Facultad de Ciencias Sociales y Jurídicas

Grado en Estadística y Empresa

**UNIVERSIDAD DE JAÉN**
*Facultad de Ciencias Sociales y Jurídicas*

Trabajo Fin de Grado

# PREDICCIÓN EN R A TRAVÉS DE REDES NEURONALES ARTIFICIALES

**Alumno: Manuel Jesús Jordán Expósito**

**Septiembre, 2017**

# Resumen

El objetivo principal a alcanzar con la elaboración de este proyecto es lograr predecir la propagación del virus del Nilo Occidental, el cual puede llegar a ser mortal causando una enfermedad mortal del sistema nervioso. Uno de los medios principales de propagación de esta enfermedad es a través de mosquitos infectados y se pretende estudiar cómo se propagan y transmiten el virus a través de la ciudad de Chicago en Estados Unidos, teniendo en cuenta distintas variables como son las climáticas. Para llevar esto a cabo se aplicará, a través de R, un modelo de Redes Neuronales Artificiales con el que se pretende predecir, bajo aprendizaje previo de dicha Red Neuronal con todos los datos de los que se dispone, cómo se propagará dicho mosquito y por ende el virus del Nilo Occidental a través de Chicago.

# Abstract

The main objective to achieve with this project is to reach a prediction of the spread of West Nile virus, which can become deadly causing a deadly disease of the nervous system. One of the main ways of propagation of this disease is through infected mosquitoes and is intended to study how the virus is spread and transmitted through the city of Chicago in the United States, taking into consideration different variables such as climate. To carry this out, a model of Artificial Neural Networks will be applied through R to predict, under previous learning of said Neural Network with all available data, how the said mosquito will propagate and thus West Nile virus through Chicago.

# Índice

# 1. INTRODUCCIÓN

Para la introducción de esta memoria vamos a comenzar explicando el cómo intentar predecir y lograr anteponerse a sucesos que pueden ocurrir se ha convertido en un objetivo realmente importante hoy en día con un campo de aplicación realmente amplio.

La predicción de sucesos para la toma de decisiones, con el fin de solucionar problemas presentes o futuros, ha sido siempre una cuestión de interés durante el transcurso de la historia. Por esto se sigue buscando mejorar la manera de manejarnos en este mundo de "incertidumbre".

El avance científico ha permitido ir evolucionando las soluciones a estos problemas y se han ido utilizando cada vez herramientas más fiables y precisas y que están en constante evolución, hasta llegar a la etapa actual en la que la tecnología se ha vuelto totalmente indispensable para casi cualquier situación a la que nos estamos enfrentando.

Desde la aplicación meteorológica, lo cual lleva utilizándose tiempo atrás como predicción de tendencias a través de modelos temporales, predicción del comportamiento de un determinado consumidor según datos personales o hábitos de compra e incluso aplicación de geomarketing a través de la aplicación de distintas técnicas con las que se llega a realizar una predicción a la hora de seleccionar el mejor emplazamiento para un determinado negocio.

Bajo esto último se va a basar este trabajo con el cual se buscará realizar una predicción con respecto a unos datos de posicionamiento teniendo en cuenta distintas variables que se expondrán más adelante con el fin de anteponerse a posibles localizaciones.

El objetivo principal y básico se va a basar en el estudio de técnicas de predicción utilizando Redes Neuronales Artificiales, y para probar dichas técnicas se va a utilizar un problema concreto con cierta complejidad para ver si dichas Redes Neuronales Artificiales son adecuadas y, por consiguiente, capaces para este tipo de predicción.

### 1.1. Datos sobre los que trabajar

Los datos de los que disponemos para llevar a cabo este trabajo han sido obtenidos de la comunidad Kaggle, que a su vez fueron publicados por *"Chicago Department of Public Health"*, buscando poder realizar el objetivo de este proyecto, predecir si el virus del Nilo Occidental está o no presente en un momento y ubicación determinada.

Cada año, desde finales de mayo hasta principios de octubre, los trabajadores de salud pública colocan trampas para mosquitos dispersos por toda la ciudad de Chicago.

Cada semana, de lunes a miércoles, estas trampas en las que se encuentran los mosquitos son recogidas y los mosquitos contenidos en ellas se analizan para determinar la presencia del virus del Nilo Occidental llevando esto a cabo antes del final de la semana. Los resultados de las pruebas incluyen el número de mosquitos, las especies de mosquitos, y si está o no el virus del Nilo Occidental está presente en la cohorte.

Los resultados de este ensayo están organizados de tal manera que cuando el número de mosquitos recogidos exceda de 50, éste se divida en otro registro (otra fila dentro del conjunto de datos), de tal manera el número de mosquitos están limitado a 50.

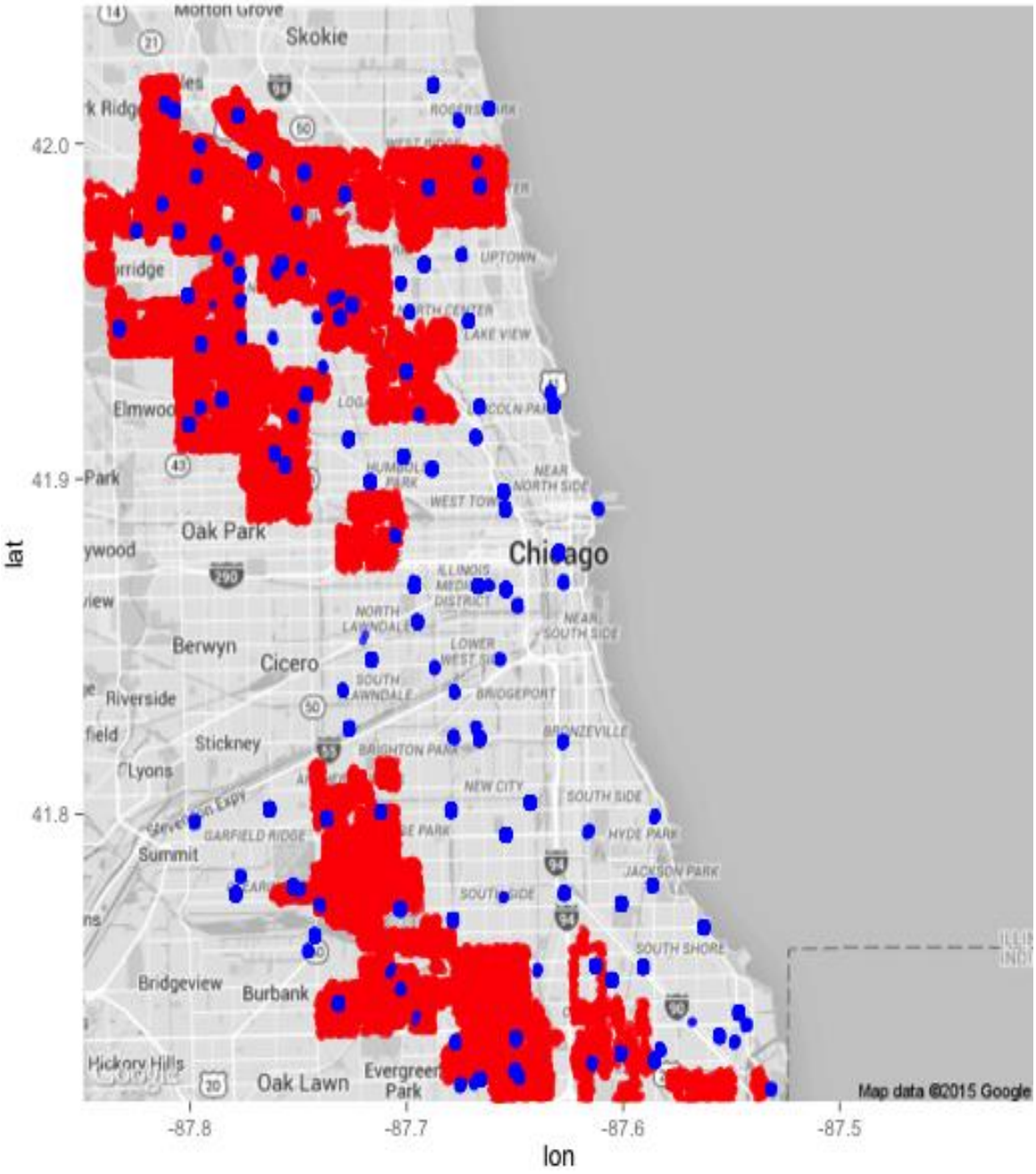
La ubicación de las trampas se describe mediante el número de bloque y el nombre de la calle. Y se han mapeado estos atributos en longitud y latitud en el conjunto de datos para organizar e identificar estos puntos geográficamente. Hay que tener en cuenta que estos son lugares aproximados.

Algunas de las trampas establecidas son trampas "satélites" las cuales se instalan para mejorar los esfuerzos de vigilancia y se establecieron con letras al final de los códigos de las trampas, por ejemplo, T220A es una trampa satélite para T220.

También hay que tener en cuenta que no todos los lugares son probados en todo momento y que existen registros en el momento en que una especie particular de los mosquitos se encuentra a una cierta trampa en un momento determinado.

La ciudad de Chicago se encuentra en una situación de dispersión de insecticida para eliminar mosquitos. La pulverización puede reducir el número de mosquitos en la zona, y por lo tanto podría eliminar la aparición del virus del Nilo Occidental.

A continuación se muestra un mapa donde se localizan las pulverizaciones de insecticidas y las trampas establecidas a lo largo de la ciudad de Chicago.




*Fuente: Chicago Department of Public Health*

Con respecto a la situación climática se estima que las condiciones cálidas y secas son más favorables para el virus del Nilo Occidental que el frío y húmedo. Tenemos a nuestra disposición datos del "National Centers for Environmental Information" donde encontramos las condiciones climáticas de 2007 a 2014, durante los meses de las pruebas.

Las distintas variables de las que vamos a disponer dentro de los ficheros de datos se enumeran en la tabla mostrada a continuación junto a una breve descripción de las mismas.

| | | | |
|---|---|---|---|
| **ID** | Identificador del registro | **Address Number And Street** | Dirección aproximada que devuelve Geocoder. |
| **Fecha** | Fecha en que se realiza el examen | **Latitud, Longitud** | Latitud y longitud que devuelve Geocoder |
| **Dirección** | Dirección aproximada de la ubicación de la trampa | **Address Accuracy** | Exactitud que devuelve Geocoder |
| **Especie** | Las especies de mosquitos | **Num Mosquitos** | Número de mosquitos atrapados en esta trampa |
| **Bloque** | Número de bloque de dirección | **WnvPresent** | Si el virus está presente. 1 es presente y 0 significa no presente |
| **Trampa** | Id de la trampa | | |

*Fuente: Elaboración propia.*

Los datos recogidos en *weather.csv* son los datos del clima de los años comprendidos entre 2007 y 2014, años durante los que se ha llevado a cabo la recogida de datos correspondiente al estudio del mosquito. Estos datos abarcan variables como

temperatura, humedad, velocidad del viento y lluvias entre otras que se muestran a continuación recogidos en una tabla.

| | | | |
|---|---|---|---|
| **Station** | Estación climática | **Depart** | Temperatura de comienzo |
| **Date** | Fecha en que se realiza el examen | **WetBulb** | Temperatura Wet-Bulb |
| **Tmax** | Temperatura máxima | **Heat** | Temperatura de comienzo a partir de Julio |
| **Tmin** | Temperatura mínima | **Cool** | Temperatura de comienzo a partir de Enero |
| **Tavg** | Temperatura media | **Sunrise** | Hora de amanecer |
| **Sunset** | Hora de puesta de sol | **AvgSpeed** | Velocidad media del viento |
| **Depth** | Profundidad de nieve en pulgadas | **Water1** | Agua procedente de lluvias y nieve derretida |
| **Snowfall** | Pulgadas de nieve caída | **PrecipTotal** | Precipitación total registrada |
| **StnPressure** | Presión atmosférica | **SeaLevel** | Nivel del mar |
| **ResultSpeed** | Velocidad resultante del viento | **ResultDir** | Dirección del viento resultante en grados |

*Fuente: Elaboración propia.*

## 2. DATA MINING

Como se ha comentado anteriormente, hoy en día las organizaciones empresariales acumulan una gran cantidad de datos a una velocidad enorme procedentes de una muy amplia variedad de fuentes, tales como transacciones de clientes, tarjetas de crédito, retiro de efectivo del banco, búsquedas por internet, datos meteorológicos entre tantos más.

El Data Mining, es un conjunto de técnicas que permiten explorar de forma automática grandes bases de datos, con la finalidad principal de encontrar patrones o tendencias con las que explicar cierto comportamiento. Para llevar a cabo esto se lleva a cabo una explotación de grandes cantidades de datos aplicando técnicas estadísticas, diferentes algoritmos e incluso aplicación de inteligencia artificial.

Simoudis (1996) afirma que el Data Mining es *"el proceso de extracción de la información previamente desconocida, comprensible y aplicable a partir de grandes bases de datos y usarlo para tomar decisiones de negocio cruciales."*

Esta definición de Data Mining está muy enfocada a entornos de negocio, pero hay que tener en cuenta que es un proceso que puede ser aplicado a cualquier tipo de datos que van desde la predicción del tiempo, diseño de un determinado producto, técnicas de segmentación o agrupamiento, análisis de regresión o redes neuronales entre otras muchas.

Antes de llevar estos procesos a cabo los datos con los que se trabajan necesitan pasar previamente por procesos de extracción, análisis, preprocesamiento, limpieza y selección de los mismos entre otros procesos, ya que trabajar con los datos brutos no permitiría aplicar la mayor parte de los análisis.

Esta parte de extracción, análisis, preprocesamiento, limpieza y selección, también llamada ETL, suele ocupar un porcentaje bastante amplio de la labor llevada a cabo sobre los datos. Con esta parte se logra obtener un proceso que consigue extraer la información potencialmente útil o desconocida a partir de los datos.

El objetivo final de proceso de Data Mining, como se ha comentado anteriormente es encontrar patrones ocultos dentro de grandes conjuntos de datos, y por consiguiente interpretarlos aportando información relevante y que aporte valor.

Personalmente podríamos definir el Data Mining también como un proceso que excava y analiza enormes conjuntos de datos y luego extrae el conocimiento de él, permitiendo automatizar las detecciones de los patrones relevantes en la base de datos. De manera resumida describe cómo se puede descubrir valor oculto dentro de un almacén de datos.

Para verlo de manera conjunta podemos describir el Data Mining como la extracción de información, la cual se encuentra oculta en grandes bases de datos, que posee capacidad analítica, descriptiva y predictiva entre otras. Es una tecnología nueva y potente con un gran potencial para ayudar a las empresas que acumulan información importante dentro de sus almacenes de datos.

Las herramientas de Data Mining son capaces de predecir futuras tendencias y comportamientos de clientes, permitiendo a estas empresas tomar decisiones basadas en el conocimiento que esto les aporta.

Las técnicas de Data Mining comenzaron a evolucionar cuando los datos de negocio se comenzaron a almacenar en primer lugar ordenadores. De manera más reciente se han desarrollado tecnologías que han permitido a los usuarios navegar a través de sus datos en tiempo real.

Podemos apuntar que principalmente el Data Mining se basa en tres tecnologías que han alcanzado un punto importante en su desarrollo, la recopilación de datos masiva, la utilización de equipos con varios procesadores potentes y la aplicación de algoritmos de Data Mining.

El Data Mining ha estado en desarrollo en áreas de investigación, como la estadística, inteligencia artificial y aprendizaje automático. Hoy en día, la madurez de estas técnicas, junto con otros factores, hacen estas tecnologías muy prácticas para entornos de almacenamiento de datos actuales.

El alcance que posee el Data Mining abarca diferentes campos como son el automatizado de predicción de comportamientos, el cual realiza una búsqueda de información predictiva en grandes bases de datos siendo hoy en día una de las razones principales por las que empresas y organizaciones encuentran una creciente necesidad de incorporarlo. Cómo explica Palma et al. 2009 hablando sobre la Inteligencia de Negocios, *"Este nuevo énfasis profesional, más que tecnológico, también se manifiesta en la actividad del Data Mining debido a la creciente necesidad de incorporar, en forma creativa, información de un entorno cada vez más cambiante y a todas luces influyente en la marcha de los negocios de cada compañía."*

Dentro de las distintas funciones encontramos también el descubrimiento automatizado de patrones desconocidos, como ejemplo práctico podemos plantear la observación de compras de productos que se llevan a cabo al comprar otros que están relacionados y que previamente ese patrón determinado de compra no se conocía.

Cuando el Data Mining se implementa en sistemas con alto rendimiento, se pueden analizar bases de datos de gran tamaño en un corto espacio de tiempo, por lo tanto un procesamiento rápido significa que los usuarios pueden trabajar con más modelos y por consiguiente, cuanto antes se tiene el modelo, mayor es el beneficio que se puede obtener debido a la inmediatez.

Como hemos visto anteriormente, existen diferentes técnicas englobadas dentro del área del Data Mining y dentro de la misma se encuentra la principal protagonista de este trabajo, aparte de que están siendo cada vez más utilizada debido a su utilidad, hablamos de las redes neuronales artificiales, las cuales son capaces de formar una serie de modelos con capacidad predictiva a través de aprendizaje. Su funcionamiento y estructura se asemeja a las redes neuronales biológicas.

Por otro lado encontramos los árboles de decisión los cuales representan en forma de árbol conjuntos de decisiones las cuales forman una serie de reglas con las que un conjunto de datos puede ser clasificado. Hay que tener en cuenta que un árbol se puede expresar con reglas, pero al contrario, las reglas no siempre se pueden expresar con árboles.

Los algoritmos genéticos es otra de las herramientas de minería que según Truyol et al. (2007), lo definen como *"un método de búsqueda que imita la teoría de la evolución biológica de Darwin para la resolución de problemas".*

En Data Mining, para llevar todo esto a cabo, se utiliza el llamado modelado basado en la construcción, en el caso de aprendizaje supervisado, de un modelo bajo una determinada situación en la que se conoce previamente la respuesta para posteriormente aplicar esto a una situación en la que se desconoce la respuesta.

Es una técnica para deducir una función a partir de datos de entrenamiento los cuales están formados por datos de entrada y por los resultados deseados. La salida obtenida puede ser un valor numérico o un valor de clasificación. El objetivo del aprendizaje supervisado es el intentar predecir el valor de una situación de entrada después de haber "aprendido" a partir de los datos de entrenamiento.

Hay ocasiones en las que se desconoce la respuesta a priori, y para estos casos existen técnicas no supervisadas basadas en el aprendizaje automático de las respuestas, por lo tanto no necesita que le mostremos los patrones objetivo que buscamos obtener en la salida, dicho de otra manera, nos encontramos en la situación en la que no disponemos de un conjunto de ejemplos, sino que únicamente, a partir de las propiedades de los ejemplos, se intenta llevar a cabo por ejemplo una clasificación o agrupación de los ejemplos que tenemos.

Una vez construido el modelo éste se puede utilizar en situaciones similares en las que se desconoce totalmente la respuesta, ya sean ventas futuras, comportamientos de clientes, realización de pedidos por parte de un tipo de cliente determinado, entre incontables aplicaciones.

Como concluye Suárez (2009) el Data Mining *"ahorra grandes cantidades de dinero a una empresa y abre nuevas oportunidades de negocios. Contribuye a la toma de decisiones tácticas y estratégicas. Proporciona poder de decisión a los usuarios del negocio, y es capaz de medir las acciones y resultados de una mejor forma. Genera modelos descriptivos: permite a empresas, explorar y comprender los datos e identificar patrones, relaciones y dependencias que impactan en los resultados finales.*

*Genera modelos predictivos: permite que relaciones no descubiertas a través del proceso del Data Mining sean expresadas como reglas de negocio."*

### 2.1. Redes Neuronales

La manera con la que vamos a presentar las redes neuronales artificiales es la comparación con las mismas redes neuronales biológicas, ya que como se comenta a continuación, han sido la base en las que se ha basado el desarrollo de las redes neuronales artificiales.

A continuación mostramos una comparativa entre las Redes Neuronales Biológicas y las Artificiales.

| **Redes Neuronales Artificiales** | **Redes Neuronales Biológicas** |
|---|---|
| Unidades de proceso | Neuronas |
| Conexiones ponderadas | Conexiones sinápticas |
| Peso de las conexiones | Efectividad de la sinapsis |
| Signo del peso de una conexión | Efecto excitatorio de una conexión |
| Función de propagación | Efecto combinado de la sinapsis |

*Fuente: Inteligencia artificial. Redes neuronales y aplicaciones. Galán y Martínez.*

Como hemos introducido, a lo largo del tiempo los expertos en computación se han inspirado en el funcionamiento del cerebro humano para llevar a cabo estos avances. A partir de una red neuronal artificial (y a partir de ahora nos referiremos simplemente como una "red neuronal") la cual constituye un modelo computacional basado en el cerebro, se ha utilizado para resolver ciertos tipos de problemas.

Como ejemplo podemos tener en cuenta problemas que resultan simples para una persona como el reconocer en una imagen a un niño o un animal e identificarlos, o reconocer a personas que se han conocido el día anterior, pero para una máquina no sería tan fácil.

Una de las aplicaciones de las redes neuronales, las cuales evolucionan de manera continua hoy en día, se resume en llevar a cabo tareas que resultan fáciles para una persona basándose en el reconocimiento de patrones. Esta aplicación de las redes neuronales va desde el reconocimiento óptico de caracteres ya sean, signos, escritos o reconocimiento facial, determinación de relaciones no lineales, predicción y clasificación entre otras.

Una red neuronal no sigue una trayectoria lineal, más bien la información se procesa en conjunto y en paralelo a lo largo de una red de nodos. Uno de los elementos clave de una red neuronal es su capacidad para aprender. Una red neuronal no es sólo un sistema complejo, sino que puede cambiar su estructura interna basada en la información que fluye a través de él y esto se logra mediante el ajuste de pesos de cada conexión.

Si la red genera un resultado "bueno" en la salida no hay necesidad de ajustar los pesos. Sin embargo, si la red genera un resultado "malo" a la salida, un error por decirlo de alguna manera, el sistema se adapta, y por consiguiente encontramos la alteración de los pesos de la red, con el fin de mejorar los resultados y aprender de estos resultados "malos". Existen varias estrategias para el aprendizaje, y a continuación vamos a examinar dos de ellas, por un lado la estrategia de aprendizaje supervisado y por otro lado el no supervisado.

Como ya hemos estado comentado anteriormente, una de las estrategias de aprendizaje que estudiamos es el aprendizaje supervisado, el cual, esencialmente, se basa en una

estrategia que implica un “maestro” que es más inteligente que la propia red y procurará enseñar a la misma con situaciones previamente establecidas.

Por ejemplo, tomemos el ejemplo del reconocimiento facial con lo que el profesor muestra la red muchas caras, y el maestro ya sabe el nombre asociado a cada cara. La red hace sus conjeturas, a continuación, el maestro proporciona la red con las respuestas. La red puede luego comparar sus respuestas a las “correctas” conocidos y hacer ajustes de acuerdo a sus errores.

Por otro lado, y como se ha visto anteriormente, encontramos el aprendizaje no supervisado dentro de las redes neuronales. Como ejemplo podemos tener en cuenta la búsqueda de un patrón oculto en un determinado conjunto de datos. Una posible aplicación de esto es la agrupación, es decir, la división de un conjunto de elementos en grupos de acuerdo a un patrón desconocido.

El aprendizaje no supervisado, por decirlo de algún modo práctico, se basa principalmente en la observación. Pongamos como ejemplo a un ratón que corre a través de un laberinto, el cual, si gira a la izquierda, encuentra un trozo de queso, por el contrario, si se gira a la derecha, recibe una pequeña descarga. Por lo tanto, se presupone que el ratón aprenderá con el tiempo únicamente girar hacia la izquierda. Su red neuronal toma una decisión con un resultado previamente obtenido a través de dichas observaciones. Si la observación es negativa, la red puede ajustar sus pesos con el fin de tomar una decisión diferente la próxima vez.

Este tipo de aprendizaje por refuerzo es común en la robótica en la que se hace de forma continua. Esta capacidad de una red neuronal para aprender, para realizar ajustes en su estructura con el tiempo, es lo que hace tan útil en el campo de la inteligencia artificial.

Tras haber comentado anteriormente las aplicaciones de las redes neuronales, vamos a profundizar un poco más en algunos de los usos habituales de las mismas que encontramos en el software de hoy en día.

Por un lado el reconocimiento de patrones, los cuales hemos mencionado anteriormente y que es probablemente la aplicación más común de las redes neuronales. Ejemplos de

ello, como hemos ejemplarizado anteriormente son el reconocimiento facial y reconocimiento óptico de caracteres entre muchas otras aplicaciones.

Otra aplicación es la predicción de series de tiempo, a través de las cuales se establecen unas redes que se pueden utilizar para hacer predicciones a lo largo del tiempo. Como ejemplo práctico utilizaremos los ejemplos de valores de acciones de una determinada empresa, y estudiaremos si aumentará el valor de la acción o si caerá mañana.

Como expone García et al. 2007 en la Revista Ingeniería Biomédica, *"en el área de la salud encontramos la aplicación de las redes neuronales en el procesamiento de señales desarrollados en implantes cocleares y audífonos, los cuales las necesitan para filtrar el ruido innecesario y amplificar los sonidos importantes. Las redes neuronales pueden ser entrenadas para procesar una señal de audio y filtrarla de manera adecuada."*

Actualmente se están desarrollando grandes avances en la investigación en los vehículos con sistemas de auto-conducción donde las redes neuronales se utilizan a menudo para tomar las decisiones de dirección de vehículos físicos (o simulados).

Las redes neuronales también se pueden aplicar bajo un sistema de sensores las cuales llevan a cabo un proceso de análisis de una colección de muchas mediciones. Por ejemplo un termómetro que puede indicar la temperatura del aire, la humedad, la presión barométrica, punto de rocío, la calidad del aire, la densidad del aire, entre otras variables.

Las redes neuronales se pueden emplear para procesar los datos de entrada procedentes de muchos sensores individuales y evaluarlos conjuntamente y poder indicar posibilidad de determinadas situaciones climáticas.

Otra aplicación de las redes neuronales es la detección de anomalías, esto se debe a que son tan potentes reconociendo patrones que pueden ser entrenadas para generar una salida (aviso) en el momento que algo no encaja con el patrón previamente aprendido. Por ejemplo en el proceso de control de calidad de una fábrica, una red neuronal sigue su rutina diaria durante un largo período de tiempo, y después de aprender los patrones

del comportamiento de la maquinaria, podría alertar cuando algo ande mal en dicho proceso.

Por lo tanto, teniendo en cuenta todo lo anteriormente comentado, podemos sacar en claro que las redes neuronales podrán funcionar bien para la identificación de posibles asociaciones o para descubrir posibles regularidades dentro de un conjunto de patrones donde el volumen, el número de variables o la diversidad de los datos es muy grande, o en situaciones donde las relaciones entre variables están vagamente identificadas o son difíciles de describir adecuadamente con un enfoque convencional.

Para las redes neuronales existen muchas ventajas y limitaciones y para discutir este tema adecuadamente tendríamos que estudiar individualmente cada tipo de red.

Podemos identificar como principales ventajas de las redes neuronales, en primer lugar la capacidad de aprendizaje adaptativo que éstas tienen, capaces de aprender a realizar tareas basadas es “experiencia” previa, por otro lado la auto-organización de las mismas, ya que una red neuronal es capaz de auto-organizar la información que recibe a través del aprendizaje.

Por otro lado tenemos en cuenta la alta tolerancia a fallos, debido a que la red neuronal es capaz de mantener algunas capacidades a pesar de una destrucción parcial de la misma. La operación en tiempo real es otra de sus ventajas, lo cual permite realizar cómputos en paralelo lo cual necesita hardware con cierto nivel de capacidad.

En referencia a los tipos de redes y sus características comenzamos refiriéndonos a las redes “backpropagation” las cuales son llamadas, en cierto sentido, “cajas negras”. De manera general se puede decir que el usuario no tiene otro papel que “alimentar” la entrada de datos, observar como ésta entrena y esperar la salida de la red. De hecho, “casi no se sabe lo que está haciendo".

Algunos paquetes de software de libre que podemos encontrar como son “NevProp” o “Mactivation” entre otros permiten ir probando las redes a intervalos de tiempo regulares, pero el aprendizaje continúa avanzando por sí mismo.

Como describe Valencia et al. 2006, en una red Backpropagation *"existe una capa de entrada con n neuronas y una capa de salida con m neuronas y al menos una capa oculta de neuronas internas. Cada neurona de una capa (excepto las de entrada) recibe entradas de todas las neuronas de la capa anterior y envía su salida a todas las neuronas de la capa posterior (excepto las de salida). No hay conexiones hacia atrás feedback ni laterales entre las neuronas de la misma capa."*

El producto final de esta actividad es una red entrenada que no proporciona ningún tipo de ecuaciones o coeficientes que definen una relación, al contrario de lo que se puede encontrar en un modelo de regresión.

Uno de los problemas que podemos encontrar en las redes "Backpropagation" es que tienden a ser más lentas de entrenar que otros tipos de redes. Esto también dependerá de con qué máquina se ejecuten este tipo de redes, y en función de esto necesitará invertir más o menos tiempo.

### 2.2. Deep Learning

El Deep Learning no toma parte relevante de este estudio, pero sí es importante tenerlo presente por múltiples razones, una de ellas que la de mostrar cómo las Redes Neuronales continúan utilizándose, dicho de otro modo, el Deep Learning es una nueva técnica de Inteligencia Artificial basada en las Redes Neuronales las cuales desarrollan un conjunto de algoritmos, que como comentaremos más adelante, integra análisis de imágenes, sistemas de solución de problemas y lenguaje natural entre otros.

Como describe L. Deng et al. 2014, el Deep Learning o aprendizaje profundo *"es un conjunto de algoritmos que intenta modelar abstracciones de alto nivel en los datos mediante el uso de arquitecturas compuestas de transformación no lineales múltiples"*. En definitiva se trata de lograr aprender modelos muy complejos con nuevas técnicas mucho más eficientes.

Como explica LeCun et al. (2015), el Deep Learning *"permite aplicar modelos computacionales que están compuestos de múltiples capas de procesamiento para aprender representaciones de datos con múltiples niveles de abstracción. Estos métodos*

*han mejorado drásticamente el estado de la técnica en reconocimiento de voz, reconocimiento de objetos visuales, detección de objetos y muchos otros dominios tales como descubrimiento de fármacos y genómica."*

El que se lleven a cabo modelos con gran número de capas permite crear modelos con relaciones más complejas, que como se ha comentado anteriormente, permite que dichos modelos lleguen a ser más eficientes.

Al igual que muchos otros muchos programas, los que utilizan el Deep Learning pasan por un proceso en el que a cada algoritmo se le aplica una transformación no lineal en su entrada y utiliza lo que aprende para crear un modelo estadístico como salida. Las iteraciones continúan hasta que la salida ha alcanzado un nivel aceptable de precisión. El número de capas durante el procesamiento por las que deben pasar los datos es lo que provocó que se calificara como "profunda".

La ventaja del aprendizaje profundo es que el sistema se basa en lo mismo pero sin llevar a cabo ningún proceso de supervisión. Esto no sólo es más rápido, por lo general es más precisa.

Inicialmente, podemos estar provistos de datos de entrenamiento, siendo estos un conjunto de imágenes para las que previamente hemos etiquetado cada imagen como "perro" o "no perro". El programa utiliza la información que recibe de los datos de entrenamiento para crear un conjunto de características para "perro" y a partir de dicho conjunto construir un modelo predictivo.

En este caso, el modelo creado podría predecir que en principio cualquier cosa en una imagen que tiene "cuatro patas" y una "cola" debe ser etiquetada como "perro". Por supuesto, no es capaz de entender que son "cuatro patas y cola", pero simplemente busca patrones de píxeles de los datos introducidos previamente.

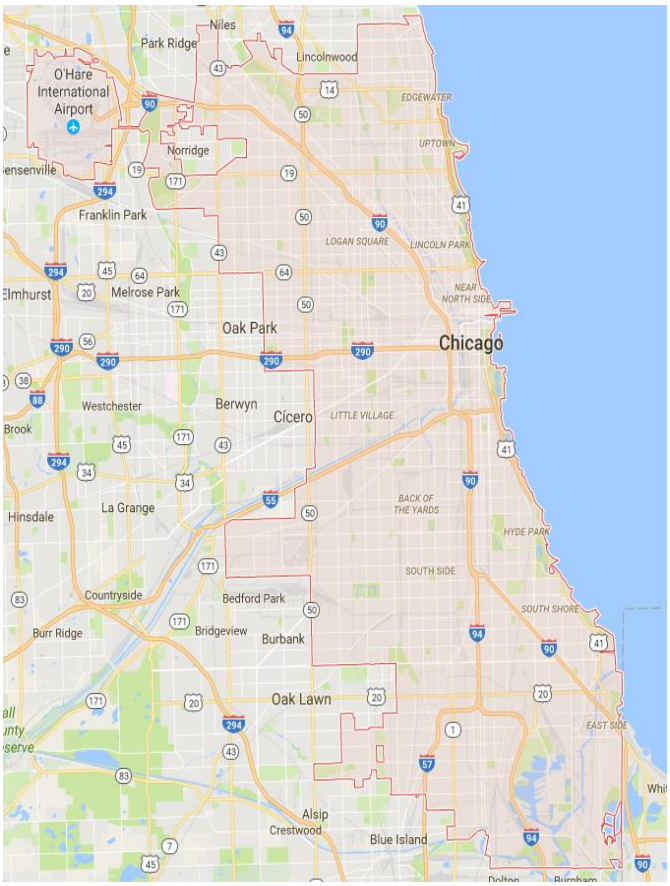
Con cada iteración, el modelo predictivo se vuelve más complejo y más preciso. De esta manera utilizando algoritmos de aprendizaje profundo, se puede establecer un conjunto de entrenamiento, ordenar a través de millones de imágenes e identificar con precisión las imágenes que tienen "perros" en ellos dentro de unos minutos.

Con el fin de lograr un nivel aceptable de precisión, los programas de aprendizaje profundo requieren acceso a enormes cantidades de datos de entrenamiento y capacidad de procesamiento, y esto ha supuesto un problema por la ingente cantidad de datos de los que disponer, pero gracias a los avances actualmente se es capaz de crear modelos estadísticos complejos directamente de la propia producción de datos. Esto es posible principalmente a tecnologías actuales como el "Internet de las Cosas" (IOT) debido a que actualmente existe la posibilidad de innumerables conexiones entre máquinas.

Hoy en día los casos de uso para el aprendizaje profundo abarcan gran número de aplicaciones, especialmente las que se centran en el procesamiento del lenguaje natural como puede ser la traducción de idiomas, diagnóstico médico, seguridad en la red o identificación de imagen entre otras muchas posibilidades.

## 3. PROBLEMA A ABORDAR

Se expone a continuación a qué caso nos enfrentamos, siendo este la predicción de cómo se va a comportar un determinado mosquito perteneciente a la familia "Culex", familia de mosquitos los cuales actúan como vector directo de importantes enfermedades, ya sea encefalitis y malaria aviar entre otras, y el cual transmite el Virus del Nilo Occidental el cual va a ser el principal objetivo de este trabajo, conocer el comportamiento/movimiento de dicho mosquito y por consiguiente, estudiar de qué forma se va a dispersar el virus del Nilo a través de la ciudad de Chicago, en Estados Unidos.




*Ciudad de Chicago. Fuente: Google Maps*

*Mosquito de la familia "culex".*

*Fuente: https://commons.wikimedia.org/wiki/File:CulexNil.jpg*

A continuación, expondremos información con el fin de conocer un poco más acerca del mosquito objeto de estudio, el cual es portador del virus del Nilo. En los primeros momentos de vida de este mosquito, encontrado en forma de larva, eclosiona en presencia de agua, viviendo durante dicho estado dentro de la misma y formando grupos compactos en aguas estancadas.

Como se ha comentado antes, gracias al comportamiento previo de dichos mosquitos portadores del virus a estudiar, en presencia ciertas variables como los distintos factores climáticos procederemos a diseñar un sistema de Redes Neuronales con el que predecir la propagación de los mismos a través del aprendizaje de las propias Redes.

A continuación vamos a presentar de manera más específica los datos con los que se va a llevar a cabo este trabajo y seguidamente profundizaremos en las herramientas que vamos a usar, desde R como el software base sobre el que desarrollar este proyecto, y de manera conjunta exponer también los distintos métodos que se van a llevar a cabo sobre R para lograr el objetivo descrito.

## 4. HERRAMIENTAS A UTILIZAR

Actualmente encontramos incontables herramientas para llevar a cabo todo tipo de cálculos, algoritmos y representación gráfica entre otras muchas herramientas, todas caracterizadas por distintas características como versatilidad, potencia, aplicabilidad o escalabilidad entre otras tantas.

Encontramos tanto herramientas de pago, las cuales han ido logrando una mejora sustancial con el paso del tiempo, como herramientas con licencia gratuita que gracias a grandes comunidades de usuarios han ido ayudando a su mejora y desarrollo, herramientas previamente estructuradas y definidas para realizar labores específicas como en este caso, basadas en el Data Mining e inteligencia artificial o incluso herramientas capaces de realizar otras muchas tareas. A continuación vamos a profundizar en esta situación.

En nuestro caso, para ser específicos, no vamos a estudiar una serie de herramientas determinadas, más bien nos vamos a centrar en una sola que no es una herramientas en

sí, sino que se puede definir como un entorno y lenguaje de programación con una gran aplicabilidad a innumerables problemas y situaciones a los que se puede aplicar para su posible resolución, estamos hablando de R, y en el cual vamos a profundizar a continuación.

### 4.1. R

A continuación analizamos los medios que se van a usar para llevar acabo la red neuronal con la que alcanzar nuestro objetivo. Para llevar a cabo este estudio predictivo se utilizará el entorno y lenguaje de R, el cual está enfocado principalmente al análisis estadístico y con el cual se pueden desarrollar gran cantidad de análisis estadísticos.

R es un entorno en el que se han implementado incontables técnicas estadísticas distribuido bajo una licencia pública y estando disponible para distintos sistemas operativos como son Windows, Unix, MacOS y Linux.

Profundizando aún más en R encontramos que éste proviene de un lenguaje desarrollado en los años 80 con la finalidad de realizar cálculos estadísticos, éste se denominaba S. A partir de dicho lenguaje se desarrolló el conocido lenguaje de código abierto llamado R.

Una de las ventajas principales de R es la capacidad de implementación de distintas librerías las cuales contienen todo lo investigado y desarrollado en el ámbito estadístico. A parte, R forma una comunidad que posee una gran magnitud la cual constantemente se sigue ampliando y lleva a cabo el mantenimiento del sistema.

Una de las ventajas más atractivas de R es que posee un lenguaje de programación orientado a objetos realmente completo brindando gran número de facilidades. Por otro lado R permite modificar las funciones y ajustarlas según se necesite o incluso tomarlas para implementarlas en alguna rutina que esté el usuario llevando a cabo.

Otro beneficio que posee R es la facilidad de conectarlo con distintos programas, como puede ser unirlo con SQL a través de los distintos paquetes enfocados para esto tales como *“sqldf”,* lo que permite realizar consultas desde R conectado a SQL pudiendo llevar a cabo tratamiento de bases de datos de gran tamaño e implementándolo en R

para trabajar con ello más adelante brindando todas las posibilidades que SQL brinda para la explotación de bases de datos.

Otro punto fuerte con el que contaremos de R en la gran variedad de gráficos implementados y posibilidad de incorporación de muchos otros que aportan ciertas mejoras que no se pueden encontrar en los previamente incorporados en R y los cuales facilitan enormemente el análisis, representación y obtención de conclusiones basadas en los mismos debido a la facilidad de extracción para su posterior presentación y edición.

Cómo comentamos anteriormente, al tener R licencia gratuita se puede acceder de manera fácil a toda la documentación y manuales referentes a R y a toda la librería de la que se dispone, resultando de gran ayuda a la hora de utilizar algún paquete determinado. Esto permite que pueda utilizarse de manera cómoda sin necesidad de conocer íntegramente las librerías que se están usando.

R nos permite realizar una cantidad enorme de estudios estadísticos de una gran variedad de datos, y además éste puede llegar a integrarse con otros lenguajes a través de los que puede ser llamado, estos lenguajes interpretados pueden ser Python, Perl y Rubi, y por supuesto existen otros proyectos que permiten utilizar R desde .net y Java.

## 5. RESOLUCIÓN DEL CASO PLANTEADO

A continuación procederemos a resolver el caso planteado al comienzo de este trabajo a través de Redes Neuronales Artificiales y llevado a cabo gracias al entorno de R y sus diferentes paquetes los cuales aportarán al nuestro entorno una riqueza de funciones necesarias para llevar a cabo la resolución del caso.

Siendo más específicos utilizaremos R bajo el entorno de RStudio, el cual facilita el uso de R gracias a la interfaz amigable de la que disponemos bajo dicho entorno, y donde introduciremos los datos disponibles, desarrollaremos un análisis de datos y su correspondiente Red Neuronal con el fin de lograr el principal objetivo de predecir las futuras localizaciones del mosquito portador del virus del Nilo.

### 5.1. Preparación de los datos

Para llevar a cabo la correspondiente preparación de los datos comenzaremos unificando las bases de datos recogida con la correspondiente información sobre los mosquitos y su localización con la base de datos del clima, uniéndolas a través de la fecha en la que se ha recogido esa información para lograr agrupar las variables de localización del mosquito con las variables de clima.

Tras llevar a cabo la correspondiente unificación en una sola base de datos donde hemos incorporado las variables recogidas sobre los mosquitos como la identificación de la trampa, localización y presencia del virus del Nilo en ellos entre otras, las variables climáticas recogidas durante el estudio de los mosquitos, llegando de este modo a la selección final de las variables de interés para llevar a cabo el entrenamiento de la red neuronal correspondiente.

Debido a esta unificación de variables logramos acumular un total de 21012 observaciones con un total de 16 variables, pero antes de introducir dichos datos a una red neuronal y comenzar a entrenarla deberemos normalizar todas estas variables.
Las variables seleccionadas finalmente para la aplicación de la red neuronal son las siguientes:

| | |
|---|---|
| **Número de mosquitos** | **Temperatura media** |
| **Temperatura mínima** | **Velocidad media del viento** |
| **Precipitación total** | **Temperatura máxima** |
| **Velocidad del viento resultante** | **Virus presente en el mosquito** |
| **Temperatura húmeda del aire** | **Bloque** |

| | |
|---|---|
| **Punto de rocío** | **Dirección resultante** |
| **Presión estándar** | **Nivel del mar** |

*Fuente: Elaboración propia.*

Como explica Mendelssohn, (1993), *"En las redes neuronales es buena práctica el preprocesar datos de entrada antes de su uso. El preprocesamiento de datos hace que la formación de la red sea más rápida, eficiente en cuanto a la memoria y obtenga resultados de pronóstico precisos. Las redes neuronales sólo funcionan con datos normalmente entre un rango especificado, p. - 1 a 1 o 0 a 1, hace necesario que los datos sean reducidos y normalizados. La escala puede ser tan simple como tomar las relaciones (normalización recíproca), calcular las diferencias (normalización del rango) y normalización multiplicativa o normalización del eje Z. La normalización garantizará que los datos se distribuyen uniformemente entre las entradas de red y las salidas."*

Con las variables a trabajar definidas y normalizadas podremos proceder a introducir los datos en la red neuronal. Debemos tener en cuenta que todas las variables explicativas con las que vamos a diseñar la red neuronal se encuentran normalizadas a excepción de las variables explicadas, las cuales son "latitud" y "longitud".

### 5.2. Técnicas aplicadas

Para comenzar mostrando que paquetes y funciones se están utilizando para llevar a cabo todo el proceso nombraremos los paquetes que hemos utilizado al comienzo para la representación de los mosquitos según fechas de mayor recuento de los mismos y sus localizaciones según longitud y latitud.

Para la carga del mapa en formato rds proporcionado en la plataforma donde se obtuvieron los datos cargamos el paquete "readRDS". Para la representación gráfica sobre el mapa se ha usado el paquete "ggmap", utilizando las funciones de este último paquete para cargar y representar sobre el mapa de Chicago las coordenadas de

localización de los mosquitos. Para esto se ha usado la función “ggmap” y “geom_point” entre otras y se han representado las correspondientes localizaciones.

A continuación nombraremos al paquete de R llamado “plyr”, el cual nos ha facilitado la unión de las bases de datos de las que disponíamos al comienzo de manera bastante ágil, utilizando la función que incorpora “join” con el fin de incorporar a la información recogida sobre la localización del mosquito las distintas variables climáticas y agrupados según el día en el que se recogieron los mismos datos.

Para la normalización de las variables explicativas se ha utilizado la función “scale” del paquete “base” de R con el cual es posible centrar y tipificar las variables con el fin de facilitar la labor a la red neuronal.

Para llevar a cabo la red neuronal hemos utilizado el paquete “nnet” y sus distintas funciones para entrenar la red neuronal y para representarla gráficamente, gracias a “devtools”, hemos instalado una actualización para los gráficos del paquete “nnet”.

Tras ello utilizaremos la función “predict” utilizando la red neuronal previamente diseñada para hacer una predicción de las próximas coordenadas.

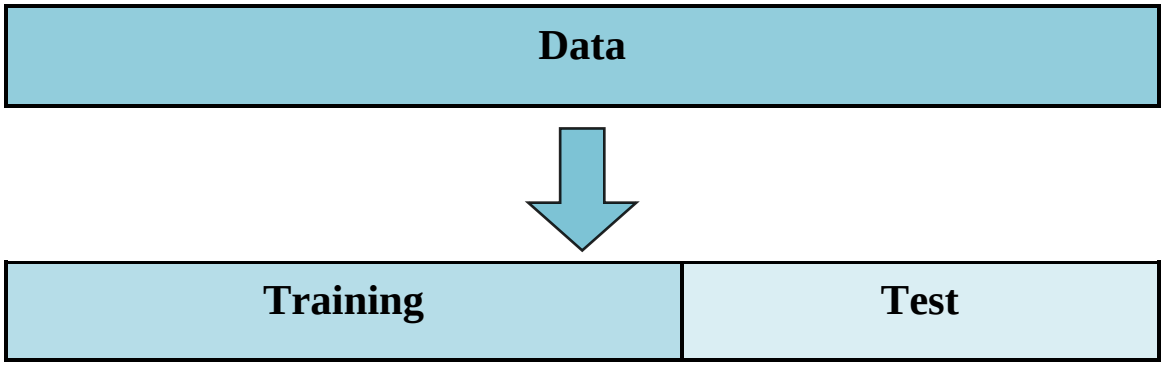
Con el fin de evaluar los resultados del proceso de predicción llevado a cabo por la red neuronal artificial vamos a dividir los datos que tenemos en dos partes con el paquete “Caret”, con lo que obtendremos dos conjuntos de datos, un “train” o datos de entrenamiento, con el que enseñaremos a la red neuronal, y otro llamado “test” o datos de prueba, con el que validaremos la predicción de la misma.

Cuando tenemos suficientes datos, se puede subdividir los datos en distintos conjuntos, y en esta ocasión los dos que se han comentado anteriormente. Durante el proceso de aprendizaje la red neuronal toma los datos de entrenamiento.

Finalmente, una vez que termina el proceso de aprendizaje y posteriormente predicción, se pueden utilizar los datos de prueba para evaluar la manera en que la red ajustada finalmente se generaliza para los datos que no jugaron ningún papel en la modelización de la misma.

Hastie et al. 2011, señalan que *"es difícil dar una regla general sobre cuántas observaciones se deben asignar a cada conjunto, aunque indican que una división típica puede ser de 50% para el entrenamiento y 25% para la validación y prueba, respectivamente."*

Como comentamos anteriormente solo utilizaremos dos subconjuntos, el de entrenamiento y el de prueba. A continuación representamos la división del conjunto principal en los dos subconjuntos.




*Fuente: Elaboración propia.*

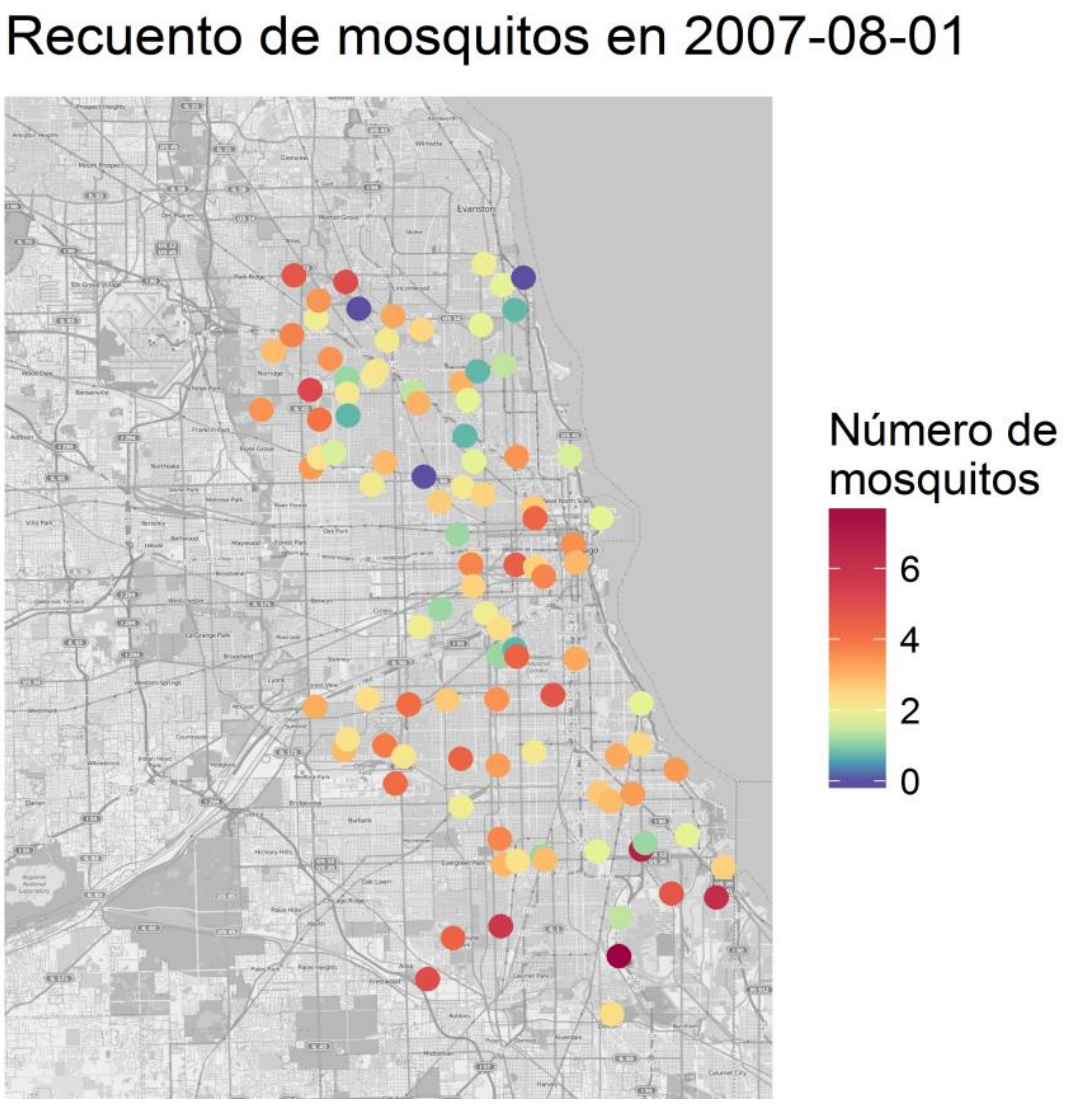
Antes de llevar a cabo la división de los datos en "train" y "test" tenemos una base de datos con variables de mosquitos y de clima con un total de 21012 observaciones. Tras la división a través de "caret" obtenemos un conjunto de entrenamiento de 15759 observaciones y un conjunto de test de 5253 observaciones.

A partir de estos dos conjuntos de datos se va a llevar a cabo el entrenamiento y validación de la red neuronal que posteriormente se va a diseñar.

### 5.3. Pruebas y resultados obtenidos

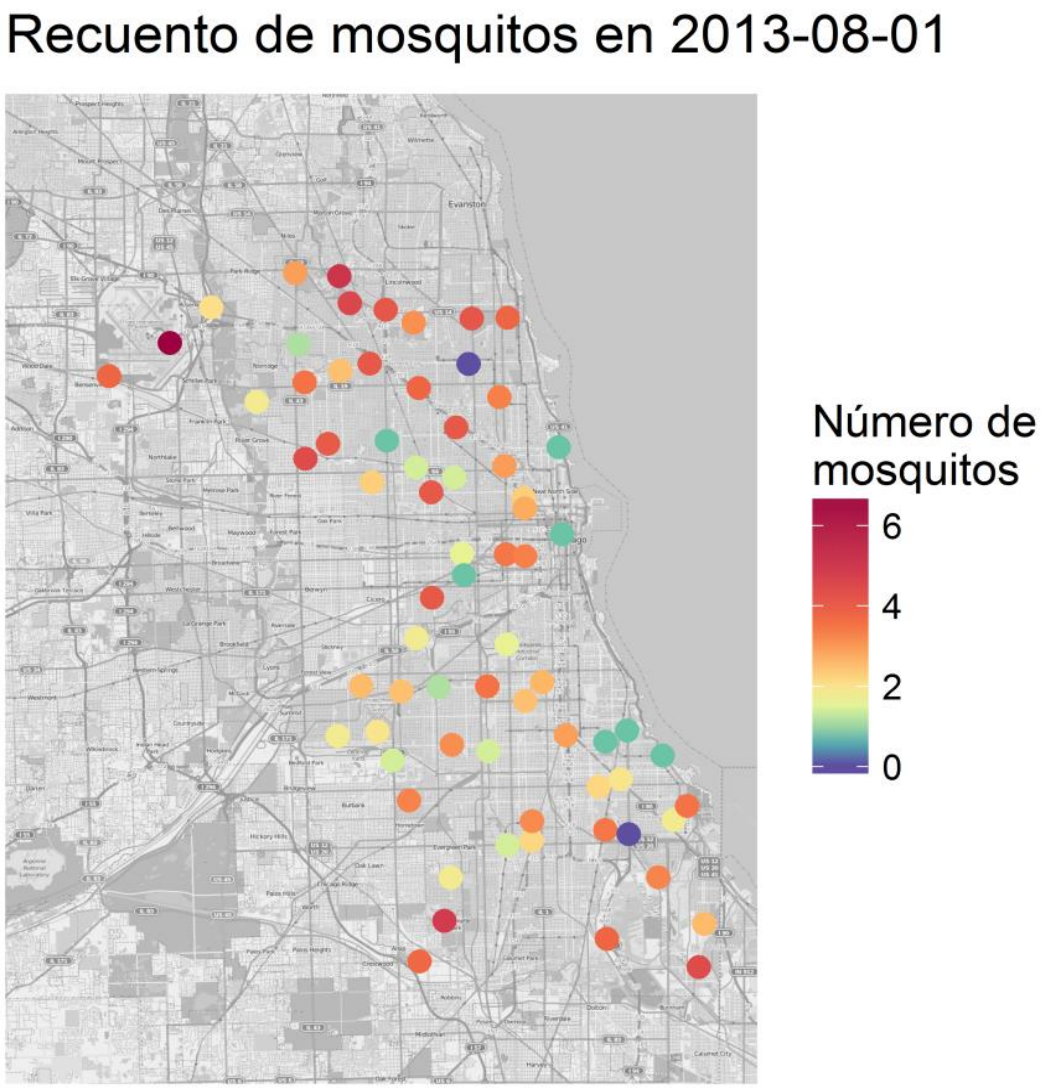
Llevaremos a cabo una serie de análisis para conocer previamente comportamientos previos de los mosquitos. En primer lugar representaremos sobre el mapa de Chicago cómo se han distribuido los mosquitos en diferentes fechas. Debido al gran número de fechas de las que disponemos vamos a agrupar las mismas según el número de mosquitos por fecha, seleccionando aquellas en las que mayor número de los mismos se encontraron y a continuación representado gráficamente formando un mapa de calor.

Con un mayor número de mosquitos contabilizados, con un total de 551 a fecha de 1/08/2007 encontramos esta distribución de los mosquitos.




*Fuente: Elaboración propia.*

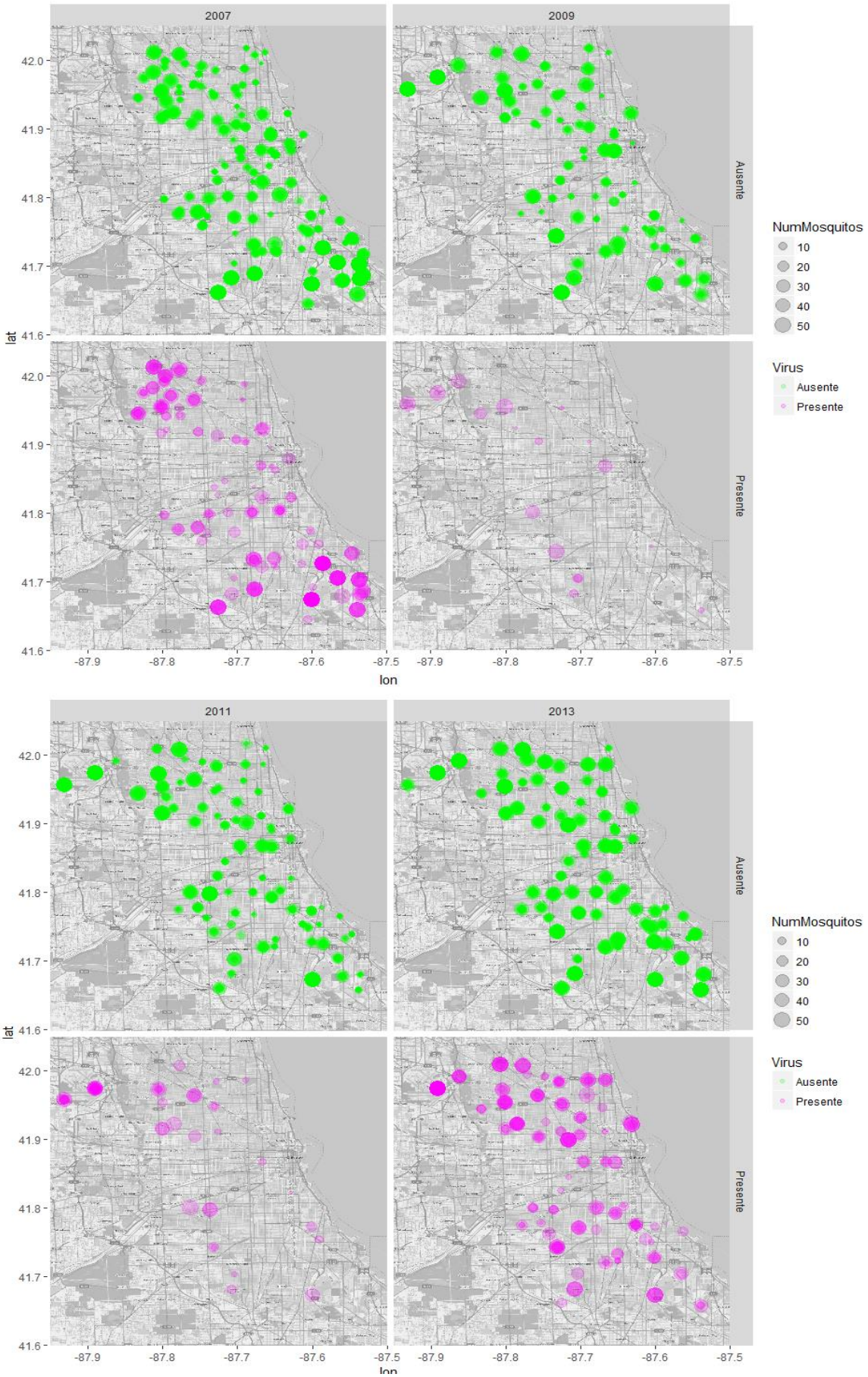
En segundo lugar, vamos a comparar el recuento de mosquitos de 2007 que observamos en el mapa anterior con el recuento de las últimas fechas de obtención de datos y con uno de los mayores recuentos de mosquitos de todas las fechas con un total de 186 a fecha de 01/08/2013 encontramos esta distribución de los mosquitos.




*Fuente: Elaboración propia.*

Con estos mapas de calor podemos observar cómo se distribuyen los mosquitos en fechas señaladas y distantes en las que se han llegado a recoger el mayor número de los mismos.

A continuación, para seguir profundizando en los datos y analizar de manera más gráfica al mosquito y la existencia de virus durante los años que se ha llevado el estudio hemos representado gráficamente sobre el mapa de Chicago gracias a “ggmap” y utilizando las variables de presencia de virus, número de mosquitos, longitud y latitud, un mapa de calor con el que observar las localizaciones dónde se ha encontrado mosquitos dividiendo en el caso de si fuera o no fuera portador del virus del Nilo. Se puede observar en el gráfico cómo según el grosor de las burbujas si el recuento de los mismos es mayor o menor.




*Fuente: Elaboración propia.*

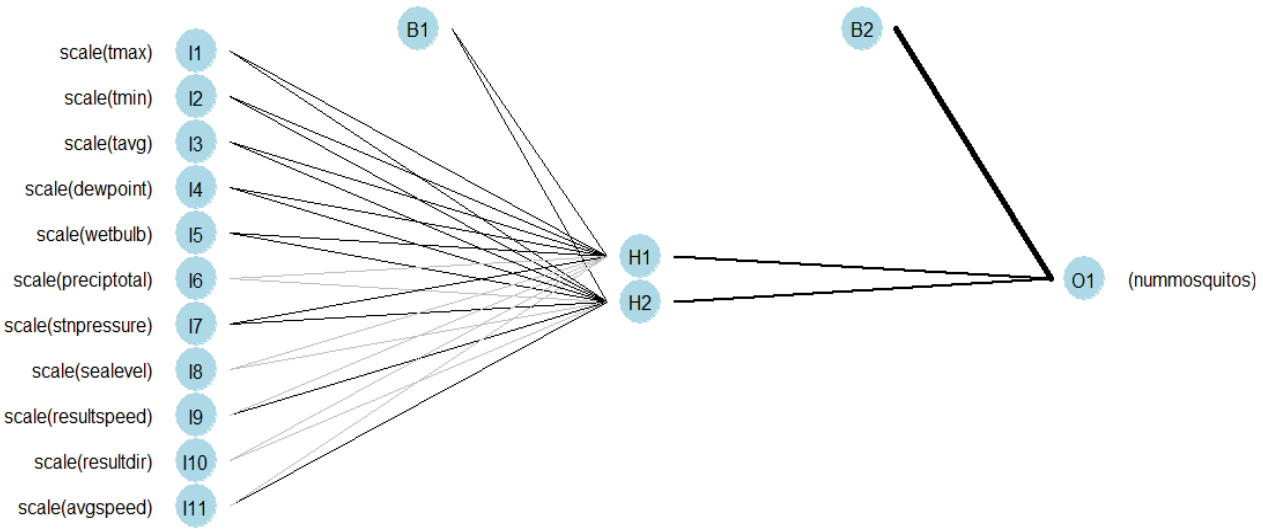
Se puede observar en el mapa anterior cómo los años 2007 y 2013 muestran a simple vista un aumento de la aparición de un mayor número de mosquitos portadores del virus del Nilo.

A continuación vamos a aplicar una red neuronal utilizando los distintos conjuntos previamente divididos, el conjunto de entrenamiento para, como su nombre indica, entrenar la red, y el conjunto de test con el que validaremos la predicción realizada por la propia red neuronal. Se ajusta una red neuronal 11-2-1 con un total de 27 pesos, de los cuales se obtendrán distintos según para que bloque (localización) se entrene la red. A continuación se muestra un pequeño número de los mismos pertenecientes al bloque número 10.

| ***b->h1*** | ***i1->h1*** | ***i2->h1*** | ***i3->h1*** | ***i4->h1*** | ***i5->h1*** | ***i6->h1*** |
|---|---|---|---|---|---|---|
| *131.32* | *25.96* | *36.82* | *27.10* | *31.50* | *-15.97* | *7.81* |

| ***b->h2*** | ***i1->h2*** | ***i2->h2*** | ***i3->h2*** | ***i4->h2*** | ***i5->h2*** | ***i6->h2*** |
|---|---|---|---|---|---|---|
| *-355.76* | *62.46* | *75.11* | *73.30* | *52.33* | *64.89* | *-15.59* |

Estableciendo un total de 2 nodos en la capa oculta obtenemos la siguiente red neuronal de estructura semejante para cada uno de los bloques.



A partir de la red neuronal previamente entrenada con el conjunto de datos “train” procedemos a predecir, por cada localización identificada como bloque, el número de mosquitos que ocuparán los mismos utilizando el conjunto de datos “test”, y de tal manera observar el error de predicción obtenido.

Para llevar esto a cabo se diseña en RStudio un bucle con el que dividir los conjuntos de datos según bloques, para así enfocar la predicción del número de mosquitos sobre cada uno de ellos y lograr conocer el siguiente movimiento de los mismos.

Tras llevar a cabo el “predict” para cada uno de los bloques utilizando el conjunto de datos “test” con el fin de comprobar la capacidad de predicción que posee la red neuronal a través del error medio.

En primer lugar obtenemos los residuos o errores observando la diferencia entre el bloque obtenido a partir de la predicción de la red neuronal y del valor real el cual encontramos en el conjunto de datos que forma “test”.

A continuación se muestra, como ejemplo, una de las tablas con los residuos obtenidos en la predicción del bloque 10 para conocer el número de mosquitos. A residuo nos referimos con la diferencia entre el valor estimado frente al real.

| **Observación** | **Error de predicción** |
|---|---|
| 117 | 1,4434174 |
| 381 | 2,2831178 |
| 383 | 1.7509277 |

*Fuente: Elaboración propia.*

A continuación, agrupamos en una tabla los errores de estimación del número de mosquitos en función del bloque en el que se encuentren. El error con el que se está trabajando es la diferencia entre el valor ajustado o predicho a través de la red neuronal previamente entrenada y el valor real tomado del conjunto de datos “test”.

Al ser una gran cantidad de bloques vamos a mostrar el error medio absoluto de las más relevantes teniendo en cuenta las que poseen un mayor número de observaciones.

| **Bloque** | **Error de predicción por bloque** |
|---|---|
| 10 | 2,4434174 |
| 11 | 4,2831178 |
| 12 | 3.7509277 |
| 52 | 4.523222 |
| 50 | 3.232652 |

*Fuente: Elaboración propia.*

## 6. Conclusiones

A la vista de los resultados podemos concluir que las redes neuronales artificiales poseen la capacidad de aprender con los datos incorporados y lograr el objetivo principal de este trabajo, el cual se ha basado en la comprobación de las técnicas de redes con el fin de aprender, de realizar predicciones y lograr anteponerse a futuros sucesos.

Aunque hay que tener cierto cuidado, con un número de entradas que tenemos para dentro de cada bloque, ya que la red neuronal ha demostrado que crea mejores predicciones, y por lo tanto reduce el error al poseer una cantidad mayor de observaciones con las que la red puede aprender mejor.

Conociendo la capacidad que poseen las redes neuronales y como su aplicabilidad sigue en auge podemos imaginar que muchos modelos estadísticos quedarán obsoletos y en desuso debido a las capacidades predictivas de estas redes. Gracias a que cada vez estas técnicas están en continuo desarrollo se logrará alcanzar objetivos mucho más específicos.

## 7. Bibliografía


- *Brachman. R.J.; Khabaza, T.; Kloesgen, W.; Piatetsky Shapiro, G.; and Simoudis, E. 1996. "Industrial Applications of Data Mining and Knowledge Discovery", in Proc. Of the Second International Conference on Knowledge Discovery and Data Mining (KDD-96), Menlo Park.*

- *Rodríguez Suárez, Yuniet; Díaz Amador, Anolandy; 2009. Herramientas de Minería de Datos. Revista Cubana de Ciencias Informáticas 3: 73-80.*

- *César Ferri Ramírez, José Hernández Orallo, Mª José Ramírez Quintana, 2004. Introducción a la Minería de Datos. Editorial Alhambra S. A.*

- *Ethem Alpaydın, 2010. Introduction to Machine Learning, Second Edition. Massachusetts Institute of Technology.*

- *Tom Michael Mitchell, 1997. Machine Learning. McGraw-Hill.*

- *Marco Antonio Valencia Reyes, Cornelio Yáñez Márquez, Luis Pastor Sánchez Fernández, 2006. Algoritmo Backpropagation para Redes Neuronales: conceptos y aplicaciones Instituto Politécnico Superior. Centro de Investigación en Computación.*

- *Arranz de la Peña, Jorge; Parra Truyol, Antonio "Algoritmos Genéticos", Universidad Carlos III.*

- *LeCun, Yann, Bengio, Yoshua, Hinton, Geoffrey (2015). Deep learning, Nature Publishing Group, Macmillan Publishers Limited.*

- *Mendelssohn, L. 1993. Preprocessing Data for Neural Networks [online document]. A Complementary Method for Preventing Hidden Neurons' Saturation in Feed Forward Neural Networks Training. Iranian Journal of Electrical and Computer Engineering, Vol. 9 (2), pp. 127-133.*

- *Felipe García Quiroz, Adriana Villa Moreno, Paula Castaño Jaramillo. Línea de Bioinstrumentación, Señales e Imágenes Médicas. Interfaces neuronales y sistemas máquina-cerebro: fundamentos y aplicaciones. Revista Ingeniería Biomédica. ISSN 1909–9762, número 1, mayo 2007, págs. 14-22. Escuela de Ingeniería de Antioquia–Universidad CES, Medellín, Colombia.*

- *L. Deng and D. Yu. Deep Learning: methods and applications. Foundations and Trends in Signal Processing. Vol. 7, Issues 3-4, 2014.*

- *Hastie, Tibshirani y Friedman. The Elements of Statistical Learning. 2011.*

- *Claudio Palma, Wilfredo Palma y Ricardo Pérez. Data Mining. EL arte de anticipar. RIL editores, 2009.*

- *Inteligencia Artificial. Redes Neuronales y Aplicaciones. Hugo Galán Asensio y Alexandra Martínez Bowen. Universidad Carlos III de Madrid.*