



**UNIVERSIDAD AUTÓNOMA DEL ESTADO DE MÉXICO**

**CENTRO UNIVERSITARIO UAEM VALLE DE MÉXICO**

**Aprendizaje incremental para la tarea de reconocimiento  
de dígitos con Redes Neuronales Artificiales**

**PROTOCOLO DE TESIS**

Que para obtener el Título de

**INGENIERO EN SISTEMAS Y COMUNICACIONES**

P r e s e n t a

**C. Fragoso García Sandra  
C. González Hernández Luis Ángel**

**Asesor: Dr. Víctor Manuel Landassuri Moreno**

**Atizapán de Zaragoza, Edo. de Méx. octubre 2022**



## Contents

<b>1</b>	<b>Introducción</b>	<b>3</b>
<b>2</b>	<b>Planeamiento</b>	<b>4</b>
<b>3</b>	<b>Objetivos</b>	<b>5</b>
3.1	Objetivos Particulares . . . . .	5
<b>4</b>	<b>Justificación</b>	<b>5</b>
<b>5</b>	<b>Delimitación</b>	<b>7</b>
<b>6</b>	<b>Consecuencias</b>	<b>7</b>
<b>7</b>	<b>Marco Teórico</b>	<b>8</b>
7.1	Redes Neuronales Artificiales . . . . .	8
7.2	Redes Neuronales de Perceptrón Multicapa . . . . .	8
7.3	Algoritmo Backpropagation . . . . .	9
7.4	Redes Neuronales Convolucionales . . . . .	9
7.5	Aprendizaje . . . . .	10
7.5.1	Aprendizaje en humanos . . . . .	10
7.5.2	Aprendizaje Humano . . . . .	10
7.5.3	Aprendizaje Activo . . . . .	11
7.6	Aprendizaje Incremental . . . . .	11
7.6.1	Algoritmos de Aprendizaje Incremental . . . . .	11
<b>8</b>	<b>Hipotesis</b>	<b>12</b>
<b>9</b>	<b>Metodología</b>	<b>12</b>
<b>10</b>	<b>Cronograma de Actividades</b>	<b>12</b>
<b>11</b>	<b>Organización del Capitulado</b>	<b>13</b>

## List of Figures

1	Red Neuronal de Perceptrón Multicapa [liu2015] . . . . .	9
2	Esquema de una Red Convolucional [duran2017] . . . . .	10

# 1 Introducción

La inteligencia artificial (IA) es una área del conocimiento que se enfoca en poder hacer máquinas que tengan un comportamiento y razonamiento humano, para que en un momento, se pueda interactuar con ellas sin darse cuenta que se está interactuando con una máquina. Así mismo, también es posible pensar que mucho del desarrollo en el área de inteligencia artificial, es el poder tener mejores herramientas que ayuden a las actividades diarias.

En este sentido, un área de la IA es el llamado Aprendizaje Máquina, donde se estudian algoritmos que permiten aprender de forma automática una tarea. Así, una de las técnicas más conocidas en la actualidad, dentro del área de IA son las Redes Neuronales Artificiales (RNAs), siendo técnicas que realizan procesos matemáticos para poder aprenderse tareas a resolver. Algunas áreas en las que son útiles las RNAs son en el aprendizaje de tareas no lineales, como la predicción de la capacidad de la red 5G, basada en el tráfico diario de este [zhao2022] o clasificación, por ejemplo la clasificación de metales y rocas por medio de RNAs y lógica difusa [salazar2013].

Las RNAs están formadas por neuronas artificiales que simulan a las biológicas. Así los procesos químicos que suceden en el cerebro, se simulan computacionalmente a través de señales que viajen a través de las neuronas artificiales, de aquí en adelante simplemente se referirá a ellas como "neuronas". Las neuronas en una RNA cuentan con una estructura distribuida en paralelo, presentando una buena habilidad de aprendizaje [liu2015].

Dentro del aprendizaje máquina, cuando una técnica, por ejemplo RNA, se enfoca en aprender una tarea, se le conoce como *algoritmo lineal sin memoria* siendo uno de los métodos más empleados desde el inicio de las RNAs [GiraudCarrier2000]. Sin embargo, si es necesario incorporar nueva información del problema, es necesario volver a entrenar todo el modelo, considerando toda la información existente, esto es la anterior y la nueva que acaba de llegar, es ahí donde nace el concepto de Aprendizaje Incremental, siendo un área enfocada en poder incorporar información del problema en cuestión, sin tener que volver a re-entrenar todo el modelo.

Derivado del aprendizaje incremental se desprende el concepto de memoria dentro de la IA, analizando como un algoritmo de aprendizaje máquina puede olvidar la información que se usó en un entrenamiento previo, al entrenar con información más reciente. Si se hace la analogía con los humanos la memoria, es un factor importante para estudiar considerando la pérdida de información aprendida, así este es un problema biológico, el cual tanto afecta a los humanos como a las máquinas. Por ello, se han elaborado distintos experimentos para poder combatir esta problemática. Uno de estos es el caso de [bullinaria2009], donde se propone el manejo de RNAs con pesos dobles, donde la primer capa de pesos esta enfocada a comportarse como memoria a corto plazo, y la segunda como memoria a largo plazo. Los experimentos mostrados en [bullinaria2009] permiten notar un mejora en tareas de aprendizaje incremental, teniendo menos perdida de información en comparación de implementaciones anteriores como el algoritmo Learn++ [li2008, Elwell2011].

Así, el presente trabajo de investigación esta enfocado en poder explorar nuevas configuraciones de pesos duplicados para poder extender el trabajo previamente presentado en [bullinaria2009].

## 2 Planeamiento

Las Redes Neuronales Artificiales tienen la habilidad de poderse aprender una tarea, y poder hacer o resolver tareas de predicción o clasificación básicamente. En este sentido, con algoritmos de aprendizaje como el Backpropagation, permite ajustar los pesos de una RNA para que esta pueda empezar a resolver una tarea particular. No obstante, muchas de las tareas que se resuelven en la vida diaria, van generando mas información con el tiempo, por ejemplo, el comportamiento de una serie financiera, o bien la predicción del clima en una determinada región. Así, se tiene el aprendizaje incremental, siendo un método poco explorado, enfocado en poder aprender nueva información del problema, sin tener que volver a entrenar todo el modelo con la información anterior, y la nueva que acaba de llegar, esto es, en los modelos actuales de aprendizaje máquina, si se usa un conjunto de datos para entrenar un modelo en específico, dicho modelo es funcional para dicho conjunto de datos y la información que ello representa. Sin embargo, si es necesario incorporar nueva información al modelo, es necesario recolectar dicha información nueva, agregarla a la que se tenía anteriormente, y volver a entrenar todo el modelo para que éste pueda incorporar la nueva información.

De esta forma, una red ya entrenada con un primer conjunto de datos  $d_1$  se desea entrenar con un nuevo conjunto de datos del mismo problema  $d_2$ , al entrenar la red con  $d_2$  se perderá el conocimiento aprendido por  $d_1$ . Si no se desea utilizar un modelo de aprendizaje incremental, se tendría que juntar el conjunto  $d_1$  y  $d_2$  en un solo conjunto y volver a entrenar la RNA para así poder incorporar el nuevo conocimiento ( $d_2$ ) a la RNA. Si se desea utilizar el método de aprendizaje incremental, se puede entrenar en un primer momento a la red con el conjunto  $d_1$ , posteriormente con  $d_2$  teniendo poca pérdida de información de  $d_1$ . Si más adelante llega mas información del problema ( $d_3$ ) que se desee incorporar a la base de conocimientos de la RNA, entonces solo habrá que entrenar la RNA con  $d_3$  usando el modelo incremental para tener una pérdida mínima de información de  $d_1$  y  $d_2$ .

De esta forma, el presente trabajo tomará como base la investigación de [bullinaria2009], en donde utiliza una configuración de pesos dobles (a una RNA se duplican todos sus pesos), donde a una capa de pesos duplicados se enlaza con una tasa de aprendizaje alta, para simular un rápido aprendizaje, y por ende simular lo que sería memoria a corto plazo. En su contraparte, la segunda capa de pesos duplicados, se enlaza con una tasa de aprendizaje baja para aprender lentamente un problema, simulando la memoria a largo plazo. Es decir, al momento de aprender una tarea nueva, una capa de pesos aprenderá muy rápido la nueva tarea (tasa alta de aprendizaje) y por ende olvidará más rápidamente los datos ya aprendidos anteriormente, por otro lado, la segunda capa de pesos duplicados, aprenderá muy lentamente los nuevos datos que llegan, y por ende olvidando poco la información que anteriormente se aprendió. Considerando ello y también que la RNA trabaja en conjunto con ambas capas de pesos duplicados, se pondera la salida de la RNA para que pueda contemplar la información nueva que se acaba de agregar a la RNA, en ambas capas, así como la información que se acaba de olvidar de ambas capas de pesos.

El problema principal del aprendizaje incremental mostrado en [bullinaria2009], es que entre más conjuntos de datos nuevos que lleguen, mas se olvidaran los

primeros conjuntos que se aprendieron, lo cual no es tan útil si se contempla que en el futuro de una RNA podrán existir 10 o 20 etapas de entrenamiento incremental con nuevos conjuntos de datos que se vayan recolectando.

Por ello es indispensable poder explorar nuevas configuraciones de RNAs que permitan mejorar los métodos actuales para permitir una menor cantidad de olvido conforme llegue nueva información al modelo. Donde al igual que en trabajos anteriores, el presente trabajo se basará en conceptos de memoria a corto y largo plazo, y en lugar de hacer una copia de los pesos actuales y tener dos tasas de aprendizaje, una rápida para simular la memoria a corto plazo, y otra tasa de aprendizaje para simular la memoria a largo plazo, se explorará por hacer mas copias de los pesos, teniendo más tasas de aprendizaje que operen en cada una de dichas copias.

### 3 Objetivos

Diseñar una red neuronal artificial para aprendizaje incremental basada en el principio de la memoria a corto y largo plazo, buscando usar más de dos capas de pesos duplicados para el reconocimiento de dígitos, y con una menor pérdida de información de trabajos previos.

#### 3.1 Objetivos Particulares

1. Implementar el algoritmo mostrado en [bullinaria2009] para el reconocimiento de dígitos con aprendizaje a corto y largo plazo con los parámetros que ahí se indican.
2. Obtener el conjunto de datos de Optical Digits, limpiar los datos y prepararlos según lo indicado con [bullinaria2009].
3. Separar el conjunto de entrenamiento y de prueba de acuerdo a lo que se explica en el artículo de Bullinaria y probar el primer código implementado en miras de comprobar los resultados previamente mostrados en [bullinaria2009].
4. Tomar como base el algoritmo implementado, y extenderlo para permitir mas de dos pesos duplicados, aplicando el conjunto de datos previamente mostrado.
5. Comparar ambas implementaciones en busca de una reducción significativa de las tasas de aprendizaje con respecto a trabajos previos en la literatura.

### 4 Justificación

Las redes neuronales permiten el aprendizaje automático y la resolución de distintos problemas, pero como se comentó anteriormente, las técnicas de aprendizaje máquina, tienen una deficiencia que es al momento de aumentar los nuevos bloques de datos que llegan para aprender, se obtiene un deterioro en el rendimiento de aprendizaje de información y olvido de la información anterior [bullinaria2009]. (REFIVEN COMO VOY PONEINDO LA REDACCIÓN

DE LO QNTERIOR, Y CO NELLO TRATEN DE CAMBIAR LO QUE VIENE PARA QUE LE MOJOREN EL ESTILO, DE AQUÍ EN ADELANTE SOLO LES HARE OBSERVACIONES EN ESPERA QUE USTEDES MEJOREN EN GENERAR LA REDACCIÓN)

Los resultados que se han obtenido no funcionan a la perfección, la memoria a corto plazo olvida poco pero va olvidando, y lo ideal sería que no olvidara. Biológicamente los humanos pueden aprender nuevas tareas, o información nueva de un problema, y no olvida de forma significativa lo que anteriormente aprendió, no obstante eso no pasa actualmente con las RNA y en general con cualquier algoritmo de aprendizaje máquina. En otro sentido, los humanos ya tienen cierta configuración en el cerebro que les permite aprender como se hace actualmente, y se puede afirmar que por el momento no hay ningún procedimiento (quirúrgico o no) que permita modificar la estructura del cerebro para aprender mas y olvidar menos.

No obstante, computacionalmente nada puede impedir que se experimente con más configuraciones y llegar al punto en donde toda la información que ingrese a un modelo (por ejemplo RNAs) se acumule, y si no hay problema de almacenamiento que se siga acumulando y que no olvide, esto podría ser bueno en diversas situaciones.

Desde el punto de vista computacional, si llega nueva información y no se ocupa aprendizaje incremental, ello implicará volver a entrenar todo el sistema con la información anterior y la actual (por ejemplo  $d_1$  y  $d_2$ ) y considerando que una de las desventajas que tienen la RNAs es que el entrenamiento es un cuello de botella, siendo este donde se lleva la mayor parte de cómputo y por consiguiente de energía. Ello implica que volver a entrenar con toda la información acumulada, gastará más energía y tiempo, que si solo se entrena con la nueva información que llega al modelo. En su contraparte, existe una gran variedad de herramientas las cuales permiten codificar una red neuronal artificial con librerías ya preexistentes, por el momento se expondrán solo 2 empresas, siendo estas las más importantes: Microsoft y Google. La primera cuenta con la plataforma de Azure que renta una maquina virtual donde se puede programar en Python. Por el contrario, Google cuenta con Google Colab que igual brinda una máquina virtual para realizar experimentos de Maching Learning, la única diferencia contra Azure es que dicha herramienta es gratuita, una similitud que tienen es que en las dos se puede programar en el mismo lenguaje.

Como se observa, ambas herramientas permiten la programación en Pyhton y esto se debe a que dicho lenguaje es una herramienta de software libre que no requiere licencia, es relativamente fácil poder depurar un código y permite acelerar más el desarrollo de aplicaciones, a diferencia de otros lenguajes más estructurados como C o Java, además tiene más librerías para el desarrollo de Maching Learning, por ejemplo, TensorFlow, Numpi, entre otras.

TensorFlow es una librería de Python que permite construir y entrenar redes neuronales para detectar patrones y razonamientos usados por los humanos, en la presente investigación se usará dado a que favorece la creación de una RNA, permite la elaboración de cualquier tipo de algoritmo de Machine Learning, cabe mencionar que también se puede usar para Deep Learning, facilita la adquisición de datos modelos de capacitación, predicciones y refinamiento de resultados, está disponible para el uso en computadores personales, pero es recomendado usarlo

en su propio editor en la nube que es Colab.

Keras (MISMO COMENTARIO QUE LES PONGO DE TENSORFLOW) es un framework de alto nivel para el aprendizaje, escrito en Python y capaz de correr sobre los frameworks TensorFlow. Fue desarrollado con el objetivo de facilitar un proceso de experimentación rápida. Diseñado para construir por bloques la arquitectura de cada red neuronal, incluyendo redes convolucionales y modelos recurrentes, que son las que permiten, junto a los bloques "más tradicionales", entrenar aprendizaje profundo.

## 5 Delimitación

En la siguiente investigación solamente se van a utilizar redes neuronales artificiales, cabe mencionar que este no es el único tipo de red, porque también se tiene lo que son Redes Neuronales Monocapa, Redes Neuronales Perceptrón Multicapa (RNPM), Redes Neuronales Convolucionales (RNC), Redes Neuronales Recurrentes (RNR), Redes de Base Radial (RBR) [royo2021]. Pero para este experimento se van a utilizar RNAs que es el algoritmo que por el momento beneficiará a la presente investigación, cabe mencionar que no se usarán algoritmos genéticos, ya que si se implementan, se estará optimizando y el objetivo principal es utilizar el aprendizaje incremental para que acepte más datos de entrenamiento.

(Esto es lo que lo que el puso en el marco teorico, y tambien deje su comentario ahí para que te ubiques va) Así como las redes convolucionales, también existen las redes profundas(deep learning), sin embargo, como se comenta en esta sección el presente trabajo solo se basará en redes del tipo multicapa perceptron usando el algoritmo de bacpkpropagation, donde de ser demostrados los principios descritos anteriormente, se podrá aplicar a cualquier otro tipo de algoritmo de aprendizaje.

## 6 Consecuencias

Si el experimento funciona a la perfección ocurrirá lo siguiente:

1. Habrá menos olvido.
2. El aprendizaje tomará menos tiempo.

Como se menciona antes, se van a poder ingresar más datos sin tener que volver a codificar el modelo.

Los modelos de predicción van a ser más precisos, porque se podrán ingresar datos mensualmente o hasta semanalmente, esto provocará que el proyecto esté trabajando con datos actuales.

Pueden funcionar para proyectos y predicciones tan simples, por ejemplo, la preiddción climatológica, predicciones de la bolsa de valores, predicción del Bitcoin, entre otras.

## 7 Marco Teórico

### 7.1 Redes Neuronales Artificiales

Las redes neuronales artificiales (RNA) son procesos los cuales contienen simples unidades de procesamientos.

Como se puede observar, al mencionar RNA lo primero que se viene a la mente es el procesamiento biológico por el que transita el cerebro humano. El método principal de las redes neuronales es sacarle el máximo poder a los algoritmos de aprendizaje maquina, ya que las redes neuronales tienen un antecedente biológico.

Presentan demasiadas utilidades las cuales ayudan a resolver problemas como los siguientes [liu2015]: no linealidad, mapeo entrada-salida, aprendizaje robusto a errores en los datos de entrenamiento, entre otros.

Existen varios tipos de Redes Neuronales tales como: Redes Neuronales de Perceptrón Multicapa, Redes Neuronales Convolucionales, entre otras.

### 7.2 Redes Neuronales de Perceptrón Multicapa

(REVISAR REDACCIÓN Y MEJORARLA) Mapear el progreso de una red sirve para conocer y buscar un camino adecuado entre la capa de entrada y salida, esto se puede hacer con problemas de una sola neurona, no obstante, los problemas no lineales no se pueden resolver, para ello se usan los perceptrones multicapas, los cuales rompen con esta limitación.

Estas neuronas están compuestas de 3 capas como se muestra en la Figura 1. Los datos de entrada se van a ir propagando capa por capa, hasta llegar a la capa final, donde la información pasará por una función de activación la cual permite tener en un rango controlado los valores que se propagan por la red. Al final, las neuronas de la capa de salida darán el resultado esperado o muy próximo a él, si el valor no es correcto, se deberá proceder a una etapa de entrenamiento [liu2015].



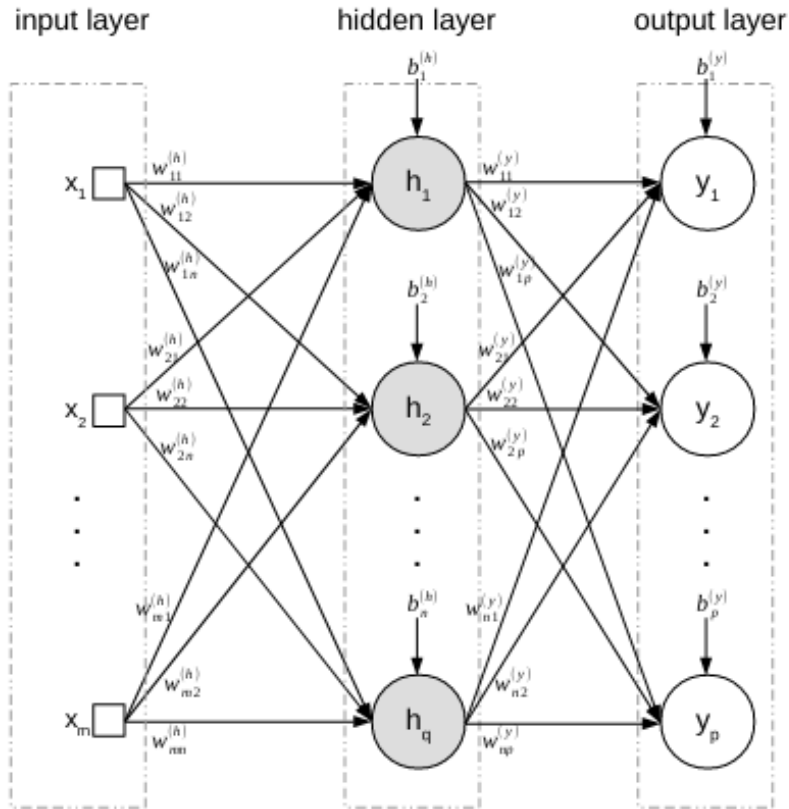


Figure 1: Red Neuronal de Perceptrón Multicapa [liu2015]

### 7.3 Algoritmo Backpropagation

(Aquí describan este algoritmo, la página que les pase de John les puede ayudar. Términos importantes de escribirlo porque sobre él se basaron para hacer la implementación de las redes con pesos duplicados)

### 7.4 Redes Neuronales Convolucionales

Este tipo de red trabaja con el uso de imágenes, por lo general de alta calidad, el único problema que se tiene al momento de que sean de alta resoluciones son que el tiempo de entrenamiento sea enorme y el tiempo de testo (Tengo duda de que la palabra testo quede, puedes checarlo? si no pues dejalo asi jsjs)sea muy tardío.

Consta de diversas multicapas alternadas, al final tiene una red perceptrón multicapa. (REVISEN LA REDACCIÓN, ESTÁ MUY SEGMENTADA LAS IDEAS, TIENEN QUE EXPLICAR UN POCO MÁS PARA QUE SE ENTIENDA MEJOR, TOMEN COMO BASE LA REDACCIÓN QUE LES PONGO AL INICIO DEL TRABAJO, VEAN CÓMO VOY A EXPLICANDO MÁS

COSAS, Y AUNQUE ME LLEVO MÁS ESPACIO, QUEDA MEJOR) La entrada de una red convolucional, con diferentes medidas en altura y anchura de imagen, para el uso de los proyectos se trabajan en escalas de grises, las cuales contienen filtros y cada filtro tiene distintos rasgos y características de tamaño. Cada capa es submuestreo de mínimo a máximo, muestra donde se toman valores desde 2 imágenes pequeñas hasta no mas de 5 imágenes grandes.

Antes o después del submuestreo se aplica la activación sigmoideal para cada mapeo de rasgos [duran2017].

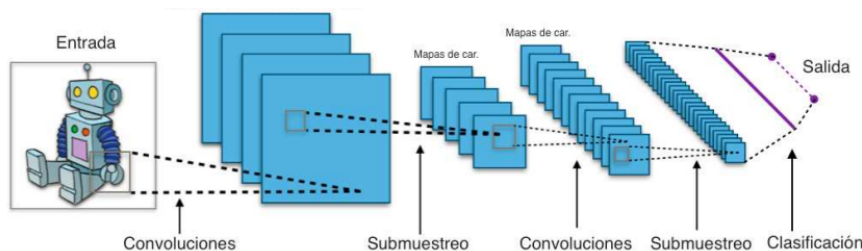


Figure 2: Esquema de una Red Convolucional [duran2017]

Así como las redes convolucionales, también existen las redes profundas(deep learning), sin embargo, como se comenta en esta sección el presente trabajo solo se basará en redes del tipo multicapa perceptron usando el algoritmo de bacpkpropagation, donde de ser demostrados los principios descritos, se podrá aplicar a cualquier otro tipo de algoritmo de aprendizaje. (LUIS Y SANDRA, AGREGUEN A LA DELIMITACIÓN LO QUE ESTOY INDICANDO AQUÍ, PARA QUE TENGA SENTIDO ESTO)

## 7.5 Aprendizaje

### 7.5.1 Aprendizaje en humanos

El humano tiene una forma de aprendizaje muy particular, la cual se basa del estudio, donde lee, escribe y practica acerca de su tema de interés, pero dicho aprendizaje se puede ir olvidando, esta es una acción muy común que a cualquier persona le sucede. Existen estudios donde se comenta que existen tres motivos del porque se olvidan las cosas, proviene parte de la regularización de las emociones, el como se adquirieron los conocimientos, y porque el olvido es un proceso por el cual el ser humano transita a lo largo de su vida [Nrby2015]. Pero cabe mencionar que esto no es lo único que causa la perdida de memoria, ya que existe la déficit de memoria.

### 7.5.2 Aprendizaje Humano

Al momento de hablar del aprendizaje humano, se debe de hablar de la ciencia cognitiva, que es quien se encarga de descubrir esta incógnita, esta ciencia lo estudia de un modo multidisciplinario, el cual abarca las áreas de [bransford2000]: antropología, lingüística, filosofía, sicología del desarrollo, ciencia de la computación, neurociencia. Con el método de esta ciencia se pueden descubrir dos tipos de aprendizaje que son: el aprendizaje con compresión y el aprendizaje Activo.

### 7.5.3 Aprendizaje con Compresión

La comprensión es una actividad la cual se ha generado al momento de realizar cualquier tipo de lectura.

Teniendo un enfocamiento en el ámbito estudiantil, ya que es donde más se maneja esta táctica, esto es una practica algo compleja, sistemática y organizada, pues da el significado de la literatura, gracias a esto se puede obtener el contexto de la literatura. (Insisto que lo anterior esta mal dicho, pero no se me ocurre de que manera expresarlo)

Al conocer esto se puede decir con seguridad que para cualquier tipo de aprendizaje la comprensión es una parte primordial [perez2014]. (Tambien siento que se debe de agregar un concepto en si o como dice Landassuri mas explicado)

### 7.5.4 Aprendizaje Activo

El aprendizaje de la forma en la que se conoce no es del todo efectiva, ya que el sistema educativo no se basa en el principio de *belongingness*, el cual esta asociado al estimulo con su respuesta, y esto es lo más importante para que el ser humano pueda aprender cualquier cosa.

Este tipo de aprendizaje se basa en la recepción de conocimientos y la práctica donde se ponen en marcha los conocimientos adquiridos.

Otro concepto importante aquí es la tautología doble (*selbsttätiges Lernen*), que en palabras informales es convertirse en autodidacta, se puede observar que esto pertenece a dicho aprendizaje, porque usa el principio mencionado anteriormente [Huber2008].

### 7.5.5 Aprendizaje Incremental

Con el pasar de los años la tecnología a evolucionado, eso quiere decir que el Aprendizaje Automático se ha actualizado y que la cantidad de datos va aumentado con más frecuencia.

Se puede verificar como "*Una tarea de aprendizaje es incremental si los ejemplos de entrenamiento usados para resolverla están disponibles en horas extras, generalmente uno a la vez*" [GiraudCarrier2000], si los resultados no se necesitan de manera urgente, este tipo de trabajos serán resueltos por algoritmos de aprendizaje no incremental.

Una área donde esto es de mucha utilidad es la *Robótica* porque este necesita estar en constante entrenamiento [GiraudCarrier2000].

Dicha forma de aprender fue inspirada en la forma en que el humano aprende y esta más rápida, fue por esto que fue adoptada por el aprendizaje máquina.

Con el paso del tiempo se ha convertido en un paradigma del aprendizaje automático, aquí el aprendizaje toma el lugar de nuevos ejemplos para juntarlos y conforme van aprendiendo estos toman el lugar de los ejemplos ya aprendidos [liu2015].

### 7.5.6 Algoritmos de Aprendizaje Incremental

"Un algoritmo de aprendizaje es incremental si, para cualquier muestra de entrenamiento dada:

$$e_1, \dots, e_s \tag{1}$$

produce una secuencia de hipótesis

$$h_0, h_1, \dots, h_n \quad (2)$$

tal que  $h_{i+1}$  depende solo de  $h_i$  (Tengo duda con la palabra  $h_i$  si lo tradujiste fijate bien que hayas copiado bien, luego los símbolos no los copia como es ) y del ejemplo actual  $e$ ” [GiraudCarrier2000], como se observa, estos son algoritmos que permiten a la inteligencia artificial poder realizar actividades de predicción de una manera más eficaz.

Un ejemplo del uso de esta rama es el proyecto *COBWEB*, donde se trata de categorizar el número de Clúster y la pertenencia de dichas categorías por medio de una métrica probabilística global, esto lo realiza por medio de que se agrega una nueva categoría, este proceso lo que realizará es actualizar todas las probabilísticas con los nuevos datos recabados [fisher1987].

AQUÍ FALTA DESCRIBIR A DETALLE CÓMO SE LLEVA ACABO EL APRENDIZAJE INCREMENTAL, SE PUEDEN BASAR EN EL TRABAJO DE BULLINARIA Y LEARN ++, EXPLIQUENLO A DETALLE PARA QUE SEA SU ESTADO DEL ARTE

## 8 Hipotesis

Al tener más de dos capas de pesos duplicados con sus respectivas tasas de aprendizaje, permite tener un menor olvido de la información previa aprendida, en un modelo de aprendizaje incremental con redes neuronales artificiales del tipo MLP, al usar el algoritmo de Backpropagation.

## 9 Metodología

El primer paso a realizar en esta investigación es recrear el código mostrado en [bullinaria2009], donde se describe la implementación de una red neuronal multicapa usando el algoritmo de entrenamiento backpropagation en el lenguaje de programación Python. Para esto se utilizará el conjunto de datos de Optical Digits, en donde se tendrá que preprocesar los datos, para eliminar registros inválidos.

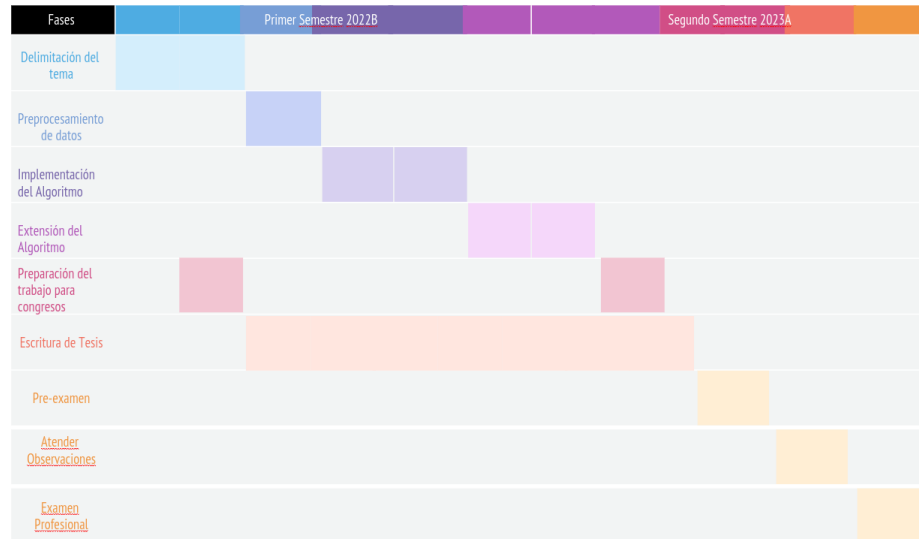
Posteriormente se implementará una nueva versión del código en donde se experimentará con mas de dos capas de pesos duplicados para mejorar la tasa de olvido de información al momento de usar el aprendizaje incremental. Para ello se explorará incrementando gradualmente el nueron de capas (esto no lo entendi,neuronas de capas?)de pesos duplicados, hasta llegar al punto en que mas capas no generen un decremento de las tasas de olvido/error.

(REVISEN LA REDACCIÓN Y MEJORENLA DE FAVOR) Cuando los dos proyectos se tengan, se realizará una comparación, donde se verá cual de estos dos experimentos es más eficaz en proyectos de la vida real.

## 10 Cronograma de Actividades

(EL CRONOGRAMA DE ACTIVIDADES DICE JOHN BULLINARIA EN LA PARTE DE ARRIBA, ELIMINAR ESO)

## Algoritmo de John Bullinaria



## 11 Organización del Capitulado

En el capítulo 2 se verá lo que es el aprendizaje humano y el aprendizaje incremental con sus algoritmos, se describirán las redes neuronales artificiales.

En el capítulo 3 se implementará el artículo de John A. Bullinaria, como funciona, resultado que da al pasar los datos que dice para comprobar que funciona como menciona en su artículo. En el capítulo 4 se explicará como se hará la modificación a su algoritmo, cuantas capas se van a poner, como se van a repartir las tazas de aprendizaje.

Posteriormente en el capítulo 5 se mostrará una comparación de los resultados de ambos trabajos. En el capítulo 6 se verán las conclusiones y trabajo futuro.