

Universidade do Minho  
Mestrado em Engenharia Informática

## Sensorização e Ambiente

# Impacto da utilização do telemóvel na qualidade de sono

## Grupo 8

Carolina Santejo  
pg47102

Filipa Pereira  
pg46978

Luís Pinto  
pg47428

Raquel Costa  
pg47600

Braga  
10 de Maio de 2022

# Conteúdo

<b>1</b>	<b>Introdução</b>	<b>5</b>
<b>2</b>	<b>Contextualização</b>	<b>5</b>
<b>3</b>	<b>Principais Objetivos</b>	<b>5</b>
<b>4</b>	<b>Arquitetura e Metodologia</b>	<b>6</b>
4.1	Arquitetura do sistema . . . . .	6
4.2	Metodologia implementada . . . . .	7
<b>5</b>	<b>Recolha de dados</b>	<b>8</b>
5.1	Monitorização do Sono . . . . .	8
5.1.1	Ambiente . . . . .	9
5.1.2	Sensores . . . . .	9
5.1.3	Colecionador . . . . .	9
5.2	Monitorização da Atividade no Telemóvel . . . . .	10
5.2.1	Ambiente . . . . .	10
5.2.2	Sensores . . . . .	10
5.2.3	Colecionador . . . . .	11
<b>6</b>	<b>Tratamento de dados</b>	<b>11</b>
6.1	Pré-Processamento . . . . .	11
6.2	Tratamento . . . . .	13
<b>7</b>	<b>Descrição e Exploração dos dados colecionados</b>	<b>14</b>
7.1	Atributos do dataset tratado . . . . .	14
7.2	Visualização de dados do dataset . . . . .	15
<b>8</b>	<b>Previsão com Modelos de Machine Learning</b>	<b>19</b>
<b>9</b>	<b>Aplicação Android: <i>X-Leap</i></b>	<b>20</b>
<b>10</b>	<b>Análise dos Resultados Obtidos</b>	<b>27</b>
<b>11</b>	<b>Conclusão e Trabalhos Futuros</b>	<b>29</b>

## Lista de Figuras

1	Arquitetura do sistema desenvolvido . . . . .	6
2	Metodologia adotada pelo grupo . . . . .	7
3	As 4 fases do sono . . . . .	9
4	Matriz de correlação . . . . .	15
5	Impacto do Activity_Time_2H no Deep_Sleep (por utilizador) . . . . .	16
6	<i>Pairplot</i> de alguns atributos do dataset . . . . .	17
7	Distribuição da Sleep_quality . . . . .	18
8	<i>Boxplot</i> do atributo Activity_Time . . . . .	18
9	Ícone da aplicação X-Leap . . . . .	20
10	Interface da aplicação X-Leap . . . . .	21
11	Opções de gráficos da aplicação X-Leap . . . . .	22
12	Gráfico relativo ao <i>Sleep Cycle</i> . . . . .	23
13	Gráfico relativo ao tempo de atividade . . . . .	24
14	Gráfico relativo à variação da qualidade do sono . . . . .	25
15	Gráfico de barras relativo às avaliações da qualidade do sono . . . . .	26

**Lista de Tabelas**

1	Tipo e Quantidade de dados recolhidos . . . . .	8
2	Representação das fases do sono . . . . .	10
3	Descrição dos atributos do Dataset tratado . . . . .	14

# 1 Introdução

O seguinte projeto foi desenvolvido no âmbito da Unidade Curricular de Sensorização e Ambiente, onde a principal motivação foi implementar um sistema inteligente capaz de obter dados e gerar informação útil de acordo com o contexto do ambiente onde está inserido. Para tal, é necessário recorrer à integração de sensores físicos ou virtuais, possibilitando, assim, a monitorização do ambiente que o rodeia. Posto isto, o sistema desenvolvido deverá focar-se em domínios relacionados com *Smart Cities* e *Internet of Things*.

Evidentemente, existem inúmeros problemas nos grandes centros urbanos, muitos relacionados com a gestão populacional, segurança rodoviária ou até mesmo com a saúde pública, tanto física como mental. Uma *Smart City* deve ser capaz de responder a este leque de problemas, recorrendo a várias técnicas. Uma delas trata-se do uso de sensores para coleção de dados, de modo a ser possível prever fenómenos futuros.

Assim sendo, o problema escolhido para o projeto centra-se no estudo do "impacto da utilização do telemóvel na qualidade de sono". Para tal, é crucial explorar os dados recolhidos, de forma a ser possível prever a qualidade de sono de um indivíduo de acordo com o seu grau de utilização do telemóvel. A principal meta do grupo é investigar como é que é possível melhorar a qualidade do sono da população e, inerentemente, a sua qualidade de vida e saúde.

## 2 Contextualização

É inegável que, na atualidade, o telemóvel é um dispositivo indispensável para a grande maioria das pessoas. Este é utilizado não só como forma de comunicação imediata, mas também como um valioso instrumento de trabalho ou entretenimento. No entanto, várias questões têm sido levantadas, relativamente ao impacto do uso de aparelhos eletrónicos no bem-estar dos indivíduos, sendo uma das mais pertinentes e faladas, a forma como a utilização do telemóvel poderá afetar a qualidade do sono. Vários estudos realizados pela organização australiana *Sleep Health Foundation* revelam que o uso de dispositivos digitais algumas horas antes de dormir é suficiente para afetar o sono dos indivíduos uma vez que reduz a produção da hormona melatonina, provocando a estimulação cerebral e, consequentemente, maior dificuldade em adormecer.

Tendo isto em conta, neste trabalho prático pretende-se averiguar como é que a qualidade do sono dos utilizadores, sendo estes elementos do grupo, varia de acordo com a utilização do seu telemóvel, tanto num período de duas horas antes de adormecer, como também ao longo do dia.

## 3 Principais Objetivos

Com a realização deste projeto pressupõe-se a aquisição de novas valências na vertente de Sensorização Ambiente em *Smart Cities*. Assim, pretende-se através de um cariz mais prático consolidar os conteúdos teóricos abordados na unidade curricular de

Sensorização e Ambiente.

Mais concretamente, os principais objetivos deste trabalho passam primeiro pela utilização de sensores para a recolha e coleção de dados, onde para este efeito serão usados tanto o *smartwatch* como o *smartphone*. De seguida, pretende-se através da aplicação de técnicas de fusão sensorial combinar os dados das diferentes fontes, de forma a que as informações resultantes possuam menos incerteza e/ou se tornem mais explícitas do que aquilo que seria possível caso essas fontes fossem usadas individualmente. Por último, é também requerida a utilização de modelos de *Machine Learning*, de forma a permitir prever fenómenos futuros. Neste caso, o foco do estudo do modelo será na previsão da qualidade do sono de um utilizador, posteriormente este tópico será abordado em maior detalhe.

De modo a cumprir com a concretização dos objetivos idealizados, podemos considerar uma abordagem em duas fases. Primeiramente procedendo-se à recolha de dados e, de seguida, numa segunda fase ocorrendo o tratamento e visualização dos mesmos, bem como a construção dos diferentes modelos de *machine learning*.

## 4 Arquitetura e Metodologia

### 4.1 Arquitetura do sistema

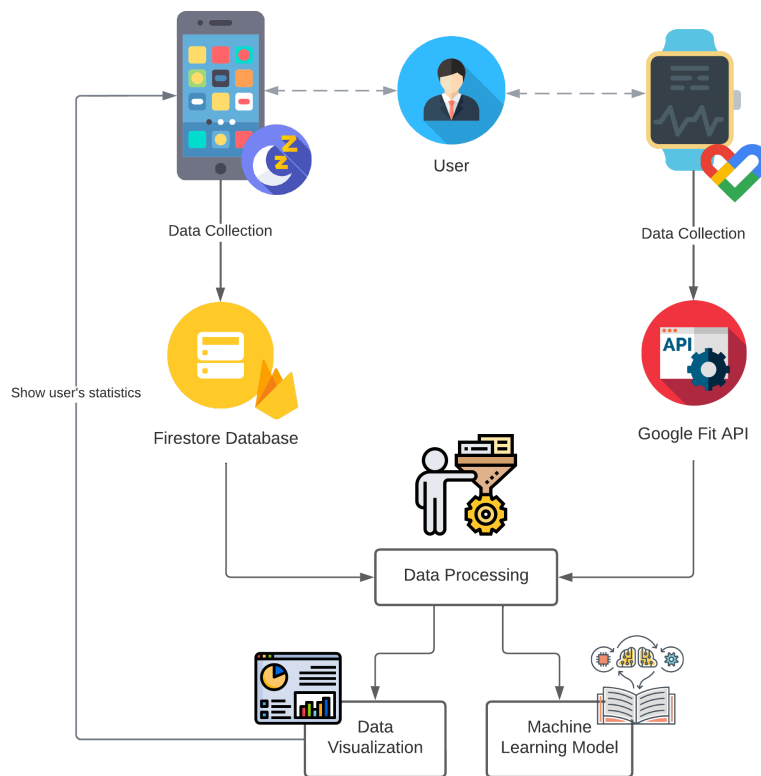


Figura 1: Arquitetura do sistema desenvolvido

Na Figura 1 encontra-se descrita a arquitetura do sistema desenvolvido pelo grupo, onde são referidas as aplicações usadas, os métodos de armazenamento usados, e os vários processos que integram este projeto. É de realçar que este trabalho encontra-se dividido em duas fases, a primeira que inclui a conceção de colecionadores de dados e, conseqüentemente, a recolha dos mesmos. Já numa segunda fase, é esperado o tratamento dos dados obtidos, a visualização e exploração dos mesmos, e, além disso, a implementação de um modelo de *Machine Learning*.

Em primeiro lugar, um utilizador que seja monitorizado por este sistema deve possuir um *smartphone* e um *smartwatch*, ou então, como alternativa a esta última condição, uma aplicação já disponível que seja capaz de supervisionar o sono e tenha a capacidade de sincronizar com o "Google Fit". Posto isto, o utilizador deve instalar no seu telemóvel a aplicação "X-Leap" desenvolvida pelo grupo, e sincronizar o seu *smartwatch* com a aplicação "Google Fit" já disponibilizada pela Google. Após a realização destes primeiros passos, o sistema já estará a recolher os dados pessoais tanto sobre o sono como sobre o grau de atividade do utilizador no telemóvel. Estes dados colecionados são armazenados no *Firebase* e na *API do Google Fit*.

Relativamente à segunda fase, é feito, primeiramente, um tratamento dos dados que foram recolhidos na fase anterior, uma vez que estes se encontram ainda "em bruto". Assim que os dados já se encontrarem num formato apropriado, é desenvolvido um modelo de *Machine Learning* capaz de prever fenómenos futuros. Além disso, este novo formato de dados é enviado de volta para a aplicação "X-Leap", de forma a que o utilizador consiga aceder a estatísticas e visualizar informações sobre o seu sono e atividade no telemóvel.

## 4.2 Metodologia implementada

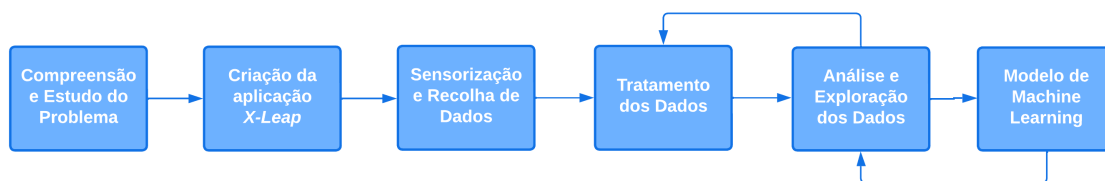


Figura 2: Metodologia adotada pelo grupo

Ao longo do projeto, o grupo seguiu a metodologia representada na Figura 2. Numa primeira etapa o grupo começou por compreender e estudar o problema em questão, de forma a entender quais os dados que deveriam ser captados, e como é que seria esse procedimento. De seguida o grupo inicializou o desenvolvimento da aplicação "X-Leap". Esta aplicação tem como objetivo monitorizar o grau de atividade de um utilizador no telemóvel e fornecer-lhe informações relacionadas com o seu sono, como por exemplo a percentagem de sono profundo e leve que teve na última noite. Além disso, o "X-Leap" analisa dados tal como o tempo que uma pessoa passa no telemóvel ou então o brilho do dispositivo e a luminosidade do ambiente. Este tipo de dados

permitem ao grupo inferir como é que o brilho poderá afetar o cansaço dos olhos e, consequentemente, a qualidade do sono.

Depois da aplicação acabada, o grupo começou a fase de sensorização, onde foram escolhidas três pessoas para usar a aplicação "X-Leap", apenas disponível para android, e o "Google Fit" para capturar os dados do *smartwatch*. Quando o grupo percebeu que já tinha dados suficientes para começar a analisar e explorar, foi começada então a etapa seguinte.

Para ser possível analisar os dados, foi preciso primeiro recorrer ao tratamento dos mesmos, uma vez que estes são dados ainda em "bruto". No final, obteu-se um formato já "limpo" e adequado para começar a explorar os dados obtidos. Depois de visualizar os dados, e explorar os mesmos, o grupo desenvolveu um modelo de *Machine Learning* capaz de prever a qualidade de sono de uma pessoa com base em diferentes atributos.

## 5 Recolha de dados

A etapa correspondente à recolha de dados está dividida em duas partes. Uma correspondente à monitorização do sono e outra à monitorização da atividade no telemóvel.

Antes demais, é necessário entender quais as condições que o grupo criou para o funcionamento desta etapa. Em primeiro lugar, foram monitorizadas três pessoas desde o dia 24 de abril até dia 9 de maio. Todos os dias estas pessoas tiveram de manter em execução a aplicação *X-Leap*, além de dormir também com um smartwatch no seu pulso, de forma a ser possível monitorizar o seu sono.

Assim sendo, a monitorização destes ambientes permite ao grupo através dos dados captados não só inferir informação útil como também prever fenómenos futuros. Na tabela seguinte é possível observar a quantidade de dados que foram recolhidos no intervalo de tempo estabelecido, tais como os dados relacionados com a atividade do utilizador no telemóvel, o *rating* que este atribuiu à qualidade do seu sono e ainda os dados referentes ao próprio sono dele.

ID	Atividade	<i>Ratings</i>	Sono
1	1110	14	207
2	1510	14	596
3	571	14	717

Tabela 1: Tipo e Quantidade de dados recolhidos

### 5.1 Monitorização do Sono

A monitorização do sono permite que o grupo consiga estudar os vários ciclos de sono de uma pessoa, estes incluem diversas fases, entre elas, a do sono leve, sono profundo, e REM. Além disso, é possível também analisar o número de horas que a pessoa dormiu, tratando-se este de um fator essencial para medir a qualidade do sono de alguém.



### 5.1.1 Ambiente

Primeiramente, foi necessário definir um ambiente no qual fosse possível monitorizar o sono de uma pessoa. Portanto, decidiu-se medir o sono dos participantes do estudo, somente durante a noite, de forma a tentar perceber unicamente como é que o sono é afetado pelo uso do telemóvel antes de dormir. Além disso, o grupo tentou escolher condições que permitissem a menor interferência possível de fatores externos que pudessem influenciar a qualidade do sono, daí a monitorização ser feita apenas durante a noite para todos os participantes, e todos eles terem rotinas diárias semelhantes. É de notar que os três participantes são estudantes universitários não-trabalhadores.

### 5.1.2 Sensores

Para concretizar esta monitorização recorreu-se ao uso de *smartwatches* e pulseiras de *fitness*. Através dos vários sensores presentes neste tipo de dispositivos, é possível analisar o tempo e a qualidade de sono. Alguns exemplos de sensores que são utilizados para este fim são: o **Acelerómetro**, uma vez que consegue captar o movimento da pessoa durante o seu sono; E o **Monitor de frequência cardíaca e respiração**, que permite estimar qual a fase de sono em que a pessoa se encontra, por exemplo se é a de "sono leve", "sono profundo" ou "REM", sendo esta última aquela onde há maior atividade cerebral, isto é, onde ocorrem os sonhos vívidos [Fig. 3].



Figura 3: As 4 fases do sono

### 5.1.3 Colecionador

De forma a armazenar os dados recolhidos através do *smartwatch* ou pulseira, foi utilizada a aplicação "Google Fit", que permite sincronizar a atividade do dispositivo com uma conta Google. Deste modo, foi possível guardar automaticamente todos os dados recolhidos do smartwatch numa API, facilitando assim o processo de acesso aos dados.

Sempre que é feito um pedido à API, esta devolve um ficheiro JSON, que contém os dados do sono da pessoa em questão desde o primeiro dia que foi feita a sincronização entre a sua conta Google e o seu smartwatch. Este ficheiro contém, por dia, os intervalos de tempo correspondentes a cada fase do sono, sendo que cada uma delas é representada por um inteiro, tal como está descrito na tabela 2.

Sleep Stage Type	Value
Awake (during sleep cycle)	1
Sleep	2
Out-of-bed	3
Light Sleep	4
Deep Sleep	5
REM	6

Tabela 2: Representação das fases do sono

## 5.2 Monitorização da Atividade no Telemóvel

A segunda fase da recolha de dados do projeto correspondeu à monitorização da atividade no telemóvel. Desta forma, o grupo desenvolveu uma aplicação *Android* chamada ***X-Leap***, na qual foram captados todos os dados relevantes utilizando os sensores do telemóvel.

### 5.2.1 Ambiente

Quando se faz a recolha de dados é importante definir o ambiente no qual os sensores estão inseridos.

Para monitorizar a atividade do telemóvel, considerou-se um ambiente mais abstrato. Assim sendo, este correspondeu a todo meio em que o utilizador esteja com o seu telemóvel, seja em casa, ou em qualquer outro local. É de notar que a recolha de dados foi feita a partir das 20 horas até ao momento em que o utilizador acorda. A hora de início da recolha foi definida para que fosse possível detetar o tempo de utilização do telemóvel, pelo menos 2 horas antes da pessoa adormecer.

### 5.2.2 Sensores

Para recolher os dados necessários, o grupo incluiu na aplicação *X-Leap* tanto sensores físicos como virtuais.

Em primeiro lugar, relativamente aos físicos, foi utilizado o sensor da luminosidade do ambiente. Desta forma, o grupo recorreu à classe `SensorManager` do Java que permitiu obter os dados do sensor pretendido. Consequentemente, também se utilizou a interface `SensorEventListener` para conseguir obter os dados do sensor de forma contínua.

Por outro lado, em relação aos virtuais recorreu-se a sensores de brilho e sensores que detetam se o ecrã está ligado ou desligado. Assim sendo, para obter o brilho implementou-se uma *thread* que a cada segundo obtém o valor atual. De seguida, para detetar se o ecrã está ligado ou desligado, foi utilizado um *BroadcastReceiver* que executa uma ação cada vez que o ecrã liga ou desliga.

### 5.2.3 Colecionador

Por fim, a ultima etapa da recolha de dados relativa à atividade do telemóvel, consistiu em definir um colecionador para guardar as informações necessárias. Desta forma, o grupo decidiu utilizar a base de dados *Cloud Firestore* da plataforma *Firebase*, uma vez que é bastante flexível e não tem custos monetários envolvidos. Por motivos de privacidade e para distinguir as informações de cada utilizador, foi criado um projeto na conta de cada um, onde serão armazenados os seus dados individualmente.

Assim sendo, os dados vão sendo continuamente recolhidos pela aplicação *X-Leap* que, por sua vez, será responsável por guardar as informações relevantes constantemente na base de dados. Em adição, é de realçar que a aplicação, utilizando o *BroadcastReceiver* referido anteriormente, foi guardando os dados cada vez que o utilizador ligasse ou desligasse o ecrã do telemóvel.

## 6 Tratamento de dados

Tendo em conta o que foi explicado anteriormente, passamos agora para a fase de processamento e tratamento dos dados colecionados. Ao longo deste tópico, irá explicar-se o processo de criação de um *dataset* CSV a partir dos dados armazenados no **Firestore** e no **Google Fit** de cada utilizador, além de se explicar também o tratamento que se efetuou, posteriormente, a esse mesmo *dataset*, com a ferramenta *Datalore*.

### 6.1 Pré-Processamento

Relativamente aos dados armazenados nos colecionadores já referidos, o objetivo principal, consistiu em colocar as informações num ficheiro CSV, sendo que seria criado um *dataset* final para cada um dos utilizadores, com os dados do sono e do uso do telemóvel. No entanto, e visto que os dados são obtidos de colecionadores diferentes, foi necessário, em primeiro lugar, criar dois *datasets* separados (um como as informações do sono e outro com as do telemóvel) sendo que depois foi feito o *merge* para obter esse *dataset* final.

Atendendo a isto, começamos com os dados do **Google Fit**. Com este colecionador, e recorrendo a um pequeno código em *Python*, foi possível obter informações relativas ao ciclo de sono de uma pessoa nomeadamente a data e hora a que acordou (*Start\_SleepTime*), a data e hora a que adormeceu (*End\_SleepTime*) e a duração dos seis estados do sono (*Deep Sleep*, *Light Sleep*, *Sleep*, *REM*, *Out of Bed*, *Awake*). Ora, cada um destes atributos corresponderia a uma coluna no *dataset*, sendo que o grupo achou que seria também relevante adicionar uma coluna extra que contivesse o tempo que durou o ciclo de sono de um utilizador. Para tal, é suficiente calcular a diferença entre a hora em que a pessoa adormeceu e a hora em que acordou. É de salientar que, estes dados, lidos de um ficheiro JSON, são armazenados num dicionário em que para cada um dos dias registados, está associada uma lista com os respetivos dados. Estes dias não são uma data, mas sim um *id* que identifica esse mesmo dia. A título de exemplo, o valor 1 representa o primeiro dia registado, o 2 o segundo e assim por

diante. Posto isto, basta fazer a conversão num *Pandas Dataframe*.

Passemos agora para a parte referente ao **Firestore**. A partir do **Firestore**, é possível obter as várias intensidades de brilho (do telemóvel) e luminosidade (do ambiente) ao longo de um dia, sendo que ambos estes atributos serão duas colunas no *dataset* referente a este colecionador. No entanto, e visto que possuímos um conjunto de valores, tanto de brilho como de luminosidade, decidiu-se calcular, para cada uma destas *features*, a sua mediana. Esta escolha deveu-se ao facto de que esta métrica não é tão afetada pela presença de valores extremos, ou seja, *outliers*. Por outro lado, visto que é possível também saber o momento em que o ecrã do telemóvel de um utilizador é ativado ou desativado, pode-se determinar o tempo de atividade, ou seja, o tempo total que esse usuário despendeu com dispositivo, num determinado dia. Para tal, basta somar todos os períodos temporais que ocorrem entre o ecrã estar ativado e passar estar desativado. É relevante também referir que, de acordo com vários estudos, nomeadamente da já falada *Sleep Health Foundation*, o impacto do uso dos aparelhos eletrónicos é maior aproximadamente uma hora e meia antes da pessoa adormecer. Desta forma, o grupo considerou que seria pertinente adicionar três novos atributos, sendo eles: o tempo de atividade, brilho e luminosidade verificados duas horas antes do ciclo de sono iniciar. Escolheu-se o período de duas horas, contrariamente a uma hora e meia, dado que, por norma, estes estudos possuem sempre uma pequena margem de erro. Outra informação armazenada no **Firestore** e que se considerou ao *dataset* é o atributo da avaliação da qualidade do sono (*rating*), sendo esta a *target feature*. No entanto, apesar de no colecionador, a *rating* do sono ter uma data associada, este atributo é relativo sono do dia anterior. Já esta data de *rating* é tida em conta num novo atributo (End Time). Por fim, considerou-se uma última *feature* no *dataset* sendo ela a data e hora em que o utilizador acordou (Start\_SleepTime). Este atributo é o mesmo que é obtido pelo **Google Fit**, tendo-se tomado a decisão de acrescentar novamente uma vez que é necessário haver uma coluna em comum entre *datasets* para se efetuar a sua união. É de salientar que tal como ocorreu o **Google Fit**, criou-se um pequeno código em *Python* para obter os dados do **Firestore**, sendo que estes foram armazenados num dicionário em que para cada um dos dias registados, estão associados as respetivas informações já detalhadas. Posto isto, basta converter este dicionário o num *Pandas dataframe*.

Considerando tudo o foi dito, o passo seguinte consiste em criar o *dataset* final para cada utilizador, efetuando-se, para isto, o *merge* entre o *dataframe* com os dados do **Google Fit** e o *dataframe* com os dados **Firestore**, sendo que esta operação é feita usando o atributo *Start\_SleepTime*, comum entre ambos. Após isto, basta fazer a conversão para CSV.

## 6.2 Tratamento

Tendo, nesta etapa, ficheiros CSV para cada um dos utilizadores, é feito o seu *import* na plataforma *Datalore*, adicionando-se para cada um dos *dataframes* uma nova coluna com o *id* que representa o utilizador. De seguida, efetua-se a concatenação destes *dataframes* num único.

Posto isto, e de forma a tornar o *dataset* mais limpo e pronto para a fase de visualização de previsão com um modelo de *Machine Learning*, houve a necessidade de tratar de algumas colunas. Em primeiro lugar, as colunas cujos atributos representam durações temporais (tipo *datetime*), nomeadamente, o tempo de atividade diário, o tempo de atividade nas últimas duas horas ou ainda as colunas dos estados do sono, foram convertidas para minutos (tipo inteiro). Esta decisão foi tomada visto que o grupo considerou que as durações em minutos seriam de mais fácil interpretação na fase de visualização.

Por outro lado, durante a análise do *dataframe* final verificou-se que as colunas com as durações dos estados do sono *Sleep* e *Out of Bed* possuíam sempre o valor 0 pelo que se decidiu retirar estas mesmas colunas.

Por fim, tanto o atributo *Start\_SleepTime* como o *End\_SleepTime*, ambos do tipo *datetime*, foram divididos em novas colunas com o mês, dia, hora e minuto de cada um dos atributos. É de realçar que, ao se fazer esta divisão, passamos assim a ter uma coluna com os dias exatos dos registos, pelo que a coluna com o *id* dos dias passa a ser desnecessária.

## 7 Descrição e Exploração dos dados coletados

### 7.1 Atributos do dataset tratado

Atributo	Descrição
User_id	Id do utilizador em estudo
Activity_Time	Tempo de utilização do telemóvel por parte do utilizador entre as 20 e a hora que acorda (em minutos)
Brightness	Mediana do brilho do ecrã do telemóvel durante o <i>activity_time</i>
Light	Mediana da luz ambiente durante o <i>activity_time</i>
Activity_Time_2H	Tempo de utilização do telemóvel por parte do utilizador duas horas antes de adormecer (em minutos)
Brightness_2H	Mediana do brilho do ecrã do telemóvel duas horas antes de adormecer
Light_2H	Mediana da luz ambiente do ecrã do telemóvel duas horas antes de adormecer
Awake	Tempo total que a pessoa esteve acordada na cama (em minutos)
Light_Sleep	Tempo total de sono leve (em minutos)
Deep_Sleep	Tempo total de sono profundo (em minutos)
REM	Tempo total de sono REM (em minutos)
Total_Minutes	Tempo total que o indivíduo esteve a dormir (em minutos)
Start_Sleep_Day	Dia em que a pessoa adormeceu
Start_Sleep_Month	Mês em que a pessoa adormeceu
Start_Sleep_Hour	Hora em que a pessoa adormeceu
Start_Sleep_Minute	Minuto em que a pessoa adormeceu
End_Sleep_Day	Dia em que a pessoa acordou
End_Sleep_Month	Mês em que a pessoa acordou
End_Sleep_Hour	Hora em que a pessoa acordou
End_Sleep_Minute	Minuto em que a pessoa acordou
Sleep_quality	Qualidade de sono avaliada pela pessoa

Tabela 3: Descrição dos atributos do Dataset tratado

## 7.2 Visualização de dados do dataset

Tendo-se obtido um *dataframe* tratado, passa-se de seguida para a fase de visualização de dados através de diversos elementos gráficos.

Em primeiro lugar, de modo a melhor entender como é que os atributos se relacionam entre si e o seu impacto no *target* (qualidade do sono), definiu-se a matriz de correlação [Fig. 4].

Analisando os valores obtidos verifica-se que a luz ambiente nas últimas duas horas antes de adormecer (Light\_2H) tem uma correlação negativa com a qualidade do sono, ou seja, há medida que uma aumentou a outra diminuiu. Por outro lado, tanto o sono leve (Light\_Sleep), como o sono profundo (Deep\_Sleep) apresentam correlação positiva com o *target*. Do mesmo modo, a qualidade do sono também aumenta à medida que o tempo de sono total (Total\_Minutes) cresce.

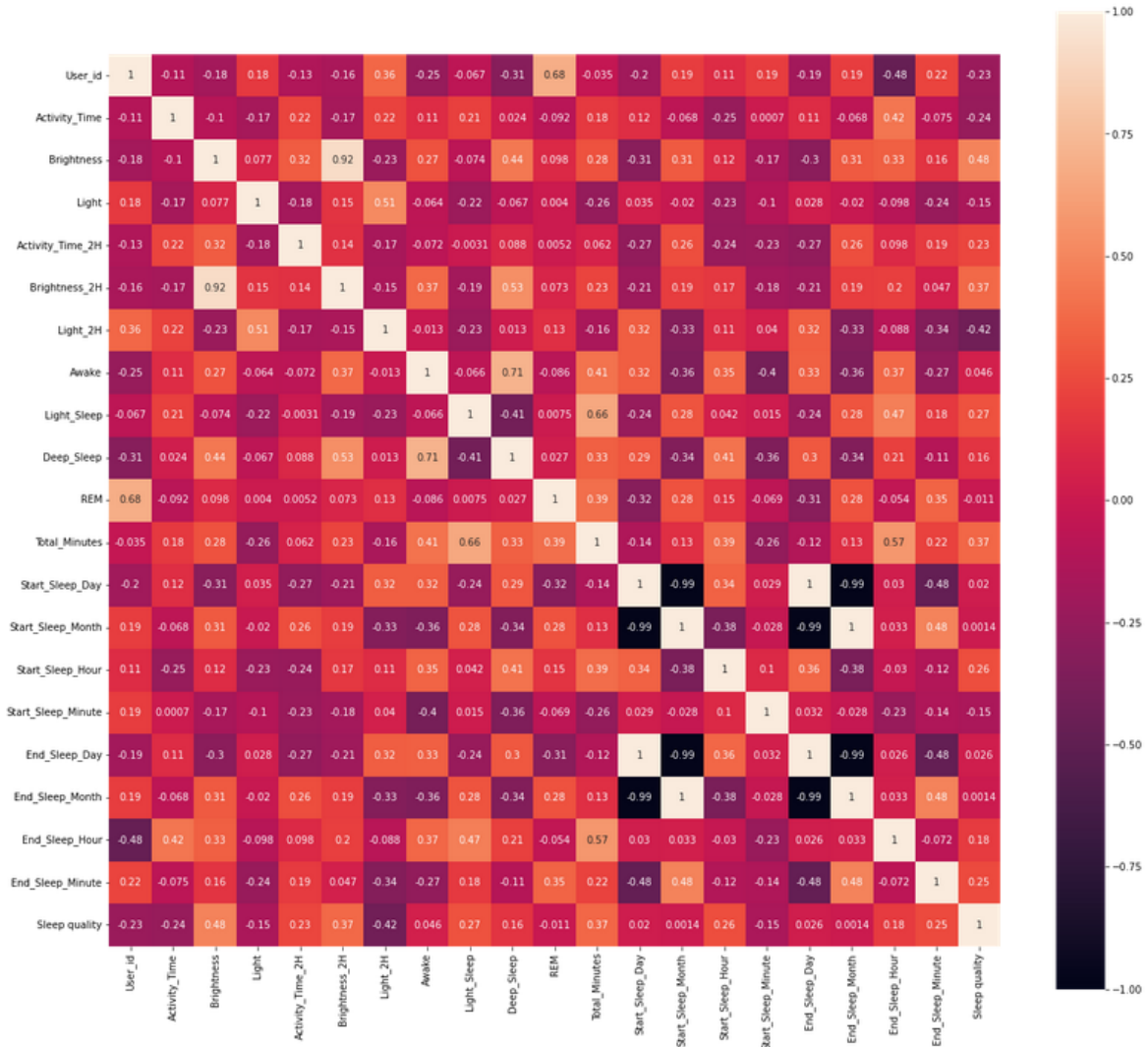


Figura 4: Matriz de correlação

Além desta matriz de correlação, também se criaram outros gráficos, de forma a analisar mais precisamente a relação entre atributos concretos.

Para verificar o impacto da utilização do telemóvel no sono foi feito um gráfico que relaciona o tempo de atividade com o tempo de sono profundo [Fig. 5]. Pela análise do gráfico pode-se verificar que o sono profundo apresenta variações pouco significativas à medida que o tempo de atividade varia. Por este motivo pode-se inferir que estas duas características não têm uma grande relação entre si

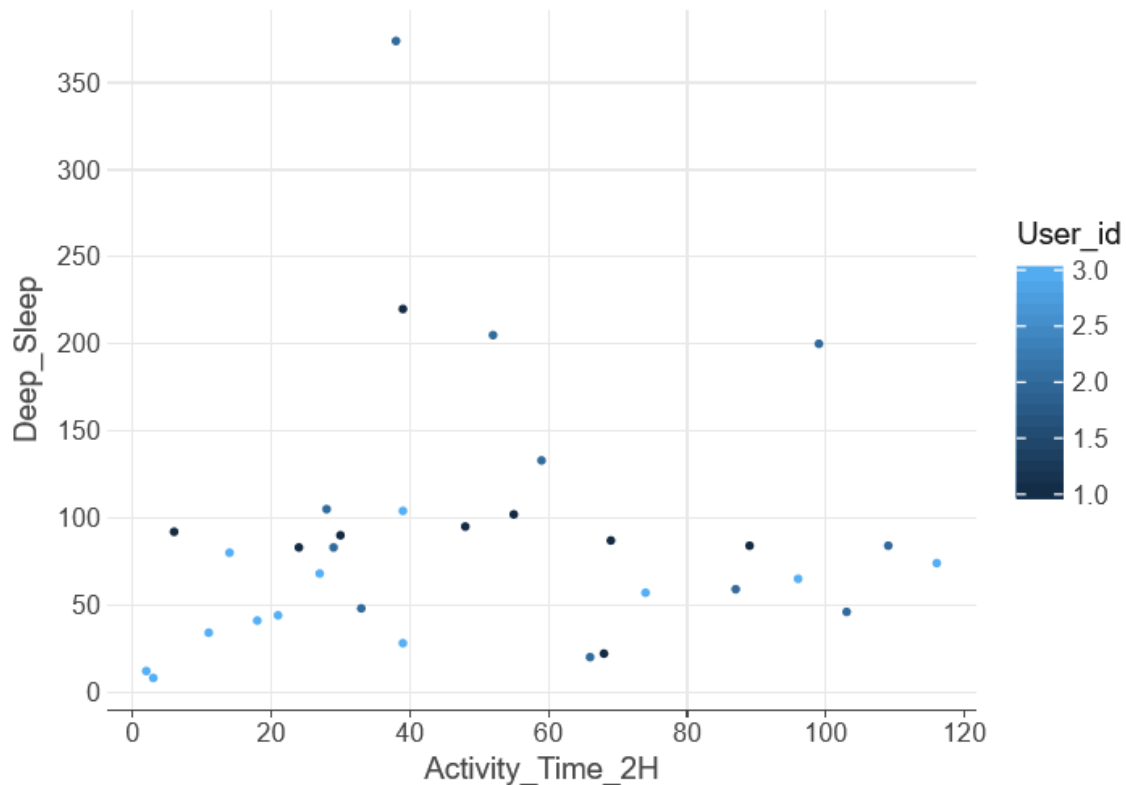


Figura 5: Impacto do Activity\_Time\_2H no Deep\_Sleep (por utilizador)

De seguida foi gerado um pairplot com as colunas 'Sleep quality', 'Total\_Minutes', 'Activity\_Time\_2H', 'Start\_Sleep\_Hour' [Fig. 6]. Pela análise dos gráficos verifica-se, em primeiro lugar, que pessoas que se deitaram mais cedo (entre as 20 e as 0) tiveram valores mais altos na qualidade de sono e com menor variabilidade do que aquelas que se deitaram após esse intervalo.



Para além disso, é possível observar que a qualidade de sono tende a crescer com o tempo total de sono da pessoa, como já se tinha verificado através da matriz de correlação

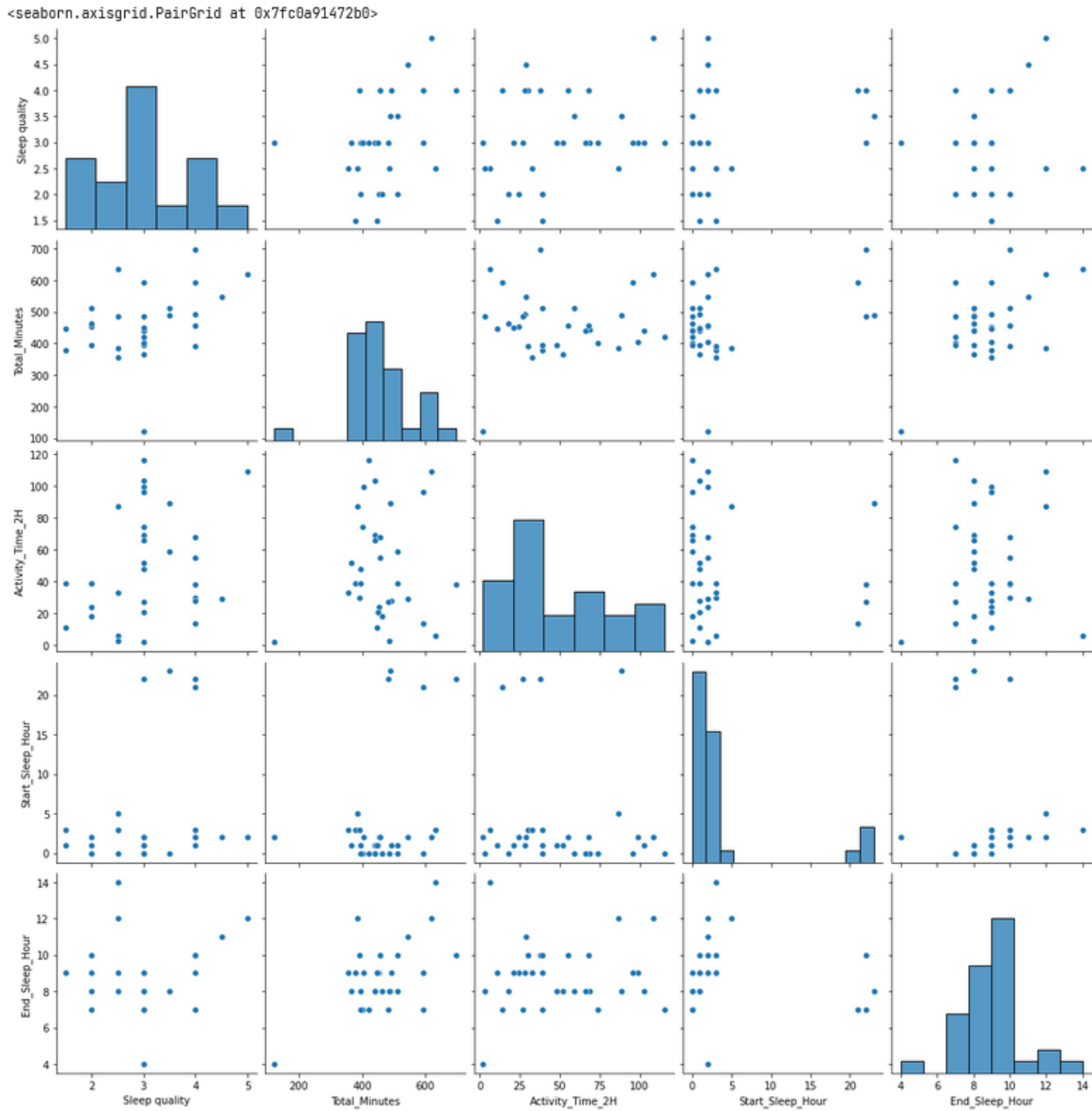


Figura 6: *Pairplot* de alguns atributos do dataset

Para ter uma noção da distribuição dos dados gerou-se gráfico de densidade de probabilidade para o atributo *Sleep\_quality* [Fig. 7]. Como se observa na figura, os dados estão distribuídos segundo uma distribuição normal, aproximadamente.

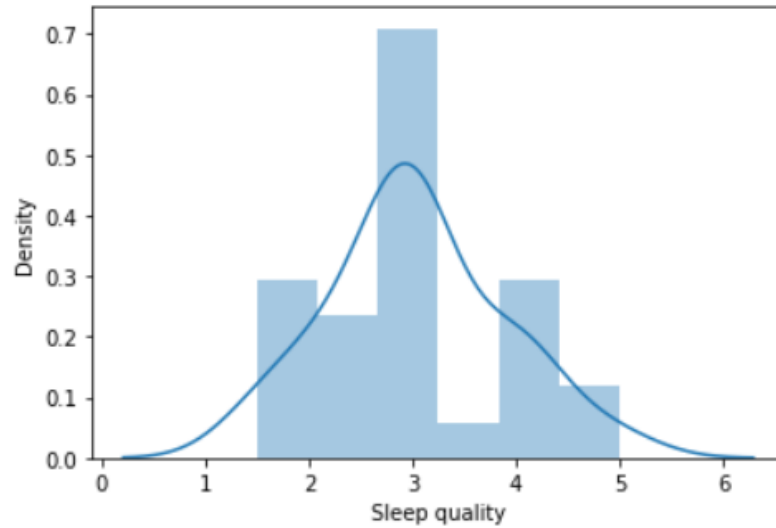


Figura 7: Distribuição da Sleep\_quality

Por fim, com o objetivo de encontrar outliers foram gerados *Boxplot* [Fig. 8]. Na figura está representado o boxplot do atributo Activity\_Time onde é possível identificar alguns outliers. Após alguma análise verificou-se que estes valores foram obtidos, uma vez que, um dos telemóveis onde eram recolhidos os dados, enquanto estivesse a carregar o seu ecrã mantinha-se ligado continuamente. Desta forma foi contabilizado como tempo de atividade esse intervalo em que o aparelho esteve apenas a carregar a bateria.

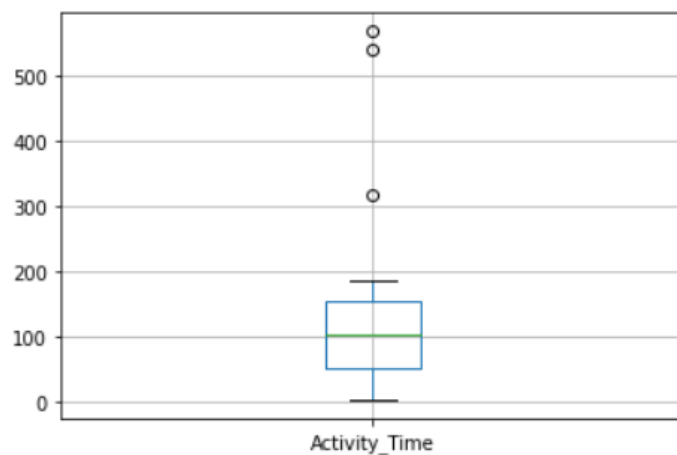


Figura 8: *Boxplot* do atributo Activity\_Time

## 8 Previsão com Modelos de Machine Learning

No que diz respeito aos modelos de *machine learning*, estes foram implementados sobre os dados recolhidos de modo a explorar como classificar a **qualidade de uma noite de sono** ( o nosso *target* / objetivo) tendo em consideração as restantes variáveis de input apresentadas anteriormente.

Assim sendo, seguimos uma abordagem clássica, começando por dividir os dados em dados de treino e teste. Os dados de treino são utilizados para treinar os modelos implementados : ***Decision Tree*** e ***Random Forest***. De seguida, os dados de teste são utilizados para testar ambos os modelos e gerar previsões. Finalmente, estes resultados gerados pelos modelos são analisados e avaliados, quer através da *accuracy* do modelo, por uma matriz de confusão e ainda por um *classification report*.

Numa nota comparativa, é possível observar que o modelo de *Random Forest* apresenta melhores resultados e, apesar do volume reduzido de dados gerados, este indício é bastante convincente. A justificação baseia-se no facto de o modelo de *Random Forest* ser mais robusto perante a presença de outliers e conseguir inferir mais eficazmente a distribuição dos dados, pelo que se torna num modelo mais completo e com melhor precisão comparativamente ao modelo de *Decision Tree*.

## 9 Aplicação Android: *X-Leap*

Como já se referiu anteriormente, de maneira a colectar dados relativos ao uso do telemóvel dos utilizadores, e enviá-los para o colecionador **Firestore**, o grupo desenvolveu uma aplicação *Android*, denominada X- Leap e cujo ícone é o demonstrado na Fig. 9 .

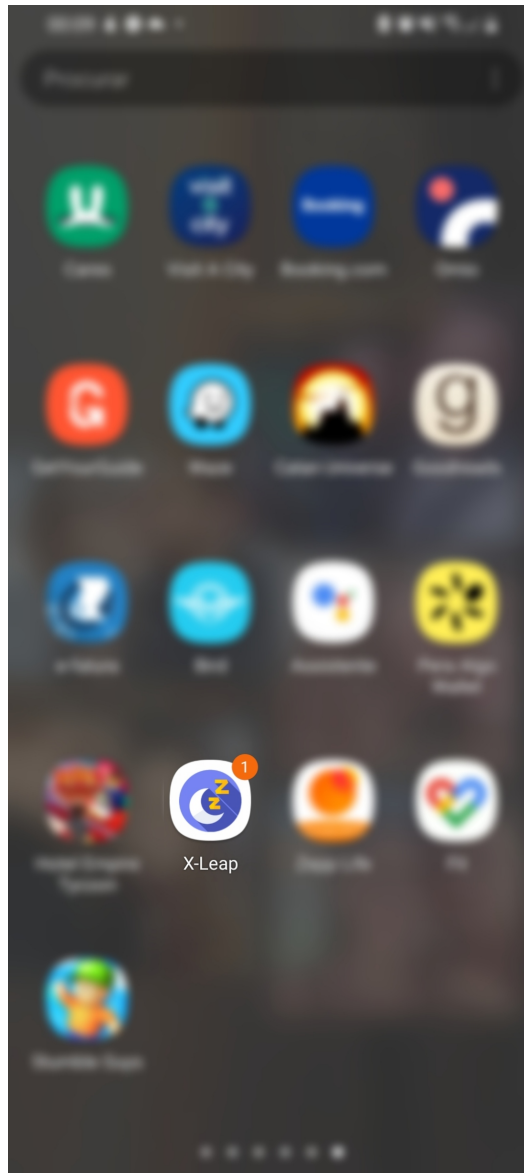


Figura 9: Ícone da aplicação X-Leap

Quando a aplicação é aberta, o seu utilizador depara-se com uma interface principal, clara e intuitiva (Fig. 10), na qual poderá ver algumas métricas tais como o número de toques efetuados no ecrã (apenas dentro da aplicação), a luminosidade e brilho do mesmo, além do momento em que o ecrã foi ativado pela primeira vez num, dado dia(*ScreenOn Start Time*). Além disto, nesta interface, o utilizador pode ainda ativar

ou desativar a recolha de dados na *app*, ou fazer a avaliação da qualidade do seu sono, selecionando, para o efeito, a quantidade de estrelas que deseja.

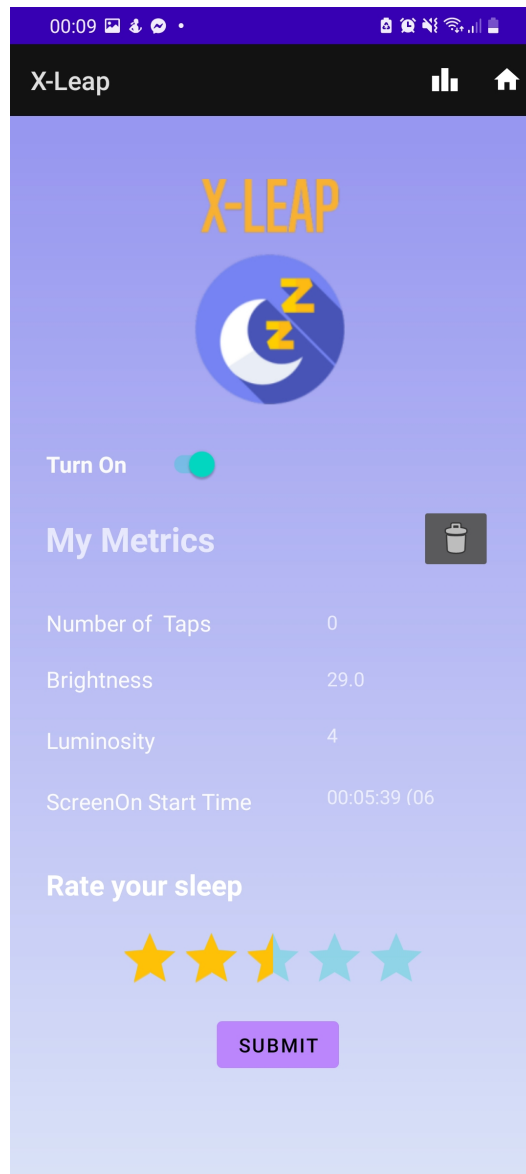


Figura 10: Interface da aplicação X-Leap

É importante também realçar que na barra preta, mais precisamente no canto superior direito, existem dois ícones, sendo que o primeiro revela aos utilizadores um conjunto de opções de gráficos estatísticos (Fig. 11), e o segundo ícone serve para retroceder para a interface principal.

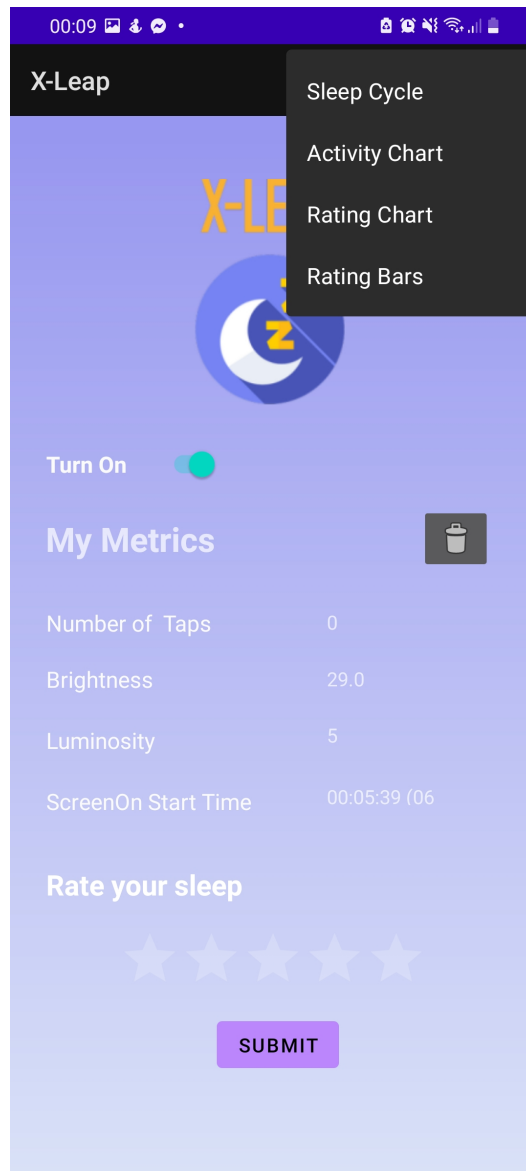


Figura 11: Opções de gráficos da aplicação X-Leap

Assim sendo, iremos agora abordar os diversos gráficos que é possível visualizar na **X-Leap**.

Em primeiro lugar, criou-se o *Sleep cycle*, sendo este um gráfico do tipo *Pie Chart* (Fig. 12), e no qual o utilizador poderá ver a percentagem que obteve para cada um dos estados do sono, na noite anterior. Ora, o grupo tomou esta decisão uma vez que considerou relevante que os utilizadores da aplicação conseguissem ver, de forma mais clara, como é que o seu ciclo de sono está dividido ao longo dos dias.

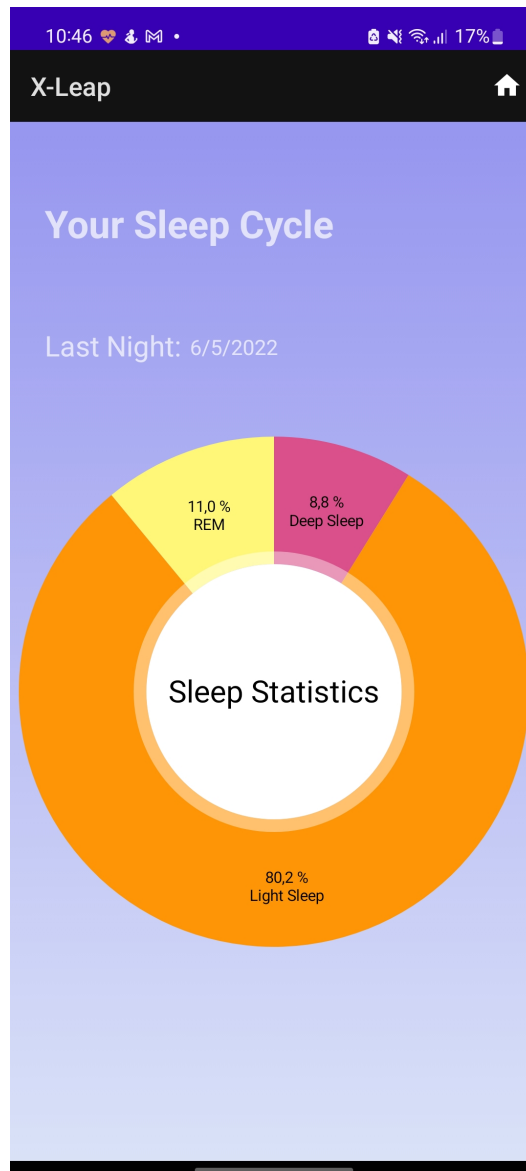


Figura 12: Gráfico relativo ao *Sleep Cycle*

Outro gráfico disponibilizado é o *Activity Chart* (Fig. 13), sendo este um gráfico do tipo *Line Chart*. Com ele, é possível observar de que forma o tempo de atividade/uso do telemóvel de um dado utilizador, tem variado ao longo de todos dias nos quais essa pessoa possui registos. O grupo colocou esta informação visto que considerou interessante o facto do utilizador conseguir observar se tem passado mais tempo no telemóvel, podendo assim concluir se este é um fator que lhe poderá estar a ser prejudicial.

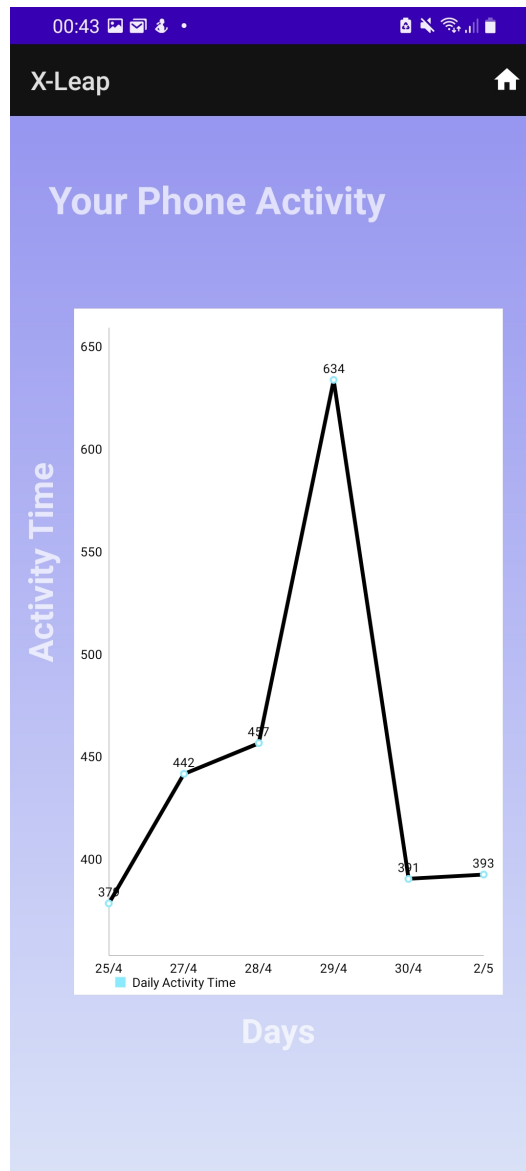


Figura 13: Gráfico relativo ao tempo de atividade

Relativamente ao *Rating Chart* (Fig. 14), este é um gráfico também do tipo *Line Chart* e no qual é possível observar de que forma a qualidade do seu sono tem variado ao longo de todos os dias nos quais essa pessoa possui registos. Com esta informação, e relacionando com o gráfico *Activity Chart*, o utilizador poderá determinar se a qualidade do seu sono tem sido, ou não, afetada pelo tempo que despende no telemóvel.



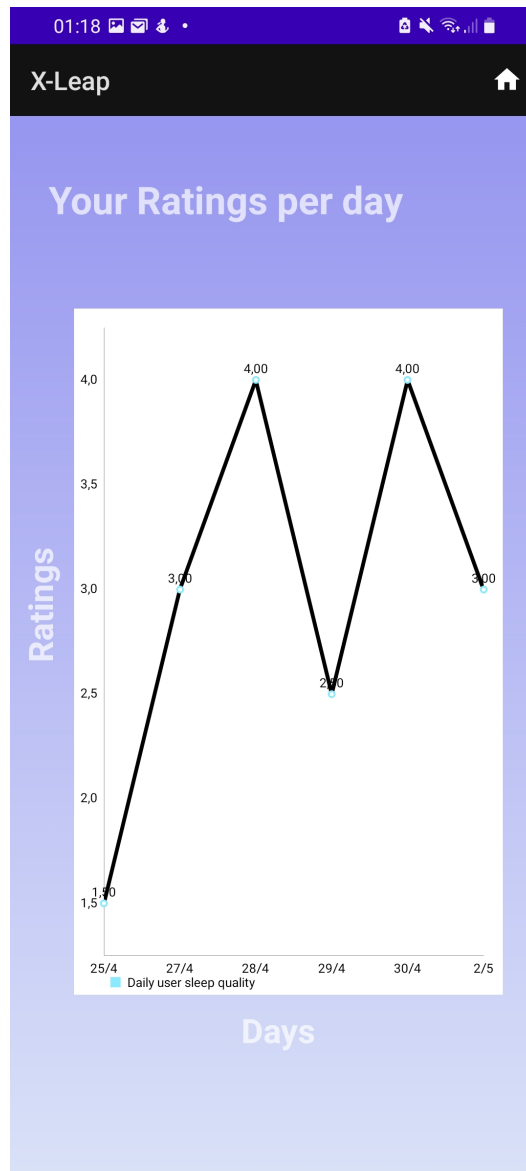


Figura 14: Gráfico relativo à variação da qualidade do sono

Por último, o gráfico *Rating Bars* (Fig. 15) é um *Bar Chart* sendo que este representa a mesma informação referida na *Rating Chart* explicado anteriormente. No entanto, com um gráfico em barras o utilizador poderá visualizar de forma mais clara a avaliação que atribuiu ao seu sono, ao longo dos dias.

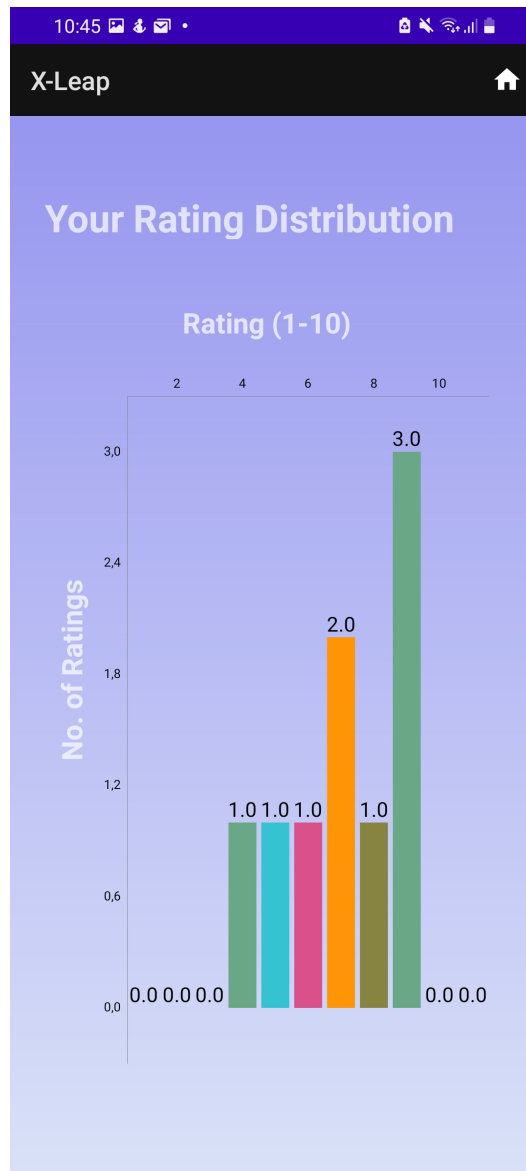


Figura 15: Gráfico de barras relativo às avaliações da qualidade do sono

Resumidamente, podemos concluir que os utilizadores têm à sua disposição uma aplicação simples, mas intuitiva na qual poderão visualizar de forma clara e esclarecedora diversas estatísticas relacionadas com o seu uso do telemóvel e com o seu ciclo de sono.

## 10 Análise dos Resultados Obtidos

Tendo por base todo o processo pelo qual os dados passaram, bem como um maior entendimento dos mesmos através da sua representação visual, fará agora sentido analisar de que forma os resultados obtidos poderão explicar quais são os principais fatores para uma boa ou má noite de sono.

Os resultados obtidos permitem comprovar alguns dos fatores que são bem conhecidos pela população mas que, no entanto, nem sempre são devidamente aplicados. Nomeadamente o facto de que o uso de dispositivos eletrónicos antes de dormir compromete a qualidade do sono, particularmente atividades que estimulem intensamente o cérebro, tais como jogos ou outros conteúdos ligados à resolução de problemas”, sendo este apenas um de outros tantos exemplos. De forma a obter resultados mais precisos seria necessário ter uma noção mais completa da vida de cada indivíduo em análise/estudo, considerando outros fatores como o seu consumo de cafeína, o tempo que passou ao telemóvel durante o dia e o fator de imersividade no dispositivo, o barulho ambiente nas horas antes de ir dormir, a exposição à luz solar e o uso ou não do filtro de luz azul antes de ir dormir, a atividade física e o batimento cardíaco diário, o estado emocional do próprio e a quantidade de stress a que o mesmo foi sujeito.

Num trabalho futuro seria interessante considerar estes diversos fatores, na medida em que permitiriam obter indicadores mais completos que auxiliariam no estudo dos resultados do sono de cada utilizador. Contudo, não sendo possível ter em conta estes fatores atualmente, a nossa análise basear-se-á nos restantes indicadores coletados - pelos nossos sensores - que, não obstante, apresentam uma boa base para a compreensão de quais as melhores práticas e padrões a adotar para um melhor sono, assim como quais a evitar.

Anteriormente, vimos através da análise da matriz de correlação [Fig. 4] que não existe apenas uma só vertente a impactar a qualidade do nosso sono mas sim um conjunto de características, algumas contribuindo com um peso maior do que outras, e aí se encontram os resultados, na subtilidade e nos pequenos detalhes que os dados nos podem transmitir.

Em adição, conforme apresentado nos gráficos das figuras [Fig. 5] a [Fig. 8], retém-se conceitos importantes como o porquê de existirem determinados *outliers* e qual a distribuição dos dados recolhidos. Com esta visualização cumpre-se o intuito de demonstrar mais explicitamente o que significam os dados recolhidos, na esperança de melhor compreender o que os mesmos pretendem transmitir e, por sua vez, tornando-se esta uma parte fundamental para a obtenção de resultados concretos.

Com base nos dados coletados, colocam-se agora várias questões de modo a concluir como se poderá obter uma melhor noite de sono.

A primeira questão relaciona-se com a ideia de que o **uso de dispositivos** afeta negativamente o sono. Para comprovar este argumento teve-se em consideração os dados do tempo de atividade do utilizador. No entanto, os resultados mostraram-se inconclusivos devido à escassez de dados, que os tornam pouco representativos do uso real do dispositivo, e à existência de *outliers* que tornam as conclusões contraditórias. Ou seja, por um lado, demonstram que o tempo de atividade diário (Activity\_Time)

apresenta uma correlação negativa com a qualidade do sono, corroborando assim com a ideia de ser mau usar o telemóvel. Por outro, parece que a utilização do dispositivo 2 horas antes de ir dormir (Activity\_Time\_2H) tem correlação positiva, o que pode dar a entender que o fator de relaxamento pode ajudar a ter um descanso melhor. Contudo esta ideia teria de ser analisada com melhor cuidado tendo em conta o tipo de conteúdo que o utilizador consome e qual o nível de envolvimento com o dispositivo antes de dormir, ambas métricas não medidas atualmente.

Em segundo lugar, procurou-se comprovar de que modo o nível de **luminosidade da tela** impacta o sono. À primeira vista os resultados parecem ir contra o expectável, parecendo que quanto maior o brilho da tela melhor a qualidade do sono. No entanto, após uma análise mais ponderada o grupo chegou a conclusão de que estes valores tem uma baixa correlação com o target (nomeadamente o Brightness\_2H), sendo que não consideramos o seu impacto muito relevante na qualidade do sono.

Outro fator importante relaciona-se com a hora de dormir. Tal como esperado, **deitar mais cedo** levou a um melhor sono. De acordo com os dados, quem se deitou entre as 20h e as 00h atingiu normalmente uma melhor qualidade de sono, ao passo de que quem se deitou depois das 00h tem uma maior variância associada a qualidade do sono.

Por último, pôde-se concluir qual o impacto do tempo total de sono e do sono profundo na qualidade do sono. Tal como era esperado, o **aumento do sono total leva a um aumento direto da qualidade de sono**. No entanto, o aumento do sono profundo não corresponde diretamente a um aumento da qualidade do sono, pelo que os dados são inconclusivos no que diz respeito a este aspeto e seria necessário uma pesquisa mais aprofundada.

## 11 Conclusão e Trabalhos Futuros

A realização deste trabalho permitiu aprofundar conhecimentos no que toca à conceção e implementação de sistemas de sensorização de ambientes tirando partido da integração de sensores físicos e virtuais. Para além disso, possibilitou o desenvolvimento de competências relativamente ao tratamento de dados em 'bruto', isto é sem qualquer tratamento prévio.

No que diz respeito ao principal objetivo do projeto, que foi avaliar o impacto da utilização do telemóvel na qualidade de sono, pode-se concluir através dos resultados obtidos que o manuseamento de dispositivos tecnológicos não afetou a qualidade do sono dos utilizadores em estudo. Apesar disso, o estudo poderia ter tido melhores resultados caso fosse utilizado um maior número de pessoas e se tivessem sido recolhidos dados por um maior intervalo de tempo. Por outro lado, o facto de terem sido utilizadas pessoas do grupo para fazer este estudo pode ter tornado os dados enviesados, na medida em que, a pessoa avaliou o seu próprio sono no entanto também teve acesso aos dados do mesmo.

Em relação ao trabalho desenvolvido, o grupo considera que foram cumpridos os objetivos pretendidos. Um dos principais pontos a destacar foi sucesso no desenvolvimento de uma aplicação *Android*, uma vez que foi algo nunca feito anteriormente. Em adição, o facto da recolha de dados ter sido feita de duas formas diferentes (smartwatch e telemóvel) tornou o trabalho mais completo no que toca à qualidade dos dados recolhidos.

No entanto, existem alguns pontos a melhorar, tais como a forma como foi monitorizada a atividade do telemóvel. Um dos objetivos inicialmente era contabilizar o número de toques que a pessoa dá no ecrã ao longo do tempo, no entanto após várias tentativas não foi possível implementar esta recolha.

Por fim, um dos principais objetivos no futuro seria acrescentar outras funcionalidades e melhorar ainda mais aparência da aplicação *X-Leap*, para que assim seja publicada numa plataforma como a *Play Store* e esteja disponível ao um maior número de pessoas.