

Experimentos con redes neuronales recurrentes LSTM para la predicción del nivel de glucosa de pacientes con diabetes

Anderson Mosquera Ruiz¹
Scotiabank
anderson.mosquera.ing@gmail.com

Fecha de recepción: 06 de julio de 2023
Fecha de aprobación: 23 de noviembre de 2023



Cómo citar este artículo: Mosquera Ruiz, A. (2023). Experimentos con redes neuronales recurrentes LSTM para la predicción del nivel de glucosa de pacientes con diabetes. *Revista Ontare*, 11, (páginas). DOI:

Resumen

La diabetes es una enfermedad en la cual el cuerpo no procesa de manera adecuada la glucosa; el tratamiento para esta enfermedad se basa en el autocuidado del paciente, sus tendencias dietarias, el ejercicio y la administración de insulina. Predecir los niveles de glucosa futuros puede ser de gran ayuda para que el paciente y el personal médico que lo atiende determinen estrategias que mantengan sus niveles de glucosa en un rango que no sea peligroso. Las técnicas de aprendizaje profundo, entre otras cosas, permiten predecir valores en una serie temporal. En la actualidad, la técnica más usada es la predicción mediante redes neuronales recurrentes tipo LSTM. Este artículo se propone realizar experimentos variando los parámetros de redes neuronales tipo LSTM para determinar si dichos parámetros tienen alguna influencia en la precisión de la predicción del modelo.

Palabras clave: redes neuronales; predicciones; LSTM; diabetes; manejo de la diabetes.

Experiments with LSTM recurrent neural networks for glucose level prediction in patients with diabetes

Abstract

Diabetes is a disease in which the body does not properly process glucose; Treatment for this disease is based on the patient's self-care, dietary tendencies, exercise, and insulin administration. Predicting future glucose levels can help the patient and their medical staff determine strategies to keep their glucose levels in a range that is not dangerous. Deep learning techniques, among other things, can predict values in a time series, currently the most widely used of these techniques is LSTM recurrent neural networks. This paper performs experiments varying the parameters of LSTM neural networks to determine whether these parameters have any influence on the accuracy of the model's prediction.

Keywords: Neural networks; predictions; LSTM; diabetes; diabetes management.

¹ Manager Analytics para Global Operations. Scotiabank. Ingeniero de Petróleos. Universidad Industrial de Santander. Ingeniero de Sistemas. Universidad Ean. ORCID: <https://orcid.org/0009-0005-0194-9378>

1. Introducción

A lo largo de los años se ha observado un incremento en el interés de la aplicación de técnicas de inteligencia artificial (AI, por sus siglas en inglés) para la detección, predicción y manejo de la diabetes, esto se evidencia en el incremento del número de revistas que entre el 2000 y el 2022 publicaron temas relacionados a la aplicación de AI al estudio de la diabetes (García-Jaramillo *et al.* 2023).

Se ha observado que el manejo de la diabetes es uno de los campos de la investigación que menos autores tiene y menos artículos produce año tras año, pero en las publicaciones encontradas por García-Jaramillo *et al.* (2023), se enfatiza el potencial del uso de técnicas de AI en el manejo de la diabetes, especialmente para la predicción de los niveles de glucosa futuros, con el fin de corregir tratamientos inefectivos, ajustar la dosis de insulina y generar alertas tempranas de eventos de hiper o hipoglucemia, que pueden llevar a cuadros de coma (Contreras y Vehi, 2018; Zainuddin *et al.*, 2009).

Algunas de las publicaciones donde se predice el nivel de glucosa de los pacientes se resaltan a continuación:

- Fiorini *et al.* (2017): estaban motivados por la idea de predecir el nivel de glucosa con métodos puramente basados en datos, y utilizaron regresiones y redes neuronales para conseguirlo.
- Zhu *et al.* (2020): usaron redes neuronales recurrentes dilatadas, en las cuales los parámetros de entrada fueron los niveles de glucosa en tiempo real, la ingesta de insulina, ingesta de comidas, género, edad, entre otros, para predecir con gran precisión los niveles de glucosa futuros de los pacientes.
- Martinsson *et al.* (2020): hicieron una red con neuronas tipo LSTM, donde variaron algunos de sus parámetros para determinar el punto óptimo donde pudieran predecir los

siguientes 30 minutos del nivel de glucosa, y encontraron que aumentar el número de neuronas LSTM disminuye el error asociado a la predicción.

- Alfian *et al.* (2020): usaron un *multilayer perceptron* acoplado con redes neuronales para predecir el nivel de glucosa 15, 30, 45 y hasta 60 minutos en el futuro.

El manejo de la diabetes se ha transformado por la integración de nuevas tecnologías, como el CGM, el desarrollo del páncreas artificial y la búsqueda de la explotación de la data adquirida por la aplicación de estas nuevas técnicas (Contreras y Vehi, 2018). Fagherazzi y Ravaud (2019), afirman que se necesitan medidas de evaluación de riesgo para saber en tiempo real qué tan propenso es un paciente de llegar a eventos de hipo o hiperglucemia, al igual que la industria financiera tiene modelos de riesgo crediticio que le permiten predecir qué tan riesgoso es que un cliente entre en periodos de mora. El objetivo de este artículo es determinar la influencia que pueden tener los parámetros de una red neuronal recurrente tipo LSTM sobre la precisión de la predicción del nivel de glucosa, para que estos hallazgos puedan ser usados por otros investigadores en el campo de la AI aplicada a la diabetes.

2. Marco teórico

2.1. La diabetes

La diabetes es una enfermedad crónica, en la que el cuerpo no disuelve la glucosa de manera apropiada (Joshi *et al.*, 2021). Existen 3 tipos bien diferenciados de diabetes: 1) diabetes tipo 1 (T1D, por sus siglas en inglés), en la que el cuerpo no genera insulina para la regulación de glucosa; 2) diabetes tipo 2 (T2D, por sus siglas en inglés), en la que el cuerpo genera pequeñas cantidades de insulina que no son suficientes para la regulación de la glucosa; 3) diabetes gestacional, en la que el cuerpo de la madre gestante no produce insulina (Joshi *et al.*, 2021).

La historia de la *diabetes mellitus* se remonta hasta el año 1500 a. C. en Egipto, a través de un papiro se describen sus síntomas característicos. En el siglo II d. C. en Grecia se usó por primera vez el término diabetes para describir una enfermedad en la que el cuerpo libera mucha orina y hay poca supervivencia (Holt *et al.*, 2017). En Europa, en 1776, Matthew Dobson ligó el acceso a la buena comida y el alto consumo de vinos con la diabetes, y concluyó que la sangre contenía azúcar y que esta se excretaba en la orina a través de los riñones (Holt *et al.*, 2017). Hacia inicios de 1800, John Rollo aplicó el adjetivo *mellitus* (de la palabra latina que significa *miel*) a la diabetes. hacia mediados del siglo XIX, Claude Bernand descubrió que la glucosa estaba presente en la sangre de todos los animales, aun en condiciones normales, y que el hígado contenía una sustancia similar al almidón que podía convertirse fácilmente en glucosa; según Holt *et al.* (2017), de estos descubrimientos derivó la teoría glucogénica, en la que enunciaba que el azúcar absorbido por el intestino se convertía por el hígado y luego era liberado constantemente en la sangre. Hacia finales del siglo XIX, Gustave Laguesse, apoyado en el trabajo de Paul Langerhans, sugirió que esta secreción interna era producida por islas de células, y en 1909, Jean de Meyer llamó a esta secreción interna insulina (derivado del vocablo latín para *isla*). Hacia finales del siglo XIX, Etienne Lancereaux identificó claramente dos grandes grupos de pacientes: unos con diabetes delgada y otros con diabetes grasa (Holt *et al.*, 2017). La diabetes delgada era más común en personas jóvenes y tenía un inicio agudo, mientras que la diabetes grasa se presentaba en personas de mediana edad con sobrepeso (Levy, 2016). Los trabajos y publicaciones de Harold Himsworth a mediados del siglo XX permitieron distinguir entre T1D (sensible a la insulina) y T2D (resistente a la insulina) (Levy, 2016).

Hoy se sabe que la homeostasis de la glucosa es un estado de equilibrio entre comidas, y se mantiene por varias hormonas interrelacionadas, especialmente la insulina, el glucagón, la adrenalina, el cortisol y las incretinas (Dunning, 2012). También juegan un papel importante el tipo de alimentos ingeridos, el ejercicio y los hábitos diarios (Fonseca, 2010). Se sabe que la insulina se secreta en 2 fases, la primera ocurre justo después de la ingestión de comidas y continúa durante los siguientes 10 a 15 minutos; en la segunda fase, la secreción de insulina se mantiene hasta que se restablecen los niveles normales de glucosa en la sangre (Dunning,

2012). Según Holt y Kumar (2015), en los pacientes con T1D la primera fase de secreción de insulina no es tan eficiente, con lo cual el paciente tiene altos niveles de glucosa en la sangre, que pueden llevar a trastornos tales como la aterosclerosis, hipertrigliceridemia, coagulopatías, disfunción endotelial, hipertensión y la hiperglucemia crónica (Dunning, 2012; Shaw y Cummings, 2012).

Un perfil glucémico es la representación gráfica de una serie de tiempo donde el eje y contiene los niveles de glucosa, y el eje x contiene el tiempo; de estos perfiles se pueden realizar observaciones para optimizar los niveles de glucosa a diferentes horas o momentos del día, dependiendo de la rutina del paciente y sus preferencias dietarias (Bando *et al.*, 2018). Para obtener un perfil glucémico se realizan mediciones continuas de los niveles de glucosa en la sangre, a través de dispositivos médicos tales como glucómetros y sistemas de monitorización continua de glucosa (CGM, por sus siglas en inglés) (Joshi *et al.*, 2021). Los primeros usan una tira radioactiva que requiere una gota de sangre para determinar el nivel de glucosa capilar. El sistema de monitoreo continuo requiere un sensor que puede tener una duración de 6 a 14 días, este se adhiere a la piel y toma muestras de glucosa intersticial cada 5 minutos (Mosquera y García-Jaramillo, 2022).

2.2. Series temporales y redes neuronales

Una serie temporal es una secuencia ordenada de valores registrados a intervalos iguales de tiempo (Blázquez-García *et al.*, 2021; Chou y Tran, 2018). Estas pueden tener una distribución interna que es aleatoria, o que responde a la relación de una o varias variables que determinan sus valores; los componentes que explican la distribución interna de una serie temporal son:

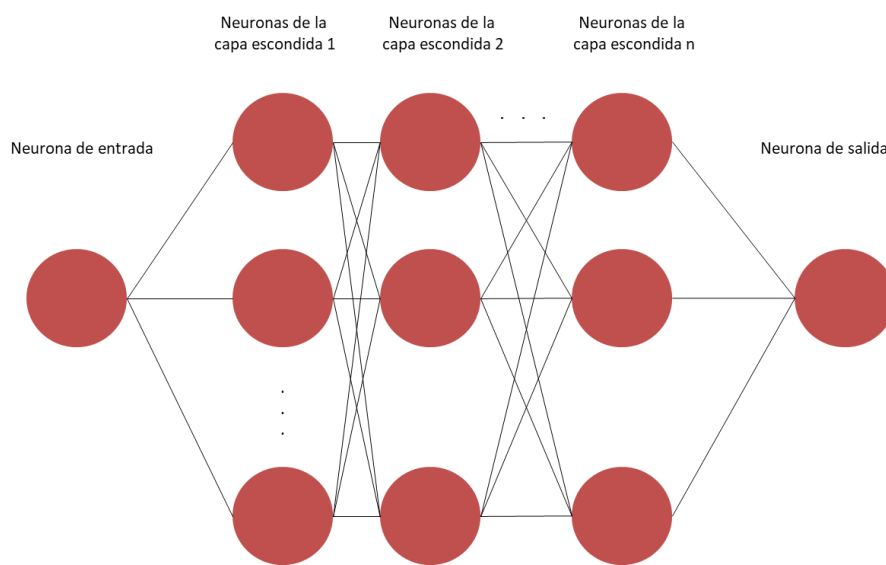
- Tendencia, que considera el patrón armónico de la serie.
- Estacionalidad, es el cambio cíclico de las variables internas de la serie temporal.
- Residual o ruido, que es la parte inexplicable de la serie temporal (Comito y Pizzuti, 2022).

Usando estos componentes se puede modelar matemáticamente el mecanismo estocástico que genera la serie temporal para generar pronósticos o predicciones (Athiyarath *et al.*, 2020).

Se ha encontrado que los datos de entrada, las variables y el horizonte de predicción afectan la precisión de los modelos (Sharma, 2018); por ejemplo, entre más largo el horizonte de predicción (qué tanto en el futuro se quiere predecir), se tienen errores de predicción más grandes (Oh *et al.*, 2015).

El aprendizaje profundo (*deep learning*) es un área del campo de la AI en la cual se utilizan datos, algoritmos y modelos estadísticos para extraer inferencias y patrones que permiten a las computadoras aprender y adaptarse automáticamente a nuevos estímulos (Comito y Pizzuti, 2022). Las redes neuronales recurrentes (RNN, por sus siglas en inglés) son un tipo de algoritmo de aprendizaje profundo, aptas para analizar datos secuenciales, están organizadas en capas o nodos, a través de los cuales se transmite la información de manera simultánea en pasos discretos de tiempo, en la figura 1 se puede apreciar la arquitectura de una red neuronal recurrente.

Figura 1. Estructura típica de una red neuronal recurrente



Fuente. Mosquera-Ruiz (2023).

Una variación de las redes neuronales recurrentes son las redes de larga duración de término corto (LSTM, por sus siglas en inglés), las cuales tienen nodos ocultos que permiten retener la salida de iteraciones anteriores, para así mejorar su precisión de predicción (Comito y Pizzuti, 2022); esta característica hace que las redes LSTM sean la técnica más comúnmente utilizada en la predicción de series temporales (Islam *et al.*, 2020; Lara-Benítez *et al.*, 2021).

Las redes LSTM fueron introducidas por Hochreiter y Schmidhuber en 1997 (Colah's Blog, 2015). Estas redes son adecuadas para procesar y predecir eventos importantes con intervalos muy largos en series de tiempo, ya que están diseñadas para evitar el problema de la dependencia a largo plazo, en el cual, la información provista hace mucho tiempo puede ser relevante para la predicción o generación de salidas (Colah's Blog, 2015; Kang *et al.*, 2023). Para lograr esto, las redes LSTM agregan un mecanismo de puertas para controlar la transmisión de información entre neuronas sobre la base de una red neuronal recursiva simple. Las puertas permiten que las transformaciones realizadas y la información agregada en cada neurona sea o no agregada a los datos que son transportados a neuronas siguientes en la red.

En una red LSTM, la información puede fluir entre cada neurona sin cambios, a menos que las puertas definan lo contrario. Existen tres puertas: 1) la puerta de entrada; 2) la puerta de olvido; 3) la puerta de salida. Estas puertas son funciones sigmoides con resultados de entre 0 y 1, donde 0 es un estado de “cierre completo” de la puerta (la información original de la red cambia) y 1 es un estado de “completa apertura” de la puerta (la información transmitida entre neuronas se mantiene) (Kang *et al.*, 2023). Según Colah's Blog (2015), el funcionamiento simplificado de una red LSTM normal es el siguiente:

1. La puerta del olvido determina si cada valor del estado anterior es relevante o no en el estado actual; si el resultado de la función sigmoide es 1 el valor original se mantiene completamente en el estado actual; si el resultado está entre 0 y 1 el valor cambiará proporcionalmente en el estado actual; y si el resultado es 0 el valor se olvidará y cambiará en el estado actual, en otras palabras, la puerta se cierra para ese valor.

2. La puerta de entrada determina cuáles valores del estado anterior se van a actualizar, y qué tanto deberían ser actualizados, también crea una serie de valores que se pueden agregar para producir el estado actual y una escala de preferencia para cada uno de ellos y, posteriormente, actualiza el estado anterior al estado actual con toda esa información.

3. Finalmente, la puerta de salida filtra los resultados de la puerta de entrada para dejar salir la parte del estado actual de la neurona que sea relevante para las neuronas siguientes de la red.

Algunas variaciones de LSTM permiten que cada puerta vea el estado de la neurona, dependiendo de los propósitos de la investigación. En otras variaciones la puerta del olvido y la puerta de entrada son una sola función sigmoide que cambia los valores que se olvidan directamente y en un solo paso (Colah's Blog, 2015).

De acuerdo con Amir (2022), algunos parámetros que pueden generar cambios en los resultados de una red neuronal recurrente son:

- Tasa de aprendizaje: indica la velocidad a la que se actualizan los coeficientes, puede ser fijo o variable.
- Número de épocas: define el número de iteraciones por las que se transmitirá el conjunto de datos de entrenamiento a la red, para que esta última ajuste gradualmente la salida.
- Número de neuronas: determina el número de neuronas conectadas en una red neuronal recurrente, a mayor número, más procesamiento de los datos de entrada para producir los datos de salida.
- Datos de entrada: son los datos de entrada al modelo; pueden ser todo el conjunto de datos o porciones del conjunto de datos que ingresan de manera secuencial, y sobre los cuales se genera la salida del modelo.
- Función de evaluación: determina cómo se evalúan las predicciones realizadas por el modelo sobre los datos de entrenamiento. Las redes neuronales buscan reducir medidas de error entre las predicciones, y los datos reales entre cada época de entrenamiento.
- Función de activación: determina si las transformaciones realizadas en una neurona pueden pasar a la siguiente o no. La más usada actualmente es Rectified Linear o ReLu por su versatilidad y facilidad de comprensión (Sharma, 2017).

3. Materiales y métodos

Se realizaron experimentos con redes neuronales de tipo LSTM para predecir el nivel de glucosa de los 5 minutos siguientes a la última medición de 5 pacientes con T1D, usando algunos de los datos obtenidos en un estudio previo realizado por Jaeb Center for Health Research (2015), cuyos resultados fueron analizados en un artículo publicado por Anderson *et al.* (2016), y cuyos datos están disponibles a través del centro coordinador del Consorcio de Páncreas Artificial de la JDRF, cuyo ID es NCT02137512.

Para realizar el análisis se desarrolló un código de Python versión 3.10 usando las librerías *numpy*, *pandas*, *matplotlib*, *datetime*, *timedelta* para importar y transformar los perfiles glucémicos de los pacientes con diabetes. Durante la importación y transformación de los datos se realizó la creación de un objeto tipo *dataframe* de la librería *pandas*, al cual se agregaron columnas adicionales a las originalmente provistas de cara a la inyección de estos datos a los experimentos con redes neuronales. Los niveles de glucosa nulos fueron reemplazados usando la interpolación lineal hacia adelante para cada paciente. Se eliminaron del *dataframe* las mediciones cuya distancia con respecto a la medición anterior fuera diferente a 5 minutos, para minimizar los errores y estandarizar los datos de entrada a los experimentos.

Python y sus librerías *sklearn* y *tensorflow*, *keras* y *pickle*, se utilizaron para la creación de una serie de redes neuronales recurrentes LSTM, en las cuales se variaron algunos parámetros para medir su influencia en la precisión de la predicción de resultados; los parámetros generales utilizados en la creación de las redes LSTM fueron los siguientes:

- Neuronas de entrada: 1.
- Número de capas: 1 de entrada, 1 de salida, 1 capa escondida tipo LSTM, 1 capa escondida tipo dense (en la cual todas las neuronas están conectadas).
- Función de activación de la segunda capa escondida: *rectified linear* (ReLU).
- Función de activación de la neurona de salida: *linear*.

- Se usó el 80 % de las mediciones por paciente como datos de entrenamiento y el resto se utilizó para pruebas.
- Función de error: media del error cuadrático entre los datos de prueba y los datos predichos sobre los datos de entrenamiento (RMSE, por sus siglas en inglés).

Para facilitar el procesamiento de datos y reducir las líneas de código necesarias, se crearon funciones con tres actividades principales: 1) entrenar las redes neuronales; 2) predecir los 5 minutos posteriores a la última medición; 3) guardar los datos reales y predichos en archivos de Excel. Los parámetros de las funciones eran el *dataframe* de cada paciente y los parámetros de red neuronal que se deseaba variar; las funciones se encapsularon en ciclos iterativos tipo *for* de varios niveles que recorrían todas las combinaciones de parámetros y pacientes ejecutando las 3 actividades descritas anteriormente.

Los parámetros y valores variados en la creación de los experimentos fueron:

- Tamaño de la ventana (capa de entrada) 5, 10, 20.
- Épocas de entrenamiento 3, 5, 7, 10.
- Neuronas de la primera capa escondida 32, 64, 128.
- Neuronas de la segunda capa escondida 4, 8, 16.
- Tasa de aprendizaje 0.1, 0.0001.

Al combinar los parámetros y valores se obtuvieron 216 experimentos. Como entrada a los modelos se utilizaron todos los datos para cada paciente, excepto los 3 últimos puntos de la serie temporal, con el fin de comparar los datos reales con el resultado predicho y que este estuviera en la tendencia de los datos reales. Se utilizó el valor absoluto de la variación porcentual entre el dato predicho por los modelos y el dato real provisto en los datos originales (APE, por sus siglas en inglés), para realizar un análisis de sensibilidad que permitiera determinar si cada parámetro tenía o no una influencia sobre la precisión de la predicción, y también para determinar el mejor modelo aplicable a cada paciente. Los resultados se describen en la siguiente sección.

4. Resultados y discusión

El modelo que obtuvo el menor APE por paciente se describe en la tabla 1.

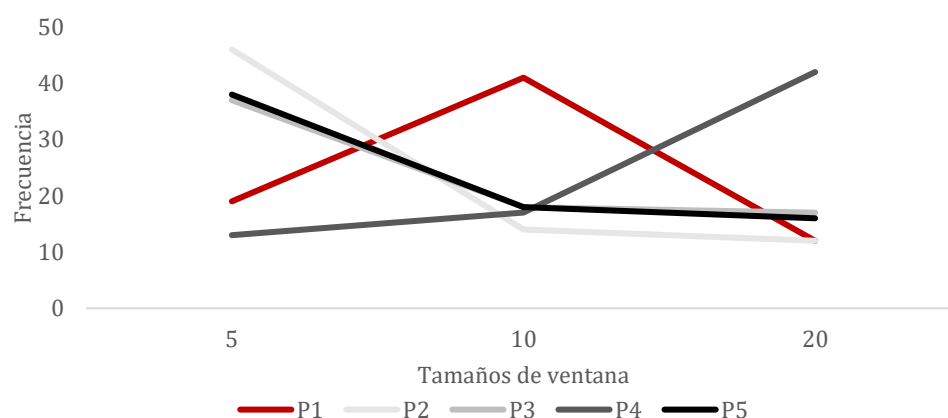
Tabla 1. Modelo con menor APE por paciente

Parámetro	P1	P2	P3	P4	P5
Tamaño de la ventana	5	5	10	20	5
Épocas de entrenamiento	3	10	3	7	3
Neuronas de la primera capa escondida	32	128	128	64	64
Neuronas de la segunda capa escondida	16	16	8	16	16
Tasa de aprendizaje	0.0001	0.1	0.0001	0.1	0.1

Fuente. Elaboración propia.

Se realizó un análisis de sensibilidad con los resultados obtenidos para determinar cuál valor de los parámetros generaba menores APE; en este análisis se observó cómo la variación de un parámetro, mientras los demás se mantenían constantes, afectaba la precisión del modelo en términos del APE. Se registró la frecuencia con la que cada parámetro obtenía el menor APE por paciente; en la figura 2 se puede apreciar el resultado para el tamaño de ventana.

Figura 2. Frecuencia en la que cada tamaño de ventana genera el menor APE por paciente



Fuente. Elaboración propia.

De la figura 2 se puede concluir que en el paciente 5 la relación entre el tamaño de ventana y el APE es directa, es decir que disminuir el tamaño de la ventana en el paciente 5 disminuye el APE, con lo cual mejora la precisión de la predicción del modelo. Se puede apreciar que, de manera consistente en los modelos del paciente 5, usar 5 como tamaño de ventana genera una mejor precisión en la predicción. En el paciente 1 se puede observar que no hay una relación inversa ni directa entre el tamaño de la ventana y la precisión del modelo; además, se puede observar que 10 es el tamaño de ventana que genera consistentemente los mejores resultados. En el paciente 4 se puede observar que la relación entre el tamaño de ventana y el APE es inversa, y que 20 es el tamaño de ventana que genera consistentemente los mejores resultados a través de los modelos de este paciente. Este mismo análisis se realizó por paciente para los demás parámetros utilizando herramientas computacionales y los resultados se describen a continuación.

En la tabla 2 se pueden observar de manera resumida los valores por parámetro que generan una mejor precisión en la predicción por paciente de manera consistente en los modelos.

Tabla 2. Valores que generan el menor APE por paciente

Parámetro	P1	P2	P3	P4	P5
Tamaño de la ventana	10	5	5	20	5
Épocas de entrenamiento	3	5	3	3	5
Neuronas de la primera capa escondida	64	128	32	128	128
Neuronas de la segunda capa escondida	16	8	4	16	16
Tasa de aprendizaje	0.1	0.1	0.1	0.1	0.1

Fuente. Elaboración propia.

Para cada paciente se obtuvieron resultados diferentes en el efecto que produce cada parámetro en la precisión del modelo, en la tabla 3 se muestran los resultados en función del APE.

Tabla 3. Efecto de los parámetros sobre el APE por paciente

Parámetro	P1	P2	P3	P4	P5
Tamaño de la ventana	-	Directa	Directa	Inversa	Directa
Épocas de entrenamiento	-	-	Directa	-	-
Neuronas de la primera capa escondida	-	Inversa	-	Inversa	Inversa
Neuronas de la segunda capa escondida	-	-	-	Inversa	Inversa
Tasa de aprendizaje	Inversa	Inversa	Inversa	Inversa	Inversa

Fuente. Elaboración propia.

De la tabla 3 se puede observar que una mayor tasa de aprendizaje genera una mayor precisión en la predicción para los cinco pacientes; de manera general, una mayor cantidad de neuronas en la primera capa escondida genera una mayor precisión en la predicción y un menor tamaño de ventana genera, al igual, una mayor precisión en la predicción. No se tienen suficientes resultados para definir el efecto general que tienen las épocas de entrenamiento y las neuronas en la segunda capa escondida.

5. Conclusiones

Este artículo realiza experimentos con algunos de los parámetros de redes neuronales tipo LSTM para determinar si tienen alguna influencia en la precisión de la predicción del nivel de glucosa de cinco pacientes con diabetes.

Los resultados experimentales muestran que algunos de los parámetros seleccionados sí tienen influencia sobre la precisión de la predicción, mientras que en los demás esta influencia no pudo ser evidenciada; se recomienda para futuros estudios hacer un análisis más exhaustivo, que incluya más pacientes y más valores para estos parámetros, que permitan determinar si tienen o no influencia en la precisión de la predicción.

Mediante los experimentos descritos en este artículo se pudo comprobar que una mayor tasa de aprendizaje mejora la precisión de la predicción de los modelos; también, se pudo comprobar que aumentar el número de neuronas tipo LSTM en la primera capa escondida de una red neuronal recurrente aumenta la precisión en la predicción y, finalmente, que de manera general disminuir el tamaño de ventana mejora la precisión de la predicción de las redes neuronales tipo LSTM.

Aviso: La fuente de los datos es el centro coordinador del Consorcio de Páncreas Artificial de la JDRF (JDRFAPPCC, por sus siglas en inglés), pero los análisis, el contenido y las conclusiones presentadas en este documento son responsabilidad exclusiva de los autores y no han sido revisados ni aprobados por la JDRFAPPCC.

6. Referencias

- Alfian, A., Syafrudin, M., Anshari, M., Benes, F., Dwi, F. T., Fahrurrozi, I., Hidayatullah A. F. & Rhee, J. (2020). Blood glucose prediction model for type 1 diabetes based on artificial neural network with time-domain features. *Biocybernetics and Biomedical Engineering*, 40(4), 1586-1599. <https://doi.org/10.1016/j.bbe.2020.10.004>
- Amir, T. (2022). Setup Neural Networks Hyperparameters for Best Results. *LinkedIn*. <https://www.linkedin.com/pulse/setup-neural-networks-hyperparameters-best-results-takbiri>
- Anderson, S. M., Raghinaru, D., Pinsker, J. E., Boscari, F., Renard, E., Buckingham, B. A., Nimri, R., Doyle, F. J., Brown, S. A., Keith-Hynes, P., Breton, M. D., Chernavvsky, D., Bevier, D. W., Bradley, P.
- Athiyarath, S., Paul, M. & Krishnaswamy, S. (2020). A comparative study and analysis of time series forecasting techniques. *SN Computer Science*, 1(175). <https://doi.org/10.1007/s42979-020-00180-5>

- Bruttomesso, D., Favero, S., Calore, R., Cobelli, C., Avogaro, A. & Farret, A. (2016). Multinational Home Use of Closed-Loop Control Is Safe and Effective. *Diabetes Care*, 39(7), 1143-1150. <https://doi.org/10.2337/dc15-2468>
- Bando, H., Ebe, K., Kato, Y., Bando, M. & Yonei, Y. (2018). Investigation of blood glucose profile by continuous glucose monitoring (CGM). *Endocrinology Research and Metabolism*, 2(1:10), 1-4. <https://www.imedpub.com/articles/investigation-of-blood-glucose-profile-by-continuous-glucose-monitoring-cgm.pdf>
- Blázquez-García, A., Conde, A., Mori, U. & Lozano, J. A. (2021). A review on outlier/anomaly detection in time series data. *ACM Computing Surveys*, 54(3), 1-33. <https://doi.org/10.1145/3444690>
- Chou, J. S. & Tran, D. S. (2018). Forecasting energy consumption time series using machine learning techniques based on usage patterns of residential householders. *Energy*, 165(part b), 709-726. <https://doi.org/10.1016/j.energy.2018.09.144>
- Colah's Blog. (2015). *Understanding LSTM Networks*. <https://colah.github.io/posts/2015-08-Understanding-LSTMs/>
- Comito, C. & Pizzuti, C. (2022). Artificial intelligence for forecasting and diagnosing COVID-19 pandemic: A focused review. *Artificial Intelligence in Medicine*, 128, 102286. <https://doi.org/10.1016/j.artmed.2022.102286>
- Contreras, I. & Vehi, J. (2018). Artificial intelligence for diabetes management and decision support: literature review. *Journal of Medical Internet Research*, 20(5), e10775. <https://doi.org/10.2196/10775>
- Dunning, T. (2012). *Diabetes education: Art, science and evidence*. John Wiley y Sons, Incorporated.
- Fagherazzi, G. & Ravaud, P. (2019). Digital diabetes: perspectives for diabetes prevention, management and research. *Diabetes & Metabolism*, 45(4), 322-329. <https://doi.org/10.1016/j.diabet.2018.08.012>

Fiorini, S., Martini, C., Malpassi, D., Cordera, R., Maggi, D., Verri, A. & Barla, A. (2017). Data-driven strategies for robust forecast of continuous glucose monitoring time-series. *39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Jeju, Korea (South), pp. 1680-1683 <https://doi.org/10.1109/EMBC.2017.8037164>

Fonseca, V. (2010). *Diabetes: improving patient care*. Oxford University Press, Incorporated.

García-Jaramillo, M., Luque, C. & León-Vargas, F. (2023). Machine learning and deep learning techniques applied to diabetes research: a bibliometric analysis. *Journal of Diabetes Science and Technology*, 18(2), 287-301. <https://doi.org/10.1177/19322968231215350>

Holt, R. I., Cockram, C., Flyvbjerg, A. & Goldstein, B. J. (2017). *Textbook of diabetes*. John Wiley y Sons, Incorporated.

Holt, T. & Kumar, S. (2015). *Abc of diabetes*. John Wiley y Sons, Incorporated.

Islam, S., Hossain, E., Rahman, A., Hossain, M. S. & Andersson, K. (2020). A review on recent advancements in FOREX currency prediction. *Algorithms*, 13(8), 186. <https://doi.org/10.3390/A13080186>

Jaeb Center for Health Research. (2015). *Pilot study 3 of outpatient control-to-range: safety and efficacy with day-and-night in-home use (CTR3)*. ID: NCT02137512. <https://www.clinicaltrials.gov/study/NCT02137512>

Joshi, A., Jain, P. & Mohanty, S. P. (2021). *Everything you wanted to know about continuous glucose monitoring*. *IEEE Consumer Electronics Magazine*, 10(6), 61-66. <https://ieeexplore.ieee.org/document/9415429>

Kang, Q., Chen, E., Li, C., Luo, H. & Liu, Y. (2023). Attention-based LSTM predictive model for the attitude and position of shield machine in tunneling. *Underground Space*, 13, 335-350. <https://doi.org/10.1016/j.undsp.2023.05.006>

- Lara-Benítez, P., Carranza-García, M. & Riquelme, J. C. (2021). An experimental review on deep learning architectures for time series forecasting. *International Journal of Neural Systems*, 31(3), 2130001. <https://doi.org/10.1142/S0129065721300011>
- Levy, D. (2016). *Type 1 diabetes*. (2nd ed.). Oxford University Press, Incorporated.
- Martinsson, J., Schliep, A., Eliasson, B. & Mogren, O. (2020). Blood glucose prediction with variance estimation using recurrent neural networks. *Journal of Healthcare Informatics Research*, 4, 1-18. <https://doi.org/10.1007/s41666-019-00059-y>
- Mosquera-Ruiz, A. y García-Jaramillo, M. (2022, septiembre 13-16). Diseño de un prototipo para la generación de alarmas a partir del análisis de tendencias del perfil glucémico de un grupo de pacientes con diabetes. *Nuevas realidades para la educación en ingeniería: Currículo, tecnología, medio ambiente y desarrollo*, Cartagena, Colombia, ACOFI EIEI. <https://acofipapers.org/index.php/eiei/article/view/2246/1880>
- Mosquera-Ruiz, A. (2023). *Diseño de un sistema para la visualización del nivel de glucosa de un grupo de pacientes con diabetes*. [Trabajo de grado]. Universidad EAN. <http://hdl.handle.net/10882/12769>
- Oh, S., Byon, Y. J., Jang, K. & Yeo, H. (2015). Short-term travel-time prediction on highway: a review of the data-driven approach. *Transport Reviews*, 35(1), 4-32. <https://doi.org/10.1080/01441647.2014.992496>
- Sharma, E. (2018). Energy forecasting based on predictive data mining techniques in smart energy grids. *Energy Informatics*, 1(suppl. 1):4, 367-428. <https://doi.org/10.1186/s42162-018-0048-9>
- Sharma, S. (2017). *Activation Functions in Neural Networks*. Towards Data Science. <https://towardsdatascience.com/activation-functions-neural-networks-1cbd9f8d91d6>
- Shaw, K. M. & Cummings, M. H. (2012). *Diabetes: chronic complications*. (2nd, ed.). John Wiley y Sons, Incorporated.

Zainuddin, Z., Ong, P. & Cemal, A. (2009). A neural network approach in predicting the blood glucose level for diabetic patients. *International Journal of Computational Intelligence*, 5(1).
https://www.researchgate.net/publication/242595444_A_Neural_Network_Approach_in_Predicting_the_Blood_Glucose_Level_for_Diabetic_Patients

Zhu, T., Li, K., Chen, J., Herrero, P. & Georgiu, P. (2020). Dilated recurrent neural networks for glucose forecasting in type 1 diabetes. *Journal of Healthcare Informatics Research*, 4, 308-324.
<https://doi.org/10.1007/s41666-020-00068-2>