Universidad Nacional Autónoma de México

Facultad de Ciencias

Una red neuronal aprende a contar en distintas bases: el caso del Perceptrón Multicapa

T E S I S

QUE PARA OBTENER EL TÍTULO DE:
LICENCIADA EN CIENCIAS DE LA COMPUTACIÓN
PRESENTA:

ABRIL BEATRIZ MUÑOZ QUIROZ

DIRECTOR DE TESIS:
JOSÉ ANTONIO NEME CASTILLO

2010

1. Datos del alumno Apellido paterno Apellido materno Nombre(s) Teléfono Universidad Nacional Autónoma de México Facultad de Ciencias Carrera Número de cuenta	1. Datos del alumno Muñoz Quiroz Abril Beatriz 56 81 88 27 Universidad Nacional Autónoma de México Facultad de Ciencias Ciencias de la Computación 0-9808398-5
2. Datos del tutor Grado Nombre(s) Apellido paterno Apellido materno	2. Datos del tutor Dr. José Antonio Neme Castillo
3. Datos del sinodal 1 Grado Nombre(s) Apellido paterno Apellido materno	3. Datos del sinodal 1 Dr. Pedro Miramontes Vidal
4. Datos del sinodal 2 Grado Nombre(s) Apellido paterno Apellido materno	4. Datos del sinodal 2 M. en C. Francisco Struck Chávez
5. Datos del sinodal 3 Grado Nombre(s) Apellido paterno Apellido materno	5. Datos del sinodal 3 L. en C. C. Francisco Lorenzo Solsona Cruz
6. Datos del sinodal 4 Grado Nombre(s) Apellido paterno Apellido materno	6. Datos del sinodal 4 Mat. Vicente Carrión Miranda
7.Datos del trabajo escrito. Título Subtitulo Número de páginas Año	7. Datos del trabajo escrito. Una red neuronal aprende a contar en diferentes bases. El caso del Perceptrón Multicapa 57 2010

A mis padres

Agradecimientos

Recuerdos y añoranzas. Lo primero que llega con una mezcla de nostalgia y alegría es la huelga; el evento más significativo, un árbol cargado de emociones y hazañas que marcaron la (a veces) monótona vida de estudiante y la pintaron con colores que aún ahora dejan ver su intensidad. Así que agradezco a ese ente amorfo, apasionante, vivo y desconcertante que fue la huelga.

Se me hace nudo la garganta cuando recuerdo a las personas que quise, quiero y conocí en la escuela. Son tantas. Invoco a algunos plasmando su nombre, como un ritual de manifiesto cariño.

Apenas si pienso tu nombre y se me ocurre escribir amor, camino, noches, sueños, vida, ternura, pasión, peleas, encuentro, espera, nostalgia, abrazo, escucha, amor, otra vez amor... y gracias, gracias Sergio por ser la certeza de un presentimiento.

Lucía y Mauro gracias por compartir mi felicidad en la conclusión de esta travesía.

Nico mi carnavalito, por el remanso, la compañía y la autenticidad que ha marcado nuestra amistad.

A mis hermanos por su insospechada enseñanza.

Claudia, Nelly, Genaro, Antonio, Juan, Ulises, Víctor, Beto, Fidel, David, Cindy, Celia, Ivan, Cuadros, Ericka, Rock, Cristi, Abel, Pris, Mónica, porque los pienso y he tejido con ustedes trozos de mi vida. Y ese tejido abriga y calienta mi andar, me hace sonreir, recordar su voz, imaginar su rostro y revivir momentos.

Los maestros que con su hacer lograron contagiarme el gusto por el saber: Anguiano, Adela, Emilio, Rodolfo, Luz, Neme y Rodrigo.

Por supuesto a M. A. la imagen móvil que poco a poco se esfuma pero no desaparece.

En fin, los quiero a todos.

Abril

Índice general

Introducción

Resulta sorprendente la cantidad de datos, conocimiento y experiencia que el cerebro es capaz de almacenar, así como la versatilidad de los temas que maneja. A través de la historia su estudio ha cautivado a filósofos, médicos, psicólogos, pedagogos y estudiosos de diversas áreas, que desean comprender el origen y procesos de las creaciones científicas, tecnológicas y artísticas que este órgano orquesta.

Las máquinas de cómputo más sofisticadas en su respectiva época[1] han servido para que el hombre compare su funcionamiento con el cerebro. En nuestra época el dispositivo más cercano a las funciones de este órgano es la computadora, incluso es llamada *cerebro electrónico*.

El conocimiento que actualmente tenemos ha sido fuente para la creación de modelos que tratan de imitar sus procesos. Los humanos nos encontramos ante algo que nos acompaña desde siempre y nos conforma, pero que nos resulta extremadamente complejo entender y analizar a profundidad.

Se han desarrollado métodos neuroquímicos, electrofisiológicos, conductuales y de imagenología para el estudio del sistema nervioso, cada método aporta información importante para el análisis global del funcionamiento de este órgano.

Aunque aún no existen teorías que expliquen totalmente su funcionamiento, los modelos que ha inspirado muestran interesantes caminos para el desarrollo de la inteligencia artificial.

Los diversos algoritmos basados en el cerebro como modelo, resuelven múltiples problemas prácticos y a la fecha son usados con éxito en diversas áreas.

Se reconoce que "[...]un sistema para el procesamiento de la información con cinco mil millones de años de experiencia en el diseño, debe ser tomado seriamente como una fuente de buenas ideas."[Anderson, 2007, p. 289]

La estructura neuronal del sistema nervioso descubierta por Santiago de Ramón y Cajal en 1888 inicia la historia de las redes neuronales, las neuronas son las unidades fundamentales del sistema nervioso y desde el punto de vista funcional son procesadores de información [Del Brío, 2001] y son la base del aprendizaje de los animales.

El estudio de sus componentes, principalmente las neuronas y sus sinapsis condujeron a la arquitectura actual de las computadoras. [Rojas, 1996]

[1] Las comparaciones se han establecido con máquinas de cálculo, relojes, redes telefónicas e incluso en estudios recientes se ha comparado su topología con sistemas complejos tales como las redes sociales. sakata

Los modelos de redes neuronales artificiales están inspirados en los componentes y funcionamiento del cerebro, por lo tanto una de las cuestiones de interés es el estudio de las semejanzas en el aprendizaje entre éstas y el cerebro.

Estos modelos han ganado un lugar a lado de epistemología, la psicología cognitiva y la lingüística en la tarea de comprender los procesos de aprendizaje.

En nuestro trabajo usamos un modelo de red neuronal[2] conocido como **perceptrón multicapa** que contiene los cuatro componentes funcionales de la representación fisiológica de la neurona: componente de entrada, componente desencadenante, componente de conducción y de salida.

Nuestra tarea es analizar los errores y aciertos de la red neuronal cuando se enfrenta a aprender a contar en distintas bases numéricas.

Las pregunta inicial en la investigación es: *¿existe una base óptima para aprender a contar?* y de ser así, *¿cuál es esa base?*

Las redes neuronales son un recurso idóneo para experimentar ya que si tratamos de implementar alguna nueva técnica de enseñanza necesitaríamos esperar varios años para poder analizar si ésta propuesta repercutió positivamente en el desarrollo de los estudiantes.

Las pruebas en redes neuronales se pueden correr las veces necesarias, monitorear los errores, hacer modificaciones pertinentes al programa, esto se logra en un periodo de tiempo muy corto y lo mejor de todo es que ningún niño tiene que sufrir los estragos de los experimentos.

Aunque existen investigaciones que han buscado identificar la similitud en el aprendizaje las redes neuronales son un camino poco explorado y explotado para su uso en pedagogía. Esto repercute de manera negativa en el conocimiento de los posibles paralelos en estas dos áreas ya que su estudio es fuente de ideas que repercuten en una y otra, pues la comprensión de los procesos cognitivos permite el desarrollo de nuevos modelos de redes neuronales y viceversa.

Consideramos que el aprendizaje de las matemáticas elementales es un factor trascendental que debe ser analizado para proponer mejoras. Ahí radica la utilidad de experimentar en redes neuronales artificiales.

Con el fin de mostrar las dificultad para enseñar y aprender matemáticas[3] en el capítulo 1 discutimos el fracaso escolar y los resultados de pruebas aplicadas a estudiantes de nivel básico, también mostramos dos teorías pedagógicas que señalan cómo el desarrollo de las habilidades matemáticas estructura y posibilita el aprendizaje en general y la última sección del capítulo trata acerca de la importancia de aprender a contar.

En el capítulo 2 se detalla el fundamento biológico de las redes neuronales en general y de los perceptrones multicapa en particular. En el capítulo 3 se presentan los experimentos y se muestran los resultados, al tiempo que se da una interpretación a la luz del aprendizaje de las redes neuronales.

[2] En lo sucesivo a las redes neuronales artificiales las denotaremos simplemente como redes neuronales.

[3] En este trabajo aprender matemáticas significará "[...] una verdadera comprensión de la matemática como un todo orgánico y como base para el pensamiento y la acción científicos."(Courant, 1955, p. IX)

Finalmente presentamos las conclusiones donde se resumen los puntos importantes y el trabajo futuro en el que enfatizamos las posibilidades de las redes neuronales como modelos de aprendizaje en humanos.

Capítulo 1

Fracaso escolar en matemáticas

Lo único que interfiere con mi aprendizaje es mi educación

Einstein

Dada la importancia que las matemáticas han tenido en el desarrollo de todas las culturas humanas podemos afirmar que es fundamental crear soluciones ante la problemática en su enseñanza y aprendizaje.

Preguntamos a un grupo de matemáticos que dedican parte de su tiempo a la enseñanza a nivel medio y superior, acerca de las razones que hay para aprender matemáticas.[1]

Cada uno de los entrevistados está convencido de la indiscutible utilidad de esta ciencia y entre sus argumentos estan los siguientes:

“Aprender matemáticas permite estructurar una forma de pensar propicia para la solución de problemas, así como desarrollar el pensamiento abstracto y razonamiento.”

Otro motivo para estudiar matemáticas es observar su interrelación con otras áreas de conocimiento[2], se considera que “forma parte intrínseca de muchas de ellas y que los avances logrados en una complementan y eventualmente repercute en la otra.”

Aunque estos argumentos son reconocidos[3] sabemos que existe un recelo que comparten muchos estudiantes y educadores hacia *la verdadera utilidad de las matemáticas.*

El trabajo en el aula muestra que es común la pregunta *¿para qué me sirven las matemáticas?*, es decir, *¿matemáticas para qué?*

Tratemos de esbozar una respuesta.

La historia nos enseña a mirar el desarrollo y las aplicaciones de ésta ciencia.

Sin embargo, ocurre algo que desconcierta en el aprendizaje de las matemáticas: la

[1]Pregunta formulada en Primer encuentro de *La Ciencia en tu escuela*, noviembre 2009, organizado por la Academia Mexicana de Ciencias (AMC) memorias en elaboración.

[2] Debemos recordar que la división en las áreas conocimiento se ha ido modificando de acuerdo al desarrollo de las sociedades y a la amplitud y desarrollo de las mismas, así nos encontramos ahora que la enseñanza superior se enfoca a la especialización ya no en alguna de las áreas que fueron resultado de la división del saber, sino las más de las veces en alguna de las ramas de las mismas.

[3]Esta pregunta tuvo casi las mismas respuestas cuando preguntamos a profesores de otras áreas durante el Diplomado *La Ciencia en tu escuela* impartido por la AMC.

sistematización lógica ha encubierto el desarrollo epistemológico que dio origen a gran parte de las matemáticas que son enseñadas actualmente en la escuela.

Dicha sistematización es hoy día una forma de *aprender matemáticas* que se encuentra extendida mundialmente, se aplica o al menos se trata de aplicar en todos los niveles escolares.

Es importante señalar que esta manera de aprender no se corresponde con la forma en la que la mayor parte de las creaciones matemáticas han tenido lugar, se observa por el contrario que dicha aspiración sistemática tiene lugar luego de un desarrollo práctico y teórico de los saberes matemáticos.

El primer ejemplo que encontramos lo proporciona Euclides quien en su tratado los *Elementos* organiza el saber matemático de su época y con ello se convierte en el referente para diversos matemáticos cuyas aportaciones son trascendentales.[4]

Se dice de su tratado: "La obra de Euclides durante mucho tiempo fue una fuente de inspiración para los matemáticos y es el fundamento y la guía para algunas partes de las matemáticas más finas incluso hasta la época presente". [Gemignani, 1967]

De esta manera encontramos desde 300 a.n.e. la expresión de una necesidad compartida por los matemáticos acerca de la sistematización que siglos más tarde encontramos en matemáticos como David Hilbert[5](1862-1943) y Giuseppe Peano (1858 –1932), con ello debemos considerar lo siguiente:

Partimos de reconocer uno de los objetivos de las matemáticas: la creación de un sistema axiomático. Este saber axiomático parte de la generalización de conceptos y algoritmos que el hombre inventó, en principio para facilitar operaciones, teorías y demás aplicaciones matemáticas.

De lo anterior podemos observar que los programas educativos pretenden exponer los saberes matemáticos de acuerdo a esta organización axiomática,[6] lo que origina conocimientos desconectados entre sí, dificultando la aprehensión de conceptos que el hombre tardó cientos de años en crear y organizar.[7]

Si reconocemos la necesidad de establecer conexiones entre los conocimientos que aprendemos estaremos en posibilidad de crear otra forma de enseñanza.[8]

---

[4]Claudio Ptolomeo (*ca.* 85-165), Girolamo Saccheri (1667-1773),Carl Friedrich Gauss (1777-1855) y Bernhard Riemann (1826-1866) por mencionar sólo algunos de quienes fueron estudiosos de su trabajo, en particular el estudio del quinto postulado.

[5] Se recorrió un largo camino antes de lograr la axiomatización de la geometría, en donde se buscaba despegar la interpretación intuitiva de los términos y hacer énfasis en las relaciones entre ellos. Hilbert es reconocido entre otras cosas por inventar la teoría de la prueba, misma que en el origen de la axiomatización no figuraba de manera clara para él y que hoy se exige sin tomar en cuenta las rupturas epistemológicas necesarias para llegar a ella. [Rodríguez-Consuegra, 2001]

[6]Esta situación es clara si se observa los planes y programas de estudio de las instituciones de educación superior (o si entramos a alguna clase en las mismas) que dan una formación matemática, en ellas el formalismo impera.

[7] Rodrigo Cambray habla de esto para el caso de la geometría "Debe causarnos admiración que la lucha por lograr que la geometría se convirtiera en una *mathema* tomara a la humanidad aproximadamente 2500 años." [Cambray, 2003]

[8]Podemos incluso agregar que desde el punto de vista de las redes neuronales el aprendizaje tiene que ver con las conexiones que son reforzadas.

Sin embargo, niños y adultos son capaces de crear sus propios algoritmos,[9] lo cual indica que la generalización surge también como necesidad y es consecuencia de las matemáticas que ellos hacen; este hecho debe apuntar hacia la cuestión de cómo esto puede ser usado en la enseñanza para contribuir a la maduración, desarrollo y evolución de sus conocimientos; con tales fundamentos el presente trabajo se enfoca en un tema crucial de las matemáticas básicas: *aprender a contar*.

La posible respuesta tendría que pasar entonces por las siguientes consideraciones:

Una mirada a la historia dará cuenta de la amplitud y utilidad que ha representado para el ser humano el conocimiento matemático.

Las matemáticas no son un evento histórico que ya pasó.

Los encargados de crear tales conocimientos matemáticos son los seres humanos y aunque la escuela no determine dicha capacidad de creación, puede favorecerla y es deseable que así lo haga, entonces es importante ocuparnos del acercamiento que los niños y adolescentes tienen con la ciencia en general y por ende con las matemáticas.

La contribución que hagamos repercutirá en beneficio de todos.

Existe, sin embargo otra razón que los apasionados de las matemáticas argumentan para su estudio y que no debe soslayarse: *entender matemáticas es una experiencia sumamente placentera.*

Lo anterior contiene otro fundamento por el cual las matemáticas se han desarrollado, es necesario enfatizar que si bien los planes y programas de estudio, así como la innovación y análisis acerca de cómo estudiamos y enseñamos matemáticas puede y debe ser mejorado también es necesario señalar enfáticamente que para entender las matemáticas es menester tomar gusto por el estudio analítico del saber científico y la experiencia personal de los estudiantes de crear este saber podría reforzar el gusto por las matemáticas.

Cuando hablamos de considerar el desarrollo epistemológico de los conceptos lo hacemos para evidenciar el esclarecimiento que el estudiante podría experimentar al contextualizarlos.

Por último compartimos con Terezhina Nunes y Peter Bryant la razón fundamental del aprendizaje de las matemáticas " [...] en nuestra opinión lo primero y -más importante- que debemos puntualizar es que los niños y las niñas necesitan aprender matemáticas para comprender el mundo que los rodea ".[Nunes y Bryant, 2003, p.10]

1.1. Matemática educativa: fracaso nacional

La enseñanza y aprendizaje de las matemáticas representa un problema educativo que no se ha podido resolver satisfactoriamente en nuestro país.

Para mostrar de esta situación podemos consultar los resultados de las diversas pruebas aplicadas por organizaciones como Programme for International Student Assessment (PISA), la prueba nacional Evaluación Nacional del logro Académico en centros nacionales (ENLACE), creada recientemente por la Secretaria de Educación Pública

[9] Ver trabajo [Nunes, 2001]

(SEP), así como las aplicadas por el también reciente Instituto Nacional para la Evaluación de la Educación (INEE).

Pero quizá la prueba más irrefutable de las dificultades a la hora de aprender matemáticas es la experiencia cotidiana en las aulas y otros centros laborales.

1.1.1. Resultados de pruebas en matemáticas

El primer organismo cuyas evaluaciones están promovidas por la Organización para la Cooperación y el Desarrollo Económicos (OCDE), en su aplicación trianual (2006), puso a prueba a jóvenes de 57 países, con edades de entre 15 años tres meses y 16 años dos meses de edad al momento de la evaluación que se encuentran inscritos en alguna institución educativa, secundaria, media superior o capacitación para el trabajo. La prueba calificó su desempeño en las áreas de ciencias, matemáticas y lectura.

Para su evaluación, la prueba PISA usa el concepto *literacy*, cuya traducción en México es competencia o aptitud, y se define como: "[...]la capacidad de los estudiantes para extrapolar lo que han aprendido y aplicar sus conocimientos y habilidades en nuevos escenarios; así como para analizar, razonar y comunicarse de manera satisfactoria al plantear, resolver e interpretar problemas en diversas situaciones del mundo real". [Díaz et al, 2007, p.18]

La competencia matemática, según este organismo, es la capacidad del individuo para "[...]identificar y comprender la función que desempeñan las matemáticas en el mundo, emitir juicios fundados, utilizar las matemáticas y relacionarse con ellas de forma que se puedan satisfacer las necesidades de la vida de los individuos como ciudadanos constructivos, comprometidos y reflexivos".[Ibid., p. 88]

Los reactivos son de opción múltiple, opción múltiple compleja, respuesta corta, respuesta abierta construida o respuesta corta construida.

A continuación, dos ejemplos de los reactivos de la prueba aplicada [10] en 2003 tomados de [2] :

> Exámenes de Ciencias
>
> En la escuela de Malena, su maestro de Ciencias les aplicó exámenes de 100 preguntas. Malena tuvo un promedio de 60 aciertos en sus primeros cuatro exámenes de Ciencias. En el quinto examen obtuvo 80 aciertos. ¿Cuál es el promedio de aciertos de Malena en Ciencias después de los cinco exámenes?
>
> Promedio:
>
> Tipo de cambio
>
> Mei-Ling, de Singapur, se estaba preparando para viajar a Sudáfrica durante 3 meses como participante en un intercambio estudiantil. Necesitó cambiar dólares de Singapur (SGD) a rands de Sudáfrica (ZAR).
>
> Mei-Ling encontró que el tipo de cambio entre los dólares de Singapur y los rands de Sudáfrica era: 1 SGD = 4.2 ZAR
>
> Mei-Ling cambió 3000 dólares de Singapur a rands sudafricanos a este tipo de cambio.
>
> ¿Cuánto dinero en rands sudafricanos recibió Mei-Ling?
>
> Respuesta:

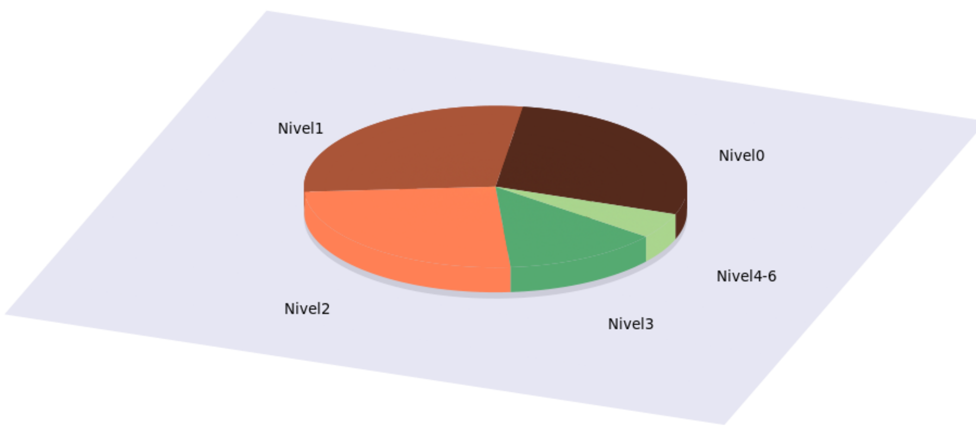
[10] Los reactivos de PISA 2006 no han sido liberados.

Los resultados se clasifican en seis niveles de acuerdo al puntaje obtenido.

La media de los resultados de los estudiantes mexicanos se ubica en el nivel 1, que refiere en este rango a los jóvenes que “[...]pueden contestar preguntas relacionadas con contextos familiares, en los que está presente toda la información relevante y las preguntas están claramente definidas. Son capaces de identificar la información y desarrollar procedimientos rutinarios conforme a instrucciones directas en situaciones explícitas. Pueden realizar acciones obvias que se deducen inmediatamente de los estímulos dados”.[Ibid. p.104]

Resultados de la prueba aplicada[11]:

	Porcentajes
Nivel 0	28
Nivel 1	28
Nivel 2	25
Nivel 3	13
Nivel 4-6	5



El informe específica que México se encuentra dentro de “los márgenes de lo esperado”. Sin embargo, es alarmante que un porcentaje significativo (28 %), se encuentre clasificado en el nivel 0, para quienes se concluye su incapacidad de tener éxito en las tareas más básicas que busca medir PISA. Mientras que sólo 5 alcanza el nivel 4-6 de la evaluación.

[11]La suma de los porcentajes no es 100, la información fue obtenida de [Díaz et al, 2007]

La prueba ENLACE se aplica desde el 2006 a estudiantes de tercero de primaria y tercero de secundaria y en 2008 se aplicó por primera vez a nivel medio superior, pretende evaluar los conocimientos en Español, Matemáticas y Ciencias, según el resultado del trabajo escolar contenido en los planes y programas de estudio oficiales vigentes. Se afirma que fue diseñada bajo estándares internacionales de calidad establecidos para la elaboración de pruebas.

Las preguntas son de opción múltiple. A continuación se muestran dos reactivos de ejemplo de las pruebas aplicadas a tercero de primaria y tercero de secundaria. [3]

Julia tenía 15 pesos. Luego Pati le dio algunos más. Ahora Julia tiene 41 pesos, ¿cuánto dinero le dio Pati a Julia?

A) 66 pesos.
B) 56 pesos.
C) 36 pesos.
D) 26 pesos.

Si una circunferencia mide 53.38 cm, ¿cuál es la medida de su radio si $\pi = 3.14$?

A) 4.25 cm

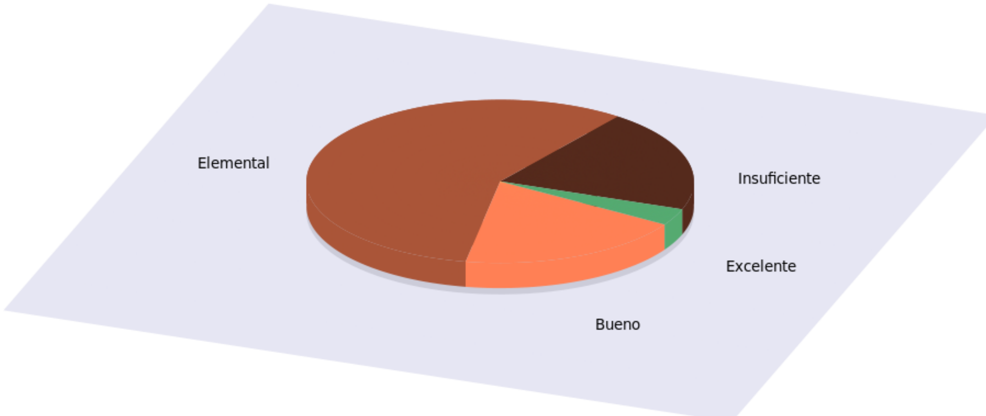
B) 8.50 cm

C) 17 cm

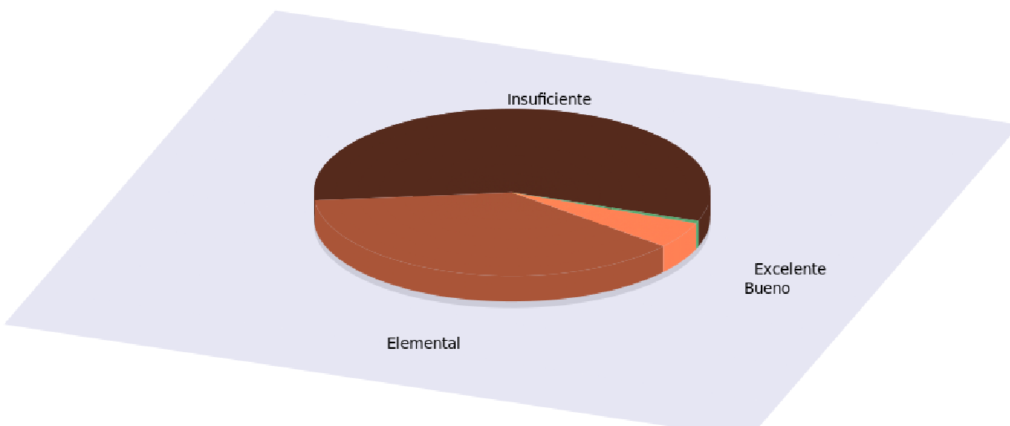
D) 34 cm

Resultados prueba ENLACE-2007 en primaria y secundaria respectivamente [1]:

Grado	Insuficiente	Elemental	Bueno	Excelente
Tercero	21.7	50.8	23.2	4.3
Cuarto	19.1	59.7	18.1	3.1
Quinto	20.4	57.7	18.9	3.0
Sexto	19.7	61.8	15.8	2.8
Medias	20.2	57.5	19.0	3.3



	Insuficiente	Elemental	Bueno	Excelente
Medias	57.1	37.3	5.1	0.5



En el ciclo escolar 2007-2008 se aplicó por primera vez una prueba que pertenece a los Exámenes de la Calidad y el Logro Educativos (EXCALE) a niños de tercero de preescolar, aplicada por el INEE. Esta prueba evalúa los conocimientos en Lenguaje y comunicación y Pensamiento matemático.

El diseño de la prueba consistió en las competencias que establece el Programa de Educación Preescolar, en el se señala que los niños deben adquirir la abstracción numérica y el razonamiento numérico y deben desarrollar su comportamiento espacial, es decir, capacidades de razonamiento que los niños utilizan para establecer relaciones con los objetos y entre los objetos.

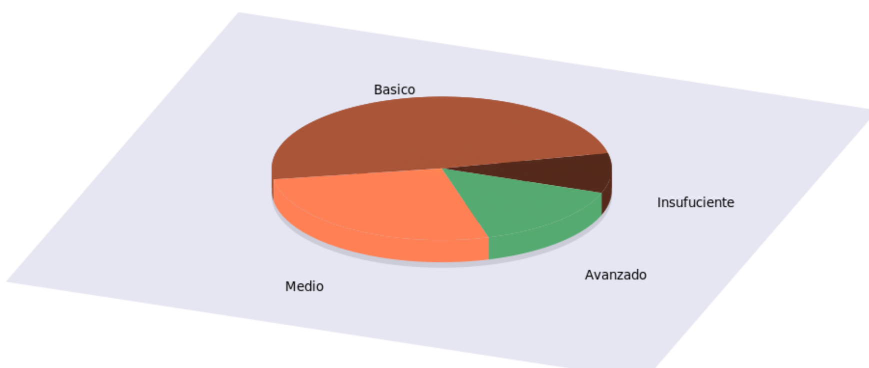
Los reactivos en esta prueba son de dos tipos, de opción múltiple, con y sin texto y de respuesta construida de ejecución manual y de respuesta verbal. La mayoría de los reactivos son leídos y explicados por una evaluadora, no es necesario que los niños sepan leer.

Dos ejemplos de los reactivos correspondientes a nivel medio:

> Se muestra el dibujo de unos niños que tienen que tomar un paraguas para cubrirse. Hay menos paraguas que niños, y se pide que el niño indique cuántos paraguas faltan para que cada quien tenga uno.
>
> Se muestran varios dibujos de vías de tren de mayor o menor longitud, en los cuales hay seis vagones de tren, distribuidos de distintas maneras. Se pide al niño que indique cuál de las vías mide exactamente lo mismo que seis vagones. [Backhoff et al, 2008, p. 47]

Resultados de la evaluación nacional en Pensamiento Matemático:

	Por debajo del básico	básico	Medio	Avanzado
Nacional	9	49	27	15



1.1.2. Discusión sobre los resultados

En las tres pruebas la mayoría de los estudiantes obtienen resultados clasificados como insuficientes, básicos, elementales. Si comenzamos por la prueba aplicada a niños de preescolar encontramos que el 9 % de los estudiantes esta clasificado por debajo del nivel básico y conforme se avanza en los grados escolares los niños obtienen resultados más bajos.

Ocurre también que la categoría de los que obtienen resultados altos en las pruebas decrece de manera drástica, en preescolar los niños en tal categoría son el 19 %, mientras que para secundaria la prueba ENLACE ubica al 0.5 % y PISA al 5 %.

El informe PISA específica que México se encuentra dentro de "los márgenes de lo esperado". Sin embargo, es alarmante que un porcentaje significativo (28 %), se encuentre clasificado en el nivel 0, para quienes se concluye su incapacidad de tener éxito en las tareas más básicas que busca medir PISA.

Los resultados de ENLACE muestran que hay un porcentaje realmente alto para los niveles insuficiente y elemental, tanto de primaria con un 77.7 %, como de secundaria con un 94.4 %.

De manera análoga en la prueba del INEE casi la mitad (49 %) de los niños evaluados se ubican en el nivel básico. Paradójicamente los niños con menos años en la escuela obtuvieron mejores resultados.

Ante los resultados descritos anteriormente cabe hacer la pregunta:

¿Quiénes son los evaluados y para qué son necesarias las pruebas?

Las pruebas señalan las deficiencias de un sistema. De ahí la necesidad de crear evaluaciones que permitan obtener información acerca de los aciertos y desatinos de la institución educativa.[12]

Se recurre a ellas para medir la magnitud de las carencias en determinadas materias. Aunque otros países no necesariamente se llevan a cabo pruebas como señal de alerta, en México se implementaron, entre otras causas, porque es más o menos evidente para todos que *algo anda mal en la educación* y aunque los resultados se prestan a distintas interpretaciones la experiencia cotidiana confirma la sospecha colectiva: existen grandes deficiencias en la educación en todos los niveles y es urgente investigar por qué y proponer soluciones.

Pese a las deficiencias en las pruebas,[13] los resultados confirman que es necesaria una reestructuración del sistema educativo sin perder de vista factores socioculturales y económicos.[14]

Los resultados involucran a toda la sociedad ya que de una u otra manera todos avalamos la institución encargada de la educación básica en el país. En apariencia existe consenso en cuanto a que la educación debe ser tal, que permita a los estudiantes el

[12]En ocasiones la forma en la que se presentan los resultados de las pruebas hace alusión únicamente a las deficiencias de los estudiantes, sin tomar en cuenta a todos los actores responsables.

[13] Como muestra de esta situación, el testimonio de Raspón Domínguez, director de la escuela telesecundaria "Bernardino de Sahagún" de la localidad de Temaxcalapa de Gabino Barreda en el municipio mixteco de Chietla –la cual fue la peor calificada en la prueba ENLACE- criticó la prueba ENLACE, pues "con los resultados que obtuvimos no sólo nosotros, sino todos las escuelas de Puebla no se hará nada y el próximo año únicamente vamos seguir igual, en el sótano de la evaluación. Me parece que las autoridades deben de cumplir con lo que prometen y realizar las mejoras a las escuelas para que los alumnos cuenten con todas las herramientas para su enseñanza. Si la prueba ENLACE evalúa por igual a todas las escuelas, entonces que el gobierno las ponga en igual de circunstancias a todas", concluyó [4].

[14]No es casualidad que los países con mejores resultados en la prueba PISA sean a su vez los que mayores recursos económicos destinan a la educación.

desarrollo de habilidades que lo lleven a razonar y resolver problemas en cualquier área, le ofrezca una perspectiva de desarrollo y calidad de vida.[15]

El problema de enseñar es indisoluble a aprender, por tanto, conviene reflexionar acerca de qué es menester enseñar a los estudiantes en términos que los conocimientos sean verdaderas herramientas en su vida en el ámbito social y personal.

Es necesario señalar que la prueba PISA en su informe 2006, arroja información acerca de los Gastos de Investigación y Desarrollo que cada país destinó en el año 2003. El caso de Finlandia es significativo, pues siendo el país que más recursos destina para esta área, 3.5 de su Producto Interno Bruto (PIB), es también el país que mejores resultados generales obtiene. En matemáticas, por ejemplo, la mayor parte de sus estudiantes se ubican en el nivel 4, lo que evidencia que el desarrollo de mejoras educativas que repercutan de manera tangible en el aprendizaje de las matemáticas, es posible [Díaz et al, 2007].

Mientras que en México se destino el 0.4 del PIB, con los resultados ya vistos. El problema educativo en nuestro país es fundamentalmente un problema de recursos.

Pese a lo anterior o precisamente por ello, es necesario hacer investigaciones que aporten material para la mejora en el ámbito educativo.

Si partimos de considerar que los caminos que han de llevar a remediar esta situación hace necesario hacer investigaciones que nos ayuden a comprender cómo aprendemos estamos en posibilidad de entender el fenómeno educativo y con ello plantear alternativas para su mejora.

1.2. Problemática en la enseñanza de las matemáticas

Tomando en cuenta lo anterior y con la finalidad de contribuir a la solución de esta problemática, es necesario que las investigaciones en matemática educativa responda la pregunta:

¿De qué manera podemos hacer más eficiente la enseñanza de las matemáticas?

Para el análisis de la pregunta no podemos hacer a un lado que uno de los determinantes al respecto tiene que ver con el factor económico.

Una de las condiciones para responder la pregunta anterior implica necesariamente el estudio del órgano encargado de los procesos cognitivos: el cerebro y una forma de acercarnos a su estudio es a través de las redes neuronales.

Por ello la pregunta clave del presente trabajo es determinar si para aprender a contar, ¿es esencial el sistema de representación[16] o es un recurso secundario para el

[15]Perspectiva que dadas nuestras condiciones socioeconómicos ha dejado de ser real, pues la preparación profesional no representa la posibilidad en una mejora de vida, ni siquiera en el ámbito laboral como lo expone Karina Avilés "Para muchos jóvenes la educación no tiene sentido, pues no les garantiza acceder a mínimos de bienestar. Siete millones de ellos, conocidos como ninis porque "ni estudian ni trabajan", son blanco potencial de la ilegalidad. Ante ese panorama, se ha gestado entre la juventud un fenómeno de desesperanza y frustración, pero también de malestar social, que de acuerdo con expertos ya deja sentir sus efectos. [6]

[16]En este caso el lugar de los sistemas de representación lo ocupan las distintas bases.

aprendizaje de las redes neuronales?

Al revisar las investigaciones encaminadas a la comprensión de los procesos de aprendizaje, observamos que la aprehensión de los conceptos abstractos constituye un hito en este campo.

En diversos momentos históricos, las matemáticas han servido como ejemplo *per se* de la ciencia y el pensamiento abstracto, de ahí la importancia de investigar uno de los primeros acercamientos del niño con un concepto que al ser despojado de los conjuntos concretos se convierte en una creación abstracta: los números.[17]

A continuación trataremos brevemente algunas teorías pedagógicas enfocadas a dilucidar por un lado la importancia del desarrollo biológico y por otro, la influencia del medio sobre el aprendizaje del hombre.

La primera teoría corresponde a Jean William Fritz Piaget (1896-1980) quien es reconocido entre otras cosas por sus aportaciones en psicología y pedagogía.

Su teoría está basada en la epistemología genética[18], que a grandes rasgos enmarca el desarrollo del conocimiento válido, [19] marcando cuatro estadios del desarrollo cognitivo en el ser humano, relacionados con procesos como pensar, reconocer, percibir, recordar y ordenar.

Para Piaget los procesos de aprendizaje son resultado del desarrollo y la lógica. La conducta está basada en la necesidad de recuperar el equilibrio del individuo con su medio, así la acción se encamina a la adaptación, entendida como las interacciones entre el sujeto y su medio y el tiempo como factor determinante para la efectividad de las readaptaciones posteriores. Con fundamento en diversos experimentos plantea la teoría de los estadios del desarrollo. Piaget midió los grados de desarrollo a través de diversas evaluaciones. [Piaget, 1969]

El trabajo de Piaget es otro ejemplo de cómo el estudio de los errores es útil para entender procesos. Él analizó los patrones de errores que se superaban de una edad a otra en el desarrollo infantil y comparó estas pruebas con otras análogas aplicadas a adultos, esto lo hizo distinguir los procesos cognitivos de los niños pues son inherentemente distintos a los procesos de los adultos.

Cabe mencionar que las pruebas que Piaget llevó a cabo, dan un marco de referencia para mostrar la necesidad de que los niños deben entender ciertos principios lógicos para llevar a cabo tareas como conservación, deducción, seriación, etc. [Piaget, 1978]

En este sentido la lógica es el criterio más sólido para definir la comprensión numérica

[17] Es interesante denotar que la pregunta acerca de la naturaleza de los números ha sido abordada por filósofos y matemáticos como Peano (1858-1932), Dedekind (1831-1916), Rusell (1872-1970), Frege (1848-1925), entre otros y no hay una respuesta que agote las objeciones filosóficas a los planteamientos de cada uno. [Rodríguez-Consuegra, 2001]

[18] Piaget describe la epistemología genética como una investigación esencialmente interdisciplinaria que se propone estudiar la significación de los conocimientos, de las estructuras operativas o de las nociones, al recurrir, por una parte a su historia y a su funcionamiento actual en una ciencia determinada, por otra parte a su aspecto lógico y finalmente a su formación psicogenética o sus relaciones con las estructuras mentales. [Piaget, 1978]

[19] Podemos encontrar muchas clasificaciones en las categorías de conocimiento: intuitivo, sensible, técnico, empírico, explicito, tácito, etc., esto hace que se tengan diversos parámetros para validar el conocimiento.

de los niños. Y existe polémica en cuanto a establecer un criterio general respecto a qué edades son "correctas" para el surgimiento de la comprensión lógica.

Los procesos de aprendizaje, al menos para esta teoría no quedan reducido a la observación del desarrollo biológico de por si insuficiente para la explicación de capacidades intelectuales.

La importancia de la teoría piagetiana es señalar que existen requerimientos lógicos que deben respetarse al pensar en términos matemáticos.

Veamos otra teoría que articula en un binomio inseparable, el desarrollo biológico y un entorno social y cultural que permita la aprehensión y creación del saber.

Ésta teoría es creada por Lev Semionovich Vygotsky (1896-1934), quien resalta la importancia de las relaciones sociales, declarando que todo lo que el niño aprende es resultado de la relación que establece con los otros. La interacción social cumple un papel formador y constructor.

Además considera la asimilación de conceptos científicos como un proceso fundamental que muestra el desarrollo humano y, que modifica profundamente la forma de pensar del niño. Por otra parte, la existencia de la escuela es usada para fundamentar sus tesis, resaltando la importancia de transformar el papel de *educación artificial* que ésta sustenta.

Una de las implicaciones de su investigación establece que tampoco es suficiente la consideración de factores hereditarios (por ejemplo, la capacidad que tenemos los seres humanos de desarrollar el lenguaje) si bien son indispensables no bastan para explicar la complejidad del aprendizaje. [Vygotsky, 1983]

Sin la intervención social el cachorro humano no tendría posibilidad, por sí sólo de desarrollar el lenguaje mismo que en calidad de instrumento de las relaciones sociales, se transforma en un instrumento de la organización psíquica interior del niño.

Vygotsky habla sobre la apropiación de los instrumentos culturales que se convierten en técnicas interiores; esto atañe a la formación de los conceptos: estudios comparados sobre los conceptos experimentales, los conceptos espontáneos y los conceptos científicos. [Tudge, 1989]

En el presente se estudian y se proponen otras teorías educativas y no existe un consenso en cuanto a la efectividad infalible en alguna de ellas.

Pero nosotros elegimos éstas dos teorías por dos razones fundamentalmente: ambas otorgan un lugar privilegiado en sus experimentos, técnicas y resultados a la emergencia del pensamiento abstracto en el niño, como parámetro de desarrollo intelectual.

Dicho lo anterior, la tarea primordial es observar los procesos que seguimos en la conformación del pensamiento matemático.

Es importante conocer la estructura y las funciones del cerebro. Es en este punto en donde la aportación de las redes neuronales es fuente de modelos y teorías que se retroalimentan con las investigaciones biológicas, ya que constituyen por sí mismas nuevas formas de comprender la actividad cerebral.

En esta búsqueda una de las principales fuentes de investigación es el análisis minucioso de la forma en la que los estudiantes conciben los conceptos, operaciones y problemas que se les presentan.

Para ejemplificar lo anterior, uno de los primeros pasos que nos acercan a la comprensión de las funciones cerebrales, es tomar en cuenta los errores en las pruebas de los estudiantes, pues sólo la comprensión de los procesos que los lleva a dar determinada respuesta arrojará señales que apunten a como ellos *equivocándose en matemáticas, aprenden matemáticas.*

La importancia de los errores radica en estos permiten reconstruir los procesos de pensamiento que los originaron, tarea comparable al estudio de la evolución de los conceptos.

Se debe intentar comprender la lógica que siguen al elegir una respuesta.[20] Es igualmente importante analizar la causa de los errores tanto como el por qué de los aciertos.

Una prueba es absurdamente desaprovechada cuando nos contentamos con obtener de ella el número de aciertos, sin intentar mirar lo que en medio se quedo.

En los cursos que imparto como profesora de bachillerato[21], en una de las clases de geometría medimos la superficie de un cilindro, los estudiantes recurrieron a lo que recordaban de la fórmula, pero no estaban seguros si la recordaban bien. Algunos preguntaron de dónde se obtenía el valor de π y al cuestionar a todo el grupo, estas fueron las respuestas:

"Es la línea que pasa por el centro y divide al círculo."

"Es la línea que va del centro al círculo."

"Es lo que mide del diámetro."

"Es 3.1416"

Cuando pregunté acerca de sus respuestas varios de ellos contestaron que *algo así*, recordaban de cursos anteriores.

Sucedió algo similar con el cálculo de área de un triangulo equilátero en el que les proporcione las medidas de la altura y el lado.

Algunos recordaban correctamente la fórmula. Otros calcularon el área multiplicando base por altura.

Les dije que la fórmula correcta para medir el área del triangulo era multiplicar la base por la altura y dividir el resultado entre dos.

Luego les pregunte: ¿por qué esta fórmula es correcta?

Estas fueron sus respuestas:

"Porque la altura divide al triángulo en otros dos triángulos. "

"Porque quitando la base quedan otros dos lados."

[20] Paulo Freire comenta lo siguiente: "En el nordeste brasileño, es común combatir la plaga de orugas, clavando tres estacas en forma de ángulo en el lugar más castigado por ellas. En la extremidad de una de las estacas hay un clavo, en el que el campesino inserta una de ellas. Está convencido de que las demás, con miedo, se retiraran en procesión, entre una estaca y otra. Mientras espera que se vaya la plaga, el campesino pierde parte o gran parte de su cosecha. ¿Qué hacer desde el punto de vista educativo, en una comunidad campesina, que se encuentra en tal nivel? La respuesta no puede estar, en la extensión mecanicista, de los procedimientos técnicos. El pensamiento mágico no es ilógico. Tiene su estructura lógica interna, y reacciona hasta donde puede, al ser mecanicistamente por otro. Este modo de pensar, como cualquier otro, esta indiscutiblemente ligado tanto a un lenguaje y a una estructura, como a una forma de actuar".(Freire, 1973, p. 31-35)

[21] Curso de Matemáticas I en la Modalidad Semiescolar del Sistema de Bachillerato del Gobierno del D.F.

Sin duda alguna, los estudiantes conservan *algo* de sus lecciones matemáticas, pero pocas veces consiguen aplicarlo correctamente y aun cuando lo logran, no significa que comprenden lo que están haciendo.

Las matemáticas ofrecen explicaciones que si son comprendidas satisfacen el entendimiento, hallando en ellas razón y sentido. Pero al no saber de dónde sale un valor, una fórmula, en fin, un resultado, el hueco que deja la explicación precisa es llenado por otra las más de las veces incorrecta u olvidada.

De esta manera aunque se habla de las matemáticas como la ciencia que posibilita y desarrolla el entendimiento, sus métodos y resultados se perciben como mágicos y asequibles en tanto son memorizables aún cuando se consiga resolver un tipo de problemas planteados.[22]

Es decir nuestro cerebro lleva a cabo una tarea específica para posibilitar el aprendizaje matemático o de otra índole.

El objetivo de este trabajo no es proponer una técnica más de enseñanza, tampoco establecemos el uso de las redes neuronales como módelo de análisis de datos. Nosotros usamos el perceptrón multicapa como *modelo de aprendizaje* para que haciendo uso de lo que ofrecen y tomando en cuenta sus limitaciones, sirvan de apoyo para experimentar con ellas en torno a un tema crucial en las matemáticas: aprender a contar.

1.3. Importancia de aprender a contar

El nombre es arquetipo de la cosa
en las letras de rosa está la rosa
y todo el Nilo en la palabra Nilo
Jorge Luis Borges, *El Golem*

Los números son uno de los primeros acercamientos que el niño tiene con las matemáticas, es decir, con el pensamiento abstracto.

Sin el lenguaje sería vano e imposible comprender la utilidad de los números. Posterior al manejo del lenguaje que permita la comprensión de los números, estos adquieren toda su importancia gracias a los símbolos escritos.

Páginas atrás hablamos del camino que los conceptos y creaciones matemáticas han tenido a través del tiempo, si rastreamos el origen de los símbolos que representan los números observamos que los primeros signos que el hombre creó fueron aritméticos. Luis Puig [Filloy, 2003] narra como 3500 años a. n. e. en un pueblo de Mesopotamia usaron bolas huecas de barro blando rellenas con guijarros de distintas formas y en su exterior encontraron signos, las marcas encontradas representan los guijarros ya que compartían con ellos su forma y su número. Se determinó que estas bolas de barro son un paso intermedio en la representación de cantidades, su antecedente fueron bolas rellenas de guijarros sin iconos y su consecuente bolas de barro sin los guijarros y con

[22] "La enseñanza de la matemática ha degenerado con frecuencia en un vacio entrenamiento de resolución de problemas, que si bien puede desarrollar una habilidad formal, no conduce en cambio a una compresión efectiva ni a una mayor independencia intelectual".(Courant, 2003, p. IX).

la representación de los iconos, para luego convertirse en tablillas grabadas pues no era necesario encerrar nada. Ahora bien, las tablillas son funcionales para los registros, pero se piensa que los comerciantes seguirían operando los guijarros. Y es hasta 2000-1600 a. n. e. que se escriben verdaderos textos matemáticos.

En el aprendizaje en general y en particular en las matemáticas, la comprensión de los símbolos y el lenguaje usados en un sistema ocupan un lugar fundamental.

Al comenzar a trabajar con los símbolos que representan los números inicia la independencia de los conjuntos de elementos antes necesarios para enumerar.

Este camino de la abstracción será el que recorrerán los objetos matemáticos con los que el niño tendrá que lidiar durante su vida y es de hecho el camino que ha seguido la humanidad .

No pocas veces nos topamos con que la problemática por comprender las matemáticas está estrechamente vinculada a la dificultad para entender el significado de los símbolos.

El problema para aprender matemáticas es un problema relacionado con la notación y ésta da forma y sentido a las creaciones matemáticas.[23]

Observamos que requirió mucho tiempo el poder separar los números para su uso como dígitos del uso original para medir cantidades tal como sucede cuando se trata de comprender cualquier noción matemática, pues es necesario establecer la diferencia entre los objetos matemáticos y su representación

La maduración necesaria para el concepto de número parte de la capacidad de uso y sentido del objeto matemático, aún cuando no haya ningún significante perceptible.

Ahora bien las distintas representaciones, en este caso sistemas numéricos, conocidas como **Sistemas de Representación Semiótica** [24], encapsulan información acerca del proceso cognitivo desarrollado que permitió la creación de dicho sistema. Ahí podemos encontrar su lógica y por tanto el proceso que se siguió al resolver un problema inscrito en dicho sistema.

Cuando se crea un sistema de representación semiótica la aprehensión conceptual de un objeto [25], puede establecerse si y sólo si se da también la aprehensión semiótica; el sistema de representación no ocupa un lugar secundario pues es síntesis y herramienta de la matemática, además es un medio que permite la comunicación y transmisión del saber.

Las conexiones que existen entre los conceptos y objetos matemáticos requieren un manejo fluido de los mismos para comprender temas nuevos.

Bajo este contexto el aprendizaje de los números es una piedra angular que permitirá el conocimiento posterior.

[23] Otro ejemplo de la importancia de los símbolos y la representación lo encontramos en el trabajo de los programadores para quienes constituye un reto simbolizar los problemas que quieren resolver.

[24] Duval comenta que una característica importante de la actividad matemática es el uso de diversos sistemas de expresión y representación además del lenguaje natural: variados sistemas de escritura para los números, escrituras algebraicas para expresar relaciones y operaciones, figuras geométricas, gráficos cartesianos, redes, diagramas, esquemas, etc. [Duval, 1996]

[25] Este proceso se conoce como noesis.

1.3.1. ¿Qué significa aprender a contar?

Terezinha Nunes y Peter Bryant señalan tres condiciones:

- Principio de correspondencia biunívoca.
- Principio de orden constante.
- Correspondencia entre el nombre del número de objetos contados y la cantidad real de objetos.

Los autores señalan que estas tres condiciones pueden estar cubiertas y, sin embargo, eso no garantiza que los niños sepan contar. [Nunes y Bryant, 2003]

El niño que aprende a contar necesita comprender inferencias lógicas y transitivas de las unidades. Hablamos de estas condiciones páginas anteriores cuando señalamos algunos aportes de Piaget, de hecho es una de las consecuencias de la teoría de Piaget respecto al lugar que ocupa el pensamiento lógico en la comprensión de las matemáticas.

Lo fundamental no reside en “recitar” los números sino en dominar el procedimiento de clasificación.

Aprender a contar sigue el modelo de todo el aprendizaje en matemáticas: saber jugar un juego en donde se conocen los sistemas de prácticas, los conceptos y las relaciones entre ellos se manejan, de esta manera observamos que los números jugarán el papel de conceptos reutilizables a lo largo de todo el conocimiento matemático que el ser humano desarrolla ya que funcionan como herramienta que posibilita el conocimiento posterior y estructuran un modo de pensar, al requerir el manejo de propiedades como conservación y transitividad, inferencias lógicas necesarias para contar.

Si el conocimiento no es resultado de la comprensión, el alumno enfrenta a un mundo lleno de confusión, donde se usan nombres, fórmulas y números, que no logra relacionar adecuadamente.

En este sentido una acercamiento “amigable” a esta comprensión lo ofrece la metodología de las redes neuronales. En el problema de aprender a contar en distintas bases, experimentaremos con las redes neuronales que serán consideradas estudiantes virtuales.

Capítulo 2

Perceptrones multicapa

Las redes neuronales pueden ser de diferentes modelos de aprendizaje y topologías dependiendo del uso que se espere darle a cada una. Pueden ser aplicadas a diferentes campos tales como: control de eficiencia de máquinas, reconocimiento de firmas, reconocimiento de blancos mediante sonares, predicciones en el tiempo, decisiones, análisis de inversiones, monitoreo, mercadotecnia, planeación estratégica, diagnostico, diseño, optimización y en aplicaciones (recientemente) de ingeniería civil como la valoración de efectos sísmicos, solución de problemas de gerencia de construcción, control activo estructural y diagnósticos de daño. [Paolucci, 2000]

Para la clasificación de las redes neuronales existen dos criterios: el tipo de aprendizaje y la arquitectura. Si la clasificación depende del tipo de aprendizaje se considera si éste es supervisado o no supervisado (autoorganizado), híbrido o reforzado.

Si la clasificación depende de la arquitectura y esta no presenta ciclos se llama unidireccional, en caso contrario se denomina recurrente o retroalimentada.

En el presente trabajo usamos una red neuronal conocida como perceptrón multicapa que pertenece al grupo de redes unidireccionales organizadas en capas y con aprendizaje supervisado.

Actualmente se considera que el perceptrón multicapa (o alguna de sus variantes) es el más usado en las aplicaciones prácticas, este modelo usa el algoritmo de *backpropagation* y su popularidad se debe a que gracias al paradigma que constituyó su capa oculta, como veremos más adelante, se considera capaz de aprender cualquier tipo de función o relación continua entre un grupo de variables de entrada y salida.

El perceptrón multicapa tiene su origen en los componentes de las neuronas biológicas por ello presentamos brevemente la estructura y funcionamiento de éstas, para luego mostrar los origenes y desarrollo del perceptrón.

2.1. Redes neuronales biológicas

El estudio de los componentes del sistema nervioso ha hecho evidente su complejidad. La clasificación más general para éste es el sistema nervioso central y el sistema nervioso autónomo. El primero está constituido por el cerebro y la médula espinal, el segundo está formado por los ganglios raquídeos y los nervios craneales.[1] o raquídeos[2]

Existen dos tipos de neuronas en el sistema nervioso central: las células nerviosas (neuronas) y las células gliales (glía).

El cerebro humano contiene aproximadamente 10^{11} neuronas, podría darse una clasificación de al menos mil tipos diferentes. La principal diferencia entre ellas es su forma, específicamente el número y la forma de las prolongaciones que salen de su cuerpo celular. Cada una establece 1000 conexiones, puede llegar incluso a las 10,000.[3] Se calcula que por cada neurona hay entre 10 y 50 células gliales y aunque no se ha comprobado la participación directa de estas en la elaboración de la información tienen otras funciones vitales.

Pese a la variedad de neuronas se ha logrado conocer el mecanismo que todas comparten para generar una conducta. Dicho mecanismo queda descrito por cuatro señales: un componente de entrada local (receptivo), un componente desencadenante (de suma o integrador), un componente de conducción de larga distancia (señalización) y un componente de salida (secretor).

2.1.1. Componentes de la neurona

La neurona está constituida por cuatro regiones morfológicas: cuerpo celular o soma, las dendritas, el axón y las terminales presinápticas.

Cuerpo celular: contiene un núcleo que almacena los genes y realiza las funciones lógicas de la célula.

Axón: es una fibra nerviosa que transmite las señales (potenciales de acción) a otras neuronas. Puede formar sinapsis hasta con 1000 neuronas.

Dendritas: ramificaciones receptoras de señales provenientes de otras neuronas. Son un grupo de fibras muy ramificadas y de forma irregular que se conectan directamente al núcleo.

Sinapsis: Son las conexiones entre neuronas y pueden ser de tipo eléctrico o mecánico. Se considera que en ellas se almacena la información y que son capaces de cambiar la polaridad de los potenciales provenientes de otras neuronas y por ello se suele hablar de su naturaleza excitadora o inhibidora[4].

[1] Originados en el encéfalo.
[2] Originados en la médula.
[3] Las células de Purkinje del cerebelo reciben hasta 100 000 conexiones.
[4] Diversos tipos de plasticidad sináptica y las distintas escalas de tiempo en las que operan sugiere que las sinapsis cumplen un papel más activo en el proceso de información, son algo más que simples conductos. [Abbott y Regehr, 2004]

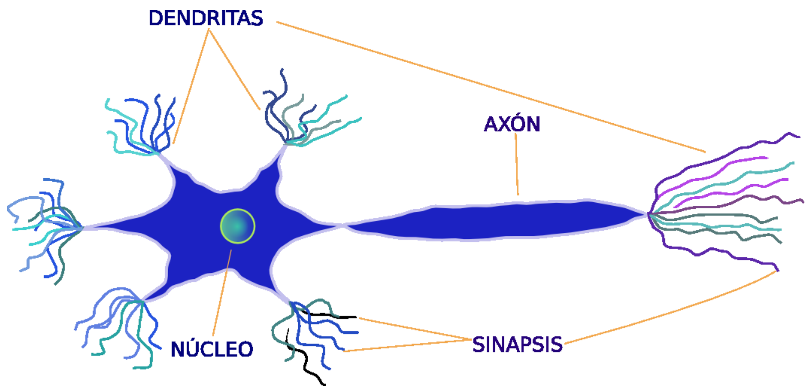


Figura 2.1: Neurona biológica genérica

Las señales que permiten el flujo de información son eléctricas pero estos impulsos están combinados con procesos químicos regulados por las sinapsis. Las señales se determinan midiendo los cambios de potencial eléctrico en ambos lados de la membrana. Mientras la neurona se encuentra en reposo mantiene una diferencia de potencial eléctrico de aproximadamente 65 mV.

La neurona concentra $\mathbf{K}^{+}$ y manda al exterior $\mathbf{Na}^{+}$, por tanto en el interior de la membrana se encuentra cargado positivamente respecto al exterior, cuando el potencial de la membrana se hace más negativo se dice que la célula esta hiperpolarizada y cuando se hace menos negativo se dice que la membrana esta depolarizada, los cambios en el potencial de la membrana generan las señales locales y de largo alcance.

2.1.2. Potencial de acción

Si las señales logran sobrepasar un umbral se genera un potencial de acción, cuando queremos enviar algún mensaje al cerebro debe ser en forma de una serie de potenciales de acción, ya que estos desencadenan que el cerebro reciba, analice y transmita información.

Si un estímulo no alcanza el umbral no produce señal, si consigue alcanzar el umbral e incluso sobrepasarlo la señal será la misma aunque la diferencia entre intensidad o duración sea grande, la amplitud y el potencial de acción son los mismos.

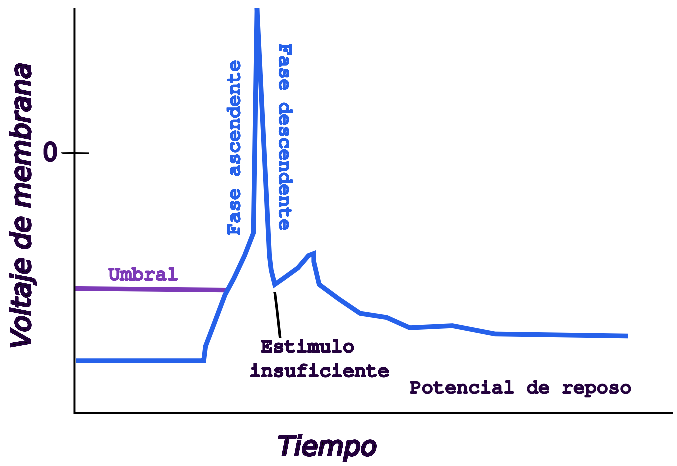


Figura 2.2: Neurona biológica genérica

El potencial de acción tiene variaciones sutiles pero la mayor parte de ellos son estereotipados, ¿qué hace entonces la diferencia entre una señal para hablar y otra que ordene al cuerpo bailar?

La clave del funcionamiento cerebral se encuentra en las redes que siguen los potenciales de acción.

El elemento clave de las funciones cerebrales es su buena organización. El cerebro distribuye la información de manera que se pueda llegar a ella a través de distintos caminos, de manera que si alguna vía sufre daño, existen rutas alternas preparadas para conservar el funcionamiento correcto. Se ha observado que existen regiones encargadas de tareas específicas tales como el lenguaje, la visión, etc.

Es una estructura del todo o nada. Mientras los estímulos por debajo del umbral no producen señal todos los estímulos que superan el umbral producen la misma señal aunque la diferencia en la intensidad o duración sea grande la amplitud y la duración del potencial de acción son prácticamente iguales. El potencial se regenera periódicamente. Los potenciales de acción son muy estereotipados y sólo varían de forma sutil.

Lo que determina la intensidad de la sensación o movimiento no es la magnitud o la duración de cada potencial sino su frecuencia. Se sabe que las señales sólo viajan en una dirección.[5]

Nuevamente aquí la clave es la vía nerviosa que transporta el potencial de acción, lo que da origen a la gama de acciones de la que son fuente son los distintos tipos de conexiones y la conformación de circuitos anatómicos precisos.

En el caso de las redes neuronales artificiales las funciones primitivas están en los nodos de la red y las reglas de composición se encuentran implícitamente en la inter-

[5] Existen otras moléculas que contribuyen en la comunicación entre células, ya sea neuronas o glía. La descripción relativamente reciente de los potentes efectos de gases como el óxido nitroso (NO) o el monóxido de carbono (CO) a nivel de la sinapsis hacen pensar que la lista de neurohumores aumentará en el futuro. A estos agentes, junto con el ácido araquidónico, se les ha denominado “terceros mensajeros”, pues son capaces de transmitir información en “sentido contrario”, es decir, desde la terminal postsináptica a la presinapsis.

conexión de los nodos, en la sincronía o asincronía de la transmisión de información y en la presencia o ausencia de ciclos.

Existe además otra propiedad que ha inspirado modelos para las redes neuronales de aprendizaje no supervisado: la influencia que una neurona ejerce sobre otras es función de la distancia entre ellas, mayor influencia mientras más cercanas estén.[6] Otra función que se estudia es el papel que juega la experiencia en la permeabilidad de las neuronas ya que la estructura de las conexiones es susceptible de modificación.

Aunque se reconoce la importancia genética en la determinación de tal organización es probable que parte de ella se origine durante el aprendizaje, lo que significa que el cerebro puede poseer la capacidad inherente de formar mapas topológicos de las informaciones recibidas del exterior.

Con este breve acercamiento a las redes neuronales biológicas podemos observar que el estudio de cada uno de los componentes resulta un gran mosaico que muestra los niveles de complejidad que el cerebro posee. Actualmente algunos de los modelos de redes neuronales están inspirados en modificaciones particulares que son fruto del estudio detallado de alguno de los componentes del cerebro.

Observamos que cada neurona por si sola es tan compleja o aún más compleja que cualquier computadora.

2.2. Antecedentes del perceptrón

Se han propuesto modelos de la arquitectura de las computadoras que copian algunas características del funcionamiento de los organismos vivos. Uno de los primeros diseños son los autómatas celulares propuestos por John von Neumann en la década de los cuarenta y este modelo a su vez estuvo inspirado en las redes neuronales propuestos por Warren McCulloch y Walter Pitts.

El ser humano ha tomado el funcionamiento de su cuerpo como modelo para la creación de diversos sistemas.[7]

Sabemos que la arquitectura usada para nuestras computadoras actuales está basada en los principios de lógica booleana y máquina de Turing. Sin embargo para abordar una amplia gama de problemas que involucran el manejo de información conocida como *del mundo real* (reconocimiento de patrones y procesamiento de señal) son mejor resueltos cuando se aplican estructuras de hardware o software inspirados en sistemas biológicos.

El modelo de Warren McCulloch y Walter Pitts propuesto en 1943, es el antecedente del perceptrón [McCulloch y Pitts, 1943]. Este modelo conocido como *neurona de McCulloch-Pitts* tiene los componentes de la neurona genérica. Y su herencia al perceptrón radica fundamentalmente en lo que se conoce como lógica de umbral.

[6] También ocurre que no sólo se activan células cercanas.

[7] Algunos ejemplos de lo anterior los encontramos en diseños de inteligencia artificial: algoritmos genéticos con los cuales se copia el proceso de evolución de las cadenas de ADN, ejecución de una respuesta predeterminada por cada entrada, análogas a actos reflejos en seres vivos y diversos sistemas de reconocimiento de habla, de escritura, entre algunas de la características que son imitadas.

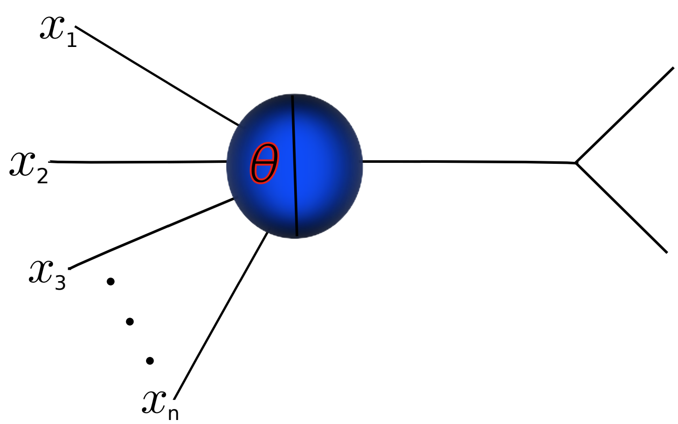


Figura 2.3: Modelo de McCulloch-Pitts

La célula de McCulloch-Pitts funciona como sigue:

- Se proporcionan entradas x_1, x_2, x_3,..., x_n
- El valor de las entradas se suma $X = x_1+x_2+x_3++x_n$
- Se compara con el umbral si $X < \theta$ la célula no se activa, si $X \geq \theta$ la célula se activa.

Las entradas son las encargadas de recibir información de sinapsis excitatorias o inhibitorias que revelan el estado de las células presinápticas. En el primer tipo de sinapsis los pesos son idénticos y su actividad se suma linealmente y en el segundo caso su actividad es absoluta, es decir, si la señal es inhibitoria, la neurona no se activa.

Luego de ser recibida la suma del de las entradas se compara con el valor del umbral θ (el cual es fijo), si es igual o mayor la neurona se activa, de no alcanzar el valor del umbral o de estar activa una sinapsis inhibitoria la neurona permanece inactiva; esta última explicación se conoce como la *ley del todo o nada*. [Rojas, 1996]

Observamos que la neurona sólo puede estar en dos estados: activa o inactiva, es un dispositivo binario y por lo tanto tiene la equivalencia al verdadero falso de la lógica booleana, por esta razón se pensó que la actividad de cualquier neurona podía ser representada como una proposición lógica.

Si tal estructura era una buena aproximación al cerebro su estudio podía quedar simplificado al análisis de procesamiento de proposiciones lógicas. Sin embargo, como observamos líneas atrás la complejidad de los procesos cerebrales hizo visible que su estudio no puede reducirse a la lógica formal.

Siguiendo la idea de estos componentes básicos de la neurona de McCulloch y Pitts surgió el perceptrón. Fue creado por Frank Rosenblantt en 1958.

Por ser una red de aprendizaje supervisado el perceptrón trabaja con un conjunto de datos de entrenamiento hecho con el total o un subconjunto de los patrones de entrada y salida conocidos y esto permite que la red actualice los pesos de las conexiones de acuerdo a la regla aplicada y finalmente produzca las respuestas deseadas que puede ser otro conjunto de datos o una clasificación.

Además por ser básicamente una neurona de McCulloch-Pitts es una buena herramienta para la representación de funciones booleanas y como decodificador.

2.2.1. Estructura del perceptrón

El modelo simple del perceptrón se basa en los procesos del aprendizaje perceptual, particularmente la visión.

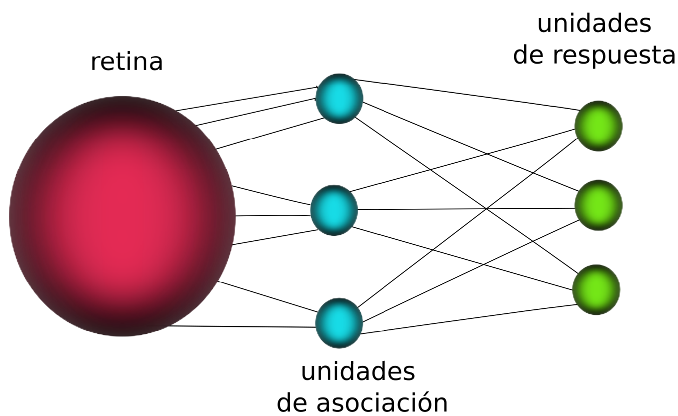


Figura 2.4: Perceptrón simple

Tenemos superficie denominada **retina** que transmite valores binarios a las unidades del área de asociación, las conexiones con la retina son determinísticas y no adaptativas. Las conexiones de la segunda capa son seleccionadas estocásticamente.

Las unidades de la primera capa obtienen información de un superficie conocida como retina, las neuronas no se encuentran totalmente conectadas a esta región se trabaja con subconjuntos de la retina por tanto la información es local, a esta capa se le conoce como sensorial o de entrada y se denota como $X = (x_1, x_2, x_3, ..., x_n)$ y la que identificamos como un vector con todos los valores de entrada; la segunda capa es de salida y se denota por $Y = (y_1, y_2, y_3, ..., y_m)$, también es un vector, en el perceptrón simple se supone que las primer y segunda capa están conectadas totalmente. Las unidades de la capa de salida comienzan con pesos aleatorios y que responda ante una clase particular de eventos producidos en la retina.

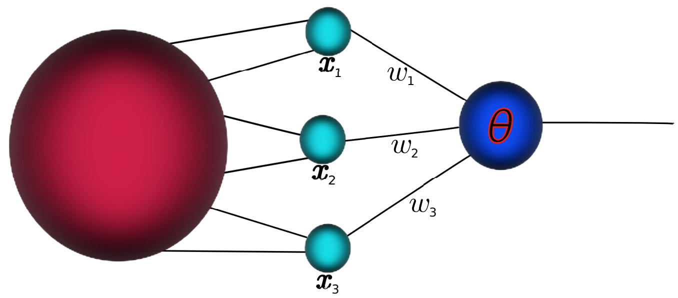


Figura 2.5: Perceptrón con umbral

La innovación respecto a la neurona McCulloch-Pitts fue la determinación automática de los pesos sinápticos y una especial interconexión de patrones, ya que en el modelo original las conexiones se determinaban estocásticamente y el aprendizaje se llevaba a cabo adaptando los pesos de la red con un algoritmo numérico.

El elemento básico de cómputo en el perceptrón es un dispositivo conocido como la **unidad de lógica de umbral**[8] que se corresponde con el umbral que ha de superar el potencial de acción de las neuronas biológicas. [Anderson, 2007]

La información para la capa asociativa proviene de una superficie sensorial conocida como retina, las conexiones establecidas entre ambas son parciales, se trabaja con subconjuntos de la retina. El perceptrón funciona ajustando los pesos de acuerdo a la diferencia existente entre la salida actual de la red y la salida deseada. El aprendizaje continua si es necesario un cambio en los pesos sinápticos; si la red ha alcanzado un punto de equilibrio el aprendizaje se detiene.

El aprendizaje del perceptrón queda descrito como sigue:

Cada entrada x_1, x_2, x_3,..., x_n tiene un peso asociado w_1, w_2, w_3,..., w_n.

El perceptrón se activa si $\displaystyle\sum_{i=1}^{n} w_i x_i \geq \theta$

Esta operación es el producto punto de las entradas y los pesos de la red y las comparaciones entre éste y el umbral da lugar a una serie de desigualdades que separan el espacio de las posibles respuestas en dos regiones.

Se ajustan los valores de los pesos para que concuerde con la salida deseada y se repite el proceso hasta lograr el resultado deseado.

[8]En inglés es conocida como *threshold logic unit o TLU.*

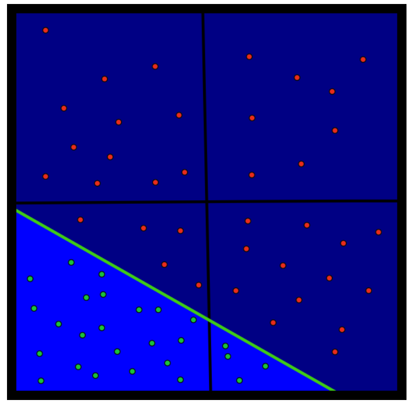
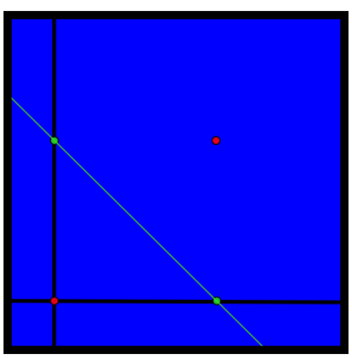


Figura 2.6: Linealmente separable

Las desigualdades organizan el espacio de las respuestas del perceptrón y por esta condición el perceptrón trabaja con funciones que son linealmente separables, de ahí proviene su uso como clasificador de patrones.

Una función es linealmente separable cuando su espacio de variables de entrada puede ser dividido en regiones de igual salida mediante una condición lineal (un hiperplano). La función AND y OR son ejemplos de funciones linealmente separables.

La principal desventaja de los perceptrones es precisamente esta característica ya que es incapaz de resolver problemas que no sean linealmente separables. El ejemplo clásico para mostrar esta limitación es la representación de la función OR Exclusiva.



p	q	OR-Exclusiva
0	0	0
0	1	1
1	0	1
1	1	0

Cuadro 2.1: La función OR-Exclusiva no es separable linealmente.

Minsky y Papert hicieron un estudio de estas limitaciones y publicaron sus resultados en un libro llamado *Perceptrons*.

Para trazar dos rectas en el espacio habría que añadir a la capa de salida otra neurona y para elegir entre una u otra zona de las tres, es necesaria una nueva capa con una neurona cuyas entradas serían las salidas de las dos neuronas anteriores, la nueva capa no es de entrada ni de salida así que se le nombró capa oculta. Pasamos de un perceptrón bicapa a uno con tres capas (neuronas de entrada, capa oculta con dos neuronas y capa de salida con una neurona). Las dos zonas o regiones que contienen los puntos (0,0) y (1,1) se asocian a una salida nula de la red y la zona central con la salida 1.

La idea geométrica en este caso que se corresponde con el aprendizaje es la posibilidad de definir un número mayor de regiones que correspondan a una clasificación particular.

Aumentando una capa se encontró una solución que mejora considerablemente la capacidad del perceptrón.

2.3. Perceptrón Multicapa

A partir del primer modelo de perceptrón propuesto por Rosenblantt y de las críticas de Minsky y Papert se crearon más modelos. Y se mostró que con la utilización de más capas se podía dar solución a problemas que el perceptrón simple no podía resolver.

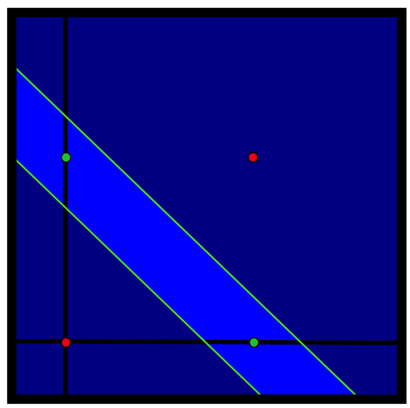


Figura 2.7: Solución para la función OR Exclusiva

De esta manera se llegó al modelo que hoy conocemos como **perceptrón multicapa**. Para el cual queda resuelto el problema de la función OR Exclusiva.

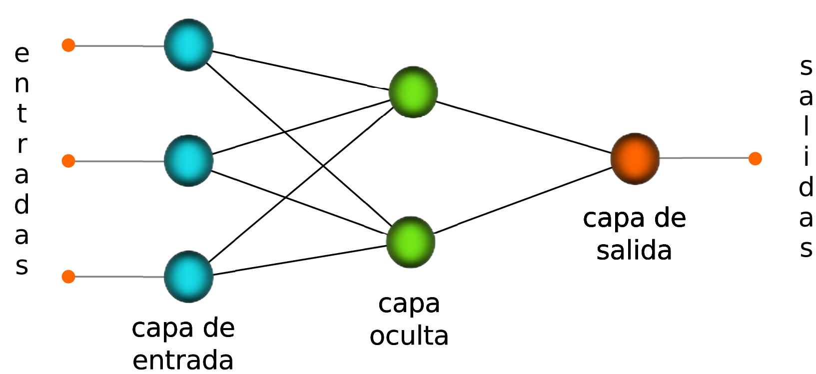


Figura 2.8: Perceptrón Multicapa

Actualmente el perceptrón multicapa es la red neuronal más utilizada en las investigaciones.

Su uso en pedagogía no es frecuente.

Una característica importante en el aprendizaje del perceptrón es que sólo aprende con los errores, de ahí vienen las semejanzas con el estudio de los procesos cognitivos; salvando una diferencia importante: en el perceptrón ocurre que una vez que los errores se terminan, el aprendizaje se queda estancado, lo que por fortuna no ocurre con el aprendizaje del ser humano. Cuando hemos aprendido alguna cosa, ese aprendizaje es susceptible de ser perfeccionado.

Actualmente cuando una red neuronal es diseñada por primera vez, se deben asignar valores a los pesos a partir de los cuales comenzar la etapa de entrenamiento. Los pesos de umbral y de conexión se pueden inicializar de forma totalmente aleatoria, si bien es conveniente seguir algunas sencillas reglas que permitirán minimizar la duración del entrenamiento.

El perceptrón refleja un tipo de aprendizaje conocido como Hebbiano en honor a Donald Hebb (1949) quien propuso en su libro *The organization of behavior: a neuropsychological theory* lo siguiente: cuando un axón de la célula A está lo suficientemente cerca para excitar a la célula B y repetida o persistentemente toma parte en su disparo, algún proceso de crecimiento o cambio metabólico toma lugar en una o ambas células, de tal manera en que la eficiencia de A como una de las células haciendo que se dispare B, se incrementa, esto vuelve a hacer énfasis en la interconexión de las neuronas y la experiencia del sujeto como factores que permiten y refuerzan el aprendizaje.

2.3.1. Algoritmo de retropropagación

El perceptrón multicapa lleva a cabo el aprendizaje por medio de un algoritmo conocido como *backpropagation error*[9] (propagación del error hacia atrás), su utilidad radica en proporcionar información de lo que sucede en las capas ocultas de la red a fin de aprender la relación que existe entre un conjunto de entradas y salidas, es además un algoritmo iterativo que permite entrenar redes multicapa. De forma similar a la regla delta, la base matemática del algoritmo backpropagation es la técnica de gradiente decreciente, basada en modificar los pesos en la dirección opuesta al gradiente.

La dirección de la línea de máxima pendiente viene dada por el vector gradiente de la derivada parcial del error con respecto al peso:

$$\left(\frac{\partial E}{\partial W_1}, \frac{\partial E}{\partial W_2}, \frac{\partial E}{\partial W_3}, \ldots, \frac{\partial E}{\partial W_n}\right)$$

Cuando la última capa de salida arroja un valor se hace una estimación del mismo, generalmente una función cuadrática de los errores individuales cometidos por cada unidad, siendo cada error individual la diferencia entre la salida deseada y la obtenida, luego comienza el proceso contrario analizando qué es lo que ha pasado en cada neurona de cada capa y dando órdenes a cada una de ellas para ver cómo puede reducir el error. El error total será:

[9] También se conoce como método de gradiente en descenso.

$$Error = \frac{1}{2}\sum_{k}(e_k)^2$$

Donde e_k es la salida deseada menos la salida obtenida correspondiente a la k-ésima entrada o patrón.

Con este entrenamiento se consigue usar la experiencia de la red para dar salidas satisfactorias a entradas que el sistema no ha visto antes pero guardan relación con alguna entrada proporcionada en el aprendizaje, de esta manera se consigue la generalización en las clasificaciones.

2.4. Otros Modelos de Redes Neuronales

El uso de las redes neuronales artificiales se dirige en dos direcciones fundamentalmente, como modelos para el análisis de sistema nervioso y **procesos cognitivos** o como herramienta para resolver problemas prácticos.

De manera general cuando una red neuronal es diseñada por primera vez, se deben asignar valores a los pesos a partir de los cuales comenzar la etapa de entrenamiento. Los pesos de umbral y de conexión se pueden inicializar de forma totalmente aleatoria, si bien es conveniente seguir algunas sencillas reglas que permitirán minimizar la duración del entrenamiento.

Además del perceptrón existen otras redes que basan su funcionamiento en la neurona biológica y ponen énfasis en distintas características de la misma. Actualmente se manejan 40 paradigmas de redes neuronales artificiales que son usados en diversos campos. [5]

Ejemplos de estas redes son Adaline, red de Hamming, red de Kohonen, red de Hopfield, red de Elman, etc.

En esta sección presentamos dos redes neuronales que pertenecen a las redes de aprendizaje no supervisado.

En las redes de aprendizaje no supervisado los datos son simplemente presentados a la red, en cierto número de categorías de clasificación.

A continuación veremos dos redes neuronales: el mapeo autoorganizado y las redes de Hopfield ambas son ejemplos de sistemas dinámicos que buscan determinar principios y patrones en la organización de la red.

2.4.1. Mapeo Autoorganizado (SOM)

El mapeo autoorganizado[10] SOM *(Self-Organizing Maps)*es un tipo de red neuronal que se caracteriza por ser de aprendizaje no supervisado, lo cual quiere decir que no existe una función de error que indique la modificación de peso necesaria encaminada al aprendizaje de la red, de aquí el adjetivo de autoorganización ya que esto es precisamente el proceso que lleva a cabo la red por ser un sistema dinámico discreto. Esta organización se establece en función de los datos procedentes del exterior.

[10] También llamado redes de Kohonen.

Una característica de los sistemas dinámicos consiste en pasar de un estado desordenado a uno ordenado con un mínimo de información compartida, esto sucede en las neuronas del SOM, esta característica permite observar que las redes neuronales tienen una arquitectura intrínsecamente asociativa que hace que la actividad de una neurona influya a otras.

La red debe descubrir por sí misma regularidades o categorías en los datos de entrada e incorporarlas a la estructura de sus conexiones. De acuerdo a estas categorías la información similar debe activar la misma neurona de salida.

En este proceso se observan dos estados uno conocido como ordenamiento, donde encontramos un ordenamiento topológico en los vectores de peso y el otro estado es llamado convergencia en donde los vectores de peso tienen otro reajuste que mejora la aproximación al estado de entrada. [Kohonen, 1997]

Una característica importante en el SOM es la conservación de propiedades topológicas. Esto resulta conveniente cuando se trabaja con objetos multidimensionales y se requiere un mapeo en donde las vecindades se conserven para una salida que tenga un espacio de menor dimensión.

2.4.2. Redes de Hopfield

El origen de las redes Hopfield radica en un análisis del aprendizaje en la red neuronal, fue propuesto por un físico del mismo nombre en 1982. En el proceso de aprendizaje se observa cómo la red minimiza una función para cada posible patrón de activación, que puede ser equiparada con el concepto de energía de un sistema físico.

Hopfield hizo un estudio de los cambios cualitativos de la red, estudió su dinámica y observó que las áreas de atracción son el conjunto de estados de activación que con una probabilidad alta hacen alcanzar un estado de equilibrio.

Este análisis concuerda con la teoría de los sistemas dinámicos no lineales, en donde los valores estables funcionan como atractores y se aleja de otro tipo de puntos llamados repulsores.

La arquitectura de la red carece de entradas o salidas en forma de neuronas y se habla del estado de la red.

La conectividad en una red de este tipo es autoasociativa, cualquier unidad se puede conectar con otra bajo la restricción de no conectarse con sí misma. Estas conexiones no implican cálculo de pesos sinápticos ni valores de umbral.

Ejecutar un patrón en la red de Hopfield consiste en proporcionar un patrón a la red y actualizar las neuronas repetidamente hasta que se estabilicen los estados de las neuronas a un patrón memorizado; es decir, las neuronas modifican su estado de acuerdo a la activación de las otras neuronas y lo hacen aleatoriamente, el estado de alguna neurona concuerda con la regla y gobierna el estado de la neurona. Este proceso se repetirá hasta encontrar un equilibrio.

Entre las características benéficas que esta red tiene, se encuentra la posibilidad de encontrar valores en un vector de estado incompleto.

Capítulo 3

Desarrollo de la investigación

Los resultados que presentamos a continuación se hiceron tomando al perceptrón multicapa como un modelo simplificado de un agente que aprende. Como antes mencionamos existen diferencias y limitaciones insalvables entre el aprendizaje de un ser vivo y la red neuronal, sin embargo, los resultados podrían ayudar a comprender algunos aspectos del aprendizaje en los niños.

Tomamos como referencia para la realización de este trabajo, además de nuestro interés personal en matemática educativa, la lectura de dos artículos así como una herramienta prehispánica de cálculo llamada *nepohualtzintzin.*

El primero de los artículos se titula, *La matemática en la vida y en la escuela: Dos décadas de investigación de Terezinha Nunes* y el segundo *Neural network models for teaching multiplication table in primary school* de Alexander L. Tatuzov.

El primer artículo es una investigación realizada en Brasil,y trata sobre las causas del fracaso escolar en matemáticas de niños provenientes de los sectores más pobres, en comparación con niños de escuelas privadas de estrato social alto.

Para realizar dicha investigación el equipo de trabajo verificó la comprensión de los niños de las operaciones básicas en ámbitos distintos al escolar, lo que dió origen a descubrir *la matemática de la calle*, en la que los niños desarrollan exitosamente sus propio métodos para resolver las operaciones básicas.

El segundo artículo consiste en la simulación del proceso de memorización de las tablas de multiplicar y para ello se usa una red neuronal. Esta red aprende usando los errores y aciertos de cada niño y propone tareas para guíar el aprendizaje.

Así pues, guardando las distancias, algunas de ellas insalvables, entre el aprendizaje humano y el de una red neuronal esta investigación muestra que trabajar con redes neuronales permite experimentar libremente, reduce los costos que supone la implementación de nuevos enfoques educativos y disminuye el tiempo que lleva un experimento de este tipo. Y lo que es todavía más importante: permite un amplio margen de error sin afectar a los estudiantes.

En nuestro trabajo la pregunta original fue:

¿La red neuronal aprende a sumar de manera más efectiva con el método tradicional[1]

[1]Nos referimos al método de colocar los sumandos en filas sucesivas ordenando las cifras en columnas,

o con el usado por el *nepohualtzintzin*.[2]

Para contestar lo anterior debíamos encontrar la representación del algoritmo usado por el nepohualtzintzin para que la red pudiera aprender. La solución a este problema frenó la investigación por algún tiempo y, nos hizo considerar un problema intrínseco al que nos proponíamos explorar.

Una de las características del nepohualtzintzin es que fue construido para usarse en base vigesimal. [3] Para usarlo debemos familiarizarnos con esta base.

Así la pregunta inicial se transformó en otra, al darnos cuenta de que el problema de aprender a sumar lleva intrínseco aprender a contar.

El ser humano usó distintas bases para los sistemas de numeración; cada cultura desarrolló diferentes conocimientos matemáticos.

Entonces, ¿hay diferencia entre las bases en las que se enseña a contar a la red o el aprendizaje es cualitativamente el mismo independientemente de la base?, ¿qué lugar ocupa en estas comparaciones la base decimal?

Para responder lo anterior tomamos distintas bases y observamos el aprendizaje en el perceptrón multicapa.

Para este análisis programamos un perceptrón multicapa. Su diseño permite proporcionar parámetros tales como conjunto de entrenamiento, conjunto de prueba, número de épocas y número de neuronas de la capa oculta (también llamada capa intermedia). La función de activación que usa nuestro perceptrón es **la función sigmoide**, los pesos se inicializan de manera aleatoria.

También necesitamos algunas bases representativas, conjuntos de distintos tamaños con sus conjuntos de prueba correspondientes y distinto número de neuronas y épocas. A continuación presentamos la selección que hicimos para cada uno de estos elementos.

3.1. Elección de las bases y conjuntos de entrenamiento

Las bases para apreciar el aprendizaje de la red son: 2, 4, 5, 10, 13, 20, 60, 73, 100. El conjunto de bases seleccionadas incluye:

- Bases usadas en diversas culturas.
- Bases primas.
- La base más usada a nivel mundial.

empezando por la derecha con la cifra de las unidades, a la izquierda las decenas, la siguiente las centenas, la siguiente los millares, etc. Se suman en primer lugar las cifras de la columna de las unidades según las tablas elementales, colocando en el resultado la cifra de unidades que resulte; cuando estas unidades sean más de 10 las decenas se acumulan como un sumando más en la fila de acarreo y así sucesivamente

[2]Para darnos una idea del uso del nepohualtzintzin basta remitirnos a las operaciones con números mayas.

[3]También puede ser usado en base 10 y de hecho algunos trabajos actuales para su difusión se llevan a cabo en esta base, sin embargo, se ha observado que usarlo en la base para la que fue creado facilita el aprendizaje de los niños. Aclaramos que no sucede lo mismo para los adultos pues estamos bastante acostumbrados a usar el sistema decimal y esto entorpece el manejo del *nepo*.

- La base usada en los sistemas computacionales.

- Por lo menos una base que ninguna cultura ha usado.

Uno de los principales problemas cuando se trabaja con redes neuronales es la representación de los datos. Ya que estamos analizando el problema de aprender a contar, proporcionamos a la red conjuntos de entrenamiento que simulen este proceso. A continuación describimos de manera general como son estos conjuntos y después damos un ejemplo de uno de los conjuntos tal y como lo dimos al perceptrón.

Los conjuntos de entrenamiento que proporcionamos a la red para cada base cumplen lo siguiente:

- Proporcionamos a la red 6 conjuntos de entrenamiento de tamaños[4] 5, 10, 20, 100 y dos de 500.[5]

- Estos conjuntos tienen 3 columnas, la primera es una codificación de la base, un valor comprendido entre [0, 1] y que varía para cada base; en la segunda listamos los números $0, 1, 2, ..., k$ y la tercera esta compuesta por los sucesores de cada uno de los números de la primera $1, 2, ..., k + 1$.

- Por cada conjunto de entrenamiento seleccionamos otro de prueba, es decir, le hacemos un examen a la red para ver que tanto aprendió y este conjunto también tiene 3 columnas organizadas de la misma forma que en los conjuntos de entrenamiento. El tamaño de la prueba es aproximadamente un tercio del conjunto de entrenamiento.

Para los conjuntos de entrenamiento de tamaño 5, 10, 20 elegimos números consecutivos comenzando con $0, 1, 2, ..., k$ donde k es el tamaño del conjunto de entrenamiento y procedimos de igual manera con los conjuntos de prueba respectivos iniciando con $k + 1, k + 2, ..., n$ hasta completar el conjunto de prueba.

Para los conjuntos de entrenamiento de tamaño 100 quitamos $\frac{1}{6}$ de los números del 0 al 99 de manera aleatoria, aunque tuvimos cuidado de no quitar los ejemplos clave: múltiplos y potencias, pues consideramos que estos ejemplos son vitales para el aprendizaje. Completamos este conjunto de tamaño con números mayores a 100; luego usamos el conjunto de $\frac{1}{6}$ (el que habíamos quitado previamente) como conjunto prueba y para completarlo tomamos $\frac{1}{6}$ de números consecutivos mayores a 100.

Para el conjunto 500a procedimos de igual manera manteniendo las proporciones requeridas para conformar los conjuntos de entrenamiento y prueba.

El conjunto 500b contiene los números del 0 al 499 y su conjunto de prueba los números del 500 al 666. Estos dos últimos conjuntos son relevantes por lo que más adelante veremos.

Para ejemplificar lo anterior proporcionamos el conjunto de entrenamiento de tamaño 5 correspondiente a la base binaria y su correspondiente conjunto prueba.

[4] El tamaño lo medimos de acuerdo al número de renglones.

[5] En adelante para diferenciarlos los denotaremos como 500a y 500b.

0.12	0.0000000000	0.0000000001
0.12	0.0000000001	0.0000000010
0.12	0.0000000010	0.0000000011
0.12	0.0000000011	0.0000000100
0.12	0.0000000100	0.0000000101

Cuadro 3.1: Conjunto de entrenamiento tamaño 5 para base binaria.

0.12	0.0000000101	0.0000000110
0.12	0.0000000110	0.0000000111

Cuadro 3.2: Conjunto de prueba correspondiente al conjunto anterior.

Los valores que proporcionamos a la red para analizar su desempeño deben estar dentro del intervalo [0,1], por lo tanto, expresamos cada conjunto de números de esta forma. Cualquier análisis llevado a cabo en el perceptrón multicapa debe cumplir con este requisito.

Es importante aclarar y establecer la diferencia entre lo que significa aprender a contar para los humanos y lo que quiere decir en el contexto de las redes neuronales. La manera en la hacemos que la red neuronal aprenda es a través de los conjuntos de entrenamiento, como podemos apreciar en los cuadros anteriores aprender a contar para la red neuronal significa que al aplicarle la prueba, ella será capaz de dar los sucesores de los números sobre los cuales le estamos preguntando.

3.2. Número de neuronas y épocas

El número de neuronas y épocas son otros factores a considerar en el aprendizaje del perceptrón. En nuestras pruebas variamos ambos.

En el programa que hicimos del perceptrón multicapa nosotros podemos variar el número de neuronas de la capa intermedia y el número de épocas

El número de neuronas para comparar el aprendizaje es :$1, 3, 5, 10$,y las épocas que consideramos son:$1, 5, 10, 15, 20, ..., 100$.

Las épocas corresponden al número de veces que mostramos el conjunto de aprendizaje a la red. Esto podría ser equivalente al número de veces que permitimos que un estudiante vea un determinado número de ejemplos.

Nos interesa determinar la combinación de neuronas y épocas para que el perceptrón mejore su desempeño. Elegimos los datos anteriores considerando queu para un primer acercamiento al desempeño del perceptrón multicapa en la tarea de aprender a contar estas combinaciones proporcionan un marco general para responder la pregunta principal de nuestro trabajo.

3.3. Datos relevantes para las pruebas

Para hacer una prueba en la red neuronal se eligen:

- Un conjunto de entrenamiento.
- Un conjunto de prueba.
- Número de neuronas.
- Número de épocas.

Estos datos conforman el conjunto de entradas para el perceptrón, las variaciones en cada uno de estos parámetros determinan el aprendizaje. La salida que obtenemos proporciona el tamaño del **error**. Como antes lo indicamos, el estudio de los errores indica el éxito o fracaso en el aprendizaje en la red.

Existen diferentes métodos para estudiar el error, entre ellos encontramos el entrenamiento **online** que es usado cuando se estudia cada patrón de entrenamiento y la actualización de los pesos sinápticos se produce tras cada ejemplo de entrenamiento, si un conjunto de entrenamiento posee n ejemplos se dan n corrección de pesos por época. Por ejemplo, a nosotros nos interesa el desempeño de la red para cada conjunto de entrenamiento.

Otro modo de estudiar el error es a través de lo que conocemos como entrenamiento **batch** o por lotes para el cual la actualización de los pesos se produce una única vez tras la presentación de todo el conjunto de entrenamiento. [Haykin, 1999]

Debido al uso que estamos haciendo del perceptrón multicapa como un agente de aprendizaje, y con el fin de determinar la combinación para la cual el error disminuye elegimos el criterio de entrenamiento online.

El error se encuentra definido como la diferencia entre lo que se espera y lo que se obtiene. Es decir:

$E_k = error$

$D_k = salidadeseada$

$O_k = salidaobtenida$

De manera general la función error queda definida como:

$E_k = [D_k - O_k]^2$

Para ejemplificar lo anterior tomemos como base el ejemplo que pusimos del conjunto de tamaño 5 en la base binaria, supongamos que queremos comparar el aprendizaje con la combinación de 1 neurona 1 época contra 10 neuronas y 100 épocas. Entonces nuestro archivo de salida es un arreglo de 2 renglones y 3 columnas. La primer columna corresponde al número de neuronas, la segunda las épocas y la tercera es el error correspondiente a cada una de estas combinaciones como se observa en la figura siguiente:

Lo que hace la red neuronal para arrojar este error es: modificar los pesos de acuerdo al conjunto de entrenamiento y en base al algoritmo de backpropagation, para luego dar sus respuestas al examen que nosotros aplicamos (conjunto prueba) luego hacemos la diferencia renglón a renglón de las respuestas obtenidas y las deseadas, las elevamos

Neuronas	Épocas	Error
10	100	0.00000000553476762
1	1	0.00143431765848000

Cuadro 3.3: Ejemplo de conjunto de errores.

al cuadrado, las sumamos y sacamos el promedio, el error que presentamos es este promedio.

Los archivos de salida que nosotros obtuvimos son iguales a los del Cuadro 3.3, con 84 renglones, que presentan el erro para cada uno de las combinaciones del número de épocas y neuronas.

Las gráficas siguientes son comparaciones entre las bases, en ellas mostramos distintos aspectos del aprendizaje.

Las gráficas representan los errores máximos de la red para cada conjunto de salida. Cada curva corresponde a un conjunto de entrenamiento[6] de tamaño fijo y separamos las gráficas de acuerdo a éste para evidenciar los resultados, la primera corresponde a los conjuntos de tamaño 5, 10 y 20 y la segunda muestra los conjuntos restantes. Las gráficas indican: el tamaño del conjunto, la base, su error máximo y cómo se ubican entre sí.

En los conjuntos 5 y 10, el error es relativamente pequeño comparado con los conjuntos mayores.

En el conjunto 20 observamos que la base 73 presenta los errores máximos, seguida en todos los casos de la base 10;la base 20 presenta los errores mínimos dentro de esta clasificación en los conjuntos 20, 500a y 500b y la base 2 para el conjunto 100.

[6]Con fines prácticos hacemos referencia al conjunto de entrenamiento, sin embargo, es importante tener presente que el error se determina de acuerdo a la "calificación" obtenida en base al **conjunto de prueba**

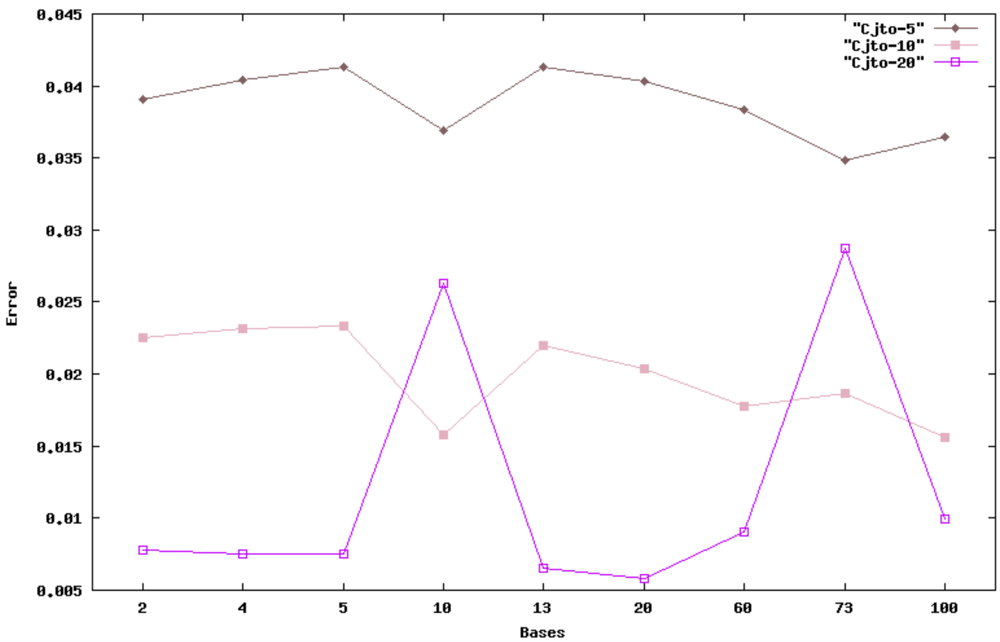


Figura 3.1: Errores máximos para los conjuntos de 5, 10 y 20 elementos.

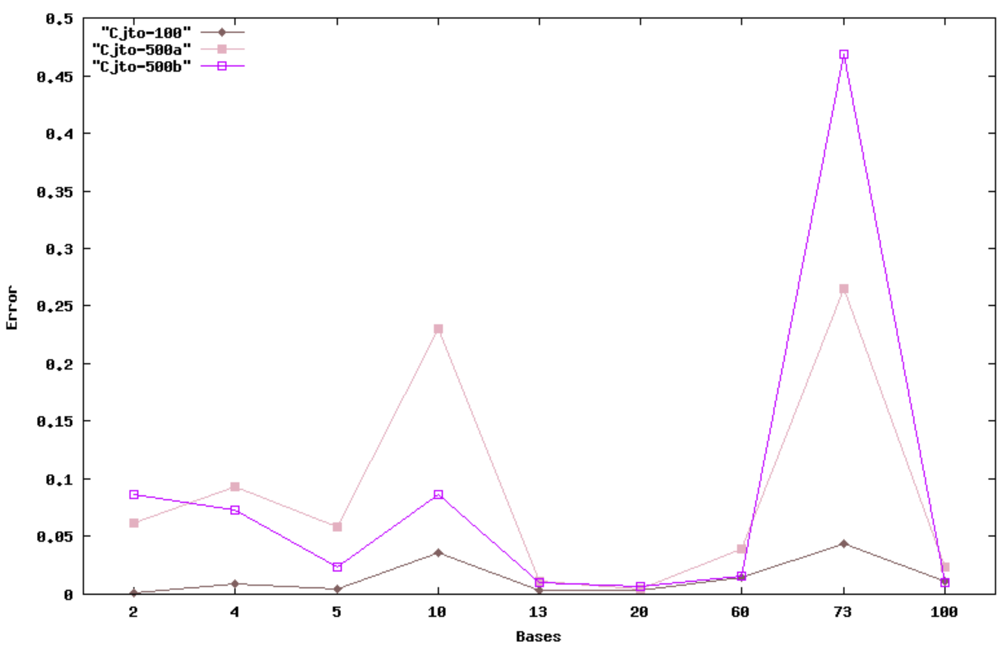


Figura 3.2: Errores máximos para los conjuntos 100, 500a y 500b elementos.

Las siguientes gráficas muestran los errores mínimos. La base 2 presenta el mejor rendimiento para los conjuntos 5, 10 y 20.

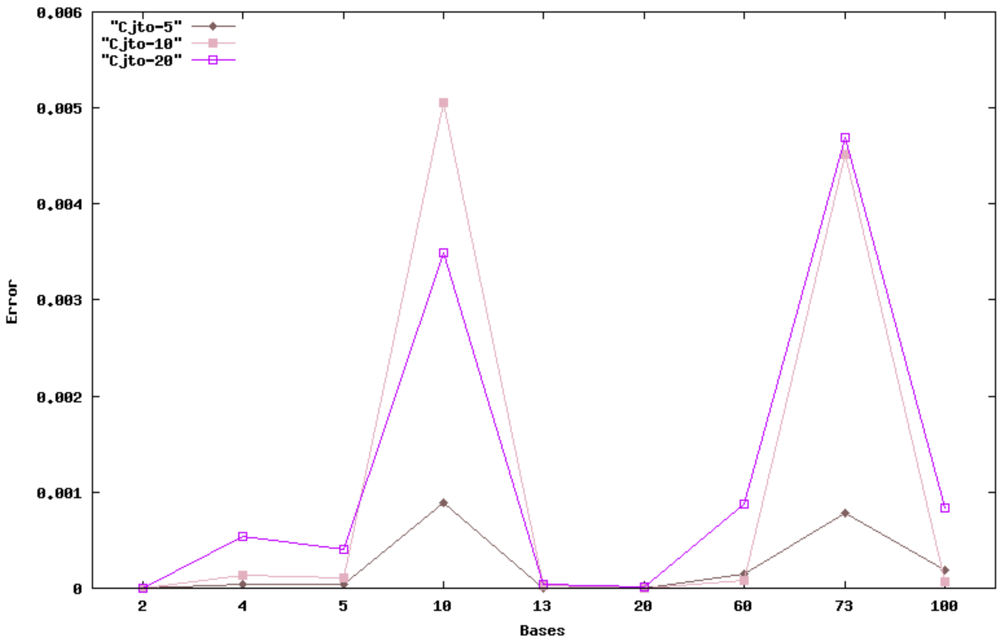


Figura 3.3: Errores mínimos para los conjuntos 5, 10 y 20 elementos.

En los otros conjuntos (figura 3.4) observamos que la base 20 obtiene el error más bajo.

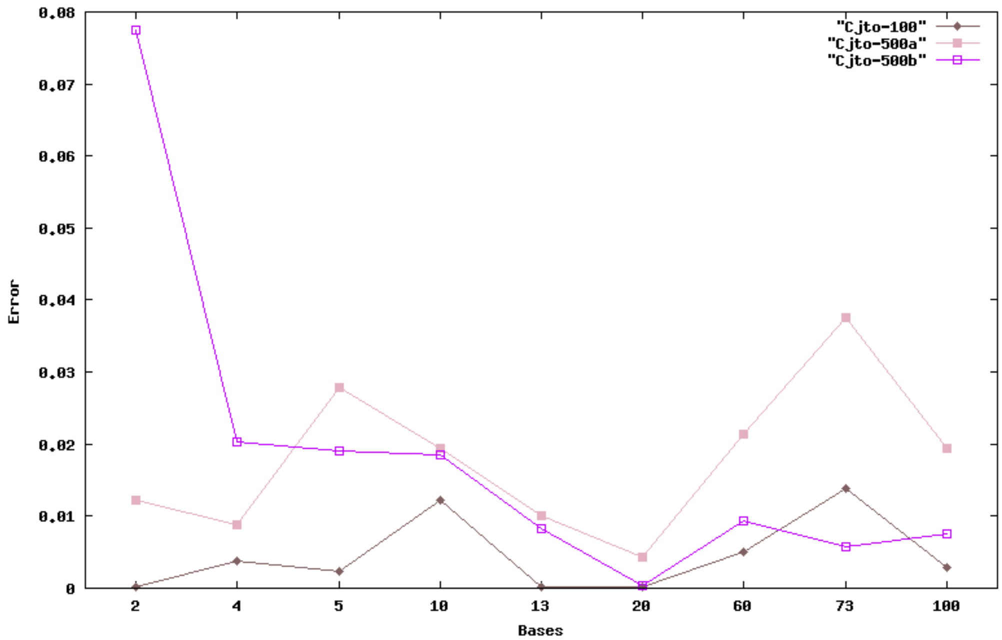


Figura 3.4: Errores mínimos para los conjuntos 100, 500a y 500b elementos.

Las siguientes gráficas indican la diferencia entre el error máximo y el mínimo, en ellas observamos cual red consigue disminuir el error de manera más efectiva.

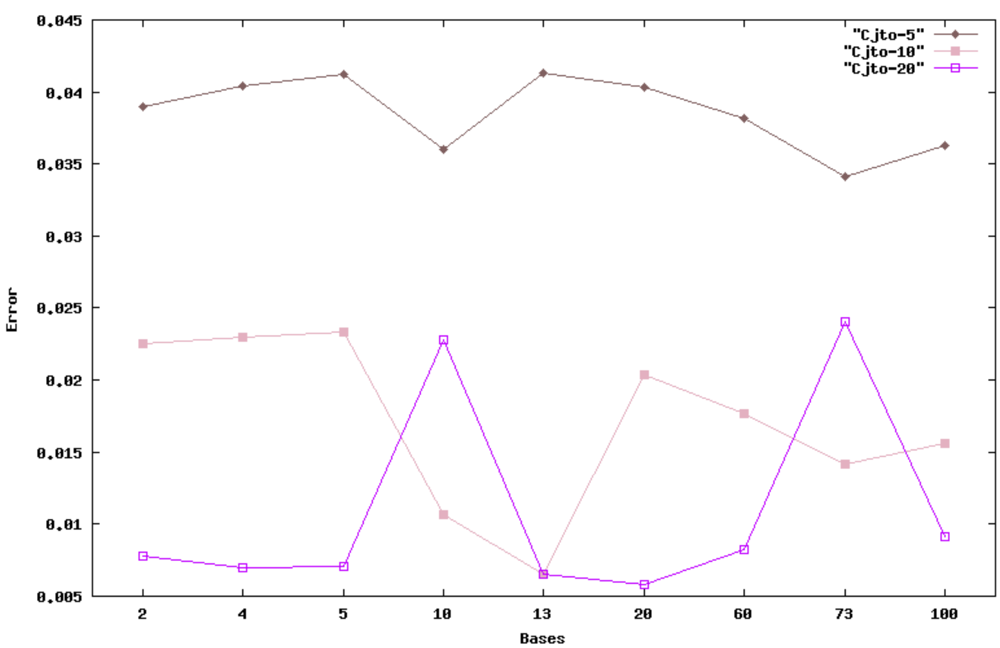


Figura 3.5: Disminución de error en los conjuntos 5, 10 y 20 elementos.

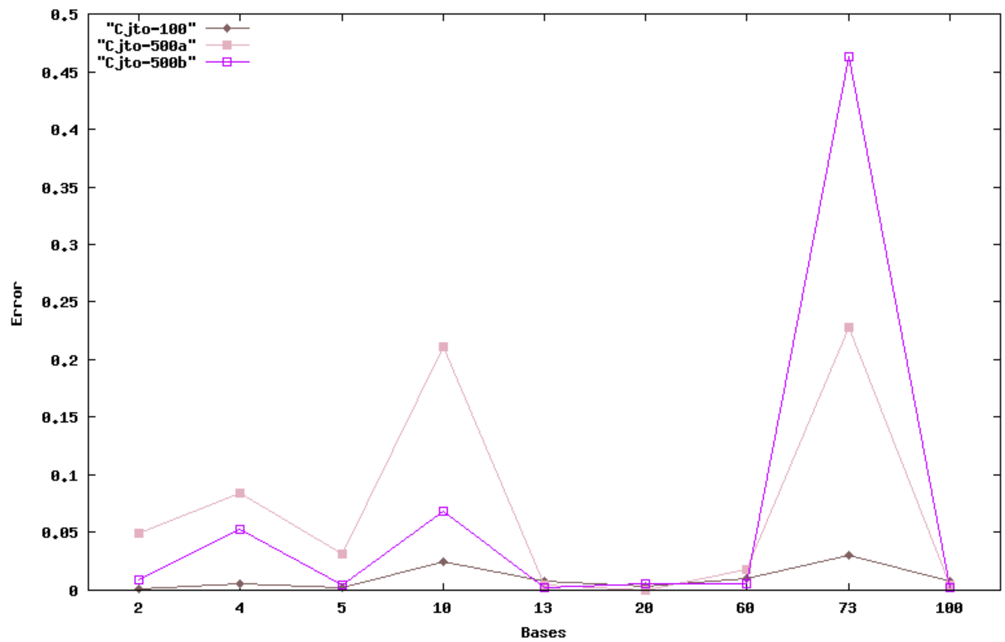


Figura 3.6: Disminución de error en los conjuntos 100, 500a y 500b elementos.

Como especificamos al principio de este capítulo, otra cuestión importante en las redes neuronales es examinar qué papel juegan en el aprendizaje el número de neuronas y de épocas.

Para los conjuntos 5, 10 y 20 se consigue el error mínimo con las siguientes combinaciones de neuronas y épocas respectivamente: 10 y 30, 10 y 15, 1 y 80.

En el conjunto 100 el error mínimo se consigue con 10 neuronas y 1 época. Para los conjuntos 500a y 500b, el error mínimo se obtiene con la capacidad máxima de la red: 10 neuronas y 10 épocas.

Los errores máximos en los primeros tres conjuntos se "ganan" con la combinación 10 neuronas y una época. Y para los conjuntos restantes con una neurona y una época.

Finalmente, después de mostrar resultados globales, queremos presentar algunas características interesantes del aprendizaje con base 20, y también algunas diferencias entre esta base y la base 2, por ser la que obtiene mejores resultados en conjuntos pequeños, y la base 10 por el interés que creemos, nos atañe a todos.

La primera gráfica muestra la diferencia entre el desempeño del perceptrón en los conjuntos 500a y 500b, con la base 20. En ambos conjuntos la base 20 consigue disminuir el error más que las otras bases, pero muestra mayor eficiencia en el conjunto 500b.

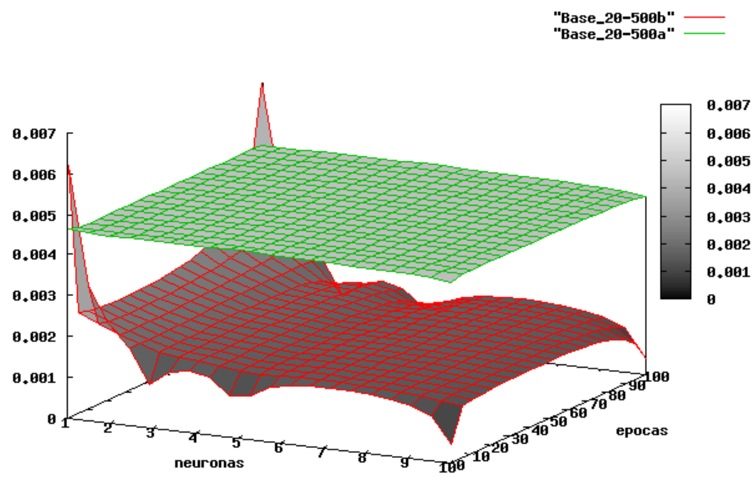


Figura 3.7: Base 20

La base 2 es la óptima para los conjuntos 5, 10 y 20. Para los conjuntos mayores esto no sucede. En el conjunto 500b esta diferencia se aprecia con mayor claridad.

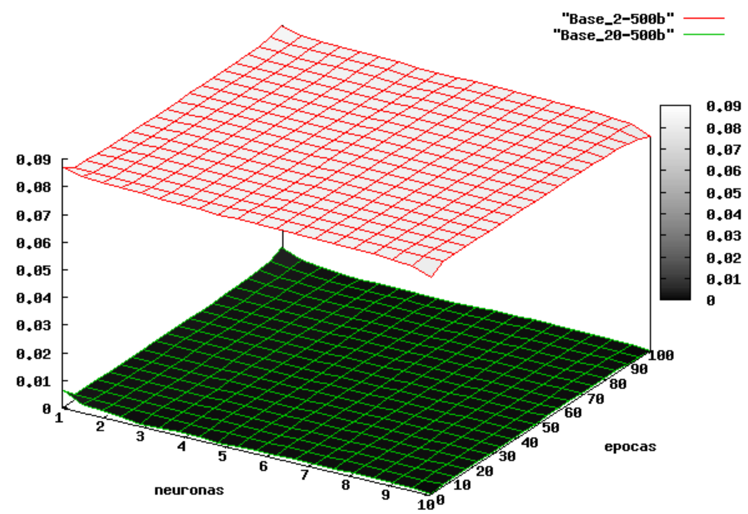


Figura 3.8: Bases 2 y 20, conjunto 500b

La figura 3.9 nos indica la diferencia entre los errores del perceptrón con las bases 10 y 20.

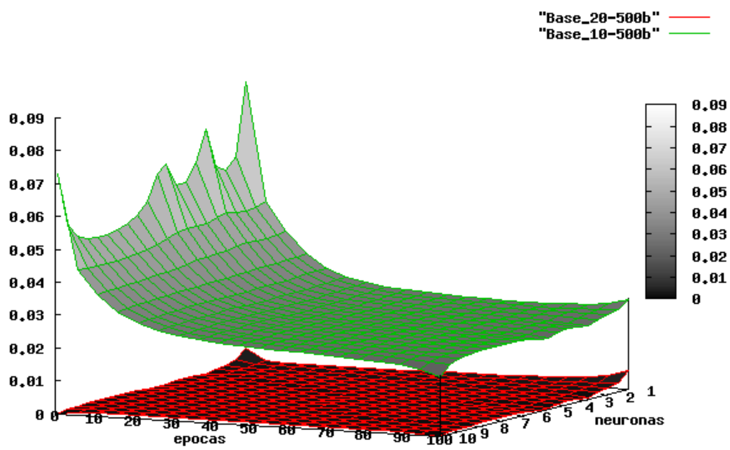


Figura 3.9: Bases 10 y 20

Finalmente la figura 3.10 dan una idea del comportamiento general para algunas de las bases más representativas en nuestro trabajo.

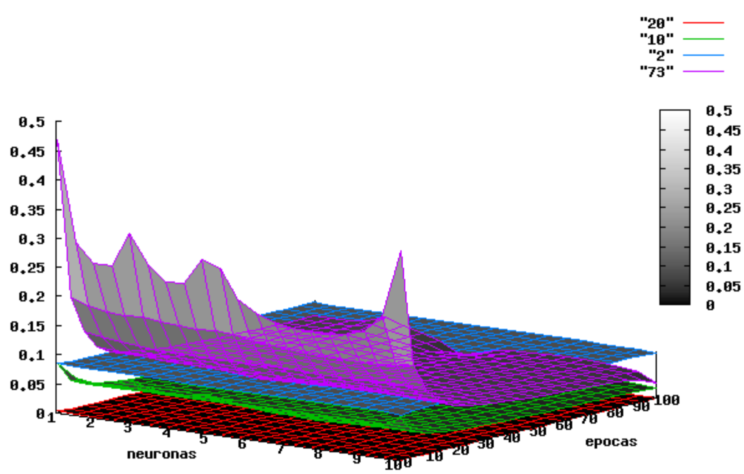


Figura 3.10: Bases 2, 10, 20 y 73 en el conjunto 500b

3.4. Discusión de los resultados

Con ayuda de las redes neuronales analizamos las diferencias que existen al contar en distintas bases numéricas para el perceptrón multicapa.

Los resultados presentamos en la sección anterior permiten analizar algunas propiedades en el aprendizaje del perceptrón multicapa. A partir de ellos podemos responder, dentro de nuestro marco experimental, la pregunta que originó la investigación, *¿existe una base óptima para aprender a contar?* y de ser así, *¿cuál es esa base?*

Comparamos el aprendizaje de la red y observamos que la base 2 consigue disminuir el error en los conjuntos 5, 10 y 20; además logra el mínimo global en el conjunto 5.

En los conjuntos 100, 500a y 500b, la base 20 consigue el mejor resultado.

Cabe resaltar que en los conjuntos 10 y 20 la base vigesimal ocupa el segundo lugar en errores mínimos, mostrando su eficiencia también en conjuntos pequeños. Otra particularidad es que el error presenta un comportamiento homogéneo, es decir, cuando variamos el número de neuronas y épocas, no se observa grandes saltos en el tamaño del error.

Gracias a las comparaciones descartamos la importancia del número de símbolos que la base debe manejar ya que no es un factor determinante. Muestra de esto es la base 100, que obtiene mejores resultados que otras bases más pequeñas.

Recordemos que nos interesa conocer el desempeño de la base decimal. Ahora sabemos que la base que usamos tiene errores comparables con la base 73 (Ver Fig.3.3 y 3.4); cuando en la primera disminuye o aumenta el error, ocurre lo mismo en la segunda, y este comportamiento se mantiene en todas las comparaciones.

Lo anterior es relevante porque hemos visto que la base 73 muestra el peor desempeño en el aprendizaje; en las bases 10 y 73 encontramos los errores máximos a partir del conjunto 20.

Lejos de lo que podíamos augurar nuestra base presenta el peor desempeño, sólo superada en éste por la base 73; esto es revelador pues intuitivamente podríamos pensar que por ser la base usada convencionalmente, podría ser eficiente también para el perceptrón, lo cual no ocurre.

Uno de los resultados más importantes en nuestro estudio fue analizar los errores en los conjuntos 500a y 500b. Como ya mencionamos al principio de este capítulo el conjunto 500a se construyó de modo que la red *completara* los huecos en la lista de números del 0 al 666. Sospechamos que los errores en este conjunto serían menores, pero encontramos que las únicas bases donde esto ocurre son para la 2 y 4.

La diferencia fundamental entre los conjuntos 500a y 500b consiste en que para el último, la red no cuenta ni con el sucesor ni el antecesor[7] de los números solicitados por el conjunto de prueba.

En este conjunto la red demuestra que puede "caminar" sola en la tarea de contar; basta darle una cantidad suficiente de ejemplos.

De esta manera mostramos que el desempeño de la red es mejor cuando construye toda la serie, que cuando tiene que completarla.

[7]Excepto, claro está, el primer elemento del conjunto de prueba.

Por ello la base 20 presenta las cualidades que permiten un mejor aprendizaje para el perceptrón multicapa. Para el conjunto 500b obtiene resultados excelentes; asimismo presenta una diferencia significativa en disminuir su error respecto a los otros conjuntos.

Para algunas bases el desempeño de la red es mejor cuando construye toda la serie, que cuando tiene que completarla.

Además es la única que aún en el peor de los casos, es decir, en los errores máximos, tiene menor error y al variar el número de neuronas y épocas el tamaño de su error no presenta saltos significativos. Es cierto que para los conjuntos de 500 elementos la red requiere de toda su capacidad para disminuir el error, sin embargo, casi todas las combinaciones de épocas y neuronas presentan resultados favorables.

Aunque la base 2 también tiene resultados satisfactorios, los mantiene sólo para conjuntos pequeños; en el conjunto 500b, sobre el cual hemos hecho énfasis, la base dos ocupa el último sitio dentro de la clasificación de los errores mínimos. Por esta razón consideramos a la base 20 como la que presenta las mejores cualidades para que el perceptrón aprenda a contar.

Conclusiones

El afán del ser humano por comprender sus procesos mentales es un motor que ha permitido la creación de modelos que, aunque no logran explicar totalmente la complejidad del funcionamiento del cerebro o para decirlo mejor, gracias a que aún no se le comprende, es una fuente que continúa inspirando la creación de algoritmos.[8]

Los problemas en la enseñanza de las matemáticas son numerosos. Para los interesados en el tema resulta importante analizar las causas de las dificultades para aprender.

El problema relativo al aprendizaje ha sido abordado por un considerable número de teorías pedagógicas. Una pregunta esencial en ellas debería ser: *¿cómo aprendemos?*. Si la teoría sostiene que es capaz de dar solución a esta interrogante propondrá técnicas dirigidas a mejorar el aprendizaje. Y en el mejor de los casos, como se ha comprobado, estas técnicas serán eficaces para algunas personas.

Si partimos desde otro punto de vista, podemos hacer otra pregunta, esta vez la respuesta la encontramos en nuestra experiencia personal, ¿cómo aprendimos alguna tarea específica? o aún más, tomemos como referencia alguna capacidad o conocimiento que hemos sido capaces de desarrollar excepcionalmente, ¿cómo lo hemos conseguido?

Suponemos que la respuesta no es única y esta peculiaridad hace sumamente interesante este fenómeno educativo.

Si el aprendizaje pudiera reducirse a una técnica, el problema hace mucho podría haberse resuelto por lo menos parcialmente. Pero es precisamente la abundancia de técnicas y teorías lo que nos permite vislumbrar que la capacidad de aprender y crear, no pueden ser simplificadas a métodos de enseñanza o técnicas de aprendizaje.

Aunque se afirme que "...el carácter relativamente reciente en el área de conocimiento de la didáctica de las matemáticas explica que no exista aún un paradigma de investigación consolidado y dominante"[9] [Godino, et al, 2006]. Algunos investigadores[10] en matemática educativa sugieren un camino: el estudio de la historia de las matemáticas.

Esta investigación contribuye al replanteamiento acerca de la importancia de los sistemas semióticos de información, creados por distintas culturas. Las producciones culturales de cada pueblo ofrecen nuevos significantes que orientan en la búsqueda hacia

[8] Como ejemplo de esto tenemos el aprendizaje Hebbiano, aprendizaje comparativo y competitivo y las variaciones en los modelos conocidos para mejorarlos.

[9] Las demás áreas del saber tampoco cuentan con el paradigma que garantice su enseñanza y aprendizaje.

[10] Un ejemplo de esto es la reunión que organiza el *Grupo Internacional de Estudio sobre las Relaciones entre la Historia y la Pedagogía de las Matemáticas.*

la comprensión de los procesos que hacen posible el saber y las creaciones culturales.

El uso de las redes neuronales demuestra que en el caso de aprender a contar en distintas bases, [11] es posible encontrar una que logra disminuir el error.Una conclusión importante es que aunque lo que se requiere aprender es lo mismo en todos los casos, no es lo mismo hacerlo en una base u otra.

Además, las redes neuronales resultan de interés para los encargados del curriculo escolar. Los recursos que se requieren para su uso son tanto cualitativa como cuantitavamente mejores para experimentar.

Probar determinadas ideas o teorías pedagógicas *in silico* es mucho más fácil y ético que hacerlo *in situ* o *in vivo*. Trabajar con redes neuronales y modelos computacionales y matemáticos pueden ayudar a los pedagogos a depurar, validar y proponer nuevas teorías.

A continuación: especulemos con los resultados obtenidos.

Los resultados obtenidos en el aprendizaje de las redes neuronales, resultan de interés por su aportación para comprender la cognición humana.

Un punto importante es la necesidad de cuestionar el currículo escolar. Cómo y para qué se enseñan las matemáticas y sobre todo, qué matemáticas se enseñan.

Supongamos que podemos extrapolar el aprendizaje del perceptrón. Entonces el uso mundial de la base decimal tendría que ser revisado cuidadosamente. La base decimal es una de entre muchas bases usadas históricamente. ¿En qué momento y por qué la adoptamos?, ¿deberíamos seguir usándola?

La infancia se caracteriza por la enorme cantidad de conocimiento que durante ella adquirimos. En ella desarrollamos capacidades físicas y también somos capaces de desarrollar una cualidad humana básica y extremadamente compleja que nos caracteriza: el lenguaje. Los primeros años de vida son fundamentales para el aprendizaje posterior. Por tal motivo el estudio de los encuentros del niño con los saberes matemáticos presentan especial importancia.

Al respecto podemos comparar los señalamientos hechos por Piaget acerca de la necesidad de comprensión lógica, con el número de épocas y neuronas que la red necesita para disminuir su error. Observamos que para los conjuntos 5 y 10, los mejores resultados se obtienen con el mayor número de neuronas y se requiere de pocas épocas, podriamos constatar que importa más la "madurez" que el entrenamiento. En los conjuntos 500a y 500b, re requiere de ambas condiciones madurez y entrenamiento, para lograr resultados exitosos.

En nuestro trabajo encontramos que la base vigesimal presenta cualidades atractivas para proponer su uso. Tenemos la ventaja de contar con un refente para el análisis en la "vida real" de este hecho.

Las culturas precolombinas usaron la base vigesimal, siendo los mayas el ejemplo más conocido. Gracias al legado de las culturas prehispánicas podríamos establecer una relación entre los resultados que obtuvimos y el uso de la base vigesimal para la creación del saber matemático.

[11] Aunque los números no constituyan por sí solos un sistema semiótico de información, si constituyen la base para los mismos.

Así mismo la historia del conocimiento matemático nos introduce en una seria reflexión acerca de las creaciones que conformaron el bagaje tecnológico, artístico y cultural de muchos pueblos. Quizá el reencuentro de este conocimiento provoque los beneficios esperados de una sociedad que *disfrute de las matemáticas*.

Trabajo futuro

Es importante señalar las posibilidades de las redes neuronales como modelos del aprendizaje en humanos. Con el desarrollo de redes más grandes, con topologías más realistas y mecanismos de aprendizaje que se parezcan más al de los humanos, se podrían diseñar experimentos para que el aprendizaje fuera más rápido, con menos error, etc.

Por ejemplo, en nuestro trabajo podemos hacer nuevos experimentos trabajar con nuevas bases, con conjuntos de entrenamiento de mayor tamaño, hacer nuevas combinaciones de épocas y neuronas.

El uso de las redes como parte de una metodología de investigación educativa es una variante importante en pedagogía.

En este punto retornamos, siempre con fundamento en las bases neuronales, a la investigación original: ¿el uso del nepohualtzintzin podría favorecer el acercamiento con los números?

Respondimos la pregunta acerca de las bases, sin embargo, queda pendiente el estudio de los algoritmos para las operaciones básicas que el nepo ofrece.

La investigación de Terezhina Nunes, acerca de los algoritmos creados por niños y adultos para las operaciones básicas, se relaciona con el algoritmo usado por el nepohualtzintzin.

En una de las teorías expuestas anteriormente, creada por Vygotsky, encontramos otra razón para el uso del nepo. El hombre crea herramientas no sólo útiles para transformar su medio sino para transformarse así mismo en sus funciones, la percepción, la memoria y el pensamiento. De ésta manera el nepo es una herramienta que permite un replanteamiento de la forma en la que se aprende a contar y, de la lógica subyacente en los algoritmos de las operaciones básicas.

El nepo señala otro camino al remitirnos al uso de los números mayas considerando que “La característica principal del sistema de numeración maya consiste en que los símbolos que se utilizan tienen un valor intrínseco. En otras palabras, que en sí mismos contienen la multiplicidad que describen” [Calderón, 2000]

En el mismo sentido podría servir para fundamentar otras discusiones como el uso *sin trámites* de la calculadora y sus repercusiones en el pensamiento y aprendizaje. Funciona como un primer acercamiento para comprender otras repercusiones en el ámbito de las nuevas tecnologías.

El estudio de los procesos cognoscentes se enriquece con el uso de las redes neuronales. Los resultados que estas proporcionan son a su vez material para vislumbrar nuevas relaciones con la pedagogía, historia y todas aquellas áreas del saber interesadas en comprender la dinámica del conocimiento.

Bibliografía

[Abbott y Regehr, 2004] Abbott, L., y Regehr, W., 2004, "Synaptic Computation," *Nature* **431** pp. 796-802.

[Anderson, 2007] Anderson, J., 2007, *Redes Neuronales*, Alfaomega Grupo Editor, México.

[Backhoff et al, 2008] Backhoff, E., Andrade, E., Sánchez A., Peon, M., 2008, *El aprendizaje en tercero de preescolar en México, Lenguaje y comunicación, Pensamiento matemático*, INNE, México.

[Blanck, 1977] Blanck, G., 1977, *La determinación social de la actividad psíquica específicamente humana.* Stokoe, Buenos Aires.

[Calderón, 2000] Calderón, H., 2000 *La Ciencia Matemática de los Mayas*, Editora Cuzamil, México.

[Cambray, 2003] Cambray, R., 2003, *Memorias de las Conferencias Plenarias*, Universidad Michoacana de San Nicolás de Hidalgo, Morelia, Michoacán , México, pp. 59-75.

[Courant y Robbins, 2002] Courant, R. y H. Robbins, 2002, *¿Qué son las matemáticas? Conceptos y métodos fundamentales*, Fondo de Cultura Económica, México.

[Esparza, 1976] Esparza, E., 1976. *Cómputo azteca* Editorial Diana, México.

[Del Brío, 2001] Del Brío, B. y Sanz, M., 2001, *Redes Neuronales y Sistemas Difusos*, Alfaomega Grupo Editor, México.

[Díaz et al, 2007] Díaz, M., Flores, G., Martínez, F., 2007, *PISA en México*, INEE, México.

[Duval, 1996] Duval, R., 1996, Quel cognitive retenir en didactique des mathématiques? *Reserches en Didactique des Mathématiques.*

[Filloy, 2003] Filloy, E. (coord.), 2003, *Matemática Educativa. Aspectos de la investigación actual*, Fondo de Cultura Económica, México, pp.174-186.

[Fogel, 1999] Fogel, D., 1999 *Blondie24.* Morgan Kaufmann, San Francisco.

[Freire, 1973] Freire, P., 1973, *¿Extensión o comunicación? La concientización en el medio rural*, Siglo XXI editores, México.

[Funahashi, 1989] Funahashi, K., 1989. "On the approximate realization of continuous mapping by neural networks." *Neural Networks*, **2**, 183-192.

[Gemignani, 1967] Gemignani, M.C., 1967, "On the Geometry of Euclid", *The Mathematics Teacher 60*, p. 160-164.

[Godino, et al, 2006] Godino,J., et al, 2006, "Una visión de la didáctica francesa desde el enfoque ontosemiótico de la cognición e instrucción matemática." *RELIME*, **Vol. 1**, 117-150.

[Haykin, 1999] Haykin, S., 1999. *Neural Networks: A Comprehensive Foundation.* Prentice Hall.

[Kandel et al, 2001] Kandel, E., Schwartz, J. y Jessell, T., 2001 *Principios de neurociencia*, McGraw-Hill Interamericana de España, Madrid.

[Kohonen, 1997] Kohonen, T., 1997, *Self-Organising Maps*,Springer Verlag, Nueva York.

[McCulloch y Pitts, 1943] McCulloch, W. y Pitts, W., 1943, "A logical calculus of the ideas immanent in neurons activity", *Bull. Math. Biophys.*, **5**, 115-133.

[Nunes, 2001] Nunes, T. 2001 "La matemática en la vida y en la escuela : dos décadas de investigación" *Pluriculturalidad y aprendizaje de la matemática en América Latina : experiencias y desafíos.* Madrid, pp. 234-252.

[Nunes y Bryant, 2003] Nunes, T. y Bryant, P., 2003, *Las matemáticas y su aplicación: la perspectiva del niño*, Siglo XXI de España Editores, Madrid.

[Paolucci, 2000] Paolucci, R., Colli, P., y Giacinto, G., 2000, Assessment of Seismic Site Effects in 2-D Alluvial Valleys Using Neural Networks. *Earthquake Spectra,* Volúmen 16, Number 3.

[Piaget, 1969] Piaget J., 1969 *Biología y conocimiento*, Siglo XXI, España.

[Piaget, 1978] Piaget, F., 1978 *La equilibración de las estructuras cognitivas*, Siglo XXI, Madrid.

[Rodríguez-Consuegra, 2001] Rodríguez-Consuegra, F., 2001, Filosofía de las matemática: Tres preguntas fundamentales, *Mathesis, Serie II, I (1)*, pp. 79-114.

[Rojas, 1996] Rojas, R., 1996 *Neural networks: a systematic introduction* Springer Verlag, Nueva York.

[Rosenblantt, 1961] Rosenblatt F., 1961 *Principles of neurodynamics: Perceptrons and the theory of brain mechanisms.* Spartan Books.

[Sakata y Yamamori, 2006] Sakata, Sh. y Yamamori, T., 2006, Topological relationships between brain and social networks *Science* pp.12-21.

[Strogatz, 2001] Strogatz, s., 2001, "Exploring complex networks", *Nature* **410** pp. 268-276.

[Tudge, 1989] Tudge, J. y Rogoff, B., 1989, "Peer influences on cognitive development: Piagetian and Vygotskian perspectives", *Interaction in human development.* Hillsdale, New York.

[Tatuzov, 1967] Tatuzov, A., 2006, "Neural network models for teaching multiplication table in primary school", *2006 International Joint Conference on Neural Networks*, Vancouver, Canada.

[Vygotsky, 1983] Vygotsky, L., 1983. *La imaginación y el arte en la infancia.* Akal, Madrid.

[1] SEP, 2008 *Resultados de la prueba ENLACE 2007* [en línea]. http://enlace2007.sep.gob.mx/ (9 sep. 2008)

[2] PISA, 2003, *Reactivos PISA 2003*, [en línea], http://www.inee.edu.mx/images/stories/ReactivosPISA/2003/m413q01-q02-q03tipodecambio.pdf (15 sep. 2008)

[3] ENLACE, 2007 *Pruebas aplicadas ENLACE*, [en línea], http://enlace.sep.gob.mx/ba/?p=descpruebas (20 sep. 2008)

[4] Alfaro, A., 2009. Maestros "sabotearon" la prueba ENLACE en el plantel que salió peor calificado. *La Jornada* [en línea] http://www.lajornadadeoriente.com.mx/2009/10/12/puebla/edu107.php (1 sep. 2009)

[5] Sarle, W., 1998, *Neural network FAQ. Periodic posting to the Usenet newsgroup comp.ai.neural-nets*, ftp://ftp.sas.com/pub/neural/FAQ.html. (15 nov. 2009)

[6] Avilés, K., 2010 Suicidio y narco, "opciones" de jóvenes por el fracaso educativo. *La Jornada* [en línea] http://www.jornada.unam.mx/2010/01/12/index.php?section=sociedad &article=033n1soc (12 ene. 2009)