



## DETECCIÓN Y CLASIFICACIÓN DE TRIGLICÉRIDOS EN SANGRE CAPILAR BASADO EN REDES NEURONALES USANDO TENSOR FLOW

**Israel Rivera Zárate**

*Instituto Politécnico Nacional-CIDETEC*  
*irivera@ipn.mx*

**Miguel Hernández Bolaños**

*Instituto Politécnico Nacional-CIDETEC*  
*mbolanos@ipn.mx*

**Patricia Pérez Romero**

*Instituto Politécnico Nacional-CIDETEC*  
*promerop@ipn.mx*

### Resumen

*One of the main causes of death in the world and in Mexican population is smoking and its consequences, high levels of cholesterol, hypertension, obesity, physical inactivity or diabetes. The measurement to be carried out is supported by the absorption and transmittance properties of tissues and substances with respect to the incidence of light radiation according to the Lambert-Beer law. This work proposes the development of a portable system that performs the acquisition of the triglyceride level and digital processing classification based on deep neural networks using Tensor Flow language. The project aims to be an element of medical support in the care, treatment and monitoring of patients with cardiovascular disease.*

*Palabras clave: Triglycerides, neural network, tensor Flow.*

De acuerdo con Harris (2007), los hidratos de carbono y los lípidos se utilizan principalmente como fuente de energía. Sus formas de almacenamiento en el organismo son el glucógeno y el tejido adiposo, respectivamente. Un nivel alto de triglicéridos puede generar pancreatitis y llevar a cuadros aterogénicos y arterioesclerosis, lo cual incrementa el riesgo de infarto agudo a miocardio y accidentes cerebrovasculares. Como lo indica García Saavedra *et al.* (1999),

el perfil lipídico lo constituye la cuantificación analítica de una serie de lípidos que son transportados en la sangre por los diferentes tipos de lipoproteínas plasmáticas. La determinación de estos parámetros es un procedimiento analítico básico para el diagnóstico y seguimiento de enfermedades metabólicas, primarias o secundarias. Entre estos parámetros analíticos que se pueden determinar están: el colesterol total, el colesterol transportado por las LDL, el



colesterol transportado por las HDL, los triglicéridos totales, ciertas apoproteínas particulares en general.

200-499 mmol/L)	mg/dL	(2.26-5.63	Alto
>=500 mmol/L)	mg/dL	(>=5.65	Muy alto

## I. Determinación de triglicéridos

Existen diferentes métodos para la determinación de triglicéridos, según expresa González de Buitrago (1998) en suero se utilizan reactivos comerciales (kit comercial de Linear Chemicals) que incluyen las enzimas y sustratos necesarios para la cuantificación por espectrofotometría visible.

Lo que ocurre es lo siguiente:

1. Una lipoproteína lipasa (LPL) hidroliza los triglicéridos dando lugar a glicerol más ácidos grasos libres (FFA).

2. El glicerol formado es sustrato de un glicerol (GK) quinasa en presencia de adenosin trifosfato (ATP) al cual fosforila a glicerol 3P (G-3-P) y adenosin difosfato (ADP).

3. El glicerol G-3-P es oxidado por la glicerofosfato oxidasa (GPO) en dihidroxiacetona fosfato (DHAP) y peróxido de hidrógeno (H<sub>2</sub>O<sub>2</sub>).

4. En presencia de la peroxidasa (POD) el fenol y la 4-aminoantipirina (4-AA) se condensan por acción del peróxido de hidrógeno formándose un cromógeno rojo (quinona roja) proporcional a la concentración de los triglicéridos presentes en la muestra.

Los valores de referencia clínicos actualizados de triglicéridos empleados para clasificar los grupos de riesgo se ilustran en la tabla 1.

**Tabla 1. Valores de referencia.**

Triglicéridos	Clasificación
<150 mg/dL (<1.7 mmol/L)	Normal
150 mg/dL-199 mg/dL (1.7-2.25 mmol/L)	Medio/Alto

### A. Transmitancia

Como lo indica Hernández-Hernández *et al.* (2002), al hacer pasar un haz de radiación luminosa a través de una capa de solución con cierto grado de concentración, y que contiene una especie molecular que posee un coeficiente de absorción ante tal longitud de onda radiante, se observa, que como consecuencia de las interacciones entre los fotones y las partículas absorbentes, la potencia del haz disminuye de  $P_0$  a  $P$ . La transmitancia  $T$  de la solución es la fracción de la radiación incidente transmitida por la solución (por lo general, la transmitancia se expresa como porcentaje):

$$T = \frac{P}{P_0} \dots (1)$$

### B. Absorbancia

La absorbancia de una solución está definida por (2):

$$A = -\log_{10} T = -\log_{10} \frac{P}{P_0} \dots (2)$$

Se debe observar que, a diferencia de la transmitancia, la absorbancia de una solución crece a medida que aumenta la atenuación del haz.

### D. Ley de Lambert-Beer

Esta ley indica que, para una cantidad de absorbente, la intensidad de luz transmitida, que previamente se ha logrado que sea paralela y plana y que entre al medio absorbente, formado ángulos rectos con el plano, disminuye logarítmicamente a medida que la longitud del trayecto aumenta en forma aritmética. Ver figura 1.

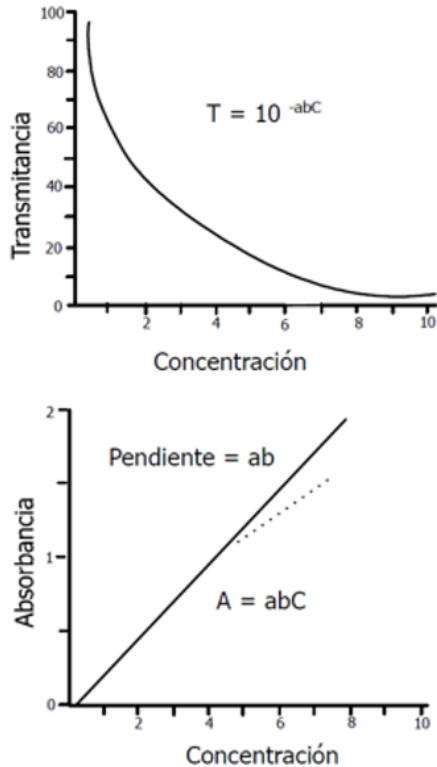


Figura 1. Rivera, et al (2021). Ley de Lambert-Beer.

Por lo tanto, se tiene que:

$$\frac{I_1}{I_0} = 10^{-\alpha l} = 10^{-\varepsilon l C} = 10^{-A} \dots (3)$$

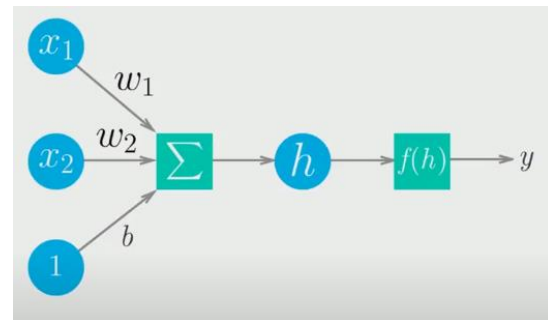
Donde:

- $I_1$  : Intensidad saliente.
- $I_0$  : Intensidad entrante.
- $A = \alpha l = \varepsilon l C$  : Absorbancia
- $\alpha$  : Coeficiente de absorción.
- $l$  : Longitud atravesada por la luz en el medio
- $C$  : Concentración del absorbente en el medio

## II. Red Neuronal

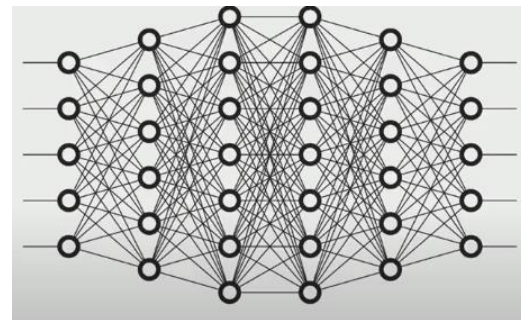
De acuerdo con Ian Goodfellow *et al.* (2016), una red neuronal (RNN) es una representación algorítmica de una regresión logística y pertenece a una rama del machine learning, están inspiradas en el cerebro humano. De manera general, el cerebro humano se conforma de muchas neuronas donde cada una de ellas se conecta a otras mediante una red. Si una neurona recibe varios estímulos eléctricos, decide si se activa o no. Al activarse, esta primera neurona envía un impulso a la siguiente neurona conectada a ella (Pesos sinápticos).

Esta comunicación define de este modo una serie de capas. La capa de entrada recibe todo el conjunto de datos iniciales con los que se desea operar. En seguida se tienen una o varias capas ocultas que son las que se encargan de hacer todo el aprendizaje de un algoritmo ya sea de clasificación o de reconocimiento de



patrones. Por último, se tiene la capa de salida, se tienen varias neuronas que mostraran el resultado buscado. Ver figura 2.

a)



b)

Figura 2. Rivera et al (2021). a) Neurona, b) Red neuronal.

Se observa que todas las neuronas se conectan punto a punto al resto de neuronas de la capa inmediata siguiente. Se cuenta con una función denominada de activación y que es la que define bajo que circunstancia una neurona comunica su salida. Ver ecuación 5.

$$Y = \text{softmax}(\text{sigmoide}(Wx + b)) \dots (5)$$

Donde Y es la salida obtenida en la última capa, la cual se alcanza luego de realizar una evaluación probabilística mediante la función softmax (comúnmente empleada en clasificación). En la etapa de entrada se reciben los datos por evaluar x, los pesos sinápticos se aprenden durante el entrenamiento W. Debido a que la RNN es una regresión logística que corresponde a una línea recta, b es el sesgo, el bias o la ordenada en el origen.

Como lo indica Alex Graves (2014), para entrenar una red neuronal se deben alimentar datos en la capa de entrada y evaluar el error de los resultados a la salida respecto del valor esperado. Se hace necesario repetir el proceso muchas veces (épocas). Para esto se requiere definir una función de error (típicamente error cuadrático medio) y un proceso de optimización que lleve el error a cero gradualmente para lo que se emplea una tasa de aprendizaje; lo que se conoce como algoritmo del descenso del gradiente. Ver figura 3.

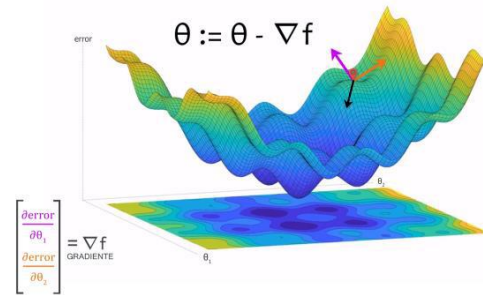


Figura 3. Rivera et al (2021). Descenso del gradiente.

De este modo se aprecia que en el proceso de aprendizaje de la red existe una etapa de entrenamiento basada en la reducción del error y otra posterior que lleva a cabo una evaluación o prueba con datos que la red no ha visto nunca; Como lo ha expresado Fei-Fei et al. (2017), el mecanismo de optimización es conocido como el algoritmo de backpropagation o de propagación hacia atrás del error que puede resumirse en los siguientes pasos:

1. Error de la última capa.

$$\delta^L = \frac{\partial C}{\partial \alpha^L} \cdot \frac{\partial \alpha^L}{\partial z^L}$$

2. Retropropagación del error a la capa anterior.

$$\delta^{L-1} = w^L \delta^L \cdot \frac{\partial \alpha^{L-1}}{\partial z^{L-1}}$$

3. Derivadas parciales respecto a los parámetros.

$$\begin{aligned} \frac{\partial C}{\partial b^{L-1}} &= \delta^{L-1} & \frac{\partial C}{\partial w^{L-1}} \\ & & = \delta^{L-1} \alpha^{L-2} \end{aligned}$$

Donde:



$$Z^L = W^L a^L + b^L . \quad \dots (4)$$

$$c(a_j^L) = \frac{1}{2} \sum_j (y_j - a_j^L)^2 \quad \dots (5)$$

Indica la salida de cada neurona Z en cada capa L y C es la función de error cuadrático medio.

### III. Metodología

La metodología adoptada en el presente artículo consiste en 3 pasos esenciales:

[Paso 1:] Preprocesamiento.

[Paso 2:] Transferencia inalámbrica.

[Paso 3:] Aprendizaje, detección y clasificación.

### Preprocesamiento

En esta etapa se establece el diseño de los circuitos que constituyen el medidor de concentración de triglicéridos, tomando como base las reacciones del método de hidrólisis enzimática de los triglicéridos. Donde, entre otros, en presencia de la peroxidasa (POD) el fenol y la 4-aminoantipirina (4-AA) se condensan por acción del peróxido de hidrógeno ( $H_2O_2$ ), formándose un cromógeno rojo (quinona roja) proporcional a la concentración de los triglicéridos presentes en la muestra. Lo cual se logra al medir satisfactoriamente a una longitud de onda de los 550 nm (longitud de la luz verde). Cabe señalar, que la intensidad de la luz se reducirá logarítmicamente con la longitud de la trayectoria conforme lo establece la ley de Beer-Lambert. Ver figura 4.

### Transferencia de la información

Se empleó un módulo HC-05 conectado a la tarjeta Arduino UNO para comunicación inalámbrica vía Bluetooth con un dispositivo móvil en formato de 8 bits de enteros sin signo.



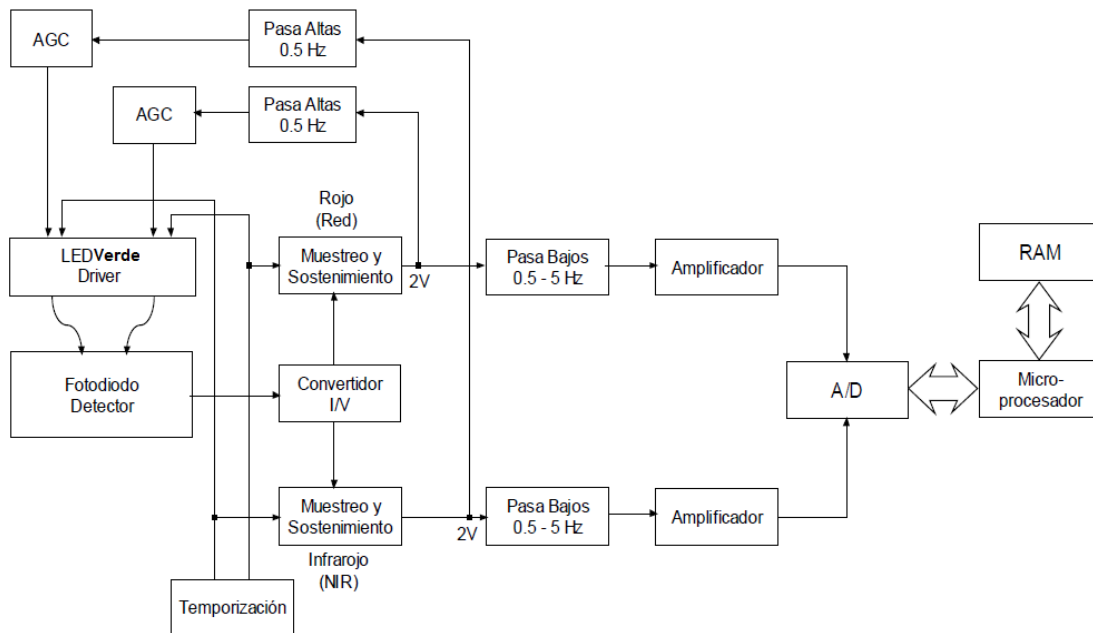
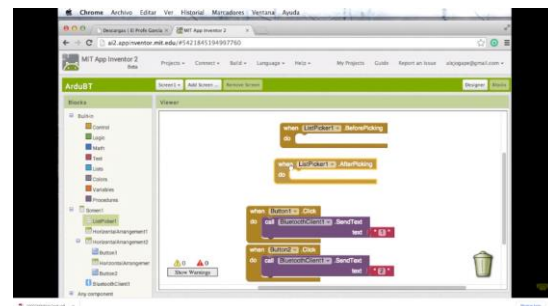


Figura 4. Rivera et al (2021). Diagrama a bloques del sistema de adquisición.

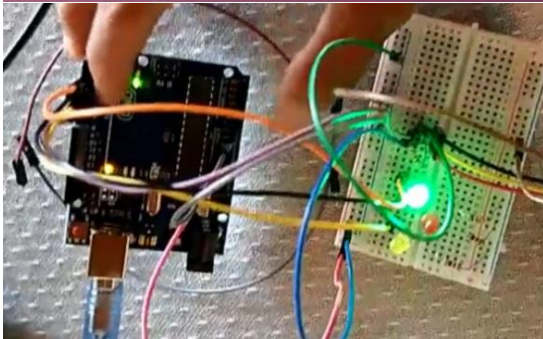
Adicionalmente mediante el módulo Bluetooth HC-05 se realizó la comunicación con el Smartphone. El alcance es de 10 m, se requieren terminales de transmisión y recepción (Tx y Rx). Se puede alimentar desde 3.6 V hasta 6 V. Su velocidad de transmisión puede emplearse de 9,600 bps hasta 11,5200 bps. El algoritmo usado se ilustra en la figura 5b. El programa desarrollado en la tarjeta determina las magnitudes de la intensidad recibida en el fototransistor TEPT 440 de Vishay Semiconductors cuyo rango de sensibilidad va de los 400 a los 800nm con pico máximo en 550nm. Así mismo se realizó un programa en Arduino para la generación de pulsos de 100 Hz. Ver figura 5.

Por último, se utilizó la aplicación MIT App Inventor el cual es un entorno de desarrollo de software creado por Google Labs para la elaboración de aplicaciones destinadas al sistema operativo Android. El usuario puede de forma visual y a partir de un conjunto de

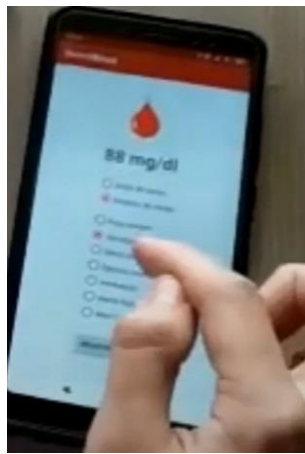
herramientas básicas ir enlazando una serie de bloques para crear la aplicación.



a)



b)



c)

**Figura 5. Rivera et al (2021). a) MIT App inventor entorno de programación, b) Control inalámbrico del led. c) Pantalla en Smartphone con Android.**

### Aprendizaje, detección y clasificación

Desde 1990 los laboratorios de análisis clínicos y microbiológicos Montecristo, en Chalco municipio del Estado de México, han llevado a cabo una base de datos de sus pacientes sobre la determinación de los niveles de colesterol de baja densidad (LDL), colesterol de alta densidad (HDL) y triglicéridos. La información ha sido estructurada en archivos que reportan mil pacientes con información de la presión arterial, temperatura, edad, género y si el paciente es diabético o no. Se tomará la

información de 700 pacientes para el entrenamiento y 300 para evaluación.

#### A. Creación del modelo

Tomaremos como base los programas desarrollados por Andrej Karpathy (2018), donde crearemos una red neuronal con una capa de entrada de 28 elementos, una capa oculta con 128 neuronas y función de activación RELU y una capa de salida con función de activación Softmax con 3 unidades que corresponden a las 3 categorías: alto, medio o bajo de los niveles de concentración de triglicéridos. Se importa la librería Keras ya incluida en Tensor Flow, y específicamente importamos los módulos: Sequential, Flatten y Dense con los que se implementará la RNN. Ver figura 6. La función Sequential se crea un contenedor, donde se agregan secuencialmente las capas del modelo. La función Flatten crea vectores de entrada en este caso de 28 elementos o pacientes. La función Dense permite agregar la capa oculta de 128 neuronas y la capa de salida con 3 neuronas. Especificando las funciones de activación RELU y Softmax respectivamente.

```

1 # TensorFlow/Keras
2 import tensorflow as tf
3 tf.random.set_seed(4)
4 from tensorflow.keras import datasets, Sequential
5 from tensorflow.keras.layers import Flatten, Dense
6
65 #
66 # 1. MODELO: red neuronal con una capa oculta de 18 neuronas y salida tipo softmax con 10 neuronas
67 #
68
69 modelo = Sequential()
70 modelo.add( Flatten(input_shape=(28,28)) )
71 modelo.add( Dense(128, activation = 'relu') )
72 modelo.add( Dense(10, activation = 'softmax') )
73
74
75

```

**Figura 6. Rivera et al (2021). Fragmento de código: Creación del modelo.**

#### B. Compilación y entrenamiento del modelo

En la etapa de compilación se define la función de error, en este caso la: *categorical cross entropy* que corresponde a los mínimos cuadrados. Del mismo modo se define el algoritmo para minimizar el error, en este



caso: Adam, que es una función que adapta el factor o tasa de aprendizaje en el algoritmo del descenso del gradiente en función de cómo estén distribuidos los parámetros.

A continuación, se define la forma en que se mide el desempeño del modelo. Se debe recordar que ésta es la base del entrenamiento, ya que al minimizar la función de error se ajustan los coeficientes de cada neurona de la red, para así lograr la mayor precisión en la clasificación de los datos. En la etapa de entrenamiento, se presentan los datos de entrenamiento que equivalen a la información de 700 pacientes con sus correspondientes categorías. Se debe decir que se obtuvo el 91% de precisión como resultado luego de varias corridas de entrenamiento. Ver figura 7.

```
74 #  
75 # 2/3. COMPILACIÓN Y ENTRENAMIENTO DEL MODELO  
76 #  
77 modelo.compile(optimizer='adam', loss='categorical_crossentropy', metrics=['accuracy'])  
78 modelo.fit(x_train, y_train, epochs=10, verbose=1)
```

**Figura 7. Rivera et al (2021). Fragmento de código: Compilación y Entrenamiento del modelo.**

### C. Evaluación y Predicción del modelo

En la etapa de evaluación se logró una precisión del 88% empleando el conjunto de prueba (datos de los 300 pacientes restantes). Ver figura 8. Finalmente se lleva a cabo la etapa de predicción donde aleatoriamente se

seleccionan datos de los 300 pacientes del set de evaluación. Se pudo observar que el comportamiento no es ideal ya que existieron varios casos donde la red no ubicó correctamente la categoría.

```
80 #  
81 # 4/5. EVALUACIÓN Y PREDICCIÓN  
82 #  
83 error, precision = modelo.evaluate(x_test, y_test, verbose=2)  
84 print('\nPrecisión con el set de validación:', precision)  
85  
86 predicciones = modelo.predict(x_test)
```

**Figura 8. Rivera et al (2021). Fragmento de código: Evaluación y predicción del modelo.**

## IV. Pruebas y Resultados

Se solicitó el apoyo de los laboratorios Montecristo para probar el sistema en sus instalaciones de los cuales nos reportaron el 90% de acierto en la clasificación de los 40 pacientes que fueron sometidos a estudio y que participaron voluntariamente. En la tabla 2 siguiente, se muestran los resultados obtenidos en la etapa de prueba donde se utilizó el sistema portátil. Se realizó la prueba sobre 40 pacientes distribuidos en grupos identificados como: diabéticos, colesterol elevado de baja y alta densidad (LDL/HDL), así como diferentes niveles de triglicéridos. Se puede observar que el reconocimiento adecuado en el entrenamiento alcanzó un valor promedio del 91% en tanto que en la fase de prueba fue del 88%.

**Tabla 2. Resultados obtenidos en las fases de entrenamiento y prueba de acuerdo con el tipo de grupo.**





Tipo de Grupo	Reconocimiento				Precisiones			
	Entrenamiento		Prueba		Entrenamiento		Prueba	
	Reconocimiento Adecuado / Total	Porcentaje	Reconocimiento Adecuado / Total	Porcentaje	Valor	Porcentaje	Valor	Porcentaje
Diabético	39/40	98.51	36/40	89.53	0.9735	97.4	0.9244	92.4
Colsterol LDL/HDL	38/40	97.53	35/40	88.13	0.9608	96.1	0.9121	91.2
Triglicéridos	36/40	91.58	37/40	92.16	0.9015	90.1	0.8501	85

## V. Conclusiones

Se tomaron las informaciones de mil pacientes, para el entrenamiento de la red neuronal; de los cuales 700 fueron utilizados para la fase de entrenamiento de la red y 300 para la fase de prueba. Se logró determinar bajo experimentación aplicando directamente el sistema portátil de medición de triglicéridos a 40 pacientes en las instalaciones del laboratorio Montecristo que fueron clasificados en tres grupos: pacientes con diabetes, pacientes con colesterol LDL/HDL alto y pacientes con diferentes niveles de concentración de triglicéridos. Con base en los resultados obtenidos se puede mencionar que se logró una precisión superior al 90%. En comparación con otros métodos existentes para la detección automática de niveles de concentración se probó y verificó el funcionamiento del sistema resultando de comportamiento igual de competitivo según se observó el desempeño en la evaluación de los individuos.

## VI. Referencias

Andrej Karpathy. (2018). The Unreasonable Effectiveness of Recurrent Neural Networks. 2015.

URL:<http://karpathy.github.io/2015/05/21/rnn-effectiveness/> (visitado 08-02-2020).

Alex Graves. (2014). Supervised sequence labelling with Recurrent Neural Network, Arxiv Preprint Arxiv: 1308.0850v5.

García-Saavedra MJ, Vicente García JC (1999): Fundamentos y Técnicas de Análisis Bioquímico, Análisis de Muestras Biológicas. Editorial Paraninfo, Madrid.

González de Buitrago JM, Arilla Ferreiro E, Rodríguez-Sedade M, Sánchez Pozo A (1998): Bioquímica Clínica. Editorial McGraw-Hill. Interamericana, Madrid.

Harris, D. C. (2007). Análisis químico cuantitativo. 3ª. Edición. Capítulo 18. Ed. Reverté.

Hernández-Hernández, L. (2002) Introducción al análisis instrumental. Capítulo 3. Ed. Ariel Ciencia.

Ian Goodfellow, Yoshua Bengio y Aaron Courville. (2016). Deep learning book in preparation for MIT Press. En: URL <http://www.deeplearningbook.org>.

Fei-Fei et al. (2017). Convolutional Neural Networks (CNNs/ConvNets).URL: <http://cs231n.github.io/convolutional-networks> (visitado 20-2-2020).