UNIVERSIDAD DE MURCIA

ESCUELA INTERNACIONAL DE DOCTORADO

Managing Type 1 Diabetes Mellitus
with IoT Devices and
Machine Learning Techniques

Gestión de la Diabetes Mellitus
Tipo 1 con Dispositivos Internet de
las Cosas y Técnicas de Aprendizaje
Automático

D. Ignacio Rodríguez Rodríguez

2020

Universidad de Murcia

Facultad de Informática

Gestión de la Diabetes Mellitus tipo 1 con Dispositivos Internet de las Cosas y técnicas de Aprendizaje Automático

Presentada por:

D. Ignacio Rodríguez Rodríguez

Supervisada por:

Prof. Dr. Miguel Ángel Zamora Izquierdo (UM)

Prof. Dr. José Víctor Rodríguez Rodríguez (UPCT)

Murcia, enero de 2020

University of Murcia

Computer Sciences Faculty

Managing Type 1 Diabetes Mellitus with IoT Devices and Machine Learning Techniques

Author:
Ignacio Rodríguez Rodríguez
Thesis Advisors:
Prof. Dr. Miguel Ángel Zamora Izquierdo (UM)
Prof. Dr. José-Víctor Rodríguez (UPCT)

Murcia, January 2020

La paciencia todo lo alcanza.

Aquello que ha sobrevivido es lo que ha ganado el justo derecho de permanecer.

Agradecimientos

El desarrollo de esta tesis ha supuesto un enorme sacrificio en múltiples aspectos de la vida. Es justo agradecer a aquellos que han sido soporte y han insuflado ánimo en los momentos bajos, y también han sido copartícipes de las alegrías.

Comienzo con mi director Miguel Ángel Zamora, por haber sido paciente y haber confiado en que llegaría a buen fin. Su apoyo y generosidad han sido clave en este trabajo. A mi más que director José Víctor Rodríguez, empuje fundamental sin el cual esta tesis no hubiese sido (o hubiese sido otra cosa). Ésta es su segunda tesis. Gracias a ambos por vuestro tiempo y vuestra guía. También quiero agradecer al profesor Ioannis Chatzigiannakis su buena disposición y ayuda durante mi estancia en La Sapienza.

A quienes han sido mis compañeros, tanto de la Facultad de Informática como del Servicio Público de Empleo Estatal. Han estado en lo cotidiano, apoyando y viviendo el día a día de esta la tesis. Gracias por la flexibilidad y facilidades.

A mi familia y amigos, hermanos y cuñadas, por haber entendido mis ausencias y desvelos. A Pepe especialmente, pues ha vivido la tesis prácticamente desde el principio y sabe y entiende lo que ha sido la construcción de este proyecto.

A mis padres, ambos sabios y doctos en el difícil arte de vivir, amar, trabajar y construir. Les debo lo que soy. Sin su esfuerzo, ejemplo, cariño y apoyo, llegar a este punto hubiese sido una ficción.

Contents

Tables

Figures

Capítulo 1

Resumen

1.1 Motivación

La Diabetes Mellitus (DM) es sin duda, en 2019, uno de los primeros problemas de salud en el mundo, sobre todo en los países más desarrollados. Factores asociados al estilo de vida occidental, tales como el sedentarismo, el sobrepeso, la carencia de actividad física, los malos hábitos alimenticios, así como otros relacionados con la genética o la edad están íntimamente conectados con el desarrollo de esta patología.

Hoy en día, se calcula que la diabetes afecta a entre el 5% y el 10% de la población española, según los estudios consultados. En este sentido, en la última Encuesta Nacional de Salud en España, relativa al año 2017, un 7.8% de toda la población había sido diagnosticada DM. Concretamente, dicha cifra subía al 21% en los mayores de 65 años, mientras que, en el año 1993, tan sólo alcanzó el 4.1% [20].

La DM se caracteriza por una elevación continua de la glucemia en sangre, ya sea por falta de producción endógena de insulina, o por una resistencia a su acción, pudiéndose distinguir la denominada Tipo 1 en el primer supuesto y la Tipo 2 en el segundo (DM1 y DM2, respectivamente). Tanto en uno como en otro caso, las consecuencias de presentar unas cifras anormalmente altas de glucemia en sangre, de manera sostenida en el tiempo, son devastadoras para el organismo.

Así, la DM es una pluripatología, esto es, genera posibles múltiples dolencias relacionadas con la evolución de la enfermedad. En este sentido, cabe destacar que se trata de la primera causa de: retinopatía, ceguera por debajo de los 65 años, trasplante renal, enfermedades cardiovasculares, etc. Al margen del drama humano que produce el padecimiento de estas complicaciones, la DM genera un creciente y enorme gasto sanitario en todo el mundo, especialmente en las sociedades occidentales, debido a las características de éstas, que vienen determinadas por un abuso de alimentos procesados excesivamente calóricos, así como por la falta de actividad física.

El coste sanitario directo medio anual de cada paciente con DM supera en 2.145 euros el coste medio de una persona sin esa patología [21]. Por tanto, la DM impone una carga económica que podría alcanzar el 2,5 % del PIB, con una estimación de 19.908,661 millones de euros en 2015. Tan sólo si se llevaran a cabo ciertos cambios actitudinales en relación con algunos factores de riesgo previamente enumerados, la sociedad podría ahorrar el 64,8 % de ese coste, en torno a 12.900,8 millones de euros (entre 2.428,5 y 17.764,2 millones). De ellos, el mayor ahorro se alcanzaría mediante un mejor control de la dieta, que aparece como responsable del 40 % de los costes sociales incrementales de la enfermedad.

No obstante, las complicaciones anteriormente descritas se pueden contener, retrasar e incluso evitar en su totalidad si se incrementa el nivel de control del paciente diabético. Los estudios Diabetes Control Complications trial (DCCT) [22] y el United Kingdom Prevention Diabetes Study (UKPDS) [23], para DM1 y DM2 respectivamente, demuestran que un buen control metabólico es básico para la prevención de dolencias asociadas al curso de la diabetes.

Además, en dichos estudios se prueba que la mejoría en el control de la glucemia es clara si el paciente planifica:

- Alimentación, según las necesidades de la persona.
- Ejercicio físico, de nuevo según las características concretas del paciente, edad y estado de salud general.
- Medicación (insulina, hipoglucemiantes), siguiendo las indicaciones del médico endocrino.
- Hábitos saludables, con una adecuada higiene del sueño y control del estrés.
- Controles periódicos, a través de las distintas opciones de medidores de glucosa existentes y por medio de revisiones médicas de carácter periódico.

De alguna forma, estas indicaciones ya están apuntando a las variables que es necesario considerar a la hora de gestionar la evolución de la DM. El paciente diabético debe tener en mente toda esa información, la cual procesa subjetivamente, y con ello decidir ajustes en sus rutinas y en su medicación para intentar controlar el curso de su glucemia, siempre con la colaboración y el debido asesoramiento de su profesional sanitario. Este control resulta singularmente complejo en el caso de la DM1 ya que, en este caso, el páncreas deja de producir insulina en las denominadas células beta de unas estructuras llamadas islotes de Langerhans. Dicho tipo de diabetes se caracteriza por desarrollarse en la infancia o adolescencia, y su origen se considera de tipo autoinmune. La insulina es la hormona necesaria para introducir la glucosa en sangre en las células que demandan energía. Así, el paciente tiene que inyectársela de forma exógena, fijando las dosis según recomendaciones médicas y corrigiéndolas, en su caso, según su propia experiencia y, tradicionalmente, a partir de mediciones puntuales de la glucosa en sangre por medio de tiras reactivas. De esta manera, la DM1 es el tipo de DM de evolución más agresiva que hay debido a sus oscilaciones más pronunciadas de glucemia, la ausencia de insulina endógena que pueda amortiguar dichas fluctuaciones, y las consecuencias para la salud que pueden darse en los y las pacientes que la sufren, a corto plazo y a largo plazo.

En la DM1, la interpretación subjetiva del contexto por parte del individuo hace que, en ocasiones, se puedan producir errores en el tratamiento debido a esa apreciación parcial. No obstante, nuevos dispositivos han irrumpido entre las opciones de gestión de la DM. La entrada de los Medidores Continuos de Glucosa (CGM, *Continuos Glucose*

Monitoring) ha supuesto una revolución en la gestión de la diabetes. Estos instrumentos permiten una medición continua de la glucemia presente en el líquido intersticial subcutáneo, con una frecuencia de muestreo normalmente entre una por minuto o cada cinco minutos de forma que, por medio de algoritmos de corrección, muestran los valores de glucosa en sangre. De esta forma, el conocimiento que tiene el paciente de su glucemia es ininterrumpido. En la Figura 1, se puede ver uno de los medidores más populares, el Freestyle Libre de Abbot[1], el cual se sirve de un sistema de monitorización por Near Field Communication (NFC) de manera que muestra la glucemia bajo demanda del usuario, siendo debido a esta característica por lo que se denomina Flash Glucose Monitoring (FGM).

Figura 1. CGM Freestyle Libre

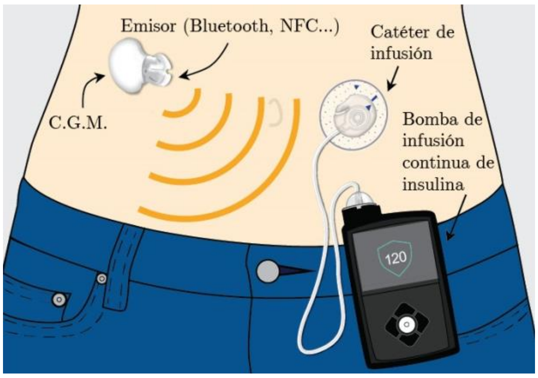

Figura 2. Esquema CGM - Bomba

Unido a esto, el paciente de DM1 dispone de otros equipos que pueden ayudar a gestionar su enfermedad. Las bombas de insulina permiten al paciente administrar de una forma más cómoda y exacta la dosificación de ésta, por medio de una infusión continua que, a través de un catéter, se absorbe bajo la piel (Figura 2).

La utilización conjunta de sensores continuos y bombas es un hecho, en parte gracias a la mejora de las comunicaciones inalámbricas de corto alcance, tales como Bluetooth Low Energy, NFC (ya nombrado), o incluso las redes WiFi. Así, se podrá generar una red personal (Body Area Network, BAN), donde se gestionará la información. El uso de teléfonos inteligentes, cada vez con más capacidad de conexión y potencia de cálculo, aporta todo tipo de posibilidades en este aspecto, haciendo uso del potencial que el Internet of Things (IoT) está demostrando tener en los últimos años. Este nuevo paradigma indica un prometedor camino investigador a seguir.

Por otro lado, se ha de resaltar que, hasta ahora, la gestión que se hace aprovechando estas comunicaciones indicadas en el párrafo anterior se realiza teniendo en cuenta únicamente la evolución de la glucemia, aportada por el medidor continuo. Si, por un lado, parece evidente que ésta es la principal variable a considerar, es también sabido que existen numerosas circunstancias que afectan a la evolución glucémica de un paciente de DM1. La alimentación, el ejercicio físico, la rutina diaria, el sueño, etc. se han demostrado como influyentes a la hora de afectar a los niveles de glucosa en sangre. Si bien hace unos años estas variables se obviaban, debido a la dificultad de obtener estos niveles como una señal continua (y, posteriormente, gestionarla), la aparición de nuevos dispositivos biométricos portátiles, capaces de registrar 24 horas al día medidas como el

[1] https://www.freestylelibre.co.uk/libre/

ritmo cardiaco, actividad física y muchas otras, hace que tenga sentido registrar dichas variables y aproximarse al estudio de su relevancia en la evolución glucémica. Con la inclusión de estas variables descritas anteriormente, se puede mejorar la predicción de la glucemia que puede hacer un determinado algoritmo de Machine Learning (ML), incrementando su precisión y pudiendo así, no ya sólo interrumpir la infusión de insulina en episodios de hipoglucemia, sino también anticiparse a situaciones de hiperglucemia y aumentar la dosis cuando se requiera. De esta manera, la mejora en el control glucémico del paciente de DM1 se haría más que evidente.

Sin embargo, la inclusión de más variables que caractericen al paciente con diabetes conlleva sin duda un incremento de las necesidades de *hardware* computacional para gestionar dicha información. En este sentido, las limitaciones de elementos portátiles como un *smartphone* o una *Raspberry Pi* pueden suponer un problema, por lo que es necesario ajustar la cantidad de datos a utilizar, consiguiendo un compromiso entre rapidez de ejecución de los algoritmos y precisión de la solución obtenida. Recurrir a oportunidades como el Cloud Computing puede ser una estrategia adecuada para solventar esta limitación, pero también se debe estudiar hasta qué punto un pequeño dispositivo puede aportar una solución "en local", incluso construyendo el modelo "en directo" bajo determinadas circunstancias que hagan este hecho necesario (falta de conexión a internet o inestabilidad de la misma, áreas rurales remotas, o circunstancias en las que los dispositivos tienen que estar fuera de línea: aviones, conferencias, etc.).

Con todo lo expuesto anteriormente se concluye como principal motivación de la presente tesis doctoral el estudio de posibles soluciones tecnológicas que faciliten una mejor gestión de la información para el control de la DM1, enfocadas a conseguir una predicción de la glucemia futura más fiable, rápida y computacionalmente asequible. Esto se ha concretado en una serie de objetivos, los cuales se expresan a continuación.

1.2 Objetivos

Una vez presentada la motivación de la presente tesis doctoral y el tema central de la misma, se enumeran los objetivos a satisfacer y que han servido de guía para el desarrollo de este trabajo:

1. Estudiar las posibilidades tecnológicas actuales con respecto a su aplicación en el tratamiento de la DM1. Para ello, se debe realizar con anterioridad una revisión completa de la literatura.
2. Aplicar el concepto de IoT. Las posibilidades de interconectividad y ubicuidad que promete este nuevo paradigma hace que su aplicación a la DM1 sea idónea.
3. Incrementar la información disponible. Hasta ahora, la caracterización del paciente diabético se ha realizado por medio de mediciones puntuales de su glucemia, la cual expresaba con un único valor el estado del paciente en ese momento.
4. Adecuar el tratamiento de la información. Añadido a lo anterior, y como consecuencia de esto, será necesario adecuar el tratamiento de los datos, incrementados por el objetivo anterior.
5. Incrementar el conjunto de datos disponibles. Se identifica la necesidad de disponer de un adecuado y confiable conjunto de datos para realizar la

experimentación; por ello, se plantea como objetivo realizar una campaña de recogida de datos entre pacientes con diabetes tipo 1 que cumplan unos requisitos mínimos de confiabilidad.

6. Efectuar un tratamiento de la naturaleza y transmisión de series de datos temporales. La información recogida va a tener una característica principal: se trata de datos unidos a una marca de tiempo, conformando series temporales que, a su vez, pueden ser transmitidos de manera inalámbrica hacia un nuevo receptor.
7. Seleccionar las variables más relevantes en la evolución de la glucemia. Se puede establecer una clasificación según la influencia en las variaciones de los valores de glucemia en sangre.
8. Realizar predicción de valores futuros de series de datos temporales. Con las características del paciente diabético conformadas y el conjunto de datos pre-procesado, se van a sopesar las posibilidades que los diferentes algoritmos de aprendizaje automático (ML) ofrecen en el objetivo de conseguir una predicción de valores de glucosa en sangre a un cierto horizonte de predicción, de forma que permita al paciente anticiparse en la gestión de sus valores de glucemia.
9. Conocer los límites de esfuerzo de los dispositivos a utilizar en la gestión de la DM1. La gestión de información y la ejecución de algoritmos de predicción son tareas de un cierto grado de exigencia que plantean dudas sobre la capacidad y solvencia de unos dispositivos u otros. Se harán pruebas comparativas para valorar esta cuestión.
10. Exponer claramente ideas y conceptos. Como objetivo transversal se pretende incrementar las competencias en cuestión de exposición de resultados, desarrollo de propuestas y concreción en las conclusiones (incluyendo la redacción de los artículos científicos publicados y de esta misma memoria, así como la exposición oral de los trabajos presentados en congresos). Además, se ha intentado que, en la medida de lo posible y dentro de los foros apropiados, cumplan también con el propósito de incrementar la divulgación científica de una forma clara, con una visión pedagógica.

1.3 Resultados

En el marco de la presente tesis doctoral, se han realizado numerosas contribuciones, todas ellas recogidas en artículos científicos o comunicaciones escritas y orales a congresos. No se ha recogido su totalidad en el compendio de publicaciones que compone este documento por estar algunas enfocadas en cuestiones que, aun conexas, escapan de la unidad científica que se presenta en este texto. No obstante, una lista exhaustiva de trabajos realizados durante el periodo temporal de desarrollo de esta tesis se presenta al final de este documento (Bibliografía, Publicaciones).

La mayoría de trabajos desarrollados se encuentran relacionados con la gestión de la diabetes mellitus; no obstante, otros se han centrados en gestión de la información en otros contextos, o en aplicaciones de las Tecnologías de la Información y la Comunicación (TIC) a otros ámbitos, aprovechando el conocimiento ganado en este sentido con la DM1. En todo caso, la experiencia adquirida en la elaboración de unos trabajos redunda, indudablemente, en la mejor ejecución de los siguientes.

Con el fin de centrarse en los retos de gestión tecnológica que ofrece la DM1, se estudiaron las posibilidades que el IoT ofrece a la gestión remota de situaciones [1] [2]. Así, el paradigma del IoT permite aportar servicios centrados en el usuario, con una monitorización continua de la persona.

La aplicación del anterior concepto ha permitido proponer una plataforma de gestión inteligente de la DM1, integrando en este entorno IoT canales de información que dirigen los datos cosechados de la persona bien a ella misma, bien a profesionales sanitarios o familiares, de forma que conozcan en todo momento el estado integral del paciente [3].

Una correcta caracterización del paciente se consigue por medio de dispositivos inteligentes que indicarán en qué situación se encuentra: *smartphones*, *smartbands* y otros sensores portátiles [4]. Ésta es la idea principal expuesta en [5], donde, además, se propone un tratamiento novedoso para las variables que se consideran en la DM1. Ya se había apuntado en otros trabajos de otros autores la acción prolongada, pero de menor intensidad, que presenta la insulina más allá de su acción a corto plazo, llamada en su momento "*insulin on-board*". En la presente tesis doctoral se ha propuesto la extensión de dicho concepto a otras variables recogidas, como el ejercicio, el sueño, o la alimentación, entendiéndolo como una acción remanente de muy baja intensidad perdurable en el tiempo a lo largo de varias horas, con una acción puntual normalmente despreciable, pero con efectos aditivos que se deben tener en cuenta ante la concurrencia de múltiples dosis de insulina en un periodo de tiempo, un ejercicio de mayor intensidad, o una carencia de descanso.

Igualmente, la idea de una completa caracterización de una persona se ha extrapolado a otros ámbitos, de forma que, en situaciones de violencia o agresión, se puedan conocer datos como la posición GPS, estado de salud, etc., con el fin de intervenir y salvaguardar la seguridad de la víctima [6].

Igualmente, se ha realizado una novedosa toma de datos. Así, durante un periodo de 14 días se monitorizó a 25 pacientes con DM1 a través de un dispositivo de monitorización continua de glucosa (CGM), obteniendo, además de su glucemia, sus dosis de insulina y cantidad de alimento ingerido. Añadido a esto, se recogieron otras características como ritmo cardiaco, ejercicio físico, sueño y horarios (a través de una *smartband*). Esto convierte este conjunto de datos en uno de los más relevantes que se pueden encontrar en la literatura, pues los anteriormente utilizados adolecen de una monitorización completa, de un número suficientemente representativo de participantes, una extensión mínimamente aceptable del ensayo o incluso incorporan anotaciones subjetivas, por parte del paciente, en algunas variables [7].

La información recogida de un paciente diabético se caracteriza por ir asociada a una marca de tiempo, resultando el conjunto de muestras tomadas en series de datos temporales. Por sus características concretas, este tipo de información debe ser tratada de forma adecuada, y esto se ha abordado, entre otros, especialmente en los trabajos [8] [9]. Así mismo, dicha información recogida por el conjunto de sensores correspondiente puede ser, a su vez, transmitida de manera inalámbrica hacia otros receptores en forma de ondas electromagnéticas. Estas ondas se propagan a lo largo de un entorno en el que puede existir toda una serie de obstáculos que dificulten la comunicación. Por tanto, con el fin de estudiar la atenuación que pueden sufrir estas señales (de cara a asegurar el establecimiento de la conexión), se ha estudiado el fenómeno de la difracción como principal mecanismo de pérdida de energía de las ondas [10] [11] [12].

Todas las variables mencionadas anteriormente influyen de forma conjunta en la glucemia, en sus valores y en su evolución a través del tiempo. Sin embargo, no todas lo hacen en la misma medida, de forma que unas son más influyentes que otras. Para analizar este hecho, existen algoritmos de selección de variables, siendo algunos específicos para datos de series temporales. Dentro de estos, se ha utilizado el *Sequential Input Selection Algorithm* (SISAL), con el cual se ha llegado a una catalogación de la influencia de las variables, así como el tiempo que tardan en resultar influyentes [7]. Esto es determinante para poder realizar una predicción de los valores de glucemia, pues se puede prever qué se debe incluir en un algoritmo predictivo y conformar así un set de variables que no sobrecargue la labor de pronosticar valores futuros. Con este estudio, se ha observado que la variable más influyente es la insulina, debiendo cogerse los 105 últimos minutos de datos. A continuación, se situaría la comida, siendo relevantes las últimas 2.91 horas y, finalmente, la propia glucemia, con unos resultados de algo más de 4 horas. El resto de las variables estudiadas, ejercicio, ritmo cardiaco, sueño y el horario completan el set de variables con ese orden de importancia.

Una vez propuesta una plataforma de gestión y generadas múltiples formas de caracterización de la persona diabética así como una gradación de su influencia, se ha conseguido una predicción de glucemia a distintos horizontes temporales. Este trabajo se ha llevado a cabo, primeramente, de la forma más sencilla posible: teniendo en cuenta únicamente los valores pasados de glucosa en sangre con el fin de predecir los siguientes. Esta aproximación univariante pretende ser el primer paso de toda una serie de estrategias predictivas. En este primer estudio con una única variable, se ha discutido la cantidad de valores pasados que se debe tener en cuenta, así como su incidencia en la precisión de la predicción. De la misma forma, la posibilidad de una menor frecuencia de muestreo (y, consiguientemente, una señal más *ligera* y fácilmente tratable) también se ha discutido, llegando a la conclusión de que es posible un compromiso entre carga de información y precisión [8].

Los algoritmos de predicción juegan aquí un papel fundamental y, en este sentido, se han comparado tres sobradamente conocidos: *AutoRegressive Integrated Moving Average* (ARIMA), *Random Forest* (RF), y *Support Vector Machines* (SVM). Las posibilidades de este último ya se habían estudiado con anterioridad, con prometedores resultados [13]. Los modelos univariantes desarrollados pudieron predecir los valores de glucosa en un horizonte predictivo de 15 minutos con un error medio de tan sólo 15.43 mg/dL, usando únicamente 24 valores pasados recogidos en un periodo de 6 horas. Aumentando la frecuencia de muestreo hasta incluir 72 valores, el error descendió hasta 10.15 mg/dL. Del estudio en cuestión se ha concluido que RF es el que ofrece, en líneas generales, una precisión mayor. Este estudio se realizó durante una estancia predoctoral de 3 meses de duración en la *Sapienza Università di Roma*, en el *Dipartimento di Ingegneria Informatica, Automatica e Gestionale 'Antonio Ruberti'*, bajo la atenta supervisión del profesor Dr. Ioannnis Chatzigiannakis.

Otros algoritmos predictivos se han probado también en el tiempo de desarrollo de esta tesis. Así, igualmente aplicado a predicción de series de datos temporales, se empleó con éxito la novedosa librería PROPHET del software matemático R, demostrando que la adición de la predicción meteorológica podía incrementar la precisión de la predicción del consumo energético de una vivienda [14] [15].

Continuando con el desarrollo de la tesis, se debe reseñar que los algoritmos utilizados pueden llegar a ser muy exigentes desde el punto de vista computacional para realizar la tarea de predicción de glucemia. Tal y como se ha comentado, se planteó evaluar hasta qué punto una predicción se puede realizar *en local*, esto es, en un pequeño dispositivo que puede llevar la persona consigo (por ejemplo un teléfono inteligente). Para ello, se han hecho pruebas de esfuerzo replicando las mismas técnicas de ML aplicadas en [8] en tres dispositivos: por un lado, un potente servidor y, de la misma manera, en un *smartphone* y una *Raspberry Pi*. Los resultados han indicado que determinadas técnicas son más livianas que otras pero, en todo caso, sin unos requerimientos técnicos mínimos, la ejecución en local resultaría inasumible. No obstante, la ejecución del algoritmo SVM sí ha resultado ser factible en un *smartphone* de características medias, y este hecho es aún más claro si se realizan estrategias para aligerar la serie temporal, como disminuir la frecuencia de muestreo. Los resultados indican que es posible modelar y predecir valores futuros de glucemia en un *smartphone* con un horizonte de predicción de 15 minutos y un RMSE de 11.65 mg/dL en sólo 16.15 segundos, muestreando cada 10 minutos las últimas 6 horas y utilizando el algoritmo RF. Con la Raspberry Pi, el esfuerzo computacional se incrementa a 56.49 segundos en las mismas circunstancias, pero se puede mejorar a 34.89 segundos si se emplea SVM, llegando en este caso a un RMSE de 19.90 mg/dL. Por lo tanto, se concluye que una predicción prácticamente en tiempo real de glucemia es posible usando pequeños dispositivos. Los resultados de este estudio se indican en [9].

De manera transversal, se ha pretendido, a lo largo del desarrollo de esta tesis doctoral, adquirir de forma simultánea unas competencias en el ámbito pedagógico-divulgativo, requisitos esenciales para poder transferir el conocimiento a la sociedad. Aunque esta intención se ha pretendido en todos los trabajos anteriormente indicados (con las limitaciones que el registro lingüístico científico impone), de forma especial se ha materializado en una serie de publicaciones de índole didáctica en las que se han abordado cuestiones como el trabajo en laboratorio (medición de la velocidad de la luz [16]), así como otras relativas al estudio del aprendizaje [17], e igualmente en la publicación de manuales de estudio docentes [18] [19]. Esta parte contribuye, indudablemente, a completar la formación investigadora, transversal y divulgativa del doctorando.

En el siguiente cuadro, se han resumido los principales resultados y las publicaciones en las que se han presentado. Se han especificado igualmente otras publicaciones que desarrollan de forma derivada dichos resultados.

Tabla 1. Resultados de la tesis doctoral y publicaciones relacionadas.

Nº.	Resultado	Publicación
1	Estudio de las posibilidades TIC en la gestión de situaciones centradas en el paciente diabético.	[1] [2] [3] [4] Derivado: [6]
2	Posibilidades del IoT como entorno para implementar las oportunidades halladas en el apartado anterior, desarrollando así una plataforma de gestión de la DM1.	[1] [2] [3] Derivado: [6]
3	Incremento de la caracterización de la persona con DM1, llegando a un conocimiento total de su situación en todo momento.	[3] [4] [5] Derivado: [6]
4	Campaña de monitorización de 25 pacientes con DM1 durante 14 días, recogiendo múltiples variables, consiguiendo una completa descripción de su estado.	[7] [8] [9]
5	Adecuación de la información generada en la monitorización de la DM1.	[8] [9] Derivado: [10] [11] [12].
6	Tratamiento de la naturaleza y transmisión de series de datos temporales relacionada con la DM1.	[8] [9] [13] Derivado: [10] [11] [12] [14] [15]
7	Selección de las variables más influyentes en la evolución de la glucemia en pacientes con DM1	[7]
8	Predicción de valores futuros de glucemia en DM1 y comparación de algoritmos de ML.	[8] [13] Derivado: [14] [15]
9	Estudio de los límites de esfuerzo y capacidad computacional de diversos dispositivos portátil comparados con otros situados en la nube.	[9]
10	Incremento de las capacidades didáctico-pedagógicas así como de difusión de resultados científicos	[16] [17] [18] [19]

1.4 Conclusiones y trabajos futuros

Los numerosos avances conseguidos en las TIC y, más concretamente, en el paradigma del IoT presentan un gran potencial en cuanto a la cantidad de información que se puede recoger de una persona. Es por esto que las posibilidades que se generan en el ámbito de la *e-salud* son tremendamente numerosas. La gestión de dicha información para obtener conocimiento se puede realizar recurriendo a otro paradigma, las técnicas de ML y tratamiento automatizado de datos. En este trabajo, se han sondeado las mejoras que la aplicación de estos conceptos puede traer al campo de la gestión de la DM1.

Así, tras un estudio teórico, se ha propuesto un modelo de plataforma IoT en el que se establecen las entradas a considerar, así como las posibilidades tecnológicas para gestionar la DM1 mediante esta recogida de datos [3]. Sin duda, la propuesta alcanzada establece un marco de trabajo en el que se encuadran los siguientes trabajos de esta tesis, así como otros futuros.

De la inclusión de dispositivos de monitorización portátil se derivan nuevas variables que han sido también estudiadas. El concepto novedoso de la existencia de ciertas variables 'on-board' [5], como extensión del ya descrito por otros autores para la insulina, resulta sin duda un aporte en el tratamiento de las variables de series temporales para el estudio de la DM1, pero también para otro tipo de dolencias, pues los resultados monitorizados también se encontrarán asociados a una marca de tiempo, y pueden presentar acciones remanentes análogas a la insulina y a otras variables descritas en esta tesis.

Tras el marco expuesto, se ha realizado una campaña de monitorización de 25 personas con diabetes tipo 1, lo cual ha supuesto un desafío por el número de participantes durante 14 días, así como por la gestión de los dispositivos utilizados y el procesamiento de los datos. Esta toma de datos ha sido un logro en sí misma.

Con todos los datos recogidos, se han categorizado las variables que se habían descrito anteriormente. A esta categorización, se ha sumado un establecimiento de un *ranking* de importancia, indicando cuáles son aquellas que más influyen en la evolución de la glucemia. Una vez adecuadas las señales de las distintas variables, se plantea la necesidad de estudiar si las nuevas señales monitorizadas son realmente relevantes para caracterizar el estado de un paciente diabético. Para ello, se ha aplicado un método de selección de variables específico para estructuras de datos temporales denominado *Sequential Input Selection Algorithm* (SISAL) [7]. Los resultados de este trabajo han indicado que, efectivamente, se puede establecer una gradación de influencia, considerando niveles de importancia entre variables.

Otro aporte de la presente tesis se ha llevado a cabo en la tarea de predecir valores futuros de glucemia en pacientes DM1. Se ha establecido una comparativa en igualdad de condiciones para tres conocidos algoritmos de ML utilizando sólo valores pasados de glucemia, abriendo una provechosa discusión sobre las circunstancias más influyentes en la precisión: cantidad de datos pasados recogidos, frecuencia de muestreo y horizonte predictivo [8]. Una comparación como ésta no se ha encontrado publicada con anterioridad en la literatura científica.

Para finalizar, ha sido también relevante la evaluación de manera práctica y real de la posibilidad de implementar la anterior predicción en un dispositivo, planteándose dos opciones: cálculo en un servidor remoto (nube), o en dispositivo local (*smartphone*, *Rasperry Pi*) [9]. Los resultados son de gran importancia para llevar a la práctica un control real y portable, por parte del paciente, de la DM1.

No obstante, todos los anteriores logros han planteado más interrogantes que han escapado de los límites de esta tesis, quedando a la espera de ser desarrollados en trabajos futuros, tal y como se indica a continuación.

Primeramente, es de reseñar que se han recogido más características de las personas participantes en nuestra fase de estudio y, aunque se ha trabajado con estos datos para categorizar su influencia, no se han utilizado para la predicción de glucemia, pues, hasta el momento, se han trabajado algoritmos univariantes. Como primer trabajo futuro, se prevé incluir otras variables para comprobar si la precisión se ve mejorada, sin dejar de lado el análisis correspondiente del esfuerzo de cálculo que esto conlleva. Para ello, además de los algoritmos ya estudiados, se realizarán intentos con otros ya empleados en otros trabajos, como es la librería del software matemático R denominada PROPHET [14] [15].

Con los datos recogidos, y ampliándolos de forma que se muestree a más pacientes, se podrían enfocar otras cuestiones. Relacionado con el concepto de 'on-board' extendido al ejercicio físico, se han referenciado numerosas experiencias de diabéticos que reportan hipoglucemias más allá de una ventana temporal inmediatamente posterior a la realización de actividad física. Se pretende aislar estas situaciones del conjunto de datos para estudiar con más detalle este fenómeno.

Igualmente, centrándose en la detección de hipoglucemias, se tiene como objetivo futuro estudiar la relación ya conocida entre la ocurrencia de éstas y las variaciones de ritmo cardiaco asociadas, así como los años de evolución del paciente y la calidad de su control, pues es conocido que una mayor cantidad de hipoglucemias conlleva la desaparición paulatina de los síntomas asociados. De esta manera, se puede estudiar estas variables para comprobar en qué momento y bajo qué control (bueno/malo) desaparecen los síntomas de bajadas de glucemia en el ritmo cardiaco.

Del mismo modo, también se ha referido la influencia de la falta de sueño en una mayor resistencia a la insulina durante el día siguiente, lo que resulta en valores más altos de glucemia. Del mismo modo que el concepto anterior, se pretende aislar estas situaciones para cuantificar esta relación. De hecho, yendo más allá, es factible establecer medias diarias por horas de valores de glucemia y comprobar las horas de sueño de la noche anterior.

Todas estas cuestiones derivan en un mejor entendimiento de la DM1 gracias a profundizar más en las posibilidades que un entorno IoT y la aplicación de algoritmos de ML nos ofrecen. Con todo esto, se plantea realimentar la señal de predicción de glucemia de forma que entre en un controlador y cierre el lazo de control. Un sistema de control en lazo cerrado, con una entrada continua de las variables de monitorización, una salida de predicción de glucosa y, a su vez, una entrada de realimentación que modifica la cantidad de insulina a inyectar por la bomba de infusión correspondiente constituiría, sin duda, un páncreas artificial que supondría, a todas luces, la —esperada por millones de personas— "cura tecnológica" de la DM1.

1.5 Organización de la Tesis

La presente tesis se encuadra bajo el esquema de compendio de publicaciones. Comienza con un resumen que incluye una breve descripción de la motivación de la tesis, así como los objetivos y las conclusiones finales de la misma, exponiéndose dicho resumen en castellano y en inglés. Corresponden a esta finalidad los capítulos primero y segundo.

En el tercer capítulo, se presenta el hilo conductor de las publicaciones que conforman el compendio y la relación entre ambas. Así, en esa parte de la tesis se presentan la aplicación del paradigma IoT y técnicas de ML a la gestión de la DM1. Se hace, pues, un recorrido a través de las investigaciones que se han desarrollado. En este caso, la tesis comienza con la propuesta de una plataforma IoT para gestionar la enfermedad, innovando con la inclusión de nuevos dispositivos para monitorizar de forma continua al paciente diabético. De esta novedad, se deriva la generación de nuevas variables y el adecuado tratamiento que se debe dar a estas series de datos temporales y, dentro de este proceso, se realiza una selección de estas variables que permita elaborar un *ranking* de importancia, así como indicar la cantidad de datos que es relevante en cada momento para una futura predicción de la glucemia. Esto último es el tema principal de la continuación del trabajo, donde se pretende comparar tres conocidos algoritmos predictivos de ML. Finalmente, termina el capítulo tercero estudiando las posibilidades de realizar dicha tarea predictiva de dos formas principales: en la nube o en pequeños dispositivos portátiles.

A continuación, el capítulo 4 recoge los cinco artículos que componen el compendio de esta tesis. El artículo presentado en la sección 4.1, con título “Towards an ICT-Based Platform for Type 1 Diabetes Mellitus Management” [3] analiza las posibilidades que el IoT aporta a la gestión de la DM1, sacando provecho de dispositivos portátiles de monitorización, vías de comunicación inalámbricas, y la capacidad de recogida de información que presentan los *smartphones*. Igualmente, se apuntan sus posibilidades para hacer, con limitaciones, una computación de los datos recogidos o dirigirla hacia la nube. El artículo continúa exponiendo las ventajas de una gestión en un entorno IoT desde el punto de vista de los cuidados y la gestión de situaciones de emergencia. Concluye el trabajo con una propuesta de plataforma integral para manejar la enfermedad.

En la sección 4.2, se incluye el artículo denominado “Variables to be Monitored via Biomedical Sensors for Complete Type 1 Diabetes Mellitus Management: An Extension of the ‘On-Board’ Concept” [5], que deriva directamente del anterior. En éste, se exponen las nuevas variables que se registran por medio de los dispositivos de monitorización portátiles IoT y el tratamiento que se debe dar a esas variables. A partir de un exhaustivo trabajo de documentación, se propone un set de variables que caracterizará de forma integral al paciente con diabetes. Como novedad, se plantea la necesidad de un múltiple tratamiento de cada variable, según sea una variable de tipo pulso, de tendencia, o acumulada (‘*on-board*’), extendiendo este concepto ya aplicado a la insulina al resto de variables.

Seguidamente, el apartado 4.3 contiene el artículo titulado “Feature Selection for Blood Glucose Level Prediction in Type 1 Diabetes Mellitus by Using the Sequential Input Selection Algorithm (SISAL)” [7], donde se estudian los respectivos aportes de las distintas variables recogidas en la fase de experimentación a la evolución de la glucemia. El método de selección de variables (SISAL) es específico para series de datos temporales y arroja una clasificación en orden de importancia de aquellas series que tienen una mayor influencia, así como cuántos datos pasados se deben tener en cuenta.

El siguiente punto 4.4, comprende el artículo "Utility of Big Data in Predicting Short-term Blood Glucose Levels in Type 1 Diabetes Mellitus through Machine Learning Techniques" [8], en el cual se comprueba la capacidad de predicción que conocidos algoritmos de ML tienen para anticipar valores de glucemia utilizando únicamente los datos pasados de glucosa en sangre. Además de comparar tres métodos, se barajan posibilidades para aligerar la información a tratar a través de reducir el número de tiempo pasado a tener en cuenta así como la cantidad de datos mediante la variación de la frecuencia de muestreo.

La 4.5 es la última sección del capítulo, presentando el artículo "On the Possibility of Predicting Glycaemia 'On The Fly' with Constrained Devices in DM1 Patients" [9], donde se estudian las posibilidades de realizar la predicción de glucemia mencionada en la anterior sección utilizando para ello dispositivos portátiles, a saber: un *smartphone* o una *Raspberry Pi 3 b+*, comparando el esfuerzo computacional con el que puede derivarse de utilizar un servidor o en un ordenador personal. Como conclusiones, se puede extraer que determinados algoritmos son más ligeros que otros, siendo asumibles en determinados dispositivos y en otros de menor potencia únicamente tras reducir la cantidad de datos. Bajo determinadas circunstancias, las exigencias computacionales son tal que en un dispositivo de bajo coste serían imposibles de plantear, pero una predicción a corto plazo con exigencias medias computacionales sí resulta abordable por un *smartphone*.

El capítulo 5 recoge las cartas de aceptación de los artículos que forman el compendio, siguiendo el capítulo 6, donde se expone la aprobación de la fase de experimentación por la Comisión de Ética de la Investigación de la Universidad de Murcia. Finalmente, el capítulo 7 incluye la bibliografía de este documento, diferenciada entre los trabajos de autoría del doctorando y otros de diversos autores.

Chapter 2

Summary

2.1 Background

In 2019, Diabetes Mellitus (DM) is undoubtedly one of the leading health problems in the world, especially in more developed countries. Factors associated with the western life, such as a sedentary lifestyle, obesity, lack of physical activity, poor eating habits, and others related to genetics or age are intimately connected with the development of this pathology.

Today, it is estimated that diabetes affects between 5% and 10% of the Spanish population, according to the studies consulted. In this sense, in the last National Health Survey in Spain, for year 2017, 7.8% of the entire population had been diagnosed with DM. Specifically, this figure rose to 21% in those over 65, while in 1993, it only reached 4.1% [20].

DM is characterized by a continuous elevation of blood glucose, either due to a lack of endogenous insulin production, or resistance to its action, referring to the former as Type 1 and the latter as Type 2 (DM1 and DM2, respectively). In both cases, the consequences of presenting abnormally high blood glucose levels, sustained over time, are devastating for the body.

Thus, DM is a multipathology, which means it generates multiple possible ailments related to the evolution of the disease. In this regard, it should be noted that DM is the primary cause of: retinopathy, blindness under 65 years of age, kidney transplantation, cardiovascular diseases, etc. Apart from the human drama suffered from these complications, DM generates a growing and enormous health expenditure worldwide, especially in Western societies, due to its characteristics, which are determined by an abuse of excessively processed foods calories, as well as the lack of physical activity.

The average annual direct health cost of each patient with DM exceeds the average cost of a person without this pathology by 2,145 euros [21]. Therefore, DM imposes an economic burden that could reach 2.5% of GDP, with an estimated 19,908,661 million euros in 2015. If only a few specific attitudinal changes are made in relation to some of the risk factors listed above, the company could save 64.8% of that cost, around 12,900.8 million euros (between 2,428.5 and 17,764.2 million). Of these possible savings, the greatest would be achieved through better diet control, which appears to be responsible for 40% of the incremental social costs of the disease.

However, the complications described above can be contained, delayed and even avoided altogether if the diabetic patient's level of control is increased. The Diabetes Control Complications trial (DCCT) [22] and the United Kingdom Prevention Diabetes Study (UKPDS) [23], for DM1 and DM2 respectively, show that good metabolic control is essential for the prevention of ailments associated with the course of diabetes.

In addition, these studies prove that the improvement in glycemic control is clear if the patient plans:

- Food, according to the needs of the person.
- Physical exercise, again according to the specific characteristics of the patient, age and general state of health.
- Medication (insulin, hypoglycemic agents), following the instructions of the endocrine doctor.
- Healthy habits, with adequate sleep and stress control.
- Periodic checks, through the various options of existing glucose meters and through periodic medical checks.

These indications are already highlighting the variables which need to be considered when managing the evolution of the DM. The diabetic patient must keep all this information in mind, which he must process subjectively, and use this information to decide on adjustments to his routines and in his medication to try to control the course of his blood glucose, always with the collaboration and due advice of his healthcare professional. This control is singularly complex in the case of DM1 since, in this case, the pancreas stops producing insulin in the so-called beta cells of structures called the islets of Langerhans. This type of diabetes is characterized by its development in childhood or adolescence, and its origin is considered autoimmune. Insulin is the hormone needed to introduce blood glucose into cells that demand energy. Thus, the patient has to inject it exogenously, establishing the doses according to medical recommendations and correcting them, if appropriate, according to their own experience and, traditionally, from specific measurements of blood glucose with the use of test strips. In this manner, DM1 is the most aggressive type of DM due to its more pronounced glycemic oscillations, the absence of endogenous insulin that can buffer these fluctuations, and the health consequences that can occur in patients who suffer from it, both short term and long term.

In DM1, the subjective interpretation of the context by the individual means that, at times, errors in treatment may occur due to this partial assessment. However, new devices have been developed to provide more DM management options. The introduction of Continuous Glucose Meters (CGM) has led to a revolution in diabetes management. These instruments allow the continuous measurement of the blood glucose present in the subcutaneous interstitial fluid, with a sampling frequency normally between once per minute or once every five minutes such that, by means of correction algorithms, they present the blood glucose values. In this manner, the patient's knowledge of his blood glucose is uninterrupted. Figure 1 presents one of the most popular meters, the Abbot

Freestyle Libre[2], which uses a Near Field Communication (NFC) monitoring system to show the user's glycemia on demand. Because of this feature, it is called Flash Glucose Monitoring (FGM).

Figure 1. CGM Freestyle Libre

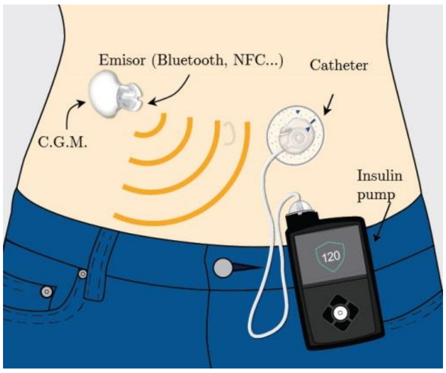

Figure 2. CGM - Pump

Together with this, the DM1 patient has other equipment that can help manage his illness. Insulin pumps allow the patient to administer their dosage more comfortably and accurately, through a continuous infusion which, through a catheter, is absorbed under the skin (Figure 2).

The joint use of continuous sensors and pumps is due, in part, thanks to the improvement of short-range wireless communications, such as Bluetooth Low Energy, NFC (cited above), or even WiFi networks. Thus, a personal network (Body Area Network, BAN) can be generated, where the information will be managed. The use of smartphones, increasingly with connection capacity and computing power, provides all sorts of possibilities in this regard, making use of the potential that the Internet of Things (IoT) has been providing in recent years. This new paradigm indicates a promising research path to follow.

In addition, it should be noted that, until now, the management that takes advantage of the communications indicated in the previous paragraph is carried out only taking into account the evolution of glycemia, provided by the continuous meter. If, on the one hand, it seems clear that this is the main variable to consider, it is also known that there are numerous circumstances that affect the glycemic evolution of a DM1 patient. Food, physical exercise, daily routine, sleep, etc. have all been shown to be influential when it comes to affecting blood glucose levels. Although a few years ago these variables were obviated due to the difficulty of obtaining these levels as a continuous signal (and, subsequently, managing it), with the appearance of new portable biometric devices capable of recording measures such as heart rate 24 hours a day, as well as physical activity and many others, it now makes sense to record these variables and approach the study of their relevance to glycemic evolution. With the inclusion of the variables described above, the glycemia predication made by a certain Machine Learning algorithm (ML) can be improved, increasing its accuracy and thus being able to not only interrupt the infusion of insulin in episodes of hypoglycemia, but also anticipate situations of hyperglycemia and increase the dose when required. In this manner, the improvement in the glycemic control of the DM1 patient would be more than evident.

[2] https://www.freestylelibre.co.uk/libre/

However, the inclusion of more variables that characterize the patient with diabetes undoubtedly entails an increase in the needs of computational *hardware* to manage such information. In this sense, the limitations of portable elements such as a *smartphone* or a *Raspberry Pi* can be a problem, so it is necessary to adjust the amount of data to be used, achieving a compromise between the speed of execution of the algorithms and the accuracy of the solution obtained. Resorting to opportunities such as Cloud Computing can be an appropriate strategy to resolve this limitation, but it should also be studied to what extent a small device can provide a solution "locally", even building the model "live" under certain circumstances that make this a necessity (lack of internet connection or its instability, remote rural areas, or circumstances in which the devices have to be offline: airplanes, conferences, etc.).

With all of the above, the main motivation of this doctoral thesis is the study of possible technological solutions to facilitate better management of the information in order to control DM1, focused on achieving a more reliable, fast and computationally affordable prediction of future blood glucose. This has been specified in a series of objectives, cited below.

2.2 Objectives

Having presented the motivation of this doctoral thesis and its central theme, the objectives to be met are listed, which have served as a guide for the development of this work:

1. Study the current technological possibilities with respect to their application in the treatment of DM1. For this, a complete review of the literature must be carried out beforehand.
2. Apply the concept of the IoT. The possibilities of interconnectivity and ubiquity promised by this new paradigm make its application to DM1 ideal.
3. Increase the available information. Until now, the characterization of the diabetic patient has been carried out by means of specific measurements of his blood glucose, which expresses as a single value the state of the patient at that time.
4. Adapt the treatment of information. Added to the above, and as a consequence, it will be necessary to adapt the processing of the data, increased by the previous objective.
5. Increase the set of available data. The need to have an adequate and reliable set of data to carry out the experimentation is identified. Therefore, the objective is to carry out a campaign to collect data among patients with type 1 diabetes who meet the minimum requirements for reliability.
6. Carry out a treatment of the nature and transmission of temporary data series. The information collected will have a main characteristic: this data is linked to a time stamp, forming time series which, in turn, can be transmitted wirelessly to a new receiver.
7. Select the most relevant variables in the evolution of blood glucose. A classification can be established according to the influence on the variations of blood glucose values.

8. Make a prediction of future values of time series data. Having formed the characteristics of the diabetic patient and the pre-processed data set, the possibilities that different machine learning algorithms (ML) offer in order to achieve a prediction of blood glucose values at a certain horizon are going to be weighed by prediction, in a manner that allows the patient to anticipate the management of their blood glucose values.
9. Learn the stress limits of the devices to be used in the management of the DM1. Information management and the execution of prediction algorithms are tasks with a certain degree of demand which raises doubts about the capacity and solvency of some devices or others. Comparative tests will be done to assess this issue.
10. Clearly present ideas and concepts. As a cross-cutting objective, it is intended to increase the competences in terms of presentation of results, development of proposals and conclusions (including the writing of the scientific articles published and of this report, as well as the oral presentation of the works presented in conferences) . In addition, the work attempts, as far as possible and within the appropriate forums, to also fulfill the purpose of increasing scientific dissemination in a clear manner, with a pedagogical vision.

2.3 Results

Within the framework of this doctoral thesis, numerous contributions have been made, all of them included in scientific articles or presented through written and oral communications to conferences. Not all have been fully collected in the compendium of publications that make up this document because some are focused on issues which, even if related, escape the scientific unit presented in this text. However, an exhaustive list of works carried out during the temporary period of development of this thesis is presented at the end of this document (Bibliography, Publications).

The majority of works developed are related to the management of Diabetes Mellitus; nevertheless, others have focused on information management in other contexts, or on Information and Communication Technologies (ICT) in other areas, taking advantage of the knowledge gained in this regard with DM1. In any case, the experience acquired in the elaboration of some works undoubtedly results in the best execution of the following.

In order to focus on the technological management challenges presented by DM1, the possibilities offered by the IoT in regard to remote situation management [1] [2] were studied. Thus, the IoT paradigm makes it possible to provide user-centered services, with continuous monitoring of the person.

The application of the previous concept has allowed us to propose an intelligent management platform for DM1, integrating in this IoT environment various channels of information that direct the data collected from the person either to themselves, or to health professionals or family, so that they know at all times the comprehensive state of the patient [3].

A proper characterization of the patient is achieved by means of intelligent devices that will indicate the situation: *smartphones*, *smartbands* and other portable sensors [4]. This is the main idea set out in [5], where, in addition, a novel treatment is proposed for the variables considered in DM1. Works of other authors have already noted the prolonged action, but of less intensity, presented by insulin beyond its short-term action, referred to at the time as "*insulin on-board*". This doctoral thesis proposes the extension of this concept to other collected variables, such as exercise, sleep, or food, understood as a remnant action of very low intensity lasting over a time of several hours, with a normally negligible action in time, but with additive effects that must be taken into account in the presence of multiple doses of insulin over a period of time, exercise of greater intensity, or a lack of rest.

Likewise, the idea of a complete characterization of a person has been extrapolated to other areas, so that, in situations of violence or aggression, data such as GPS position, health status, etc., can be known in order to intervene and safeguard the safety of the victim [6].

Likewise, novel data collection has been carried out. Thus, over a period of 14 days, 25 patients with DM1 were monitored through a continuous glucose monitoring device (CGM), obtaining, in addition to their blood glucose, their insulin doses and amount of food ingested. Added to this, other characteristics such as heart rate, physical exercise, sleep and schedules were collected (through a *smartband*). This makes this set of data one of the most relevant that can be found in the literature, since those used previously suffer from complete monitoring, a sufficiently representative number of participants, a minimally acceptable extension of the essay or even incorporate subjective annotations by the patient in some variables [7].

The information collected from a diabetic patient is characterized by being associated with a time stamp, resulting in the set of samples taken in time series. Due to its specific characteristics, this type of information must be treated appropriately, and this has been especially addressed, among others, in the works [8] [9]. Likewise, said information collected by the corresponding set of sensors can, in turn, be transmitted wirelessly to other receivers in the form of electromagnetic waves. These waves propagate throughout an environment in which there can be a whole series of obstacles that make communication difficult. Therefore, in order to study the attenuation that these signals may suffer (in order to ensure the establishment of the connection), the phenomenon of diffraction has been studied as the main mechanism of wave energy loss [10] [11] [12].

All the aforementioned variables jointly influence glycemia, its values and its evolution over time. However, not all do so to the same extent, as some are more influential than others. To analyze this fact, there are algorithms for selecting variables, some being specific for time series data. Of these, the *Sequential Input Selection Algorithm* (SISAL) has been selected, which has made it possible to catalogue the influence of the variables, as well as the time they take to be influential [7]. This is decisive in being able to make a prediction of blood glucose values, since it is possible to predict what should be included in a predictive algorithm and thus form a set of variables is not overloaded with the work of forecasting future values. With this study, it has been observed that the most influential variable is insulin, and the last 105 minutes of data should be taken. The next most-influential variable would be food, with the last 2.91 hours being relevant and, finally, the blood glucose itself, with results from just over 4 hours. The rest of the variables studied (exercise, heart rate, sleep and schedule) complete the set of variables in that order of importance.

Once a management platform has been proposed and multiple forms of characterization of the diabetic person have been generated, as well as a gradation of their influence, a blood glucose prediction can be generated at different time horizons. This work has been carried out, first of all, in the simplest way possible: taking into account only the past blood glucose values in order to predict the following. This univariate approach aims to be the first step in a whole series of predictive strategies. In this first study with a single variable, the amount of past values that should be taken into account, as well as their influence on the accuracy of the prediction, has been discussed. In the same manner, the possibility of a lower sampling frequency (and, consequently, a *lighter* and easily treatable signal) has also been discussed, concluding that a compromise between the information load and accuracy is possible [8].

The prediction algorithms play a fundamental role here and, in this sense, three well-known algorithms have been compared: *AutoRegressive Integrated Moving Average* (ARIMA), *Random Forest* (RF), and *Support Vector Machines* (SVM). The possibilities of the latter had already been studied previously, with promising results [13]. The univariate models developed were able to predict glucose values in a 15-minute predictive horizon with an average error of only 15.43 mg/dL, using only 24 past values collected in a 6-hour period. Increasing the sampling frequency to include 72 values, the error decreased to 10.15 mg/dL. The study in question has concluded that RF is the algorithm which provides the greatest accuracy, in general. This study was carried out during a 3-month predoctoral stay at the *Sapienza Università di Roma*, in the *Computer Science, "Antonio Ruberti" Automation and Management Division*, under the close supervision of Professor Dr. Ioannnis Chatzigiannakis.

Other predictive algorithms have also been tested during the development of this thesis. Thus, also applied to the prediction of time series, the novel PROPHET library of the mathematical software R was successfully used, demonstrating that the addition of meteorological prediction could increase the accuracy of the prediction of the energy consumption of a dwelling [14] [15].

Continuing with the development of the thesis, it should be noted that the algorithms used can become very demanding from the computational point of view when performing the task of blood glucose prediction. As mentioned, it was proposed to evaluate the extent to which a prediction can be made *locally*, that is, in a small device that the person can take with them (for example, a smartphone). For this, stress tests have been done replicating the same ML techniques applied in [8] on three devices: on the one hand, a powerful server and, in the same manner, on a *smartphone* and a *Raspberry Pi*. The results have indicated that certain techniques are require less processing than others but, in any case, without minimum technical requirements, local execution would be unattainable. However, the execution of the SVM algorithm has proved to be feasible in a *smartphone* with average characteristics, and this fact is even clearer if strategies are made to lighten the time series, such as reducing the sampling frequency. The results indicate that it is possible to model and predict future blood glucose values on a *smartphone* with a prediction horizon of 15 minutes and an RMSE of 11.65 mg/dL in only 16.15 seconds, sampling the last 6 hours every 10 minutes and using the RF algorithm. With the Raspberry Pi, the computational effort is increased to 56.49 seconds in the same circumstances, but it can be improved to 34.89 seconds if SVM is used, in this case reaching an RMSE of 19.90 mg/dL. Therefore, it is concluded that practically real-time glycemic prediction is possible using small devices. The results of this study are indicated in [9].

In a transversal manner, it has been attempted, throughout the development of this doctoral thesis, to simultaneously acquire competencies in the pedagogical-informative field, essential requirements in transferring knowledge to society. Although this intention has been sought in all the works indicated above (with the limitations imposed by the scientific linguistic registry), in a special way it has materialized in a series of publications of a didactic nature in which issues such as laboratory work have been addressed (measurement of the speed of light [16]), as well as others related to the study of learning [17], and also in the publication of teaching study manuals [18] [19]. This undoubtedly contributed to the completion of the transveral and informative research training of the doctoral student.

The following table summarizes the main results and the publications in which they were presented. Other publications that develop these results in a derivative manner have also been specified.

Table 1. Results of the doctoral thesis and related publications.

No.	Result	Publication
1	Study of the ICT possibilities in the management of situations focused on the diabetic patient.	[1] [2] [3] [4] Derivative: [6]
2	Possibilities of the IoT as an environment for implementing the opportunities found in the previous section, thus developing a DM1 management platform.	[1] [2] [3] Derivative: [6]
3	Increase in the characterization of the person with DM1, reaching a total knowledge of their situation at all times.	[3] [4] [5] Derivative: [6]
4	Campaign to monitor 25 patients with DM1 for 14 days, collecting multiple variables, getting a complete description of their status.	[7] [8] [9]
5	Adequacy of the information generated in the DM1 monitoring.	[8] [9] Derivative: [10] [11] [12].
6	Treatment of the nature and transmission of temporary data series related to DM1.	[8] [9] [13] Derivative: [10] [11] [12] [14] [15]
7	Selection of the most influential variables in the evolution of blood glucose in patients with DM1	[7]
8	Prediction of future blood glucose values in DM1 and comparison of ML algorithms.	[8] [13] Derivative: [14] [15]
9	Study of the stress limits and computational capacity of various portable devices compared to others located in the cloud.	[9]
10	Increase of teaching-pedagogical capacities as well as dissemination of scientific results	[16] [17] [18] [19]

2.4 Conclusions and future work

The numerous advances achieved in ICT and, more specifically, in the IoT paradigm have great potential in terms of the amount of information that can be collected from a person. This is why the possibilities that are generated in the field of *e-health* are tremendously numerous. The management of this information to obtain knowledge can be done using another paradigm, the techniques of ML and automated data processing. In this work, the improvements that the application of these concepts can bring to the DM1 management field have been examined.

Thus, after a theoretical study, an IoT platform model has been proposed in which the inputs to be considered are established, as well as the technological possibilities for managing DM1 through this data collection [3]. Undoubtedly, the proposal reached establishes a framework in which the following works of this thesis are framed, as well as other future work.

New variables which have also been studied are derived from the inclusion of portable monitoring devices. The novel concept of the existence of certain 'on-board' variables [5], as an extension of that already described by other authors for insulin, is undoubtedly a contribution in the treatment of time series variables for the study of DM1, but also for other types of ailments, since the monitored results will also be associated with a timestamp, and may present remaining actions analogous to insulin and other variables described in this thesis.

After the cited framework, a monitoring campaign of 25 people with type 1 diabetes has been carried out, which has been a challenge for the number of participants over 14 days, as well as for the management of the devices used and the data processing. This data collection has been an achievement in itself.

With all the data collected, the variables described above were categorized. With this categorization, an establishment of a *ranking* of importance has been added, indicating which variables are the ones that most influence the evolution of blood glucose. Once the signals of the different variables are adequate, there is a need to study whether the new monitored signals are really relevant to characterize the condition of a diabetic patient. To this end, a specific variable selection method has been applied for temporary data structures called the *Sequential Input Selection Algorithm* (SISAL) [7]. The results of this work have indicated that, indeed, a gradation of influence can be established, considering levels of importance between variables.

Another contribution of this thesis has been carried out in the task of predicting future blood glucose values in DM1 patients. A comparative equality of conditions has been established for three known ML algorithms using only past blood glucose values, opening a helpful discussion on the most influential circumstances in regard to accuracy: the amount of past data collected, the sampling frequency and the predictive horizon [8] . A comparison like this has not been found to be published previously in the scientific literature.

Finally, it has also been relevant to evaluate in a practical and real manner the possibility of implementing the previous prediction on a device, considering two options:

calculation on a remote server (cloud), or on a local device (*smartphone*, *Rasperry Pi*) [9]. The results are of great importance to the implementation of a real and portable control, by the patient, for DM1.

However, all the previous achievements have raised more questions that have escaped the limits of this thesis, waiting to be developed in future work, as indicated below.

First, it is worth noting that more characteristics of the people participating in our study phase have been collected and, although we have worked with these data to categorize their influence, they have not been used for the prediction of glycemia, because, so far, Univariate algorithms have worked well. As the first future work, it is planned to include other variables to check if accuracy is improved, without neglecting the corresponding analysis of the calculation effort that this entails. For this, in addition to the algorithms already studied, attempts will be made with others already employed in other works, such as the library of R mathematical software, called PROPHET [14] [15].

With the data collected, and expanding the data set so that more patients are sampled, other issues could be addressed. Related to the concept of 'on-board', extended to physical exercise, numerous experiences of diabetics who report hypoglycemia beyond a time window immediately following the performance of physical activity have been referenced. It is intended to isolate these situations from the data set to study this phenomenon in more detail.

Likewise, focusing on the detection of hypoglycemias, a future objective is established to study the already known relationship between their occurrence and the associated heart rhythm variations, as well as the years of evolution of the patient and the quality of their control, as it is known that a greater amount of hypoglycemia leads to the gradual disappearance of the associated symptoms. In this manner, these variables can be studied to check at what time and under what control (good/bad) the symptoms of blood glucose drops in the heart rhythm disappear.

Similarly, the influence of lack of sleep on increased insulin resistance during the next day has also been reported, resulting in higher blood glucose levels. As in the previous concept, it is intended to isolate these situations to quantify this relationship. In fact, going further, it is feasible to establish daily averages per hour of blood glucose values and check the number of hours of slept during the previous night.

All these questions lead to a better understanding of DM1 thanks to delving deeper into the possibilities that an IoT environment and the application of ML algorithms can provide. With all this, it is proposed to feed the blood glucose prediction signal back into the system so that it enters a controller and closes the control loop. A closed-loop control system, with a continuous input of the monitoring variables, a glucose prediction output and, in turn, a feedback input that modifies the amount of insulin to be injected by the corresponding infusion pump would constitute, no doubt, an artificial pancreas that would, obviously, be the one - expected by millions of people - "technological cure" of DM1.

2.5 Organization of the Thesis

This thesis is framed under the scheme of a compendium of publications. It begins with a summary that includes a brief description of the motivation of the thesis, as well as the objectives and the final conclusions of the thesis, presenting said summary in Spanish and English. The first and second chapters are related to this purpose.

In the third chapter, the main thread of the publications that make up the compendium and the relationship between them is presented. Thus, this part of the thesis presents the application of the IoT paradigm and ML techniques to the management of the DM1. It is, therefore, a journey through the research that has been developed. In this case, the thesis begins with the proposal of an IoT platform to manage the disease, innovating with the inclusion of new devices to continuously monitor the diabetic patient. From this novelty, the generation of new variables and the appropriate treatment that should be given to these series of temporal data is derived and, within this process, a selection of these variables is made to make it possible to create a *ranking* of importance, as well as indicate the amount of data that is relevant at all times for a future glycemic prediction. The latter is the main theme for the continuation of the work, which is intended to compare three well-known predictive algorithms of ML. Finally, the third chapter ends by studying the possibilities of performing this predictive task using two principal methods: in the cloud or in small portable devices.

Next, chapter 4 includes the five articles that make up the compendium of this thesis. The article presented in section 4.1, entitled “Towards an ICT-Based Platform for Type 1 Diabetes Mellitus Management” [3] analyzes the possibilities that the IoT provides to the management of DM1, taking advantage of portable monitoring devices, routes of wireless communication, and the ability to collect information presented by *smartphones*. Likewise, its possibilities are noted in order to perform, with limitations, a computation of the collected data or direct it towards the cloud. The article continues to expose the advantages of management in an IoT environment from the point of view of care and emergency management. The work concludes with a proposal for a comprehensive platform to manage the disease.

Section 4.2 includes the article called “Variables to be Monitored via Biomedical Sensors for Complete Type 1 Diabetes Mellitus Management: An Extension of the 'On-Board' Concept” [5], which derives directly from the previous. In it, the new variables recorded by means of the portable IoT monitoring devices and the treatment that should be given to those variables are presented. Based on an exhaustive review of the documentation, a set of variables is proposed that will comprehensively characterize the patient with diabetes. As a novelty, there is a need for multiple treatments of each variable, depending on whether it is a pulse, trend, or accumulated ('*on-board*') variable, extending this concept (already applied to insulin) to the remaining variables.

Next, section 4.3 contains the article entitled “Feature Selection for Blood Glucose Level Prediction in Type 1 Diabetes Mellitus by Using the Sequential Input Selection Algorithm (SISAL)” [7], where the respective contributions of the different variables included in the experimentation phase are studied in regard to the evolution of blood glucose. The variable selection method (SISAL) is specific for time series of data and yields a ranking in order of importance of those series that have the greatest influence, as well as how much past data should be taken into account.

The following point 4.4 includes the article " Utility of Big Data in Predicting Short-term Blood Glucose Levels in Type 1 Diabetes Mellitus through Machine Learning Techniques" [8], in which the prediction capacity of known ML algorithms is tested to anticipate blood glucose values using only the past blood glucose data. In addition to comparing three methods, possibilities are being considered to lighten the information to be treated by reducing the number of samples to consider, as well as the amount of data by varying the sampling frequency.

The 4.5 is the last section of the chapter, presenting the article "On the Possibility of Predicting Glycaemia 'On The Fly' with Constrained Devices in DM1 Patients" [9], where the possibilities of making the blood glucose prediction using portable devices mentioned above are studied, namely: a *smartphone* or a *Raspberry Pi 3 b +*, comparing the computational stress derived from using a server or a personal computer. In conclusion, it can be extracted that certain algorithms are lighter than others, being acceptable in certain devices and in other less-powerful devices, only after reducing the amount of data. Under certain circumstances, the computational requirements are such that in a low-cost device, they would be impossible to raise, but a short-term prediction with average computational requirements is approachable with the use of a *smartphone*.

Chapter 5 collects the letters of acceptance of the articles that form the compendium, followed by Chapter 6, which presents the approval of the experimentation phase by the Research Ethics Commission of the University of Murcia. Finally, Chapter 7 includes the bibliography for this thesis, differentiating between the works of authorship of the doctoral student and others of various authors.

Chapter 3

Introduction

3.1 Related works

In the scientific literature, some approaches that ease the management of some illnesses by taking advantage of the possibilities of ICT can be found, especially those that need permanent monitoring, as is the case in DM1.

Currently, it is not possible to think of an approach to an artificial pancreas without a CGM device. This mechanism constitutes a revolution in diabetes care, since it can provide for the magnitude, tendency, frequency and duration of the fluctuations of the glucose levels in diabetic patients [24]. Compared with conventional glucose monitoring (fingerstick, capillary blood glucose monitoring), which provide between three to ten measurements of glucose level per day, CGM can deliver up to one measurement per minute (i.e. 1440 data points per day). So, in order to get a complete monitoring of a diabetic person, the first decision is to choose a proper CGM.

In another vein, groundbreaking advances in electronics have introduced miniaturization and more powerful innovations in biometrics, the field in which measurable biological characteristics are studied. It is now possible to measure certain variables, most of them vital signs (such as the heart rate, exercise and others), in a continuous mode. All of these measurements play a part in the complex balance of blood glucose and must be taken into account. Some years ago, it was difficult to obtain a full-day's data but, nowadays, devices such as smartbands, smartwatches and other fitness and medical wearables make this objective easily attainable, offering a lot of useful information [25]. These innovations have potential in the health field due to their ability to collect different types of data and present them in a user-friendly way. To complete the monitoring, some other medical devices could be worn by the patient, in order to measure temperature, blood presure etc.

With all the above mentioned, we can build a body area network (BAN). This is a wireless network of wearable computing devices that are widely used in remote monitoring in telemedicine. The idea of a BAN is introduced in [26], focusing on epilepsy sufferers. Via a 24-h monitoring method, the system aims to notify the person of the possibility of an epileptic attack. In this way, such paper could suggest, in a DM1 context, the need to predict glucose evolution and alert not only the patient but also health professionals via the Internet or mobile phone in case of emergency. However, this proposal could be completed with a proper information-exchange platform, as also occurs in [27].

When using a BAN, we need a device that is able to play the role of a Body Gateway and a Network Hub. This device is easily identified with a smartphone. The possibilities of this device in diabetes management have already been pointed out initially in [28]. A review of how new applications designed for smartphones can help in diabetes management is presented in [29].

The context aforementioned needs a proper environment to be development and in this sense, an innovate concept is becoming colloquial in our daily lives: The Internet of Things, which could be described as the interconnection via the Internet of computing devices embedded in everyday objects, enabling them to send and receive data. All in all, the IoT is proposed as a good environment for diabetes management, as can be seen in [30]. Smartphones are proposed for remote collection and monitoring of data, with feedback for the patient and medical caregivers. Although such work is a complete proposal, some of the features are not monitored, and, what is more, CGM is not included. In another vein, the proposed model tries to modify, through advice, the patient's insulin dosage, but it should be noted that, without 24-h monitoring, this could not be carried out 100% successfully. Advancements in cloud computing are proposed in [31] for diabetes control. In this case, the cloud provides an easy way to exchange information, but unfortunately full monitoring and management of emergency situations is missing.

Taking advantage of new continuous biometric sensors will create a technological structure in the patient that will generate a huge amount of data. Initially, it has to be ensured that the systems have enough compatibility between clinical devices, platform management, and computerized records in medical institutions in order to promote the sharing of data. This interoperability includes biosensors, which have to be ready to send collected data in an understandable way. In this sense, an accepted criterion is Health Level 7 (HL7) [32], which is the most important standard for the transfer of clinical and administrative data between software applications used by various healthcare providers. It defines a method of moving clinical data between independent medical applications in near real time.

Managing data is a distinctive issue in Wireless Sensor Networks, where the main feature is the mobility of the source of the information [33]. Low-power body area networks can be a solution, because of their low power consumption and their mobility features. Today, these wireless communications can be hosted on smartphones, namely Bluetooth Low Energy (BLE), as well as Near Field Communication (NFC) or WiFi.

As described above, several types of new sensors have burst onto the technological landscape, and this means there is a broad range of different communications requirements in the IoT environment. Therefore, the chosen method of transmitting data must allow for high continuous data transmission, low consumption, high payload size, and sufficient bandwidth. To achieve this, some approaches have been followed previously. The problem of energy saving in a health monitoring system with smartphones is studied in [34]. Taking advantage of Zigbee and Bluetooth communications, low wastage of energy is achieved. However, such work is not focused on any disease in particular, and the study of forecasted situations is omitted. ZigBee and Bluetooth have been implemented successfully in low-power wireless body area networks with discrete sensors, such as blood pressure monitors and weighing scales [35]. Recently, NFC has been proven as a proper choice to monitor glycemia in the successful device Abbot Freestyle Libre, a flash blood glucose monitoring (FGM) system with a popularity constantly raising in the past two years [36].

So, blood glucose information has been dramatically improved thanks to CGM/FGM devices. These types of devices have been increasingly included in diabetes treatment over the last few years and their utility is undoubted in the improvement of glycemia control, as many studies have shown [37]. On the other hand, glycemia change is, unluckily, a complex process and cannot be totally described solely with the previously mentioned variables. Physical activity, quality of rest, emotions, and lifestyle can intuitively help to define the process but, unfortunately, have not always been included in the glycemia-prediction problem. In fact, some of these features have an undisputed role in glucose metabolism. The absence of the abovementioned variables in a predictor model has been due to difficulties in continuously registering these features but, nowadays, in addition to CGM devices, many new electronic gadgets related to health and sports are offered on the technological market. These new tools inform users, in an attractive way, about the intensity of their physical exercise, the quality of their sleep, and their heart rate at every single moment of the day. These novel resources usually adopt the shape of a smart band and, due to their sufficient accuracy, open new routes to complete a full characterization of the status of the patient and therefore to achieve better management of diabetes.

The importance of the proper set of variables and the data collection procedure used is cardinal in blood glucose prediction methods, as is the previous treatment made of them; they should be expressed as time-series data because the number of historical events taken into account is important. Feature selection in time series is different from feature selection in normal static data. The target value of the latter only relates to the current values of features, while the target value of the former relates to the values of features in the previous timestamps as well as in the current timestamp. Hence in feature selection, removing irrelevant and/or redundant features/variables as well as choosing the proper past values is a critical task in order to build the dataset well. A proper method is an improved filter method reported by Tikka and Hollmén [38], the Sequential Input Selection Algorithm (SISAL).

Having drafted an overview about hardware and communications added to variables and their importance, one of the most important tasks to manage DM1 is to anticipate to situations like hypo or hyperglycaemia. Numerous attempts have been made to develop models for predicting blood glucose levels in people with diabetes. Today, the most well-studied approach for blood glucose prediction is based on detailed physiological models that try to capture the dynamics of glucose-relevant variables within different systems in the body. More recently, a new approach has been proposed to address this problem by applying machine learning techniques. The key concept behind this approach is that there exist repetitive cycles in glucose-insulin dynamics, e.g., before/after meals, before/after bedtime, and predictable insulin sensitivity changes throughout the day because of circadian variations in hormone levels. Remarkable studies have evaluated patterns by time of day, even showing the presence of variability according to the day of the week [39]. Several data-driven models have been developed to explore the repetitive nature of glucose-insulin dynamics, using diverse techniques for time series forecasting such as autoregressive (AR), impulse-response (IR), autoregressive exogenous input (ARX), autoregressive moving average exogenous input (ARMAX), and autoregressive integrated moving average (ARIMA). For example, Estrada et al. developed an ARX model that relies on blood glucose levels and insulin dosages to forecast future blood glucose levels within a 45-minute prediction horizon [40]. Other approaches, as performed by Nuryani et al., explored the use of support vector machines (SVM) to predict hypoglycemia using electrocardiograms (ECG) in addition to blood glucose levels and insulin injections [41]. Marling et al. used the SVM models, combining data gathered from wearable activity trackers, monitoring heart rate, galvanic skin response, and skin/air temperatures [42].

Another class of empirical models uses artificial neural networks (ANNs) to learn the relationship between past and future blood glucose levels, also taking into consideration other data sources. For example, Pappada et al. developed an approach that combined the fingerstick method to measure blood glucose levels with insulin dosages, meals, and in some ways lifestyle conditions and emotions [43]. Zecchin et al. developed a predictor based on a neural network (NN) model and a first-order polynomial extrapolation algorithm that combines past CGM sensor readings with data on carbohydrate intake [44]. Unfortunately, NN-based models require a large volume of data to properly calculate the internal parameters of the networks and avoid overfitting. In addition, the majority of these studies assessed the performance of the models either using simulated datasets or real-world data collection campaigns of very limited duration. It is therefore important to make sure that data used for the training of such an empirical model is by design long-term and involves real individuals.

So, we can find many examples of glucose forecasting including past glucose values, but also using other features such as insulin regimen, meals, exercise, etc. The possibility of forecasting using only past glucose values has been studied, and this could have some benefits: relying only on one device (CGM), avoiding human error due to subjectivity, and simplifying the calculation process of the algorithm. With this, we look for the possibility of doing this task locally, i.e. in a smartphone, which could be useful under some conditions (unstable internet connection, remote areas).Although modeling and forecasting glycaemia using cloud computing present many advantages, we will evaluate if it could be possible to translate this double task to a constrained device (simplifying the work by using univariate approaches) in order to use the limited computational power existing in small devices under unavailability of the cloud resources.

The concept of developing data-centric prediction models relying solely on CGM technologies has been studied in the past [45] [46] [47]. Different machine learning techniques have been used to develop data-centric prediction models that can be used for early hypoglycemic/hyperglycemic alarms and for closing the glucose regulation loop with an insulin pump. For example, Sparacino et al. evaluated an AR-based model within a hospital environment involving 28 DM1 patients for 48 hours [45]. Eren-Oruklu et al. developed a model using ARIMA based on data collected from monitoring 22 DM1 patients for 48 hours [46]. Hamdi et al. monitored 12 patients to develop a model combining support vector regression and a differential evolution algorithm [47]. In any case, to the best of the authors' knowledge, a proper and complete comparison between glucose prediction algorithms has not been carried out previously. At best, in some works a few methods were compared, using glycemia but also (depending on the paper) a few other factors, assuming some parameters, and it is not possible to extract overall conclusions. In addition, the origin of the data was fairly different from one study to another, so developing a comparison among the methods studied in different works becomes a task only achievable in a subjective way.

Once we have studied the diverse algorithms to predict, we need to note that the glycaemia forecasting task needs to be performed in a specific hardware and, as it will be made several times per day –'on-the-fly' (recalculating the model and then predicting)– it is necessary to take into account the execution time. Depending on the ML algorithm that we choose, we will face at some point a computer limitation, not only due to hardware, but also to software, including the Operating System [48]. This idea of restrictions in devices has been explored previously in some more demanding processes, like the Internet Protocol traffic flow [49]. In this case, some methods applied to the classification task were studied and compared.

It is possible to find an ML performance comparison in the field of medicine, as can be seen in [50], regarding Functional Magnetic Resonance Imaging (fMRI), which compares the accuracy of six different ML algorithms (among them, RF and SVM) applied to neuroimaging data of people engaged in a bivariate task, asserting their belief or disbelief of a variety of propositional statements. In this interesting study, RF was found to be more accurate than SVM. The performance of each algorithm was treated, reducing each feature set in order to reduce the amount of data to deal with, therefore lowering the computational burden without losing accuracy.

In [51], we can find a general comparison of ML algorithms, among them RF and SVM. In this case, the algorithms were tested using the free Gisette Dataset[3], a handwritten digit recognition problem. With this, SVM is one or multiple orders of magnitude faster, but RF is more precise.

Concluding with this review, to the best of the authors' knowledge, no data about the computational burden have been found in glucose prediction for DM1; so we have the aim of improving the computational performance of different glycaemia forecasting methods, then, the possibility of running such algorithms in constrained devices will be analyzed.

3.2 The problem of managing type 1 Diabetes Mellitus from an ICT view.

In this section, the structure of an IoT platform that is aimed at managing diabetes, modeled in layers where different technological issues are arranged, is presented. This structure can also be found in several areas, such as energy savings or in buildings, but has not been employed in a diabetes management system up to now. Therefore, as shown in Figure 3, the configuration of an integral diabetes management platform should be built as follows:

[3] https://archive.ics.uci.edu/ml/datasets/Gisette

- Biological substrate. This is the layer where the patient is located. His or her physical changes, reflected in skin, blood, movement, and so on, will be disaggregated as variables to be measured. This layer is where the inception of the data occurs, where it is generated, and, finally, where the system's outputs will actuate (via therapeutic decisions).
- Sensorization layer. Input data are acquired from a plethora of sensors, all of them connected to an IoT framework. Sensors can be configured and controlled remotely through the Internet, enabling a variety of monitoring applications and creating a technological structure.
- Communication layer. This tier allows for data permeation until the next stage. All of the sensors must support several communication channels in order to connect easily. Avoiding direct input/output (I/O) through regular wiring, Wi-Fi, 4G, ZigBee (or 6LowPAN), and Bluetooth connections need to be available to support direct access to sensors, mainly via another smart device (smartphone, tablet) used as a gateway, following an IoT approach. All of the elements are connected to each other by using a little Local Area Network (LAN), and, likewise, to a smartphone (or the cloud).
- Middleware layer. Given the heterogeneity of data sources and the necessity of a seamless integration of devices and networks covered by the sensors and communication layer of our architecture, a middleware mediator is needed to deal with this task. Therefore, the transformation of the collected data from the different data sources into a common language representation is performed in the middleware layer. As mentioned before, HL7 is a proper candidate to deal with this task.
- The ontology implemented to represent the knowledge of the bloodstream glucose level clearly follows a user-centric approach, which takes into account all of the information characterizing the situations, the devices and their status, and the patient himself of herself.
- Computing and management layer. In this layer, data collected are handled in order to carry out a data analysis, obtain a glucose prediction, and chose an optimal therapeutic solution. Here, is where a data-processing center and modeling core are placed. Pervasive computing has to be done on two levels: one local, in the smartphone, and another in the cloud. In this way, it is possible to avoid the risk of lack of connection to the Internet or battery failure. Data harvested are sent, via LAN, to a smartphone, or maybe, in some circumstances (e.g., the absence of a smartphone), directly from the sensors to the cloud, if it is provided. Therefore, ubiquitous computing allows for a powerful and safe way to address the problem.
- Display layer (interface). Access to the system will be via browser. Thus, the parameters can be adjusted either by a smartphone or by a computer, local, or external, which is close to the patient or remote for healthcare staff. Data collected are also accessible, as well as statistics and the general status of the glycemia control system. It should be noted that the previous layer must have reached a blood glucose prediction and an optimized solution of insulin input, and this can be shown to the user either for information only or to await confirmation or variation.

- Outputs. With the goal of integral management, the platform, thanks to all of the data collected and available resources, needs to be ready to cover the following exposed requirements.

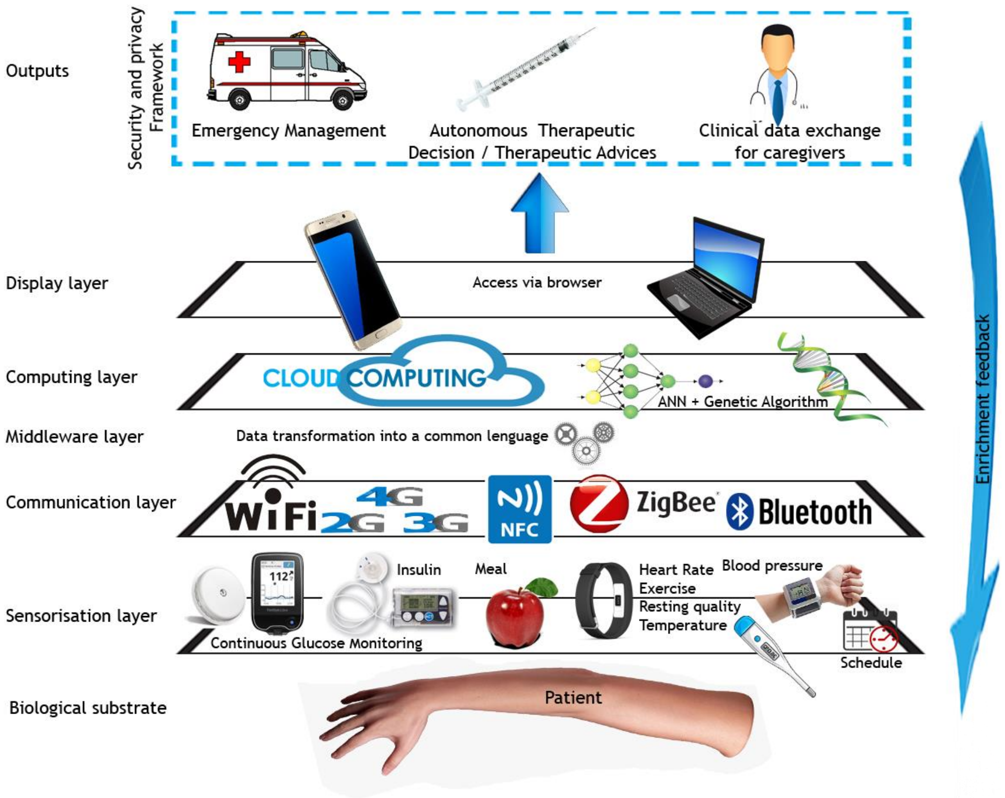

Figure 3. Proposed layered structure for an artificial pancreas

And as a result, it is possible to obtain:

- Concerning therapeutic decisions.
- Emergency management.
- Clinical data exchange for caregivers.

This structure needs to be networked properly. In this sense, as can be seen in Figure 4, the system architecture would be composed of a local gateway (smartphone) that is connected to the Internet via 4G or through a domestic WiFi. The smartphone will be exclusive for each patient. It is on the Internet that it is possible to place the data warehouse and a processing center, as well as the modeling core. Moreover, by using the Internet, the management of the system via a Web browser will be possible.

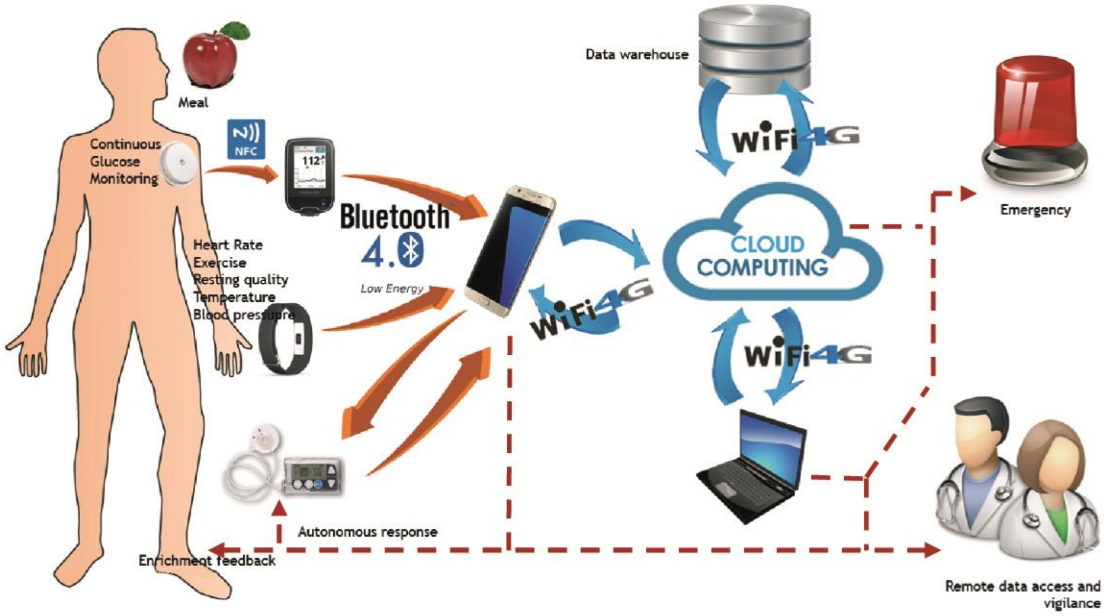

Figure 4. Diagram of the flows of data

The smartphone is responsible for collecting data from the sensors. It also has to be reliable to establish and manage connections with devices that are attached to the patient. By using several methods of communication (BLE is a good candidate, but 6LoWPAN, NFC, and so on can also be used), the exchange of data becomes possible so that a plethora of biometric meters can be interlinked.

3.3 Introducing new devices

The main device which is the cornerstone of a DM1 management plan is CGM. In this sense, Dexcom is probably the most accurate CGM system available. In its most recent version (Dexcom G6) a MARD of 9% is reached thanks to the new 505 software, which features the same advanced algorithm as that used in artificial pancreas research. The life of the sensor is seven days, according to the manufacturer's specifications, but users can usually restart it when this expires. The G5 model has several features, such as the possibility of displaying the glucose levels on a smartphone, without the necessity of carrying the Dexcom receiver. The previous version (Dexcom G4 Platinum) cannot transmit to a smartphone (just to a Dexcom receiver) but both are able to be connected to an Animas Vibe insulin pump. Other combinations can also be found, like the integration of a T:Slim insulin pump with the Dexcom G4 and G5. Dexcom will launch its next CGM (G7) in 2020. Medtronic also offers CGM sensors: Medtronic Enlite, together with their insulin pumps, although it is also possible to use the sensor separately. Its MARD (13.60%) is a little bit higher than that of the Dexcom devices, and the approved life of the sensor is 6 days, although it can be reactivated for up to twelve days. By adding a specific Bluetooth device, it is possible to show the results on a smartphone.

At this moment, the US Food and Drug Administration (FDA) has approved a device with a CGM gadget and an insulin pump, capable of stopping the insulin infusion in case of hypoglycemia (Medtronic 530G) and, recently, the FDA has also approved the

Medtronic 670G, which is also able to minimize hyperglycemia thanks to its connection to the CGM gadget.

Abbott has offered the Freestyle Libre from the end of 2014 [36]. Strictly speaking, it is not a real CGM device since the glucose information is not transmitted continuously (although it is stored and can be read by scanning the sensor with the receiver), so it is not possible to set alarms for low or high values of glucose. The sensor is placed on the arm and the information can then be transmitted via NFC, known as the "Flash System" (FGM). It achieves a MARD of 11.4% and the sensors have a longer life (14 days) but they cannot be restarted. One of its perks is that it does not need to be calibrated.

At the moment, smartphones present a level of adaptability that is unachievable by any other device. They allow us to:

- Keep the software responsible to model the dynamics of the system, make a glucose prediction, optimize a solution and control the process.
- Gather information between CGM devices and an insulin pump. The possibilities of connectivity are enormous. Today, a mobile phone is not only able to communicate via 4G but also through Bluetooth, Wi-Fi, NFC, Ant+, etc. which offers a plethora of different choices.
- Send data and information to the Cloud, to be either stored or computed (cloud computing).
- Forward emergency calls in case the patient is at risk.
- Update its software when required.

Furthermore, smartphones provide a wide range of possibilities in the field of monitoring. Nowadays, the inclusion of accelerometer sensors, gyroscopes and pedometers allows physical activity to be quantified. These possibilities have been tested in experiments performed by Place et al. [52]; in real time, patients' evolution has been followed from different locations. Other publications have studied the use of a smartphone as a device capable of sending an emergency call, providing the location via GPS in case the patient is at risk [53].

However, there are some doubts as to whether smartphones will be approved as medical devices (class III, high risk). The controller application installed in the phone has to be reliable, so as to avoid conflicts with other applications running on the device. Other circumstances, such as a discharge of the battery or lost connectivity, would have to be prevented. These pros and cons are exposited in Rigla [54].

In another vein, groundbreaking advances in electronics have introduced miniaturization and more powerful innovations in biometrics, the field in which measurable biological characteristics are studied. It is now possible to measure certain variables, most of them vital signs (such as the heart rate, exercise and others), in a continuous mode. All of these measurements play a part in the complex balance of blood glucose and must be taken into account. Some years ago, it was difficult to obtain a full-day's data but, nowadays, devices such as smartbands, smartwatches and other fitness and medical wearables make this objective easily attainable, offering a lot of useful information [25]. These innovations have potential in the health field due to their ability to collect different types of data and present them in a user-friendly way. In addition, they have potential due to their connectivity; they are usually provided with a Bluetooth connection. However, they are restricted by their size, length of battery life, and the fact that these devices are sometimes conceived for non-health care, professional uses. These factors cast some doubts on their accuracy [55] but, in any case, they can provide data of an appropriate order of magnitude. A compilation of the most popular and affordable fitness smartbands is presented in Table 2, along with a summary of their main features. As we can see, these devices can provide a wide range of physiological features and, therefore, providing a good characterization of the patient. To complete the monitoring, some other medical devices could be worn by the patient, in order to measure temperature, blood presure etc.

Therefore, although the use of novel devices (accelerometers, electrocardiograms, thermisors, etc.) has been tested, they have only been studied in isolated trials and not included in DM1 management systems. Some of them (but not all) have been used together with CGM devices, sometimes with the sole aim of studying the relationship between BG and other features, and sometimes to predict glycemia levels. It should be said that 24 hour, simultaneous monitoring of all the features has not been carried out, nor has a study of the types of variables, the correlations between them, ranks of influence on the BG, and a valuable set of features used for modeling the evolution of glycaemia.

3.4 Introducing new variables

At present, most of the previous literature on diabetes management systems has only taken into account glycaemia and insulin levels, and sometimes an estimation of meals, but it seems reasonable to incorporate additional variables that could also influence glucose levels as far as it is possible to measure or estimate them. It seems that there is a general agreement in using glycaemia, insulin, and meal as remarkable variables [56] [57], and the studies with this set of variables are a majority, but it is also possible to find works involving just previous glycaemia data as the only variable in use (like can be found in auto-regressive models approaches) or in [40], just adding insulin to glycaemia and using an autoregressive with exogenous terms model. There is also a lot of studies in the last years taking into account other variables, mainly exercise, both in silico and in vivo [58], and also considering the possibility of heart rate, temperature, etc. These ideas will be presented in the following lines. In any case, to the best of the authors' knowledge, there is no previous work using a comprehensive characterization of the diabetic patient, examining a global and integral overview.

In this section, a comprehensive list of the significant variables for a complete DM1 control system is analyzed. Some of them have been previously discussed in the scientific literature but others are presented in this paper for the first time.

- *Current and previous glucose bloodstream level*: This is one of the principal variables. In addition to traditional capillary blood monitoring, which consists of discrete glucose values, new CGM devices offer enough accuracy and a good sample frequency [59] to be one of the inputs to a continuous control algorithm.
- *Insulin*: This hormone has an exogenous origin in a diabetic patient, who has to make a decision about the dose to be injected. As it is the variable which primarily governs how much glucose levels will decrease due to hypoglycemic

response, this could be the main variable requiring optimization. It is necessary to differentiate three ways of taking insulin into account:

- Instantaneous insulin input.
- Basal insulin.
- Accumulated Insulin (Insulin on-Board, IOB): It involves basal insulin and remaining fast-insulin, which can be acting with a low-intensity (but still noticeable) for several hours. It is assumed to have a remarkable effect for the previous eight hours [60].
- *Exercise*: Physical activity increases the muscles' demand for glucose; it is necessary to a normal work out. It also raises blood circulation, so insulin is used up faster and its effectiveness is increased because exercise temporarily makes the cellular walls more permeable, which lets glucose enter the cells more effectively [61]. It is possible to make a distinction between three aspects regarding exercise:
- Recent Exercise.
- Intensity of exercise.
- Accumulated exercise (Exercise on-board, EOB): Although the increased effect of exercise on both blood glucose and insulin requirements starts to decrease at the same time the physical activity is being performed, a remaining effect can act for up to 48 hours afterwards [62] [63].
- *Meals*: What we eat and when we do it has a great influence on blood glucose levels. Food is mainly transformed and absorbed as glucose, which is dumped into the blood, raising glycaemia almost instantly [64]. Three aspects should be noted, regarding meals:
- Notification of ingestion.
- Meal (carbohydrate counting).
- Accumulated intakes (Meal on-board, MOB): Regarding the above, the idea of a 'meal on-board' concept has arisen and seems reasonable. Unfortunately, to the best of the authors' knowledge, there is no research which takes this idea into account. In this sense, this variable is first presented in the present paper.
- *Stress*: Stress hormones also have a hyperglycemic influence [65]. Adrenaline and cortisol are secreted for several reasons. In this case, two important aspects should be noted:
- Sleep quality at night.
- Heart rate.
- *Temperature*: This parameter can be a symptom of hypoglycemia, becoming lower under these circumstances [66].
- *Perspiration*: This is another expression of hypoglycemia. As with temperature, this symptom can go unnoticed in long-term patients [67].
- *Blood pressure*: Therefore, high blood pressure could point to a poor control of glycaemia [68] and lead to complications involving the heart, eyes, kidneys and blood vessels.
- *Schedule (time)*: By identifying both the hour and the day of the week, it should be possible to anticipate and predict the evolution of the system. Another advantage of recording the schedule is the identification of the basal

insulin evolution. New types of slow-acting insulin [69] have an almost flat profile but it should be noted that it is not 100% flat. Therefore, the hourly absorption and its alterations should be taken into account.

Other features that could be considered in the system, in order to reach personalized solutions, could be: age, sex, height, or weight.

There are also some others patient's features, situations, and concurrent (chronic or temporary) illnesses that could affect to glucose levels. So, although it would be reasonable to add other variables if they have any significance, only the ones that can easily be measured with non-invasive devices are listed above.

Figure 5 summarizes all of the points analyzed in this section.

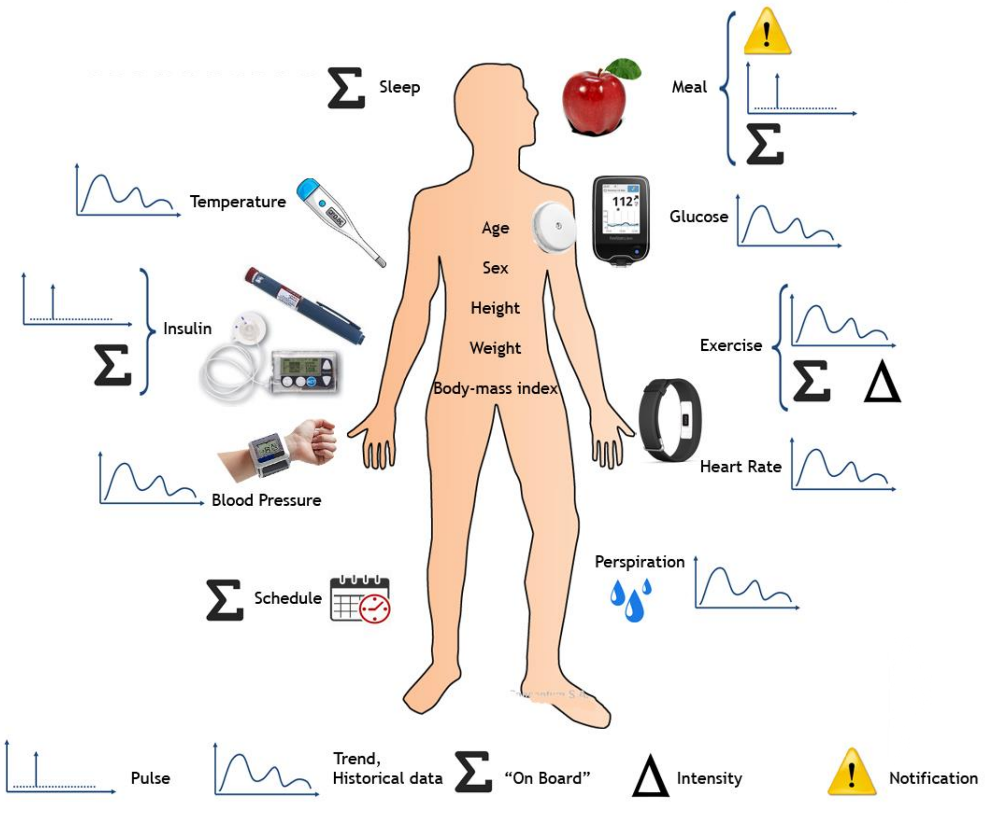

Figure 5. How to characterize a diabetic person using continuous monitoring devices

3.5 The necessity of a proper data set

We conducted a study utilizing the above system involving 25 DM1 patients with diabetes during 2018 under the supervision of the Endocrinology Departments of the Morales Meseguer and Virgen de la Arrixaca Hospitals, of the city of Murcia (Spain). The study was conducted in accordance with the Helsinki Declaration. The study was approved by the Ethics Committee of Universidad de Murcia, Spain. Data storage complied with the stricter data protection rules for protecting personal information. The clinical characteristics of the patients considered in the study are summarized in Table 2. All participants were fully informed about the purpose of the experiment and were provided written informed consent and assent according to national regulations. The participants were informed about the data collected during the study and how they were stored.

Participants were recruited via advertisements and word-of-mouth. The group was composed of 14 men and 11 women, all of them under medical treatment and professional supervision. Patients were students or office workers of the Universidad de Murcia, 18 to 56 years of age (average 24.51); that is, most of them were young adults. Patients were chosen with an illness duration of at least 5 years to guarantee familiarity with the course of the disease and, moreover, all of them were familiarized with the use of the Abbott Freestyle Libre FGM sensor.

Table 2. Clinical characteristics of monitored diabetic population

Population Feature	Value		
Subjects (number)	25		
Sex	14 men – 11 women		
Occupation	16 students – 9 office workers		
Population Feature	Median	Min	Max
Age (years)	24.51	18	56
Body mass index (BMI, kg/m^2)	22.20	19.42	24.80
Duration of diabetes (years)	9	5	29
Fingersticks per day (usages)	7	5	12
Insulin units per day (fast insulin + slow insulin, median)	47	36	59
HbA1c (%)	6.8	6.3	7.8

Data set was completed with the smart band *Fitbit Charge HR*. Each person wore a smart watch which registered, continuously, the exercise done (number of steps), heart rate, and minutes of sleep. Although these devices are not specifically medical devices, the accuracy is good enough for the input to be considered as valid, and the energy consumption requirements are becoming lower and lower. Further characteristics can be found on the manufacturer's website.

All the patients reported leading a healthy life, and all of them practiced sports at least 3 times per week. There was also a control of their schedules, ensuring that all of them followed reasonably well-regulated daily routines without abrupt changes to their daily timetables. Patients' DM1 was usually well controlled, all having glycated hemoglobin (HbA1c) values between 6% and 7% at the beginning of the experiment.

During the passive monitoring period, patients were encouraged to follow their daily routines and apply a balanced diet according to their caloric necessities. All patients were encouraged not to deviate from their doctors' advice during the monitoring period. All participants followed a basal-bolus regimen, using slow insulins such as Levemir, Tresiba, or Lantus, which achieve a flat-action curve, and fast insulin such as Humalog Lispro. The former provides more than 24 hours as a basal coverage, and the latter is used to compensate for a rise in glucose, which can be due to the intake of a meal or a hyperglycemia caused by other factors.

3.6 Feature selection: A rank of importance of variables using SISAL

With the collected data, we focused the idea of making a rank of influence, discovering also the most influential amount of past data of each feature in glycaemia evolution. In order to do this, SISAL [38] was put into effect using the SISAL package, (https://cran.r-project.org/web/packages/sisal/index.html) a library specially programmed for the so-called mathematical software R. It has been applied to a feature set composed of each of the 25 samples of the lagged variables: glycemia, insulin, meal, exercise, heart rate, and sleep, taking into account the past 75 values (as the sample frequency is 1 value per 5 min, we are considering a past window of 375 min, 6.25 h). In addition to the lagged values, we introduced schedule; this feature was registered as the hour and minute when glycemia was recorded.

The results are shown in Figure 6 and Table 3. For each lagged feature and each patient, we choose an optimum value, taking into account the significance, expressed by the p-value obtained using SISAL after evaluating a Z-score test [70], and the corresponding associated lag of that feature. The lower the p-value, the more significant the feature, considering a specific range of past data values. As schedule is not a lagged feature, for each patient the associated timestamp is taken into account, resulting in a p-value only.

Considering the p-values, the most influential feature is insulin, followed by meal. Glycemia takes third position in the ranking because it reflects the influence of all the relevant features at the same time. Continuing, and belonging to a separate influential group, we can find heart rate and sleep, which are significantly less important. Schedule is in this group, with a discrete influence.

With this, SISAL offered results on the subject of importance but also on the number of past values that need to be taken into account. This is a crucial idea in order to ease the execution of the forecasting algorithm and its computational effort. In this sense, we have results that indicate that the most remarkable variables are insulin, meal, and past glycemia followed by exercise and then heart rate, sleeping time, and schedule. The set of features will depend on how demanding the chosen algorithm is and the power of the device which is computing the prediction.

In addition, the conclusions express that some variables noticeably show the same behavior between diabetic people (like insulin) and others display characteristics depending on the specific person. This is interesting in order to presume the behavior of a variable or study a particular performance according to each patient.

Glycaemia: In the importance ranking, glycemia takes third position. Although some autoregressive models take into account only glycemia and simply the past nine values (barely half an hour, in [45]), these choices are not properly contended. In our study, the results indicate that around 4 h (49 past measures) can be a proper past window to take into account.

Insulin: Insulin is the most influential feature, according to the p-values, followed by meal. In this sense, we can appreciate that the standard deviation is smaller in the case of insulin, which is because the behavior of this one feature is more constant among people. Insulin increases importance at 21 past values, which means 1.75 h. This is reasonable, since we are referring to fast-acting insulin (boluses), and this type of insulin usually has its main effect during two and a half hours, with a maximum at 90 min [71].

Meal: Meal is the second most important variable. In this case, it is seen to be more variable from one patient to another, which makes sense as the absorption of meals is different depending on the nutritional composition and on the metabolism of each person.

The results offer a good past window with 35 past values, in other words, 2.91 h. This is consistent with previous literature, which estimates it in 3.27 h [72].

Exercise: Its influence on glycemia evolution has been studied so far as it increases the demand for glucose and raises the insulin sensitivity [73]. Relevant past values of exercise data include 1.75 h (21 past values). This result seems to be coherent with literature related to the influence of sports [61] and it presents a wide variability, as it covers a wide range of different sports and intensities. Anyway, exercise still has an effect up to 48 h

Heart Rate: Heartbeat is the fifth variable. If an increase in heart rate is not related to physical activity, it could be generated by stress situations [74]. Heart-rate reactions are very variable from a person to other, which is the reason why results that offer relevant past data are included in a window of 2.08 h.

Sleep: As the influence is low, according to the p-values obtained, and is fairly variable from person to person, sleep deprivation has been pointed to as a cause of insulin resistance [75], which leads to hyperglycemia in diabetic patients.

Schedule: Circadian rhythms and the importance of having a schedule and a routine in order to manage diabetes are well known but have not been included in a prediction algorithm yet. At least, from one week to another, it is possible to identify a pattern, especially in the case of diabetic patients, since this is generally helpful to control their diabetes. Disruptions in the circadian cycle could lead to insulin resistance [76].

Table 3. Past values and *p*-value for each feature. Median values and standard deviation.

Feature	Past Values (Measures)		Past Values (Minutes)		*p*-Value	
	Median	Std. D.	Median	Std. D.	Median	Std. D.
Insulin	21	7.08	105 (1.75 h)	35.40	0.00360	0.01011
Meal	35	13.88	175 (2.91 h)	69.40	0.00540	0.01030
Glycemia	49	11.27	245 (4.08 h)	56.35	0.00621	0.00992
Exercise	20	13.88	100 (1.66 h)	69.40	0.02340	0.01030
Heart Rate	25	17.19	125 (2.08 h)	85.95	0.02682	0.01154
Sleep	51	12.54	255 (4.25 h)	62.70	0.03721	0.00896
Schedule	–	–	–	–	0.02864	0.01254

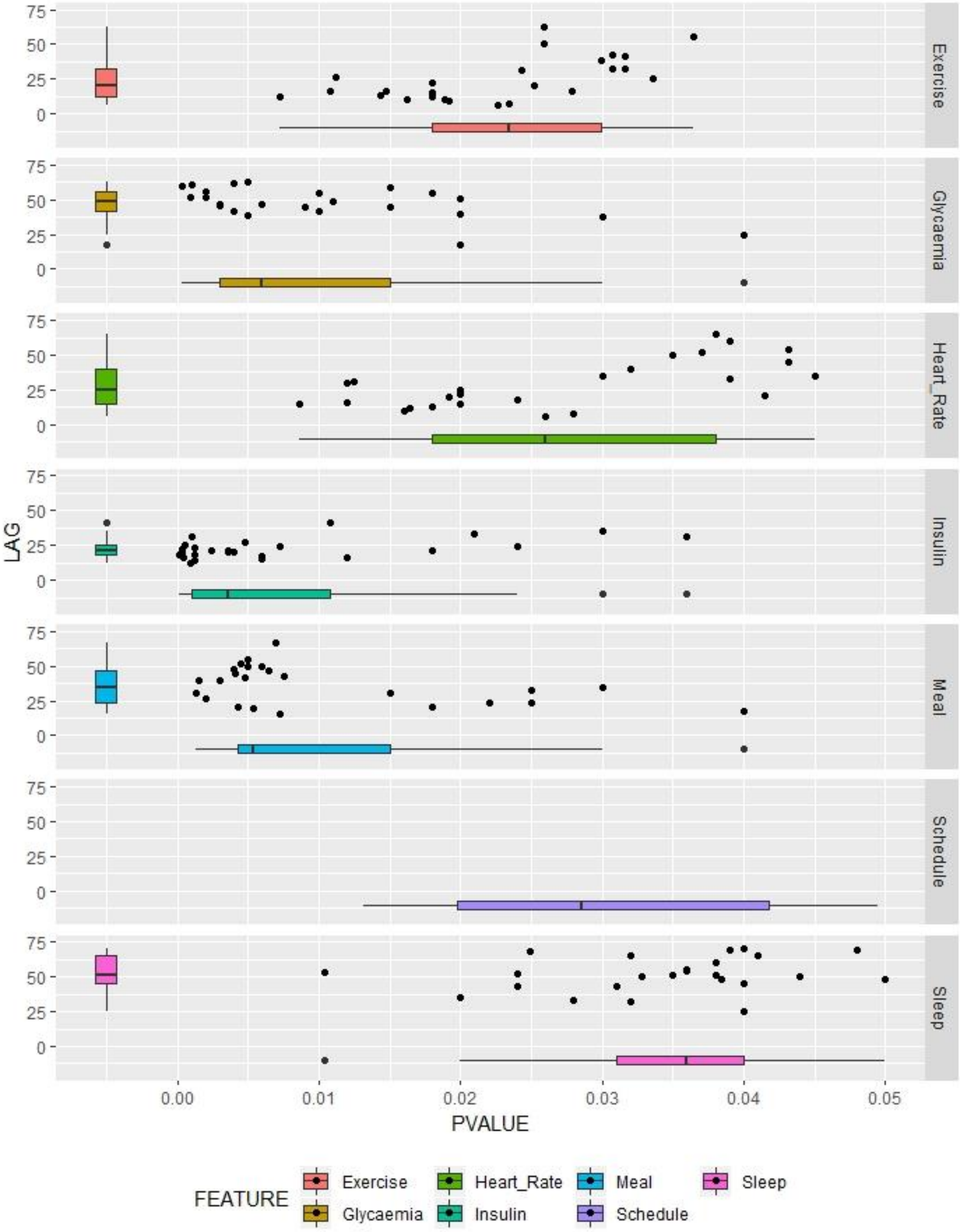

Figure 6. Feature selection results.

3.7 The glycaemia forecasting task

As we mentioned, one of the main characteristics of a DM1 management structure is a forecasting blood glucose model. In order to begin with a simplified problem, our goal was to develop reliable prediction models that can estimate the future glucose level with high accuracy based only on current and past data collected from the FGM sensor. We developed patient-centric models by using different machine learning methods to capture the properties of the blood glucose time series of each individual patient. Therefore, for each patient, we generated a separate prediction model.

The data provided by the FGM sensor were sampled with a *sample frequency* (SF) of 1 measurement every 5 minutes, 10 minutes or 15 minutes. Therefore, the SF controlled the *velocity* of the data that we are taking into account. The values sampled were used to create a *past sliding window* (PSW) including historic values varying from 3 to 36 hours. The PSW controlled the *volume* of the data the model used for the prediction. Given the sliding window data, the model continuously predicted the glucose levels at preset *prediction horizons* (PH) at 15, 30, 45, and 60 minutes ahead from the present time. Figure 6 provides a graphical representation of the fragmentation of the data collected into windows that are used as input for the patient-centric prediction model.

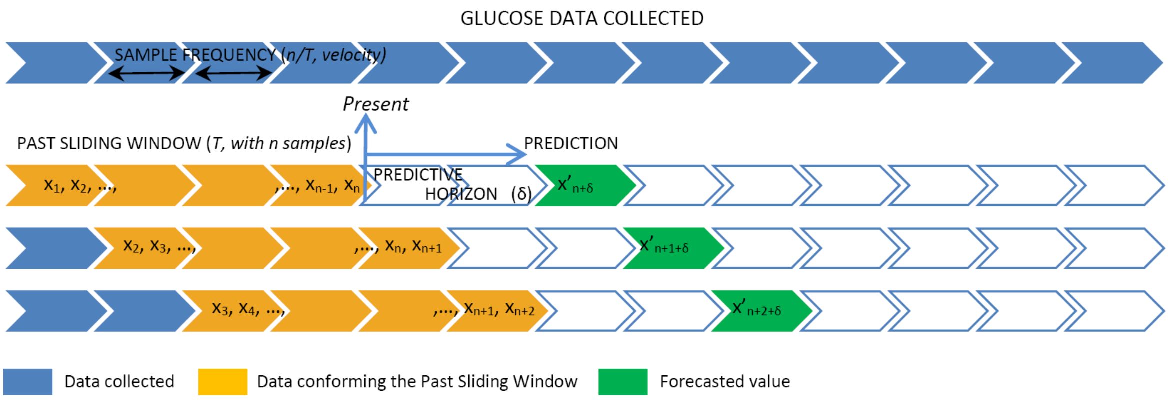

Figure 7. Time series analysis and cross-validation with slide-window.

In more detail, at a given time t the data collected from the FGM sensor over the past period T were used to generate the training set of n data points $\{x_i\}_{i=1}^{n}$, where $x_i \in R$ is each individual data point received from the FGM sensor. We say that n characterizes the *volume* of the data set and $\frac{n}{T}$ the *velocity* of the data. Given this training set, the goal of the prediction model is to approximate the real underlying mapping accurately enough such that it can predict the next δ value, that is, the output value $y(t + \delta) \in R$ for that time series. We say that δ is the prediction horizon.

The sliding window works as follows. As soon as a new FGM value is received, the training set is reorganized by removing the oldest observation (x_1), shifting all the values 1 position up (i.e., x_i becomes x_{i-1}) and finally adding the new value received as the newest one (x_n). In this way, the dataset size and ordering of the observations is always preserved.

The methods considered in this first approach for glucose level prediction are the following:

- Autoregressive integrated moving average (ARIMA) [77].
- Random forest (RF) [78].
- Support vector machines (SVM) [79].

We start by looking into the achieved performance in predicting future blood glucose levels when we use only historical data on the blood glucose levels. Table 3 shows the results obtained in terms of root mean square error (RMSE) between the predicted values and the observed values. We first observe that for all the patient models developed, the error of the prediction is greater as the prediction horizon (PH) increases, which seems reasonable since the data collected is progressively further from the prediction. In any case, in general, for the three different methods used to develop the patient models, acceptable predictions are achieved for PHs of 15 minutes and 30 minutes, with average RMSEs practically smaller than 20 *mg/dL* for ARIMA and RF, and smaller than 28 *mg/dL* for SVM.

We continue at Table 4 by looking into the importance of historical data volume on the prediction accuracy of future glycemic levels. The size of the window of past data allows the model to capture the temporal structure in the time series. One expects that increasing the volume of historical data will improve the prediction accuracy. For this reason, we designed the first experiments by training the patient models using different past sliding window (PSW) sizes that contained the observations of the past 3, 6, 12, 24, and 36 hours. In this experiment, the data velocity is fixed to 1 sample every 5 minutes and a prediction horizon of 15, 30, 45, and 60 minutes.

Looking into the volume of historical data needed to achieve an accurate prediction, as expressed by the different PSW sizes used, the results achieved indicate that all the patient models developed have a reduced RMSE with a window size of 6 hours.

Experimental evidence indicates that regardless of the method used to develop the models, and for all the prediction horizons, there is a limit moving backwards (6 hours) when it comes to considering past data in order to improve accuracy of the prediction. Previous works have introduced this order of magnitude, linking the idea of circadian cycles [80] and the slots of morning/afternoon/night.

Table 4. Root mean square error (RMSE in mg/dL) for different past sliding window sizes (PSW) and for different predictive horizon (PH) using a fixed sampling frequency (SF = 5 minutes)

SF = 5 min					
RMSE (mg/dl)		PH=15	30	45	60 min
ARIMA	3	11.64	17.62	23.21	28.17
	6	11.53	17.31	22.75	27.60
PSW (hours)=	12	13.00	18.64	23.92	28.64
	24	13.93	19.45	24.55	29.05
	36	14.78	20.21	25.24	29.68
RF	3	10.59	15.07	19.04	22.56
	6	10.15	14.63	18.60	22.12
PSW (hours)=	12	11.18	15.66	19.64	23.17
	24	11.65	16.11	20.06	23.56
	36	11.98	16.43	20.38	23.89
SVM	3	18.14	22.55	26.46	29.74
	6	17.65	20.82	23.74	26.36
PSW (hours)=	12	19.73	22.57	25.17	27.48
	24	21.45	26.73	30.71	22.68
	36	23.08	27.90	31.30	33.90

Table 5. Root mean square error (RMSE in mg/dL) for different sample frequencies (SF) and for different predictive horizon (PH) using a fixed past sliding window size (PSW = 6 hours).

PSW = 6 hour					
RMSE (mg/dl)		PH=15	30	45	60 min
ARIMA	5	11.53	17.31	22.75	27.60
SF (min)=	10	13.14	20.13	23.64	29.81
	15	15.08	21.10	26.32	30.82
RF	5	10.15	14.63	18.60	22.12
SF (min)=	10	11.65	17.37	19.84	24.26
	15	15.43	19.41	22.87	25.92
SVM	5	17.65	20.82	23.74	26.36
SF (min)=	10	19.90	23.97	25.73	28.90
	15	23.26	26.03	28.44	30.57

Looking into the performance achieved by the three different methods used to develop the patient models, the results indicate that RF method is the most accurate one, achieving better predictions with all the PSWs and for all the PHs considered.

We now proceed by looking into the effect of the velocity of the data collected from the GCM sensor on the prediction ability of the patient models. In the previous set of experiments, the data velocity, (i.e., the SF), was set to 1 sample every 5 minutes. In this set of experiments, we evaluated the effect of using different sampling frequencies between

5 and 15 minutes on the achieved RMSE. Table 4 depicts the results obtained in terms of RMSE between the predicted values and the observed values when the past sliding window is fixed to 6 hours, as this was the optimum value identified in the previous experiment. It is evident that the higher the SF, the higher the resulting RMSE, regardless of the PH. However, as the PH increases, the effect of the SF is bigger: in a PH of 15 minutes, when increasing the SF to 10 minutes, RMSE increased by 50%; but with the same change in SF, in a PH of 60 minutes, RMSE increased only around 10% to 12%. We conclude that the decision of reducing SF depends on how far we want to predict. An interesting point is that, just by changing SF from 5 minutes to 10 minutes, acceptable RMSEs are still achieved. We also observed that the patient models that are based on the SVM method reached a very high RMSE when SF is equal to 15 minutes. We therefore conclude that models based on the SVM method should use an SF of 5 minutes.

As with the previous experiment, we observed again that the patient models developed using the random forest method achieve a better performance for each different SF considered.

3.8 Exploring the limits of constrained devices in DM1 management. Local computing vs. cloud computing

After these considerations regarding accuracy, if we want to implement these algorithms in a small device with limited computing power, testing a smartphone (Samsung S8+) and a Raspberry Pi 3b+. We have to choose a light algorithm, in order to obtain in each iteration an affordable execution time, even sacrificing some accuracy. In Table 5, second column, we can see the performance of each algorithm developed in the above mentioned desktop/personal computer (PC). As can be seen, it is remarkable that the execution time is practically independent of the PH, the PSW being more influential and, as might be expected, for ARIMA and RF algorithms, the more we increase the PSW, the longer is the required time per each iteration since more data are considered and processed. However, regarding this issue, this time, SVM offers the best performance with another advantage: it is almost independent of the length of the PSW due to the capabilities of this hardware environment. On the other hand, both ARIMA and RF seem to increase exponentially the computational burden when expanding the PSW, making at some point the computation unaffordable. Anyway, the execution time seems to be affordable with the above-mentioned optimum PSW (six hours) for the three algorithms, and pretty similar to the smallest one (three hours), with values that range from 31.22 seconds (ARIMA) to 9.00 seconds (SVM) for a 15-minute prediction. Therefore, we consider that in a desktop computer environment a good compromise is choosing RF as the algorithm with the best performance under these conditions since, as previously mentioned, it achieved the most accurate results with a small standard deviation.

Table 5 also shows the computational performance of RF and SVM in the two constrained devices considered. This time the ARIMA method is discarded because of its high requirements even in a desktop computer and for small PSWs. In view of the results, it can be noticed that, in constrained devices, the data processing requirements soar, up the point that in RF, its exponential growth makes it unaffordable to use a Raspberry Pi for high PSWs: with a 36 h PSW, the 15-minute ahead prediction takes more than 13 minutes to be calculated, which seems unreasonable for an 'on-the-fly' prediction. Focusing on the six-hour PSW, which, as previously mentioned, in the end is the most accurate window, we can observe a better performance with SVM, with a very fast execution with the smartphone (27.47 seconds for a 15-minute prediction) as well as good enough with the Raspberry Pi (89.83 seconds). In any case, it can be concluded that the possibility of executing glucose prediction methods in constrained devices can be considered as realistic.

Table 6 shows different results lowering SF. As can be observed, it is possible to see that our expectations were right. All the algorithms reduce their execution times when lowering the SF (especially when using a Rasperry Pi, where the improvement with this strategy is very remarkable). Specifically, when moving from SF = 5 min to SF = 15 min, RF decreases by more than 75% the execution times in a desktop computer (13.20 seconds to 3.06, in a 15-minute forecast with RF, and from 9.00 to 2.76 seconds in the same case with SVM), and five times in constrained devices (56.27 seconds to 10.19 seconds for the Samsung S8+, in a PH of 15 minutes, and from the exaggerated 172.72 seconds to the restrained 26.52 seconds in a 15-minute forecast in a Raspberry Pi). SVM, the fastest algorithm, also makes an improvement but it is not as prominent since it was already very fast before the SF variation; in any case, using a smartphone like the Samsung s8+, when moving from an SF of five minutes to an SF of fifteen minutes, the execution time decreases from 27.47 to 8.12 seconds (for a 15-minute prediction), which is also an interesting point.

In any case, the interesting point is that just by changing the SF from five minutes to ten minutes, still acceptable RMSEs are achieved while at the same time obtaining a considerable reduction in the CE, this reduction being more marked the less powerful is the hardware used. On the other hand, lowering the SF to one-third of its value results in the SVM, which offered a reasonable accuracy, reaches an unaffordable RMSE when the SF is equal to 15 minutes (23.26 mg/dl), although the CE is really low using a desktop computer, and pretty good with constrained devices. So, SVM is advised to be used with an SF of five minutes. But, even in this case, RF offers a better performance in every SF under discussion, with execution times in the same range when considering SFs of ten and even fifteen minutes.

Table 6. RMSE and Computational Effort of the ARIMA, RF and SVM glycaemia prediction methods, for different PSWs and PHs, when executed in a Personal Computer, a Samsung S8+ smartphone and a Raspberry Pi (all the simulations with SF = 5 min)

***SF* = 5 min**		**RMSE (mg/dl)**				**Comput. Effort PC (seg)**				**Comput. Effort S8+ (seg)**				**Comput. Effort R. Pi (seg)**			
		PH=15	30	45	60 min	***PH***=15	30	45	60 min	***PH***=15	30	45	60 min	***PH***=15	30	45	60 min
ARIMA	3	11.64	17.62	23.21	28.17	30.28	30.32	30.54	30.54								
	6	11.53	17.31	22.75	27.60	31.22	31.32	31.35	31.43								
PSW	12	13.00	18.64	23.92	28.64	36.13	36.23	36.30	36.35								
(hours)=	24	13.93	19.45	24.55	29.05	45.21	45.47	45.39	45.45								
	36	14.78	20.21	25.24	29.68	53.41	53.47	53.27	53.22								
RF	3	10.59	15.07	19.04	22.56	10.20	10.19	9.98	10.00	64.85	47.16	45.64	44.35	130.04	116.88	116.70	119.04
	6	10.15	14.63	18.60	22.12	13.20	13.17	13.03	12.90	56.27	55.06	53.11	53.89	172.72	173.24	174.45	175.96
PSW	12	11.18	15.66	19.64	23.17	19.34	19.59	19.27	19.18	100.20	99.66	99.82	100.33	332.92	337.47	332.71	328.83
(hours)=	24	11.65	16.11	20.06	23.56	32.16	32.33	32.32	32.28	205.38	175.43	179.32	178.74	577.61	593.09	594.56	576.60
	36	11.98	16.43	20.38	23.89	48.07	47.78	47.72	47.38	263.92	248.95	252.90	259.74	851.90	890.00	840.23	850.28
SVM	3	18.14	22.55	26.46	29.74	8.60	9.01	8.66	8.26	29.91	27.44	27.74	27.66	91.19	89.44	89.41	89.28
	6	17.65	20.82	23.74	26.36	9.00	8.77	8.62	9.14	27.47	27.01	26.96	27.06	88.93	89.73	89.26	89.28
PSW	12	19.73	22.57	25.17	27.48	9.20	9.06	8.92	8.98	27.92	27.71	27.69	27.50	91.93	92.12	92.04	92.13
(hours)=	24	21.45	26.73	30.71	22.68	9.63	10.02	10.07	10.09	33.52	33.21	33.08	32.77	109.16	109.08	108.85	108.27
	36	23.08	27.90	31.30	33.90	11.63	11.47	11.56	11.35	41.04	41.86	40.69	41.20	142.83	142.87	143.54	143.28

Table 7. RMSE and Computational Effort of the RF and SVM glycaemia prediction methods, for different SFs and PHs, when executed in a Personal Computer, a Samsung S8+ smartphone and a Raspberry Pi (all the simulations with PSW = 6 hour)

***PSW* = 6h**		**RMSE (mg/dl)**				**Comput. Effort PC (seg)**				**Comput. Effort S8+ (seg)**				**Comput. Effort R. Pi (seg)**			
		PH=15	30	45	60 min	***PH***=15	30	45	60 min	***PH***=15	30	45	60 min	***PH***=15	30	45	60 min
RF	5	10.15	14.63	18.60	22.12	13.20	13.17	13.03	12.90	56.27	55.06	53.11	53.89	172.72	173.24	174.45	175.96
SF	10	11.65	17.37	19.84	24.26	4.64	4.53	4.61	4.51	16.15	16.00	16.10	16.03	56.49	56.83	56.76	57.36
(min)=	15	15.43	19.41	22.87	25.92	3.06	3.00	2.96	3.06	10.19	10.22	9.86	9.77	26.52	26.59	26.57	26.50
SVM	5	17.65	20.82	23.74	26.36	9.00	8.77	8.62	9.14	27.47	27.01	26.96	27.06	89.83	89.73	89.26	89.28
SF	10	19.9	23.97	25.73	28.9	4.10	3.90	3.98	3.92	12.05	12.03	12.07	12.09	34.89	35.47	37.24	35.32
(min)=	15	23.26	26.03	28.44	30.57	2.76	2.79	2.80	2.75	8.12	8.07	8.10	8.27	25.02	25.11	25.03	24.96

3.9 Lessons learned

The proliferation of ICT solutions (IoT among them) represents new opportunities for the development of new intelligent services, contributing to more efficient and sustainable management of many different aspects of our society. Helping people in the field of health is a remarking issue to be taken into account. Being in these days Diabetes Mellitus reaching heights increasingly, it is critical to lead our efforts to apply ICT to help people with diabetes.

Given the constant evolution of technologies and protocols that are emerging for the IoT, in the scope of this thesis, we have proposed the definition of an architectural framework that abstracts from the underlying technologies for managing different requirements of a person with diabetes. The proposed design enables connectivity between patient and caregivers, and allows the patient to take advantage of optimized solutions. In addition, within this framework, we have introduced novel features to be taken into account.

Another contribution of this thesis is the development of a proper data set, which is a novelty itself, since it engaged a quantity of subjects, for such a period of time, and with a completed monitoring, that makes this novel data set, to the best of the authors' knowledge, as the most complete until now. This data set will open the door to future works.

At the end, the study of forecasting glycaemia in a simple and affordable way was another novelty exposed in this thesis, considering the idea of performing data modeling and prediction in a constrained device. For this purpose, we used simple and well-known univariate algorithms. Experimental evidences show that such forecasting is possible under some circumstances.

All in all, this work describes novel ways to continue developing paths to properly manage diabetes by using technologies and machine learning techniques, making patients' lives easier while at the same time lowering the risks associated to this illness. For such purposes, in addition to the efforts that have been taken place in many universities and research centers in the world, we have humbly contributed with the results presented in this PhD thesis that here concludes.

Chapter 4

Publications composing the PhD Thesis

4.1 Towards an ICT-Based Platform for Type 1 Diabetes Mellitus Management

Title	Towards an ICT-Based Platform for Type 1 Diabetes Mellitus Management.
Authors	Rodríguez-Rodríguez, Ignacio; Zamora-Izquierdo, Miguel Ángel and José-Víctor Rodríguez.
Type	Journal
Journal	Applied Sciences
Impact factor (2018)	2.217
Publisher	MDPI
Year	2018
ISSN	2076-3417
DOI	https://doi.org/10.3390/app8040511
URL	https://www.mdpi.com/2076-3417/8/4/511
State	Published

Journal details:	
Academic Editor:	Scofield Wang
ISSN:	2076-3417
Publisher:	MDPI
Impact factor (2018):	2.217
Classification	Applied Physics
Rank and Quartile	67/148 (Q2)
Website:	https://www.mdpi.com/journal/applsci

Authors – Personal details:	
Name Position University	Ignacio Rodríguez-Rodríguez PhD student of the Department of Information and Communications Engineering University of Murcia
Name Position University	Dr. Miguel Ángel Zamora Izquierdo Full Professor of the Department of Information and Communications Engineering University of Murcia
Name Position University	Dr. José-Víctor Rodríguez Associate Professor of the Department of Information and Communications Technologies Technical University of Cartagena

Contribution of the PhD student
The PhD student, Ignacio Rodríguez-Rodríguez, declares to be the main author and the major contributor of the paper Towards an ICT-Based Platform for Type 1 Diabetes Mellitus Management.

Abstract
Type 1 Diabetes Mellitus (DM1) is a metabolic disease that is characterized by chronic hyperglycemia due to a lack of pancreatic insulin production. This forces patients to perform several blood glucose measurements per day—by means of capillary glucometers—in order to infer a trend and try to predict future values. In this way, a decision about the insulin dosage that has to be exogenously injected to maintain glycemia within the desirable levels is made. Unfortunately, this method usually suffers from relatively high imprecision. However, recent advances in information and communication technologies (ICT), along with novel biosensors that could provide a real-time comprehensive condition of the patient, offer a new perspective in DM1 management. In this sense, new disruptive technologies like Big Data, the Internet of Things (IoT), and Cloud Computing, as well as Machine Learning (ML) can play an important role in managing DM1. In this work, firstly, an analysis of previously published ICT-based methods for the management of diabetes continuous monitoring is carried out. In this way, an assessment of the possible lack of such proposals is presented, along with the challenges to be overcome in forthcoming smart DM1 management systems. Finally, an overview of a holistic ICT-based platform for DM1 management that try to solve the limitations of previous works, while at the same time, taking advantage of the abovementioned disruptive technologies is hereby proposed.

4.2 Variables to Be Monitored via Biomedical Sensors for Complete Type 1 Diabetes Mellitus Management: An Extension of the "On-Board" Concept

Title	Variables to Be Monitored via Biomedical Sensors for Complete Type 1 Diabetes Mellitus Management: An Extension of the "On-Board" Concept.
Authors	Rodríguez-Rodríguez, Ignacio; Rodríguez, José-Víctor; and Zamora-Izquierdo, Miguel Ángel.
Type	Journal
Journal	Journal of Diabetes Research
Impact factor (2018)	3.040
Publisher	Hindawi
Year	2018
ISSN	2314-6745
DOI	https://doi.org/10.1155/2018/4826984
URL	https://www.hindawi.com/journals/jdr/2018/4826984/
State	Published

Journal details:	
Academic Editor:	Michaelangela Barbieri
ISSN:	2314-6745
Publisher:	Hindawi
Impact factor (2018):	3.040
Classification	Medicine, Research and Experimental
Rank and Quartile	62/136 (Q2)
Website:	https://www.hindawi.com/journals/jdr/

Authors – Personal details:	
Name Position University	Ignacio Rodríguez-Rodríguez PhD student of the Department of Information and Communications Engineering University of Murcia
Name Position University	Dr. José-Víctor Rodríguez Associate Professor of the Department of Information and Communications Technologies Technical University of Cartagena
Name Position University	Dr. Miguel Ángel Zamora Izquierdo Full Professor of the Department of Information and Communications Engineering University of Murcia

Contribution of the PhD student
The PhD student, Ignacio Rodríguez-Rodríguez, declares to be the main author and the major contributor of the paper Variables to Be Monitored via Biomedical Sensors for Complete Type 1 Diabetes Mellitus Management: An Extension of the “On-Board” Concept.

Abstract

Type 1 diabetes mellitus (DM1) is a growing disease, and a deep understanding of the patient is required to prescribe the most appropriate treatment, adjusted to the patient's habits and characteristics. Before now, knowledge regarding each patient has been incomplete, discontinuous, and partial. However, the recent development of continuous glucose monitoring (CGM) and new biomedical sensors/gadgets, based on automatic continuous monitoring, offers a new perspective on DM1 management, since these innovative devices allow the collection of 24-hour biomedical data in addition to blood glucose levels. With this, it is possible to deeply characterize a diabetic person, offering a better understanding of his or her illness evolution, and, going further, develop new strategies to manage DM1. This new and global monitoring makes it possible to extend the "on-board" concept to other features. This well-known approach to the processing of variable "insulin" describes some inertias and aggregated/remaining effects. In this work, such analysis is carried out along with a thorough study of the significant variables to be taken into account/monitored—and how to arrange them—for a deep characterization of diabetic patients. Lastly, we present a case study evaluating the experience of the continuous and comprehensive monitoring of a diabetic patient, concluding that the huge potential of this new perspective could provide an acute insight into the patient's status and extract the maximum amount of knowledge, thus improving the DM1 management system in order to be fully functional.

4.3 Feature Selection for Blood Glucose Level Prediction in Type 1 Diabetes Mellitus by Using the Sequential Input Selection Algorithm (SISAL)

Title	Feature Selection for Blood Glucose Level Prediction in Type 1 Diabetes Mellitus by Using the Sequential Input Selection Algorithm (SISAL)
Authors	Rodríguez-Rodríguez, Ignacio; Rodríguez, José-Víctor; González-Vidal, Aurora and Zamora-Izquierdo, Miguel Ángel.
Type	Journal
Journal	Symmetry
Impact factor (2018)	2.143
Publisher	MDPI
Year	2019
ISSN	2073-8994
DOI	https://doi.org/10.3390/sym11091164
URL	https://www.mdpi.com/2073-8994/11/9/1164
State	Published

Journal details:	
Academic Editor:	Teresa Zhang
ISSN:	2073-8994
Publisher:	MDPI
Impact factor (2018):	2.143
Classification	Multidisciplinary Sciences
Rank and Quartile	30/69 (Q2)
Website:	https://www.mdpi.com/journal/symmetry

Authors – Personal details:	
Name Position University	Ignacio Rodríguez-Rodríguez PhD student of the Department of Information and Communications Engineering University of Murcia
Name Position University	Dr. José-Víctor Rodríguez Associate Professor of the Department of Information and Communications Technologies Technical University of Cartagena
Name Position University	Aurora González-Vidal PhD student of the Department of Information and Communications Engineering University of Murcia
Name Position University	Dr. Miguel Ángel Zamora Izquierdo Full Professor of the Department of Information and Communications Engineering University of Murcia

Contribution of the PhD student
The PhD student, Ignacio Rodríguez-Rodríguez, declares to be the main author and the major contributor of the paper Feature Selection for Blood Glucose Level Prediction in Type 1 Diabetes Mellitus by Using the Sequential Input Selection Algorithm (SISAL)

Abstract

Feature selection is a primary exercise to tackle any forecasting task. Machine learning algorithms used to predict any variable can improve their performance by lessening their computational effort with a proper dataset. Anticipating future glycemia in type 1 diabetes mellitus (DM1) patients provides a baseline in its management, and in this task, we need to carefully select data, especially now, when novel wearable devices offer more and more information. In this paper, a complete characterization of 25 diabetic people has been carried out, registering innovative variables like sleep, schedule, or heart rate in addition to other well-known ones like insulin, meal, and exercise. With this ground-breaking data compilation, we present a study of these features using the Sequential Input Selection Algorithm (SISAL), which is specially prepared for time series data. The results rank features according to their importance, regarding their relevance in blood glucose level prediction as well as indicating the most influential past values to be taken into account and distinguishing features with person-dependent behavior from others with a common performance in any patient. These ideas can be used as strategies to select data for predicting glycemia depending on the availability of computational power, required speed, or required accuracy. In conclusion, this paper tries to analyze if there exists symmetry among the different features that can affect blood glucose levels, that is, if their behavior is symmetric in terms of influence in glycemia.

4.4 Utility of Big Data in Predicting Short-term Blood Glucose Levels in Type 1 Diabetes Mellitus through Machine Learning Techniques

Title	Utility of Big Data in Predicting Short-term Blood Glucose Levels in Type 1 Diabetes Mellitus through Machine Learning Techniques.
Authors	Rodríguez-Rodríguez, Ignacio; Chatzigiannakis, Ioannis; Maranghi, Marianna; Rodríguez, José-Víctor; Gentili, Michele and Zamora-Izquierdo, Miguel Ángel.
Type	Journal
Journal	Sensors
Impact factor (2018)	3.031
Publisher	MDPI
Year	2019
ISSN	1424-8220
DOI	https://doi.org/10.3390/s19204482
URL	https://www.mdpi.com/1424-8220/19/20/4482
State	Published

Journal details:	
Academic Editor:	Grace Feng
ISSN:	1424-8220
Publisher:	MDPI
Impact factor (2018):	3.031
Classification	Instruments and Instrumentation
Rank and Quartile	15/61 Q1
Website:	https://www.mdpi.com/journal/sensors

Authors – Personal details:	
Name Position University	Ignacio Rodríguez-Rodríguez PhD student of the Department of Information and Communications Engineering University of Murcia
Name Position University	Dr. Ioannis Chatziginnakis Lecturer of the Dipartimento di Ingegneria Informatica, Automatica e Gestionale Sapienza Università di Roma
Name Position University	Dra. Marianna Marianghi Assistant Professor of the Dipartamento de Medicina Traslazionale e di Precisione Polytechnical University of Cartagena
Name Position University	Dr. José-Víctor Rodríguez Associate Professor of the Department of Information and Communications Technologies Technical University of Cartagena
Name Position University	Michele Gentili PhD student of the Dipartimento di Ingegneria Informatica, Automatica e Gestionale Sapienza Università di Roma
Name Position University	Dr. Miguel Ángel Zamora Izquierdo Full Professor of the Department of Information and Communications Engineering University of Murcia

Contribution of the PhD student
The PhD student, Ignacio Rodríguez-Rodríguez, declares to be the main author and the major contributor of the paper Utility of Big Data in Predicting Short-term Blood Glucose Levels in Type 1 Diabetes Mellitus through Machine Learning Techniques.

Abstract

Machine learning techniques combined with wearable electronics can deliver accurate short-term blood glucose level prediction models. These models can learn personalized glucose–insulin dynamics based on the sensor data collected by monitoring several aspects of the physiological condition and daily activity of an individual. Until now, the prevalent approach for developing data-driven prediction models was to collect as much data as possible to help physicians and patients optimally adjust therapy. The objective of this work was to investigate the minimum data variety, volume, and velocity required to create accurate person-centric short-term prediction models. We developed a series of these models using different machine learning time series forecasting techniques suitable for execution within a wearable processor. We conducted an extensive passive patient monitoring study in real-world conditions to build an appropriate data set. The study involved a subset of type 1 diabetic subjects wearing a flash glucose monitoring system. We comparatively and quantitatively evaluated the performance of the developed data-driven prediction models and the corresponding machine learning techniques. Our results indicate that very accurate short-term prediction can be achieved by only monitoring interstitial glucose data over a very short time period and using a low sampling frequency. The models developed can predict glucose levels within a 15-min horizon with an average error as low as 15.43 mg/dL using only 24 historic values collected within a period of sex hours, and by increasing the sampling frequency to include 72 values, the average error is reduced to 10.15 mg/dL. Our prediction models are suitable for execution within a wearable device, requiring the minimum hardware requirements while at simultaneously achieving very high prediction accuracy.

4.5 On The Possibility Of Predicting Glycaemia 'On The Fly' With Constrained IoT Devices In Type 1 Diabetes Mellitus Patients

Title	On The Possibility Of Predicting Glycaemia 'On The Fly' With Constrained IoT Devices In Type 1 Diabetes Mellitus Patients
Authors	Rodríguez-Rodríguez, Ignacio; Rodríguez, José-Víctor; Chatzigiannakis, Ioannis and Zamora-Izquierdo, Miguel Ángel.
Type	Journal
Journal	Sensors
Impact factor (2018)	3.031
Publisher	MDPI
Year	2019
ISSN	1424-8220
DOI	https://doi.org/10.3390/s19204538
URL	https://www.mdpi.com/1424-8220/19/20/4538
State	Published

Journal details:	
Academic Editor:	Jayleen Chen
ISSN:	1424-8220
Publisher:	MDPI
Impact factor (2018):	3.031
Classification	Instruments and Instrumentation
Rank and Quartile	15/61 Q1
Website:	https://www.mdpi.com/journal/sensors

Authors – Personal details:	
Name Position University	Ignacio Rodríguez-Rodríguez PhD student of the Department of Information and Communications Engineering University of Murcia
Name Position University	Dr. José-Víctor Rodríguez Associate Professor of the Department of Information and Communications Technologies Technical University of Cartagena
Name Position University	Dr. Ioannis Chatziginnakis Lecturer of the Dipartimento di Ingegneria Informatica, Automatica e Gestionale Sapienza Università di Roma
Name Position University	Dr. Miguel Ángel Zamora Izquierdo Full Professor of the Department of Information and Communications Engineering University of Murcia

Contribution of the PhD student
The PhD student, Ignacio Rodríguez-Rodríguez, declares to be the main author and the major contributor of the paper On The Possibility Of Predicting Glycaemia ‘On The Fly’ With Constrained IoT Devices In Type 1 Diabetes Mellitus Patients

Abstract

Type 1 Diabetes Mellitus (DM1) patients are used to checking their blood glucose levels several times per day through finger sticks and, by subjectively handling this information, to try to predict their future glycaemia in order to choose a proper strategy to keep their glucose levels under control, in terms of insulin dosages and other factors. However, recent Internet of Things (IoT) devices and novel biosensors have allowed the continuous collection of the value of the glucose level by means of Continuous Glucose Monitoring (CGM) so that, with the proper Machine Learning (ML) algorithms, glucose evolution can be modeled, thus permitting a forecast of this variable. On the other hand, glycaemia dynamics require that such a model be user-centric and should be recalculated continuously in order to reflect the exact status of the patient, i.e., an 'on-the-fly' approach. In order to avoid, for example, the risk of being disconnected from the Internet, it would be ideal if this task could be performed locally in constrained devices like smartphones, but this would only be feasible if the execution times were fast enough. Therefore, in order to analyze if such a possibility is viable or not, an extensive, passive, CGM study has been carried out with 25 DM1 patients in order to build a solid dataset. Then, some well-known univariate algorithms have been executed in a desktop computer (as a reference) and two constrained devices: a smartphone and a Raspberry Pi, taking into account only past glycaemia data to forecast glucose levels. The results indicate that it is possible to forecast, in a smartphone, a 15-min horizon with a Root Mean Squared Error (RMSE) of 11.65 mg/dL in just 16.15 s, employing a 10-min sampling of the past 6 h of data and the Random Forest algorithm. With the Raspberry Pi, the computational effort increases to 56.49 s assuming the previously mentioned parameters, but this can be improved to 34.89 s if Support Vector Machines are applied, achieving in this case an RMSE of 19.90 mg/dL. Thus, this paper concludes that local on-the-fly forecasting of glycaemia would be affordable with constrained devices.

Chapter 5

Acceptance letters

Towards an ICT-Based Platform for Type 1 Diabetes Mellitus Management

De: Scofield Wang
Enviado: sábado, 24 de marzo de 2018 18:16
Para: Ignacio Rodriguez-Rodriguez
CC: Miguel A. Zamora; José-Víctor Rodríguez; Applied Sciences Editorial Office; Scofield Wang
Asunto: [Applied Sciences] Manuscript ID: applsci-263093 - Accepted forPublication

Dear Mr. Rodriguez-Rodriguez,

We are pleased to inform you that the following paper has been officially accepted for publication:

Manuscript ID: applsci-263093
Type of manuscript: Review
Title: Towards an ICT-Based Platform for Type 1 Diabetes Mellitus Management
Authors: Ignacio Rodriguez-Rodriguez *, Miguel A. Zamora *, José-Víctor Rodríguez *
Received: 30 December 2017
E-mails: ignacio.rodriguez1@um.es, mzamora@um.es, jvictor.rodriguez@upct.es
Submitted to section: Computer Science and Electrical Engineering,
http://www.mdpi.com/journal/applsci/sections/computer_sci
Advanced Internet of Things for Smart Infrastructure System
http://www.mdpi.com/journal/applsci/special_issues/IoT_Smart
http://susy.mdpi.com/user/manuscripts/review_info/af0f94a09776b70a43f7db094626cf3d

We will now make the final preparations for publication, then return the manuscript to you for your approval.

Kind regards,
Mr. Scofield Wang
Assistant Editor
E-Mail: scofield.wang@mdpi.com

MDPI
Applied Sciences Editorial Office
Postfach, CH-4020 Basel, Switzerland
Office: St. Alban-Anlage 66, 4052 Basel, Switzerland
Tel.: +41 61 683 77 34; Fax: +41 61 302 89 18
E-Mail: applsci@mdpi.com
http://www.mdpi.com/journal/applsci/

Full reference:
Rodríguez-Rodríguez, I., Zamora-Izquierdo, M. Á., & Rodríguez, J. V. (2018). Towards an ICT-based platform for type 1 diabetes mellitus management. *Applied Sciences*, *8*(4), 511.

Variables to Be Monitored via Biomedical Sensors for Complete Type 1 Diabetes Mellitus Management: An Extension of the "On-Board" Concept

De: Michaelangela Barbieri
Enviado: jueves, 9 de agosto de 2018 13:32
Para: ignacio.rodriguez1@um.es
CC: michelangela.barbieri@unicampania.it; jvictor.rodriguez@upct.es; mzamora@um.es
Asunto: 4826984: Your manuscript has been accepted

Dear Dr. Rodríguez-Rodríguez,

The review process of Review Article 4826984 titled "Variables to be monitored via biomedical sensors for complete type 1 Diabetes Mellitus management. An extension of the 'On-Board' concept. " by Ignacio Rodríguez-Rodríguez, José-Víctor Rodríguez and Miguel A. Zamora submitted to Journal of Diabetes Research has been completed. I am pleased to inform you that your manuscript has now been accepted for publication in the journal.

The publication process of your manuscript will be initiated upon the receipt of electronic files. Please log in to the Manuscript Tracking System at the link below using your username and password, and upload the electronic files of your final accepted version within the next 2-3 days.

http://mts.hindawi.com/author/4826984/upload.files/

The electronic files should include the following:

1- Source file of the final accepted manuscript (Word or TeX/LaTeX).
2- PDF file of the final accepted manuscript.
3- Editable figure files (each figure in a separate EPS/PostScript/Word file) if any, taking into consideration that TIFF, JPG, JPEG, BMP formats are not editable.

Thank you again for submitting your manuscript to Journal of Diabetes Research.

Best regards,

Michaelangela Barbieri
michelangela.barbieri@unicampania.it

Full reference:
Rodríguez-Rodríguez, I., Rodríguez, J. V., & Zamora-Izquierdo, M. Á. (2018). Variables to Be Monitored via Biomedical Sensors for Complete Type 1 Diabetes Mellitus Management: An Extension of the "On-Board" Concept. *Journal of diabetes research, 2018.*

Feature Selection for Blood Glucose Level Prediction in Type 1 Diabetes Mellitus by Using the Sequential Input Selection Algorithm (SISAL)

De: Teresa Zhang
Enviado: miércoles, 11 de septiembre de 2019 10:00
Para: Ignacio Rodriguez-Rodriguez
CC: José-Víctor Rodríguez; Aurora González-Vidal; Miguel-Ángel Zamora; Symmetry Editorial Office; Teresa Zhang
Asunto: [Symmetry] Manuscript ID: symmetry-587405 - Accepted for Publication

Dear Mr. Rodriguez-Rodriguez,

We are pleased to inform you that the following paper has been officially accepted for publication:

Manuscript ID: symmetry-587405
Type of manuscript: Article
Title: Feature Selection for Blood Glucose Level Prediction in Type 1 Diabetes Mellitus by using the Sequential Input Selection Algorithm (SISAL)
Authors: Ignacio Rodriguez-Rodriguez *, José-Víctor Rodríguez, Aurora González-Vidal, Miguel-Ángel Zamora
Received: 19 August 2019
E-mails: ignacio.rodriguez1@um.es, jvictor.rodriguez@upct.es, auroragonzalez@um.es, m.zamora@um.es

https://susy.mdpi.com/user/manuscripts/review_info/cc7de48a50152832de6caaf95215d889

We will now make the final preparations for publication, then return the manuscript to you for your approval.

If, however, extensive English edits are required to your manuscript, we will need to return the paper requesting improvements throughout.

We encourage you to set up your profile at SciProfiles.com, MDPI's researcher network platform. Articles you publish with MDPI will be linked to your SciProfiles page, where colleagues and peers will be able to see all of your publications, citations, as well as your other academic contributions.

We also invite you to contribute to Encyclopedia (https://encyclopedia.pub), a scholarly platform providing accurate information about the latest research results. You can adapt parts of your paper to provide valuable reference information for others in the field.

Kind regards,
Teresa Zhang
Assistant Editor

Full reference:
Rodríguez-Rodríguez, I., Rodríguez, J. V., González-Vidal, Aurora & Zamora-Izquierdo, M. Á. (2019). Feature Selection for Blood Glucose Level Prediction in Type 1 Diabetes Mellitus by Using the Sequential Input Selection Algorithm (SISAL). *Symmetry*, *2019*, 11, 1164.

Utility of Big Data in Predicting Short-term Blood Glucose Levels in Type 1 Diabetes Mellitus through Machine Learning Techniques

De: Grace Feng
Enviado: martes, 15 de octubre de 2019 4:14
Para: Ignacio Rodriguez-Rodriguez
CC: Ioannis Chatzigiannakis; José-Víctor Rodríguez; Marianna Maranghi; Michele Gentili; Miguel-Ángel Zamora-Izquierdo; Sensors Editorial Office; Grace Feng
Asunto: [Sensors] Manuscript ID: sensors-593890 - Accepted for Publication

Dear Mr. Rodriguez-Rodriguez,

We are pleased to inform you that the following paper has been officially accepted for publication:

Manuscript ID: sensors-593890
Type of manuscript: Article
Title: On The Utility of Big Data in Predicting Short-term Blood Glucose Levels in Type 1 Diabetes Mellitus through Machine Learning Techniques
Authors: Ignacio Rodriguez-Rodriguez *, Ioannis Chatzigiannakis, José-Víctor Rodríguez, Marianna Maranghi, Michele Gentili, Miguel-Ángel Zamora-Izquierdo
Received: 28 August 2019
E-mails: ignacio.rodriguez1@um.es, ichatz@diag.uniroma1.it, jvictor.rodriguez@upct.es, marianna.maranghi@uniroma1.it, gentili@diag.uniroma1.it, mzamora@um.es
Submitted to section: Intelligent Sensors,
https://www.mdpi.com/journal/sensors/sections/Intelligent_Sensors
https://susy.mdpi.com/user/manuscripts/review_info/dc4ea7a504d1054d9c8efcb7e0ea27f6

We will now make the final preparations for publication, then return the manuscript to you for your approval.

If, however, extensive English edits are required to your manuscript, we will need to return the paper requesting improvements throughout.

We encourage you to set up your profile at SciProfiles.com, MDPI's researcher network platform. Articles you publish with MDPI will be linked to your SciProfiles page, where colleagues and peers will be able to see all of your publications, citations, as well as your other academic contributions.

We also invite you to contribute to Encyclopedia (https://encyclopedia.pub), a scholarly platform providing accurate information about the latest research results. You can adapt parts of your paper to provide valuable reference information for others in the field.

Full reference:
Rodríguez-Rodríguez, I.; Chatzigiannakis, I.; Rodríguez, J.-V.; Maranghi, M.; Gentili, M.; Zamora-Izquierdo, M.-Á. Utility of Big Data in Predicting Short-Term Blood Glucose Levels in Type 1 Diabetes Mellitus Through Machine Learning Techniques. *Sensors 2019*, 19, 4482.

On The Possibility Of Predicting Glycaemia 'On The Fly' With Constrained IoT Devices In Type 1 Diabetes Mellitus Patients

De: Jayleen Chen
Enviado: jueves, 17 de octubre de 2019 5:39
Para: Ignacio Rodriguez-Rodriguez
CC: José-Víctor Rodríguez; Ioannis Chatzigiannakis; Miguel-Ángel Zamora; Sensors Editorial Office; Jayleen Chen
Asunto: [Sensors] Manuscript ID: sensors-600875 - Accepted for Publication

Dear Mr. Rodriguez-Rodriguez,

We are pleased to inform you that the following paper has been officially accepted for publication:

Manuscript ID: sensors-600875
Type of manuscript: Article
Title: On The Possibility Of Predicting Glycaemia 'On The Fly' With Constrained IoT Devices In Type 1 Diabetes Mellitus Patients
Authors: Ignacio Rodriguez-Rodriguez *, José-Víctor Rodríguez *, Ioannis Chatzigiannakis *, Miguel-Ángel Zamora *
Received: 6 September 2019
E-mails: ignacio.rodriguez1@um.es, jvictor.rodriguez@upct.es, ichatz@diag.uniroma1.it, mzamora@um.es
Submitted to section: Intelligent Sensors,
https://www.mdpi.com/journal/sensors/sections/Intelligent_Sensors
https://susy.mdpi.com/user/manuscripts/review_info/ce4871eae70927fde4dbfc5750ad5a82

We will now make the final preparations for publication, then return the manuscript to you for your approval.

If, however, extensive English edits are required to your manuscript, we will need to return the paper requesting improvements throughout.

We encourage you to set up your profile at SciProfiles.com, MDPI's researcher network platform. Articles you publish with MDPI will be linked to your SciProfiles page, where colleagues and peers will be able to see all of your publications, citations, as well as your other academic contributions.

We also invite you to contribute to Encyclopedia (https://encyclopedia.pub), a scholarly platform providing accurate information about the latest research results. You can adapt parts of your paper to provide valuable reference information for others in the field.

Kind regards,
Ms. Jayleen Chen

Full reference:
Rodríguez-Rodríguez, I.; Rodríguez, J.-V.; Chatzigiannakis, I.; Zamora Izquierdo, M.Á. (2019) On the Possibility of Predicting Glycaemia 'On the Fly' with Constrained IoT Devices in Type 1 Diabetes Mellitus Patients.. *Sensors 2019*, 19, 4538.

Chapter 6

Favorable Report of the Research Ethics Committee of the University of Murcia

UNIVERSIDAD DE MURCIA | Vicerrectorado de Investigación

CEI Comisión de Ética de Investigación

CAMPUS MARE NOSTRUM

INFORME DE LA COMISIÓN DE ÉTICA DE INVESTIGACIÓN DE LA UNIVERSIDAD DE MURCIA

Jaime Peris Riera, Catedrático de Universidad y Secretario de la Comisión de Ética de Investigación de la Universidad de Murcia

CERTIFICA:

Que D. Ignacio Rodríguez Rodríguez ha presentado la Tesis Doctoral titulada *"Sistemas de gestión integral inteligente para el control de la Diabetes Mellitus T1 basado en las tecnologías de Internet de las cosas",* dirigida por el Dr. D. Miguel Ángel Zamora Izquierdo y el Dr. D. José Víctor Rodríguez Rodríguez, a la Comisión de Ética de Investigación.

Que dicha Comisión analizó toda la documentación presentada, y de conformidad con lo acordado el día 30 de octubre de 2017, por unanimidad, se emite INFORME FAVORABLE, desde el punto de vista ético de la investigación.

Y para que conste y tenga los efectos que correspondan, firmo esta certificación, con el visto bueno del Presidente de la Comisión

Vº Bº
EL PRESIDENTE DE LA COMISIÓN
DE ÉTICA DE INVESTIGACIÓN DE LA
UNIVERSIDAD DE MURCIA

Fdo.: Antonio Juan García Fernández

ID: 1683/2017

Chapter 7

Bibliography

7.1 Publications by the author

[1] I. Rodríguez-Rodríguez, M. Á. Zamora and J.-V. Rodríguez, "Towards a New Diabetes Mellitus Management by Means of Novel Biosensors and Information and Communication Technologies," in *Proceedings of the 2017 4th International Conference on Biomedical and Bioinformatics Engineering*, 2017.

[2] I. Rodríguez-Rodríguez, M. Á. Zamora and J.-V. Rodríguez, "Una propuesta de gestión de la diabetes mellitus tipo 1 a través de comunicaciones inalámbricas," in *XXXII Simposium Nacional de la Unión Científica Internacional de Radio (URSI)*, Cartagena, 2017.

[3] I. Rodríguez-Rodríguez, M.-Á. Zamora-Izquierdo and J.-V. Rodríguez, "Towards an ICT-based platform for type 1 diabetes mellitus management," *Applied Sciences,* vol. 8, p. 511, 2018.

[4] I. Rodríguez-Rodríguez, M. A. Z. Izquierdo and J. V. R. Rodríguez, "Biosensores integrados en un entorno de internet de las cosas. Una gestión inteligente de la diabetes mellitus tipo 1," in *Jornadas doctorales de la Universidad de Murcia*, 2019.

[5] I. Rodríguez-Rodríguez, J.-V. Rodríguez and M.-Á. Zamora-Izquierdo, "Variables to Be Monitored via Biomedical Sensors for Complete Type 1 Diabetes Mellitus Management: An Extension of the “On-Board” Concept," *Journal of diabetes research,* vol. 2018, 2018.

[6] I. Rodríguez-Rodríguez and J.-V. Rodríguez, "How can we tackle gender-based violence in cyberspace?," in *UNIRE Gender Violence is also a Cultural Issue*, Trento, 2019.

[7] I. Rodríguez-Rodríguez, J.-V. Rodríguez, A. González-Vidal and M.-Á. Zamora, "Feature Selection for Blood Glucose Level Prediction in Type 1 Diabetes Mellitus by Using the Sequential Input Selection Algorithm (SISAL)," *Symmetry,* vol. 11, p. 1164, 2019.

[8] I. Rodríguez-Rodríguez, I. Chatzigiannakis, J.-V. Rodríguez; M. Maranghi, M. Gentili, M.-Á. Zamora-Izquierdo. "Utility of Big Data in Predicting Short-Term Blood Glucose Levels in Type 1 Diabetes Mellitus Through Machine Learning Techniques", *Sensors*, vol. 19, p. 4482, 2019.

[9] I. Rodríguez-Rodríguez, J.-V. Rodríguez, I. Chatzigiannakis, M.-Á. Zamora Izquierdo. "On the Possibility of Predicting Glycaemia 'On the Fly' with Constrained IoT Devices in Type 1 Diabetes Mellitus Patients " *Sensors,*vol. 19, p. 4538, 2019.

[10] I. Rodríguez-Rodríguez, M.-T. M. Inglés and J.-V. Rodríguez, "On the possibility of estimating the multiple-diffraction losses of a rectangular obstacle at mm-wave frequencies from the corresponding double knife-edge results," *AEU-International Journal of Electronics and Communications,* vol. 82, pp. 516-519, 2017.

[11] I. Rodríguez-Rodríguez, J.-V. Rodríguez and C. Sanchis-Borrás, "On the impact of the type of wave incidence in multiple-cylinder diffraction analysis at millimeter-wave frequencies," *Journal of Electromagnetic Waves and applications,* vol. 32, pp. 572-578, 2018.

[12] I. Rodríguez-Rodríguez, J.-V. Rodríguez and L. Juan-Llácer, "Validation with measurements of plane and spherical-wave UTD-PO propagation models which assume flat-topped obstacles," *AEU-International Journal of Electronics and Communications,* vol. 85, pp. 174-178, 2018.

[13] I. Rodríguez-Rodríguez, M. Á. Zamora and J.-V. Rodríguez, "On predicting glycaemia in type 1 diabetes mellitus patients by using support vector machines," in *Proceedings of the 1st International Conference on Internet of Things and Machine Learning*, 2017.

[14] I. Rodríguez-Rodríguez, A. González Vidal, A. Ramallo González and M. Zamora, "Commissioning of the controlled and automatized testing facility for human behavior and control (CASITA)," *Sensors,* vol. 18, p. 2829, 2018.

[15] I. Rodríguez-Rodríguez, A. González Vidal, A. Ramallo González and M. Zamora, "On the possibilities of Controlled and Automatised teStIng FaciliTy for human behAviour and control (CASITA) An overview," in *2017 1st International Conference of Data for Low Energy Buildings*, Murcia, 2018.

[16] J.-V. Rodríguez, I. Rodríguez-Rodríguez and J. L. S. Martínez, "Propuestas para Medir la Velocidad de la Luz en Diferentes Medios dentro de un Ámbito Docente," *Elektron,* vol. 1, pp. 66-74, 2017.

[17] A. Eliozondo-Moreno, J.-V. Rodríguez and I. Rodríguez-Rodríguez, "La importancia de la emoción en el aprendizaje: Propuestas para mejorar la motivación de los estudiantes," *Cuaderno de Pedagogía Universitaria,* vol. 15, no. 29, pp. 3-11, 2018.

[18] J. V. R. Rodríguez, I. R. Rodríguez and D. C. Rebenaque, Tests resueltos de teoría y prácticas de comunicaciones ópticas (II)., Universidad Politécnica de Cartagena, 2017.

[19] J. V. R. Rodríguez, I. R. Rodríguez and D. C. Rebenaque, Problemas resueltos de comunicaciones ópticas (II)., Universidad Politécnica de Cartagena, 2017.

7.2 References

[20] M. Ruiz-Ramos, A. Escolar-Pujolar, E. Mayoral-Sánchez, F. Corral-San Laureano and I. Fernández-Fernández, "La diabetes mellitus en España: mortalidad, prevalencia, incidencia, costes económicos y desigualdades," *Gaceta Sanitaria,* vol. 20, pp. 15-24, 2006.

[21] A. Sicras-Mainar, R. Navarro-Artieda and J. Ibanez-Nolla, "Clinical and economic characteristics associated with type 2 diabetes," *Revista Clínica Española (English Edition),* vol. 214, pp. 121-130, 2014.

[22] J. Q. Purnell, B. H. Braffett, B. Zinman, R. A. Gubitosi-Klug, W. Sivitz, J. P. Bantle, G. Ziegler, P. A. Cleary, J. D. Brunzell, D. D. I. C. R. Group and others, "Impact of excessive weight gain on cardiovascular outcomes in type 1 diabetes: results from the Diabetes Control and Complications Trial/Epidemiology of Diabetes Interventions and Complications (DCCT/EDIC) study," *Diabetes care,* vol. 40, pp. 1756-1762, 2017.

[23] A. I. Adler, R. J. Stevens, S. E. Manley, R. W. Bilous, C. A. Cull, R. R. Holman, U. K. P. D. S. Group and others, "Development and progression of nephropathy in type 2 diabetes: the United Kingdom Prospective Diabetes Study (UKPDS 64)," *Kidney international,* vol. 63, pp. 225-232, 2003.

[24] G. Cappon, G. Acciaroli, M. Vettoretti, A. Facchinetti and G. Sparacino, "Wearable continuous glucose monitoring sensors: a revolution in diabetes treatment," *Electronics,* vol. 6, p. 65, 2017.

[25] S. Ding and M. Schumacher, "Sensor monitoring of physical activity to improve glucose management in diabetic patients: a review," *Sensors,* vol. 16, p. 589, 2016.

[26] T. Broens, A. Van Halteren, M. Van Sinderen and K. Wac, "Towards an application framework for context-aware m-health applications," *International Journal of Internet Protocol Technology,* vol. 2, pp. 109-116, 2007.

[27] M. R. Yuce, "Implementation of wireless body area networks for healthcare systems," *Sensors and Actuators A: Physical,* vol. 162, pp. 116-129, 2010.

[28] D. Preuveneers and Y. Berbers, "Mobile phones assisting with health self-care: a diabetes case study," in *Proceedings of the 10th international conference on Human computer interaction with mobile devices and services*, 2008.

[29] J. C. Sieverdes, F. Treiber, C. Jenkins and K. Hermayer, "Improving diabetes management with mobile health technology," *The American journal of the medical sciences,* vol. 345, pp. 289-295, 2013.

[30] M. A. Al-Taee, W. Al-Nuaimy, A. Al-Ataby, Z. J. Muhsin and S. N. Abood, "Mobile health platform for diabetes management based on the Internet-of-Things," in *2015 IEEE Jordan Conference on Applied Electrical Engineering and Computing Technologies (AEECT)*, 2015.

[31] W. C. Hsu, K. H. K. Lau, R. Huang, S. Ghiloni, H. Le, S. Gilroy, M. Abrahamson and J. Moore, "Utilization of a cloud-based diabetes management program for insulin initiation and titration enables collaborative decision making between healthcare providers and patients," *Diabetes technology & therapeutics,* vol. 18, pp. 59-67, 2016.

[32] R. H. Dolin, L. Alschuler, S. Boyer, C. Beebe, F. M. Behlen, P. V. Biron and A. Shabo, "HL7 clinical document architecture, release 2," *Journal of the American Medical Informatics Association,* vol. 13, pp. 30-39, 2006.

[33] H. B. Lim, Y. M. Teo, P. Mukherjee, V. T. Lam, W. F. Wong and S. See, "Sensor grid: integration of wireless sensor networks and the grid," in *The IEEE Conference on Local Computer Networks 30th Anniversary (LCN'05) l*, 2005.

[34] M. S. H. Talpur, M. Z. A. Bhuiyan and G. Wang, "Energy-efficient healthcare monitoring with smartphones and IoT technologies," *International journal of high performance computing and networking,* vol. 8, pp. 186-194, 2015.

[35] A. H. Omre and S. Keeping, "Bluetooth low energy: wireless connectivity for medical monitoring," *Journal of diabetes science and technology,* vol. 4, pp. 457-463, 2010.

[36] L. Leelarathna and E. G. Wilmot, "Flash forward: a review of flash glucose monitoring," *Diabetic Medicine,* vol. 35, pp. 472-482, 2018.

[37] J. D. R. F. C. G. Monitoring, "Effectiveness of continuous glucose monitoring in a clinical care environment: evidence from the Juvenile Diabetes Research Foundation continuous glucose monitoring (JDRF-CGM) trial," *Diabetes Care,* vol. 33, pp. 17-22, 2010.

[38] J. Tikka and J. Hollmén, "Sequential input selection algorithm for long-term prediction of time series," *Neurocomputing,* vol. 71, pp. 2604-2615, 2008.

[39] D. Rodbard, "Optimizing display, analysis, interpretation and utility of self-monitoring of blood glucose (SMBG) data for management of patients with diabetes," *Journal of diabetes science and technology,* vol. 1, pp. 62-71, 2007.

[40] G. C. Estrada, L. Re and E. Renard, "Nonlinear gain in online prediction of blood glucose profile in type 1 diabetic patients," in *49th IEEE Conference on Decision and Control (CDC)*, 2010.

[41] N. Nuryani, S. S. H. Ling and H. T. Nguyen, "Electrocardiographic signals and swarm-based support vector machine for hypoglycemia detection," *Annals of biomedical engineering,* vol. 40, pp. 934-945, 2012.

[42] C. Marling, L. Xia, R. Bunescu and F. Schwartz, "Machine learning experiments with noninvasive sensors for hypoglycemia detection," in *Proceedings of IJCAI Workshop on Knowledge Discovery in Healthcare Data. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA*, 2016.

[43] S. M. Pappada, B. D. Cameron, P. M. Rosman, R. E. Bourey, T. J. Papadimos, W. Olorunto and M. J. Borst, "Neural network-based real-time prediction of glucose in patients with insulin-dependent diabetes," *Diabetes technology & therapeutics,* vol. 13, pp. 135-141, 2011.

[44] C. Zecchin, A. Facchinetti, G. Sparacino, G. De Nicolao and C. Cobelli, "Neural network incorporating meal information improves accuracy of short-time prediction of glucose concentration," *IEEE transactions on biomedical engineering,* vol. 59, pp. 1550-1560, 2012.

[45] G. Sparacino, F. Zanderigo, S. Corazza, A. Maran, A. Facchinetti and C. Cobelli, "Glucose concentration can be predicted ahead in time from continuous glucose monitoring sensor time-series," *IEEE Transactions on biomedical engineering,* vol. 54, pp. 931-937, 2007.

[46] M. Eren-Oruklu, A. Cinar, L. Quinn and D. Smith, "Estimation of future glucose concentrations with subject-specific recursive linear models," *Diabetes technology & therapeutics,* vol. 11, pp. 243-253, 2009.

[47] T. Hamdi, J. B. Ali, V. Di Costanzo, F. Fnaiech, E. Moreau and J.-M. Ginoux, "Accurate prediction of continuous blood glucose based on support vector regression and differential evolution algorithm," *Biocybernetics and Biomedical Engineering,* vol. 38, pp. 362-372, 2018.

[48] M. A. Taie, I. El-Faramawy and M. Elmawazini, "Methods for Prediction, Simulation and Verification of Real-Time Software Architectural Design based on Machine Learning Algorithms," 2015.

[49] N. Williams, S. Zander and G. Armitage, "A preliminary performance comparison of five machine learning algorithms for practical IP traffic flow classification," *ACM SIGCOMM Computer Communication Review,* vol. 36, pp. 5-16, 2006.

[50] P. K. Douglas, S. Harris, A. Yuille and M. S. Cohen, "Performance comparison of machine learning algorithms and number of independent components used in fMRI decoding of belief vs. disbelief," *Neuroimage,* vol. 56, pp. 544-553, 2011.

[51] J. P. Vink and G. Haan, "Comparison of machine learning techniques for target detection," *Artificial Intelligence Review,* vol. 43, pp. 125-139, 2015.

[52] J. Place, A. Robert, N. Ben Brahim, P. Keith-Hynes, A. Farret, M.-J. Pelletier, B. Buckingham, M. Breton, B. Kovatchev and E. Renard, *DiAs web monitoring: a real-time remote monitoring system designed for artificial pancreas outpatient trials,* SAGE Publications Sage CA: Los Angeles, CA, 2013.

[53] E. Dassau, L. Jovanovic, F. J. Doyle III and H. C. Zisser, "Enhanced 911/global position system wizard: a telemedicine application for the prevention of severe hypoglycemia—monitor, alert, and locate," *Journal of diabetes science and technology,* vol. 3, pp. 1501-1506, 2009.

[54] M. Rigla, *Smart telemedicine support for continuous glucose monitoring: the embryo of a future global agent for diabetes care,* SAGE Publications, 2011.

[55] M. A. Case, H. A. Burwick, K. G. Volpp and M. S. Patel, "Accuracy of smartphone applications and wearable devices for tracking physical activity data," *Jama,* vol. 313, pp. 625-626, 2015.

[56] H. Kirchsteiger, R. Johansson, E. Renard and L. d. Re, "Continuous-time interval model identification of blood glucose dynamics for type 1 diabetes," *International Journal of Control,* vol. 87, pp. 1454-1466, 2014.

[57] J. Bondia and J. Vehi, "Physiology-based interval models: a framework for glucose prediction under intra-patient variability," in *Prediction Methods for Blood Glucose Concentration*, Springer, 2016, pp. 159-181.

[58] S. K. Garg, S. A. Weinzimer, W. V. Tamborlane, B. A. Buckingham, B. W. Bode, T. S. Bailey, R. L. Brazg, J. Ilany, R. H. Slover, S. M. Anderson and others, "Glucose outcomes with the in-home use of a hybrid closed-loop insulin delivery system in adolescents and adults with type 1 diabetes," *Diabetes technology & therapeutics,* vol. 19, pp. 155-163, 2017.

[59] D. Rodbard, "Continuous glucose monitoring: a review of successes, challenges, and opportunities," *Diabetes technology & therapeutics,* vol. 18, pp. S2--3, 2016.

[60] M. Gentili, D. Caltabiano, R. Sannino, C. Toffanin, F. Di Palma, L. Magni and S. Lane, "Embedded implementation of modular closed-loop control of diabetes and in silico validation," in *2013 IEEE 15th International Conference on e-Health Networking, Applications and Services (Healthcom 2013)*, 2013.

[61] L. B. Borghouts and H. A. Keizer, "Exercise and insulin sensitivity: a review," *International journal of sports medicine,* vol. 21, pp. 1-12, 2000.

[62] J. T. Devlin, M. Hirshman, E. D. Horton and E. S. Horton, "Enhanced peripheral and splanchnic insulin sensitivity in NIDDM men after single bout of exercise," *Diabetes,* vol. 36, pp. 434-439, 1987.

[63] D. King, P. Baldus, R. Sharp, L. Kesl, T. Feltmeyer and M. Riddle, "Time course for exercise-induced alterations in insulin action and glucose tolerance in middle-aged people," *Journal of applied physiology,* vol. 78, pp. 17-22, 1995.

[64] C. Dalla Man, R. A. Rizza and C. Cobelli, "Meal simulation model of the glucose-insulin system," *IEEE Transactions on biomedical engineering,* vol. 54, pp. 1740-1749, 2007.

[65] M. Housiaux, H. Dorchy and O. Luminet, "Influence of an emotional level of conscience on the glycemic equilibrium in type 1 diabetic children and adolescents," *Diabetes & Metabolism,* vol. 34, p. A33, 2008.

[66] K. Johansen, S. Ellegaard and S. Wex, "Detection of Nocturnal Hypoglycemia in Insulin-treated Diabetics by a Skin Temperature-Skin Conductance Meter," *Acta Medica Scandinavica,* vol. 220, pp. 213-217, 1986.

[67] V. McAulay, I. J. Deary and B. M. Frier, "Symptoms of hypoglycaemia in people with diabetes," *Diabetic Medicine,* vol. 18, pp. 690-705, 2001.

[68] A. Lurbe, J. Redon, J. M. Pascual, J. Tacons, V. Alvarez and D. C. Batlle, "Altered blood pressure during sleep in normotensive subjects with type I diabetes.," *Hypertension,* vol. 21, pp. 227-235, 1993.

[69] NovoNordisk, *Novo Nordisk receives Complete Response Letter in the US for Tresiba® and Ryzodeg®.[Press release],* 2013.

[70] J. Friedman, T. Hastie and R. Tibshirani, The elements of statistical learning, vol. 1, Springer series in statistics New York, 2001.

[71] F. Holleman and J. B. L. Hoekstra, "Insulin lispro," *New England Journal of Medicine,* vol. 337, pp. 176-183, 1997.

[72] E. K. Eskaf, O. Badawi and T. Ritchings, "Predicting blood glucose levels in diabetics using feature extraction and Artificial Neural Networks," in *2008 3rd International Conference on Information and Communication Technologies: From Theory to Applications*, 2008.

[73] L. J. Goodyear and B. B. Kahn, "Exercise, glucose transport, and insulin sensitivity," *Annual review of medicine,* vol. 49, pp. 235-261, 1998.

[74] J. Taelman, S. Vandeput, A. Spaepen and S. Van Huffel, "Influence of mental stress on heart rate and heart rate variability," in *4th European conference of the international federation for medical and biological engineering*, 2009.

[75] E. Donga, M. Dijk, J. G. Dijk, N. R. Biermasz, G.-J. Lammers, K. W. Kralingen, E. P. M. Corssmit and J. A. Romijn, "A single night of partial sleep deprivation induces insulin resistance in multiple metabolic pathways in healthy subjects," *The Journal of Clinical Endocrinology & Metabolism,* vol. 95, pp. 2963-2968, 2010.

[76] S.-q. Shi, T. S. Ansari, O. P. McGuinness, D. H. Wasserman and C. H. Johnson, "Circadian disruption leads to insulin resistance and obesity," *Current Biology,* vol. 23, pp. 372-381, 2013.

[77] G. E. Box, G. M. Jenkins, G. C. Reinsel and G. M. Ljung, *Time series analysis, control, and forecasting . Hoboken,* New Jersey: John Wiley & Sons, 2015.

[78] L. Breiman, "Bagging predictors," *Machine learning,* vol. 24, pp. 123-140, 1996.

[79] V. Vapnik, The nature of statistical learning theory, Springer science & business media, 2013.

[80] E. Van Cauter, K. S. Polonsky and A. J. Scheen, "Roles of circadian rhythmicity and sleep in human glucose regulation," *Endocrine reviews,* vol. 18, pp. 716-738, 1997.