

# Abstract

A long, long time ago...



# Resumo

Há muito, muito tempo



# Agradecimentos

Obrigado a todos, obrigado ...

**Dedico a ...**

# Conteúdo

<b>Abstract</b>	<b>i</b>
<b>Resumo</b>	<b>iii</b>
<b>Agradecimentos</b>	<b>v</b>
<b>Conteúdo</b>	<b>viii</b>
<b>Lista de Tabelas</b>	<b>ix</b>
<b>Lista de Figuras</b>	<b>xi</b>
<b>Lista de Blocos de Código</b>	<b>xiii</b>
<b>1 Introdução</b>	<b>1</b>
1.1 Projeto . . . . .	4
1.1.1 Objetivos . . . . .	5
1.2 Second Section example . . . . .	5
1.2.1 SubSection example . . . . .	5
<b>2 Estado da Arte</b>	<b>7</b>
2.1 Dispositivos para monitorizar a diabetes . . . . .	7
2.1.1 Bombas infusoras de insulina . . . . .	7
2.1.2 Glicosímetros . . . . .	8
2.1.3 Monitor contínuo de glicose . . . . .	8

2.2	Aplicações para smartphones Android . . . . .	8
2.2.1	Diário da Diabetes mySugr . . . . .	8
2.2.2	Diabetes:M . . . . .	9
2.2.3	OnTrack Diabetes . . . . .	9
2.2.4	Diabetes - Diário Glucose . . . . .	10
2.2.5	Glucose Buddy: Diabetes Log . . . . .	10
2.3	Medicina personalizada e data mining na saúde . . . . .	11
<b>3</b>	<b>MyDiabetes</b>	<b>13</b>
3.1	Objetivo da aplicação . . . . .	13
3.2	Arquitetura . . . . .	13
3.3	Variáveis recolhidas . . . . .	13
<b>4</b>	<b>Análise de dados</b>	<b>15</b>
4.1	Descrição do estudo . . . . .	15
<b>5</b>	<b>Conclusões</b>	<b>17</b>
5.1	Trabalho Futuro . . . . .	17
	<b>Bibliografia</b>	<b>19</b>
<b>A</b>	<b>Acrónimos</b>	<b>19</b>



# Lista de Tabelas



# Lista de Figuras



# Lista de Blocos de Código



# Capítulo 1

## Introdução

A diabetes, também conhecida por *diabetes mellitus*, é uma doença crónica e bastante comum, conhecida por fazer com que os seus portadores tenham níveis de glicose (açúcar) no sangue mais elevados que o normal. Isto deve-se ao facto de o pâncreas não funcionar da forma devida ou nem sequer funcionar, de todo. Antes de nos aprofundarmos sobre a doença em si, eis alguns factos preocupantes:

Segundo a International Diabetes Federation (IDF), em 2014, cerca de 387 milhões de pessoas tinham diabetes. Em 2035, este número aumentará para 592 milhões e, de acordo com a Organização Mundial da Saúde, em 2030 a diabetes será a sétima causa de morte no planeta.

Como se pode perceber, esta doença afeta muita gente e, por isso, cada vez mais se torna importante conseguir adiar, ou até mesmo prevenir, o seu aparecimento. No entanto, nem sempre isso é possível e, uma vez tendo a doença, esta é crónica. De seguida, explicaremos, de forma resumida, o porquê de a diabetes aparecer.

A diabetes ocorre quando uma pessoa tem valores de glicose no sangue demasiado elevados. A glicose vem dos carboidratos e é o único monossacarídeo que produz ATP. Resumindo, a glicose produz energia que vai ser utilizada pelas células. Há diferentes razões para que os níveis de glicose no sangue sejam mais elevados que o desejável: a produção de insulina pode não ser suficiente ou, mesmo que seja, as células não respondem à insulina como deveriam. É portanto natural que os níveis de glicose tendam a aumentar depois das refeições. Embora a diabetes esteja relacionada com a insuficiência de insulina no corpo, há diferentes tipos de diabetes. Os principais são:

- Diabetes Mellitus Tipo 1: A diabetes tipo 1, também conhecida como diabetes insulino-dependente ou diabetes juvenil, caracteriza-se pela produção deficiente de insulina pelo pâncreas e requer que o paciente tome doses de insulina diariamente. Normalmente aparece em jovens e é impossível de prevenir. Estima-se que apenas 5% dos diabéticos tenha este tipo. Entre os sintomas, incluem-se sede insaciável, fome constante ou perda de peso.

- Diabetes Mellitus Tipo 2: A diabetes tipo 2, ou diabetes não-insulino-dependente, ocorre

porque as células não usam a insulina de forma adequada. Neste tipo de diabetes, o problema não está necessariamente no pâncreas, mas sim na forma como as células utilizam a insulina. Cerca de 90% dos diabéticos tem este tipo de diabetes que, normalmente, é associado a um estilo de vida pouco saudável. Por isso mesmo, este tipo de diabetes é frequentemente resultado de excesso de peso ou falta de exercício físico. É comum os doentes de DM2 não necessitarem de tomar insulina e a medicação é feita através de comprimidos. Este tipo de diabetes surge, geralmente, em pessoas mais velhas, mas tem-se vindo a manifestar também em jovens.

- Diabetes gestacional: Este tipo de diabetes pode aparecer durante a gravidez. Caracteriza-se por ter valores de glicose superiores aos normais mas, ainda assim, abaixo dos valores diagnosticados na diabetes. Normalmente, este tipo de diabetes é descoberto nas consultas e não devido a sintomas. Há também o risco de mulheres que sofram deste tipo de diabetes, poderem no futuro sofrer de diabetes tipo 2.

- Diabetes LADA: "Latent Autoimmune Diabetes in Adults", traduzido para "diabetes auto-imune latente em adultos". Este tipo de diabetes é considerado uma variação de diabetes tipo 1, embora com uma evolução mais lenta. No entanto, muitas vezes é erradamente diagnosticado como tipo 2.

Uma vez que esta doença é crónica, ou seja, não há cura, é extremamente importante que os doentes recebam tratamento adequado. No caso dos portadores de diabetes tipo 1 e de uma pequena percentagem de portadores de tipo 2, uma parte fundamental do tratamento passa pela administração de insulina. No entanto, para todos os tipos de diabéticos, o importante para levar uma vida normal, passa pelo registo regular e constante dos níveis de glicose no sangue. Isto porque a falta de eficácia do pâncreas leva a um aumento natural dos níveis de glicose no sangue e, por outro lado, a toma de insulina, quando em quantidades não adequadas, pode ter o efeito contrário e provocar níveis demasiado baixos de açúcar no sangue. Estes dois estados são conhecidos, respetivamente, por hiper e hipoglicemia.

Tanto a hiper como a hipoglicemia são estados que podem fazer parte do dia-a-dia dos diabéticos. São ambos de evitar e potencialmente perigosos. A hiperglicemia pode trazer complicações a longo prazo, como doenças cardíacas e renais. Podem fazer parte dos sintomas de hiperglicemia sede intensa, cansaço, dores de cabeça, visão turva, entre outros.

Por sua vez, a hipoglicemia apresenta alguns sintomas tais como tremores, fome, nervosismo, calor, mas não só. Embora menos preocupante a longo prazo do que a hiperglicemia, a curto prazo a hipoglicemia pode ser mais perigosa. Isto porque o cérebro necessita de açúcar para poder funcionar de forma correta e, quando não o tem, pode levar a que o doente perca a consciência. A partir do momento em que o doente desmaia, não consegue contrariar a hipoglicemia, que normalmente se combate pela ingestão imediata de açúcar.

Como se pode perceber, a medição dos níveis de glicose no sangue e também o seu registo, são fatores fundamentais para que um diabético possa ter uma rotina normal. Em plena era da tecnologia, essa medição é feita através de pequenos dispositivos, os glucómetros, que, numa questão de segundos, conseguem dar uma medição precisa dos níveis de açúcar no sangue. O



registro é também importante para comparar valores, de forma a que os médicos consigam acompanhar o historial do doente. Também para o registro há dispositivos próprios. No entanto, num período em que os *smartphones* estão cada vez mais em voga, podem ser a ferramenta mais natural para se efetuarem estes registos. É importante para um médico poder observar os valores de glicose dos seus pacientes, que têm de ser registados pelos próprios pacientes. Um registro em papel torna-se rapidamente confuso e até passível de se perder, enquanto que um registro informático traz mais segurança e organização. Embora hajam dispositivos próprios para o registro dos níveis de glicose, isso pode implicar andar sempre com um dispositivo a mais. Por outro lado, como, em princípio, uma pessoa anda sempre com o seu telemóvel, este registro torna-se mais fácil, visto que oferece grande portabilidade e não necessita de nenhum dispositivo extra. É portanto uma vantagem da aliança da tecnologia à saúde. Não é a única, no entanto.

*Data Mining* é uma área da ciência de computadores que permite, através da análise de grandes quantidades de dados, descobrir padrões e regras que uma análise mais simples pode não detetar. A área de *Data Mining* usa diversos métodos de outras áreas, tais como matemática, inteligência artificial e *machine learning*, para tratar, explorar e obter conclusões acerca dos dados. *Data Mining* pode ter diversos fins, como por exemplo, deteção de anomalias, associação e classificação.

Como se pode perceber pelo nome, deteção de anomalias tem como objetivo a identificação de valores anormais, que podem apenas ser erros mas podem também ter interesse para uma determinada área.

A associação procura relações entre variáveis e pode quantificar essas relações. Um exemplo deste tipo poderia ser a relação entre o emprego de alguém e o seu carro.

Finalmente, a classificação tem como objetivo estudar conjuntos de dados para depois, ao observar novos dados, conseguir classificá-los corretamente. Por exemplo, ao estudar um *dataset* com alguns dados e tendo uma variável "diagnóstico" como sendo "diabético" ou "não-diabético", ao fazer uma aprendizagem e analisar um outro *dataset* com os mesmos tipos de dados, conseguirá, com alguma precisão, classificar corretamente como "diabético" ou "não-diabético".

Nos últimos anos, a área de *Data Mining* tem-se tornado bastante popular e consequentemente bastante usada em diversas áreas, tais como economia, educação e saúde. É precisamente nesta última área que *Data Mining* se pode tornar especialmente útil. Tal como já discutido anteriormente, a diabetes pode provocar oscilações nos valores de glicose no sangue, que podem não ser perceptíveis pelos doentes ou, mesmo que sejam, podem não ser tidas como importantes. No entanto, ao fazer aprendizagem sobre os dados dos pacientes, pode ser possível descobrir que afinal essas oscilações podem ter origem em comportamentos rotineiros, ou seja, fazem parte de um padrão. Um dos objetivos da *Data Mining* é precisamente descobrir padrões.

Torna-se portanto óbvio que esta área pode ser útil na medicina, em particular para os portadores de diabetes, ao permitir descobrir rotinas que provocam oscilações nos níveis de açúcar no sangue, e que de outra forma poderiam passar despercebidas. Se pensarmos que, ao desborir a causa de tais oscilações, pode descobrir-se a solução, isto ganha ainda mais importância.

De forma a perceber melhor de que forma a análise de dados pode melhorar o controlo da diabetes, tomemos como exemplo o seguinte caso: um determinado paciente tem como rotina fazer exercício às segundas-feiras. Às terças-feiras tem sempre níveis de glicose mais baixo. No entanto, pode acontecer que nos outros dias até tenha valores normais e portanto, não se preocupar com esta situação. Contudo, este valor mais baixo pode ser justificado: no dia anterior praticou exercício e não se alimentou devidamente, tendo hipoglicemia na manhã de terça-feira.

Tendo uma ferramenta que seja capaz de analisar dados, descobrir padrões e regras, o doente não só fica mais consciente de erros que pode cometer no controlo da diabetes como também como os combater.

Algumas referências para se ver a ordenação [? ]. Aqui outra ainda [? ].

Deve-se ler Strunk [? ]

Hypertext Transfer Protocol ([HTTP](#)) é um protocolo como diz em [? ] e também na secção 2.

Temos aqui o Clausen [? ]. E para ver várias [? ? ? ]

Deve-se acrescentar os acrónimos no ficheiro `acros.tex` e ordená-los alfabeticamente nesse ficheiro.

## 1.1 Projeto

Esta dissertação integra-se no projeto "Smart Diabetes Self-Management" que conta com uma aplicação Android chamada "My Diabetes". De momento, a aplicação permite ao utilizador registar refeições, que podem ser acompanhadas com fotos, registar medições de glicose, definir objetivos para glicemias, registar doenças, exercícios, entre outros. Contudo, permite apenas registar e visualizar os dados inseridos.

No âmbito desta dissertação, propõe-se desenvolver um sistema capaz de analisar dados introduzidos pelos doentes, que consiga detetar padrões e anormalidades nos dados introduzidos, avisando o utilizador. Para o conseguir, pretende-se ampliar as funcionalidades da aplicação "My Diabetes" para que esta consiga "aprender" o que é normal e o que é "anormal", gerando regras que traduzam as anormalidades, com base no *input* obtido.

Para o conseguir, tornou-se necessário obter dados de pacientes com diabetes insulino-dependente. Deste modo, em parceria com o Hospital de São João do Porto, foi levado a cabo uma sensibilização dos doentes para utilizarem a aplicação de forma voluntário, sendo que, em última análise, estes serão os maiores beneficiados.

A utilização da aplicação por parte de voluntários tem diversos objetivos: em primeiro lugar, obter *feedback* da aplicação em si, tal como críticas ou sugestões; em segundo, poder construir *datasets* de dados reais de registos diabéticos, algo escasso na web. A obtenção de dados reais permite ter uma ideia do tipo de dados que a aplicação vai obter por parte de utilizadores, bem

como o tipo de padrões que podem ser descobertos ou regras que podem ser geradas.

Portanto, a análise será feita aos dados que os utilizadores inserirem na aplicação. Podem ser registados dados tão diferentes como peso, altura, colesterol, exercício ou doenças. Para a descoberta de regras, alguns dados serão mais relevantes que outros, por terem um impacto direto nos valores de glicose, tais como hidratos de carbono, exercício ou doenças, e portanto a análise será feita principalmente com base nos registos destes dados. Isto servirá para dotar a aplicação de "inteligência", isto é, a aplicação ao longo do tempo aprende o que são valores normais e o que podem ser situações anormais. Ao saber que ocorreram situações anormais, também é possível o que as originaram. Uma vez feita a aprendizagem, sempre que situações semelhantes ocorram, a aplicação mostrará um aviso.

Para a aprendizagem serão utilizadas diferentes técnicas de *Data Mining* que gerarão regras. Essas regras serão então traduzidas para Prolog. A linguagem escolhida foi Prolog por se tratar de uma linguagem de programação lógica fortemente associada a inteligência artificial e *machine learning*. Além disso, e ao contrário de outras linguagens, o Prolog é declarativo. Isto significa que um programa em Prolog é expresso em termos de relações, que são definidas por factos e regras.

### 1.1.1 Objetivos

O objetivo da dissertação é integrar um sistema capaz de gerar regras e avisos a partir dos dados inseridos. De um modo sumário, os objetivos são:

- Obter dados através da participação de pessoas de diabetes;
- Analisar esses dados para conseguir detetar padrões ou anomalias;
- Criar regras a partir da análise dos dados;
- Integrar o sistema desenvolvido na aplicação para que esta, de forma autónoma, consiga gerar avisos.

[ESTRUTURA DO DOCUMENTO]

## 1.2 Second Section example

### 1.2.1 SubSection example

Sed commodo posuere pede. Mauris ut est. Ut quis purus. Sed ac odio. Sed vehicula hendrerit sem. Duis non odio. Morbi ut dui. Sed accumsan risus eget odio. In hac habitasse platea dictumst. Pellentesque non elit. Fusce sed justo eu urna porta tincidunt. Mauris felis odio, sollicitudin sed, volutpat a, ornare ac, erat. Morbi quis dolor. Donec pellentesque, erat ac sagittis

semper, nunc dui lobortis purus, quis congue purus metus ultricies tellus. Proin et quam. Class aptent taciti sociosqu ad litora torquent per conubia nostra, per inceptos hymenaeos. Praesent sapien turpis, fermentum vel, eleifend faucibus, vehicula eu, lacus.

Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Donec odio elit, dictum in, hendrerit sit amet, egestas sed, leo. Praesent feugiat sapien aliquet odio. Integer vitae justo. Aliquam vestibulum fringilla lorem. Sed neque lectus, consectetur at, consectetur sed, eleifend ac, lectus. Nulla facilisi. Pellentesque eget lectus. Proin eu metus. Sed porttitor. In hac habitasse platea dictumst. Suspendisse eu lectus. Ut mi mi, lacinia sit amet, placerat et, mollis vitae, dui. Sed ante tellus, tristique ut, iaculis eu, malesuada ac, dui. Mauris nibh leo, facilisis non, adipiscing quis, ultrices a, dui.

Morbi luctus, wisi viverra faucibus pretium, nibh est placerat odio, nec commodo wisi enim eget quam. Quisque libero justo, consectetur a, feugiat vitae, porttitor eu, libero. Suspendisse sed mauris vitae elit sollicitudin malesuada. Maecenas ultricies eros sit amet ante. Ut venenatis velit. Maecenas sed mi eget dui varius euismod. Phasellus aliquet volutpat odio. Vestibulum ante ipsum primis in faucibus orci luctus et ultrices posuere cubilia Curae; Pellentesque sit amet pede ac sem eleifend consectetur. Nullam elementum, urna vel imperdiet sodales, elit ipsum pharetra ligula, ac pretium ante justo a nulla. Curabitur tristique arcu eu metus. Vestibulum lectus. Proin mauris. Proin eu nunc eu urna hendrerit faucibus. Aliquam auctor, pede consequat laoreet varius, eros tellus scelerisque quam, pellentesque hendrerit ipsum dolor sed augue. Nulla nec lacus.

Suspendisse vitae elit. Aliquam arcu neque, ornare in, ullamcorper quis, commodo eu, libero. Fusce sagittis erat at erat tristique mollis. Maecenas sapien libero, molestie et, lobortis in, sodales eget, dui. Morbi ultrices rutrum lorem. Nam elementum ullamcorper leo. Morbi dui. Aliquam sagittis. Nunc placerat. Pellentesque tristique sodales est. Maecenas imperdiet lacinia velit. Cras non urna. Morbi eros pede, suscipit ac, varius vel, egestas non, eros. Praesent malesuada, diam id pretium elementum, eros sem dictum tortor, vel consectetur odio sem sed wisi.

## Capítulo 2

# Estado da Arte

Neste capítulo, pretende-se analisar as tecnologias existentes utilizadas na diabetes. Vamos dividir este capítulo em dois grupos: dispositivos utilizados para monitorizar e registar os vários parâmetros existentes e o uso de medicina personalizada para controlar a diabetes, através de técnicas de *data mining*. Medicina personalizada é a prática de tratar cada doente de forma individualizada, de acordo com as suas características, necessidades e preferências a cada momento, em vez de um tratamento generalizado para todos os pacientes. [PersMed.pdf]

### 2.1 Dispositivos para monitorizar a diabetes

Quanto a dispositivos para monitorizar a glicose, há várias soluções: bombas infusoras de insulina, monitores contínuos de glicose e glicosímetros. Com a constante evolução da tecnologia, nomeadamente dos telemóveis, os *smartphones* tornaram-se também uma ferramenta que pode ajudar nesta doença. De seguida, vamos comparar os diferentes dispositivos, com especial ênfase nas aplicações para *smartphone*, uma vez que este projeto se assenta numa aplicação.

#### 2.1.1 Bombas infusoras de insulina

Uma bomba infusora de insulina é um pequeno dispositivo liberta insulina de ação rápida 24 horas por dia. A quantidade de insulina libertada é ajustada de acordo com as necessidades do utilizador. Existem várias marcas e modelos no mercado, e, apesar de todas terem o mesmo objetivo fundamental, têm algumas diferenças nas funcionalidades que oferecem. Um exemplo de bomba é a Accu-Chek Combo: é composta pela bomba e por um monitor de glicemia, que comunicam entre si através de *bluetooth* para que a insulina injetada seja de acordo com os níveis de glicemia. [akku]

### 2.1.2 Glicosímetros

O glicosímetro é o dispositivo base para qualquer diabético: permite medir os níveis de glicemia a qualquer instante, através de uma pequena quantidade de sangue. São uma importante ferramenta pois permitem ao doente saber qual o seu nível de glicose no sangue a dada altura para que possa assim ajustar a insulina a tomar.

### 2.1.3 Monitor contínuo de glicose

É um pequeno aparelho que o utilizador usa a toda a hora e que está constantemente a medir os níveis de glicemia. Assim, quando estes valores forem demasiado altos ou baixos, emite um aviso para que o utilizador possa tomar a medida mais adequada. Um exemplo de um dispositivo deste tipo é o da Dexcom. [continuous]

## 2.2 Aplicações para smartphones Android

Estima-se que em 2016 o número de utilizadores de *smartphones* seja, em todo o mundo, de 2.08 mil milhões. Por outro lado, os *smartphones* têm cada vez mais poder de processamento, o que os torna úteis para inúmeras atividades, como a saúde, nomeadamente a diabetes. Uma parte importante do controlo da doença passa pela monitorização dos níveis de glicemia e do seu registo, para que se possa manter um controlo mais rigoroso. Torna-se, portanto, relevante analisar o que existe no mercado em termos de aplicações que possam ser úteis no controlo da diabetes.

Para esta comparação foram tidas em conta apenas aplicações para Android, por ser o sistema operativo móvel mais usado no mundo [mercadomovel] e porque a aplicação na qual este projeto se baseia é também para Android. Foram escolhidas cinco aplicações da *Google Play*. Estas aplicações foram escolhidas com base no número de *downloads* e número de *ratings*. Cada aplicação foi instalada e testada com o intuito de perceber aquilo que oferece ao utilizador. Todas as aplicações escolhidas são grátis para *download*.

### 2.2.1 Diário da Diabetes mySugr

Esta aplicação permite ao utilizador adicionar registos. Cada registo permite especificar alguns parâmetros, como o nível de glicemia, hidratos de carbono consumidos, tipo de insulina e tipo de refeição. Cada registo pode ser acompanhado para uma foto, caso seja uma refeição, e pode ser também escolhido um tipo para cada registo, como por exemplo "almoço", "jantar", "hipoglicemia", entre outros. Para cada registo é ainda possível escolher um outro tipo que dá mais informação, como "Stressado", "Doente", "Álcool", mas não só. De nota também que é possível especificar o tipo de alimentos caso o registo se trate de uma refeição. Entre os tipos de

alimentos existem, entre outros, "Legumes", "Carne", "Peixe", "Ovos", etc.

Esta aplicação permite a sincronização com um glicómetro, o "iHealth BG5". [ref] É ainda possível definir metas como limite para hipo e hiperglicemia, e metas de peso ou exercício. Uma característica interessante da aplicação é ter um sistema de pontos e de desafios. Os desafios são diversos, como por exemplo "Caminhada para a cura", que incentiva o utilizador a registar pelo menos 30 minutos de exercícios em 24 horas. Desafios completos desbloqueiam novos desafios.

Por cada registo efetuado ganha-se uma quantidade de pontos, que é maior quantos mais parâmetros forem preenchidos em cada registo. A aplicação tem um pequeno boneco animado que vai sendo desbloqueado com pontos. Estes dois sistemas são interessantes porque podem funcionar como um incentivo extra para o uso regular da aplicação.

Por fim, a aplicação possibilita a exportação dos registos efetuados para três formatos possíveis: xls, pdf ou csv. Esta característica, no entanto, está disponível apenas na versão paga.

### 2.2.2 Diabetes:M

Esta aplicação permite o registo de glicose, hidratos de carbono consumidos, insulina de efeito rápido e longo, peso, colesterol, pressão arterial, atividade física e hemoglobina glicosilada (também conhecida por hemoglobina glicada ou HbA1c). À primeira vista, nota-se logo o ecrã principal que se pode tornar confuso pela grande quantidade de botões que oferece. As funções disponibilizadas são bastante semelhantes às da aplicação anterior. Uma função nova é a de alarme, que ajuda os utilizadores a não se esquecerem de medir a glicose. Em termos de visualização dos dados inseridos, a aplicação mostra os mesmos em forma de gráficos para se poder acompanhar os registos num determinado intervalo de tempo. É possível verificar que se podem usar unidades de medida diferentes para os vários parâmetros. Por exemplo, para a glicemia pode-se usar mg/dL ou mmol/L. Uma vantagem do ecrã principal é mostrar a quantidade de insulina ativa presente num dado momento. Isto é, se um utilizador tomar 5 doses de insulina, a aplicação mostra, ao longo do tempo, um valor denominado "Insulina Ativa", ou seja, a previsão da insulina que "sobra" desde a última toma.

Uma outra característica interessante é a de possibilitar sincronizaçã com aplicações externas, como Dropbox, Google Drive e Google Fit. A aplicação permite ainda fazer *backup* dos dados.

É também possível exportar e importar dados nos formatos csv e xls, bem como importar dados de glicómetros de diferentes modelos, tais como OneTouch, Dexcom ou Accu-Chek.

### 2.2.3 OnTrack Diabetes

Esta aplicação permite registar glicose, refeições, exercício, medicação, peso, pressão arterial, pulsação e HbA1c. Tem uma interface bastante simples relativamente às outras aplicações experimentadas. Tem apenas três menus no ecrã principal, que permite ver relatórios, o histórico

e alguns gráficos relativamente aos dados inseridos. O ecrã principal mostra também as médias dos níveis de glicose diários, semanais e mensais. Ao explorar a aplicação foi possível verificar que esta oferece vários gráficos. Por exemplo, é possível visualizar, através de gráficos, valores de glicose, média diária de glicose, glicose por hora do dia, exercício, etc.

Ao consultar o menu "Histórico" os dados aparecem na forma de lista e por ordem de refeição, ou seja, para um mesmo dia, os dados relativamente ao pequeno almoço aparecem antes do jantar. Este menu apresenta, portanto, todos os dados registados em cada dia. No menu "Relatórios", podemos observar médias de glicose, que são diárias, semanais, mensais ou trimestrais. Existe uma outra opção chamada "glicose por categoria", que mostra os valores médios da glicose registados em cada tipo de refeição. Uma outra funcionalidade, "Logbook", permite a visualização dos dados através de gráficos, permitindo ver qualquer parâmetro registado e partilhar esses mesmos gráficos por *e-mail*.

É possível exportar os dados para csv, xml ou html. É também possível criar *backup* ou apagar todos os dados num determinado intervalo de tempo.

#### 2.2.4 Diabetes - Diário Glucose

De todas as aplicações analisadas, esta é a mais simples. É a que menos funções oferece, permitindo registar apenas o peso e a glicose, que é feito no ecrã principal. A aplicação é composta por outros três separadores que permitem visualizar os níveis de glicose em lista e em gráfico. É possível exportar os dados registados para um ficheiro pdf ou partilhar por *e-mail*.

#### 2.2.5 Glucose Buddy: Diabetes Log

Esta aplicação permite registar o tipo de diabetes, peso, altura, pressão arterial, glicose, HbA1c, exercício, refeições e a atividade do registo (refeição, antes de exercício, depois de exercício, etc.).

Pode-se observar os registos de glicose em forma de lista, utilizando o menu "Logs" ou em forma de gráfico usando o menu "Graphs". No gráfico pode-se visualizar apenas o parâmetro da glicose bem como a média de todos os valores registados por dia.

A aplicação oferece ainda um alarme que pode ser ativado para uma determinada hora ou então pode ser coordenado com um evento. Por exemplo, o utilizador pode definir um alarme para 30 minutos depois do almoço, sendo que quando fizer um registo com o tipo de refeição "almoço", ativará o alarme para o tempo definido.

É possível exportar os registos selecionando intervalos pré-estabelecidos pela aplicação e enviar para o *e-mail*.

Como se pode perceber, as aplicações não diferem muito entre si e todas elas oferecem praticamente as mesmas funcionalidades.



## 2.3 Medicina personalizada e data mining na saúde

Nesta secção pretende-se abordar de que forma a área de *Data Mining* pode ser útil para a saúde. Vamos analisar algum trabalho feito na área da saúde utilizando técnicas de *Data Mining*, de uma forma geral, e também o que já foi feito em específico para a diabetes.

Estas técnicas podem ser utilizadas para fins diferentes: fazer aprendizagem analisando dados já existentes para que se possam criar modelos, que por sua vez irão classificar novos dados; encontrar relações entre variáveis e causas; detetar padrões.

Em [tiago.pdf], os autores usaram diferentes algoritmos para tentar prever a sobrevivência ao cancro da mama. Neste caso, define-se por sobrevivência o paciente estar vivo pelo menos 5 anos após o diagnóstico do cancro. Foram usados três algoritmos de classificação diferentes: redes neuronais artificiais, árvores de decisão e regressão logística. Os autores usaram um *data set* já existente e, depois de todo o pré-processamento, como limpeza de dados, obtiveram um *data set* com 17 variáveis (16 variáveis de previsão e 1 variável de classe, isto é, a variável a ser prevista). Gerando modelos através dos três algoritmos utilizados, conseguiram classificar, com alta percentagem de precisão, se um dado paciente teria sobrevivido ou não. Além disso, conseguiram também descobrir quais as variáveis mais importantes para a classificação, e, portanto, atribuir importâncias diferentes a diferentes variáveis. Os diferentes algoritmos conseguiram diferentes precisões: a rede neuronal teve uma precisão de 0.9121; a regressão logística teve uma precisão de 0.8920 e a árvore de decisão teve uma precisão de 0.9362. De notar que estes resultados foram obtidos usando *cross-validation*. *Cross-validation* é um método que divide um *data set* em duas partes: treino e teste. Neste caso, foi usada *10 fold cross-validation* o que significa que o *data set* foi dividido em 10 partes, ou seja, nove partes são usadas para treino e gerar um modelo. Esse modelo vai ser usado na parte restante para classificação e este processo é repetido dez vezes. Em cada repetição, o conjunto de teste é diferente. A precisão obtida nestes testes foram a média das dez repetições.

Em [palaniappan2008.pdf], os autores criaram uma aplicação web para prever o risco de um dado paciente ter doença cardíaca. A partir de um *data set* com 909 registos, com 15 variáveis, usaram três algoritmos diferentes para calcular a probabilidade de um dado paciente ter uma doença cardíaca: Árvores de Decisão, *Naive Bayes* e Redes neuronais. Os registos foram divididos, em igual proporção, num conjunto de treino (455 registos) e conjunto de teste (454 registos). Obteram diferentes precisões para os modelos: *Naive Bayes* foi o modelo com maior precisão, 86.12%, seguido da rede neuronal com 85.68% e Árvores de decisão com 80.4%. Neste estudo, os autores conseguiram também encontrar relações entre variáveis. Por exemplo, conseguiram concluir que a variável "Tipo de dor no peito" é a mais influente relativamente a uma doença cardíaca. Conseguiram também obter algumas regras que ajudam a prever, com alta percentagem de correção, se um dado paciente tem doença cardíaca ou não. Uma das regras geradas foi "Chest Pain Type = 4 and CA = 0 and Exang = 0 and Trest Blood Pressure  $\geq$  146.362 and  $<$  158.036", que diz que 99.61% dos doentes cardíacos cumprem estes requisitos.

Em [associationDiabetes.pdf], os autores aplicaram o algoritmo *apriori* num *data set* com 100 registos de pacientes diabéticos, para tentar gerar regras de associação. Uma das regras geradas é "If diabetes mellitus type = 2 AND special condition = no AND target = good AND unstable diabetes = no THEN regime = 2". Neste caso, regime é a proposta de insulina por dia, sendo que "2" corresponde a duas injeções de insulina mista, com ação curta e intermédia, uma ao pequeno-almoço e uma à tarde.

Finalmente, em [SVMDiabetes.pdf], os autores usaram algoritmos de classificação para gerar um modelo de diagnóstico da diabetes. Neste caso, usaram-se SVM's (*support vector machines* e um *data set* com 56 variáveis que foi dividido em duas partes: 90% para o conjunto de treino e 10% para conjunto de teste. Foi usada *10 fold cross-validation* como método de treino para obter os parâmetros ideais para os modelos. Depois deste processo, a melhor *fold* é escolhida para gerar conjuntos de regras e para ser usada na classificação do conjunto de teste. Contudo, SVM's têm uma natureza *black-box*, são capazes de classificar dados mas não são capazes de explicar o porquê dessa mesma classificação. Isto significa que, usando apenas SVM's, não é possível extrair regras. Face a este inconveniente, os autores decidiram combinar SVM's com outros dois algoritmos: *Random Forests* (RF) e C4.5, um algoritmo para árvores de decisão. A combinação de SVM's com outros algoritmos *white-box* já vem sendo utilizada noutros estudos. [svm1.pdf][svm2.pdf] Neste caso, conseguiram-se gerar regras que ajudam a classificar dados como pertencendo a pacientes diabéticos ou não-diabéticos. Uma das regras geradas é, por exemplo, "If HBA1C > 7.15 and HDL > 1.57 and CHOL > 5.9 and AGE > 77, then diabetic". Outra é "If HBA1C > 7.25, then diabetic". Os dois algoritmos usados, SVM + RF e SVM + C4.5 conseguiram, respetivamente, 89.6% e 86.3% de precisão.

No próximo capítulo vamos abordar a aplicação utilizada nesta dissertação, a *MyDiabetes*.

## Capítulo 3

# MyDiabetes

Neste capítulo vamos analisar com mais detalhe a aplicação utilizada neste projeto de dissertação. A aplicação chama-se *MyDiabetes* e, além de ser utilizada para recolher os dados que posteriormente vão ser utilizados, é também nela que vai ser implementado o sistema definido nos objetivos desta dissertação.

### 3.1 Objetivo da aplicação

Esta aplicação tem como objetivo ajudar o doente diabético, ao disponibilizar uma ferramenta que permita registar e visualizar os vários parâmetros importantes, como glicose, insulinas e hidratos de carbono. A visualização pode ser feita através de gráficos ou de lista. A aplicação facilita o dia-a-dia do diabético pois permite ter um maior controlo sobre os seus registos, através de uma fácil visualização ou navegação para datas anteriores de forma a ver valores num determinado intervalo de tempo.

### 3.2 Arquitetura

[é o que exatamente??]

A aplicação é bastante intuitiva: tem menus de registo, como registo de refeições, insulinas, exercício ou doenças. Esses registos são guardados no *smartphone*, numa base de dados *sqlite*. Tem também um menu para visualização dos registos efetuados, o "Logbook".

### 3.3 Variáveis recolhidas

A aplicação permite recolher dados sobre doentes diabéticos através dos registos que estes vão efetuando. Além de valores de glicose, hidratos de carbono e insulina, como já mencionado

anteriormente, permite também a recolha de outros dados como doença ou exercício. A data e hora de cada registo também são guardadas, isto porque pode ser útil descobrir padrões temporais, associados a um dia da semana ou hora do dia, por exemplo. Permite ainda recolher outras variáveis como pressão arterial, colesterol e peso. Estas variáveis não são tão relevantes, no entanto, por não haver uma relação direta entre estas e os valores de glicose. A recolha destas variáveis é feita através de uma funcionalidade presente na aplicação, que permite ao utilizador mandar a base de dados ao gestor do projeto.

## Capítulo 4

# Análise de dados

Este capítulo foca a análise de dados obtidos. Será feita uma descrição do estudo, desde o processo de recolha de dados até ao tratamento desses mesmos dados para uma análise posterior. Serão também descritos os vários passos efetuados na análise desses dados, bem como os diferentes métodos ou algoritmos utilizados.

### 4.1 Descrição do estudo

O objetivo da dissertação é integrar, na aplicação *MyDiabetes*, um sistema que seja capaz de analisar, em tempo real, os dados inseridos pelo utilizador, para que se possam detetar anormalidades. Essas anormalidades serão apresentadas na forma de regras, que serão obtidas a partir de algoritmos de associação. No entanto, estes algoritmos geram todo o tipo de regras: algumas relevantes e outras não. Para que o sistema saiba quais as regras relevantes e quais as regras descartáveis é preciso saber que tipo de regras vão ser obtidas. Para isso, o primeiro passo seria aplicar as diferentes técnicas a dados já existentes para fazer uma filtragem de regras úteis. Neste caso, o tipo de dados necessários seriam medições contínuas dos vários parâmetros ao longo de algumas semanas, para que se possam detetar padrões temporais. Por exemplo, um padrão deste tipo poderia ser "Um dado utilizador pratica exercício às terças à noite. Às quartas de manhã costuma ter hipoglicemia".

No entanto, os *data sets* disponíveis *online* relativamente a dados de pacientes diabéticos são escassos e não estão no formato pretendido. Por exemplo, um *data set* disponível tem dados relativos a uma tribo indígena norte-americana com algumas variáveis, entre as quais idade, índice de massa corporal e pressão arterial, em que cada registo pertence a um paciente. Este tipo de dados é para classificação de cada paciente como diabético ou não-diabético, e portanto não corresponde ao que se pretende nesta dissertação.

O tipo de *data set* ideal para este projeto seria com dados de vários doentes ao longo do tempo, em vez de dados pontuais. Assim seria possível analisar cada paciente de forma individual, ou seja, aplicar o conceito de medicina personalizada. Face à falta de *data sets* com estas condições,

a solução encontrada foi recolhermos os nossos próprios dados. Para o conseguirmos, contámos com a ajuda do serviço de endocrinologia do Hospital de S. João, através do Dr. Celestino Neves, que serviu de intermediário entre a faculdade e o hospital. O objetivo nesta fase foi de apresentar a aplicação aos pacientes do hospital, mostrando o que a aplicação permite fazer e explicando que, a longo prazo, os pacientes seriam os principais beneficiários. A apresentação da aplicação passava por mostrar as funcionalidades mais importantes e explicar que o objetivo seria fazer com que a aplicação fosse capaz de detetar anormalidades e mostrar alertas ou conselhos para esses casos. Os pacientes eram também informados que os dados por eles registados seriam usados para fins de investigação e, caso aceitassem usar a aplicação de forma voluntária, assinavam um consentimento informado.

Este processo era feito depois de o Dr. Celestino dar a conhecer, de forma resumida, do que se tratava. Desta forma, como era o médico a fazer uma primeira abordagem ao projeto, o paciente sentia-se mais seguro em falar connosco e mais recetivo a um diálogo. Quando o paciente aceitava participar, a aplicação era instalada na hora, para que pudesse começar a usá-la o mais rapidamente possível.

Este processo foi feito ao longo de sensivelmente três meses, duas vezes por semana. Ao longo deste tempo, falámos com várias dezenas de pacientes. Praticamente todos acharam a aplicação bastante interessante e importante e o *feedback* foi bastante positivo. Apesar disto, nem todos os pacientes com quem falámos participaram no estudo, por variadas razões: ou impedimentos técnicos (não ter um *smartphone* Android, sendo que neste caso podíamos emprestar um telemóvel), ou por não querer ou por não poder (por exemplo, pacientes com idade mais avançada ou com outros problemas de saúde). Ainda assim, conseguimos angariar 31 voluntários, o que seria suficiente para o que se pretendia.

Infelizmente, como vimos a comprovar, a percentagem de utilizadores que de facto usaram a aplicação e enviaram os dados foi bastante pequena: apenas 7 pacientes enviaram dados, sendo que apenas 4 enviaram dados relativamente a mais de uma semana. Foram feitas algumas tentativas para incentivar os outros utilizadores a enviar os dados, como mostrarmos algumas atualizações da aplicação e consequentemente novas funcionalidades, sem sucesso, no entanto. Esta pouca adesão no envio de dados foi o principal problema encontrado para se fazer o projeto nos moldes desejados.

## Capítulo 5

# Conclusões

Teste de citações e sua ordem [ ? ? ? ]

Quisque ullamcorper placerat ipsum. Cras nibh. Morbi vel justo vitae lacus tincidunt ultrices. Lorem ipsum dolor sit amet, consectetur adipiscing elit. In hac habitasse platea dictumst. Integer tempus convallis augue. Etiam facilisis. Nunc elementum fermentum wisi. Aenean placerat. Ut imperdiet, enim sed gravida sollicitudin, felis odio placerat quam, ac pulvinar elit purus eget enim. Nunc vitae tortor. Proin tempus nibh sit amet nisl. Vivamus quis tortor vitae risus porta vehicula.

Fusce mauris. Vestibulum luctus nibh at lectus. Sed bibendum, nulla a faucibus semper, leo velit ultricies tellus, ac venenatis arcu wisi vel nisl. Vestibulum diam. Aliquam pellentesque, augue quis sagittis posuere, turpis lacus congue quam, in hendrerit risus eros eget felis. Maecenas eget erat in sapien mattis porttitor. Vestibulum porttitor. Nulla facilisi. Sed a turpis eu lacus commodo facilisis. Morbi fringilla, wisi in dignissim interdum, justo lectus sagittis dui, et vehicula libero dui cursus dui. Mauris tempor ligula sed lacus. Duis cursus enim ut augue. Cras ac magna. Cras nulla. Nulla egestas. Curabitur a leo. Quisque egestas wisi eget nunc. Nam feugiat lacus vel est. Curabitur consectetur.

Suspendisse vel felis. Ut lorem lorem, interdum eu, tincidunt sit amet, laoreet vitae, arcu. Aenean faucibus pede eu ante. Praesent enim elit, rutrum at, molestie non, nonummy vel, nisl. Ut lectus eros, malesuada sit amet, fermentum eu, sodales cursus, magna. Donec eu purus. Quisque vehicula, urna sed ultricies auctor, pede lorem egestas dui, et convallis elit erat sed nulla. Donec luctus. Curabitur et nunc. Aliquam dolor odio, commodo pretium, ultricies non, pharetra in, velit. Integer arcu est, nonummy in, fermentum faucibus, egestas vel, odio.

### 5.1 Trabalho Futuro

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget,

consectetuer id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

Nam dui ligula, fringilla a, euismod sodales, sollicitudin vel, wisi. Morbi auctor lorem non justo. Nam lacus libero, pretium at, lobortis vitae, ultricies et, tellus. Donec aliquet, tortor sed accumsan bibendum, erat ligula aliquet magna, vitae ornare odio metus a mi. Morbi ac orci et nisl hendrerit mollis. Suspendisse ut massa. Cras nec ante. Pellentesque a nulla. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Aliquam tincidunt urna. Nulla ullamcorper vestibulum turpis. Pellentesque cursus luctus mauris.

Nulla malesuada porttitor diam. Donec felis erat, congue non, volutpat at, tincidunt tristique, libero. Vivamus viverra fermentum felis. Donec nonummy pellentesque ante. Phasellus adipiscing semper elit. Proin fermentum massa ac quam. Sed diam turpis, molestie vitae, placerat a, molestie nec, leo. Maecenas lacinia. Nam ipsum ligula, eleifend at, accumsan nec, suscipit a, ipsum. Morbi blandit ligula feugiat magna. Nunc eleifend consequat lorem. Sed lacinia nulla vitae enim. Pellentesque tincidunt purus vel magna. Integer non enim. Praesent euismod nunc eu purus. Donec bibendum quam in tellus. Nullam cursus pulvinar lectus. Donec et mi. Nam vulputate metus eu enim. Vestibulum pellentesque felis eu massa.



# Apêndice A

## Acrónimos

**BS** Base Station

**BSN** Body Sensor Network

**HTTP** Hypertext Transfer Protocol

**TCP** Transmission Control Protocol

**UDP** User Datagram Protocol