



Samaritana Isabel Barbosa da Silva

Proposta e desenvolvimento de soluções para aferir estados emocionais em aplicações IoT

Dissertação de Mestrado em Engenharia Biomédica,
apresentada à Faculdade de Ciências e Tecnologia da Universidade de Coimbra

Julho 2018



UNIVERSIDADE DE COIMBRA

• U



C •

FCTUC

FACULDADE DE CIÊNCIAS
E TECNOLOGIA

UNIVERSIDADE DE COIMBRA

Samaritana Isabel Barbosa da Silva

Proposta e desenvolvimento de soluções para aferir estados emocionais em aplicações IoT

Tese submetida à

Universidade de Coimbra para o grau de

Mestre em Engenharia Biomédica

Supervisores:

Prof. Dr. Jorge Sá Silva (Coordenador)

Prof. Dr. Luís Macedo (Coordenador)

Coimbra, 2018

Esta cópia da tese é fornecida na condição de que quem a consulta reconhece que os direitos de autor são pertença do autor da tese e que nenhuma citação ou informação obtida a partir dela pode ser publicada sem a referência apropriada.

This copy of the thesis has been supplied on condition that anyone who consults it is understood to recognize that its copyright rests with its author and that no quotation from the thesis and no information derived from it may be published without proper acknowledgement.

Agradecimentos

Apesar desta dissertação ser um trabalho maioritariamente individual não seria possível sem a contribuição de algumas pessoas.

Quero expressar o meu sincero agradecimento ao Prof. Jorge Sá Silva e ao Prof. Luís Macedo por todo o apoio dado ao longo do projeto. Um agradecimento especial por toda a revisão e críticas construtivas ao manuscrito, bem como pela partilha de conhecimento e experiência científica.

Quero agradecer também aos excelentes colegas de laboratório que tive, pela intensa partilha de boas histórias e sobretudo, pela disponibilidade sem medida.

Quero agradecer aos meus amigos e ao meu namorado pelas demonstrações de amor, pela partilha diária de tudo, pela criação de momentos para a vida, pela ajuda constante em tudo o que precisei e pelos momentos cheios de juízo e os outros sem juízo nenhum.

Por último, o maior agradecimento que poderia fazer, aos meus pais. Sem eles não teria a possibilidade de realizar o sonho de vir para a Universidade. Por serem os melhores do mundo e nunca desistirem de mim, pela confiança depositada a 100% e pelo orgulho demonstrado a cada objetivo alcançado. Por serem o meu modelo a seguir e pelo apoio incondicional em todos os momentos da minha vida. Quero ainda agradecer ao meu irmão que partilhou comigo grandes momentos em Coimbra, e por ser o meu maior apoio quando os nossos pais não estão presentes. Claro está, não poderia deixar de agradecer à minha nova irmã, que apesar de ser criança é a pessoa que põe sempre um sorriso na minha cara com as palhaçadas que faz, e que me vê como um exemplo a seguir.

Resumo

A Internet das Coisas (IoT), compreende todos os aparelhos e objetos que têm a capacidade de estarem permanentemente ligados à Internet, sendo capazes de se indentificar na rede e de comunicar entre si. A IoT torna os Sistemas Ciber Físicos (CPS) possíveis. Estes sistemas permitem muitas funcionalidades inovadoras por meio das suas redes e do acesso ao mundo virtual, alterando significativamente o dia a dia do utilizador. Neste contexto, irão surgir novos modelos de negócios, processos de trabalho e métodos de desenvolvimento que atualmente parecem impossíveis. Estas mudanças também influenciarão fortemente a sociedade e as pessoas. Existe ainda pouco trabalho científico que se foque nos efeitos do contexto humano no controlo dos CPS. De facto, um importante elemento frequentemente não considerado nos atuais CPS é o utilizador humano.

Um típico Human-in-the-Loop Cyber-Physical System (HITLCPS) consiste num ciclo que envolve um humano, um sistema incorporado (componente cibernética) e o ambiente físico. Basicamente, o sistema aumenta a interação do ser humano com o mundo físico. A presença e comportamento humanos já não serão vistos como fatores externos e desconhecidos, mas tornar-se-ão numa parte fulcral do sistema.

O principal objetivo deste projeto é desenvolver soluções que permitam aferir estados físicos e emocionais dos utilizadores através de informação recolhida por dispositivos IoT na aplicação IoT Student Advisor & BEst Lifestyle Analyser (ISABELA). Utilizando técnicas de aprendizagem automática, pretende-se compreender o impacto das situações diárias no *stress*, padrões de sono, actividade física, sociabilidade e bem-estar emocional de vários estudantes e relacionar estes parâmetros com o seu desempenho académico.

Através da aplicação no telemóvel, dos sensores sociais e de caixas IoT é recolhida a informação para obter a atividade, a localização, a sociabilidade e os padrões de sono para inferir o comportamento do utilizador.

Abstract

Internet-of-Things (IoT) comprises all devices and objects that have the ability to be permanently connected to the Internet, being able to identify themselves in the network and to communicate with each other. IoT makes Cyber-Physical Systems (CPS) possible. These systems allow many innovative features through their networks and access to the virtual world, significantly changing the day to day of the user. In this context, new business models, work processes and development methods that today seems impossible will appear. These changes will also strongly influence society and people. There is still a lack of scientific work focusing on the effects of the human context on CPS control. In fact, an important element often not considered in the current CPS is the human user.

A typical HITLCPS consists of a loop involving a human, an embedded system (the cyber component), and the physical environment. Basically, the embedded system augments a human's interaction with the physical world. Human presence and behavior will no longer be seen as external factors and unknown, but they will become a central part of the system.

The main goal of this project is to develop solutions that allow users to measure their physical and emotional states through information collected by IoT devices in the application ISABELA. Using machine learning techniques, it is intended to understand the impact of daily situations on stress, sleep patterns, physical activity, sociability and emotional well-being of various users and to relate these parameters to the student's academic performance.

Through the application on the mobile phone, social sensors and IoT boxes, information is collected to obtain the activity, location, sociability and sleep patterns to infer the behavior of the user.

Lista de Abreviaturas

API Application Programming Interfaces.

CISUC Centro de Informática e Sistemas da Universidade de Coimbra.

CPS Sistemas Ciber Físicos.

DEI Departamento de Engenharia Informática.

EPN Escuela Politécnica Nacional do Equador.

GE Generic Enablers.

GPS Sistema de Posicionamento Global.

HITLCPS Human-in-the-Loop Cyber-Physical System.

HTTP Hypertext Transfer Protocol.

IA Inteligência Artificial.

IC Intervalo de Confiança.

IoT Internet das Coisas.

ISABELA IoT Student Advisor & BEst Lifestyle Analyser.

ISCAC Instituto Superior de Contabilidade e Administração de Coimbra.

JSON JavaScript Object Notation.

KDN Kit de Desenvolvimento Nativo.

LCT Laboratório de Comunicação e Telemática.

M2M Máquina-para-Máquina.

REST Representational State Transfer.

SO Sistema Operativo.

SVC Sistema de Versão Controlada.

Lista de Figuras

1.1	Ciclo HITLCPS [1]	4
2.1	Janela de questões do GitLab	10
2.2	Planeamento original do projeto	11
2.3	Planeamento final do projeto	11
3.1	Arquitetura do projeto	19
3.2	Fontes de informação do ISABELA	20
3.3	Sociabilidade	21
3.4	Caixa IoT	22
3.5	Arquitetura FIWARE da aplicação ISABELA	23
3.6	Exemplo de conversa do ChatBot	28
4.1	Fluxo de informação	32
4.2	Exemplo de um gráfico do sono	33
4.3	Ecrã do módulo da atividade	34
5.1	Quantidade de dados obtida por semana	39
5.2	Número de questionários obtidos por semana	39
5.3	Boxplot para a variável "Estudo"	41
5.4	Boxplot para a variável "Número de horas dormidas diariamente"	41
5.5	Boxplot para a variável "Qualidade do Sono"	42
5.6	Boxplot para a variável "Sociabilidade"	43
5.7	Boxplot para a variável "Atividade Física"	43
5.8	Relação entre a variável "Estudo" e a variável "Qualidade do sono" para o grupo 1	44
5.9	Relação entre a variável "Estudo" e a variável "Número de horas de sono dormidas diariamente" para o grupo 1	45
5.10	Relação entre a variável "Estudo" e a variável "Sociabilidade" para o grupo 1	45

5.11	Relação entre a variável "Estudo" e a variável "Atividade Física" para o grupo 1	46
5.12	Relação entre a variável "Número de horas de sono dormidas diariamente" e a variável "Qualidade do sono" para o grupo 1	46
5.13	Relação entre a variável "Qualidade do sono" e a variável "Atividade Física" para o grupo 1	47
5.14	Relação entre a variável "Qualidade do sono" e a variável "Sociabilidade" para o grupo 1	47
5.15	Relação entre a variável "Sociabilidade" e a variável "Atividade Física" para o grupo 1	48
5.16	Relação entre a variável "Estudo" e a variável "Qualidade do sono" para o grupo 2	48
5.17	Relação entre a variável "Estudo" e a variável "Número de horas dormidas diariamente" para o grupo 2	49
5.18	Relação entre a variável "Estudo" e a variável "Sociabilidade" para o grupo 2	49
5.19	Relação entre a variável "Estudo" e a variável "Atividade Física" para o grupo 2	50
5.20	Relação entre a variável "Número de horas dormidas diariamente" e a variável "Qualidade do sono" para o grupo 2	50
5.21	Relação entre a variável "Qualidade do sono" e variável "Atividade Física" para o grupo 2	51
5.22	Relação entre a variável "Qualidade do sono" e variável "Sociabilidade" para o grupo 2	51
5.23	Relação entre a variável "Sociabilidade" e variável "Atividade Física" para o grupo 2	52

Lista de Tabelas

2.1	Descrição dos parâmetros dos requisitos funcionais	14
3.1	Parâmetros lidos pelos componentes do ISABELA	21
4.1	Melhor desempenho do classificador Random Forest	31
5.1	Relevância dos atributos para a detecção do sono	55
5.2	Valor da correlação dos atributos	56
5.3	Desempenho dos diferentes classificadores, recorrendo ao software WEKA	56
5.4	Desempenho do classificador Random Forest, variando o número de árvores de decisão	57
5.5	Desempenho da rede neuronal, variando o número de neurónios das camadas escondidas	58
5.6	Desempenho do classificador Random Forest no novo conjunto de dados	59
5.7	Relevância dos atributos para a "Qualidade do sono"	60
5.8	Valor da correlação dos atributos para a "Qualidade do sono"	61
5.9	Desempenho do classificador Random Forest para a "Qualidade do sono"	61
5.10	Relevância dos atributos para a "Sociabilidade"	62
5.11	Valor da correlação dos atributos para a "Sociabilidade"	63
5.12	Desempenho do classificador Random Forest para a "Sociabilidade"	63

Conteúdo

Lista de Abreviaturas	ix
Lista de Figuras	xi
Lista de Tabelas	xiii
1 Introdução	1
1.1 Enquadramento e motivação	1
1.2 Objetivos	2
1.3 Conceitos	3
1.3.1 IoT	3
1.3.2 Human in the Loop Cyber-Physical Systems	4
1.3.3 Inteligência Artificial	4
1.4 Tecnologias Utilizadas	5
1.4.1 Android	5
1.4.2 Encog	6
1.4.3 WEKA	6
1.4.4 FIWARE	6
1.4.5 Python	7
1.4.6 GIT- Sistema de Versão Controlada	7
1.5 Estrutura do documento	7
2 Metodologia e Planeamento	9
2.1 Metodologia	9
2.2 Planeamento	10
2.3 Análise de Riscos	12
2.3.1 Principais Riscos	12
2.3.2 Estratégias de Mitigação	13
2.4 Critérios de Sucesso	13

2.5	Requisitos	14
2.5.1	Requisitos funcionais	14
2.5.2	Requisitos não funcionais	16
3	Plataforma de estudo ISABELA	19
3.1	Arquitetura Geral	19
3.2	Contexto	19
3.3	Caixas IoT	22
3.4	Arquitetura do FIWARE	23
3.4.1	ORION	23
3.4.2	IDAS	24
3.4.3	CYGNUS	24
3.4.4	KEYROCK	24
3.4.5	COMET	25
3.4.6	CKAN	25
3.5	Trabalho relacionado - Aplicações semelhantes	25
3.5.1	StudentLife	25
3.5.2	RescueTime	26
3.5.3	Life Cycle	26
3.5.4	ATracker	26
3.5.5	Sleep Cycle Alarm clock	27
3.5.6	ISABELA vs Outra Aplicações	27
4	Novos módulos do ISABELA	29
4.1	Módulo do Sono	29
4.1.1	Trabalhos Relacionados	29
4.1.2	Métodos Utilizados	30
4.1.3	Automatização do processo de deteção do sono	31
4.2	Módulo da Atividade no telemóvel	33
4.3	Comunicação com o FIWARE	35
5	Validação e Testes	37
5.1	Preparação dos testes ISABELA	37
5.2	Análise da aceitação ao ISABELA por parte dos alunos	37
5.3	Teste do ISABELA	38
5.3.1	Análise da participação no teste	38
5.3.2	Grupos do teste	40
5.3.3	Análise das variáveis	40

5.3.4	Relação entre as variáveis	44
5.3.5	Teste de Significância: T-Student	52
5.4	Módulo do sono	54
5.4.1	Primeiro tratamento dos dados	54
5.4.2	Segundo tratamento dos dados	57
5.5	Validação do sistema ISABELA	59
5.5.1	Qualidade do Sono	60
5.5.2	Sociabilidade	61
6	Conclusões	65
	Bibliografia	67
	Anexos	73
A	Manual de Instruções ISABELA	75
B	Manual final de intruções do ISABELA	89
C	Declaração de Consentimento	100

Introdução

1.1 Enquadramento e motivação

A IoT compreende os objetos conectados à Internet e entre si através da rede, munidos de sensores, circuitos eletrónicos e *softwares* capazes de obter, processar e trocar dados. A chegada da IoT está a tornar-se uma realidade. De acordo com recentes estudos, espera-se que os mercados de comunicação Máquina-para-Máquina (M2M) atinjam quase os 35,16 mil milhões de dólares em 2020 [2]. Da utilização destes diversos elementos computacionais surge o conceito de CPS [3], que consiste na conexão de objetos com software integrado. Esta conexão pode ser diretamente entre os objetos ou através da Internet, cujo objetivo é formar um sistema único em rede. Estes CPS são usados onde quer que seja necessária uma comunicação entre estes sistemas e o mundo digital. Os CPS desempenham uma importância crescente no processo industrial e no controlo da produção, em particular, no contexto da IoT.

Apesar destas ferramentas interligadas comunicarem umas com as outras sem intervenção humana, a tecnologia é sempre feita por humanos e para humanos. Desta forma, para promover a criação de sistemas que sejam úteis, não basta simplesmente considerar a heterogeneidade e a integração de ferramentas inteligentes. Fora da área de *e-health* [4], cujo objetivo principal é o de monitorizar pacientes, existe ainda pouco trabalho científico que se foque nos efeitos do contexto humano no controlo dos CPS. De facto, um importante elemento frequentemente não considerado nos atuais CPS é o utilizador humano.

Os sistemas que considerem o contexto humano irão tornar-se progressivamente mais importantes, e a maioria das futuras tecnologias irão convergir para uma consciência deste contexto humano. Quando todos estes sensores e dispositivos móveis começarem a ser usados para detetar e compreender a natureza humana, os humanos tornar-se-ão parte integral da IoT e dos CPS convergindo-se, então, para os

HITLCPS [5], aumentando a interação do humano com o mundo físico. A presença e comportamento humanos já não serão vistos como fatores externos e desconhecidos, mas tornar-se-ão numa parte fulcral do sistema.

A interação entre humanos e o ambiente digital tornou-se não apenas uma atividade técnica, mas também uma forma de companheirismo empático. Este tipo de relacionamento fornece funcionalidades passivas, mas também funções baseadas na percepção das necessidades atuais do utilizador, respostas, preferências, coerência e intenção. Os sistemas devem-se adaptar completamente e individualmente ao utilizador, e retratar os estados emocionais deste [6]. Os estados emocionais fornecem informação importante, e permitem desenvolver aplicações móveis que ajudam a melhorar a qualidade de vida do ser humano.

Ao longo dos anos, assumiu-se que os pacientes com doenças mentais não estavam dispostos a usar tecnologias e muitos profissionais de saúde diziam que para o tratamento ser mais eficaz era necessário este ser cara a cara. Estudos recentes de diversas instituições de ensino, mostram que os pacientes podem gerir a sua condição psicológica, com sucesso, usando telemóveis, computadores e outros sensores [7].

Ortony, Clore & Collins (OCC) propuseram o modelo OCC em 1988 [8]. Este modelo fornece uma estrutura clara das emoções e das variáveis que afetam a intensidade das emoções. Steunebrink et al. fizeram uma revisão ao modelo OCC de modo a tornar as emoções computáveis [9]. O modelo OCC descreve uma hierarquia que classifica 22 tipos de emoções. Cada emoção é classificada com base nas condições que a despertam, numa lista de palavras associada a esta e numa lista de variáveis que afetam a intensidade da emoção. Mas este modelo apresenta diversas ambiguidades que foram resolvidas recorrendo a um novo modelo baseado na herança da estrutura das emoções. O modelo OCC e o modelo *Pleasure-Arousal-Dominance* [10] são os modelos mais utilizados para detetar e simular estados emocionais.

1.2 Objetivos

O ISABELA é uma aplicação que tem como objetivo ajudar os estudantes a melhorar o seu desempenho académico. O ISABELA é uma aplicação desenvolvida por alguns membros do Laboratório de Comunicação e Telemática (LCT), do Centro de Informática e Sistemas da Universidade de Coimbra (CISUC), no Departamento de Engenharia Informática (DEI).

Enquanto membro da equipa do ISABELA, os objetivos do meu trabalho foram:

- Estudar o sistema ISABELA e todos os seus componentes;
- Desenvolver dois novos módulos da aplicação, mais concretamente o módulo do sono e o módulo da atividade no telemóvel
- Preparar os testes ISABELA, mais propriamente o manual de instruções e a declaração de consentimento;
- Fazer a análise estatística dos dados obtidos;
- Analisar o impacto das situações diárias nos padrões de sono, actividade física e sociabilidade dos vários estudantes e relacionar estes parâmetros com o desempenho académico.

1.3 Conceitos

1.3.1 IoT

O termo IoT foi proposto em 1999 por Kevin Ashton, co-fundador e diretor executivo do Auto-ID Center no MIT, durante uma apresentação que fez à Procter & Gamble [11]. Kevin Ashton criou um sistema de sensores, conectando o mundo físico à Internet, enquanto trabalhava na identificação por rádio frequência (RFID).

A IoT evoluiu a partir da convergência de tecnologias sem fio, sistemas microeletromecânicos e Internet. Esta convergência ajudou a destruir as barreiras existentes entre a tecnologia operacional e a tecnologia da informação, permitindo que os dados obtidos fossem analisados.

A IoT é uma rede de objetos físicos (por exemplo, veículos) que possuem uma tecnologia capaz de analisar o mundo com exatidão e comunicar entre si em tempo real, sem necessidade de interação humana ou de um computador.

Com a IoT os produtos industriais e os objetos de uso diário poderão vir a ter identidades eletrónicas ou poderão ser equipados com sensores que detetam mudanças físicas à sua volta. Estas mudanças transformarão objetos estáticos em dinâmicos, misturando inteligência e estimulando a criação de produtos inovadores e novos serviços.

1.3.2 Human in the Loop Cyber-Physical Systems

Quando os dispositivos móveis começarem a ser utilizados para detetar e compreender a natureza humana, os humanos tornar-se-ão parte integral da IoT e dos CPS convergindo para os HITLCPS. Os HITLCPS englobam uma série de aplicações com imenso potencial para criar impacto no quotidiano das pessoas.

Na figura 1.1 está representado um típico HITLCPS que consiste num utilizador, no ambiente físico e num sistema associado. O HITLCPS regista a intenção do utilizador, medindo a atividade cognitiva deste. O sistema associado traduz essa intenção em sinais que interagem com o ambiente físico. Por fim, o utilizador fecha o ciclo tomando novas decisões com base nas observações do mundo físico [1].

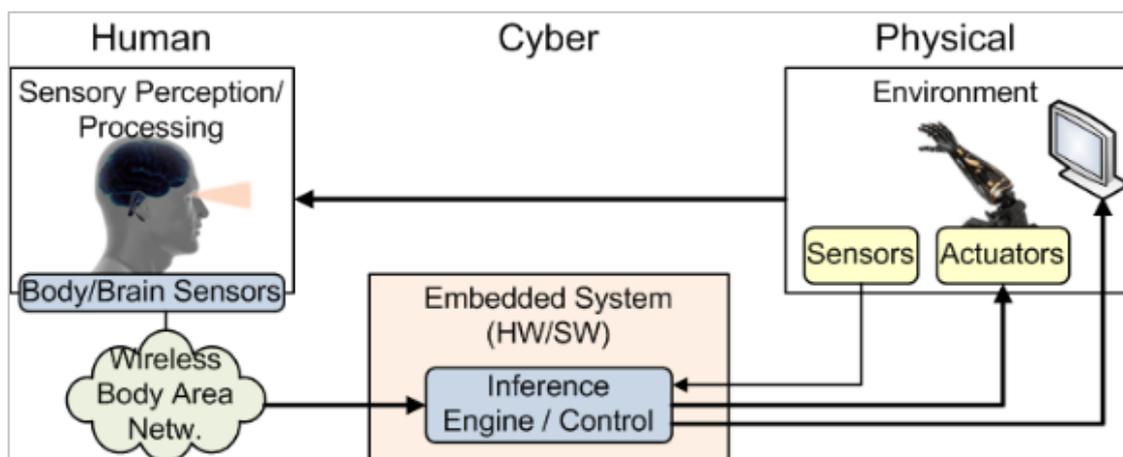


Figura 1.1: Ciclo HITLCPS [1]

1.3.3 Inteligência Artificial

A Inteligência Artificial (IA) é a simulação de processos de inteligência humana em máquinas [12]. Esses processos são compostos por aprendizagem, aquisição de informação, regras para o uso da informação e raciocínio (uso das regras para chegar a conclusões e auto-correção). As aplicações particulares da IA incluem sistemas periciais, reconhecimento de fala e visão computacional. Em algumas abordagens, IA pode resumir-se ao estudo e desenvolvimento de agentes inteligentes, ou seja, dispositivos que compreendam o ambiente que os rodeia e que conduzam ações que maximizem a probabilidade de sucesso.

Da SIRI aos carros automáticos, a IA está a progredir muito rápido. Normalmente a ficção científica retrata a IA como um robô com características humanas, mas a IA

pode abranger as mais diversas áreas, como por exemplo algoritmos de pesquisa do Google [13].

Atualmente, a IA é denominada como IA fraca, na medida em que é projetada para executar uma tarefa simples (por exemplo, reconhecimento facial ou pesquisas na Internet). No entanto, o objetivo a longo prazo é criar a IA forte. Enquanto a IA fraca pode superar os seres humanos numa tarefa específica, como jogar xadrez ou resolver equações, a IA forte irá superar os humanos em quase todas as tarefas cognitivas [14].

Enquanto IA é uma inteligência humana exibida em máquinas, a aprendizagem automática é um pilar, sendo mesmo considerada uma área da IA, muito importante para o desenvolvimento de sistemas com IA [15].

Aprendizagem automática é a ciência que faz com que os computadores executem o seu papel de maneira natural sem que pareçam programados para tal [16], ou seja, é a prática de usar algoritmos para recolher dados, aprender com eles, e então fazer uma escolha ou previsão sobre alguma coisa no mundo. Uma máquina é “treinada” usando uma grande quantidade de dados e algoritmos que lhe dão a capacidade de aprender como executar uma determinada tarefa. Usando aprendizagem automática é possível gerar de forma rápida modelos que permitam analisar um conjunto de dados e fornecer resultados de forma mais rápida e precisa, o que ajuda a tomar decisões e ações mais inteligentes em tempo real sem a intervenção humana.

1.4 Tecnologias Utilizadas

1.4.1 Android

O *Android* é um Sistema Operativo (SO), com base em *Linux Kernel* [17]. Atualmente é desenvolvido pela empresa de tecnologia Google. Este SO é desenvolvido especialmente para dispositivos móveis como telemóveis e *tablets*. Fornece ao utilizador uma interface visual para que este possa interagir com o dispositivo sem que necessariamente saiba o que acontece dentro deste. A principal interação com o dispositivo é através do toque no ecrã. Atualmente existem versões do Android para TV (Android TV), para carros (Android Auto) e para relógios de pulso (Android Wear) [18].

Neste projeto, é utilizado o Android Wear versão 1.5 no relógio de pulso que atualmente está na versão 2.0. Este relógio de pulso tem como função monitorizar o

batimento cardíaco, a atividade e a contagem de passos. O utilizador pode interagir com o relógio através do comando de voz ou simplesmente tocar no ecrã. A versão 2.0 vem com um teclado virtual para introduzir texto. O Android Wear pode ser utilizado como complemento do telemóvel, caso este tenha uma versão Android 4.3+.

O Android é o SO móvel mais utilizado no mundo, e em 2013 possuía a maior percentagem de dispositivos móveis vendidos [19].

As aplicações são desenvolvidas sobretudo na linguagem Java mas existem outras ferramentas de desenvolvimento disponíveis como o Kit de Desenvolvimento Nativo (KDN) que permite a programação em C ou C++.

1.4.2 Encog

Encog é uma estrutura utilizada para aprendizagem automática em JAVA e .NET, o que nos permite usar esta estrutura na plataforma Android[20]. Inicialmente o Encog foi criado para suportar apenas redes neuronais. Versões posteriores expandiram o Encog para aprendizagem automática no geral. Atualmente suporta uma variedade de algoritmos avançados como *Support Vector Machines*, *Neural Networks*, *Bayesian Networks*, *Hidden Markov Models* entre outros. Possui também classes para normalização e processamento dos dados [21].

1.4.3 WEKA

O WEKA [22] é um *software* desenvolvido pela Universidade de Waikato, que fornece diversos algoritmos de aprendizagem automática. Há duas possibilidades para chamar os algoritmos: uma é aplicar os algoritmos diretamente ao nosso conjunto de dados e outra é chama-los através do nosso código JAVA. O WEKA pode ser utilizado para pré-processamento, classificação, clustering, seleção de atributos (*features*), associação e visualização dos dados. Uma das grandes vantagens é que se pode aplicar o WEKA a grandes conjuntos de dados.

1.4.4 FIWARE

O FIWARE representa uma nova plataforma Europeia para o mercado de Informação e Comunicação Tecnológica que apresenta novas vantagens para a sociedade. O nome

FIWARE provém de *Future Internet Ware*. O FIWARE fornece novos componentes de *software*, disponíveis através de Application Programming Interfaces (API), e oferece aos utilizadores novas funcionalidades [23]. A plataforma FIWARE serve de apoio para aplicações IoT, pois tem como base uma biblioteca de componentes denominada Generic Enablers (GE). Os GE oferecem diversas funções através de APIs, facilitando o desenvolvimento de diversas aplicações e estabelecem as bases da arquitectura associadas à aplicação [24].

1.4.5 Python

O *Python* tornou-se uma das linguagens de programação mais dinâmicas para análise e visualização de dados [25]. Nos últimos anos, as bibliotecas de suporte à linguagem *Python* foram melhoradas (principalmente o *pandas*) o que faz desta linguagem uma forte alternativa para trabalhos que envolvam mineração de dados.

1.4.6 GIT- Sistema de Versão Controlada

O Sistema de Versão Controlada (SVC) é uma ferramenta que permite gerir e ver diferentes versões de um projeto [26], é muito utilizado em desenvolvimento de *software*. O SVC permite reverter ficheiros para um estado anterior, comparar as mudanças ao longo do tempo, ver quem modificou pela última vez algum ficheiro que pode ser a causa de algum problema e muito mais. Usar um SVC também significa que se alguém estragar o projeto ou perder ficheiros, estes podem ser recuperados facilmente.

O SVC usado durante este projeto foi uma instância local do GitLab. Ao usar uma instância local com autenticação garante-se que todas as aplicações e código são protegidos e privados. Esta ferramenta permite também a gestão do trabalho da equipa, pois acompanha o trabalho de todos, bem como coordena o trabalho dos diferentes membros do projeto.

1.5 Estrutura do documento

Este documento está organizado em seis capítulos: Introdução, Metodologia e Planeamento, Plataforma de estudo ISABELA, Implementações, Validação e Testes e por fim Conclusões.

1. Introdução

No primeiro capítulo é explicado o enquadramento do projeto, os objetivos deste, os conceitos abordados durante a realização do mesmo e ainda são referidas as tecnologias utilizadas.

No segundo capítulo é feita uma apresentação sobre a metodologia adotada durante a realização do projeto e sobre o seu planeamento, é também efetuada uma análise dos riscos e critérios de sucesso.

No terceiro capítulo é apresentada a plataforma de estudo ISABELA, onde é elaborada uma introdução ao sistema ISABELA e à arquitetura deste.

No quarto capítulo são explicados os novos módulos implementados na aplicação, mais concretamente o módulo do sono e da atividade no telemóvel.

No quinto capítulo são apresentados os resultados das validações e testes efectuados.

E por último, no sexto capítulo são apresentadas as conclusões deste projeto.

Metodologia e Planeamento

2.1 Metodologia

Ao longo do desenvolvimento do projeto, às quintas-feiras, havia uma reunião com as pessoas envolvidas no projeto ISABELA onde se definia o planeamento do projeto, e onde se discutiam os resultados obtidos. Também às quinta-feiras havia uma reunião com o Prof. Jorge Sá Silva, com o Prof. Luís Macedo e com o Prof. Hugo Oliveira, onde se falava essencialmente sobre questões de aprendizagem automática. Às segundas-feiras havia uma reunião com o Prof. Jorge Sá Silva, onde era feito o planeamento semanal, e onde se falavam das dificuldades sentidas.

Para facilitar a comunicação entre todos os membros do grupo LCT, usava-se o SLack [27], onde é possível ter um canal para cada projeto, e também iniciar conversas com apenas uma pessoa. Quando algum membro do grupo não podia estar presente fisicamente numa reunião, utilizava-se como meio de comunicação o Skype [28].

Para melhorar a colaboração entre os membros do grupo, usava-se uma instância do GitLab [29] (explicado no capítulo 3) que funciona não só como um SVC mas também como uma ferramenta de gestão de trabalho de equipa. Os membros do grupo LCT tinham acesso a todos os projetos que estavam a ser desenvolvidos pelo grupo. Isto é importante pois permite ter uma ideia do trabalho feito por cada membro e ainda é possível criar novas questões sobre os projetos, atribuir estas questões a uma membro específico ou pôr uma data limite. Na figura 2.1 é possível ver uma janela com algumas questões do GitLab.

2. Metodologia e Planeamento

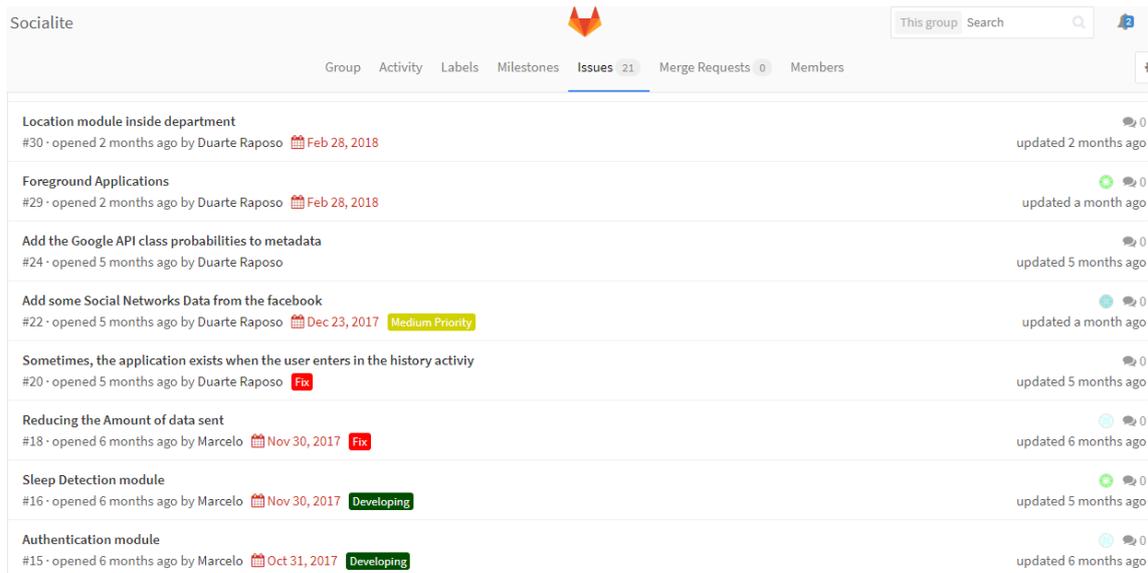


Figura 2.1: Janela de questões do GitLab

2.2 Planeamento

Na figura 2.2 está representado o planeamento original do projeto sob a forma de um Diagrama de Gantt. Ao longo da realização do projeto o planeamento sofreu algumas alterações e foi necessário fazer um novo Diagrama de Gantt, o que pode ser visualizado na figura 2.3. Estas alterações no planeamento ocorreram devido a atrasos no desenvolvimento do *software* e, principalmente com a necessidade de criar mecanismos de privacidade de acordo com o Regulamento Geral de Protecção de Dados [30] que entrou em vigor.

Como se pode ver na figura 2.3, os meses de setembro e outubro foram reservados para o estudo do problema e da plataforma existente. Este período serviu também para conhecer as tecnologias usadas ao longo do projeto.

Em dezembro deu-se início ao desenvolvimento do módulo do sono, em que na primeira parte se elencou e estudou o que foi feito na área da deteção automática do sono. Em seguida fez-se uma comparação do desempenho de vários classificadores, recorrendo ao *software* WEKA de modo a perceber qual o melhor. De acordo com o estado de arte para a classificação do sono, criou-se uma rede neuronal em JAVA recorrendo à livreria Encog para depois comparar o resultado desta com o melhor classificador obtido no WEKA. Esta análise terminou no final do mês de fevereiro.

Como o mês de janeiro foi uma altura de exames, só foi possível voltar a trabalhar no projeto a partir do meio deste mês. A última metade do mês foi reservada à

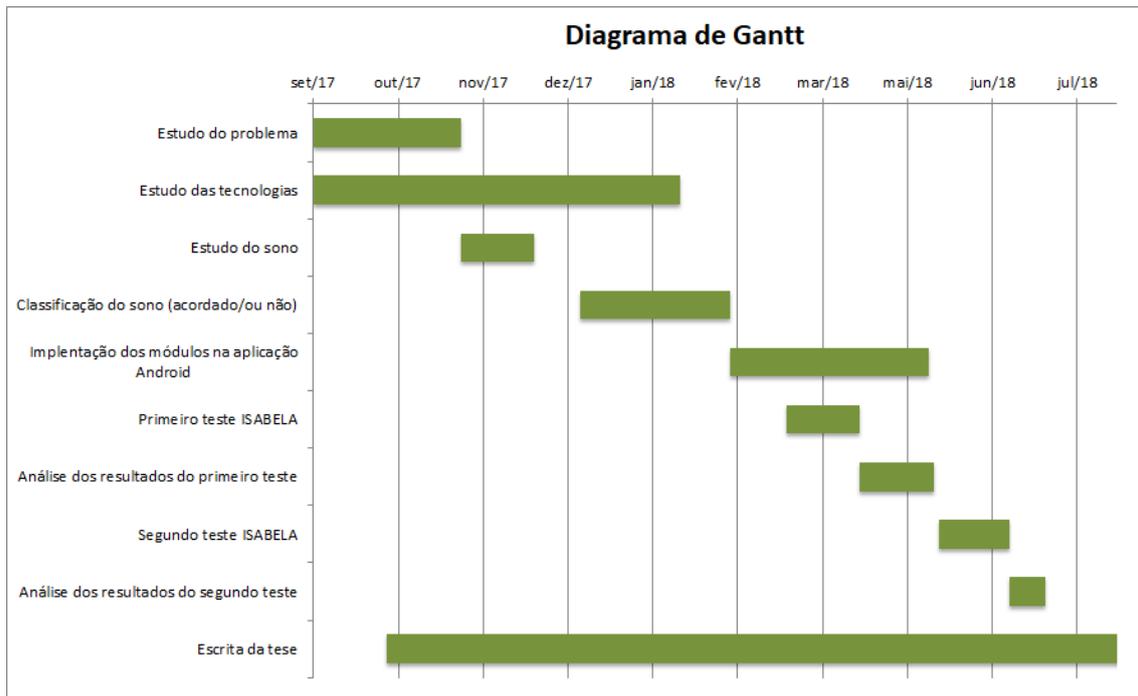


Figura 2.2: Planejamento original do projeto

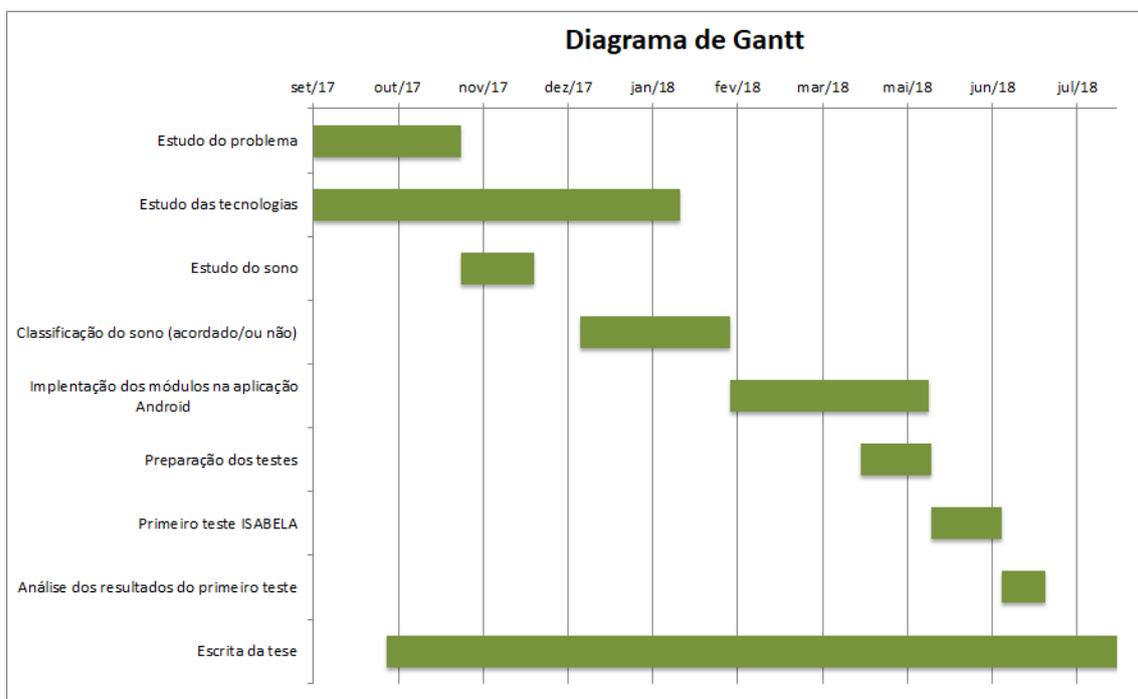


Figura 2.3: Planejamento final do projeto

escrita do capítulo 2 da dissertação.

O mês de fevereiro, março, abril e maio foram dedicados à implementação dos novos módulos na aplicação ISABELA, mais propriamente o módulo do sono e o módulo da atividade no telemóvel.

A última metade do mês de abril e o início do mês de maio foram dedicados à preparação para o início dos testes, mais propriamente da criação de um manual de instruções e de uma declaração de consentimento. Também foram realizados testes no grupo LCT para detetar eventuais erros. Durante o mês de maio foi escrito também o capítulo a explicar o módulo do sono e da atividade no telemóvel.

A última metade do mês de maio e a primeira metade do mês de junho foram dedicadas à realização de testes com alunos de cinco turmas: três turmas do DEI, uma do Instituto Superior de Contabilidade e Administração de Coimbra (ISCAC) e uma da Escuela Politécnica Nacional do Equador (EPN).

O final do mês de junho e o início do mês de julho foram dedicados à análise dos testes do ISABELA e à escrita da dissertação.

2.3 Análise de Riscos

A análise de riscos foi feita no início do projeto, esta é essencial para a realização de um projeto pois estabelece o contexto de um risco e como atuar sobre o mesmo.

2.3.1 Principais Riscos

Qualquer projeto tem riscos associados e, como tal, deve ser feita uma análise com o objetivo de os determinar. Um risco é um evento ou circunstância indesejada que pode impedir o cumprimento de objetivos definidos para o projeto. É um conceito abstrato pois em muitos casos é difícil de medir com precisão [31].

Os principais riscos associados ao projeto eram:

1. Os alunos não instalarem a aplicação ISABELA, não tendo o número de medições necessárias para trabalhar;
2. O facto de só existirem 3 caixas IoT para os estudantes;
3. O facto de os testes apenas terem começado em maio.

2.3.2 Estratégias de Mitigação

É importante criar um plano de mitigação dos riscos, para que nenhum risco identificado impeça a realização do projeto.

O risco 2 é provavelmente o mais difícil de contornar, uma vez que depende da "boa-vontade" dos alunos. O facto de os alunos terem que preencher um pequeno questionário todos os dias e de fornecerem informações sobre a sua vida poderá fazer com que não tenham muito interesse em participar neste estudo. É necessário criar algo que os motive a contribuir para este estudo. Uma maneira de contornar este risco é recorrer a incentivos financeiros para os alunos que participam como caso de estudo. Se realmente os alunos não contribuírem para o projeto, pode-se utilizar um conjunto de dados público como o DEAP [32] ou o StudentLife [33].

O facto de só existirem, até ao momento, 3 caixas IoT para colocar na casa dos estudantes, não permite recolher informação sobre o ambiente da casa dos estudantes que não tiverem caixas. Como é referido no capítulo 3, as caixas IoT dão informação sobre a luz, temperatura, ruído e humidade. A aplicação *Android* também dá informação sobre a luz e o ruído. Para contornar este problema pode-se inserir no questionário diário duas questões: uma sobre se a casa está húmida ou não e outra sobre se a casa está muito quente ou não.

O facto de os testes terem começado apenas em maio, atrasou um pouco o planeamento do projeto. Mas como os alunos da EPN participaram bastante e os dados obtidos foram bons, permitiu começar logo a trabalhar nestes dados e assim inferir o desempenho académico destes.

2.4 Critérios de Sucesso

Os estudantes ao terem uma possível previsão do seu desempenho académico vão poder ver quais os parâmetros que são necessários corrigir para que o seu desempenho melhore. Esta previsão do desempenho académico será uma mais valia para qualquer estudante.

Os principais critérios de sucesso deste projeto são:

- Conseguir estabelecer uma relação entre os parâmetros físicos, parâmetros emocionais e estilo de vida com o desempenho académico do estudante;

- Conseguir fechar o ciclo (*Human in the Loop*). Nomeadamente dar conselhos/recomendações, em tempo útil, aos estudantes para que eles consigam ter um bom desempenho académico.

2.5 Requisitos

2.5.1 Requisitos funcionais

Os requisitos funcionais descrevem explicitamente as funcionalidades e serviços do sistema. Fornecem consistência ao sistema, impedindo informação contraditória.

Os parâmetros usados para descrever os requisitos funcionais são: prioridade, descrição, pré-condições, fluxo do evento e resultado. Na tabela 2.1 é possível ver uma descrição dos parâmetros usados para descrever os requisitos funcionais.

Parâmetros	Descrição dos parâmetros
Prioridade	Obrigatória.
Descrição	Exposição geral do requisito funcional.
Pré-condições	Condições necessárias para o funcionamento do requisito.
Fluxo do evento	Descrição das ações que o utilizador necessita de realizar para obter o resultado esperado.
Resultado	Resultado esperado da ação.

Tabela 2.1: Descrição dos parâmetros dos requisitos funcionais

Os requisitos funcionais do sistema são:

1) Recolha de informação para o módulo do sono

- Prioridade: obrigatória;
- Descrição: a aplicação deve recolher a informação necessária dos diferentes sensores e se o estudante está acordado ou não;
- Pré-condições: o utilizador deve fornecer todas as autorizações necessárias;
- Fluxo do evento: após o *Login* a aplicação inicia os serviços necessários para recolher informação;
- Resultado: permitir à aplicação inferir se o estudante está acordado ou não.

2) Recolha de informação para o módulo da atividade

- Prioridade: obrigatória;
- Descrição: a aplicação deve recolher a informação necessária dos diferentes sensores e inferir o tempo gasto noutras aplicações;
- Pré-condições: o utilizador deve fornecer todas as autorizações necessárias;
- Fluxo do evento: após o *Login* a aplicação inicia os serviços necessários para recolher informação;
- Resultado: permitir à aplicação inferir o tempo gasto noutras aplicações;

3) Armazenamento de Informação

- Prioridade: obrigatória;
- Descrição: a aplicação deve guardar a informação até o telemóvel ter uma conexão à Internet, para então a informação ser enviada para o servidor central;
- Pré-condições: nenhuma;
- Fluxo do evento: quando a informação é recolhida dos sensores, fica armazenada na base de dados do telemóvel até ser enviada para o FIWARE;
- Resultado: se não existir uma ligação à Internet a informação fica armazenada na base de dados do telemóvel. Quando a informação é enviada para o FIWARE é apagada da base de dados do telemóvel.

4) Envio de informação para o FIWARE

- Prioridade: obrigatória;
- Descrição: a aplicação deve enviar a informação recolhida para o FIWARE;
- Pré-condições: conexão estável à Internet;
- Fluxo do evento: Quando a aplicação tem uma ligação estável à Internet, a informação é enviada para o FIWARE e é eliminada do telemóvel quando a confirmação é recebida;
- Resultado: se a informação for enviada com sucesso para o FIWARE, é recebida uma resposta e a informação é eliminada da base de dados do telemóvel. Se o pedido falhar, o serviço tenta enviar a informação uns segundos depois.

5) Desempenho académico

- Prioridade: obrigatória;

- Descrição: estabelecer uma relação de entre os parâmetros físicos, parâmetros emocionais e estilo de vida com o desempenho académico do estudante;
- Pré-condições: ligação à Internet;
- Fluxo do evento: quando o utilizador entra na aplicação os gráficos são atualizados no ecrã principal e a aplicação está a correr serviços que vão prever o desempenho académico do estudante;
- Resultado: a aplicação envia uma notificação quando algum dos parâmetros lidos não está dentro do normal, e o *ChatBot* deve enviar uma mensagem a dizer qual o problema, ajudando deste modo o estudante a controlar o seu desempenho académico.

6) Aconselhar o estudante

- Prioridade: obrigatória;
- Descrição: dar conselhos através do *ChatBot*, em tempo útil, aos estudantes para que eles consigam ter um bom desempenho académico;
- Pré-condições: ligação à Internet;
- Fluxo do evento: quando algum problema é detetado o estudante deve receber uma mensagem com a descrição do problema. A aplicação deve conseguir responder às questões do utilizador sobre esse problema;
- Resultado: quando o utilizador vê uma mensagem da aplicação com a indicação de um determinado problema, esta deve ser capaz de responder a qualquer questão do utilizador.

2.5.2 Requisitos não funcionais

Os requisitos não-funcionais definem propriedades e restrições do sistema, como por exemplo segurança e desempenho. Podem ser do sistema todo ou de partes do sistema. Demonstram qualidade acerca dos serviços ou funções disponibilizadas pelo sistema.

Os requisitos não funcionais são:

- Desempenho: é importante que a aplicação não consuma demasiada bateria ao telemóvel;

- Acessibilidade: a aplicação não deve ficar limitada pelo número de utilizadores conectados;
- Segurança: como se está a lidar com questões pessoais dos estudantes é importante que estes dados fiquem seguros;
- Fiabilidade: é necessário assegurar com algum grau de certeza as conclusões obtidas sobre o desempenho académico dos estudantes.
- Suporte: a aplicação deve suportar a maioria dos aparelhos *Android* existentes no mercado.

O desempenho é assegurado, pois os serviços que executam as tarefas mais complexas são executados em segundo plano.

A segurança dos dados é assegurada através da autenticação e da encriptação da base de dados do telemóvel. Os dados apenas permanecem na base de dados do telemóvel por curtos intervalos de tempo.

A fiabilidade e escalabilidade são asseguradas pelo FIWARE e não pela aplicação em si. A arquitetura do FIWARE assegura que o sistema está sempre disponível, independentemente do número de utilizadores. É também no FIWARE que estão todos os métodos necessários para a análise dos dados.

Qualquer telemóvel *Android* com versão 4.3 ou superior consegue suportar a aplicação ISABELA. Atualmente mais de 98,2% dos telemóveis *Android* disponíveis no mercado possuem uma versão superior à 4.3, sendo assim assegurada a execução da aplicação na maioria dos telemóveis existentes.

3

Plataforma de estudo ISABELA

3.1 Arquitetura Geral

Foi criada a arquitetura presente na figura 3.1 para a aplicação ISABELA.

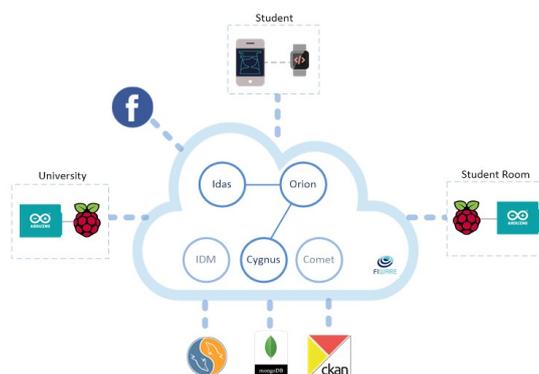


Figura 3.1: Arquitetura do projeto

Para criar um sistema que implementa uma arquitetura HITLCPS são necessárias quatro funcionalidades: aquisição de dados, inferência, inferência futura e atuação. Os dados são adquiridos através do telemóvel, do relógio de pulso, das caixas IoT e do sensor social. O FIWARE é estrutura responsável pelo armazenamento de dados e pela comunicação entre os dispositivos. É no telemóvel que são apresentados os estados e onde a atuação é feita através de notificações e mensagens. O relógio de pulso é o único dispositivo que não comunica diretamente com o FIWARE.

3.2 Contexto

Como já foi dito anteriormente a ISABELA é uma aplicação HITLCPS que tem como objetivo ajudar os estudantes a melhorar o seu desempenho académico. O

3. Plataforma de estudo ISABELA

ISABELA tem por base o estudo StudentLife da Universidade de Dartmouth [34]. No entanto o ISABELA e o StudentLife diferem em vários pontos: o ISABELA é um sistema HITLCPS, ou seja, permite dar feedback ao aluno e evitar o mau desempenho académico, usa uma arquitetura IoT, obtendo dados de mais fontes do que apenas do telemóvel (*smartwatch* e caixa IoT), utiliza o ChatBot como meio de interação com os alunos e usa sensores virtuais/sociais, enquanto o estudo da Universidade de Dartmouth foi principalmente estatístico;

A aplicação recolhe informação de várias fontes, o que pode ser visualizado na figura 3.2.

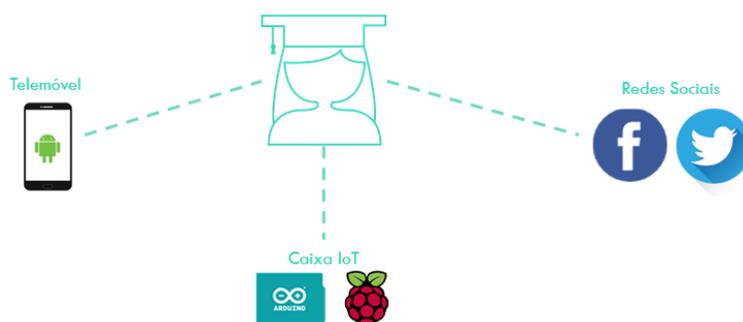


Figura 3.2: Fontes de informação do ISABELA

Na tabela 3.1 estão representados os parâmetros lidos pelas diferentes fontes de informação do ISABELA, como os sensores do telemóvel (por exemplo o som através do microfone ou a posição através do GPS), os sensores presentes na caixa IoT (por exemplo o sensor de temperatura), e ainda o sensor social que apresenta a média das emoções expressas nas publicações do Facebook e do Twitter.

Esta aplicação faz um acompanhamento diário do estudante para obter os dados dos sensores do telemóvel. Os dados obtidos são enviados para o FIWARE quando existe uma conexão estável à Internet. Os dados são retirados do FIWARE e depois de processados são visualizados pelo utilizador permitindo ao utilizador acompanhar o seu dia-a-dia.

De modo a ajudar o estudante a perceber ainda melhor o seu estado emocional, foi implementado um sensor social. Este sensor vai buscar as publicações feitas no *Facebook* e no *Twitter* e apresenta a média da intensidades das emoções expressas nas publicações dos últimos cinco dias. Quando os conteúdos têm subjacentes palavras positivas aparece uma cara com um sorriso feliz, quando o conteúdo é negativo aparece uma cara com um sorriso triste e quando a média do conteúdo é neutro aparece uma cara com um sorriso neutro, isto pode ser visualizado na figura 3.3.

Telemóvel	Sensor Social	Caixa IoT
SMS recebidos SMS enviados Destinatários SMS Chamadas recebidas Chamadas efetuadas Destinatários chamadas Duração chamadas Atividade Luminosidade Localização Som Proximidade Ecrã Ecrã Bloqueado Dia da semana Atividade no telemóvel	Publicações do Facebook e Twitter	Luz Temperatura Som

Tabela 3.1: Parâmetros lidos pelos componentes do ISABELA

O estudante tem que preencher um inquérito diário o "Sleep Form" que demora cerca de 30 segundos. Neste inquérito o utilizador tem que dizer a que horas se deitou, a que horas acordou, como foi a noite de sono (calma ou inquieta), como se sente (feliz ou triste) e se estudou muito ou não. Este inquérito tem como objetivo qualificar o dia a dia do utilizador para uma posterior classificação do seu estado físico e emocional.

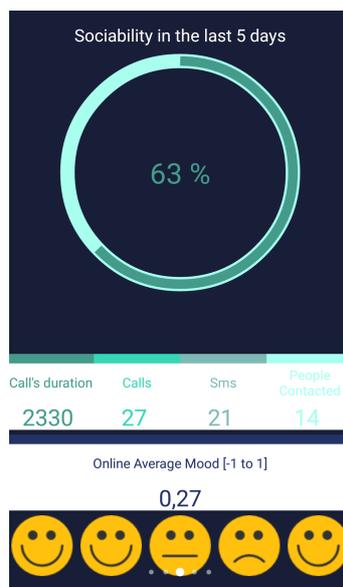


Figura 3.3: Sociabilidade

Quando algum comportamento irregular é detetado (por exemplo, o estudante pas-

sar pouco tempo na Universidade), a aplicação envia uma notificação ao utilizador a dizer que este tem uma nova mensagem da aplicação. Estas mensagens são personalizadas para o problema em questão e o utilizador pode solicitar dicas para o mesmo e receber sugestões do *ChatBot*. Se o utilizador implementar as mudanças de comportamento sugeridas pela aplicação, o problema fica resolvido e o ciclo de controlo é fechado pelo utilizador, implementando desta forma uma arquitetura HITLCPS.

3.3 Caixas IoT

A caixa IoT não foi feita por mim, mas sim por alguns elementos da equipa do ISABELA. Esta caixa é uma componente do sistema ISABELA.

A caixa IoT é composta por um sensor de som, um sensor de temperatura e um sensor de luz. O aspecto da caixa ISABELA pode ser visto na figura 3.4. O sensor de luz está colocado na parte superior da caixa, os sensores de som (esquerda) e de temperatura (direita) estão colocados na frente da caixa.

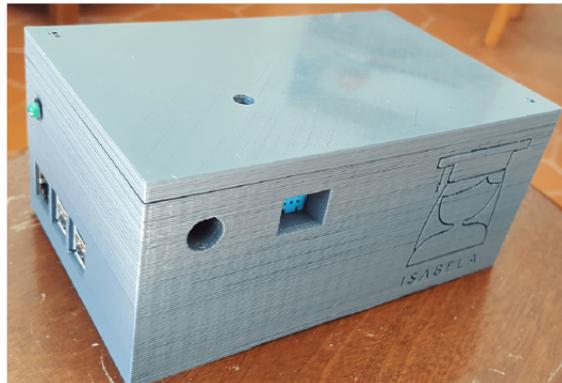


Figura 3.4: Caixa IoT

O sistema de aquisição de dados realiza-se através de um sistema Arduino. O sensor de som é um sensor analógico e, como tal, precisa de ser conectado aos pinos de entrada analógica (pinos laterais esquerdos) no Arduino, os sensores de luz e temperatura são digitais e estão conectados ao pinos digitais (pinos laterais direito).

O Arduino por si só não é capaz de fazer a comunicação com a Internet. Deste modo, os dados adquiridos pelo Arduino, são enviados para um microcomputador (Raspberry Pi 3). Este microcomputador envia os dados dos sensores para a plataforma FIWARE através da conexão disponível/configurada, cabo ou WiFi.

3.4 Arquitectura do FIWARE

Como foi dito anteriormente, o FIWARE é a estrutura responsável pelo armazenamento dos dados e pela comunicação entre os dispositivos. Neste projeto são utilizados seis módulos do FIWARE, nomeadamente, ORION, IDAS, CYGNUS, KEYROCK, COMET e CKAN. Na figura 3.5 é possível visualizar os diferentes módulos do FIWARE.

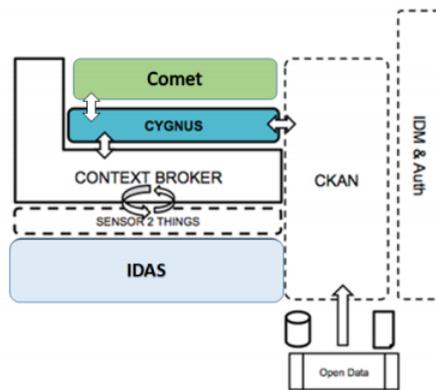


Figura 3.5: Arquitectura FIWARE da aplicação ISABELA

3.4.1 ORION

O ORION ou *Context Broker* permite gerir o ciclo de vida da informação de contexto [35]. Este módulo permite registar aplicações que produzem informação de contexto, atualizar essa informação, receber notificações quando ocorrem mudanças na informação ou receber a informação com uma dada frequência e por último permite fazer consultas. Para isso ser possível, este módulo implementa uma API Representational State Transfer (REST).

É o módulo da arquitetura que liga os produtores (por exemplo, sensores) aos consumidores (por exemplo, aplicação móvel), o que é importante para uma arquitetura IoT. Permite criar entidades virtuais para representar objetos do mundo real ou até mesmo pessoas. Estas entidades podem ter diferentes atributos e serem do mesmo tipo.

Apesar deste módulo permitir criar entidades e guardar os seus atributos, só é capaz de armazenar a última instância desses dados. Ou seja, se o valor de um atributo for atualizado, o módulo só guarda o último valor. Para guardar os dados para

o contexto histórico, existe a necessidade de implementar outros módulos como o CYGNUS e o COMET.

3.4.2 IDAS

Este módulo é responsável por registrar um novo aparelho e criar uma subscrição desse aparelho para uma entidade. Mas a principal função deste módulo é assegurar a comunicação com dispositivos incorporados que não usam protocolo Hypertext Transfer Protocol (HTTP) [36]. Isto permite ao sistema integrar mais dispositivos, aumentando assim a possível heterogeneidade do mesmo. O IDAS também suporta *Ultra Light 2.0* (UL2.0) um protocolo JavaScript Object Notation (JSON), que permite enviar mensagens curtas.

3.4.3 CYGNUS

Cygnus é o módulo responsável por persistir certas fontes de dados em certos armazenamentos de terceiros como o MySQL [37] ou o PostgreSQL [38], criando uma visão histórica desses dados [39]. Neste módulo são criadas subscrições a partir de entidades para armazenamentos de terceiros. Estas subscrições são feitas por tipo, e quando uma subscrição de um certo tipo é criada, atualizada ou alterada, o CYGNUS automaticamente guarda essas alterações na base de dados onde a subscrição foi feita.

3.4.4 KEYROCK

O KEYROCK é o módulo responsável pelo acesso dos utilizadores a redes, serviços e aplicações, incluindo a autenticação segura e privada a dispositivos, redes, entre outros [40]. Este módulo permite gerir todos os módulos de acesso, podendo restringir informação e funcionalidades a contas específicas. Este módulo implanta um serviço simples *sign-on*, e as credenciais do utilizador (email e palavra-passe) são as mesmas para todos os módulos.

Este módulo é muito importante, pois a privacidade dos dados é um dos principais requisitos para este projeto e com este módulo podemos evitar que os dados pessoais sejam libertados para terceiros, ou mesmo serem acedidos por utilizadores que não devem ter acesso aos dados.

3.4.5 COMET

O COMET é o módulo responsável pela gestão (armazenamento e consulta) da informação do contexto histórico [41]. Este módulo comunica através de APIs REST com o ORION e com aplicações externas, neste caso a aplicação ISABELA. Esta API permite agregar os dados por intervalos de tempos específicos. Esta API permite também agregar os dados como somas ou ocorrências (valores discretos).

Este módulo fornece a maioria dos dados para a aplicação, é um componente muito importante para manter a plataforma em funcionamento. Este módulo também aborda os requisitos para armazenar informações no projeto bem como os requisitos para ter informações contextualizadas no tempo.

3.4.6 CKAN

O CKAN não é um módulo do FIWARE. É uma das mais extensas plataformas de publicação de dados abertos e está-se a tornar padrão para a publicação de dados na Europa. Além disso, o CKAN é uma plataforma de código aberto, o que significa que este pode ser facilmente adaptado e expandido [42].

Esta plataforma é muito semelhante ao COMET pois permite armazenar e recuperar dados, mas não tem uma API tão “poderosa” como a do COMET para recuperar dados com contexto histórico. No entanto, esta plataforma tem as suas vantagens, como por exemplo o facto de apresentar uma *interface* da *Web* que pode ser usada para visualizar os dados.

3.5 Trabalho relacionado - Aplicações semelhantes

3.5.1 StudentLife

O *StudentLife* é uma aplicação *Android* que foi desenvolvida na Universidade de Dartmouth [33]. O *StudentLife* é o primeiro estudo que usa dados dos sensores do telemóvel dos estudantes para avaliar a saúde mental, o desempenho académico e as tendências comportamentais destes.

Foram utilizados 5 sensores: acelerómetro, microfone, luz, Sistema de Posicionamento Global (GPS) e *bluetooth* para obter os dados sobre a atividade do utilizador, padrões de conversação, períodos de sono e localização. A aplicação *StudentLife* acompanha os telefones dos alunos e mede automaticamente os comportamentos humanos 24/7 sem qualquer interação do utilizador.

Este estudo foi realizado com 48 estudantes durante um período de 10 semanas. Foram obtidos resultados interessantes com este estudo, por exemplo, quando os estudantes voltam das férias estão mais relaxados (os níveis de *stress* são baixos), dormem melhor e vão ao ginásio regularmente, todo este padrão muda quando os exames finais estão a chegar.

O *StudentLife* foi a principal inspiração para o projeto ISABELA.

3.5.2 RescueTime

O *RescueTime* é uma aplicação para gestão do tempo, que regista como e onde o utilizador gasta seu tempo *online*. Ajuda o utilizador a entender os seus hábitos diários de modo a este ser mais produtivo [43]. Esta aplicação contabiliza o tempo gasto em *websites* e noutras aplicações, dando ao utilizador uma imagem de como foi o seu dia. O *RescueTime* ajuda o utilizador a melhorar a sua perspectiva, a definir objetivos, por exemplo passar menos de uma hora por dia no email e a tomar uma atitude, por exemplo definir um alarme para quando se utiliza o Facebook mais do que duas horas.

3.5.3 Life Cycle

O *Life Cycle* é uma aplicação para iOS e pertence à categoria *Health & Fitness* [44]. O *Life Cycle* permite ao utilizador acompanhar o seu dia a dia, e representa a vida deste num diagrama de fatias, em que cada fatia representa uma atividade diferente. Esta aplicação permite ao utilizador ver a sua vida, dia após dia, semana após semana ou mesmo ano após ano.

3.5.4 ATracker

O *ATracker* é uma aplicação que permite ao utilizador acompanhar o seu tempo [45], apenas é necessário dar um nome a uma atividade e depois pode-se começar

ou parar a atividade com apenas um toque no ecrã. Esta aplicação apresenta um histórico do tempo, cria relatórios que podem ser enviados para o email do utilizador e permite também a integração do calendário.

3.5.5 Sleep Cycle Alarm clock

O *Sleep Cycle Alarm Clock* [46] é uma aplicação que analisa o sono do utilizador e acorda-o na fase em que o sono é mais leve. Esta aplicação faz uma análise do som enquanto o utilizador dorme, utilizando o microfone e o acelerómetro para perceber em que fase do sono o utilizador se encontra. O *Sleep Cycle* encontra a melhor fase para acordar o utilizador de modo a que este não se sinta tão cansado, utilizando uma janela de tempo de 30 minutos que termina no horário do alarme configurado.

3.5.6 ISABELA vs Outra Aplicações

A maior diferença entre o ISABELA e as restantes aplicações é a implementação de uma arquitetura HITLCPS onde é dado *feedback* ao utilizador. O ISABELA dá *feedback* aos utilizadores através de mensagens no *ChatBot*, permite que o estudante crie uma conversa com o *ChatBot* perguntando o que se pode fazer para mudar um certo comportamento, isto pode ser visualizado na figura 3.6.

Outra diferença prende-se com a arquitectura do nosso sistema. Este possui uma arquitectura IoT com uma plataforma *middleware* IoT que implementa todos os protocolos de comunicação necessários para comunicar com todo o tipo de dispositivos conectados. Isto permite que no futuro, se use mais aparelhos e se crie novos cenários.

Enquanto cada aplicação trata de um tema diferente, o ISABELA junta tudo numa só aplicação. No ISABELA é possível ver o número de horas dormidas, acompanhar a atividade física, ver o número de horas gastas diariamente nas aplicações, ver a localização e a sociabilidade, entre outras funcionalidades.

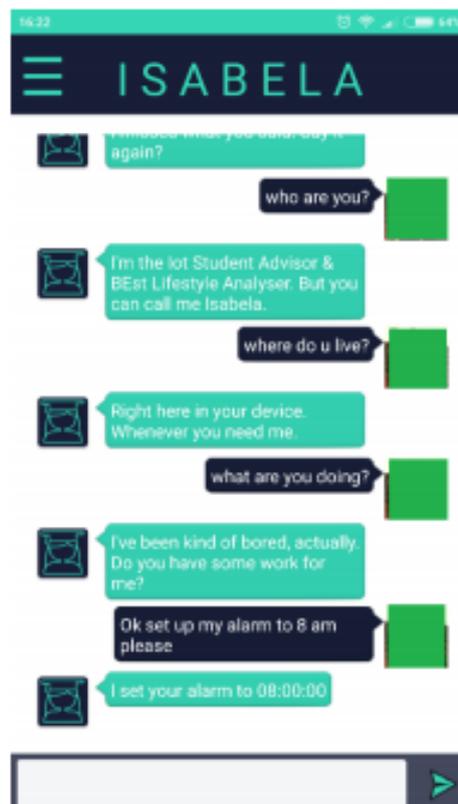


Figura 3.6: Exemplo de conversa do ChatBot

Novos módulos do ISABELA

Neste capítulo é explicado o trabalho desenvolvido no módulo do sono e da atividade no telemóvel.

4.1 Módulo do Sono

O objetivo deste módulo é criar um ecrã que mostre ao utilizador o número de horas dormidas nos últimos cinco dias.

4.1.1 Trabalhos Relacionados

O sono é muito importante para que o corpo humano e a mente estejam bem, é durante este período que ocorre a restauração do cérebro, consolidação da memória e a correta gestão das funções metabólicas [47]. O sono inadequado prejudica a qualidade de vida e resulta num aumento do risco da mortalidade [48].

Ao longo dos anos têm surgido muitas aplicações que ajudam a controlar e a monitorizar o sono [49], também há alguns trabalhos na área da classificação para deteção do sono como por exemplo Chen Z et al [50] e Min J et al[51].

Na experiência de Chen Z et al [50] os participantes utilizaram um sensor no pulso da mão dominante que media a condutância e a temperatura da pele, uma aplicação *Android* que recolhia o tempo da chamada, o tempo das mensagens, a localização e o tempo em que o ecrã estava ligado e utilizaram ainda um sensor de actigrafia no pulso da mão não dominante que recolhia a atividade e a luz. Os participantes tinham ainda que preencher um diário todas as manhãs no qual diziam a que horas se deitavam, a que horas acordavam e se acordavam durante a noite. Os autores utilizaram redes neuronais recorrentes, mais precisamente *Long Short-Term Memory networks* para classificar o problema. Este método obteve uma *accuracy* de 96,5%.

A experiência de Min J et al[51] teve como objetivo descobrir se um telemóvel conseguia modelar a qualidade do sono sem necessidade de mudanças significativas na vida das pessoas. Através de uma aplicação *Android* foi recolhida a amplitude do som, a luz, o movimento, a proximidade ao ecrã, o estado do ecrã (ligado ou desligado), aplicações a correrem no telemóvel e se o telemóvel estava a carregar ou não. Os participantes tinham ainda que preencher um diário onde relatavam a que horas tinham ido dormir, se dormiram bem ou se tiveram insónias e onde tinham o telemóvel durante a noite. Para detectar o sono, os autores criaram um algoritmo com base numa série de classificações sono/não sono e obtiveram uma *accuracy* de 93,06 %.

4.1.2 Métodos Utilizados

Através da aplicação *Android* são recolhidos seis atributos [51] para classificar se uma pessoa ou está a dormir ou acordada. Estes atributos são recolhidos de 30 em 30 segundos, mas só são enviados para o FIWARE quando o telemóvel tem uma conexão estável à Internet. Os atributos são os seguintes:

- Luz
- Som
- Atividade
- Bloqueio do ecrã
- Tempo para o alarme seguinte
- Dia da semana

Todos estes atributos podem ser facilmente explicados pelos fatores que estão associados ao sono. Geralmente as pessoas tendem a dormir em ambientes silenciosos e com menos intensidade de luz. Quando o utilizador está a dormir ele não utiliza o telemóvel, ou seja, o telemóvel está bloqueado. O atributo Atividade é utilizado para detetar se o telemóvel está a ser movido ou não.

As pessoas definem diariamente a que horas o despertador vai tocar. Isto faz com que o atributo tempo para o alarme seguinte tenha uma forte correlação com a duração do sono. Quando o alarme é definido no telemóvel, verifica-se a hora em que o alarme é definido e subtrai-se o tempo atual para isso. Se nenhum alarme for definido, damos o valor -1 a este atributo.

Também se utiliza o atributo dia da semana porque os padrões de sono do utilizador estão relacionados com este atributo. Normalmente as pessoas tendem a dormir mais no fim de semana do que nos dias da semana.

Alguns dos atributos, dia da semana e atividade, eram dados por *strings*, mas estes atributos eram um número discreto de *strings*, e deste modo facilmente se normalizou estes atributos para números. O bloqueio do ecrã também era uma *string*. Esta foi convertida para 0 ou 1, 0 quando o telemóvel não está bloqueado e 1 quando o telemóvel está bloqueado. Os outros atributos já eram valores numéricos e foram apenas normalizados.

Foram utilizados dados de nove utilizadores do grupo do LCT para construir o conjunto de dados. A classe "Dormir" apresenta o valor 1 quando a pessoa está a dormir e 0 quando está acordada.

Implementou-se uma rede neuronal simples em Java recorrendo à livreria Encog [20], onde se foi variando o número de neurónios nas duas camadas escondidas. Também se procedeu à implementação do classificador Random Forest em Java, para comparar o desempenho da rede neuronal com o classificador Random Forest. Depois de vários testes conclui-se que o classificador Random Forest era o classificador que apresentava melhores resultados. Na tabela 4.1 é possível ver o melhor resultado para o classificador Random Forest.

Todo o tratamento dos dados e os testes realizados com os diferentes classificadores está explicado em pormenor no capítulo 5.

Nº de árvores de decisão	<i>Precision</i>	<i>Recall</i>	<i>F-measure</i>
50	0,941	0,939	0,939

Tabela 4.1: Melhor desempenho do classificador Random Forest

4.1.3 Automatização do processo de deteção do sono

Na figura 4.1 está representado o fluxo de informação desde a captação dos diferentes atributos pela aplicação ISABELA até ao envio do valor correspondente ao estado do utilizador (1 caso os dados sejam classificados como "dormir" ou 0 caso seja "acordado") novamente para a aplicação.

No ponto 1 da figura 4.1 está representado o envio dos atributos utilizados para a classificação do sono, da aplicação para o módulo ORION do FIWARE. Este envio só é feito quando o telemóvel tem uma conexão estável à Internet.

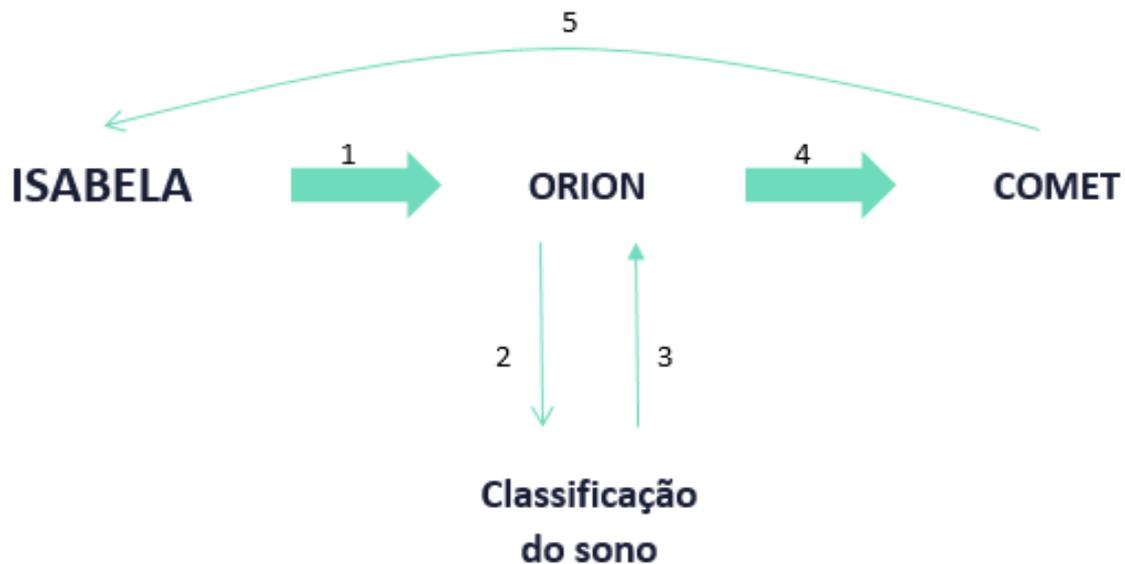


Figura 4.1: Fluxo de informação

Sempre que o ORION recebe dados novos estes são enviados para o ficheiro "Classificação do sono", o que está representado no ponto 2. Os dados que chegam ao ORION são ainda guardados numa base de dados, no PostgreSQL [38], para estar sempre a treinar o classificador. O ficheiro "Classificação do sono" foi feito em JAVA, e recebe como entrada um objeto JSON, com os atributos necessários para a classificação (atividade, luz, som e telemóvel bloqueado). O objeto JSON é lido, e cada atributo é normalizado. Depois dos valores normalizados, estes são processados pelo classificador.

O resultado da classificação, 1 ou 0, a data do processamento e a data de origem dos dados são colocados noutra objeto JSON e enviados novamente para o ORION, o que pode ser visto no ponto 3.

O módulo COMET está em comunicação com o ORION, e quando o novo objeto JSON chega ao ORION é também enviado para o COMET. O ponto 4 representa a comunicação entre o ORION e o COMET.

O ponto 5 representa a comunicação entre o COMET e a aplicação ISABELA. A aplicação recolhe a soma diária do número de segundos que a pessoa dormiu e esta soma é processada para horas. Na aplicação é apresentado um gráfico ao estudante que mostra o número de horas dormidas nos últimos 5 dias.

Na figura 4.2 é possível ver um gráfico que apresenta ao estudante o número de horas dormidas nos últimos cinco dias. No eixo do x estão representados os dias e no eixo do y o número de horas dormidas.



Figura 4.2: Exemplo de um gráfico do sono

4.2 Módulo da Atividade no telemóvel

O objetivo deste módulo é criar um ecrã que apresente o ícone da aplicação, o nome da aplicação e o tempo diário gasto nessa aplicação pelo estudante.

Para esse efeito, utilizou-se a API *UsageStats* [52]. Esta API foi introduzida no *Android 5.0 Lollipop* (API nível 21) e permite ter acesso a estatísticas sobre o uso das aplicações instaladas no telemóvel.

A API *UsageStats* permite obter o nome do *package* da aplicação e o tempo que a aplicação esteve em primeiro plano em milissegundos. O sistema junta os dados em quatro intervalos diferentes:

- Intervalo Diário
- Intervalo Semanal
- Intervalo Mensal
- Intervalo Anual

Os registos do sistema são limitados no tempo, pode-se obter os dados de uso da aplicação até sete dias no intervalo diário, até quatro semanas por intervalo semanal,

até seis meses por intervalo mensal e por fim até dois anos por intervalo anual.

Depois de lidar com a permissões necessárias pode-se obter as estatísticas de uso da aplicação chamando o método `queryUsageStats` (int intervalo_tempo, long tempo_inicial, long tempo_final) que devolve uma lista de `UsageStats`. Como o objetivo é obter o uso diário das aplicações, o intervalo_tempo é diário, o tempo_inicial é o dia anterior e o tempo_final é o dia atual. Um `UsageStats` contém a hora inicial e final para a estatística escolhida, o tempo que uma aplicação teve primeiro plano, e a última vez em que a aplicação foi usada. É claro que também tem o nome do `package` da aplicação.

Para não aparecer o nome do `package` no ecrã, utiliza-se apenas o nome da aplicação e o tempo foi convertido em horas. Esta API também permite obter o ícone da aplicação. Na figura 4.3 é possível ver uma representação do ecrã que mostra o uso diário das aplicações do telemóvel.

Para a informação sobre o uso das aplicações ficar armazenada, é enviado um objeto JSON que tem como entrada o número de horas gastas na aplicação, o nome da aplicação e a data de origem para o módulo ORION do FIWARE.



Figura 4.3: Ecrã do módulo da atividade

4.3 Comunicação com o FIWARE

O FIWARE tem muitos módulos como foi referido no capítulo 3, mas para os novos módulos implementados apenas se utiliza o ORION e o COMET. Todas as comunicações são feitas utilizando pedidos HTTP/REST [35].

A comunicação com o ORION é feita com três objetivos: criar entidades, atualizar ou adicionar valores de atributos e recuperar o valor de um atributo. Para poupar recursos computacionais e diminuir o tráfego da rede na comunicação com o módulo ORION, faz-se primeiro um *PATCH* (*Update*). O ORION serve como um gestor de contexto e uma vez que as entidades estejam criadas, raramente (ainda que possa acontecer se for desejado) elas são apagadas. Ao fazer primeiro um *PATCH* previne-se falhas de comunicação, que de outra forma iriam acontecer uma vez que cada vez que o módulo corresse no servidor seria devolvido um erro, uma vez que a entidade já tinha sido criada. Caso a entidade ainda não tenha sido criada, ao tentar fazer o *PATCH* dá erro, e faz-se então um *POST* que cria então uma nova entidade.

O outro módulo que está em comunicação com a aplicação é o COMET. O COMET implementa uma API REST que permite fazer pedidos *GET* (consultar informação). Nestes pedidos pode-se definir que informação se quer, em que intervalo de tempo e como a queremos agregar. A comunicação com este módulo é feita em várias partes da aplicação, mas a maioria dos pedidos são feitos para obter a informação para mostrar no ecrã do telemóvel. Para mostrar o gráfico com o número de horas dormidas a aplicação faz um pedido *GET* ao COMET.

Validação e Testes

5.1 Preparação dos testes ISABELA

Antes dos alunos efetuarem os testes, foi necessária uma preparação para tal acontecer.

De modo a facilitar a utilização da aplicação ISABELA foi feito um manual de instruções. Neste manual é explicado em que consiste a aplicação e qual o objetivo deste estudo, são referidos os termos e condições gerais, estão explicados os requisitos para a instalação da aplicação e quais as funcionalidades desta (Anexo A e B).

Também foi realizada uma declaração de consentimento (Anexo C) onde se informa que os dados são todos anonimizados e que estão de acordo com a Lei de Proteção dos Dados 2016/679 [30]. Esta lei garante a proteção dos dados pessoais e tem como finalidade contribuir para um mercado único europeu de dados e harmonizar a legislação de todos os Estados Membros da União Europeia. Esta lei aplica-se a todas as pessoas singulares e coletivas que efetuem tratamento de dados pessoais a residentes da União Europeia, entrou em vigor a 25 de Maio de 2018.

5.2 Análise da aceitação ao ISABELA por parte dos alunos

Os nossos testes foram realizados a uma turma de Engenharia Informática, duas de Design e Multimedia do DEI, uma turma do ISCAC e uma turma do EPN. No geral a participação dos alunos, com a exceção do EPN, foi muito reduzida.

No DEI, o total das três turmas era quase 200 alunos, mas apenas 7 participaram no estudo até ao fim. Mesmo com um prémio de 10 Arduínos para os 10 alunos que

tivessem mais dados, a participação no estudo foi bastante reduzida.

No ISCAC a turma era de 8 alunos, mas apenas 5 alunos participaram no estudo sendo que mesmo assim estes alunos não usaram a aplicação todos os dias, pelo que a quantidade de dados obtida foi reduzida.

No caso do EPN a turma era de 30 alunos e participaram 28 alunos, e estes usavam a aplicação diariamente. Este uso da aplicação deveu-se sobretudo à insistência do professor que integrou estes testes como um trabalho prático que contava para a nota da disciplina.

Em todas as turmas a utilização do Facebook e do Twitter foi quase nula, pelo que a análise das emoções expressas nas publicações não foi realizada. Isto deve-se principalmente ao facto de atualmente a aplicação que está mais na moda ser o Instagram e onde o conteúdo é essencialmente fotos e não texto.

No futuro terá de ser repensada a melhor maneira de atrair os alunos a participarem neste estudo. Provavelmente uma das maneiras a seguir passa pelo professor da disciplina onde os testes vão ser realizados, seguir a mesma abordagem que foi seguida pelo professor do EPN.

5.3 Teste do ISABELA

O objetivo destes testes foi, a partir dos parâmetros em bruto lidos, aferir: o *mood*, o estudo, a actividade e sociabilidade e depois relacionar estes valores com o desempenho académico.

Os testes foram realizados em cinco turmas, uma de Engenharia Informática do DEI, duas de Design e Multimédia do DEI, uma do ISCAC e uma do EPN. Os alunos apenas utilizaram a aplicação móvel, pois não havia relógios suficientes para todos os alunos. No caso dos estudantes do DEI foi instalada uma caixa IoT no bar, pois é uma zona que os alunos frequentam muito. Esta caixa media o nível de ruído através do som, a luz e a temperatura do bar.

5.3.1 Análise da participação no teste

Como já foi referido anteriormente a adesão ao estudo não foi a esperada. A figura 5.1 mostra a quantidade de dados obtida semanalmente durante o período de teste. Como se pode ver a participação do DEI e do ISCAC foi bastante reduzida, sendo

que na última semana foi quase nula. Na última semana a participação do EPN diminuiu quase para metade pois era uma semana de exames, mas mesmo assim foi bastante superior à do DEI e do ISCAC.

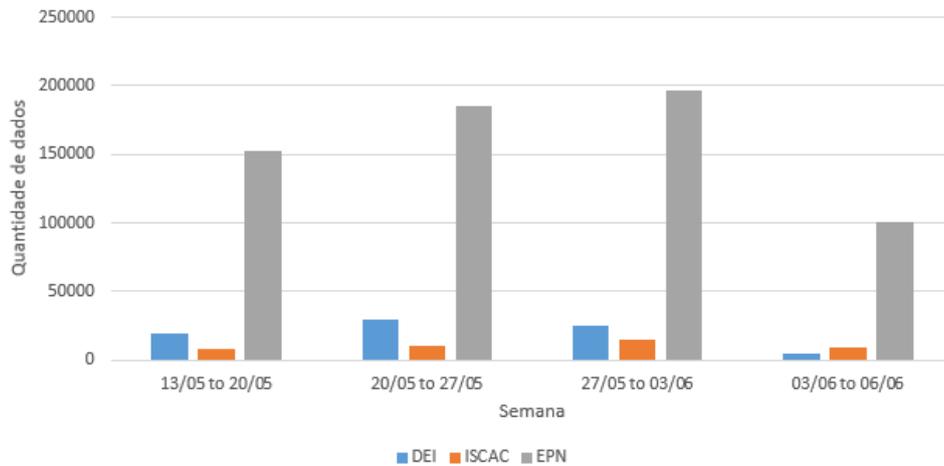


Figura 5.1: Quantidade de dados obtida por semana

A figura 5.2 mostra o número de Sleep Forms preenchidos semanalmente, durante o período de teste. O preenchimento dos Sleep Forms segue o mesmo padrão que a figura 5.1. Apesar da participação do EPN ser bastante boa, os alunos não responderam todos os dias ao questionário. Mas mesmo assim a quantidade de questionários obtida é suficiente para avaliar o dia a dia dos estudantes do EPN, e tentar relacionar o comportamento destes alunos com o seu estado físico e emocional.

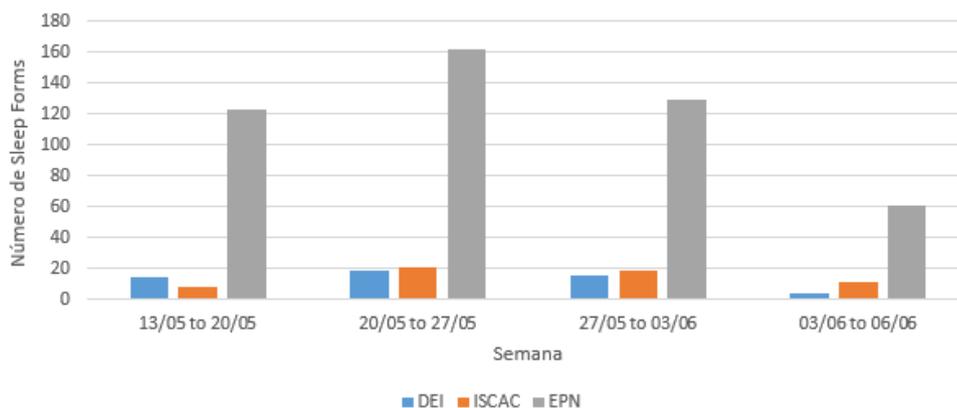


Figura 5.2: Número de questionários obtidos por semana

Como os dados obtidos são maioritariamente do EPN, apenas se analisaram os dados desta turma.

5.3.2 Grupos do teste

Os 28 alunos do Equador que participaram no estudo foram divididos em dois grupos de acordo com a classificação obtida no exame. Para os alunos passarem à disciplina têm que ter um mínimo de 6 em cada um dos dois exames que realizaram. De acordo com isto no primeiro grupo constam os alunos que tiveram uma classificação no exame entre 0 e 6, num total de 10 alunos, e no segundo grupo constam os alunos que tiveram uma classificação no exame entre 6,1 e 10, num total de 18 alunos.

5.3.3 Análise das variáveis

Com o objetivo de relacionar o desempenho académico dos estudantes do Equador com as variáveis, foram analisadas algumas das variáveis do questionário que era respondido diariamente. As variáveis analisadas foram:

- Número de horas dormidas diariamente
- Qualidade do Sono
- Sociabilidade
- Atividade Física
- Estudo

Para cada variável foi calculada a média por grupo. A escala das variáveis "Estudo", "Qualidade do Sono", "Sociabilidade" e "Atividade Física" varia entre 0 e 4.

Para cada variável foi feito um *boxplot* que representa a variação dos dados observados da variável através de quartis. O *boxplot* apresenta o mínimo, o primeiro quartil (Q1), a mediana (linha central no retângulo), o terceiro quartil (Q3) e o máximo. A mediana corresponde ao valor que separa metade superior da metade inferior de um conjunto de dados. A vantagem da mediana em relação à média é que a mediana dá uma ideia melhor do valor médio porque não é tão influenciada pelos extremos da amostra.

Na figura 5.3 está representado o *boxplot* da variável "Estudo" para o grupo 1 e para o grupo 2. Da análise da figura 5.3 pode-se concluir que para o grupo 1 a variável "Estudo" tem um máximo de 4 (muito estudo) e um mínimo de 1 (pouco estudo). O Q1 é 1,91 e o Q3 é 2,79. A mediana é 2,07 e esta está mais próxima do Q1 o que indica que a distribuição dos dados é assimétrica positiva. A média é $\bar{x} = 2,33$ e o desvio-padrão é $\sigma = 0,69$. A amplitude interquartil ($Q3 - Q1$) permite ver a

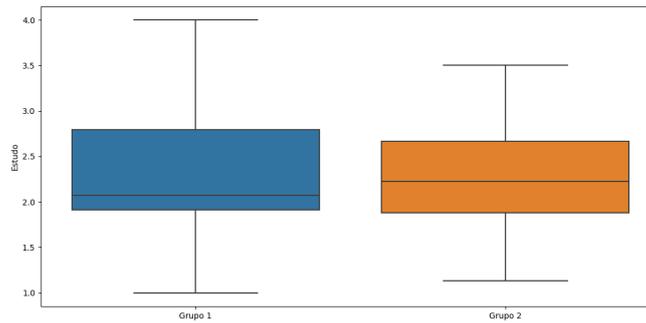


Figura 5.3: Boxplot para a variável "Estudo"

variabilidade dos dados e é uma medida mais robusta que a diferença entre o máximo e o mínimo uma vez que não está sujeita a *outliers*. Para o grupo 1 a amplitude interquartil é de 0,88.

Para o grupo 2 a variável "Estudo" apresenta um máximo de 3,5 e um mínimo de 1. O Q1 é 1,88 e o Q3 é 2,67. A mediana é 2,23 e está mais próxima do Q1. A média é $\bar{x} = 2,27$ e o desvio-padrão é $\sigma = 0,59$. A amplitude interquartil é de 0,79. Como a amplitude interquartil para o grupo 2 é menor que para o grupo1, pode-se concluir que os dados do grupo 1 estão mais dispersos.

Na figura 5.4 está representado o *boxplot* para a variável "Número de horas dormidas diariamente". Da análise desta figura pode-se concluir que para o grupo 1 a variável "Número de horas dormidas diariamente" apresenta um máximo de 8,85 horas e um mínimo de 4,2 horas. O Q1 é 6,14 e o Q3 é 7,52. Apresenta uma mediana de 6,68 e esta está mais próxima do Q1. A média é $\bar{x} = 6,76$ e o desvio-padrão é $\sigma = 1,14$. A amplitude interquartil é 1,39.

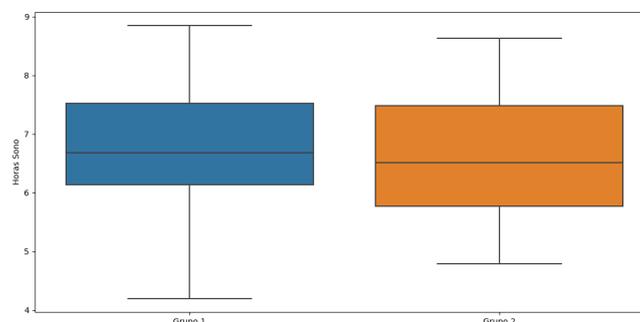


Figura 5.4: Boxplot para a variável "Número de horas dormidas diariamente"

Para o grupo 2 a variável "Número de horas dormidas diariamente" apresenta um

máximo de 8,63 horas e um mínimo de 4,79 horas. O Q1 é 5,78 e o Q3 é 7,49. A mediana é 6,51 e está também mais próxima do Q1. A média é $\bar{x} = 6,58$ e o desvio-padrão é $\sigma = 1,11$. A amplitude interquartil é 1,71. Como a amplitude interquartil para o grupo 2 é maior que para o grupo 1, pode-se concluir que os dados do grupo 2 estão mais dispersos.

Na figura 5.5 está representado o *boxplot* para a variável "Qualidade do Sono". Observado esta figura pode-se concluir que para o grupo 1 o máximo é 3,67 e o mínimo é 1,75. O Q1 é 2,06 e o Q3 é 2,75. A mediana é 2,4 e está mesmo no centro do retângulo, o que indica que a distribuição dos dados é simétrica. A média é $\bar{x} = 2,48$ e o desvio-padrão é $\sigma = 0,45$. A amplitude interquartil é 0,69.

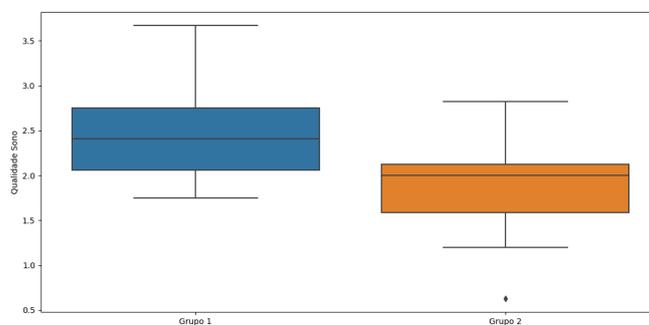


Figura 5.5: Boxplot para a variável "Qualidade do Sono"

Para o grupo 2 a variável "Qualidade do Sono" apresenta um máximo de 2,82 e um mínimo de 1,20. O Q1 é 1,59 e o Q3 é 2,12. Apresenta ainda um *outlier* abaixo do limite inferior ($L_{inf} = \bar{x} + 1,5 * AmplitudeInterquartil$) do retângulo. O *outlier* é o 0,63. A mediana é 2 e está mais próxima do Q3 o que indica que a distribuição dos dados é assimétrica negativa. A média é $\bar{x} = 1,91$ e o desvio-padrão é $\sigma = 0,47$. A amplitude interquartil é 0,54. Como a amplitude interquartil para o grupo 2 é menor que para o grupo 1, pode-se concluir que os dados do grupo 1 estão mais dispersos.

Na figura 5.6 está representado o *boxplot* para a variável "Sociabilidade". Analisando a figura, pode-se observar que para o grupo 1 o máximo é 3,33 e o mínimo é 1,37. O Q1 é 1,92 e o Q3 é 2,66. A mediana é 2,34 e está mais próxima do Q3. A média é $\bar{x} = 2,30$ e o desvio-padrão é $\sigma = 0,47$. A amplitude interquartil é 0,74.

Para o grupo 2 a variável "Sociabilidade" apresenta um máximo de 2,4 e um mínimo de 1,13. O Q1 é 1,63 e o Q3 é 1,98. A mediana é 1,78 e está mais próxima do Q1. A média é $\bar{x} = 1,81$ e o desvio-padrão é $\sigma = 0,31$. A amplitude interquartil é 0,35.

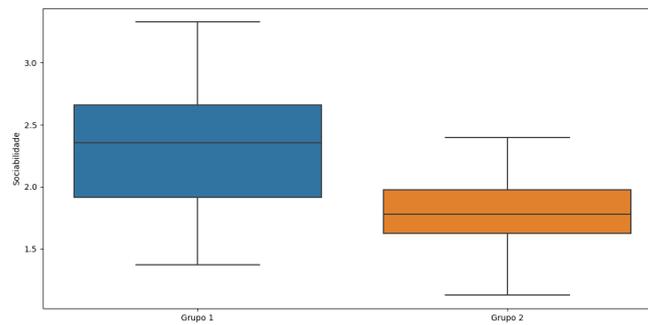


Figura 5.6: Boxplot para a variável "Sociabilidade"

Como a amplitude interquartil para o grupo 2 é menor que para o grupo 1, pode-se concluir que os dados do grupo 1 estão mais dispersos.

Na figura 5.7 está representado o *boxplot* para a variável "Atividade Física". Analisando a figura pode-se observar que para o grupo 1 o máximo é 3 e o mínimo é 1,43. O Q1 é 2,0 e o Q3 é 2,5. A mediana é 2,27 e está mais próxima do Q3. A média é $\bar{x} = 2,22$ e o desvio-padrão é $\sigma = 0,38$. A amplitude interquartil é 0,5.

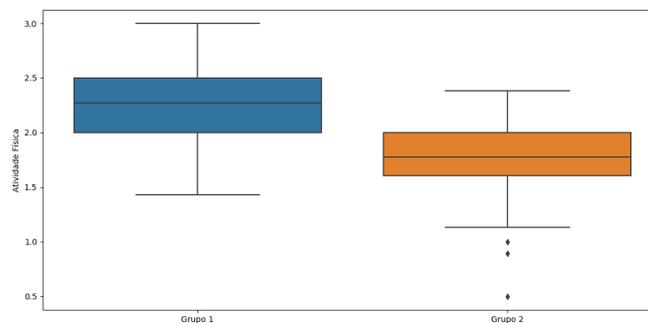


Figura 5.7: Boxplot para a variável "Atividade Física"

Para o grupo 2 a variável "Atividade Física" apresenta um máximo de 2,38 e um mínimo de 1,4. Apresenta ainda 3 *outliers* abaixo do limite inferior do retângulo. Os outliers são o 1, 0,89 e 0,5. O Q1 é 1,61 e o Q3 é 2,0. A mediana é 1,77 e está mais próxima do Q1. A média é $\bar{x} = 1,74$ e o desvio-padrão é $\sigma = 0,42$. A amplitude interquartil é 0,39. Como a amplitude interquartil para o grupo 2 é menor que para o grupo 1, pode-se concluir que os dados do grupo 1 estão mais dispersos.

5.3.4 Relação entre as variáveis

Nesta secção é apresentada a relação entre as variáveis. Para cada variável foi calculada a média por dia e por grupo. Para cada relação foi calculado o coeficiente de correlação de Pearson. Este coeficiente mede a relação linear entre dois conjuntos de dados. O valor deste coeficiente varia entre -1 e +1, onde 0 implica nenhuma correlação. Correlações de -1 ou +1 implicam uma relação linear exata. Correlações positivas implicam que conforme x aumenta, o mesmo acontece com y. Correlações negativas implicam que, quando x aumenta, y diminui.

A figura 5.8 mostra a relação entre a variável "Estudo" e a variável "Qualidade do sono" para o grupo 1.

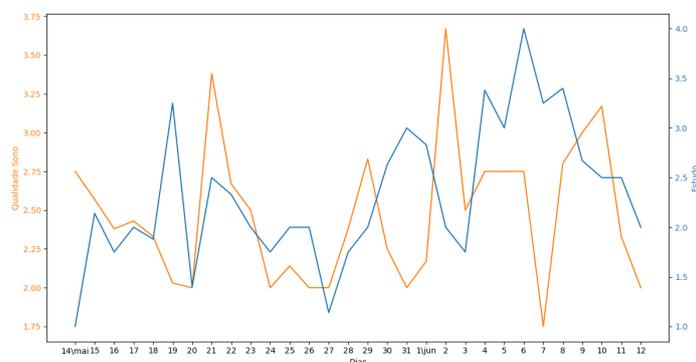


Figura 5.8: Relação entre a variável "Estudo" e a variável "Qualidade do sono" para o grupo 1

Estas variáveis apresentam um coeficiente de correlação de Pearson de 0,115, pode-se considerar que a correlação é desprezível. Nota-se que em alguns dias a qualidade do sono foi pior quando os alunos estudaram mais. Também se verifica que os dias que antecederam o teste (5 e 6 de junho) e o dia do teste (7 de junho) foram os dias em que os alunos mais estudaram e foi também nestes dias que a qualidade do sono foi a pior.

A figura 5.9 mostra a relação ente a variável "Estudo" e a variável "Número de horas dormidas diariamente" para o grupo 1. Estas variáveis apresentam um coeficiente de correlação de Pearson de -0,578, este valor negativo indica que quando uma variável aumenta a outra diminui. Pode se considerar que a correlação é moderada. Observa-se que quando os alunos dormem menos horas é quando estudam mais. Os dias que antecederam o teste foi quando os alunos dormiram menos, e quando estudaram mais.

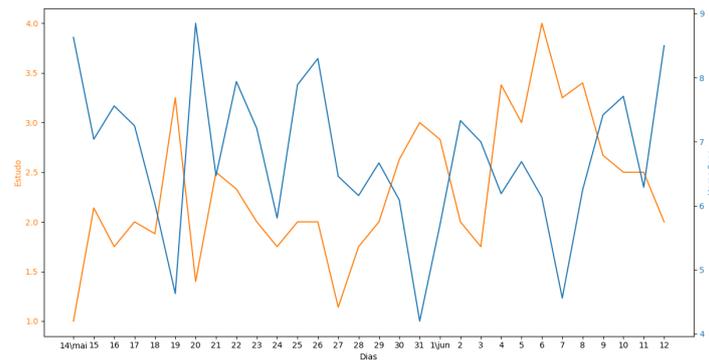


Figura 5.9: Relação entre a variável "Estudo" e a variável "Número de horas de sono dormidas diariamente" para o grupo 1

A figura 5.10 mostra a relação entre a variável "Estudo" e a variável "Sociabilidade" para o grupo 1. Estas variáveis apresentam um coeficiente de correlação de Pearson de 0,185. A correlação entre estas variáveis é desprezível. Apesar do baixo coeficiente de correlação nota-se que em alguns dia quando a sociabilidade apresenta um valor mais elevado o estudo apresenta os valores mais baixos.

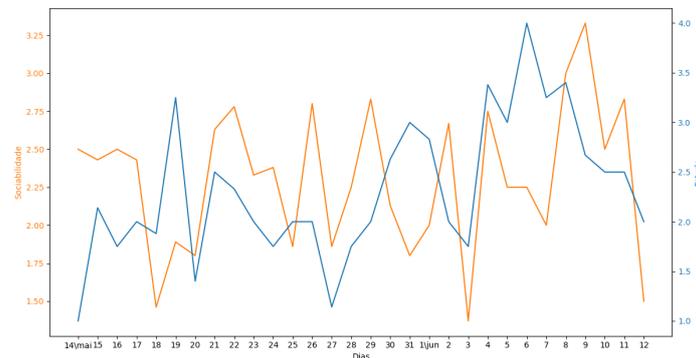


Figura 5.10: Relação entre a variável "Estudo" e a variável "Sociabilidade" para o grupo 1

A sociabilidade é uma variável complicada de analisar, pois depende muito do que cada aluno entende por sociabilidade. No questionário apenas era perguntado o quão sociável os alunos se sentiam. Mas enquanto para uns alunos sociabilidade pode implicar apenas sair de casa para outros pode implicar falar muito através das redes sociais, é portanto uma variável muito subjetiva.

A figura 5.11 mostra a relação entre a variável "Estudo" e a variável "Atividade Física" para o grupo 1. Estas variáveis têm um coeficiente de correlação de Pearson de 0,328. A correlação entre estas variáveis é fraca. Observa-se, por exemplo, que os dias 4,5 e 6 de junho os alunos estudaram muito mas também praticaram atividade

física.

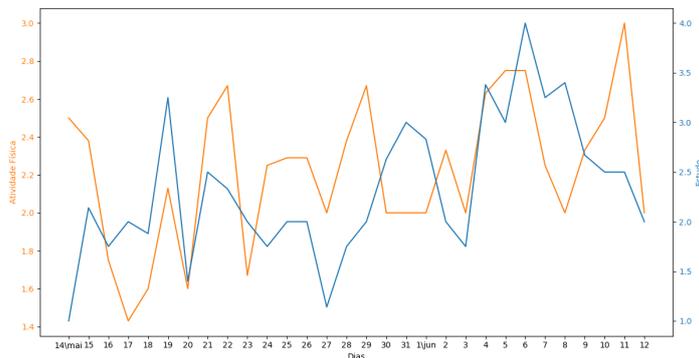


Figura 5.11: Relação entre a variável "Estudo" e a variável "Atividade Física" para o grupo 1

A figura 5.12 mostra a relação entre a variável "Número de horas de sono dormidas diariamente" e a variável "Qualidade do sono" para o grupo 1. Estas variáveis apresentam um coeficiente de correlação de Pearson de 0,240. Considera-se que a correlação entre estas variáveis é desprezível. Observa-se que quando os alunos dormem menos horas a qualidade do sono é pior. Mas também se observa que quando os alunos dormem muito (os picos da figura) a qualidade do sono é baixa. Muitas vezes dormir muito não implica qualidade.

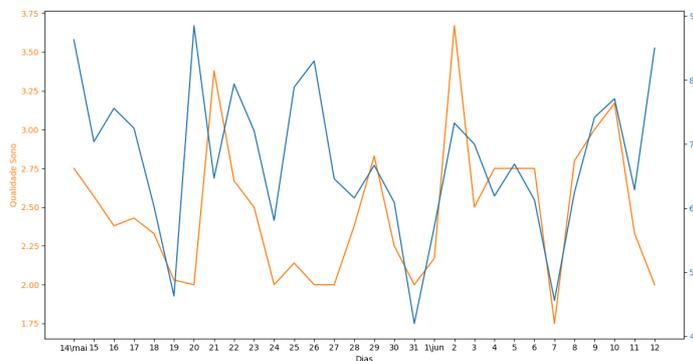


Figura 5.12: Relação entre a variável "Número de horas de sono dormidas diariamente" e a variável "Qualidade do sono" para o grupo 1

A figura 5.13 mostra a relação entre a variável "Qualidade do sono" e a variável "Atividade Física" para o grupo 1. Estas variáveis apresentam um coeficiente de correlação de Pearson de 0,394, considera-se que é uma correlação moderada. Observa-se que de um modo geral que quando os alunos dormem melhor durante a noite, ou seja com mais qualidade, também praticaram mais atividade física durante o dia. Houve dias em que a qualidade do sono não foi muito boa, mas mesmo assim os

alunos praticaram alguma atividade física.

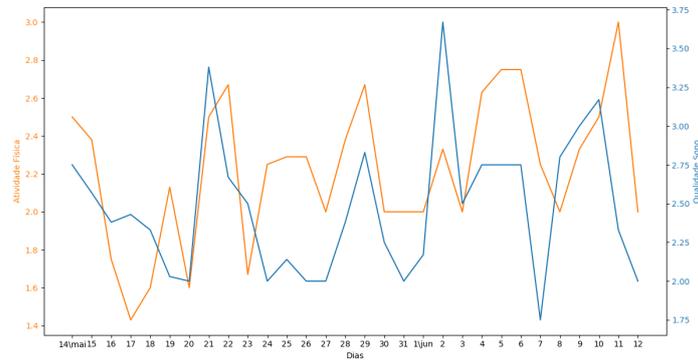


Figura 5.13: Relação entre a variável "Qualidade do sono" e a variável "Atividade Física" para o grupo 1

A figura 5.14 mostra a relação entre a variável "Qualidade do sono" e a variável "Sociabilidade" para o grupo 1. Estas variáveis apresentam um coeficiente de correlação de Pearson de 0,563, a correlação é moderada. De um modo geral, observa-se que quando a qualidade do sono foi boa, os alunos também se sentiram mais sociais durante o dia. Entre os dias 24 e 27 de maio a qualidade do sono foi baixa, mas a sociabilidade até foi alta.

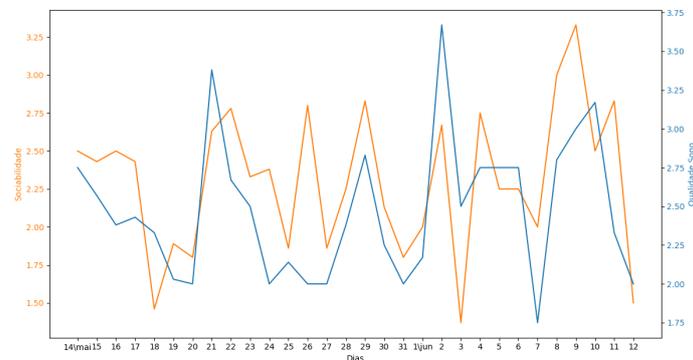


Figura 5.14: Relação entre a variável "Qualidade do sono" e a variável "Sociabilidade" para o grupo 1

A figura 5.15 mostra a relação entre a variável "Sociabilidade" e a variável "Atividade Física" para o grupo 1. Estas variáveis apresentam um coeficiente de correlação de Pearson de 0,476, a correlação é moderada. Analisando a figura, observa-se de um modo geral que nos dias em que a sociabilidade é maior os alunos também praticaram mais atividade física. Quando a sociabilidade é menor a atividade física também é menor.

A figura 5.16 mostra a relação entre a variável "Estudo" e a variável "Qualidade do

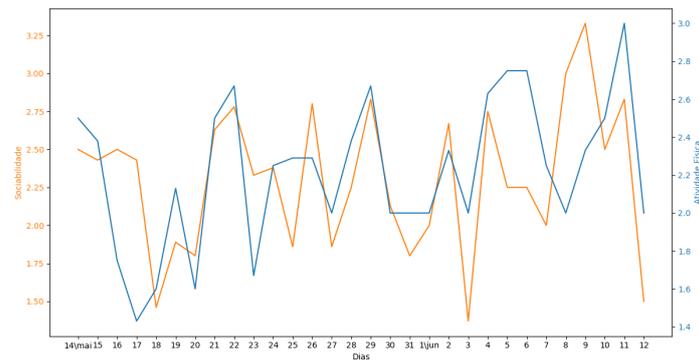


Figura 5.15: Relação entre a variável "Sociabilidade" e a variável "Atividade Física" para o grupo 1

sono" para o grupo 2. Estas variáveis apresentam um coeficiente de correlação de Pearson de $-0,541$, ou seja, a correlação entre estas variáveis é moderada. Da análise da figura 5.16 observa-se que a uma qualidade do sono boa está associado menos estudo. Vê-se claramente que o pico de estudo foi entre os dias 4 e 7 de junho, e foi também entre estes dias que a qualidade do sono foi a pior. Também se observa que a qualidade do sono foi sempre a diminuir desde o dia 14 de maio até ao dia do teste (7 de junho) e que depois voltou a aumentar.

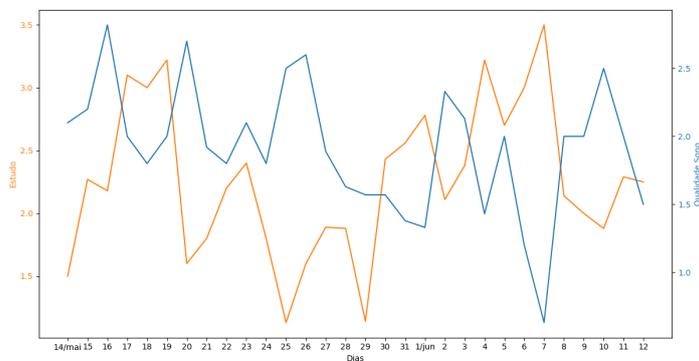


Figura 5.16: Relação entre a variável "Estudo" e a variável "Qualidade do sono" para o grupo 2

A figura 5.17 mostra a relação entre a variável "Estudo" e a variável "Número de horas dormidas diariamente" para o grupo 2. A correlação entre estas variáveis é de $-0,642$ ou seja, a correlação entre estas variáveis é moderada. Os dias em que os estudantes dormem mais corresponde aos dias em que os alunos estudam menos. Mais uma vez os dias em que os alunos dormiram menos foram os dias antes e o dia do teste, e foi também nesses dias que eles mais estudaram. Pode se observar que os alunos dormem mais ao fim-de-semana e estudam menos.

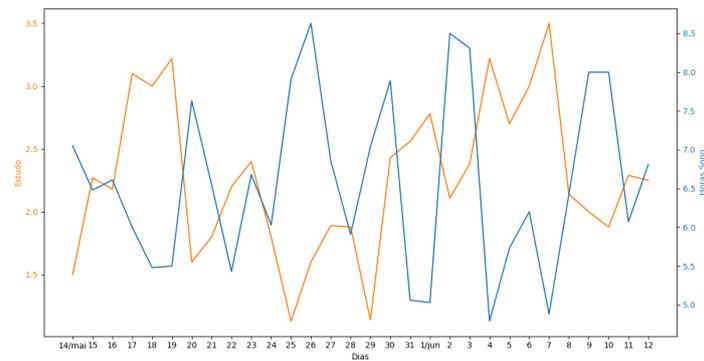


Figura 5.17: Relação entre a variável "Estudo" e a variável "Número de horas dormidas diariamente" para o grupo 2

A figura 5.18 mostra a relação entre a variável "Estudo" e a variável "Sociabilidade" para o grupo 2. Estas variáveis apresentam uma correlação de $-0,377$, ou seja, a correlação entre estas variáveis é fraca. Observa-se que a sociabilidade diminuiu rapidamente nos dias que antecederam o teste, enquanto o aumento do estudo não foi tão acentuado, uma vez que o estudo já estava elevado. Também se pode concluir que os dias que os estudantes se sentiam mais sociais ao fim-de-semana.

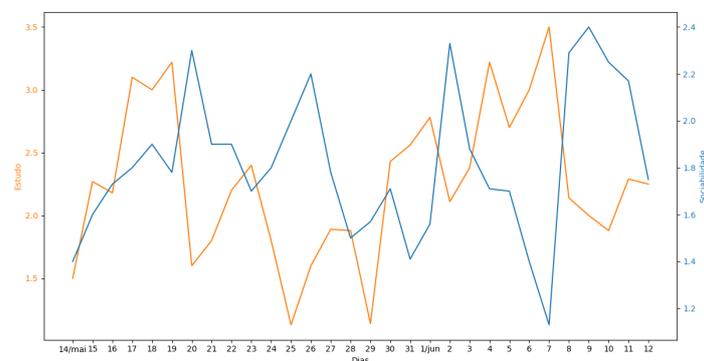


Figura 5.18: Relação entre a variável "Estudo" e a variável "Sociabilidade" para o grupo 2

A figura 5.19 mostra a relação entre a variável "Estudo" e a variável "Atividade Física" para o grupo 2. Estas variáveis apresentam uma correlação de $-0,440$, ou seja, a correlação entre estas variáveis é fraca. Observa-se que no dia 7 de junho foi quando os alunos praticaram menos atividade física e foi quando estudaram mais. De um modo geral, também se observa que nos dias em que os alunos praticam mais atividade física correspondem aos dias em que estudam menos.

A figura 5.20 mostra a relação entre a variável "Número de horas dormidas diariamente" e a variável "Qualidade do sono" para o grupo 2. Estas variáveis têm uma

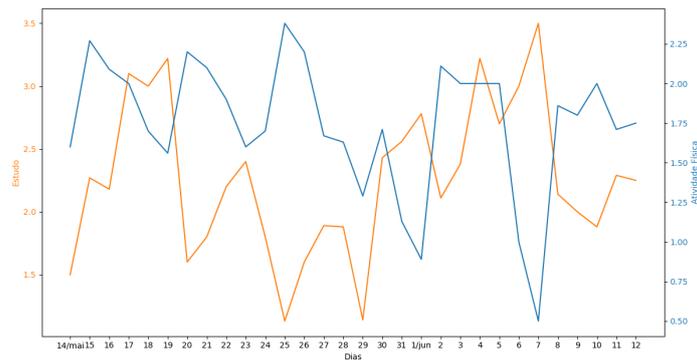


Figura 5.19: Relação entre a variável "Estudo" e a variável "Atividade Física" para o grupo 2

correlação de 0,631, ou seja, estas variáveis apresentam uma correlação moderada. De um modo geral, observa-se nos dias que os alunos dormem mais, são os dias que dormem com mais qualidade. Nos dias antes do teste dá para observar claramente que foram os dias em que os alunos dormiram menos e com pior qualidade.

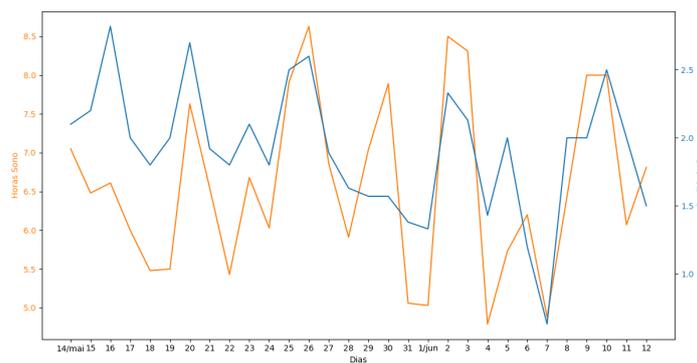


Figura 5.20: Relação entre a variável "Número de horas dormidas diariamente" e a variável "Qualidade do sono" para o grupo 2

A figura 5.21 mostra a relação entre a variável "Qualidade do sono" e variável "Atividade Física" para o grupo 2. Estas variáveis têm uma correlação de 0,822, ou seja, estas variáveis apresentam uma correlação forte. Observa-se que os dias em que a qualidade do sono é melhor corresponde aos dias em que os alunos praticam mais atividade física. Nos dias em que os alunos dormem pior não praticam tanta atividade física. O dia do teste foi o dia em que os alunos praticaram menos atividade física e dormiram pior.

A figura 5.22 mostra a relação entre a variável "Qualidade do sono" e variável "Sociabilidade" para o grupo 2. Estas variáveis apresentam uma correlação de 0,669, ou seja, a correlação entre estas variáveis é moderada. De um modo geral, os dias em

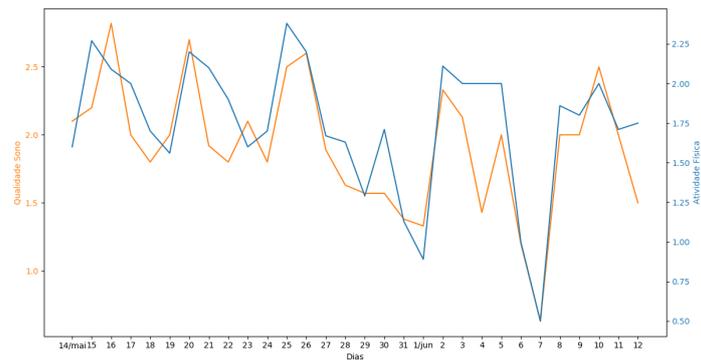


Figura 5.21: Relação entre a variável "Qualidade do sono" e variável "Atividade Física" para o grupo 2

que os alunos dormiram melhor corresponde aos dias em que se sentem mais sociais. O dia do teste foi o dia em que os alunos dormiram pior e se sentiram menos sociais.

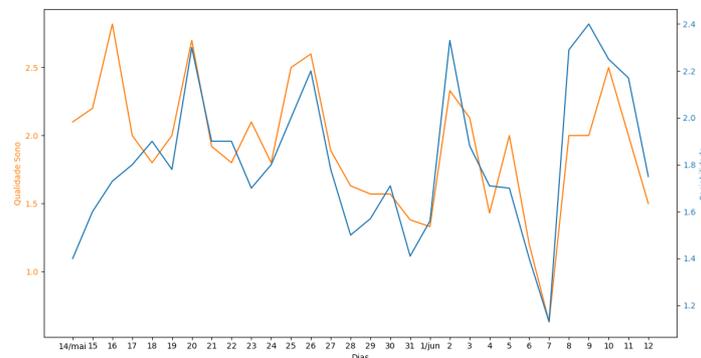


Figura 5.22: Relação entre a variável "Qualidade do sono" e variável "Sociabilidade" para o grupo 2

A figura 5.23 mostra a relação entre a variável "Sociabilidade" e variável "Atividade Física" para o grupo 2. Estas variáveis apresentam uma correlação de 0,661, ou seja, a correlação entre estas variáveis é moderada. De um modo geral, observa-se que os dias em que os alunos se sentiram mais sociais corresponde aos dias em que praticaram mais atividade física. Os dias em que os alunos praticaram mais atividade física e se sentiram mais sociais foi ao fim-de-semana.

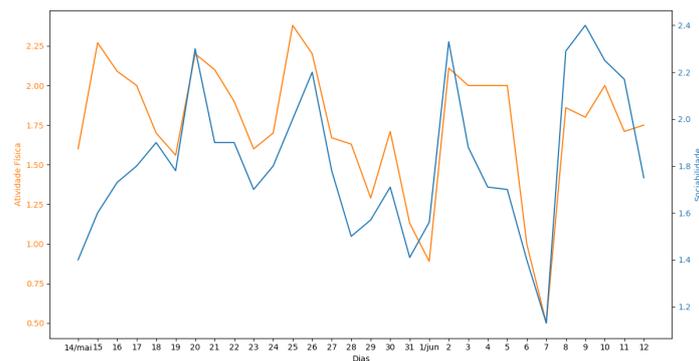


Figura 5.23: Relação entre a variável "Sociabilidade" e variável "Atividade Física" para o grupo 2

5.3.5 Teste de Significância: T-Student

O teste t-Student é utilizado para avaliar se existe alguma diferença significativa entre as médias de duas amostras [53]. As observações feitas com uma amostra da população deve ser independente da outra. Indivíduos de uma população não devem ser individualmente comparados com os da outra população. Os indivíduos dentro de uma amostra não devem estar uns relacionados com os outros.

Para cada variável, a população foi dividida em duas amostras de acordo com o intervalo correspondente. Para as variáveis "Estudo", "Qualidade do Sono", "Sociabilidade" e "Atividade Física" a primeira amostra pertence ao intervalo $[0;2]$ e a segunda amostra ao intervalo $[2,1;4]$. Para a variável "Número de horas dormidas diariamente" a primeira amostra pertence ao intervalo $[0;6,5]$ e a segunda amostra pertence ao intervalo $[6,6;10]$. Para cada aluno foi calculada a média para cada variável. Depois de calculada a média, a nota da classificação no teste foi colocada na amostra correspondente ao intervalo onde pertence a média.

Para cada variável a Hipótese Nula é $H_0 : \mu_1 = \mu_2$, ou seja, que a média das duas amostras é igual, neste caso que a média da classificação nos testes é igual para as duas amostras. A Hipótese Alternativa é $H_1 : \mu_1 \neq \mu_2$, ou seja, que a média das duas amostras é diferente.

O primeiro passo foi calcular o Intervalo de Confiança (IC) para cada variável.

O IC escolhido foi de 95% para a diferença entre duas médias. A equação 5.1 mostra como se calcula o IC.

$$IC = ((\bar{x}_1 - \bar{x}_2) - t \times SE; (\bar{x}_1 - \bar{x}_2) + t \times SE) \quad (5.1)$$

O valor de t é visto na tabela de distribuição t . E o SE é o erro padrão da diferença das médias.

A equação 5.2 mostra como calcular o erro padrão da diferença das médias (SE).

$$SE = s_p \times \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} \quad (5.2)$$

A equação 5.3 mostra como calcular o desvio-padrão da variância combinada (s_p).

$$s_p = \sqrt{\frac{(n_1 - 1) \times \sigma_1^2 + (n_2 - 1) \times \sigma_2^2}{n_1 + n_2 - 2}} \quad (5.3)$$

Onde n_1 é o número de elementos da amostra 1, e σ_1 é o desvio-padrão da amostra 1. E n_2 é o número de elementos da amostra 2, e σ_2 é o desvio-padrão da amostra 2.

O IC de confiança escolhido foi de 95%, então $\alpha = 0,05$. O α dá-nos o risco de se concluir se existe alguma diferença, quando na verdade não existe nenhuma diferença real, neste caso esse risco é de 5%. O $p - value$ depende diretamente de uma dada amostra e tenta fornecer uma medida da força dos resultados de um teste.

Se $p - value \leq \alpha$, significa que a diferença entre a média das amostras é estatisticamente significativa, ou seja, rejeita-se a hipótese nula.

Se $p - value > \alpha$, significa que a diferença entre as médias das amostras não é estatisticamente significativa, ou seja, aceita-se a hipótese nula.

Como $\alpha = 0,05$ e os graus de liberdade são 26 (28-2), então o valor de t na equação 5.1 é $t=2,0555$.

Para a variável "Número de horas dormidas diariamente" o IC= (0,33;1,37), ou seja há 95% de confiança, que em média, os alunos da amostra 1 têm entre 0,33 e 1,37 a mais na classificação do teste em relação à amostra 2. O $p - value$ obtido foi de 0,061, que é maior do que 0,05, logo aceita-se a hipótese nula, ou seja, a diferença da média das duas amostras não é estatisticamente significativa.

Para a variável "Qualidade do Sono" o IC= (-0,43;0,69), ou seja há 95% de confiança, que em média, os alunos da amostra 1 têm entre -0,43 e 0 a menos na classificação do teste em relação à amostra 2. E têm também entre 0 e 0,69 a mais na classificação do teste em relação à amostra 2. O $p - value$ obtido foi de 0,65, que é maior do que 0,05, logo aceita-se a hipótese nula, ou seja, a diferença da média das duas amostras

não é estatisticamente significativa.

Para a variável "Estudo" o IC= (-0,14;0,30), ou seja há 95% de confiança, que em média, os alunos da amostra 1 têm entre -0,14 e 0 a menos na classificação do teste em relação à amostra 2. E têm também entre 0 e 0,30 a mais na classificação do teste em relação à amostra 2. O $p - value$ obtido foi de 0,36, que é maior do que 0,05, logo aceita-se a hipótese nula, ou seja, a diferença da média das duas amostras não é estatisticamente significativa.

Para a variável "Sociabilidade" o IC= (-0,52;0,60). Ou seja há 95% de confiança, que em média, os alunos da amostra 1 têm 0,52 e 0 a menos na classificação do teste em relação à amostra 2. E têm também entre 0 e 0,60 a mais na classificação do teste em relação à amostra 2. O $p - value$ obtido foi de 0,89, que é maior do que 0,05, logo aceita-se a hipótese nula, ou seja, a diferença da média das duas amostras não é estatisticamente significativa.

Para a variável "Atividade Física" o IC= (-0,29;0,80). Ou seja há 95% de confiança, que em média, os alunos da amostra 1 têm entre -0,29 e 0 a menos na classificação do teste em relação à amostra 2. E têm também entre 0 e 0,80 a mais na classificação do teste em relação à amostra 2. O $p - value$ obtido foi de 0,36, que é maior do que 0,05, logo aceita-se a hipótese nula, ou seja, a diferença da média das duas amostras não é estatisticamente significativa.

5.4 Módulo do sono

Esta seção tem como objetivo explicar todo o tratamento dos dados e os testes realizados com os diferentes classificadores para o módulo do sono.

5.4.1 Primeiro tratamento dos dados

Os termos utilizados para ver o desempenho de um classificador estão em inglês para evitar ambiguidades.

Depois de construído o conjunto de dados, foi necessário proceder à normalização dos mesmos. A normalização é feita para garantir que cada atributo pertence ao mesmo intervalo de valores. Os valores foram normalizados de acordo com a equação 5.4 [20].

$$f(x) = \frac{(x - d_L)(n_H - n_L)}{(d_H - d_L)} + n_L \quad (5.4)$$

A equação 5.4 normaliza um valor x , onde a variável \mathbf{d} , representa o valor mais alto e mais baixo de cada atributo e a variável \mathbf{n} representa o valor mais alto e mais baixo do intervalo de normalização. O intervalo de normalização escolhido foi $[0,1]$.

Os classificadores de aprendizagem automática, não conseguem lidar com conjuntos de dados de treino não balanceados, uma vez que estes são sensíveis às proporções das diferentes classes. Como consequência, estes algoritmos tendem a favorecer a classe com a maior proporção de observações (classe maioritária) o que pode conduzir a conclusões erradas. Isso é particularmente problemático quando estamos interessados na classificação correta da classe minoritária, mas encontramos precisões elevadas, que são na verdade o resultado da classificação correta da classe maioritária [54].

De modo a não ter conclusões erradas, o conjunto de dados foi balanceado recorrendo ao filtro *Resample* do WEKA. Para este projeto foi feita amostragem com substituição de modo a ficar com uma distribuição uniforme das duas classes.

Depois do conjunto de dados estar normalizado e balanceado, utilizou-se o método *InfoGainAttributeEval* do software WEKA [22], para perceber quais os atributos mais relevantes para a classificação do sono. A tabela 5.1 apresenta o *ranking* dos atributos por ordem decrescente de relevância.

Atributo	Ganho de Informação
Som	0,151
Tempo para o alarme seguinte	0,149
Luz	0,137
Atividade	0,029
Bloqueio do ecrã	0,013
Dia da semana	0,002

Tabela 5.1: Relevância dos atributos para a deteção do sono

Utilizou-se também o método *CorrelationAttributeEval* do software WEKA para ver quais os atributos que melhor se correlacionavam com a classificação. O resultado pode ser visto na tabela 5.2.

O atributo "Dia da semana" tem uma relevância baixa, quando comparado com os restantes atributos, e não está muito correlacionado com a classificação, por estas razões eliminou-se este atributo.

Numa primeira abordagem recorreu-se ao software WEKA para testar o desempenho de vários classificadores e determinar qual o melhor.

Atributo	Correlação
Som	0,165
Tempo para o alarme seguinte	0,125
Luz	0,090
Atividade	0,045
Bloqueio do ecrã	0,025
Dia da semana	0,013

Tabela 5.2: Valor da correlação dos atributos

Para ver o desempenho de cada classificador, comparou-se o valor da *precision*, do *recall* e área ROC de cada classificador. Os resultados podem ser vistos na tabela 5.3 .

Utilizou-se *Cross Validation* para a classificação dos dados. A *Cross Validation* [55] é um recurso importante pois ajuda a desenvolver e a ajustar os modelos de mineração de dados. Além disso é importante pois valida a robustez de um determinado modelo e constrói vários modelos identificando o melhor através de uma análise estatística.

A *precision* é a percentagem dos elementos que são realmente positivos de todos que foram classificados como positivos.[56].

O *recall* é a percentagem dos verdadeiros positivos dentro de todos os exemplos cuja classe esperada é a classe positiva [56].

A *F-measure* faz o balanceamento entre a *precision* e o *recall*.

A área sob a curva ROC (*Receiver Operating Characteristic*) é obtida pela representação da taxa de verdadeiros positivos versus a taxa de falsos positivos. Quanto mais perto de 1 for a área melhor é o teste [57].

Classificador	<i>Precision</i>	<i>Recall</i>	ROC
Naive Bayes	0,689	0,750	0,653
Bayes Net	0,795	0,799	0,795
Logistic	0,754	0,760	0,708
ZeroR	0,575	0,758	0,500
J48	0,764	0,774	0,588
Random Forest	0,906	0,907	0,957
Random Tree	0,904	0,905	0,930

Tabela 5.3: Desempenho dos diferentes classificadores, recorrendo ao software WEKA

O Random Forest [58] foi o classificador que apresentou melhores resultados no *software* WEKA. É um classificador que cria várias árvores de decisão e procura relações

não lineares nos dados, o que permite fazer correlações que não seriam possíveis com outros classificadores. Para além disso é dos métodos não lineares menos computacionalmente exigentes.

O classificador que apresentou os piores resultados é o ZeroR [59]. Este classificador tem como base apenas o alvo e ignora os preditores, ou seja, apenas prevê qual a classe maioritária.

Como o classificador *Random Forest* foi o que apresentou melhores resultados do software WEKA, procedeu-se à implementação deste em JAVA, onde se foi variando o número de árvores de decisão. Os resultados do desempenho do classificador Random Forest podem ser vistos na tabela 5.4. O número de árvores de decisão não foi aumentado acima de 50 pois o processamento ficava muito lento. Para ver o desempenho do Random Forest foi utilizada a validação cruzada.

Nº de árvores de decisão	<i>Precision</i>	<i>Recall</i>	<i>F-measure</i>
1	0,933	0,932	0,932
2	0,940	0,940	0,939
5	0,943	0,941	0,41
10	0,945	0,944	0,944
20	0,945	0,977	0,961
50	0,948	0,987	0,967

Tabela 5.4: Desempenho do classificador Random Forest, variando o número de árvores de decisão

Na experiência de Chen Z et al [50] foi utilizada uma rede neuronal e a *accuracy* obtida foi bastante boa. Deste modo procedeu-se também à implementação de uma rede neuronal simples em Java recorrendo à livreria Encog [20], onde se foi variando o número de neurónios nas duas camadas escondidas. Foi utilizada a função de ativação *Log-Sigmoid* em cada camada escondida. O número de neurónios máximo utilizados nas camadas escondidas foi 20 e este valor não foi aumentado, pois o processamento ficava muito lento. O resultado do desempenho da rede neuronal pode ser visto na tabela 5.5

Como o classificador Random Forest apresentava melhores resultados que a rede neuronal, foi utilizado este classificador para a deteção do sono.

5.4.2 Segundo tratamento dos dados

Após alguns testes para ver se realmente o modelo de classificação estava correto, concluiu-se que o modelo não estava a classificar muito bem novos dados. As prin-

Estrutura da rede	<i>Precision</i>	<i>Recall</i>	<i>F-measure</i>
1-1	0,501	0,428	0,462
2-2	0,567	0,620	0,592
3-3	0,865	0,641	0,736
5-5	0,863	0,821	0,841
10-10	0,911	0,963	0,936
15-15	0,921	0,973	0,946
20-20	0,941	0,977	0,959

Tabela 5.5: Desempenho da rede neuronal, variando o número de neurónios das camadas escondidas

principais razões encontradas para este erro na classificação foram:

- os valores do som não estavam normalizados de forma correta;
- o atributo do tempo para o alarme seguinte estava a influenciar muito a classificação.

A escala de decibéis é uma escala logarítima, usada mundialmente para medir o nível do som. Com o tratamento de dados feito anteriormente não se estava a utilizar o valor do som em decibéis. E ao dividir o valor do som pelo valor máximo, os valores estavam todos muito perto de zero, o que não permitia uma distinção clara entre a classe "Dormir" e "Acordado".

Antes de normalizar os valores do som entre 0 e 1 é necessário obter o valor do som em Decibel. Para resolver este problema utilizou-se a equação 5.5 [60], a equação 5.6 e a equação 5.7.

$$som_dB = 20 \times \log_{10}(som/32768) \quad (5.5)$$

Na equação 5.5 o valor do som é transformado em Decibel. O valor máximo do som que o microfone deteta é 32768. Quando o valor do som é zero ou inferior a zero significa que o microfone está a ser utilizado noutra aplicação. Nestes casos assume-se que o som_dB é 0.

Na equação 5.6 o valor do som_dB é somado ao logaritmo do valor máximo, para os valores do som ficarem positivos.

$$som_dB_positivo = 20 \times \log_{10}(32768) + som_dB \quad (5.6)$$

Na equação 5.7 os valores do som são normalizados entre 0 e 1. Os valores são

divididos por 90, pois $20 \times \log_{10}(32768) = 90$.

$$\text{valor_som_normalizado} = \frac{\text{som_dB_positivo}}{90} \quad (5.7)$$

O atributo do tempo para o alarme seguinte foi eliminado do conjunto de dados. Pois mesmo quando não existe nenhum alarme ativado a aplicação deteta sempre um alarme. E muitas vezes os utilizadores têm alarmes ligados só para se lembrarem de algo e não propriamente para acordar.

Como para a primeira classificação dos dados o classificador Random Forest apresentava melhores resultados, foi utilizado apenas este para classificar os dados com o novo tratamento. Na tabela 5.6 está representado o desempenho do classificador Random Forest.

Nº de árvores de decisão	<i>Precision</i>	<i>Recall</i>	<i>F-measure</i>
1	0,827	0,825	0,825
2	0,832	0,830	0,830
5	0,834	0,832	0,832
10	0,835	0,833	0,833
20	0,901	0,897	0,899
50	0,941	0,939	0,939

Tabela 5.6: Desempenho do classificador Random Forest no novo conjunto de dados

Apesar do desempenho do classificador não ser tão bom como os resultados da tabela 5.4, verificou-se que o classificador estava a detetar melhor se uma pessoa ou está a dormir ou acordada.

5.5 Validação do sistema ISABELA

No futuro pretende-se que o sistema ISABELA dê *feedback* constante ao aluno sobre o seu estado físico e emocional através de mecanismos de IA.

Para validar os dados do questionário, foi utilizada a informação da aplicação *Android* ISABELA. A informação da caixa IoT não foi utilizada porque foram utilizados apenas os dados dos alunos da EPN, e no Equador não foi colocada nenhuma caixa IoT. A informação do módulo social (Facebook e Twitter) também não foi utilizada porque os alunos não utilizam muito as redes sociais e não havia dados para tratar.

5.5.1 Qualidade do Sono

O objetivo deste módulo é aferir a qualidade do sono, para depois informar o aluno e este ficar a saber da sua qualidade de sono diária.

Para classificar a "Qualidade do Sono" foram utilizados os seguintes atributos:

- Atividade
- Luz
- Som
- Bloqueio do ecrã
- Proximidade ao telemóvel

No questionário diário a escala da variável "Qualidade do Sono" era entre 0 e 4. De modo a tornar esta classe binária para a classificação, assumiu-se que entre o 0 e 2 a "Qualidade do Sono" era 0 e entre 2,1 e 4 a "Qualidade do Sono" era 1.

Primeiro juntou-se os dados dos alunos da EPN todos num ficheiro. Em seguida os atributos foram normalizados entre 0 e 1. Os atributos "Atividade", "Luz", "Bloqueio do ecrã" e "Proximidade ao telemóvel" foram normalizados de acordo com a equação 5.4. O atributo "Som" foi normalizado de acordo com as equações 5.5, 5.6 e 5.7. Para classificar a "Qualidade do sono" utilizou-se apenas os dados da aplicação quando o estudante estava a dormir (recorrendo ao módulo do sono, anteriormente implementado).

Utilizou-se o método *InfoGainAttributeEval* do software WEKA para perceber quais o atributos mais relevantes para a qualidade do sono. A tabela 5.7 apresenta o *ranking* dos atributos por ordem decrescente de relevância.

Atributo	Ganho de Informação
Som	0,140
Proximidade ao ecrã	0,120
Atividade	0,069
Luz	0,021
Bloqueio do ecrã	0,011

Tabela 5.7: Relevância dos atributos para a "Qualidade do sono"

Utilizou-se também o método *CorrelationAttributeEval* do software WEKA para ver quais os atributos que melhor se correlacionavam com a classificação. O resultado pode ser visto na tabela 5.8.

Atributo	Correlação
Proximidade ao ecrã	0,393
Som	0,263
Atividade	0,201
Bloqueio do ecrã	0,188
Luz	0,101

Tabela 5.8: Valor da correlação dos atributos para a "Qualidade do sono"

Da análise do ganho de informação e da correlação dos atributos com a classificação, não se eliminou nenhum atributo, pois não há nenhum atributo com uma relevância e com uma correlação muito baixa.

Na classificação para o módulo do sono, o classificador Random Forest foi o classificador que apresentou os melhores resultados, então para classificar a "Qualidade do sono" foi utilizado este classificador. Na tabela 5.9 está representado o desempenho do classificador Random Forest.

Nº de árvores de decisão	<i>Precision</i>	<i>Recall</i>	<i>F-measure</i>
1	0,795	0,793	0,794
2	0,793	0,796	0,794
5	0,812	0,809	0,810
10	0,817	0,813	0,814
20	0,820	0,816	0,818
50	0,829	0,825	0,827
100	0,915	0,910	0,912

Tabela 5.9: Desempenho do classificador Random Forest para a "Qualidade do sono"

Como se pode ver o classificador apresenta o melhor desempenho para 100 árvores de decisão. Depois de alguns testes verificou-se que o classificador estava a detetar se a qualidade do sono era boa (1) ou má (0). No futuro poderá usar-se este modelo para informar o estudante da sua qualidade do sono, não sendo necessário o estudante responder ao questionário diário.

5.5.2 Sociabilidade

O objetivo deste módulo é de informar ao estudante como está a sua sociabilidade.

Como não existiam dados associados às redes sociais (Facebook e Twitter) devido à falta de uso destas, para classificar a "Sociabilidade" foram utilizados os seguintes atributos da aplicação:

- Chamadas recebidas
- Chamadas efetuadas
- Duração das chamadas
- Destinatários chamadas
- SMS recebidos
- SMS enviados
- Destinatários SMS

No questionário diário a escala da variável "Sociabilidade" era entre 0 e 4. De modo a tornar esta classe binária para a classificação, assumiu-se que entre 0 e 2 a "Sociabilidade" era 0 e entre 2,1 e 4 a "Sociabilidade" era 1.

Novamente todos os dados dos alunos do EPN relativos à sociabilidade foram juntos num ficheiro. Foi criada uma variável "destinatário" e sempre que os destinatários das chamadas e dos SMS eram diferente era adicionado 1 a essa variável. Os atributos "Chamadas recebidas" e "Chamadas efetuadas" foram adicionados ficando com a variável "Número total de chamadas". Os atributos "SMS recebidos" e "SMS enviados" foram adicionados ficando com a variável "Número total de SMS".

Utilizou-se o método *InfoGainAttributeEval* do software WEKA para perceber quais o atributos mais relevantes para a classificação da sociabilidade. A tabela 5.10 apresenta o *ranking* dos atributos por ordem decrescente de relevância.

Atributo	Ganho de Informação
Número total de SMS	0,183
Destinatários	0,108
Tempo Chamadas	0,002
Número total de chamadas	0,001

Tabela 5.10: Relevância dos atributos para a "Sociabilidade"

Utilizou-se também o método *CorrelationAttributeEval* do software WEKA para ver quais os atributos que melhor se correlacionavam com a classificação. O resultado pode ser visto na tabela 5.11.

Da análise do ganho de informação e da correlação dos atributos com a classificação, pode-se ver que o "Tempo das Chamadas" e o "Número total de Chamadas" são os atributos com menos ganho de informação e menos correlacionados com a classificação. Apesar disso não se eliminou estes atributos para não ficar apenas com o

Atributo	Correlação
Número total de SMS	0,453
Destinatários	0,385
Tempo das Chamadas	0,012
Número total de chamadas	0,005

Tabela 5.11: Valor da correlação dos atributos para a "Sociabilidade"

"Número total de SMS" e os "Destinatários". Ao analisar os dados podia-se ver que mesmo quando os estudantes falavam muito ao telemóvel, não implicava que eles se sentissem mais sociais, outras vezes eles não falavam nenhuma vez ao telemóvel mas sentiam-se muito sociais. É pelo facto de o "Número total de chamadas" não contribuir muito para a classificação, que o "Número de SMS" tem uma correlação mais elevada com a classificação. Ao fazer a análise dos dados deu para perceber que havia alunos que não enviavam mensagens e se sentiam bastante sociais e outros que enviavam bastantes mensagens e não se sentiam sociais. Como já foi dito anteriormente o conceito de sociabilidade depende muito da interpretação pessoal de cada indivíduo.

Em seguida recorreu-se ao classificador Random Forest para classificar os dados da "Sociabilidade". O desempenho do classificador pode ser visto na tabela 5.12

Nº de árvores de decisão	<i>Precision</i>	<i>Recall</i>	<i>F-measure</i>
1	0,659	0,652	0,653
2	0,706	0,671	0,671
5	0,672	0,661	0,662
10	0,677	0,662	0,663
20	0,673	0,660	0,661
50	0,708	0,680	0,694
100	0,715	0,683	0,698

Tabela 5.12: Desempenho do classificador Random Forest para a "Sociabilidade"

Como se pode ver os resultados do classificador não são muito bons. Para medir a sociabilidade são necessários mais atributos. Como já foi dito anteriormente o número de chamadas e a duração de chamadas não quer dizer nada, porque houve alunos que num dia não realizaram nenhuma chamada e se sentiram muito sociais e outros alunos fizeram imensas chamadas e não se sentiram muito sociais. Também o facto de não termos dados das redes sociais não permite tirar muitas conclusões. É necessário pensar noutra estratégia para definir a sociabilidade.

6

Conclusões

O objetivo deste projeto foi desenvolver soluções que permitissem aferir estados físicos e emocionais dos utilizadores através da informação recolhida pelos diferentes sensores da aplicação ISABELA. Também era pretendido compreender o impacto das situações diárias no *stress*, padrões de sono, atividade física, sociabilidade e bem-estar emocional do utilizador e relacionar estes parâmetros com o desempenho académico.

Além destes objetivos iniciais foram também desenvolvidos dois módulos para a aplicação ISABELA, o módulo do sono e o módulo da atividade no telemóvel. O desenvolvimento destes módulos foi muito importante pois para além de recolher mais informação sobre o dia-a-dia do estudante permitiu-me ter mais conhecimento sobre o desenvolvimento da aplicação e perceber melhor a comunicação com o FIWARE.

Os atrasos na realização dos testes, devido a toda a questão da privacidade dos dados, não permitiu fechar o ciclo (*Human in the Loop*), ou seja, dar conselhos/recomendações, sobre as variáveis do questionário, em tempo útil, aos estudantes para que estes consigam melhorar o seu desempenho académico. Este é um aspeto fundamental da aplicação, e agora que já existem dados é possível tentar fechar este ciclo, fica como sugestão para trabalho futuro.

É necessário definir uma nova estratégia para obter a sociabilidade do estudante. Atualmente os alunos passam muito tempo nas redes sociais, mas não são feitas muitas publicações escritas, normalmente são feitas publicações com fotos e uma breve descrição desta e são colocados "gostos" noutras publicações e páginas. Uma possível estratégia para obter a sociabilidade pode passar por analisar os "gostos" colocados nas fotos, em que tipo de fotos e páginas são colocados os "gostos". Também se pode analisar as fotos publicadas pelos estudantes, e tentar perceber se uma foto é mais feliz ou triste. Além do Facebook e do Twitter, tentar utilizar também o Instagram e o Whatsapp (a maioria das pessoas atualmente utiliza o Whatsapp para se comunicar e não chamadas ou mensagens) para obter mais informação. Apesar

de no ano anterior terem havido reuniões com a Faculdade de Psicologia, estas não deram frutos para já. Mas é importante ter alguém desta área para se perceber melhor o conceito de sociabilidade, e que parâmetros devem realmente ser usados.

A realização deste projeto foi muito importante para o meu desenvolvimento pessoal, uma vez que me ajudou a melhorar muitas das habilidades que adquiri durante o curso, ajudou-me também a não desistir apesar das dificuldades encontradas, adquiri novas habilidades e aprofundei os meus conhecimentos em desenvolvimento de *software*.

Bibliografia

- [1] G. Schirner, D. Erdogmus, K. Chowdhury, and T. Padir, “The future of human-in-the-loop cyber-physical systems,” *Computer*, vol. 46, no. 1, pp. 36–45, 2013.
- [2] “Mercado M2M.” <https://pt.linkedin.com/pulse/mercado-de-conex%C3%B5es-m2m-valer%C3%A1-us-35-bilh%C3%B5es-em-2020-ramundo>. Acedido: 08-12-2017.
- [3] “Sistemas Ciber-físicos.” <https://www.hbm.com/pt/6267/sistemas-ciber-fisicos-o-fio-que-une-toda-a-iiot/>. Acedido: 06-12-2017.
- [4] G. Eysenbach, “What is e-health?,” *Journal of medical Internet research*, vol. 3, no. 2, 2001.
- [5] G. Schirner, D. Erdogmus, K. Chowdhury, and T. Padir, “The future of human-in-the-loop cyber-physical systems,” *Computer*, vol. 46, no. 1, pp. 36–45, 2013.
- [6] J.-W. Tan, A. O. Andrade, H. Li, S. Walter, D. Hrabal, S. Rukavina, K. Limbrecht-Ecklundt, H. Hoffman, and H. C. Traue, “Recognition of intensive valence and arousal affective states via facial electromyographic activity in young and senior adults,” *PloS one*, vol. 11, no. 1, p. e0146691, 2016.
- [7] J. Torous, “The case for digital psychiatry,” *IEEE Spectrum*, 2017.
- [8] A. Ortony, G. L. Clore, and A. Collins, *The cognitive structure of emotions*. Cambridge university press, 1990.
- [9] B. R. Steunebrink, M. Dastani, and J.-J. C. Meyer, *The OCC model revisited*. Department of Information and Computing Sciences, Utrecht University, The Netherlands, 2009.
- [10] A. Mehrabian, “Analysis of affiliation-related traits in terms of the pad temperament model,” *The Journal of psychology*, vol. 131, no. 1, pp. 101–117, 1997.

- [11] “Internet das Coisas.” <http://internetofthingsagenda.techtarget.com/definition/Internet-of-Things-IoT>. Acedido: 17-10-2017.
- [12] “O que é a Inteligência Artificial?.” <http://searchcio.techtarget.com/definition/AI>. Acedido: 15-10-2017.
- [13] “Inteligência Artificial.” <https://futureoflife.org/background/benefits-risks-of-artificial-intelligence/>. Acedido: 17-10-2017.
- [14] “IA Fraca vs IA Forte.” <http://iaexpert.com.br/index.php/2017/01/17/ia-forte-x-ia-fraca/>. Acedido: 13-12-2017.
- [15] “Qual a diferença entre Inteligência Artificial e Machine Learning.” https://medium.com/@sergio_henr_/qual-a-diferen%C3%A7a-entre-intelig%C3%A7%C3%A3o-artificial-machine-learning-e-deep-learning-c33d8910bf31. Acedido: 13-12-2017.
- [16] “Machine Learning.” <https://www.coursera.org/learn/machine-learning>. Acedido: 13-12-2017.
- [17] “Android.” <http://www.techtudo.com.br/artigos/noticia/2011/01/afinal-o-que-e-android.html>. Acedido: 15-10-2017.
- [18] “Android Wear.” <https://www.android.com/wear/>. Acedido: 15-10-2017.
- [19] “O que é o Android?.” <https://tecnoblog.net/145067/vendas-smartphones-terceiro-trimestre-2013/>. Acedido: 15-10-2017.
- [20] H. Jeff, “Programming neural networks with encog3 in java,” 2011.
- [21] “Encog Machine Learning Framework.” <http://www.heatonresearch.com/encog/>. Acedido: 16-10-2017.
- [22] “WEKA.” <https://www.cs.waikato.ac.nz/ml/weka/index.html>. Acedido: 2017-27-11.
- [23] M. Fazio, A. Celesti, F. G. Márquez, A. Glikson, and M. Villari, “Exploiting the fiware cloud platform to develop a remote patient monitoring system,” in *Computers and Communication (ISCC), 2015 IEEE Symposium on*, pp. 264–270, IEEE, 2015.
- [24] “Generic Enablers.” <https://catalogue.fiware.org/>. Acedido: 05-02-2018.
- [25] W. McKinney, *Python for data analysis: Data wrangling with Pandas, NumPy, and IPython*. ”O’Reilly Media, Inc.”, 2012.

-
- [26] J. Loeliger and M. McCullough, *Version Control with Git: Powerful tools and techniques for collaborative software development.* "O'Reilly Media, Inc.", 2012.
- [27] "Slack." <https://slack.com/>. Acedido: 26-01-2018.
- [28] "Skype." <https://www.skype.com/pt/>. Acedido: 26-01-2018.
- [29] "GitLab." <https://about.gitlab.com/>. Acedido: 26-01-2018.
- [30] "Regulamento (UE) 2016/679 do Parlamento e do Conselho, de 27.04.2016." <https://www.anacom.pt/render.jsp?contentId=1385132>. Acedido: 25-03-2018.
- [31] J. Raftery, *Risk analysis in project management.* Routledge, 2003.
- [32] S. Koelstra, C. Muhl, M. Soleymani, J.-S. Lee, A. Yazdani, T. Ebrahimi, T. Pun, A. Nijholt, and I. Patras, "Deap: A database for emotion analysis; using physiological signals," *IEEE Transactions on Affective Computing*, vol. 3, no. 1, pp. 18–31, 2012.
- [33] "studentLife." <http://studentlife.cs.dartmouth.edu/>. Acedido: 15-10-2017.
- [34] R. Wang, F. Chen, Z. Chen, T. Li, G. Harari, S. Tignor, X. Zhou, D. Ben-Zeev, and A. T. Campbell, "Studentlife: assessing mental health, academic performance and behavioral trends of college students using smartphones," in *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, pp. 3–14, ACM, 2014.
- [35] "ORION." <https://fiware-orion.readthedocs.io/en/master/>. Acedido: 05-03-2018.
- [36] "IDAS." <https://catalogue.fiware.org/enablers/backend-device-management-idas>. Acedido: 05-03-2018.
- [37] "MySQL." <https://www.mysql.com/>. Acedido: 05-03-2018.
- [38] "PostgreSQL." <https://www.postgresql.org/>. Acedido: 05-03-2018.
- [39] "CYGNUS." <https://github.com/telefonicaid/fiware-cygnus>. Acedido: 05-03-2018.
- [40] "KEYROCK." <https://catalogue.fiware.org/enablers/identity-management-keyrock>. Acedido: 05-03-2018.

- [41] “COMET.” <https://fiware-orion.readthedocs.io/en/master/https://github.com/telefonicaid/fiware-sth-comet>. Acedido: 05-03-2018.
- [42] “CKAN.” <https://catalogue.firmware.org/enablers/firmware-ckan-extensions>. Acedido: 05-03-2018.
- [43] “RescueTime.” <https://www.rescuetime.com/>. Acessado: 21-11-2017.
- [44] “Life Cycle.” <https://itunes.apple.com/us/app/life-cycle-track-your-time/id1064955217?mt=8>. Acedido: 27-11-2017.
- [45] “ATracker.” http://www.wonderapps.se/atracker_android/#home. Acedido: 27-11-2017.
- [46] “Sleep Cycle alarm clock.” <https://itunes.apple.com/us/app/sleep-cycle-alarm-clock/id320606217?mt=8>. Acedido: 27-11-2017.
- [47] “A ciência do Sono.” <https://jamesclear.com/sleep#ThePurposeofSleep>. Acedido: 30-11-2017.
- [48] M. A. Grandner, L. Hale, M. Moore, and N. P. Patel, “Mortality associated with short sleep duration: the evidence, the possible mechanisms, and the future,” *Sleep medicine reviews*, vol. 14, no. 3, pp. 191–203, 2010.
- [49] “As melhores 10 aplicações para o sono.” <https://www.medicalnewstoday.com/articles/317816.php>. Acedido: 30-11-2017.
- [50] W. Chen, A. Sano, D. L. Martinez, S. Taylor, A. W. McHill, A. J. Phillips, L. Barger, E. B. Klerman, and R. W. Picard, “Multimodal ambulatory sleep detection,” in *Biomedical & Health Informatics (BHI), 2017 IEEE EMBS International Conference on*, pp. 465–468, IEEE, 2017.
- [51] J.-K. Min, A. Doryab, J. Wiese, S. Amini, J. Zimmerman, and J. I. Hong, “Toss’n’turn: smartphone as sleep and sleep quality detector,” in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 477–486, ACM, 2014.
- [52] “UsageStats.” <https://developer.android.com/reference/android/app/usage/UsageStats.html>. Acedido: 20-03-2018.
- [53] S. Landau, *A handbook of statistical analyses using SPSS*. CRC, 2004.
- [54] “Importância de ter um conjunto de dados balanceado.” <http://amsantac.co/blog/en/2016/09/20/balanced-image-classification-r.html>. Acedido: 13-12-2017.

- [55] P. Refaeilzadeh, L. Tang, and H. Liu, “Validação cruzada,” in *Encyclopedia of database systems*, pp. 532–538, Springer, 2009.
- [56] “Precisão e Sensibilidade.” http://scikit-learn.org/stable/auto_examples/model_selection/plot_precision_recall.html. Acedido: 13-12-2017.
- [57] “Área sob a curva ROC.” <http://gim.unmc.edu/dxtests/roc3.htm>. Acedido: 15-12-2017.
- [58] “Classificador Random Forest.” <http://scikit-learn.org/stable/modules/generated/sklearn.ensemble.RandomForestClassifier.html>. Acedido: 13-12-2017.
- [59] “Classificador ZeroR.” <http://chem-eng.utoronto.ca/~datamining/dmc/zeror.htm>. Acedido: 13-12-2017.
- [60] “Som em Decibel para Android.” <https://stackoverflow.com/questions/9597767/decibel-sound-meter-for-android>. Acedido: 27-03-2018.

Anexos

A Manual de Instruções ISABELA

Manual de Instruções

ISABELA

Duarte Raposo

Marcelo Fernandes

Samaritana Silva

Caixa IoT

A caixa ISABELA é composta por um sensor de som, um sensor de temperatura e um sensor de luz. O sistema de aquisição de dados realiza-se através de um sistema Arduino. O Arduino por si só não é capaz de fazer a comunicação com Internet. Deste modo, os dados adquiridos pelo Arduino, são assim enviados para um microcomputador (Raspberry Pi 3). Este microcomputador envia os dados dos sensores para a plataforma FIWARE através da conexão disponível/configurada, cabo ou WiFi.

Requisitos técnicos para o funcionamento e configuração caixa ISABELA

Para o funcionamento da caixa na casa do estudante é necessário:

- Uma conexão sem fios à Internet;
- Tomada de energia elétrica.

Para o processo de configuração é necessário:

- Conhecer o SSID e a password do ponto de acesso sem fios à Internet;
- Cabo de rede;
- Computador pessoal com entrada de rede para Ethernet;

Procedimento para a configuração da caixa ISABELA

- 1) Ligar o carregador de alimentação na caixa ISABELA;
- 2) Ligar o cabo de rede entre as interfaces de rede da caixa e do computador. O cabo deve ser um cabo cruzado (o cabo é fornecido juntamente com a caixa).



- 3) (A caixa usa um endereço IP 192.168.10.1/24, pelo que o computador pessoal deve ter um endereço nesse intervalo, por exemplo 192.168.10.1.2/255.255.255.0). Caso não saiba como definir um IP estático veja no Anexo I os passos a seguir.

- 4) Verificar se os dois equipamentos estão conectados, através do comando ping na linha de comandos:

Linux:

```
pi@device21b7:~ $ sudo ping 192.168.10.1
```

Windows:

```
C:\Users\User_name> ping 192.168.10.1
```

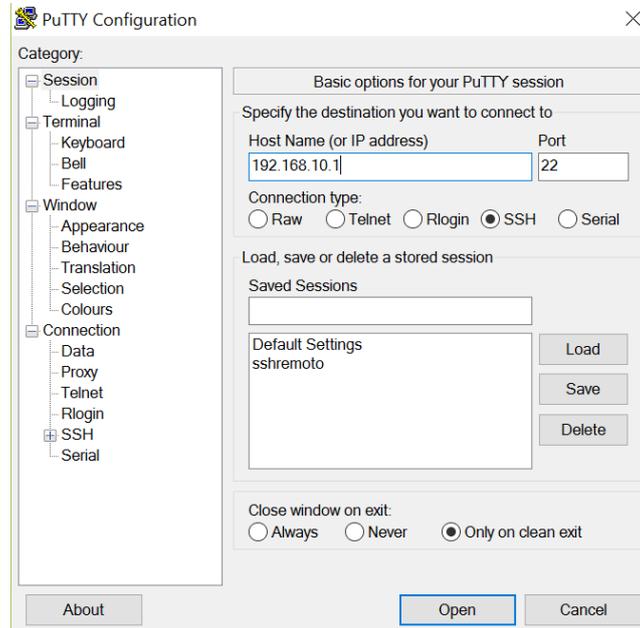
Caso não exista comunicação, irá ver uma mensagem como a seguinte. Uma das possíveis falhas deve-se ao firewall interno do computador estar ativo, este tem que ser desativado temporariamente. Caso desligar não resolva o problema contacte a equipa do projeto.

```
PS C:\Users\Marcelo> ping 192.168.10.1
Pinging 192.168.10.1 with 32 bytes of data:
Request timed out.
Request timed out.
Ping statistics for 192.168.10.1:  Packets: Sent = 2, Received = 0, Lost = 2
(100% loss)
```

Caso exista conexão irá ver, uma mensagem como a seguinte e pode avançar para o passo 5:

```
PS C:\Users\Marcelo> ping 192.168.10.1
Pinging 192.168.10.1 with 32 bytes of data:
Reply from 192.168.10.1 bytes=32 time=2ms TTL=62
Reply from 192.168.10.1: bytes=32 time=2ms TTL=62
Ping statistics for 192.168.10.1:  Packets: Sent = 4, Received = 4, Lost = 0
(0% loss),
```

- 5) Estabelecer uma sessão SSH do computador para o endereço de IP 192.168.10.1, onde se pode usar o programa [PuTTY](#) (basta inserir o IP como mostra a figura e pressionar “OK”).



Após isso ira aparecer no seu ecrã uma Janela de comandos onde deverá inserir os seguintes dados:

- a) Username: pi
 - b) Password: qazplm123
- 6) Configurar a interface WiFi do raspberry Pi, editando o arquivo `wpa_supplicant.conf`, com os dados da rede Wifi de casa. Para tal, utiliza-se o seguinte comando:

```
pi@device42c7:~$ sudo nano /etc/wpa_supplicant/wpa_supplicant.conf
```

Por exemplo: se se tem um SSID: Vodafone-4B5333 e uma Password: TnjT527Y2K, faz-se o seguinte:

```
network={
    ssid="Vodafone-4B5333"
    psk="TnjT527Y2K"
    key_mgmt=WPA-PSK
}
```

- 7) Guardas as mudanças (`ctrl+o`) e sair do editor (`ctrl+x`).
- 8) Por último, reinicia-se o equipamento e o processo está terminado.

```
pi@device42c7:~ $sudo reboot
```

Verificação do funcionamento da caixa ISABELA

Para verificar o correto funcionamento da caixa, tem que se estabelecer novamente uma sessão ssh e proceder como está demonstrado na tabela 1:

Comandos de verificação	Descrição
<pre>pi@device42c7:~ \$sudo ping www.yahoo.com Pinging atsv2-fp.wg1.b.yahoo.com [46.228.47.115] with 32 bytes of data: Reply from 46.228.47.115: bytes=32 time=52ms TTL=51 Reply from 46.228.47.115: bytes=32 time=51ms TTL=51 Reply from 46.228.47.115: bytes=32 time=51ms TTL=51 Reply from 46.228.47.115: bytes=32 time=53ms TTL=51 Ping statistics for 46.228.47.115: Packets: Sent = 4, Received = 4, Lost = 0 (0% loss), Approximate round trip times in milli-seconds: Minimum = 51ms, Maximum = 53ms, Average = 51ms</pre>	Verificação da conectividade a-a Internet. Em caso de não existir resposta, verificar se o serviço de Internet se encontra funcional.
<pre>pi@device42c7:~ \$ i2cdetect -y 1 0 1 2 3 4 5 6 7 8 9 a b c d e f 00: -- 04----- 10: -----</pre>	A saída deve apresentar o número 04, que é o identificador do Arduino para o Raspberry Pi. Se a saída não for 04, tem que se verificar as conexões físicas entre o Raspberry Pi, Arduino e o módulo BSS138.

Comando de verificação do funcionamento da caixa ISABELA

Também se pode observar o LED da caixa que, quando está a funcionar normalmente, está a piscar.

Se não conseguir verificar o correto funcionamento, deve alertar o suporte técnico do Projeto ISABELA.

Aplicação Smartphone

Requisitos

Antes de correr a aplicação tenha em atenção os seguintes pontos:

- É necessária uma conexão estável à Internet para que tudo funcione corretamente.
- A aplicação apenas funciona em smartphones em que a versão Android seja 4.3 ou superior.
- Na versão Android 6.0 ou superior, o utilizador tem que dar as permissões necessárias antes de correr a aplicação. Nas outras versões as permissões são pedidas quando a aplicação está a correr.

Introdução

Para a aplicação começar o utilizador tem que clicar no ícone da aplicação. Vai aparecer um ecrã com uma introdução à aplicação, *Figura 1*, e 2 segundos depois aparece a animação “Made by University of Coimbra”, *Figura 2*.



Figura 1- Ecrã antes da animação

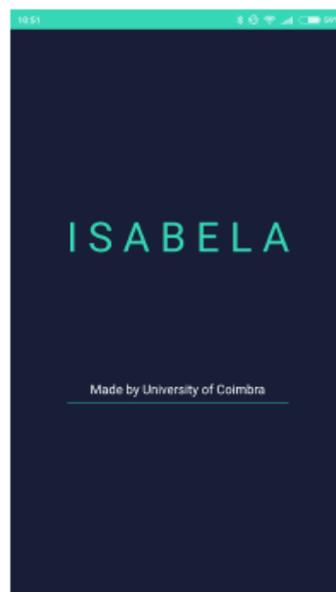


Figura 2- Ecrã depois da animação

Login

Se o utilizador quiser passar diretamente para o ecrã do Login basta clicar duas vezes no ecrã da *Figura 1* ou então esperar 5 segundos.

No ecrã de Login, se o utilizador tiver feito o Login anteriormente pode entrar diretamente na aplicação. Caso contrário ou faz o Login normalmente, *Figura 3* ou se não tiver conta tem que criar uma nova através do Sign Up, *Figura 4*.

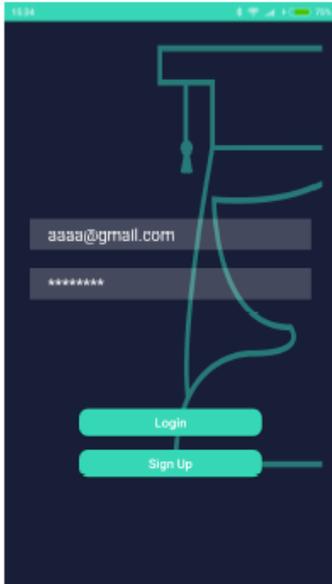


Figura 3- Ecrã para fazer o Login

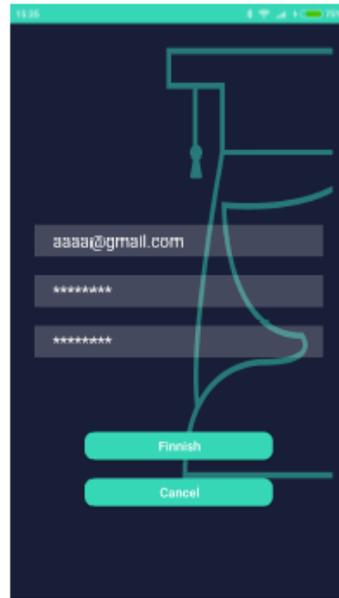


Figura 4- Ecrã para criar uma nova conta

Configurações

Se for a primeira vez que o utilizador usa a aplicação, vai ter que efetuar umas configurações iniciais para que a aplicação funcione. Tem que dizer qual o SSID do WiFi de casa, o ID da caixa IoT que tem em casa, e se quer apenas enviar dados através do WiFi. Na *Figura 5* pode-se ver que é pedida uma quarta configuração, o Departamento, mas esta configuração não é usada por agora.

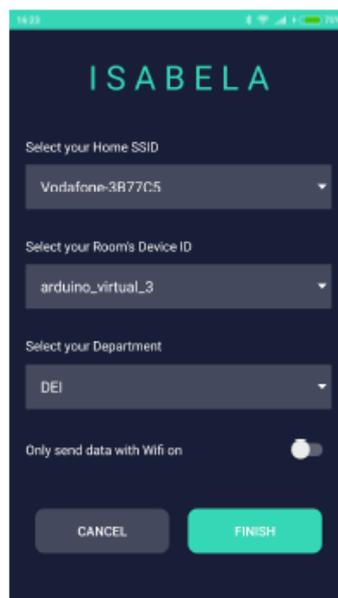


Figura 5- Ecrã com as configurações iniciais

Quando terminar estas configurações, o utilizador pode navegar nas funcionalidades da aplicação.

Ecrãs de Informação

No ecrã principal, existe uma opção de slide horizontal, onde o utilizador pode navegar entre os diferentes ecrãs, deslizando para a direita ou esquerda.

O primeiro ecrã é o resumo da atividade do utilizador, em que aparece a percentagem do tempo que o utilizador gasta em cada atividade. Na *Figura 6* podemos ver que esta percentagem de tempo está representada na forma de um donut. Quando por exemplo, o utilizador passa muito tempo sentado o alarme vai ser ativado, e as cores vão mudar, representado na *Figura 7*.

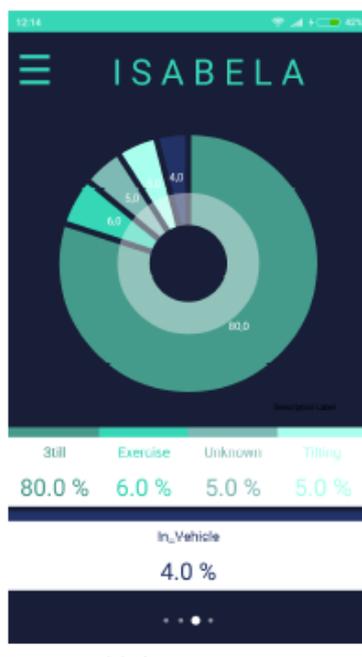


Figura 6- Ecrã da atividade



Figura 7- Ecrã da atividade com alarme

O segundo ecrã diz respeito ao ecrã de Localização, onde se pode ver a distribuição de tempo do utilizador em 3 lugares (Casa, Universidade e Outro). Esta distribuição está representada na forma de uma pirâmide invertida, que faz lembrar os pins do google, *Figura 8*.



Figura 8- Ecrã da Localização

O terceiro ecrã representa a Sociabilidade, o valor desta é dado como um score e é representado um uma barra de progresso que vai dos 0% aos 100%. É utilizado o número total de SMS, o tempo de chamadas, o número de chamadas efetuadas e as pessoas contactadas para obter este score, como mostra a *Figura 9*.

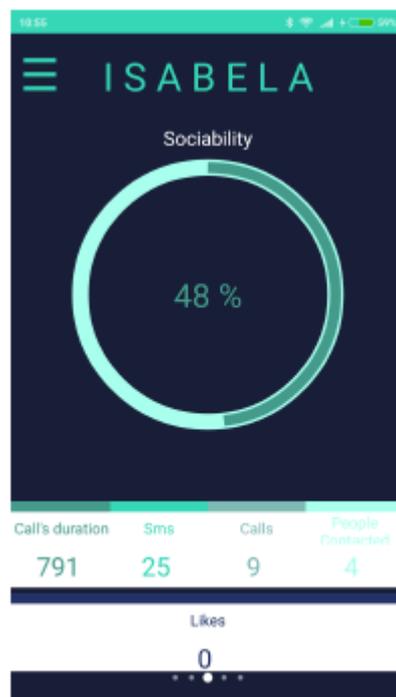


Figura 9- Ecrã da Sociabilidade

Existe também um botão de navegação, que se pode aceder no canto superior esquerdo do ecrã, *Figura 10*. Neste botão existe a opção de mudar as configurações iniciais caso seja necessário, *Figura 11*.

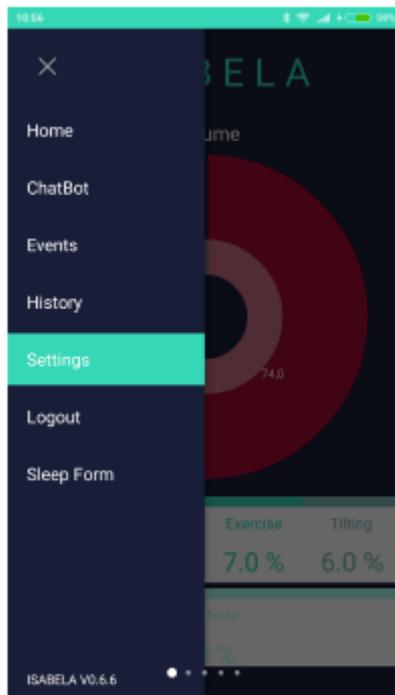


Figura 10- Opções de navegação

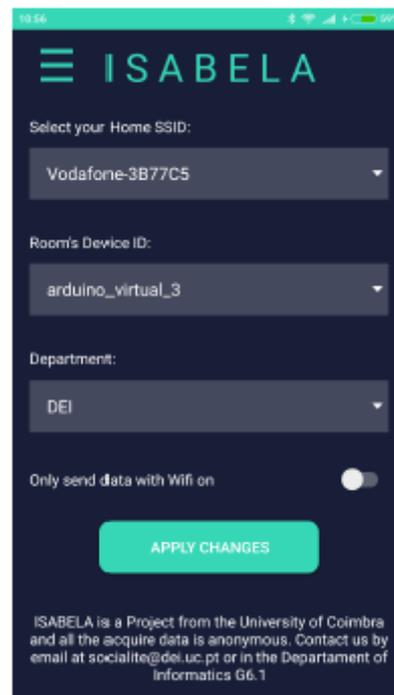


Figura 11- Ecrã para alterar as configurações iniciais

A aplicação tem também o ecrã ChatBot. Este ecrã é uma maneira de interagir com os estudantes e dizer lhes aquilo que está mal. Na *Figura 12* pode-se ver as funcionalidades se uma Small talk, onde o utilizador pode colocar questões. Na *Figura 13* pode-se ver uma conversa onde o utilizador pergunta os valores dos sensores da caixa IoT.

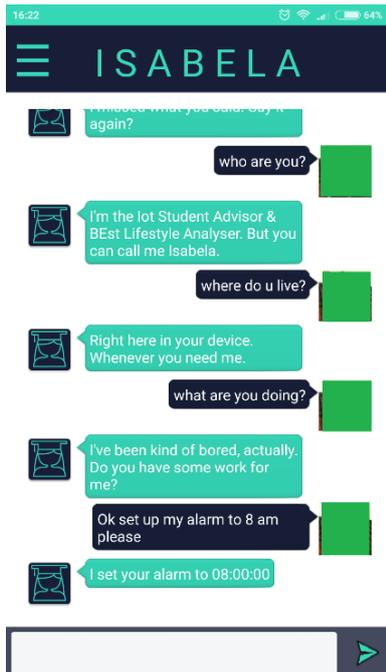


Figura 12- Exemplo de conversa

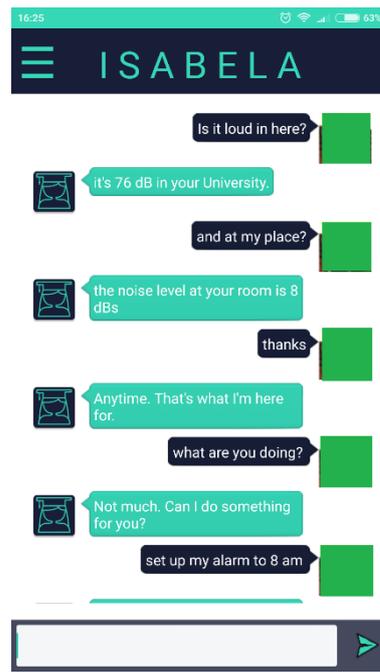


Figura 13- Exemplo de conversa

Smartwatch

Como o ecrã do smartwatch é muito mais pequeno, apenas aparece o nome da aplicação e o valor dos 3 sensores em tempo real (batimento cardíaco, contagem de passos e a atividade), isto está representado na *Figura 14*. A *Figura 15* representa visualização dos valores do smartwatch na aplicação Android.



Figura 14- Ecrã do smartwatch

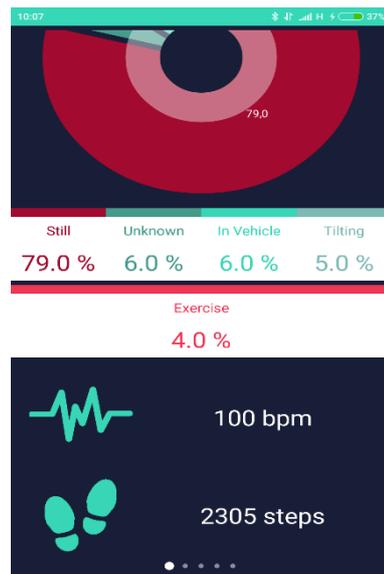


Figura 15- Valores do smartwatch na aplicação

Nota: caso existam dúvidas ou problemas contacte os e-mails presentes na capa do documento.

ANEXO I

Para definir um IP estático deve abrir a sua Janela do Centro de Rede e Partilha (**Painel de Controlo\Rede e Internet\Centro de Rede e Partilha**), como mostra na figura 2. Deve seleccionar a rede Ethernet, o que ira abrir a janela da figura 3.

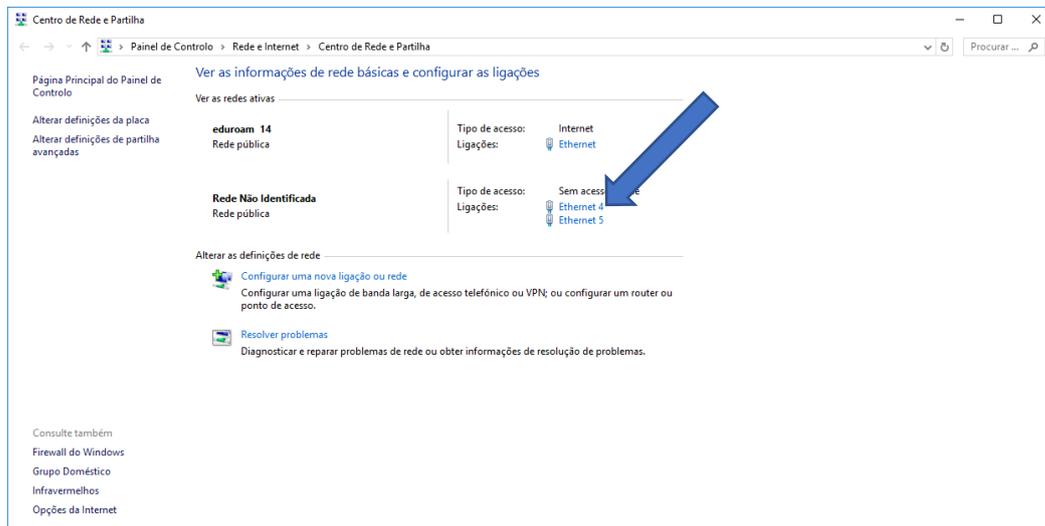


Figura 2 Centro de rede e partilha.

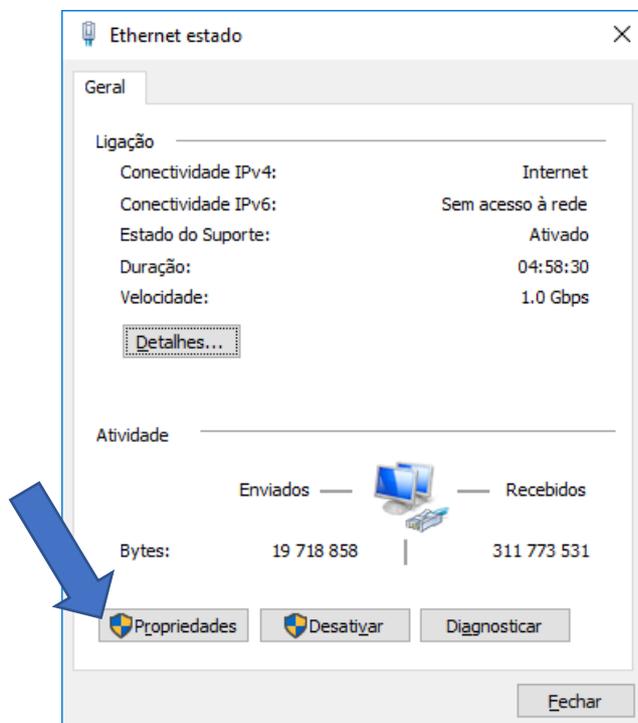
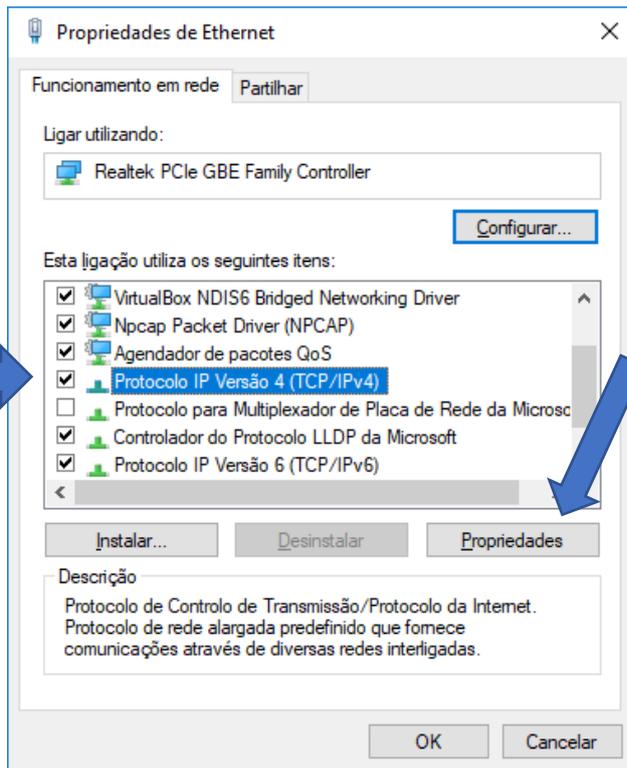
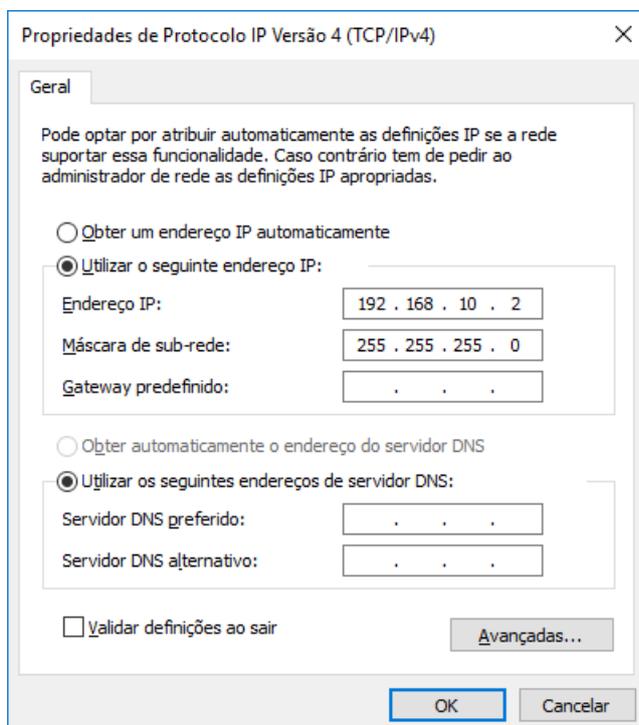


Figura 3 Propriedades da rede

Como indica na figura 3 deve seleccionar a opção das propriedades da rede o que ira abrir o Menu da figura 4. Deve ativar ou seleccionar o Protocolo IPV4 como indica a figura 4, e seleccionar o botão *Propriedades*, o que ira abrir o menu da figura 5.



Tal como indica a figura 5 deve seleccionar “Utilizar o seguinte endereço IP:”, e inserir os dados como mostrado na figura e pressionar “OK”.



Nota: Após fazer os passos necessários deve reverter estes passos, isto é, voltar a escolher a opção para obter um IP automaticamente. E caso tivesse o protocolo IPV4 desligado pode voltar a desligá-lo.

B Manual final de instruções do ISABELA

IoT Student Advisor & BEst Lifestyle Analyser

Conteúdo

1	ISABELA	3
2	Termos e Condições gerais	3
3	Requisitos	3
4	Aplicação Móvel	3
4.1	Introdução	3
4.2	Login	4
4.3	Configurações	4
4.4	Ecrãs de Informação	5
5	Caixa IoT	10
	Referências	10

1 ISABELA

ISABELA é um sistema Human-in-the-loop Cyber-Physical Systems (HITLCPS) que é constituído por uma aplicação Android, uma caixa IoT e um servidor baseado na tecnologia FIWARE. O sistema ISABELA permite obter vários parâmetros físicos e emocionais dos estudantes, e parâmetros físicos do ambiente, de forma a relacionar hábitos e estilos de vida com o desempenho escolar. O principal objetivo desta aplicação é melhorar os resultados escolares dos alunos e, como tal, destina-se a ser usado por estudantes universitários. Nesta aplicação recolhe-se a atividade dos alunos, o sono, a sociabilidade e a localização para inferir o comportamento do aluno.

O ISABELA tem por base o estudo StudentLife da Universidade de Dartmouth [1]. No entanto, o ISABELA e o StudentLife diferem em vários pontos: o ISABELA é um sistema HITLCPS, ou seja, permite dar feedback ao aluno e evitar o fraco desempenho académico, enquanto o estudo da Universidade de Dartmouth foi principalmente estatístico. Existem também outras diferenças, como o uso de uma arquitetura IoT, obtendo dados de mais fontes do que apenas do telemóvel, o uso do ChatBot como meio de interação com os alunos e o uso de sensores virtuais/sociais.

2 Termos e Condições gerais

Os dados recolhidos durante a utilização da aplicação ISABELA serão utilizados para efeitos científicos. Este estudo tem como objectivo avaliar a performance dos estudantes durante diferentes períodos lectivos, especificamente, durante períodos de aulas e períodos de avaliação.

Para protecção dos dados dos utilizadores, técnicas de anonimização serão utilizadas com vista à protecção de informações sensíveis. Estas técnicas visam aplicar a lei de protecção de dados (Regulamento Geral de Protecção de dados (UE) 2016/679*). Para mais informação sobre riscos e outros temas relacionados com a lei de protecção de dados, por favor consulte a declaração de consentimento.

3 Requisitos

Antes de correr a aplicação tenha em atenção os seguintes pontos:

- Para a fase de envio dos dados é necessária uma conexão estável à Internet.
- A aplicação apenas funciona em telemóveis cuja versão *Android* seja 4.3 ou superior.
- Na versão *Android* 6.0 ou superior, o utilizador tem que dar as permissões necessárias antes de correr a aplicação. Nas outras versões as permissões são pedidas quando a aplicação está a executar.

4 Aplicação Móvel

4.1 Introdução

Para a aplicação ser iniciada o utilizador tem que clicar no ícone da aplicação no telemóvel. Irá aparecer um ecrã com uma introdução à aplicação, figura 1, e 2 segundos depois aparece a animação “Made by University of Coimbra”, figura 2.



Figura 1: Ecrã antes da animação

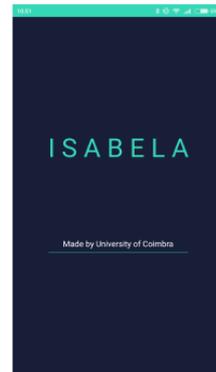


Figura 2: Ecrã depois da animação

4.2 Login

Se o utilizador quiser passar diretamente para o ecrã do Login basta clicar duas vezes no ecrã da Figura 1 ou então esperar 5 segundos. É necessário fazer o Login com o Facebook para ter acesso ao conteúdo da aplicação. O ID do Facebook é encriptado e deste modo os dados são todos anonimizados e a **privacidade** é assegurada.



Figura 3: Login

4.3 Configurações

Quando se utiliza a aplicação pela primeira vez é necessário proceder a algumas configurações iniciais para que a aplicação funcione. Nomeadamente, é necessário informar qual o SSID da rede WiFi de casa, se quer apenas enviar dados através do WiFi e qual o departamento em que estuda.

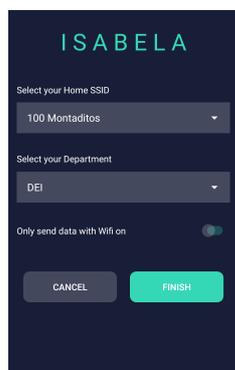


Figura 4: Configurações iniciais

Quando terminar estas configurações, pode navegar nas funcionalidades da aplicação.

4.4 Ecrãs de Informação

No ecrã principal existe uma opção de slide horizontal, onde o utilizador pode navegar entre os diferentes ecrãs, deslizando para a direita ou esquerda. O primeiro ecrã é o resumo da atividade do utilizador, em que aparece a percentagem do tempo que o utilizador gasta em cada atividade. Na figura 5 é possível observar a percentagem de tempo na forma de um gráfico circular. Quando por exemplo, o utilizador passa muito tempo sentado o alarme vai ser ativado, e as cores vão mudar, como é representado na figura 6.

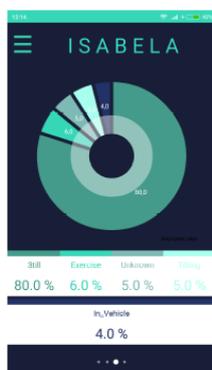


Figura 5: Atividade



Figura 6: Atividade com alarme

O segundo ecrã diz respeito à Localização, onde se pode ver a distribuição do tempo do utilizador em 3 lugares (Casa, Universidade e Outro). Esta distribuição está representada na forma de uma pirâmide invertida, que faz lembrar os pins do Google Maps, figura 7.

O terceiro ecrã representa a Sociabilidade, figura 8 . O valor desta é representado num círculo de progresso que vai dos 0% aos 100%. É utilizado o número total de SMS enviados, o tempo de chamadas, o número de chamadas efetuadas e as pessoas contactadas para obter este valor. Apresenta também a média das emoções expressas nas publicações feitas no Facebook e no Twitter dos últimos cinco dias. Quando os conteúdos têm subjacentes palavras positivas aparece uma cara com um sorriso feliz,



Figura 7: Localização

quando o conteúdo é negativo aparece uma cara com um sorriso triste e quando a média do conteúdo é neutro aparece uma cara com um sorriso neutro.

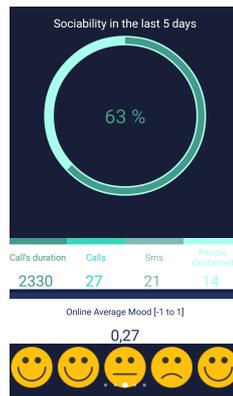


Figura 8: Sociabilidade

O quarto ecrã dá informação sobre a atividade, o score, a localização, o valor da luz, o valor do som e o tempo, como é possível observar na figura 9 e na figura 10. Apenas é registado o valor do som e nenhuma conversa é registada, como mostra a figura 10.

Existe também um botão de navegação, que se pode aceder no canto superior esquerdo do ecrã, figura 11. Neste botão existe a opção de mudar as configurações iniciais caso seja necessário, figura 12.



Figura 9: Informação do estudante



Figura 10: Informação do estudante



Figura 11: Opções de navegação

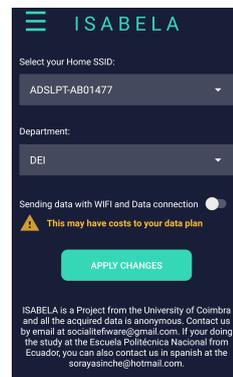


Figura 12: Ecrã para alterar as configurações iniciais

Este botão permite também aceder à LeaderBoard, figura 13, que faz um ranking por ordem decrescente dos utilizadores que estão a participar no estudo e a enviar mais dados. A cor azul clara representa o próprio utilizador e a cor amarela o utilizador que está em primeiro. É possível ainda aceder ao botão Friends, figura 14, que mostra quanto tempo estamos sozinhos e com que utilizadores mais convivemos. Para saber os amigos com quem mais convivemos é utilizado o *Bluetooth*, mas apenas vê os que também têm a aplicação instalada.



Figura 13: LeaderBoard

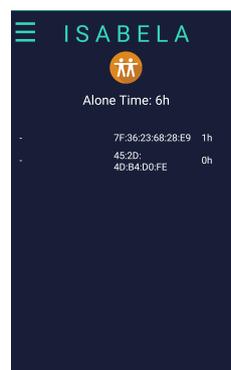


Figura 14: Friends

Este botão de navegação permite também aceder ao Sleep Form, figura 15 e figura 16. O Sleep Form é um questionário diário, que os utilizadores têm que preencher. Os utilizadores têm que informar a que horas foram para a cama e a que horas acordaram, se a noite foi tranquila ou não, como se sentem, quão sociais se sentem neste dia e quanto é que estudaram no dia anterior.

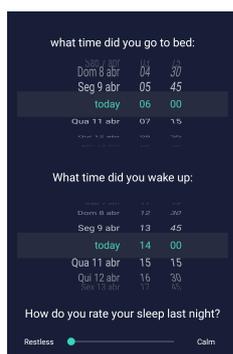


Figura 15: Sleep Form

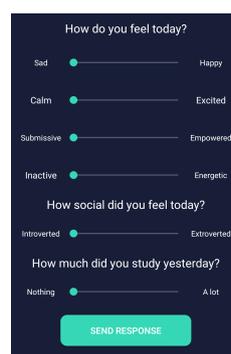


Figura 16: Sleep Form

Se no botão de navegação for escolhida a opção History é apresentado um ecrã com 4 opções *Activity*, *Location*, *Sociability* e *Sleep*. Em cada opção é mostrado um gráfico dos últimos 5 dias com a evolução da actividade, da localização, da sociabilidade ou do sono. Na figura 17 está representado um gráfico que apresenta o número de horas dormidas nos últimos 5 dias.

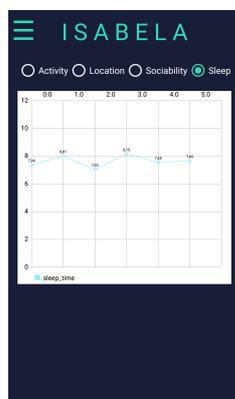


Figura 17: Número de horas dormidas nos últimos 5 dias

A aplicação apresenta também o ecrã ChatBot. Este ecrã permite interagir com os estudantes e dizer-lhes aquilo que, eventualmente, está mal. Na figura 18 pode-se ver as funcionalidades de uma Small talk, onde o utilizador pode colocar questões. Na figura 19 pode-se ainda observar uma conversa onde o utilizador pergunta os valores dos sensores da caixa IoT.



Figura 18: Exemplo de conversa

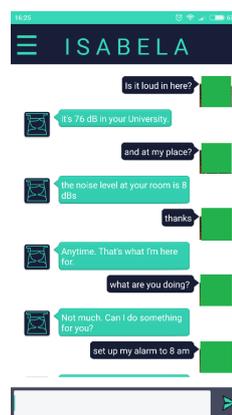


Figura 19: Exemplo de conversa

O botão de navegação permite ainda aceder a um último ecrã que apresenta a média diária do uso das aplicações instaladas no telemóvel, o que pode ser visto na figura 20.



Figura 20: Uso diário das aplicações

5 Caixa IoT

No bar do Departamento de Engenharia Informática vai ser colocada uma caixa Internet das Coisas (Internet of Things - IoT) composta por um sensor de som, um sensor de temperatura e um sensor de luz, baseados em Arduino e RaspberryPi. O sensor do som apenas deteta o nível de ruído, pois nenhuma conversa é registada.

Referências

- [1] R. Wang, F. Chen, Z. Chen, T. Li, G. Harari, S. Tignor, X. Zhou, D. Ben-Zeev, and A. T. Campbell, "StudentLife," in *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing - UbiComp '14 Adjunct, 2014*.

C Declaração de Consentimento

ISABELA - IoT Student Advisor & BEst Lifestyle Analyser

INTRODUÇÃO:

Está a ser convidado para participar numa pesquisa para analisar o impacto diário da carga de trabalho no stress, sono, atividade, humor, sociabilidade, bem-estar mental e desempenho académico de uma turma. O nosso sistema estudará a correlação entre os dados de deteção automática e de várias medidas físicas e de bem-estar.

Antes de decidir participar, é importante que entenda o objetivo deste estudo e as suas implicações. Por favor, reserve algum tempo para ler atentamente as informações a seguir. Se algum ponto parecer impreciso ou se alguma informação adicional for necessária, não hesite em perguntar. Então, tome o tempo necessário para decidir se deseja participar ou não.

OBJECTIVO DO ESTUDO:

O objetivo deste estudo é avaliar o desempenho académico do estudante utilizando dispositivos de Internet das Coisas e telemóveis para medir os parâmetros do ambiente e do nível de ansiedade dos alunos antes de eles terem uma avaliação.

POTENCIAIS RISCOS: Nenhum.

POTENCIAS BENEFÍCIOS: Desconhecido.

COMPENSAÇÃO POR DOENÇA OU LESÃO:

Não é esperado que a participação neste estudo cause alguma doença ou lesão.

USO DE DADOS PESSOAIS/CONFIDENCIALIDADE:

Ao participar neste estudo, os dados obtidos serão utilizados e processados de forma anónima.

Os seus dados pessoais são guardados para fins de pesquisa científica para obter informações sobre os dispositivos e seu desempenho durante e após este estudo. Podem ser usados para pesquisas científicas adicionais, propósitos educacionais e publicações, bem como para estudos futuros ou para obter avaliações atuais ou futuras para aprovações para os dispositivos.

As suas informações pessoais confidenciais serão anonimizadas e codificadas por chave. Em todos os casos, as suas informações pessoais serão tratadas em todos os momentos, de acordo com os padrões de confidencialidade apropriados e com todas as leis de privacidade e proteção de dados aplicáveis (Lei de Proteção dos Dados 2016/679**).

Tem o direito de aceder às informações pessoais guardadas sobre você e de corrigir alguma informação que esteja errada.

*** Lei 2016/679, sobre Proteção dos Dados Pessoais; os dados pessoais necessários (por exemplo, idade, sexo, dados de saúde) são apenas os necessários para atingir os objetivos deste estudo. O seu nome não aparecerá em nenhum dos relatórios do estudo e sua identidade não será revelada a ninguém, exceto para cumprir os propósitos do estudo e no caso de emergências médicas ou requisitos legais. Qualquer informação de natureza pessoal que possa ser identificável será armazenada e processada eletronicamente com medidas de segurança apropriadas. O acesso a essas informações será restrito ao pessoal designado para esse fim ou a outro pessoal autorizado que será obrigado a manter a confidencialidade dessas informações. Os resultados do estudo podem ser comunicados às autoridades de saúde e possivelmente à comunidade científica por meio de convenções e / ou publicações. De acordo com a legislação em vigor, você tem o direito de aceder, corrigir ou remover seus registros de dados pessoais.*

PARTICIPAÇÃO VOLUNTÁRIA:

A sua participação neste estudo é inteiramente voluntária. Está livre de recusar participar e pode interromper a sua participação a qualquer momento, sem medo de penalidade.

O responsável pode decidir terminar a sua participação no estudo a qualquer momento sem o seu consentimento prévio. Se isso acontecer, você será notificado e as razões serão explicadas.

QUESTÕES:

Se você tiver mais dúvidas ou quiser mais informações ou explicações sobre este estudo e os direitos dos participantes do mesmo, entre em contato:

NOME DO INVESTIGADOR: Professor Jorge Sá Silva

TELEFONE DO INVESTIGADOR: +351 239790000

ASSINATURA:

- Eu li e compreendi a informação deste estudo.
- Eu tive tempo suficiente para considerar a minha participação neste estudo e estou consciente que a participação neste estudo é completamente voluntária.
- Eu percebi que posso decidir interromper a participação a qualquer momento.
- Eu percebi e concordo que a informação pessoal será usada e processada (manualmente e por computador) pelo investigador ou por qualquer outra parte designada que esteja envolvida neste estudo (por exemplo, hospital, médico, autoridades reguladoras, comités de ética).
- Eu entendo que tenho o direito de aceder às informações pessoais obtidas sobre mim e de corrigir qualquer informação que esteja errada.

- Eu recebi uma cópia desta informação e por este meio concordo participar voluntariamente no lot Student Advisor & BEst Lifestyle Analyser.

Assinatura do Estudante

Data

Para a pessoa que obtém o consentimento informado:

Certifico que expliquei o conteúdo deste Termo de Consentimento Informado e informei o indivíduo acima mencionado antes de realizar quaisquer procedimentos relacionados ao estudo. Eu perguntei se havia alguma pergunta e respondi a todas as perguntas que foram levantadas.

Investigador Principal ou outra pessoa autorizada a obter o consentimento informado:

Nome

Assinatura

Data