ESCUELA SUPERIOR POLITÉCNICA DEL LITORAL

FACULTAD DE INGENIERÌA EN ELECTRICIDAD

Y COMPUTACIÓN

RESUMEN

TESIS "CLASIFICACIÒN DE DATOS Y MANEJO DE INFORMACIÓN USANDO UN MODELO DE RED NEURONAL"

Trabajo previo a la obtención del tìtulo de:

Ingeniero en Electricidad

Especialización Computación

Autor:

JORGE CABELLO

Guayaquil – Ecuador

2000

RESUMEN

La Escuela Superior Politécnica del Litoral, en su afán de establecer un mecanismo de pago de matrículas, que permita asignar valores más justos de acuerdo al nivel económico de los estudiantes, ha implantado una fórmula que contempla el promedio de calificaciones y el nivel socioeconómico del estudiante.

El nivel socioeconómico que se refleja en un valor o factor que permite a la Universidad evaluar los ingresos económicos del estudiante y su familia y el entorno social en el que se desenvuelve, se denomina el Factor P.

Este Factor P pondera a los estudiantes con valores que van de 3 para los estudiantes de nivel socioeconómico más bajo a 40 para los estudiantes de nivel socioeconómico más alto. Cada uno de los cinco criterios impactan de cierta manera al valor de este factor y estos valores asignados inciden directamente sobre el valor que el estudiante paga a la Universidad por concepto de matrícula para un determinado semestre. ¿Pero realmente este valor asignado es el adecuado?.

Esta tesis presenta un estudio que permite establecer el impacto real de cada uno de los criterios que formula el factor P y entrega una herramienta basada en una **red neuronal** que permite asignar de diferente maneras estos criterios y ver la incidencia sobre el factor P. Por ejemplo, si un estudiante recibe un factor P de 35, podemos evaluar cual es el impacto real sobre este valor por concepto de ingreso económico del grupo familiar del estudiante, institución de estudios primarios, ubicación domiciliaria, etc.

Todos sabemos que el cerebro humano supera en muchas tareas a los sistemas de computo digital, tareas tal como el reconocimiento de patrones. Una persona reconoce un patrón como una cara mucho más rápido y de manera sencilla que un computador, solo en las tareas matemáticas un ordenador supera a un cerebro humano.

Las redes neuronales intentan simular el procesamiento de un cerebro humano para aplicaciones donde la inteligencia humana no pueda ser emulada de forma satisfactoria por algoritmos aritméticos que pueden ser implementados en ordenadores. Estas redes neuronales deben cumplir características muy similares a las funciones del cerebro:

- Robustez: Mantener su funcionamiento de forma satisfactoria.
- Flexibilidad: Debe adaptarse a las nuevas circunstancias o situaciones.
- Inteligencia: Mantener su conocimiento mediante un autoaprendizaje.
- Consistencia: Debe lograr trabajar con información borrosa, incompleta, probabilística, con ruido o inconsistente.
- Integración: Debe mantener un comportamiento colectivo mediante neuronas interconectadas.

Las Redes Neuronales Artificiales son aplicadas a un gran número de problemas reales de considerable complejidad, tales como reconocimiento de patrones, clasificación de datos, predicciones, etc. Su principal orientación es la solución de problemas que son demasiado complejos para las técnicas convencionales: problemas que no tienen un algoritmo especifico para su solución, o cuyo algoritmo es demasiado complejo para ser encontrado.

Existe actualmente una tendencia a establecer un nuevo campo de las ciencias de la computación que integraría los diferentes métodos de resolución de problemas que no pueden ser descritos fácilmente mediante un enfoque algorítmico tradicional. Estos métodos, de una forma u otra, tienen su origen en la emulación, más o menos inteligente, del comportamiento de los sistemas biológicos.

Con las Redes Neuronales se intentará expresar la solución de problemas complejos, no como una secuencia de pasos, sino como la evolución de unos sistemas de computación, inspirados en el funcionamiento del cerebro humano y dotados, por tanto, de una cierta "inteligencia"; los cuales, no son sino la combinación de una gran cantidad de elementos simples de proceso (neuronas) interconectados que, operando de forma masivamente y paralela, consiguen resolver problemas relacionados con el reconocimiento de formas o patrones, predicción, codificación, clasificación, control y optimización.

El concepto de una red neuronal no es nuevo. El objetivo de crear una red neuronal fue originalmente concebida como un intento por modelar la fisiología del cerebro. El objetivo era crear un modelo capaz de realizar procesos del pensamiento humano, inclusive muchos de los primeros trabajos fueron realizados por fisiólogos y psicólogos.

Nuestro Universo entrega una diversidad de fuentes de entradas a nuestro cerebro, el cual las procesa de tal manera que se crean respuestas a estos estímulos. De forma muy organizada y sistemática, nuestro cerebro entreteje todos los estímulos posibles y los relaciona creando un sistema neuronal de información conformado por billones de neuronas, las cuales recepta la información externa y se encargan de ordenar al resto de nuestro cuerpo a realizar acciones.

Han sido muchos los intentos del hombre por comprender y construir una máquina que se asemeje a este sistema neuronal y realice tareas inteligentes.

Entonces, una red neuronal es un grupo de elementos interconectados, unidades o nodos, los cuales funcionalmente están basados en una neurona animal. En otras palabras, una Red Neuronal artificial esta compuesta de un grupo de elementos de procesamiento altamente interconectados (Neuronas) trabajando al mismo tiempo para la solución de procesos específicos. Es importante recalcar que las redes neuronales a igual que los humanos aprenden en base a la experiencia y entrenamiento. De esta combinación de conceptos nació una nueva forma de computación que trata de resolver problemas reales con soluciones automatizadas y fáciles de implementar.

Una red neuronal es capaz de aprender. Las redes neuronales no ejecutan instrucciones secuenciales; así tampoco contienen memoria de almacenamiento, de instrucciones o de datos. En vez de esto, las redes neuronales son entrenadas, presentándoles ejemplos de entradas y salidas, los que son memorizados alterando los vectores de pesos.

Tienen la capacidad de Generalizar, son capaces de acomodar los pesos de interconexión de modo de lograr una salida correcta frente a una entrada determinada. Esto sucede aún cuando una de las neuronas esté inhabilitada o alterada (tolerancia a fallas), lo que en la computación clásica puede ser logrado sólo con algoritmos complejos y de alto costo.

Son capaces de Abstraer, siendo así un ente ideal desde un conjunto de entrenamiento no ideal, y recordar algo que no necesariamente se le haya enseñado. Son Veloces y cada neurona de la red es un procesador que opera sobre sus entradas independientemente de los otros procesadores y la convergencia ocupa a todas ellas, aunque se agreguen más procesadores. Esto contrasta con los problemas que presenta la programación paralela convencional.

Podemos concluir que el procesamiento es diferente al de una red humana, pues la combinación de las señales produce una nueva señal, en contraste a la ejecución de una instrucción guardaban en memoria.

La información se almacena una serie de pesos y no en un programa. Los pesos se adaptan a la red en un proceso de entrenamiento y aprendizaje.

No son susceptibles a ruido, pues cuando aparecen pequeños cambios en las señales de entrada no cambian drásticamente la señal de salida.

La construcción de la red neuronal comprende 6 etapas comenzando con la conceptualización del modelo a usar. Se debe conocer claramente el modelo con sus entradas y salidas. Después de la conceptualización del modelo, se debe hacer una recopilación de los datos. Los datos son procesados y formateados en una etapa de procesamiento, para que luego se cree la arquitectura del modelo. Se entrena la red con los valores mas indicados para el buen funcionamiento de la misma, y finalmente se analiza la red con valores reales.

El entrenamiento de una red es todo el proceso de aprendizaje que realiza una red neuronal. Su objetivo es lograr que la aplicación de un conjunto de entradas produzca el conjunto de salidas deseado (o uno al menos consistente).

Existen una gran variedad de tipo de modelos neuronales, los cuales se pueden clasificar dependiendo de su estructura, de su manera de aprender u operar y de su aplicación.

Podemos tener tres tipos de clasificaciones:

1. Clasificación por estructura.
2. Clasificación por aprendizaje.
3. Clasificación por aplicación.

En cuanto a la **clasificación por estructura**, nos concentramos en los componentes de la red, tales como, las capas, sus neuronas y conexión y tipos de funciones de activación. Entonces, encontraremos diversos modelos que se clasifican de acuerdo al número de sus componentes. Por ejemplo, si hablamos de redes neuronales de una capa, se las conoce como redes monocapas; y si hablamos de redes que poseen capas ocultas, hablamos de redes multicapas.

En cuanto a la **clasificación por aprendizaje**, la parte más importante de la modulación de la red neuronales determinar formas para ajustar los pesos de forma que la red realice un buen mapeo entre entradas y salidas: a esto se le llama "entrenamiento" de la red.

Por último, su **clasificación por aplicación** puede ser enfocada enfocada hacia la función que le asigne el diseñador a la red (predicción, clasificación, asociación, filtraje, conceptualización y optimización).

En general, debido a que son parecidas al cerebro humano, las redes neuronales son bien nombradas ya que son buenas para resolver problemas que el humano puede resolver pero las computadoras no. Estos problemas incluyen el reconocimiento de patrones y la predicción del tiempo. De cualquier forma, el humano tiene la capacidad para el reconocimiento de patrones, pero la capacidad de las redes neuronales no se ve afectada por la fatiga, condiciones de trabajo, estado emocional y compensaciones.

Un área en la que las redes neuronales artificiales han tenido éxito ha sido el reconocimiento óptico de caracteres (ROC) ya sea impresos o manuscrito. El proceso de reconocimiento óptico no es trivial e implica el uso de técnicas sofisticadas de varios tipos. Por ejemplo, se requiere de técnicas de procesamiento de imágenes para convertir las escalas de gris en blanco y negro y representar estas imágenes con números binarios. Es muy probable que se requiera también de técnicas para traslapados

Dentro de las múltiples áreas de aplicación para las redes neuronales artificiales, podemos citar algunos ejemplos tales como:

Control de la eficiencia de una máquina

El comportamiento de la máquina de un carro es influenciada por un gran número de parámetros como temperatura, mezcla de combustible, viscosidad de lubricante, etc. Rolls Royce ha usado redes neuronales para ajustar dinámicamente una máquina de acuerdo a su estado actual.

Predicción de bolsa de valores

El triunfo o no en el uso real en la predicción de mercados es difícil de alcanzar desde que las compañías usando estas técnicas son comprensiblemente reacias a mostrar información. Algunos creen que el uso de redes neuronales en este ambiente es un truco de mercadeo.

Reconocimiento de firmas

Cada persona tiene una firma distinta. Aunque cada vez que firmamos algo, la firma es ligeramente distinta, la forma en general es similar y puede ser reconocida por un experto humano. Una compañía ha creado una máquina que reconoce firmas con un gran nivel de precisión. Toma en cuenta la velocidad en que se firma, además de la presión, grosor y otros factores.

Otras aplicaciones tales como reconocimiento óptico de caracteres, control de procesos industriales, aplicaciones de predicción del tiempo, decisiones sobre otorgamientos de préstamos, Análisis de inversiones, Análisis de firmas, Monitoreo y Mercadotecnia.

Esta tesis presenta una aplicación llamada **FPSys 1.0** hacia un modelo de asignación de valores (factor P), el cual se ha implementado a través de una **red neuronal** basada en un esquema de entrenamiento no supervisado que se conoce como **modelo de Kohonen.**

La red neuronal de Kohonen es una técnica ordenada de asociación que permite proyectar puntos multidimensionales a una red de dos dimensiones. Hay dos conceptos dominantes importantes en la comprensión de un modelo de Kohonen; son el aprendizaje competitivo y la autoorganización.

El aprendizaje competitivo es encontrar simplemente una neurona que sea la más similar al modelo de la entrada de información. La capacidad de autoorganización implica que la red tiene la facultad de modificar a la nueva neurona y a la vecindad de la neurona tal que sea aún más similar a ella.

Las evidencias sugieren que esta agrupación u organización neuronal está predeterminada genéticamente, pero es muy probable que este tipo de funcionalidad se origine mediante el aprendizaje. Esto sugiere, por tanto, que el cerebro podría poseer la capacidad inherente de formar mapas topológicos de las informaciones recibidas de exterior.

De hecho, esta teoría podría explicar su poder de operar con elementos semánticos: algunas áreas del cerebro simplemente podrían crear y ordenar neuronas especializadas o grupos con características de alto nivel y sus combinaciones. Se trataría, en definitiva, de construir mapas espaciales para atributos y características.

Teuvo Kohonen presentó en 1982, basándose en estas ideas, un sistema que incorpora este comportamiento de agrupación y organización. Se trataba de un modelo de red neuronal con capacidad para formar mapas de características de manera similar a como ocurre en el cerebro.

Kohonen trata de demostrar con este tipo de sistema o red que una entrada externa por si sola, suponiendo una estructura propia y una descripción funcional del comportamiento de la red, era suficiente para forzar la formación de mapas.

El modelo de Mapas topológicos SOM trata de establecer una correspondencia entre los datos de entrada y un espacio bidimensional de salida, creando mapas topológicos de dos dimensiones, de tal forma que ante datos de entrada con características comunes se deben activar neuronas situadas en zonas próximas de la capa de salida.

Esta tesis presenta, entonces, un modelo que permite asignar mapas topográficos para un sistema que posee 5 entradas y una salida. La idea fundamental es generar un conjunto de datos para la asignación del valor P de los estudiantes mediante este proceso de autoaprendizaje a través de una red de Kohonen.

Dado que la problemática del factor P involucra 5 variables como entradas a la red neuronal, se ha decidido desarrollar un paquete de software que permita visualizar una red de Kohonen con 5 entradas y convertirlas a un sistema de salida con una única variable.

La principal ventaja de desarrollar la aplicación **FPSys 1.0**, es la opción de personalizar y hacer un sistema a la medida que permita visualizar lo mejor posible los resultados de la red neuronal partiendo de un diseño basado en el problema planteado.

El factor limitante en este desarrollo, el cual también lo encontramos en las aplicaciones existentes en el mercado, es que los mapas topográficos se muestran en planos de 2 o 3 variables. En nuestro caso, nuestros mapas topográficos contienen 5 variables. Debido a esto se hace difícil esquematizar un escenario que grafique el sistema con este número de variables. Para tratar de llevar un control sobre los mapas se implementa en el sistema un registro que almacena todas las variables con sus variaciones durante la ejecución de la red.

Uno de los objetivos principales de esta tesis es conocer las aplicaciones existentes en el mercado para la solución de sistemas neuronales.

Hemos hecho un estudio de las principales herramientas existentes, dividiéndolas en paquetes de software que son de libre uso y paquetes de software con un valor comercial. Entre las más conocidas, destacando en su interface con el usuario, tenemos 2 herramientas muy utilizadas para crear un sistema neuronal: NEUROSOLUTIONS, y SPSS.

En cuanto a los resultados obtenidos, el rango del Factor P asignado por la ESPOL, es simulado por **FPsys 1.0**, el cual entrega resultados muy similares a los obtenidos por la Universidad, lo cual facilitaría hacer ciertas modificaciones en los rangos y obtener resultados muy cercanos a la realidad. Este estudio permite verificar que las asignaciones hechas para el factor P por la Universidad, son reproducidos con un sistema a través de un mecanismo sistematizado, lo cual implica su correcta asignación y ponderación para cada uno de los criterios tomados para asignar el Valor del Factor P.

Los resultados presentados por la herramienta se encuentran detallados en el capítulo de resultados de esta tesis.