



Tesis Doctoral

**PREDICCIÓN DE GLUCOSA MEDIANTE
TRANSFORMACIONES DEL ESPECTRO AL
CAMPO DE VISIÓN, COMPUTACIÓN EVOLUTIVA
Y APRENDIZAJE PROFUNDO**

JORGE ALVARADO DÍAZ

Programa de Doctorado en Tecnologías Informáticas

Dr. Francisco Chávez de la O

Dr. José Ignacio Hidalgo Pérez

Esta tesis cuenta con la autorización del director/a y codirector/a de la misma y de la Comisión Académica del programa. Dichas autorizaciones constan en el Servicio de la Escuela Internacional de Doctorado de la Universidad de Extremadura.

2024

Agradecimientos

En primer lugar, me gustaría agradecer a mis directores, Francisco Chávez y José Ignacio Hidalgo, por su dedicación en este trabajo así por su conocimiento aportado. Sin ellos no hubiese sido posible este trabajo de tesis.

También me gustaría agradecer a los miembros del Grupo de Evolución Artificial (GEA) de la Universidad de Extremadura, en especial mención a Francisco Fernández de Vega, así como a los miembros del grupo Adaptive and Bioinspired Systems (ABSys) de la Universidad Complutense de Madrid.

Al profesor José Manuel Velasco Cabo, de la Universidad Complutense de Madrid, por la aportación de la idea de la aplicación de la transformada de wavelet para la predicción de la glucosa.

Al profesor Rolando Miragaia, del Instituto Politécnico de Leiria, por su aportación a este trabajo de tesis. Sin su ayuda, no podría haber alcanzado uno de los objetivos de este trabajo de tesis.

También quiero agradecer a todos los compañeros y amigos, que durante estos 4 años que he estado realizando el doctorado, me han ayudado a continuar, sobre todo en los momentos mas difíciles. En especial, quiero agradecer a Abel y a Pacopy.

A las instituciones que parcialmente han financiado mi trabajo:

- Ministerio de Ciencia e Innovación del Gobierno de España, a través del Fondo Europeo de Desarrollo Regional, "Una manera de hacer Europa" y Mecanismo de Resiliencia y Recuperación de la Unión Europea, fondos "NextGenerationEU" - proyectos PID2021-585125549OB-I00, PDC2022-133429-I00, AEI/10.13039/501100011033 y PID2020-115570GB-C21.
- Consejería de Economía e Infraestructuras de la Junta de Extremadura, a través del Fondo Europeo de Desarrollo Regional, "Una manera de hacer Europa" - proyecto GR21108.
- Programas Operativos del Fondo Social Europeo y del Fondo Europeo de Desarrollo Regional de la Comunidad de Madrid - proyecto Y2018/NMT-4668 (Micro-Stress-MAP-CM).

Por ultimo, no menos importante, quiero agradecer este trabajo a mi madre por su apoyo incondicional.

RESUMEN

La Diabetes Mellitus es una enfermedad autoinmune que se caracteriza por unos niveles de glucosa en sangre elevados si no se actúa de manera adecuada. Existen dos tipos, la Diabetes Tipo 1 en la que el páncreas no produce la insulina necesaria para que las células puedan procesar la glucosa y la Diabetes Tipo 2, donde el páncreas sí produce insulina, pero las células receptoras son incapaces de procesarla correctamente para absorber la glucosa. Tanto la carencia de insulina como la imposibilidad de procesarla por parte de las células es un problema de salud para las personas que la padecen. Es necesario mantener los niveles de glucosa en lo que se conoce como rango saludable, no llegando a los extremos, tanto inferiores (Hipoglucemia) como superiores (Hiperglucemia). Para las personas con diabetes es de vital importancia mantener los niveles de glucosa en un rango aceptable, que si es alterado deben tomar acciones correctoras con el fin de evitar situaciones potencialmente peligrosas. Los avances en los últimos años en el ámbito de la Inteligencia Artificial, atendiendo a técnicas de Machine Learning y aprendizaje profundo, están consiguiendo resolver problemas complejos donde es necesario predecir situaciones y/o valores que hasta ahora no eran posible. Esto nos lleva a plantear en este trabajo de tesis la hipótesis de poder generar modelos que puedan predecir situaciones potencialmente peligrosas a personas con diabetes. Se toma como objetivo principal la generación de modelos predictivos de glucosa desde dos enfoques diferentes. Un primer enfoque, donde se busca la predicción del valor exacto de glucosa mediante el uso modelos basados en sistemas de lógica difusa, y un segundo enfoque, donde el objetivo es más ambicioso, predecir situaciones de hipoglucemias e hiperglucemias a futuro, con una ventana temporal superior a la de los modelos existentes en la actualidad. Para este segundo enfoque se han aplicado combinaciones de transformadas de espectro al campo de visión, en concreto funciones de wavelet, y modelos de redes neuronales. De esta forma es posible determinar situaciones potencialmente peligrosas con varias horas de antelación, proporcionando modelos que pueden informar a las personas con diabetes de la posibilidad de sufrir tanto hipoglucemias como hiperglucemias en un futuro. Esto permite tomar acciones correctoras para revertir la situación y no convertirla en una situación crítica para su salud. Los modelos propuestos en esta tesis serán la base para una aplicación móvil que cuente con un sistema de alertas tempranas con el objetivo de que la persona pueda realizar las acciones necesarias para corregir su nivel de glucosa y mejorar su gestión de la diabetes y por lo tanto su calidad de vida.

PALABRAS CLAVE

- Redes Neuronales
- Inteligencia Artificial
- Computación Evolutiva
- Predicción de series temporales
- Predicción de glucosa
- Detección Hipoglucemia y Hiperglucemia

ABSTRACT

Diabetes Mellitus is an autoimmune disease characterized by elevated blood glucose levels if not treated appropriately. There are two types: Type 1 Diabetes, in which the pancreas does not produce the insulin necessary for the cells to process glucose, and Type 2 Diabetes, where the pancreas does produce insulin, but the receptor cells are unable to process it correctly to absorb glucose. The lack of insulin and the inability of the cells to process it is a health problem for people who suffer from it. Maintaining glucose levels in what is known as a healthy range is necessary, not reaching extremes, both lower (hypoglycemia) and higher (hyperglycemia). For people with diabetes, it is of vital importance to maintain glucose levels in an acceptable range, which, if altered, must take corrective actions to avoid potentially dangerous situations. Advances in recent years in Artificial Intelligence, using Machine Learning and deep learning techniques, are managing to solve complex problems where it is necessary to predict situations and/or values that were not possible until now. This leads us to propose in this thesis work the hypothesis of generating models that can predict potentially dangerous situations for people with diabetes. The main objective is the generation of glucose predictive models from two different approaches. A first approach, where the prediction of the exact glucose value is sought by using models based on fuzzy logic systems, and a second approach, where the objective is more ambitious, predicting situations of hypoglycemia and hyperglycemia in the future, with a time window longer than that of currently existing models. For this second approach, combinations of spectrum transforms to the field of view, specifically wavelet functions and neural network models, have been applied. In this way, it is possible to determine potentially dangerous situations several hours in advance, providing models that can inform people with diabetes of the possibility of suffering both hypoglycemia and hyperglycemia in the future. This allows corrective actions to be taken to reverse the situation and not turn it into a critical situation for their health. The models proposed in this thesis will be the basis for a mobile application with an early warning system so that the person can take the necessary actions to correct their glucose level and improve their diabetes management and, therefore, their quality of life.

KEYWORDS

- Neural Networks
- Artificial Intelligence
- Evolutionary Computing
- Time series forecasting

- Glucose prediction
- Hypoglycemia and Hyperglycemia Detection

Lista de acrónimos

- Adam** Adaptive Moment Estimation. 68
- ANN** Artificial Neural Network. 27, 31, 53, 54, 57, XV
- APE** Absolute Percentage Error. 27
- ARIMA** Autoregressive Integrated Moving. 26–28
- AUC** Area Under the ROC Curve. 22, 30, 31, 72–75
- BG** Glucosa en Sangre. 1–3, 10, 11, 14–16, 18, 20, 23, 27–29, 37, 44, 46, 50, 52, 65, 71, 73, 83
- Bi-LSTM** Bidirectional Long Short-Term Memory. 28
- CF** Factor Corrección Insulina. 9, 10
- CNN** Convolutional Neural Network. 27, 28, 31, 32, 53, 56, 57, 66, 68, 72, 73, 80, 84, XIII, XV
- ConvLSTM** Convolutional Long Short-Term Memory. 28
- CRNN** Convolutional Recurrent Neural Network. 28
- CWT** Continuous Wavelet Transform. 62, 63
- DA** Data Augmentation. 63, 64, 69, 74–76, 79, 84
- DM** Diabetes Mellitus. 1, 2, 14–16, XV
- DT** Decision Trees. 26
- EA** Evolutionary Algorithm. 38, 40
- EGA** Error Grid Analysis. 17, 18

FN False Negative. 21, 22

FNN Feed-forward Neural Network. 29

FP False Positive. 21, 22

FPGA Field Programable Gate Array. 29

FPR False Positive Rate. 22

FRBS Fuzzy Rule Based System. 36–42, 51, XV

GA Genetic Algorithm. 26, 37, 40

GE Grammatical Evolution. 26, 27

GFS Genetic Fuzzy Systems. 40

GP Genetic Programming. 26

GRU Gated Recurrent Unit. 32

ISCI Infusión Subcutánea Continua de Insulina. 8, 9, 69, 70, XIII

KB Knowledge Base. 38–40, 42, 43

KNN K-Nearest Neighbors. 26, 31, 32

KS Kolmogorov-Smirnov. 22, 75, 77, 78

LR Logistic Regression. 31

LSTM Long Short-Term Memory. 27–29, 31

MAE Mean Absolute Error. 16, 17, 26, 28, 29

MARD Mean Absolute Relative Difference. 16, 17, 29

MCG Monitor Continuo de Glucosa. 9, 11–13, 16, 23, 27–32, 45, 69, 71, 76, 79

MDI Múltiple Dosis de Insulina. 8, 9, 13, 69, 70, XIII

METSK-HD Multi-Objective Evolutionary Learning of TSK systems for high-dimensional problems with estimated error. 39–42, 46, 48, 50, 52, 83

MF Membership Function. 35, 39, 40, 43, 44, XV

ML Machine Learning. 26, 27, 29–31, 33, 52, 74, 83, 84

- MLP** Multilayer Perceptron. 27
- MOEA** Multi-Objective Evolutionary Algorithm. 40, 42, 43
- NN** Neural Network. 27–29, 33, 74, 84
- RB** Rules Database. 38–40, 42
- RF** Random Forest. 26, 30, 31
- RMSE** Root Mean Square Error. 16, 17, 26–29, 43, 46, 50, 51
- RNN** Recurrent Neural Network. 27, 28, 31, 32
- ROC** Receiver Operating Characteristic. 22, 75, 77, 78, XVI
- RT** Random Trees. 26
- SGD** Stochastic Gradient Descent. 68, 69
- SGE** Structured Grammatical Evolution. 27
- STFT** Short-time Fourier transform. 61
- SVM** Support Vector Machine. 26, 30, 31
- SVR** Support Vector Regression. 28
- T1DM** Diabetes Mellitus Tipo 1. 1–3, 7, 18, 19, 23, 26–29, 32, 45, 69, XV
- T2DM** Diabetes Mellitus Tipo 2. 1–3, 7, 18, 29, 32
- TCN** Temporal Convolutional Network. 28
- TIR** Time in Range. 11
- TITR** Time in Tight Range. 11
- TN** True Negative. 21, 22
- TP** True Positive. 20–22
- TPR** True Positive Rate. 21, 22
- TSK** Takagi-Sugeno-Kang. 37–44, 48, XV
- VG** Variabilidad Glucémica. 14, 15, XV

Índice general

Índice general	XI
Índice de tablas	XIII
Índice de figuras	XV
1 Introducción	1
2 Predicción de niveles de Glucosa en personas con diabetes	7
2.1. Monitorización Continua de Glucosa	12
2.2. Predicción del valor exacto de la Glucosa	14
2.2.1. Escenarios de predicción	15
2.2.2. Métricas	16
2.3. Predicción de Hipoglucemias e Hiperglucemias	18
2.3.1. Métricas	20
2.4. Conclusiones	23
3 Estado del Arte	25
3.1. Predicción de valores de Glucosa	26
3.2. Predicción de Hipoglucemias e Hiperglucemias	30
3.3. Transformada de Wavelet	32
3.4. Conclusiones	33
4 Predicción de glucosa mediante un sistema basado en reglas difusas	
TSK de dos fases: METSK-HD	35
4.1. FRBS-TSK	37
4.2. Método Propuesto	39
4.3. Fases de optimización del método METSK-HD	41
4.4. Resultados	44

4.4.1.	Configuración Experimental	44
4.4.2.	Conjunto de Datos	45
4.4.3.	Escenarios What-If y Agnóstico	46
4.4.4.	Resultados Experimentales	46
4.4.5.	Análisis de las soluciones obtenidas	48
4.5.	Conclusiones	51
5	Predicción de Hipoglucemia e Hiperglucemias	53
5.1.	Redes Neuronales Convolucionales	53
5.1.1.	Transfer Learning	57
5.1.2.	Arquitectura AlexNet	57
5.1.3.	Arquitectura DenseNet	58
5.2.	Transformada de Wavelet	61
5.3.	Técnicas de Aumentado de Datos	63
5.3.1.	Aumentado de Datos mediante Ventana Deslizante	63
5.3.2.	Aumentado de Datos mediante Ruido Gaussiano	64
5.4.	Método Propuesto	65
5.5.	Resultados	68
5.5.1.	Configuración experimental	68
5.5.2.	Conjunto de Datos	69
5.5.3.	Resultados en predicción de Hiperglucemias	71
5.5.4.	Resultados de predicción de Hipoglucemias	73
5.5.5.	Discusión	73
5.5.6.	Análisis de Incertidumbre	75
5.5.7.	Análisis de Sensibilidad	76
5.6.	Conclusiones	79
6	Conclusiones	83
7	Publicaciones e Indicadores de Calidad	87
	Bibliografía	91

Índice de tablas

2.1. Características de los diferentes tipos de insulinas	8
4.1. Mejor modelo en la fase de entrenamiento	47
4.2. Mejor modelo Ensemble.	47
4.3. Mejor modelo en la fase de validación	48
4.4. Mejor modelo en la fase de test (con fines comparativos).	48
4.5. Clasificación general del BGLP Challenge.	49
5.1. Parámetros de las diferentes capas del modelo AlexNet	59
5.2. Parámetros de las diferentes capas del modelo Densenet-121	60
5.3. Hiperparámetros usados para entrenar los modelos con las distintas arquitecturas CNN.	68
5.4. Caracterización clínica de cada paciente. ID, Sexo, HbA1c, Edad, Años de enfermedad, Peso, Altura, Tratamiento (MDI: Múltiple Dosis de Insulina; ISCI: Infusión Subcutánea Continua de Insulina).	70
5.5. Características de los datasets por paciente. Registros indica el número de registros de glucosa, Hipo indica número de eventos de hipoglucemias, Imágenes Train indica el número de imágenes para entrenamiento, e Imágenes Test el número de imágenes utilizadas para test. Tanto para training como para test, se utilizan un 50 % de imágenes de hipoglucemias.	70
5.6. Comparación de los modelos generales entrenados con las diferentes arquitecturas CNN y wavelets. Umbral de hiperglucemia: $> 240 \text{ mg/dL}$. Accuracy: Exactitud, Sensitivity: Sensibilidad, Specificity: Especificidad, Precision: Precisión.	72
5.7. Comparación de los modelos generales entrenados con las diferentes arquitecturas CNN y wavelets. Umbral de hipoglucemia: $< 70 \text{ mg/dL}$. Accuracy: Exactitud, Sensitivity: Sensibilidad, Specificity: Especificidad, Precision: Precisión.	73

- 5.8. Comparación de modelos personales entrenados con DenseNet-121 y utilizando la wavelet Mexican Hat. Los resultados se muestran como la media de todas las ejecuciones. Umbral de hipoglucemia: $< 70 \text{ mg/dL}$. El texto en rojo indica valores inferiores a 50%. Accuracy: Exactitud, Sensitivity: Sensibilidad, Specificity: Especificidad, Precision: Precisión. 74
- 5.9. Comparación de modelos personales entrenados con DenseNet-121 y utilizando la wavelet Morlet. Los resultados se muestran como la media de todas las ejecuciones. Umbral de hipoglucemia: $< 70 \text{ mg/dL}$. El texto en rojo indica valores inferiores a 50%. Accuracy: Exactitud, Sensitivity: Sensibilidad, Specificity: Especificidad, Precision: Precisión. 74
- 5.10. Comparación de método propuesto con otras técnicas de aprendizaje automático y redes neuronales. Los resultados se muestran como la media de todas las ejecuciones. El texto en rojo indica valores inferiores al 50%. Accuracy: Exactitud, Sensitivity: Sensibilidad, Specificity: Especificidad, Precision: Precisión. 75

Índice de figuras

1.1. Comparativa de los tipos de DM	2
2.1. Curvas de acción de los diferentes tipos de insulina	8
2.2. Bomba de insulina	9
2.3. Pluma de insulina	9
2.4. MCG Abbott FreeStyle Libre de primera generación	13
2.5. MCG Dexcom ONE	14
2.6. Ejemplo VG de una persona con diabetes durante 5 días	15
2.7. Ejemplos de Clarke (a) y Parkes (b) Error Grid (T1DM) de un modelo predictivo de glucosa con horizonte de $T = 60$ minutos	19
2.8. Matriz de Confusión	21
4.1. Ejemplo de lógica difusa: Temperatura	36
4.2. Flujo del modelo FRBS-TSK.	39
4.3. Proceso de optimización del FRBS-TSK.	41
4.4. Procesos implicados en el módulo FRBS.	42
4.5. Desplazamiento lateral MF en $[-0,1,0,1]$	43
4.6. MF triangular clásico.	44
5.1. Esquema básico de una ANN	54
5.2. Ejemplo de operación de convolución de tamaño 3×3 sobre una imagen de entrada de tamaño 6×6	55
5.3. Ejemplo de operaciones Max Pooling y Average Pooling de tamaño 2×2 sobre un mapa de características de 4×4	55
5.4. Esquema básico de ejemplo de una CNN	56
5.5. Arquitectura AlexNet	58
5.6. Arquitectura DenseNet-121	60
5.7. Mexican Hat Wavelet	62
5.8. Morlet Wavelet	62

5.9. Aumentado de datos mediante una ventana deslizante con tamaño una hora	64
5.10. Método propuesto. La Figura (a) muestra el método para la detección de hipoglucemias y la Figura (b) muestra el método para la detección de hiperglucemias.	66
5.11. Clasificación de las imágenes en dos grupos. La Figura (a) muestra la clasificación de los datos diarios de glucosa para la detección de hipoglucemias y la Figura (b) muestra la clasificación de los datos diarios de glucosa para la detección de hiperglucemias.	67
5.12. Porcentaje de tiempo que el paciente tiene un nivel de glucosa muy bajo ($< 54 \text{ mg/dL}$), bajo ($[54, 70) \text{ mg/dL}$), en rango ($[70, 180] \text{ mg/dL}$), alto ($(180, 250] \text{ mg/dL}$) y muy alto ($> 250 \text{ mg/dL}$).	71
5.13. Rangos intercuartílicos de glucosa por paciente. El valor medio se muestra con un punto rojo.	72
5.14. Porcentaje de imágenes generadas con hipoglucemia (barras naranjas) y sin hipoglucemia (barras verdes) y la precisión en test (línea azul) obtenida por los modelos de cada paciente. Umbral de hipoglucemia: $< 70 \text{ mg/dL}$, Wavelet: Mexican Hat.	76
5.15. Curvas ROC de los modelos generales de hiperglucemias	77
5.16. Curvas ROC de los modelos generales de hipoglucemias	78
5.17. Análisis de Incertidumbre. Gráficos de la media (línea continua) y la desviación estándar (sombra azul) de la precisión en test de los modelos personales en diferentes umbrales de hipoglucemia entre 65 y 75 mg/dL	79
5.18. Análisis de Sensibilidad: la línea de puntos rojos marca la precisión del modelo entrenado en ausencia de ruido. El resto de líneas indican las diferentes métricas (Accuracy: Exactitud, Sensitivity: Sensibilidad, Specificity: Especificidad, Precision: Precisión) al añadir distintos porcentajes de ruido al conjunto de datos.	80

Capítulo 1

Introducción

La Diabetes Mellitus (DM) es una enfermedad cada vez más relevante en el mundo, especialmente en los países menos desarrollados y de renta media. Según la Organización Mundial de la Salud, será una de las principales causas de muerte en 2030. Se estima que, actualmente más del 8,5% de la población mayor de 18 años padece algún tipo de diabetes, tratándose así de un problema de salud pública.

La DM se puede definir como una enfermedad autoinmune que se caracteriza por unos niveles elevados de glucosa en sangre (BG, de las siglas en inglés Blood Glucose) si no se toman las medidas adecuadas. Existen dos tipos principales de DM:

- **Diabetes Mellitus de Tipo 1 (T1DM):** el páncreas de una persona que padece T1DM no produce suficiente insulina, por lo que las células son incapaces de regular y procesar la glucosa de la sangre. Esto produce que la concentración de glucosa en sangre aumente hasta niveles potencialmente insalubres.
- **Diabetes Mellitus de Tipo 2 (T2DM):** es el tipo de diabetes más extendido y a diferencia de la T1DM, el páncreas produce insulina pero ésta no es absorbida por las células. Este fenómeno se denomina resistencia a la insulina.

En la Figura 1.1 se muestra una comparativa de los dos tipos de diabetes con respecto a una persona sana. Se puede observar como un páncreas sano produce la insulina necesaria para que las células puedan absorber y transformar la glucosa en la energía necesaria que requiere el organismo. Un páncreas con T1DM no produce la insulina necesaria debido a un fallo provocado por la destrucción de las células beta, siendo estas células las encargadas

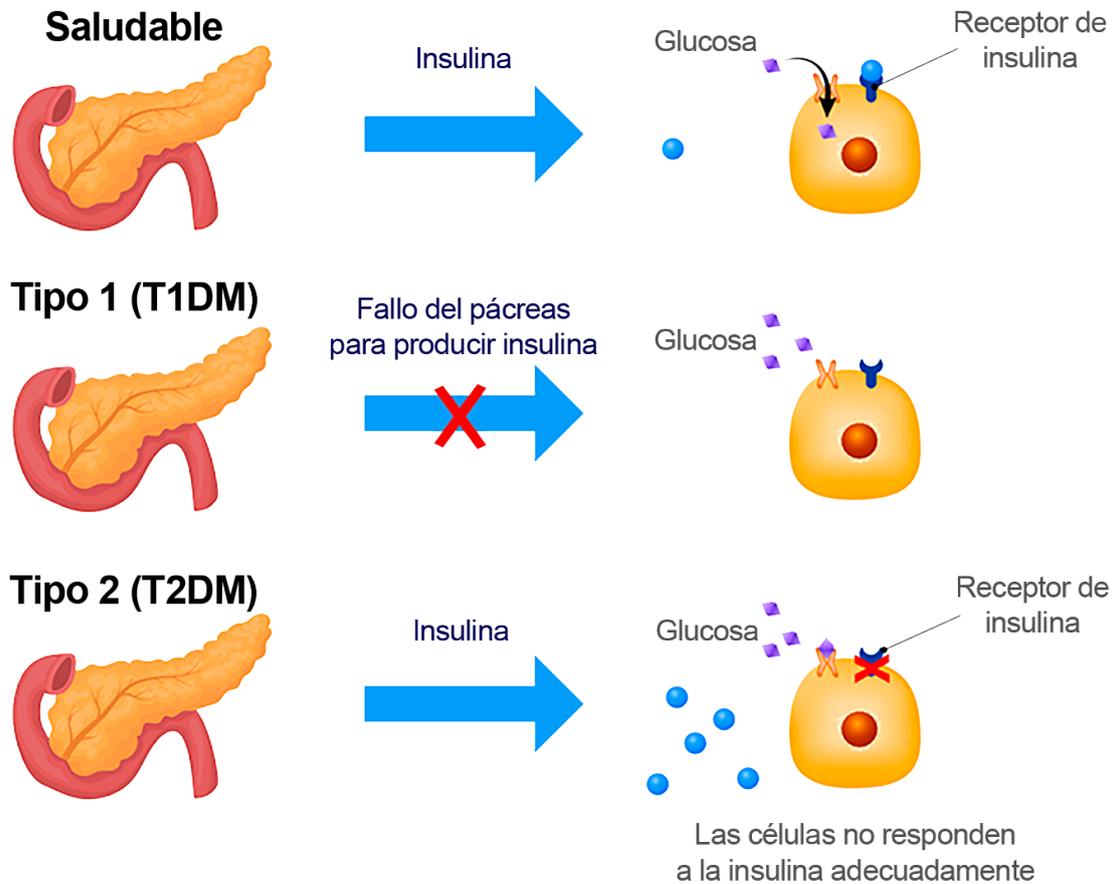


Figura 1.1: Comparativa de los tipos de DM

de generar la insulina. Por otro lado, un páncreas con T2DM si produce insulina, pero las células presentan resistencia a la insulina, es decir, son incapaces de absorberla.

Más del 95% de las personas con diabetes tienen T2DM, siendo el tipo de diabetes más común. Existen otros tipos de diabetes que no entraremos en detalle en este trabajo de tesis, como puede ser la *diabetes gestacional*. Este tipo de diabetes se puede producir durante el embarazo de una mujer, incluso si la persona no es diabética, y generalmente desaparece tras el parto. Por último, existen otros casos de diabetes que no pueden ser clasificados como T1DM o T2DM.

No llevar un control adecuado de la DM puede suponer que los niveles de BG alcancen niveles potencialmente insalubres. El intervalo de valores de glucosa que se considera saludable es $[70 - 180] \text{ mg/dL}$, es decir, tanto los valores de glucosa por debajo de 70 mg/dL como valores por encima de 180 mg/dL son peligrosos para la salud. El principal objetivo de una persona con diabetes es mantener el mayor tiempo posible el valor de BG dentro de este rango.

El control de la diabetes depende del tipo. Para solucionar la falta de insulina, las personas insulino dependientes (T1DM) se aplican inyecciones de insulina farmacéutica para sustituir la que no produce el páncreas. La administración de estas dosis de insulina es una tarea crítica para la persona, porque un cálculo erróneo de la cantidad o usar un tipo de insulina incorrecta, puede hacer que la glucosa alcance niveles peligrosos. En el caso de la T2DM, se suele controlar mediante un cambio en los hábitos de dieta y ejercicio combinado con medicación no inyectable. En estados más avanzados de la enfermedad, y en ciertas ocasiones, es necesario administrar alguna cantidad de insulina, como en la T1DM.

Es primordial para una persona con diabetes tener sus niveles de glucosa dentro del rango adecuado. Para conseguirlo, es necesario realizar una correcta estimación de la BG futura que permita a la persona tomar las acciones correctoras necesarias, con el fin de evitar que sus niveles de glucosa alcancen valores que se encuentren fuera del rango saludable y evitar así situaciones potencialmente peligrosas para la salud. Por ello, la hipótesis de este trabajo de tesis es la generación de modelos de predicción de glucosa, tanto para predicción del valor exacto de la glucosa como para predicción de hipoglucemias e hiperglucemias. Para alcanzar esta meta, se han abordado el problema desde varios enfoques:

- Generar modelos para predecir los valores BG en escenarios *What-If* usando un sistema basado en reglas difusas. Estos modelos deben estar compuestos por un conjunto de reglas de forma que sean interpretables por parte de un humano.
- Generar modelos para predecir los valores BG en escenarios *Agnósticos* usando un sistema basado en reglas difusas. Al igual que en el punto anterior, los modelos deben ser interpretables y estar formados por un conjunto de reglas.
- Generar modelos para predecir episodios de Hiperglucemias (niveles BG altos). Para llegar a este objetivo, es necesario estudiar técnicas de aumento de datos, y se propone el estudio de la combinación de transformaciones del espectro al campo de visión y modelos de aprendizaje profundo.
- Generar modelos para predecir episodios de Hipoglucemias (niveles BG bajos). De forma similar con respecto a las hiperglucemias, es necesario estudiar técnicas de aumento de datos, y la combinación de transformaciones del espectro al campo de visión y modelos de aprendizaje profundo.

El resto de este trabajo de tesis está organizado de la siguiente manera. En el Capítulo 2 se presenta una introducción más profunda del problema de la diabetes. Se explican las

diferentes terapias y tipos de insulinas, así como las variables que influyen en el cálculo de la dosis a administrar. También se detallan los diferentes dispositivos que existen para monitorizar la glucosa. A continuación, se muestran los dos enfoques seguidos en este trabajo de tesis. El primer enfoque, se centra en el problema de la predicción del valor exacto de glucosa. En el segundo enfoque, profundiza en el problema de la predicción de valores críticos de glucosa, como son las hipoglucemias y las hiperglucemias.

En el Capítulo 3 se realiza una revisión del estado del arte del problema de la predicción de la glucosa. Se comienza con una breve introducción a los diferentes tipos de modelos que se utilizan para la predicción de la glucosa. Continuamos con una breve revisión del estado del arte sobre el problema de la predicción del valor exacto de glucosa mediante el uso de técnicas de Machine Learning y aprendizaje profundo. Además, se ha realizado una revisión de la literatura más centrada en el problema de la detección de las hipoglucemias e hiperglucemias. Y, por último, se ha realizado una breve revisión sobre la aplicación de la transformada de wavelet en problemas de series temporales. Esto se ha realizado con el objetivo de ver la viabilidad de aplicar la transformada de wavelet al problema de la predicción de la glucosa, al tratarse también de series temporales de datos.

El Capítulo 4 aborda la metodología del primer enfoque de este trabajo de tesis. En primer lugar, se ha realizado una breve introducción a la lógica difusa, así como a los sistemas basados en reglas difusas. A continuación, detallaremos en profundidad el método propuesto, que consiste en la generación de modelos predictivos de glucosa en escenarios *What-If* y *Agnóstico* en diferentes horizontes temporales (30, 60, 90 y 120 minutos) haciendo uso de un sistema basado en reglas difusas TSK de dos fases. Los resultados muestran que los modelos obtienen buenos resultados en líneas generales, destacando en este primer enfoque la interpretabilidad de los modelos, algunos de ellos formados únicamente por 2 reglas difusas.

Por otro lado, el Capítulo 5 aborda la metodología del segundo enfoque. En primer lugar, se realiza una introducción a los modelos de redes neuronales así como una introducción a la transformada de wavelet y su aplicación a series temporales de datos. A continuación, se detalla el método propuesto, que consiste en la predicción de hiperglucemias e hiperglucemias mediante la aplicación de transformaciones del espectro al campo de visión, en concreto la transformada de wavelet, y modelos de aprendizaje profundo basados en redes neuronales convolucionales. Para entrenar estos modelos, se han empleado diferentes técnicas de aumento de datos, dado que el conjunto de datos original no tenía la dimensionalidad necesaria para obtener modelos robustos. Los resultados muestran que el método propuesto funciona mejor para detectar las hipoglucemias con respecto a las hiperglucemias.

Por último, el Capítulo 6 aborda las conclusiones y trabajo futuro de este trabajo de

CAPÍTULO 1. INTRODUCCIÓN

tesis.

Capítulo 2

Predicción de niveles de Glucosa en personas con diabetes

Como se ha introducido en el capítulo anterior, las personas con diabetes deben llevar un control glucémico con el objetivo de mantener sus niveles de glucosa dentro un rango considerado como saludable.

Existen diferentes formas de tratar la diabetes dependiendo del tipo. Las personas insulino dependientes (T1DM que necesitan inyectarse insulina artificial para sustituir la acción de la insulina natural que generaría un páncreas sano, junto con algunos casos de personas T2DM). La insulina permite transformar la glucosa que viene de los alimentos para que se pueda utilizar las células del cuerpo humano. Si la insulina no funciona correctamente, es necesario suplir la carencia mediante insulinas artificiales. Existen diferentes tipos de insulina artificial. Atendiendo al tiempo de acción podemos clasificarlas de la siguiente forma:

- **Insulinas de acción rápida/corta:** también denominadas bolos de insulinas, sirven para corregir rápidamente el valor de glucosa. Se suelen administrar con las comidas debido a que los hidratos de carbono ingeridos son descompuestos por el cuerpo en glucosa, aumentando así los niveles en sangre. La rápida acción de este tipo de insulina ayuda a regresar el nivel de glucosa a valores saludables, aunque su efecto no es prolongado en el tiempo. Además de su uso con las comidas, se pueden usar también para corregir picos inesperados de niveles altos de glucosa.
- **Insulinas de acción intermedia y prolongada:** también denominada insulina basal, se emplea para controlar la glucosa entre las comidas y durante la noche.

El hígado genera glucosa para que el cuerpo disponga de la energía necesaria para su correcto funcionamiento y este tipo de insulina evita que aumente a niveles no saludables. Como su nombre indica, la acción de estas insulinas no es inmediata y sus efectos son prolongados en el tiempo. Además no tienen picos de acción a diferencia de las insulinas rápidas.

La Tabla 2.1 detalla las características de los diferentes tipos de insulina [1]. Las curvas de acción de cada tipo de insulina se muestran en la Figura 2.1.

Tipo	Inicio	Pico	Duración
Insulina de acción muy rápida: Fiasp	5 minutos	30 minutos	3-5 horas
Insulina de acción rápida: lispro	15-30 minutos	0,5-2 horas	2-5 horas
Insulina de acción rápida: aspart	15 minutos	1-3 horas	3-5 horas
Insulina de acción rápida: glulisina	12-30 minutos	1,5 horas	5-6 horas
Insulina de acción corta: Regular	0,5-1 hora	2-4 horas	5-8 horas
Insulina de acción intermedia: NPH	2-4 horas	4-10 horas	8-16 horas
Insulina de acción prolongada: detemir	1-2 horas	6-12 horas	20-24 horas
Insulina de acción prolongada: glargine	2-4 horas		24 horas
Insulina de acción muy prolongada: glargine U300	6 horas		24-36 horas
Insulina de acción muy prolongada: degludec	0,5-1,5 horas		42 horas

Tabla 2.1: Características de los diferentes tipos de insulinas

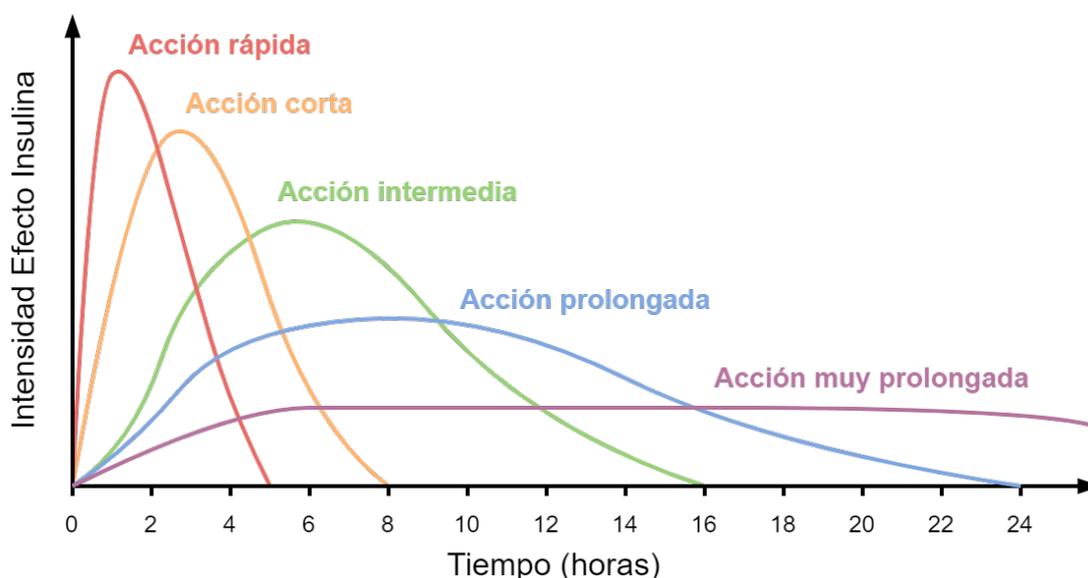


Figura 2.1: Curvas de acción de los diferentes tipos de insulina

A la hora de administrar las insulinas, las dos terapias principales se denominan *In-fusión Subcutánea Continua de Insulina* (ISCI) y *Múltiple Dosis de Insulina* (MDI). El



Figura 2.2: Bomba de insulina



Figura 2.3: Pluma de insulina

tratamiento ISCI consiste en el uso de una bomba de insulina (véase Figura 2.2) conectada a un MCG. El usuario ajusta las cantidades de las dosis para cada instante de tiempo y necesidades. Por otro lado, la terapia MDI consiste, habitualmente, en la inyección de una dosis de insulina de acción prolongada al día y varias dosis de insulinas de acción corta, normalmente una por cada comida, y en ocasiones algún bolo corrector. Para las inyecciones de insulina, habitualmente se utiliza un dispositivo denominado pluma de insulina (véase Figura 2.3).

Inyectar la cantidad adecuada de insulina es todo un reto, al existir diferentes parámetros y variables que pueden afectar a esta decisión. Los principales parámetros que se utilizan para el cálculo de la dosis de insulina son las siguientes:

- **Factor de sensibilidad:** también denominado factor de corrección (CF, de sus

siglas en inglés Correction Factor), se puede definir como la cantidad de nivel de glucosa (mg/dL) reducida por cada unidad de bolos de insulina administrada. El factor de sensibilidad habitualmente se calcula usando la fórmula denominada regla de 1700 [2], aunque este valor puede variar:

$$CF = 1700 / TDD \quad (2.1)$$

dónde TDD es la dosis total de insulina diaria.

- **Ratio basal:** definida como la dosis continua de insulina basal necesaria para contrarrestar la glucosa generada por el hígado.
- **Ratio de insulina por ración de hidratos:** tal como su nombre indica, es la cantidad de bolos de insulina necesaria para metabolizar una ración de hidratos de carbono. Habitualmente suele ser de 1:10, es decir, un bolo de insulina por cada 10 gramos de hidratos (una ración).
- **Bolo corrector:** se define como la cantidad de insulina rápida adicional necesaria en base al factor de sensibilidad para alcanzar el nivel de BG objetivo. La fórmula para su cálculo es la siguiente:

$$Bolo\ corrector = (Glucosa\ actual - Glucosa\ objetivo) / CF \quad (2.2)$$

Hay que tener en cuenta que esta serie de parámetros son habitualmente establecidos en base a las observaciones realizadas por el endocrinólogo, por lo que son valores individuales para cada persona.

Por otro lado, tenemos las variables que se utilizan para hacer este cálculo de manera manual. Estas variables son la glucosa actual que tiene la persona, los valores de las insulinas y las raciones de carbohidratos que se van a ingerir, es decir, la persona debe estimar las raciones de hidratos que va a consumir durante la comida. Una vez obtenidas estas variables y aplicando los parámetros anteriores, se puede calcular la dosis de insulina necesaria.

Un cálculo incorrecto de la dosis puede provocar que los niveles de BG puedan alcanzar niveles fuera del intervalo recomendable. Cuando el nivel de glucosa en sangre desciende por debajo de un umbral, normalmente 70 mg/dL , se dice que la persona sufre hipoglucemia. En el caso opuesto, la hiperglucemia, se produce cuando el nivel de glucosa supera los 180 mg/dL . El objetivo principal de los protocolos de control de la glucosa es permanecer entre esos valores el mayor tiempo posible, es decir, maximizar el tiempo en el intervalo $[70, 180]\text{ mg/dL}$, o simplemente lo que se denomina *tiempo en rango*. Los

CAPÍTULO 2. PREDICCIÓN DE NIVELES DE GLUCOSA EN PERSONAS CON DIABETES

eventos de hipoglucemia e hiperglucemia son potencialmente peligrosos para la salud si sus efectos se mantienen en el tiempo y pueden incluso causar la muerte de la persona. Actualmente se está girando a Time in Tight Range (TITR) [3, 4, 5]. A diferencia de TIR, el intervalo dónde deben permanecer los valores de glucosa el mayor tiempo posible en TITR es más ajustado. En concreto, este intervalo es de $[70, 140]$ mg/dL , siendo un 70 % con respecto a TIR. Al considerarse un umbral de hiperglucemia de $140 mg/dL$, esto puede suponer una mejor esperanza de vida al minimizar las consecuencias negativas de las hiperglucemias para la salud, sobre todo en personas jóvenes.

Uno de los factores más importantes para determinar los valores futuros de los niveles de glucosa y prevenir los episodios de hipoglucemia e hiperglucemia es conocer el valor de la glucosa en el momento de la predicción. Hoy en día, las personas con diabetes utilizan ampliamente los sistemas de monitorización continua de la glucosa (MCG), que consisten en un sensor que registra y monitoriza los valores de glucosa intersticial de manera regular. Estos sistemas los detallamos en la Sección 2.1.

Las inyecciones de insulina se registran mediante el dispositivo utilizado para esta tarea, una bomba de insulina automática que lleva el paciente o de manera manual, o mediante algún otro dispositivo auxiliar. Las ingestas de alimentos son registradas por el paciente manualmente como estimaciones de la cantidad de gramos de hidratos de carbono ingeridos. Además, la inclusión de una bomba de insulina permite automatizar parcialmente la administración de la insulina basal acorde a los parámetros introducidos por el usuario.

Hay otras tareas importantes que deben realizar las personas con diabetes como estimar los hidratos de carbono ingeridos, evaluar el ejercicio realizado, el estrés, etc. De alguna manera, lo que hacen es predecir el valor de glucosa en sangre bajo unas condiciones de ingesta de carbohidratos y suministro de insulina. Dicho de otra forma, tratan de ajustar estos elementos para que, transcurridos unos 90 minutos, la glucosa en sangre permanezca dentro de los niveles saludables. Nuestra propuesta es ayudar a las personas con diabetes en esta tarea, proporcionando modelos predictivos generados mediante distintas técnicas de inteligencia artificial. El objetivo de una buena predicción de la glucosa es corregir las dosis de insulinas y las raciones de hidratos, de forma que la cantidad de estas sea la adecuada para alcanzar un nivel óptimo de BG.

En este trabajo de tesis, hemos abordado el problema de la predicción de la glucosa en dos enfoques diferentes:

- **Predicción exacta del valor de Glucosa**
- **Predicción de Hipoglucemias e Hiperglucemias**

El primer enfoque se centra en estimar el valor exacto de glucosa que va a tener una persona con diabetes dado un horizonte temporal. El segundo enfoque es diferente porque tiene como objetivo anticiparse a los posibles eventos de hipoglucemias e hiperglucemias que pueden ocurrir en un horizonte temporal dado.

2.1. Monitorización Continua de Glucosa

Es de vital importancia conocer el nivel de glucosa actual, por ello, las personas con diabetes habitualmente necesitan usar medidores de glucosa para tomarse varias muestras de sangre al día. Gracias a los avances tecnológicos producidos en el ámbito de la medicina, actualmente existen los sistemas de Monitorización Continua de la Glucosa (MCG), muy utilizados en la actualidad y que miden los niveles de glucosa en la fase intersticial cada cinco o quince minutos.

Gracias a estos sistemas, se puede llevar un mejor control de la enfermedad al poderse conocer el nivel de glucosa en tiempo real, siendo este dato importante para llevar a cabo las acciones requeridas en caso de que sea necesario corregirlo. Al disponer de mecanismos de alarma, pueden avisar en todo momento si el nivel de glucosa se encuentra fuera del rango saludable.

Existen en el mercado multitud de soluciones de sistemas MCG, siendo Abbott, Medtronic y Dexcom los fabricantes más extendidos en España en el momento de escribir este trabajo de tesis.

Abbott cuenta con la gama FreeStyle Libre que utiliza sensores desechables que permiten registrar la glucosa intersticial en intervalos de 15 minutos y una duración del sensor de hasta 14 días. Existen tres generaciones del sensor. La diferencia más destacable radica que el sistema de primera generación requiere de un dispositivo lector para registrar los datos del sensor. En cambio, los sistemas FreeStyle Libre 2 y 3 no es necesario contar con el dispositivo lector, es decir, el sensor se conecta mediante Bluetooth a cualquier Smartphone y la aplicación móvil FreeStyle LibreLink¹, que permite registrar y visualizar los datos. Además, estos sistemas permiten la conexión con bombas de insulinas compatibles. En la Figura 2.4 se muestra un MCG FreeStyle Libre de primera generación.

Por otro lado, Medtronic cuenta con la gama de sistemas MCG Medtronic Guardian. Este sistema usa sensores desechables con duración de hasta 6 días que permite tomar mediciones de la glucosa intersticial en intervalos de 5 minutos. Se utiliza la aplicación móvil Guardian Connect² para registrar y visualizar los datos del sensor mediante Bluetooth.

¹<https://www.freestyle.abbott/es-es/productos/conectividad/freestylelibrelinkapp.html>

²<https://www.medtronicdiabetes.com/products/sistema-guardian-connect>



Figura 2.4: MCG Abbott FreeStyle Libre de primera generación

Además, este sistema permite la conexión con las bombas de insulina de la gama Medtronic MiniMed. Dentro de esta gama, nos encontramos principalmente con los sistemas MiniMed 780G, 770G y 630G. Estas bombas permiten una administración automática de la insulina en base a los datos recogidos por el sensor y registrados por el usuario. Además, Medtronic cuenta con el sistema InPen, que consiste en una pluma de insulina usada habitualmente en las terapias MDI conectada a la aplicación móvil del MCG Guardian. Gracias a los datos registrados por el sensor y la pluma, la aplicación permite calcular la dosis de insulina, así como ver el histórico de inyecciones, ingestas de hidratos de carbono, etc.

Por último, nos encontramos con los sistemas MCG Dexcom. Este fabricante utiliza sensores desechables en sus sistemas MCG que permiten tomar mediciones de la glucosa intersticial cada 5 minutos y una duración del sensor de hasta 10 días. Existen varias generaciones del sensor, siendo el sistema Dexcom G7 el más reciente a fecha de escritura de este trabajo de tesis. Desde la quinta generación (Dexcom G5), los datos se transmiten mediante Bluetooth hacia cualquier dispositivo móvil compatible sin necesidad de usar un dispositivo lector dedicado, aunque Dexcom cuenta en su catálogo con dispositivos lectores opcionales compatibles con las últimas generaciones del sensor. Además, se pueden conectar bombas de insulinas compatibles para complementar el MCG. Por otro lado, Dexcom cuenta con una gama de sensores más accesibles denominada Dexcom ONE. Estos sistemas están formados únicamente por el sensor desechable conectado a la aplicación móvil correspondiente, teniendo funcionalidades más limitadas con el resto de las gamas y sin posibilidad de usar bombas de insulina. En la Figura 2.5 se muestra un sensor Dexcom ONE.



Figura 2.5: MCG Dexcom ONE

2.2. Predicción del valor exacto de la Glucosa

Llevar un control óptimo de la DM supone para las personas con diabetes estimar sus niveles de BG futuros. Esta tarea habitualmente es realizada manualmente en base a una serie de variables, lo que supone un reto diario al que se enfrentan estas personas. En concreto, consiste en calcular el nivel de glucosa futuro en base a las ingestas de hidratos de carbonos y las inyecciones de insulinas administradas.

Es importante mencionar en este punto, además de las variables antes mencionadas, el valor de glucosa futuro va a depender de una serie de factores como puede ser la edad, el género, la actividad física, medicación entre otras. Además, existe otro factor a tener muy en cuenta a la hora de predecir la glucosa denominado Variabilidad Glucémica (VG). La VG se puede definir como la fluctuación de los niveles de glucosa que se producen a lo largo del día. En la Figura 2.6 se observa un ejemplo de la VG de una persona con diabetes durante 5 días consecutivos.

Realizar una estimación correcta del valor de BG permite a las personas con diabetes llevar a cabo las acciones necesarias para corregir el nivel de glucosa. Por ejemplo, durante las comidas, deben calcular la dosis de bolos de insulina en base a las raciones ingeridas de carbohidratos para contrarrestar la glucosa producida por la descomposición de los hidratos de carbonos y así volver a un nivel de BG dentro del rango saludable. En otro caso, una persona con diabetes estima que va a tener un nivel de BG muy bajo, entonces toma la decisión de ingerir una cantidad de azúcar para restablecer su glucosa a niveles óptimos.

La hipótesis de este trabajo de tesis es la generación de modelos predictivos que per-

CAPÍTULO 2. PREDICCIÓN DE NIVELES DE GLUCOSA EN PERSONAS CON DIABETES

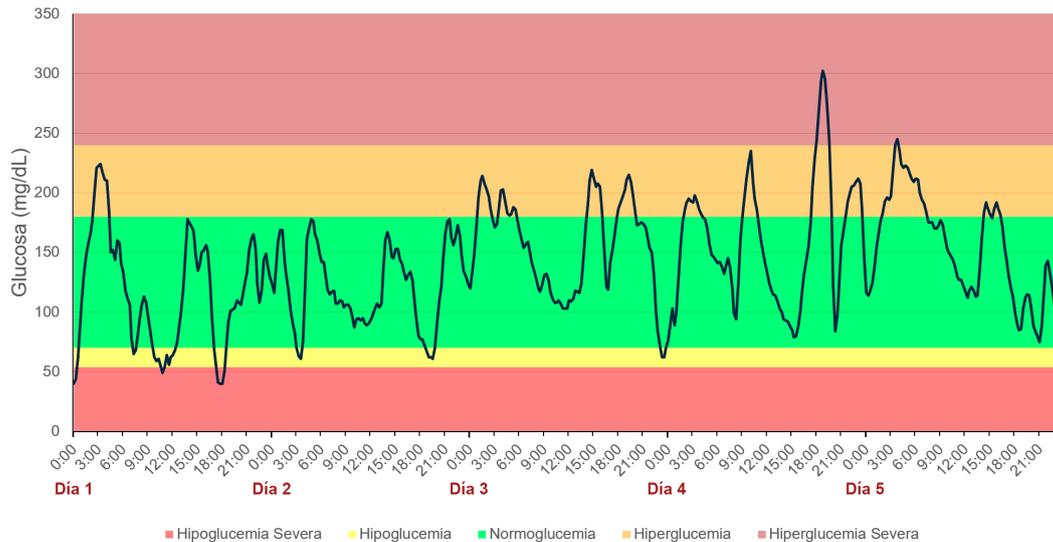


Figura 2.6: Ejemplo VG de una persona con diabetes durante 5 días

mitan predecir la glucosa en un horizonte temporal determinado. Para ello se define el problema de la siguiente forma:

- *Predecir el valor del nivel de glucosa en sangre (BG en un horizonte temporal de M minutos, conociendo la información de los niveles actuales de glucosa en sangre, las ingestas de alimentos (CH), las inyecciones de insulina (I), y la actividad física en el momento de la predicción incluyendo los eventos de las dos últimas horas.*

En resumen, estos modelos permiten dar una estimación de la glucosa con horizontes temporales a 30, 60, 90 y 120 minutos conociendo únicamente los datos de la glucosa, hidratos, insulinas y actividad física de las dos horas anteriores al momento de la predicción.

2.2.1. Escenarios de predicción

A la hora de entrenar modelos predictivos de glucosa, existen varias formas de generar estos modelos dependiendo de la información disponible en el momento de su generación. en este punto se deben contemplar los escenarios *Agnóstico*, *What-If* e *Inercial*.

El primer escenario, *Agnóstico*, únicamente se toman los valores de las variables inmediatamente anteriores al momento de la petición de construir el modelo. Por ejemplo, una persona con DM dado un instante de tiempo desea predecir su BG a un horizonte de predicción de $T + 60$ minutos y únicamente conoce los valores de glucosa, insulinas administradas y raciones de hidratos de carbono ingeridas de las 2 horas anteriores al momento de la predicción.

El segundo escenario, *What-If*, además de considerar la información anterior al momento de solicitud de generar el, tiene en cuenta los valores futuros de algunas de las variables durante la ventana de predicción, es decir, desde el momento actual hasta el horizonte de la predicción. En concreto, es habitual incluir las raciones de hidratos de carbono ingeridas y las dosis insulinas administradas desde el momento actual hasta el horizonte de predicción deseado. Este escenario se suele utilizar a la hora de generar modelos para sistemas recomendadores de insulinas y carbohidratos. Un ejemplo de modelo *What-If* sería el de una persona con DM que desea predecir el nivel de BG que va a tener a un horizonte de predicción de $T + 120$ minutos, conoce los valores de glucosa, carbohidratos e insulina de las últimas 2 horas y además va a comer, es decir, ha calculado las raciones futuras de hidratos de carbono y las dosis de bolos de insulina desde el momento actual hasta el horizonte de la predicción.

Por último, el escenario *Inercial*, los modelos se entrenan usando, además de la información anterior al momento de la predicción, únicamente con los valores que provienen del MCG durante la ventana de predicción, descartándose cualquier evento vital futuro por parte de la persona como la ingesta de carbohidratos, administración de bolos de insulina, ejercicio, etc. Al no poderse usar eventos de otras variables durante la ventana de predicción, conlleva la necesidad de seleccionar cuidadosamente los datos a utilizar durante el entrenamiento del modelo. En resumen, los modelos inerciales predicen la BG teniendo en cuenta que la persona con DM no va a realizar ninguna acción durante el momento de solicitud hasta el horizonte de la predicción. Este escenario es útil para ciertas situaciones como la predicción de la glucosa nocturna, al no producirse ningún otro evento que pueda afectar a la predicción.

Este trabajo de tesis se centra únicamente en los escenarios *Agnóstico* y *What-If* al ser los más habituales en la gestión del día a día de la DM. El escenario inercial corresponde a situaciones normalmente nocturnas, mucho más peligrosas por el sueño, las cuales serán objeto de un trabajo posterior.

2.2.2. Métricas

Para evaluar el rendimiento de los modelos predictivos, habitualmente se utilizan una serie de métricas que indican el desempeño y la robustez del modelo. Las más habituales son la raíz del error cuadrático medio (RMSE, de sus siglas en inglés Root Mean Square Error), error absoluto medio (MAE, de sus siglas en inglés Mean Absolute Error) y diferencia relativa absoluta media (MARD, de sus siglas en inglés Mean Absolute Relative Difference).

- **RMSE** se puede definir como la raíz cuadrada de la media de los errores cuadráticos

CAPÍTULO 2. PREDICCIÓN DE NIVELES DE GLUCOSA EN PERSONAS CON DIABETES

entre los valores estimados y los valores actuales, es decir, mide la diferencia de error entre ambos valores. Esta métrica penaliza los errores más grandes, de forma que cuando menores sean éstos, mejor será el ajuste del modelo. La fórmula para calcular este error es la siguiente (2.3):

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (2.3)$$

dónde n es el número total de muestras, y_i el valor real e \hat{y}_i el valor estimado por el modelo.

- **MAE** se define como el error promedio absoluto entre los valores estimados y los valores reales. A diferencia de RMSE, los errores penalizan de forma proporcional. Se calcula usando la siguiente fórmula (2.4):

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (2.4)$$

dónde n es el número total de muestras, y_i el valor real e \hat{y}_i el valor estimado por el modelo.

- **MARD** es la diferencia relativa absoluta entre el valor real y el valor predicho por el modelo. Esta métrica es útil en situaciones donde se desea conocer la precisión del modelo en términos proporcionales, es decir, cómo difieren los valores predichos en relación con los valores reales, sin considerar la dirección de la diferencia. A diferencia de las otras métricas, MARD se expresa como un porcentaje en lugar del valor del error. La fórmula es la siguiente (2.5):

$$MARD = \frac{1}{n} \sum_{i=1}^n \frac{|y_i - \hat{y}_i|}{|y_i|} \cdot 100 \quad (2.5)$$

dónde n es el número total de muestras, y_i el valor real e \hat{y}_i el valor estimado por el modelo.

Por otro lado, se puede utilizar la familia de Error Grid Analysis EGA. Esta técnica de análisis del desempeño de modelos predictivos, en el caso que nos concierne, modelos predictivos de glucosa. A diferencia de las métricas anteriores, clasifican las predicciones realizadas por los modelos en zonas o cuadrantes dependiendo del riesgo clínico que supone el error en la predicción, es decir, la discrepancia entre los valores de la glucosa estimado y el de referencia. Existen 5 zonas diferentes:

- **Zona A:** las predicciones que caen dentro de esta zona se consideran seguras clínicamente al no diferir sustancialmente del valor real.
- **Zona B:** aunque las predicciones que están dentro de esta zona tienen un error mayor con respecto a la Zona A, no tienen consecuencias clínicas destacables.
- **Zona C:** las predicciones de esta zona presentan discrepancias con respecto los valores de referencia y puede suponer que la persona no tome las decisiones correctas pero teniendo únicamente consecuencias clínicas mínimas.
- **Zona D:** esta zona se considera peligrosa porque las predicciones tienen un error considerable con respecto al valor real. Estas discrepancias conllevan que la persona tome decisiones inapropiadas que pueden tener consecuencias clínicas graves.
- **Zona E:** es la zona más peligrosa. Las predicciones que caen en esta zona son completamente irreales al presentar un error bastante considerable y esto puede provocar que la persona tome acciones que pueden poner en peligro su vida.

Los EGA más usados son el Clarke Error Grid [6] y el Parkes Error Grid [7]. La principal diferencia entre ellos es la distribución de las zonas. Además, Clarke Error Grid se utiliza sobre todo para estudios de la práctica clínica al enfocarse en la seguridad de las predicciones, mientras que Parkes Error Grid se utiliza habitualmente en investigaciones y estudios de validación, al centrarse en la precisión en relación con los valores reales de glucosa. Además, existen dos variantes diferentes de Parkes Error Grid, una específica para la T1DM y otra para la T2DM. La única diferencia entre ambas versiones es la ubicación de las zonas. En la Figura 2.7 se muestra un ejemplo de cada tipo de EGA detallados en esta sección.

2.3. Predicción de Hipoglucemias e Hiperglucemias

Uno de los principales objetivos que tienen las personas con diabetes es mantener los niveles de BG dentro del intervalo óptimo para la salud, pero existen situaciones que el valor de la glucosa se encuentra fuera de ese rango. Cuando el nivel de glucosa es bajo, se denomina *Hipoglucemia* y en el caso de un valor elevado, se denomina *Hiperglucemia*. Dependiendo del nivel de glucosa, podemos definir la siguiente escala:

- **Hipoglucemia severa:** $glucosa < 54 \text{ mg/dL}$
- **Hipoglucemia:** $glucosa < 70 \text{ mg/dL}$
- **Normoglucemia:** $glucosa \geq 70 \text{ mg/dL}$ y $glucosa \leq 180 \text{ mg/dL}$

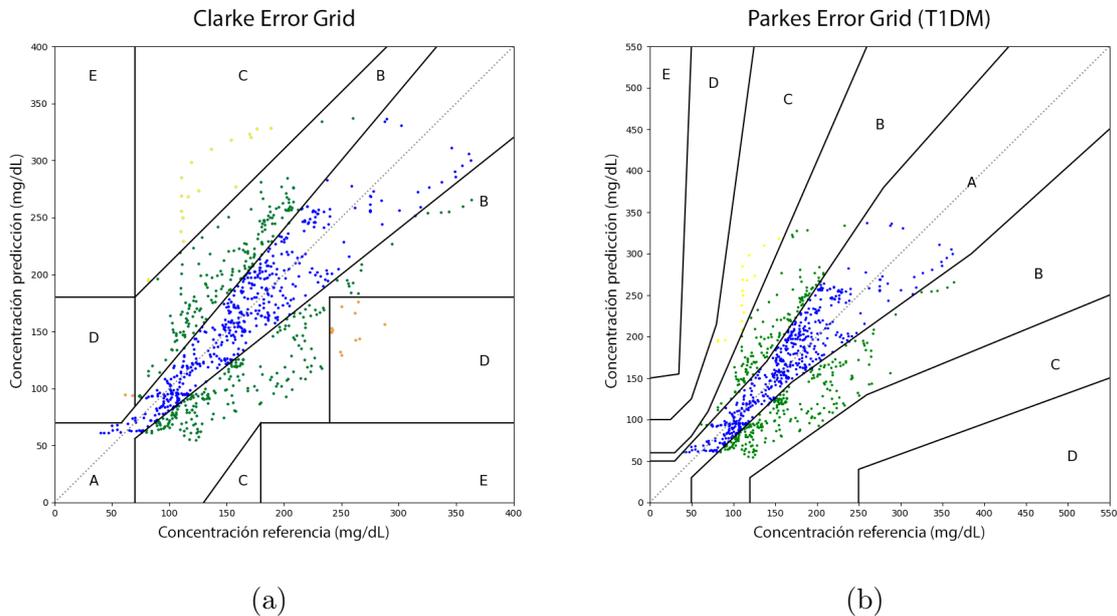


Figura 2.7: Ejemplos de Clarke (a) y Parkes (b) Error Grid (T1DM) de un modelo predictivo de glucosa con horizonte de $T = 60$ minutos

- **Hiperglucemia:** $glucosa > 180$ mg/dL
- **Hiperglucemia severa:** $glucosa > 240$ mg/dL

Es muy importante evitar estos eventos debido a que son potencialmente peligrosos para la salud si no se toman las acciones correctivas necesarias de manera inmediata y en el caso de sus efectos se mantengan en el tiempo pueden provocar complicaciones de salud severas y daños permanentes.

La hipoglucemia es una de las principales complicaciones de la diabetes [8]. Estos episodios de niveles bajos de glucosa son peligrosos para la salud, ya que pueden causar daños en múltiples órganos, mayor riesgo de enfermedad renal, problemas en el sistema nervioso, etc. Si el nivel de glucosa es inferior a 54 mg/dL, se produce un episodio de hipoglucemia severa que, si no se corrige, puede ser potencialmente peligroso, ya que puede producirse una pérdida de consciencia e incluso la muerte de la persona. Los episodios de hipoglucemia suelen clasificarse en función del momento del día en que se producen y suelen establecerse en tres clases: hipoglucemia diaria, posprandial y nocturna. Los episodios de hipoglucemia nocturna son probablemente la situación más peligrosa, ya que suelen pasar desapercibidos debido a que el paciente está dormido. Al tratarse de episodios asintomáticos, sus efectos pueden mantenerse en el tiempo, afectando gravemente a la salud y, eventualmente, provocando el síndrome de *muerto en la cama*.

En el caso de las hiperglucemias, los niveles altos de BG pueden favorecer la acumulación de ácidos tóxicos denominados cetonas. Los principales síntomas de las hiperglucemias son sed, visión borrosa, dolor abdominal, etc. Además, si los efectos se mantienen en el tiempo, las principales complicaciones de las hiperglucemias son riesgo de enfermedad cardiovascular, daños renales, daños en el sistema nervioso, etc. Si el nivel de glucosa es muy elevado (hiperglucemia grave) se pueden alcanzar niveles muy peligrosos de cetonas (Cetoacidosis), siendo las principales complicaciones el coma diabético. El coma diabético puede conllevar la pérdida del conocimiento, daños cerebrales permanentes e incluso la muerte de la persona.

El principal objetivo de este enfoque es la creación de modelos que permitan predecir el rango donde se situará el valor de la BG, es decir, consiste en estimar los posibles eventos de Hipoglucemias e Hiperglucemias a un horizonte temporal definido. En este punto, podemos definir el problema:

- *Predecir los eventos de Hipoglucemia ($< 70 \text{ mg/dL}$) e Hiperglucemias ($> 180 \text{ mg/dL}$) en un horizonte temporal de M minutos, conociendo la información de los niveles actuales y de las últimas 24 horas de la glucosa en sangre.*

Una vez construidos estos modelos predictivos, pueden ser la base para un sistema de alarmas de hipoglucemias e hiperglucemias. Gracias a las notificaciones generadas por el sistema, las personas con diabetes pueden tomar las acciones adecuadas para así evitar que sus niveles de glucosa salgan del umbral saludable.

2.3.1. Métricas

En este enfoque, los modelos propuestos generarán salidas que representan distintas clases en lugar de valores numéricos reales. Es decir, estos modelos indicarán si va a ocurrir una hipoglucemia, hiperglucemia o normoglucemia. Por ello, nos enfrentamos a un problema de clasificación. Para evaluar el rendimiento de los modelos de clasificación, tenemos que introducir en primer lugar, el concepto de matriz de confusión. Una matriz de confusión no es una métrica como tal, sino que se trata de una herramienta que se utiliza para evaluar el rendimiento de los modelos de clasificación. Esta matriz organiza la cantidad de instancias de un conjunto de datos clasificadas correctamente o incorrectamente por un modelo. La matriz de confusión se organiza en cuatro secciones principales, tal como se puede observar en la Figura 2.8:

- **Verdaderos positivos (TP):** cantidad de instancias clasificadas correctamente como positivas por el modelos .

CAPÍTULO 2. PREDICCIÓN DE NIVELES DE GLUCOSA EN PERSONAS CON DIABETES

- **Verdaderos negativos (TN):** cantidad de instancias clasificadas correctamente como negativas por el modelo.
- **Falsos positivos (FP):** cantidad de instancias clasificadas incorrectamente como positivas por el modelo. También conocidos como errores de Tipo I.
- **Falsos negativos (FN):** cantidad de instancias clasificadas incorrectamente como negativas por el modelo. También conocidos como errores de Tipo II.

		Observaciones	
		Positivo	Negativo
Predicciones	Positivo	TP	FP
	Negativo	FN	TN

Figura 2.8: Matriz de Confusión

A partir de esta matriz, se sacan las métricas habitualmente utilizadas para evaluar el desempeño de los modelos de clasificación, siendo las más importantes:

- **Exactitud:** también conocida como *Accuracy*, es la métrica más utilizada para evaluar los modelos de clasificación. Se define como la proporción de las predicciones correctas (verdaderos positivos y verdaderos negativos) con respecto al total de las predicciones realizadas por el modelo. Se calcula de la siguiente manera (2.6):

$$Exactitud = \frac{TP + TN}{TP + TN + FP + FN} \quad (2.6)$$

- **Sensibilidad:** también conocida como *Recall* o tasa de verdaderos positivos (TPR, de las siglas en inglés True Positive Rate), es una métrica que mide la proporción de instancias positivas que fueron correctamente clasificadas por el modelo. Se calcula como (2.7):

$$\text{Sensibilidad} = \frac{TP}{TP + FN} \quad (2.7)$$

- **Especificidad:** también conocida como tasa de verdaderos negativos (TNR), es una métrica complementaria a la Sensibilidad. Mientras que la sensibilidad se enfoca en la capacidad del modelo en detectar correctamente las instancias positivas, la especificidad mide la capacidad del modelo para evitar clasificar incorrectamente las instancias negativas. Se calcula como (2.8):

$$\text{Especificidad} = \frac{TN}{TN + FP} \quad (2.8)$$

- **Precisión:** se define como la capacidad de un modelo para predecir correctamente las instancias positivas entre todas las instancias que predijo como positivas, es decir, indica la proporción de las predicciones positivas que fueron correctas. Se calcula como (2.9):

$$\text{Precisión} = \frac{TP}{TP + FP} \quad (2.9)$$

Además existen otras métricas también ampliamente utilizadas, como son las curvas ROC (Receiver Operating Characteristic) [9], AUC [10], y la prueba estadística KS (Kolmogorov-Smirnov) [11, 12].

Las curvas ROC describen el equilibrio entre la tasa de verdaderos positivos (TPR) y la tasa de falsos positivos (FPR, del inglés False Positive Rate) de un modelo para varios umbrales de clasificación. La TPR es la proporción de verdaderos positivos clasificados correctamente como positivos, mientras que la FPR es la proporción de verdaderos negativos clasificados incorrectamente como positivos. Una curva ROC representa la TPR frente a la FPR con distintos umbrales, y la curva ilustra lo bien que el modelo puede distinguir entre las clases positivas y negativas. La línea discontinua representa un clasificador aleatorio, de modo que cuanto más se aleje la línea ROC de la línea discontinua, mejor será el clasificador. Otra forma de decirlo es que cuanto más cerca esté la curva ROC de la esquina superior izquierda, mejor.

AUC es una métrica comúnmente utilizada para evaluar el rendimiento global de un modelo de clasificación binaria. El AUC oscila entre 0 y 1, donde 0,5 indica adivinación aleatoria y 1 indica un clasificador perfecto. La prueba estadística KS es otra medida de uso común del rendimiento del modelo en la clasificación binaria. Es la diferencia máxima entre las funciones de distribución acumulativa (FDA) de las clases positiva y negativa. El estadístico KS oscila entre 0 y 1, donde 0 indica que no hay separación entre las clases y 1 indica una separación perfecta.

2.4. Conclusiones

Estimar la glucosa es un reto diario que deben realizar las personas con diabetes ya que involucra una serie de decisiones, variables y factores. Además, las personas con T1DM tienen que calcular las dosis de insulina en base a una serie de factores, algunos de ellos individuales dependiendo de cada persona.

Gracias a los avances tecnológicos producidos en los últimos años, han surgido los sistemas MCG. Con estos dispositivos, las personas con diabetes pueden llevar un mejor control de la enfermedad, pero los MCG no evitan la tarea de predecir los niveles de glucosa. Por ello, es necesario hacer uso de modelos predictivos que permitan una estimación automática de la BG con el objetivo de permitir a las personas con diabetes llevar a cabo las acciones necesarias para evitar, en la medida de lo posible, que la glucosa alcance niveles fuera del intervalo saludable.

Además, es importante contar con modelos que permitan una detección temprana de hipoglucemias e hiperglucemias para que las personas con diabetes puedan actuar de forma anticipada y evitar así estos eventos glucémicos. Es importante recalcar que las hipoglucemias e hiperglucemias son peligrosas para la salud si sus efectos se mantienen en el tiempo.

En resumen, éste trabajo de tesis está dividido en dos enfoques diferentes a la hora de abordar este problema. El primer enfoque, la predicción exacta del valor de glucosa lo explicamos en el capítulo 4 y el segundo enfoque, la predicción de las hipoglucemias e hiperglucemias, lo detallamos en el capítulo 5.

Capítulo 3

Estado del Arte

Existen multitud de propuestas en la literatura que abordan el problema de la predicción de la glucosa desde diferentes puntos de vista. Este problema se puede abordar con diferentes metodologías a la hora de construir los modelos, pero principalmente se usan tres técnicas diferentes. La primera, denominada modelos fisiológicos, hace uso de modelos matemáticos lineales para simular el comportamiento de la interacción glucosa-insulina. Estos modelos son bastantes complejos y para su modelado se requiere un alto conocimiento de las dinámicas de la glucosa con la insulina. La segunda técnica, denominada modelos Data-Driven, utilizan la información de entrada y salida, así como las relaciones existente entre los datos para predecir la glucosa. Una de las principales ventajas de estos modelos es que permiten la inclusión de nuevas variables con facilidad, pero tienen como principal desventaja que únicamente se basan en la información y no en el sentido fisiológico de los datos. Por último, tenemos un enfoque híbrido que hace uso de una parte de los datos fisiológicos (glucosa-insulina) a la que se añade información adicional para desarrollar los modelos.

Una vez seleccionada la metodología a abordar, es necesario contar con los suficientes datos para entrenar los modelos. En este punto se suma la dificultad de obtener datos de pacientes reales, en gran medida, al considerarse información de carácter sensible por la naturaleza médica de la misma. En numerosas ocasiones es necesario contar con los respectivos permisos tanto por parte de los pacientes como del comité ético del hospital para llevar a cabo el estudio médico, con la finalidad de obtener los datos. Además, una vez registrados los datos, habitualmente hay que descartar información por errores de lecturas del sensor, muchas veces por desconocimiento de uso por parte del paciente. Por ello, existen una serie de repositorios de datos tanto de acceso público como restringido.

Ejemplos de repositorios públicos son OhioT1DM [13], Pima Indians Diabetes Database¹, UCI Diabetes Data Set [14], etc. Además existe el repositorio Awesome-CGM [15] contiene una recopilación de varios conjuntos de datos de acceso público.

Por otro lado, existen técnicas para generar datos in-silico, es decir, datos de pacientes generados mediante simulación. Existen numerosos simuladores para generar datos, siendo los más usados UVA/PADOVA [16] desarrollado por la Universidad de Virginia y aprobado por la Food and Drug Administration (FDA) de EEUU, y AIDA Simulator [17].

3.1. Predicción de valores de Glucosa

En primer lugar, nos encontramos con los trabajos que aplican técnicas tradicionales de modelización, es decir, modelos o perfiles definidos por ecuaciones lineales con un conjunto limitado de entradas. En [18] nos encontramos con una completa revisión de la literatura sobre modelos lineales glucosa-insulina para predecir la glucosa. Además, los principales grupos de investigación sobre páncreas artificial han presentado otros enfoques de control personalizado [19, 20, 21, 22] que siguen la práctica clínica. Algunas propuestas proporcionan modelos para el caso medio [23] y otras para las particularidades de cada paciente.

Existen diferentes formas de modelar la dinámica de la glucosa, incluidos los modelos basados en series temporales de datos. Algunas de las técnicas utilizadas en estos modelos hacen uso de algoritmos de aprendizaje automático (ML) como Random Forest (RF) [24], K-Nearest Neighbors (KNN) [25], Support Vector Machine (SVM) o Autoregressive Integrated Moving Average (ARIMA) [26]. Por otro lado, también se han aplicado métodos computacionales evolutivos como la Programación Genética (GP, del inglés Genetic Programming), las Gramáticas Evolutivas (GE, del inglés Grammatical Evolution) [27], y los Algoritmos Genéticos (GA, del inglés Genetic Algorithm).

Dentro de las aproximaciones que usan técnicas clásicas de ML, en [28] los autores utilizan diferentes algoritmos de ML para crear un modelo predictivo con horizonte a 120 minutos. En concreto comparan un Decision Trees (DT), Random Forest (RF) y Random Trees (RT). Han usado el conjunto de datos Aleppo [29] del repositorio público Awesome-CGM, compuesto por 225 pacientes T1DM y 6 meses de datos. El rendimiento de los modelos es evaluado mediante RMSE y MAE. Según los resultados, los modelos RT y DT ofrecen un rendimiento similar con un RMSE de 13,54 y 14,05 respectivamente, superando al modelo RF.

Otros trabajos analizan el problema desde diferentes variantes de la GP [30, 31, 32,

¹<https://www.kaggle.com/datasets/uciml/pima-indians-diabetes-database>

33, 34, 35]. En [36], los autores han construido un modelo ensemble basado en GE para predecir la glucosa con 30, 60, 90 y 120 minutos de antelación. El conjunto de datos utilizado para entrenar los modelos está formado por los datos de 5 pacientes con T1DM y 12 días de datos recogidos por un MCG. El rendimiento del modelo propuesto es evaluado con la métrica Clarke Error Grid y lo comparan con otras técnicas clásicas de ML como ARIMA y otras variantes de GE. El método propuesto ofrece un rendimiento superior con respecto al resto de técnicas de ML comparadas. Además, con una tasa de predicciones peligrosas (zonas D+E) bastante aceptable, incluso en el horizonte a 120 minutos (5,5%).

Dentro de las GE, nos encontramos con las denominadas Gramáticas Evolutivas Estructuradas (SGE). En [37], los autores proponen la combinación de técnicas de clústering de datos con un modelo basado en (SGE) para predecir la BG postprandial con un horizonte temporal de hasta 2 horas después de las comidas. Al igual que muchos otros trabajos en la literatura, esta propuesta hace uso del conjunto de datos público OhioT1DM para el entrenamiento de los modelos. El método propuesto combina una técnica de preprocesado de datos consistente en particionar las series temporales de datos en segmentos de 4 horas y la aplicación de una técnica de clústering. Una vez preprocesados los datos, los autores comparan diferentes modelos, en cuestión, un modelo basado en SGE, un modelo de redes neuronales (NN) [38] y el algoritmo Sparse Identification of Nonlinear Dynamics (SINDy). Los modelos han sido evaluados usando el Parkes Error Grid y demuestran que el modelo propuesto en este trabajo ofrece predicciones clínicamente seguras, sin apenas estimaciones que caigan en las zonas peligrosas (D+E). Una de las principales ventajas de esta propuesta, es que el método propuesto está enfocado en obtener modelos interpretables. Este enfoque en la interpretabilidad ha dado como resultado en la precisión obtenida por los modelos, ya que no representa diferencias relevantes con respecto a las técnicas basadas NN y SINDy.

Aunque gracias a los avances producidos en los últimos años en el ámbito del aprendizaje profundo o Deep Learning, cada vez existen más trabajos que hacen uso de técnicas basadas en redes neuronales artificiales (ANN) e incluso hoy en día se dedica un workshop a este problema [39]. Los principales tipos de redes neuronales son el Multilayer Perceptron (MLP), Recurrent Neural Network (RNN), Long Short-Term Memory (LSTM) y Redes Neuronales Convolucionales (CNN, del inglés Convolutional Neural Network).

En [40] se enfocan en el desarrollo de un modelo personalizado por paciente basado en RNN para estimar la glucosa con diferentes horizontes temporales de 15 hasta 60 minutos. Utilizan los datos en bruto del MCG de 38 pacientes con T1DM, previo preprocesado para eliminar los datos de glucosa incoherentes debido a posibles errores en las mediciones. Comparan su aproximación con técnicas de ML clásicas y evalúan el rendimiento de los modelos con las métricas del Error Porcentual Absoluto (APE) y RMSE. Se puede

observar en los resultados que el modelo propuesto es bastante consistente en todos los horizontes temporales analizados.

Zhu et al. [41] propone un modelo basado en RNN para predecir la BG con horizontes de 30 y 60 minutos. Para entrenar los modelos, han usado los conjuntos de datos públicos OhioT1DM, ABC4D y ARISES. En total, han utilizado los datos de 49 pacientes reales con T1DM. Una vez entrenado el modelo, la inferencia del modelo se realiza en un dispositivo wearable basado en un microcontrolador de bajo consumo que obtiene los datos directamente del sensor MCG. Un punto interesante de esta propuesta ha sido el desarrollo de una aplicación móvil que permite visualizar el histórico de la BG y la posible trayectoria futura. Los autores comparan su método con modelos basado en TCN, CRNN, LSTM, Bi-LSTM, SVR y ARIMA. Evalúan el rendimiento de los modelos con las métricas habituales RMSE y MAE. Los resultados muestran que la técnica propuesta tiene un rendimiento superior en líneas generales al resto de técnicas comparadas en los tres conjuntos de datos probados.

En [42] tenemos un ejemplo de trabajo que hace uso de modelos basados en NN y datos generados mediante simulador. En concreto, se emplean modelos basados en LSTM y CNN para predecir el valor de glucosa a 30 y 60 minutos. El conjunto de datos lo forman 10 pacientes pediátricos virtuales generados in-silico mediante el simulador UVA/PADOVA. Se han simulado las variables de glucosa, insulinas administradas e ingestas de carbohidratos para los 10 pacientes. Los modelos han sido adaptados para que la inferencia se realice en dispositivos Edge Computing (computación en el borde), en concreto, han usado una Raspberry Pi 4 y Coral Devboard TPU. Los modelos han sido implementados usando la librería TensorFlow Lite y dependiendo del dispositivo, se ha usado el formato de cuantización (.tflite) para la Raspberry Pi 4 y el formato de cuantización de 8 bits (uint8) para la Coral Devboard TPU. Los resultados muestran que los modelos LSTM ofrecen mejor rendimiento usando el formato *.tflite* y los modelos CNN usando el formato *uint8*, pero ambos ofrecen resultados clínicamente aceptables. El trabajo concluye que el uso de dispositivos Edge Computing permite predicciones en tiempo real, sin una caída de rendimiento significativa con respecto a otros dispositivos.

Otra propuesta que hace uso modelos basados en LSTM es [43]. Los autores proponen un modelo predictivo del valor exacto de la BG empleando modelos LSTM, Bi-LSTM y ConvLSTM con horizontes de predicción a 30, 60, 90 y 180 minutos. Los modelos han sido entrenados usando los datos del conjunto de datos público *T1DiabetesGranada* [44]. Este conjunto de datos lo componen 736 pacientes con T1DM y más de 4 años de recogida de datos, los cuales han sido registrados con el MCG FreeStyle Libre 2. Los diferentes modelos han sido entrenados usando el histórico de datos a 120 y 360 minutos anteriores al momento de la predicción en los cuatros horizontes temporales. Un punto

a favor de este trabajo es que realiza predicciones a largo plazo de hasta 180 minutos, no siendo lo habitual en este tipo de trabajos. Los resultados muestran que la mayoría de las predicciones entran en el 95 % de las zonas consideradas como seguras (A+B) del Clarke Error Grid, lo que se puede considerar clínicamente aceptable. Además, el trabajo concluye que no existe un modelo perfecto para todos los pacientes, siendo unos mejores que otros dependiendo del paciente.

Por otro lado, Tena et al. [45] también utiliza modelos basados en LSTM, pero a diferencia de otras propuestas, los autores han realizado una implementación del modelo a nivel hardware, a diferencia como habitualmente se suelen implementar a nivel software. En concreto, han desarrollado un modelo predictivo de glucosa implementado en un dispositivo hardware FPGA Xilinx Virtex-7. Además, han realizado las correspondientes optimizaciones de la arquitectura, así como de los hiperparámetros de la NN, con el objetivo de optimizar lo máximo posible el modelo para integrarlo en hardware de bajo consumo. Los modelos generados permiten predecir el nivel de glucosa con horizonte temporal a 30 minutos y han sido entrenados usando el conjunto de datos público OhioT1DM. Para comparar el rendimiento de modelo LSTM propuesto basado en hardware, lo han comparado con otras implementaciones basadas en FPGA e implementaciones software en diferentes dispositivos móviles, tanto con sistema operativo Android como iOS. El análisis de resultados demuestra que el modelo desarrollado en este trabajo es más eficiente que las implementaciones software y ofrece mejor rendimiento por vatio consumido con respecto a otras implementaciones FPGA que existen en la literatura.

En [46] los autores comparan varias técnicas, en concreto dos modelos basados en ML (*Last Observation Carried Forward* y *Regresión Lineal*), dos modelos ensemble basados en ML (*Least-squares Boosting* y *Bagging*) y dos modelos basados en NN. Para todos los modelos, establece diferentes horizontes temporales de entre 10 y 60 minutos. Han usado el conjunto de datos Aleppo, formado por los datos reales de 225 pacientes con T1DM recogidos usando el MCG Dexcom G4. Los resultados muestran que los modelos avanzados basados en técnicas de ML ofrecen unos resultados bastantes competitivos, siendo *Bagging* el mejor modelo en cuestión, obteniendo un RMSE de 34,33 y MARD de 18,32 con horizonte de predicción a 60 minutos.

Otras propuestas se centran en predecir la glucosa en un momento determinado del día. En [47] se enfocan en estimar la BG posprandial con horizontes temporales de 60, 120 y 180 minutos. El modelo está basado en una Feed-forward Neural Network (FNN) cuyas entradas son los valores de glucosa, insulinas administradas y carbohidratos, lípidos y fibras absorbidos. Los datos provienen de un conjunto de 8 pacientes con T1DM y T2DM. Han evaluado los modelos usando las métricas habituales de RMSE y MAE. El rendimiento de los modelos es bastante aceptable en los horizontes de 60 y 120 minutos,

obteniéndose una precisión media del 96 %.

3.2. Predicción de Hipoglucemias e Hiperglucemias

Se han hecho varios intentos de resolver el problema de la predicción de la glucosa en general, y más en particular, en predecir los eventos de valores críticos de glucosa, como son las hipoglucemias e hiperglucemias. Al igual que con la predicción del valor exacto de la glucosa, existen numerosas propuestas que hacen uso de modelos basados en técnicas de ML y aprendizaje profundo.

Dentro de las propuestas que hacen uso de técnicas ML, en [48] los autores proponen el uso de un SVM para predecir eventos de hipoglucemia postprandial a un horizonte temporal de 240 minutos tras la comida. Se ha usado un conjunto de datos reales con diez pacientes con diabetes tipo 1 obtenidos mediante un MCG. Los resultados muestran una sensibilidad del 79 % y 71 % para las hipoglucemias y las hipoglucemias severas, respectivamente. La principal limitación de esta propuesta es que se centra únicamente en la hipoglucemia posprandial y no en los eventos hipoglucémicos que puedan producirse durante las 24 horas del día.

En este otro trabajo [49], los autores proponen un modelo basado en ML para predecir hipoglucemias e hiperglucemias con horizonte temporal de 60 minutos. En concreto, han desarrollado un modelo basado en el algoritmo XGBoost. Además, han usado explicaciones de aprendizaje automático (SHAP, del inglés SHapley Additive exPlanations) para identificar las características más relevantes a la hora de predecir los riesgos por paciente. En este trabajo, han usado un conjunto de datos compuesto por los datos reales de 153 pacientes con T1DM. Los resultados muestran un rendimiento bastante decente, con un AUC promedio de 0,95 para las hipoglucemias y un AUC promedio de 0,93 para las hiperglucemias. Como otros trabajos en la literatura, el principal límite de esta propuesta es el horizonte de predicción a corto plazo de 60 minutos.

Vu et al. [50] proponen modelos basados en RF para predecir eventos de hipoglucemia nocturna, con un horizonte temporal ampliado de hasta 360 minutos. Sólo se han utilizado como datos de entrada los valores de glucemia obtenidos por un MCG. El conjunto de datos es grande, unos 9800 pacientes reales y una media de 100 días registrados por paciente. Las predicciones se han dividido en varios enfoques en función de la hora de la noche: noche completa (desde medianoche a 6 de la mañana), noche temprana (desde medianoche a 3 de la mañana) y noche tardía (desde las 3 de la mañana a 6 de la mañana). Los resultados muestran que el modelo tiene una buena precisión a primera hora de la noche, con un AUC de 0,90, y un rendimiento de la predicción después de las 3 de la madrugada con un AUC de 0,75, puede mejorarse. En comparación con otros trabajos,

esta propuesta se limita únicamente a detectar la hipoglucemia nocturna.

En [51] los autores proponen un modelo predictivo de eventos de hipoglucemia con un horizonte temporal a corto plazo de 30 y 60 minutos. Los modelos han sido desarrollados usando los algoritmos de Regresión Logística (LR) y RF. Como entrada, se han utilizado características extraídas de los datos brutos de un MCG. El conjunto de datos comprende los datos reales de 112 pacientes con diabetes de tipo 1 y con datos de 90 días de media por cada paciente. Los resultados muestran que el modelo de predicción propuesto tiene una sensibilidad superior al 91 % en el horizonte de 60 minutos y tiene un alto rendimiento para la hipoglucemia nocturna con una sensibilidad del 95 %.

Seo et al [52], realizan una comparativa de cuatro modelos ML diferentes (RF, KNN, SVM, LR) para predecir eventos de hipoglucemia posprandial con un horizonte temporal de 30 minutos. El conjunto de datos consta de 104 pacientes reales, y para cada paciente se registró el nivel de glucemia durante una media de 3 días utilizando un MCG. Se compararon los diferentes modelos de predicción, y el algoritmo RF obtuvo la mejor precisión en la detección de eventos de hipoglucemia con un rendimiento de predicción medio con un AUC de 0,97, una sensibilidad del 89,6 % y una especificidad del 91,3 %. Este enfoque tiene varias limitaciones; en primer lugar, se centra únicamente en los eventos de hipoglucemia después de las comidas y un horizonte temporal de detección a corto plazo.

Por otro lado, tenemos las propuestas que hacen uso de modelos de ANN. En [53] los autores utilizan modelos basados en LSTM para predecir eventos de hipoglucemia. El horizonte temporal para detectar estos eventos se limita a solo 30 minutos. Los datos se han obtenido utilizando una lente de contacto con un electrodo que mide el nivel de glucosa en sangre contenido en el líquido lagrimal. El conjunto de datos es relativamente pequeño, ya que no se menciona el número de pacientes utilizados; sólo se han utilizado 2.000 mediciones de glucosa. Los resultados obtenidos por el modelo predictivo muestran una precisión del 80 %.

Asimismo, Porumb et al. [54] utilizan redes neuronales, precisamente dos enfoques diferentes: una CNN y una combinación de CNN + RNN. Esta propuesta pretende detectar eventos de hipoglucemia nocturna con una probabilidad de ocurrencia sin un horizonte temporal definido. El conjunto de datos es limitado porque sólo hay datos de cuatro pacientes reales. La señal de cardiograma (ECG) a intervalos de 5 minutos y los niveles de glucemia se han registrado durante 14 días consecutivos para cada paciente. En el modelo CNN, la red se estructura en 15 capas convolucionales más una capa totalmente conectada. El modelo CNN + RNN se compone de una red CNN con cinco capas convolucionales y una red RNN formada por una única capa LSTM. Los resultados obtenidos de los modelos muestran un rendimiento predictivo medio para todos los pacientes con una precisión del 82,4 %, una sensibilidad del 87,5 % y una especificidad del 81,7 %.

En [55], los autores proponen el uso de un modelo Transformer para predecir tanto el valor exacto de glucosa como la detección de episodios de hipoglucemias e hiperglucemias. Los modelos Transformer [56] se pueden definir como una arquitectura de red neuronal que utiliza el mecanismo de atención para procesar secuencias de datos, especialmente eficaz en tareas de procesamiento de lenguaje natural, y son modelos que están empezando a remplazar a las RNN y CNN. Los autores han abordado el problema desde varios enfoques, es decir, han abordado la predicción de la glucosa como un problema de regresión y la detección de las hipoglucemias e hiperglucemias como un problema de clasificación. El conjunto de datos de este trabajo lo forman 226 pacientes con T1DM y 104 pacientes con T2DM, los cuales han sido registrados usando el MCG Dexcom G5. Como punto a destacar, se ha aplicado una técnica de aumentado de datos basada en un modelo Generative Adversarial Networks (GAN), en cuestión el modelo ECGGAN [57]. Se ha realizado una comparación con diferentes tamaños de entrada al modelo, o sea, realizando la inferencia con los datos de 60, 120 y 180 minutos previos al momento de la predicción. Además, los autores han comparado el método propuesto con un modelo basado en redes Gated Recurrent Unit (GRU). El análisis de los resultados muestra que el método propuesto ofrece mejores resultados con respecto al modelo GRU, pero únicamente prueban con un horizonte temporal de 30 minutos, siendo una ventana de predicción muy corta y se desconoce el rendimiento del modelo con predicciones a largo plazo (más de 30 minutos).

Por último, en [58] se puede encontrar una revisión sistemática de la literatura sobre algoritmos y modelos basados en datos que utilizan datos reales de personas con diabetes para la predicción de la glucemia y la hipoglucemia. Como muestran Felizardo et al., todos los enfoques se basan en el uso de datos registrados directamente o con los típicos pasos de preprocesamiento de datos.

3.3. Transformada de Wavelet

La clasificación de series temporales se ha abordado desde distintos puntos de vista en la literatura científica. Inicialmente, la idea era encontrar una función de distancia para establecer un sentido de proximidad entre distintas series temporales que permitiera agruparlas. La idea básica es utilizar la distancia euclídea y, a continuación, agrupar las series mediante un algoritmo como KNN. Sin embargo, la distancia euclídea no puede captar la estructura y los patrones de comportamiento de las series temporales complejas. Los investigadores han probado diferentes funciones de distancia para abordar este problema, siendo quizá la más conocida y exitosa la técnica Dynamic Time Warping [59, 60]. Un estudio que recopila el conocimiento de las técnicas de clasificación de series temporales antes de la llegada del aprendizaje profundo es el trabajo de Wang et al. [61].

En el trabajo de Ye et al. [62], encontramos la idea de transformar formas complejas

encontradas en la naturaleza (como las hojas de las plantas) en series temporales y aplicarles después técnicas de clasificación de series temporales. Sin embargo, este enfoque también puede funcionar en la otra dirección, es decir, transformando las series temporales a otros dominios para utilizar diferentes funciones de distancia y técnicas de clasificación. Un ejemplo de esta idea es la transformación de las series temporales en una secuencia de caracteres para su posterior uso de técnicas de procesamiento del lenguaje [63]. Sin embargo, la opción más adoptada es utilizar una transformación matemática como paso previo al uso de técnicas de clasificación, como podemos encontrar en [64]. El análisis de series temporales mediante transformaciones matemáticas tiene más de medio siglo de antigüedad [65]. Un resumen de las transformaciones utilizadas para clasificar series temporales que no incluyen la transformada wavelet se puede encontrar en la encuesta de Ji et al. [64].

De todas las transformaciones matemáticas empleadas con éxito en los últimos años, la transformada wavelet [66] merece una mención especial. Los investigadores del mundo financiero fueron probablemente los primeros en vislumbrar las enormes posibilidades de aplicar la transformada wavelet a las series temporales. Como ejemplo, tenemos los trabajos de Masset (2015) [67], y Bolman et al. [68]. Esta investigación se ha extendido a otras series temporales, como en Grant e Islam [69], o en Li et al., 2016 [70]. También podemos encontrar ejemplos que mezclan wavelets y técnicas modernas como las redes neuronales convolucionales [71].

Recientemente, podemos encontrar ejemplos de investigadores interesados en la clasificación de series temporales multivariantes, Pasos et al. [72], y Middlehurst et al. [73]. El segundo estudia la inclusión de un conjunto de clasificadores, una opción que también podría utilizarse para series temporales univariantes.

En el campo de la diabetes, encontramos el ejemplo de Ashok et al. [74]. Este enfoque consiste en un sistema para medir la glucosa de forma no invasiva utilizando un rayo láser y un fotodetector. Se utiliza la transformada wavelet de Haar para descomponer las señales recibidas por el sensor. Un modelo predictivo basado en una red neuronal con backpropagation se encarga de la predicción de la glucosa.

3.4. Conclusiones

Se ha abordado ampliamente en la literatura el problema de la predicción de la glucosa, así como la predicción de hipoglucemias e hiperglucemias. Existen diferentes técnicas a la hora de construir los modelos predictivos de glucosa, siendo las más usadas a día de hoy las basadas en modelos de ML y NN.

Una de las principales dificultades con la que se encuentran los investigadores a la hora de afrontar este problema es la falta de datos disponibles dado el carácter sensible de los mismos. Esto es importante, porque cuando más información reciba el modelo, tendrá más robustez a la hora de realizar las predicciones. Para solucionar este punto, existen repositorios de datos tanto de acceso público o mediante petición, así como simuladores de datos.

Tal como se ha introducido en el Capítulo 2, este trabajo de tesis parte con el objetivo de predecir el valor exacto de Glucosa. En este enfoque, la mayoría de los trabajos presentes en la literatura se enfocan en horizontes de predicción a corto plazo dada la dificultad de predecir la glucosa a largo plazo. Nuestro trabajo parte de esta premisa y el uso de una técnica de lógica difusa que no ha sido aplicada anteriormente al problema de la predicción de la glucosa.

Además, como continuación del trabajo previo anterior, hemos ido un paso más allá y hemos enfocado la segunda parte de este trabajo en la predicción de eventos de hipoglucemias e hiperglucemias. En la literatura se encuentran numerosas propuestas que, al igual que la predicción de la glucosa, se centran en horizontes de predicción a corto plazo o en predecir únicamente estos eventos glucémicos en un momento determinado del día, es decir, posprandiales, nocturnos, etc. Partiendo de este punto, este trabajo de tesis ha tenido como objetivo desarrollar modelos predictivos de hipoglucemias e hiperglucemias con un horizonte de predicción de hasta 24 horas mediante el uso de técnicas que no han sido aplicadas previamente a este problema.

Capítulo 4

Predicción de glucosa mediante un sistema basado en reglas difusas TSK de dos fases: METSK-HD

La lógica difusa, propuesta por Lotfi Zadeh [75], es un paradigma que difiere del enfoque conocido de la lógica clásica o booleana. Mientras que la lógica booleana se basa en una distinción clara entre lo que es verdadero y lo que es falso, la lógica difusa es una lógica multivaluada, es decir, permite que las variables puedan tomar valores intermedios entre estos dos extremos. En lugar de afirmar categóricamente si algo es verdadero o falso, la lógica difusa permite expresar grados de pertenencia (grados de verdad), de esta manera se puede reflejar formalmente la imprecisión e incertidumbre de muchas situaciones de la vida real.

El principal concepto de la lógica difusa son los denominados Conjunto Difusos. Un conjunto difuso se puede definir como el grado de pertenencia que puede tener cada elemento al conjunto, es decir, el valor que puede tomar la función de pertenencia (MF, del inglés Membership Function) de cada objeto entre la pertenencia total (valor uno) o no pertenencia (valor cero). En cambio, en la lógica clásica, la función de pertenencia de los elementos al conjunto únicamente puede tomar los valores de pertenencia (valor uno) y no pertenencia (valor cero). Estos conjuntos clásicos (o conjuntos crisp) se pueden considerar un tipo particular de conjuntos difusos con únicamente dos grados de pertenencia (uno y cero). Gracias a estos conjuntos, se puede representar la imprecisión que representan ciertos tipos de objetos, lo cual es fundamental para abordar problemas del mundo real donde las categorías no siempre están claramente definidas. En la Figura 4.1 se muestra un ejemplo de lógica difusa, donde se puede observar que existen tres términos asociados a

la temperatura, siendo estos “Frío”, “Templado”, “Caliente”. Estos términos se denominan variables lingüísticas y el significado se determina mediante los conjuntos difusos.

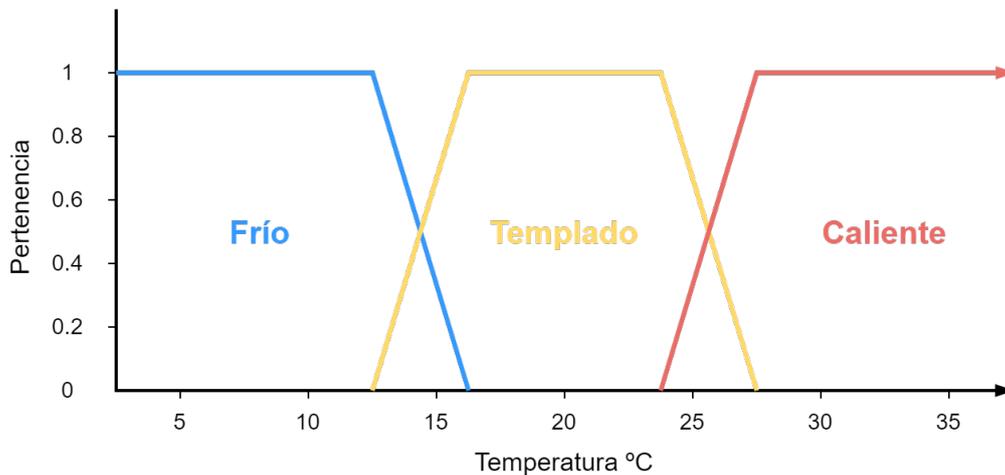


Figura 4.1: Ejemplo de lógica difusa: Temperatura

Dentro de la lógica difusa, nos encontramos con los denominados sistemas basados en reglas difusas (FRBS o Fuzzy rule-based systems de sus siglas en inglés). Estos sistemas habitualmente se han usado como sistemas de control, demostrando su capacidad para dar muy buenas respuestas en situaciones en las que se necesita un control automático. Tal como su nombre indica, estos sistemas están basados en reglas y tienen un comportamiento similar al humano en la forma de la toma de decisiones. En concreto, estos sistemas realizan las decisiones siguiendo una serie de reglas con cierto grado de incertidumbre.

Un ejemplo de FRBS puede ser un termostato, al contar con una serie de reglas basadas en función al valor de la temperatura ambiente. Estas reglas pueden ser, de forma muy simplificada, las siguientes:

- Si la temperatura es baja, entonces enciende la calefacción a máxima potencia
- Si la temperatura es templada, entonces enciende la calefacción a media potencia
- Si la temperatura es alta, entonces apaga la calefacción

Este sistema de ejemplo toma las decisiones basándose en términos lingüísticos como temperatura "Baja", "Templada" y "Alta", en lugar de tomar las decisiones en base al valor exacto de la temperatura. De esta forma, se facilita la adaptación del sistema a diferentes condiciones y reduce la necesidad de una modelización precisa del sistema.

CAPÍTULO 4. PREDICCIÓN DE GLUCOSA MEDIANTE UN SISTEMA BASADO EN REGLAS DIFUSAS TSK DE DOS FASES: METSK-HD

Una de las principales dificultades de los FRBS, es la parte de su diseño siendo necesario contar con amplios comedimientos del problema a abordar para su desarrollo. Habitualmente se diseñan ad-hoc, es decir, de forma específica para un problema en concreto sin necesidad de seguir el enfoque estándar. Gracias a la incorporación de GA al proceso de diseño, se han podido abordar problemas más complejos, pero a costa de perder interpretabilidad a la hora de entender el funcionamiento del modelo. Esto ha permitido que los FRBS puedan ser empleados en entornos dónde no se habían utilizado anteriormente. En concreto, se ha demostrado que estos sistemas son válidos para ser aplicados en problemas de predicción, al obtenerse predicciones muy afinadas [76].

Existen diferentes tipos de FRBS, entre los más conocidos están los sistemas Mandani y Takagi-Sugeno-Kang (TSK). Los sistemas Mandani fueron introducidos por Ebrahim Mamdani [77] y su funcionamiento se basa en cuatro etapas: fuzificación, evaluación de las reglas, agregación de las salidas y desfuzificación. La fuzificación consiste en trasladar los valores reales de las variables de entrada a cada una de las funciones de pertenencia, para así determinar el valor de pertenencia de la variable a la función difusa. La segunda parte es la evaluación de las reglas, se debe obtener el valor de respuesta del sistema según las reglas activadas, que dependerán a su vez de la pertenencia de cada una de las variables a las funciones de pertenencia. La tercera fase sería la agregación de las salidas, es decir, con las reglas que se disparan, determinar el valor de las salidas de su consecuente. Por último, se necesita conocer el valor real que debe arrojar el sistema, para ello se desfuzifica la salida. Aplicando diferentes métodos de desfuzificación se obtiene el valor de respuesta del sistema. Por otro lado, los sistemas TSK [78], a diferencia de los Mandani, no requieren realizar ninguna operación de defuzificación de las variables y el procesamiento de las reglas es diferente. Mientras en un sistema Mandani cada regla especifica la relación de las variables de entrada con la salida difusa, en los sistemas TSK la salida de cada regla se puede representar como una función lineal en base a las variables entradas, con coeficientes específicos para cada regla. El valor de respuesta del sistema TSK se calcula según la ponderación de las reglas disparadas en el proceso de agregación. Los sistemas Mandani están pensados para resolver problemas dónde no existe una relación lineal entre las variables, mientras que los sistemas TSK son más adecuados en problemas donde existe una relación lineal entre las variables de entrada y salida.

4.1. FRBS-TSK

En este primer enfoque de este trabajo de tesis, proponemos el uso de un modelo basado en un FRBS-TSK al ser el que mejor se adapta a las particularidades del problema de la predicción del valor de BG. En concreto, consideramos la hipótesis inicial que el sistema FRBS va a ser capaz de encontrar la relación lineal de las variables de entrada de este

problema con la variable final a predecir. De esta forma, esperamos obtener modelos con un comportamiento similar a los modelos fisiológicos.

El modelo predictivo de glucosa, tal como hemos introducido, será generado en base a reglas difusas siguiendo el modelo TSK. Además, la base de conocimiento (KB, del inglés Knowledge Base) y la base de datos de reglas (RB, del inglés Rule Base) necesarias para dicho modelo se optimizarán con un algoritmo evolutivo (EA, del inglés Evolutionary Algorithm). Un EA es un algoritmo inspirado en los procesos de evolución biológica y su funcionamiento está basado en la teoría de la evolución y selección natural, aplicando los conceptos de reproducción, cruce, mutación y elitismo. Estos algoritmos son ampliamente utilizados en problemas de optimización y búsqueda. Se ha abordado la optimización del modelo predictivo mediante un algoritmo genético perteneciente al grupo de los EA para llevar a cabo la optimización del modelo. Este proceso es realizado en dos fases:

- **Fase de aprendizaje del modelo:** en esta fase se optimizan las funciones de pertenencia y las reglas.
- **Fase de ajuste del modelo:** se optimizan el uso de las reglas, así como realización de mejoras en las funciones de pertenencia.

Una vez ha sido optimizado el modelo, este puede ser empleado para predecir el valor de la glucosa para diferentes pacientes con horizontes temporales a 30, 60, 90 y 120 minutos. Pero antes de poder realizar esta predicción, estos modelos tomarán una serie de variables de entrada, las cuales serán usadas para llevar a cabo la inferencia fuzzy que tendrá con resultado el valor de glucosa a predecir. El listado de variables de entrada y salida del modelo FRBS-TSK son las siguientes:

- **Variables de entrada**
 - Nivel de glucosa en el momento de la predicción.
 - Valores de glucosa de las dos horas anteriores al momento de la predicción t .
 - Cantidad de hidratos de carbono ingeridos en el momento de la predicción (si los hay).
 - Cantidad de raciones de hidratos ingeridas en las dos horas anteriores al momento de la predicción.
 - Unidades de insulina administradas en el momento de la predicción (si las hay).
 - Unidades de insulina administradas en las dos horas anteriores a la predicción.
- **Variable de salida**

- Valor estimado de glucosa a los H minutos después del momento de la predicción, siendo $H \in \{30, 60, 90, 120\}$ el horizonte de predicción en minutos.

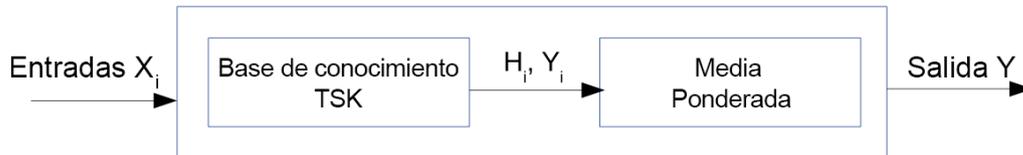


Figura 4.2: Flujo del modelo FRBS-TSK.

4.2. Método Propuesto

El método propuesto en este trabajo está basado en el sistema METSK-HD (Multi-Objective Evolutionary Learning of TSK systems for high-dimensional problems with estimated error) [79], el cuál es una derivación de FRBS-TSK estándar. En concreto, se trata de una evolución que permite trabajar con problemas de alta dimensionalidad y datos masivos, ajustándose perfectamente al problema de la pérdida de la glucosa en sangre.

Este sistema al derivar de un TSK, mantiene la estructura básica de los mismos. En general, estos sistemas están formados por una KB y una RB. La KB se encarga de almacenar el conocimiento que ha sido extraído del problema y la RB está compuesta por las relaciones entre las variables establecidas por las funciones de pertenencia (MF). El conocimiento del problema se representa mediante reglas difusas. Estas tienen una estructura simple del tipo (Si-Entonces) formadas por el antecedente (Si) y el consecuente (Entonces). El antecedente lo forman las variables lingüísticas, es decir, variables que describen la condición que debe cumplirse para aplicar la regla en cuestión. Por otro lado, el consecuente es la acción que se llevará a cabo si se cumple las condiciones del antecedente. Dicha acción tiene como salida el resultado de una función polinómica con las variables de entrada. En resumen, el formato de las reglas de un FRBS-TSK tiene la siguiente estructura:

$$\begin{array}{l} \text{Si } X_1 \text{ es } A_1 \dots \text{ y } X_n \text{ es } A_n \\ \text{Entonces } Y = p_1 * X_1 + \dots + p_n * X_n + p_0 \end{array}$$

donde las variables de entrada del sistema se denotan como X_i , Y es la variable de salida del sistema, p_i representa coeficientes de valores reales y A_i son conjuntos difusos.

En la Figura 4.2 se puede observar el flujo del modelo propuesto FRBS-TSK. Teniendo en cuenta que la KB contiene m reglas difusas TSK, la salida del sistema TSK se calcula como la \bar{x} ponderada de la salida de cada regla individual Y_i como se observa en (4.1):

$$\frac{\sum_{i=1}^m h_i \cdot Y_i}{\sum_{i=1}^m h_i} \quad (4.1)$$

dónde $i = 1 \dots m$, $h_i = T(A_1(x_1) \dots A_n(x_n))$ representa el grado de coincidencia entre el antecedente de la regla i th y las entradas del sistema $x = (x_1 \dots x_n)$, y siendo T una T-norma, es decir, la operación que se utiliza para combinar las activaciones de las reglas en la etapa de inferencia difusa.

Los modelos FRBS-TSK se han aplicado con éxito a un gran número de problemas, en los que existe una alta dimensionalidad de los datos. La novedad del sistema propuesto es el uso de una ecuación de sistema compacta para la estimación del parámetro p_i , mediante el empleo de métodos clásicos. Esto tiene como resultado un sistema con una precisión que resulta muy útil para la modelización difusa precisa.

Al trabajar con conjuntos de datos de alta dimensionalidad, los sistemas basados en reglas difusas ad-hoc, es decir, diseñados específicamente para el problema en cuestión, conllevan una dificultad añadida por el elevado tamaño de los datos a tratar. Por este motivo es necesario contar con sistemas que permitan diseñar de forma automática el conjunto de KB y RB. Se ha demostrado que el uso de los denominados Sistemas Difusos Genéticos (GFS, del inglés Genetic Fuzzy Systems) ofrece muy buenos resultados [76] a la hora de diseñar estos sistemas. Estos sistemas están basados en EA y son capaces de diseñar de forma conjunta la KB y RB, optimizando en el proceso de aprendizaje las respectivas MF de los conjuntos difusos.

En la Figura 4.4 se muestra un resumen del proceso de diseño y optimización empleado para generar los modelos mediante el algoritmo METSK-HD. Este proceso se divide en dos fases, *Aprendizaje* y *Ajuste*. En la fase de Aprendizaje, se genera la KB inicial, formada por un conjunto de reglas TSK de orden cero, es decir, se inicializan las reglas de forma muy simplificada. Este aprendizaje es realizado mediante el uso de un Algoritmo Evolutivo Multiobjetivo (MOEA, del inglés Multi-Objective Evolutionary Algorithm) [80, 81]. En la segunda fase, se lleva a cabo un ajuste fino de las KB y RB. En este post-procesado se lleva a cabo una selección de las reglas, es decir, mediante un GA se realiza un ajuste de forma que el sistema sea funcional con el mínimo de reglas posibles. El proceso descrito anteriormente se puede observar en la Figura 4.3.

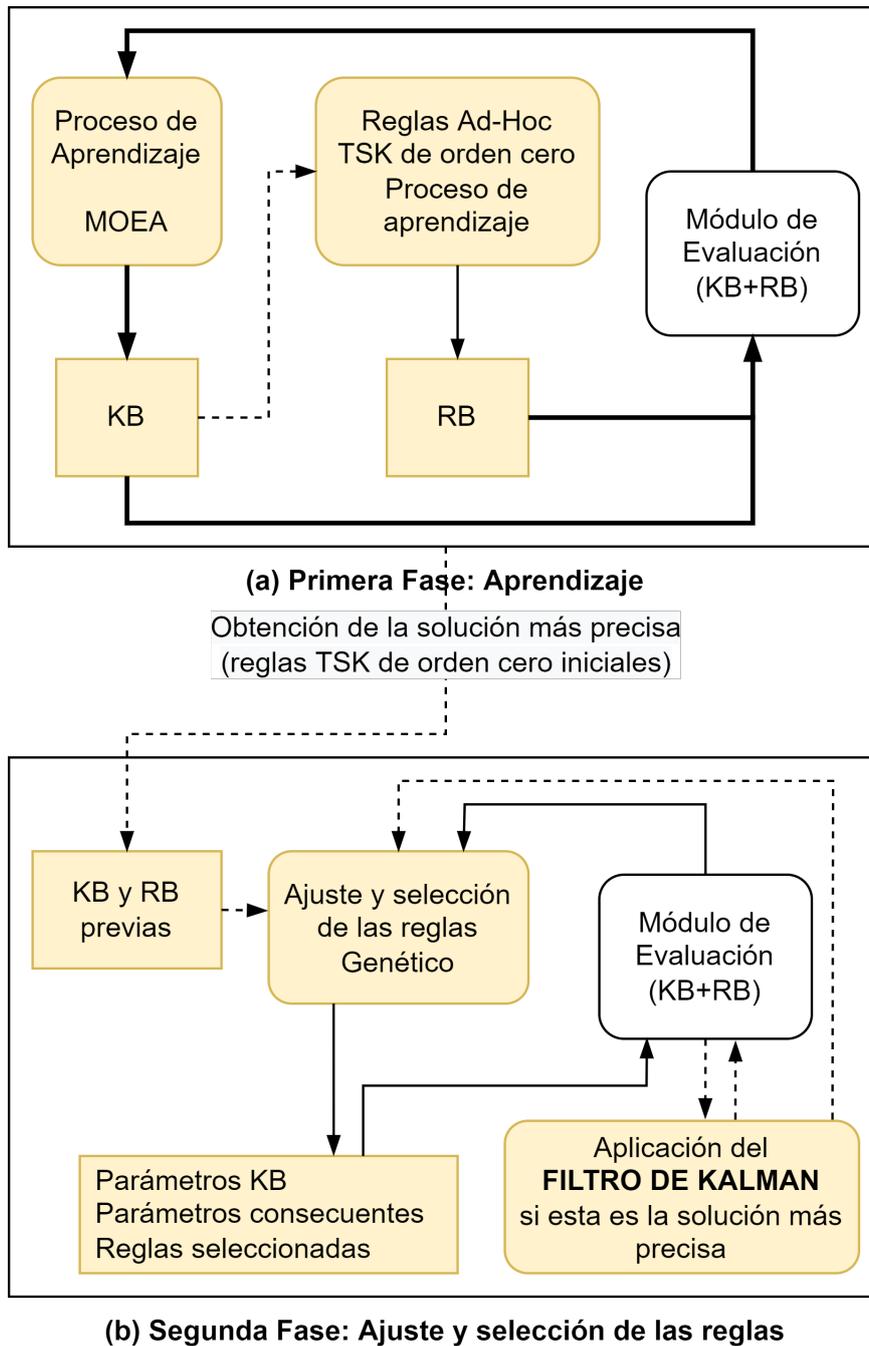


Figura 4.3: Proceso de optimización del FRBS-TSK.

4.3. Fases de optimización del método METSK-HD

En esta sección detallaremos en las fases en las que se divide el proceso de optimización del método METSK-HD.

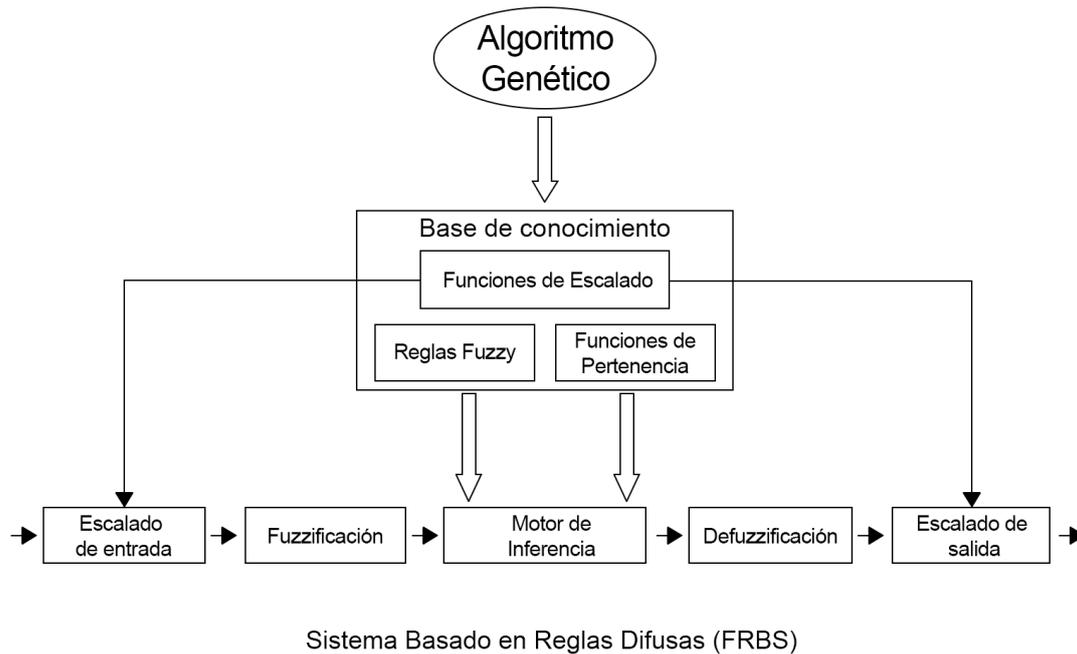


Figura 4.4: Procesos implicados en el módulo FRBS.

En la primera fase de *aprendizaje*, la KB inicial es generada usando un MOEA. Con este algoritmo se diseña el sistema inicial con el objetivo de reducir la dimensionalidad de forma que se usen mecanismos eficaces que garanticen una convergencia rápida en conjuntos de datos grandes. Este proceso se divide en:

- **Generación de la KB.** Para obtener la KB completa, el algoritmo METSK-HD utiliza el conocido método de Wang y Mendel (WM) [82]. En la generación de la KB inicial, se ha puesto un límite de un máximo de 100 reglas. Dado que son reglas para un sistema TSK, estas son generadas con todos los coeficientes con valor 0 y los términos lingüísticos independientes calculados como la media de los ejemplos cubiertos por la regla ponderados por su coincidencia.
- **Codificación de RB.** Se ha considerado un sistema de representación lingüística formado por 2 tuplas (véase la Figura 4.5). Gracias a esto, se puede utilizar un único parámetro en el proceso de optimización, siendo el desplazamiento lateral de las funciones de pertenencia. Esto difiere de otros sistemas que requieren el uso de los tres parámetros para describir la forma triangular de las funciones de pertenencia usadas. En este enfoque, se utiliza el mismo desplazamiento $[-0, 1, 0, 1]$ para todas las variables.

En resumen, los principales objetivos en este proceso de optimización son:

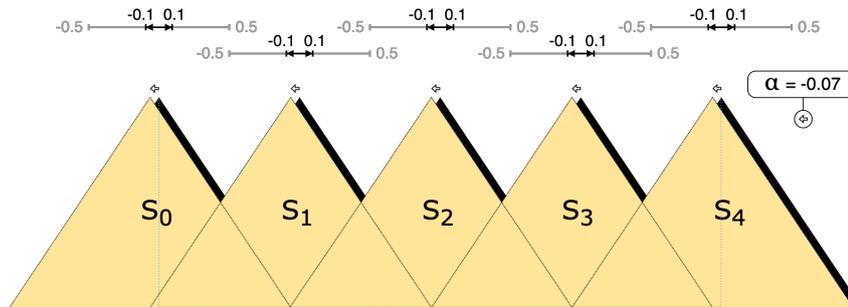


Figura 4.5: Desplazamiento lateral MF en $[-0,1,0,1]$

- Minimizar el RMSE.
- Minimizar el número de reglas.
- Maximizar el grado de cobertura de los ejemplos.

Para ello, se parte de una población inicial donde cada cromosoma tiene el mismo número de etiquetas lingüísticas para todas las variables del sistema (siendo mínimo 2 y máximo 7), junto con un desplazamiento lateral inicial calculado aleatoriamente. Se ha usado como operador de cruce el Parent Centric BLX (PCBLX) [83] y por otro lado el operador de mutación se encargará de disminuir la granularidad en un gen seleccionado aleatoriamente. Además, se ha incluido un mecanismo para prevenir el incesto en las partes siguiendo los conceptos del CHC [84].

En la segunda fase, fase de *ajuste*, una vez obtenida la KB completa de orden cero en la fase anterior, se aplica un post-procesamiento para realizar el ajuste de MF y la selección de las reglas mediante el uso de un MOEA. El objetivo de este ajuste es la mejora significativa de la precisión del modelo al reducirse el número de reglas necesarias. El esquema de codificación C es el siguiente:

$$C = C_t + C_s + C_c$$

dónde:

- Esquema clásico de funciones de pertenencia representado por un triángulo (véase Figura 4.6) y ajustados en la segunda fase. C_t
- Codificación de la selección de reglas (1 variable asignada 0 no asignada) C_s
- Codificación de los consecuentes de las reglas TSK. C_c

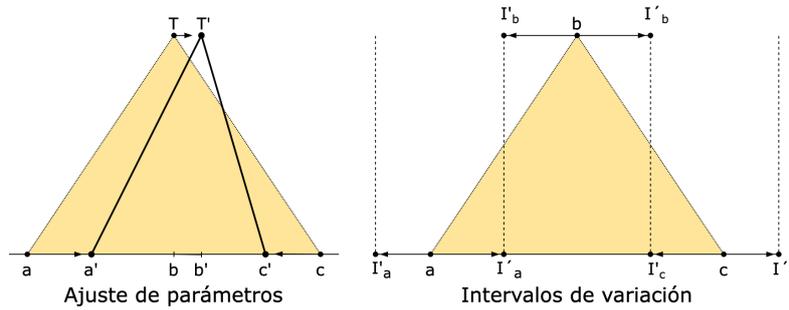


Figura 4.6: MF triangular clásico.

El control de incesto es aplicado en la parte del cromosoma C_t , el cruce BLX-0.5 a C_c y la mutación se aplica a C_s .

Los coeficientes de la función consecuyente polinómica en las reglas TSK son estimados mediante la técnica clásica del filtro de Kalman [85]. Por último, el algoritmo de optimización cuenta con un operador de reinicio en caso de que un valor de umbral L tome un valor inferior a 0. Este operador de reinicio es aplicado copiando sólo el mejor individuo para cada uno de los tres objetivos, y estos tres individuos formarán parte de una nueva población inicial.

4.4. Resultados

4.4.1. Configuración Experimental

Los conjuntos de datos utilizados (véase Sección 4.4.2) han sido preparados siguiendo la metodología clásica de división en dos subconjuntos: 70 % de los datos para entrenamiento y 30 % de los datos para test. Los modelos han sido entrenados por paciente, es decir, se tratan de modelos personales usando únicamente los datos de cada paciente.

Para cada modelo, se ha realizado una validación cruzada K-Fold. Este tipo de validación sirve para evaluar el rendimiento de los modelos y consiste en dividir todo el conjunto de datos en k subconjuntos denominados pliegues o *fold*s, donde k es un número entero mayor que 1. A continuación, el modelo se entrena k veces, utilizando un *fold* diferente como conjunto de validación en cada iteración y el resto de los *fold* como conjunto de entrenamiento. En concreto, se ha usado una validación cruzada 8-Fold, es decir, se ha dividido el conjunto de datos en 8 pliegues diferentes.

Se han llevado a cabo ocho tipos de experimentos usando los escenarios (*What-If* y *Agnóstico*) y cuatro horizontes temporales diferentes para la predicción del valor BG: 30, 60, 90 y 120 minutos. Por cada experimento se han realizado 80 ejecuciones, 10 por cada

“fold”. Además, se han considerado las siguientes tres opciones para la selección del mejor modelo a probar con los datos del subconjunto de test:

- El modelo que ha obtenido mejor rendimiento en la fase de entrenamiento.
- Un conjunto de los mejores modelos durante la fase de entrenamiento. La predicción final en el subconjunto de prueba es la media de todas las predicciones de todos los modelos del conjunto.
- En la tercera opción, nos quedamos con el modelo que ha obtenido el mejor rendimiento durante la fase de validación. Como se ha realizado una validación cruzada 8-Fold, el modelo seleccionado será el que use los datos del pliegue o “fold” que mejor resultado ha obtenido dentro de todos los folds.

4.4.2. Conjunto de Datos

Tal como hemos introducido en capítulos anteriores, la obtención de datos clínicos dentro del ámbito de la medicina no es una tarea sencilla. Se trata de información de carácter sensible y además para su recogida, es necesario contar con la aprobación del estudio a realizar por parte del comité ético del hospital, así como el consentimiento de los pacientes. Además, los pacientes deben llevar dispositivos, registrar la información y tomar notas durante los días que dure el estudio con las menores molestias posibles. Esto conlleva que sea habitual descartar datos por fallos en los dispositivos de medición, así como errores por parte del paciente. Por estos motivos, los estudios en el campo de la medicina y en concreto, de la diabetes, suelen tratar con conjuntos de datos pequeños, habitualmente de diez pacientes o incluso menos.

Para el primer enfoque de este trabajo de tesis, se ha usado un conjunto de datos recogido gracias a la colaboración del grupo ABSys¹ de la Universidad Complutense de Madrid con el Hospital Príncipe de Asturias de Alcalá de Henares (España). El comité ético del hospital aprobó la recogida de datos para los estudios y todos los pacientes firmaron previamente un consentimiento.

El conjunto de datos está constituido por los datos de diez pacientes con T1DM. Los datos se han registrados en intervalos de cinco minutos utilizando el sistema MCG Guardian de Medtronic durante al menos diez días completos. Las mediciones no han sido necesariamente realizadas en días consecutivos ni durante el mismo número de días por paciente. La media de días con datos es de 44,80 (+/- 30,73). Las características de los pacientes son mujeres (80%), edad media 42,30 (+/- 11,07), peso 64,78 (+/- 13,31)

¹<http://absys.dacya.ucm.es/>

kg, años de enfermedad 27,20 (+/- 10,32), años con terapia de bomba 10,00 (+/- 4,98), hemoglobina glucosilada (HbA1c) promedia de 7,27 % (+/- 0,5 %).

4.4.3. Escenarios What-If y Agnóstico

Se han usado los escenarios *What-If* y *Agnóstico* para la predicción del valor exacto de BG (véase Sección 2.2.1). La principal diferencia entre ambos escenarios radica en la información empleada a la hora de construir los modelos. En el escenario *What-If* además de incluir la información anterior al momento de la predicción, se pueden incluir eventos futuros en algunas de las variables de entrada. En concreto, hemos incluido la información futura de las ingestas de hidratos de carbono y las dosis de insulinas desde el momento de la predicción hasta el horizonte temporal de la predicción. En el caso del escenario *Agnóstico*, sólo se tiene en cuenta la información de las variables de entrada que son anteriores al instante de tiempo inmediatamente posterior al momento actual, es decir, al momento de la petición de la predicción.

4.4.4. Resultados Experimentales

En las Tablas 4.1, 4.2 y 4.3, podemos encontrar los resultados obtenidos usando el subconjunto de test de las diferentes opciones explicadas en la Sección 4.4.1. Las tablas muestran el RMSE para los diez pacientes en los ocho tipos de experimentos explicados en la sección anterior (dos escenarios y cuatro horizontes temporales). Los valores expresados en las columnas de la tabla indican el promedio del RMSE de las 80 ejecuciones para cada tipo de modelo. Las dos filas inferiores muestran la media y la desviación estándar del valor RMSE de los 10 pacientes.

Tal como se puede observar, seleccionar el mejor modelo durante la fase de validación (Tabla 4.3) mejora los resultados obtenidos con respecto a utilizar el mejor modelo durante la fase de entrenamiento (Tabla 4.1). Y a su vez, las predicciones del modelo Ensemble (Tabla 4.2) mejoran ligeramente las cifras obtenidas por el mejor modelo en la fase de validación. Esta situación no cambia para los cuatro horizontes de predicción y los dos escenarios presentados en este trabajo de tesis.

La Tabla 4.4 muestra en la que todos los modelos han sido ejecutados en la fase de test siendo seleccionado el que ha obtenido el mejor resultado. Obviamente, esto no es una opción realista, pero simplemente se han incluido estos resultados para mostrar que el método tiene margen de mejora.

Por último, en la Tabla 4.5 se ha realizado una comparativa el método propuesto frente a los mejores modelos registrados en el Blood Glucose Level Prediction (BGLP) Challenge [94]. Los resultados obtenidos por los diferentes modelos METSK-HD son similares con

CAPÍTULO 4. PREDICCIÓN DE GLUCOSA MEDIANTE UN SISTEMA BASADO EN REGLAS DIFUSAS TSK DE DOS FASES: METSK-HD

Paciente	What-If				Agnóstico			
	t+30	t+60	t+90	t+120	t+30	t+60	t+90	t+120
1	31,62	46,30	47,86	51,04	30,58	30,93	42,02	56,21
2	19,51	33,01	43,76	51,22	20,04	32,97	44,19	51,18
3	30,27	31,62	41,99	40,53	18,11	28,48	45,27	48,94
4	21,44	31,57	42,88	51,47	21,99	33,97	39,39	46,27
5	31,63	43,67	59,23	56,17	31,43	48,53	55,83	60,66
6	27,32	41,37	52,44	53,14	28,09	44,32	49,45	60,38
7	46,55	55,91	67,37	65,53	46,96	59,76	65,98	52,48
8	30,01	41,60	42,49	43,20	31,77	42,81	49,39	55,29
9	24,58	39,27	53,41	64,74	24,08	43,50	52,81	65,46
10	37,38	57,32	79,12	92,59	36,97	59,83	88,39	144,41
Media	30,03	42,16	53,36	56,96	29,00	42,51	53,27	64,13
Desv.	7,84	9,16	12,25	14,82	8,66	11,22	14,52	28,80

Tabla 4.1: Mejor modelo en la fase de entrenamiento

Paciente	What-If				Agnóstico			
	t+30	t+60	t+90	t+120	t+30	t+60	t+90	t+120
1	21,24	32,63	38,26	42,17	21,10	33,10	41,44	47,37
2	18,28	30,11	38,44	43,43	18,30	29,93	38,30	44,39
3	19,15	27,91	32,79	35,53	19,05	28,07	33,90	37,10
4	19,59	31,74	40,02	44,76	19,57	31,87	40,56	45,80
5	29,49	43,30	50,12	54,05	29,41	43,59	50,81	54,53
6	24,98	38,68	45,27	48,66	25,00	39,39	47,43	51,78
7	25,13	38,54	46,59	50,91	25,27	38,87	47,99	52,50
8	27,79	35,75	37,64	38,38	29,48	40,89	43,42	44,16
9	23,21	36,35	42,80	45,19	23,16	36,94	46,37	48,38
10	24,34	43,42	56,50	63,64	24,43	43,39	57,15	67,76
Media	23,32	35,84	42,84	46,67	23,48	36,60	44,74	49,38
Desv.	3,74	5,29	6,96	8,12	4,01	5,56	6,67	8,17

Tabla 4.2: Mejor modelo Ensemble.

respecto al uso de otras técnicas, pero se puede observar que la aplicación de esta técnica al problema de la predicción de la glucosa no ofrece los mejores resultados. Cabe destacar que el conjunto de datos utilizado en el BGLP está formado por 6 pacientes con características completamente diferentes a los 10 pacientes utilizados en este trabajo, es decir, no se puede considerar como una comparación de resultados del todo válida. Simplemente se ha realizado esta comparativa para tener una idea general del rendimiento del algoritmo propuesto.

Paciente	What-If				Agnóstico			
	t+30	t+60	t+90	t+120	t+30	t+60	t+90	t+120
1	22,00	34,35	40,27	43,67	22,33	34,98	44,67	50,93
2	18,28	30,35	38,81	44,66	18,52	30,48	39,06	45,52
3	19,41	27,89	33,36	36,13	19,30	28,61	34,28	37,40
4	19,55	31,92	40,53	45,22	19,72	32,10	40,64	46,00
5	29,58	43,96	51,75	55,34	29,83	44,66	52,18	55,86
6	25,38	39,44	46,31	49,05	25,36	39,96	47,62	52,44
7	25,75	39,58	48,22	51,59	25,73	39,56	48,09	53,76
8	28,22	36,11	37,93	38,69	30,38	41,60	44,46	46,33
9	23,78	37,31	45,14	48,22	23,89	38,59	46,37	53,18
10	26,31	45,60	60,64	67,68	26,04	45,24	62,42	76,58
Media	23,83	36,65	44,30	48,02	24,11	37,58	45,98	51,80
Desv.	3,90	5,72	7,91	8,95	4,18	5,80	7,70	10,26

Tabla 4.3: Mejor modelo en la fase de validación

Paciente	What-If				Agnóstico			
	t+30	t+60	t+90	t+120	t+30	t+60	t+90	t+120
1	16,18	26,37	30,87	33,92	16,14	27,91	34,39	38,36
2	15,91	26,10	32,78	36,03	15,88	26,52	33,52	37,60
3	16,57	24,71	28,13	29,48	16,63	25,03	28,57	31,11
4	17,12	30,02	36,36	41,15	17,05	30,21	37,85	42,76
5	22,48	36,13	44,38	47,72	22,07	35,94	44,17	46,23
6	21,67	33,34	36,90	39,53	20,39	33,43	40,24	43,99
7	22,30	36,52	40,15	42,70	23,44	37,04	43,07	48,14
8	17,82	25,51	29,38	30,96	18,22	29,03	34,56	35,94
9	18,63	31,97	37,47	37,92	19,14	31,46	39,12	40,13
10	20,07	41,16	52,96	58,43	21,52	40,81	54,32	62,33
Media	18,88	31,18	36,94	39,78	19,05	31,74	38,98	42,66
Desv.	2,57	5,60	7,54	8,55	2,70	5,03	7,15	8,57

Tabla 4.4: Mejor modelo en la fase de test (con fines comparativos).

4.4.5. Análisis de las soluciones obtenidas

Como se ha explicado en secciones anteriores, los modelos generados con el algoritmo METSK-HD están formados por una serie de reglas difusas del tipo TSK. En la primera fase del método se generan las reglas iniciales (máximo 100) y en la segunda fase se realiza una selección y ajuste de las reglas, generándose en el mejor de los casos modelos con muy pocas reglas difusas. Por ejemplo, se muestra un modelo personalizado para un paciente compuesto únicamente por 5 reglas difusas:

CAPÍTULO 4. PREDICCIÓN DE GLUCOSA MEDIANTE UN SISTEMA BASADO EN REGLAS DIFUSAS TSK DE DOS FASES: METSK-HD

	Trabajo	30 minutos		60 minutos		Total
		RMSE	MAE	RMSE	MAE	
	Harry Rubin-Falcone [86]	18.22	12.83	31.66	23.60	86.31
	Hadia Hameed [87]	19.21	13.08	31.77	23.09	87.15
	Taiyu Zhu [88]	18.34	13.37	32.21	24.20	88.12
	Tao Yang [89]	19.05	13.50	32.03	23.83	88.41
	Robert Bevan [90]	18.23	14.37	31.10	25.75	89.45
	Xiaoyu Sun [91]	19.37	13.76	32.59	24.64	90.36
	David Joedicke [92]	19.60	14.25	34.12	25.99	93.96
	Ning Ma [93]	20.03	14.52	34.89	26.41	95.85
Agnóstico	Mejor modelo en fase de entrenamiento	29.00	21.01	42.51	32.02	124.54
	Modelo Ensemble	23.48	17.26	36.60	27.89	105.25
	Mejor modelo en fase de validación	24.11	17.54	37.58	28.47	107.70
	Mejor modelo en fase de test	19.05	14.40	31.74	24.81	90.00
What-If	Mejor modelo en fase de entrenamiento	30.03	21.22	42.16	32.13	125.54
	Modelo Ensemble	23.32	17.20	35.84	27.43	103.79
	Mejor modelo en fase de validación	23.83	17.39	36.65	27.87	105.74
	Mejor modelo en fase de test	18.88	14.31	31.18	24.64	89.01

Tabla 4.5: Clasificación general del BGLP Challenge.

IF

Gluct IS -431.119345792438 140.042060447636 343.758053918418 AND

Cht3015 IS -1.392142156644 -0.006904784107 4.951113815262 AND

Inst150 IS -4.854728905816 0.113456493810 6.594804019427

THEN

Gluct30 = 0.803506721623*Gluct + 0.000000000000*Cht3015 + 13.255802077073*Inst150
+ 17.677216029411

IF

Gluct IS -130.774544844824 140.988659415715 178.103242645382 AND

Cht3015 IS -0.253826538967 1.322834619254 7.013473563511 AND

Inst150 IS -4.563665959429 -0.477271427584 4.833029769003

THEN

Gluct30 = 0.540590886283*Gluct + 7.039652193178*Cht3015 + -0.392312957770*Inst150
+ 65.953614434726

IF

Gluct IS 35.459944600393 172.807438236155 337.168689515264 AND

Cht3015 IS -2.165187988886 0.004000347064 4.561740858872 AND

Inst150 IS -6.047064328093 0.169485154530 8.753947182327

THEN

Gluct30 = 0.837415924038*Gluct + 0.000000000000*Cht3015 + 8.217824598981*Inst150
+ 24.648893338722

IF

```

Gluct IS 18.512955154659 209.862803608202 436.652313437837 AND
Cht3015 IS 0.757699077219 4.302158597560 6.583163618792 AND
Inst150 IS -3.141005737658 -0.982369580180 3.258936835540
THEN
Gluct30 = 0.000000000000*Gluct + 0.000000000000*Cht3015 + 0.000000000000*Inst150
+ 222.000000000000

IF
Gluct IS 110.979103581559 276.406060756646 696.931375034793 AND
Cht3015 IS -4.039929827056 0.550634712441 3.856565494654 AND
Inst150 IS -5.738937029478 0.542122677877 8.342454714391
THEN
Gluct30 = 0.911644027046*Gluct + 41.828316372623*Cht3015 + -3.753239982435*Inst150
+ 23.858765057309

Regla por defecto: 0.911644027046*Gluct + 41.828316372623*Cht3015
+ -3.753239982435*Inst150 + 23.858765057309

MIN and MAX output limits: -30.400000000000 471.733333333333
RMSE_tra: 18.5028510745484. RMSE_tst: 16.7313362899570
Used variables: Gluct Cht3015 Inst150 Gluct30

```

dónde *Gluc* (Glucosa actual), *Cht3015* (Raciones de hidratos de carbono del intervalo [30, 15) minutos anteriores al momento de la predicción), *Inst150* (Dosis de insulina del intervalo [15, 0) anteriores al momento de la predicción) son las variables de entrada al sistema y *Gluc30* es la variable de salida, siendo un modelo con horizonte a 30 minutos.

Se puede observar claramente la estructura del modelo, formado por las reglas con su antecedente (IF) y consecuente (THEN). Además, existe una regla por defecto en caso de no activarse ninguna. En el modelo anterior, únicamente se han usado 3 variables de entrada, siendo éstas la glucosa actual, la ingesta total de carbohidratos del intervalo [30, 15) minutos anteriores y la dosis de insulinas del intervalo [15, 0) minutos anteriores al momento de la predicción. Nos encontramos en este ejemplo con un modelo *agnóstico* al no usarse información futura en ninguna de sus variables. El modelo además ofrece un buen rendimiento, obteniendo un RMSE de 16,73 en la fase de prueba.

En este otro ejemplo, se muestra hasta dónde puede llegar la reducción en el número de reglas durante la segunda fase del algoritmo METSK-HD. Nos encontramos así con un modelo personalizado para un paciente para predecir la BG con horizonte a 120 minutos formado únicamente por 2 reglas difusas:

CAPÍTULO 4. PREDICCIÓN DE GLUCOSA MEDIANTE UN SISTEMA BASADO EN REGLAS DIFUSAS TSK DE DOS FASES: METSK-HD

IF

Gluct IS -881.862605488103 215.910459227967 353.572920253486 AND

Cht4530 IS -5.997891112123 0.041904463183 1.692501313855

THEN

Gluct120 = 0.470538903175*Gluct + 8.889000754388*Cht4530 + 79.457909302792

IF

Gluct IS -374.033071649091 186.360408664890 452.078043723754 AND

Cht4530 IS -0.115536680569 3.776258404072 4.941114915088

THEN

Gluct120 = 0.598125498837*Gluct + -31.090610443804*Cht4530 + 75.306831169460

Regla por defecto: 0.455686813660*Gluct + -31.297761299644*Cht4530

+ 79.413519918176

MIN and MAX output limits: -19.866666666667 427.200000000000

RMSE_tra: 48.7125733129452. RMSE_tst: 52.8963971108857

Used variables: Gluct Cht4530 Gluct120

dónde *Gluc* (Glucosa actual), *Cht4530* (Raciones de carbohidratos en el intervalo [45, 30) minutos anteriores al momento de la predicción) son las variables de entrada al sistema y *Gluc120* es la variable de salida, siendo un modelo con horizonte a 120 minutos.

Este modelo únicamente toma dos variables de entrada, siendo la glucosa actual y las ingestas de raciones de hidratos en el intervalo [45, 30) minutos anteriores al momento de la predicción, tratándose así de un modelo *agnóstico*. Además, este modelo ofrece un rendimiento RMSE de 52,90 en fase de test. Hay que aclarar en este punto que el número de reglas es independientemente si es la mejor o peor solución, es decir, que este modelo este formado únicamente por dos reglas no significa que sea el mejor modelo.

Vistos los ejemplos anteriores, el funcionamiento de los modelos generados con este algoritmo son interpretables por parte de una persona. Se trata así, de una técnica de caja blanca, es decir, ofrece una visión completa y detallada del funcionamiento de la lógica del modelo. Hay que recalcar aquí, que estos modelos son interpretables si están formados por unas 15 reglas o menos.

4.5. Conclusiones

Los sistemas FRBS son usados habitualmente para sistemas de control automático, aunque por sus particularidades se adaptan perfectamente al problema de la predicción del valor exacto de la glucosa. Esto es debido a la idoneidad de estos sistemas para encontrar la relación matemática entre las variables de entrada y el valor de glucosa a predecir.

Una de las principales ventajas de esta técnica, es la interpretabilidad de los modelos. Gracias a la fase de selección y ajuste, se obtienen modelos formados por muy pocas reglas donde se plasma el conocimiento del modelo. En el mejor de los casos, se han conseguido modelos compuestos únicamente por dos reglas difusas. Esto facilita la comprensión de los modelos, es decir, a menor número de reglas, más interpretables son los modelos por parte de un humano. Además, la simplicidad de estos modelos posibilita su implementación en dispositivos móviles y pulseras de actividad.

En cuanto a los resultados obtenidos con los modelos METSK-HD, aunque son comparables con respecto a otras técnicas de ML, la aplicación del algoritmo propuesto al problema de la predicción del valor exacto de BG no ofrece un rendimiento del todo aceptable. Visto los resultados obtenidos con esta técnica, decidimos cambiar de técnica en este trabajo de tesis, centrándonos únicamente en predecir eventos de hipoglucemias e hiperglucemias mediante la aplicación de transformaciones del espectro al campo de visión y modelos de aprendizaje profundo. Este enfoque se explicará en el siguiente capítulo de este trabajo de tesis.

Capítulo 5

Predicción de Hipoglucemia e Hiperglucemias

En el segundo enfoque de esta tesis, nos hemos centrado en un objetivo más ambicioso dentro de la predicción de la glucosa, siendo la predicción de eventos de hipoglucemias e hiperglucemias. Para la detección de estos episodios, vamos a hacer uso de técnicas basadas en redes neuronales artificiales (ANN). Las ANN han supuesto un gran avance en materia de inteligencia artificial en estos últimos años y en concreto, dentro del denominado aprendizaje profundo o Deep Learning. Las ANN son modelos inspirados en el funcionamiento de las neuronas de un cerebro biológico. Estos modelos están formados por las denominadas neuronas artificiales interconectadas entre sí. Estas neuronas reciben una serie de entradas, realizan un cálculo dependiendo de la función utilizada y los pesos asociados a las conexiones entre neuronas, y en base a lo anterior producen una salida que es transmitida a otras neuronas. Como se puede observar en la Figura 5.1, las neuronas están organizadas en capas. Habitualmente, todos los modelos de ANN al menos cuentan con una capa de entrada, una o más capas ocultas y una capa de salida.

Los modelos de ANN se utilizan en una amplia variedad de problemas de aprendizaje automático y procesamiento de datos, como reconocimiento de imágenes, predicción de valores, clasificación, procesamiento de lenguaje natural, entre otros. Existen numerosos tipos de modelos de ANN, pero en este trabajo de tesis nos hemos centrado en las denominadas redes neuronales convolucionales (CNN).

5.1. Redes Neuronales Convolucionales

Las redes neuronales convolucionales (CNN) son un tipo de redes neuronales cuya disposición de las capas se asemeja a la corteza visual primaria de un cerebro biológico y están englobadas dentro de la denominada visión por computador. Estas redes son ampliamente utilizadas para el

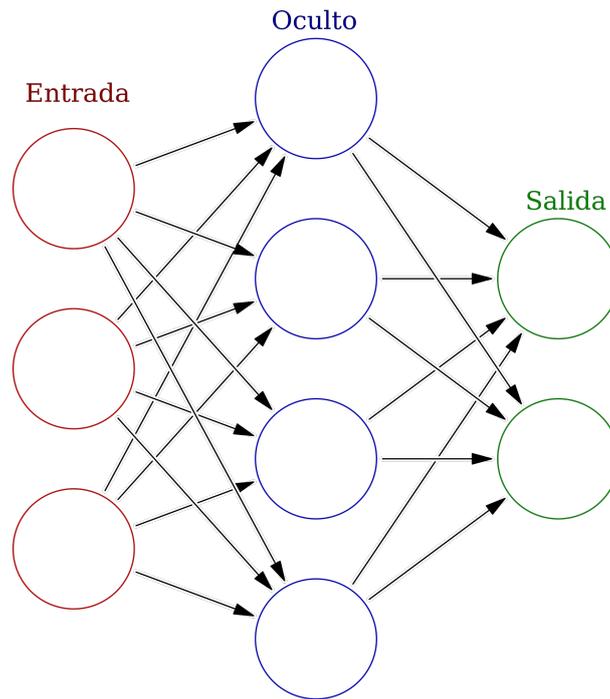


Figura 5.1: Esquema básico de una ANN

reconocimiento de imágenes, la detección de objetos, la clasificación, la detección de anomalías y la predicción basada en series temporales, entre otros.

Estas redes están formadas por varios tipos de capas diferentes, siendo las principales:

- **Capas de convolución:** estas capas son las encargadas de extraer los mapas de características de una imagen. Para ello, llevan a cabo operaciones de convolución que consisten en la aplicación de un filtro (kernel) deslizante de tamaño determinado, que recorre la imagen, extrayendo así las características relevantes de la imagen. Se muestra ejemplo de operación de convolución en la Figura 5.2.
- **Capas de pooling:** utilizadas habitualmente tras las capas de convolución, su principal función es reducir la dimensionalidad de los mapas de características con el objetivo de ganar robustez ante pequeñas variaciones en la posición de las características. Las operaciones más utilizadas son las siguientes:
 - **Max Pooling:** operación que divide el mapa de características en varias secciones de tamaño determinado y obtiene el valor máximo de cada una de ellas
 - **Average Pooling:** operación que divide el mapa de características en varias secciones de tamaño determinado y obtiene el valor promedio de cada una de ellas

Un ejemplo de las operaciones de Max Pooling y Average Pooling se muestra en la Figura 5.3.

- **Capas completamente conectadas:** también denominadas capas densas, su función es transformar las características extraídas por las capas convolucionales en una forma adecuada para el problema en cuestión, normalmente clasificación y regresión.

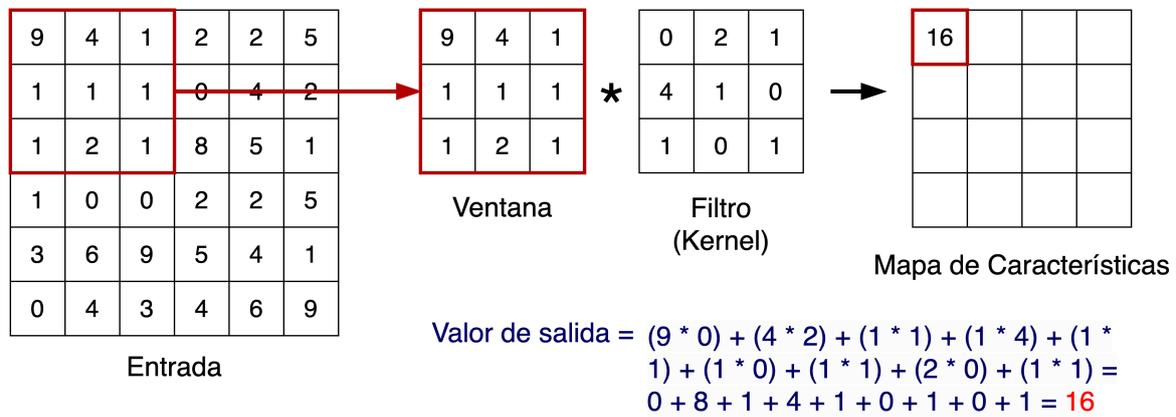


Figura 5.2: Ejemplo de operación de convolución de tamaño 3×3 sobre una imagen de entrada de tamaño 6×6

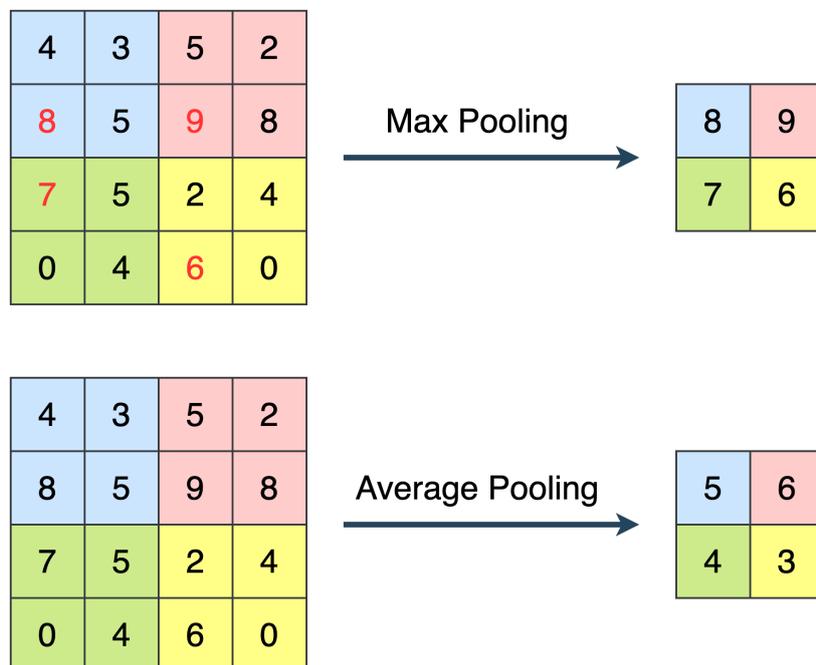


Figura 5.3: Ejemplo de operaciones Max Pooling y Average Pooling de tamaño 2×2 sobre un mapa de características de 4×4

Además de las capas principales, existen otra serie de capas que pueden ser utilizadas como son las capas de *normalización* y *dropout*. Las capas de *normalización* son las encargadas de

realizar operaciones de normalización sobre los mapas de activaciones de una capa anterior, siendo la más habitual la denominada Batch normalization. Por otro lado, las capas de *dropout* se utilizan para evitar el sobreentrenamiento de la red, es decir, un sobreajuste del modelo que puede causar una pérdida en la precisión del mismo. Para ello, se desconectan aleatoriamente un porcentaje de neuronas de la capa anterior. En este punto cabe mencionar un tercer tipo de capa utilizado en algunas CNN, las capas *flatten*. Ésta capa va colocada generalmente entre las capas convolucionales y las capas completamente conectadas. Su principal función es la de transformar los mapas de características multidimensionales provenientes de las capas convolucionales a un vector unidimensional que servirá como entrada para las capas completamente conectadas.

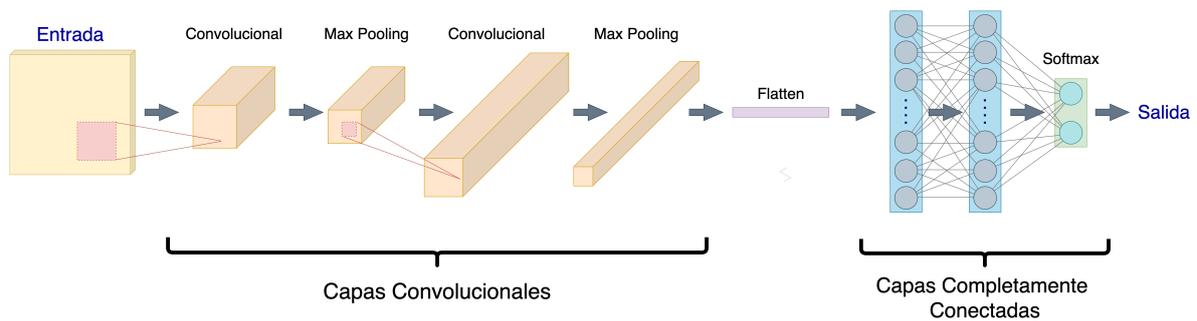


Figura 5.4: Esquema básico de ejemplo de una CNN

En resumen, las CNN se suelen organizar en dos grandes bloques tal como se puede observar en la Figura 5.4. En primer lugar, tenemos las capas del tipo convolucional (convolucionales, pooling, normalización, etc) y por otro lado tenemos las capas completamente conectadas, que habitualmente se ubican al final de la red. Como hemos mencionado anteriormente, la capa de salida se adaptará según el problema a abordar, sea de clasificación como de regresión. Habitualmente, en esta última capa se utiliza la función denominada *softmax*. Se trata de una función matemática que toma un vector de valores reales como entrada y produce un número real como salida, con valores comprendidos entre 0 y 1. Esta función se utiliza ampliamente en tareas de clasificación y se emplea frecuentemente como capa final en modelos de redes neuronales para transformar puntuaciones brutas en probabilidades asociadas a clases. La función softmax se define matemáticamente como (5.1):

$$\sigma(\vec{z})_i = \frac{e^{z_i}}{\sum_{j=1}^N e^{z_j}} \quad (5.1)$$

Existen multitud de arquitecturas CNN diferentes, pero en este trabajo de tesis nos hemos centrado en dos arquitecturas en concreto, siendo las siguientes:

- AlexNet

- **DenseNet**

En la Sección 5.1.2 se detalla la arquitectura AlexNet y en la Sección 5.1.3 detallamos la arquitectura DenseNet.

5.1.1. Transfer Learning

Para poder utilizar cualquier modelo de aprendizaje profundo, es necesario realizar un proceso de entrenamiento previo, el cual varía dependiendo del problema a abordar. Este entrenamiento generalmente implica el uso de conjuntos de datos diseñados específicamente para el problema en cuestión. En términos más precisos, el proceso de aprendizaje de un modelo de aprendizaje profundo implica ajustar de manera óptima los pesos de las neuronas para que la salida coincida con la esperada en función de los datos de entrada proporcionados.

Habitualmente, cuantos más datos se utilicen para el entrenamiento de un modelo, más robusto será el mismo. Pero esto tiene una serie de inconvenientes debido a que no siempre se disponen de los datos suficientes o los recursos computacionales necesarios. Por esta razón, en las Redes Neuronales Artificiales (ANN) se emplea comúnmente la técnica conocida como Transfer Learning. Esta estrategia implica transferir el conocimiento obtenido durante el entrenamiento de un modelo en un problema específico a otro problema relacionado. De este forma, no es necesario entrenar de un modelo desde cero para cada nuevo problema.

En concreto, en las CNN, la técnica de Transfer Learning consiste en realizar un ajuste fino de los pesos de las últimas capas (habitualmente las completamente conectadas) dejando el resto de las capas congeladas. De esta forma, se mantiene el conocimiento previamente extraído en las capas convolucionales a la vez que se adaptan las capas encargadas de realizar la clasificación o regresión al nuevo problema a abordar. Con esta técnica se ahorra tiempo de computación al no ser necesario entrenar el modelo desde cero y permite adaptar modelos existentes para resolver nuevos problemas. Además, se pueden utilizar modelos bastantes robustos que han sido entrenados con una gran cantidad de datos y recursos de los que no siempre se disponen para adaptarlos a las necesidades del problema a resolver.

5.1.2. Arquitectura AlexNet

En primer lugar, se ha utilizado la arquitectura AlexNet [95], desarrollada en 2012 por Alex Krizhevsky et al. Para la implementación de los modelos, se ha empleado el módulo de python *Keras Applications*¹. Como el repositorio de modelos de *Keras Applications* no dispone de modelos AlexNet pre-entrenados, los modelos han sido entrenados desde cero sin recurrir a Transfer Learning, es decir, la totalidad de los parámetros ha sido entrenada de cero. Como se puede ver en la Figura 5.5, la arquitectura de AlexNet incorpora cinco capas convolucionales, tres capas de pooling y dos capas totalmente conectadas. La entrada de esta red son imágenes RGB de

¹<https://keras.io/api/applications/>

227×227 píxeles, a las que se aplica una primera convolución de 11×11 , seguida de un Max Pooling de 3×3 . A continuación, se realiza la segunda convolución de tamaño 5×5 y un Max Pooling de 3×3 . El mapa de características resultante se somete a tres convoluciones sucesivas de 3×3 y a una operación de Max Pooling de 3×3 . Por último, dos capas totalmente conectadas de 4096 neuronas y la capa de clasificación. Esta última capa utiliza la función softmax y se ha adaptado a nuestro problema estableciendo dos salidas, una para cada categoría a predecir.

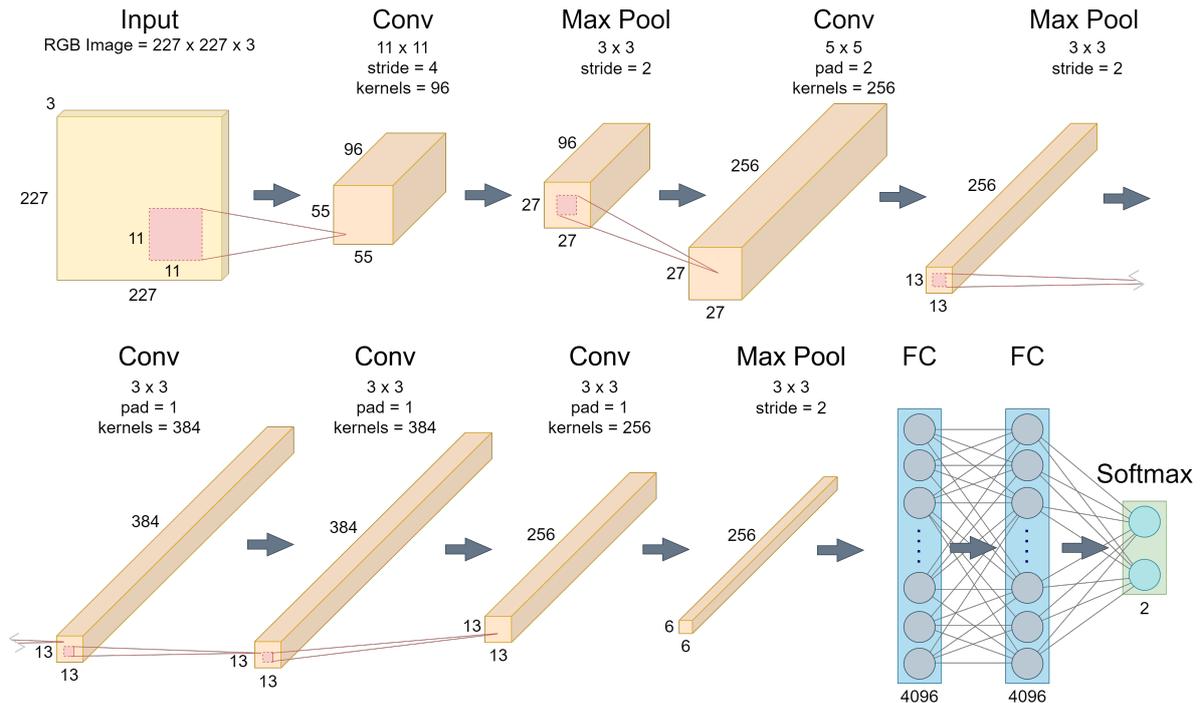


Figura 5.5: Arquitectura AlexNet

En la Tabla 5.1 se muestra en detalle el número de parámetros de cada una de las capas del modelo AlexNet utilizado. Se puede observar que el modelo cuenta con más de 52 millones de parámetros, siendo en casi su totalidad entrenables al no utilizarse Transfer Learning.

5.1.3. Arquitectura DenseNet

También hacemos uso de la arquitectura DenseNet [96], desarrollada en 2017 por Gao Huang et al. Al igual que los modelos AlexNet, se ha utilizado el módulo de python *Keras Applications* para la implementación de los modelos. En este caso, los modelos han sido entrenados mediante Transfer Learning utilizando una red pre-entrenada con el conjunto de datos ImageNet [97] disponible en el repositorio de modelos de *Keras Applications*. Hemos realizado un ajuste fino de las capas completamente conectadas y la última capa, dejando el resto de las capas congeladas para mejorar la precisión de los modelos. La Figura 5.6 muestra la configuración de DenseNet utilizada en este trabajo, conocida como DenseNet-121. La red tiene como entrada imágenes RGB de tamaño 224×224 píxeles y luego se aplica una convolución 7×7 . DenseNet se organiza

Identificador capa	Parámetros
Input	0
Conv 1	34.944
Batch normalization 1	384
Max Pooling	0
Conv 2	614.656
Batch normalization 2	1.024
Max Pooling	0
Conv 3	885.120
Batch normalization 3	1.536
Conv 4	1.327.488
Batch normalization 4	1.536
Conv 5	884.992
Batch normalization 5	1.024
Max Pooling	0
Flatten	0
Dense 1	31.461.376
Dropout 1	0
Dense 2	16.781.312
Dropout 2	0
Output (Softmax)	8194
Parámetros totales	52.003.586
Parámetros entrenables	52.000.834
Parámetros no entrenables	2.752

Tabla 5.1: Parámetros de las diferentes capas del modelo AlexNet

en bloques densos, y entre estos bloques se colocan capas de transición. Un bloque denso combina varias capas (entre 6 y 24) formadas por la función compuesta de Batch normalization (ReLU) y una operación de convolución de 3×3 o 1×1 . Estas capas están interconectadas de modo que los mapas de características producidos se concatenan al final del bloque, añadiendo nueva información que puede reutilizarse. Las capas de transición se encargan de submuestrear los mapas de características entre los bloques densos mediante la operación Batch normalization tras una convolución de 1×1 y un Max Pooling de 2×2 . Al final del último bloque denso, se aplica un Average Pooling de 7×7 , y la salida se conecta a una capa densa de tamaño 512. Por último, la capa de clasificación se ha adaptado a nuestro problema utilizando una capa de tamaño 2 con una función softmax.

En la Tabla 5.2 se puede observar en detalle el número de parámetros de cada una de las capas del modelo DenseNet-121 utilizado. El modelo cuenta con más de 8 millones de parámetros, siendo un número menor en comparación con el modelo Alexnet. Además, al usarse Transfer Learning, únicamente son entrenables los parámetros de las capas completamente conectadas y la salida.

5.1. REDES NEURONALES CONVOLUCIONALES

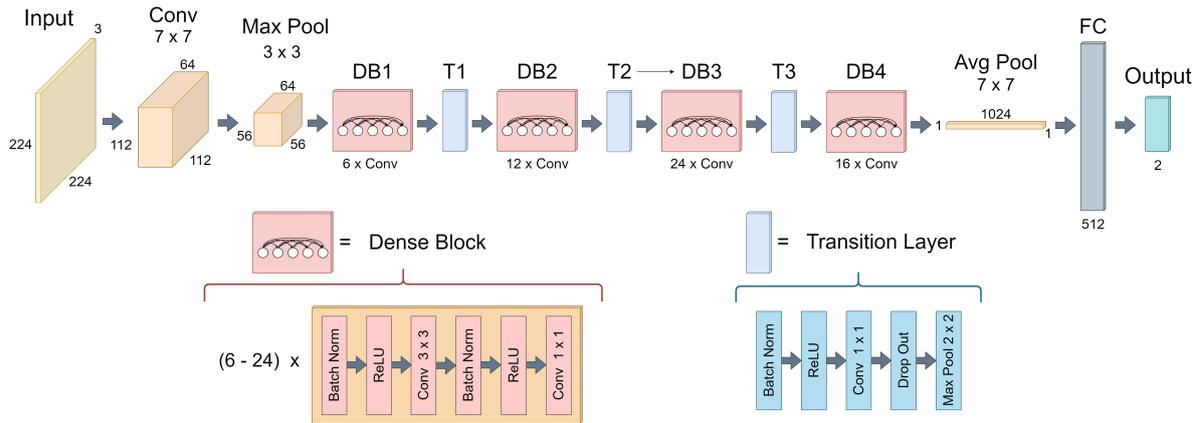


Figura 5.6: Arquitectura DenseNet-121

Identificador Capa	Parámetros
Input	0
Conv 1	9.408
Batch normalization 1	256
Max Pooling	0
DB1	338.304
T1	33.792
DB2	930.048
T2	133.120
DB3	2.873.856
T3	528.384
DB4	2.186.240
Batch normalization 2	4.096
Average Pooling	0
Batch normalization 3	4.096
Dropout 1	0
Dense 1	1.049.600
Dense 2	524.800
Batch normalization 4	2.048
Dropout 2	0
Output (Softmax)	1.026
Parámetros totales	8.619.074
Parámetros entrenables	1.578.498
Parámetros no entrenables	7.040.576

Tabla 5.2: Parámetros de las diferentes capas del modelo Densenet-121

5.2. Transformada de Wavelet

La evolución de los niveles de glucosa en sangre se ve afectada por numerosos biorritmos cuyos ciclos abarcan diferentes escalas (semanales, diarios, relacionados con la ingesta, etc.) pero no son estacionarios. En otras palabras, no esperamos que la presencia de estos ciclos sea constante, sino que se limite a segmentos temporales específicos. Necesitamos utilizar una transformación matemática que pueda caracterizar las series temporales de glucosa con la información de las frecuencias actuales, sus magnitudes y su variación en el tiempo. La herramienta más adecuada para esta tarea es la transformada wavelet.

Mientras que la Transformada de Fourier de Tiempo Reducido (STFT) [98] nos permite descomponer una señal en múltiples señales dentro de rangos de frecuencia de igual anchura, la transformada de wavelet va más allá y proporciona una descomposición basada en señales de igual ancho de banda en una escala logarítmica, es decir, tenemos una resolución cuyo grano depende de la frecuencia. Con la transformada wavelet, tenemos altas resoluciones temporales y de baja frecuencia para eventos de alta frecuencia (esperamos pocos eventos de este tipo en la dinámica de la glucosa). Ofrece una buena resolución de frecuencia, pero una resolución temporal pobre para los eventos de baja frecuencia (más comunes en las series temporales de glucosa). Esto último convierte a la wavelet en la transformada más adecuada para caracterizar la evolución de los niveles de glucosa en el plano de la frecuencia temporal.

La wavelet es una función con un pequeño conjunto de oscilaciones no estacionarias. Si estas oscilaciones fueran estacionarias, podríamos decir que tienen una frecuencia determinada. Si cambiamos el tamaño, es decir, la escala de la wavelet, cambiamos también la frecuencia estacionaria hipotética correspondiente de las oscilaciones. Así, cuando realizamos una convolución de la wavelet con una escala determinada a través de nuestra serie temporal, buscamos segmentos de la serie temporal que muestren un comportamiento frecuencial igual a la frecuencia hipotética de la wavelet. Cambiando el tamaño (es decir, la escala) de la wavelet y repitiendo el proceso de convolución (es decir, desplazando la wavelet a través de nuestra serie temporal y calculando la convolución de las dos funciones) obtendremos los coeficientes de la transformada wavelet para distintas frecuencias. El conjunto de estos coeficientes puede representarse en forma de “escalograma” o de espectro de potencia. Utilizaremos esta representación visual de la transformada wavelet para caracterizar la serie temporal de los niveles de glucosa en sangre.

Así con los diferentes valores de “escala” y “desplazamiento”, tendremos un conjunto de funciones generadoras de la transformada de wavelet con la forma (5.2):

$$\Psi_{scale,shift}(t) = \frac{1}{\sqrt{scale}} \Psi\left(\frac{t - shift}{scale}\right) \quad (5.2)$$

Además, la forma de las oscilaciones da lugar a distintas familias de wavelets: Mexican Hat, Morlet, Shannon o Spline, entre otras. Siguiendo los resultados publicados en [99], las funciones wavelet utilizadas en este trabajo de tesis son las siguientes:

- **Mexican Hat wavelet** (también llamada wavelet Ricker). Esta función se representa en la Figura 5.7 y se define matemáticamente como (5.3):

$$\Psi(t) = \frac{2}{\sqrt{3}\sqrt[4]{\pi}} e^{-\frac{r^2}{2}} (1 - r^2) \quad (5.3)$$

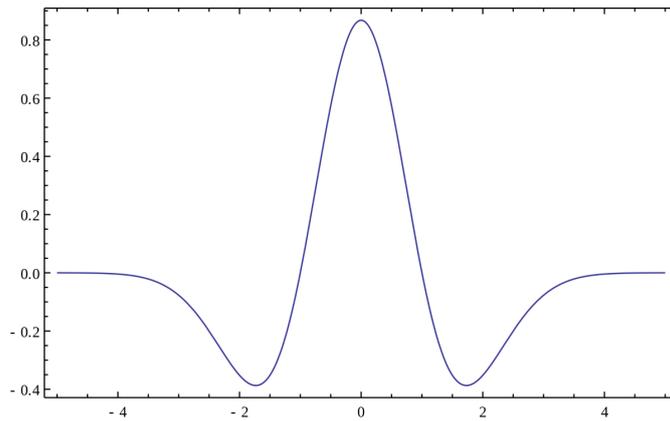


Figura 5.7: Mexican Hat Wavelet

- **Morlet Wavelet**. Esta función se representa en la Figura 5.8 y se define matemáticamente como (5.4):

$$\Psi(t) = e^{-\frac{r^2}{2}} \cos(5t) \quad (5.4)$$

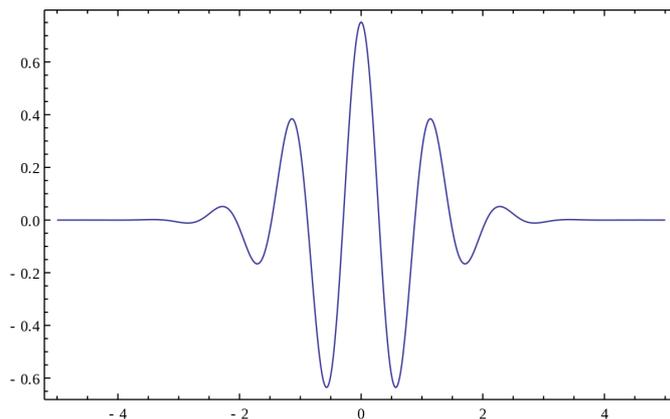


Figura 5.8: Morlet Wavelet

Se han elegido las wavelets Mexican Hat y Morlet porque son las más utilizadas de la familia de las transformadas continuas de wavelet (CWT, del inglés Continuous Wavelet Transform).

Las ventajas de las CWT son que permiten realizar análisis de tiempo-frecuencia, lo que es importante en el campo del procesamiento de señales biomédicas, como las series temporales de glucosa. Estas dos familias de wavelets se complementan, y su uso conjunto mejora tanto el análisis temporal como el análisis frecuencial [99].

5.3. Técnicas de Aumentado de Datos

5.3.1. Aumentado de Datos mediante Ventana Deslizante

En las aplicaciones modernas de aprendizaje profundo, el rendimiento de las redes neuronales está profundamente influenciado por el tamaño y la diversidad del conjunto de datos sobre el que se entrenan. Sin embargo, en ámbitos específicos como la endocrinología, la adquisición de una gran cantidad de datos reales de pacientes con fines de entrenamiento puede ser extremadamente difícil, debido principalmente a problemas de privacidad y a la limitada disponibilidad de bases de datos completas. En el contexto de esta situación, se ha utilizado el aumentado de datos (DA, de sus siglas en inglés) [100], que ofrece una solución ingeniosa para mejorar la utilidad de conjuntos de datos limitados, como es el caso de los conjuntos de datos de glucosa en sangre [101]. DA es una técnica fundamental de aprendizaje automático y visión por computador ampliamente adoptada para ampliar el tamaño y la diversidad de los conjuntos de datos de entrenamiento de forma artificial. Consiste en crear nuevos puntos de datos modificados a partir del conjunto de datos existente aplicando diversas transformaciones, como rotación, escalado, recorte o, en nuestro caso, una técnica de ventana deslizante. Estos puntos de datos transformados diversifican el conjunto de datos, reduciendo el riesgo de sobreajuste y mejorando la capacidad del modelo para generalizar a datos desconocidos.

En el contexto de este trabajo de tesis, el DA es esencial debido a la escasez de eventos, tanto de hipoglucemias como de hiperglucemias, en el conjunto de datos. Para superar esta limitación, se ha ideado una técnica innovadora de DA que utiliza un enfoque de ventana deslizante. Al emplear esta técnica, se generan de forma efectiva, puntos de datos adicionales. Cada uno de estos datos generados, representa una porción diferente de los datos de la serie temporal original, simulando así varios episodios de hipoglucemia e hiperglucemias. Gracias a esta aumento de datos, se amplía el conjunto de datos e introduce diversidad en los patrones de los episodios de hipoglucemias e hiperglucemias, lo que hace que el modelo sea más robusto y capaz de manejar una gama más amplia de escenarios.

El problema es el siguiente:

Si se utiliza un día entero para la predicción de cada paciente, sólo se genera un espectrograma para cada día de datos. Además, la detección de eventos de hipoglucemias e hiperglucemias debería realizarse una vez al día. Además, no todos los días presentan eventos de hipoglucemias e hiperglucemias, por lo que se producirían muy pocas imágenes y, en consecuencia, los modelos serían incapaces de detectarlas correctamente.

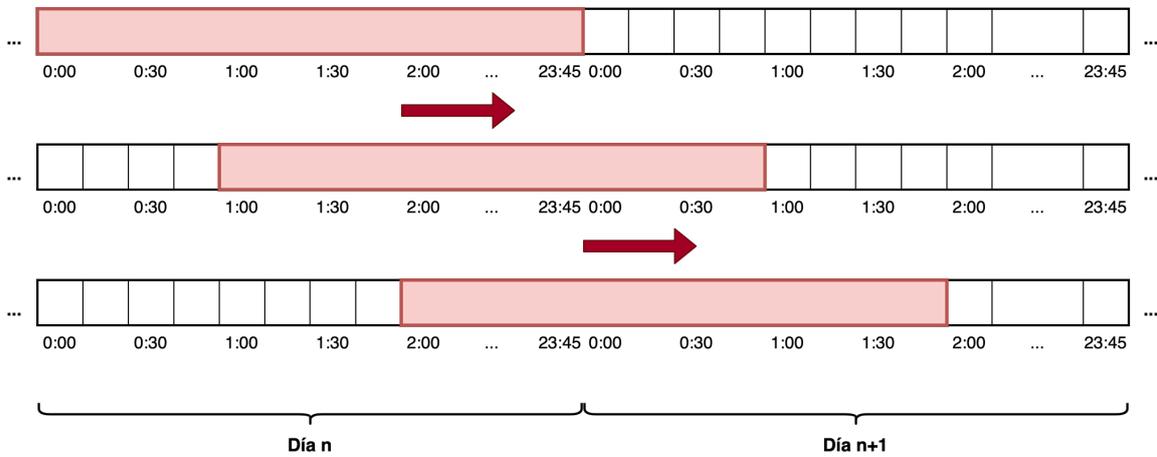


Figura 5.9: Aumentado de datos mediante una ventana deslizante con tamaño una hora

En la Figura 5.9 se muestra el funcionamiento de la técnica de DA. Como se puede observar, se ha aplicado una ventana deslizante de tamaño 96 registros (un día de datos) a la serie temporal de datos de glucosa. Esta ventana cuenta con un desfase de 1 hora, es decir, es desplazada una hora cada vez que se obtienen los nuevos valores dentro de la ventana. De esta forma se pueden generar hasta 24 series temporales de glucosa por paciente y día, permitiendo aumentar en 24 veces el conjunto de datos original.

5.3.2. Aumentado de Datos mediante Ruido Gaussiano

Además de la técnica de DA detallada en el apartado anterior, se ha aplicado un aumento de datos introduciendo ruido a los datos, de forma que se genera una nueva serie temporal de glucosa en base a la original. Este aumento de datos pretende simular el ruido que metería un sensor en su normal funcionamiento. En concreto, se ha aplicado ruido gaussiano a cada uno de los valores de las series temporales de glucosa pertenecientes al conjunto de entrenamiento, alterando su valor aleatoriamente siguiendo distribución de gauss.

Para ello, se utiliza la distribución de gauss para generar una serie temporal de valores de ruido en base a la media de los valores de glucosa pertenecientes a serie temporal original y un porcentaje de ruido previamente establecido. Esta serie de valores de ruidos se suman a los valores de la serie temporal original para obtener la nueva serie temporal de glucosa. La fórmula matemática empleada para calcular los valores de ruido es la siguiente (5.5):

$$p(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad (5.5)$$

dónde μ es la media y σ es la desviación estándar. Para aplicar esta técnica, se ha establecido el valor $\mu = 0$ y $\sigma = \text{porcentaje_ruido} \cdot \text{media_serie_temporal}$. Por ejemplo, para aplicar

un 2% de ruido y tenemos una media de valores de glucosa de 134 mg/dL , daría un valor $\sigma = 0,02 \cdot 134 = 2,68$.

Como los nuevos valores generados son una alteración de los originales, se puede dar el caso que éstos se salgan del rango realista, es decir, se den valores por debajo de 40 mg/dL o por encima de 550 mg/dL . Para solucionar este problema, una vez generado los nuevos valores de ruido, se aplican las siguientes reglas:

- Si nuevo valor BG $< 40 \text{ mg/dL}$: nuevo valor BG = 40 mg/dL
- Si nuevo valor BG $> 550 \text{ mg/dL}$: nuevo valor BG = 550 mg/dL

Con esto evitamos que se generen valores fuera del intervalo $[40, 550]$, siendo estos valores irreales en la vida real. Por debajo ($< 40 \text{ mg/dL}$) se trataría de una hipoglucemia muy severa y por encima ($> 550 \text{ mg/dL}$) se daría el caso de un coma diabético. En resumen, estos límites son críticos, es decir, los valores que no se encuentren dentro de este intervalo son incompatibles con la vida.

5.4. Método Propuesto

El método propuesto para el enfoque de la predicción de hipoglucemias e hiperglucemias lo hemos denominado *predicción de valores críticos de glucosa mediante transformaciones del espectro al campo de visión y aprendizaje profundo*, y se puede observar detalladamente en la Figura 5.10. El proceso se inicia a partir de series temporales formadas por registros de glucosa a intervalos de 15 minutos, 96 registros en total por día:

1. Se aplica un proceso de Aumentado de Datos a los datos (véase la Sección 5.3).
2. Los datos de cada paciente se dividen en series temporales que contienen los registros de los valores de glucosa durante 24 horas.
3. Se establece un umbral para el nivel de hipoglucemia T_{hipo} y hiperglucemia T_{hiper} . El valor predeterminado es 70 mg/dL para las hipoglucemias y 180 mg/dL para las hiperglucemias.
4. Las series temporales con los registros de 24 horas de glucosa se clasifican en dos grupos de categorías, (ver Figura 5.11). El primer grupo, para detectar las hipoglucemias, se siguen los siguientes criterios:
 - **Hipoglucemia:** al menos un valor en las siguientes 24 horas es inferior al T_{hipo} establecido.
 - **No hipoglucemia:** que los valores en las siguientes 24 horas sean iguales o superiores a T_{hipo} o iguales e inferiores a T_{hiper} .

El segundo grupo, para detectar las hiperglucemias, los criterios son los siguientes:

- **Hiperglucemia:** al menos un valor en las siguientes 24 horas es superior al T_{hiper} establecido.
- **No hipoglucemia:** que los valores en las siguientes 24 horas sean iguales o superiores a T_{hipo} o iguales e inferiores a T_{hiper} .

5. Una vez realizada la clasificación, se genera un espectrograma a partir de cada serie temporal aplicando la transformada wavelet. El espectrograma es una imagen resultante de la aplicación de la función wavelet correspondiente a cada valor de glucosa. Las funciones wavelet utilizadas se detallan en la Sección 5.2. Se seleccionan 24 h de registros para obtener imágenes con suficiente información tras aplicar la transformada wavelet.
6. Se entrena una CNN de aprendizaje profundo con las imágenes (espectrogramas) generadas en el paso anterior. Se han utilizado dos arquitecturas CNN diferentes para comparar el rendimiento, las cuales han sido detalladas en la Sección 5.1.

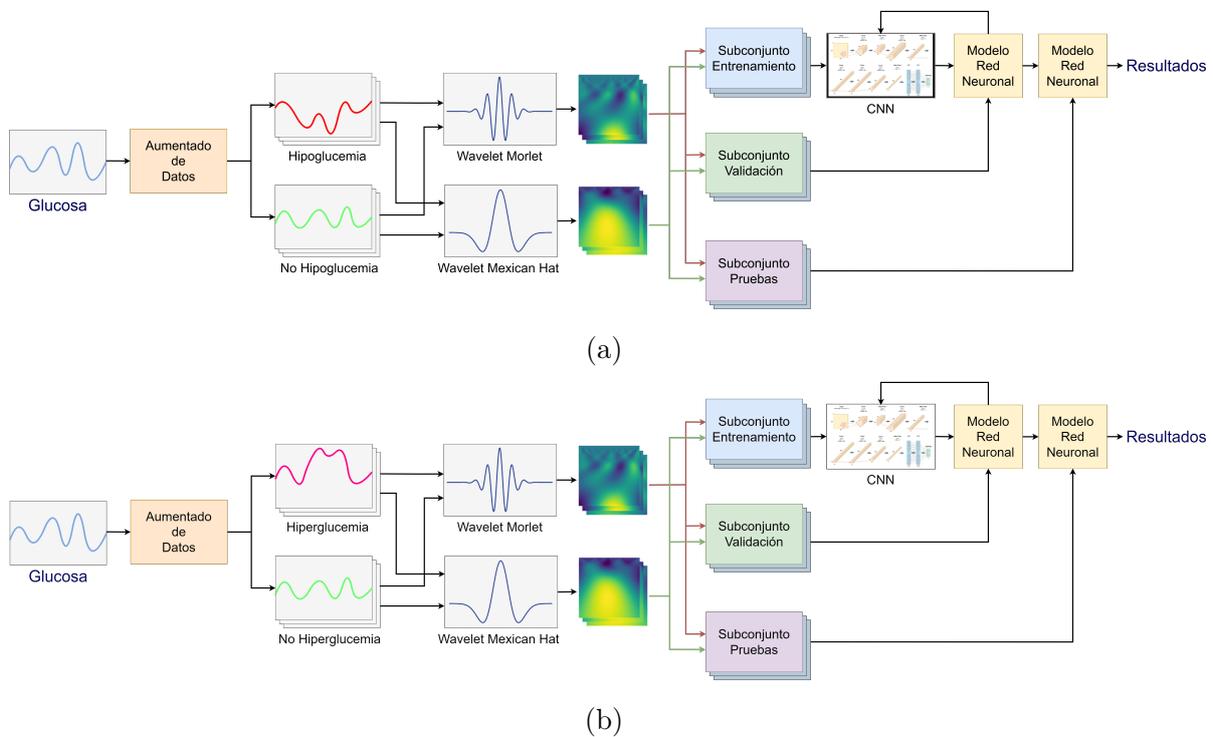
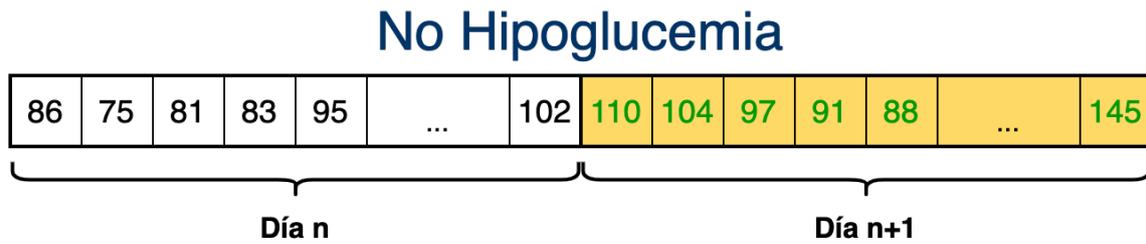
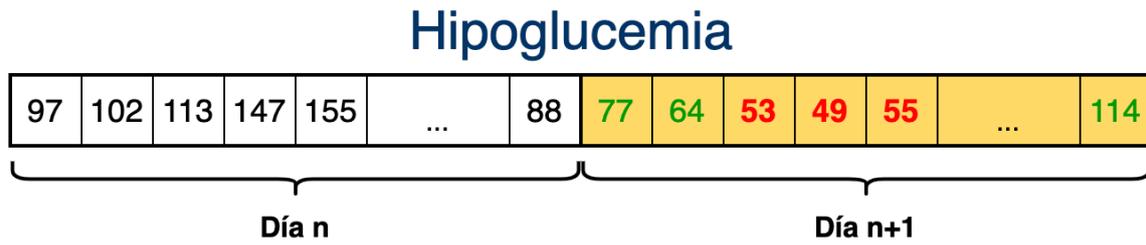
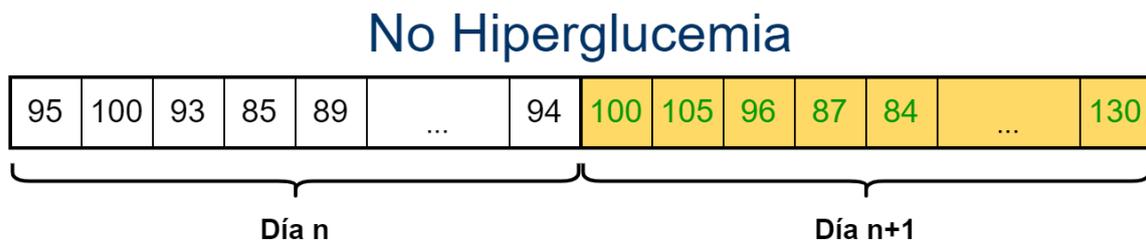
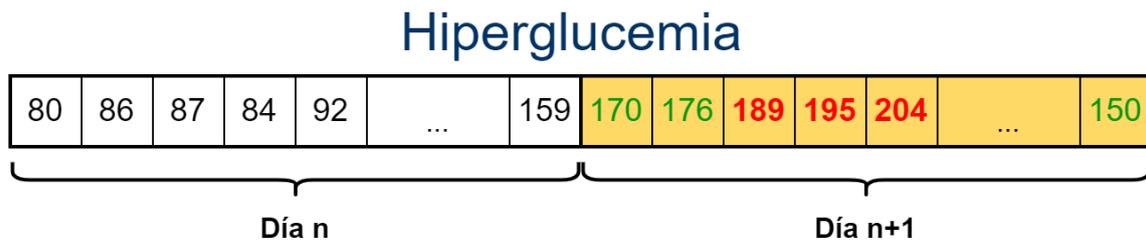


Figura 5.10: Método propuesto. La Figura (a) muestra el método para la detección de hipoglucemias y la Figura (b) muestra el método para la detección de hiperglucemias.



(a)



(b)

Figura 5.11: Clasificación de las imágenes en dos grupos. La Figura (a) muestra la clasificación de los datos diarios de glucosa para la detección de hipoglucemias y la Figura (b) muestra la clasificación de los datos diarios de glucosa para la detección de hiperglucemias.

5.5. Resultados

5.5.1. Configuración experimental

El método propuesto se ha desarrollado utilizando Python con las librerías TensorFlow y Keras, junto con el paquete denominado PyWavelets, que incluye una selección de funciones para la aplicación de las transformadas wavelet necesarias en el trabajo propuesto.

La Tabla 5.3 muestra los valores de los hiperparámetros aplicados para entrenar los modelos CNN. Para mejorar los resultados, se ha utilizado un método de ajuste de la tasa de aprendizaje (lr , de las siglas en inglés Learning Rate) basado en el tiempo denominado *Time-based Decay*, que permite reducir gradualmente el valor de la tasa de aprendizaje a lo largo del tiempo de entrenamiento del modelo. La tasa de aprendizaje (lr) para cada época se calcula mediante la siguiente expresión (5.6):

$$lr = \frac{lr_0}{(1 + kt)} \quad (5.6)$$

Donde lr_0 es la tasa de aprendizaje inicial, k es el hiperparámetro de decaimiento, y t es el número de épocas. En este trabajo, k es el resultado de dividir la tasa de aprendizaje inicial por el número total de épocas.

Parámetro/CNN	AlexNet	DenseNet-121
Épocas	300	300
Intervalo de validación (épocas)	30	30
lr inicial	0.001	0.001
Método de ajuste lr	Time-based Decay	Time-based Decay
Tamaño Batch	64	64
Optimizador	SGD	Adam
Función de pérdida	Categorical Crossentropy	Categorical Crossentropy

Tabla 5.3: Hiperparámetros usados para entrenar los modelos con las distintas arquitecturas CNN.

Además, se han usado dos de los principales optimizadores utilizados en redes CNN: Stochastic Gradient Descent (SGD) y Adaptive Moment Estimation (Adam) [102]. Para la arquitectura AlexNet se ha elegido el optimizador SGD, dado que es el optimizador recomendado [95] por los autores de esta arquitectura. También se puede utilizar el optimizador Adam, que facilita el entrenamiento de la red a costa de detectar peor los patrones en las imágenes [103]. En el caso de DenseNet-121, se pueden utilizar ambos optimizadores, pero se ha elegido el optimizador Adam

porque ofrece mejores resultados [104, 105] en líneas generales con respecto a SGD, aunque la diferencia entre ambos optimizadores es mínima.

Para realizar los experimentos, validar el aprendizaje del algoritmo utilizado y probar los resultados, es necesario dividir el conjunto de datos de la siguiente manera: el 75 % de las imágenes se han utilizado para el entrenamiento, el 15 % para la validación y el 10 % para test. Además, se ha comprobado de que no existen fugas de datos de los conjuntos de entrenamiento a los de test y de que las imágenes del mismo paciente no estuvieran presentes tanto en los conjuntos de entrenamiento como en los de validación/prueba.

Se han entrenado dos tipos diferentes de modelos: modelos generales con los datos de todos los pacientes y modelos personales con los datos de cada paciente. En ambos tipos de modelos se han generado imágenes para cada paciente. A continuación, se han eliminado aleatoriamente las imágenes de la clase mayoritaria para equilibrar ambas clases con el mismo número de imágenes. En el caso de los modelos generales, se componen poniendo las imágenes de todos los modelos personales sin validación cruzada.

5.5.2. Conjunto de Datos

Para este segundo enfoque de este trabajo de tesis, se ha usado un conjunto de datos recogido gracias a la colaboración del grupo ABSys² de la Universidad Complutense de Madrid con el Hospital Príncipe de Asturias de Alcalá de Henares (España). El comité ético del hospital aprobó la recogida de datos para los estudios y todos los pacientes firmaron previamente un consentimiento. El conjunto de datos está formado por los datos de 20 pacientes con T1DM y es diferente al presentado en el Capítulo 4.

Los datos se han obtenido utilizando el sistema MCG Abbott FreeStyle Libre, que registran los valores de glucosa en intervalos de 15 minutos. La Tabla 5.4 muestra las características clínicas de los participantes en el estudio. Las características de los pacientes son mujeres (50 %), edad media 40,16 (+/- 11,86), peso 69,31 (+/- 13,62) kg, altura 169,31 (+/- 9,69) cm, hemoglobina glucosilada (HBA1c) promedia de 7,45 % (+/- 0,93 %), años de enfermedad 17,9 (+/- 9,7), tratamiento: 60 >= 60 % (ISCI) y 40 >= 40 % (MDI). Además, la Tabla 5.5 muestra la distribución de eventos de hipoglucemia por paciente. Aunque, en general, el proceso DA consigue generar un elevado número de imágenes, en conjuntos de datos específicos como los de los pacientes HUPA0002P y HUPA0003P, el número de imágenes generadas es menor debido a que los eventos de hipoglucemia se concentran en el tiempo.

La Figura 5.12 muestra el porcentaje de tiempo en rango de cada paciente. Este valor se traduce como el tiempo que el paciente mantiene el valor de glucosa entre 70 y 180 *mg/dL*. El valor se muestra junto con los porcentajes de tiempo que los pacientes están en otras situaciones, que son las siguientes: hipoglucemia grave (< 54 *mg/dL*), hipoglucemia (54–70 *mg/dL*), hiperglucemia (180–250 *mg/dL*) e hiperglucemia grave (> 250 *mg/dL*). Puede observarse que la mayoría

²<http://absys.dacya.ucm.es/>

5.5. RESULTADOS

Medida	HUPA0001P	HUPA0002P	HUPA0003P	HUPA0004P	HUPA0005P	HUPA0006P	HUPA0007P	HUPA0008P	HUPA0011P	HUPA0014P
Sexo	Mujer	Hombre	Mujer	Hombre	Mujer	Hombre	Mujer	Mujer	Mujer	Mujer
HbA1c [%]	8.2	7.1	7.3	7.8	6.9	7.8	6.6	8.0	6.0	8.5
Edad [años]	56.3	48.6	43.4	41.2	41.9	22.1	37.6	18.5	41.9	50.0
Años desde debut	15.5	36.5	12.5	8.5	39.5	13.5	10.1	8.7	15.2	12.9
Peso [kg]	59.0	82.4	62.0	88.0	58.5	71.0	102.6	83.5	51.0	61.0
Altura [cm]	161	186	182	180	161	170	183	161	164	155
ISCI/MDI	ISCI	MDI								

Medida	HUPA0015P	HUPA0016P	HUPA0017P	HUPA0020P	HUPA0022P	HUPA0023P	HUPA0024P	HUPA0025P	HUPA0026P	HUPA0027P
Sexo	Mujer	Mujer	Mujer	Hombre	Hombre	Hombre	Hombre	Hombre	Mujer	Hombre
HbA1c [%]	6.4	6.5	8.2	9.7	6.7	7.7	8.3	7.0	7.2	7.0
Edad [años]	43.1	29.9	26.3	45.7	59.6	22.9	47.9	38.1	61.8	26.4
Años desde debut	11.2	20.1	24.2	13.5	14.6	0.8	35.9	20.3	21.5	23.7
Peso [kg]	58.6	64.9	61.8	71.6	77.6	55.5	80.5	104.7	80.0	76.0
Altura [cm]	162	157	167	168	179	173	174	188	165	185
ISCI/MDI	MDI	ISCI	MDI	MDI	ISCI	MDI	MDI	ISCI	MDI	MDI

Tabla 5.4: Caracterización clínica de cada paciente. ID, Sexo, HbA1c, Edad, Años de enfermedad, Peso, Altura, Tratamiento (MDI: Múltiple Dosis de Insulina; ISCI: Infusión Subcutánea Continua de Insulina).

ID Paciente	Registros	Hipo	Imágenes Training	Imágenes Test
HUPA0001P	1312	9	190	26
HUPA0002P	1312	34	32	6
HUPA0003P	1312	22	56	10
HUPA0004P	1027	12	92	14
HUPA0005P	1306	16	202	28
HUPA0006P	725	12	30	4
HUPA0007P	1304	18	84	12
HUPA0008P	1304	15	82	12
HUPA0011P	1248	10	170	24
HUPA0014P	1304	10	150	22
HUPA0015P	1248	14	134	20
HUPA0016P	1309	32	42	6
HUPA0017P	1293	11	202	28
HUPA0020P	1268	13	150	20
HUPA0022P	1248	21	56	10
HUPA0023P	1248	12	172	24
HUPA0024P	1248	15	106	16
HUPA0025P	1247	29	74	12
HUPA0026P	16409	173	2568	346
HUPA0027P	12574	177	1220	164
Media	2562.30	32.75	290.60	40.20
Desv.	148.13	7.29	60.64	8.02

Tabla 5.5: Características de los datasets por paciente. Registros indica el número de registros de glucosa, Hipo indica número de eventos de hipoglucemias, Imágenes Train indica el número de imágenes para entrenamiento, e Imágenes Test el número de imágenes utilizadas para test. Tanto para training como para test, se utilizan un 50% de imágenes de hipoglucemias.

de los pacientes se encuentran la mayor parte del tiempo en el intervalo de 70 y 180 mg/dL . Esto puede deberse a que los participantes no estaban acostumbrados a usar MCG con regularidad en el momento del estudio. Estudios recientes demuestran que el uso de MCG suele aumentar el porcentaje de tiempo en rango de las personas con diabetes. Además, la Figura 5.13 muestra el rango intercuartílico y la mediana del valor de glucosa de los pacientes. Se puede observar que existe una gran variabilidad en los datos, ya que la mediana de glucosa difiere entre los pacientes. Junto con estas observaciones, podemos destacar una importante variabilidad glucémica intrapaciente, ya que la desviación estándar de la glucosa es elevada, lo que puede considerarse como una dificultad añadida en la predicción.

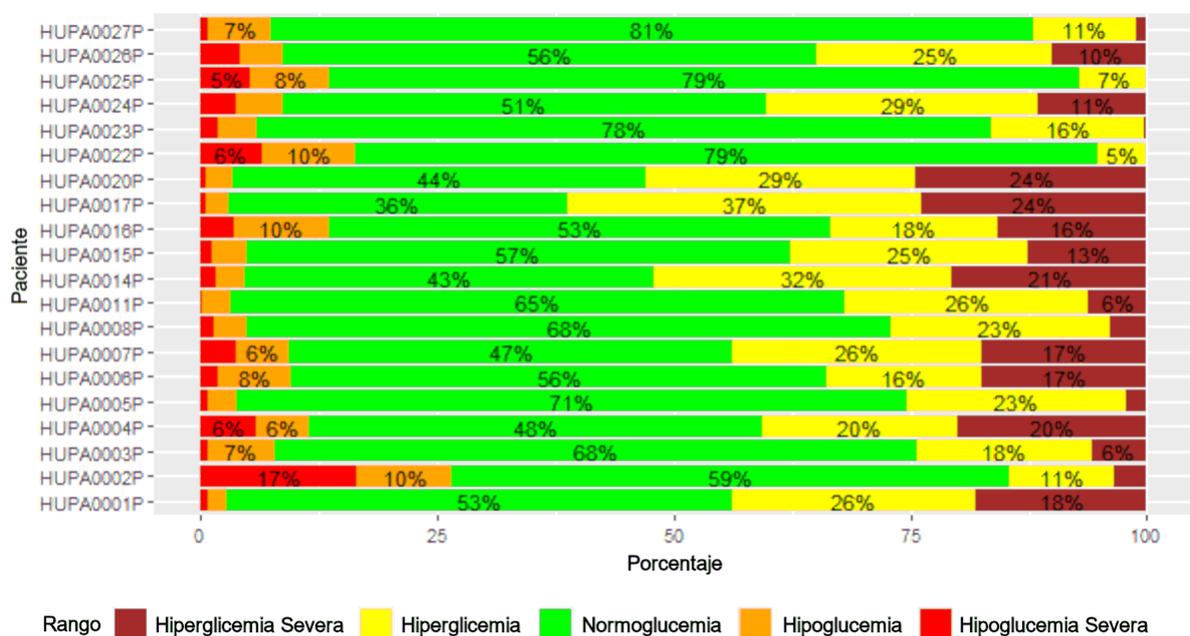


Figura 5.12: Porcentaje de tiempo que el paciente tiene un nivel de glucosa muy bajo ($< 54 mg/dL$), bajo ($[54, 70) mg/dL$), en rango ($[70, 180] mg/dL$), alto ($(180, 250] mg/dL$) y muy alto ($> 250 mg/dL$).

5.5.3. Resultados en predicción de Hiperglucemias

Para entrenar los modelos predictivos de hiperglucemias, se ha establecido como umbral de hiperglucemia $> 240 mg/dL$. Se ha elegido este umbral debido a que al probar con los umbrales de $> 180 mg/dL$ y $> 200 mg/dL$ no se han generado suficientes imágenes de la clase *No Hiperglucemia* para poder generar modelos válidos. Esto es debido a la detección en casi todas las ventanas de datos (Véase Figura 5.11b) al menos un valor o más de BG por encima de los umbrales de hiperglucemia, lo que ha causado que la mayoría de las imágenes generadas hayan sido etiquetadas con la categoría *Hiperglucemia*. El modelo general entrenado es parcial, es decir, hay pacientes que no ha sido posible obtener conjuntos de datos válidos por el motivo anteriormente expuesto. En la Tabla 5.6 se puede observar la precisión de los modelos generales

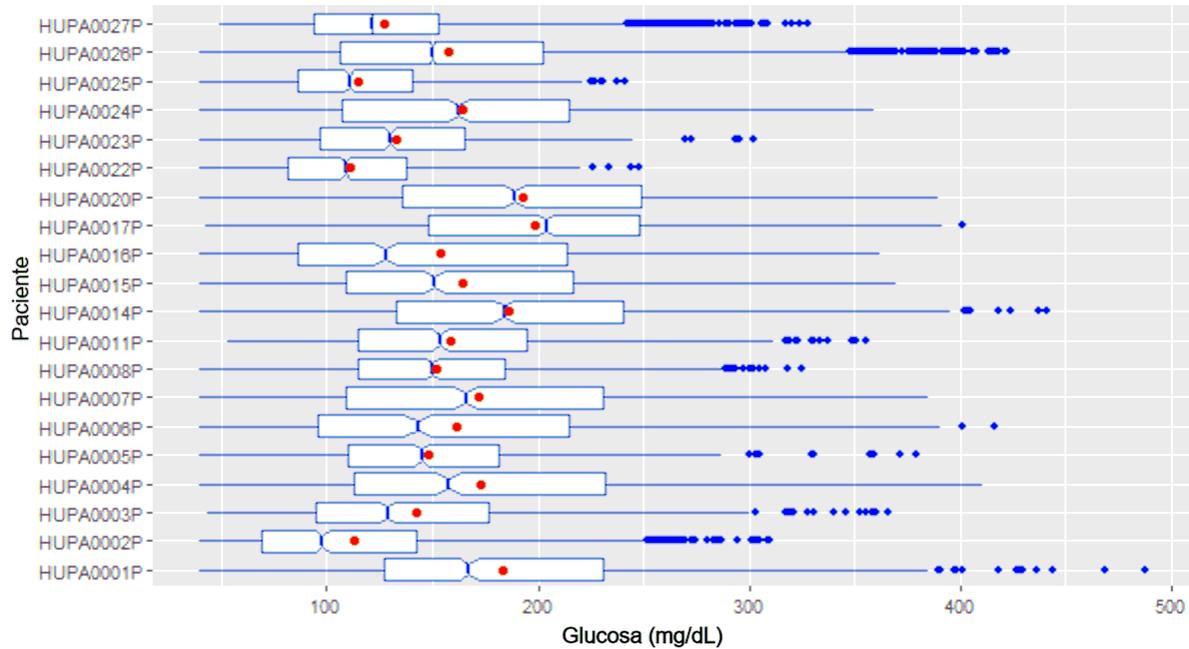


Figura 5.13: Rangos intercuartílicos de glucosa por paciente. El valor medio se muestra con un punto rojo.

(parciales) entrenados con las diferentes arquitecturas CNN y tipos de wavelet, estableciendo un umbral de hiperglucemia de $> 240 \text{ mg/dL}$. El modelo con la combinación de DenseNet-121 y la wavelet Mexican Hat es la que ofrece los mejores resultados, con una precisión del 78 %, aunque el resto de modelos también ofrecen buenos resultados.

Mexican Hat Wavelet							
	Entrenamiento	Validación	Test				
CNN	Accuracy	Accuracy	Accuracy	Sensitivity	Specificity	Precision	AUC
AlexNet	99.73 (+/- 0.17)	72.61 (+/- 1.50)	69.78 (+/- 2.43)	72.13 (+/- 4.17)	69.36 (+/- 4.82)	66.25 (+/- 11.60)	0.23 (+/- 0.02)
DenseNet-121	98.32 (+/- 0.12)	81.53 (+/- 0.83)	78.52 (+/- 1.53)	78.74 (+/- 1.89)	78.37 (+/- 1.92)	78.18 (+/- 2.49)	0.13 (+/- 0.02)
Morlet Wavelet							
	Entrenamiento	Validación	Test				
CNN	Accuracy	Accuracy	Accuracy	Sensitivity	Specificity	Precision	AUC
AlexNet	99.81 (+/- 0.01)	69.37 (+/- 0.88)	65.04 (+/- 1.19)	65.71 (+/- 1.14)	64.52 (+/- 1.70)	62.98 (+/- 3.67)	0.30 (+/- 0.01)
DenseNet-121	98.37 (+/- 0.14)	75.49 (+/- 0.80)	74.19 (+/- 1.16)	75.75 (+/- 1.79)	72.87 (+/- 1.25)	71.26 (+/- 2.16)	0.18 (+/- 0.01)

Tabla 5.6: Comparación de los modelos generales entrenados con las diferentes arquitecturas CNN y wavelets. Umbral de hiperglucemia: $> 240 \text{ mg/dL}$. Accuracy: Exactitud, Sensitivity: Sensibilidad, Specificity: Especificidad, Precision: Precisión.

Debido al problema de generar los conjuntos de imágenes explicado con anterioridad, no ha sido posible entrenar los modelos personales de todos los pacientes. Además, los pacientes que, si se han podido generar las imágenes, el número de las mismas es bajo, lo que conduce a modelos poco fiables y robustos. Por estos motivos, no se muestran las tablas de resultados de los modelos personales.

Visto lo anterior, la detección de hiperglucemias usando el método propuesto no ha dado los

resultados esperados. Debido al incorrecto control glucémico de algunos pacientes, sobre todo con valores de BG elevados a lo largo del tiempo, el método no ha funcionado como esperábamos. Por los motivos expresados anteriormente, hemos decidido poner el enfoque en la detección de hipoglucemias.

5.5.4. Resultados de predicción de Hipoglucemias

En el caso de las hipoglucemias, se ha establecido el umbral de $< 70 \text{ mg/dL}$. Los resultados obtenidos por los modelos entrenados con las diferentes arquitecturas CNN y tipos de wavelet se muestran en la Tabla 5.7. Al igual que con las hiperglucemias, la combinación de DenseNet-121 y la wavelet Mexican Hat es la que ofrece los mejores resultados, con una precisión del 81 %, aunque también de forma general, se obtienen buenos resultados utilizando la wavelet Morlet.

A diferencia de lo ocurrido a la hora de generar los conjuntos de datos para la detección de hiperglucemias, en el caso de las hipoglucemias se han podido generar las imágenes de todos los pacientes. Gracias a ello, se han podido entrenar modelos personalizados, es decir, modelos entrenados sólo con los datos de cada paciente. Los resultados de los modelos personales se pueden observar en las Tablas 5.8 y 5.9. Para establecer las comparaciones necesarias, los modelos generales y de cada paciente se han entrenado utilizando las wavelets Mexican Hat y Morlet. Las diferentes características de cada paciente y la falta de datos suficientes en algunos de ellos explican las variaciones entre los resultados obtenidos por los modelos. Además, al igual que los modelos generales, la wavelet Mexican Hat obtiene mejores resultados, pero la diferencia es insignificante. Cabe destacar el caso del paciente con ID *HUPA0003P*, la baja precisión obtenida por el modelo se debe a que el paciente ha tenido un mal control glucémico, lo que ha provocado que la mayoría de las imágenes generadas se encuentren dentro del rango de hipoglucemia, como se puede observar en la Figura 5.14. En términos generales, los modelos de predicción personalizados por cada paciente tienen una precisión cercana al 80 %.

Wavelet Mexican Hat							
	Entrenamiento	Validación	Test				
CNN	Accuracy	Accuracy	Accuracy	Sensitivity	Specificity	Precision	AUC
AlexNet	99.86 (+/- 0.01)	70.79 (+/- 1.13)	70.91 (+/- 3.06)	71.50 (+/- 2.31)	71.06 (+/- 3.82)	69.99 (+/- 9.99)	0.77 (+/- 0.02)
DenseNet-121	98.61 (+/- 0.09)	79.81(+/- 0.49)	80.15 (+/- 0.82)	79.55 (+/- 1.22)	80.83 (+/- 1.23)	81.21 (+/- 1.70)	0.88 (+/- 0.01)
Wavelet Morlet							
	Entrenamiento	Validación	Test				
CNN	Accuracy	Accuracy	Accuracy	Sensitivity	Specificity	Precision	AUC
AlexNet	99.95 (+/- 0.00)	64.59 (+/- 0.71)	64.99 (+/- 1.10)	64.88 (+/- 1.46)	65.18 (+/- 1.11)	65.54 (+/- 2.26)	0.71 (+/- 0.01)
DenseNet-121	98.79 (+/- 0.07)	73.91 (+/- 0.68)	72.98 (+/- 0.84)	72.98 (+/- 1.27)	73.08 (+/- 1.46)	73.08 (+/- 2.63)	0.81 (+/- 0.01)

Tabla 5.7: Comparación de los modelos generales entrenados con las diferentes arquitecturas CNN y wavelets. Umbral de hipoglucemia: $< 70 \text{ mg/dL}$. Accuracy: Exactitud, Sensitivity: Sensibilidad, Specificity: Especificidad, Precision: Precisión.

5.5.5. Discusión

La Tabla 5.10 muestra una comparación del modelo general para la detección de hipoglucemias con y sin ruido con otras técnicas de aprendizaje automático y redes neuronales. Los modelos se han entrenado con datos de todos los pacientes, y se ha utilizado el umbral de hipoglucemia

Wavelet Mexican Hat							
ID Paciente	Entrenamiento	Validación	Test				
	Accuracy	Accuracy	Accuracy	Sensitivity	Specificity	Precision	AUC
HUPA0001P	100.00 (+/- 0.00)	94.12 (+/- 1.47)	98.72 (+/- 1.81)	100.00 (+/- 0.00)	97.62 (+/- 3.37)	97.44 (+/- 3.63)	1.00 (+/- 0.00)
HUPA0002P	100.00 (+/- 0.00)	87.50 (+/- 7.05)	76.67 (+/- 18.60)	75.00 (+/- 45.83)	73.25 (+/- 18.03)	53.33 (+/- 37.20)	0.99 (+/- 0.04)
HUPA0003P	100.00 (+/- 0.00)	82.33 (+/- 4.23)	43.33 (+/- 9.07)	43.78 (+/- 12.15)	43.30 (+/- 7.57)	36.00 (+/- 8.00)	0.41 (+/- 0.09)
HUPA0004P	100.00 (+/- 0.00)	80.74 (+/- 3.43)	71.19 (+/- 3.44)	71.43 (+/- 4.98)	71.15 (+/- 2.05)	71.43 (+/- 0.00)	0.81 (+/- 0.04)
HUPA0005P	100.00 (+/- 0.00)	96.25 (+/- 1.25)	97.74 (+/- 3.26)	97.35 (+/- 4.20)	98.44 (+/- 3.69)	98.33 (+/- 3.99)	1.00 (+/- 0.01)
HUPA0006P	100.00 (+/- 0.00)	70.56 (+/- 8.26)	100.00 (+/- 0.00)	100.00 (+/- 0.00)	100.00 (+/- 0.00)	100.00 (+/- 0.00)	1.00 (+/- 0.00)
HUPA0007P	100.00 (+/- 0.00)	94.38 (+/- 3.73)	86.39 (+/- 5.48)	98.06 (+/- 5.96)	79.78 (+/- 6.06)	74.44 (+/- 9.36)	0.94 (+/- 0.05)
HUPA0008P	100.00 (+/- 0.00)	93.33 (+/- 4.82)	86.39 (+/- 4.56)	100.00 (+/- 0.00)	79.01 (+/- 5.67)	72.78 (+/- 9.11)	0.89 (+/- 0.05)
HUPA0011P	100.00 (+/- 0.00)	98.73 (+/- 1.64)	85.42 (+/- 2.34)	99.67 (+/- 1.80)	77.64 (+/- 2.85)	71.11 (+/- 4.68)	0.96 (+/- 0.02)
HUPA0014P	100.00 (+/- 0.00)	88.89 (+/- 2.17)	91.36 (+/- 4.75)	95.95 (+/- 4.66)	88.49 (+/- 4.66)	86.67(+/- 9.31)	0.99 (+/- 0.01)
HUPA0015P	100.00 (+/- 0.00)	93.97 (+/- 2.37)	85.67 (+/- 4.03)	94.20 (+/- 7.61)	80.29 (+/- 3.21)	76.67 (+/- 4.71)	0.94 (+/- 0.04)
HUPA0016P	100.00 (+/- 0.00)	83.33 (+/- 8.12)	67.22 (+/- 8.03)	93.89 (+/- 15.80)	61.22 (+/- 6.32)	38.89 (+/- 12.42)	0.73 (+/- 0.09)
HUPA0017P	100.00 (+/- 0.00)	94.83 (+/- 2.03)	94.76 (+/- 4.49)	91.39 (+/- 6.86)	99.48 (+/- 1.94)	99.52 (+/- 1.78)	1.00 (+/- 0.00)
HUPA0020P	100.00 (+/- 0.00)	91.89 (+/- 2.05)	91.17 (+/- 2.79)	98.59 (+/- 3.59)	85.95 (+/- 3.67)	83.67 (+/- 4.82)	0.98 (+/- 0.01)
HUPA0022P	100.00 (+/- 0.00)	77.00 (+/- 9.00)	77.00 (+/- 7.18)	75.02 (+/- 7.56)	75.06 (+/- 7.16)	75.33 (+/- 8.46)	0.78 (+/- 0.05)
HUPA0023P	100.00 (+/- 0.00)	96.18 (+/- 1.72)	93.19 (+/- 2.28)	88.16 (+/- 3.53)	100.00 (+/- 0.00)	100.00 (+/- 0.00)	0.99 (+/- 0.01)
HUPA0024P	100.00 (+/- 0.00)	91.33 (+/- 2.21)	92.08 (+/- 3.20)	87.76 (+/- 2.67)	98.01 (+/- 5.30)	97.92 (+/- 5.67)	0.95 (+/- 0.03)
HUPA0025P	100.00 (+/- 0.00)	84.29 (+/- 2.86)	84.29 (+/- 2.86)	94.72 (+/- 5.04)	93.17 (+/- 7.32)	97.46 (+/- 5.70)	0.99 (+/- 0.02)
HUPA0026P	99.17 (+/- 0.08)	85.53 (+/- 0.67)	84.50 (+/- 1.09)	86.87 (+/- 1.33)	82.45 (+/- 1.39)	81.31 (+/- 1.79)	0.92 (+/- 0.01)
HUPA0027P	97.59 (+/- 0.25)	84.30 (+/- 1.27)	85.93 (+/- 2.16)	89.06 (+/- 3.41)	83.62 (+/- 3.27)	82.20 (+/- 4.70)	0.94 (+/- 0.01)
Media	99.84	88.47	85.06	88.97	83.61	79.71	0.92
Desv.	0.55	7.17	13.01	13.41	14.46	18.85	0.13

Tabla 5.8: Comparación de modelos personales entrenados con DenseNet-121 y utilizando la wavelet Mexican Hat. Los resultados se muestran como la media de todas las ejecuciones. Umbral de hipoglucemia: $< 70 \text{ mg/dL}$. El texto en rojo indica valores inferiores a 50%. Accuracy: Exactitud, Sensitivity: Sensibilidad, Specificity: Especificidad, Precision: Precisión.

Wavelet Morlet							
ID Paciente	Entrenamiento	Validación	Test				
	Accuracy	Accuracy	Accuracy	Sensitivity	Specificity	Precision	AUC
HUPA0001P	100.00 (+/- 0.00)	91.14 (+/- 1.73)	91.92 (+/- 3.76)	92.22 (+/- 5.99)	92.16 (+/- 3.77)	92.05 (+/- 4.21)	0.98 (+/- 0.01)
HUPA0002P	100.00 (+/- 0.00)	81.11 (+/- 5.67)	95.56 (+/- 7.37)	93.33 (+/- 11.06)	100.00 (+/- 0.00)	100.00 (+/- 0.00)	1.00 (+/- 0.00)
HUPA0003P	100.00 (+/- 0.00)	96.00 (+/- 4.90)	54.00 (+/- 5.54)	54.17 (+/- 6.07)	53.89 (+/- 5.24)	60.00 (+/- 0.00)	0.72 (+/- 0.04)
HUPA0004P	100.00 (+/- 0.00)	70.74 (+/- 3.78)	83.10 (+/- 6.52)	77.10 (+/- 6.22)	94.10 (+/- 9.66)	94.76 (+/- 8.64)	0.91 (+/- 0.06)
HUPA0005P	100.00 (+/- 0.00)	86.58 (+/- 1.51)	78.69 (+/- 4.07)	87.09 (+/- 7.91)	73.74 (+/- 3.41)	68.10 (+/- 5.13)	0.85 (+/- 0.03)
HUPA0006P	100.00 (+/- 0.00)	71.11 (+/- 7.37)	87.50 (+/- 12.50)	100.00 (+/- 0.00)	83.33 (+/- 16.67)	75.00 (+/- 25.00)	1.00 (+/- 0.00)
HUPA0007P	100.00 (+/- 0.00)	93.33 (+/- 1.56)	83.06 (+/- 4.56)	84.98 (+/- 7.36)	81.96 (+/- 4.01)	81.11 (+/- 5.67)	0.90 (+/- 0.04)
HUPA0008P	100.00 (+/- 0.00)	85.42 (+/- 4.06)	96.67 (+/- 5.09)	96.03 (+/- 6.60)	97.94 (+/- 5.28)	97.78 (+/- 5.67)	0.99 (+/- 0.02)
HUPA0011P	100.00 (+/- 0.00)	80.00 (+/- 2.67)	90.28 (+/- 3.93)	89.12 (+/- 5.42)	92.31 (+/- 5.81)	92.22 (+/- 6.06)	0.98 (+/- 0.01)
HUPA0014P	100.00 (+/- 0.00)	82.89 (+/- 2.23)	72.58 (+/- 4.31)	76.61 (+/- 6.84)	70.34 (+/- 5.25)	66.06 (+/- 8.44)	0.82 (+/- 0.05)
HUPA0015P	100.00 (+/- 0.00)	78.46 (+/- 3.53)	82.67 (+/- 5.73)	79.65 (+/- 7.28)	88.25 (+/- 8.21)	89.00 (+/- 8.31)	0.92 (+/- 0.04)
HUPA0016P	100.00 (+/- 0.00)	98.75 (+/- 3.75)	83.33 (+/- 6.09)	97.78 (+/- 8.31)	76.11 (+/- 6.71)	68.89 (+/- 8.31)	0.91 (+/- 0.10)
HUPA0017P	100.00 (+/- 0.00)	90.25 (+/- 1.75)	89.52 (+/- 3.19)	87.22 (+/- 5.15)	92.70 (+/- 3.18)	93.10 (+/- 3.44)	0.96 (+/- 0.02)
HUPA0020P	100.00 (+/- 0.00)	90.11 (+/- 2.51)	92.00 (+/- 2.77)	100.00 (+/- 0.00)	86.40 (+/- 4.10)	84.00 (+/- 5.54)	0.98 (+/- 0.02)
HUPA0022P	100.00 (+/- 0.00)	86.33 (+/- 4.82)	74.67 (+/- 7.18)	99.17 (+/- 4.49)	67.14 (+/- 6.57)	50.00 (+/- 14.38)	0.96 (+/- 0.04)
HUPA0023P	100.00 (+/- 0.00)	80.29 (+/- 2.96)	84.86 (+/- 2.94)	85.51 (+/- 5.44)	85.17 (+/- 4.60)	84.72 (+/- 5.73)	0.94 (+/- 0.01)
HUPA0024P	100.00 (+/- 0.00)	82.00 (+/- 4.00)	59.17 (+/- 6.40)	62.61 (+/- 9.76)	57.65 (+/- 5.57)	47.08 (+/- 11.03)	0.64 (+/- 0.05)
HUPA0025P	100.00 (+/- 0.00)	78.57 (+/- 0.00)	88.06 (+/- 5.13)	84.62 (+/- 1.64)	93.10 (+/- 9.49)	92.78 (+/- 10.26)	0.87 (+/- 0.03)
HUPA0026P	99.29 (+/- 0.10)	82.19 (+/- 0.93)	76.83 (+/- 1.45)	76.71 (+/- 1.50)	77.06 (+/- 2.17)	77.11 (+/- 3.01)	0.85 (+/- 0.01)
HUPA0027P	98.04 (+/- 0.24)	78.39 (+/- 0.84)	79.51 (+/- 2.25)	79.66 (+/- 3.07)	79.74 (+/- 3.35)	79.55 (+/- 4.85)	0.87 (+/- 0.02)
Media	99.87	84.18	82.20	85.18	82.15	79.67	0.90
Desv.	0.45	7.31	10.71	11.82	12.48	15.10	0.09

Tabla 5.9: Comparación de modelos personales entrenados con DenseNet-121 y utilizando la wavelet Morlet. Los resultados se muestran como la media de todas las ejecuciones. Umbral de hipoglucemia: $< 70 \text{ mg/dL}$. El texto en rojo indica valores inferiores a 50%. Accuracy: Exactitud, Sensitivity: Sensibilidad, Specificity: Especificidad, Precision: Precisión.

$< 70 \text{ mg/dL}$. Además, se ha utilizado la técnica DA descrita en la sección 5.3.1 pero sin generar imágenes con el resto de técnicas ML y NN, es decir, se han utilizado series temporales de glucosa

como entrada a los modelos de las otras técnicas. Los resultados de las Tablas 5.7, 5.8, 5.9 y 5.10 son la media de 30 ejecuciones realizadas para cada modelo. Nuestro modelo con un 1 % de ruido, aplicando la técnica DA descrita en la Sección 5.3.2, es el que obtiene mejores resultados, con un 89 % de exactitud, un 88 % de sensibilidad, un 90 % de especificidad y un 90 % de precisión. Los resultados muestran que las demás técnicas obtienen una sensibilidad mayor que nuestro enfoque, pero alcanzan una especificidad muy baja, incluso por debajo del 20 %.

Modelo	Accuracy	Sensitivity	Specificity	Precision
Método propuesto con 1 % ruido	89.42 (+/- 0.55)	88.75 (+/- 0.94)	90.14 (+/- 0.74)	90.31 (+/- 0.84)
Método propuesto sin ruido	80.15 (+/- 0.82)	79.55 (+/- 1.22)	80.83 (+/- 1.23)	81.21 (+/- 1.70)
Extra-Trees Classifier	72.25 (+/- 1.20)	98.52 (+/- 0.44)	19.39 (+/- 1.94)	71.09 (+/- 1.26)
Bagging Classifier	71.91 (+/- 1.26)	97.27 (+/- 0.65)	20.91 (+/- 1.88)	71.22 (+/- 1.26)
Random Forest Classifier	71.84 (+/- 1.32)	97.06 (+/- 0.60)	21.13 (+/- 2.15)	71.23 (+/- 1.37)
Gradient Boosting	69.33 (+/- 1.30)	94.64 (+/- 0.64)	18.43 (+/- 2.05)	70.01 (+/- 1.31)
Gaussian Process Classifier	68.52 (+/- 0.91)	91.81 (+/- 1.02)	21.68 (+/- 1.69)	70.23 (+/- 1.08)
Decision Tree Classifier	68.44 (+/- 1.39)	92.97 (+/- 3.03)	19.14 (+/- 4.73)	69.84 (+/- 1.47)
AdaBoost Classifier	62.26 (+/- 1.24)	70.41 (+/- 1.29)	45.88 (+/- 2.28)	72.36 (+/- 1.52)
Gaussian Naive Bayes	60.25 (+/- 1.55)	61.06 (+/- 1.79)	58.60 (+/- 2.16)	74.81 (+/- 1.11)
Multilayer Perceptron Classifier	56.40 (+/- 10.16)	60.69 (+/- 30.90)	48.40 (+/- 32.62)	72.50 (+/- 5.26)
Stochastic Gradient Descent Classifier	54.69 (+/- 13.53)	63.35 (+/- 41.36)	37.19 (+/- 42.64)	66.02 (+/- 19.92)

Tabla 5.10: Comparación de método propuesto con otras técnicas de aprendizaje automático y redes neuronales. Los resultados se muestran como la media de todas las ejecuciones. El texto en rojo indica valores inferiores al 50 %. Accuracy: Exactitud, Sensitivity: Sensibilidad, Specificity: Especificidad, Precision: Precisión.

La Figura 5.14 muestra el porcentaje de imágenes generadas con hipoglucemia y sin hipoglucemia y la precisión en la fase de test obtenida por los modelos de cada paciente. Estos modelos han sido entrenados con un umbral de hipoglucemia $< 70 \text{ mg/dL}$ y wavelet Mexican Hat. Se puede observar que la precisión del modelo no está relacionada con el porcentaje de imágenes de hipoglucemia debido a las diferentes características así como a la cantidad de datos disponibles para cada paciente.

Las Figuras 5.15 y 5.16 muestran las curvas ROC, AUC, y la prueba estadística KS para los modelos generales de hiperglucemias e hipoglucemias respectivamente. En el caso de la detección de hipoglucemias, podemos ver en la Figura 5.16 que los clasificadores que utilizan DenseNet tienen valores de AUC superiores a 0,8, claramente mejores que los de los clasificadores que utilizan AlexNet. Este hecho también puede apreciarse visualmente en las líneas ROC (columna izquierda) y en la separación entre las funciones de distribución acumulativa (columna derecha). Del mismo modo, hay una clara mejora cuando se utiliza la wavelet Mexican Hat frente a la Morl tanto para DenseNet como para AlexNet. Por el otro lado, se puede observar en la Figura 5.15 que los valores de AUC están por debajo de 0,25, saliendo a la luz la falta de robustez de los modelos de hiperglucemias por el problema mencionado anteriormente.

5.5.6. Análisis de Incertidumbre

Como se ha visto anteriormente, los modelos se han entrenado utilizando un umbral de hipoglucemia fijo. En la vida real, el umbral puede cambiar debido a las características de cada

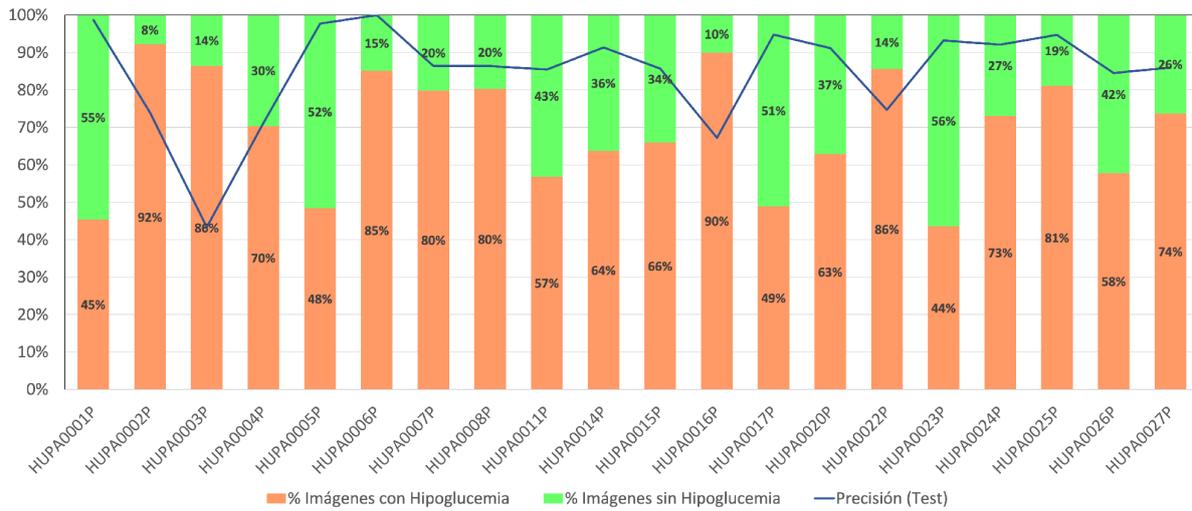


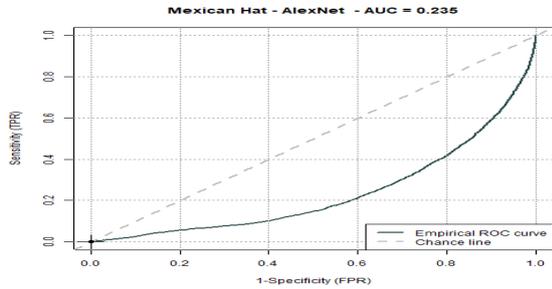
Figura 5.14: Porcentaje de imágenes generadas con hipoglucemia (barras naranjas) y sin hipoglucemia (barras verdes) y la precisión en test (línea azul) obtenida por los modelos de cada paciente. Umbral de hipoglucemia: $< 70 \text{ mg/dL}$, Wavelet: Mexican Hat.

paciente. Por ello, se ha realizado un análisis de incertidumbre para comprobar si el método propuesto se ve afectado en un rango de umbrales de hipoglucemia, en concreto, entre 65 y 75 mg/dL . La Figura 5.17 muestra los gráficos de la media y la desviación estándar de la precisión de la prueba de los modelos personales en diferentes umbrales de hipoglucemia. Las figuras 5.17a y 5.17b muestran los resultados después de 30 ejecuciones para la wavelet Mexican Hat y la wavelet Morlet, respectivamente. Como puede verse, no hay cambios significativos en la precisión en los diferentes umbrales de hipoglucemia. Hay poca variabilidad en los resultados del modelo cuando se modifica el umbral de hipoglucemia.

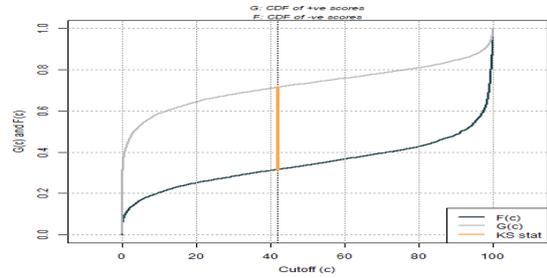
5.5.7. Análisis de Sensibilidad

Se ha realizado un análisis de sensibilidad para determinar cuánto ruido puede manejar el método propuesto sin perder fiabilidad mediante la aplicación de la técnica DA detallada en la Sección 5.3.2. Durante este análisis, se añade deliberadamente ruido a las series temporales originales de glucosa en los datos de entrenamiento, oscilando entre el 1% y el 100% de ruido. El objetivo es medir la resistencia del método y su rendimiento en presencia de distintos niveles de ruido. La razón de introducir este ruido es evaluar la capacidad del método propuesto para funcionar eficazmente frente a la incertidumbre y la variabilidad inherentes a los datos de glucosa del mundo real. Aumentando sistemáticamente los niveles de ruido, se pretende entender cómo afecta esto a la precisión de las predicciones. Es importante destacar que el método está diseñado para su uso en el mundo real, y la introducción intencionada de ruido controlado ayudó a evaluar la capacidad para manejar datos de entrada ruidosos, manteniendo al mismo tiempo la fiabilidad de las predicciones.

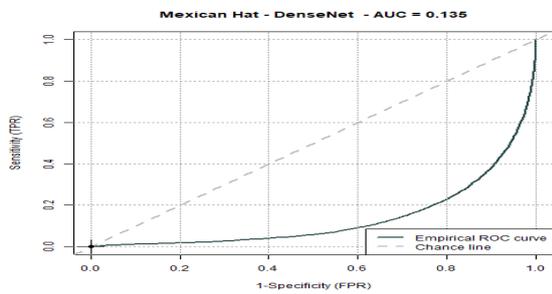
En cuanto a los datos provenientes del sistema MCG, cabe señalar que esta fuente de datos



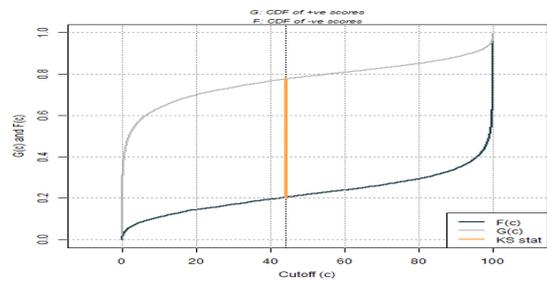
(a) Mexican Hat. AlexNet. Curva ROC.



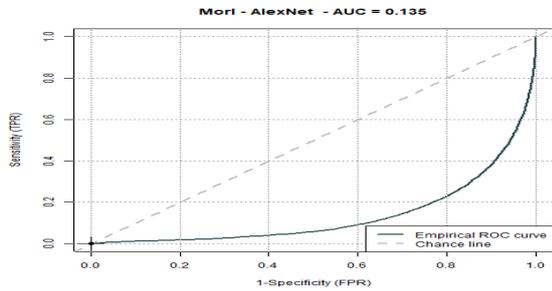
(b) Mexican Hat. AlexNet. Curva KS.



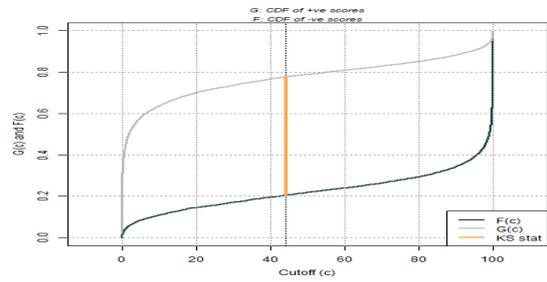
(c) Mexican Hat. DenseNet. Curva ROC.



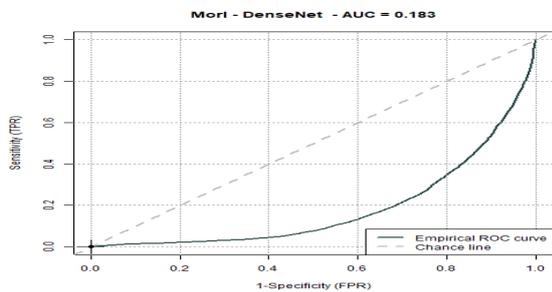
(d) Mexican Hat. DenseNet. Curva KS.



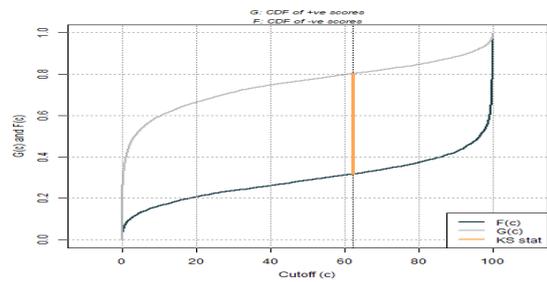
(e) Morl. AlexNet. Curva ROC.



(f) Morl. AlexNet. Curva KS.

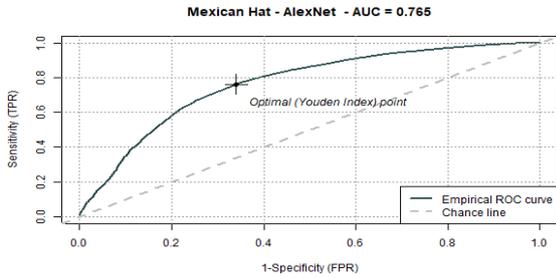


(g) Morl. DenseNet. Curva ROC.

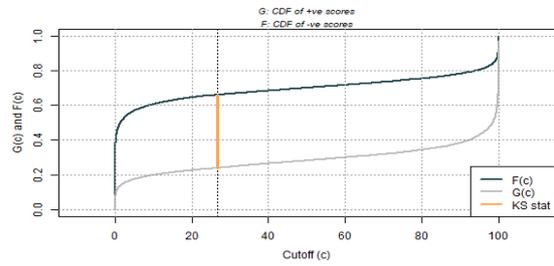


(h) Morl. DenseNet. Curva KS.

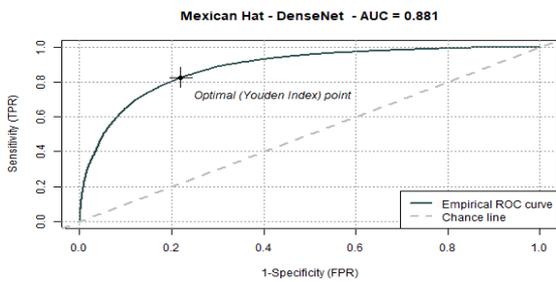
Figura 5.15: Curvas ROC de los modelos generales de hiperglucemias



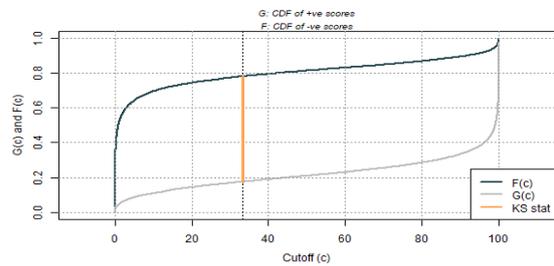
(a) Mexican Hat. AlexNet. Curva ROC.



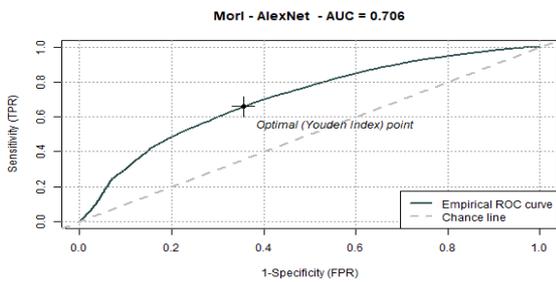
(b) Mexican Hat. AlexNet. Curva KS.



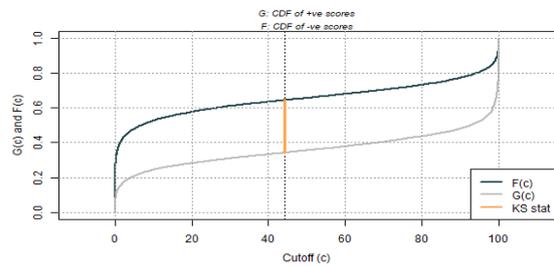
(c) Mexican Hat. DenseNet. Curva ROC.



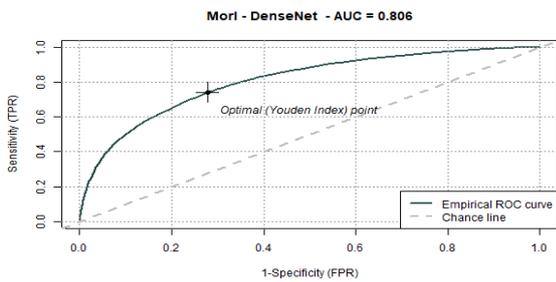
(d) Mexican Hat. DenseNet. Curva KS.



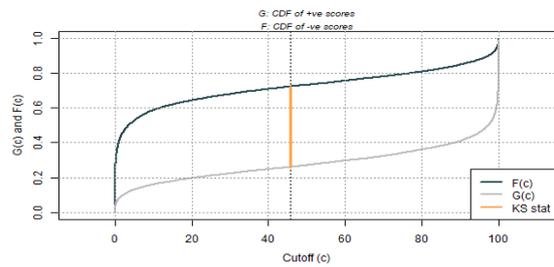
(e) Morl. AlexNet. Curva ROC.



(f) Morl. AlexNet. Curva KS.



(g) Morl. DenseNet. Curva ROC.



(h) Morl. DenseNet. Curva KS.

Figura 5.16: Curvas ROC de los modelos generales de hipoglucemias

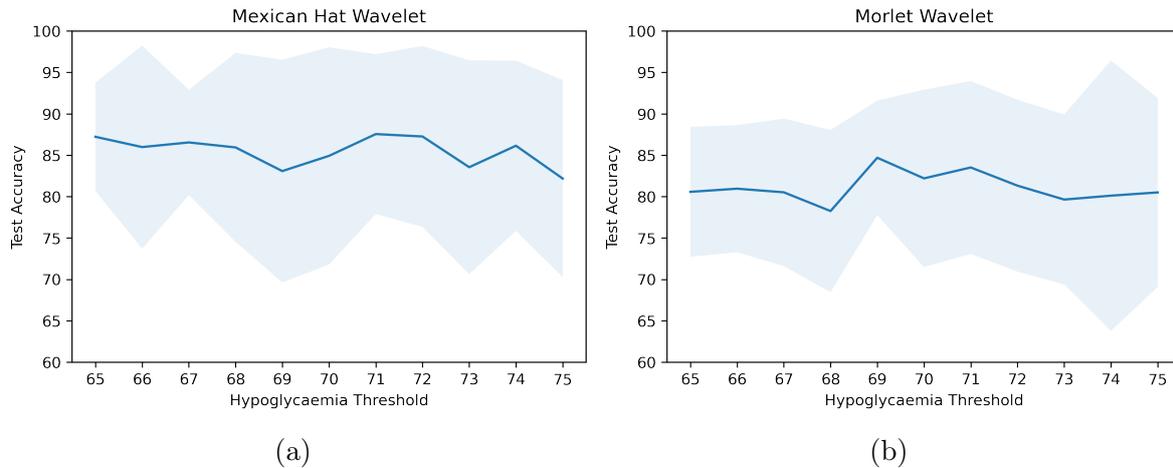


Figura 5.17: Análisis de Incertidumbre. Gráficos de la media (línea continua) y la desviación estándar (sombra azul) de la precisión en test de los modelos personales en diferentes umbrales de hipoglucemia entre 65 y 75 mg/dL .

ya introduce cierto nivel de ruido e incertidumbre debido a factores como la precisión del sensor y las variaciones fisiológicas, como se documenta en la bibliografía existente [106, 107, 108]. Sin embargo, el análisis de sensibilidad se centró en la introducción de una capa de ruido adicional y controlada para evaluar el rendimiento del método con distintos niveles de ruido, más allá de los presentes de forma natural en los datos MCG. Esto permitió comprender mejor cómo responde la técnica propuesta a distintos niveles de ruido y su capacidad para realizar predicciones precisas en escenarios reales. Una vez obtenidas las series temporales con ruido, se incorporan al conjunto de entrenamiento. Las imágenes se generan utilizando la transformada wavelet del mismo modo que en la Sección 5.4.

La Figura 5.18 muestra el análisis de sensibilidad para las diferentes métricas (Exactitud, Sensibilidad, Especificidad y Precisión). Puede observarse que al aplicar ruido, los valores de las diferentes métricas aumentan debido al incremento del número de imágenes de hipoglucemias generadas, es decir, se ha aumentado el tamaño del conjunto de datos como resultado de la aplicación de la técnica de DA antes descrita (véase Sección 5.3.2). A partir de un nivel de ruido del 10%, se produce un degradamiento en las diferentes métricas debido a que se generan más imágenes de la clase de *no hipoglucemia* porque hay menos probabilidad de generar una imagen que esté dentro del rango de hipoglucemia (40 – 70 mg/dL) que del rango de no hipoglucemia (> 70 mg/dL).

5.6. Conclusiones

En esta segunda parte de este trabajo, hemos propuesto el uso de una técnica novedosa para la detección de hipoglucemias e hiperglucemias. Este método combina la aplicación de la transformada de wavelet a las series de datos y la creación de modelos predictivos basados en

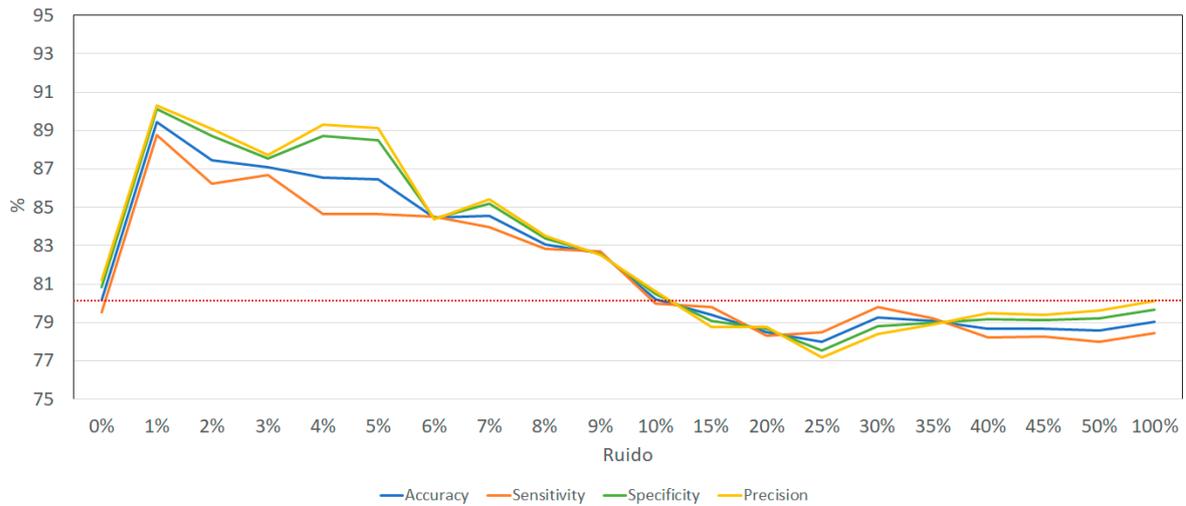


Figura 5.18: Análisis de Sensibilidad: la línea de puntos rojos marca la precisión del modelo entrenado en ausencia de ruido. El resto de líneas indican las diferentes métricas (Accuracy: Exactitud, Sensitivity: Sensibilidad, Specificity: Especificidad, Precision: Precisión) al añadir distintos porcentajes de ruido al conjunto de datos.

CNN. La transformada de wavelet es la encargada de transformar las series temporales de glucosa en imágenes o espectrogramas que plasman las escalas de frecuencias temporales de los valores de glucosa. A partir de estas imágenes, clasificadas dependiendo si hay hipoglucemia, hiperglucemia o normoglucemia, se procede al entrenamiento de modelos CNN. Una vez entrenados, éstos modelos permiten detectar éstos episodios glucémicos con hasta 24 horas de antelación.

Dado que el conjunto de datos original carecía de la dimensionalidad necesaria para generar modelos robustos, los primeros resultados que obtuvimos con esta técnica no fueron lo suficientemente aceptables. Por este motivo, hemos aplicado dos técnicas de aumentado de datos personalizadas. Gracias a estas dos técnicas de aumentado de datos, se han conseguido aumentar el conjunto de datos original en más de 24 veces, generándose de esta manera modelos bastantes precisos.

Los resultados muestran que esta técnica permite detectar correctamente las hipoglucemias con un umbral $< 70 \text{ mg/dL}$ con una precisión superior al 90%, pero no funciona del todo bien con las hiperglucemias. Tal como hemos explicado, para las hiperglucemias, el número generado de imágenes en algunos pacientes no es el suficiente para obtener modelos válidos. El deficiente control glucémico que han llevado algunos pacientes ha sido la principal causa de este problema. Además, los resultados muestran que los modelos personalizados por paciente (hipoglucemias) ofrecen mejores resultados a los modelos generales. Se afirma, de esta manera, que la personalización de los modelos para cada paciente mejora la detección de las hipoglucemias

Por último, se ha realizado un análisis de sensibilidad para determinar la cantidad de ruido que puede soportar el método propuesto. Tal como se ha explicado en la Sección 5.3.2, al aplicar

ruido a las series temporales, se produce una mejora de los modelos porque aumentamos el número de espectrogramas y, por tanto, la información utilizada para el entrenamiento de los modelos. Al aplicar más de un 10 % de ruido, la precisión disminuye, porque con estos niveles de ruido, se pierde información sobre los eventos de hipoglucemia, lo que significa que el modelo no dispone de la información requerida para ajustarse. Esta disminución se produce porque baja la probabilidad de que se genere una imagen de hipoglucemia al haber un umbral de valores más pequeño ($40 - 70 \text{ mg/dL}$).

Capítulo 6

Conclusiones

Estimar la glucosa es una tarea que realizan diariamente las personas con diabetes con el principal objetivo de mantener los niveles de BG dentro de un rango óptimo. Una buena estimación del valor de glucosa es importante para evitar los eventos de hipoglucemias e hiperglucemias, perjudiciales para la salud. Por este motivo, el objetivo de este trabajo de tesis ha sido la creación de modelos predictivos de glucosa bajo dos enfoques diferentes. Por un lado, la creación de modelos que estimen el valor exacto de la glucosa y por otro lado, el desarrollo de modelos que permitan una detección de episodios de hipoglucemias e hiperglucemias.

En el Capítulo 4 se ha desarrollado la técnica denominada METSK-HD basada en un sistema de reglas difusas TSK de dos fases para predecir el valor exacto de glucosa. Habitualmente estos sistemas se utilizan para aplicaciones de control automático, pero dada las particularidades de estos para encontrar la relación entre las variables de entrada y salida, se han adaptado perfectamente al problema de la predicción de la BG. Además, al tratarse de una técnica de caja blanca, los modelos obtenidos por estas técnicas son interpretables. Esto es posible, dado que estos modelos están compuestos por una serie de reglas difusas que pueden ser comprendidas por parte de un experto. En el mejor de los casos, se han obtenido modelos formados únicamente por dos reglas.

Tras el análisis de los resultados obtenidos por la técnica METSK-HD, se pueden comparar con los obtenidos con otras técnicas ML, pero no llegan a ofrecer un rendimiento del todo aceptable. Esto es debido a que este primer enfoque de este trabajo de tesis se puede considerar como un trabajo previo, porque para explotar todo el potencial del algoritmo METSK-HD, era necesario una optimización más profunda del algoritmo para adaptarlo a las particularidades del problema de la predicción de la glucosa. Esta optimización se puede considerar como una línea de investigación completamente nueva, no siendo uno de los objetivos de este trabajo de tesis. Por los motivos expresados anteriormente, desechamos la idea de seguir trabajando en este enfoque. Además, gracias a los avances producidos estos últimos años en el ámbito de la inteligencia arti-

ficial y el buen desempeño que están ofreciendo los modelos de aprendizaje profundo, pudiendo ser aplicados al problema de la predicción de la glucosa, decidimos hacer uso de modelos basados en NN en la segunda parte de este trabajo de tesis.

En el segundo enfoque, se ha creado una técnica completamente novedosa, detallada en el Capítulo 5, para la detección de eventos de hipoglucemias e hiperglucemias. Esta técnica combina la aplicación de transformaciones del espectro al campo de visión (transformada de wavelet) a la series de datos temporales y el uso de modelos predictivos basados en CNN. La transformada de wavelet nos ha permitido convertir las series temporales provenientes de los datos del sensor en imágenes válidas, que luego han sido usadas para entrenar modelos basados en CNN.

Además, se han desarrollado dos técnicas de aumentado de datos completamente innovadoras. Por un lado, la primera técnica DA está basada en una ventana deslizante que nos ha permitido aumentar en 24 veces el número de imágenes disponibles para entrenar el modelo. Por otro lado, se ha aplicado otra técnica DA consistente en aplicar ruido gaussiano a los valores de glucosa de las series temporales de glucosa de forma que se obtienen nuevas series, éstas con ruido. La idea de aplicar ruido es simular las posibles variaciones que se pueden producir en las lecturas debido al ruido introducido por el sensor y la incertidumbre producida por las variaciones fisiológicas. Con esta técnica, hemos podido duplicar el número de imágenes disponibles.

Los modelos desarrollados con esta técnica ofrecen, en el mejor de los casos, una precisión por encima del 90 %, siendo unos resultados bastantes prometedores al hablar de un horizonte de predicción de 24 horas. Además, los modelos propuestos ofrecen un rendimiento superior con respecto a las técnicas ML analizadas en este trabajo. Por este motivo, el análisis de estos resultados, nos invita a seguir mejorando la técnica propuesta en este trabajo de tesis.

Dado los buenos resultados obtenidos, el método desarrollado en el segundo enfoque de este trabajo consistente en la combinación de la transformada de wavelet con modelos CNN, puede aplicarse a cualquier otro problema que haga uso de series temporales con un comportamiento similar. Por ejemplo, en la vigilancia del medio ambiente, las series temporales de datos procedentes de sensores medioambientales, como las condiciones meteorológicas o los niveles de contaminación, podrían beneficiarse de nuestro enfoque para predecir e identificar acontecimientos extremos, como patrones meteorológicos severos o picos de contaminación. Por otro lado, el método presentado es adaptable a la previsión energética, facilitando la predicción de la demanda de energía y detectando pautas de consumo inusuales. Por otra parte, el sector financiero, incluidos los datos bursátiles, los tipos de cambio y los precios de las criptomonedas, exige el análisis de series temporales para la predicción y la detección de anomalías. Aunque las series temporales financieras suelen presentar menos periodicidad que los datos de glucosa en sangre, nuestro enfoque podría resultar valioso para predecir las fluctuaciones de los precios de las acciones y detectar anomalías financieras.

Como trabajo futuro, se debe focalizar en el ajuste y mejora de los modelos, en concreto en la parte de la detección de hiperglucemias. Por ejemplo, a la hora de realizar la clasificación

CAPÍTULO 6. CONCLUSIONES

de las imágenes, se debería tomar varios valores en lugar de un único valor para clasificar las imágenes. Otro punto de mejora, sería la creación de modelos que permitan detectar ambos tipos de episodios, es decir, que permitan alertar si se va a producir una hipoglucemia o hiperglucemia, en lugar de contar con dos tipos de modelos separados centrados en cada tipo de evento.

Por último, tal como hemos mencionado anteriormente, los modelos desarrollados en este trabajo de tesis pueden ser la base de una aplicación móvil que cuente con un sistema de alertas tempranas de hipoglucemias e hiperglucemias. Gracias a estas alertas, las personas con diabetes pueden anticiparse y tomar las acciones correctoras necesarias con el objetivo de evitar en la medida de lo posible, que sus niveles de glucosa alcancen niveles insalubres. De esta forma, se mejoraría la gestión de la diabetes y en consecuencia, la calidad de vida de estas personas.

Capítulo 7

Publicaciones e Indicadores de Calidad

Se presenta una relación de las publicaciones realizadas durante el desarrollo de este trabajo de tesis.

Revistas Indexadas

- Jorge Alvarado, J Manuel Velasco, Francisco Chavez, Francisco Fernández-de-Vega, J Ignacio Hidalgo. “Combining wavelet transform with convolutional neural networks for hypoglycemia events prediction from CGM data”. *Chemometrics and Intelligent Laboratory Systems*, 2023, vol. 243, p. 105017
DOI: <https://doi.org/10.1016/j.chemolab.2023.105017>
- Jorge Alvarado Díaz, José Manuel Velasco Cabo, Marta Botella Serrano, Esther Maqueda, Francisco Fernández de Vega, José Ignacio Hidalgo Pérez. “Blood glucose prediction using a two phase TSK Fuzzy Rule Based System”. *Diabetes Technology & Therapeutics*, 2022, vol. 24, pp. 231 - 231
DOI: <https://doi.org/10.1089/dia.2022.2525.abstracts>

Congresos Internacionales

- Marina De La Cruz, Carlos Cervigón, Jorge Alvarado, Marta Botella-Serrano, J Ignacio Hidalgo, “Evolving Classification Rules for Predicting Hypoglycemia Event”. *Proceedings of the IEEE Congress on Evolutionary Computation (CEC)*, pp. 1-8. Padua, Italia. IEEE.

2022

DOI: <https://doi.org/10.1109/CEC55065.2022.9870380>

- Jorge Alvarado, J Manuel Velasco, F Chávez, J Ignacio Hidalgo, F Fernández de Vega, "Blood Glucose Prediction Using a Two Phase TSK Fuzzy Rule Based System". *Proceedings of the IEEE Congress on Evolutionary Computation (CEC)*, pp. 712-719. Cracovia, Polonia. IEEE. 2021

DOI: <https://doi.org/10.1109/CEC45853.2021.9504992>

Congresos Nacionales

- J Alvarado, J Manuel Velasco, F Chávez, Almudena Sánchez, F Fernández de Vega, "Predicción de la glucosa en sangre mediante un sistema basado en reglas difusas TSK de dos fases". *Actas de la XIX Conferencia de la Asociación Española para la Inteligencia Artificial (CAEPIA 20/21)*, pp. 537-542. Málaga, España. CAEPIA. 2021

Pósters Científicos presentados en Congresos

- J. Alvarado Díaz, J. Manuel Velasco, M. Botella, E. Maqueda, F. Fernández de Vega, J.I. Hidalgo, "Blood glucose prediction using a two phase TSK Fuzzy Rule Based System". *15th International Conference on Advanced Technologies & Treatments for Diabetes. Barcelona, España. ATTD. 2022*

Otras Publicaciones

- Javier Bussons Gordo, Mario Fernández Ruiz, Manuel Prieto Mateo, Jorge Alvarado Díaz, Francisco Chávez de la O, J. Ignacio Hidalgo, Christian Monstein. "Automatic Burst Detection in Solar Radio Spectrograms Using Deep Learning: deARCE Method". *Solar Physics*, 2023, vol. 298, 82

DOI: <https://doi.org/10.1007/s11207-023-02171-0>

- Francisco Fernández De Vega, Jorge Alvarado, Abel Sánchez, Manuel Serrano, Elia Pacioni. "Evolutionary Algorithms: A new hope for the future of music teaching". *Proceedings of the Companion Conference on Genetic and Evolutionary Computation*, pp. 65-66. Lisboa, Portugal. ACM. 2023

DOI: <https://doi.org/10.1145/3583133.3596945>

- Josefa Díaz-Álvarez, Pedro A. Castillo, Francisco Fernández de Vega, Francisco Chávez, Jorge Alvarado. “Population size influence on the energy consumption of genetic programming”. *Measurement and Control*, 2022, vol. 55, issue 1-2, pp. 102-115
DOI: <https://doi.org/10.1177/00202940211064471>
- Francisco Fernández De Vega, Jorge Alvarado, Juan Villegas Cortez, “Optical Music recognition and Deep Learning: An application to 4-part harmony”. *Proceedings of the IEEE Congress on Evolutionary Computation (CEC)*, pp. 1-7. Padua, Italia. IEEE. 2022
DOI: <https://doi.org/10.1109/CEC55065.2022.9870357>
- Francisco Fernández de Vega, Jorge Alvarado Díaz, Martín Morita Hernández, “CAEPIA-App Competition: Sharpmony: A Computational Intelligence based tool for 4-part harmony”. *Actas de la XIX Conferencia de la Asociación Española para la Inteligencia Artificial (CAEPIA 20/21)*, pp. 1009-1012. Málaga, España. CAEPIA. 2021
- Jorge Alvarado Díaz, Francisco Chávez de la O, Pedro Ángel Castillo, Josefa Díaz Álvarez, Francisco Fernández de Vega, “En busca de algoritmos genéticos eficientes desde el punto de vista energético”. *Actas de la XIX Conferencia de la Asociación Española para la Inteligencia Artificial (CAEPIA 20/21)*, pp. 505-510. Málaga, España. CAEPIA. 2021

Patente

- José Ignacio Hidalgo Pérez, Javier Hidalgo García, Juan Lanchares Dávila, Jorge Alvarado Díaz, José Manuel Velasco Cabo, Óscar Garnica Alcázar, Francisco Fernández de Vega, Francisco Chávez de la O.
“Método y sistema para predicción de valores de glucosa y generación de alertas de hipoglucemia e hiperglucemia”.
Números de publicación: *ES2900099A1*, *WO2023135346A1*, *WO2023135346A9*
<http://invenes.oepm.es/InvenesWeb/detalle?referencia=P202230016>
<https://patentscope.wipo.int/search/es/detail.jsf?docId=W02023135346>

Bibliografía

- [1] Irl B Hirsch, Rattan Juneja, John M Beals, Caryl J Antalis, and Eugene E Wright Jr. The evolution of insulin and how it informs therapy and treatment choices. *Endocrine reviews*, 41(5):733–755, 2020. [cited at p. 8]
- [2] Paul C. Davidson, Harry R. Hebblewhite, Robert D. Steed, and Bruce W. Bode. Analysis of guidelines for basal-bolus insulin dosing: Basal insulin, correction factor, and carbohydrate-to-insulin ratio. *Endocrine Practice*, 14(9):1095–1101, 2008. [cited at p. 10]
- [3] Roy W Beck. Is it time to replace time-in-range with time-in-tight-range? maybe not. *Diabetes Technology and Therapeutics*, (ja). [cited at p. 11]
- [4] Stefano Passanisi, Claudia Piona, Giuseppina Salzano, Marco Marigliano, Bruno Bombaci, Anita Morandi, Angela Alibrandi, Claudio Maffei, and Fortunato Lombardo. Aiming for the best glycemic control beyond time in range: Time in tight range as a new cgm metric in children and adolescents with type 1 diabetes using different treatment modalities. *Diabetes Technology and Therapeutics*, (ja), 2023. [cited at p. 11]
- [5] Javier Castañeda, Arcelia Arrieta, Tim van den Heuvel, Tadej Battelino, and Ohad Cohen. Time in tight glucose range in type 1 diabetes: Predictive factors and achievable targets in real-world users of the minimed 780g system. *Diabetes Care*, page dc231581, 2023. [cited at p. 11]
- [6] William L Clarke, Daniel Cox, Linda A Gonder-Frederick, William Carter, and Stephen L Pohl. Evaluating clinical accuracy of systems for self-monitoring of blood glucose. *Diabetes care*, 10(5):622–628, 1987. [cited at p. 18]
- [7] Joan L Parkes, Stephen L Slatin, Scott Pardo, and Barry H Ginsberg. A new consensus error grid to evaluate the clinical significance of inaccuracies in the measurement of blood glucose. *Diabetes care*, 23(8):1143–1148, 2000. [cited at p. 18]
- [8] Javier Morales and Doron Schneider. Hypoglycemia. *The American journal of medicine*, 127(10):S17–S24, 2014. [cited at p. 19]

-
- [9] M. S. Pepe. *The Statistical Evaluation of Medical Tests for Classification and Prediction*. Oxford Statistical Sciences Series, 2003. [cited at p. 22]
- [10] Kelly H Zou, A James O'Malley, and Laura Mauri. Receiver-operating characteristic analysis for evaluating diagnostic tests and predictive models. *Circulation*, 115(5):654–657, February 2007. [cited at p. 22]
- [11] Dimitrina S. Dimitrova, Vladimir K. Kaishev, and Senren Tan. Computing the kolmogorov-smirnov distribution when the underlying cdf is purely discrete, mixed, or continuous. *Journal of Statistical Software*, 95(10):1–42, 2020. [cited at p. 22]
- [12] Md Riaz Ahmed Khan and Thomas Brandenburger. *ROCit: Performance Assessment of Binary Classifier with Visualization*, 2020. R package version 2.1.1. [cited at p. 22]
- [13] Cindy Marling and Razvan Bunescu. The ohiot1dm dataset for blood glucose level prediction: Update 2020. *CEUR workshop proceedings*, 2675:71–74, 09 2020. [cited at p. 26]
- [14] Michael Kahn. Diabetes. UCI Machine Learning Repository. DOI: <https://doi.org/10.24432/C5T59G>. [cited at p. 26]
- [15] Mary Martin, David Buchanan, Elizabeth Chun, Rucha Bhat, Shaun Cass, Eric Wang, Sangaman Senthil, and Irina Gaynanova. irinagain/Awesome-CGM: Updated release with simulators and enhanced processing, April 2021. [cited at p. 26]
- [16] Chiara Dalla Man, Francesco Micheletto, Dayu Lv, Marc Breton, Boris Kovatchev, and Claudio Cobelli. The uva/padova type 1 diabetes simulator: new features. *Journal of diabetes science and technology*, 8(1):26–34, 2014. [cited at p. 26]
- [17] ED Lehmann, T Deutsch, ER Carson, and PH Sönksen. Aida: an interactive diabetes advisor. *Computer methods and programs in biomedicine*, 41(3-4):183–203, 1994. [cited at p. 26]
- [18] J. D. Hoyos, M. F. Villa-Tamayo, C. E. Builes-Montaña, A. Ramirez-Rincón, J. L. Godoy, J. Garcia-Tirado, and P. S. Rivadeneira. Identifiability of control-oriented glucose-insulin linear models: Review and analysis. *IEEE Access*, 9:69173–69188, 2021. [cited at p. 26]
- [19] Roman Hovorka, Janet M Allen, Daniela Elleri, Ludovic J Chassin, Julie Harris, Dongyuan Xing, Craig Kollman, Tomas Hovorka, Anne Mette F Larsen, Marianna Nodale, Alessandra De Palma, Malgorzata E Wilinska, Carlo L Acerini, and David B Dunger. Manual closed-loop insulin delivery in children and adolescents with type 1 diabetes: a phase 2 randomised crossover trial. *The Lancet*, 375:743 – 751, 2010. [cited at p. 26]
- [20] B. Kovatchev, C. Cobelli, E. Renard, S. Anderson, M. Breton, S. Patek, W. Clarke, D. Bruttomesso, A. Maran, S. Costa, A. Avogaro, C. Dalla Man, A. Facchinetti, L. Magni, G. De Nicolao, J. Place, and A. Farret. Multinational study of subcutaneous model-predictive closed loop control in type 1 diabetes mellitus: Summary of the results. *Diabetes Sci Technol*, 4:1374–1381, 2010. [cited at p. 26]

BIBLIOGRAFÍA

- [21] Firas El-Khatib, Steven Russell, David Nathan, Robert Sutherlin, and Edward Damiano. A bihormonal closed-loop artificial pancreas for type 1 diabetes. *Science translational medicine*, 2:6–16, 04 2010. [cited at p. 26]
- [22] Lalo Magni, Marco Forgione, Chiara Toffanin, Chiara Dalla Man, Boris Kovatchev, Giuseppe De Nicolao, and Claudio Cobelli. Run-to-run tuning of model predictive control for type 1 diabetes subjects: In silico trial. *Journal of Diabetes Science and Technology*, 3(5):1091–1098, 2009. [cited at p. 26]
- [23] Lucy Mays. Diabetes mellitus standards of care. *Nursing Clinics of North America*, 50(4):703 – 711, 2015. [cited at p. 26]
- [24] Mohammad Reza Vahedi, Koenrad B MacBride, Woo Wunsik, Yosep Kim, Chris Fong, Andrew J Padilla, Mohammad Pourhomayoun, Alex Zhong, Sameer Kulkarni, Siddharth Arunachalam, et al. Predicting glucose levels in patients with type1 diabetes based on physiological and activity data. In *Proceedings of the 8th ACM MobiHoc 2018 Workshop on Pervasive Wireless Healthcare Workshop*, pages 1–5, 2018. [cited at p. 26]
- [25] Nichole S Tyler, Clara M Mosquera-Lopez, Leah M Wilson, Robert H Dodier, Deborah L Branigan, Virginia B Gabo, Florian H Guillot, Wade W Hilts, Joseph El Youssef, Jessica R Castle, et al. An artificial intelligence decision support system for the management of type 1 diabetes. *Nature metabolism*, 2(7):612–619, 2020. [cited at p. 26]
- [26] Ignacio Rodríguez-Rodríguez, Ioannis Chatzigiannakis, José-Víctor Rodríguez, Marianna Maranghi, Michele Gentili, and Miguel-Ángel Zamora-Izquierdo. Utility of big data in predicting short-term blood glucose levels in type 1 diabetes mellitus through machine learning techniques. *Sensors*, 19(20):4482, 2019. [cited at p. 26]
- [27] J Ignacio Hidalgo, J Manuel Colmenar, Gabriel Kronberger, Stephan M Winkler, Oscar Garnica, and Juan Lanchares. Data based prediction of blood glucose concentrations using evolutionary methods. *Journal of medical systems*, 41(9):1–20, 2017. [cited at p. 26]
- [28] S. Alexiou, E. Dritsas, O. Kocsis, K. Moustakas, and N. Fakotakis. An approach for personalized continuous glucose prediction with regression trees. 2021. cited By 1. [cited at p. 26]
- [29] Grazia Aleppo, Katrina J. Ruedy, Tonya D. Riddlesworth, Davida F. Kruger, Anne L. Peters, Irl Hirsch, Richard M. Bergenstal, Elena Toschi, Andrew J. Ahmann, Viral N. Shah, Michael R. Rickels, Bruce W. Bode, Athena Philis-Tsimikas, Rodica Pop-Busui, Henry Rodriguez, Emily Eyth, Anuj Bhargava, Craig Kollman, Roy W. Beck, and for the REPLACE-BG Study Group. REPLACE-BG: A Randomized Trial Comparing Continuous Glucose Monitoring With and Without Routine Blood Glucose Monitoring in Adults With Well-Controlled Type 1 Diabetes. *Diabetes Care*, 40(4):538–545, 02 2017. [cited at p. 26]
- [30] Sergio Contador, J Manuel Velasco, Oscar Garnica, and J Ignacio Hidalgo. Profiled glucose forecasting using genetic programming and clustering. In *Proceedings of the 35th Annual ACM Symposium on Applied Computing*, pages 529–536, 2020. [cited at p. 26]

-
- [31] Jose Ignacio Hidalgo, Jose Manuel Velasco, Juan Lanchares, Sergio Contador, and Oscar Garnica. *An analysis of solutions based on Genetic Programming to solve problems of symbolic regression of data from continuous glucose monitoring*. AEPIA, Granada, Spain, 2018. [cited at p. 26]
- [32] J. Ignacio Hidalgo, Esther Maqueda, Jose L. Risco-Martan, Alfredo Cuesta-Infante, J. Manuel Colmenar, and Javier Nobel. glucmodel: A monitoring and modeling system for chronic diseases applied to diabetes. *Journal of Biomedical Informatics*, 48:183 – 192, 2014. [cited at p. 26]
- [33] Mirko Messori, Chiara Toffanin, Simone Del Favero, Giuseppe De Nicolao, Claudio Cobelli, and Lalo Magni. Model individualization for artificial pancreas. *Computer Methods and Programs in Biomedicine*, 171:133–140, 2016. [cited at p. 27]
- [34] Christina-Maria Kastorini, George Papadakis, Haralampos J. Milionis, Kallirroï Kalantzi, Paolo-Emilio Puddu, Vassilios Nikolaou, Konstantinos N. Vemmos, John A. Goudevenos, and Demosthenes B. Panagiotakos. Comparative analysis of a-priori and a-posteriori dietary patterns using state-of-the-art classification algorithms: A case/case-control study. *Artificial Intelligence in Medicine*, 59(3):175–183, 2013. [cited at p. 27]
- [35] Ivanoe De Falco, Antonio Della Cioppa, Tomas Koutny, Michal Krcma, Umberto Scafuri, and Ernesto Tarantino. Genetic programming-based induction of a glucose-dynamics model for telemedicine. *Journal of Network and Computer Applications*, 119:1–13, 2018. [cited at p. 27]
- [36] Jose Velasco, Oscar Garnica, Juan Lanchares, Marta Botella, and Ignacio Hidalgo. Combining data augmentation, edas and grammatical evolution for blood glucose forecasting. *Memetic Computing*, 10, 06 2018. [cited at p. 27]
- [37] Daniel Parra, David Joedicke, J Manuel Velasco, Gabriel Kronberger, and J Ignacio Hidalgo. Learning difference equations with structured grammatical evolution for postprandial glycaemia prediction. *IEEE Journal of Biomedical and Health Informatics*, 2024. [cited at p. 27]
- [38] S. Mirshekarian, H. Shen, R. Bunescu, and C. Marling. Lstms and neural attention models for blood glucose prediction: Comparative experiments on real and synthetic data. pages 706–712, 2019. cited By 21. [cited at p. 27]
- [39] Kerstin Bach, Razvan C. Bunescu, Cindy Marling, and Nirmalie Wiratunga, editors. *Proceedings of the 5th International Workshop on Knowledge Discovery in Healthcare Data co-located with 24th European Conference on Artificial Intelligence, KDH@ECAI 2020, Santiago de Compostela, Spain & Virtually, August 29-30, 2020*, volume 2675 of *CEUR Workshop Proceedings*. CEUR-WS.org, 2020. [cited at p. 27]
- [40] M. Armandpour, B. Kidd, Y. Du, and J.Z. Huang. Deep personalized glucose level forecasting using attention-based recurrent neural networks. volume 2021-July, 2021. cited By 0. [cited at p. 27]

- [41] T. Zhu, L. Kuang, J. Daniels, P. Herrero, K. Li, and P. Georgiou. Iomt-enabled real-time blood glucose prediction with deep learning and edge computing. *IEEE Internet of Things Journal*, 2022. cited By 1. [cited at p. 28]
- [42] Federico D’Antoni, Lorenzo Petrosino, Fabiola Sgarro, Antonio Pagano, Luca Voller, Vincenzo Piemonte, and Mario Merone. Prediction of glucose concentration in children with type 1 diabetes using neural networks: An edge computing application. *Bioengineering*, 9(5):183, 2022. [cited at p. 28]
- [43] Ciro Rodriguez Leon, Oresti Banos, Oscar Fernandez Mora, Alex Martinez Bedmar, Fernando Rufo Jimenez, and Claudia Villalonga. Prediction of blood glucose levels in patients with type 1 diabetes via lstm neural networks. In *International Work-Conference on Artificial Neural Networks*, pages 563–573. Springer, 2023. [cited at p. 28]
- [44] Ciro Rodriguez-Leon, Maria Dolores Aviles-Perez, Oresti Banos, Miguel Quesada-Charneco, Pablo J Lopez-Ibarra Lozano, Claudia Villalonga, and Manuel Munoz-Torres. T1diabetesgranada: a longitudinal multi-modal dataset of type 1 diabetes mellitus. *Scientific Data*, 10(1):916, 2023. [cited at p. 28]
- [45] Felix Tena, Oscar Garnica, Juan Lanchares Davila, and J. Ignacio Hidalgo. An lstm-based neural network wearable system for blood glucose prediction in people with diabetes. *IEEE Journal of Biomedical and Health Informatics*, pages 1–12, 2023. [cited at p. 29]
- [46] S.L. Cichosz, T. Kronborg, M.H. Jensen, and O. Hejlesen. Penalty weighted glucose prediction models could lead to better clinically usage. *Computers in Biology and Medicine*, 138, 2021. cited By 0. [cited at p. 29]
- [47] Rebaz A. H. Karim, Istvan Vassanyi, and Istvan Kosa. After-meal blood glucose level prediction using an absorption model for neural network training. *COMPUTERS IN BIOLOGY AND MEDICINE*, 125, OCT 2020. [cited at p. 29]
- [48] Silvia Oviedo, Ivan Contreras, Carmen Quirós, Marga Giménez, Ignacio Conget, and Josep Vehi. Risk-based postprandial hypoglycemia forecasting using supervised learning. *International journal of medical informatics*, 126:1–8, 2019. [cited at p. 30]
- [49] Christopher Duckworth, Matthew J Guy, Anitha Kumaran, Aisling Ann O’Kane, Amid Ayobi, Adriane Chapman, Paul Marshall, and Michael Boniface. Explainable machine learning for real-time hypoglycemia and hyperglycemia prediction and personalized control recommendations. *Journal of Diabetes Science and Technology*, 18(1):113–123, 2024. [cited at p. 30]
- [50] Long Vu, Sarah Kefayati, Tsuyoshi Idé, Venkata Pavuluri, Gretchen Jackson, Lisa Latts, Yuxiang Zhong, Pratik Agrawal, and Yuan-chi Chang. Predicting nocturnal hypoglycemia from continuous glucose monitoring data with extended prediction horizon. In *AMIA Annual Symposium Proceedings*, volume 2019, page 874. American Medical Informatics Association, 2019. [cited at p. 30]

-
- [51] Darpit Dave, Daniel J DeSalvo, Balakrishna Haridas, Siripoom McKay, Akhil Shenoy, Chester J Koh, Mark Lawley, and Madhav Erraguntla. Feature-based machine learning model for real-time hypoglycemia prediction. *Journal of Diabetes Science and Technology*, 15(4):842–855, 2021. [cited at p. 31]
- [52] Wonju Seo, You-Bin Lee, Seunghyun Lee, Sang-Man Jin, and Sung-Min Park. A machine-learning approach to predict postprandial hypoglycemia. *BMC medical informatics and decision making*, 19(1):1–13, 2019. [cited at p. 31]
- [53] Tran Minh Quan, Takuyoshi Doike, Dang Cong Bui, Shigeki Arata, Atsuki Kobayashi, Md Zahidul Islam, Kiichi Niitsu, et al. Ai-based edge-intelligent hypoglycemia prediction system using alternate learning and inference method for blood glucose level data with low-periodicity. In *2019 IEEE international conference on artificial intelligence circuits and systems (AICAS)*, pages 201–206. IEEE, 2019. [cited at p. 31]
- [54] Mihaela Porumb, Saverio Stranges, Antonio Pescapè, and Leandro Pecchia. Precision medicine and artificial intelligence: a pilot study on deep learning for hypoglycemic events detection based on ecg. *Scientific reports*, 10(1):1–16, 2020. [cited at p. 31]
- [55] Sang-Min Lee, Dae-Yeon Kim, and Jiyoung Woo. Glucose transformer: Forecasting glucose level and events of hyperglycemia and hypoglycemia. *IEEE Journal of Biomedical and Health Informatics*, 27(3):1600–1611, 2023. [cited at p. 32]
- [56] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. Attention is all you need. *Advances in neural information processing systems*, 30, 2017. [cited at p. 32]
- [57] Anne Marie Delaney, Eoin Brophy, and Tomas E Ward. Synthesis of realistic ecg using generative adversarial networks. *arXiv preprint arXiv:1909.09150*, 2019. [cited at p. 32]
- [58] Virginie Felizardo, Nuno M Garcia, Nuno Pombo, and Imen Megdiche. Data-based algorithms and models using diabetics real data for blood glucose and hypoglycaemia prediction—a systematic literature review. *Artificial Intelligence in Medicine*, 118:102120, 2021. [cited at p. 32]
- [59] Chotirat Ann Ratanamahatana and Eamonn Keogh. *Making Time-series Classification More Accurate Using Learned Constraints*, pages 11–22. [cited at p. 32]
- [60] Che-Jui Hsu, Kuo-Si Huang, Chang-Biau Yang, and Yi-Pu Guo. Flexible dynamic time warping for time series classification. *Procedia Computer Science*, 51:2838–2842, 2015. International Conference On Computational Science, ICCS 2015. [cited at p. 32]
- [61] Xiaoyue Wang, Abdullah Al Mueen, Hui Ding, Goce Trajcevski, Peter Scheuermann, and Eamonn J. Keogh. Experimental comparison of representation methods and distance measures for time series data. *Data Mining and Knowledge Discovery*, 26:275–309, 2012. [cited at p. 32]

- [62] Lexiang Ye and Eamonn J. Keogh. Time series shapelets: a new primitive for data mining. In *KDD*, 2009. [cited at p. 32]
- [63] Jessica Lin, Eamonn Keogh, Stefano Lonardi, and Bill Chiu. A symbolic representation of time series, with implications for streaming algorithms. In *Proceedings of the 8th ACM SIGMOD Workshop on Research Issues in Data Mining and Knowledge Discovery*, DMKD '03, pages 2–11, New York, NY, USA, 2003. ACM. [cited at p. 33]
- [64] Cun Ji, Xiunan Zou, Yupeng Hu, and Shijun Liu. A 2d transform based distance function for time series classification. In *CollaborateCom*, 2018. [cited at p. 33]
- [65] G. E. P. Box and D. R. Cox. An analysis of transformations. *Journal of the Royal Statistical Society. Series B (Methodological)*, 26(2):pp. 211–252, 1964. [cited at p. 33]
- [66] Sachin P. Nanavati and Prasanta K. Panigrahi. Wavelet transform. *Resonance*, 9(3):50–64, Mar 2004. [cited at p. 33]
- [67] Philippe Masset. *Analysis of Financial Time Series Using Wavelet Methods*, pages 539–573. Springer New York, New York, NY, 2015. [cited at p. 33]
- [68] Reginald Bolman. and Thomas Boucher. Data mining using morlet wavelets for financial time series. In *Proceedings of the 8th International Conference on Data Science, Technology and Applications - DATA,,* pages 74–83. INSTICC, SciTePress, 2019. [cited at p. 33]
- [69] Paul Grant and Md Zahidul Islam. Signal classification using smooth coefficients of multiple wavelets. *ArXiv*, abs/2109.09988, 2021. [cited at p. 33]
- [70] Daoyuan Li, Tegawendé F. Bissyandé, Jacques Klein, and Yves Le Traon. Time series classification with discrete wavelet transformed data: Insights from an empirical study. In *SEKE*, 2016. [cited at p. 33]
- [71] Pengju Liu, Hongzhi Zhang, Wei Lian, and Wangmeng Zuo. Multi-level wavelet convolutional neural networks. *CoRR*, abs/1907.03128, 2019. [cited at p. 33]
- [72] Alejandro Pasos Ruiz, Michael Flynn, James Large, Matthew Middlehurst, and Anthony J. Bagnall. The great multivariate time series classification bake off: a review and experimental evaluation of recent algorithmic advances. *Data Min. Knowl. Discov.*, 35(2):401–449, 2021. [cited at p. 33]
- [73] Matthew Middlehurst, James Large, Michael Flynn, Jason Lines, Aaron Bostrom, and Anthony J. Bagnall. HIVE-COTE 2.0: a new meta ensemble for time series classification. *CoRR*, abs/2104.07551, 2021. [cited at p. 33]
- [74] Vajravelu Ashok and Nirmal Kumar. Determination of blood glucose concentration by using wavelet transform and neural networks. *Iranian journal of medical sciences*, 38(1):51, 2013. [cited at p. 33]
- [75] L.A. Zadeh. Fuzzy sets. *Information and Control*, 8(3):338–353, 1965. [cited at p. 35]

-
- [76] Francisco Herrera. Genetic fuzzy systems: taxonomy, current research trends and prospects. *Evolutionary Intelligence*, 1(1):27–46, 2008. [cited at p. 37, 40]
- [77] E.H. Mamdani and S. Assilian. An experiment in linguistic synthesis with a fuzzy logic controller. *International Journal of Man-Machine Studies*, 7(1):1–13, 1975. [cited at p. 37]
- [78] Tomohiro Takagi and Michio Sugeno. Fuzzy identification of systems and its applications to modeling and control. *IEEE transactions on systems, man, and cybernetics*, (1):116–132, 1985. [cited at p. 37]
- [79] María José Gacto, Marta Galende, Rafael Alcalá, and Francisco Herrera. Metsk-hde: A multiobjective evolutionary algorithm to learn accurate tsf-fuzzy systems in high-dimensional and large-scale regression problems. *Information Sciences*, 276:63–79, 2014. [cited at p. 39]
- [80] Carlos A Coello Coello, Gary B Lamont, David A Van Veldhuizen, et al. *Evolutionary algorithms for solving multi-objective problems*, volume 5. Springer, 2007. [cited at p. 40]
- [81] Kalyanmoy Deb. Multi-objective optimisation using evolutionary algorithms: an introduction. In *Multi-objective evolutionary optimisation for product design and manufacturing*, pages 3–34. Springer, 2011. [cited at p. 40]
- [82] L-X Wang and Jerry M Mendel. Generating fuzzy rules by learning from examples. *IEEE Transactions on systems, man, and cybernetics*, 22(6):1414–1427, 1992. [cited at p. 42]
- [83] Manuel Lozano, Francisco Herrera, Natalio Krasnogor, and Daniel Molina. Real-coded memetic algorithms with crossover hill-climbing. *Evolutionary computation*, 12(3):273–302, 2004. [cited at p. 43]
- [84] Larry J Eshelman. The chc adaptive search algorithm: How to have safe search when engaging in nontraditional genetic recombination. In *Foundations of genetic algorithms*, volume 1, pages 265–283. Elsevier, 1991. [cited at p. 43]
- [85] Rudolph Emil Kalman. A new approach to linear filtering and prediction problems. 1960. [cited at p. 44]
- [86] Harry Rubin-Falcone, Ian Fox, and Jenna Wiens. Deep residual time-series forecasting: Application to blood glucose prediction. In *Proceedings of the 5th International Workshop on Knowledge Discovery in Healthcare Data co-located with 24th European Conference on Artificial Intelligence, KDH@ ECAI*, pages 29–30, 2020. [cited at p. 49]
- [87] Hadia Hameed and Samantha Kleinberg. Investigating potentials and pitfalls of knowledge distillation across datasets for blood glucose forecasting. In *Proceedings of the 5th Annual Workshop on Knowledge Discovery in Healthcare Data*, 2020. [cited at p. 49]
- [88] Taiyu Zhu, Xi Yao, Kezhi Li, Pau Herrero, and Pantelis Georgiou. Blood glucose prediction for type 1 diabetes using generative adversarial networks. In *CEUR Workshop Proceedings*, volume 2675, pages 90–94, 2020. [cited at p. 49]

BIBLIOGRAFÍA

- [89] Tao Yang, Ruikun Wu, Rui Tao, Shuang Wen, Ning Ma, Yuhang Zhao, Xia Yu, and Hongru Li. Multi-scale long short-term memory network with multi-lag structure for blood glucose prediction. [cited at p. 49]
- [90] Robert Bevan and Frans Coenen. Experiments in non-personalized future blood glucose level prediction. In *CEUR Workshop Proceedings*, volume 2675, pages 100–104, 2020. [cited at p. 49]
- [91] Xiaoyu Sun, Mudassir Rashid, Mert Sevil, Nicole Hobbs, Rachel Brandt, Mohammad Reza Askari, Andrew Shahidehpour, and Ali Cinar. Prediction of blood glucose levels for people with type 1 diabetes using latent-variable-based model. [cited at p. 49]
- [92] David Joedicke, Oscar Garnica, Gabriel Kronberger, J Manuel Colmenar, Stephan Winkler, J Manuel Velasco, Sergio Contador, and J Ignacio Hidalgo. Analysis of the performance of genetic programming on the blood glucose level prediction challenge 2020. [cited at p. 49]
- [93] Ning Ma, Yuhang Zhao, Shuang Wen, Tao Yang, Ruikun Wu, Rui Tao, Xia Yu, and Hongru Li. Online blood glucose prediction using autoregressive moving average model with residual compensation network. [cited at p. 49]
- [94] 2020 The 5th International Workshop on Knowledge Discovery in Healthcare Data. 24th European conference in Artificial Intelligence (ECAI). Blood glucose level prediction challenge, 2020. [cited at p. 46]
- [95] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. Imagenet classification with deep convolutional neural networks. *Advances in neural information processing systems*, 25:1097–1105, 2012. [cited at p. 57, 68]
- [96] Gao Huang, Zhuang Liu, Laurens Van Der Maaten, and Kilian Q Weinberger. Densely connected convolutional networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 4700–4708, 2017. [cited at p. 58]
- [97] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei. ImageNet: A Large-Scale Hierarchical Image Database. In *CVPR09*, 2009. [cited at p. 58]
- [98] Ervin Sejdić, Igor Djurović, and Jin Jiang. Time–frequency feature representation using energy concentration: An overview of recent advances. *Digital Signal Processing*, 19(1):153–183, 2009. [cited at p. 61]
- [99] Xiangcheng Mi, Haibao Ren, Zisheng Ouyang, Wei Wei, and Keping Ma. The use of the mexican hat and the morlet wavelets for detection of ecological patterns. *Plant Ecology*, 179:1–19, 2004. [cited at p. 61, 63]
- [100] Martin A. Tanner and Wing H. Wong. From em to data augmentation: The emergence of mcmc bayesian computation in the 1980s. *Statistical Science*, 25(4):506–516, 2010. [cited at p. 63]

-
- [101] Jose Manuel Velasco, Oscar Garnica, Sergio Contador, Jose Manuel Colmenar, Esther Maqueda, Marta Botella, Juan Lanchares, and J. Ignacio Hidalgo. Enhancing grammatical evolution through data augmentation: Application to blood glucose forecasting. In Giovanni Squillero and Kevin Sim, editors, *Applications of Evolutionary Computation*, pages 142–157, Cham, 2017. Springer International Publishing. [cited at p. 63]
- [102] P Kingma Diederik. Adam: A method for stochastic optimization. (*No Title*), 2014. [cited at p. 68]
- [103] Merope Manataki, Antonis Vafidis, and Apostolos Sarris. Comparing adam and sgd optimizers to train alexnet for classifying gpr c-scans featuring ancient structures. In *2021 11th International Workshop on Advanced Ground Penetrating Radar (IWAGPR)*, pages 1–6. IEEE, 2021. [cited at p. 68]
- [104] Maneet Kaur Bohmrah and Harjot Kaur. Classification of covid-19 patients using efficient fine-tuned deep learning densenet model. *Global Transitions Proceedings*, 2(2):476–483, 2021. [cited at p. 69]
- [105] Faisal Mehmood, Shabir Ahmad, and Taeg Keun Whangbo. An efficient optimization technique for training deep neural networks. *Mathematics*, 11(6):1360, 2023. [cited at p. 69]
- [106] Xi Xie, Joshua C. Doloff, Volkan Yesilyurt, Atieh Sadraei, James J. McGarrigle, Mustafa Omami, Omid Veisheh, Shady Farah, Douglas Isa, Sofia Ghani, Ira Joshi, Arturo Vegas, Jie Li, Weiheng Wang, Andrew Bader, Hok Hei Tam, Jun Tao, Hui-jiuan Chen, Boru Yang, Katrina Ann Williamson, Jose Oberholzer, Robert Langer, and Daniel G. Anderson. Reduction of measurement noise in a continuous glucose monitor by coating the sensor with a zwitterionic polymer. *Nature Biomedical Engineering*, 2(12):894–906, Dec 2018. [cited at p. 79]
- [107] O. Garnica, J. Lanchares, J.M. Velasco, J.I. Hidalgo, and M. Botella. Noise spectral analysis and error estimation of continuous glucose monitors under real-life conditions of diabetes patients. *Biomedical Signal Processing and Control*, 61:101934, 2020. [cited at p. 79]
- [108] Hong Zhao and Chunhui Zhao. An automatic denoising method with estimation of noise level and detection of noise variability in continuous glucose monitoring. *IFAC-PapersOnLine*, 49(7):785–790, 2016. 11th IFAC Symposium on Dynamics and Control of Process Systems Including Biosystems DYCOPS-CAB 2016. [cited at p. 79]