



## DETECCIÓN Y CLASIFICACIÓN DEL HÍGADO GRASO NO ALCOHÓLICO POR MEDIOS FOTOMÉTRICOS BASADO EN REDES NEURONALES USANDO PYTHON

**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

*Nonalcoholic fatty liver disease (NAFLD) is produced by an accumulation of triglycerides in hepatocytes related to insulin resistance in the liver and muscles. This project proposes the development of a system that allows classifying degree of progression of liver damage based on the information provided by a clinical laboratory in the State of Mexico and digital processing classification based on deep neural networks using Python language. 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 Beer-Lambert law. In this work, alanine-aminotransferase and aspartate-aminotransferase molecules in blood were measure. This is due because both substances present in a high level in blood and along other conditions, provide a well stablish diagnosis of liver diseases.*

*Key words: Nonalcoholic fatty liver, neural network, severidad, Python.*

En el hígado se pueden presentar varias enfermedades como infecciones, diferentes tipos de hepatitis, etc. Asimismo, se puede dar lugar en este órgano a una condición metabólica donde sin necesidad de consumo crónico de alcohol se logran niveles elevados de los lípidos, a esta enfermedad se le denomina enfermedad del hígado graso no alcohólico HGNA o en sus siglas en inglés como NAFLD (Nonalcoholic Fatty Liver

Disease). Este padecimiento exhibe varias fases y grados de severidad, de acuerdo con Díaz Fonden et al. (2015), la patogénesis de la esteatosis no alcohólica y sus complicaciones queda descrita en la secuencia de eventos:

1. Se reduce el efecto antilipolítico de la insulina e incrementa la lipólisis.
2. Aumenta el flujo de ácido grasos libres (AGL) hacia el hígado.

3. Se incrementa la lipogénesis y la reesterificación, ocasionando una acumulación de triglicéridos (TG).

4. Se Reduce la exportación de triglicéridos, por lo que se acumula la TG.

5. La acumulación excesiva de triglicéridos por arriba del 5% del peso total del hígado origina la esteatosis hepática no alcohólica.

6. La acumulación excesiva de TG provoca un lipotoxicidad y un estrés oxidativo.

7. La inflamación crónica desencadena la esteatohepatitis.

8. Si la esteatohepatitis no es controlada adecuadamente, puede progresar a fibrosis, cirrosis y cáncer hepático.

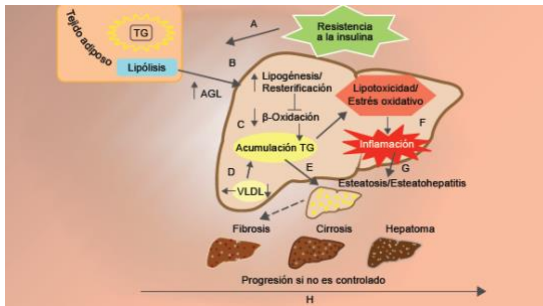


Figura 1. Patogénesis de la esteatosis no alcohólica y sus complicaciones. Con base a Aguilera Méndez, (2017).

### Determinación cuantitativa de alanina aminotransferasa GPT (ALT)

Existen diferentes métodos para la determinación de la alanina aminotransferasa (ALT) inicialmente llamada transaminasa glutámico-pirúvica (GPT), según expresa González de Buitrago et al. (1998), en suero se utilizan reactivos comerciales (kit comercial de Spinreact Chemicals) que incluyen las enzimas y sustratos necesarios para la cuantificación por espectrofotometría visible. Las reacciones que tienen lugar se ilustran en la figura 2:

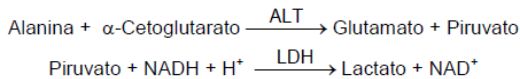


Figura 2. Reacciones del método de determinación cuantitativa del ALT. Con base a González de Buitrago et al. (1998).

La velocidad de disminución de la concentración de NADH (al oxidarse y convertirse en NAD) en el medio, determinado fotométricamente, es proporcional a la concentración catalítica de ALT en la muestra ensayada que debe satisfacer una longitud de onda de 340 nm y un paso de luz de 1 cm bajo temperatura constante de 25°C, 30°C o 37°C. La ALT es una enzima intracelular, se encuentra principalmente en las células del hígado y el riñón. Su mejor aplicación es en el diagnóstico de las enfermedades del hígado. Se observan niveles elevados en enfermedades hepáticas, enfermedades de los músculos y traumatismos. Cuando se emplean en conjunción con el aspartato aminotransferasa (AST) ayuda en el diagnóstico de infartos al miocardio. Los valores de referencia clínicos empleados para clasificar los grupos se ilustran en la tabla 1.

Tabla 1. Valores de referencia de ALT y AST.

Alanina aminotransferasa (U/L)		Temperatura °C
Hombres	Mujeres	
22	18	25
29	22	30
40	32	37

### Determinación cuantitativa de aspartato aminotransferasa GOT (AST)

La determinación cuantitativa del aspartato aminotransferasa (AST) se obtiene de forma análoga a la de ALT. Las reacciones que tienen lugar se ilustran en la figura 3:

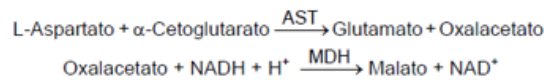




Figura 3. Reacciones del método de determinación cuantitativa del GOT (AST). Con base a González de Buitrago et al. (1998).

concentración del analito en la solución. Es decir:

$$A = a b c \dots (3)$$

Donde  $a$  es una constante de proporcionalidad llamada absorptividad. Resulta evidente que la magnitud de  $a$  depende de las unidades utilizadas para  $b$  y  $c$ . Cuando se expresa la concentración en moles por litro y la trayectoria a través de la celda en centímetros, la absorptividad se denomina absorptividad molar.

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.

Por lo tanto, se tiene que:

$$\frac{I_1}{I_0} = 10^{-\alpha l} = 10^{-\varepsilon l C} = 10^{-A} \dots (4)$$

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.

## 1. Descripción de conceptos

### 1.1 Definición de 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  $T$  (1) de la solución es la fracción de la radiación incidente transmitida por la solución (por lo general, la transmitancia se expresa como porcentaje):

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

### 1.2 Definición de 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.

### 1.3 Ley de Lambert-Beer

Cuando un haz de luz monocromática (de determinada longitud de onda) atraviesa una solución, la absorbancia es directamente proporcional a la distancia recorrida por la luz atravesando la solución absorbente y a la

### 1.4 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, Existen una serie de capas donde 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 muestran el resultado buscado. Ver figura 4.

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.

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

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.

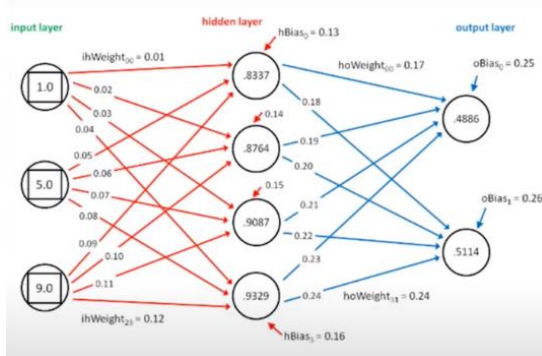


Figura 4. Red neuronal. Elaboración propia.

Como lo indica Graves Alex et al. (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 5.

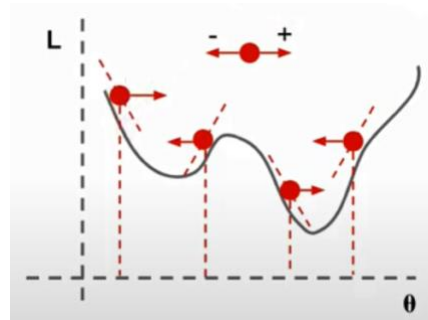


Figura 5. Descenso del gradiente. Elaboración propia.

## 2. Desarrollo

En el presente trabajo se realizó un circuito electrónico basado en la tarjeta Arduino que active la fuente de luz, en este caso el led UV (340 nm) UV5TZ de la compañía BIVAR. Posteriormente mediante una adaptación, se prepara el reactivo de la compañía SPINREACT que se combina con la muestra sanguínea del orden de 100  $\mu$ L (equivalente a dos gotas de sangre) dando lugar a las reacciones; donde luego, por medio de un fotodetector ML8511 se realiza la medición de la absorbancia correspondientes a las sustancias: alanina aminotransferasa (ALT) y aspartato aminotransferasa (AST) respectivamente.

### 2.1 Metodología

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

[Paso 1:] Preprocesamiento.

[Paso 2:] Transferencia inalámbrica.

[Paso 3:] Aprendizaje, detección y clasificación.

## 2.2. Preprocesamiento

En esta etapa se establece el diseño de los circuitos que constituyen el medidor de concentración de dos sustancias, primero: la alanina aminotransferasa y segundo: el aspartato aminotransferasa. Tomando como base las reacciones del método donde, la alanina reacciona formándose piruvato y glutamato. La actividad de la ALT y AST se mide determinando la disminución de absorbancia a 340 nm. Ver figura 6.

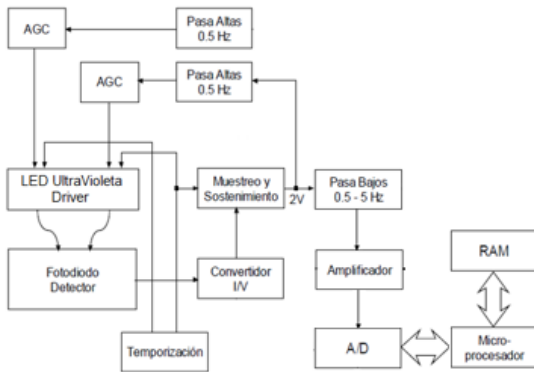
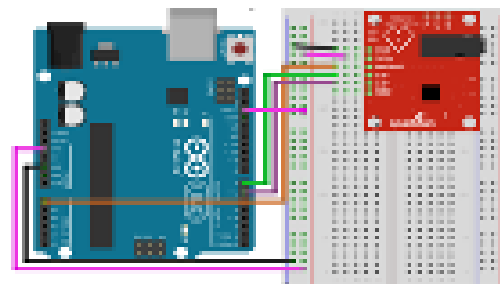


Figura 6. Diagrama a bloques del sistema de adquisición. Elaboración propia.

## 2.3. 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 (ver figura 7). La señal se digitalizó a un formato de 8 bits de enteros sin signo. 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 115,200 bps. El algoritmo usado se ilustra en la figura 7b. El programa desarrollado en la tarjeta determina las magnitudes de la intensidad recibida entre el led BIVAR UV5TZ y el fotodetector ML8511, cuyo pico máximo se encuentra en los 390 nm.



a)

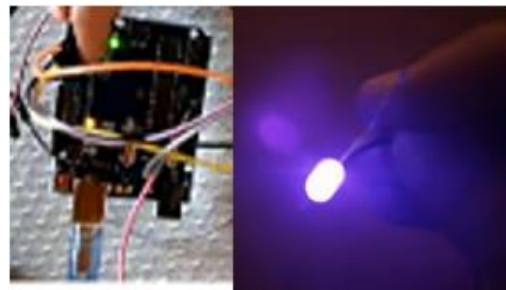
### Algoritmo 1: Adquisición de señal cardiaca

```

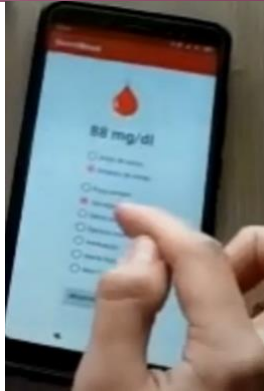
1: procedure COMUNICACIÓN SPARKFUN(estadoBT)
2:   pinLomen = entrada;
3:   pinLomas = entrada;
4:   pinEntradaAnalógica = entrada;
5:   if DatoADC >= Lomen y DatoADC <= Lomas then
6:     Reducir dato a formato de 8 bits entero sin signo
7:     Dato8bit = map(datoADC, 0, ADC33, 0, 255);
8:     Envía por Bluetooth pinEntradaAnalógica;
9:     Serial.write(datoSerie);
7:   end if
8:   if DatoADC >= Lomen then
8:     Señal a cero del rango DatoADC= ADC33/2;
9:   end procedure

```

b)



c)



d)

Figura 7. a) Tarjeta Arduino UNO y módulo HC05, b) Algoritmo de adquisición y comunicación, c) Led UV y d) Aplicación Android. Elaboración propia.

### 3. 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 diversas sustancias a efecto de prevenir, detectar o diagnosticar diferentes enfermedades. Tomando como referencia estos valores es posible establecer criterios de análisis para desarrollar un sistema de apoyo médico para el diagnóstico de la enfermedad de hígado graso no alcohólico.

Para la variable de entrada (X):

1. Edad (años).
2. Índice de masa corporal (kg/m<sup>2</sup>).
3. Presión arterial (mm Hg).
4. Concentración de colesterol (mg/dL).
5. Concentración de triglicéridos (mg/dL).
6. Concentración de glucosa en plasma (mg/dL).
7. Concentración de alanina aminotransferasa ALT/GPT (U/L).
8. Concentración de aspartato aminotransferasa AST/GOT (U/L).

Para la variable de salida (Y):

1. Clase, determinación de hígado graso no alcohólico en un plazo de cinco años.

#### A) Datos de carga de la Red Neuronal

Se va a trabajar con el conjunto de datos sobre pacientes con enfermedad de hígado graso no alcohólico. Se puede cargar el archivo directamente usando la función `loadtxt()` de NumPy (ver figura 8). Hay ocho variables de entrada y una variable de salida (la última columna). Una vez cargado podemos dividir el conjunto de datos en variables de entrada (X) y la variable de clase de salida (Y).

```

1 # carga los datos
2 dataset = numpy.loadtxt("higado_graso.csv", delimiter=",")
3 # dividido en variables de entrada (X) y salida (Y)
4 X = dataset[:,0:8]
5 Y = dataset[:,8]
```

Figura 8. Fragmento de código: Carga de datos. Elaboración propia.

#### B) Creación del modelo

Se toma como base los programas desarrollados por Andrej Karpathy (2018), donde crearemos una red neuronal con base en los modelos en Keras en el que se definen como una secuencia de capas. Creamos un modelo secuencial y añadimos capas una a una hasta que estamos satisfechos con nuestra topología de red. Ver figura 9.

```

1 # crea el modelo
2 model = Sequential()
3 model.add(Dense(12, input_dim=8, activation='relu'))
4 model.add(Dense(8, activation='relu'))
5 model.add(Dense(1, activation='sigmoid'))
```

Capa visible (8 entradas)>>Capa oculta (12 neuronas)>>

Capa oculta (8 neuronas)>>Capa de salida (1)

Figura 9. Fragmento de código: Creación del modelo. Elaboración propia.



Figura 10. Fragmento de código: Compilación y entrenamiento del modelo. Elaboración propia.

Las capas son completamente conectadas para lo cual se emplea la clase Dense. Como función de activación en las dos primeras capas se eligió la función de rectificación lineal (relu) y la función de activación sigmoide en la capa de salida. La función de activación sigmoide en la capa de salida permite que la salida de la red se encuentre entre los valores 0 y 1. El umbral por defecto será de 0,5. Cabe mencionar que la primera capa oculta tiene 12 neuronas y espera 8 variables de entrada, en tanto que la segunda capa oculta tiene 8 neuronas y la capa de salida cuenta con 1 neurona para predecir la clase.

### C) Entrenamiento del modelo

La compilación del modelo utiliza las eficientes librerías numéricas de backend como Theano o TensorFlow. El backend elige automáticamente la mejor manera de representar la red para entrenar y hacer predicciones para que se ejecuten en su hardware, como CPU o GPU o incluso distribuidas. Se debe especificar la función de pérdida en este caso, utilizaremos la pérdida logarítmica, que para un problema de clasificación binaria se define en Keras como “binary\_crossentropy”. También utilizaremos el algoritmo de descenso de gradiente eficiente “adam” por su alta eficiencia en estos problemas. 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 90% de precisión como resultado luego de varias corridas de entrenamiento. Ver figura 10.

```
1 | # Compila el modelo
2 | model.compile(loss='binary_crossentropy',
  | optimizer='adam', metrics=['accuracy'])
```

### D) Modelo de ajuste de la Red Neuronal

Para este problema se utiliza un pequeño número de iteraciones o epochs (150) y un número de instancias o batch\_size relativamente pequeño (10). Una vez más, estos pueden ser elegidos experimentalmente por ensayo y error. Ver figura 11.

```
1 |
2 | # Ajusta el modelo
  | model.fit(X, Y, epochs=150, batch_size=10)
```

Figura 11. Fragmento de código: Ajuste de la Red Neuronal. Elaboración propia.

### E) 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 12). Finalmente se lleva a cabo la etapa de predicción donde aleatoriamente se seleccionan datos de los 300 pacientes del set de evaluación.

```
1 |
2 | # evalua el modelo
3 | scores = model.evaluate(X, Y)
  | print("\n%s: %.2f%%" % (model.metrics_names[1],
  |                       scores[1]*100))
22 | # calcula las predicciones
23 | predictions = model.predict(X)
24 | # redondeamos las predicciones
25 | rounded = [round(x[0]) for x in predictions]
  | print(rounded)
[1.0, 0.0, 1.0, 0.0, 1.0, 0.0, 0.0, 1.0, 1.0, 0.0, 0.0, 1.0, 0.0, 1.0]
```

Figura 12. Fragmento de código: Evaluación y predicción del modelo. Elaboración propia.

## 4. Pruebas y resultados

Se solicitó el apoyo de los laboratorios Montecristo para probar el sistema en sus instalaciones donde 40 pacientes que fueron



sometidos a estudio y participaron voluntariamente. Se llevó a cabo tanto el análisis clínico convencional (muestra de sangre extraída vía intravenosa) en el que se empleó el analizador semiautomatizado BTS-350 que es un fotómetro capaz de obtener medidas de absorbancia, así como el uso del sistema portátil de este proyecto.

En las tablas 2 y 3, se muestran los resultados obtenidos en la etapa de prueba donde se utilizó el sistema portátil.

**Tabla 2. Resultados obtenidos en la medición de la ALT/GPT. Elaboración propia.**

LONGITUD DE ONDA (Å)	TEMPERATURA (°C)	VOLUMEN DE LA MUESTRA (µL)	RESULTADOS OBTENIDOS (U/L)			MEDIA (m)	DESVIACIÓN ESTANDAR (s)
340nm	No programada	300	17	17	18	17,333	0,577
		400	18	20	21	19,666	1,527
		500	20	19	19	19,333	0,577
	25	300	26	23	23	24	1,732
		400	24	22	21	22,333	1,527
		500	26	29	26	27	1,732
	37	300	38	35	34	35,666	2,081
		400	34	33	32	33	1
		500	37	37	36	36,666	0,577

**Tabla 3. Resultados obtenidos en la medición de la AST/GOT. Elaboración propia.**

LONGITUD DE ONDA (Å)	TEMPERATURA (°C)	VOLUMEN DE LA MUESTRA (µL)	RESULTADOS OBTENIDOS (U/L)			MEDIA (m)	DESVIACIÓN ESTANDAR (s)
340nm	No programada	300	25	25	27	25,666	1,154
		400	17	20	19	18,666	1,527
		500	20	22	19	20,333	1,527
	25	300	29	30	35	31,333	3,214
		400	21	19	25	21,666	3,055
		500	28	27	32	29	2,645
	37	300	39	51	42	44	6,244
		400	43	45	46	44,666	1,527
		500	37	48	40	41,666	5,666

Para la precisión se determina a partir de la población de datos obtenidos, bajo las mismas condiciones experimentales, se proceden a determinar los parámetros de calidad, tales como desviación estándar absoluta, desviación

estándar relativa, coeficiente de variación y varianza. Los valores obtenidos aquí nos dan un estimativo de la precisión con la cual se podrá realizar cada análisis. Se puede observar que el reconocimiento adecuado en el



entrenamiento alcanzó un valor promedio del 88%.

La gráfica de la figura 13 muestra la relación de la concentración de las muestras sanguíneas expresadas en mg/dL contra las medidas determinadas de Absorbancia. Se puede ver que la red neuronal ajusta la recta

para la serie de puntos de entrada lo que corresponde con una gráfica lineal que se expresa para:

a) ALT/GPT:  $Y = 1.041x + 1.447$ .

b) AST/GOT:  $Y = 1.01x - 0.26$ .

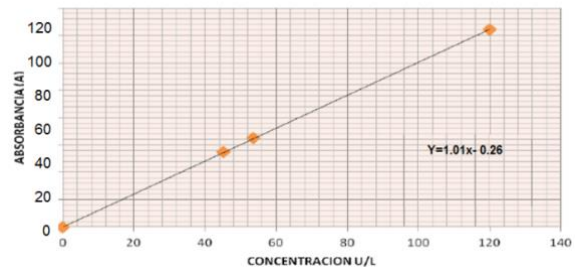
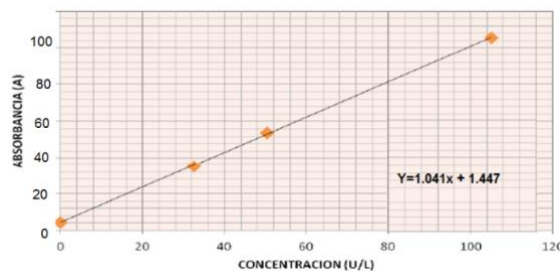


Figura 13. Curva de concentración vs Absorbancia en ALT y AST. Elaboración propia.

## 5. Conclusiones

Existe una clara asociación entre la diabetes tipo 2 y la HGNA, se estima que hay una prevalencia entre 60-80% de HGNA en los pacientes con diabetes tipo 2. También se ha demostrado que la obesidad, la acumulación de grasa visceral, la resistencia a la insulina, la hipertensión y las dislipidemias son factores de riesgo para la esteatosis hepática. Por otra parte, se logró interpretar adecuadamente la información de la base de datos proporcionada por los laboratorios Montecristo de Chalco Estado de México, a quienes sea de paso les agradecemos su colaboración amplia y diligente. Se pudo determinar bajo experimentación aplicando directamente el sistema portátil de medición los niveles de concentración de la ALT /GPT como de la AST/GOT a 40 pacientes en las instalaciones del laboratorio Montecristo que fueron clasificados en tres grupos con una precisión en los resultados del orden de 88%. En comparación con otros métodos existentes para la detección automática de niveles de concentración se verificó el funcionamiento

del sistema resultando competitivo según la evaluación de los individuos.

## Referencias

- Álvarez Cedillo, J. A., Herrera Lozada, J. C., & Pérez Romero, P. (2010). Sistema informático para análisis de cardiopatía Holter. *Revista Polibits*(41), 59-66.
- Andrej Karpathy. (2018). The Unreasonable Effectiveness of Recurrent Neural Networks. 2015.URL:<http://karpathy.github.io/2015/05/21/rnn-effectiveness/> (Consultado 08-02-2020).
- Aguilera Méndez Asdrúbal (2017). Esteatosis hepática no alcohólica: una enfermedad silenciosa. *Revista médica IMSS*. 2018; 56(6):544-9.
- Díaz Fonden Jesús et al (2015). Relación entre los hallazgos ecográficos, laparoscópicos e histológicos en pacientes con esteatosis hepática no alcohólica. *Revista MEDISAN* 2015; 19(3):345. Cuba.
- Graves Alex. (2014). Supervised sequence labelling with Recurrent Neural Network, Arxiv Preprint Arxiv: 1308.0850v5.



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

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>.