103	0.05778	0.177	0.0534	1
460	19.18	22.49	127.5	1148	0.08523	0.1428	0.1114	0.06772	0.1767	0.05529	1
461	14.45	20.22	94.49	642.7	0.09872	0.1206	0.118	0.0598	0.195	0.06466	1
462	17.54	19.32	115.1	951.6	0.08968	0.1198	0.1036	0.07488	0.1506	0.05491	1
463	23.29	26.67	158.9	1685	0.1141	0.2084	0.3523	0.162	0.22	0.06229	1
464	13.81	23.75	91.56	597.8	0.1323	0.1768	0.1558	0.09176	0.2251	0.07421	1
465	15.12	16.68	98.78	716.6	0.08876	0.09588	0.0755	0.04079	0.1594	0.05986	1
466	17.01	20.26	109.7	904.3	0.08772	0.07304	0.0695	0.0539	0.2026	0.05223	1
467	20.58	22.14	134.7	1290	0.0909	0.1348	0.164	0.09561	0.1765	0.05024	1
468	28.11	18.47	188.5	2499	0.1142	0.1516	0.3201	0.1595	0.1648	0.05525	1
469	17.42	25.56	114.5	948	0.1006	0.1146	0.1682	0.06597	0.1308	0.05866	1
470	14.19	23.81	92.87	610.7	0.09463	0.1306	0.1115	0.06462	0.2235	0.06433	1
471	13.86	16.93	90.96	578.9	0.1026	0.1517	0.09901	0.05602	0.2106	0.06916	1
472	19.8	21.56	129.7	1230	0.09383	0.1306	0.1272	0.08691	0.2094	0.05581	1
473	19.53	32.47	128	1223	0.0842	0.113	0.1145	0.06637	0.1428	0.05313	1
474	15.75	20.25	102.6	761.3	0.1025	0.1204	0.1147	0.06462	0.1935	0.06303	1
475	12.83	22.33	85.26	503.2	0.1088	0.1799	0.1695	0.06861	0.2123	0.07254	1
476	17.05	19.08	113.4	895	0.1141	0.1572	0.191	0.109	0.2131	0.06325	1
477	20.51	27.81	134.4	1319	0.09159	0.01074	0.1554	0.0834	0.1448	0.05592	1
478	23.21	26.97	153.5	1670	0.09509	0.1682	0.195	0.1237	0.1909	0.06309	1
479	20.48	21.46	132.5	1306	0.08355	0.08348	0.09042	0.06022	0.1467	0.05177	1

480	17.46	39.28	113.4	920.6	0.09812	0.1298	0.1417	0.08811	0.1809	0.05966	1
481	19.4	23.5	129.1	1155	0.1027	0.05971	0.04831	0.0307	0.1737	0.0644	1
482	20.94	23.56	138.9	1364	0.1007	0.1606	0.2712	0.131	0.2205	0.05898	1
483	19.73	19.82	130.7	1206	0.1062	0.1849	0.2417	0.0974	0.1733	0.06697	1
484	17.3	17.08	113	928.2	0.1008	0.1041	0.1266	0.08353	0.1813	0.05613	1
485	19.45	19.33	126.5	1169	0.1035	0.1188	0.1379	0.08591	0.1776	0.05647	1
486	13.96	17.05	91.43	602.4	0.1096	0.1279	0.09789	0.05246	0.1908	0.0613	1
487	19.55	28.77	133.6	1207	0.0926	0.2063	0.1784	0.1144	0.1893	0.06232	1
488	15.32	17.27	103.2	713.3	0.1335	0.2284	0.2448	0.1242	0.2398	0.07596	1
489	15.66	23.2	110.2	773.5	1109	0.3114	0.3176	0.1377	0.2495	0.08104	1
490	15.53	33.56	103.7	744.9	0.1063	0.1639	0.1751	0.08399	0.2091	0.0665	1
491	20.31	27.06	132.9	1288	0.1	0.1088	0.1519	0.09333	0.1814	0.05572	1
492	17.35	23.06	111	933.1	0.08662	0.0629	0.02891	0.02837	0.1564	0.05307	1
493	17.29	22.13	114.4	947.8	0.08999	0.1273	0.09697	0.07507	0.2108	0.05464	1
494	15.61	19.38	100	758.6	0.0784	0.05616	0.04209	0.02847	0.1547	0.05443	1
495	17.19	22.07	111.6	928.3	0.09726	0.08995	0.09061	0.06527	0.1867	0.0558	1
496	20.73	31.12	135.7	1419	0.09469	0.1143	0.1367	0.08646	0.1769	0.05674	1
497	21.75	20.99	147.3	1491	0.09401	0.1961	0.2195	0.1088	0.1721	0.06194	1
498	17.93	24.48	115.2	998.9	0.08855	0.07027	0.05699	0.04744	0.1538	0.0551	1
499	18.81	19.98	120.9	1102	0.08923	0.05884	0.0802	0.05843	0.155	0.04996	1
500	19.16	26.6	126.2	1138	0.102	0.1453	0.1921	0.09664	0.1902	0.0622	1
501	19.4	18.18	127.2	1145	0.1037	0.1442	0.1626	0.09464	0.1893	0.05892	1
502	16.24	18.77	108.8	805.1	0.1066	0.1802	0.1948	0.09052	0.1876	0.06684	1
503	11.76	18.14	75	431.1	0.09968	0.05914	0.02685	0.03515	0.1619	0.06287	1
504	19.53	18.9	129.5	1217	0.115	0.1642	0.2197	0.1062	0.1792	0.06552	1
505	20.09	23.86	134.7	1247	0.108	0.1838	0.2283	0.128	0.2249	0.07469	1
506	18.22	18.87	118.7	1027	0.09746	0.1117	0.113	0.0795	0.1807	0.05664	1

507	20.16	19.66	131.1	1274	0.0802	0.08564	0.1155	0.07726	0.1928	0.05096	1
508	20.34	21.51	135.9	1264	0.117	0.1875	0.2565	0.1504	0.2569	0.0667	1
509	16.27	20.71	106.9	813.7	0.1169	0.1319	0.1478	0.08488	0.1948	0.06277	1
510	16.26	21.88	107.5	826.8	0.1165	0.1283	0.1799	0.07981	0.1869	0.06532	1
511	16.03	15.51	105.8	793.2	0.09491	0.1371	0.1204	0.07041	0.1782	0.05976	1
512	17.06	21	111.8	918.6	0.1119	0.1056	0.1508	0.09934	0.1727	0.06071	1
513	18.77	21.43	122.9	1092	0.09116	0.1402	0.106	0.0609	0.1953	0.06083	1
514	23.51	24.27	151.5	1747	0.1069	0.1283	0.2308	0.141	0.1797	0.05506	1
515	19.68	21.68	129.9	1194	0.09797	0.1339	0.1863	0.1103	0.2082	0.05715	1
516	15.75	19.22	107.1	758.6	0.1243	0.2364	0.2914	0.1242	0.2375	0.07603	1
517	25.73	17.46	174.2	2010	0.1149	0.2363	0.3368	0.1913	0.1956	0.06121	1
518	15.08	25.74	98	716.6	0.1024	0.09769	0.1235	0.06553	0.1647	0.06464	1
519	20.44	21.78	133.8	1293	0.0915	0.1131	0.09799	0.07785	0.1618	0.05557	1
520	20.2	26.83	133.7	1234	0.09905	0.1669	0.1641	0.1265	0.1875	0.0602	1
521	21.71	17.25	140.9	1546	0.09384	0.08562	0.1168	0.08465	0.1717	0.05054	1
522	22.01	21.9	147.2	1482	0.1063	0.1954	0.2448	0.1501	0.1824	0.0614	1
523	16.35	23.29	109	840.4	0.09742	0.1497	0.1811	0.08773	0.2175	0.06218	1
524	21.37	15.1	141.3	1386	0.1001	0.1515	0.1932	0.1255	0.1973	0.06183	1
525	20.64	17.35	134.8	1335	0.09446	0.1076	0.1527	0.08941	0.1571	0.05478	1
526	11.08	18.83	73.3	361.6	0.1216	0.2154	0.1689	0.06367	0.2196	0.0795	1
527	14.6	23.29	93.97	664.7	0.08682	0.06636	0.0839	0.05271	0.1627	0.05416	1
528	19.55	23.21	128.9	1174	0.101	0.1318	0.1856	0.1021	0.1989	0.05884	1
529	15.49	19.97	102.4	744.7	0.116	0.1562	0.1891	0.09113	0.1929	0.06744	1
530	21.61	22.28	144.4	1407	0.1167	0.2087	0.281	0.1562	0.2162	0.06606	1
531	17.91	21.02	124.4	994	0.123	0.2576	0.3189	0.1198	0.2113	0.07115	1
532	17.99	20.66	117.8	991.7	0.1036	0.1304	0.1201	0.08824	0.1992	0.06069	1
533	15.13	29.81	96.71	719.5	0.0832	0.04605	0.4686	0.02739	0.1852	0.05294	1

534	15.5	21.08	102.9	803.1	0.112	0.1571	0.1522	0.08481	0.2085	0.06864	1
535	14.9	22.53	102.1	685	0.09947	0.2225	0.2733	0.09711	0.2041	0.06898	1
536	20.18	19.54	133.8	1250	0.1133	0.1489	0.2133	0.1259	0.1724	0.06053	1
537	18.82	21.97	123.7	1110	0.1018	0.1389	0.1594	0.08744	0.1943	0.06132	1
538	13.98	19.62	91.12	599.5	0.106	0.1133	0.1126	0.06463	0.1669	0.06544	1
539	17.27	25.42	112.4	928.8	0.08331	0.1109	0.1204	0.05736	0.1467	0.05407	1
540	18.03	16.85	117.5	990	0.08947	0.1232	0.109	0.06254	0.172	0.0578	1
541	17.75	28.03	117.3	981.6	0.09997	0.1314	0.1698	0.08293	0.1713	0.05916	1
542	21.1	20.52	138.1	1384	0.09684	0.1175	0.1572	0.1155	0.1554	0.05661	1
543	19.59	25	127.7	1191	0.1032	0.09871	0.1655	0.09063	0.1663	0.05391	1
544	17.08	27.15	111.2	930.9	0.09898	0.111	0.1007	0.06431	0.1793	0.06281	1
545	27.42	26.27	186.9	2501	0.1084	0.1988	0.3635	0.1689	0.2061	0.05623	1
546	17.6	23.33	119	980.5	0.09289	0.2004	0.2136	0.1002	0.1696	0.07369	1
547	16.25	19.51	109.8	815.8	0.1026	0.1893	0.2236	0.09194	0.2151	0.06578	1
548	19.44	18.82	128.1	1167	0.1089	0.1448	0.2256	0.1194	0.1823	0.06115	1
549	16.69	20.2	107.1	857.6	0.07497	0.07112	0.03649	0.02307	0.1846	0.05325	1
550	18.01	20.56	118.4	1007	0.1001	0.1289	0.117	0.07762	0.2116	0.06077	1
551	18.49	17.52	121.3	1068	0.1012	0.1317	0.1491	0.09183	0.1832	0.06697	1
552	20.59	21.24	137.8	1320	0.1085	0.1644	0.2188	0.1121	0.1848	0.06222	1
553	13.82	24.49	92.33	595.9	0.1162	0.1681	0.1357	0.06759	0.2275	0.07237	1
554	23.09	19.83	152.1	1682	0.09342	0.1275	0.1676	0.1003	0.1505	0.05484	1
555	15.46	23.95	103.8	731.3	0.1183	0.187	0.203	0.0852	0.1807	0.07083	1
556	13.4	20.52	88.64	556.7	0.1106	0.1469	0.1445	0.08172	0.2116	0.07325	1
557	15.05	19.07	97.26	701.9	0.09215	0.08597	0.07486	0.04335	0.1561	0.05915	1
558	18.31	20.58	120.8	1052	0.1068	0.1248	0.1569	0.09451	0.186	0.05941	1
559	19.89	20.26	130.5	1214	0.1037	0.131	0.1411	0.09431	0.1802	0.06188	1
560	24.63	21.6	165.5	1841	0.103	0.2106	0.231	0.1471	0.1991	0.06739	1

561	20.47	20.67	134.7	1299	0.09156	0.1313	0.1523	0.1015	0.2166	0.05419	1
562	20.55	20.86	137.8	1308	0.1046	0.1739	0.02085	0.1322	0.2127	0.06251	1
563	14.27	22.55	93.77	629.8	0.1038	0.1154	0.1463	0.06139	0.1926	0.05982	1
564	15.22	30.62	103.4	716.9	0.1048	0.2087	0.255	0.09429	0.2128	0.07152	1
565	20.92	25.09	143	1947	0.1099	0.2236	0.3174	0.1474	0.2149	0.06879	1
566	21.56	22.39	142	1479	0.111	0.1159	0.2439	0.1389	0.1726	0.05623	1
567	20.13	28.25	131.2	1261	0.0978	0.1034	0.144	0.09791	0.1752	0.05533	1
568	16.6	28.08	108.3	858.1	0.08455	0.1023	0.09251	0.05302	0.159	0.05648	1
569	20.6	29.33	140.1	1265	0.1178	0.277	0.3514	0.152	0.2397	0.07016	1

Anexo 1B

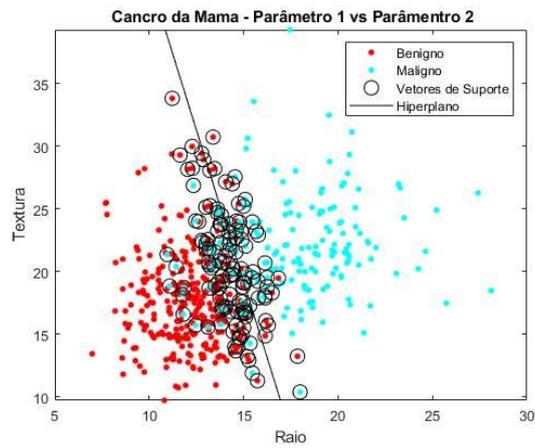


Figura 1: Gráfico representativo da relação entre o parâmetro 1 (Raio) e o parâmetro 2 (Textura).

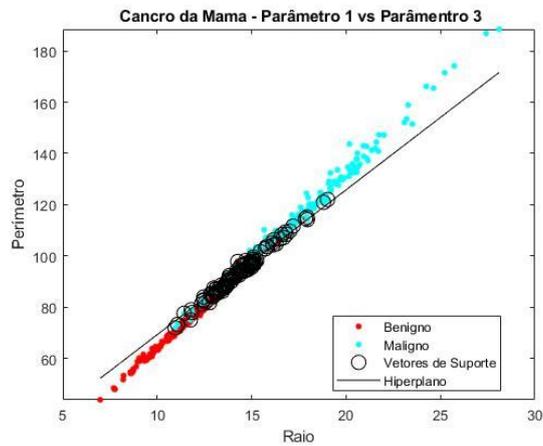


Figura 2: Gráfico representativo da relação entre o parâmetro 1 (Raio) e o parâmetro 3 (Perímetro).

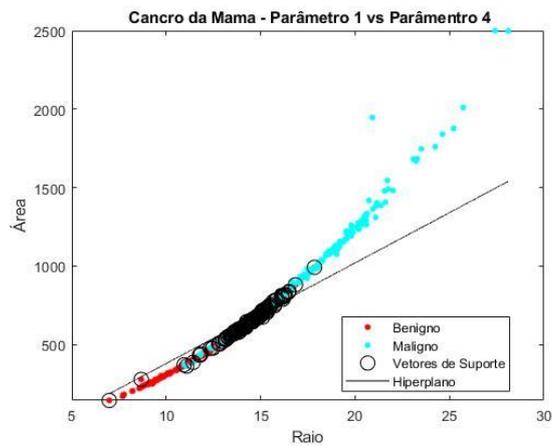


Figura 3: Gráfico representativo da relação entre o parâmetro 1 (Raio) e o parâmetro 4 (Área).

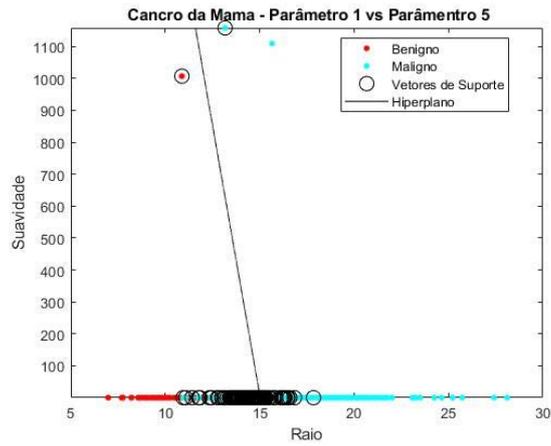


Figura 4: Gráfico representativo da relação entre o parâmetro 1 (Raio) e o parâmetro 5 (Suavidade).

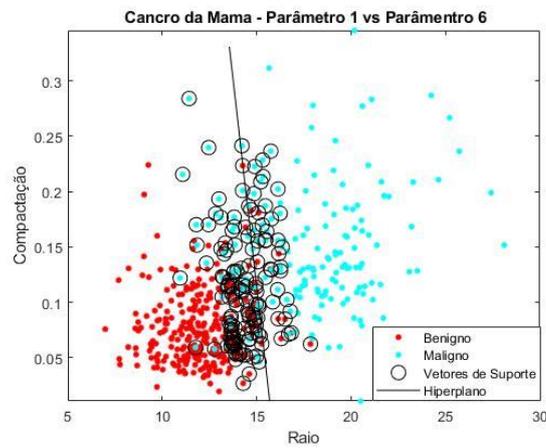


Figura 5: Gráfico representativo da relação entre o parâmetro 1 (Raio) e o parâmetro 6 (Compactação).

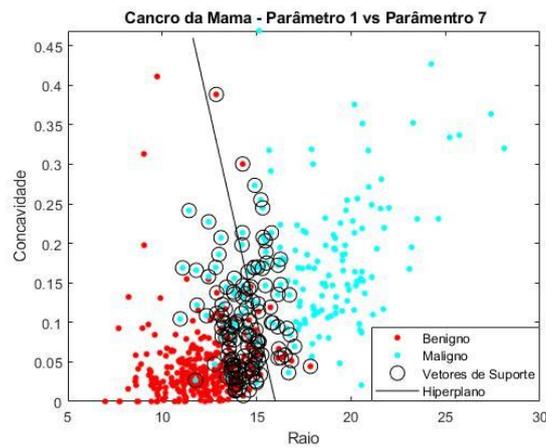


Figura 6: Gráfico representativo da relação entre o parâmetro 1 (Raio) e o parâmetro 7 (Concavidade).

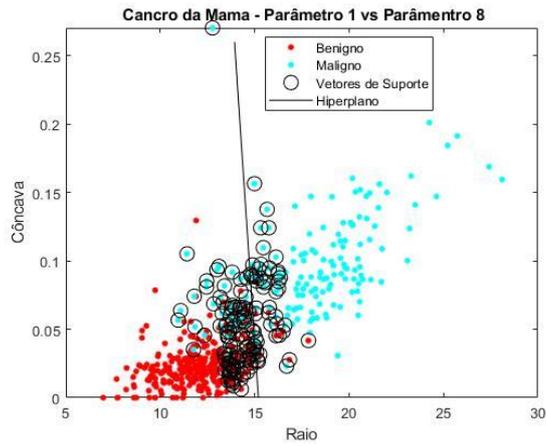


Figura 7: Gráfico representativo da relação entre o parâmetro 1 (Raio) e o parâmetro 8 (Cõncava).

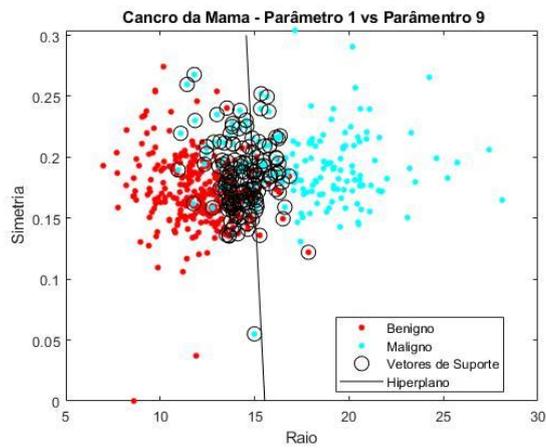


Figura 8: Gráfico representativo da relação entre o parâmetro 1 (Raio) e o parâmetro 9 (Simetria).

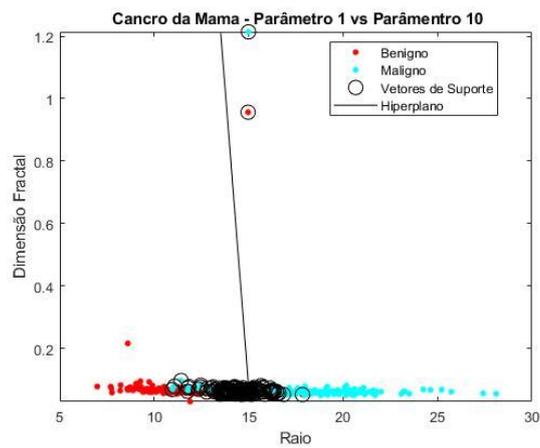


Figura 9: Gráfico representativo da relação entre o parâmetro 1 (Raio) e o parâmetro 10 (Dimensão Fractal).

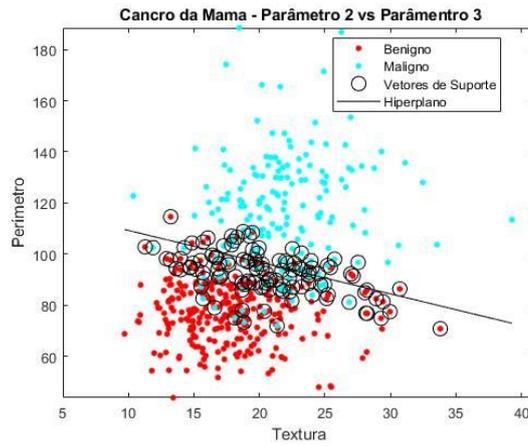


Figura 10: Gráfico representativo da relação entre o parâmetro 2 (Textura) e o parâmetro 3 (Perímetro).

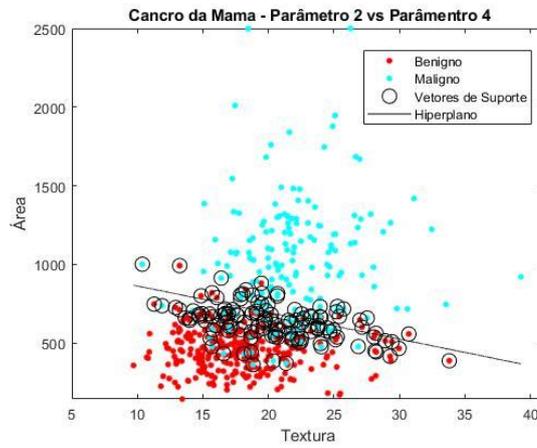


Figura 11: Gráfico representativo da relação entre o parâmetro 2 (Textura) e o parâmetro 4 (Área).

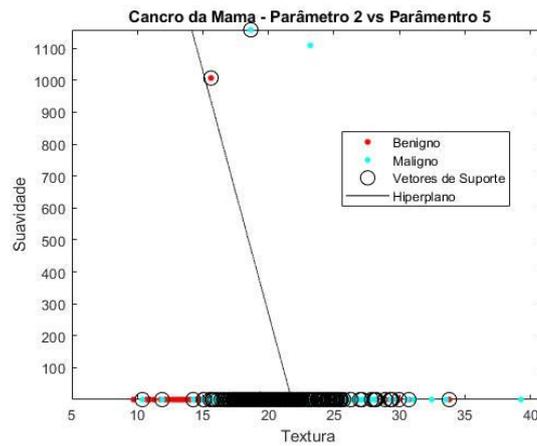


Figura 12: Gráfico representativo da relação entre o parâmetro 2 (Textura) e o parâmetro 5 (Suavidade).

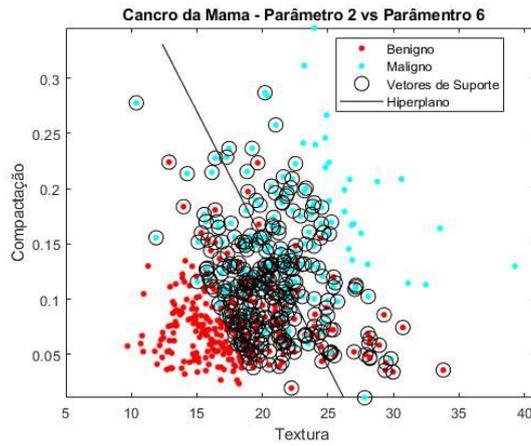


Figura 13: Gráfico representativo da relação entre o parâmetro 2 (Textura) e o parâmetro 6 (Compactação).

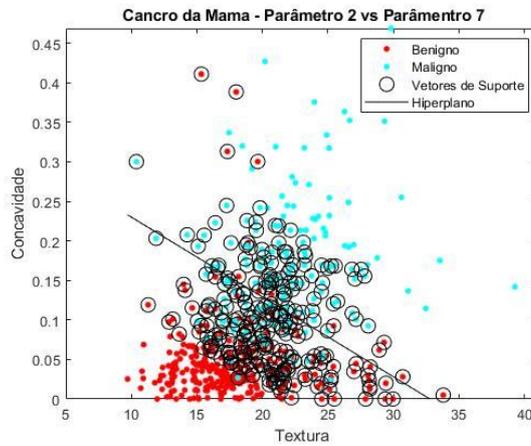


Figura 14: Gráfico representativo da relação entre o parâmetro 2 (Textura) e o parâmetro 7 (Concavidade).

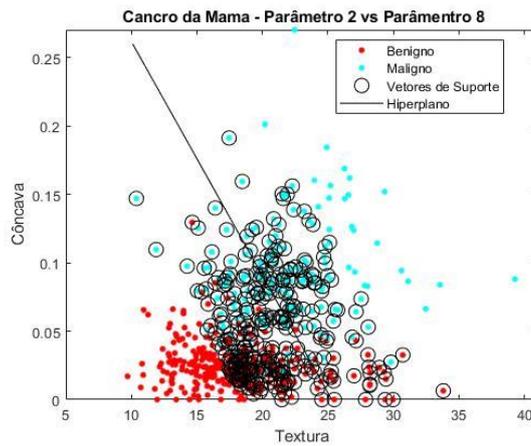


Figura 15: Gráfico representativo da relação entre o parâmetro 2 (Textura) e o parâmetro 8 (Côncava).

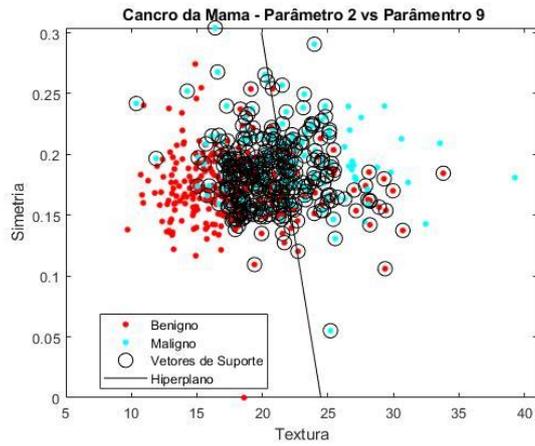


Figura 16: Gráfico representativo da relação entre o parâmetro 2 (Textura) e o parâmetro 9 (Simetria).

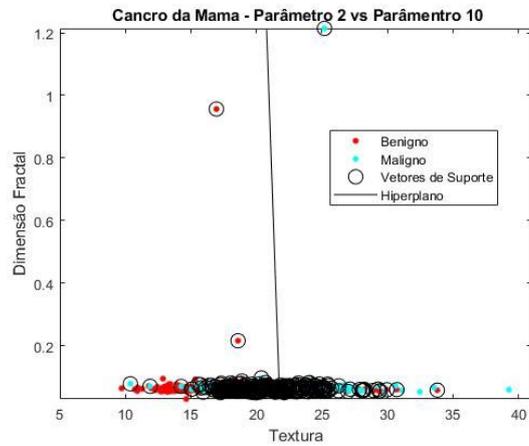


Figura 17: Gráfico representativo da relação entre o parâmetro 2 (Textura) e o parâmetro 10 (Dimensão Fractal).

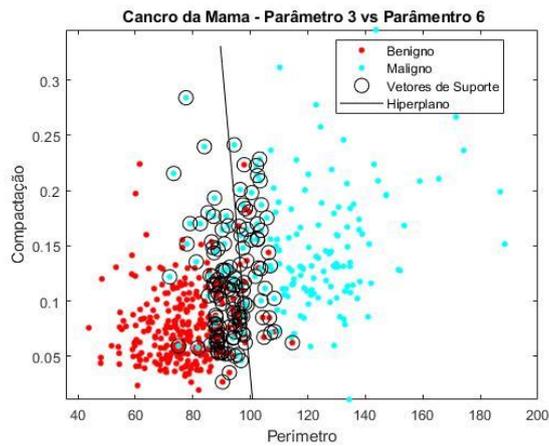


Figura 18: Gráfico representativo da relação entre o parâmetro 3 (Perímetro) e o parâmetro 6 (Compacção).

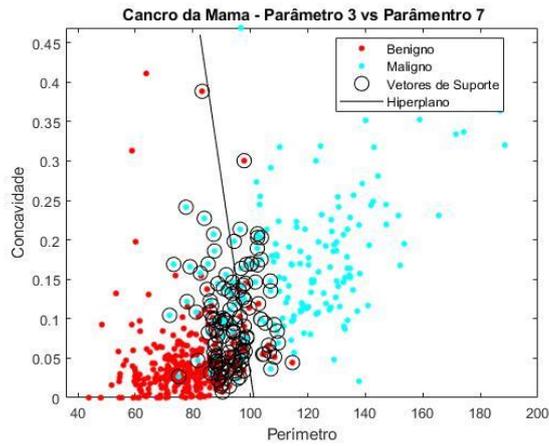


Figura 19: Gráfico representativo da relação entre o parâmetro 3 (Perímetro) e o parâmetro 7 (Concavidade).

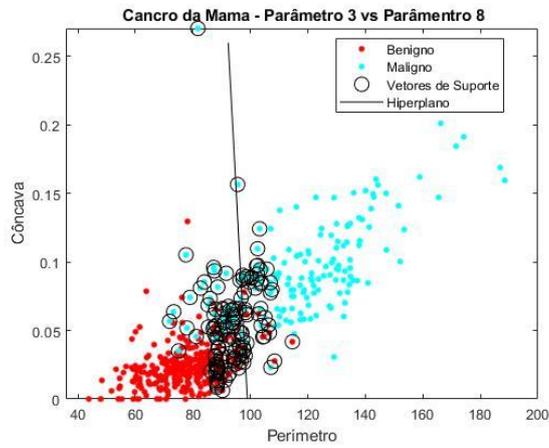


Figura 20: Gráfico representativo da relação entre o parâmetro 3 (Perímetro) e o parâmetro 8 (Côncava).

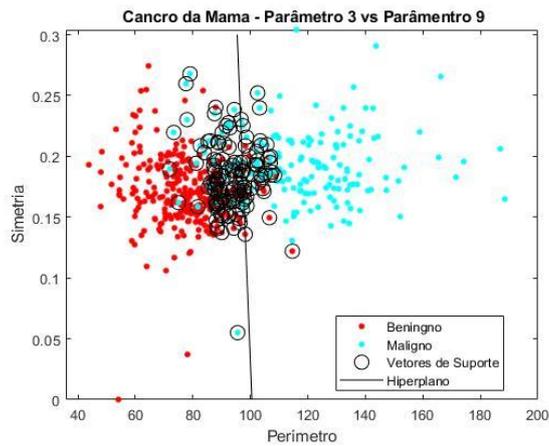


Figura 21: Gráfico representativo da relação entre o parâmetro 3 (Perímetro) e o parâmetro 9 (Simetria).

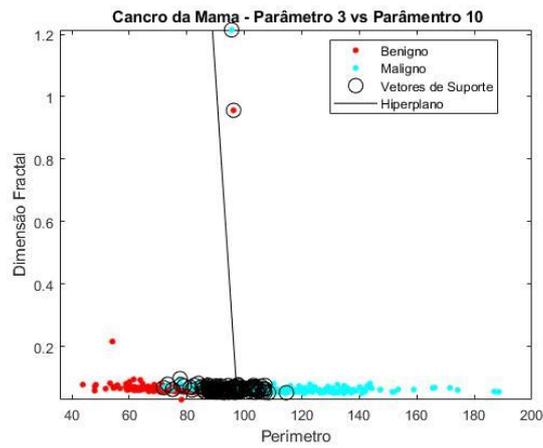


Figura 22: Gráfico representativo da relação entre o parâmetro 3 (Perímetro) e o parâmetro 10 (Dimensão Fractal).

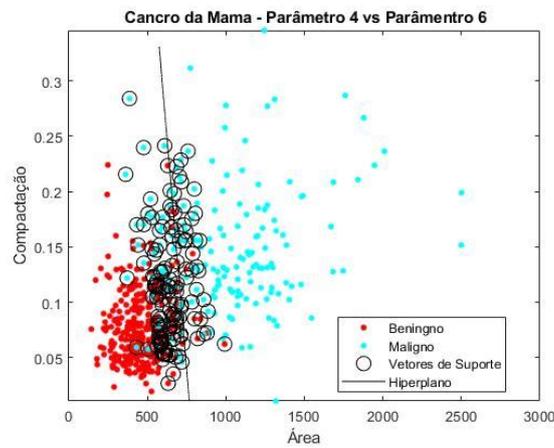


Figura 23: Gráfico representativo da relação entre o parâmetro 4 (Área) e o parâmetro 6 (Compactação).

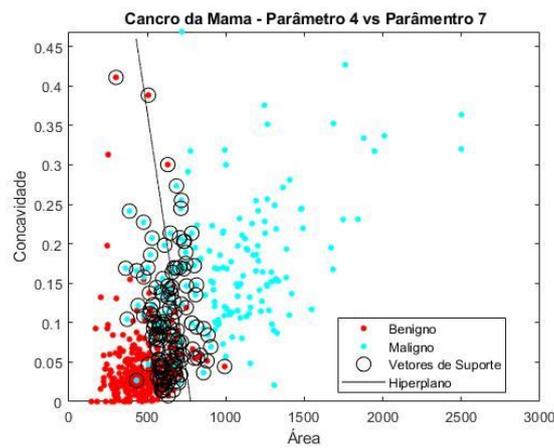


Figura 24: Gráfico representativo da relação entre o parâmetro 4 (Área) e o parâmetro 7 (Concavidade).

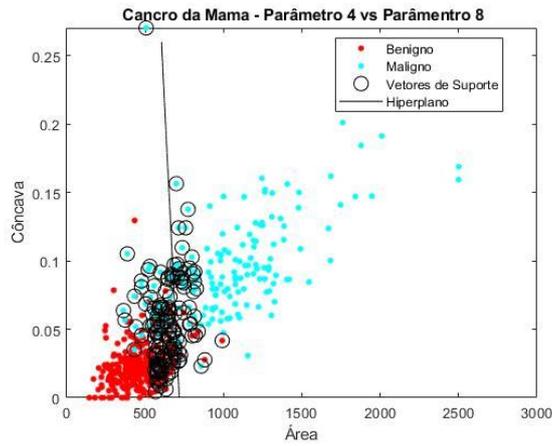


Figura 25: Gráfico representativo da relação entre o parâmetro 4 (Área) e o parâmetro 8 (Côncava).

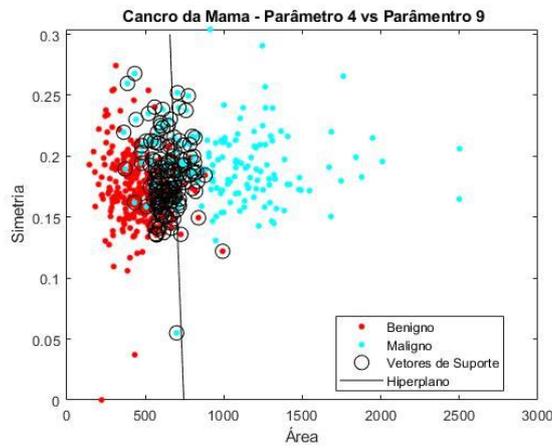


Figura 26: Gráfico representativo da relação entre o parâmetro 4 (Área) e o parâmetro 9 (Simetria).

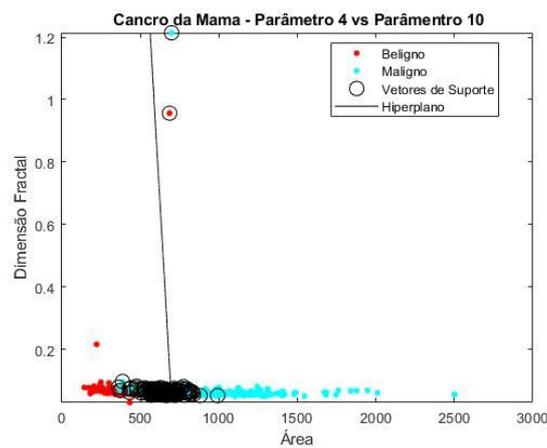


Figura 27: Gráfico representativo da relação entre o parâmetro 4 (Área) e o parâmetro 10 (Dimensão Fractal).

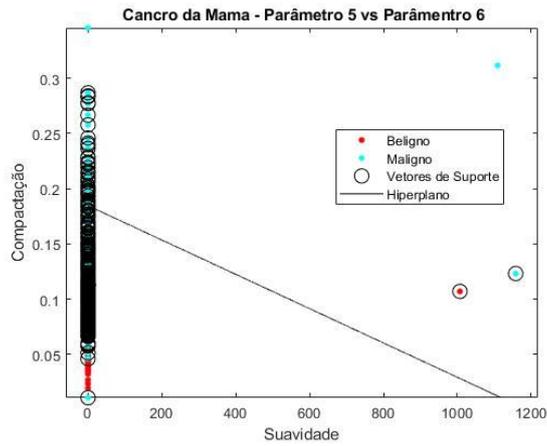


Figura 28: Gráfico representativo da relação entre o parâmetro 5 (Suavidade) e o parâmetro 6 (Compactação).

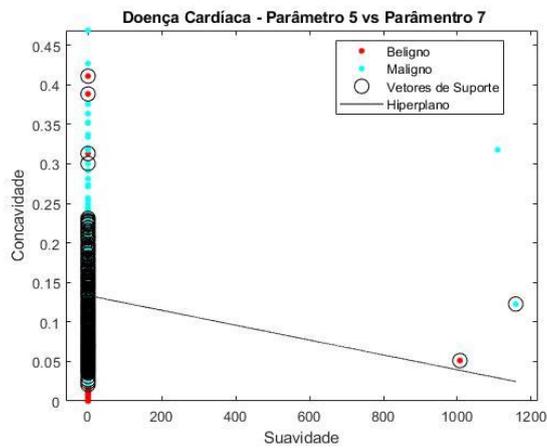


Figura 29: Gráfico representativo da relação entre o parâmetro 5 (Suavidade) e o parâmetro 7 (Concavidade).

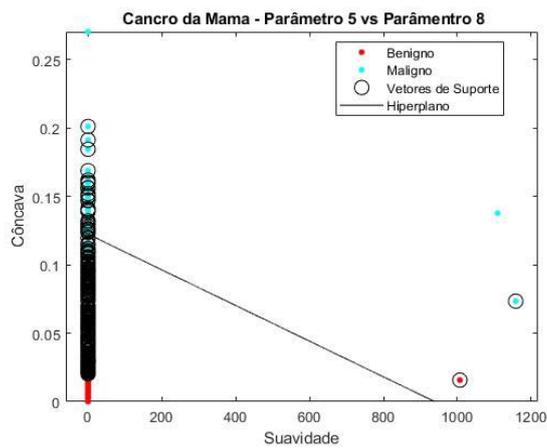


Figura 30: Gráfico representativo da relação entre o parâmetro 5 (Suavidade) e o parâmetro 8 (Cônica).

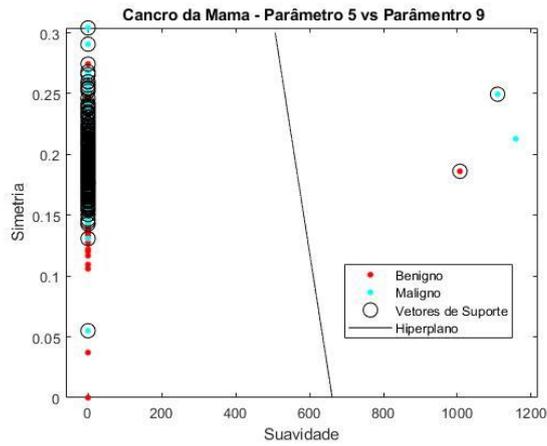


Figura 31: Gráfico representativo da relação entre o parâmetro 5 (Suavidade) e o parâmetro 9 (Simetria).

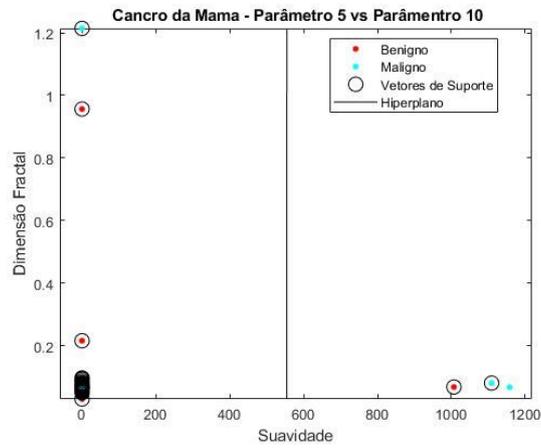


Figura 32: Gráfico representativo da relação entre o parâmetro 5 (Suavidade) e o parâmetro 10 (Dimensão Fractal).

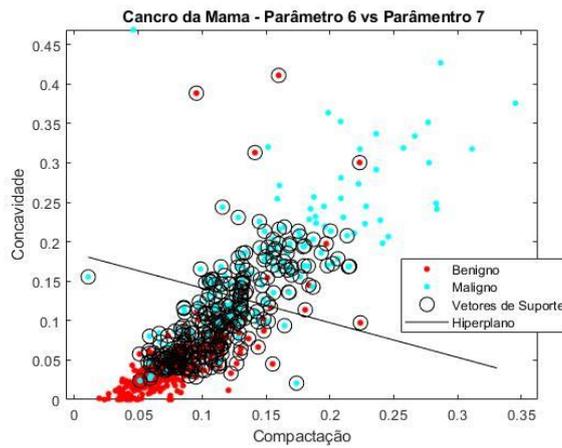


Figura 33: Gráfico representativo da relação entre o parâmetro 6 (Compactação) e o parâmetro 7 (Concavidade).

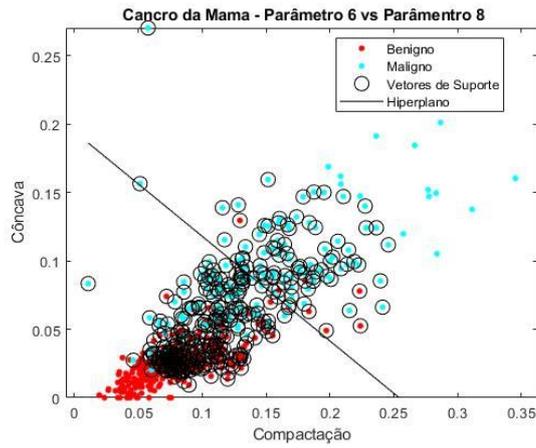


Figura 34: Gráfico representativo da relação entre o parâmetro 6 (Compactação) e o parâmetro 8 (Côncava).

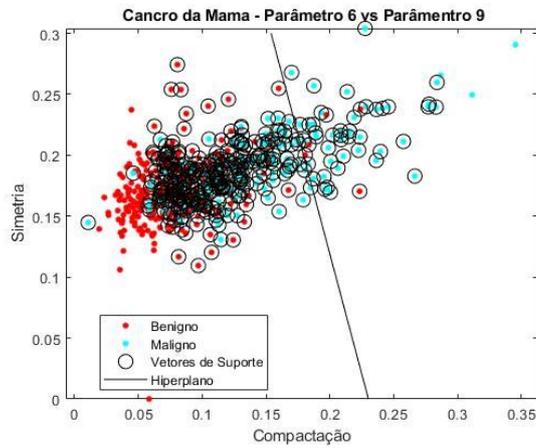


Figura 35: Gráfico representativo da relação entre o parâmetro 6 (Compactação) e o parâmetro 10 (Dimensão Fractal).

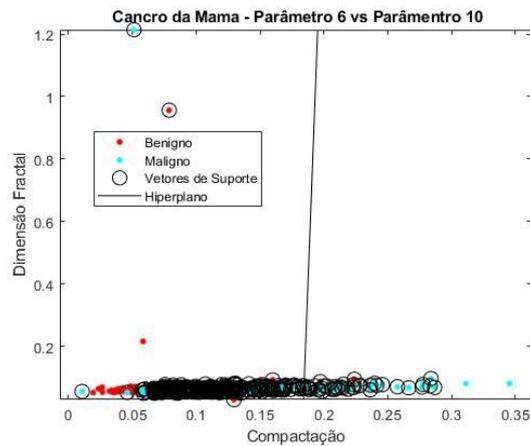


Figura 36: Gráfico representativo da relação entre o parâmetro 6 (Compactação) e o parâmetro 10 (Dimensão Fractal).

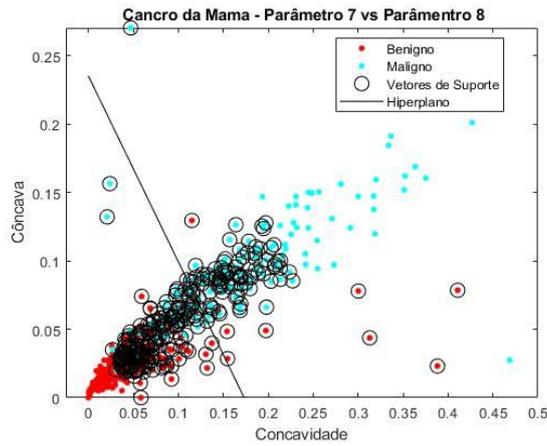


Figura 37: Gráfico representativo da relação entre o parâmetro 7 (Concavidade) e o parâmetro 8 (Côncava).

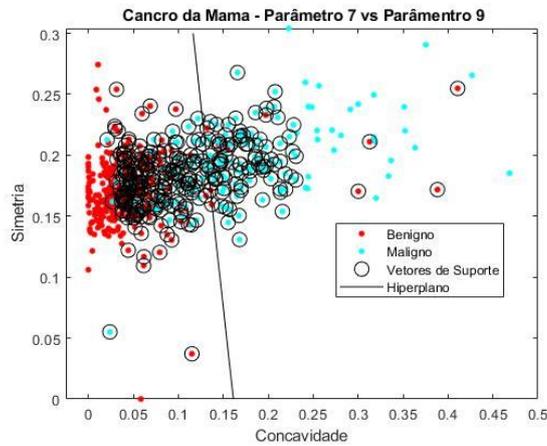


Figura 38: Gráfico representativo da relação entre o parâmetro 7 (Concavidade) e o parâmetro 9 (Simetria).

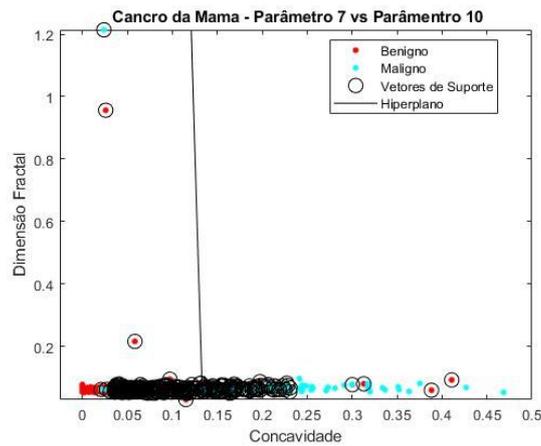


Figura 39: Gráfico representativo da relação entre o parâmetro 7 (Concavidade) e o parâmetro 10 (Dimensão Fractal).

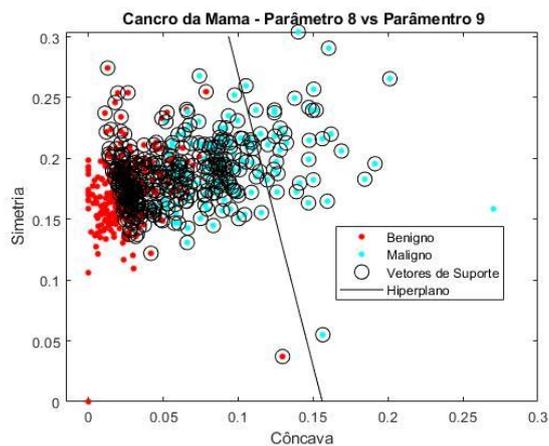


Figura 40: Gráfico representativo da relação entre o parâmetro 8 (Côncava) e o parâmetro 9 (Simetria).

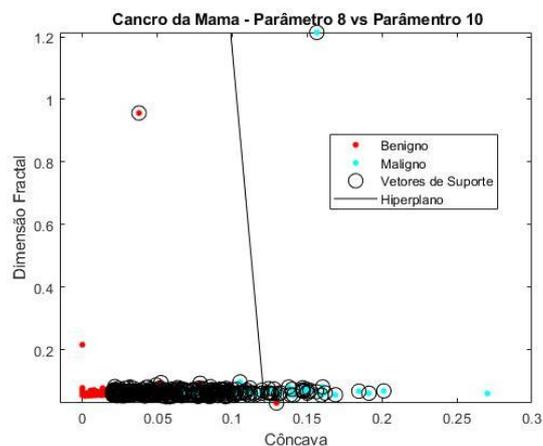


Figura 41: Gráfico representativo da relação entre o parâmetro 8 (Côncava) e o parâmetro 10 (Dimensão Fractal).

Anexo 1C

Tabela 2: Resultados de Previsão Obtidos.

ID Sujeito	Diagnóstico Real	Diagnóstico Previsto	Probabilidade (%)	
			Benigno (0)	Maligno (1)
7	0	0	86.32	13.68
21	0	0	99.90	0.10
29	0	1	29.65	70.35
43	0	0	99.97	0.03
51	0	0	91.78	8.22
56	0	0	67.04	32.96
58	0	0	99.86	0.14
62	0	0	98.75	0.12
73	0	0	83.10	16.90

79	0	0	97.27	27.31
110	0	0	99.96	0.04
113	0	0	98.44	15.60
117	0	1	49.00	51.00
118	0	0	98.07	19.32
129	0	0	96.30	36.96
136	0	0	99.96	0.04
144	0	1	43.28	56.72
150	0	0	99.88	0.12
155	0	0	96.71	3.29
160	0	0	100	0.00
179	0	0	99.69	0.31
184	0	0	81.13	18.87
194	0	0	99.96	0.04
196	0	0	65.22	34.78
201	0	0	94.91	5.09
202	0	0	99.12	0.89
209	0	0	52.11	47.89
217	0	0	99.80	0.20
219	0	0	99.53	0.47
220	0	0	99.98	0.02
222	0	0	97.86	0.21
227	0	0	99.70	0.30
230	0	0	99.35	0.65
239	0	1	38.92	61.08
244	0	0	99.04	0.96
254	0	0	97.92	2.08
255	0	0	97.35	2.65
259	0	0	99.27	0.73
260	0	0	99.69	0.31
265	0	0	94.81	5.19
277	0	0	99.30	0.70
280	0	0	99.98	0.02
301	0	0	99.97	0.03
304	0	0	94.43	5.57
305	0	0	99.70	0.30
309	0	0	96.55	34.46
316	0	0	99.12	0.89
317	0	0	91.39	8.61
318	0	0	95.86	4.14
336	0	1	25.14	74.86
341	0	0	99.96	0.04
350	0	0	98.21	1.79
354	0	0	96.72	3.28
362	1	1	0.05	99.95

368	1	0	64.34	35.66
372	1	1	1.41	98.59
374	1	0	72.70	27.30
387	1	1	0.00	100
388	1	1	0.02	99.98
393	1	0	66.43	33.57
404	1	1	35.12	96.49
406	1	1	25.84	74.16
411	1	1	13.53	86.47
417	1	1	22.09	77.91
419	1	1	7.37	92.63
425	1	1	0.00	100
432	1	1	1.44	98.56
435	1	1	2.96	97.04
443	1	1	0.00	100
444	1	1	0.00	100
447	1	0	96.72	3.28
450	1	1	0.00	100
452	1	0	50.06	49.94
453	1	0	61.36	38.64
474	1	1	8.30	91.70
485	1	1	0.18	99.82
492	1	1	41.84	58.16
494	1	0	90.77	9.23
501	1	1	0.08	99.92
504	1	1	0.02	99.98
511	1	1	22.74	77.26
534	1	1	0.76	99.24
538	1	0	56.98	43.02
558	1	1	0.12	99.88
567	1	1	0.00	100

Anexo 2 – Demência

Anexo 2A

Tabela 3: Conjunto de dados utilizado.

ID Paciente	Idade	CDR	MMSE	MR Delay	n-WBV	Diagnóstico
1	87	0	27	0	0.7	0
2	88	0	28	0	0.7	0
3	90	0	27	538	0.7	0
4	85	0	30	1603	0.7	0
5	78	0	29	0	0.7	0
6	80	0	29	730	0.7	0
7	85	0	29	1456	0.7	0
8	81	0.5	26	617	0.8	0
9	86	0	27	2400	0.8	0
10	64	0	29	828	0.8	0
11	71	0	30	609	0.8	0
12	73	0	30	1234	0.8	0
13	77	0	29	0	0.7	0
14	60	0	30	0	0.8	0
15	88	0	30	0	0.7	0
16	89	0	27	405	0.7	0
17	75	0	29	2369	0.8	0
18	75	0	30	0	0.7	0
19	78	0	30	1510	0.7	0
20	84	0	28	842	0.7	0
21	67	0	27	726	0.8	0
22	71	0	28	0	0.8	0
23	85	0	30	1340	0.7	0
24	70	0	30	873	0.7	0
25	72	0	30	1651	0.7	0
26	79	0	29	0	0.7	0
27	83	0	29	1351	0.7	0
28	68	0	29	451	0.7	0
29	71	0	29	1438	0.7	0
30	73	0	28	2163	0.7	0
31	90	0	29	743	0.7	0
32	82	0	30	432	0.7	0
33	82	0	29	672	0.7	0
34	84	0	29	1415	0.7	0
35	86	0	30	2386	0.7	0
36	70	0	29	0	0.8	0
37	72	0	28	580	0.8	0
38	66	0	30	0	0.7	0

39	73	0	29	1393	0.7	0
40	89	0	28	0	0.7	0
41	75	0	29	680	0.8	0
42	76	0	30	1345	0.8	0
43	61	0	30	0	0.8	0
44	77	0	29	0	0.8	0
45	76	0	30	1631	0.8	0
46	87	0	30	675	0.7	0
47	74	0	30	576	0.8	0
48	78	0	29	1927	0.7	0
49	81	0	28	0	0.8	0
50	74	0	30	647	0.7	0
51	86	0	30	0	0.7	0
52	88	0	30	597	0.7	0
53	76	0	28	0	0.8	0
54	78	0	30	0	0.7	0
55	82	0	29	1591	0.6	0
56	74	0	30	0	0.7	0
57	78	0	27	1146	0.7	0
58	73	0	29	0	0.8	0
59	65	0	30	0	0.8	0
60	91	0	30	561	0.7	0
61	93	0	29	1553	0.7	0
62	82	0	30	1806	0.7	0
63	81	0	29	0	0.7	0
64	66	0	29	0	0.8	0
65	68	0	29	790	0.8	0
66	77	0	28	791	0.7	0
67	77	0	29	0	0.7	0
68	73	0	30	691	0.7	0
69	77	0	30	493	0.8	0
70	68	0	26	0	0.8	0
71	70	0	28	665	0.8	0
72	89	0	29	0	0.8	0
73	90	0	28	600	0.7	0
74	66	0	30	182	0.7	0
75	61	0	30	0	0.8	0
76	63	0	30	763	0.8	0
77	62	0	26	1180	0.7	0
78	87	0	30	0	0.7	1
79	88	0	29	489	0.7	1
80	92	0.5	27	1933	0.7	1
81	86	0	30	0	0.7	1
82	73	0	28	756	0.8	1
83	85	0	29	0	0.7	1

84	87	0.5	24	846	0.7	1
85	75	0.5	30	2002	0.7	1
86	79	0	29	0	0.7	1
87	81	0.5	29	1042	0.7	1
88	84	0.5	29	2153	0.7	1
89	86	0.5	30	2639	0.7	1
90	65	0.5	30	0	0.8	1
91	68	0	30	0	0.8	1
92	87	0	30	774	0.7	1
93	80	0.5	22	1895	0.7	2
94	71	0.5	28	0	0.7	2
95	75	1	27	1281	0.7	2
96	68	0.5	27	0	0.8	2
97	66	0.5	30	0	0.8	2
98	68	0.5	29	854	0.8	2
99	82	0.5	27	0	0.7	2
100	90	0.5	21	0	0.7	2
101	85	0.5	29	1123	0.7	2
102	89	0.5	26	2508	0.7	2
103	83	0.5	25	486	0.7	2
104	86	0.5	27	567	0.7	2
105	66	1	21	248	0.7	2
106	68	1	19	647	0.7	2
107	69	1	4	1233	0.7	2
108	79	0.5	26	212	0.7	2
109	81	0.5	27	490	0.7	2
110	81	0.5	26	830	0.7	2
111	82	0.5	18	1282	0.7	2
112	62	0.5	30	497	0.7	2
113	84	1	28	365	0.7	2
114	75	0.5	22	567	0.7	2
115	71	1	16	584	0.7	2
116	66	0.5	25	0	0.7	2
117	68	0.5	30	580	0.7	2
118	69	0.5	28	1209	0.7	2
119	82	0.5	26	0	0.7	2
120	78	1	21	0	0.7	2
121	72	1	27	563	0.7	2
122	67	0.5	28	661	0.8	2
123	80	0.5	27	0	0.8	2
124	82	0.5	29	0	0.7	2
125	86	0.5	30	1387	0.7	2
126	70	1	22	0	0.7	2
127	65	0.5	17	881	0.7	2
128	78	0.5	20	558	0.7	2

129	75	0.5	28	504	0.7	2
130	76	0.5	27	0	0.7	2
131	71	0.5	27	472	0.7	2
132	74	0.5	28	0	0.7	2
133	75	0.5	30	636	0.7	2
134	67	0.5	29	0	0.8	2
135	76	0.5	26	0	0.7	2
136	73	0.5	30	0	0.7	2
137	75	1	18	764	0.7	2
138	73	0.5	29	0	0.8	2
139	76	0.5	28	759	0.8	2
140	82	0.5	23	0	0.7	2
141	84	0.5	22	621	0.7	2
142	77	1	23	0	0.8	2
143	79	2	25	580	0.8	2
144	75	0.5	30	0	0.7	2
145	70	0.5	26	0	0.7	2
146	73	0.5	28	1343	0.7	2
147	79	0.5	26	0	0.7	2
148	74	0.5	26	0	0.7	2
149	73	0.5	23	0	0.7	2
150	86	0.5	26	2297	0.7	2

Anexo 2B

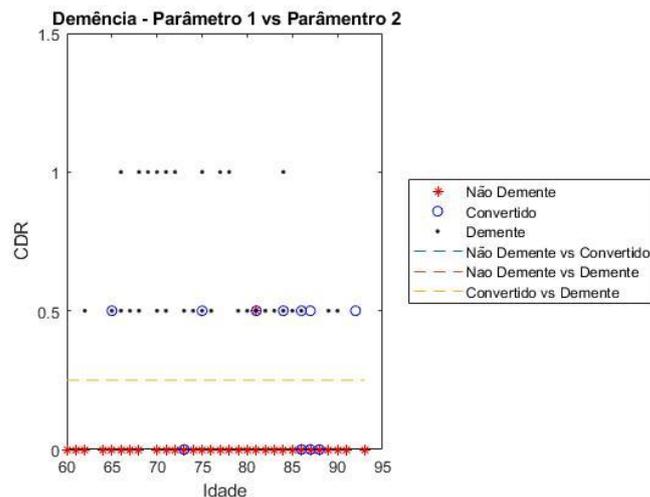


Figura 42: Gráfico representativo da relação entre o parâmetro 1 (Idade) e o parâmetro 2 (CDR).

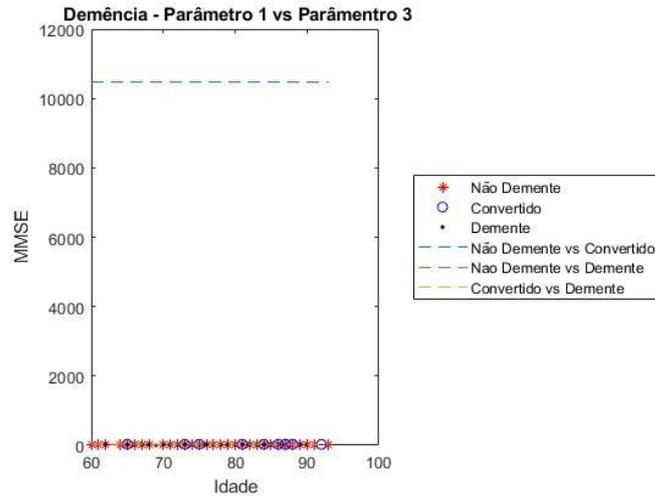


Figura 43: Gráfico representativo da relação entre o parâmetro 1 (Idade) e o parâmetro 3 (MMSE).

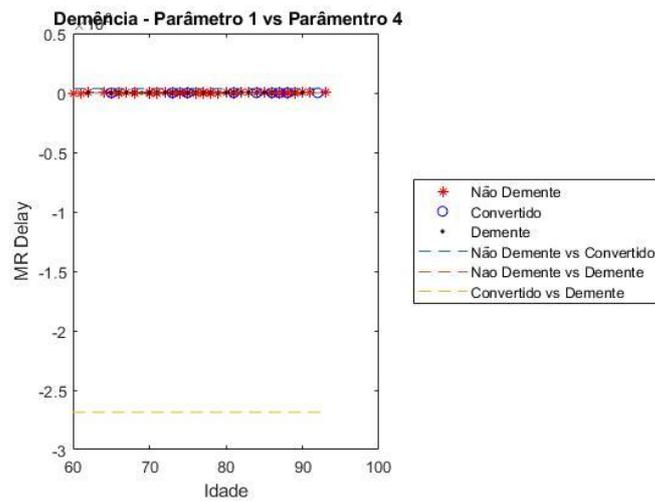


Figura 44: Gráfico representativo da relação entre o parâmetro 1 (Idade) e o parâmetro 4 (MR Delay).

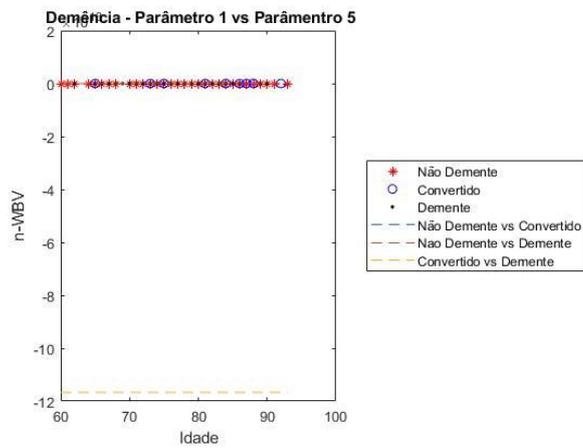


Figura 45: Gráfico representativo da relação entre o parâmetro 1 (Idade) e o parâmetro 5 (n-WBV).

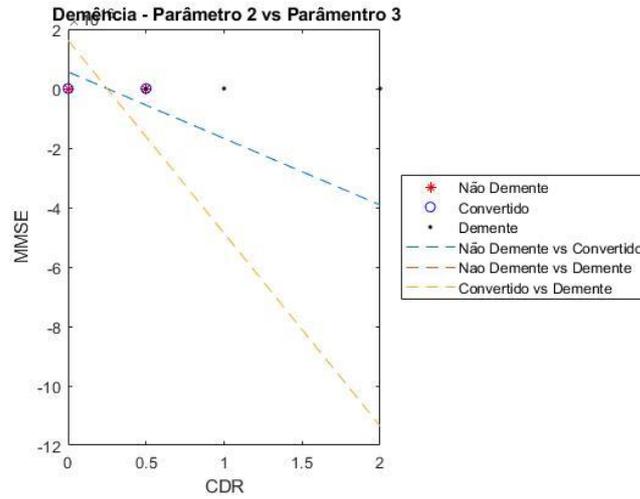


Figura 46: Gráfico representativo da relação entre o parâmetro 2 (CDR) e o parâmetro 3 (MMSE).

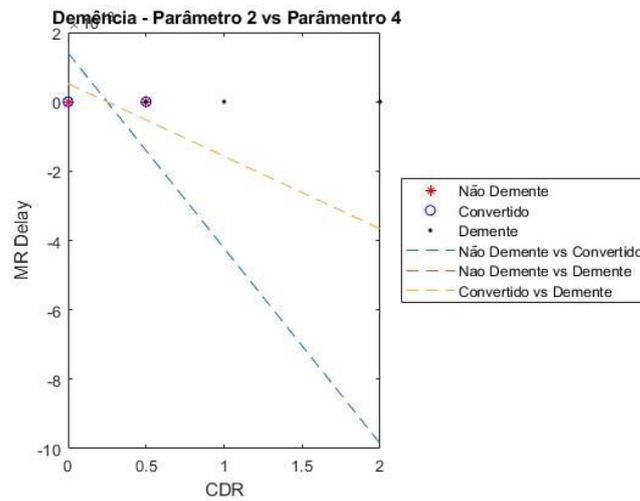


Figura 47: Gráfico representativo da relação entre o parâmetro 2 (CDR) e o parâmetro 4 (MR Delay).

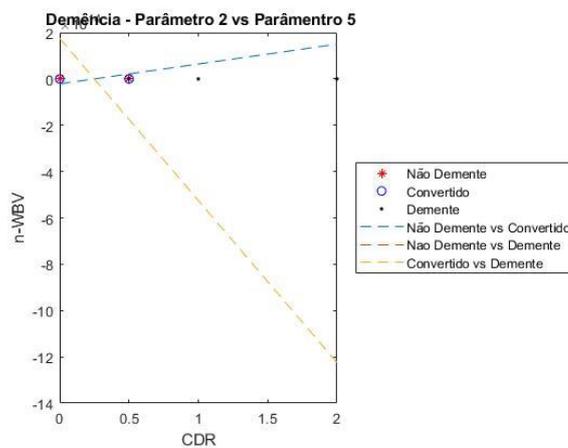


Figura 48: Gráfico representativo da relação entre o parâmetro 2 (CDR) e o parâmetro 5 (n-WBV).

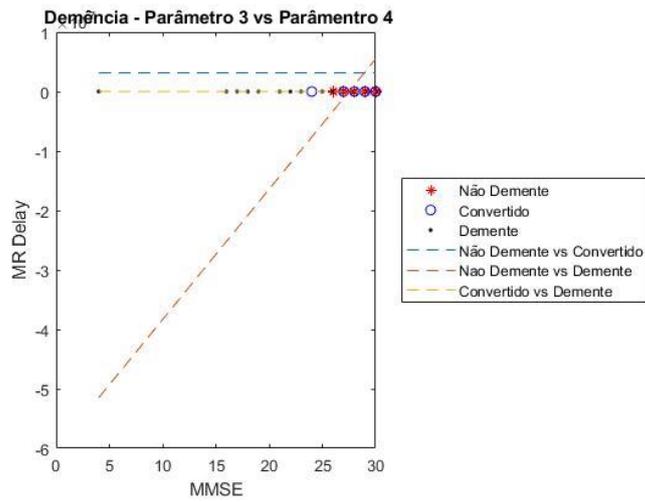


Figura 49: Gráfico representativo da relação entre o parâmetro 3 (MMSE) e o parâmetro 4 (MR Delay).

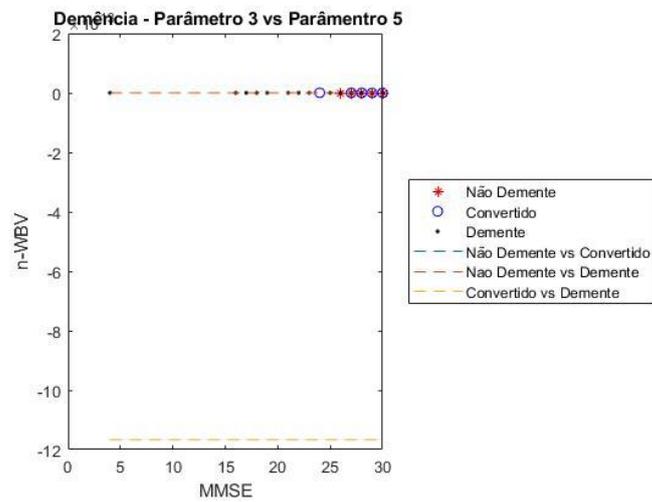


Figura 50: Gráfico representativo da relação entre o parâmetro 3 (MMSE) e o parâmetro 5 (n-WBV).

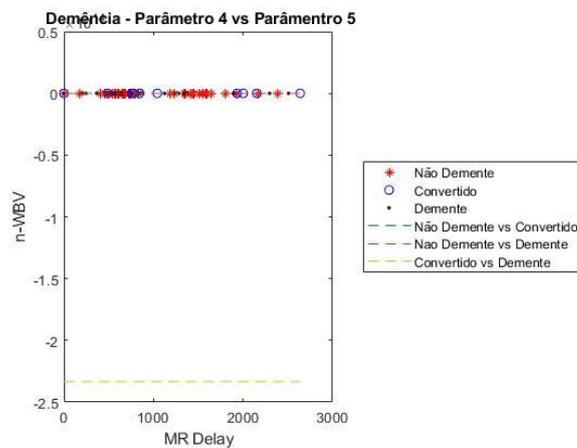


Figura 51: Gráfico representativo da relação entre o parâmetro 4 (MR Delay) e o parâmetro 5 (n-WBV).

Anexo 2C

Tabela 4: Resultados de Previsão Obtidos

ID Paciente	Diagnóstico Real	Diagnóstico Previsto	Probabilidade (%)		
			Não Demente (0)	Convertido (1)	Demente (2)
9	0	0	91.70	7.96	0.33
17	0	0	91.70	7.96	0.33
20	0	0	92.24	6.68	1.08
23	0	0	91.80	7.72	1.28
24	0	0	92.27	6.61	1.12
32	0	0	92.26	6.62	1.11
33	0	0	92.25	6.65	1.10
38	0	0	92.97	4.83	2.21
39	0	0	92.05	7.14	0.81
45	0	0	91.84	7.65	0.52
47	0	0	92.39	6.32	1.29
57	0	0	92.33	64.71	1.20
63	0	0	92.71	5.48	1.81
65	0	0	92.54	5.94	1.52
76	0	0	92.54	5.95	1.51
83	1	0	92.61	5.75	1.64
87	1	0	92.77	5.34	1.89
91	1	0	92.94	4.90	2.16
118	2	2	2.98	15.76	81.26
128	2	2	0.92	3.61	95.47
130	2	2	1.39	6.08	92.53
133	2	2	3.02	16.62	80.36

Anexo 3 – Doença Cardíaca

Anexo 3A

Tabela 5: Conjunto de dados utilizado.

ID Paciente	Idade	Género	Angina	Pressão Arterial	Colesterol	Açúcar em Jejum	ECG em repouso	Frequência Cardíaca	Angina induzida	Pico do Segmento ST	Segmento ST	Número de vasos	Talassemia	Diagnóstico
1	67	1	0	160	286	0	0	108	1	1.5	1	3	2	0
2	67	1	0	120	229	0	0	129	1	2.6	1	2	3	0
3	62	0	0	140	268	0	0	160	0	3.6	0	2	2	0
4	63	1	0	130	254	0	0	147	0	1.4	1	1	3	0
5	53	1	0	140	203	1	0	155	1	3.1	0	0	3	0
6	56	1	2	130	256	1	0	142	1	0.6	1	1	1	0
7	48	1	1	110	229	0	1	168	0	1.0	0	0	3	0
8	58	1	1	120	284	0	0	160	0	1.8	1	0	2	0
9	58	1	2	132	224	0	0	173	0	3.2	2	2	3	0
10	60	1	0	130	206	0	0	132	1	2.4	1	2	3	0
11	40	1	0	110	167	0	0	114	1	2.0	1	0	3	0
12	60	1	0	117	230	1	1	160	1	1.4	2	2	3	0
13	64	1	2	140	335	0	1	158	0	0.0	2	0	2	0
14	43	1	0	120	177	0	0	120	1	2.5	1	0	3	0
15	57	1	0	150	276	0	0	112	1	0.6	1	1	1	0
16	55	1	0	132	353	0	1	132	1	1.2	1	1	3	0
17	65	0	0	150	225	0	0	114	0	1.0	1	3	3	0
18	61	0	0	130	330	0	0	169	0	0.0	2	0	2	0
19	58	1	2	112	230	0	0	165	0	2.5	1	1	3	0

20	50	1	0	150	243	0	0	128	0	2.6	1	0	3	0
21	44	1	0	112	290	0	0	153	0	0.0	2	1	2	0
22	60	1	0	130	253	0	1	144	1	1.4	2	1	3	0
23	54	1	0	124	266	0	0	109	1	2.2	1	1	3	0
24	50	1	2	140	233	0	1	163	0	0.6	1	1	3	0
25	41	1	0	110	172	0	0	158	0	0.0	2	0	3	0
26	51	0	0	130	305	0	1	142	1	1.2	1	0	3	0
27	58	1	0	128	216	0	0	131	1	2.2	1	3	3	0
28	54	1	0	120	188	0	1	113	0	1.4	1	1	3	0
29	60	1	0	145	282	0	0	142	1	2.8	1	2	3	0
30	60	1	2	140	185	0	0	155	0	3.0	1	0	2	0
31	59	1	0	170	326	0	0	140	1	3.4	0	0	3	0
32	46	1	2	150	231	0	1	147	0	3.6	1	0	2	0
33	67	1	0	125	254	1	1	163	0	0.2	1	2	3	0
34	62	1	0	120	267	0	1	99	1	1.8	1	2	3	0
35	65	1	0	110	248	0	0	158	0	0.6	2	2	1	0
36	44	1	0	110	197	0	0	177	0	0.0	2	1	2	0
37	60	1	0	125	258	0	0	141	1	2.8	1	1	3	0
38	58	1	0	150	270	0	0	111	1	0.8	2	0	3	0
39	68	1	2	180	274	1	0	150	1	1.6	1	0	3	0
40	62	0	0	160	164	0	0	145	0	6.2	0	3	3	0
41	52	1	0	128	255	0	1	161	1	0.0	2	1	3	0
42	59	1	0	110	239	0	0	142	1	1.2	1	1	3	0
43	60	0	0	150	258	0	0	157	0	2.6	1	2	3	0
44	49	1	2	120	188	0	1	139	0	2.0	1	3	3	0
45	59	1	0	140	177	0	1	162	1	0.0	2	1	3	0
46	57	1	2	128	229	0	0	150	0	0.4	1	1	3	0

47	61	1	0	120	260	0	1	140	0	1.2	1	0	3	0
48	39	1	0	118	219	0	1	140	0	1.2	1	1	2	0
49	61	0	0	145	307	0	0	146	1	1.0	1	0	3	0
50	56	1	0	125	249	1	0	144	1	1.2	1	1	3	0
51	43	0	0	132	341	1	0	136	1	3.0	1	0	3	0
52	62	0	2	130	263	0	1	97	0	1.2	1	1	3	0
53	63	1	0	130	330	1	0	132	1	1.8	2	3	3	0
54	65	1	0	135	254	0	0	127	0	2.8	1	1	3	0
55	48	1	0	130	256	1	0	150	1	0.0	2	2	3	0
56	63	0	0	150	407	0	0	154	0	4.0	1	3	3	0
57	55	1	0	140	217	0	1	111	1	5.6	0	0	3	0
58	65	1	3	138	282	1	0	174	0	1.4	1	1	2	0
59	56	0	0	200	288	1	0	133	1	4.0	0	2	3	0
60	54	1	0	110	239	0	1	126	1	2.8	1	1	3	0
61	70	1	0	145	174	0	1	125	1	2.6	0	0	3	0
62	62	1	1	120	281	0	0	103	0	1.4	1	1	3	0
63	35	1	0	120	198	0	1	130	1	1.6	1	0	3	0
64	59	1	3	170	288	0	0	159	0	0.2	1	0	3	0
65	64	1	2	125	309	0	1	131	1	1.8	1	0	3	0
66	47	1	2	108	243	0	1	152	0	0.0	2	0	2	0
67	57	1	0	165	289	1	0	124	0	1.0	1	3	3	0
68	55	1	0	160	289	0	0	145	1	0.8	1	1	3	0
69	64	1	0	120	246	0	0	96	1	2.2	0	1	2	0
70	70	1	0	130	322	0	0	109	0	2.4	1	3	2	0
71	51	1	0	140	299	0	1	173	1	1.6	2	0	3	0
72	58	1	0	125	300	0	0	171	0	0.0	2	2	3	0
73	60	1	0	140	293	0	0	170	0	1.2	1	2	3	0

74	77	1	0	125	304	0	0	162	1	0.0	2	3	2	0
75	35	1	0	126	282	0	0	156	1	0.0	2	0	3	0
76	70	1	2	160	269	0	1	112	1	2.9	1	1	3	0
77	59	0	0	174	249	0	1	143	1	0.0	1	0	2	0
78	64	1	0	145	212	0	0	132	0	2.0	1	2	1	0
79	57	1	0	152	274	0	1	88	1	1.2	1	1	3	0
80	56	1	0	132	184	0	0	105	1	2.1	1	1	1	0
81	48	1	0	124	274	0	0	166	0	0.5	1	0	3	0
82	56	0	0	134	409	0	0	150	1	1.9	1	2	3	0
83	66	1	1	160	246	0	1	120	1	0.0	1	3	1	0
84	54	1	1	192	283	0	0	195	0	0.0	2	1	3	0
85	69	1	2	140	254	0	0	146	0	2.0	1	3	3	0
86	51	1	0	140	298	0	1	122	1	4.2	1	3	3	0
87	43	1	0	132	247	1	0	143	1	0.1	1	4	3	0
88	62	0	0	138	294	1	1	106	0	1.9	1	3	2	0
89	67	1	0	100	299	0	0	125	1	0.9	1	2	2	0
90	59	1	3	160	273	0	0	125	0	0.0	2	0	2	0
91	45	1	0	142	309	0	0	147	1	0.0	1	3	3	0
92	58	1	0	128	259	0	0	130	1	3.0	1	2	3	0
93	50	1	0	144	200	0	0	126	1	0.9	1	0	3	0
94	62	0	0	150	244	0	1	154	1	1.4	1	0	2	0
95	38	1	3	120	231	0	1	181	1	3.8	1	0	3	0
96	66	0	0	178	228	1	1	165	1	1.0	1	2	3	0
97	52	1	0	112	230	0	1	160	0	0.0	2	1	2	0
98	53	1	0	123	282	0	1	95	1	2.0	1	2	3	0
99	63	0	0	108	269	0	1	169	1	1.8	1	2	2	0
100	54	1	0	110	206	0	0	108	1	0.0	1	1	2	0

101	66	1	0	112	212	0	0	132	1	0.1	2	1	2	0
102	55	0	0	180	327	0	2	117	1	3.4	1	0	2	0
103	49	1	2	118	149	0	0	126	0	0.8	2	0	3	0
104	54	1	0	122	286	0	0	116	1	3.2	1	2	2	0
105	56	1	0	130	283	1	0	103	1	1.6	0	0	3	0
106	46	1	0	120	249	0	0	144	0	0.8	2	0	3	0
107	61	1	3	134	234	0	1	145	0	2.6	1	2	2	0
108	67	1	0	120	237	0	1	71	0	1.0	1	0	2	0
109	58	1	0	100	234	0	1	156	0	0.1	2	1	3	0
110	47	1	0	110	275	0	0	118	1	1.0	1	1	2	0
111	52	1	0	125	212	0	1	168	0	1.0	2	2	3	0
112	58	1	0	146	218	0	1	105	0	2.0	1	1	3	0
113	57	1	1	124	261	0	1	141	0	0.3	2	0	3	0
114	58	0	1	136	319	1	0	152	0	0.0	2	2	2	0
115	61	1	0	138	166	0	0	125	1	3.6	1	1	2	0
116	42	1	0	136	315	0	1	125	1	1.8	1	0	1	0
117	52	1	0	128	204	1	1	156	1	1.0	1	0	0	0
118	59	1	2	126	218	1	1	134	0	2.2	1	1	1	0
119	40	1	0	152	223	0	1	181	0	0.0	2	0	3	0
120	61	1	0	140	207	0	0	138	1	1.9	2	1	3	0
121	46	1	0	140	311	0	1	120	1	1.8	1	2	3	0
122	59	1	3	134	204	0	1	162	0	0.8	2	2	2	0
123	57	1	1	154	232	0	0	164	0	0.0	2	1	2	0
124	57	1	0	110	335	0	1	143	1	3.0	1	1	3	0
125	55	0	0	128	205	0	2	130	1	2.0	1	1	3	0
126	61	1	0	148	203	0	1	161	0	0.0	2	1	3	0
127	58	1	0	114	318	0	2	140	0	4.4	0	3	1	0

128	58	0	0	170	225	1	0	146	1	2.8	1	2	1	0
129	67	1	2	152	212	0	0	150	0	0.8	1	0	3	0
130	44	1	0	120	169	0	1	144	1	2.8	0	0	1	0
131	63	1	0	140	187	0	0	144	1	4.0	2	2	3	0
132	63	0	0	124	197	0	1	136	1	0.0	1	0	2	0
133	59	1	0	164	176	1	0	90	0	1.0	1	2	1	0
134	57	0	0	140	241	0	1	123	1	0.2	1	0	3	0
135	45	1	3	110	264	0	1	132	0	1.2	1	0	3	0
136	68	1	0	144	193	1	1	141	0	3.4	1	2	3	0
137	57	1	0	130	131	0	1	115	1	1.2	1	1	3	0
138	57	0	1	130	236	0	0	174	0	0.0	1	1	2	0
139	63	1	3	145	233	1	0	150	0	2.3	0	0	1	1
140	37	1	2	130	250	0	1	187	0	3.5	0	0	2	1
141	41	0	1	130	204	0	0	172	0	1.4	2	0	2	1
142	56	1	1	120	236	0	1	178	0	0.8	2	0	2	1
143	57	0	0	120	354	0	1	163	1	0.6	2	0	2	1
144	57	1	0	140	192	0	1	148	0	0.4	1	0	1	1
145	56	0	1	140	294	0	0	153	0	1.3	1	0	2	1
146	44	1	1	120	263	0	1	173	0	0.0	2	0	3	1
147	52	1	2	172	199	1	1	162	0	0.5	2	0	3	1
148	57	1	2	150	168	0	1	174	0	1.6	2	0	2	1
149	54	1	0	120	239	0	1	160	0	1.2	2	0	2	1
150	48	0	2	130	275	0	1	139	0	0.2	2	0	2	1
151	49	1	1	130	266	0	1	171	0	0.6	2	0	2	1
152	64	1	3	110	211	0	0	144	1	1.8	1	0	2	1
153	58	0	3	150	283	1	0	162	0	1.0	2	0	2	1
154	50	0	2	120	219	0	1	158	0	1.6	1	0	2	1

155	58	0	2	120	340	0	1	172	0	0.0	2	0	2	1
156	66	0	3	150	226	0	1	114	0	2.6	0	0	2	1
157	43	1	0	150	247	0	1	171	0	1.5	2	0	2	1
158	69	0	3	140	239	0	1	151	0	1.8	2	2	2	1
159	59	1	0	135	234	0	1	161	0	0.5	1	0	3	1
160	44	1	2	130	233	0	1	179	1	0.4	2	0	2	1
161	42	1	0	140	226	0	1	178	0	0.0	2	0	2	1
162	61	1	2	150	243	1	1	137	1	1.0	1	0	2	1
163	40	1	3	140	199	0	1	178	1	1.4	2	0	3	1
164	71	0	1	160	302	0	1	162	0	0.4	2	2	2	1
165	59	1	2	150	212	1	1	157	0	1.6	2	0	2	1
166	51	1	2	110	175	0	1	123	0	0.6	2	0	2	1
167	65	0	2	140	417	1	0	157	0	0.8	2	1	2	1
168	53	1	2	130	197	1	0	152	0	1.2	0	0	2	1
169	41	0	1	105	198	0	1	168	0	0.0	2	1	2	1
170	65	1	0	120	177	0	1	140	0	0.4	2	0	3	1
171	44	1	1	130	219	0	0	188	0	0.0	2	0	2	1
172	54	1	2	125	273	0	0	152	0	0.5	0	1	2	1
173	51	1	3	125	213	0	0	125	1	1.4	2	1	2	1
174	46	0	2	142	177	0	0	160	1	1.4	0	0	2	1
175	54	0	2	135	304	1	1	170	0	0.0	2	0	2	1
176	54	1	2	150	232	0	0	165	0	1.6	2	0	3	1
177	65	0	2	155	269	0	1	148	0	0.8	2	0	2	1
178	65	0	2	160	360	0	0	151	0	0.8	2	0	2	1
179	51	0	2	140	308	0	0	142	0	1.5	2	1	2	1
180	48	1	1	130	245	0	0	180	0	0.2	1	0	2	1
181	45	1	0	104	208	0	0	148	1	3.0	1	0	2	1

182	53	0	0	130	264	0	0	143	0	0.4	1	0	2	1
183	39	1	2	140	321	0	0	182	0	0.0	2	0	2	1
184	52	1	1	120	325	0	1	172	0	0.2	3	0	2	1
185	44	1	2	140	235	0	0	180	0	0.0	2	0	2	1
186	47	1	2	138	257	0	0	156	0	0.0	2	0	2	1
187	53	0	2	128	216	0	0	115	0	0.0	2	0	0	1
188	53	0	0	138	234	0	0	160	0	0.0	2	0	2	1
189	51	0	2	130	256	0	0	149	0	0.5	2	0	2	1
190	66	1	0	120	302	0	0	151	0	0.4	1	0	2	1
191	62	1	2	130	231	0	1	146	0	1.8	1	3	3	1
192	44	0	2	108	141	0	1	175	0	0.6	1	0	2	1
193	63	0	2	135	252	0	0	172	0	0.0	2	0	2	1
194	52	1	1	134	201	0	1	158	0	0.8	2	1	2	1
195	48	1	0	122	222	0	0	186	0	0.0	2	0	2	1
196	45	1	0	115	260	0	0	185	0	0.0	2	0	2	1
197	34	1	3	118	182	0	0	174	0	0.0	2	0	2	1
198	57	0	0	128	303	0	0	159	0	0.0	2	1	2	1
199	71	0	2	110	265	1	0	130	0	0.0	2	1	2	1
200	54	1	1	108	309	0	1	156	0	0.0	2	0	3	1
201	52	1	3	118	186	0	0	190	0	0.0	1	0	1	1
202	41	1	1	135	203	0	1	132	0	0.0	1	0	1	1
203	58	1	2	140	211	1	0	165	0	0.0	2	0	2	1
204	35	0	0	138	183	0	1	182	0	1.4	2	0	2	1
205	51	1	2	100	222	0	1	143	1	1.2	1	0	2	1
206	45	0	1	130	234	0	0	175	0	0.6	1	0	2	1
207	44	1	1	120	220	0	1	170	0	0.0	2	0	2	1
208	62	0	0	124	209	0	1	163	0	0.0	2	0	2	1

209	54	1	2	120	258	0	0	147	0	0.4	1	0	3	1
210	51	1	2	94	227	0	1	154	1	0.0	2	1	3	1
211	29	1	1	130	204	0	0	202	0	0.0	2	0	2	1
212	51	1	0	140	261	0	0	186	1	0.0	2	0	2	1
213	43	0	2	122	213	0	1	165	0	0.2	1	0	2	1
214	55	0	1	135	250	0	0	161	0	1.4	1	0	2	1
215	51	1	2	125	245	1	0	166	0	2.4	1	0	2	1
216	59	1	1	140	221	0	1	164	1	0.0	2	0	2	1
217	52	1	1	128	205	1	1	184	0	0.0	2	0	2	1
218	58	1	2	105	240	0	0	154	1	0.6	1	0	3	1
219	41	1	2	112	250	0	1	179	0	0.0	2	0	2	1
220	45	1	1	128	308	0	0	170	0	0.0	2	0	2	1
221	60	0	2	102	318	0	1	160	0	0.0	2	1	2	1
222	52	1	3	152	298	1	1	178	0	1.2	1	0	3	1
223	42	0	0	102	265	0	0	122	0	0.6	1	0	2	1
224	67	0	2	115	564	0	0	160	0	1.6	1	0	3	1
225	68	1	2	118	277	0	1	151	0	1.0	2	1	3	1
226	46	1	1	101	197	1	1	156	0	0.0	2	0	3	1
227	54	0	2	132	288	1	0	159	1	0.0	2	1	2	1
228	58	0	0	100	248	0	0	122	0	1.0	1	0	2	1
229	48	1	2	124	255	1	1	175	0	0.0	2	2	2	1
230	57	1	0	132	207	0	1	168	1	0.0	2	0	3	1
231	52	1	2	138	223	0	1	169	0	0.0	2	4	2	1
232	54	0	1	132	288	1	0	159	1	0.0	2	1	2	1
233	45	0	1	112	160	0	1	138	0	0.0	1	0	2	1
234	53	1	0	142	226	0	0	111	1	0.0	2	0	3	1
235	62	0	0	140	394	0	0	157	0	1.2	1	0	2	1

236	52	1	0	108	233	1	1	147	0	0.1	2	3	3	1
237	43	1	2	130	315	0	1	162	0	1.9	2	1	2	1
238	53	1	2	130	246	1	0	173	0	0.0	2	3	2	1
239	42	1	3	148	244	0	0	178	0	0.8	2	2	2	1
240	59	1	3	178	270	0	0	145	0	4.2	0	0	3	1
241	63	0	1	140	195	0	1	179	0	0.0	2	2	2	1
242	42	1	2	120	240	1	1	194	0	0.8	0	0	3	1
243	50	1	2	129	196	0	1	163	0	0.0	2	0	2	1
244	68	0	2	120	211	0	0	115	0	1.5	1	0	2	1
245	69	1	3	160	234	1	0	131	0	0.1	1	1	2	1
246	45	0	0	138	236	0	0	152	1	0.2	1	0	2	1
247	50	0	1	120	244	0	1	162	0	1.1	2	0	2	1
248	50	0	0	110	254	0	0	159	0	0.0	2	0	2	1
249	64	0	0	180	235	0	1	154	1	0.0	2	0	2	1
250	57	1	2	150	126	1	1	173	0	0.2	2	1	3	1
251	64	0	2	140	313	0	1	133	0	0.2	2	0	3	1
252	43	1	0	110	211	0	1	161	0	0.0	2	0	3	1
253	55	1	1	130	262	0	1	155	0	0.0	2	0	2	1
254	37	0	2	120	215	0	1	170	0	0.0	2	0	2	1
255	41	1	2	130	214	0	0	168	0	2.0	1	0	2	1
256	56	1	3	120	193	0	0	162	0	1.9	1	0	3	1
257	46	0	1	105	204	0	1	172	0	0.0	2	0	2	1
258	46	0	0	138	243	0	0	152	1	0.0	1	0	2	1
259	64	0	0	130	303	0	1	122	0	2.0	1	2	2	1
260	59	1	0	138	271	0	0	182	0	0.0	2	0	2	1
261	41	0	2	112	268	0	0	172	1	0.0	2	0	2	1
262	54	0	2	108	267	0	0	167	0	0.0	2	0	2	1

263	39	0	2	94	199	0	1	179	0	0.0	2	0	2	1
264	34	0	1	118	210	0	1	192	0	0.7	2	0	2	1
265	47	1	0	112	204	0	1	143	0	0.1	2	0	2	1
266	67	0	2	152	277	0	1	172	0	0.0	2	1	2	1
267	52	0	2	136	196	0	0	169	0	0.1	1	0	2	1
268	74	0	1	120	269	0	0	121	1	0.2	2	1	2	1
269	54	0	2	160	201	0	1	163	0	0.0	2	1	2	1
270	49	0	1	134	271	0	1	162	0	0.0	1	0	2	1
271	42	1	1	120	295	0	1	162	0	0.0	2	0	2	1
272	41	1	1	110	235	0	1	153	0	0.0	2	0	2	1
273	41	0	1	126	306	0	1	163	0	0.0	2	0	2	1
274	49	0	0	130	269	0	1	163	0	0.0	2	0	2	1
275	60	0	2	120	178	1	1	96	0	0.0	2	0	2	1
276	62	1	1	128	208	1	0	140	0	0.0	2	0	2	1
277	57	1	0	110	201	0	1	126	1	1.5	1	0	1	1
278	64	1	0	128	263	0	1	105	1	0.2	1	1	3	1
279	51	0	2	120	295	0	0	157	0	0.6	2	0	2	1
280	43	1	0	115	303	0	1	181	0	1.2	1	0	2	1
281	42	0	2	120	209	0	1	173	0	0.0	1	0	2	1
282	67	0	0	106	223	0	1	142	0	0.3	1	1	2	1
283	76	0	2	140	197	0	2	116	0	1.1	1	0	2	1
284	70	1	1	156	245	0	0	143	0	0.0	2	0	2	1
285	44	0	2	118	242	0	1	149	0	0.3	1	1	2	1
286	60	0	3	150	240	0	1	171	0	0.9	2	0	2	1
287	44	1	2	120	226	0	1	169	0	0.0	2	0	2	1
288	42	1	2	130	180	0	1	150	0	0.0	2	0	2	1
289	66	1	0	160	228	0	0	138	0	2.3	2	0	1	1

290	71	0	0	112	149	0	1	125	0	1.6	1	0	2	1
291	64	1	3	170	227	0	0	155	0	0.6	1	0	3	1
292	66	0	2	146	278	0	0	152	0	0.0	1	1	2	1
293	39	0	2	138	220	0	1	152	0	0.0	1	0	2	1
294	58	0	0	130	197	0	1	131	0	0.6	1	0	2	1
295	47	1	2	130	253	0	1	179	0	0.0	2	0	2	1
296	35	1	1	122	192	0	1	174	0	0.0	2	0	2	1
297	58	1	1	125	220	0	1	144	0	0.4	1	4	3	1
298	56	1	1	130	221	0	0	163	0	0.0	2	0	3	1
299	56	1	1	120	240	0	1	169	0	0.0	0	0	2	1
300	55	0	1	132	242	0	1	166	0	1.2	2	0	2	1
301	41	1	1	120	157	0	1	182	0	0.0	2	0	2	1
302	38	1	2	138	175	0	1	173	0	0.0	2	4	2	1
303	38	1	2	138	175	0	1	173	0	0.0	2	4	2	1

Anexo 3B

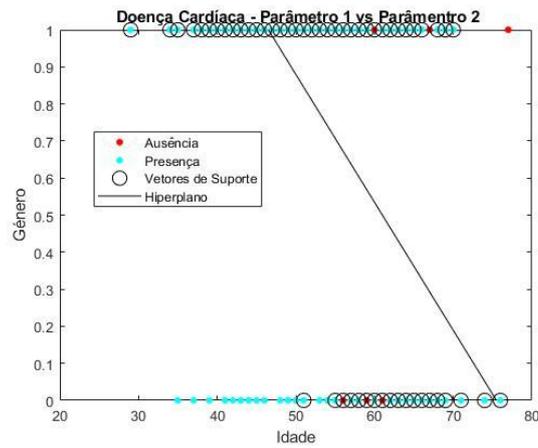


Figura 52: Gráfico representativo da relação entre o parâmetro 1 (Idade) e o parâmetro 2 (Gênero).

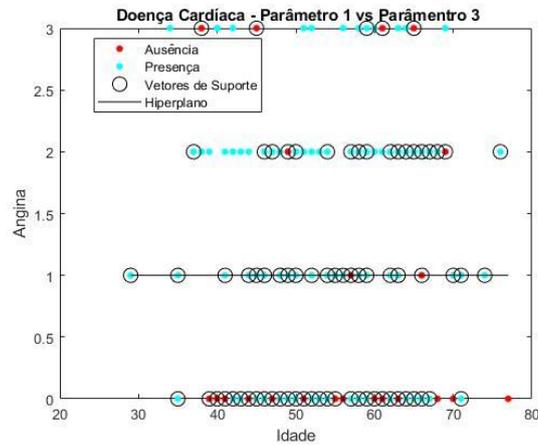


Figura 53: Gráfico representativo da relação entre o parâmetro 1 (Idade) e o parâmetro 3 (Angina).

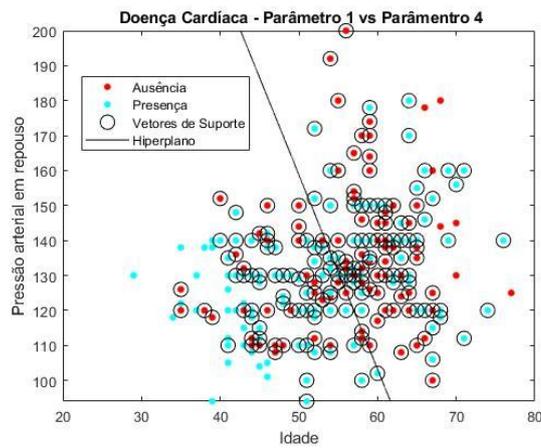


Figura 54: Gráfico representativo da relação entre o parâmetro 1 (Idade) e o parâmetro 4 (Pressão arterial em repouso).

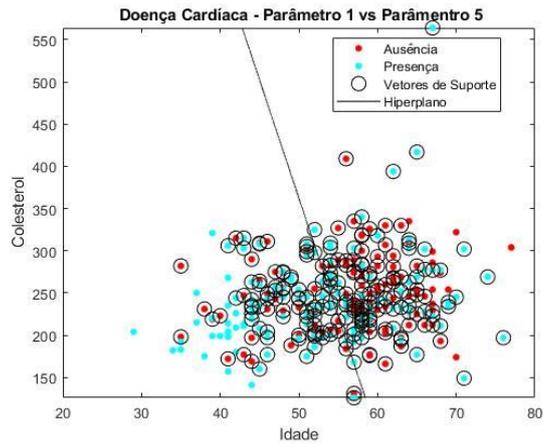


Figura 55: Gráfico representativo da relação entre o parâmetro 1 (Idade) e o parâmetro 5 (Colesterol).

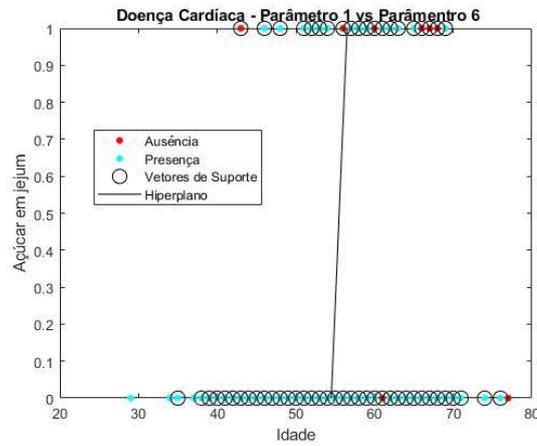


Figura 56: Gráfico representativo da relação entre o parâmetro 1 (Idade) e o parâmetro 6 (Açúcar em jejum).

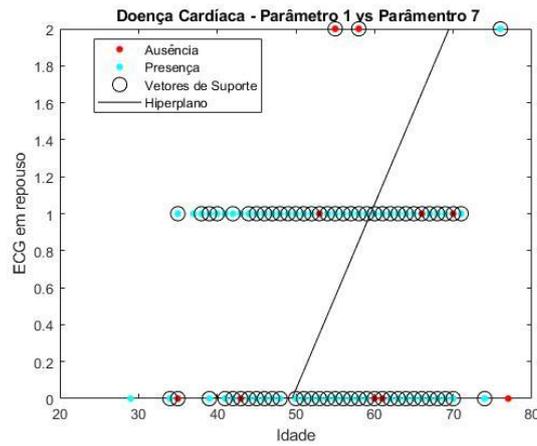


Figura 57: Gráfico representativo da relação entre o parâmetro 1 (Idade) e o parâmetro 7 (ECG em repouso).

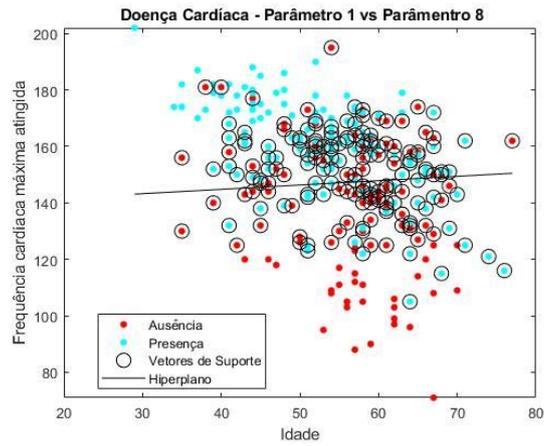


Figura 58: Gráfico representativo da relação entre o parâmetro 1 (Idade) e o parâmetro 8 (Frequência cardíaca máxima atingida).

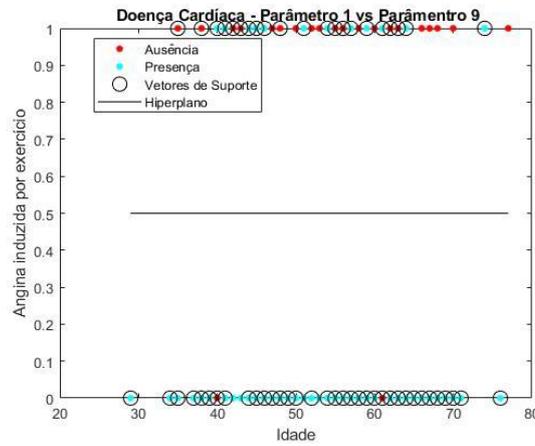


Figura 59: Gráfico representativo da relação entre o parâmetro 1 (Idade) e o parâmetro 9 (Angina induzida por exercício).

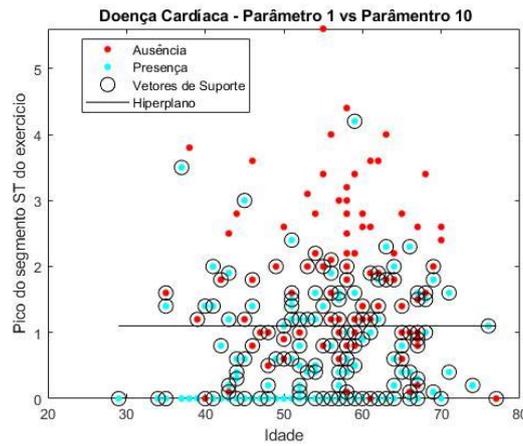


Figura 60: Gráfico representativo da relação entre o parâmetro 1 (Idade) e o parâmetro 10 (Pico do segmento ST do exercício).

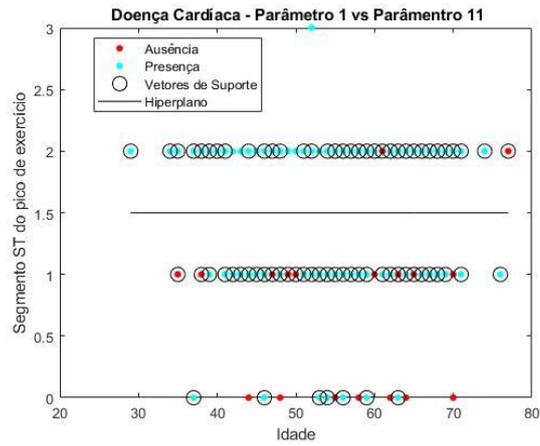


Figura 61: Gráfico representativo da relação entre o parâmetro 1 (Idade) e o parâmetro 11 (Segmento ST do pico de exercício).

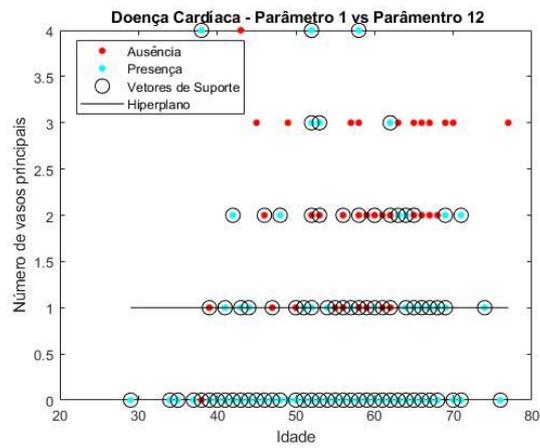


Figura 62: Gráfico representativo da relação entre o parâmetro 1 (Idade) e o parâmetro 12 (Número de vasos principais).

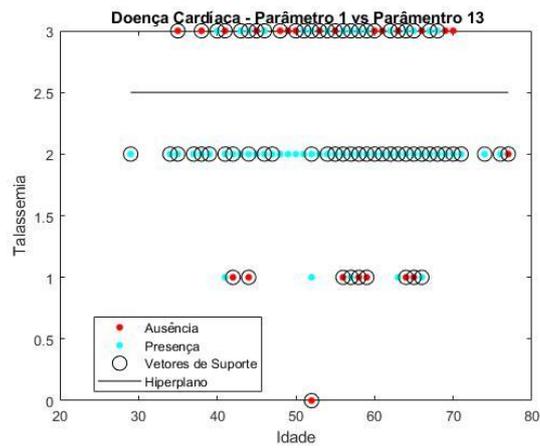


Figura 63: Gráfico representativo da relação entre o parâmetro 1 (Idade) e o parâmetro 13 (Talassemia).

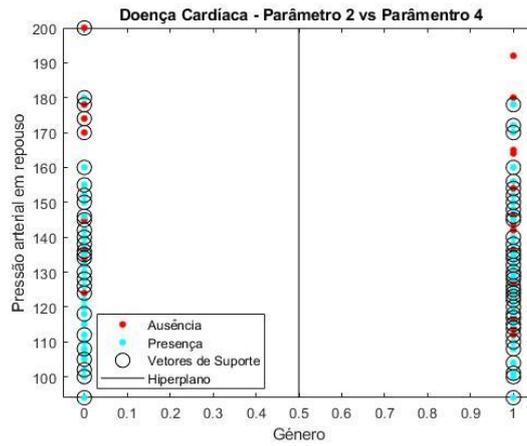


Figura 64: Gráfico representativo da relação entre o parâmetro 2 (Gênero) e o parâmetro 4 (Pressão arterial em repouso).

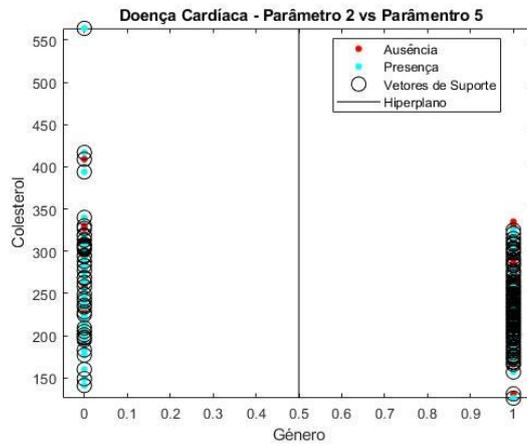


Figura 65: Gráfico representativo da relação entre o parâmetro 2 (Gênero) e o parâmetro 4 (Colesterol).

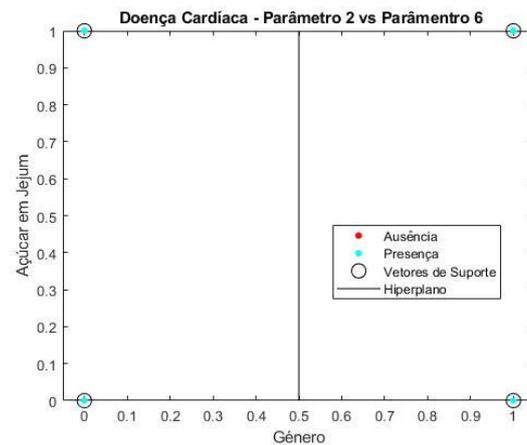


Figura 66: Gráfico representativo da relação entre o parâmetro 2 (Gênero) e o parâmetro 5 (Açúcar em Jejum).

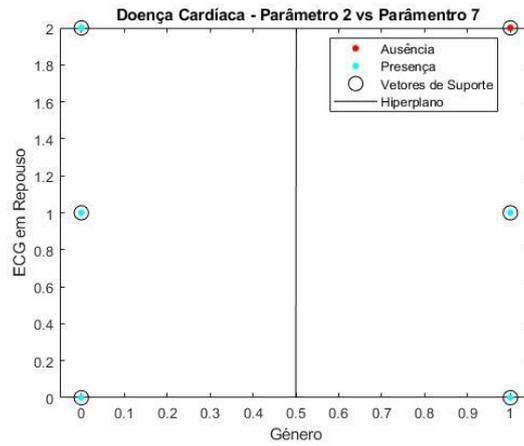


Figura 67: Gráfico representativo da relação entre o parâmetro 2 (Gênero) e o parâmetro 8 (ECG em repouso).

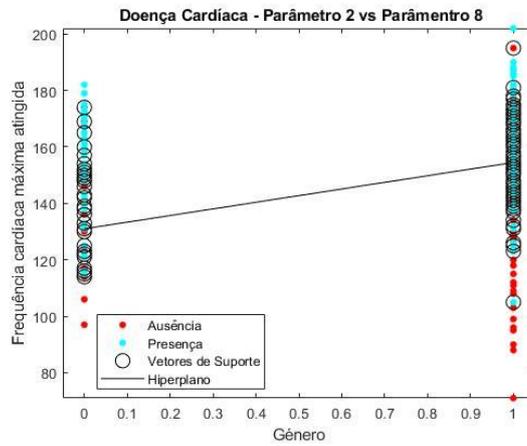


Figura 68: Gráfico representativo da relação entre o parâmetro 2 (Gênero) e o parâmetro 8 (Frequência cardíaca máxima atingida).

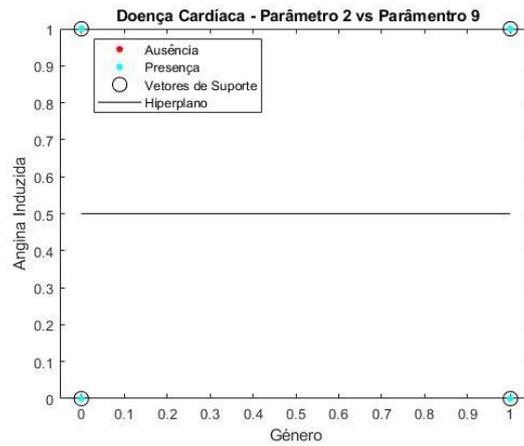


Figura 69: Gráfico representativo da relação entre o parâmetro 2 (Gênero) e o parâmetro 9 (Angina induzida).

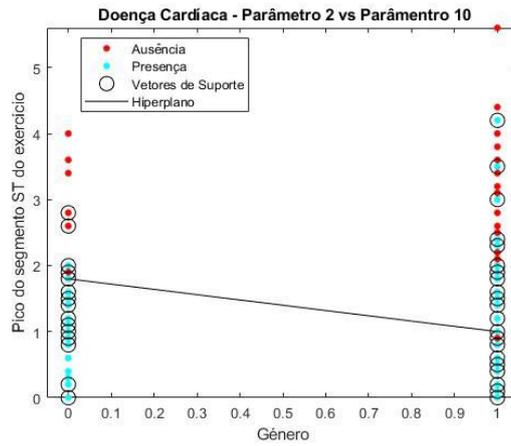


Figura 70: Gráfico representativo da relação entre o parâmetro 2 (Gênero) e o parâmetro 10 (Pico do Segmento ST).

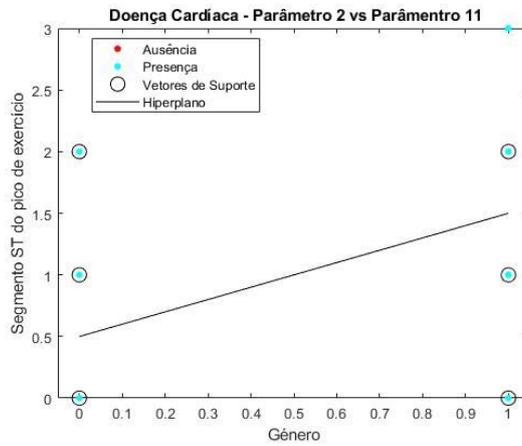


Figura 71: Gráfico representativo da relação entre o parâmetro 2 (Gênero) e o parâmetro 11 (Segmento ST).

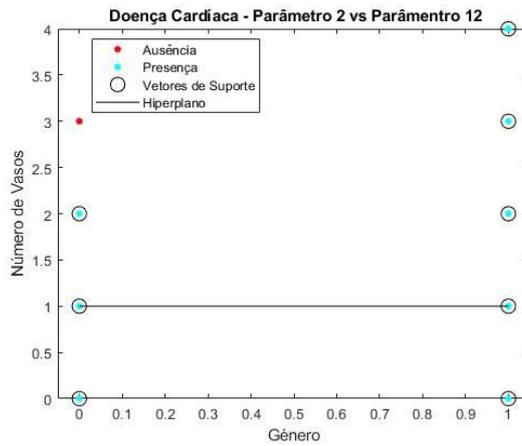


Figura 72: Gráfico representativo da relação entre o parâmetro 2 (Gênero) e o parâmetro 12 (Número de Vasos).

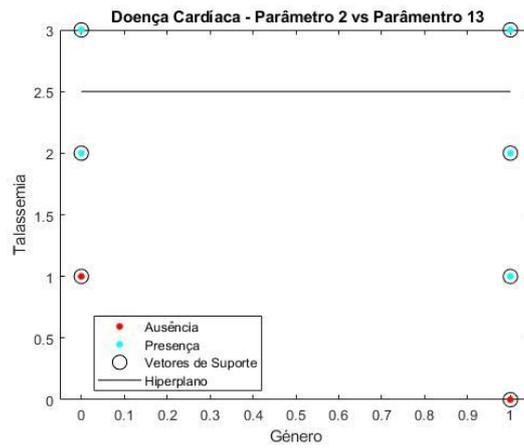


Figura 73: Gráfico representativo da relação entre o parâmetro 2 (Gênero) e o parâmetro 13 (Talassemia).

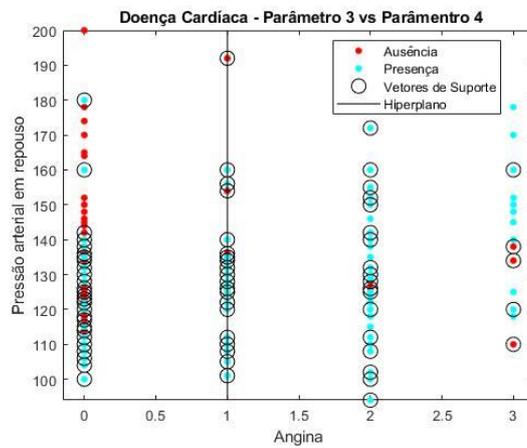


Figura 74: Gráfico representativo da relação entre o parâmetro 3 (Angina) e o parâmetro 4 (Pressão arterial em repouso).

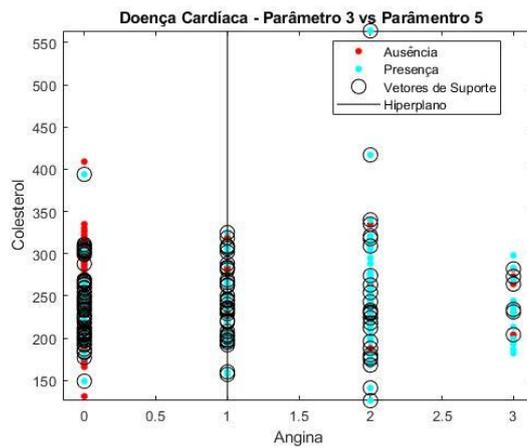


Figura 75: Gráfico representativo da relação entre o parâmetro 3 (Angina) e o parâmetro 5 (Colesterol).

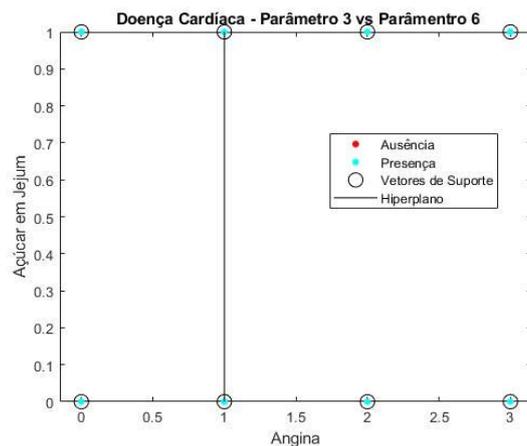


Figura 76: Gráfico representativo da relação entre o parâmetro 3 (Angina) e o parâmetro 6 (Açúcar em jejum).

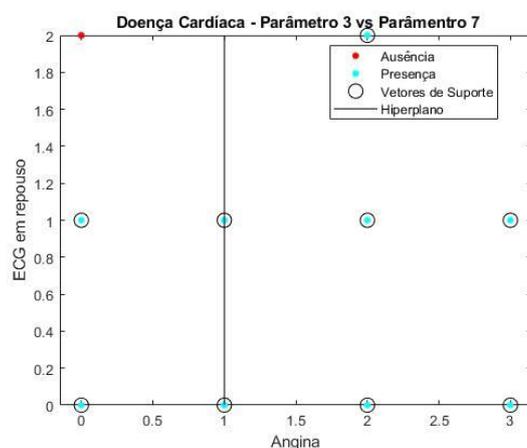


Figura 77: Gráfico representativo da relação entre o parâmetro 3 (Angina) e o parâmetro 7 (ECG em repouso).

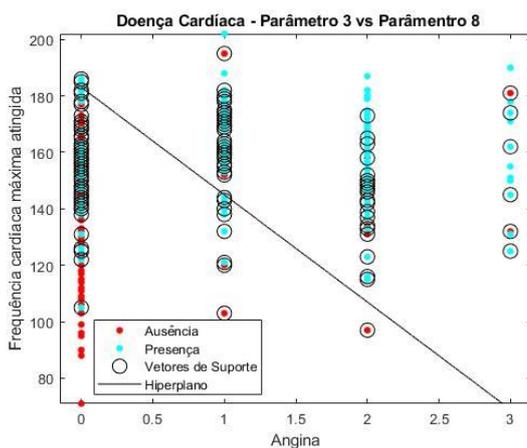


Figura 78: Gráfico representativo da relação entre o parâmetro 3 (Angina) e o parâmetro 8 (Frequência cardíaca máxima atingida).

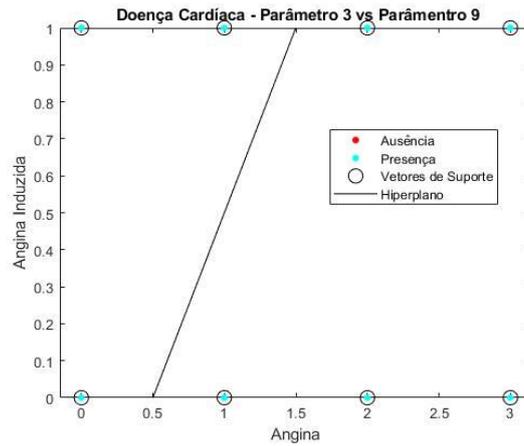


Figura 79: Gráfico representativo da relação entre o parâmetro 3 (Angina) e o parâmetro 9 (Angina induzida).

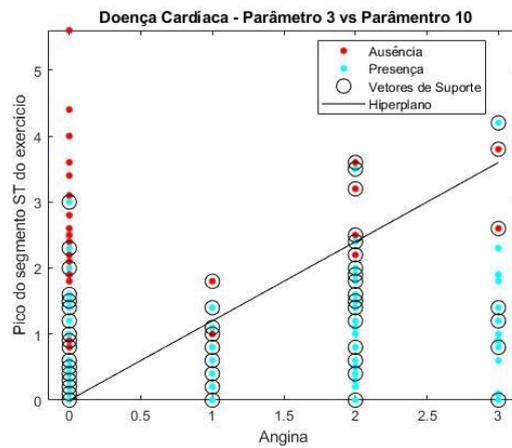


Figura 80: Gráfico representativo da relação entre o parâmetro 3 (Angina) e o parâmetro 10 (Pico do Segmento ST).

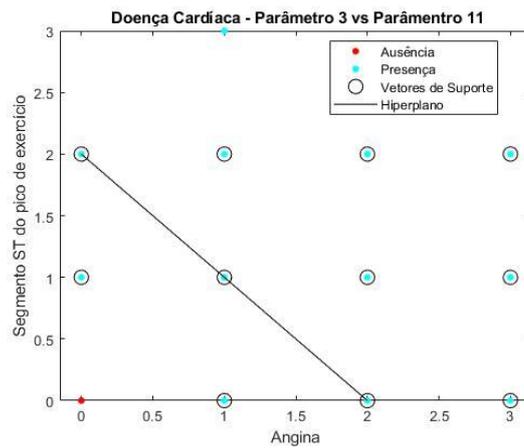


Figura 81: Gráfico representativo da relação entre o parâmetro 3 (Angina) e o parâmetro 11 (Segmento ST).

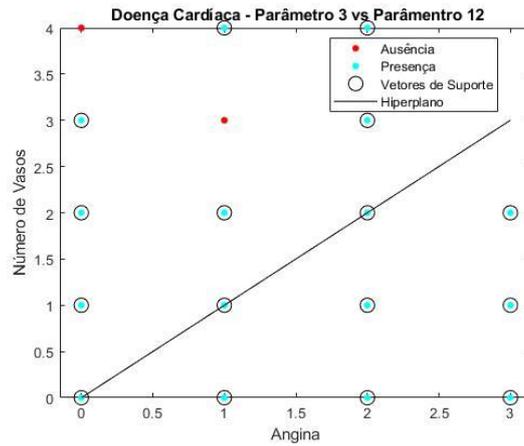


Figura 82: Gráfico representativo da relação entre o parâmetro 3 (Angina) e o parâmetro 12 (Número de Vasos).

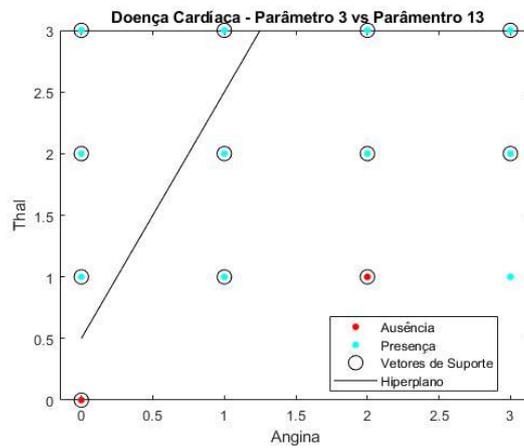


Figura 83: Gráfico representativo da relação entre o parâmetro 3 (Angina) e o parâmetro 13 (Talassemia).

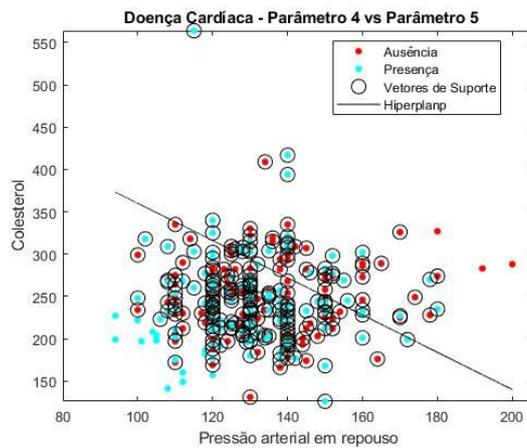


Figura 84: Gráfico representativo da relação entre o parâmetro 4 (Pressão arterial em repouso) e o parâmetro 5 (Colesterol).

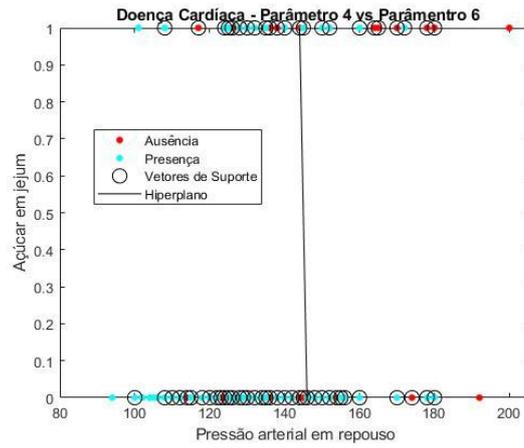


Figura 85: Gráfico representativo da relação entre o parâmetro 4 (Pressão arterial em repouso) e o parâmetro 6 (Açúcar em jejum).

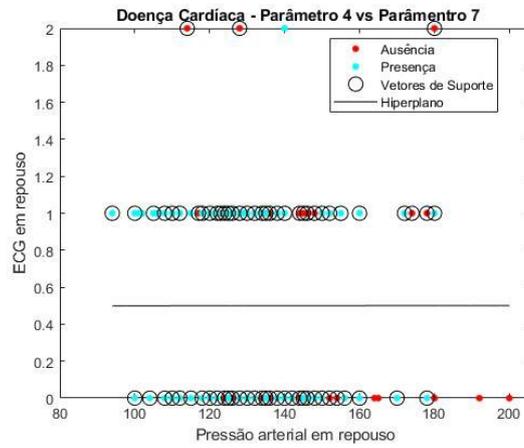


Figura 86: Gráfico representativo da relação entre o parâmetro 4 (Pressão arterial em repouso) e o parâmetro 7 (ECG em repouso).

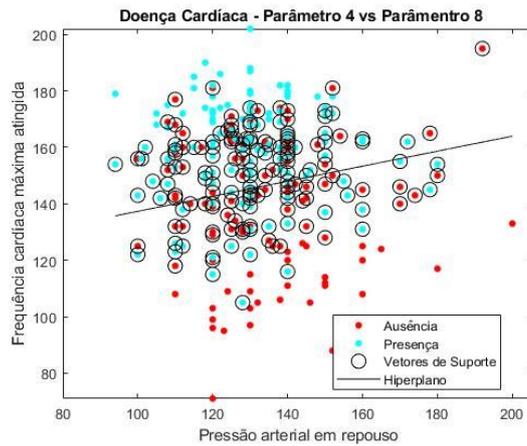


Figura 87: Gráfico representativo da relação entre o parâmetro 4 (Pressão arterial em repouso) e o parâmetro 8 (Frequência cardíaca máxima atingida).

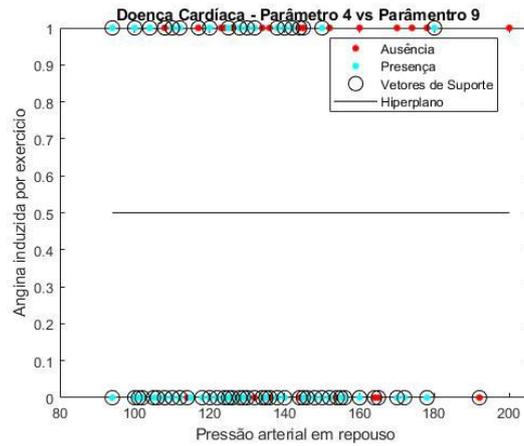


Figura 88: Gráfico representativo da relação entre o parâmetro 4 (Pressão arterial em repouso) e o parâmetro 9 (Angina induzida por exercício).

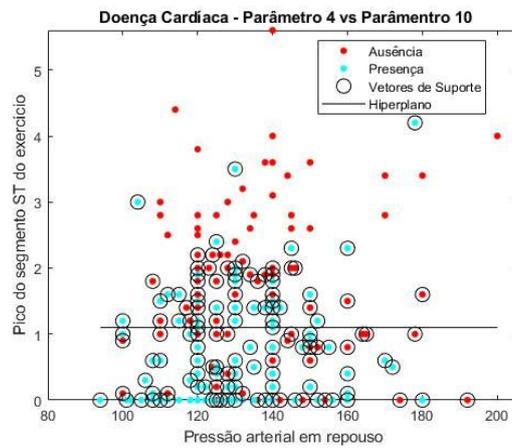


Figura 89: Gráfico representativo da relação entre o parâmetro 4 (Pressão arterial em repouso) e o parâmetro 10 (Pico do segmento ST do exercício).

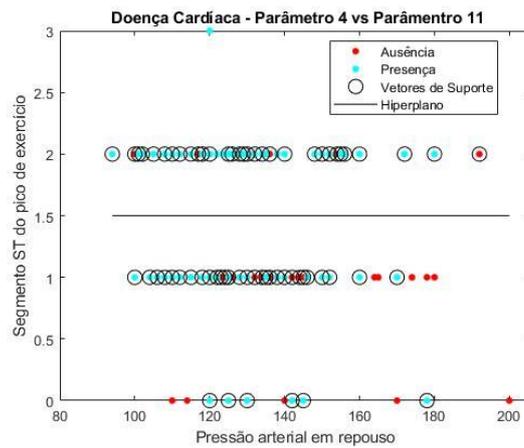


Figura 90: Gráfico representativo da relação entre o parâmetro 4 (Pressão arterial em repouso) e o parâmetro 11 (Segmento ST do pico de exercício).

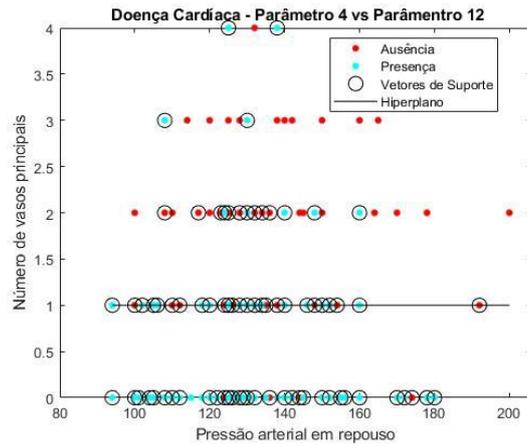


Figura 91: Gráfico representativo da relação entre o parâmetro 4 (Pressão arterial em repouso) e o parâmetro 12 (Número de vasos principais).

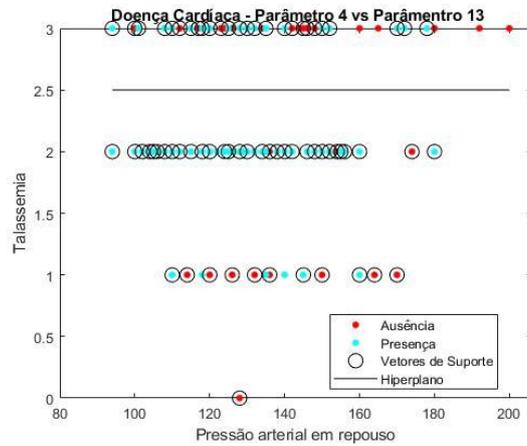


Figura 92: Gráfico representativo da relação entre o parâmetro 4 (Pressão arterial em repouso) e o parâmetro 13 (Talassemia).

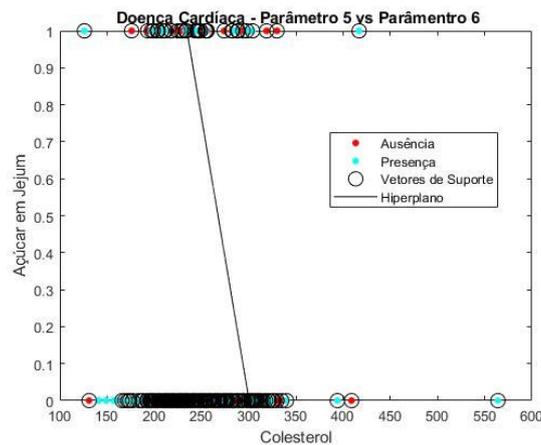


Figura 93: Gráfico representativo da relação entre o parâmetro 5 (Colesterol) e o parâmetro 6 (Açúcar em jejum).

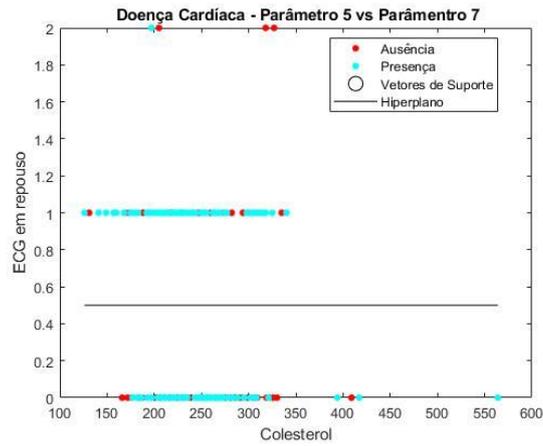


Figura 94: Gráfico representativo da relação entre o parâmetro 5 (Colesterol) e o parâmetro 7 (ECG em repouso).

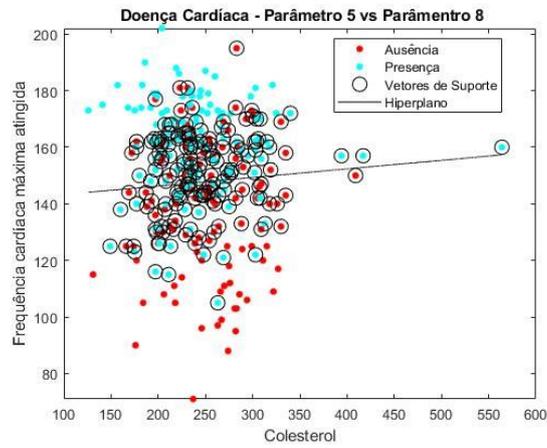


Figura 95: Gráfico representativo da relação entre o parâmetro 5 (Colesterol) e o parâmetro 8 (Frequência cardíaca máxima atingida).

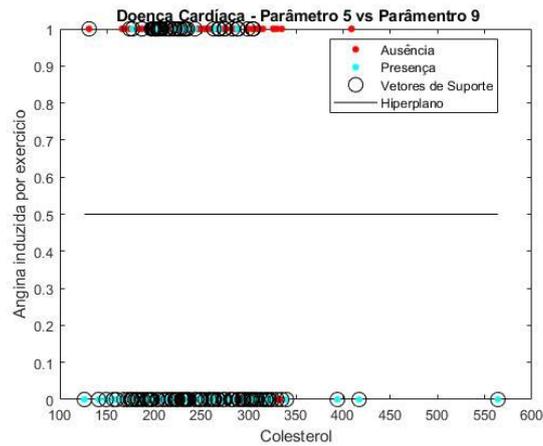


Figura 96: Gráfico representativo da relação entre o parâmetro 5 (Colesterol) e o parâmetro 9 (Angina induzida por exercício).

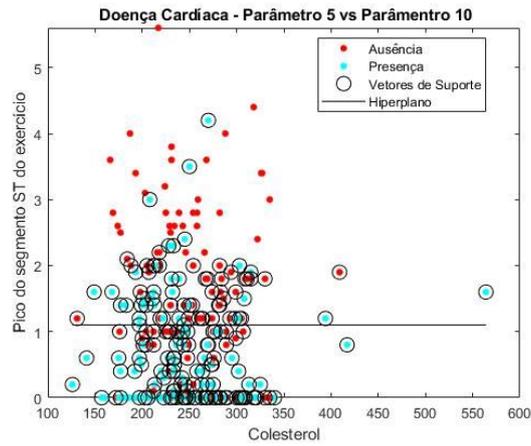


Figura 97: Gráfico representativo da relação entre o parâmetro 5 (Colesterol) e o parâmetro 10 (Pico do segmento ST do exercício).

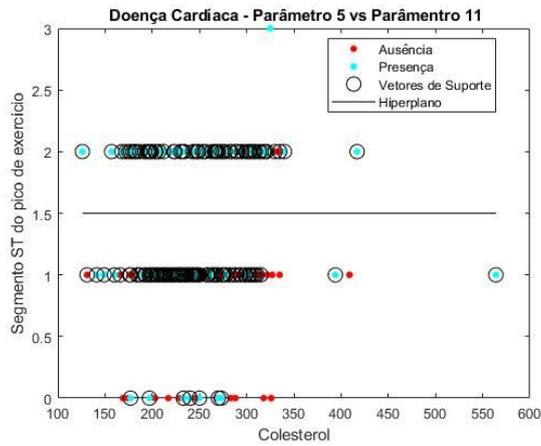


Figura 98: Gráfico representativo da relação entre o parâmetro 5 (Colesterol) e o parâmetro 11 (Segmento ST do pico de exercício).

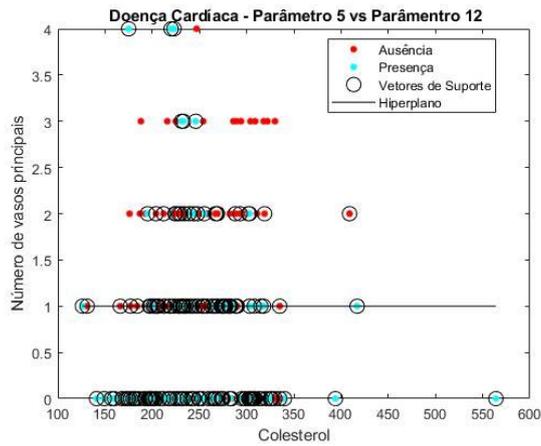


Figura 99: Gráfico representativo da relação entre o parâmetro 5 (Colesterol) e o parâmetro 12 (Número de vasos principais).

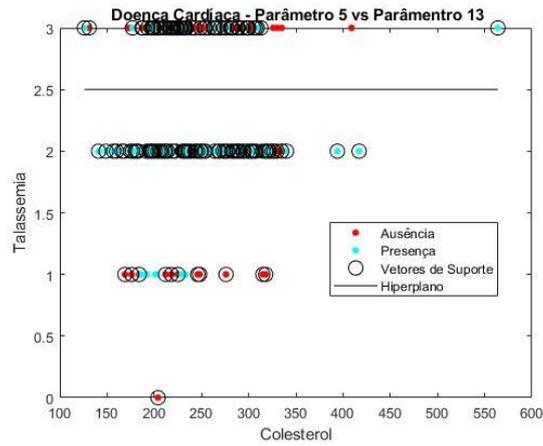


Figura 100: Gráfico representativo da relação entre o parâmetro 5 (Colesterol) e o parâmetro 13 (Talassemia).

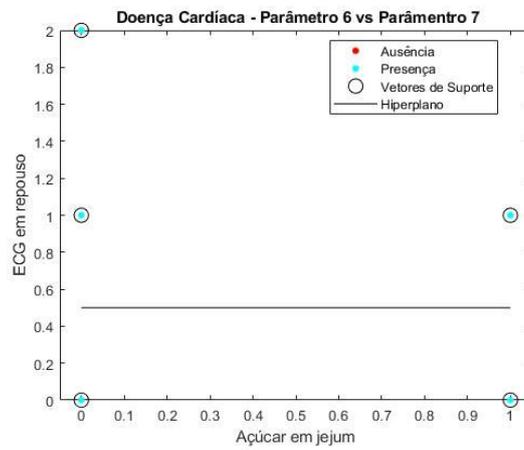


Figura 101: Gráfico representativo da relação entre o parâmetro 6 (Açúcar em jejum) e o parâmetro 7 (ECG em repouso).

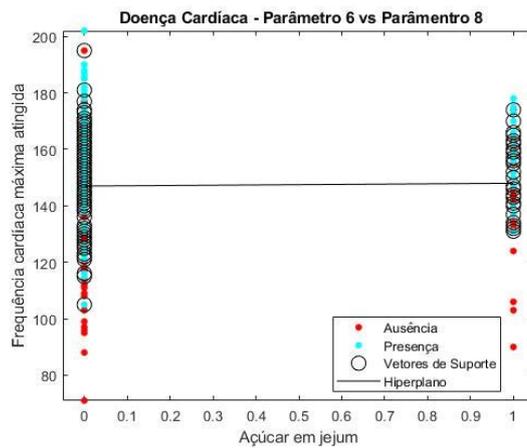


Figura 102: Gráfico representativo da relação entre o parâmetro 6 (Açúcar em jejum) e o parâmetro 8 (Frequência cardíaca máxima atingida).

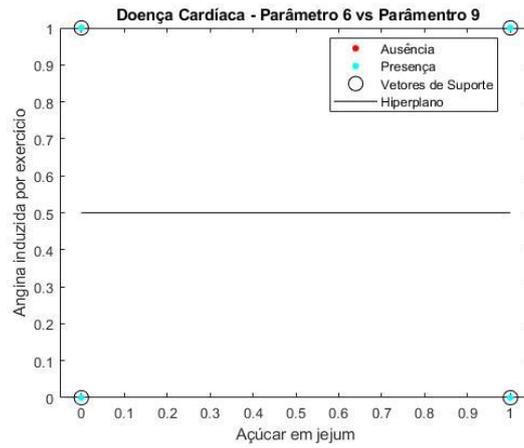


Figura 103: Gráfico representativo da relação entre o parâmetro 6 (Açúcar em jejum) e o parâmetro 9 (Angina induzida por exercício).

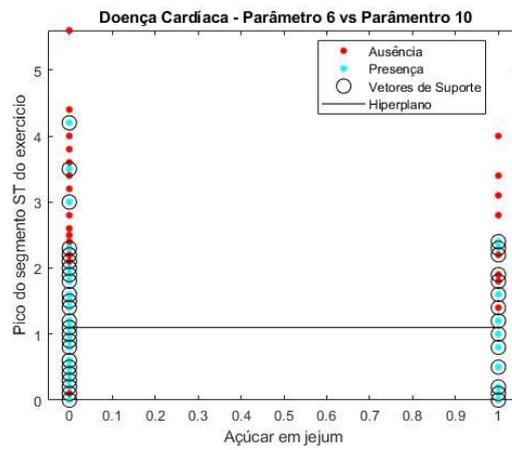


Figura 104: Gráfico representativo da relação entre o parâmetro 6 (Açúcar em jejum) e o parâmetro 10 (Pico do segmento ST do exercício).

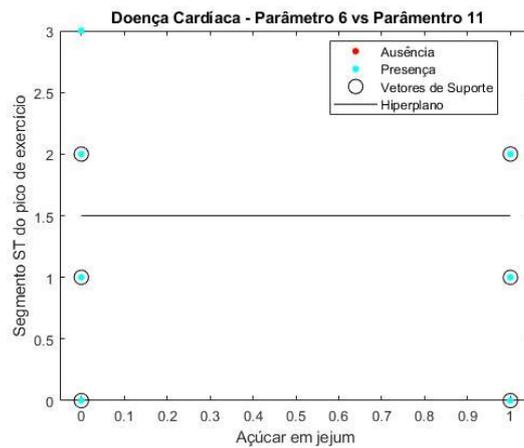


Figura 105: Gráfico representativo da relação entre o parâmetro 6 (Açúcar em jejum) e o parâmetro 11 (Segmento ST do pico de exercício).

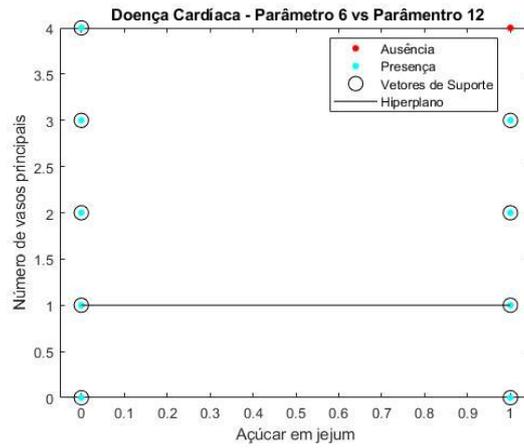


Figura 106: Gráfico representativo da relação entre o parâmetro 6 (Açúcar em jejum) e o parâmetro 12 (Número de vasos principais).

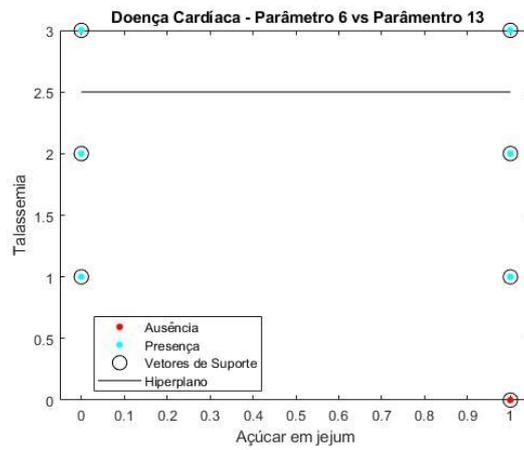


Figura 107: Gráfico representativo da relação entre o parâmetro 6 (Açúcar em jejum) e o parâmetro 13 (Talassemia).

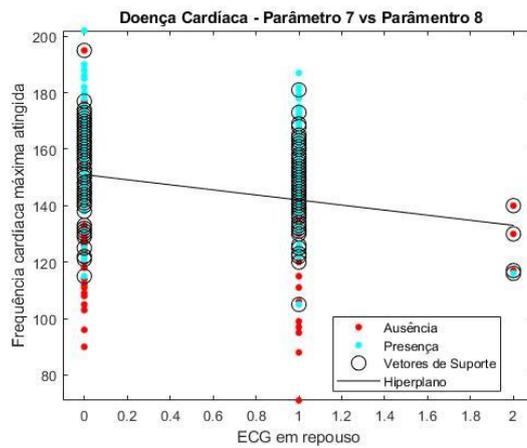


Figura 108: Gráfico representativo da relação entre o parâmetro 7 (ECG em repouso) e o parâmetro 8 (Frequência cardíaca máxima atingida).

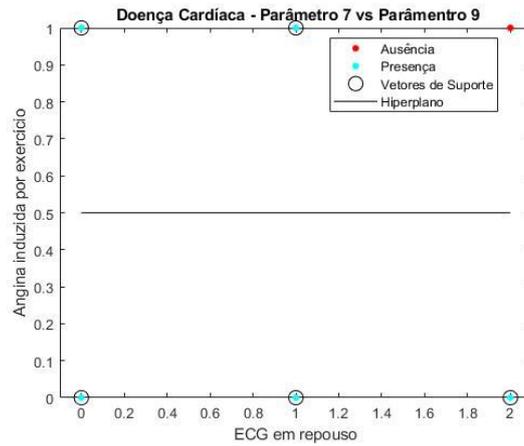


Figura 109: Gráfico representativo da relação entre o parâmetro 7 (ECG em repouso) e o parâmetro 9 (Angina induzida por exercício).

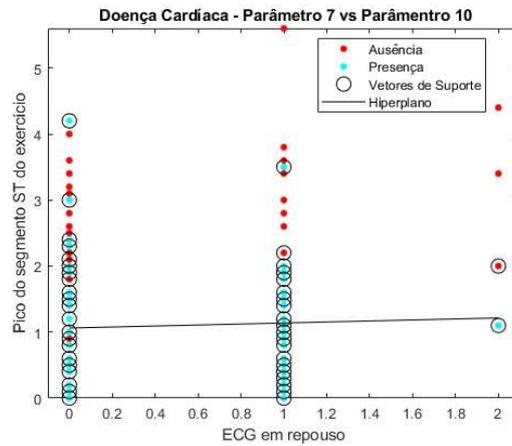


Figura 110: Gráfico representativo da relação entre o parâmetro 7 (ECG em repouso) e o parâmetro 10 (Pico do segmento ST do exercício).

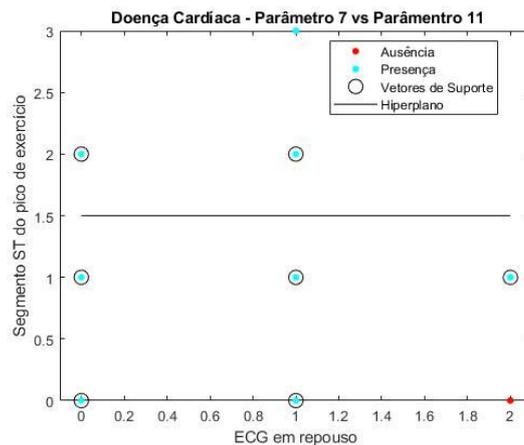


Figura 111: Gráfico representativo da relação entre o parâmetro 7 (ECG em repouso) e o parâmetro 11 (Segmento ST do pico de exercício).

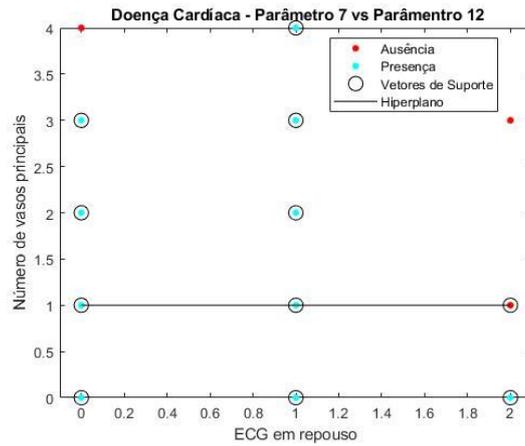


Figura 112: Gráfico representativo da relação entre o parâmetro 7 (ECG em repouso) e o parâmetro 12 (Número de vasos principais).

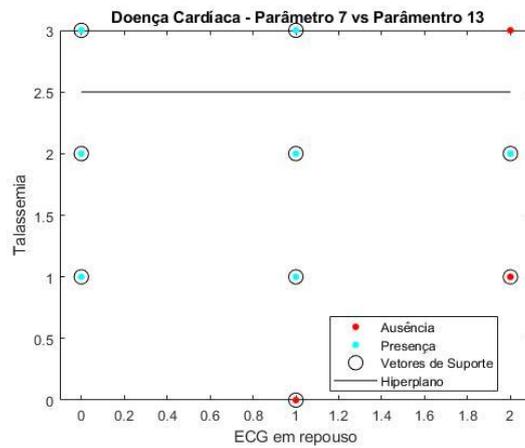


Figura 113: Gráfico representativo da relação entre o parâmetro 7 (ECG em repouso) e o parâmetro 13 (Talassemia).

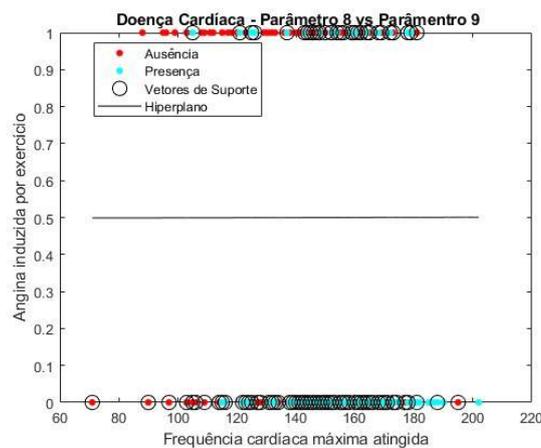


Figura 114: Gráfico representativo da relação entre o parâmetro 8 (Frequência cardíaca máxima atingida) e o parâmetro 9 (Angina induzida por exercício).

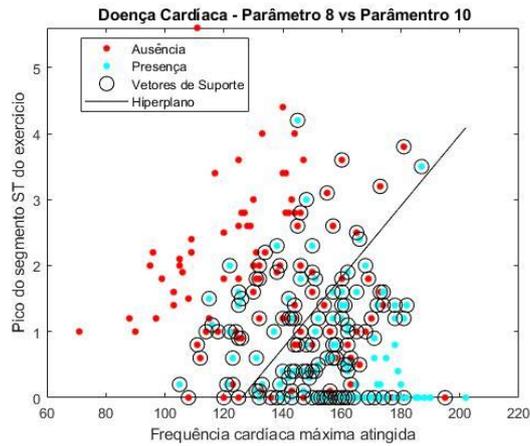


Figura 115: Gráfico representativo da relação entre o parâmetro 8 (Frequência cardíaca máxima atingida) e o parâmetro 10 (Pico do segmento ST do exercício).

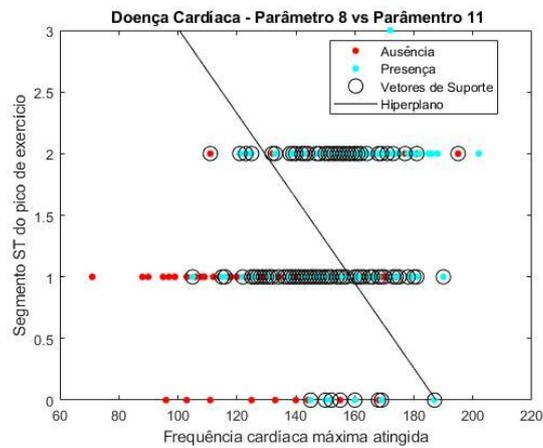


Figura 116: Gráfico representativo da relação entre o parâmetro 8 (Frequência cardíaca máxima atingida) e o parâmetro 11 (Segmento ST do pico do exercício).

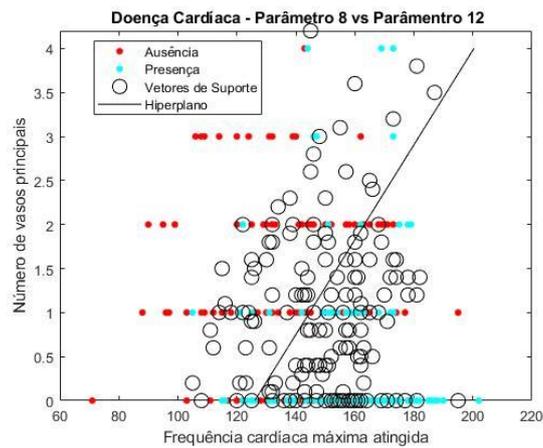


Figura 117: Gráfico representativo da relação entre o parâmetro 8 (Frequência cardíaca máxima atingida) e o parâmetro 12 (Angina induzida por exercício).

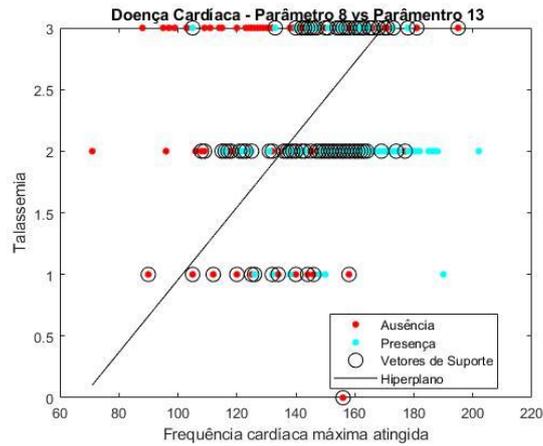


Figura 118: Gráfico representativo da relação entre o parâmetro 8 (Frequência cardíaca máxima atingida) e o parâmetro 13 (Talassemia).

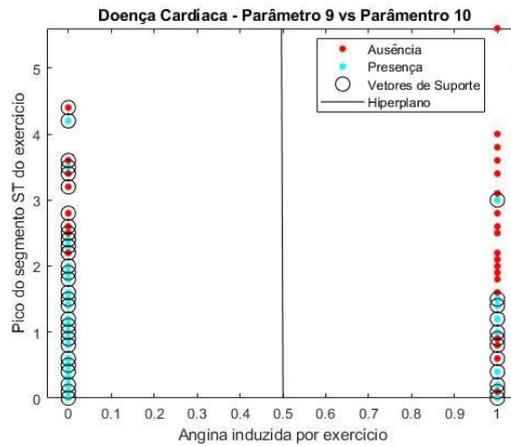


Figura 119: Gráfico representativo da relação entre o parâmetro 9 (Angina induzida por exercício) e o parâmetro 10 (Pico do segmento ST do exercício).

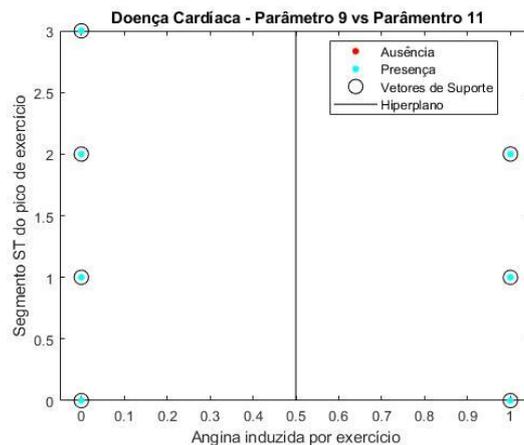


Figura 120: Gráfico representativo da relação entre o parâmetro 9 (Angina induzida por exercício) e o parâmetro 11 (Segmento ST do pico de exercício).

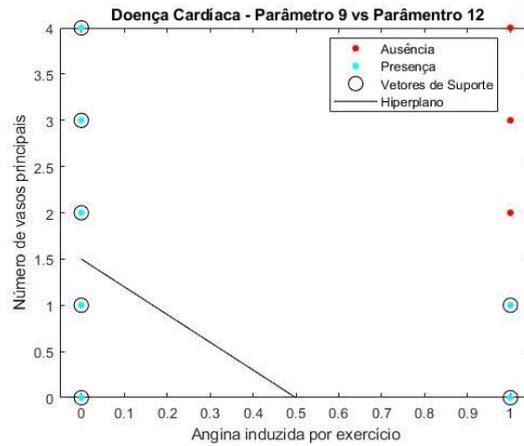


Figura 121: Gráfico representativo da relação entre o parâmetro 9 (Angina induzida por exercício) e o parâmetro 12 (Número de vasos principais).

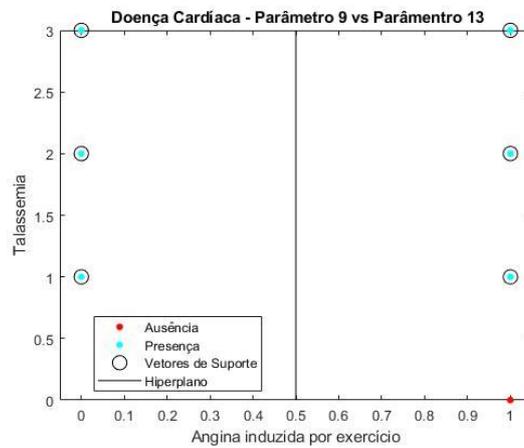


Figura 122: Gráfico representativo da relação entre o parâmetro 9 (Angina induzida por exercício) e o parâmetro 13 Talassemia).

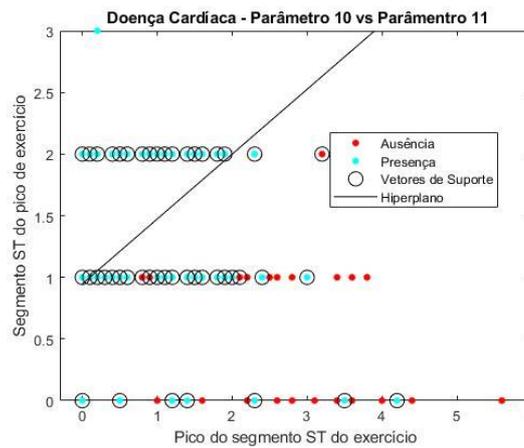


Figura 123: Gráfico representativo da relação entre o parâmetro 10 (Pico do segmento ST do exercício) e o parâmetro 11 (Segmento ST do pico de exercício).

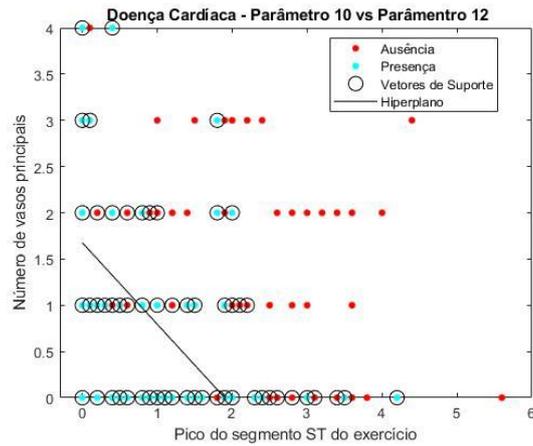


Figura 124: Gráfico representativo da relação entre o parâmetro 10 (Pico do segmento ST do exercício) e o parâmetro 12 (Número de vasos principais).

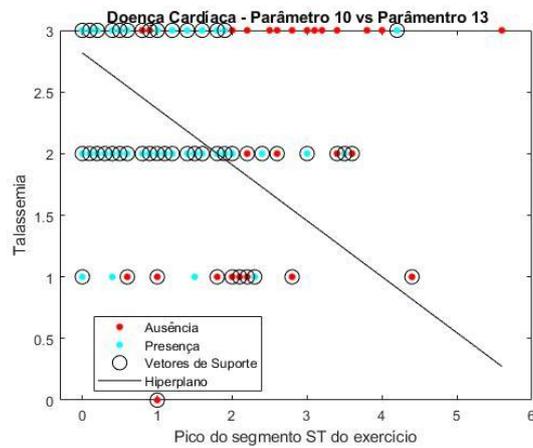


Figura 125: Gráfico representativo da relação entre o parâmetro 10 (Pico do segmento ST do exercício) e o parâmetro 13 (Talassemia).

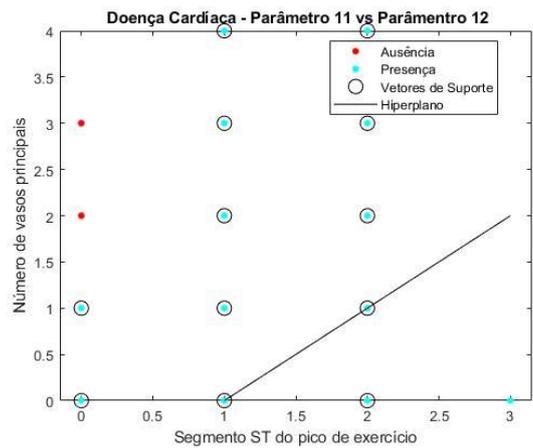


Figura 126: Gráfico representativo da relação entre o parâmetro 11 (Segmento ST do pico de exercício) e o parâmetro 12 (Número de vasos principais).

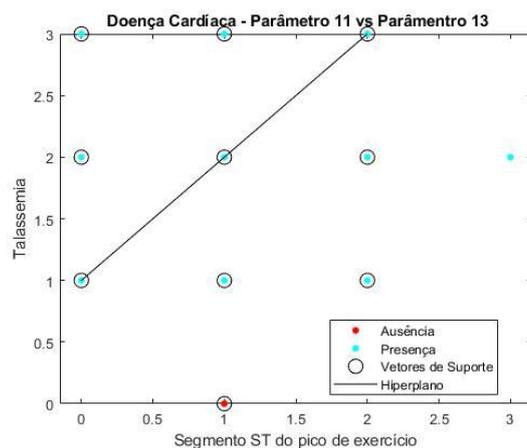


Figura 127: Gráfico representativo da relação entre o parâmetro 11 (Segmento ST do pico de exercício) e o parâmetro 12 (Talassemia).

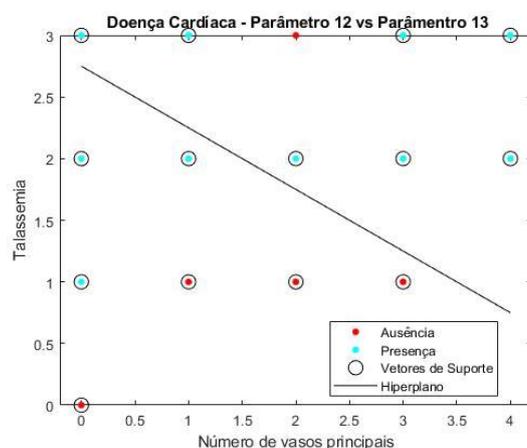


Figura 128: Gráfico representativo da relação entre o parâmetro 12 (Número de vasos principais) e o parâmetro 13 (Talassemia).

Anexo 3C

ID Paciente	Diagnóstico Real	Diagnóstico Previsto	Probabilidade (%)	
			Ausência (0)	Presença (1)
4	0	0	91.89	8.12
6	0	1	26.04	73.96
11	0	0	99.20	0.80
12	0	0	96.17	3.83
16	0	0	97.97	2.03
28	0	0	91.79	8.21
30	0	1	34.90	65.10
40	0	0	99.70	0.30
51	0	0	94.53	5.47
56	0	0	99.00	1.00
64	0	1	22.23	77.77
76	0	0	95.77	4.23

86	0	0	99.91	0.09
103	0	1	21.87	78.13
104	0	0	99.06	0.94
113	0	1	36.43	63.57
143	1	1	30.52	69.48
145	1	1	18.90	81.10
152	1	1	17.64	82.36
154	1	1	4.01	95.99
156	1	1	11.63	88.37
157	1	0	54.39	45.61
178	1	1	5.99	94.01
187	1	1	0.36	99.65
189	1	1	2.54	97.46
199	1	1	2.55	97.45
207	1	1	9.22	90.78
208	1	1	6.23	93.77
212	1	0	55.24	44.76
217	1	1	6.57	93.43
218	1	1	44.74	55.26
219	1	1	3.07	96.93
223	1	1	27.27	72.73
234	1	0	90.92	9.08
242	1	1	23.32	76.68
246	1	1	48.32	51.68
260	1	1	29.18	70.82
264	1	1	2.63	97.37
265	1	1	26.32	73.68
267	1	1	2.31	97.69
270	1	1	6.71	93.29
271	1	1	13.28	86.72
275	1	1	1.75	98.25
300	1	1	5.77	94.23
302	1	1	40.46	59.54

