



LAS REDES NEURONALES CUÁNTICAS Y ENSEÑANZA DE ASPECTOS BÁSICOS DE LA MÉCANICA CUÁNTICA

QUANTUM NEURAL NETWORKS AND TEACHING BASIC ASPECTS OF QUANTUM MECHANICS

REDES NEURAIS QUÂNTICAS E ENSINO DE ASPECTOS BÁSICOS DA MECÂNICA QUÂNTICA

Alejandra León *, Martin Arias** Juan Giraldo***, Carlos Jácome****

León, A.; Arias, M.; Giraldo, J.; Jacome, C. (2023). Las redes neuronales cuánticas y enseñanza de aspectos básicos de la mecánica cuántica. *Góndola, Enseñanza y Aprendizaje de las Ciencias*, Número especial, v18, pp.1-13.

Resumen

La enseñanza de la física en la actualidad se liga a los desarrollos que se producen en el campo de la computación, por tanto, no se puede desconocer que desde las redes neuronales se crean nuevas tecnologías y evoluciona en términos de la Inteligencia artificial. La física adoptó estos sistemas con la finalidad de poder hacer procesamiento de datos, análisis de sistemas, procesamiento de imágenes y clasificación de señales de manera más óptima, que se aplican en áreas tales como la óptica, astronomía, mecánica cuántica, entre otros. De forma paralela se crea una relación entre las redes neuronales y la computación cuántica, por lo cual, este documento profundiza dentro de dichas estructuras e identifica los lazos que hay con el formalismo cuántico. La hoja de ruta se desarrolló a partir de la construcción y análisis de algoritmos de redes neuronales, los cuales, se fundamentan en dos componentes principales, las funciones de activación y la autorregulación de pesos estadísticos que estructuran las bases del aprendizaje automático, lo que permite la investigación multidisciplinar entre la física computacional y la mecánica cuántica para la optimización y evolución del concepto de red neuronal “clásica”, se buscó un sistema cuántico de red neuronal las bases conceptuales y matemáticas para el desarrollo a futuro del concepto de la red neuronal cuántica, de esta manera, a partir del uso de conceptos fundamentales de la mecánica cuántica, se encontró alternativas para desarrollar la enseñanza de la teoría y reflejar la aplicabilidad en el *machine learning* un campo en potencia, para el desarrollo de la inteligencia artificial.

* Estudiante de Licenciatura en física, Universidad Distrital Francisco José de Caldas, Colombia. mleon@udistrital.edu.co - ORCID <https://orcid.org/0000-0003-4326-0227>.

** Estudiante de Licenciatura en física, Universidad Distrital Francisco José de Caldas, Colombia. lmariasc@udistrital.edu.co - ORCID <https://orcid.org/0000-0002-6937-6642>

*** Magister en filosofía, Universidad Distrital Francisco José de Caldas, Colombia. jgiraldo@udistrital.edu.co - ORCID <https://orcid.org/my-orcid?orcid=0000-0001-5631-6192>

**** Magister en física, Universidad Distrital Francisco José de Caldas, Colombia. cjacome@udistrital.edu.co - ORCID

Palabras-Clave: Percepción. Proceso de aprendizaje. Memoria. Asociación patronal.

Abstract

The teaching of physics is currently linked to the developments that occur in the field of computing; therefore, it cannot be ignored that new technologies are being developed from neural networks and evolving in terms of artificial intelligence. Physics has adopted these systems in order to be able to perform data processing, system analysis, image processing and signal classification in a more optimal way, which are applied in areas such as optics, astronomy, quantum mechanics, among others. In parallel, a relationship has been created between neural networks and quantum computing, for which this document seeks to investigate and delve into these structures and identify the links with quantum formalism. The established roadmap is developed from the construction and analysis of neural network algorithms, which are based on two main components, the activation functions and the self-regulation of statistical weights that structure the bases of machine learning, which allows multidisciplinary research between computational physics and quantum mechanics for the optimization and evolution of the concept of "classical" neural network, affirming to date the inexistence of a quantum system of the neural network, but the conceptual and mathematical association bases for the future development of the concept of the quantum neural network, in this way, based on the use of fundamental concepts of quantum mechanics, it is possible to blur the abstract nature of said physical theory and reflect the applicability in machine learning, a potential field for the development of artificial intelligence.

Keywords: Perception. Learning process. Memory. Employer association.

Resumo

Resumen del artículo em português. O ensino de física atualmente está atrelado aos desenvolvimentos que ocorrem no campo da computação, portanto, não se pode ignorar que novas tecnologias estão sendo desenvolvidas a partir de redes neurais e evoluindo em termos de inteligência artificial. A física adotou esses sistemas para poder realizar processamento de dados, análise de sistemas, processamento de imagens e classificação de sinais de forma mais otimizada, que são aplicados em áreas como óptica, astronomia, mecânica quântica, entre outras. Em paralelo, foi criada uma relação entre redes neurais e computação quântica, para a qual este documento busca investigar e aprofundar essas estruturas e identificar as ligações com o formalismo quântico. O roteiro estabelecido é desenvolvido a partir da construção e análise de algoritmos de redes neurais, que se baseiam em dois componentes principais, as funções de ativação e a autorregulação de pesos estatísticos que estruturam as bases do aprendizado de máquina, o que permite pesquisas multidisciplinares entre física computacional e mecânica quântica para a otimização e evolução do conceito de rede neural "clássica", afirmando até hoje a inexistência de um sistema quântico da rede neural, mas as bases conceituais e matemáticas de associação para o desenvolvimento futuro do conceito de rede neural quântica, desta forma, com base no uso de conceitos fundamentais da mecânica quântica, é possível borrar a natureza abstrata da referida teoria física e refletir a aplicabilidade no

aprendizado de máquina, um campo potencial para o desenvolvimento da inteligência artificial.

Palavras-Chave: Percepção, Processo de aprendizagem, Memória, Associação patronal.

1. Introducción

El mundo de la computación avanza a un ritmo exponencial desde 1936 con la definición de máquina de Turing (1936) que plasma el punto de partida de la computación hasta la actualidad, donde se habla de una incursión del campo de la física cuántica en varios algoritmos que ha permitido una optimización en los procesos de funcionamiento y análisis de datos. Sin embargo, la idea principal que da origen a evolución sin freno de la computación es la reproducción del funcionamiento del cerebro humano a través de instrucciones que fueran codificadas en un lenguaje de máquina, dedicado en esta idea Rosenblatt (1957) crea el concepto del perceptrón modelo de funcionamiento de una neurona biológica que es el principal componente del sistema nervioso, dicho concepto mencionado en 1957 ha evolucionado a través del punto de vista de diferentes investigadores que se han interesado en este campo de la física computacional hasta llegar a algoritmos como el “*gradient descent*” y el “*backpropagation*” que empiezan a definir el límite de evolución por medio de análisis clásicos, es por lo que, se crea la necesidad de buscar una colaboración interdisciplinar entre el formalismo cuántico y la computación de las redes neuronales o aprendizaje automático, que permita encontrar desde la definición matemática una optimización y desde la definición física un acercamiento hacia los procesos y relaciones que resultan de la comprensión del cerebro, es decir, su objetivo es mostrar las relaciones físicas que se puede generar en conceptos de percepción, asociación, aprendizaje y memorización desde el formalismo cuántico y a través de la computación, de esa manera orientar un avance de lo que será la incursión definitiva del física cuántica y la computación de los próximos años, por el momento se habla desde el software ya

que, se es de conocimiento que el hardware aún no está definido para su uso comercial. De tal manera, el lenguaje de computación que se está en constante adaptación a esta nueva forma de ver el mundo, Python, cuenta con librerías de suma importancia como Numpy (Tipo de datos de matrices multidimensionales), Pandas (Manejo y análisis de estructuras de datos), Tensor-Flow (Aprendizaje automático), Keras (Redes neuronales clásicas) y la más importante para el desarrollo de este artículo la librería Penny-Lane (Programación diferenciable de computadoras cuánticas) (Wilkinson, 2020). De tal manera se desarrolla la descripción general del modelo clásico el perceptrón como punto de inicio, para luego observar las analogías de circuitos cuánticos con la definición de red neuronal artificial y evidenciar la optimización del procesamiento de datos. Por último, el objetivo es desarrollar de forma conceptual las asociaciones que se han encontrado desde el formalismo cuántico y la teoría clásica para generar espacios en la enseñanza de la mecánica cuántica y ver la aplicabilidad en campos modernos en consolidación como la computación e información cuántica.

2. Redes neuronales artificiales

El primer concepto de neurona nace de la mente de Rosenblatt en 1957, donde crea un circuito electromecánico capaz de reconocer similitudes entre diferentes patrones de entrada con un modelo simple llamado perceptrón Fig. 1, que consistía de tres aspectos fundamentales. Primer aspecto, unos datos los cuales eran figuras geométricas sencillas como círculo y triángulo que se ingresaban a la maquinas en orden mientras se tenía una asignación encendida, es decir, mientras se tenía la luz que significaba clase triángulo, se pasaban a través de la máquina diferentes tipos de triángulos bajo transformaciones de semejanza, de esta manera se

realiza un proceso de memorización forzada y generalización que permite en un futuro reconocer cualquier patrón nuevo, se apagaba y se encendía la luz perteneciente a la clase círculo y se repetía el mismo procedimiento, a este primer aspecto se le conoce como un proceso de enseñanza y se encuentra su análogo biológico en las dendritas las cuales son el sistema sensorial que recibe la información (Kandel, 1955).

El segundo aspecto donde el sistema de entrada transmite pulsos para que la unidad se active o se inhibe de acuerdo a la intensidad de la transmisión, luego de la activación del sistema de asociación y realizando una suma algebraica de todos los pulsos recibidos para llegar al umbral de excitación el cual le permite enviar un pulso al sistema de respuesta, todo este proceso posee su análogo biológico que está descrito por lo que se comprende por intensidad sináptica y sucede en el cuerpo de la célula o el soma (Kandel, 1955).

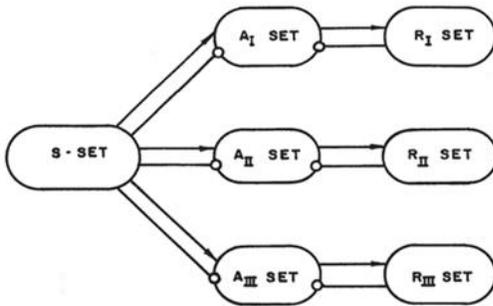


Figura 1. Organización del perceptrón con 3 salidas independientes.

Fuente: (Rosenblatt, 1957)

El tercer aspecto y último, es el sistema de respuesta si la figura es perteneciente a una clase o no, es importante resaltar que la indicación final se da por la probabilidad mayor a que la figura de entrada sea perteneciente o no a la clase triángulo o círculo, de modo que su resultado depende del proceso de aprendizaje.

El concepto principal de perceptrón dado por Rosenblatt no presenta cambio alguno durante la evolución de las redes neuronales, sin embargo, lo que sí cambia es el proceso para dar una

clasificación a un dato de entrada, ya que lo que se busca es que se tenga el mínimo error, se demore un menor tiempo en dar respuesta y la máquina llegue a procesar patrones con una complejidad más allá que unas figuras geométricas (1957).

2.1. Función de activación

En busca de mejorar el modelo se crea el diagrama que hoy se conoce como red neuronal artificial Fig. 2, donde su expresión matemática estará dada por una superposición lineal de las variables de entrada x_i junto con sus respectivos pesos sinápticos w_i , dichos pesos entran a reemplazar las matrices de asociación junto a la función de generar la intensidad al dato de entrada, por lo que el parámetro puede tomar valores positivos, negativos y cero que permite el paso de información o la inhibición de la característica de entrada (Izaurieta & Saavedra, 2000; Li et al., 2020).

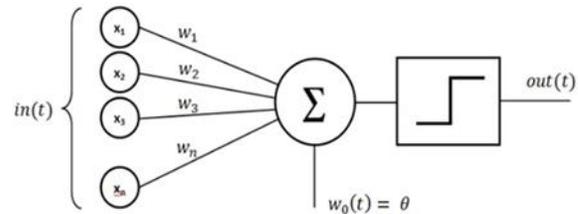


Figura 2. Diagrama de una red neuronal artificial. **Fuente:** (Escámez, 2022)

Para alcanzar la activación de la neurona se realiza una superposición lineal entre las entradas y los pesos sinápticos, que genera así un resultado, sin embargo, se puede condicionar a la neurona que se active si se alcanza un umbral requerido, para ello se maneja una variable adicional que es conocida como el *bias* b la cual se encarga de la excitación umbral de la neurona artificial Ec. 1.

$$h_i(x_n, w_{in}) = \sum_{i=1}^n w_{ij}x_j + b \quad (1)$$

Luego de la superposición lineal, la respuesta de la neurona está dada por una función de activación, hasta el día de hoy con base a las necesidades de la clasificación o predicción de datos y el modelo a usar existen varios tipos de función, sin embargo, para el

desarrollo de este artículo, se utilizará la función sigmoide Ec. 2, que es una función continua y derivable, permitiendo los métodos de aprendizaje como el *backpropagation* Fig. 3 (Hopfield, 1988).

$$y_i(h_i) = \frac{1}{1 + e^{-h_i}} \quad (2)$$

$$y_i(h_i) = \frac{1}{1 + e^{-\sum_{j=1}^n w_{ij}x_j + b}} \quad (3)$$

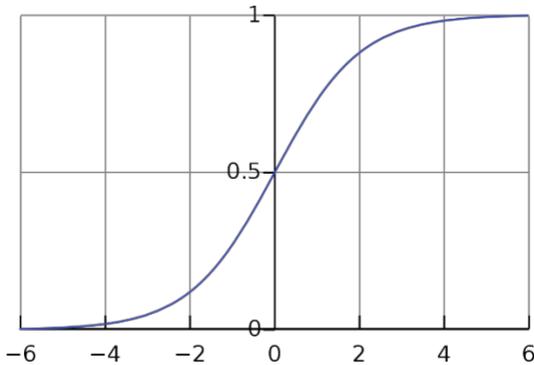


Figura 3. Gráfica de la función sigmoide con rango [0,1].

Fuente: Gnuplot

Por consiguiente, después que se obtiene una respuesta de la red, se debe validar si es correcta o no, en el caso de no ser correcta, la red tiene que entrar en el proceso de un ajuste de parámetros, es decir, los pesos sinápticos y el *bias*.

2.2. Función de pérdida-Cross Entropy

Esta función evalúa la desviación entre los resultados de la red neuronal que salen de la función activación y los valores reales de los datos de entrenamiento, primero se evalúa una función de pérdida Ec. (3). Donde el parámetro y_j estará dado por un one-vector y el parámetro correcto de cada clase tendrá el valor de 1, los demás parámetros cero (Corte, 2021).

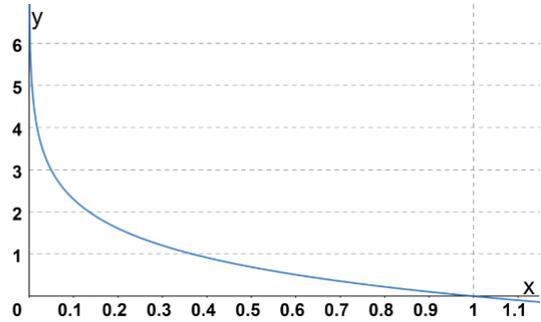


Figura 4. Grafica función de coste.

Fuente: Autores.

función de coste, que solo es dividir por los parámetros del patrón, a esta función se le conoce como "*cross-Entropy*" Ec. 4, encargada de encontrar la precisión de las variables categóricas, de acuerdo a este resultado empieza el ajuste de los parámetros de la función (Corte, 2021).

$$C = -\frac{1}{n} \sum_x [y \ln a + (1 - y) \ln(1 - a)] \quad (4)$$

La Ec 4 a través del parámetro n indica la cantidad total de los datos y el valor esperado y a el vector salida.

2.3. Gradient descent, backpropagation y la concepción de aprendizaje.

Se llama aprendizaje a la capacidad de interiorizar cualquier tipo de conocimiento, en el caso de las redes neuronales se realiza una similitud al proceso de enseñanza, ya que, se escogen diferente tipo de entrada con su respectiva clasificación, el cual es mostrado a la red en busca de que a futuro se pueda recibir cualquier tipo de entrada y realizar su asignación correcta. Es claro, como se ha visto anteriormente que la red a priori no dará un resultado válido, por lo que busca minimizar el error asociado a ese resultado principal, de tal manera que el *Gradient Descent* se representa como una función de concavidad positiva, donde el resultado principal estará en algún punto de la función, y lo que buscará el algoritmo es llevar ese resultado al punto mínimo Fig. 5, entendiendo este proceso como el valor de aprendizaje de una red neuronal,

para luego realizar sus respectivos ajustes a los parámetros de los pesos sinápticos.

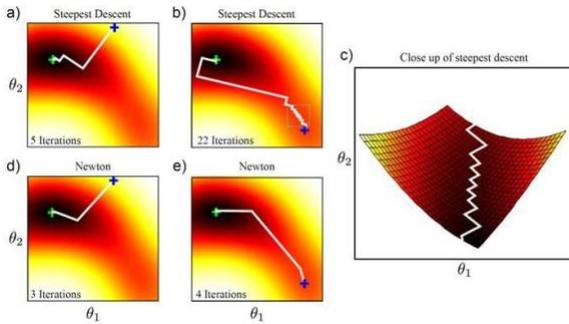


Figura 5. Graficas algoritmo *gradient descent*.

Fuente: (Caparrini, 2023)

En la búsqueda de un valor mínimo en una función se da a través del concepto de derivada parcial de la función de coste con los parámetros w_i y b , se busca el límite cuando la distancia entre dos puntos de la función tienden a un valor h , dicha distancia entre el punto del primer dato, al punto del siguiente dato busca el valor mínimo de la función estará dada por una constante la cual recibe el nombre de razón de aprendizaje α (5), su valor estará determinado por el intérprete, aunque siempre se busca un número pequeño cercano a la unidad (Li et al., 2020).

$$w_{i+1} = w_i - \alpha \nabla J(w_{in}, b) \quad (5)$$

Con el resultado del gradiente descendiente se empieza la corrección de los pesos sinápticos.

En definitiva, no hay un número definido para la repetición del ciclo, pues este cambia de acuerdo a la complejidad del conjunto de datos ingresados, lo que genera una necesidad en relación a la construcción de nuevos algoritmos, que puedan desarrollar y procesar datos de mayor complejidad. En el campo de la inteligencia artificial, se generan algoritmos que dentro de su arquitectura tienen una composición de más de una neurona, a este tipo de algoritmos se les conoce como redes multicapa y como resultado, se logra obtener clasificadores optimizados y mejoras en la predicción de datos. En ese sentido, la inteligencia artificial toma renombre

en la comunidad científica y se prevé el potencial de esta herramienta para los desarrollos tecnológicos y las aplicaciones en los estudios científicos.

2.4. Desarrollo del algoritmo

Referente a la construcción del algoritmo, se tomó un Dataset que explica el movimiento de un sujeto y el objetivo es clasificar, dos tipos de estados de movimiento. El primero determina si el sujeto se encuentra caminando y el segundo determina si el sujeto está corriendo, de esta forma, para llevar a cabo la clasificación del Dataset se plantea un modelo de perceptrón y analizar su rendimiento además de discutir sobre los efectos del aprendizaje. Ahora bien, se hace el uso las librerías Tensor-Flow y Keras, debido a la completitud de herramientas de entrenamiento y el amplio espectro de funciones de activación.

El Dataset reúne 90000 datos por cada columna, las de mayor relevancia son las siguientes, tres columnas para la aceleración y otras tres para indicar la posición en todas las direcciones cartesianas para cada variable mencionada. En efecto, para el diseño de las redes crean 6 variables de entradas, el cual se conectan a una única neurona, que cumpla con el rol de clasificador, para cumplir con dicho fin es importante la elección de la función de activación, como se mencionó anteriormente cada función cumple con un rol diferente, por tanto, para el caso de la clasificación es conveniente usar la función sigmoide, pues este cumple con todos los rubros para desarrollar una óptima clasificación.

Ahora bien, con la descripción del modelo a utilizar para analizar el conjunto de datos, se realizó la división de la misma, para separar los elementos entrenados y los elementos el test de prueba, en concreto, el 70 % del Dataset se destinó para entrenar la red, este grueso de los datos entra repetidamente por el algoritmo, y los datos restantes se usaron para colocar a prueba la red.

Por otra parte, se define la inicialización de los pesos con valores cercanos a cero para desarrollar el primer ciclo de los datos a través de la red, siguiente se define la cantidad de repeticiones por el cual la red procesa el Dataset, es decir, la cantidad de iteraciones para poder observar, el porcentaje de error que arroja la red.

Por último, se aplicó la regla de aprendizaje, basados en los parámetros del *Gradient Descent*, y allí se observa varios factores importantes, el primero el ajuste automático de los pesos a medida que realiza las respectivas iteraciones, y segundo la precisión de clasificación cercana al 100%, así bien, la concepción de aprendizaje se torna al seguimiento de instrucciones a partir del *Gradient Descent* y posteriormente a la toma de decisiones, pues la configuraciones tratadas en cuestión permite que el algoritmo no solamente realice una tarea de tratamiento de datos el cual refleja tendencias, en contraste, filtra caminos que permite la toma de decisiones en relación a un problema de clasificación.

3. Redes neuronales cuánticas

En cuanto a las redes neuronales cuánticas, es oportuno mencionar que el desarrollo en la computación cuántica, viene en ascenso y cada vez toma más fuerza para realizar investigaciones en esta área, cada vez hay más acercamientos entre las redes neuronales clásicas y la posibilidad de desarrollar algoritmos y sistemas aplicando los fundamentos de la mecánica cuántica, pues esta teoría física, ha mostrado alto grado de precisión a la hora de analizar fenómenos físicos, lo que resulta atractivo para aplicar en la redes neuronales artificiales, además se crea una oportunidad para comprender algunos fundamentos clave que se están aplicando para el desarrollo de la computación cuántica, a pesar de ser una teoría muy potente también es una de las que más conceptos abstractos contiene, así bien, al desarrollar una aplicación a un cuerpo en auge como la inteligencia artificial, da la satisfacción de desentrañar el mundo de la cuántica en aspectos de alto interés en la

ciencia moderna y de esta forma reducir el nivel de abstracción de la mencionada teoría (EzhoV & Ventura, 2000).

3.1. Conceptos fundamentales de la mecánica cuántica

La enseñanza de la mecánica cuántica se encamina a las aplicaciones que se desarrollan dentro de la computación cuántica, para comprender a detalle este campo en surgimiento es necesario abordar varios conceptos fundamentales de la teoría física además de hacer un contraste con la computación clásica.

En primer lugar, es necesario mencionar la unidad básica de la información a través de los aspectos clásicos y cuánticos, a partir de allí se desarrolla toda la construcción de cada tipo de computación, para la forma clásica se codifica a través de bit, su representación tiene carácter binario y siempre se puede observar el estado de cada unidad. En el caso cuántico la unidad básica de la información se define a través del qubit o *quantum bit*, a diferencia del primer caso se puede establecer los estados básicos binarios y a su vez estar en estados en superposición, es decir pueden estar varios estados al mismo tiempo.

De esta manera se adentra en el primer concepto clave utilizada en notación de Dirac la superposición lineal de estados, allí se expresan los estados mediante vectores columna $|z\rangle$ o también vectores ket, el otro tipo de vector es análogo a la compleja conjugada y transpuesta del ket conocida como bra $\langle z|$. En los sistemas cuánticos se describe mediante una función de onda que existe en un espacio de Hilbert, esta tiene un conjunto de estados $|\phi_i\rangle$ que forman una base y el conjunto se expresa de la siguiente forma (EzhoV & Ventura, 2000).

$$|\psi\rangle = \sum_i \alpha_i |\phi_i\rangle \quad (6)$$

Los coeficientes correspondientes a α_i en casos generales pueden ser complejos, en la mecánica cuántica el espacio de Hilbert y sus estados base, contribuye un aspecto fundamental, mientras se esté en escala cuántica, el sistema puede estar en varios estados a la vez, pero no es observable y al momento de realizar la observación el sistema siempre colapsa en uno de sus estados base. Al aplicar los conceptos anteriores, el qubit se puede manifestar en dos estados base $|0\rangle$ y $|1\rangle$ que cumplen con las características de los estados observables y forman base en el espacio de Hilbert. Dejan de ser observables cuando estos entran en una combinación lineal y correspondientemente se configuran en dos conceptos nuevos: la coherencia y la decoherencia, el cual están estrechamente relacionados con la superposición lineal (EzhoV & Ventura, 2000).

$$\alpha_0|0\rangle + \alpha_1|1\rangle \quad (7)$$

Un sistema cuántico es coherente cuando está en una superposición lineal de sus estados base Ec.(7), cuando el sistema interactúa con un entorno dicha superposición tiene una pérdida de coherencia lo cual termina en el colapso de uno de sus estados base, a esta pérdida se conoce como decoherencia y termina gobernada por el sistema cuántico (qubit).

Los coeficientes α_i se designan amplitudes de probabilidad lo cual $|\alpha_i|^2$ es la probabilidad de colapsar en uno de los estados base, la suma de los coeficientes debe ser igual a la unidad como única condición, para determinar la probabilidad del qubit en superposición y colapso en un estado propio se da por la relación $|\langle \phi_i | 0 \rangle|^2$ análogo al producto punto de dos vectores (EzhoV & Ventura, 2000).

Para la construcción de compuertas lógicas cuánticas, el análogo correspondiente en la teoría cuántica, se da a través de los operadores representados por matrices que actúan sobre los vectores y al ejecutarse crean bases en un espacio de Hilbert que permite la transformación de un estado cuántico en otro.

$$\hat{A}|\phi_i\rangle = a_i|\phi_i\rangle \quad (8)$$

El uso de un operador se puede ver reflejado en la ecuación (8), donde a_i es un valor propio y se obtienen soluciones de estados propios que permiten construir bases en el espacio de Hilbert (EzhoV & Ventura, 2000).

3.2. Sistema diferenciable cuántico análogo al modelo del perceptrón clásico

En la sección anterior se definió el perceptrón clásico, que se puede resumir en tres aspectos, datos de entrada, creación del modelo y el dato de salida del modelo de la red neuronal, en cuántica, se tienen presente los mismos tres ítems Fig 6. donde los parámetros de ingreso serán datos de carácter clásicos x_i que se transforman a través de ángulos de rotaciones Ec. 9 en estados cuánticos $|x\rangle$ listos para ser ingresados al estado de preparación de la estructura neuronal, dichas rotaciones son conocidas como las puertas de Pauli (Gomez, 2010) que representan un estado de rotación de π radianes respecto a cada uno de los ejes.

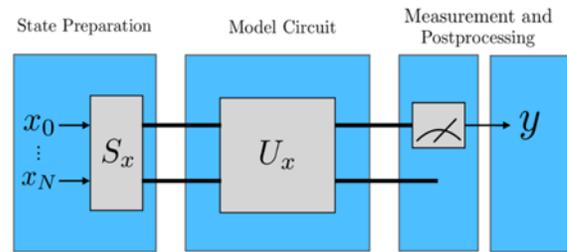


Figura 6. Modelo de red neuronal cuántico.

Fuente: (Schuld et al., 2020)

El número de qubits estará definido por el número de características del patrón de entrada, por lo que el estado conjunto de un sistema formado por N qubits se describe como un punto en el espacio de Hilbert de dimensión 2^N , de esa manera se puede definir estados de varios qubits como productos tensoriales (López, 2022).

$$\begin{aligned}
 R_x(\phi) &= \begin{bmatrix} \cos\left(\frac{\phi}{2}\right) & -i\sin\left(\frac{\phi}{2}\right) \\ -i\sin\left(\frac{\phi}{2}\right) & \cos\left(\frac{\phi}{2}\right) \end{bmatrix} & \sigma_x &= \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \\
 R_y(\phi) &= \begin{bmatrix} \cos\left(\frac{\phi}{2}\right) & -\sin\left(\frac{\phi}{2}\right) \\ \sin\left(\frac{\phi}{2}\right) & \cos\left(\frac{\phi}{2}\right) \end{bmatrix} & \sigma_x &= \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} \\
 R_z(\phi) &= \begin{bmatrix} e^{-\frac{i\phi}{2}} & 0 \\ 0 & e^{\frac{i\phi}{2}} \end{bmatrix} & \sigma_x &= \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}
 \end{aligned} \tag{9}$$

Luego del estado de preparación se encuentra como segundo parámetro el circuito cuántico, donde el modelo a aplicar varía dependiendo las necesidades de clasificación, es aquí, donde se reemplaza el modelo de red neuronal por un circuito cuántico compuesto por puertas cuánticas.

Se debe recordar que la representación de un estado se describe por 2^N qubits, de tal manera, que, al intentar modificar su estado con una puerta cuántica, dichas compuertas deben estar representadas como matrices cuadradas $2^N \times 2^N$, es decir de esta manera, existen tantas compuertas cuánticas, como posibles combinaciones se puedan aplicar bajo dicha representación. Para este artículo, se explicará las puertas más conocidas para la creación de un circuito diferenciable, empezando con la puerta de Hadamard, ésta actúa sobre un único qubit que transforma el estado base $|0\rangle \rightarrow |+\rangle$ y el estado $|1\rangle \rightarrow |-\rangle$, su representación está dada por la siguiente matriz (EzhoV & Ventura, 2000).

$$H = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \tag{10}$$

Seguido por las puertas de Pauli, que definirán una rotación ciertos grados dependiendo su eje, las cuales están definidas por las siguientes matrices.

También se encuentran las puertas controladas, en este caso una de las más reconocidas es la puerta CNOT, que se encarga de crear una relación o controlador entre los qubits para su posterior medida en uno de los cables que representa cada qubit, en otras palabras, cuando el primer qubit es cero el segundo qubit queda igual (el control entre $|00\rangle \rightarrow |00\rangle$, el control entre $|01\rangle \rightarrow |01\rangle$), sin embargo cuando el primer qubit es uno su valor cambia en base a la puerta X(NOT) (el control entre $|10\rangle \rightarrow |11\rangle$, el control entre $|11\rangle \rightarrow |10\rangle$

$$CNOT = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix} \tag{12}$$

Por último, se encuentra el aspecto de la medida y del resultado. La medida en el formalismo cuántico es un problema que aún no se define, gracias a ello, hay diferentes tipos de interpretaciones, sin embargo, para este caso, un perceptrón cuántico, se utilizara el concepto de medida de Max Born, el cual postula, que la probabilidad de encontrar un objeto cuántico en un lugar determinado es proporcional al cuadrado de la función de onda para un observable en este caso σ_z (Kaye et al., 2007).

$$\langle \sigma_z \rangle = \langle \psi | U^\dagger(\theta) \sigma_z U(\theta) | \psi \rangle \tag{13}$$

3.3. Circuito cuántico

Como se explicaba en la sección anterior cualquier circuito cuántico es aplicable, siempre y cuando, cumpla con los parámetros de clasificación, hasta la fecha se han implementado algunos tipos de circuitos, siendo los más eficientes los que utilizan clasificaciones por kernel, clasificadores variacionales, modelos recurrentes, modelos de redes neuronales convulsionales entre otros, sin embargo, para llegar hablar de una red neuronal, se debe primero definir el perceptrón cuántico parte fundamental de la red, de tal modo que, el circuito utilizado para la reemplazar el modelo clásica está dado por Fig. 7, donde los características del patrón ingresadas por qubits pasan por una primera capa donde están las compuertas de rotación y las compuertas CNOT que definen en qué qubit se realizará la medición bajo el observable de σ_z .

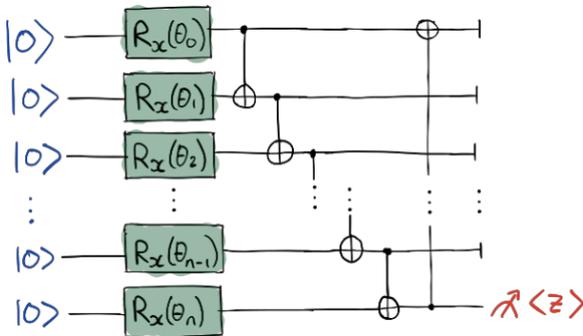


Figura 7. Circuito neuronal aplicado

Fuente: (PennyLane, 2023)

3.4. Algoritmo cuántico

Cómo el modelo que se representa es un modelo híbrido, es decir, una mezcla de computación clásica y cuántica la librería PennyLane maneja un concepto llamado nodo cuántico, donde un nodo consta de una función cuántica y un dispositivo donde se ejecuta.

Ahora bien, para la construcción de un modelo análogo a la red neuronal, se plasma el modelo de la Ec. 14. donde, R_x el operador sigma que actúe sobre

los datos los parámetros a entrar, es decir los pesos sinápticos y el *bias* (Wittek, 2014).

$$f_{NN}(x; \theta) = \sigma_z(Wx + b) \quad \theta = \{W, b\} \quad (14)$$

La función de coste, que estará dada por Ec.9, la cual empieza a determinar lo lejos que se encuentra el modelo de predecir un valor esperado.

$$C = \sum_{data\ x,y}^n |f(x; \theta) - y|^2 \quad (15)$$

El “*Gradient descent*” Ec. 15, el cual optimiza los parámetros en la dirección del gradiente buscando el punto de convergencia a un punto mínimo en cada iteración, reescribiendo de esa manera los parámetros θ para el ajuste del modelo. Se debe encontrar un análogo cuántico, ya que el “*Gradient Descent Clásico*” está definido en un espacio euclidiano, por tal razón se es de completa necesidad agregar un valor g^+ que representa un tensor métrico invariante conocido como el tensor métrico de Fubini-Study (Stokes et al., 2020), esta expresión da actualización de pesos, también tendrá una razón de aprendizaje o “*Learning rate*” que se puede expresar como un valor aleatorio.

$$\theta^{(t+1)} = \theta^{(t)} - \eta g^+(\theta_t) \nabla_{\theta} C \quad (16)$$

Ahora bien, bajo un sistema de clasificación binario de 200 datos, encontramos que con un circuito cuántico se obtiene una mejor clasificación, ya que, se guarda los parámetros trigonométricos en la medida, es decir, que se podría llegar a realizar un paralelismo con la función de onda $f(x) = \text{sen}(x)$ Fig. 8. Lo que da como resultado son regiones de ondas superpuestas, donde los pesos sinápticos entran ajustar la longitud de onda en cada iteración o época, de esa manera, se ajusten los máximos y los mínimos de las funciones y logren una clasificación binaria entre una clase A y una clase B.

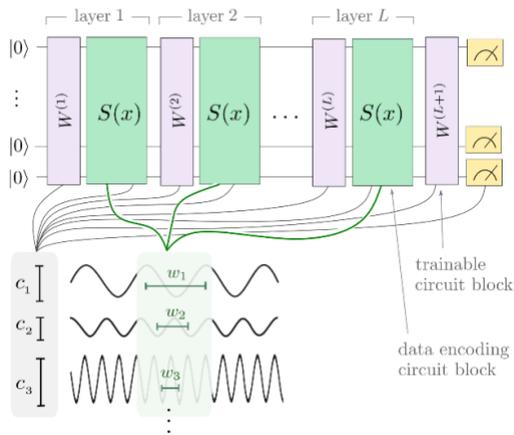


Figura 8. Clasificación de Fourier.

Fuente: (Pennylane, 2023)

Por lo que, para unos datos que presentan una distribución no calificable en un modelo lineal clásico Fig. 9. se aplica el algoritmo de circuito diferenciable presentando en líneas anteriores, donde la superposición constructiva entre dos ondas representase los datos de la clase A, y la superposición de manera destructiva representa la clasificación de clase B.

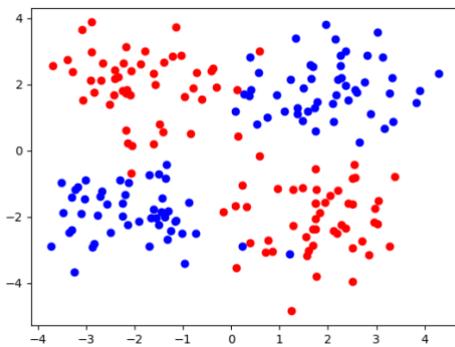


Figura 9. Datos de entrenamiento.

Fuente: Python

4. Resultados

Luego de aplicado el código se realiza un entrenamiento representando las diferentes épocas, en el caso de la época cero Fig. 10, se observa que el modelo aún no tiene ningún tipo de ajuste, por lo que su clasificación presentara una función de pérdida más elevada, luego de 100 épocas Fig. 11 se representa un ajuste mejor a los datos con una función de pérdida de 0.18, por lo cual al ser cercana a cero, se puede llegar a afirmar que el modelo da una clasificación correcta a los datos de entrada.

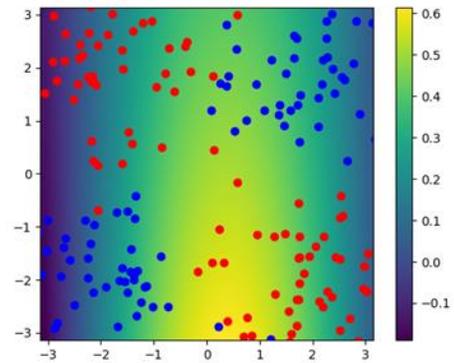


Figura 10. Datos en época 0.

Fuente: Python

Comparando de esta manera una clasificación binaria entre el perceptrón clásico y el circuito diferencial Fig. 12, se tiene que para el circuito cuántico un mejor aprendizaje por cada época y un menor valor de pérdida para la época final de entrenamiento.

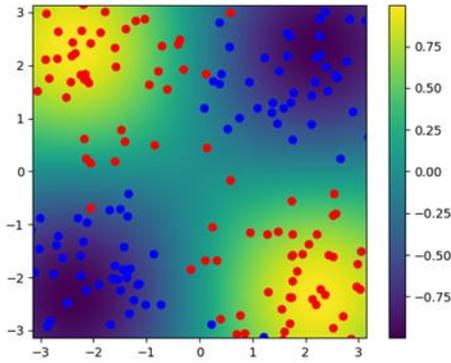


Figura 11. Datos de entrenamiento Época 100

Fuente: Python

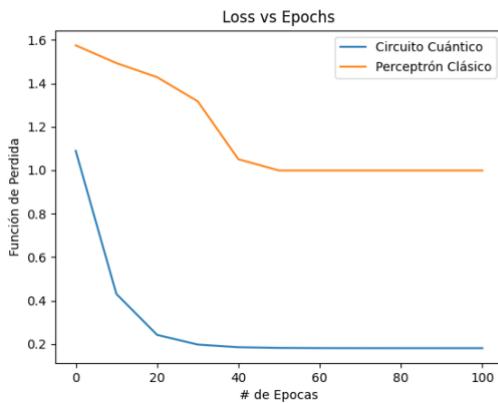


Figura 12. Perceptrón clásico vs circuito cuántico. Fuente: Python

4. Conclusiones y consideraciones finales

Luego de hacer una revisión al marco teórico sobre las redes neuronales, se pudo evidenciar varios factores que acercan al entendimiento del cerebro en términos de procesamiento de información, la estructura matemática de la redes neuronales artificiales permite comprender procesos similares, por caso la clasificación de datos y su capacidad predictiva permite obtener resultados en tiempo real y con alto índice de eficacia, por tanto, al examinar los aspectos básicos de un perceptrón se logra conectar con los aspectos teóricos

desarrollados desde la computación cuántica, la manipulación de información que parte de un sistema cuántico qubit y la superposición lineal, el cual, optimiza y amplía el espectro de manejo de datos en forma exponencial.

La definición de red neuronal cuántica es debatible por muchos sectores que trabajan en la ciencia de datos, ya que, lo que define una red neuronal, es su modelo matemático presentado en la primera sección, de modo que al cambiar la función modelo por un circuito cuántico, se plantea es un sistema diferenciable capaz de procesar algoritmos equivalentes a una red neuronal clásica como, la función de pérdida, la función de costo y el descenso por gradiente, que permite un ajuste a los parámetros que se ingresan a las compuertas cuánticas para la clasificación de datos en distintas clases, ahora bien, otro parámetro es el dato de salida, en la parte clásica, se crea una función sigmoide que permite la activación de la neurona artificial, sin embargo, en su análogo cuántico se tiene una medida de su estado siempre a través del colapso de la función de onda, con la probabilidad de la amplitud al cuadrado, que se convierte en el clasificador, por lo tanto no se podría llegar hablar de un inhibición o activación de una neurona cuántica.

La construcción de los algoritmos clásico, como cuántico, permitió comprender a detalle el funcionamiento, ejecución de los conceptos y parámetros matemáticos, las muestra de clasificación y predicción permite ver la reducción de errores y la eficacia de los modelo para aplicaciones en el campo científico, por lo cual para la enseñanza de la física teórica y la física computacional, observar los conceptos aplicados al campo de la inteligencia artificial genera una mejor comprensión de los conceptos que tienen carácter abstracto y matemático, además la implementación del formalismo cuántico ofrece una perspectiva diferente en torno a la concepción básica de la información, la capacidad almacenamiento y su procesamiento exponencial permite que la proyección a futuro, se solidifique ya que desde la

concepción teórica se acerca a al postulado de Turing de crear una máquina capaz de replicar el procesamiento de un cerebro humano.

6. Referencias

- Caparrini, F. S. (18 de Mayo de 2023). *Entrenamiento de Redes Neuronales: mejorando el Gradiente Descendente*.
<http://www.cs.us.es/~fsancho/?e=165>
- Corte, I. R. (2021). *Modelado de sistemas magnéticos*. [Tesis de grado. Universidad Nacional De La Plata].
- Escámez, M. R. (2022). *CLASIFICACIÓN AUTOMÁTICA DE FRUTAS A PARTIR DE IMÁGENES CON PHYTON. TÉCNICAS DE TRANSFERENCIA DE APRENDIZAJE A TRANSFERENCIA DE APRENDIZAJE A PARTIR DE CONJUNTOS DE DATOS PEQUEÑOS*. [Trabajo fin de grado, Universidad Miguel Hernandez de Elche].
- EzhoV, A. A., & Ventura, D. (2000). Quantum Neural Networks. *Future Directions for Intelligent Systems and Information Sciences: The Future of Speech and Image Technologies, Brain Computers, WWW, and Bioinformatics*, 213-235.
- Gomez, M. (2010). *Introducción al formalismo de la mecánica cuántica no relativista*. Universidad Nacional de Colombia.
- Google inc. (18 de mayo de 2023). *Pennylane*.
https://pennylane.ai/qml/demos/tutorial_expressivity_fourier_series.html
- Hopfield, J. (1988). Artificial Neuronal Networks. *IEEE CIRCUITS AND DEVICES MAGAZINE*, 4(5), 3-10.
- Izaurieta, F., & Saavedra, C. (2000). Redes Neuronales Artificiales. *Departamento de física, Univerisdad concepcion de chile*.
- Kandel, E. (1955). Kandel, E. (1995). Essentials of neural science and behavior. *Appleton&Lange Norwalk*, 6(23), 425-451.
- Kaye, P., Laflamme, R., & Mosca, M. (2007). *An Introduction to Quantum*. Great Britain: Oxford Express.
- Li, L., DOROSLOVAČKI, M., & LOEW, M. H. (2020). Approximating the Gradient of. *Approximating the Gradient of Cross-Entropy Loss Function*, 8, 111626-111635.
<https://doi.org/10.1109/ACCESS.2020.3001531>
- López, R. (2022). *INTRODUCCIÓN A LA COMPUTACIÓN*. [Trabajo fin de grado, UNIVERSIDAD COMPLUTENSE DE MADRID].
- Pennylane. (17 de mayo de 2023). *Pennylane*.
https://pennylane.ai/qml/demos/tutorial_expressivity_fourier_series.html
- Rosenblatt, F. (1957). The perceptron, a perceiving and recognizing automaton Project Para. *Cornell Aeronautical Laboratory*.
- Schuld, M., Bocharov, A., Svore, K. M., & Wiebe, N. (2020). Circuit-centric quantum classifiers. *PHYSICAL REVIEW A*, 101(3).
<https://doi.org/https://doi.org/10.1103/PhysRevA.101.032308>
- Stokes, J., Izaac, J., Killoran, N., & Carleo, G. (2020). Quantum Natural Gradient. *Quantum*, 4, 269.
<https://doi.org/https://doi.org/10.22331/q-2020-05-25-269>
- Turing, A. M. (1936). On computable numbers, with an application to the Entscheidungsproblem. *J. of Math*, 58(5), 345-363.
- Wilkinson, C. (2020). *Ciencia de Datos Python: Una guía definitiva para que los principiantes aprendan los fundamentos de la ciencia de datos con Python (Libro En Español/Self Publishing Spanish Book Version)*. Independently Published.
- Wittek, P. (2014). *Quantum Machine Learning What Quantum Computing Means to Data Mining*. Academic Press.