

Entrenamiento del Perceptrón

Perceptron Training

Autor 1: David Cediél

Facultad de Ingenierías, Universidad Tecnológica de Pereira, Pereira, Colombia

Correo-e: David.cediél@utp.edu.co

Resumen— En este artículo se revisará que es un perceptrón, cuál fue su impacto en las ciencias de la computación, específicamente en la inteligencia artificial y cómo es el proceso de entrenamiento de este, examinando las variables y valores claves para su funcionamiento.

Palabras clave—Perceptrón, entrenamiento, aprendizaje, convergencia, error, clasificación, umbral.

Abstract— This article will review what is a perceptron, what was its impact on computer science, specifically on artificial intelligence and how is the training process of this, examining the variables and key values for its operation

Key Word —Perceptron, training, learning, convergence, error, classification, threshold.

I. INTRODUCCIÓN

El auge de la inteligencia artificial de la última década ha dado pie a desarrollos tecnológicos que eran inimaginables hace unos años, desde procesamiento del lenguaje natural hasta vehículos autónomos. Una gran tendencia en la inteligencia artificial es el uso de redes neuronales, donde, el núcleo de estas es el perceptrón.

El perceptrón es la unidad fundamental de una red neuronal, es la representación de una neurona.

Primero se hablará de la historia del perceptrón, luego se abordarán los valores y variables necesarias para su funcionamiento, y, por último, se revisará las relaciones entre estas variables para llevar a cabo el entrenamiento.

II. CONTENIDO

A. HISTORIA

El perceptrón fue creado por Frank Rosenblatt, un psicólogo estadounidense entre 1950 y 1960, el cual introdujo la noción de los pesos para las variables, el cual es un valor real que representa la importancia de cada una de las entradas.

Los principales usos del perceptrón fueron para crear funciones lógicas, pero tenía la limitación de no poder resolver el problema de una compuerta lógica XOR.

Luego, en 1965, se confeccionó el término de perceptrón multicapa, el cual no se limitaba a una sola neurona, sino que eran varias neuronas interconectadas entre si dispuestas en varias capas.

Luego de este año se produjeron muchos otros avances en las redes neuronales, tales como: neuronas sigmoides, backpropagation, redes neuronales convolucionales, entre otros.

B. VARIABLES Y VALORES DEL PERCEPTRÓN

El perceptrón tiene ciertos valores importantes para su funcionamiento.

Las entradas son la información que recibe el perceptrón, pueden ser vistas como las características del modelo que se está estudiando

Los pesos son valores numéricos que se le asignan a cada variable, los cuales representan la importancia que tiene una entrada con respecto a la salida

La función de activación determina el valor de la salida a partir de la sumatoria de cada variable multiplicada por su peso, esta función tiene un valor de umbral que determina si la salida será cero o uno.

C. ENTRENAMIENTO DEL PERCEPTRÓN

Para el entrenamiento se seleccionan las variables de entrada que tendrá el perceptrón, luego se seleccionan unos pesos de manera aleatoria para empezar el proceso. A continuación, se realiza la suma ponderada de la multiplicación de cada una de las variables por su peso.

Para ajustar los valores de los pesos que nos lleven a un resultado correcto, se usa una técnica llamada convergencia del perceptrón.

En la convergencia del perceptrón, se escribe la suma ponderada en forma de ecuación matricial.

$$W^T X = 0 \quad (1)$$

Donde W son los pesos y X son los valores de entrada, esta ecuación representa una superficie de decisión entre dos clases de entrada, donde la dimensión de la superficie está definida por el número de variables de entrada.

Se tienen dos tipos de entrada X_1 y X_2 , las cuales hacen parte de la clase C_1 y C_2 respectivamente [1]. El proceso de entrenamiento es ajustar los valores de los pesos dado que al ingresar una entrada X_1 , esta sea clasificada como perteneciente a la clase C_1 , para esto es necesario ajustar los valores de W .

Para determinar si una variable pertenece a la entrada C_1 o C_2 , se debe procurar que

$$\begin{aligned} W^T X &\geq 0 \text{ para una entrada que pertenezca a } C_1, \text{ y} \\ W^T X &\leq 0 \text{ para una entrada que pertenezca a } C_2 \end{aligned} \quad (2)$$

Las ecuaciones anteriores pueden reescribirse en términos de magnitud y dirección, quedando de la forma

$$\begin{aligned} \|W^T\| \|X\| \cos(\alpha) &> 0 \text{ para una entrada que pertenezca a } C_1, \text{ y} \\ \|W^T\| \|X\| \cos(\alpha) &\leq 0 \text{ para una entrada que pertenezca a } C_2 \end{aligned} \quad (3)$$

Donde, el factor que hará que el valor sea mayor o menor a cero, es el ángulo entre los dos vectores, por lo que los otros factores no son necesarios, por lo tanto

$$\begin{aligned} \cos(\alpha) &> 0 \text{ para una entrada que pertenezca a } C_1, \text{ y} \\ \cos(\alpha) &\leq 0 \text{ para una entrada que pertenezca a } C_2 \end{aligned} \quad (4)$$

Así, se puede hacer una comparación de los valores de entrada dados con su clasificación, si X es clasificado de manera correcta, el vector de pesos se mantiene igual, si no, el vector de pesos se actualiza de la siguiente manera

$$\begin{aligned} W_{n+1} &= W_n + X \text{ si } X \text{ pertenece a } C_1 \\ W_{n+1} &= W_n - X \text{ si } X \text{ pertenece a } C_2 \end{aligned} \quad (5)$$

Al realizar este proceso de forma iterativa, llegará el momento en el que no se le deba hacer ninguna nueva modificación al vector W , por lo cual el hiperplano habrá separado los dos espacios; la manera en que W converge está dada por el teorema de convergencia del perceptrón.

III. CONCLUSIONES

El perceptrón fue el punto de partida para todas las opciones de redes neuronales que tenemos en la actualidad, y entender qué es y la forma en que se entrena es fundamental para comprender las redes neuronales de la actualidad.

IV. REFERENCIAS

- [1]. <http://bibing.us.es/proyectos/abreproy/11084/fichero/Memoria+por+cap%C3%ADtulos+%252FCap%C3%ADtulo+4.pdf>
- [2]. http://www.itnuevolaredo.edu.mx/takeyas/Apuntes/Inteligencia%20Artificial/Apuntes/tareas_alumnos/RNA/Redes%20Neuronales2.pdf
- [3]. <https://platzi.com/tutoriales/1157-ia-2017/2619-entrenamiento-del-perceptron/>
- [4]. <https://es.wikipedia.org/wiki/Perceptr%C3%B3n>
- [5]. <ftp://decsai.ugr.es/pub/usuarios/castro/Actividades/Redes-Neuronales/Apuntes/Apuntes%20Javier%20Rodriguez%20Blazquez/Redes%20de%20una%20capa.pdf>