

Estrategias de Entrenamiento de un Modelo de Predicción de Glucemia Basado en Redes Neuronales

C. Pérez-Gandía^{1,2}, G. García-Sáez^{2,1}, A. Rodríguez-Herrero^{2,1}, E.J. Gómez^{2,1}, M.E. Hernando^{2,1}

¹ Centro de Investigación en Red en Bioingeniería, Biomateriales y Nanomedicina (CIBER-BBN), Zaragoza, España

² Grupo de Bioingeniería y Telemedicina, Universidad Politécnica de Madrid, España

{cperez, ggarcía, arodri, [egomez](mailto:egomez@gbt.tfo.upm.es), elena.}@gbt.tfo.upm.es

Resumen

El conocimiento en las redes neuronales se incorpora mediante un proceso de aprendizaje a partir de ejemplos. Muchas veces resulta difícil obtener los registros necesarios para construir el conjunto de entrenamiento más adecuado, especialmente en aplicaciones médicas, cuando los experimentos a realizar para obtener los datos se ven limitados por razones éticas. En estos casos, sería muy positivo que los datos registrados con un grupo de pacientes pudieran utilizarse para entrenar un modelo universal, aplicable a cualquier otro paciente. En este trabajo comparamos la precisión de un modelo de predicción de glucemia basado en redes neuronales, utilizando dos estrategias de entrenamiento: personalizado vs. universal. Dicha comparativa se ha llevado a cabo con datos de nueve pacientes monitorizados con el sensor de medidas continuas de glucosa Guardian[®] Real-Time.

1. Introducción

Las redes neuronales artificiales están especialmente recomendadas para el modelado de magnitudes no lineales, problemas difíciles de describir mediante ecuaciones matemáticas y cuando las variables implicadas o las relaciones entre ellas pueden evolucionar a lo largo del tiempo. Todos estos rasgos son característicos del complejo equilibrio que permite un adecuado control de la glucosa en pacientes diabéticos. Por este motivo, modelos basados en redes neuronales han sido aplicados para la predicción y el control del perfil glucémico.

Los primeros modelos de predicción de glucemia [1-3] tenían como entradas la medida capilar, las dosis de insulina administradas y la ingesta. Debido a la escasa frecuencia en la medida de glucosa capilar (3-4 muestras al día), estos modelos solían considerar otra información relevante, como el ejercicio físico realizado, episodios de hipoglucemia, estrés, enfermedades, etc. para lograr una aproximación más precisa. Actualmente, gracias a los avances llevados a cabo en el campo de los sensores de medida continua, es posible disponer de la glucosa subcutánea cada pocos minutos. Se abre un nuevo escenario en el cual la medida de glucosa se convierte en una serie temporal. Las redes neuronales han sido utilizadas también en este nuevo escenario.

Pappada et al. [4] presentaron un modelo basado en redes neuronales entrenado con datos registrados mediante el sensor de medidas continuas CGMS[®] (Medtronic-

Minimed) en combinación con otros datos de interés registrados de forma manual por el paciente (glucosa capilar, dosis de insulina, ingesta de carbohidratos, síntomas de hipo/hiper glucemia, actividad diaria, etc.).

Nuestro grupo ha desarrollado así mismo un modelo basado en redes neuronales para predicción en tiempo real. Dicho modelo ha sido validado en base a los datos registrados por el sensor de medidas continuas Guardian[®] Real-Time (Medtronic-Minimed), con resultados muy positivos incluso para amplios horizontes de predicción (hasta 45 minutos) [5].

El problema que suelen presentar los modelos basados en redes neuronales es que en ellos el conocimiento se incorpora mediante un proceso de aprendizaje a partir de la experiencia. Incluso trabajando con medida continua de glucosa, los datos necesarios para el correcto entrenamiento del modelo a veces no son suficientes o están incompletos. Cada paciente necesitaría su propio conjunto de entrenamiento, ya que el comportamiento de cada organismo es diferente. Por tanto la puesta a punto del modelo para un paciente supondría un trabajo previo de adquisición de datos, que puede llevar semanas si se desea que en el conjunto de entrenamiento estén representadas todas las situaciones. Sin embargo resulta difícil registrar todos los casos posibles, y más aún las situaciones extremas, ya que los experimentos a realizar para obtener ese conjunto de entrenamiento óptimo, se ven limitados por razones éticas.

Otra aproximación al problema de la predicción de glucemia utilizando medida continua, son los modelos basados en análisis de series temporales. Estos modelos también adquieren el conocimiento a partir de la experiencia, aunque presentan una ventaja frente a los modelos basados en redes neuronales: suelen requerir una cantidad significativamente más pequeña de datos para su aprendizaje.

Sparacino et al. [6] y Reifman et al. [7] implementaron sendos modelos autorregresivos (ARM) cuya única entrada era la medida de glucosa registrada por un sensor de medidas continuas: el GlucoDay[®] CGM (A. Menarini Diagnostics) en el primer caso y el iSense[®] CGM (iSense Corporation) en el segundo.

Palerm et al. [8] propusieron un algoritmo de predicción basado en la estimación del nivel de glucosa y su velocidad de cambio, mediante la utilización de un filtro de Kalman. Aplicaron dicho modelo para predicción de

episodios de hipoglucemia, basándose en series de hipoglucemias registradas con un sensor de medidas continuas CGMS[®] (Medtronic-Minimed).

Un estudio comparativo [9] entre nuestro modelo basado en redes neuronales, NNM [5] y uno de los modelos basado en series temporales con mejores resultados, ARM [6], demuestra que el primero presenta una mayor precisión que el ARM, sin un deterioro significativo en el retardo del modelo. La diferencia en la precisión de ambos modelos se incrementa para amplios horizontes de predicción (30-45 minutos).

Dado que la experiencia de dicho estudio [9] indica que la precisión de los modelos basados en redes neuronales no solo es muy adecuada sino que a veces es mejor que la de modelos basados en técnicas de análisis de series temporales, sería muy importante superar la dificultad que presentan los primeros en cuanto a la necesidad de un conjunto de entrenamiento amplio y completo, que puede resultar de difícil adquisición. Una forma sencilla de salvar esa dificultad sería utilizar una estrategia de entrenamiento universal vs. el habitual entrenamiento personalizado.

Autores de modelos de predicción basados en series temporales han tenido presente esta preocupación por la obtención de un conjunto de entrenamiento amplio y completo, a pesar de tratarse de modelos menos limitados por el conjunto de datos de entrenamiento. En [7] se realiza un estudio comparativo de la precisión de un ARM cuando se entrena con los datos del propio paciente o con perfiles de otros pacientes.

En este trabajo se realiza una comparativa de las prestaciones del modelo desarrollado en nuestro grupo, NNM cuando se utilizan las dos estrategias de entrenamiento planteadas: universal vs. individualizado. Se entiende por entrenamiento universal la construcción de un conjunto de entrenamiento completo con los datos registrados para uno o varios pacientes, aplicable a cualquier paciente, haya o no participado de dicho entrenamiento.

Se demostrará que es posible el entrenamiento de un modelo universal, cuya precisión es muy similar a la de los modelos entrenados de forma individualizada, manteniendo el mismo número de perfiles en el conjunto de entrenamiento en ambos casos. Se demostrará así mismo que la precisión de dicho modelo universal no está sujeta a cuáles son concretamente los perfiles del conjunto de entrenamiento, para una dimensión fija del conjunto de entrenamiento.

2. Métodos

2.1. Adquisición de Datos

La evaluación ha sido realizada con los datos registrados en nueve pacientes con diabetes tipo 1. Los pacientes usaron el sensor de medidas continuas Guardian[®] Real-Time (Medtronic-Minimed), que proporciona medidas de la glucosa intersticial cada cinco minutos (288 muestras al día), durante 72 horas.

Se dispone de datos de 12 días por paciente: 6 de ellos abarcan días completos (24 horas) y los otros 6 son días no completos (falta el registro de alguna/s hora/s). Salvo por la longitud de las series de datos, los días completos e incompletos son equivalentes, no existe ninguna diferencia significativa entre ellos. Los primeros se han utilizado para entrenar los modelos por disponer de más muestras, y el resto para la validación de los mismos.

2.2. Descripción del Modelo de Predicción

La arquitectura del modelo inicial está descrita con detalle en [5], aunque se han realizado pequeños ajustes para mejorar la precisión y optimizar el tiempo de entrenamiento. Se trata de una red totalmente conectada, feedforward, de tres capas con 10, 5 y 1 nodo respectivamente. La función de transferencia de las neuronas es lineal.

La entrada de la red es la información de las 5 últimas muestras de glucosa (últimos 20 minutos) según el sensor de medidas continuas. La salida de la red es la predicción de glucosa para un horizonte de predicción (PH) determinado. Para comprobar la influencia del PH se ha evaluado la precisión de la red para tres casos: 15, 30 y 45 minutos.

Los parámetros de la red (pesos y constantes aditivas) se inicializan de forma aleatoria y posteriormente son optimizadas según un algoritmo de entrenamiento *backpropagation* de Levenberg Marquardt.

2.3. Metodología de Evaluación

Para poder comparar en igualdad de condiciones las dos estrategias de entrenamiento, se mantendrá fija la dimensión del conjunto de entrenamiento. Se han utilizado siempre 6 perfiles de días completos, por contener más muestras. El conjunto de validación para cada paciente ha sido siempre el mismo, no habiéndose utilizado nunca dichos perfiles durante el entrenamiento.

En el caso del entrenamiento individualizado, los perfiles de entrenamiento de cada una de las redes especializadas en cada paciente, fueron los 6 registros completos pertenecientes a dicho paciente. En adelante los modelos obtenidos con esta estrategia de entrenamiento serán referidos como INN-6 (6 es el número de perfiles utilizados durante el entrenamiento).

En el caso del entrenamiento generalizado, los perfiles del modelo universal fueron elegidos de forma aleatoria, de entre los 54 registros de días completos pertenecientes a cualquiera de los 9 pacientes. Se entrenaron 20 modelos en las mismas condiciones, para evitar que los resultados y las conclusiones extraídas sean fruto del azar. Dichos modelos serán referidos en adelante como UN-6.

Para caracterizar las prestaciones de los modelos y realizar la comparativa entre las dos estrategias de entrenamiento, se han calculado los parámetros empleados habitualmente en la bibliografía [6, 7, 9].

Para establecer la precisión del modelo se ha calculado el error cuadrático medio (RMSE) entre los perfiles originales y la predicción realizada por el modelo.

El retado medio (MD) se ha estimado como el valor medio de las diferencias de tiempo relativas al cruzar unos umbrales, entre el perfil original y la predicción. Los umbrales se situaron al 25%, 50% y 75% de la distancia entre máximos y mínimos para cada tramo creciente y decreciente del perfil original. Para facilitar el algoritmo de localización de tramos y umbrales, se ha aplicado un filtro paso bajo tanto al perfil original como a la predicción de cada uno de los modelos.

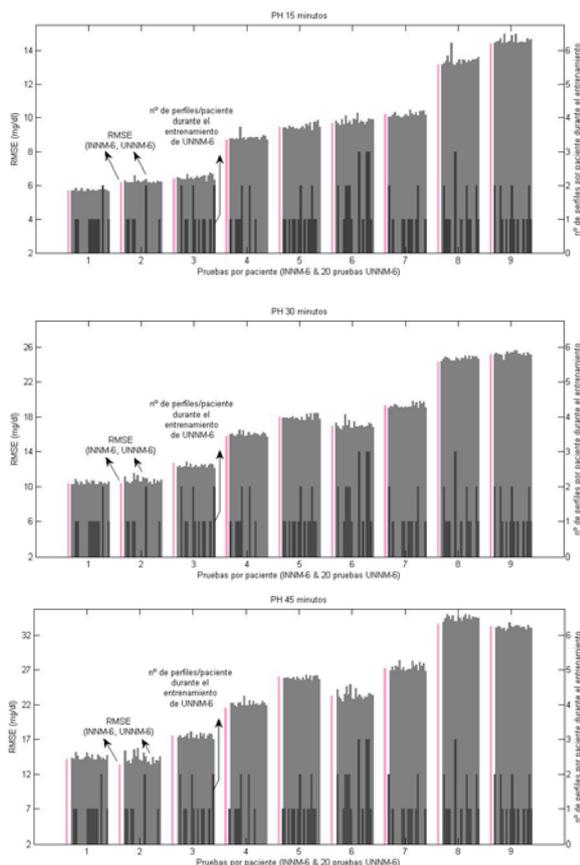


Figura 1. RMSE: INN-6 (magenta) y UNNM-6 (gris) vs. nº de perfiles de entrenamiento: UNNM-6 (gris oscuro)

3. Resultados

En la figura 1 se muestra el valor del RMSE obtenido en cada prueba (INN-6 y las 20 pruebas del UNNM-6) para los diferentes PH (15, 30 y 45 minutos). Para las pruebas realizadas con el modelo UNNM-6, se muestra también el número de perfiles de cada paciente que participó del entrenamiento durante cada prueba. (El nº de perfiles de cada paciente usados para entrenar el modelo INN-6, es fijo e igual a 6).

Las tablas 1 y 3 muestran respectivamente el RMSE y MD en los tramos crecientes (\nearrow) y decrecientes (\searrow) por paciente, obtenidos con el modelo INN-6, así como el valor medio y desviación típica de dichos parámetros, tras las 20 pruebas realizadas con los modelos UNNM-6.

La tabla 2 muestra el perfil del conjunto de validación (del 1 al 6) que obtuvo la precisión máxima y mínima, con cada una de las estrategias de entrenamiento (INN-6, UNNM-6) y para cada paciente. En el caso del modelo UNNM-6 se muestran los perfiles que presentaron mejor y peor precisión de manera más frecuente, durante las 20 pruebas realizadas, así como el número de veces que puntuaron como tales, si fue distinto de 20.

	15		30		45	
	INN-6	UNNM-6	INN-6	UNNM-6	INN-6	UNNM-6
#1	5.7	5.7 (+0.1)	10.4	10.5 (+0.2)	14.2	14.4 (+0.3)
#2	6.2	6.2 (+0.1)	10.4	10.7 (+0.3)	13.3	14.3 (+0.7)
#3	6.4	6.5 (+0.1)	12.7	12.4 (+0.2)	17.5	17.5 (+0.3)
#4	8.7	8.8 (+0.2)	15.8	16.0 (+0.2)	21.6	22.2 (+0.3)
#5	9.5	9.5 (+0.2)	18.1	18.0 (+0.2)	26.0	25.8 (+0.2)
#6	9.7	9.8 (+0.2)	17.0	17.1 (+0.4)	23.2	23.4 (+0.6)
#7	10.2	10.2 (+0.1)	19.3	19.3 (+0.3)	27.2	27.4 (+0.5)
#8	13.2	13.4 (+0.3)	24.3	24.7 (+0.2)	33.6	34.4 (+0.3)
#9	14.4	14.6 (+0.2)	25.2	25.2 (+0.2)	33.3	33.2 (+0.2)
Media	9.3	9.4 (+0.1)	17.0	17.1 (+0.1)	23.3	23.6 (+0.2)

Tabla 1. RMSE (mg/dl): INN-6 y UNNM-6

	15				30				45			
	INN-6		UNNM-6		INN-6		UNNM-6		INN-6		UNNM-6	
	Max	Min	Max	Min	Max	Min	Max	Min	Max	Min	Max	Min
#1	2	3	2	3	6	3	6	3(18)	6	3	6	3(19)
#2	1	3	1	3	1	3	1	3	2	3	2	3
#3	3	6	3	6	3	6	3	6	3	6	3	6
#4	5	3	4(12)	3	4	3	4	3	4	3	4	3
#5	6	4	6(17)	4(10)	3	5	3(18)	6(11)	2	5	3	5(15)
#6	2	5	2(19)	5	2	5	2(12)	5	3	5	4(16)	5
#7	6	3	6	3(16)	6	3	6	3	6	3	6	3
#8	5	4	5	4	5	4	5	4	2	3	5(11)	3(19)
#9	1	4	1	4	1	4	1(17)	4	2	4	2(13)	4(18)

Tabla 2. Perfil del conjunto de validación con la máxima/mínima precisión: INN-6 y UNNM-6

En la figura 2 se muestran los dos perfiles con valores extremos del RMSE para el paciente #2, con las diferentes estrategias de entrenamiento y un PH de 30.

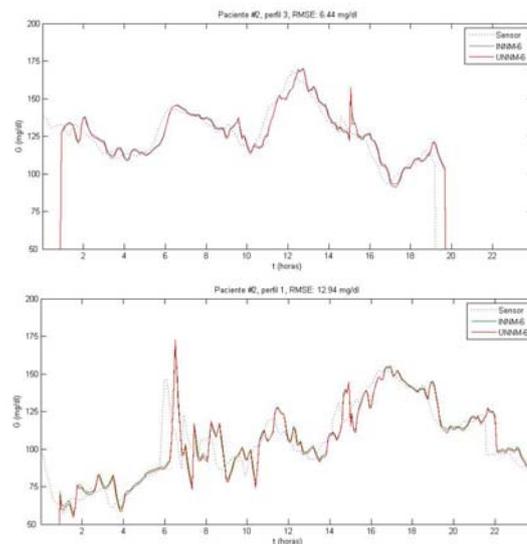


Figura 2. Perfiles con valores extremos de RMSE

4. Conclusiones

Todos los modelos entrenados con las dos estrategias presentan un RMSE muy similar (figura 1). Esto se confirma en la tabla 1 donde se observa para que: para el modelo UNNM-6, la desviación típica con respecto al

RMSE medio (para cada paciente), es muy pequeña y que el valor medio de dicho parámetro para las pruebas realizadas con la estrategia UNNM-6 es muy similar al valor del RMSE con el modelo INNM-6.

Todos los modelos entrenados con las dos estrategias presentan también un retardo muy similar. En la tabla 3 se observa que: para el modelo UNNM-6, la desviación típica con respecto al MD medio (para cada paciente), es muy pequeña y el valor medio del MD para las pruebas realizadas con la estrategia UNNM-6 es muy similar al valor de dicho parámetro con el modelo INNM-6.

En la figura 1 se observa que algunas de las pruebas para las que el conjunto de entrenamiento estaba formado por un número más elevado de perfiles de un paciente determinado, presentan un RMSE mayor, mientras que otras que presentan un menor RMSE no son necesariamente aquellas con más perfiles del paciente en el conjunto de entrenamiento. Por tanto, podemos concluir que las prestaciones del modelo de predicción no parecen depender del número de perfiles del propio paciente que participaron del entrenamiento.

Los resultados muestran que hay perfiles más sencillos y más complejos de predecir. En la tabla 2 se observa que para cada paciente los perfiles con valores de RMSE extremos son casi siempre los mismos, sea cual sea la estrategia de entrenamiento.

La figura 2 sugiere que la clave para que un perfil presente un RMSE pequeño parece estar en la variabilidad de los datos, tal como concluimos en un estudio previo [10], llevado a cabo con la estrategia de entrenamiento individualizado.

Por tanto, la estrategia de entrenamiento universal no solo es viable, sino que en ocasiones proporciona mejores prestaciones para el modelo de predicción, que la estrategia individualizada. Esto supone una gran ventaja ya que permite el entrenamiento de modelos universales, válidos para cualquier paciente aunque no se disponga de datos del mismo. También permite plantear la obtención de conjuntos de entrenamientos muy completos a partir de los datos de diferentes pacientes.

	15				30				45			
	INNM-6		UNNM-6		INNM-6		UNNM-6		INNM-6		UNNM-6	
	↗	↘	↗	↘	↗	↘	↗	↘	↗	↘	↗	↘
#1	5.4	6.4	5.3 (+1.1)	5.2 (+1.6)	15.0	14.7	14.2 (+2.2)	13.6 (+1.9)	23.3	23.0	23.2 (+4.0)	20.8 (+3.4)
#2	4.9	9.9	3.8 (+1.2)	8.6 (+1.0)	11.2	20.2	9.8 (+3.3)	20.1 (+1.4)	16.5	27.7	15.5 (+4.8)	25.1 (+5.5)
#3	5.2	3.4	6.3 (+1.0)	5.7 (+1.6)	15.9	12.0	14.8 (+2.9)	12.3 (+3.5)	23.5	17.7	21.3 (+5.3)	17.3 (+3.6)
#4	5.8	9.1	5.4 (+0.8)	6.6 (+1.0)	15.7	19.3	15.3 (+2.6)	16.9 (+2.5)	21.4	27.2	22.7 (+3.2)	20.5 (+4.6)
#5	7.2	9.8	6.2 (+0.6)	8.2 (+1.6)	17.2	18.3	16.3 (+1.6)	19.7 (+1.8)	19.2	15.0	25.2 (+2.6)	26.3 (+5.5)
#6	6.9	7.6	5.9 (+1.6)	6.0 (+1.2)	20.0	19.3	14.8 (+2.7)	16.2 (+2.0)	26.0	26.4	16.7 (+5.8)	23.8 (+1.6)
#7	6.4	3.6	6.7 (+1.0)	6.1 (+0.8)	18.2	10.9	15.3 (+2.9)	16.0 (+1.7)	9.4	14.2	19.8 (+6.5)	24.2 (+3.0)
#8	6.9	7.0	7.3 (+0.7)	5.2 (+1.2)	16.3	15.5	18.8 (+1.7)	12.2 (+3.0)	25.9	21.2	29.6 (+2.3)	14.6 (+5.3)
#9	8.3	4.8	8.2 (+1.1)	5.4 (+1.0)	17.5	15.5	18.6 (+1.7)	13.8 (+1.9)	24.2	17.8	28.6 (+3.6)	19.8 (+4.2)
Media	6.3	6.8	6.1 (+1.0)	6.3 (+1.2)	16.3	16.2	15.2 (+2.4)	15.6 (+2.2)	21.0	21.1	22.5 (+4.2)	21.4 (+4.1)

Tabla 3. MD (minutos): INNM-6 y UNNM-6

Referencias

- [1] Sandham WA, Nikolettou D, Hamilton DJ, Patterson K, Japp A, MacGregor C. Blood glucose prediction for diabetes therapy using a recurrent artificial neural network. *IX European Signal Processing Conference, Rhodes*, 1998, pp 673-676.
- [2] Mougiakakou SG, Prountzou K, Nikita KS. A real time simulation model of glucose-insulin metabolism for type 1 diabetes patients. *27th Annual International Conference of the Engineering in Medicine and Biology Society*, 2005, pp 298-301.
- [3] Tresp V, Briegel T, Moody J. Neural-network models for the blood glucose metabolism of a diabetic. *Neural Networks, IEEE Transactions on*, vol 10, 1999, pp 1204-1213.
- [4] Pappada SM, Cameron BD, Rosman PM. Development of a Neural Network for Prediction of Glucose Concentration in Type I Diabetes Patients. *J Diabetes Sci Technol*, vol 2, 2008, pp 792-801.
- [5] Pérez-Gandía C, Hernando ME, Sorzano COS, Otero-Quintana A, García-Sáez G, Rodríguez-Herrero A, Rigla M, de Leiva A, Gómez EJ. Neural networks for glucose prediction in continuous glucose monitoring. *1st Conference on Advanced Technologies & Treatments for Diabetes (ATTD), Prague*, 2008, pp 159.
- [6] Sparacino G, Zanderigo F, Corazza S, Maran A, Facchinetti A, Cobelli C. Glucose Concentration can be Predicted Ahead in Time From Continuous Glucose Monitoring Sensor Time-Series. *Biomedical Engineering, IEEE Transactions on*, vol 54, 2007, pp 931-937.
- [7] Reifman J, Rajaraman S, Gribok A, Ward WK. Predictive monitoring for improved management of glucose levels. *Journal of Diabetes Science and Technology*, vol 1, 2007, pp 478-486.
- [8] Palerm CC, Willis JP, Desemone J, Bequette BW. Hypoglycemia Prediction and Detection Using Optimal Estimation. *Diabetes Technology & Therapeutics*, vol 7, 2005, pp 3-14.
- [9] Pérez-Gandía C, Hernando ME, Facchinetti A, Sparacino G, Cobelli C, Gómez EJ. A methodology to compare prediction algorithms from continuous glucose monitoring data. *2nd Conference on Advanced Technologies & Treatments for Diabetes (ATTD), Athens*, 2009, pp 202.
- [10] Pérez-Gandía C, Hernando ME, García-Sáez G, Rodríguez-Herrero A, Rigla M, de Leiva A, Gómez EJ. Factors that influence the precision of a glucose predictor for type 1 diabetes based on neural networks. *2nd Conference on Advanced Technologies & Treatments for Diabetes (ATTD), Athens*, 2009, pp 203.