



UNIVERSIDAD AUTÓNOMA DE SAN LUIS POTOSÍ

FACULTAD DE INGENIERÍA

CENTRO DE INVESTIGACIÓN Y ESTUDIOS DE POSGRADO

Modelo para el procesamiento y análisis de datos de incidencia delictiva en la zona metropolitana de San Luis Potosí a través de redes neuronales.

T E S I S

Que para obtener el grado de:
Maestro en Planeación Estratégica e Innovación

Presenta:

Rodrigo José Hidrobo Rodríguez

Asesor:

Dra. Liliana Margarita Félix Ávila

San Luis Potosí, S. L. P.

Febrero 2020





UASLP
Universidad Autónoma
de San Luis Potosí



FACULTAD DE
INGENIERÍA

20 de junio de 2019

ING. RODRIGO JOSÉ HIDROBO RODRÍGUEZ
P R E S E N T E

En atención a su solicitud de Temario, presentada por la **Dra. Liliana Margarita Félix Ávila**, Asesor de la Tesis que desarrollará Usted, con el objeto de obtener el Grado de **Maestro en Planeación Estratégica e Innovación**, me es grato comunicarle que en la Sesión del H. Consejo Técnico Consultivo celebrada el día 20 de junio del presente año, fue aprobado el Temario propuesto:

TEMARIO:

"Modelo para el procesamiento y análisis de datos de incidencia delictiva en la zona metropolitana de San Luis Potosí a través de redes neuronales"

Introducción.

1. Estudios relacionados con la aplicación de inteligencia artificial y conceptos de redes neuronales a temas de seguridad e incidencia delictiva.
 2. Unidad de Análisis de Información de una institución de procuración de justicia y productos de inteligencia
 3. Metodología del procesamiento de información para pronosticar incidencia de delitos patrimoniales.
 4. Análisis de datos, pronósticos obtenidos y presentación de resultados.
- Conclusiones.
Referencias.
Anexos

"MODOS ET CUNCTARUM RERUM MENSURAS AUDEBO"

A T E N T A M E N T E

M. I. JORGE ALBERTO PÉREZ GONZÁLEZ
DIRECTOR

UNIVERSIDAD AUTÓNOMA
DE SAN LUIS POTOSÍ
FACULTAD DE INGENIERÍA
DIRECCIÓN

www.uaslp.mx

Copia. Archivo
*etn.

Av. Manuel Nava 8
Zona Universitaria • CP 78290
San Luis Potosí, S.L.P.
tel. (444) 826 2330 al39
fax (444) 826 2336

"2019, aniversario 160 de la fundación del Instituto Científico y Literario de San Luis Potosí"

A mi familia por enseñarme el valor de la unidad.

Por otorgarme la libertad de construirme a mí mismo.

Por ser y permanecer.

Agradecimientos

A mi asesora, Dra. Liliana Margarita Félix Ávila, por su paciencia y tenacidad al reencauzarme en las lides de la metodología científica y permitirme focalizar mis ideas y asentarlas en un proyecto tangible.

A las autoridades de la Fiscalía General de Justicia del Estado de San Luis Potosí que me dieron la oportunidad de ordenar un cúmulo sensible de datos en información sustantiva y por su esfuerzo cotidiano en procurar justicia a la sociedad potosina.

A los maestros del posgrado ya que, sin duda, cada uno de ellos contribuyó a fortalecer mi curiosidad académica.

A mis compañeros de generación por apoyarme en momentos en los que no pensé que pudiera asistirme de ellos; y aquellos que como yo hemos permanecido hasta el último instante al pie del cañón.

Si cada uno limpia su acera, la calle estará limpia.

Goethe

*Aunque nuestra visión hacia adelante es muy corta, podemos darnos cuenta de que hay
mucho por hacer.*

Alan Turing

Resumen

El análisis de la incidencia delictiva tiene como propósito fundamental proporcionar información para la identificación de reincidencias y patrones, y a través de estos generar modelos que permitan detectar y, en su caso, prever el comportamiento específico de los mismos con el fin último de orientar la toma de decisiones para la aplicación de recursos que ayuden a mitigar la comisión de delitos. El modelo propuesto está enfocado a generar pronósticos confiables que permitan elaborar líneas de acción y asignación de recursos para la apropiada prevención de los delitos. El enfoque de aplicación de redes neuronales para la interpretación del comportamiento de la delincuencia, permite, entre otras cosas, la adaptación constante del modelo a través del aprendizaje proporcionado por la información de origen, ajustando su funcionamiento y dinamizando su desempeño.

Palabras clave: Incidencia delictiva; Redes neuronales; Sistemas basados en reglas; Policía; Apoyo en toma de decisiones.

Abstract

The main purpose of crime rate analysis lays in providing information for the recognition of patterns and trends, and through these, generating models that allow detecting and, if possible, foresee a specific behavior with the ultimate goal of orienting decision making in order to apply resources to mitigate commission of crime. The proposed model aims to generate reliable forecasts that allow building lines of action and allocation of resources for an appropriate crime prevention. The approach of applying neural network to interpret crime behavior, allows, among other issues, a constant adaptation of the model through the learning provided by the original data source, adjusting its operation and boosting its performance.

Keywords: Crime rate; Neural networks; Rule based systems; Police; Decision-support

Contenido

Resumen.....	I
Abstract.....	II
Contenido.....	III
Tabla de Figuras.....	IX
Índice de Tablas.....	XII
Introducción.....	1
Antecedentes.....	1
Planteamiento del problema.....	4
Delimitación del problema.....	5
Enunciado formal del problema.....	6
Justificación.....	6
Objetivo General.....	8
Objetivos específicos.....	8
Alcance de la investigación.....	8
Limitaciones de la investigación.....	9
Capítulo 1: Estudios relacionados con la aplicación de inteligencia artificial y conceptos de redes neuronales a temas de seguridad e incidencia delictiva.	10
1.1 Acercamiento de estudio previos a aplicación de inteligencia artificial en el ámbito de impartición de justicia e investigación policial.....	10

1.1.1 Sistema de administración y análisis criminal con inteligencia artificial.	11
1.1.2 Modelo estadístico de comportamiento criminal (Short, y otros, 2008)	16
1.1.3 Analítica del crimen: Análisis delictivo a través de artículos periodísticos (Jayaweera, Sajeewa, Liyanage, Wijewardane, & Perera, 2015)	21
1.1.4 Análisis delictivo y predicción utilizando minería de datos (Shiju & Surya, 2014)	25
1.1.5 Análisis de variación en tendencias delictivas (Flatley & Bradley, 2013)	32
1.1.6 Dinámicas de puntos de concentración de robo y violencia (Chaney y Ratcliff)	36
1.1.7 Hacia la conformación de un modelo propio basado en redes neuronales	40
1.2 Conceptos básicos de redes neuronales	45
1.2.1 Definición	45
1.2.2 Orígenes y desarrollo de las redes neuronales	45
1.2.3 Composición y funcionamiento de las redes neuronales	48
1.2.4 Aspectos característicos y elementos de las redes neuronales	55
1.2.6 Principales topologías de redes neuronales	56

1.2.7 Mecanismos de aprendizaje	60
1.2.8 Propagación de la información en las redes neuronales.....	67
1.2.9 Funciones de transferencia (Activación y salida)	67
1.3 Hipótesis	72
1.4 Variables de relevancia para evaluación como parámetros de impacto en el modelo	72
1.5 Base de la metodología a implementar	73
Capítulo 2: Unidad de Análisis de Información de una institución de procuración de justicia y productos de inteligencia.	75
2.1 La Unidad de Análisis de Información	75
2.1.1 Antecedentes	75
2.1.2 Conceptualización de Unidad de Análisis de Información.....	75
2.2 La inteligencia	76
2.3 Productos de inteligencia.....	78
2.3.1 Parte de novedades	78
2.3.2 Registro de ingreso a separos	79
2.3.3 Matriz de asociación	79
2.3.4 Red de vínculos	79
2.3.5 Red técnica y red de cruces	80
2.3.6 Mapas de incidencia delictiva	80

2.3.7 Informes estadísticos delictivos	80
2.3.8 Informe de inspección y/o análisis	80
2.4 Plan de acción	81
Capítulo 3: Metodología del procesamiento de información para pronosticar incidencia de delitos patrimoniales.....	82
3.1 Introducción	82
3.2 Contexto.....	82
3.2.1 Fuentes de información	83
3.2.2 Sistema de emergencias	85
3.2.3 Instituto Nacional de Estadística Geografía e Informática (INEGI) ..	86
3.2.4 Consejo Nacional de Evaluación Política de Desarrollo Social.....	87
3.2.5 Instituto Nacional de Investigaciones Forestales, Agrícolas y Pecuarias.	88
3.3 Definición de Variables	89
3.4 Población y muestra	91
3.5 Instrumentos	91
3.6 Estructura procedimental.....	92
3.6.1 Metodología general del procesamiento de información de incidencia delictiva a través de RNA.....	92
3.6.2 Modelo estructural de procesamiento.....	94

3.6.3 Procedimientos	95
3.7 Método de investigación	97
3.8 Procesamiento de la información	97
3.8.1 Vector de entrada.....	98
3.8.2 Algoritmo de entrenamiento.....	99
3.9 Análisis de la información	103
3.10 Presentación de la información, técnicas empleadas	103
Capítulo 4: Análisis de datos, pronósticos obtenidos y presentación de resultados	104
4.1 Diseño experimental	104
4.2 Conformación de conjunto de datos y experimentación	113
4.2.1 Condiciones iniciales y preprocesamiento de los datos	113
4.2.2 Experimentación	123
4.2.3 Resultados de la prueba de las configuraciones de red.....	128
4.3 Discusión de resultados.....	142
Capítulo 5: Conclusiones	145
5.1 Conclusiones generales	145
5.2 Futuras líneas de investigación	146
5.3 Reflexiones finales.....	147
Referencias	149

Anexos 155

Tabla de Figuras

Figura 1: Método y sistema de monitoreo para monitoreo basado en reglas de una arquitectura orientada a servicio.	13
Figura 2: Funciones del Sistema de administración y análisis criminal con inteligencia artificial (AICAMS por sus siglas en inglés).....	14
Figura 3: Diagrama de flujo resumiendo las simulaciones discretas.	19
Figura 4: Arquitectura de alto nivel del sistema propuesto.	24
Figura 5: Pasos en el análisis delictivo.	28
Figura 6: Ejemplo de un árbol de decisiones.	30
Figura 7: Relación incidentes reportados en la Encuesta Delictiva para Inglaterra y Gales (por sus siglas en inglés CSEW) y delitos registrados por la policía.	34
Figura 8: Análisis temporal de un clúster, el color gris indica baja incidencia de robos, mientras el amarillo y naranja una incidencia media, siendo el rojo el índice más alto de ocurrencia.	39
Figura 9: Relación de estudios revisados con el modelo a desarrollar.	44
Figura 10: Elementos de una neurona.	50
Figura 11: Esquema de una unidad de proceso típica.....	52
Figura 12: Principio de funcionamiento de una neurona artificial	54
Figura 13: Etapas dentro de una neurona artificial	57
Figura 14: Estructura simplificada de una red en capas.....	58
Figura 15: Estructura extendida de una red multicapa.	59
Figura 16: Función Escalón.....	68
Figura 17: Función lineal-mixta.	69

Figura 18: Función tangente hiperbólica	70
Figura 19: Función sigmoïdal.	71
Figura 20: Función de Gauss.	71
Figura 21: El ciclo de Inteligencia.....	73
Figura 22: Representación del ciclo de inteligencia de acuerdo al CISEN (Centro de Investigación y Seguridad Nacional)	77
Figura 23: Metodología general del procesamiento de información de incidencia delictiva a través de RNA.	93
Figura 24: Modelado neuronal para el procesamiento de datos.....	94
Figura 25: Conformación del vector de datos de entrada para el modelo neuronal.....	98
Figura 26: Interfaz principal de Espacio y Datos de México (INEGI).....	106
Figura 27: Composición de la llamada de la función neuralnet con los argumentos utilizados más frecuentemente.....	111
Figura 28: Estructura de la función compute.	112
Figura 29: Representación gráfica de la base de datos de colonias.	114
Figura 30: Representación gráfica de la base de AGEB's.....	115
Figura 31: Representación gráfica de la conjunción de las bases de datos de colonias y AGEB's.	116
Figura 32: Diagrama de Venn de las variables componentes de la probabilidad de evento.	122
Figura 33:Desempeño de las configuraciones más efectivas.....	132
Figura 34: Visualización de la red neuronal de la configuración 6.....	133
Figura 35: Gráfica comparativa de valores reales vs calculados por la RNA en la configuración 6	134

Figura 36: Gráfica de acercamiento a 30 valores de la muestra procesada	135
Figura 37: Representación gráfica de los valores reales y su distribución según su probabilidad.	136
Figura 38: Representación gráfica de los valores procesados por la RNA y su respectiva distribución según su probabilidad.	137
Figura 39: Comparativa de datos reales con los generados por la red neuronal artificial.	138
Figura 40: Mapa DIP de probabilidades de robo comercial según los datos de la RNA.	140
Figura 41: Puntos de interés de acuerdo a la probabilidad calculada por la RNA	141
Figura 42: Desempeño de configuración 12. Dos capas (6n, 3n) con umbral 0.01.	144

Índice de Tablas

Tabla 1. Lista de parámetros presentes en el modelo discreto.	20
Tabla 2. Tabla de decisión para la ciudad de Delhi.	30
Tabla 3. Volumen y porcentaje de reducción en categorías delictivas comparables, de 2002-2003 a 2011-2012.	34
Tabla 4. Delitos violentos, zona, población, y porcentajes por cada uno de los cinco bloques vecinales del condado de la Ciudad de New York para 2006-2010.	38
Tabla 5. Medición de la pobreza por AGEB's de acuerdo al Grado de Rezago Social, de acuerdo a estudio de CONEVAL para el año 2015	107
Tabla 6. Base de datos de temperaturas de 2011 a 2017 para la zona metropolitana de San Luis Potosí, de acuerdo a registros de INIFAP.....	108
Tabla 7. Lista de argumentos de la función neuralnet	111
Tabla 8. Descripción de argumentos y valores de salida de la función "compute"	112
Tabla 9. Variables de la matriz de Incidencia	118
Tabla 10. Tabla de incidencia incluyendo temperatura	119
Tabla 11. Tabla de variables socioeconómicas	120
Tabla 12. Tiempo de procesamiento para cada configuración de la red.....	128
Tabla 13. Relación de configuraciones de red con valores de desempeño	130

Introducción.

Antecedentes.

La inteligencia es una herramienta estratégica para la toma de decisiones que tiene como objetivo convertir la información en conocimiento operable por medio del análisis (Torres Estrada, 2015).

Dentro de las instituciones relacionadas a la seguridad pública y a la procuración de justicia, el suministro, administración, procesamiento y explotación de la información ha estado sujeta a cambios evolutivos que exigen un marco procedimental para el manejo y generación de productos de alto valor informativo que faciliten la toma de decisiones y el despliegue efectivo de los recursos para completar los cometidos inherentes a la labor investigativa.

El nuevo sistema de justicia penal demanda un mejor manejo de la información y aplicación en el proceso de investigación, tanto en la persecución de los delitos como en su prevención. Debido a esto se han gestado diversos proyectos a lo largo del territorio nacional, con una composición y protocolos de acción diversos y sujetos a las legislaciones locales, pero con la meta común de proveer inteligencia a las instituciones encargadas de la procuración de justicia y de la seguridad pública.

A raíz de estos esfuerzos, una figura institucional ha sido esbozada, la Unidad de Análisis de Información, la cual en sus últimas conceptualizaciones es descrita como “una estructura altamente capacitada, acorde con una dinámica delincencial diversificada que basa su operación en la premisa de descentralización operativa y concentración de la

información, soportada por una base tecnológica que le permite realizar su función” (Torres Estrada, 2015).

En el caso de San Luis Potosí, la Unidad de Análisis de Información de la Fiscalía General funge como una entidad generadora de productos de Inteligencia que guían la aplicación de recursos operativos para mitigar los estragos producidos por el incremento de la incidencia delictiva y su correspondiente nivel de impacto.

Sin embargo, el tratamiento de la información dentro de dicha entidad requiere la incorporación de elementos tecnológicos que optimicen la generación de los productos resultantes y de métodos estructurados aplicables para su adecuado procesamiento.

El propósito de esta tesis reside en el diseño e implementación de un modelo metodológico, sustentado en procedimientos de validación y tratamiento de datos a través de la aplicación de una herramienta tecnológica, como pueden ser los algoritmos de redes neuronales. Y que, a su vez, este modelo proporcione al agente de Inteligencia un mecanismo que facilite su labor en la interpretación de los datos que se obtengan y pueda proporcionar lineamientos de acción para facilitar la toma de decisiones de los rangos jerárquicos superiores.

A finales de agosto del año 2004, el entonces Director General de la Policía Ministerial del Estado de San Luis Potosí, obtiene la autorización para la generación de un órgano interno dentro de la Fiscalía General del Estado encargado del procesamiento de información estadística para la generación de líneas de investigación y definición de estrategias para encarar la problemática delictiva de aquella época, que a la postre resultaría la primera instrumentación en el Estado de una Unidad de Análisis Táctico

dentro de los cuerpos de seguridad pública y procuración de justicia y que sería denominada en su momento como Unidad de Inteligencia.

La conformación de dicha Unidad fue lograda a través de la incorporación de profesionales de diversas ramas y ajenos al quehacer cotidiano de la investigación de delitos buscando un inicio “en blanco” con personal con nivel mínimo de licenciatura y abierto a la adquisición y desarrollo de habilidades para el combate a la delincuencia.

Como punto de partida se reclutó a un grupo de profesionistas que incluía abogados, ingenieros en sistemas y criminólogos que orientados por un grupo de la ahora extinta Agencia Federal de Investigaciones fue encaminado de manera empírica sobre los conceptos del ciclo de inteligencia y sobre la implementación de técnicas para la generación de productos de inteligencia.

Además de estos productos, se constituyó una base estructural organizativa con un encargado al mando y agrupando secciones de atención por naturaleza del delito: contra la vida, patrimoniales, sexuales y contra la paz y seguridad pública.

Algunos de los productos de inteligencia incluyen: redes de vínculos, redes técnicas, matrices delictivas, mapas criminológicos, partes informativos, relojes delictivos, etc.

Sin embargo, la mayor parte de las herramientas mencionadas están limitadas por la capacidad del analista, dado que la mayor parte del procesamiento se hace de manera manual y se generan presentaciones ejecutivas de manera periódica para la exposición del comportamiento de incidencias y eventualidades.

Debido a esto se produce una constante disyuntiva organizacional, que radica de origen en la falta de métodos de operación y asentamiento de los procedimientos de manera plausible, combinándose esto con la dependencia de las órdenes del encargado del área para procesar, generar, almacenar los productos obtenidos derivando esto en la generación de inconsistencias y falta de homogeneidad tanto en la manera en que se procesan los datos como en los productos finales.

La problemática primordial radica en el incremento del acervo informativo y en el aumento de la complejidad del análisis, que exige que el analista reprocese continuamente datos previos a los generados al momento, reduciendo el alcance del análisis únicamente a ese preciso conjunto de datos (y a la capacidad del analista); los datos obtenidos terminan siendo restringidos por la temporalidad y tipo especificado, esto sin ahondar más allá de una tendencia que pudiera no ser percibida por el analista a primera instancia, es decir, la información debe ser procesada nuevamente desde el inicio para obtener los resultados a exponer exclusivamente para un punto específico en el tiempo, reduciendo el panorama general que pudiera estar manifestándose.

Planteamiento del problema.

Los flujos de información de incidencia delictiva son de una densidad considerable en las ciudades grandes y en desarrollo. En San Luis Potosí, esta situación se agudiza con el pasar del tiempo y en el caso de la Fiscalía General, cuenta con la Unidad de Inteligencia para procesar los datos generados de los eventos delictivos. Dicha información se reúne de distintas fuentes y es procesada de manera artesanal desde la fundación de este

órgano. La problemática que enfrentan los analistas que conforman la Unidad es la falta de un mecanismo homogéneo apropiado para generar líneas de acción que ayuden a disminuir la ocurrencia de eventos delictivos

Plantear un eje sólido, basado en un modelo de procesamiento de datos de incidencia delictiva a través del uso de redes neuronales que coadyuve en la mejora de la toma de decisiones, constituye un hito que permitiría formalizar la manera de operación en cuanto al proceso de información estadística y la trascendencia de la interpretación de sus resultados.

Delimitación del problema.

Actualmente, el procesamiento de datos relacionados a incidencia criminal, está limitado a la improvisación de métodos no estandarizados, es decir, acciones acordes al momento y necesidades vigentes, y a la interpretación empírica de los datos observados en la incidencia delictiva en la zona metropolitana de San Luis Potosí por parte de la Unidad de Inteligencia y su impacto en la toma de decisiones a niveles jerárquicos superiores.

Una limitante respecto al procesamiento de los datos es la amplitud de la naturaleza de cada delito, es difícilmente equiparable el comportamiento de los delitos contra la vida y la integridad corporal, a los que van contra la paz y la seguridad de las personas o bien aquellos que corresponden a los delitos contra el patrimonio. En este caso, dada la naturaleza de los delitos patrimoniales, la cantidad de datos accesibles y el comportamiento de los mismos, se opta por los mismos, en particular el robo domiciliario, el robo a comercio y el robo de vehículo, cuyo seguimiento ha sido continuado por las políticas de las distintas administraciones de Gobierno.

La razón primordial para establecer esta delimitación radica en que el procesamiento, presentación e interpretación de la información, influyen directamente en la toma de decisiones y designación de recursos para la implementación de operativos y líneas de acción de mandos medios y superiores.

Enunciado formal del problema.

Existe una falta de especificación de una metodología formal de procesamiento de datos de incidencia delictiva en la zona metropolitana de San Luis Potosí en la Unidad de Análisis Táctico de la Fiscalía General del Estado de San Luis Potosí para el tratamiento de información estadística y la correspondiente toma de decisiones.

Justificación.

La Unidad de Inteligencia de la Fiscalía General del Estado es la responsable de procesar los datos relacionados con la actividad delictiva, ministerial y judicial para generar productos que permitan facilitar la toma de decisiones a los mandos medios y superiores de las entidades de procuración de justicia y seguridad pública con el propósito de garantizar la estabilidad social y sostener el Estado de Derecho de manera adecuada.

En la actualidad, bajo el rubro de análisis de incidencia delictiva, no se cuenta con un proceso automatizado para el manejo de este tipo de información. La mayor parte de los datos son recabados de manera manual, bajo una inspección general a los reportes generados por el sistema de emergencias 911, a partir del cual se filtran los datos corroborados por personal operativos como genuinos y estos a su vez son compilados

por analistas que los procesan de manera tal que se compilan las incidencias semanales en hojas de cálculo Microsoft Excel© y se desarrolla una descripción cartográfica en mapas plasmados en la aplicación Microsoft Visio© o Microsoft PowerPoint©.

La implementación de un modelo estructurado y ajeno a prácticas obsoletas, basado en la recolección, administración, análisis, procesamiento y presentación lógica y racional de los datos desprendidos de la ocurrencia de eventos delictivos, permitirá la automatización de las tareas descritas, reduciendo los tiempos de acceso y mejorando la generación de reportes de incidencia y aplicación de técnicas estadísticas para apoyo en la toma de decisiones permitiendo establecer métricas y esquemas de evaluación sobre el modelo mismo.

El producto a obtener consiste en un modelo general para el procesamiento y análisis de datos de incidencia delictiva a través de redes neuronales, lo cual repercutirá de fondo sobre la fundamentación en la toma de decisiones de los mandos medio y superiores de las autoridades ministeriales, orientando las actividades del personal y permitiendo construir una base de conocimiento que procure una apropiada proyección de las tendencias delictivas para poder establecer líneas de acción en la aplicación de recursos y fortalecer la seguridad pública en la zona metropolitana de la ciudad de San Luis Potosí.

El propósito final consiste en la generación de un modelo estructurado y formal que funja como un eje sólido en la producción de información válida obtenida de un proceso lógico de análisis que constituya la base en la toma de decisiones en cuanto a la aplicación de líneas de acción.

Objetivo General.

El objetivo general establecido para este trabajo es el siguiente:

Generar un modelo formal y estructurado para el tratamiento integral de la información de incidencia delictiva; desde su recolección, procesamiento y presentación como producto de inteligencia, utilizando algoritmos de redes neuronales como eje del procesamiento de la información.

Objetivos específicos.

Para el cumplimiento del objetivo general, se establecen los siguientes objetivos específicos:

- Establecer un mecanismo procedimental para el procesamiento de información.
- Alinear las actividades de los analistas de la Unidad de Análisis Táctico a un proceso lógico-racional.
- Identificar las variables de mayor relevancia dentro de la dinámica de la incidencia delictiva.
- Generar pronósticos del comportamiento delincuencia con base en las tendencias y patrones detectados.

Alcance de la investigación.

El resultado de la presente tesis está enfocado en primera instancia a la formalización del tratamiento de información y a la automatización de los procesos estadísticos de la Unidad de Análisis Táctico de la Fiscalía General en cuanto a incidencia delictiva se

refiere, fortaleciendo, a través de datos procesados sistemáticamente y cifras pronosticadas sólidas, la toma de decisiones en los mandos y buscando mejorar la expectativa de seguridad pública que atañe a la sociedad potosina.

Limitaciones de la investigación.

El planteamiento de este trabajo se limita principalmente a aspectos formales para el establecimiento de una manera estructurada para el procesamiento de datos basada en algoritmos de redes neuronales, con el objeto de obtener pronósticos consistentes de la ocurrencia de eventos delictivos a corto plazo en la zona metropolitana de San Luis Potosí, teniendo como base los registros de incidencia criminal de los años 2014 a 2016 sobre delitos patrimoniales consistentes en robo de vehículo, robo domiciliario y robo a comercio, considerando la cifra de denuncias reportadas ante la Fiscalía General del Estado de San Luis Potosí y descartando por su informalidad los reportes al Sistema de Emergencias. Se omite la extensión de aplicación a otro tipo de delitos como pueden ser los homicidios, lesiones, secuestros, extorsiones o fraudes, dado que en particular los robos tienden a acotar de manera precisa el espacio geográfico y temporal, situación que no ocurre de manera concisa en los otros tipos de evento y su naturaleza suele ser de origen diverso.

Capítulo 1: Estudios relacionados con la aplicación de inteligencia artificial y conceptos de redes neuronales a temas de seguridad e incidencia delictiva.

1.1 Acercamiento de estudio previos a aplicación de inteligencia artificial en el ámbito de impartición de justicia e investigación policial.

En este apartado del documento se analiza un conjunto de propuestas previas que abordan problemáticas similares a la del presente estudio. En su mayoría son artículos de investigación derivados de estudios que abordan la problemática criminal desde distintos enfoques y con distintos medios para el análisis de la información, la consideración de estas propuestas consiste en la contribución que pudieran aportar al presente trabajo, desde los métodos para la recopilación de datos y su tratamiento, como su análisis y métodos para afrontar la problemática particular que cada uno de ellos procura solventar.

El primero de los casos cuya revisión es de interés, consiste en la integración de sistemas expertos en la labor policial; un sistema experto es un campo de investigación de la inteligencia artificial que intenta codificar el conocimiento y los procesos de razonamiento de un experto humano dentro de un sistema computacional (Jackson, 1998); para la concreción de un aplicativo que permita la replicación de los procesos de deducción concebidos por los investigadores a través del tiempo, esto combinado con el uso de bases de datos construidas desde distintas fuentes de información.

1.1.1 Sistema de administración y análisis criminal con inteligencia artificial.

Es un esquema desarrollado por Brahan, Lam, Hilton y Leung en 1998, en el cual los sistemas expertos y las tecnologías basadas en inteligencia artificial tienen un rol significativo en el desarrollo de herramientas de soporte para las operaciones de las instituciones policiales.

Problemática: Las operaciones policiales se basan en el uso intensivo de información. Los datos son reunidos de una gran variedad de fuentes para después ser comparados con datos adquiridos con antelación, analizados y tomar las acciones apropiadas, para finalmente todo este conjunto sea almacenado como futura referencia. La pericia necesaria para interpretar los datos recae en el investigador y la eficiencia del uso de las fuentes de datos disponibles depende sobre todo de su habilidad y experiencia, así como en la facilidad de acceso a las distintas bases de datos que pudiera ser requerida.

Solución propuesta:

Brahan, Lam, Hilton, y Leung, destacan la relevancia que ha adquirido el procesamiento a través de computadoras y en específico sobre la aplicación de tecnología basada en conocimiento, a través de la cual es posible capturar la pericia de un investigador experimentado y a su vez facilitar el acceso al usuario a las bases de datos disponibles. De esta manera, se puede construir un repositorio de técnicas y habilidades al paso del tiempo y ponerse a disposición de todos los investigadores en la unidad. Así, cuando un individuo parte, sus habilidades no se marchan con él.

El estudio busca satisfacer dos requerimientos: en primer lugar, debe cubrir la necesidad operacional de la policía y en segundo lugar cumplir con la calificación como proyecto apropiado para la universidad. Por experiencia previa, los investigadores asumen que un sistema basado en reglas tendría una posibilidad razonable de alcanzar éxito. Sin embargo, también era viable realizar mejoras mediante actualizaciones al sistema a través de componentes de aprendizaje.

La propuesta se enfocó en crímenes con volumen alto de incidencia, utilizando la experiencia del investigador para desarrollar una regla inicial de base seguida por una investigación de máquinas de aprendizaje y tecnologías de redes neuronales que permitieran al sistema evolucionar a través del aprendizaje por experiencia.

Integración del modelo que esquematiza la solución del problema:

En la Figura 1: Método y sistema de monitoreo para monitoreo basado en reglas de una arquitectura orientada a servicio. Se plantea de manera general el comportamiento de un sistema general basado en reglas.

El mecanismo representado señala una etapa de aprendizaje en la que un arreglo de variables de conjunto de datos alimenta al sistema proporcionando información de la relación de dichas variables, repitiéndose el proceso hasta alcanzar un desempeño satisfactorio aproximado al comportamiento real, esto último evaluado contra la regla propuesta durante la fase de pruebas, que, una vez superada, valida la regla preestablecida.

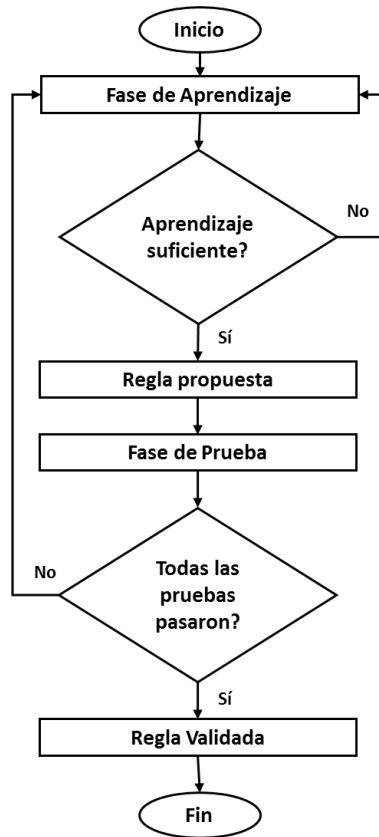


Figura 1: Método y sistema de monitoreo para monitoreo basado en reglas de una arquitectura orientada a servicio.

Fuente: Elaboración propia basada en el diagrama de flujo de Patentdocs, website.

Los responsables del estudio proyectan la idea general mostrada en la Figura 1, aplicada como núcleo de un sistema que contempla mayor cantidad de elementos como datos de investigación, bases de datos de modus operandi, investigador a cargo, información de testigos, incidentes relacionados, eventos sospechosos próximos, presuntos responsables, construcción de retratos, etc. Todo esto integrado en un sistema denominado Sistema de Administración y Análisis Criminal con Inteligencia Artificial (SAACIA) cuya integración de módulos se muestra en la Figura 2.

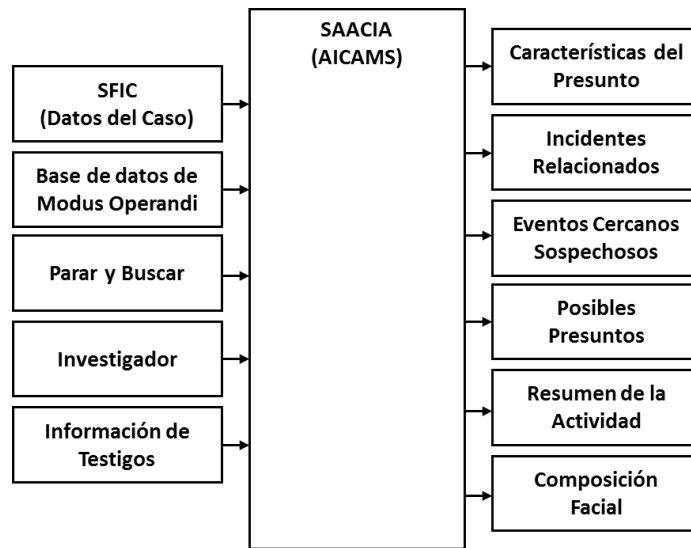


Figura 2: Funciones del Sistema de administración y análisis criminal con inteligencia artificial (AICAMS por sus siglas en inglés).

Fuente: Elaboración propia basada en el diagrama de John W. Brahan, Kai P. Lamb, Hilton Chan, William Leung.

Los componentes integrados en el sistema fueron montados con base en el análisis de casos reales para la adquisición de conocimiento, obteniendo aportación de información trascendente desde la base de modus operandi dependiente de la naturaleza del delito, así como de la tasa de incidencia delictiva. Por otra parte, sostuvieron sesiones de tormenta de ideas con oficiales de la fuerza policial para determinar parámetros relevantes de los casos y que deberían ser registrados para describir los detalles de los incidentes, sesiones posteriores ayudaron a refinar la importancia relativa de cada parámetro y a identificar la relación entre los mismos y las características de los agresores obteniendo una regla base que permitiera la descripción del agresor basada en la información reunida en una escena del crimen.

De igual manera, llevaron a cabo una revisión de las bases de datos, incluyendo la base de casos para ratificar los parámetros identificados en sesiones frente a frente con los

oficiales y obtener parámetros adicionales que se agregarían posteriormente al sistema; el sistema de modus operandi, que incluye una lista extensa de códigos que describen ampliamente la mayoría de los delitos y que añade parámetros para definir detalles tales como delito, horario, lugar, presunto, víctima, propiedad involucrada, acciones durante la comisión del delito; y también la base de “parar y buscar”, la cual no es una fuente de información de casos específicos, pero provee información relacionada al presunto, dado que registra información de incidentes donde la gente ha sido detenida por comportamiento sospechoso.

El proceso de la información queda delimitado por cuatro criterios principales (Brahan, Lam, Hilton, & Leung, 1998) :

1. El sistema debe responder a una necesidad identificada por el usuario.
2. Debe haber una posibilidad de demostración significativa de la efectividad del sistema.
3. El volumen de incidentes debe ser suficientemente alto como para permitir la identificación de patrones que puedan ser usados para identificar sospechosos a través del modus operandi.
4. Los datos deben estar disponibles para el sistema lo suficiente para proporcionar un vínculo sobre el presunto responsable.

Relación y posibles aportaciones del documento al modelo propuesto.

El estudio realizado por los investigadores de la aplicación de un sistema basado en reglas, sirve de base para el presente estudio, ya que sienta los elementos del uso de tecnología de redes neuronales y algoritmos de aprendizaje para la interpretación del

comportamiento de incidencias de delitos de alto volumen, así como la detección de parámetros que constituyen el conjunto de variables de impacto en la comisión de delitos, pero enfocado a la resolución de investigaciones y no al comportamiento de la propia incidencia delictiva.

1.1.2 Modelo estadístico de comportamiento criminal (Short, y otros, 2008)

El modelo desarrollado por Martin Short y otros, de la Universidad de California Los Angeles, fue motivado por observaciones de clústeres espacio-temporales de crímenes en una variedad de conjuntos urbanos; los investigadores presentan un modelo para estudiar la urgencia, dinámica y propiedades de estado de estabilidad de puntos de concentración de incidencia delictiva.

Problemática:

Un aspecto desafortunado de la vida moderna es la presencia de crimen en áreas urbanas grandes. Sin embargo, aunque el crimen posee ubicuidad, no aparece distribuido de manera uniforme en espacio y tiempo. Mientras algunas colonias tienden a ser razonablemente seguras, otras parecen mucho más peligrosas y despliegan clústeres densos de delitos patrimoniales y hechos violentos. Asimismo, correlaciones temporales entre delitos han sido documentados, a menudo con afectados o vecinos cercanos siendo repetidamente objetivo de la delincuencia en periodos cortos de tiempo (Short, y otros, 2008).

Los conjuntos espacio-temporales de ocurrencia delictiva son comúnmente denominados como puntos de concentración (hotspots), y gracias a los avances recientes en tecnología

de georreferenciación es posible rastrear su evolución en escalas espacio-temporales sutiles. Los tiempos de vida y escalas de duración de puntos de concentración criminal pueden ser observados dependiendo de las condiciones particulares presentes ya sean geográficas, económicas o de temporada. Y dependiendo de la categoría del delito, los puntos de concentración emergen, se desvanecen o desaparecen de tal manera que sugieren una estructurada, aunque compleja, dinámica subyacente.

A grandes rasgos, los investigadores guardan bajo consideración que, los delitos toman lugar una vez que un delincuente encuentra una persona como víctima idónea bajo la ausencia de medidas de seguridad efectivas. Bajo este contexto, la estructura del ambiente urbano juega un rol destacado acotando el campo de acción del agresor, sobresalen variables como el nivel de tráfico, propiedades abandonadas, densidad de la población y distribución de los patrones delictivos generadores de impacto criminal.

Solución propuesta:

El principal propósito de la investigación consiste en la generación de un modelo matemático cuantitativo que reúna las dinámicas básicas de los puntos de concentración delictiva. Para tal efecto, los investigadores se basan en un modelo matricial de dos dimensiones para robo domiciliario, donde cada sitio se caracteriza por una variable dinámica de atracción y donde cada delincuente es representado como un paseante aleatorio. Las dinámicas de delincuentes y del campo de atracción son acopladas la una a la otra a través de mecanismos de tendencia y retroalimentación. Dependiendo de la selección de parámetros, se observan y describen muchos regímenes de agregación, incluyendo puntos de concentración de alta actividad delictiva. Sobre la base del sistema

discreto se deriva a un modelo continuo; ambos en concordancia para grandes tamaños de sistemas. Por medio de un análisis de estabilidad lineal se pueden determinar los valores de los parámetros que conduzcan a la creación de puntos de concentración estables.

Integración del modelo que esquematiza la solución del problema:

Los investigadores parten de un modelo discreto, el cual está compuesto por dos elementos, los hogares afectados por robo domiciliario y los delincuentes que cometen los robos en esos domicilios. Los hogares son entonces representados en una matriz bidimensional con un espaciado de celda constante ℓ y condiciones periódicas del entorno, en conjunto a lo anterior se elige también una unidad discreta de tiempo δt en la que el acto criminal habrá de ocurrir. Cada hogar es descrito por su lugar en la celda $s=(i,j)$ y una cantidad $A_s(t)$, la cual señalará el atractivo del lugar, esto es modelado como equivalente a la tasa estadística de robo domiciliario en el lugar s cuando un robo de este tipo ocurre. Los investigadores no hacen hincapié en otras características subyacentes de las residencias tales como el valor, la seguridad o la ubicación. En su lugar, el atractivo es tratado bajo el espíritu del comportamiento colectivo, modelándolo en relación al fenómeno sociológico de la victimización repetida o de proximidad de ocurrencia y la teoría de las ventanas rotas, consistente en la idea de mantener los entornos urbanos en condiciones buenas para disminuir la incidencia de actos criminales (Short, y otros, 2008).

Con base en lo anterior, los autores del estudio obtienen:

$$A_s(t) = A_s^0 + B_s(t) \quad \text{Ec. (1)}$$

Donde A_s^0 representa un componente de atractivo estático, aunque posiblemente variable respecto al espacio, y $B_s(t)$ representa el componente dinámico asociado a la repetición de la victimización o proximidad de ocurrencia.

En el modelo construido, los agentes criminales pueden llevar a cabo dos acciones durante el intervalo de tiempo simulado, ya sea robar en el domicilio en el que se encuentran o bien trasladarse a un domicilio vecino. En caso de que lleven a cabo el robo, se asume que los delincuentes abandonan el vecindario con los objetos robados y permanecer inactivos durante cierto tiempo. Si deciden no robar, los delincuentes se desplazan a otro punto vecino dentro de la matriz.

El modelo es complementado con fórmulas subsecuentes para determinar la probabilidad de ocurrencia hasta derivar en un modelo continuo cuya verificación es documentada con los datos obtenidos a través de la aplicación de simulación por computadora. Los elementos involucrados y la interacción descrita con antelación se muestran en la figura

3.

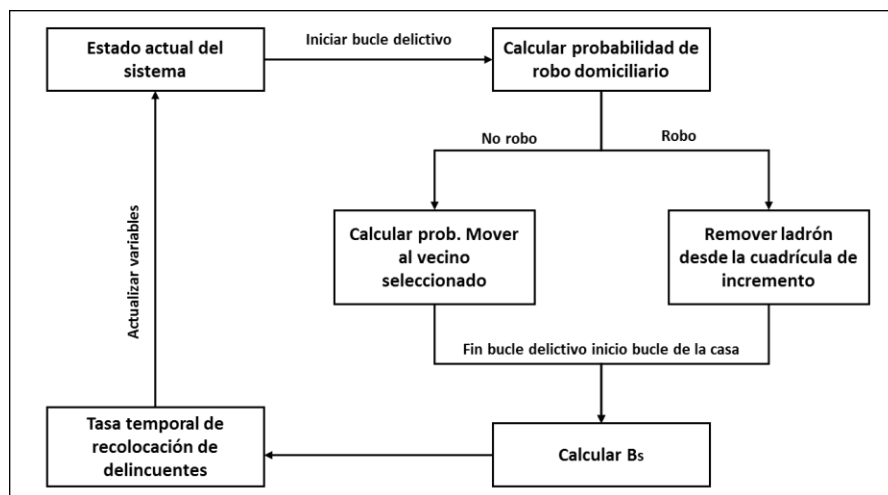


Figura 3: Diagrama de flujo resumiendo las simulaciones discretas.

Fuente: M. B. Short, M. R. D'Oorsogna†, V. B. Pasour.

Aunado a la representación de flujos del sistema, Short, destaca las variables principales en su modelo derivadas de la problemática planteada, el supuesto del comportamiento de los componentes y los estudios de la representación estadística para la exploración y formulación del mismo modelo, dichas variables son denominadas parámetros del modelo discreto y se muestran en la Tabla 1.

Relación y posibles aportaciones del documento al modelo propuesto.

El establecimiento de un modelo discreto, transportado a un modelo continuo a través de la simulación computarizada con la modificación de parámetros permite representar de manera ordenada el comportamiento de la incidencia de un delito en particular, es una base fundamental que se busca a través del estudio planteado, si bien el modelo está basado en un esquema matricial contribuye en gran manera a la identificación de variables que pudieran ser de utilidad durante el desarrollo del presente estudio.

Tabla 1. Lista de parámetros presentes en el modelo discreto.

Nombre parámetro	
l	Espaciado de la cuadrícula
δt	Paso de tiempo
Ω	Tasa dinámica de la disminución del atractivo
η	Medición del efecto del vecindario (Desde 0 hasta 1)
θ	Aumento del atractivo debido a un evento de robo
A_s^0	Atractivo intrínseco del sitio s
Γ	Tasa de generación de ladrones en cada sitio

Fuente: M. B. Short, M. R. D'Orsogna†, V. B. Pasour.

1.1.3 Analítica del crimen: Análisis delictivo a través de artículos periodísticos (Jayaweera, Sajeewa, Liyanage, Wijewardane, & Perera, 2015)

El análisis delictivo es una de las actividades más importantes de la mayor parte de las organizaciones de inteligencia y de procuración de justicia alrededor del mundo. De manera general, se colecta la información regional y foránea relacionada a la criminalidad para prevenir ataques futuros y aplicar un número limitado de recursos de la manera óptima (Jayaweera, Sajeewa, Liyanage, Wijewardane, & Perera, 2015).

Problemática:

El mayor reto encarado por la mayoría de las organizaciones de justicia e inteligencia es el análisis eficiente y preciso de los volúmenes en continuo crecimiento de datos relacionados a la delincuencia. La gran diversidad geográfica y la complejidad de patrones delictivos han hecho más difícil el análisis y registro de datos criminales.

A diferencia de países del primer mundo como Reino Unido o Estados Unidos, Sri Lanka aún utiliza registros manuales de los delitos y un sistema que aporta muy poco soporte para la toma de decisiones.

Solución propuesta:

Los investigadores resaltan el uso de la minería de datos como una herramienta potente que puede ser usada de manera efectiva para analizar grandes bases de datos y obtener resultados analíticos importantes. Se propone una aplicación basada en web compuesta

de las técnicas de análisis delictivo, tales como la detección de puntos de concentración delictiva, comparación de delitos y visualización de patrones criminales.

El sistema propuesto consiste en siete grandes componentes. Un rastreador web, que es una aplicación que sistemáticamente navega una página basada en web para crear un índice de datos, clasificación de información, un extractor de entidades, un detector de duplicados, un manejador de base de datos, un analizador y una interfaz gráfica de usuario basada en interfaz web.

Integración del modelo que esquematiza la solución del problema:

La principal responsabilidad del rastreador web es sondear artículos de noticias de un diario dado. El contenido requerido de los artículos rastreados es almacenado en una base de datos para procesamiento posterior.

La responsabilidad del clasificador es categorizar los artículos rastreados del diario como criminales o no criminales. Los artículos clasificados serán almacenados en una base de datos para la extracción de entidades. Los documentos deben ser transformados antes de usarlos en el proceso de clasificación.

El módulo de extracción de entidades obtiene las entidades de los artículos del diario ya clasificados. Por cada artículo delictivo se obtienen datos tales como fecha del delito, ubicación, policía, corte, cuenta de afectados, etc.

Se utiliza la Arquitectura General para Ingeniería de Textos (GATE por sus siglas en inglés) para lograr este cometido.

El módulo de detección de duplicados tiene como propósito identificar duplicados próximos o exactos de los artículos de los diarios y quitarlos de la base de datos.

El manejador de base de datos administra todas las transacciones de información.

El analizador desempeña operaciones de análisis delictivo en los artículos sobre crimen ya procesados, aplica sobre los mismos los siguientes pasos:

- Detección de puntos de concentración delictiva.
- Comparación de delitos.
- Visualización de patrones delictivos.

La interfaz gráfica de usuario basada en web es utilizada para presentar detalles estadísticos de años previos y resultados de los análisis de la información. Incluye herramientas de graficación y mapas.

El análisis de la información consiste en la obtención de artículos noticiosos de páginas web de diarios electrónicos específicos, posteriormente se evalúan medidas de precisión de la clasificación de la información, como son: exactitud, sensibilidad, especificidad y una medida de precisión estadística. Después se lleva a cabo la extracción de elementos distintivos del evento tales como ubicación del delito, fecha del delito y tipo de delito y denominados entidades; siendo estos los parámetros o variables que, una vez evaluados y obtenidos sus valores de precisión, que el sistema consume para proveer sus resultados. Posterior a la obtención de los datos, la extracción de entidades y el descarte de duplicados los resultados procesados tienen valores aceptables de precisión y comprensión. Esta información es utilizada para alimentar una herramienta de puntos

destacados que agrupan la mayor cantidad de incidencias identificadas en una zona específica.

La arquitectura general del sistema, mostrado en la figura 4, involucra los siete componentes mayores del sistema:

1. Rastreador.
2. Clasificador.
3. Extractor de entidades.
4. Detector de duplicados.
5. Manejador de base de datos.
6. Analizador.
7. Interfaz gráfica de usuario.

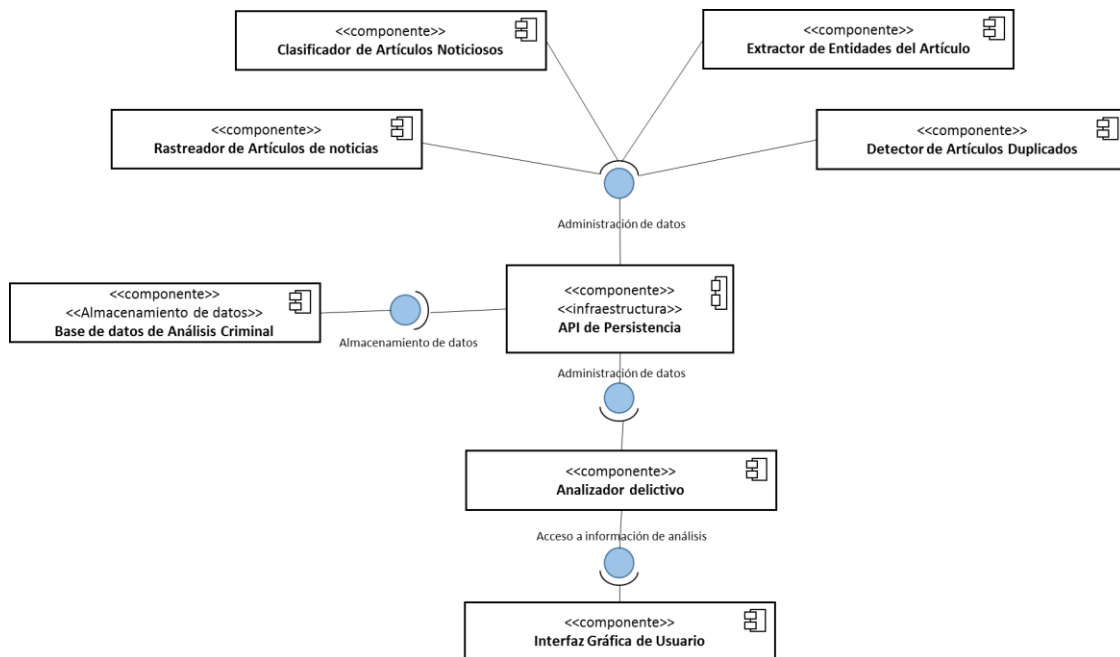


Figura 4: Arquitectura de alto nivel del sistema propuesto.

Fuente: Elaboración propia basada en el diagrama de Isuru Jayaweera, Chamath Sajeewa, Sampath Liyanage, Tharindu Wijewardane, Indika Perera.

Relación y posibles aportaciones del documento al modelo propuesto.

El estudio propone una serie de componentes que pueden contribuir a través de herramientas automatizadas al tratamiento de la información, procurando la optimización de recopilación y organización de la información a través de mecanismos previamente establecidos.

La contribución de un esquema como el presentado por estos investigadores radica en la modularización de procesos que afectan el tratamiento de información y el respectivo valor que se puede atribuir a la estructuración de los procedimientos, así como la manera indirecta de tratar el manejo de información sin acceso a fuentes oficiales.

1.1.4 Análisis delictivo y predicción utilizando minería de datos (Shiju & Surya, 2014)

El análisis y la prevención criminal utilizan un enfoque sistemático para la identificación y análisis de patrones y tendencias de la delincuencia. Con el advenimiento de los sistemas computacionales, los analistas de datos delincuenciales pueden ayudar a los oficiales del orden y la justicia a acelerar el proceso de resolución de eventos delictivos (Shiju & Surya, 2014).

Problemática:

El incremento de la incidencia delictiva alrededor del mundo es innegable, el delito no se puede predecir dado que no es ni sistemático, pero tampoco aleatorio. De igual manera, las tecnologías modernas y sus métodos ayudan al delincuente a lograr completar sus

fechorías. Los investigadores establecen que, aunque no es posible aún predecir quien será víctima de un delito, sí se puede inferir una zona de alta probabilidad de ocurrencia.

Las principales limitantes que los investigadores debían encarar son:

- Incremento de información delictiva que debe ser almacenada y analizada.
- Dificultad en el análisis de datos cuando la información es incompleta e inconsistente.
- Limitantes en la obtención de registros delictivos por parte de autoridades policiales.
- La precisión del programa depende de la precisión del set de datos ingresados.

Descubrir los patrones y tendencias de la incidencia delictiva es un factor desafiante. Para lograr la identificación de un patrón, los analistas delictivos requieren de una gran cantidad de tiempo, rastrear a través de los datos para encontrar si un delito en particular encaja dentro de un patrón conocido. Si no encaja dentro de un patrón existente, entonces los datos deben clasificarse en un nuevo patrón. Después de haber detectado un patrón, éste puede ser usado para predecir, anticipar y prevenir el delito.

Solución propuesta:

Aunque los resultados no pueden ser asegurados con una precisión al ciento por ciento, sí pueden ayudar a reducir el índice de criminalidad en cierta medida, ajustando el nivel de seguridad preventiva en áreas sensibles. Para tal efecto, es necesario recolectar registros criminales y evaluarlos.

Para ser capaces de detectar nuevos patrones en el futuro, se optó por técnicas de procesamiento por conjuntos o conglomerados.

En la manera habitual de investigación, elementos descriptivos del presunto responsable son recabados a través de entrevistas, en un sitio de ocurrencia determinado indicando características fisionómicas del delincuente, en otro lugar, la edad aproximada del responsable, en un tercer sitio que el presunto lleva tatuajes o señas particulares características, estos elementos individuales ofrecen una descripción poco homogénea del delincuente. Los reportes criminales son procesados manualmente y dan la pauta para establecer características de delitos resueltos, lo cual tampoco permite proporcionar una buena calidad respecto a la predicción, incluso la naturaleza del delito va cambiando conforme pasa el tiempo, es por esto que para detectar patrones nuevos o desconocidos en el futuro las técnicas de procesamiento por conjunto funcionan mejor.

Los pasos en la realización de análisis delictivo para este estudio son:

1. Recolección de datos.
2. Clasificación.
3. Identificación de patrones.
4. Predicción.
5. Visualización.

Por lo que los módulos fueron establecidos en base a estas etapas. La integración del modelo que esquematiza la solución del problema se representa en la figura 5:

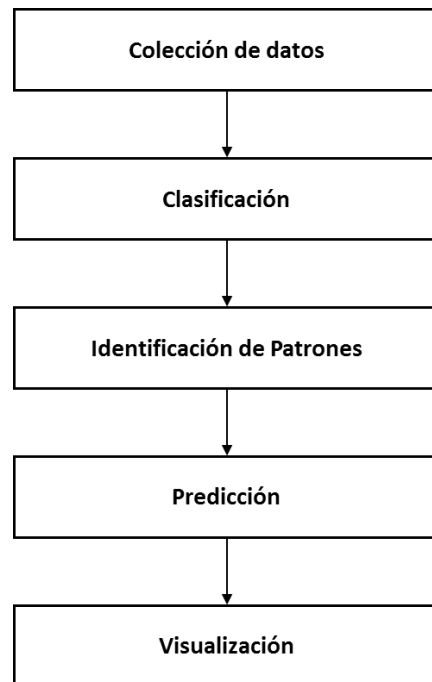


Figura 5: Pasos en el análisis delictivo.

Fuente: Elaboración propia basado en el diagrama de Shiju Sathyadevan, Surya Gangadharan.

- A. Recolección de datos. Se obtuvo información de distintas fuentes de datos, la cual sería almacenada en una base de datos para su posterior procesamiento.
- B. Clasificación. Para la clasificación de la información se utiliza un algoritmo denominado Bayesiano ingenuo, el cual es un método de aprendizaje supervisado al igual que un método estadístico de clasificación. El uso de este algoritmo permite la generación de un modelo mediante datos delictivos de entrenamiento relacionados a vandalismo, robo, robo domiciliario, abuso sexual, incendiarios, robo a mano armada, robo en carretera, etc. Por datos de entrenamiento se entiende que debemos enseñarles sobre entradas particulares de manera tal que podamos probarlos para entradas desconocidas.

C. Identificación de patrones. La tercera fase consiste en la identificación de tendencias y patrones delictivos. Para lograrlo, los investigadores aplicaron un algoritmo *A priori* (utilizado en minería de datos para encontrar Reglas de asociación en un conjunto de datos), éste se usa para determinar asociación de reglas las cuales resaltan tendencias generales en la base de datos, el resultado de la fase es el patrón delictivo para un lugar en particular.

D. Predicción. Para el proceso de predicción se aplica un esquema de probabilidades basado en un árbol de decisiones. La ventaja del uso del árbol de decisiones es que es fácil de interpretar. A grandes rasgos, el árbol posee tres tipos de elementos:

- Un nodo raíz, que tiene aristas de entrada y cero o más aristas de salida.
- Nodos internos, cada uno de los cuales tiene 1 arista de entrada y 2 o más de salida.
- Nodo hoja o nodo final, cada uno de los cuales tiene exactamente 1 arista de entrada y ninguna de salida.

La figura 6 muestra el ejemplo de un modelo de árbol de decisión, el cual es una técnica que permite establecer un análisis de decisiones con base en la utilización de resultados y probabilidades asociadas a los mismos, puede ser usado para conformar un algoritmo que de manera anticipada determine matemáticamente la mejor opción.

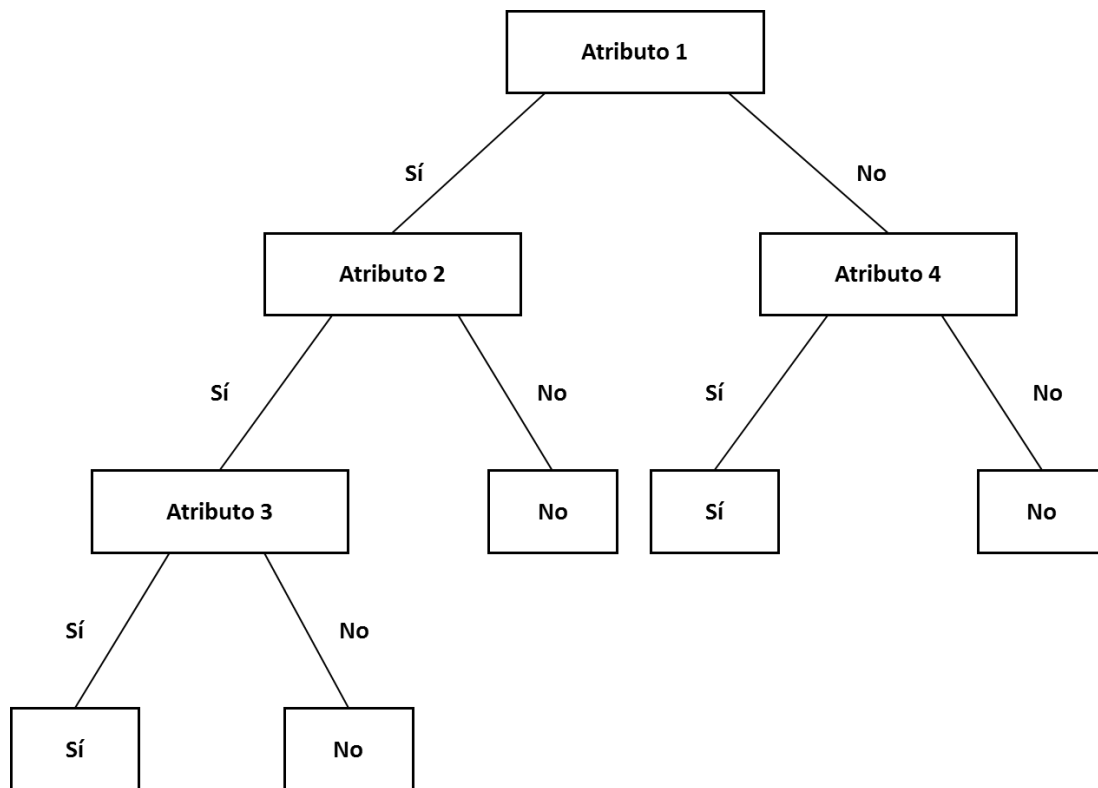


Figura 6: Ejemplo de un árbol de decisiones.

Fuente: Elaboración propia basada en el diagrama de Shiju Sathyadevan, Surya Gangadharan.

La tabla 2 es una representación de una tabla de decisión, utilizada por el estudio (Shiju & Surya), para establecer las condiciones de determinadas zonas de la ciudad de Delhi y con dichos datos construir el árbol de decisión correspondiente.

Tabla 2. Tabla de decisión para la ciudad de Delhi.

Área de Sensitividad	Evento Notable	Presencia VIP	Grupo Criminal	Delito
Sí	Sí	Sí	No	Sí
Sí	Sí	No	Sí	No
No	No	No	Sí	No
Sí	No	No	No	No
Sí	Sí	Sí	Sí	Sí
No	Sí	No	No	No

Fuente: Elaboración propia con datos de Shiju Sathyadevan, Surya Gangadhara.

Esta técnica de máquina de aprendizaje supervisado construye un árbol de decisión de un conjunto de muestras de capacitación etiquetadas por clase y con base en este árbol prueba muestras nuevas. Es un modelo predictivo el cual usa un conjunto de reglas binarias para calcular el valor de la clase. El árbol determina:

- Qué variable es viable para dividir un nodo.
- Decisión de parar o dividir.
- Asignación de nodos terminales

E. Visualización. Las zonas propensas a la delincuencia pueden representarse gráficamente utilizando un mapa térmico que indica el nivel de actividad. Por lo general se usan los colores más oscuros para indicar una baja actividad y colores más brillantes para indicar una alta actividad.

Relación y posibles aportaciones del documento al modelo propuesto.

El modelo aplicado en este estudio esquematiza de manera modular la forma de abordar la problemática y, además, resalta el uso de distintos algoritmos para cada etapa, formalizando de una manera integral la solución propuesta. Por otra parte, la aplicación de árboles de decisión fortalece la idea de recurrir a herramientas que usan el aprendizaje para la automatización de conclusiones que permitan llevar a cabo un proceso de toma de decisiones mucho más oportuno.

1.1.5 Análisis de variación en tendencias delictivas (Flatley & Bradley, 2013)

El uso apropiado de estadísticas delictivas y policiales es fundamental si se desea sostener la confianza pública en dichas estadísticas, en la policía y mandos comisionados, y en el servicio policial en general, de la misma manera, si se presenta un uso inapropiado o engañoso puede mermar dicha confianza (Flatley & Bradley, 2013).

Problemática:

La Estadística Nacional sobre Criminalidad para Inglaterra y Gales (EDIG) está basada principalmente en dos fuentes: la Encuesta Criminal para Inglaterra y Gales y la serie de registros delictivos de la policía que se compila de datos proporcionados por la misma policía derivados de los registros generados que recaen en la categoría de denuncia obligatoria o de persecución de oficio.

Una de las razones principales en llevar a cabo la encuesta en el año de 1982 fue la de evaluar la magnitud de la brecha entre los delitos registrados por la autoridad y la percibida por la población residente.

Un hallazgo de la encuesta ha sido la proporción sustancial de delitos sufridos por el público y que no se reportan a la policía, siendo en el periodo 2011-2012 de alrededor del 61%. La variación en las tasas de denuncia refleja parcialmente la seriedad de la ofensa percibida por la víctima. También se asume que se debe a cuestiones prácticas como puede ser obtener un documento referencial de la autoridad para hacer válida una queja al seguro. (Flatley & Bradley, 2013)

Solución propuesta:

No es posible comparar los microdatos obtenidos de la encuesta con los registros policiales de incidentes delictivos, dado que no es posible determinar si un delito manifestado en la encuesta en realidad aparece en alguna bitácora policial, y si lo fue, identificar cómo fue registrado.

Ambas herramientas cubren diferentes poblaciones y diferentes delitos, por lo que dificultan la realización de las comparaciones, para mejorarlas, los analistas se han centrado en un subconjunto de delitos comparables que están cubiertos por las dos series. Este subconjunto comparable se obtiene al hacer varios ajustes en las categorías de delitos registrados para maximizar la compatibilidad con la encuesta. Sin embargo, en las categorías en las que no es posible distinguir las víctimas que residen en hogares de otras víctimas, no están ajustadas para excluir entidades comerciales, delitos cometidos contra los menores y visitantes extranjeros. Por lo tanto, se debe reconocer que esta serie comparable permanece de manera más amplia que la directa entre herramientas y que el sistema de clasificación de delitos utilizado en la encuesta sólo puede aproximarse a aquella utilizada por la policía.

Integración del modelo que esquematiza la solución del problema:

El modelo comparativo se representa en la figura 7, donde se puede observar que la integración comparativa no es posible debido a la irregularidad sino hasta el periodo 2002/03. La Tabla 3 muestra una diferencia entre los delitos registrados por la policía, la cual muestra una reducción del 41% y la encuesta solo refleja una apreciación de 26% de reducción. Diferencia impulsada por las tendencias de los últimos 5 años evaluados.

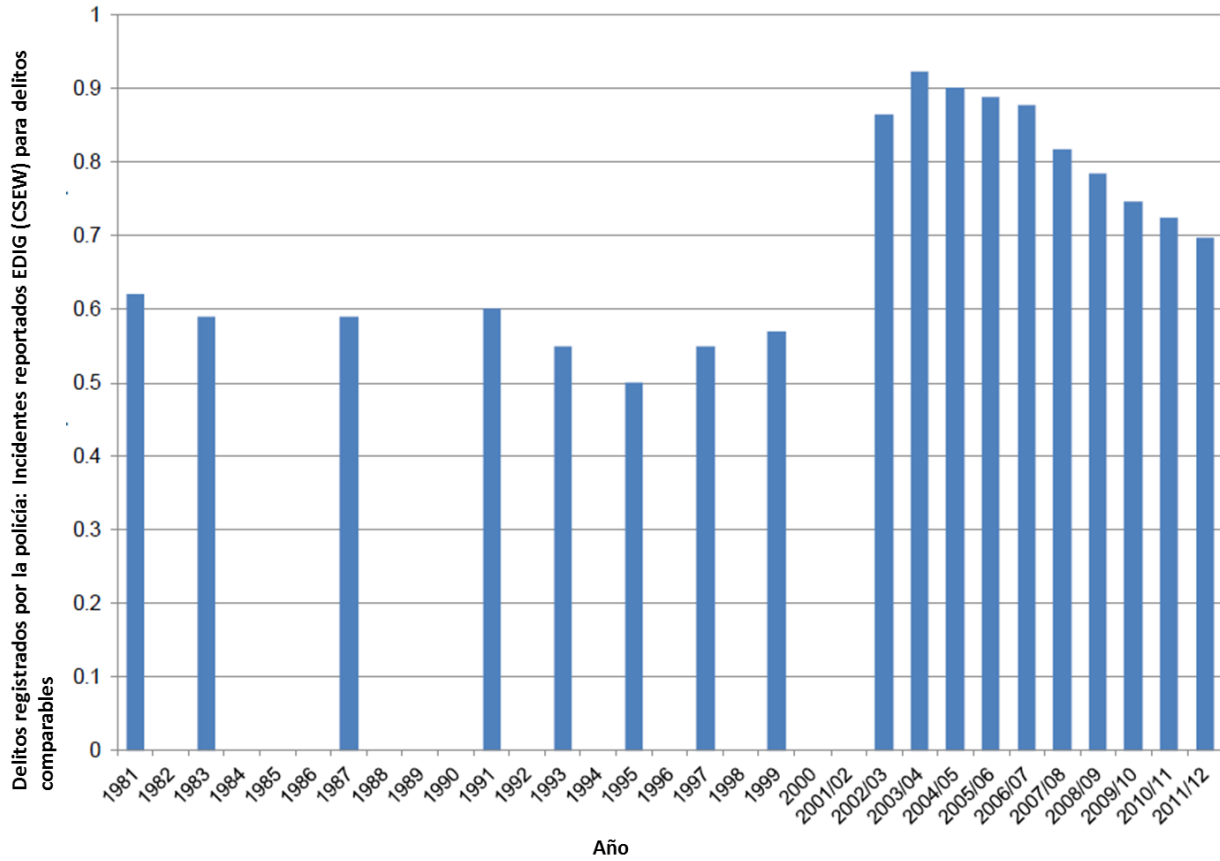


Figura 7: Relación incidentes reportados en la Encuesta Delictiva para Inglaterra y Gales (por sus siglas en inglés CSEW) y delitos registrados por la policía.

Fuente: Elaboración propia con datos de la Oficina de Estadísticas Nacionales, Inglaterra.

Tabla 3. Volumen y porcentaje de reducción en categorías delictivas comparables, de 2002-2003 a 2011-2012.

	2002/03	2006/07	2011/12	Porcentaje de cambio 2002/03- 2006/07	Porcentaje de cambio 2006/07- 2011/12	Porcentaje de cambio 2002/03- 2011-12
Delitos registrados por Policía	3,231,367	2,881,327	1,922,643	-11	-33	-41
EDIG	3,727,000	3,305,000	2,749,000	-11	-17	-26

Fuente: Elaboración propia con datos de la Oficina de Estadísticas Nacionales, Inglaterra.

Relación y posibles aportaciones del documento al modelo propuesto.

En cuestiones metodológicas y de argumentos para la implementación del presente estudio, la información arrojada por la encuesta y sus métodos, aportarían poco al mismo, dado que se trata de una comparativa entre la cifra de delitos denunciados y los registros obtenidos a través de la aplicación de una encuesta, en primera instancia el aporte es muy básico, sin embargo, contribuye a la consideración del factor de delitos sin denunciar que al menos debido a cuestiones culturales está arraigado de manera relevante en la sociedad potosina, sin embargo, sería interesante añadir un rubro de delitos no denunciados para ponderar el valor efectivo de los delitos que sí han sido registrados. La implementación de este elemento involucraría la inclusión de un mecanismo para recabar los datos tal como lo es un cuestionario.

1.1.6 Dinámicas de puntos de concentración de robo y violencia (Chaney y Ratcliff)

La ciencia criminal y las comunidades de análisis delictivo se han vuelto bastante competentes en la creación, rastreo y manejo de puntos de concentración delictiva (Chaney & Ratcliff, 2005). Investigaciones recientes indican que los científicos en el ámbito delincuenciales desglosan los niveles micro de la geografía tales como calles, estacionamientos, edificios, etc., puntos de concentración de incidencia delictiva para formar nuevas figuras, tamaños y patrones espacio temporales (Herrman, 2015).

Problemática:

Para esta investigación, los investigadores señalan que algunos estudios (Taylor, 1997), (Groff, Weisburd, & Yang, 2010) indican que los delitos están agrupados a niveles de vecindario, pero rara vez el vecindario completo es criminógeno y sólo partes de él presentan alta concentración de actividad delincuenciales. Estudios previos (Bursik & Grasmik, 1993), (Morenoff & Sampson, 1997) asumían que las relaciones entre delito, población, uso de suelo, y tipos de establecimiento de negocios son homogéneos y espacialmente estacionarios. Incluso a nivel vecindario, los delitos son altamente agrupados y los vecindarios con mayor cantidad de delitos tienen porcentajes de baja actividad o calles con actividad criminal nula.

Un amplio grupo de investigación ha identificado que un porcentaje pequeño de ubicaciones contiene un porcentaje significativo de ubicaciones criminales. (Herrman, 2015).

Una tendencia actual en criminología ambiental y análisis delictivo es el estudio del delito a lugares de nivel micro.

Solución propuesta:

Una de las corrientes actuales en el ambiente de la criminología y el análisis delictivo es el estudio del delito en lugares a micro nivel, una escala geográfica más debajo de nivel que el de vecindario.

Parte de este concepto de análisis de micro nivel es conocido como la regla 80/20 o (Koch, 1999), y aplica no sólo en prevención delictiva y control delincuencia, sino en otros campos de justicia criminal.

La regla 80/20 sugiere que al enfocarse en el 20% más alto de puntos de concentración delictiva puede influenciar dramáticamente en la mayoría de la criminalidad total en el área de estudio (Herrman, 2015).

La investigación se enfoca en el cambio de puntos de concentración para explorar, medir e ilustrar los cambios espacio-temporales que ocurren en la zona del Bronx en la ciudad de Nueva York, elegida por su alta tasa de ocurrencia de delitos violentos, comparada con los otros cuatro condados de la ciudad.

El área de investigación y la información procesada incluye varios conjuntos de datos de sistemas de información geográfica (SIG). El condado del Bronx se compone de 38 vecindarios, 36 residenciales los otros dos, clasificados como zona pública e industrial y cuenta con 12 recintos policiales. La población del Bronx es de 1332650 un 17% del total de la ciudad de Nueva York, datos acordes al censo de los Estados Unidos del año 2000.

La tabla 4 indica la cantidad desproporcionada de delitos violentos en el Bronx considerando su tamaño geográfico y su población.

Tabla 4. Delitos violentos, zona, población, y porcentajes por cada uno de los cinco bloques vecinales del condado de la Ciudad de New York para 2006-2010.

Delitos violentos (2006-2010)	Bronx	Brooklyn	Manhattan	Queens	Staten Island	Total NY
Homicidio	657 (25%)	1074 (41%)	371 (14%)	434 (17%)	86 (3%)	2622
Violación	1512 (23%)	1873 (28%)	1388 (21%)	1624 (24%)	278 (4%)	6675
Robo	23,018 (22%)	36,616 (35%)	1,745 (21%)	22,029 (20%)	2181 (2%)	105,589
Asalto	2791 (31%)	28,958 (34%)	16,015 (19%)	15,486 (18%)	2240 (3%)	84,263
Disparo de arma de fuego	2791 (31%)	3613 (40%)	1094 (12%)	1311 (15%)	222 (2%)	9031
Área (en millas cuad.)	42.41	71.46	22.78	109.67	58.50	304.82
Porcentaje de área NY	14%	23%	8%	36%	19%	100%
Población	1,332,650	2,465,326	1,537,195	2,229,379	443,728	8,008,278
Porcentaje de población	17%	31%	19%	28%	5%	100%

Fuente: CompStat del Depto. de Policía de NY; Oficina de Administración, Análisis y Planeación del NYPD; 2012.

Integración del modelo que esquematiza la solución del problema:

- Recopilación de información.
- Categorización de los datos.
- Identificación de puntos de concentración delincriminal.
- Enfoque de recursos sobre los delitos dentro del 20% más alto de incidencia.

Hay dos procesos elementales para construir, examinar y geovisualizar los aspectos temporales de los puntos de concentración delictiva. En primer lugar, un análisis espacio-tiempo, donde los especialistas llevan a cabo un análisis espacial rutinario para determinar los patrones temporales de los puntos de concentración obteniendo datos delictivos que posteriormente son importados a hojas de cálculo o programas estadísticos especializados y así detectar patrones temporales significativos respecto a cada punto de concentración. Esto puede ayudar a los analistas criminales a determinar qué tipos de objetivos o víctimas son susceptibles de afectación y como varía respecto al tiempo en sus distintos periodos como hora del día, día de la semana, mes del año, etc. La Figura 8 muestra la gráfica del análisis temporal de un conjunto o clúster del delito de robo,

donde los tonos grises indican actividad nula o baja, los amarillos a naranja indican cantidades medias de robos y el color rojo oscuro muestra una cuenta alta de estos delitos.

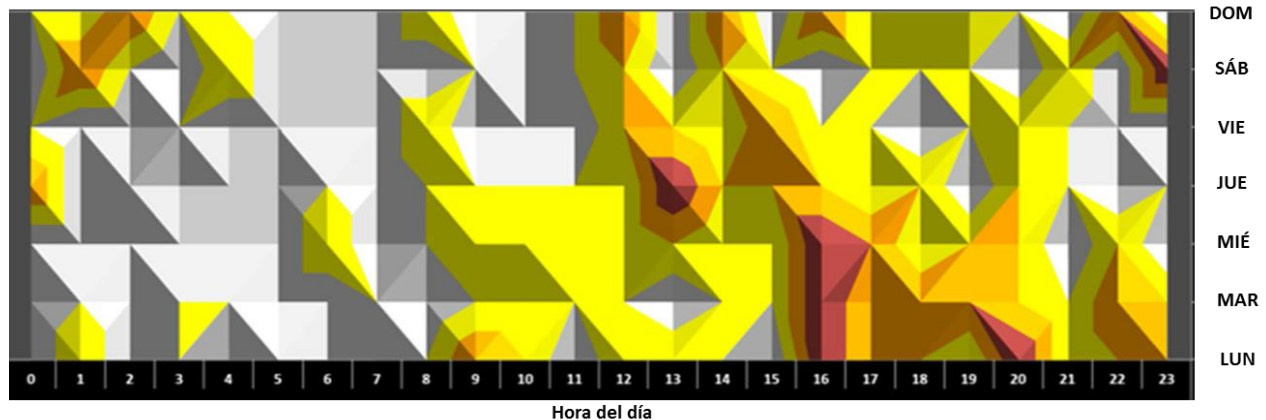


Figura 8: Análisis temporal de un clúster, el color gris indica baja incidencia de robos, mientras el amarillo y naranja una incidencia media, siendo el rojo el índice más alto de ocurrencia.

Fuente: Las dinámicas de los puntos de concentración de robo y violencia (Chainey & Ratcliff, 2005).

Relación y posibles aportaciones del documento al modelo propuesto.

Un aporte factible que se pudiera considerar de utilidad para la fase final del modelo que abordará el estudio, estaría representado en el despliegue de líneas de investigación y aplicación de recursos con base en la utilización del Principio de Pareto, para mitigar de manera fehaciente los estragos producidos por el delito que se aborde debido a su frecuencia y en la representación de puntos de concentración entre variables específicas de tiempo como pueden ser día de la semana y hora del día en que ocurre el evento.

1.1.7 Hacia la conformación de un modelo propio basado en redes neuronales

Vistos los elementos aportados por los diversos estudios revisados, es notablemente legítima la noción de la incorporación del uso de redes neuronales como núcleo de procesamiento de datos para una nueva metodología formalizada adecuadamente, dado que en la mayoría de los casos, la aplicación de inteligencia artificial es un factor común, es manifiesto que el uso de esta tecnología en asuntos de análisis de investigación delictiva no es un asunto nuevo, ni mucho menos reciente, sino un fenómeno que debido a la generación excesiva de información y a su naturaleza de identificación de patrones con independencia de intervención humana favorece su desarrollo y refinamiento constante.

Recapitulando los apartados previos para tener una perspectiva sólida y concisa respecto de las aportaciones, métodos y objetivos de cada uno de los estudios tenemos que, en el primer caso, el esquema desarrollado por Brahan, Lam, Hilton y Leung, parte de la premisa de captar el flujo de información y conservar el conocimiento de los investigadores experimentados y de alguna manera modelar la pericia que estos últimos poseen, de modo que, al momento de su partida toda esa capacidad no partiera con ellos; el método empleado en este caso ressemble un sistema experto y hace uso de elementos de inteligencia artificial como el de un sistema basado en reglas, usando máquinas de aprendizaje y redes neuronales artificiales considerando los delitos de alta incidencia, características del presunto, datos específicos del incidente, incidentes relacionados, eventos cercanos y composición facial, esto con el propósito de que el sistema arrojase un posible presunto responsable.

El modelo de Martin Short no apunta a la conservación de la experiencia del investigador, sino que se enfoca en la construcción de un modelo estadístico, una fórmula matemática capaz de representar la relación espacio-tiempo de la ocurrencia de un evento delictivo, considerando la concentración de incidencia delictiva, el factor tiempo y sobre todo el ambiente urbano con elementos específicos como el tráfico, abandono de propiedades, densidad de población, espacio de ocurrencia, aumento y disminución de atractivo, exclusivamente enfoque en el delito de robo domiciliario y la tasa de delincuentes de este tipo en la zona. El propósito último desarrollar una fórmula matemática para describir el comportamiento delictivo de zonas específicas de una comunidad urbana.

Por otra parte, el estudio de Jayaweera, Sajeewa, Liyanage, Wijewardane y Perera, tiene un enfoque de recopilación de información un tanto menos convencional, debido a la restricción de acceso a la información policial, y debiendo recurrir a la información publicada en artículos periodísticos, establecen un modelo bien estructurado para la recolección, clasificación, normalización y limpieza de datos, generando entidades para ser procesadas por un analizador basado en minería de datos, considerando variables como aquellas intrínsecas al delito, el tipo de delito y la fecha para presentar estadísticas históricas, resultados de análisis, gráficos y mapas.

El modelo de análisis delictivo y predicción a través de minería de datos de Shiju y Surya se orienta más al enfoque de predictibilidad que nos atañe, parte de elementos como el incremento constante de información, limitaciones de los detalles aportados oficialmente respecto a dicha información, busca la conformación de un conjunto de datos preciso, considerando una gama amplia y diversa de delitos y variables como: sitio de ocurrencia, características fisionómicas, edad, señas particulares y la evolución temporal del delito;

debiendo este conjunto de datos ser clasificable dentro de un patrón, esto a través de un modelo de recolección de datos, clasificación, identificación de patrones, predicción y visualización, con el uso de herramientas como son los árboles de toma de decisiones y la construcción de tablas de decisiones.

Estudios de otra naturaleza, como el análisis de variación en tendencias delictivas de Flatley y Bradley, muestran un enfoque completamente alternativo a lo visto previamente, sin embargo permite ratificar problemáticas comunes dentro de la mayoría de los estudios como son el acceso y consistencia de las fuentes de información, que en la mayoría de los casos, al ser estudios realizados por personal ciento por ciento académico no cuentan con facilidades para la obtención de información oficial y deben recurrir a métodos alternativos para la recolección de datos y a aproximación de cálculos para la presentación de resultados. Sin embargo, este estudio también permite observar variables paralelas, al considerar una encuesta se incluyen factores de percepción que, si bien es cierto no son objetivos, manifiestan un elemento relevante para la consideración del impacto social de la incidencia delictiva y buscan complementar información que no está incluida en los datos duros, como pueden ser aquellos omitidos por la falta de denuncia.

Un aporte complementario es el presentado por el estudio de Chainey y Ratcliff, el cual acota de manera minúscula el espacio de ocurrencia de los eventos delictivos, aplicando un análisis espacio temporal al fenómeno para obtener puntos de concentración delictiva, y a la par aplicando el principio de Pareto o Regla 80/20, en el cual, para un amplio número de fenómenos, aproximadamente el 80 por ciento de las consecuencias proviene del 20 por ciento de las causas. El estudio se enfoca a la recopilación de información,

categorización de datos, identificación de concentración delictiva y aplicación práctica de los recursos sobre el 20 por ciento de las zonas con mayor incidencia.

Los modelos revisados incluyen elementos de interés, aunque ninguno aborda de manera directa el propósito del presente estudio que es la predicción de incidencia delictiva con base a características del delito y del entorno utilizando redes neuronales.

De todo lo anterior, y como parte fundamental de base para la conformación del modelo es observable la relevancia de las etapas de tratamiento de la información como son:

- Determinación de las fuentes de información.
- Recopilación de la información,
- Clasificación y normalización de datos.
- Selección de herramienta para el procesamiento de los datos masivos (IA).
- Análisis de la información procesada.
- Visualización de los resultados generados.
- Interpretación y toma de decisiones.

El enfoque hacia un objetivo específico, como puede ser la identificación de un delincuente, el análisis de la información, la relación espacio-tiempo, la descripción del fenómeno o bien la inferencia y predicción de futuros eventos, esto último de interés para el modelo a construir.

Y de la misma manera, las variables a considerar, tales como: tipo de delito, delito específico, características del delito, características del ambiente, densidad de población, variables temporales como fecha y horario, valores históricos y percepción social del fenómeno, aportaciones de cada modelo al propio, figura 9.

1.2 Conceptos básicos de redes neuronales

1.2.1 Definición

Existen bastantes acepciones de lo que es una red neuronal, todas con ligeras diferencias en la definición, a grandes rasgos se puede decir que una red neuronal artificial es un preprocesador distribuido en paralelo de forma masiva que tiene una tendencia natural para almacenar conocimiento de forma experimental y lo hace disponible para su uso (Aleksander & Morton, 1990).

Otra perspectiva, distingue a las redes neuronales artificiales como sistemas de procesamiento de información cuya estructura y funcionamiento están basados en sus homologas biológicas. Se componen de elementos simples denominados nodos o neuronas interconectadas entre sí por conexiones que tienen un valor numérico modificable denominado peso (Montaño Moreno, 2002).

1.2.2 Orígenes y desarrollo de las redes neuronales

En el año de 1943, Warren S. McCulloch, un neurocientífico, y Walter Pitts, un estudioso de la lógica, desarrollaron el primer modelo conceptual de una red neuronal. En su estudio, describen el concepto de neurona, una célula viva simple en una red celular que recibe entradas, procesa estas entradas y genera una salida. Su trabajo, y el trabajo posterior de muchos científicos e investigadores, no estaba enfocado a describir cómo funciona un cerebro biológico, más bien, cómo una red neuronal artificial era diseñada como un modelo computacional basado en la manera en la que el cerebro resuelve ciertos problemas (Shiffman, 2012).

En 1949, Donald Hebb, destacó en el desarrollo de su teoría la relación del comportamiento humano combinado con su conocimiento a través de su desempeño como neurocirujano, sentando las bases de los procesos de aprendizaje y desarrollando reglas de cómo éste ocurría, posteriormente influirían en la concreción de las bases de la teoría de redes neuronales artificiales (Hebb, 1949).

Hacia el año de 1957, Frank Rosenblatt, un psicólogo, desarrolló el modelo de la red neuronal artificial más sencillo, el perceptrón. Aunque con sus limitaciones, este modelo persiste y es actualmente es un esquema muy utilizado en el reconocimiento de patrones. Rosenblatt consideraba que la conectividad en las redes biológicas tiene un elevado porcentaje de aleatoriedad, difiriendo del modelo más simple de McCulloch-Pitts, llevándolo a una teoría de separabilidad estadística para distinguir las propiedades más relevantes de estas redes.

Otro modelo, el ADALINE (Adaptative Linear Elements) Elementos Lineales Adaptativos, fue desarrollado hacia 1960, por Bernard Widrow y Marcial Hoff, su primera aplicación fue para dar solución a un problema real, la construcción de filtros adaptativos para reducir al mínimo los ecos en las líneas telefónicas. Esta red presenta la misma limitante del Perceptrón en cuanto al tipo de problemas que es capaz de resolver, ambas sólo pueden solucionar problemas linealmente separables, pero el algoritmo Mínimo Cuadrado Medio que utiliza es mucho más efectivo que la regla de aprendizaje del Perceptrón, reduciendo el error medio cuadrático.

Todavía en los años sesenta, Marvin Minsky y Seymour Papert, llevaron a cabo estudios que demostraban que el Perceptrón era una red débil y poco versátil dado que no podía

resolver problemas sencillos de funciones no lineales, lo que dio la pauta para la búsqueda de un enfoque más efectivo en el estudio de la teoría de redes neuronales (Minsky, 1961).

Posteriormente, James Anderson desarrollaría su Asociador Lineal ampliando el campo en el desarrollo de las redes neuronales y permitiendo continuar su estudio (Anderson, 1995). Posteriormente, Teuvo Kohonen, de manera independiente, desarrolló un modelo similar al de Anderson.

Para 1974, Paul Werbos propone un algoritmo de aprendizaje denominado de retropropagación para un perceptrón multicapa que más adelante sería predominante (Werbos, 1974), pero que no adquiriría popularidad sino hasta mediados de los años 80's al ser retomado por RumelHart y McClelland en sus estudios de procesamiento paralelo distribuido (RumelHart & McClelland, 1986).

Un modelo aplicado para el reconocimiento de patrones visuales fue desarrollado en los años 80's por Kunihiko Fukushima, la red era auto-organizada y podía "aprender sin profesor", adquiriendo una capacidad para reconocer estímulos de patrones basado en similitudes geométricas de las figuras sin importar su disposición espacial, la cual sería denominada como neocognitron por su autor (Fukushima, 1980).

Desde finales de la década de los 90 hasta la actualidad, el desarrollo de la computación neuronal ha cruzado el umbral de la teoría y se ha ido asentando en el ámbito de la tecnología práctica con buena recepción en distintos ámbitos. La mayoría de las aplicaciones se involucran en problemas de reconocimiento de patrones, haciendo uso de redes neuronales en sus distintas implementaciones (Bishop, 1995).

Para el presente estudio, es en la parte del reconocimiento de patrones donde el uso de redes neuronales puede permitir la construcción de un mecanismo fiable para el procesamiento de datos de incidencia delictiva, extendiendo la capacidad de técnicas más tradicionales como lo son los métodos numéricos, aproximaciones estadísticas o en el caso de la operatividad de la Fiscalía General del Estado, la interpretación empírica del comportamiento delincuencia para la implementación de operativos y campañas de prevención.

1.2.3 Composición y funcionamiento de las redes neuronales

Partiendo del comportamiento biológico en la naturaleza, la comunicación entre neuronas tanto del hombre como de los animales, está formado por el sistema nervioso y hormonal, trabajando en conjunto con los órganos de los sentidos y los órganos efectores, su propósito es recoger información, transmitirla y elaborarla, y en parte almacenarla y redirigirla en su forma ya elaborada (Isasi Viñuela & Galván León, 2004). En resumen, el sistema de comunicación neuronal está compuesto por tres partes:

1. Los receptores, recogen información del ambiente en forma de estímulos.
2. El sistema nervioso, elabora, almacena y reenvía la información.
3. Órganos efectores, reciben información elaborada y la interpretan desempeñando acciones.

La unidad básica estructural y funcional dentro del sistema de comunicación neuronal es la neurona. La neurona biológica (ver figura 10) es una célula orientada específicamente al proceso de información y está constituida por un cuerpo principal denominado soma y

dos tipos de ramificaciones: el axón y las dendritas. La neurona recibe señales de otras neuronas a través de las dendritas y emite información procesada por el cuerpo de la célula a través del axón (Ponce Cruz, 2010).

Una neurona utiliza productos de secreción para la transmisión de la información. La transmisión de la información de origen, obtenida por estímulos, se propaga entre otras neuronas, a través de prolongaciones, construyendo redes; y es en estas redes que la información es elaborada y almacenada, mientras otra parte de la red permanece en comunicación con los efectores, quienes reciben la información a manera de órdenes. A una de las prolongaciones, la responsable de la conducción de impulsos se le denomina axón. Existe también una zona de contacto de las neuronas determinada por los botones terminales de las células nerviosas y su interacción con otras neuronas o células efectoras la cual se denomina sinapsis (Isasi Viñuela & Galván León, 2004).

Esta interconectividad crea los procesos que permiten al hombre o al animal resolver las tareas cotidianas a las que debe enfrentarse.

Estableciendo una analogía entre la actividad neuronal en el proceso sináptico con el mecanismo de una red neuronal artificial, se puede considerar que los estímulos que llegan a la neurona son las entradas de información, las cuales son ponderadas, a través de un parámetro específico llamado peso, que se asocia a la sinapsis correspondiente, las señales pueden inducir una reacción en la neurona o en su defecto inhibirla (Ponce Cruz, 2010).

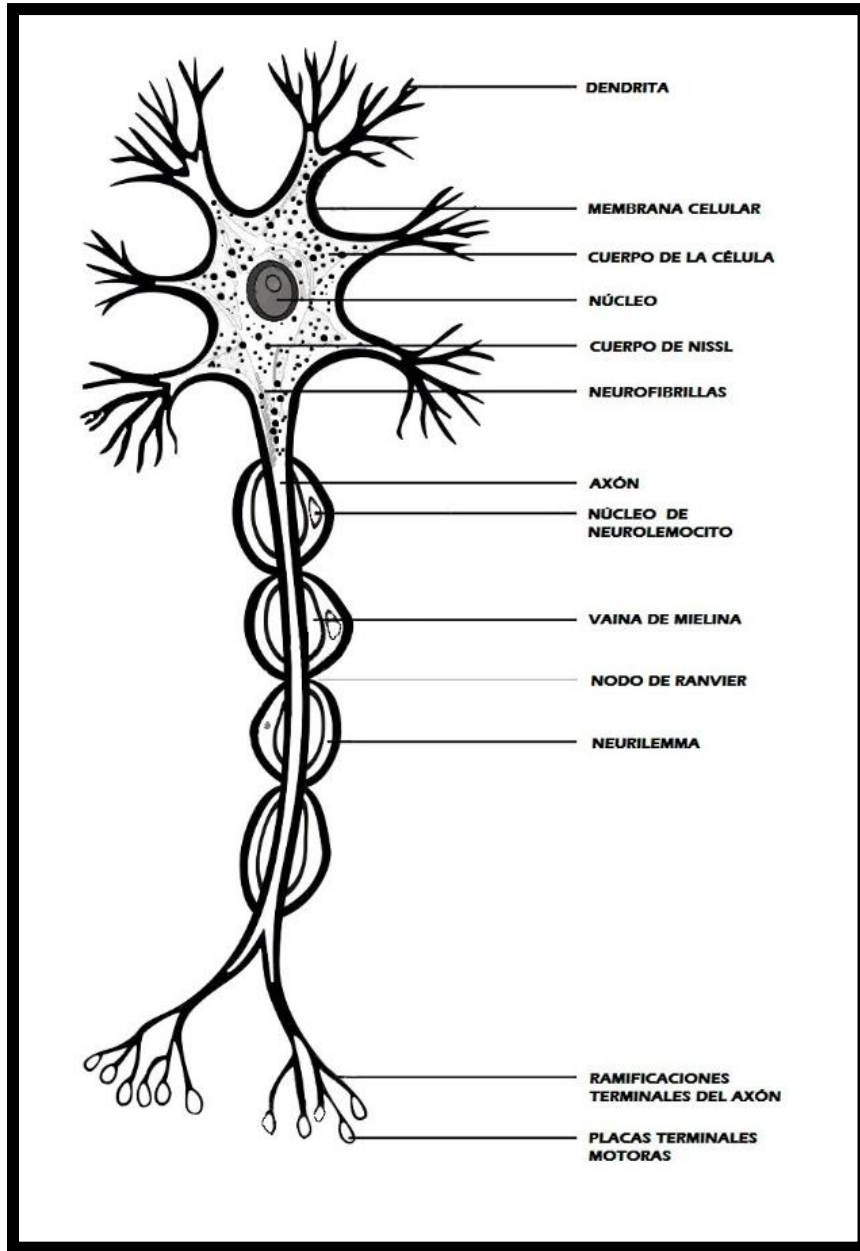


Figura 10: Elementos de una neurona.

Fuente: Shutterstock

La acción resultante es la adición de las entradas ponderadas, si esta suma es igual o mayor que el umbral de la neurona, ésta última se activa generando una salida. La capacidad de transmisión de señales se modifica con base en la actividad del sistema

nervioso. Las sinapsis son susceptibles a la fatiga, debido a la escasez de oxígeno, o intervención de agentes externos tales como los fármacos. La habilidad del ajuste de las señales se constituye como un mecanismo de aprendizaje.

Por otra parte, las Redes Neuronales Artificiales, pueden definirse como sistemas de mapeos no lineales basando su estructura en principios detectados en los sistemas nerviosos biológicos. Se conforman de un gran número de simples procesadores vinculados por interconexiones ponderadas o pesos. La unidad de procesamiento es llamada neurona, célula o autómeta; como elemento básico posee un estado interno denominado nivel de activación, cada una de éstas recibe información de otros nodos que, dependiendo el caso, le permiten cambiar de estado y generar como salida un valor simple escalar.

Una Red Neuronal representa un sistema dinámico con una topología de grafo dirigido, que puede recibir información de salida como resultado de una reacción de su estado respecto a acciones de entrada. Los elementos procesadores y los canales dirigidos son denominados nodos de la red neuronal (Galushkin, 2007).

Dentro de este contexto, la tarea que una neurona desempeña es simple. Dado un conjunto de entradas, procesa una salida cuando sus valores combinados exceden un determinado umbral, satisfaciendo el requerimiento original del problema. (Figura 11).

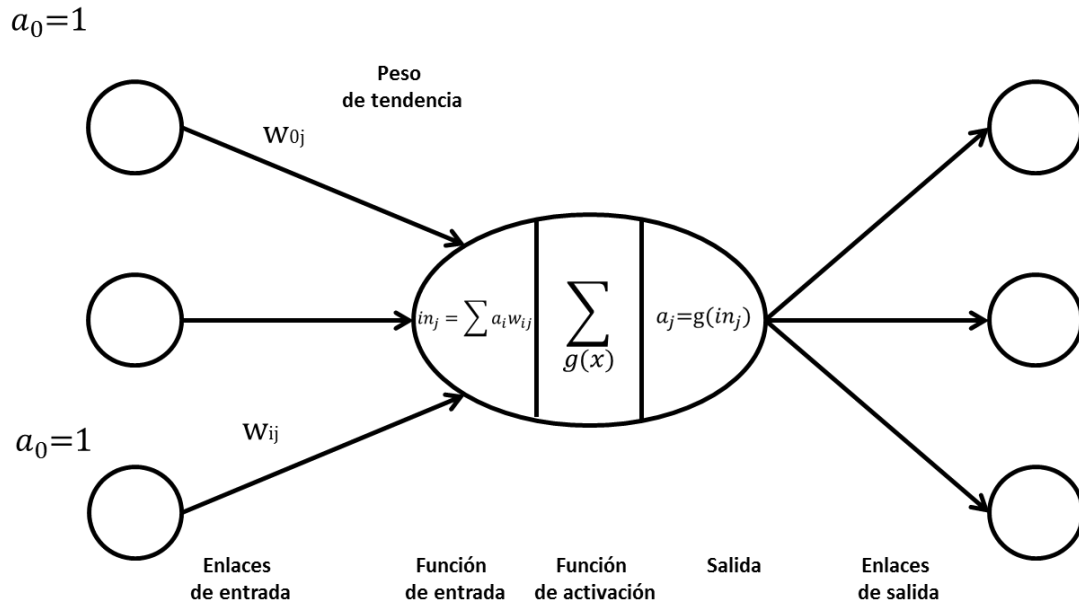


Figura 11: Esquema de una unidad de proceso típica.

Fuente: Elaboración propia basada en un diseño de Javier Béjar.

La neurona posee una tendencia de entrada (a_0) con valor 1. Los pesos de la neurona (w_{ij}) son parámetros definidos y ajustables para el modelo.

Las entradas son procesadas y calculadas como la suma ponderada de todas las entradas originales.

La función de activación $g(x)$ determina la salida y puede ser de distintos tipos de función: función de umbral (perceptrón) o función logística (función sigmoidea).

La salida a_j es obtenida aplicando la función de activación sobre la entrada.

$$a_j = g \left(\sum_{i=0}^n w_{ij} a_i \right) = h_{\alpha}(\varepsilon)$$

Ec. (2)

En resumen, una red neuronal artificial es un modelo computacional distribuido, el cual está basado en la estructura del sistema nervioso de los seres humanos. Su arquitectura se conforma al interconectar múltiples procesadores elementales, creando un sistema adaptativo que incluye un algoritmo para ajustar sus pesos (parámetros libres) para lograr alcanzar los requerimientos de desempeño del problema planteado, esto basado en muestras representativas (Salas, 2004). El bloque básico de toda red artificial es la neurona artificial, siendo un modelo matemático simple, el cual tiene tres reglas elementales: multiplicación, suma y activación. A la entrada de la neurona artificial las entradas se ponderan lo que significa que cada valor de entrada se multiplica por un peso individual. En la parte media de la neurona artificial, es la función de suma la que suma las entradas ponderadas con el sesgo. En la salida de la neurona artificial, la suma previa de las entradas ponderadas y el sesgo se pasa a la función de activación también denominada función de transferencia, el modelo puede observarse en la figura 12.

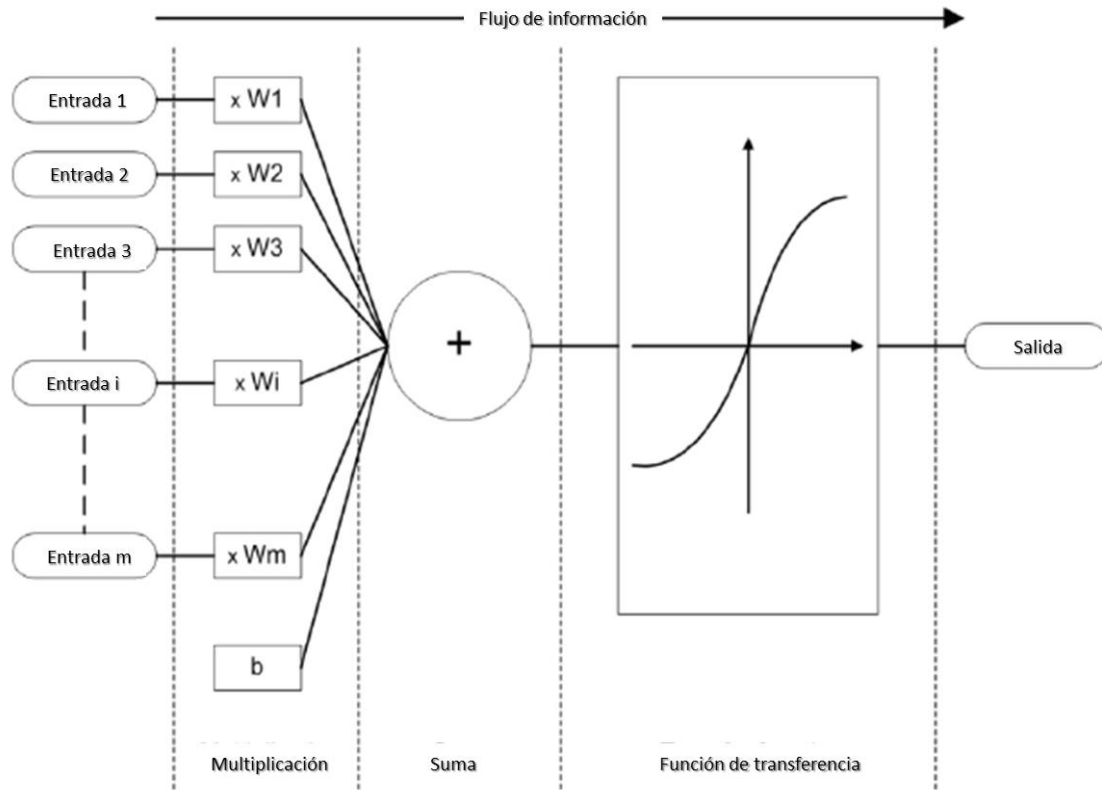


Figura 12: Principio de funcionamiento de una neurona artificial

Fuente: Elaboración propia basada en el esquema de Andrej Krenker.

Una red neuronal artificial se distingue por incluir las siguientes características:

- Contiene un conjunto de unidades elementales, las cuales individualmente poseen una baja capacidad de funcionamiento.
- Se vincula creando una estructura de interconexión muy densa, utilizando enlaces ponderados.
- Los pesos o parámetros libres se deben ajustar para cumplimentar los requerimientos de desempeño planteados por el problema.
- Posee un alto grado de paralelismo.

1.2.4 Aspectos característicos y elementos de las redes neuronales

Respecto a la estructura de las redes neuronales, se distinguen por cuatro aspectos fundamentales, siendo estos su topología, su mecanismo de aprendizaje, asociación entre entrada y salida y la representación de dicho resultado.

La topología: Considera la organización y disposición de las neuronas dentro de la red, creando grupos denominados capas, los parámetros básicos son la cantidad de neuronas por capa, grado de conectividad y el tipo de conexión entre ellas. La cantidad de capas permite determinar si se trata de una red monocapa o bien una multicapa. Por otra parte, el tipo de conexión ayuda a determinar si se trata de redes de propagación hacia adelante o hacia atrás.

El mecanismo de aprendizaje: Es el proceso mediante el cual una red neuronal ajusta sus pesos en respuesta a la información proporcionada por la entrada. A través de este proceso los pesos se van modificando y cuando llegan a un estado estable se dice que la red ha aprendido.

El tipo de asociación entre entrada y salida: Consiste en la información que la red adquiere y la manera que asocia los datos de entrada con la salida que le corresponde. Distinguiéndose dos maneras de llevar a cabo dicha asociación. Primeramente, la hetero-asociación donde la red aprende parejas de datos de la forma $[(A_1, B_1), (A_2, B_2) \dots (A_n, B_n)]$, donde para cada entrada A se genera su correspondiente salida B . Y, por otra parte, la auto-asociación, donde la red adquiere determinada información $A_1, A_2 \dots A_n$, de manera tal que cuando un dato de entrada se presente, la red realiza una auto-correlación, devolviendo el dato almacenado más parecido a la entrada.

La forma de presentación de la información: Tanto los datos de entrada y salida de una red neuronal pueden ser representados de formas diferentes: analógicas y discretas. Si la información es analógica le corresponde una función de activación continua, por ejemplo, una función lineal o sigmoidea; en el caso de entradas con valores discretos o binarios, la función de activación es de tipo escalón. Sin embargo, existen también redes en las que las entradas pueden ser continuas y las salidas discretas.

1.2.6 Principales topologías de redes neuronales

- El Perceptrón.

El perceptrón es el modelo más simple de una red neuronal. Este modelo consiste en sumar los valores de las señales de entrada y multiplicar cada uno por los valores de los pesos determinados de manera aleatoria; el valor resultante es comparado con un patrón para determinar si la neurona se activa o no, si el valor es superior, la salida toma el valor de 1, en caso contrario será 0. La forma más general de una neurona artificial puede describirse en dos etapas descritas en la figura 13. En la primera, se calcula la combinación lineal de las entradas. Cada valor del vector de entrada se asocia al valor del peso correspondiente, normalmente valores entre 0 y 1. Además, la función sumatoria a menudo toma un valor extra de entrada que representa el umbral o sesgo de una neurona.

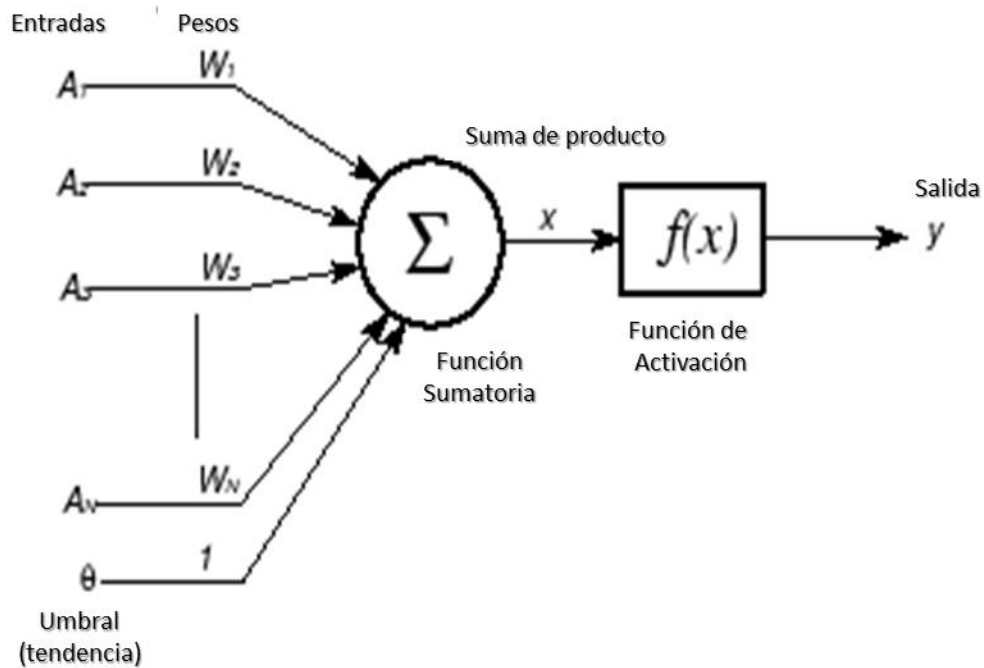


Figura 13: Etapas dentro de una neurona artificial

Fuente: Elaboración propia basada en el esquema empleado por Rodrigo Salas.

- Redes monocapa.

Al conectar múltiples neuronas, aparece el verdadero poder de cálculo de las redes neuronales, aunque, si bien es cierto, incluso una neurona simple puede desarrollar un nivel sustancial de capacidad de cómputo. La estructura más común para conectar neuronas dentro de una red es a través de capas. La forma más simple de una red en capas se puede observar en la figura 14, los nodos sombreados a la izquierda constituyen la denominada capa de entrada. Las neuronas de la capa de entrada existen solo para pasar y distribuir las entradas sin desarrollar ningún cálculo. De ahí, la única capa real de neuronas es la que se

encuentra a la derecha. Cada una de las entradas $X_1, X_2, X_3, \dots, X_n$ es conectada a cada una de las neuronas en la capa de salida a través del peso de conexión. Dado que cada valor de las salidas $Y_1, Y_2, Y_3, \dots, Y_n$ es calculado a partir del mismo conjunto de valores de entrada, cada salida varía en función de los pesos de conexión, el valor de peso de cero puede ser representado como “sin conexión” (Kawaguchi, 2000).

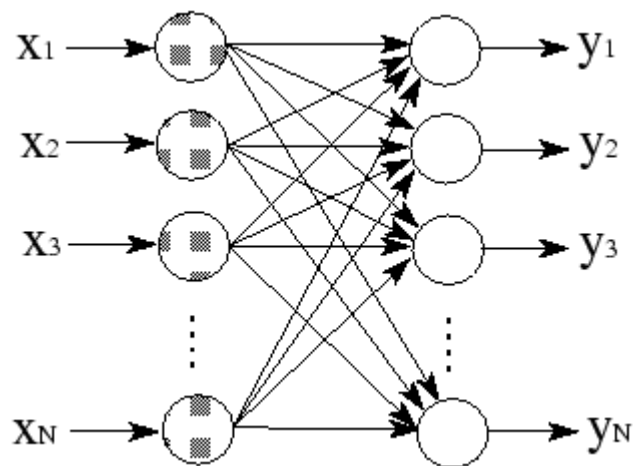


Figura 14: Estructura simplificada de una red en capas.

Fuente: Esquema empleado por Kyoshi Kawaguchi.

Entre las redes monocapa, algunas permiten que las neuronas tengan conexiones a sí mismas y se denominan autorrecurrentes.

Las redes monocapa se utilizan bastante en circuitos eléctricos, esto debido a su topología, dado que permiten su implementación a través de hardware, empleando matrices de diodos que representan las conexiones entre neuronas (Ballesteros González, 2000).

- Redes multicapa

Para alcanzar un grado superior en el cálculo computacional se requiere una estructura de red neuronal mucho más compleja, la figura 15 muestra una red neuronal que se distingue de la red monocapa al poseer una o más capas ocultas. En esta estructura multicapa, los nodos de entrada pasan la información a las unidades en la primera capa oculta, enseguida las salidas de dicha capa se pasan a la siguiente y así sucesivamente.

La red multicapa también puede verse como una cascada de grupos de redes de una sola capa. El nivel de complejidad en los cálculos puede ser visto por el hecho de que muchas redes de una sola capa se combinan en esta red multicapa. El diseñador de una red neuronal artificial debe considerar cuántas capas ocultas se requieren dependiendo de la complejidad de cálculos deseada (Kawaguchi, 2000).

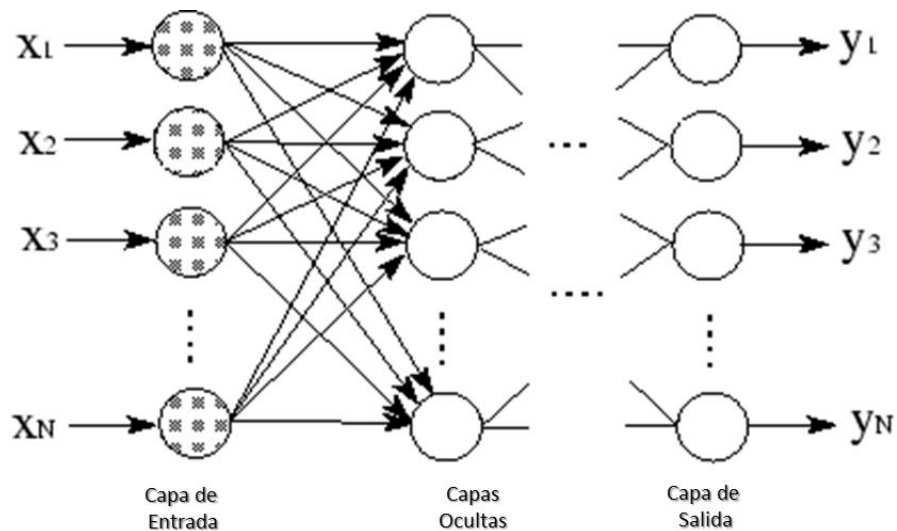


Figura 15: Estructura extendida de una red multicapa.

Fuente: Esquema empleado por Kyoshi Kawaguchi.

De manera usual, el procesamiento dentro de una red multicapa va dirigido en el orden en que reciben el valor de entrada hasta la salida y están unidas en ese mismo orden. A este tipo de conexión se le denomina feedforward o hacia adelante. Las más representativas son el perceptrón multicapa, Adaline, Madaline, Backpropagation y los modelos LQV y TMP de Kohonen (Ballesteros González, 2000).

Por otro lado, existen algunas redes en que las capas, aparte del orden normal, también se unen desde la salida hacia la entrada en el orden inverso en que los valores de entrada son transportados. A estas conexiones se les denomina hacia atrás, retroalimentadas o feedback (Ballesteros González, 2000). Algunos ejemplos representativos son las red ART, redes de Memoria Asociativa Bidireccional y el Cognitron (Ballesteros González, 2000). En la figura 15 pueden observarse los distintos tipos de redes neuronales.

1.2.7 Mecanismos de aprendizaje

Muy probablemente, la característica más importante de una red neuronal es la habilidad de asimilar (aprender) la información de entrada y mejorar el funcionamiento al procesar la información. El concepto de aprendizaje se refiere a muchas acepciones bajo varios puntos de vista y es complicado llegar a plantear una definición precisa del término. En redes neuronales, el aprendizaje se define como la secuencia de eventos listada a continuación (Haykin, 1999):

1. Estimulación del entorno en el que la red se desenvuelve.

2. Cambios en los parámetros libres de la red como resultado de la estimulación.
3. Asimilación de una forma nueva en las respuestas al medio ambiente para mejorar su rendimiento.

Un algoritmo de aprendizaje es un conjunto predeterminado de reglas bien definidas para el aprendizaje de una red neuronal. Hay muchos tipos de algoritmos de aprendizaje; el objetivo común del aprendizaje es el ajuste de pesos de conexión.

Hay dos clases de aprendizaje: el aprendizaje supervisado y no supervisado. El aprendizaje supervisado requiere una fuente de información externa para ajustar la red. Por otro lado, en el aprendizaje no supervisado, no hay un agente externo que vigile el proceso de aprendizaje. En su lugar, la red se ajusta mediante un control interno del rendimiento. (Kawaguchi, 2000).

A. Aprendizaje supervisado.

El aprendizaje supervisado puede ser visto como aquel que requiere un “maestro”, quien de hecho conoce la respuesta correcta para una amplia gama de ejemplares. Ésta no es la manera en la que la mayoría de los animales incluyendo al hombre aprenden.

Es, de hecho, una técnica de aprendizaje automatizado que establece los parámetros de una red artificial a través de datos de entrenamiento. La tarea de la red neuronal artificial de aprendizaje es establecer el valor de sus parámetros para cualquier valor de entrada válido después de haber visto el valor de salida. Los

datos de entrenamiento consisten en pares de entrada y valores de salida deseados que se representan tradicionalmente en vectores de datos. El aprendizaje supervisado también puede ser referido como una clasificación, donde se tiene una amplia gama de clasificadores, cada uno con sus fortalezas y debilidades. Sin embargo, elegir un clasificador adecuado (perceptrón multicapa, máquinas de vectores de apoyo, k-vecino más cercano, modelo mixto Gaussiano, Gaussiano, Bayes ingenuo, árbol de decisión, clasificadores funcionales de base radial, etc.) para un problema dado es aún más un arte que una ciencia.

Para resolver un problema de aprendizaje supervisado, hay que considerar varios pasos.

En el primer paso tenemos que determinar el tipo de ejemplos de entrenamiento. En el segundo paso necesitamos reunir un conjunto de datos de entrenamiento que describa satisfactoriamente un problema dado. En el tercer paso debemos describir los datos de entrenamiento reunidos en forma comprensible para una red neuronal artificial elegida. En el cuarto paso hacemos el aprendizaje y después del aprendizaje podemos probar el rendimiento de la red neuronal artificial aprendida con el conjunto de datos de prueba (validación). Conjunto de datos de prueba consisten en datos que no han sido introducidos a la red neuronal artificial mientras aprenden (Krenker, Bester, & Kos, 2011).

En este tipo de aprendizaje se suelen considerar, a su vez, tres formas de llevarlo a cabo, que dan lugar a los siguientes aprendizajes supervisados:

1. Aprendizaje por corrección de error

Este tipo de aprendizaje considera que al generarse una salida de la red neuronal se compara con un valor esperado, la diferencia de estos valores es utilizada para llevar a cabo el aprendizaje. El efecto que se tiene es que los pesos de las conexiones se ajustan de acuerdo a dicha diferencia, en función al error cometido y modificándose hasta obtener la respuesta que se desea. Se busca minimizar el error entre la salida esperada y la obtenida, un ejemplo claro de este tipo de aprendizaje es el Perceptrón.

2. Aprendizaje por refuerzo

Las bases de este aprendizaje son muy parecidas al aprendizaje supervisado pero la información que se proporciona a la red es mínima y está limitada a señalar si la respuesta de la red es correcta o incorrecta.

Este tipo de aprendizaje se basa en la noción de condicionamiento por refuerzo, esto es se aprenden las conductas reforzadas positivamente y las conductas castigadas o reforzadas negativamente. En nuestro mundo esto se traduce en premiar los pesos sinápticos cuando se acierta la salida y penalizarlos cuando no se acierta (reconocimiento-castigo) (Kawaguchi, 2000).

3. Aprendizaje estocástico

Este tipo de aprendizaje se fundamenta en realizar cambios aleatorios en los valores de los pesos de las conexiones de la red y evaluar su efecto a partir del objetivo deseado y a través de distribuciones de probabilidad.

Para comprender el funcionamiento de este aprendizaje, se puede hacer una analogía en términos termodinámicos, asociando a la red neuronal con un sólido físico que tiene cierto estado energético. En el ámbito de la red neuronal, la energía que posee representaría el grado de estabilidad de la red, de manera tal que el estado de mínima energía correspondería a una condición en la que los pesos de las conexiones optimizan su funcionamiento ajustándose al objetivo deseado.

De acuerdo con lo previo, el aprendizaje consistiría en efectuar un cambio aleatorio de los valores de los pesos y determinar el nivel de energía de la red. Si la energía fuera menor después del cambio, en otras palabras, si el comportamiento de la red se aproxima al esperado, se acepta el cambio; si, en caso contrario, la energía no es menor, se admitiría el cambio en función de una distribución de probabilidades determinada y preestablecida.

B. Aprendizaje no supervisado.

El aprendizaje no supervisado es una técnica de aprendizaje automatizado que establece los parámetros de una red neuronal artificial basada en datos dados y una función de coste que debe ser minimizada. La función de coste puede ser cualquier función y está determinada por la formulación de la tarea. El aprendizaje no supervisado se utiliza principalmente en aplicaciones que entran dentro del dominio de los problemas de estimación, tales como el modelado estadístico, la compresión, el filtrado, la separación ciega de fuentes y la agrupación.

En el aprendizaje no supervisado se busca determinar cómo se organizan los datos. Se diferencia del aprendizaje supervisado y el aprendizaje por refuerzo en

que la red neural artificial se da sólo ejemplos no etiquetados. Una forma común de aprendizaje no supervisado es la agrupación en la que tratamos de clasificar los datos en diferentes grupos por su similitud. Entre los modelos de redes neuronales artificiales descritos anteriormente, los mapas de auto-organización son los que usan con mayor frecuencia algoritmos de aprendizaje sin supervisión (Krenker, Bester, & Kos, 2011).

Se puede decir que, en este tipo de aprendizaje, se le permite a la máquina darle sentido al mundo lo mejor que pueda. Aprende a dividir entradas en categorías. Debe existir el sentido de que un conjunto de definiciones sea mejor que otro. El algoritmo básico es simplemente aprender a crear una representación del mundo:

1. Entrada x .
2. Ejecutar la red para obtener la salida y .
3. Comparar y con x .
4. Retroalimentar el error.

El simple hecho de agrupar las entradas es útil. Mientras menos unidades representen y reconstruyan la entrada significa que se tiene una red más eficiente. Queda claro que, una red que para consista de una unidad oculta por cada entrada posible, con pesos 1 o 0, la reconstrucción sería perfecta, pero esto solo sería una tabla de consulta. La idea es que una red sea más eficiente que la propia tabla de consulta, porque se desea la habilidad predictiva.

C. Aprendizaje Hebbiano

En su postulado, Hebb señaló que este tipo de aprendizaje es la base para otros, y consiste en medir la familiaridad o extraer las características de la información de entrada, se ajustan los pesos de las conexiones de acuerdo con la correlación entre los valores de entrada y de salida de cada neurona. El fundamento consiste en lo siguiente: si dos neuronas N_i y N_j toman el mismo estado simultáneamente (ambas activas o ambas inactivas), el peso de la conexión entre ambas se incrementa.

Las entradas y salidas permitidas a la neurona son: $\{-1, 1\}$ o $\{0, 1\}$ (neuronas binarias). Esto puede explicarse porque la regla de aprendizaje de Hebb se originó a partir de la neurona biológica clásica, que solamente puede tener dos estados: activa o inactiva (Hebb, 1949).

D. Aprendizaje competitivo y cooperativo

En este tipo de aprendizaje, las neuronas compiten entre sí para representar un patrón de entrada. La neurona elegida es aquella cuyos pesos incidentes se asemejan más al patrón de entrada. El aprendizaje consiste en reforzar las conexiones de la unidad ganadora y debilitar las otras, para que los pesos de la unidad ganadora se asemejen cada vez más al patrón de entrada, llevando a cabo la reconstrucción del patrón con base en los pesos de la neurona ganadora siendo estos lo que más se parecen (Kawaguchi, 2000).

1.2.8 Propagación de la información en las redes neuronales

De acuerdo al tipo de conexión, las redes neuronales artificiales se pueden clasificar de dos tipos: de propagación hacia adelante o hacia atrás.

En las redes de propagación hacia adelante o denominadas Feedforward, las señales neuronales se desplazan hacia adelante a través de las capas de la red. No hay conexiones hacia atrás ni conexiones auto recurrentes, las salidas de las neuronas en ningún momento son entradas de neuronas del mismo nivel o precedentes.

Otro caso es, cuando las salidas de las neuronas pueden ser transportadas como entradas de alguna capa previa o incluyendo el mismo nivel, la red se denomina como red de propagación hacia atrás.

1.2.9 Funciones de transferencia (Activación y salida)

Las neuronas artificiales tienen diferentes estados de activación, algunas cuentan con dos, como es el caso de su contrapartida biológica, pero otras pueden tomar un valor determinado dentro de un conjunto predefinido.

La función de activación $z_i=f(x_i,w_i)$ y la función de salida $y_i=f(z_i)$ se suman en el término funciones de transferencia.

La función de activación se usa para calcular un nuevo estado de activación, partiendo del estado original y la combinación de las entradas con los pesos de las conexiones.

Las funciones de salida definen el valor de salida de una neurona dependiendo del valor obtenido en la activación. De manera general, se aplican funciones monótonamente crecientes (Debes, Koenig, & Gross, 2005).

A continuación, se enlistan las funciones más comúnmente usadas:

Función Escalón (hardlim, hardlims)

Esta función restringe los valores entre 1 y -1, es utilizada en el Perceptrón. Se asocia a neuronas binarias en las que cuando la suma de las entradas es mayor o igual que el umbral de la neurona, el valor de activación es 1; y si es menor la activación es 0, o preferentemente -1 dado que en algunos casos al ser 0 las neuronas tienden a no activarse. La figura 16 muestra el comportamiento de dicha función.

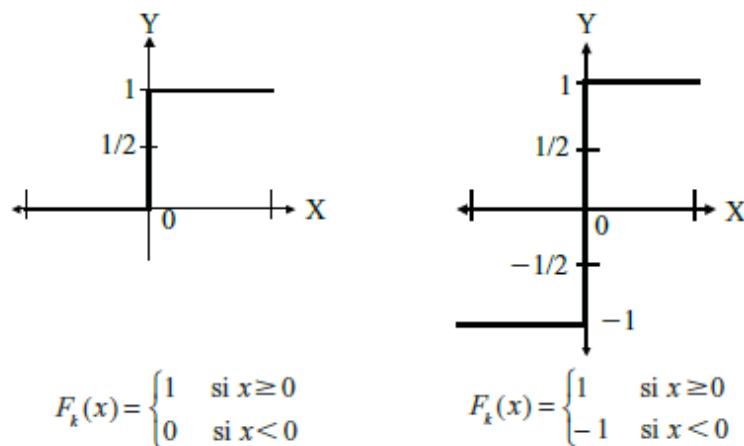


Figura 16: Función Escalón

Fuente: Diseño original de Pedro Ponce Cruz.

Función Lineal-Mixta

La función lineal o identidad es prácticamente equivalente a no aplicar función de salida (Ponce Cruz, 2010), y la salida de la misma es igual a su entrada, esto se puede ver representado en la figura 17. La red de tipo Adaline utiliza este tipo de funciones de activación.

En aquellas neuronas con función mixta, si la suma de entrada comprende entre ambos límites, tanto superior como inferior, la activación se define como una función lineal de la suma de las señales de entrada.

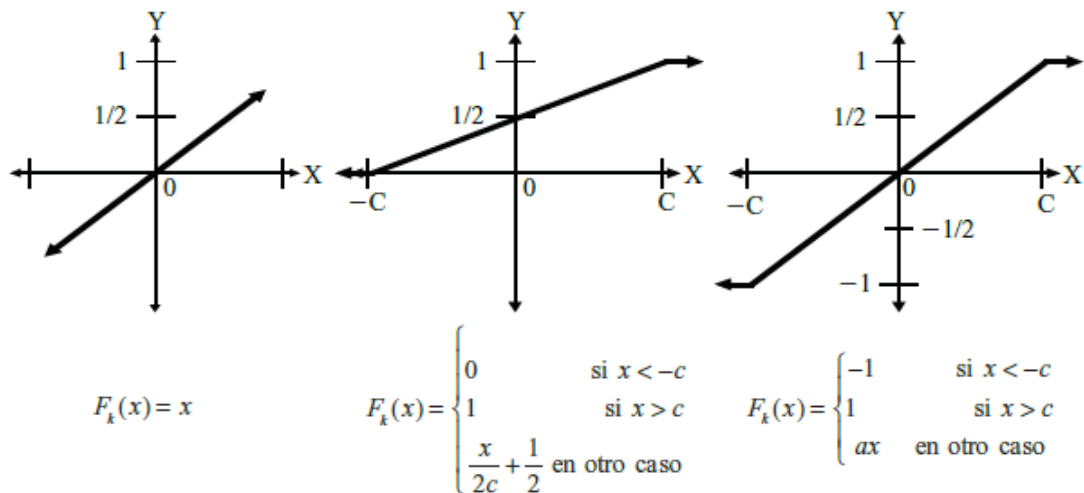


Figura 17: Función lineal-mixta.

Fuente: Diseño original de Pedro Ponce Cruz.

Función Tangente Hiperbólica

Este tipo de función es utilizada en aquellas situaciones donde se presentan variaciones suaves de valores positivos y negativos de la señal que se pretende

clasificar. Su gráfica puede observarse en la figura 18, ésta es una de las funciones mayormente usadas en entrenamientos supervisados, como es el caso de entrenamiento por retropropagación de error (Ponce Cruz, 2010).

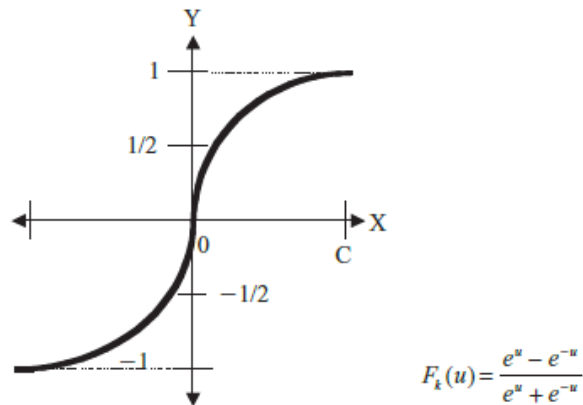


Figura 18: Función tangente hiperbólica

Fuente: Diseño original de Pedro Ponce Cruz.

Función Sigmoidal

A través de la función sigmoideal el valor dado por la función es próximo a uno de los valores asíntóticos, provocando con esto que el valor de salida se encuentre en la zona alta o en la baja del sigmoide. Y en el caso de que la pendiente sea elevada la función tiende a la función escalón. Por otra parte, la importancia de la función sigmoideal radica en que su derivada es siempre positiva y cercana a cero para los valores grandes positivos o negativos, además adquiere su máximo valor cuando $x=0$. De esta manera permitiendo que se puedan usar reglas de aprendizaje definidas para las funciones escalón con la ventaja de que la derivada está definida en todo el intervalo (Ponce Cruz, 2010). Pudiendo observarlo en la

figura 19. Su uso es particularmente frecuente en redes multicapa como la de retropropagación.

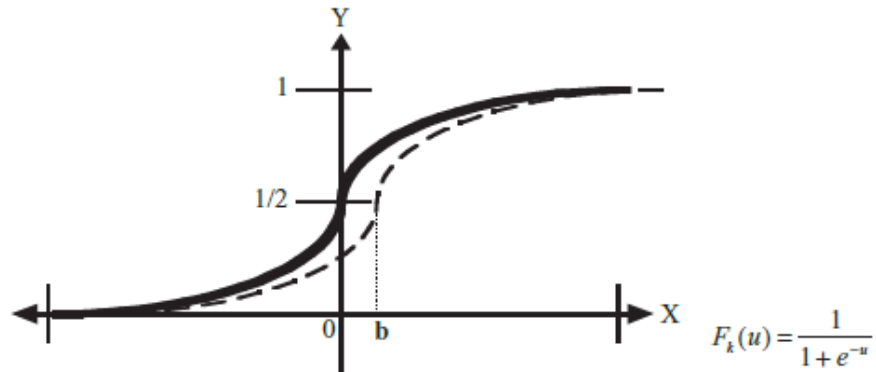


Figura 19: Función sigmoideal.

Fuente: Diseño original de Pedro Ponce Cruz.

Función de Gauss

Hay mapeos que pueden realizarse con un solo nivel de neuronas mediante la aplicación de funciones de activación del tipo Gauss, en la figura 20 se puede ver su comportamiento.

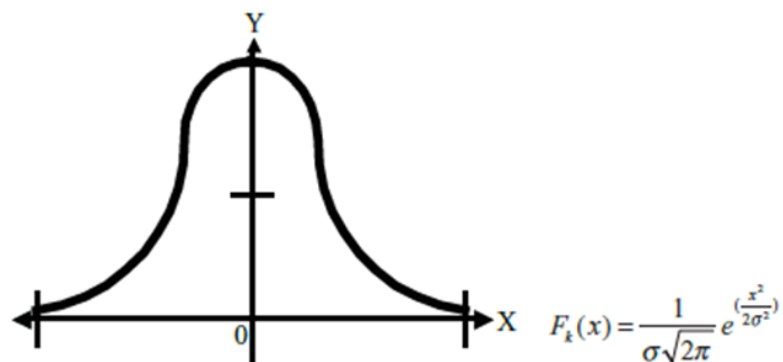


Figura 20: Función de Gauss.

Fuente: Diseño original de Pedro Ponce Cruz.

1.3 Hipótesis

La implementación de un modelo basado en redes neuronales permitirá definir un pronóstico efectivo del comportamiento de la incidencia delictiva superior al 80 %.

1.4 Variables de relevancia para evaluación como parámetros de impacto en el modelo

Para el tratamiento y resolución de cada uno de los objetivos específicos planteado, en este trabajo se definen cuatro variables dependientes:

1. Mecanismo procedimental. Definido por la descripción de la interacción de dos conjuntos de variables independientes representado en las condiciones socioeconómicas y las particularidades del delito o fenómeno antijurídico.
2. Alineación de actividades a un proceso lógico-racional. Consistente en la construcción de una guía procedimental para los analistas focalizada en una herramienta computacional como lo son las redes neuronales y dependiente en el mecanismo definido en el punto anterior.
3. Elementos de significancia. La determinación de aquellas variables que impacten de mayor manera en la generación del pronóstico con base en el modelo obtenido, así como de las funciones internas de la red neuronal que permitan un grado mayor de certeza.
4. Pronóstico del comportamiento. El valor efectivo en la aplicación práctica del modelo y comparado con los valores reales de incidencia delictiva. Esto dependiente del desempeño del modelo.

1.5 Base de la metodología a implementar

En primera instancia, la aplicación de un algoritmo de redes neuronales a un conjunto de datos, con base en la configuración previa de reglas bajo las cuales analizar dicha información parece ofrecer un resultado bastante específico y, hasta cierto punto, sólido.

Sin embargo, para dicho procesamiento es preciso establecer una serie de mecanismos que permitan conjuntar los datos en bloques procesables, incluso desde antes la recopilación de los mismos.

En su estudio (Shiju & Surya, 2014), aplican una serie de pasos para la realización del análisis delictivo, dicho procedimiento es compatible con las bases del ciclo de inteligencia utilizado ampliamente por las instituciones policiales alrededor del mundo. Véase la figura 21.



Figura 21: El ciclo de Inteligencia.

Fuente: Castellanos, Luis; Apuntes de Inteligencia militar y policial.

Los pasos aplicados por Shiju y Surya contemplan:

1. Recolección de datos.
2. Clasificación.
3. Identificación de patrones.
4. Predicción.
5. Visualización.

Y para cada uno de ellos se integran algoritmos que dan un formato específico a la información. El modelo de Shiju y Surya puede ser aplicado en el presente estudio, combinado con el ciclo de inteligencia.

El punto de Identificación de patrones sería el núcleo sobre el cual trabajar la aplicación de un algoritmo de redes neuronales. Las redes neuronales se pueden utilizar para la predicción con diversos niveles de éxito. Una de las ventajas de su implementación incluye el aprendizaje automático de las dependencias basado únicamente en los datos de la medición sin ninguna necesidad de añadir más información. La red neuronal es entrenada a partir de los datos históricos con la esperanza de que se descubra dependencias ocultas y que será capaz de utilizarlos para predecir en el futuro. En otras palabras, la red neural no está representada por un modelo dado de forma explícita. Es más parecida a una caja negra que es capaz de aprender algo.

Capítulo 2: Unidad de Análisis de Información de una institución de procuración de justicia y productos de inteligencia.

2.1 La Unidad de Análisis de Información

2.1.1 Antecedentes

Como se planteó al comienzo de este estudio, a mediados de la década pasada, se logra obtener los recursos y la autorización para la generación de un órgano interno dentro de la Fiscalía General del Estado de San Luis Potosí encargado del procesamiento de información estadística para la generación de líneas de investigación y definición de estrategias para encarar la problemática delictiva de aquella época, el proyecto derivó en lo que hoy en día continua funcionando como una entidad facultada dentro de la Fiscalía y conocida como Unidad de Inteligencia.

Su conformación fue más bien un proceso empírico prolongado derivado de esfuerzos continuos para dar respuesta a las necesidades de información y productos de valor informativo que la procuración de justicia demanda.

2.1.2 Conceptualización de Unidad de Análisis de Información

Respecto a la formalización tanto procedimental como jurídica y atendiendo los requerimientos del Nuevo Sistema de Justicia Penal, cuya implementación se homogenizó en todo el territorio nacional el 18 de junio de 2016, las Unidades de Análisis de Información toman forma dentro del esquema de seguridad pública y procuración de justicia, y hoy en día se canalizan esfuerzos para concretar dicha formalización.

Las unidades de análisis delictual se conforman por un equipo de profesionales que realizan tareas específicas, empleando tecnología y procedimientos característicos,

prestando servicios a otros, ayudando a la organización a ser más eficaz, gracias a una mejor gestión de información (Tudela, 2016).

En la práctica, dicha definición se queda un tanto corta al limitar el alcance que realmente poseen, de esta manera se puede establecer que una Unidad de Análisis de Información es una estructura altamente capacitada, acorde con una dinámica delincencial diversificada que basa su operación en la premisa de descentralización operativa y concentración de la información soportada por una base tecnológica que le permita desarrollar su función (Torres Estrada, 2015).

2.2 La inteligencia

Dentro del ámbito de Seguridad y Procuración de Justicia, la inteligencia es un proceso sistemático que permite establecer un objetivo y un sistema de toma de decisiones que permita reducir, interrumpir y prevenir la delincuencia por medio de una gestión estratégica (Ratcliffe, 2011).

La inteligencia es el proceso que una unidad de análisis debe seguir para cumplir su objetivo, que consiste en hacer eficiente el proceso de investigación por medio de los productos que genera.

Al ser un proceso sistemático, la inteligencia está conformada por una serie de pasos. A esta metodología se le conoce como ciclo de inteligencia y lo constituyen “cinco pasos orientados a la generación de conocimiento estratégico útil y ajustado a los requerimientos de información preestablecidos por un destinatario final a quien se le difunde selectivamente el resultado final plasmado en un instrumento determinado” (Navarro, 2004).

Este ciclo es una representación, figura 22, del proceso de inteligencia, donde se enlazan estas cinco fases, que se presentan a continuación:



Figura 22: Representación del ciclo de inteligencia de acuerdo al CISEN (Centro de Investigación y Seguridad Nacional)

Fuente: CISEN (Centro de Investigación y Seguridad Nacional)

1. Planeación. Aquí se plantea el propósito del sistema, identificando y priorizando los temas críticos y definir los elementos esenciales de información.
2. Investigación y recolección. Se conforma el plan de búsqueda para obtener los elementos esenciales de información y se procede a elegir y verificar las fuentes y redes de información, delimitando así el recurso humano, material y las técnicas de investigación necesarios.
3. Análisis y producción. En esta instancia se determina la pertinencia de la información recolectada para producir conocimiento a través de su transformación

por medio de herramientas metodológicas con el propósito de generar estrategias, líneas de acción, escenarios o alternativas e identificar requerimientos adicionales.

4. Difusión y explotación. En esta etapa se busca dar a conocer el producto generado a los usuarios apropiados, en el momento oportuno y a través de medios seguros para que sea utilizado de manera efectiva.
5. Retroalimentación. Contribuye a mejorar el sistema de inteligencia ayudando a adaptarse a nuevas circunstancias por medio de elementos de corrección.

2.3 Productos de inteligencia

La labor de la Unidad de Análisis de Información no concluye con el procesamiento de información, sino que además tiene la tarea de generar productos de inteligencia. Un producto de inteligencia se define como el conjunto de datos y métodos resultado de la aplicación de herramientas de análisis en el ciclo de inteligencia, permitiendo tomar decisiones en la organización para mejorar sus acciones respecto al entorno globalizado y competitivo en el que se desenvuelve (Serralde, 2011). Un producto de inteligencia termina siendo el resultado que se obtiene al aplicar el ciclo de inteligencia, permitiendo a los tomadores de decisiones elegir estrategias y líneas de acción para la resolución de problemas y anticipación de riesgos (Torres Estrada, 2015).

2.3.1 Parte de novedades

Es una compilación informativa de todos los eventos ocurridos que implican un impacto a la seguridad pública o a la procuración de justicia, tales como: detenciones en flagrancia, ordenamientos judiciales y ministeriales cumplimentados, ocurrencia de eventos delictivos, informes de acciones de impacto por las autoridades relacionadas.

2.3.2 Registro de ingreso a separos

Como una de las obligaciones de las instituciones de Seguridad Pública y con el objeto de garantizar el cumplimiento de los principios constitucionales de legalidad, objetividad, eficiencia, profesionalismo, honradez y respeto a los derechos humanos, en el artículo 40, párrafo XIX, de la Ley General del Sistema Nacional de Seguridad Pública, se deben inscribir las detenciones en el Registro Administrativo de Detenciones conforme a las disposiciones aplicables (Cámara de Diputados del H. Congreso de la Unión, 2016).

2.3.3 Matriz de asociación

Es una técnica para ordenar y vincular información como son: domicilios, empresas, vehículos, personas, y organizaciones delictivas, que facilita el análisis e interpretación al investigador.

Una matriz es un arreglo triangular de números o símbolos en el cual la información es almacenada en columnas y filas. No hay límites en el número de columnas o filas que aparecen dentro de la matriz. Su número y contenidos están determinados por el tipo y calidad de datos que están al alcance de la investigación.

2.3.4 Red de vínculos

Dentro de los productos de inteligencia, es una técnica que representa gráficamente la inteligencia evaluada, mediante matrices y diagramas que facilitan la interpretación de los diferentes vínculos detectados por la interacción de los elementos involucrados, como pueden ser: personas, teléfonos, domicilios, compañías, asociados, lugares frecuentados, cuentas bancarias, vehículos, etc.

2.3.5 Red técnica y red de cruces

La red técnica es una técnica que permite mostrar las comunicaciones de un objetivo en una investigación para determinar sus enlaces con otros actores, además de confirmar o desvirtuar una hipótesis respecto a las interacciones entre esos elementos (Auditoria Superior de la Federación, 2013).

La red de cruces es una representación gráfica derivada de las redes técnicas, donde puede observarse la comunicación e interacción telefónica entre números fijos y móviles, con la finalidad de afirmar o descartar una línea no confirmada, pero con un alto índice de probabilidad de ocurrencia, dentro de una investigación.

2.3.6 Mapas de incidencia delictiva

Son documentos en los que se hace alusión al comportamiento, en forma georreferenciada, de uno o varios tipos delictivos en una zona específica, dentro de un periodo determinado (Auditoria Superior de la Federación, 2013).

2.3.7 Informes estadísticos delictivos

Son compendios de datos procesados, derivados de la integración recabada de las fuentes oficiales que forman el registro de incidencias y su comportamiento a través de un periodo, conformando tablas y gráficos que permiten la apreciación de las tendencias de la incidencia delictiva.

2.3.8 Informe de inspección y/o análisis

Es un documento que contiene información de relevancia para la investigación y que generalmente involucra la descripción, relación e integración de uno o varios productos de inteligencia, o bien, el resultado de la inspección de un indicio o la respuesta a una

solicitud de una autoridad competente para su integración en una carpeta de investigación.

2.4 Plan de acción

Para desarrollar el presente estudio, el plan de acción se dirige a la formalización del procesamiento para la implementación de un nuevo producto de inteligencia denominado informe de incidencia delincidencial, buscando establecer un método bien estructurado, compatible con los patrones establecidos en los protocolos del Nuevo Sistema de Justicia Penal, modificando las actividades empíricas y compilándolas en una secuencia clara y asentándolas como un procedimiento lógico y formal dentro del manual de procedimientos. El alcance aplica únicamente al producto de inteligencia en comento, dado que la extensión a la formulación de un manual de procedimientos completo para la unidad implicaría un esfuerzo superior a los horizontes del presente estudio.

El plan utiliza como guía el ciclo de inteligencia que orienta de manera ideal la construcción todos los productos de información elaborados por toda unidad de análisis de información.

Incorpora la innovación de integrar un modelo computacional como las redes neuronales para el procesamiento automatizado de la información generada a través de las entradas de datos generadas por el sistema de emergencia y las denuncias captadas por la Fiscalía General del Estado.

Capítulo 3: Metodología del procesamiento de información para pronosticar incidencia de delitos patrimoniales.

3.1 Introducción

A través de este capítulo se asienta el método por medio del cual se propone el modelo afín para resolver la problemática planteada, partiendo del tipo de investigación a efectuar, abordando los parámetros de información relevantes y la manera en que dicha información debe ser recopilada y procesada, las técnicas y herramientas a utilizar, y la manera en que serán presentados los resultados.

3.2 Contexto

Dentro de los productos de inteligencia no se manejan herramientas que permitan proveer un pronóstico del comportamiento de la incidencia delictiva, sino que este proceso se genera al momento en el que, utilizando otros productos de inteligencia, las autoridades responsables de implementar acciones para contrarrestar los efectos del comportamiento delincuenciales deducen con base en la información que tienen a su alcance.

Actualmente, de manera periódica se presentan informes estadísticos y mapas delincuenciales que plasman el comportamiento de las actividades criminales en los principales centros urbanos del estado. Aunque es posible observar un reloj criminológico con base en los registros, gráficas y mapas, en realidad, la integración y suposición del comportamiento posterior se deduce a través de la experiencia de las autoridades, y no por medio de una técnica lógica y documentada que permita observar a un futuro a corto

y mediano plazo las acciones de los delincuentes para, de esta manera, plantear acciones que contrarresten el efecto de inseguridad social generado por la incidencia delictiva.

La integración de una nueva herramienta como puede ser el modelo para el procesamiento y análisis de datos de incidencia delictiva puede influir de manera positiva para la creación de un nuevo producto de inteligencia capaz de asentar de manera formal la proyección del comportamiento delincuenciales ayudando a la toma de decisiones con la incorporación de tecnología y los respectivos métodos procedimentales, compatibles con el ciclo de inteligencia que orienta las tareas de la Unidad de Análisis de Información.

3.2.1 Fuentes de información

Se entiende por información criminal a todos aquellos datos generados a partir de la presunta comisión de un delito o de sus componentes, que tenga relevancia en la toma de decisiones a futuro, ya sea para la prevención, detección o esclarecimiento del delito o bien, en la persecución de delincuentes, la mejora de procedimientos de justicia y la creación de leyes (Perversi, 2007).

Las fuentes de información respecto al comportamiento delincuenciales pueden ser muchas y de distintos tipos, desde las fuentes abiertas y públicas como son las noticias en diarios, televisión, radio e Internet a las cerradas como son las fuentes oficiales como el sistema de emergencias 911, las denuncias y querellas, o los registros del sistema penitenciario, siendo estas últimas, entendidas como el conjunto de instituciones y procedimientos en el proceso que transita un hecho delictivo que el Estado registra, y de las cuales el presente estudio echa mano para la obtención de datos para procesamiento y conformación del modelo planteado, en particular el sistema de emergencias y los registros de denuncia por un hecho delictivo ante la Fiscalía General del Estado.

Además de la información intrínseca al delito, y como se pudo observar en los estudios revisados antes de iniciar la conformación del presente, es preciso contar con información complementaria que ofrezca datos del ambiente en el que se manifiesta el fenómeno, tanto del entorno urbano como del social e incluso del ambiental.

Respecto a la determinación geoespacial la información provista por el delito es demasiado básica y general, consistiendo generalmente en un domicilio o incluso en ocasiones únicamente el nombre de una calle y colonia con alguna referencia de la zona, por lo que se precisa una fuente más sólida y uniforme como la construida por el Instituto Nacional de Estadística Geografía e Informática, que adhiere mayor información a los segmentos geoespaciales definidos y con los cuales construye las estadísticas que son utilizadas para la generación de políticas de desarrollo a nivel federal.

En cuestión social el Consejo Nacional de Evaluación Política de Desarrollo Social construye indicadores de marginación y pobreza compatibles con la distribución geográfica utilizada por INEGI.

Como complemento, y originado por la percepción empírica durante la labor de investigación delictiva del autor del presente estudio y corroborado por literatura relacionada, se considera el factor ambiental, en específico al incremento de temperatura, con una correlación al aumento de la violencia e incidencia delictiva como señala Ellen G. Cohn en su artículo *Wheater and crime* publicado en el *British Journal of Criminology*, Vol.30. (Cohn, 1990).

3.2.2 Sistema de emergencias

Concebido como un sistema de atención telefónica, originalmente a través del número 066, el sistema de emergencias 911 es el número único para la atención de urgencias por medio del despacho y canalización a servicios médicos, policiales y de protección civil, soportados en una plataforma constituida por personal capacitado, tecnología de cómputo y radiocomunicaciones con el objeto de brindar el apoyo a la ciudadanía en situaciones de riesgo tanto de afectación contra la integridad física o de su patrimonio como de salud.

El servicio a través del número 911 quedó homologado a través del acuerdo 03/XL/2016 del Consejo Nacional de Seguridad Pública para su aplicación en todo el territorio mexicano, publicado en el Diario Oficial de la Federación el 09 de septiembre de 2016.

El artículo 221 del Código Nacional de Procedimientos Penales establece las condiciones para dar inicio a una investigación de hechos que revistan características de un delito. Y la Constitución Política de los Estados Unidos Mexicanos en el artículo 21 señala que “la investigación de los delitos corresponde al Ministerio Público y a las policías, las cuales actuarán bajo la conducción y mando de aquél en el ejercicio de esta función”.

Dentro de dicha responsabilidad se establecen las normas para el registro, investigación y apertura de la Carpeta de Investigación cuando existan elementos que confirmen la posibilidad de la comisión de un delito. Estos registros son el compendio informativo con los cuales se busca alimentar el modelo propuesto en el presente estudio.

3.2.3 Instituto Nacional de Estadística Geografía e Informática (INEGI)

De acuerdo con la descripción difundida dentro de la página web del organismo (https://www.inegi.org.mx/inegi/quienes_somos.html), el 25 de enero de 1983 se creó, por decreto presidencial, el Instituto Nacional de Estadística, Geografía e Informática (INEGI), que integró en su estructura a:

- La Dirección General de Estadística, en funciones desde 1882, cuando pertenecía a la Secretaría de Fomento, Colonización, Industria y Comercio.
- La Dirección General de Geografía, establecida en 1968 y que estaba adscrita a la Secretaría de la Presidencia.
- La Dirección General de Política Informática.
- La Dirección General de Integración y Análisis de la Información.

Con su creación, el INEGI modernizó la valiosa tradición que tenía nuestro país en materia de captación, procesamiento y difusión de información acerca del territorio, la población y la economía. Conjuntó en una sola institución la responsabilidad de generar la información estadística y geográfica.

Debido a esto, INEGI proporciona no sólo acceso a datos históricos y estadísticos sino a métodos y herramientas que permiten organizar de manera efectiva la relación geoestadística producida en cada ámbito de interés gubernamental a través de elementos geográficos definidos como lo es el AGEB (Área Geoestadística Básica), que es una extensión territorial ocupada por un conjunto de manzanas que generalmente son de 1 a 50, delimitadas perfectamente por calles, avenidas, andadores o cualquier otro rasgo de fácil identificación en el terreno y cuyo uso del suelo sea principalmente habitacional, industrial, de servicios, comercial, etc. (INEGI, s.f.)

3.2.4 Consejo Nacional de Evaluación Política de Desarrollo Social.

El Consejo Nacional de Evaluación de la Política de Desarrollo Social (CONEVAL), fundado en el año 2005, es un organismo público descentralizado de la Administración Pública Federal, con autonomía y capacidad técnica para generar información objetiva sobre la situación de la política social y la medición de la pobreza en México, que permita mejorar la toma de decisiones en la materia. Dentro de las actividades del consejo, se encarga de normar y coordinar la evaluación de la Política Nacional de Desarrollo Social y las políticas, programas y acciones que ejecuten las dependencias públicas. Además de establecer los lineamientos y criterios para la definición, identificación y medición de la pobreza, garantizando la transparencia, objetividad y rigor técnico en dicha actividad.

Como tarea específica CONEVAL tiene la responsabilidad de establecer los criterios de construcción de indicadores para la medición de la pobreza y situación de las políticas sociales, generando información objetiva para facilitar la toma de decisiones.

Como parte de esos instrumentos, CONEVAL elabora un cuadro estadístico que incluye las variables principales, derivadas del estudio de indicadores, para determinar el grado de rezago social definido territorialmente a través de AGEB's.

3.2.5 Instituto Nacional de Investigaciones Forestales, Agrícolas y Pecuarias.

El Instituto Nacional de Investigaciones Forestales, Agrícolas y Pecuarias (INIFAP) fue creado por decreto presidencial el 23 de agosto de 1985, integrado por:

- El Instituto Nacional de Investigaciones Agrícolas (INIA).
- El Instituto Nacional de Investigaciones Forestales (INIF).
- El Instituto Nacional de Investigaciones Pecuarias (INIP).

La actividad continua y consistente del INIFAP en el desarrollo y divulgación científica ha permitido la conformación de conjuntos estadísticos ambientales, que, como parte de la experiencia previa, la cuestión ambiental afecta de manera directa las actividades humanas, incluyendo aquellas involucradas al caso del presente estudio, por lo que implicaría la inclusión de al menos una variable ambiental con el propósito de observar el comportamiento del modelo.

Cabe resaltar que se utilizaría una variable ambiental (temperatura, lluvia, velocidad del viento, etc.) y no una estacional o de actividad específica (periodos vacacionales, festividades, etc.), debido a que, aunque hay registro de una Institución de labor científica como el INIFAP. este último tipo de variables no presenta consistencia temporal, es decir, no se manifiesta en un periodo específico e inmutable en el año.

3.3 Definición de Variables

Las variables definidas en un principio con el propósito de dar respuesta a los objetivos específicos y al objetivo general del presente estudio son variables dependientes de la ejecución del modelo.

1. Mecanismo procedimental.
2. Alineación de actividades a un proceso lógico-racional.
3. Elementos de significancia.
4. Pronóstico del comportamiento.

El modelo propuesto requiere de la utilización de variables definidas respecto a la aplicación e intervención directa con la realidad que se aborda, es decir, se busca encontrar aquellos elementos de significancia que, por su interacción, impactan en el comportamiento de la incidencia delictiva. Durante la revisión de las variables se llevó a cabo dicha observación, obteniendo dos grupos de influencia, siendo las variables socioeconómicas y las variables que asocian las particularidades del delito.

Las variables socioeconómicas destacadas son:

1. Zona geoestadística de influencia (Área Geoestadística Básica). El Marco Geoestadístico es un sistema único y de carácter nacional diseñado por el INEGI, para referenciar correctamente la información estadística de los censos y encuestas con los lugares geográficos correspondientes, proporciona la ubicación de las localidades, municipios y entidades del país, utilizando coordenadas geográficas. Divide al territorio nacional en áreas con límites identificables en campo, denominadas Áreas geoestadísticas, con tres niveles de desagregación:

Estatad (AGEE), Municipal (AGEM) y Bsica (AGEB), sta puede ser urbana o rural.

2. Influencia econmica de la zona (Ingreso Per Cpita, Producto Interno Bruto). Medida total econmica y sociolgica combinada de la preparacin laboral de una persona y de la posicin econmica y social individual o familiar en relacin a otras personas, basada en sus ingresos, educacin, y empleo. Procesada por CONEVAL.
3. Nivel acadmico. Grado de desarrollo acadmico promedio, alcanzado por los habitantes de la zona. Procesada por CONEVAL.

Las variables con particularidades del delito relevantes son:

1. Nmero de eventos. Presunta ocurrencia de delitos registrados en averiguaciones previas iniciadas o carpetas de investigacin, reportadas por las Procuraduras de Justicia y Fiscalas Generales de las entidades federativas en el caso del fuero comn y por la Procuradura General de la Repblica en el fuero federal.
2. Da de la semana. Periodo de 24 horas con nombre definido comprendido dentro de un ciclo de siete jornadas consecutivas.
3. Hora del da. Periodo de 60 minutos con nombre cardinal dentro de un ciclo de 24 periodos consecutivos o da.
4. Sector de operacin. rea de actividad de operaciones de investigacin dentro de la zona metropolitana de San Luis Potos.
5. Delito cometido. Elemento de clasificacin de conductas antijurdicas perseguidas y sancionadas por el Estado.

6. Grado de violencia durante la comisión del delito. Nivel de agresión o ausencia de la misma para perpetrar una actividad antijurídica.

Las variables ambientales o de entorno:

1. Temperatura. Es una magnitud física relacionada a la noción de calor de un objeto, cuerpo o en el caso del estudio al medio ambiente.

3.4 Población y muestra

El conjunto total de registros incluye datos de los delitos patrimoniales principales como lo son Robo de Vehículo, Robo Comercial y Robo Domiciliario, durante el periodo comprendido entre enero de 2013 y diciembre de 2016, alrededor de 10000 registros, contrastando contra una muestra significativa deducida del mismo universo. La muestra es representativa y está definida por el vector de entrada a la red neuronal en uso.

3.5 Instrumentos

Para el procesamiento de la red neuronal se ha optado por utilizar el lenguaje R el cual es un lenguaje de programación interpretado, de distribución libre, bajo Licencia GNU, y se mantiene en un ambiente para el cómputo estadístico y gráfico. Este software corre en distintas plataformas Linux, Windows, MacOS. El término ambiente pretende caracterizarlo como un sistema totalmente planificado y coherente, en lugar de una acumulación gradual de herramientas muy específicas y poco flexibles, como suele ser con otro software de análisis de datos (Santana & Mateos Farfán, 2014).

R-Studio es un entorno de desarrollo integrado libre y abierto, es decir, su utilización no implica costo alguno por el uso de licencias, aunque también cuenta con versiones de uso comercial. Incluye una consola, editor de sintaxis que apoya la ejecución de código, así como herramientas para el trazado, la depuración y la gestión del espacio de trabajo. Se desempeña sobre distintas plataformas como son Linux, OSX y Windows.

Se utiliza el entorno R-Studio para facilitar el uso del lenguaje R y aprovechar las distintas rutinas de implementación de algoritmos de redes neuronales con que cuenta dentro de sus librerías y evitando abordar la parte de la programación desde cero de un algoritmo específico.

3.6 Estructura procedimental.

3.6.1 Metodología general del procesamiento de información de incidencia delictiva a través de RNA.

La metodología para establecer el proceso del modelo propuesto se basará principalmente en la distribución del esquema planteado por Torres Estada (Véase Figura 23), en su enfoque respecto a la generación de productos de inteligencia en su “Diseño e implementación de una Unidad de Análisis de la información para la persecución del delito”. Integrar el modelo procesamiento de datos con redes neuronales dentro de dicho esquema se plantea como un módulo tácito posterior a las fases de recolección, tratamiento e integración de información. El modelo presentado en este trabajo es una adaptación basada en los procedimientos requeridos para poder recopilar, normalizar y procesar la información a través de la herramienta de redes neuronales artificiales elegida.

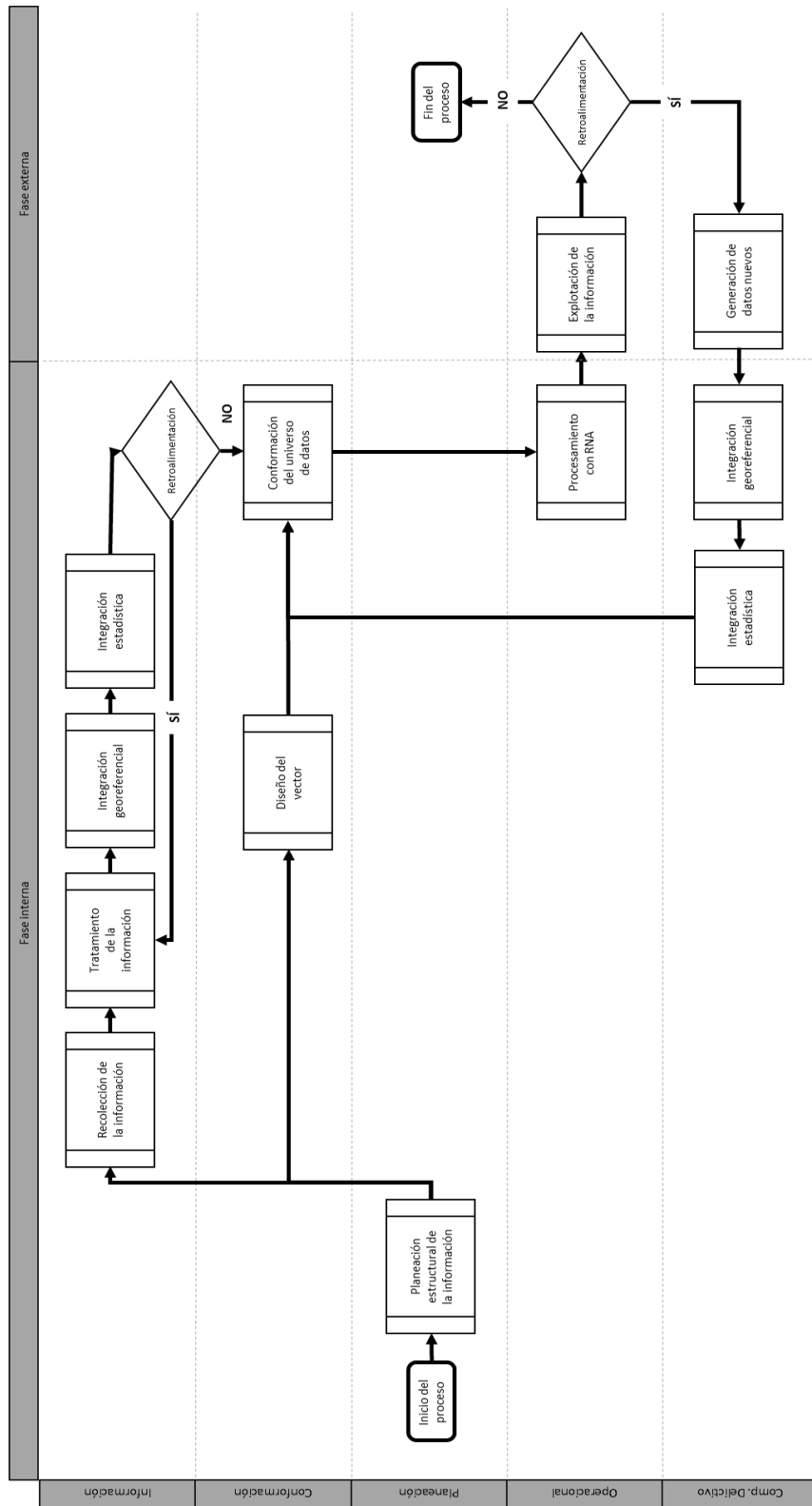


Figura 23: Metodología general del procesamiento de información de incidencia delictiva a través de RNA.

Fuente: Diagrama propio basado en el planteado por Pedro Torres Estrada.

Se propone ajustar el modelo completo dentro de la fase interna del flujo de información en una Unidad de Análisis de Información, y a través de la retroalimentación obtenida en la fase externa refinar tanto los datos como el funcionamiento de la red.

3.6.2 Modelo estructural de procesamiento

El módulo nuclear preliminar para el procesamiento de datos de incidencia delictiva Se establece como se muestra en la figura 24.

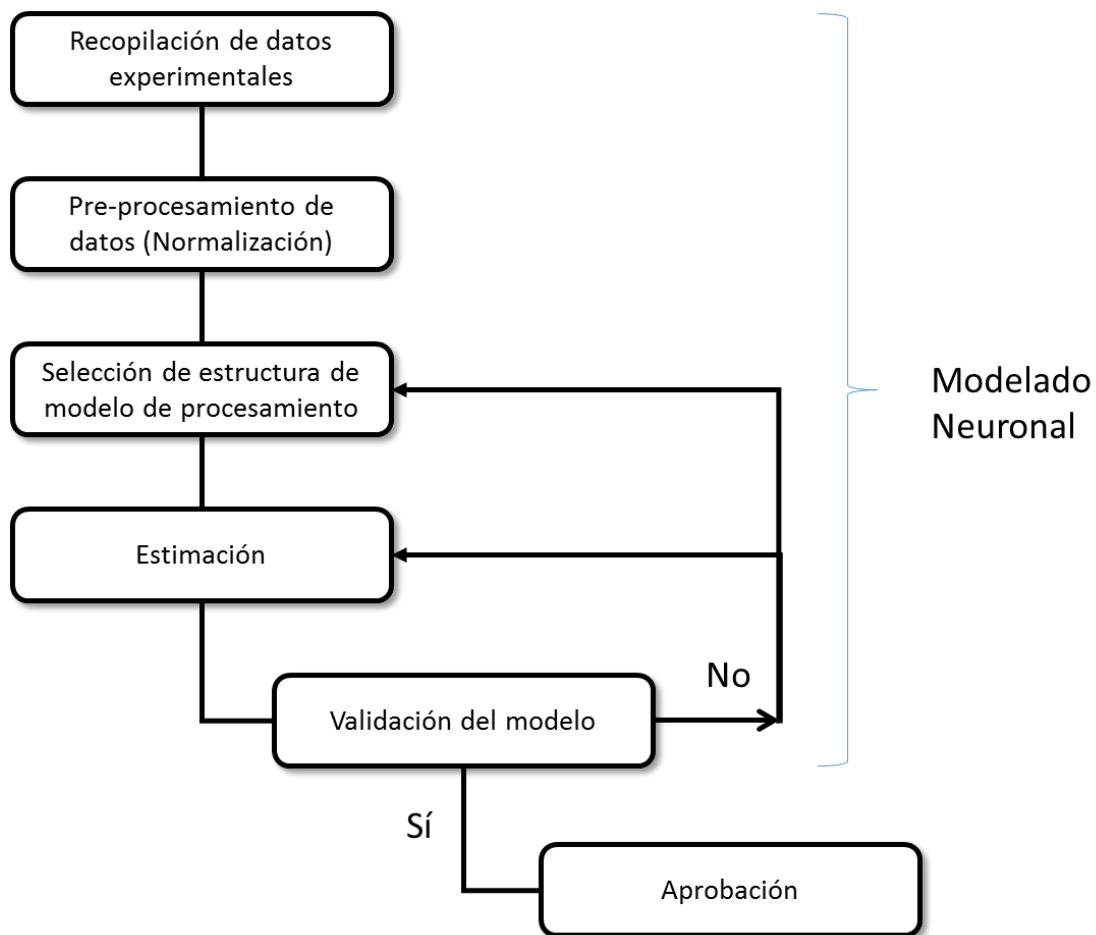


Figura 24: Modelado neuronal para el procesamiento de datos.

Fuente: Elaboración propia basado en el esquema de Alejandro Zambrano.

3.6.3 Procedimientos

Se propone efectuar un procesamiento a través de redes neuronales sobre los registros de tres categorías de delitos patrimoniales: Robo de Vehículo, Robo Comercial y Robo Domiciliario entre los años 2013 y 2016, a través de ejercicios separados y exposición de resultados de manera independiente; esto con el fin de identificar el comportamiento exclusivo en cada caso y su respuesta al modelo de red con una capa oculta planteado con cada una de las funciones de activación a probar. El proceso propuesto consiste en los siguientes pasos:

1. Planeación estructural de la información. Se revisó la conformación estructural de las distintas fuentes de datos, comprendidas por las matrices de incidencia, las variables evaluadas por CONEVAL para la evaluación de condiciones socioeconómicas de la población de la zona metropolitana de San Luis Potosí, y los registros generales ambientales de la zona.
2. Para la recopilación de datos se extraen, en primera instancia, los datos de los delitos de Robo de Vehículo, Robo Comercial y Robo Domiciliario de los años 2013 a 2016 de los registros procesados por la Unidad de Análisis, en esta instancia los datos llevan ya un ajuste en la estructura en que son almacenados.
3. Diseño y conformación de un vector con las variables definidas como elementos significativos de cada uno de los rubros de origen: incidencia, condiciones socioeconómicas y condiciones ambientales.
4. Tratamiento de la información, en inicio ha sido preciso normalizar la estructura de los datos de incidencia delictiva, cuyo origen matricial ha ido evolucionando

agregando nuevas variables y aumentando su especificidad, selección de variables relevantes de los registros socioeconómicos y ambientales.

5. Determinación de zonas geográficas específicas, compendio de zonas geográficas de la zona metropolitana, siendo este un conjunto particular para el estudio, dado que los orígenes de datos no contemplan una misma representación.
6. Conformar un universo de datos con una estructura homogénea.
7. Procesar el vector a través del lenguaje R con la rutina predefinida con la función de activación establecida: logística. Llevar a cabo la selección del Umbral, la recomendación habitual para clasificación de valores binarios es 0.5, en este caso se trata de predicción de valores con base en valores estadísticos diversos, a lo que la documentación del lenguaje recomienda valores más pequeños, por lo que se probarían los valores de 0.1, 0.05 y 0.01. Y en su caso determinar la más efectiva para los propósitos del estudio.
8. Explotación de los datos.
 - Generar las gráficas correspondientes e interpretar los resultados obtenidos y la respectiva comparativa de comportamiento de datos reales con los calculados.
 - Plasmar en puntos de concentración georreferenciados, los resultados obtenidos previamente.
 - Evaluar el rendimiento de la información obtenida a través de la retroalimentación de la primera línea de tomadores de decisiones.

3.7 Método de investigación

La naturaleza de la investigación que este estudio presenta es de carácter correlacional, se busca asociar las variables de los grupos señalados y su interacción entre sí para detectar un patrón de comportamiento para los datos registrados históricamente y a través de la definición de dicho patrón, pronosticar una conducta específica a corto y mediano plazo.

Para abordar la problemática se propone la generación de una guía procedimental asentada en un manual y ya esquematizada en la figura 22, cuyo motor de solución sería la aplicación de una red neuronal de una o varias capas, probando distintas funciones de activación para compararlas entre sí y la similitud con el comportamiento real del fenómeno.

La investigación en cuanto a su diseño es puramente documental, se basa en la recopilación de información generada a través de fuentes oficiales, su procesamiento a través de la metodología descrita, su análisis comparativo y la definición y selección de un mecanismo que proporcione de manera sostenida un pronóstico fiable.

3.8 Procesamiento de la información

El procesamiento de la información se llevará a cabo prácticamente con la definición del vector de entrada de elementos de significancia siendo introducido como parámetro a la rutina del lenguaje R, el algoritmo de entrenamiento es retropropagación, sin conexiones hacia atrás, ni laterales. Aplicando corridas por cada configuración propuesta y documentando los resultados. En una segunda instancia se plasmarán los puntos de concentración y se llevará a cabo una comparativa visual y cuantitativa de los resultados

obtenidos con respecto a la incidencia de muestra obtenida del universo de datos y en su caso con datos reales del periodo posterior próximo al evaluado.

3.8.1 Vector de entrada

Construido con las variables que constituyen los elementos de significancia, representa los parámetros de entrada que alimentarán la red neuronal, dando forma al comportamiento de la red que permitirá realizar las proyecciones necesarias para inferir un patrón, la conformación del vector de entrada puede observarse en la figura 25.

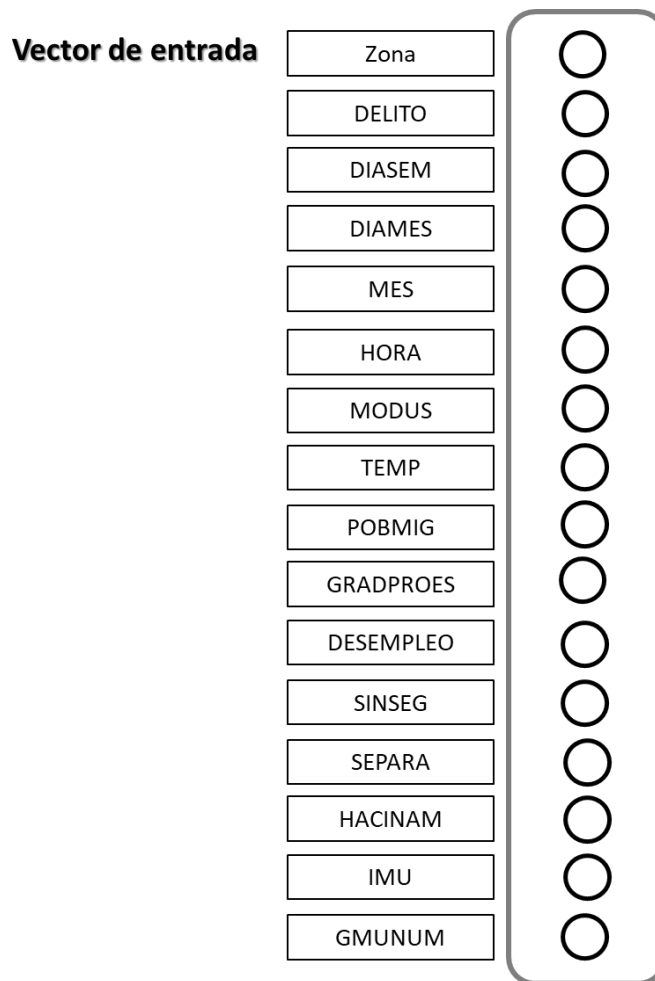


Figura 25: Conformación del vector de datos de entrada para el modelo neuronal.

Fuente: Elaboración propia.

3.8.2 Algoritmo de entrenamiento

La consecución de pasos para el procesamiento del vector de entrada y que define el entrenamiento de la red se describe a continuación.

1. Inicializar los pesos de la red con valores aleatorios pequeños.
2. Presentar el vector de entrada p y especificar la salida deseada.
3. Calcular la salida de la red, esto una vez habiendo presentado las entradas a la red, se debe calcular la salida que devuelve la capa hasta obtener los valores de salida en la capa del mismo nombre. Considerando los siguientes pasos:
 - a. Calcular los valores de entrada netos para la capa oculta procedente de la capa de entrada; para cada neurona oculta:

$$\text{net}_{pj}^{(h)} = \sum_i w_{ji}^{(h)} x_{pi} + \theta_j^{(h)}$$

Ec.(3)

Donde:

$w_{ji}^{(h)}$ es el peso que une el elemento i de la entrada con la neurona j .

x_{pi} es el elemento i del vector de entrada que contiene el patrón de entrenamiento.

$\theta_j^{(h)}$ es el umbral, sesgo o ganancia de la neurona j .

(h) representa la capa a la que pertenece cada parámetro.

- b. Calcular las salidas de las neuronas ocultas (función de activación).

$$y_{pj} = f_j^{(h)}(\text{net}_{pj}^{(h)})$$

Ec.(4)

Donde:

$f_j^{(h)}$ es la función de activación aplicada sobre la entrada neta de cada neurona.

y_{pj} es el valor de salida que representara el valor inicial de estado transmitido a la siguiente capa.

- c. Llevar a cabo los mismos cálculos para obtener los valores de las neuronas de salida.

$$\text{net}_{pk}^{(o)} = \sum_j w_{kj}^{(o)} y_{pj} + \theta_k^{(o)}$$

$$y_{pk} = f_k^{(o)}(\text{net}_{pk}^{(o)})$$

Ec.(5)

4. Realizar el cálculo de los valores de error de todas las neuronas. Si se trata de una neurona en la capa de salida el error se calcula:

$$\delta_{pk}^{(o)} = (d_{pk} - y_{pk}) f_k^{(o)'}(\text{net}_{pk}^{(o)})$$

Ec.(6)

Donde:

$f_k^{(o)'}(\text{net}_{pk}^{(o)})$ (net), es la tasa de cambio o pendiente de la función de transferencia.

d_{pk} , es el valor deseado o que se busca alcanzar.

Por otra parte, si la neurona j no pertenece a la capa de salida, el error es:

$$\delta_{pj}^{(h)} = f_j^{(h)'}(\text{net}_{pj}^{(h)}) \sum_k \delta_{pk}^{(o)} w_{kj}^{(o)} \quad \text{Ec.(7)}$$

Este valor depende de todos los términos de error de la capa de salida.

5. Se actualizan los pesos. Para esto, se comienza de las neuronas de salida y continua hacia atrás hasta la capa de entrada. Los pesos se ajustan de la manera siguiente:

- Para los pesos de las neuronas de salida:

$$w_{kj}^{(o)}(t+1) = w_{kj}^{(o)}(t) + \Delta w_{kj}^{(o)}(t+1)$$

$$\Delta w_{kj}^{(o)}(t+1) = \alpha \delta_{pk}^{(o)} y_{pj} \quad \text{Ec.(8)}$$

- Para los pesos de una capa oculta:

$$w_{ji}^{(h)}(t+1) = w_{ji}^{(h)}(t) + \Delta w_{ji}^{(h)}(t+1)$$

$$\Delta w_{ji}^{(h)}(t+1) = \alpha \delta_{pj}^{(h)} x_{pi} \quad \text{Ec.(9)}$$

Donde:

α , es la razón de aprendizaje, la cual tiende a ser un número diminuto.

A una mayor tasa de aprendizaje, la modificación en los pesos es mayor en cada iteración, lo que representa un aprendizaje más rápido.

6. El proceso se repite continuamente hasta que el error resulta lo suficientemente pequeño para cada uno de los patrones aprendidos.

El fundamento matemático a detalle excede el ámbito del presente documento, sin embargo, las fórmulas son las generalizadas en la mayor parte de la bibliografía consultada. Por otra parte, Javier García (García, 2015) ofrece la explicación punto a punto del cálculo de los valores de estado de las neuronas, los pesos y del error del algoritmo de retropropagación utilizado.

3.9 Análisis de la información

Para el análisis de información resultante se tomarán en cuenta las gráficas generadas a través del lenguaje R, generalmente gráficos de dispersión, con las cuales se buscará reconstruir matrices montando los datos resultantes del procesamiento que representen inferencias del patrón generado a través del entrenamiento de la red neuronal.

Una vez hecho esto, se procederá a la comparación de los resultados de los registros procesados contra los datos de muestra excluidos durante el proceso de aprendizaje de la red y en su caso con aquellos datos reales generados de la incidencia presentada en el compendio de datos de 2017.

3.10 Presentación de la información, técnicas empleadas

La información obtenida de la herramienta de análisis se incorporará como primer acercamiento a la representación de los datos, esto es, los gráficos de líneas apiladas y su línea de comportamiento.

Así mismo, se emplearán matrices de datos para visualizar la distribución de un determinado elemento de significancia respecto a los atributos de las otras variables.

Y se considera la implementación de mapas de calor o bien de puntos de concentración para comparar las inferencias con los registros de prueba.

Con lo anterior será posible conformar un reporte final que contraste la diferencia entre la información procesada y su par real. Adjuntando los resultados del procesamiento, su representación, comparativa, contraste y margen de error.

Capítulo 4: Análisis de datos, pronósticos obtenidos y presentación de resultados

4.1 Diseño experimental

Como se describe al comienzo del presente documento, la propuesta busca el establecimiento de un modelo formal para el tratamiento de información relacionado a la incidencia delictiva, y que dentro de las limitantes naturales de la misma, permita el pronóstico del comportamiento de eventos criminales en determinada región dentro de la zona metropolitana de San Luis Potosí, a través del análisis de los factores históricos, socioeconómicos, temporales y ambientales de la misma, esto a través del procesamiento del compendio informacional por medio de un mecanismo automatizado.

Para lograr establecer el mecanismo procedimental para el procesamiento de la información se estableció la aplicación de algoritmos de redes neuronales a la información compilada por la Unidad de Análisis de información de la Fiscalía General del Estado de San Luis Potosí, así como aquella relacionada a variables socio económicas y ambientales que pudieran influir en el modelo.

El enfoque abordado por este trabajo consiste en la integración de un vector de información compuesto por variables de influencia sobre la ocurrencia de eventos delictivos.

Los factores comúnmente abordados por la literatura especializada incluyen factores urbanos relacionados a la comisión de delitos tales como los demográficos, económicos; factores espacio temporales como la ubicación (distritos comerciales, residenciales, suburbios) y de tiempo como pueden ser horario del día, día de la semana, estacionalidad

y Factores de información contextual del ambiente, que incluyen la relación entre la comisión de actividades delictivas con la apariencia del vecindario y la seguridad que representa el nivel de desorganización del ambiente social (Kang & Kang, 2017)

Respecto a estos factores, y al caso que corresponde el presente estudio fue posible encontrar las fuentes necesarias para factores urbanos, factores espacio temporales y de condiciones ambientales como tal, no así para los factores de información contextual del ambiente social.

Para lo cual, a continuación, se detalla el tipo de información con la que el estudio nutre el modelo y de la cual se espera encontrar la correlación suficiente para poder esquematizar una metodología fiable para el procesamiento de datos y un grado confiable de predictibilidad de eventos en el futuro a corto plazo.

- Colección de información

Las fuentes de datos empleadas son tanto de carácter público como restringido. En primera instancia se han obtenido la base de registros de Área Geoestadística Básica, a través de la aplicación en línea “Espacio y Datos México” que se agrupa en dos grandes conjuntos: El estadístico El conformado por los datos del Directorio Estadístico Nacional de Unidades Económicas (DENUE), el Inventario Nacional de Viviendas y los indicadores de interés gubernamentales (Agricultura, ganadería y forestal, Población, Empleo y ocupación, Hogares y Vivienda, Gobierno, Seguridad Pública y Justicia y Medioambiente). Y, por otro lado, el Geográfico, que incluye demarcaciones territoriales con su respectivo trazado, tales como: División territorial, Topografía, Nombres Geográficos, Sitios de Interés y Recursos Naturales. La figura 26 muestra el aspecto general de la aplicación de Espacio y Datos de México, en línea.

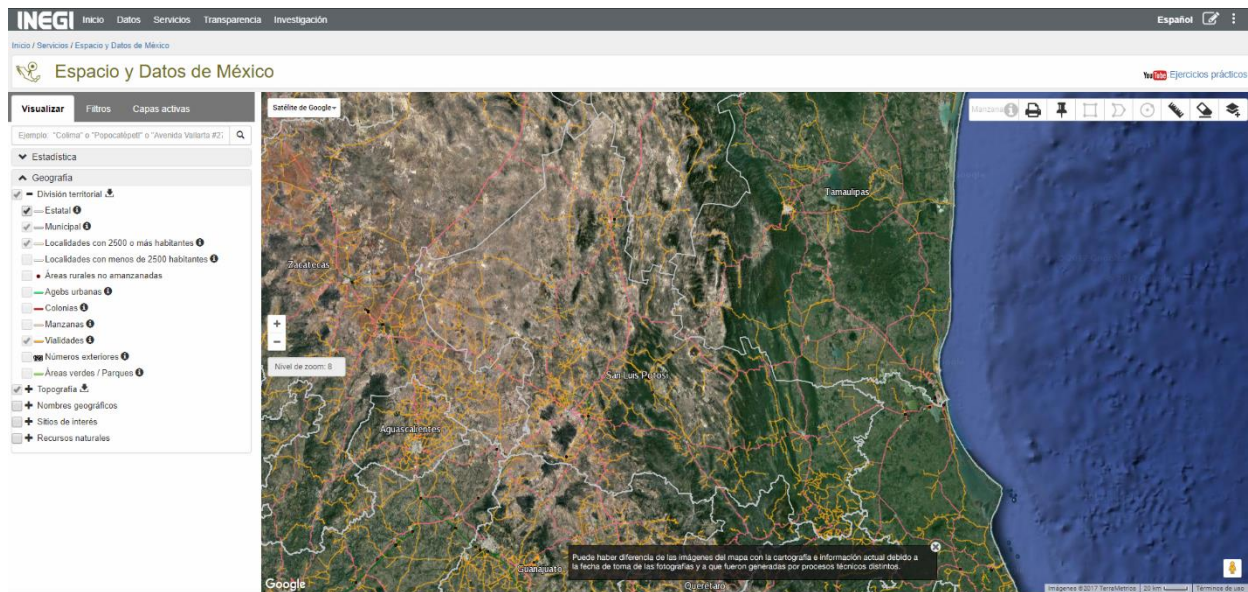


Figura 26: Interfaz principal de Espacio y Datos de México (INEGI)

Fuente: INEGI.

Información contenida en el enlace:

<http://www.beta.inegi.org.mx/app/mapa/espacioydatos/>

Además, ha sido posible obtener una base de datos de clasificación por Área Geoestadística Básica (AGEB) para la medición de la pobreza, proveyendo una escala de clasificación de tres niveles (baja, media, alta), siendo esta una fuente generada por el Consejo Nacional de Evaluación de la Política de Desarrollo Social (CONEVAL).

La tabla de información está compuesta por distintos rubros como son: claves de identidad geográfica, población, viviendas habitadas, porcentajes respecto a acceso a la educación, salud, vivienda, calidad de vivienda y servicios en vivienda para determinar una variable categórica dependiente de tres niveles denominada Grado de Rezago Social. La composición de dicha tabla se muestra en la tabla 5.

Tabla 5. Medición de la pobreza por AGEB's de acuerdo al Grado de Rezago Social, de acuerdo a estudio de CONEVAL para el año 2015

Folio de la AGEB	Población total	Viviendas particulares habitadas	Población de 15 años y más con educación básica incompleta	Población de 15 a 24 años que no asiste a la escuela	Población sin derecho a asistencia a servicios de salud	Personas que viven en hacinamiento	Viviendas que no disponen de excusado o sanitario	Viviendas que no disponen de lavadora	Viviendas que no disponen de refrigerador	Viviendas que no disponen de teléfono fijo	Población de 15 años o más analfabeta	Población de 6 a 14 años que no asiste a la escuela	Viviendas con piso de tierra	Viviendas que no disponen de agua entubada de la red pública	Viviendas que no disponen de drenaje	Viviendas que no disponen de energía eléctrica	Grado de rezago social (clases latentes)
2402800010172	3154	883	19.6	36.7	20.7	0.3	0.1	10.9	1.6	15.2	1.2	2.7	4.2	0.0	0.0	0.0	Bajo
2402800010187	1948	581	19.4	33.7	26.6	0.2	0.3	7.4	2.6	13.6	0.8	2.8	1.2	0.0	0.0	0.0	Bajo
2402800010191	1682	465	21.5	26.3	28.0	0.2	0.0	7.1	1.7	12.5	0.4	2.3	0.2	0.2	0.0	0.0	Bajo
2402800010223	1432	437	20.0	34.9	20.9	0.0	0.0	7.6	0.7	13.3	1.4	1.2	0.2	0.2	0.0	0.2	Bajo
2402800010238	5274	1418	17.2	31.7	23.6	0.4	0.2	8.0	2.1	19.9	1.1	1.8	1.0	0.3	0.1	0.1	Bajo
2402800010242	5313	1323	29.4	43.4	25.6	1.2	0.2	13.1	4.8	28.6	2.8	1.9	1.0	0.0	0.0	0.2	Bajo
2402800010257	4423	1072	31.7	49.9	24.0	1.5	0.2	15.9	7.6	41.7	3.3	3.1	1.8	0.8	0.2	0.2	Bajo
2402800010261	2404	638	33.3	44.2	28.9	1.4	0.2	24.0	11.1	42.9	2.8	1.5	0.3	0.0	0.3	0.3	Bajo

Fuente: Consejo Nacional de Evaluación de la Política de Desarrollo Social (CONEVAL).

El contenido se encuentra disponible en el enlace:

http://www.coneval.org.mx/Informes/Pobreza/Rezago_Social/Rezago_Social_2010/Rez_soc_AGEB/RS_e_dos_2010/San%20Luis%20Potosi.zip

Además, fue posible obtener una base de datos respecto de las temperaturas presentadas en territorio potosino, esta información se extrajo por acceso público a través del portal del Instituto Nacional de Investigaciones Forestales, Agrícolas y Pecuarias (INIFAP) de la Secretaría de Agricultura, Ganadería, Desarrollo Rural, Pesca y Alimentación (SAGARPA). Información registrada en la estación INIFAP San Luis.

La información fue concentrada en un archivo único dado que las consultas generadas a través del acceso en línea solamente permiten consultas mensuales, por lo que se llevaron a cabo consultas desde el 1 de enero de 2011 hasta el 17 de septiembre de 2017, aplicables a la estación local en la zona metropolitana de San Luis Potosí.

La base está compuesta por variables tales como fecha del registro, temperatura máxima, temperatura mínima, temperatura media, velocidad del viento máxima, dirección de la velocidad del viento máxima, velocidad promedio del viento, dirección promedio del viento, radiación global, Humedad relativa, evapotranspiración de referencia y evaporación potencial. La tabla 6 ejemplifica una muestra de la composición de dicha base.

Tabla 6. Base de datos de temperaturas de 2011 a 2017 para la zona metropolitana de San Luis Potosí, de acuerdo a registros de INIFAP

Fecha	Pp	T. Máx.	T. Min.	T. Med.	VV Máx.	VV	DVV Máx.	DV	Rad. G.	HR	ET	EP
01/01/2011	0	25	-4.9	11.19	21.7	5.08	248.20 (Oeste)	236.20 (Suroeste)	446.76	28.51	3.7	6.43
02/01/2011	0.3	19.4	2.1	9.64	12.3	3	136.20 (Sureste)	275.80 (Oeste)	242.86	73.26	1.5	3.16
03/01/2011	0.1	24.1	6.1	11.79	9.7	3.83	198.90 (Sur)	227.30 (Suroeste)	405.42	74.88	2.5	4.53
04/01/2011	0	25.2	-1.2	12.82	8.5	2.19	310.90	213.20 (Suroeste)	456.26	59.5	3.5	5.17
05/01/2011	0	26.9	-0.6	13.76	13.1	2.58	337.10	329.10 (Noroeste)	467.02	48.34	3.6	5.59
06/01/2011	0	25.7	2.1	12.75	15	2.81	115.90 (Sureste)	123.40 (Sureste)	403.06	60.81	2.9	4.73
07/01/2011	0	27.5	0	12.15	14.6	2.65	301.70	225.70 (Suroeste)	342.39	56.09	2.9	4.46
08/01/2011	0	25.8	-0.5	14.38	25.3	9.53	265.80 (Oeste)	247.80 (Oeste)	441.31	34.54	5.1	ND
09/01/2011	0	25.6	7.9	16.74	24.8	10.61	264.00 (Oeste)	250.30 (Oeste)	509.33	24.92	5.4	ND
10/01/2011	0	25.9	-3.8	10.47	8.5	2.61	214.60	242.30 (Suroeste)	505.25	35.93	4	5.97
11/01/2011	0	18.9	-2.1	9.27	23.6	11.04	111.40 (Este)	127.40 (Sureste)	261.25	64.54	1.9	4.22
12/01/2011	0	12.6	6.6	8.65	24.2	15.36	123.30 (Sureste)	125.00 (Sureste)	278.98	76.22	1.6	4.42
13/01/2011	0	13.7	6.7	9.05	20.2	11.5	103.00 (Este)	119.60 (Sureste)	260.72	73.66	1.5	4.3
14/01/2011	0	21.9	3.4	11.51	11.5	3.65	1.10 (Norte)	112.40 (Este)	470.25	61.54	3.2	5.55

Fuente: Instituto Nacional de Investigaciones Forestales, Agrícolas y Pecuarias (INIFAP).

Información histórica alojada en el enlace de Internet:

<http://clima.inifap.gob.mx/redinifap/historicos.aspx?est=36755&edo=23&d=10&m=10&an=2017>

Por otro lado, la información procesada tanto por el sistema de emergencias como la compilada por las denuncias presentadas ante el Ministerio Público, este conjunto es recopilado por el personal de la Unidad de Análisis de la Fiscalía General del Estado, y almacenado en tablas de datos por delito específico, las cuales representan el insumo básico para la generación del producto de inteligencia denominado Informe Semanal de

Incidencia y el cual concentra los principales delitos que ocurren en la mancha urbana de San Luis Potosí.

Los eventos de robo de vehículo, por ejemplo, son reportados al sistema de emergencias, el cual genera una base de datos la cual es consultada y se obtiene una tabla con los registros de reporte de robo y recuperación de vehículos, estos a su vez son cotejados con las denuncias efectivas ante el Ministerio Público, obteniendo una tabla depurada con datos como fecha de ocurrencia, hora, día de la semana, colonia, sector, latitud y longitud, tipo vehículo (categorías), marca, submarca, modelo y color.

Se incluyen otros dos delitos patrimoniales como son robo comercial y robo domiciliario por representar los índices más altos de incidencia, dando un total de 10222 registros.

La construcción del vector de entrada está sujeta a cambios por la adaptación del algoritmo de entrenamiento y la influencia de las variables independientes sobre el vector de salida. Lo que se observara en las primeras etapas del proceso de experimentación.

- Selección de datos

La incorporación de la totalidad de los datos se torna complicada dada la cantidad de variables que habría que introducir al modelo.

Esto debido a la gran cantidad de campos que cada una de las bases de datos recolectada contiene y que no necesariamente impacta en la ocurrencia de algún evento delictivo o que simplemente no cumple con los elementos necesarios para ser añadidas al modelo.

Además, debe tomarse bajo consideración la expansión natural del vector de entrada al buscar normalizar las variables categóricas y homogenizar las entradas de la red neuronal, por lo cual se seleccionaron los siguientes campos de las tablas correspondientes para conformar una matriz general de procesamiento. Es decir, las entradas tener de una conformación homogénea para poder incidir en el resultado producido.

- Modelo de procesamiento.

Para el procesamiento de la información se hace uso de un paquete de redes neuronales del lenguaje R denominado “neuralnet”. La descripción del paquete indica que consiste en “el entrenamiento de redes neuronales usando algoritmos de propagación hacia atrás, propagación hacia atrás resiliente con o sin retroceso de pesos, o la versión modificada globalmente convergente de Anastasiadis. Permite la configuración flexible a través de la selección personalizada del error y de la función de activación. Además, el cálculo de pesos generalizados está implementado” (Fritsch & Guenther, 2016).

La implementación del código requiere pocas líneas para la elaboración de un programa efectivo, las funciones principales son *neuralnet* para el procesamiento del entrenamiento de la red a través de un conjunto de datos indicando además los parámetros básicos para que se ejecute la red de entrenamiento, la función devuelve una matriz que contiene el umbral alcanzado, los pasos necesarios, el error, los criterios de información como AIC y BIC, y los pesos para cada repetición. La figura 27 muestra la composición básica de la función *neuralnet*, donde RED es una variable que recibe la salida de la función y la tabla 7 la lista de argumentos que puede recibir la función *neuralnet*.



Figura 27: Composición de la llamada de la función *neuralnet* con los argumentos utilizados más frecuentemente.

Fuente: Elaboración propia.

Tabla 7. Lista de argumentos de la función *neuralnet*

Lista de argumentos de la función *neuralnet*

- **Formula** Descripción simbólica del modelo a adaptar.
- **Datos** Un conjunto de datos que contiene las variables especificadas en la fórmula.
- **#Ocultas** Un vector de enteros que especifica la cantidad de neuronas ocultas en cada capa.
- **Umbral** Valor numérico que especifica el umbral para las derivadas parciales de la función de error como parámetro de interrupción.
- **PasosMáx** El número máximo de pasos para el entrenamiento de la red neuronal.
- **Rep** Cantidad de repeticiones para el entrenamiento de la red neuronal.
- **PesosIniciales** Vector que contiene los valores iniciales para los pesos.
- **Algoritmo** Cadena que indica el tipo de algoritmo para calcular la red neuronal: "backprop", "rprop+", "rprop-", "sag", "slr".
- **FuncErr** Función diferencial usada para el cálculo del error: "sse", para suma de errores cuadrados y "ce" para entropía cruzada.
- **FuncAct** una función diferenciable que se utiliza para suavizar el resultado del producto cruzado de la covariable o las neuronas y los pesos. Además, las cadenas, 'logistic' y 'tanh' son posibles para la función logística y la tangente hiperbólica.

Fuente: Elaboración propia.

Así mismo, la función *compute* es un método para objetos de la clase nn (red neuronal), generalmente producidos por la función *neuralnet*. La cual procesa y calcula el objeto dado para vectores covariables dados. La cantidad de variables en la nueva matriz o dataset debe corresponderse con la cantidad original de variables con que la red neuronal original fue entrenada. La composición de la función 'compute' se muestra en la figura 28 y las listas de argumentos y valores de salida en la tabla 8.

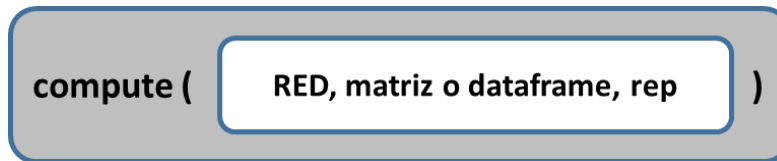


Figura 28: Estructura de la función compute.

Fuente: Elaboración propia.

Tabla 8. Descripción de argumentos y valores de salida de la función “compute”

Lista de argumentos de la función compute

- RED Objeto de clase nn (red neuronal).
- Covariate Un dataframe o matriz que contiene las variables que han sido usadas para entrenar la red.
- rep Número entero que indica la cantidad de repeticiones a usar por la red neuronal.

Lista de valores de salida de la función compute

- neurons Lista de neuronas de salida para cada capa de la red neuronal.
- net.result Matriz que contiene el resultado global de la red neuronal.

Fuente: Elaboración propia.

Como se planteó en la sección 3.8, la selección de datos que conforman el vector inicial de entrada debe ser introducido como un conjunto de datos para obtener de la función neuralnet un objeto nn, el cual representa el entramado de base producido por el entrenamiento de la red neuronal.

La función ‘compute’ recibe tanto el objeto nn, como un conjunto de datos a evaluar, dando como resultado un listado, representando el mapa neuronal generado, así como la matriz de resultados globales del procesamiento de los datos evaluados.

El proceso es prácticamente transparente para el usuario radicando la efectividad del modelo en la construcción del vector de entrada con los datos de entrenamiento, la correcta selección de variables, la correlación de las mismas con el resultado y la distribución de los datos determina la calidad del desempeño de la red.

4.2 Conformación de conjunto de datos y experimentación

Antes de proceder al procesamiento de datos a través de la red neuronal es preciso establecer un conjunto homogéneo de los mismos estructuralmente hablando, es decir, la información recabada debe presentar características similares para cada rubro a evaluar (variables), tanto en el tipo de valor que presenta cada uno como en los rangos en los que confluyen dichos valores.

4.2.1 Condiciones iniciales y preprocesamiento de los datos

Si bien es cierto se tuvo acceso a fuentes primarias y de carácter oficial en los entornos de interés para el presente desarrollo. La información obtenida de las distintas fuentes (datos espaciales por INEGI, registros de incidencia delictiva por la Fiscalía General, condiciones socioeconómicas conformadas por CONEVAL y registros de temperatura promedio en la zona metropolitana de San Luis Potosí recabados por INIFAP), fue conformada en una matriz general.

Para las cuestiones geoespaciales se contaba de inicio con un archivo shape (formato de archivo informático propietario de datos espaciales desarrollado por la compañía ESRI y es bastante popular en el manejo de datos geoespaciales) en el cual vienen definidos los límites de las colonias que componen la zona metropolitana de San Luis Potosí, éste fue originalmente creado por la Dirección de Catastro y aportado por INEGI a la Fiscalía General, el mismo no cuenta con soporte por este último debido a que el rubro de colonia no es un factor de medida para la labor del Instituto, por lo que debió ser actualizado de manera manual con apoyo de información pública como lo es la albergada en el servicio Maps de Google Inc., o la información de códigos postales y colonias del Servicio Postal Mexicano (SEPOMEX), lo que representó parte importante del tiempo invertido en el

desarrollo del estudio, debiendo verificarse y actualizarse la presencia de nuevas colonias y fraccionamientos que se habrían observado al extraer entidades únicas del dato colonia de las matrices de incidencia delictiva generadas por la Fiscalía General. Por otra parte, INEGI permite descargar el aplicativo Datos México y cuenta con el servicio en línea Mapa Digital de México que permite obtener archivos Shape con datos espaciales y estadísticos de regiones de todo el país, basándose en un elemento distinto a la colonia, pero de mayor precisión y utilidad como lo es el AGEB (Área Geoestadística Base).

La aportación de INEGI respecto a registro de colonias de la zona metropolitana de San Luis Potosí se observa en la Figura 29.

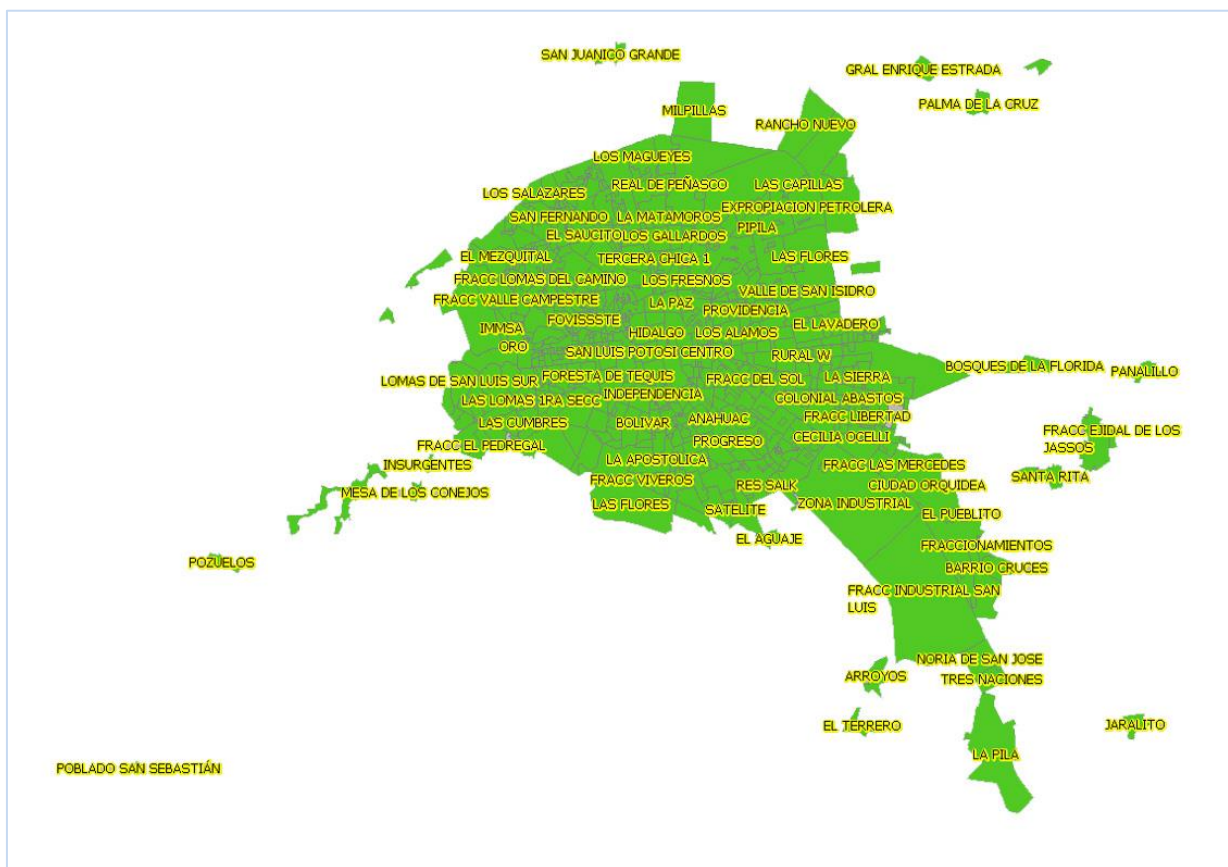


Figura 29: Representación gráfica de la base de datos de colonias.

Fuente: INEGI.

Y la otra conteniendo los registros de AGEB's en la Figura 30, obtenida a través del portal de la misma institución, en la aplicación denominada Mapa Digital de México, en el enlace:

<http://gaia.inegi.org.mx/mdm6/?v=bGF0OjlyLjEyNyY1LGxvbjotMTAwLjkzMzQwLHo6OCxsOmNkZW51ZT>

E0

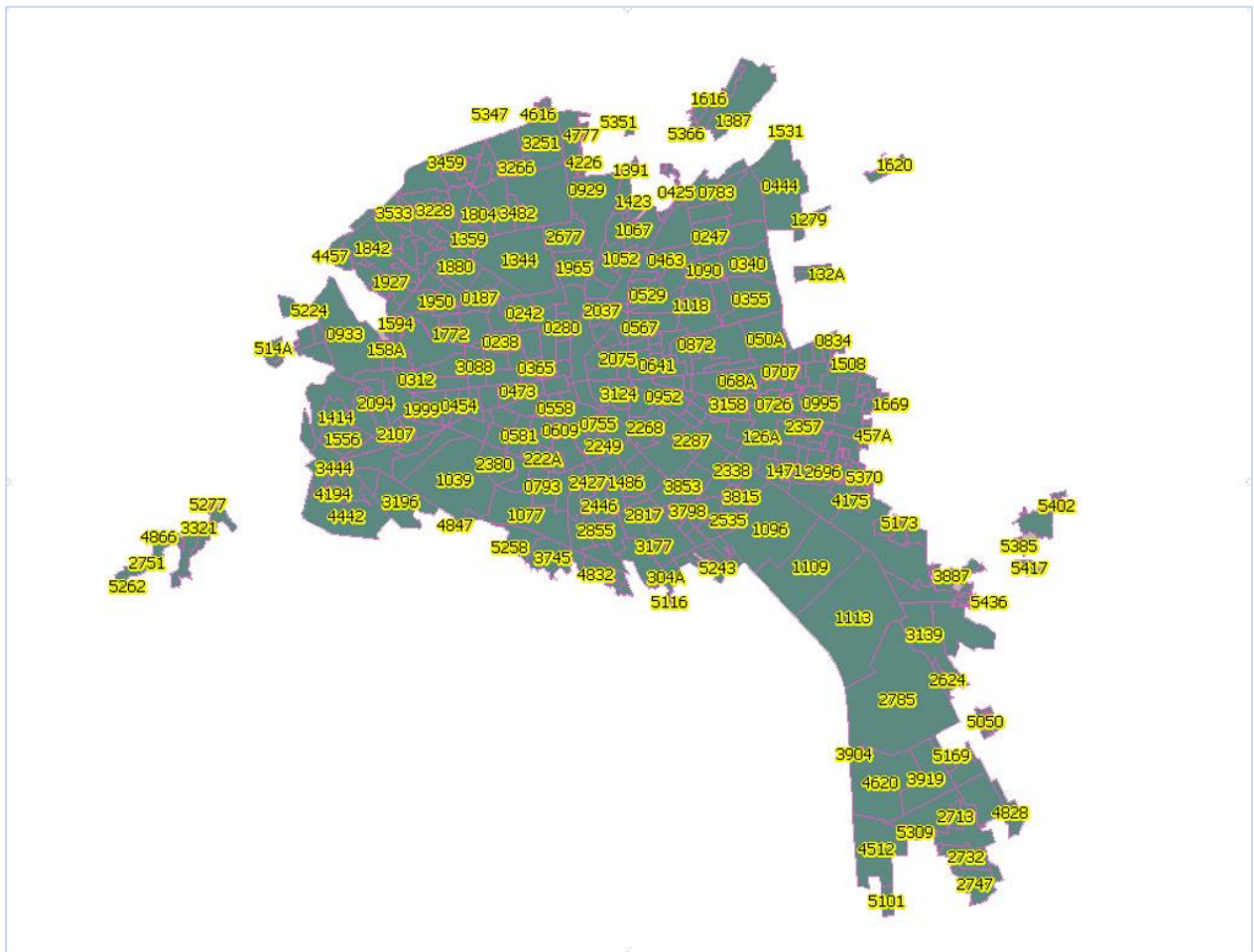


Figura 30: Representación gráfica de la base de AGEB's.

Fuente: INEGI.

Se llevó a cabo una aproximación de la correspondencia entre ambos tipos de zonificación a través de una consulta cruzada entra ambas bases de datos con la herramienta de software libre QGIS, Figura 31.

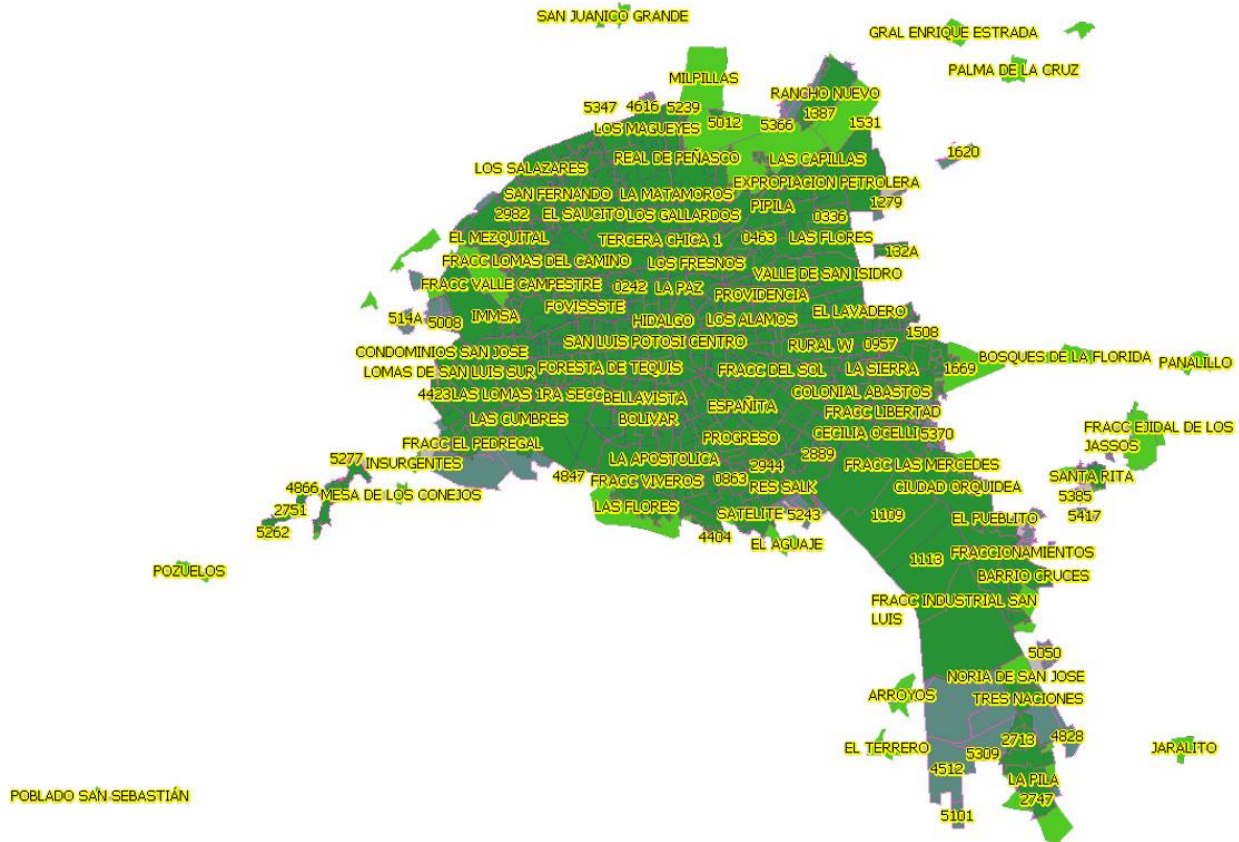


Figura 31: Representación gráfica de la conjunción de las bases de datos de colonias y AGEB's.

Fuente: Elaboración propia con datos de INEGI.

El procedimiento de conjunción espacial en un sistema de información geográfica (SIG), es un proceso que requiere que los archivos manejen un sistema de despliegue de coordenadas compatible, para poder lograr que, por decirlo de una forma, ambos se superpongan en la misma ubicación, como si de dos capas de transparencias sobre una

superficie común se tratara, sin embargo, internamente se logra también una unión de los atributos que las componen.

Cabe resaltar que de la conjunción de los archivos geospaciales se obtuvo que, al revisar el conjunto de colonias contemplado en la base de eventos de la Fiscalía, de los 10222 registros que contenía originalmente, se generó correspondencia en 10078 registros, excluyendo 144 registros que pertenecen a zonas rurales conurbadas (delegación de Bocas, delegación de Pozos, y Cerro de San Pedro) que aun que teniendo colonias urbanizadas, no se cuenta con datos socioeconómicos atribuibles a un AGEB para completar su procesamiento.

Respecto a los datos de la Fiscalía General, en el caso del material de datos de incidencia delictiva, la conformación de una matriz estable que pudiera ser evaluada resultó en una actividad muy laboriosa, esto debido a la falta de una estructura estandarizada de los registros recabados desde el año 2013 al año 2016, aunque los datos de dicho material se registran desde 2011, fue evidente la carencia de un modelo general para poder evaluar la información o realizar comparativas entre periodos de distintos años, esto a la par de otras problemáticas como son los parámetros específicos de algunas variables como puede ser en el caso de las variables de tiempo como fecha y hora en los cuales pudieron observarse registros de rangos horarios y horario determinado sobre la misma variable, incluso acrónimos para determinar secciones del día, o bien el registro de fecha en formatos no compatibles, para citar solo algunos.

Posteriormente, una vez identificadas las variables comunes y el tipo y rango de datos que pueden recibir, se procedió a la conformación de una matriz general de incidencia

delictiva en la cual convergen los delitos patrimoniales de Robo Comercial, Robo Domiciliario y Robo de Vehículos, constituyendo un conjunto total de 10222 registros, de los cuales sólo eran analizables 10078 y estableciendo las variables generales que se muestran en la tabla 9. Las variables seleccionadas son aquellas que son cuantificables, descartando aquellas que contienen valores alfanuméricos, códigos de referencia, estatus e información personal de los afectados, ésta última eliminada desde el primer paso de la normalización debido a que no es posible difundir ese tipo de datos. Las variables elegidas fueron normalizadas y estandarizadas bajo un mismo tipo de datos en todas las matrices y al final vaciadas a una tabla concentradora donde posteriormente se anexarían los datos socioeconómicos y ajustado el dato de referencia espacial (colonia y municipio) por el de AGEB.

Tabla 9. Variables de la matriz de Incidencia

Variable	Descripción
DELITO	Delito específico, valores respectivos de 1:Robo comercial, 2:Robo domiciliario, 3: Robo de vehículo.
COLONIA	Denominación de una región geográfica mínima utilizada de manera general para la ubicación un conjunto de domicilios dentro de una ciudad o población específica.
MUNICIPIO	Entidad administrativa que agrupa una o más localidades y que puede hacer referencia a una ciudad, pueblo o aldea.
DIASEM	Denominación del día específico dentro de una semana, valores entre 1 y 7, siendo 1 el correspondiente al Domingo.
DIAMES	Número correspondiente del mes de un día específico, valores entre 1 y 31.
MES	Cada uno de los 12 periodos de tiempo, de entre 28 y 31 días, en que se divide un año. Valores de 1 a 12, correspondiendo el 1 al mes de Enero y así sucesivamente.
HORA	Unidad de tiempo que se corresponde con la vigésimo-cuarta parte de un día solar medio. Valores de 1 a 24.
MODUS	Valor binario (0 ó 1) para establecer si el evento fue llevado a cabo con violencia o no.

Fuente: Elaboración propia.

A la tabla anterior, se adjuntó una sola variable de los registros de condiciones ambientales, la temperatura, utilizando como punto de convergencia la fecha específica. La determinación a utilizar solamente el valor de temperatura fue debido a que una cantidad considerable de los registros por día de los datos obtenidos por INIFAP no presentaba valores evaluables para las variables de velocidad y dirección del viento, radiación global, humedad relativa, evapotranspiración y punto de rocío, y considerando únicamente la temperatura media, descartando la máxima y la mínima para reducir la complejidad del modelo al momento de ser procesado. Incorporando la variable de temperatura se obtuvo la tabla 10.

Tabla 10. Tabla de incidencia incluyendo temperatura

Variable	Descripción
DELITO	Delito específico, valores respectivos de 1:Robo comercial, 2:Robo domiciliario, 3: Robo de vehículo.
COLONIA	Denominación de una región geográfica mínima utilizada de manera general para la ubicación un conjunto de domicilios dentro de una ciudad o población específica.
MUNICIPIO	Entidad administrativa que agrupa una o más localidades y que puede hacer referencia a una ciudad, pueblo o aldea.
DIASEM	Denominación del día específico dentro de una semana, valores entre 1 y 7, siendo 1 el correspondiente al Domingo.
DIAMES	Número correspondiente del mes de un día específico, valores entre 1 y 31.
MES	Cada uno de los 12 periodos de tiempo, de entre 28 y 31 días, en que se divide un año. Valores de 1 a 12, correspondiendo el 1 al mes de Enero y así sucesivamente.
HORA	Unidad de tiempo que se corresponde con la vigésimo-cuarta parte de un día solar medio. Valores de 1 a 24.
MODUS	Valor binario (0 ó 1) para establecer si el evento fue llevado a cabo con violencia o no.
TEMP	Temperatura promedio en la zona metropolitana de San Luis Potosí, valores en grados centígrados.

Fuente: Elaboración propia.

En el caso de las variables socioeconómicas evaluadas por CONEVAL, del total de 29 variables, se descartaron las correspondientes a condiciones específicas relacionadas a características de la vivienda y acceso a servicios y aquellas relacionadas a estado civil de la población y nivel de mortalidad de la progenie. Se consideran en el estudio la ubicación geográfica (AGEB), porcentaje de población migrante, grado promedio de estudios, tasa de desempleo, porcentaje de población sin acceso a servicios de salud, porcentaje de población separada, porcentaje de viviendas con hacinamiento, índice de marginación urbana y el grado de marginación urbana (Tabla 11).

Tabla 11. Tabla de variables socioeconómicas

Variable	Descripción
AGEB	Área Geoestadística Base. Valor Alfanumérico.
POBMIG	Porcentaje de población urbana no nacida en la localidad.
GRAPROES	Grado promedio de escolaridad.
DESEMPLEO	Porcentaje de población sin empleo.
SINSEG	Porcentaje de la población sin derechohabiencia a servicios de salud.
SEPARA	Porcentaje de la población separada o divorciada.
HACINAM	Porcentaje de viviendas habitadas con algún nivel de hacinamiento.
IMU	Índice de marginación urbana.
GMUNUM	Grado de marginación urbana (1:Muy bajo, 2: Bajo, 3: Medio, 4: Alto, 5: Muy Alto).

Fuente: Elaboración propia.

Finalmente, una vez realizados los ajustes geospaciales y asociados a coordenadas físicas, dieron cabida a la inclusión del concepto de zona, esto debido a que la variable AGEB aunque permite una identificación clara para el análisis, siendo una variable alfanumérica produciría errores al momento de ser procesada por el algoritmo de redes neuronales por lo que se llevó a cabo una consecución numérica de los AGEB correspondiéndose de manera incremental del 1 hasta 277, siendo este el total de zonas. Por lo que al final se constituyó un vector único de entradas.

Para establecer la probabilidad general de cada evento fue preciso aplicar la Regla General de la adición o regla de la suma de probabilidades para eventos no mutuamente excluyentes que establece que para obtener la probabilidad de ocurrencia de cualquier evento se aplica la regla para calcularla definida en la ecuación 10.

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

Ec. 10

Sin embargo, se consideró el conjunto de datos intrínsecos al delito para llevar a cabo la conformación correspondiente, presentar la fórmula desarrollada para siete variables resulta poco práctico e incluso gráficamente resulta complicado mostrar la conjunción de dichos elementos, a grandes rasgos, la ecuación considera la suma de todas probabilidades de cada variable individual y de la ocurrencia simultánea de todas ellas menos la sumatoria de las probabilidades de ocurrencia simultánea de sus combinaciones parciales, véase la figura 32.

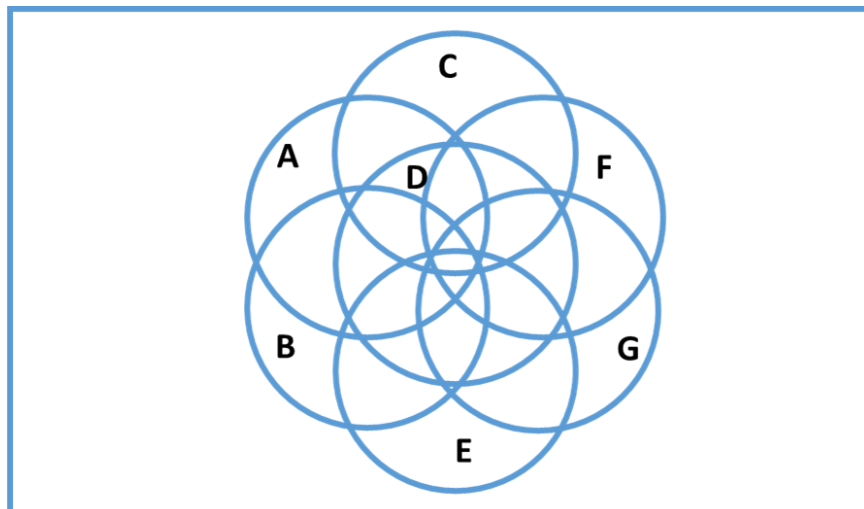
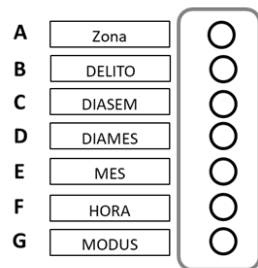


Figura 32: Diagrama de Venn de las variables componentes de la probabilidad de evento.

Fuente: Elaboración propia.

A través de la probabilidad de cada variable particular correspondiente con la incidencia de cada evento, como características no mutuamente excluyentes del mismo, se llevó a cabo el cálculo específico de cada probabilidad del evento particular realizando la adición de las probabilidades de las variables que lo componen.

En este caso, fue preciso hacer ajustes en el caso de variables como lo es el día del mes bajo el considerando de que los días 29, 30 y 31 no están presentes en todos los meses y de que existe en el periodo evaluado al menos un año bisiesto.

El valor obtenido para la probabilidad por evento en la etapa de experimentación complementa el vector de entrada como valor deseado al entrenar la red neuronal.

4.2.2 Experimentación

Dado que se desconoce la correlación de los datos de incidencia delictiva se procede a probar el modelo elaborado variando únicamente dos parámetros de la función neuralnet, en este caso la cantidad de capas ocultas, parámetro “hidden” y el umbral, parámetro “threshold”, utilizando para esto distintas configuraciones de redes neuronales.

La literatura relacionada al tema, recomienda la aplicación de la regla de la pirámide para una red de una sola capa oculta. Véase ecuación 11.

$$h = m * n \quad \text{Ec.(11)}$$

Donde:

- n, es el número de neuronas de entrada.
- m, es el número de neuronas de salida.
- h, es el número inicial de neuronas de la capa oculta.

En este caso, la cantidad de neuronas ocultas a utilizar dentro del modelo de una sola capa, siendo 16 las neuronas y una la de salida, deben considerarse 16.

Para redes con dos capas ocultas las ecuaciones son las que se muestran. Véanse ecuaciones 12 y 13.

Para la primera capa oculta:

$$H_1 = m * r^2 \quad \text{Ec.(12)}$$

Y para la segunda capa oculta:

$$H_2 = m * r \quad \text{Ec.(13)}$$

Donde:

- H1, es el número de neuronas en la primera capa oculta.
- H2, es el número de neuronas en la segunda capa oculta.
- Y el valor de r se calcula con la ecuación 14.

$$r = \sqrt[3]{\frac{n}{m}} \quad \text{Ec.(14)}$$

Para el caso de dos capas ocultas, la primer capa 6 capas neuronas, y en la segunda 3 neuronas ocultas.

Sin embargo, las fórmulas sólo son aproximaciones imprecisas del tamaño ideal de las capas ocultas. En casos donde las funciones son complicadas, se puede requerir una cantidad considerable de neuronas ocultas sobre una misma sola capa.

La regla basada en algoritmos genéticos determina la cantidad de neuronas a utilizar con la ecuación 15.

$$H = \frac{2}{3} n \quad \text{Ec.(15)}$$

Donde:

- H , Es la cantidad de neuronas ocultas.
- n , la cantidad de neuronas de entrada.

Bajo este criterio deben utilizarse entre 10 y 11 neuronas ocultas para el modelo propuesto.

La mayor parte de los casos prácticos se resuelven con una capa oculta, por lo que se recomienda el uso de una sola capa oculta como elección inicial, en caso de usar demasiadas neuronas en la capa oculta se debe incorporar una segunda capa oculta, reduciendo la cantidad de neuronas en cada una de las capas ocultas.

Posterior a pruebas preliminares y para efectos prácticos se han considerado las configuraciones siguientes:

1. Configuración: 1 capa, 1 neurona, umbral 0.1
2. Configuración: 1 capa, 1 neurona, umbral 0.05.
3. Configuración: 1 capa, 1 neurona, umbral 0.01.
4. Configuración: 1 capa, 4n, umbral 0.1.
5. Configuración: 1 capa, 4n, umbral 0.05.
6. Configuración: 1 capa, 4n, umbral 0.01.
7. Configuración: 1 capa, 16n, umbral 0.1.
8. Configuración: 1 capa, 16n, umbral 0.05.
9. Configuración: 1 capa, 16n, umbral 0.01.
10. Configuración: 2 capas, (6n,3n), umbral 0.1.
11. Configuración: 2 capas, (6n,3n), umbral 0.05.
12. Configuración: 2 capas, (6n,3n), umbral 0.01.
13. Configuración: 1 capa, 10n, umbral 0.1. (Algoritmos genéticos).
14. Configuración: 1 capa, 10n, umbral 0.05. (Algoritmos genéticos).
15. Configuración: 1 capa, 20n, umbral 0.1.
16. Configuración: 1 capa, 20n, umbral 0.05.
17. Configuración: 2 capas, (10n,5n), umbral 0.1.
18. Configuración: 2 capas, (10n,5n), umbral 0.05.
19. Configuración: 2 capas, (10n,5n), umbral 0.01.
20. Configuración: 3 capas, (20n,5n,2n), umbral 0.05.

Aunque se probaron más configuraciones que incluían mayor complejidad, los resultados arrojaban demasiados errores, bajo esta consideración, al final de las pruebas, se puede observar que el conjunto de datos no es tan complejo y que las configuraciones medianamente elaboradas resultarían más efectivas, para ofrecer elementos de comparación se acotó la selección de configuraciones a la cifra de 20.

De esta manera, se procede a procesar la información consistente en 10078 registros, de los cuales se toma una muestra de 370 elementos, basando dicha cantidad en el cálculo de muestra con la ecuación 16. Esto considerando un nivel de confianza de 95%, y un intervalo de confianza de 5%.

$$n = \frac{z^2 pq}{e^2 + \frac{z^2 pq}{N}} \text{ Ec. (16)}$$

Donde:

- n , es el tamaño de la muestra.
- Z , es el nivel de confianza deseado, en este caso 95%.
- p , es la proporción de la población con la característica deseada.
- q , es la proporción de la población sin la característica deseada.
- e , es el nivel de error dispuesto a cometer: 5%.
- N , es el tamaño de la población.

Utilizando esos 370 registros como datos a evaluar por la red neuronal y la diferencia con la población, consistente en 9708 registros, se utiliza para llevar a cabo el entrenamiento de la red.

Las rondas de prueba de cada configuración se aplican en un equipo de cómputo con un procesador AMD Ryzen 5 2600, Six-Core a 3.85 GHz, con 16 GB de memoria RAM.

4.2.3 Resultados de la prueba de las configuraciones de red

El tiempo de procesamiento varía respecto a la capacidad del equipo utilizado, como base documental se muestra la tabla 12 con los valores de tiempo obtenidos al ejecutar el modelo con cada una de las configuraciones de red neuronal planteadas.

Tabla 12. Tiempo de procesamiento para cada configuración de la red

#	Configuración	Tiempo de procesamiento (segs)
1	Configuración: 1 capa, 1 neurona, umbral 0.1	0.07
2	Configuración: 1 capa, 1 neurona, umbral 0.05	0.10
3	Configuración: 1 capa, 1 neurona, umbral 0.01	4.94
4	Configuración: 1 capa, 4n, umbral 0.1	1.39
5	Configuración: 1 capa, 4n, umbral 0.05	1.66
6	Configuración: 1 capa, 4n, umbral 0.01	91.78
7	Configuración: 1 capa, 16n, umbral 0.1	37.74
8	Configuración: 1 capa, 16n, umbral 0.05	45.56
9	Configuración: 1 capa, 16n, umbral 0.01	327.55
10	Configuración: 2 capa, (6n,3n), umbral 0.1	0.20
11	Configuración: 2 capa, (6n,3n), umbral 0.05	0.28
12	Configuración: 2 capa, (6n,3n), umbral 0.01	141.13
13	Configuración: 1 capa, 10n, umbral 0.1	3.20
14	Configuración: 1 capa, 10n, umbral 0.05	4.00
15	Configuración: 1 capa, 20n, umbral 0.1	47.05
16	Configuración: 1 capa, 20n, umbral 0.05	65.37
17	Configuración: 2 capas, (10n,5n), umbral 0.1	3.00
18	Configuración: 2 capas, (10n,5n), umbral 0.05	4.69
19	Configuración: 2 capas, (10n,5n), umbral 0.01	290.32
20	Configuración: 3 capas, (20n,5n,2n), umbral 0.05	8.15

Fuente: Elaboración propia.

En primera instancia, las configuraciones se corrieron en un procesador más antiguo y, en el mismo, era evidente que a mayor cantidad de neuronas o capas ocultas el tiempo de procesamiento se elevaba considerablemente. Sin embargo, el equipo donde se corrieran finalmente dichas configuraciones redujo considerablemente los tiempos, haciéndose notorios otros rubros de impacto en el tiempo como puede ser el nivel de umbral aplicado a los datos.

Por otra parte, para comparar el comportamiento de las configuraciones de entrenamiento de la red neuronal, se consideran en primera instancia el coeficiente de variación de los valores arrojados por la red neuronal, y su distribución comparada con el comportamiento de los datos reales, considerando primeramente que si este coeficiente es muy bajo, los valores en cuestión serán demasiado homogéneos representando únicamente un valor promedio que satisface de manera general el modelo, esto es, que no representaran un comportamiento similar al de los valores de la incidencia real.

Asimismo, se debe considerar el porcentaje promedio de similitud entre los valores de incidencia real y aquellos obtenidos a través de la red neuronal, este porcentaje promedio es una generalización del valor que se espera alcance aquel planteado en la hipótesis de este trabajo que es de 80 por ciento, resultando en todos los casos superior al 90 por ciento.

La tabla 13 muestra los valores comentados previamente para cada una de las configuraciones planteadas, especificando la denominación de la configuración y los valores respectivos.

Para la evaluación de dichos valores se tomó en cuenta la muestra de 370 registros evaluados.

La tabla 13, es el compendio de los experimentos realizados, considerando el porcentaje promedio de similitud entre valores reales y calculados, y haciendo uso del promedio y desviación estándar de los valores generados por la red neuronal para el cálculo del coeficiente de variación para determinar el grado de dispersión de los mismos.

Tabla 13. Relación de configuraciones de red con valores de desempeño

#	Configuración	Promedio de similitud	Promedio valores RNA	Desv. Est. valores RNA	Coef. de variación valores RNA
1	Configuración: 1 capa, 1 neurona, umbral 0.1	92.0008%	51.4561%	0.00389679	0.00757303
2	Configuración: 1 capa, 1 neurona, umbral 0.05	92.0092%	51.4460%	0.00447025	0.00868920
3	Configuración: 1 capa, 1 neurona, umbral 0.01	94.5594%	51.3830%	0.03194597	0.06217221
4	Configuración: 1 capa, 4n, umbral 0.1	98.0894%	51.5606%	0.04564621	0.08852922
5	Configuración: 1 capa, 4n, umbral 0.05	98.0903%	51.5579%	0.04565908	0.08855884
6	Configuración: 1 capa, 4n, umbral 0.01	98.4429%	51.4706%	0.04622318	0.08980495
7	Configuración: 1 capa, 16n, umbral 0.1	97.9479%	51.6036%	0.04816164	0.09332997
8	Configuración: 1 capa, 16n, umbral 0.05	97.9470%	51.6017%	0.04783204	0.09269471
9	Configuración: 1 capa, 16n, umbral 0.01	97.9470%	51.6017%	0.04783204	0.09269471
10	Configuración: 2 capa, (6n,3n), umbral 0.1	91.9014%	51.4642%	0.00152509	0.00296339
11	Configuración: 2 capa, (6n,3n), umbral 0.05	91.9107%	51.4579%	0.00188154	0.00365646
12	Configuración: 2 capa, (6n,3n), umbral 0.01	91.9132%	51.4563%	0.00211012	0.00410081
13	Configuración: 1 capa, 10n, umbral 0.1	98.0172%	51.5564%	0.04562851	0.08850211
14	Configuración: 1 capa, 10n, umbral 0.05	98.0295%	51.5542%	0.04562453	0.08849815
15	Configuración: 1 capa, 20n, umbral 0.1	97.7334%	51.3898%	0.05460284	0.10625227
16	Configuración: 1 capa, 20n, umbral 0.05	97.6795%	51.3555%	0.05728698	0.11154981
17	Configuración: 2 capas, (10n,5n), umbral 0.1	98.1236%	51.5742%	0.04558270	0.08838278
18	Configuración: 2 capas, (10n,5n), umbral 0.05	98.1544%	51.5338%	0.04602062	0.08930175
19	Configuración: 2 capas, (10n,5n), umbral 0.01	99.0557%	51.4831%	0.04683377	0.09096921
20	Configuración: 3 capas, (20n,5n,2n), umbral 0.05	98.1246%	51.5490%	0.04601196	0.08925861

Fuente: Elaboración propia.

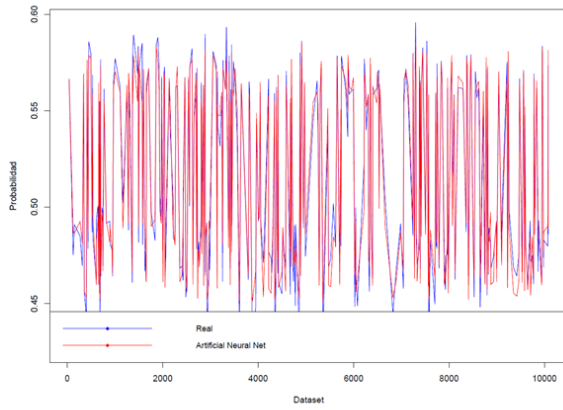
Después modificaciones en el cálculo probabilidades y ajustes en las variables preliminares, la mayor parte de las configuraciones produjo valores satisfactorios en las pruebas de desempeño, superando el objetivo planteado en la hipótesis de este estudio,

y que establecía el 80% de correspondencia de los valores obtenidos a través del cálculo por medio de redes neuronales. En las distintas corridas de prueba de configuraciones, las más eficientes en realizar el cálculo aproximado más afín a los valores reales resultaron ser la 6, 9, 15, 16, 17 y 18, variando ligeramente respecto a la semilla (número utilizado para inicializar un generador de números pseudoaleatorios, que producirá siempre la misma secuencia aleatoria) utilizada para la aleatoriedad de la muestra, y cuyas gráficas se observan en la figura 33.

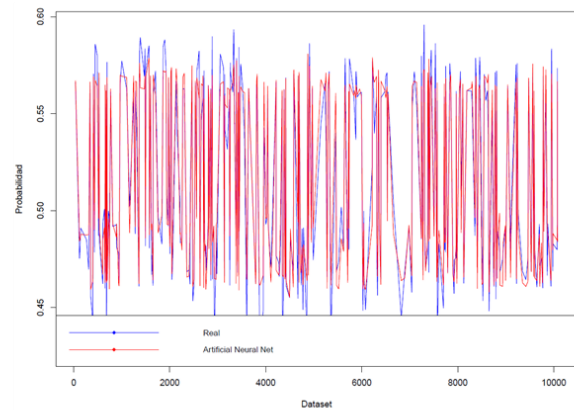
Los criterios de descarte de las otras configuraciones que fueron evaluadas consistieron en primer lugar a aquellas que contenían valores que excedían los parámetros planteados por los valores reales, es decir, aquellos que en su representación tuvieron valores extremos en sus probabilidades apreciables en valles o crestas al momento de ser graficadas. El segundo criterio fue el coeficiente de variación de los valores generados que al ser muy bajos permitían observar que los mismos se encontraban muy cerca del promedio de la totalidad de los valores, al ser graficados se podía observar una franja prácticamente paralela al promedio de los valores reales. En tercer lugar, se consideró el tiempo de procesamiento, tomando aquellos valores cuya evaluación fuera más breve.

De las configuraciones obtenidas, la más constante en cuanto a desempeño y velocidad de procesamiento es la configuración número 6, Configuración: 1 capa, 4 neuronas en la primera capa oculta, con un umbral de 0.01, cuya dispersión de los datos alcanza un coeficiente de variación de 0.08980495 y tiempo de procesamiento es de aproximadamente 91.78 segundos.

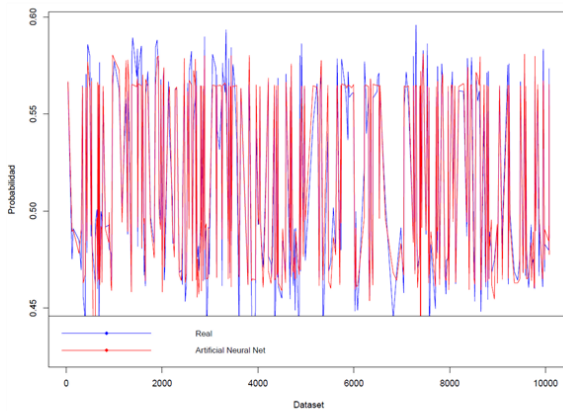
6 Configuración: 1 capa, 4n, umbral 0.01



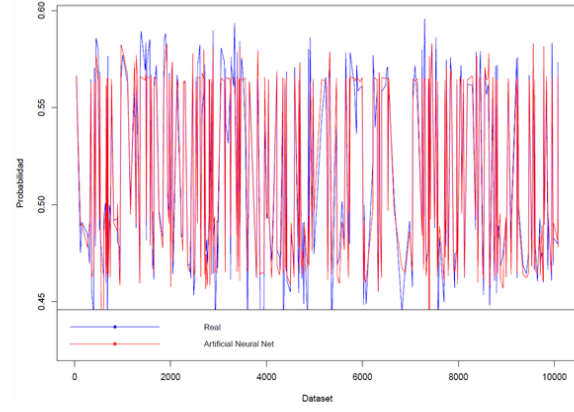
9 Configuración: 1 capa, 16n, umbral 0.01



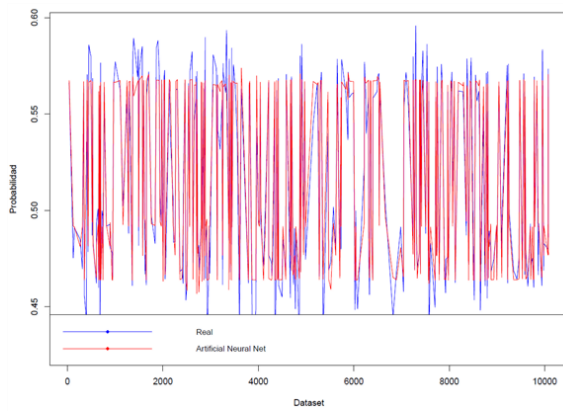
15 Configuración: 1 capa, 20n, umbral 0.1



16 Configuración: 1 capa, 20n, umbral 0.05



17 Configuración: 2 capas, (10n,5n), umbral 0.1



18 Configuración: 2 capas, (10n,5n), umbral 0.05

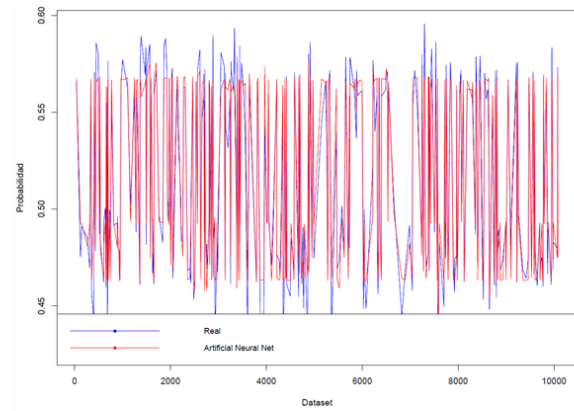


Figura 33: Desempeño de las configuraciones más efectivas

Fuente: Gráficas generadas con R.

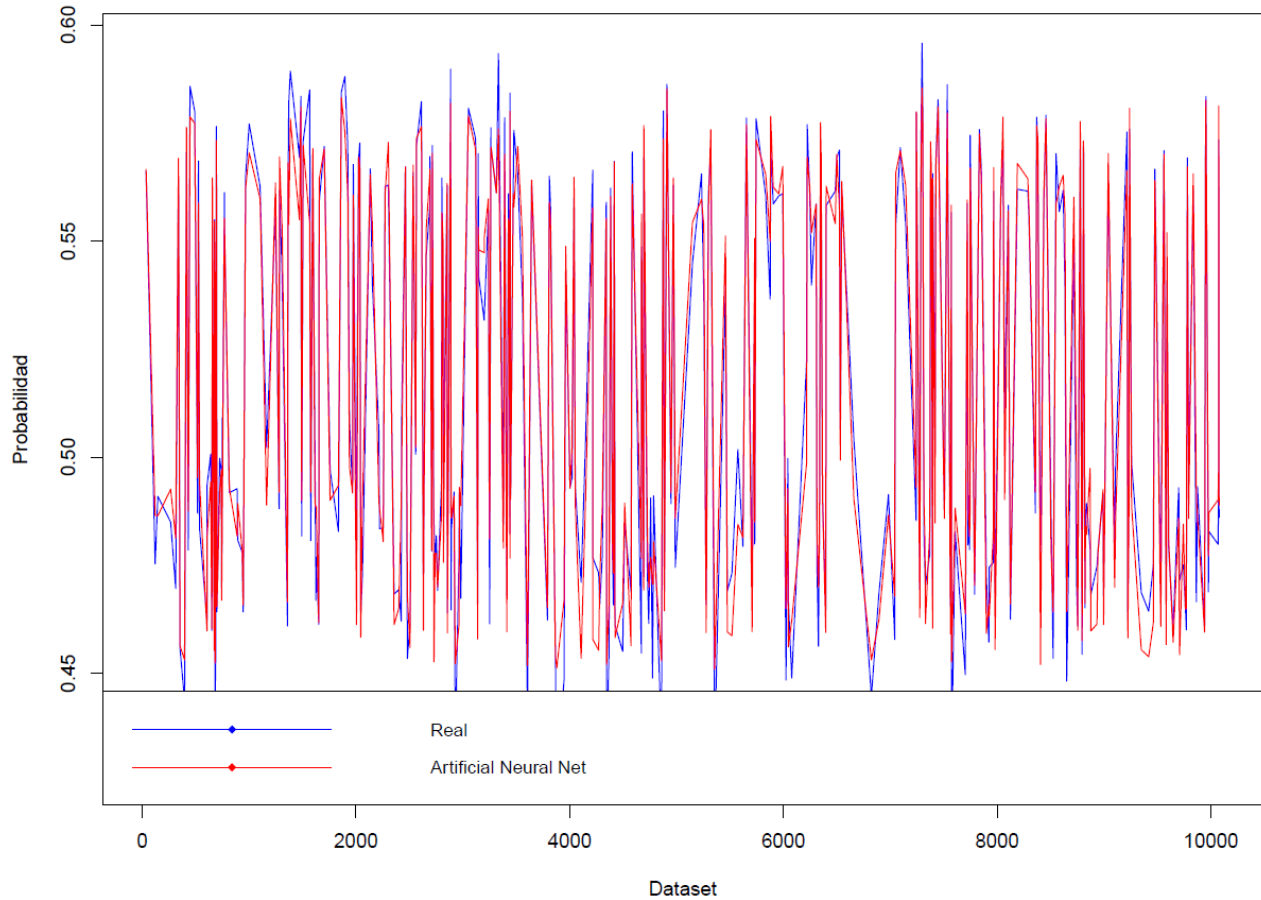


Figura 35: Gráfica comparativa de valores reales vs calculados por la RNA en la configuración 6

Fuente: Gráfica generada con R.

Se puede observar que la mayoría de los valores se concentran en la parte intermedia de la gráfica, lo que de manera explícita manifiesta que los valores calculados tienden a buscar alcanzar los valores reales, sin embargo, algunos quedan cortos en magnitud, mientras que, por otro lado, y de manera menos obvia, en algunas configuraciones resaltaron valores atípicos que se desbordan de la escala comparativa contra los valores reales, razón por la cual fueron excluidos de la selección presentada.

Al comparar los valores específicos es evidente que la red se ajusta de manera adecuada a los datos reales. La figura 36 muestra el comportamiento de los valores arrojados por la RNA respecto de los valores de incidencia real.

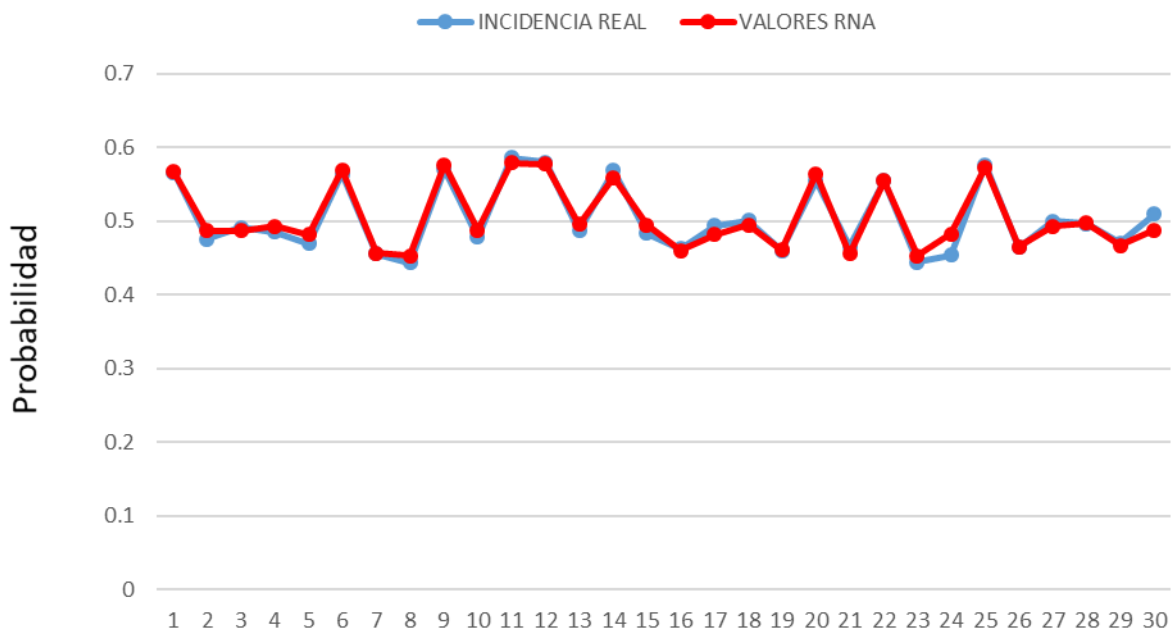


Figura 36: Gráfica de acercamiento a 30 valores de la muestra procesada

Fuente: Gráfica generada con Excel.

Para ejemplificar de manera visual el desempeño del modelo se presentan dos ejercicios con representación geoespacial. El primero de ellos consiste en la representación gráfica de la muestra utilizada para el procesamiento a través de la red neuronal con sus valores probabilísticos originales, así como la representación para el comportamiento de los datos generados por la red neuronal, adjuntando a dichas representaciones geoespaciales una gráfica con la distribución de sus respectivos valores.

En las figuras 37 y 38, se aprecian los mismos valores geoespaciales utilizados como muestra en la configuración 6, el mapa agrupa en clústeres los puntos cercanos entre sí y, en cada figura, se establece la gráfica respectiva representando la distribución de los valores en la escala de probabilidad. Si bien es cierto, se puede observar cierta diferencia en la conformación de las gráficas, sin embargo, a detalle los valores guardan la correspondencia dentro del grado de aceptabilidad planteado originalmente. .

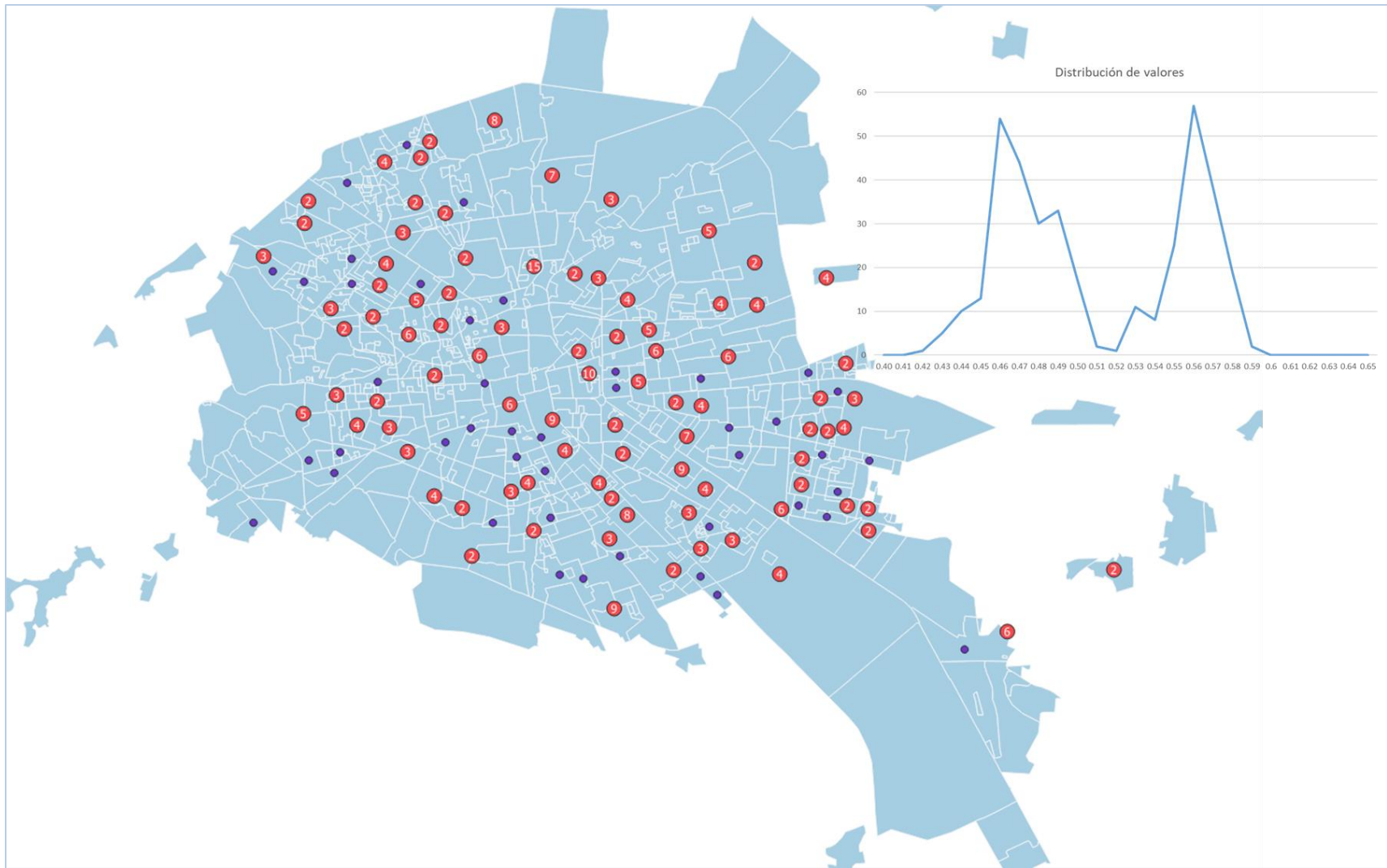


Figura 37: Representación gráfica de los valores reales y su distribución según su probabilidad.

Fuente: Elaboración propia.

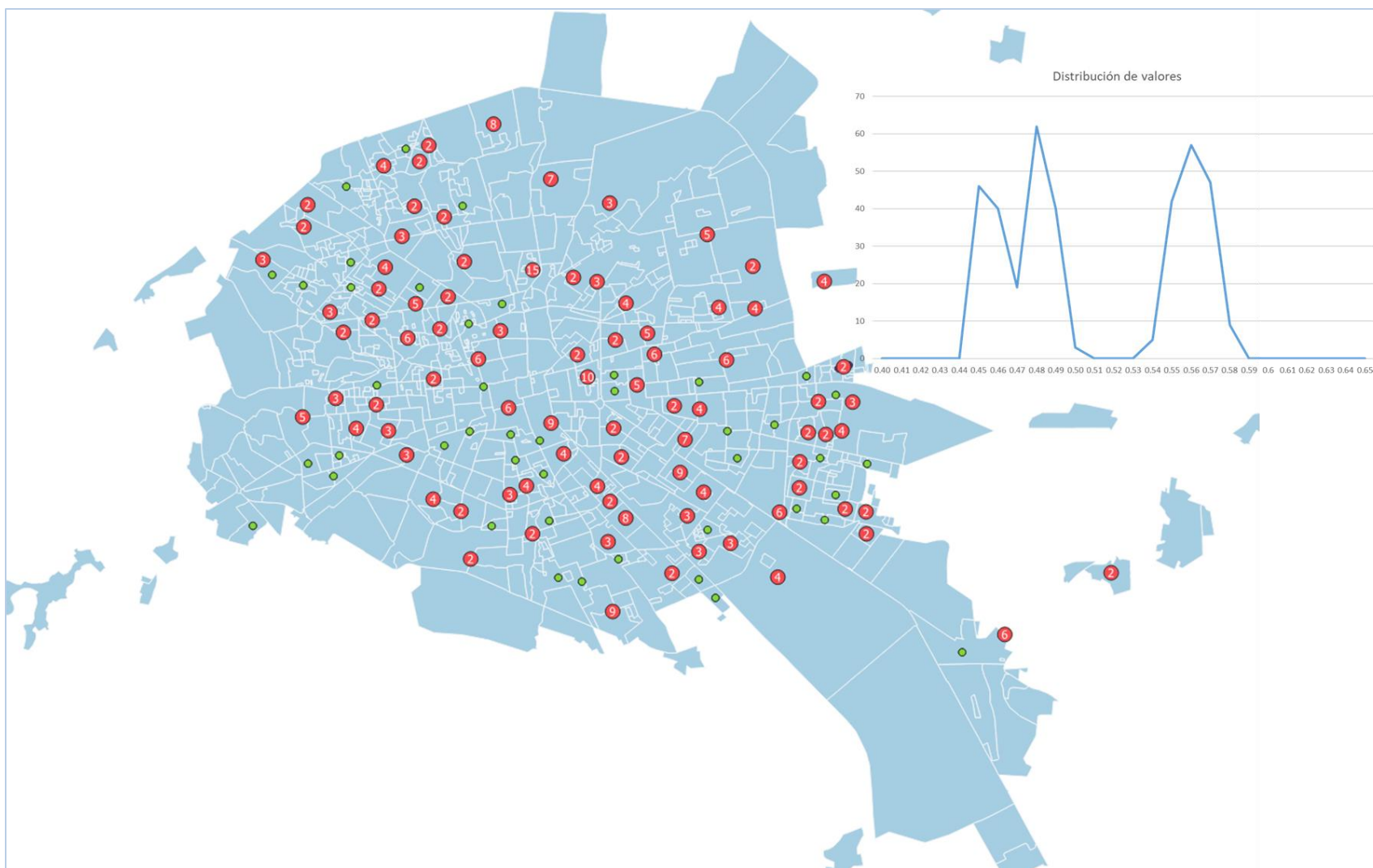


Figura 38: Representación gráfica de los valores procesados por la RNA y su respectiva distribución según su probabilidad.

Fuente: Elaboración propia.

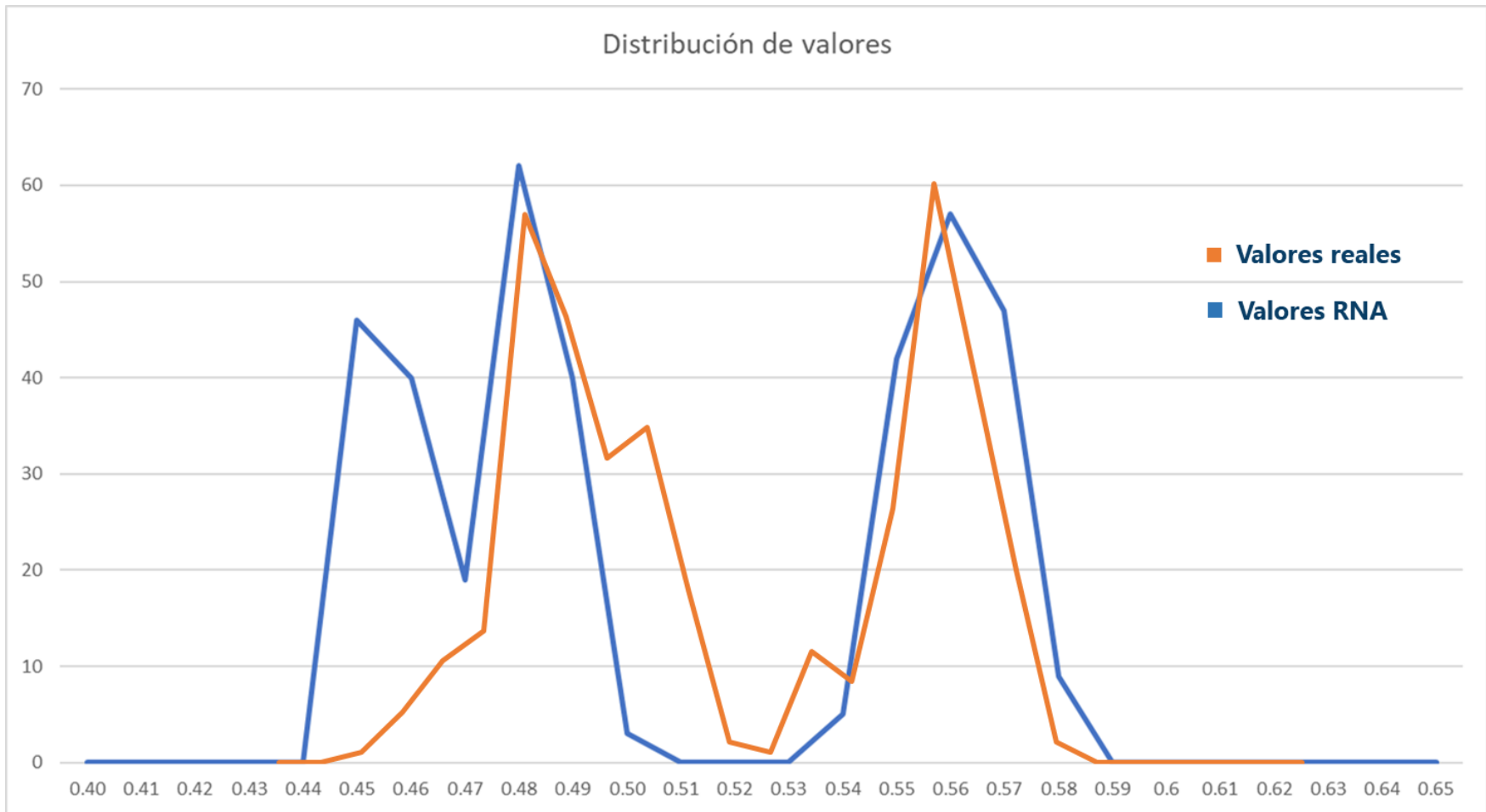


Figura 39: Comparativa de datos reales con los generados por la red neuronal artificial.

Fuente: Elaboración propia.

En la figura 39, las gráficas de cada representación geoespacial se contrastan para observar que se presentan diferencias en la distribución de los valores y que, aunque el propósito es que sean prácticamente idénticas, existen diferencias evidentes que resaltan que no son iguales y sin embargo constatan que el modelo asemeja un comportamiento similar al de los valores reales.

Por otra parte, dado que una comparativa geoespacial con un único conjunto de datos no sería apreciable en un plano, se ejemplifica estableciendo una serie de 277 registros, que incluye la totalidad de las zonas geográficas establecidas que componen la zona metropolitana, construyendo un vector de entrada específico para observar el comportamiento con datos ajenos al universo de datos con que se entrenó la red.

Para este caso, se estableció un momento determinado en el tiempo con las siguientes características, las variables modificadas son: ZONA (consecutivo de 1 a 277, constituyen 277 registros a evaluar), DELITO, establecido como robo comercial (1), DIASEM, establecido como domingo (1), DIAMES, establecido como día primero de mes (1), MES, establecido como enero (1), HORA, establecida como ocho de la noche (20), MODUS, establecido como sin violencia (0), TEMP, temperatura establecida como 18 grados (18), el resto de las variables permanece constante y correspondiente a cada zona geográfica definida.

Para efectos de visualizar los valores procesados, una vez que el algoritmo basado en la configuración 6 de procesamiento (red neuronal de una capa con cuatro neuronas y con 0.01 de valor de umbral), terminó en 90.75 segundos, se cargaron las probabilidades generadas para cada uno de los 277 registros que representan cada zona geográfica con

las condiciones previamente planteadas en el párrafo previo, además se incorporó el punto geográfico representativo de la zona para poder ser mostrado en el plano.

La figura 40 muestra un mapa generado con interpolación mediante distancia inversa ponderada, que ofrece a la vista un gráfico que resalta los valores numéricamente mayores en un tono más oscuro al resto de los mismos.

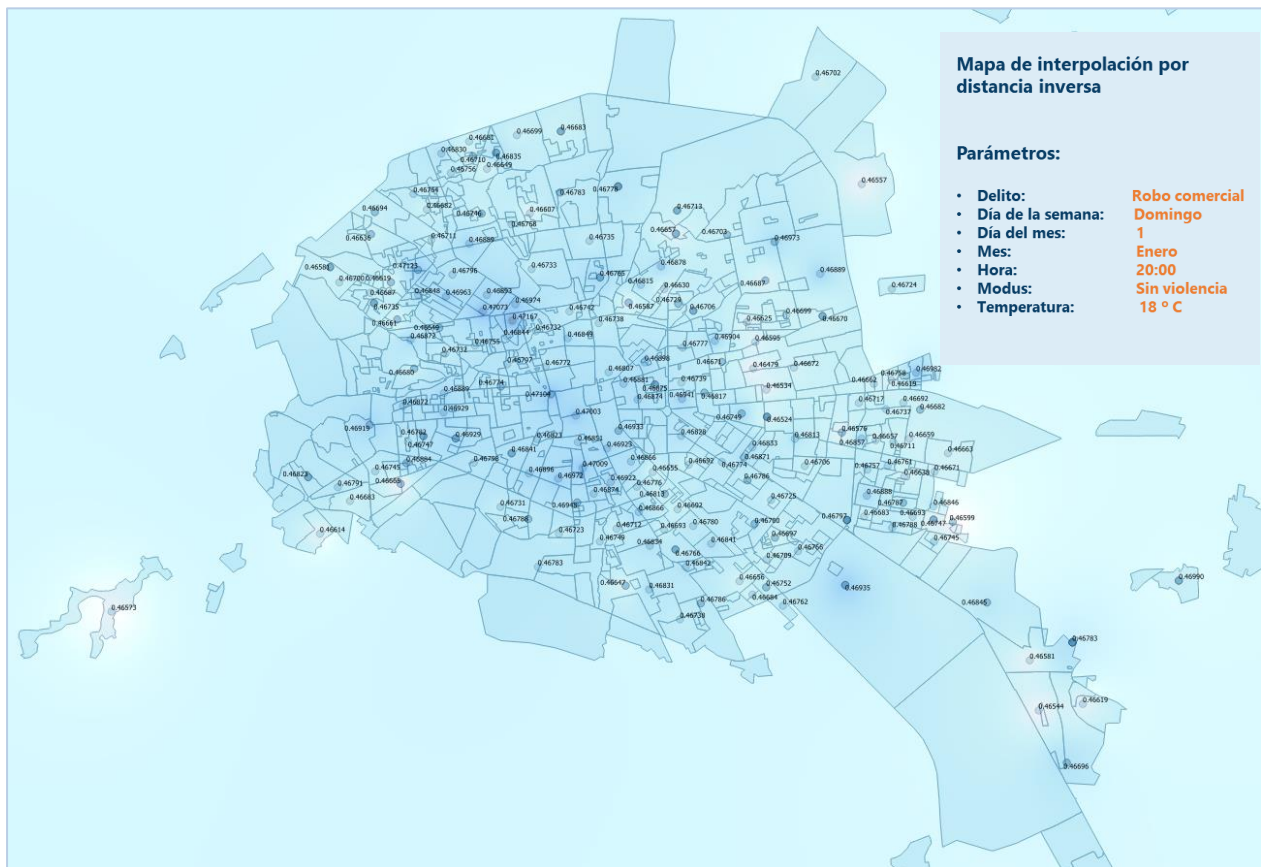


Figura 40: Mapa DIP de probabilidades de robo comercial según los datos de la RNA.

Fuente: Elaboración propia.

Los valores más bajos son observables debido al tono blanquecino que presentan. La interpolación mediante distancia inversa ponderada determina los valores de celda a través de una combinación ponderada linealmente de un conjunto de puntos de muestra. La ponderación es una función de la distancia inversa. La superficie que se interpola debe

ser la de una variable dependiente de la ubicación, en este caso la probabilidad de ocurrencia del evento. Este procedimiento permite evocar el desempeño de un mapa de calor. Por otra parte, para favorecer la interpretación de los datos, en primera instancia, se etiquetaron los valores de probabilidad generados por la RNA en cada punto indicativo de la zona geográfica.

Para destacar aún más los valores que requieren atención se señalan en el mismo mapa, Figura 41, esta vez omitiendo las etiquetas autogenerated para permitir una visualización más clara de los datos, y marcando con distintivos los puntos de interés.

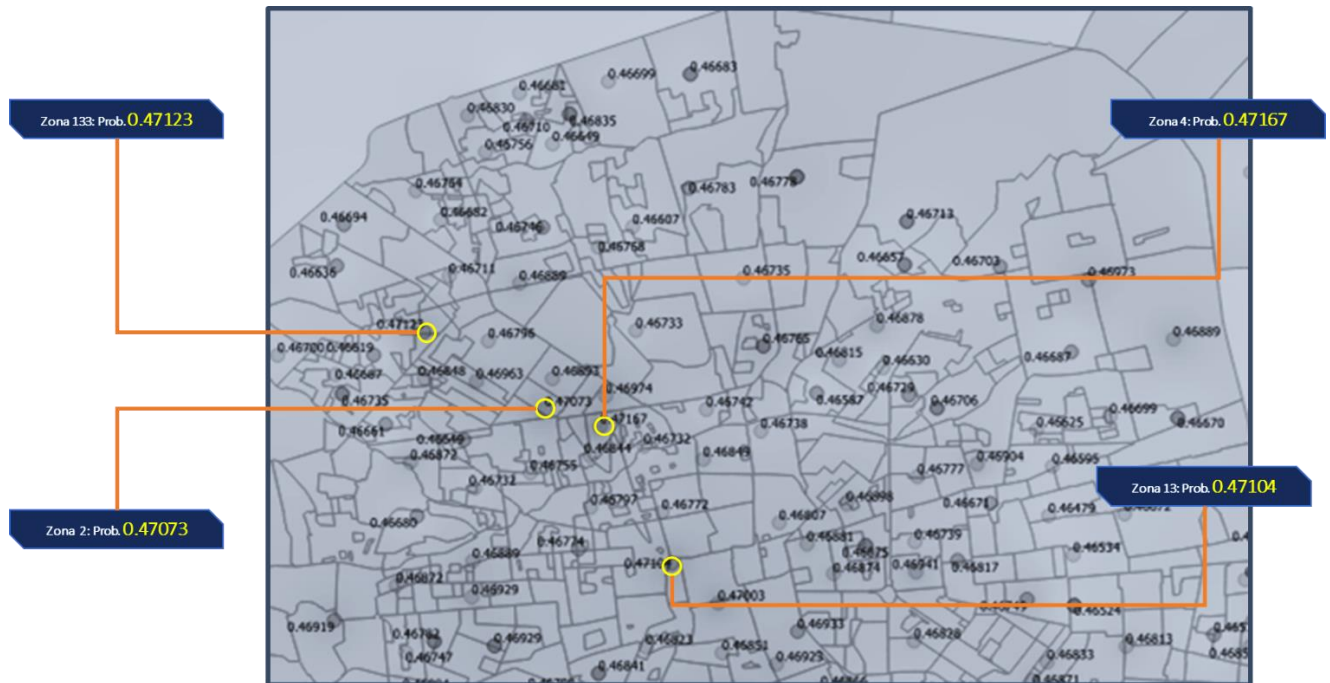


Figura 41: Puntos de interés de acuerdo a la probabilidad calculada por la RNA

Fuente: Elaboración propia.

De acuerdo a los datos arrojados por la RNA, las zonas de interés para las condiciones establecidas serían en orden de importancia las zonas 4, 133, 13 y 2. Ubicadas en la zona centro noroeste de la ciudad.

4.3 Discusión de resultados

Ante los resultados obtenidos, que fehacientemente cubren la expectativa asentada en la Hipótesis inicial, es también importante revisar ciertos cabos que se abrieron durante el planteamiento del estudio.

Si bien es cierto que varios de los modelos que fueron abordados por los estudios revisados al inicio del documento hacen uso de inteligencia artificial e incluso directamente de redes neuronales, no están enfocados a resaltar el valor intrínseco de los datos numéricos generados por la incidencia delictiva y su influencia por la correlación entre ellos para la predicción de eventos de naturaleza similar, siendo esta la razón de la conformación del modelo finalmente utilizado.

La predicción con base en datos estadísticos, con cifras oficiales y con respaldo de fuentes de datos consolidadas que robustecen la confiabilidad del material de entrada para el modelo fue una oportunidad que de manera afortunada se conjuntó para presentarse y poder ser una realidad a través de la construcción del nuevo modelo.

Otra ventaja es que el modelo puede refinarse y escalarse hasta el desarrollo de una aplicación o bien empotrarse en la tecnología bajo desarrollo actualmente por las instancias responsables de las tecnologías de información de la Fiscalía General.

Por otra parte, durante la conformación de configuraciones de redes a probar, destaca la metodología propuesta por la literatura relacionada para la generación de una topología inicial consistente, aplicando la regla de pirámide geométrica, dado que funciona bien con algunos modelos probados puesto que se basa en el supuesto de que el número de neuronas en la capa oculta es menor al número de variables de entrada y mayor al total

de variables de salida, y por otra parte, señala Richard Lippman que una capa oculta es suficiente para problemas arbitrariamente complejos debiendo ésta incluir hasta tres veces la cantidad de nodos de entrada, por lo que queda siempre bajo la duda la cantidad óptima de capas y neuronas que debe contener. Esto llevo a la conformación de más de treinta configuraciones variadas en topología y en determinación del umbral, sin embargo sólo se incluyeron veinte de ellas porque cada una debía ser documentada y contrastada con el resto, viendo los resultados de las configuraciones descartadas, era evidente que al aumentar la complejidad se generaban más errores que proporcionaban valores que se desbordaban o bien se refinaba tanto el umbral que el procesamiento se aletargaba o bien simplemente no concluía. Por lo que fue preciso mantener las 20 configuraciones candidatas para ser evaluadas primeramente con la totalidad de los valores para comparar los valores calculados contra los valores reales y posteriormente con la uniformidad de su distribución y el tiempo de procesamiento.

La validación de estas tres características llevó a selección de una configuración óptima, y es que, retomando una de las configuraciones recomendadas, es evidente que el tipo de problema que se aborda no se puede satisfacer con dicha configuración, donde los valores generados presentan un comportamiento bastante raquítico comparado con el seleccionado al final de las validaciones véase figura 42.

En dicha gráfica, puede observarse como apenas se inmutan los valores generados por la red neuronal, básicamente sólo siguen el promedio de los valores por poco menos de dos puntos porcentuales por encima del mismo.

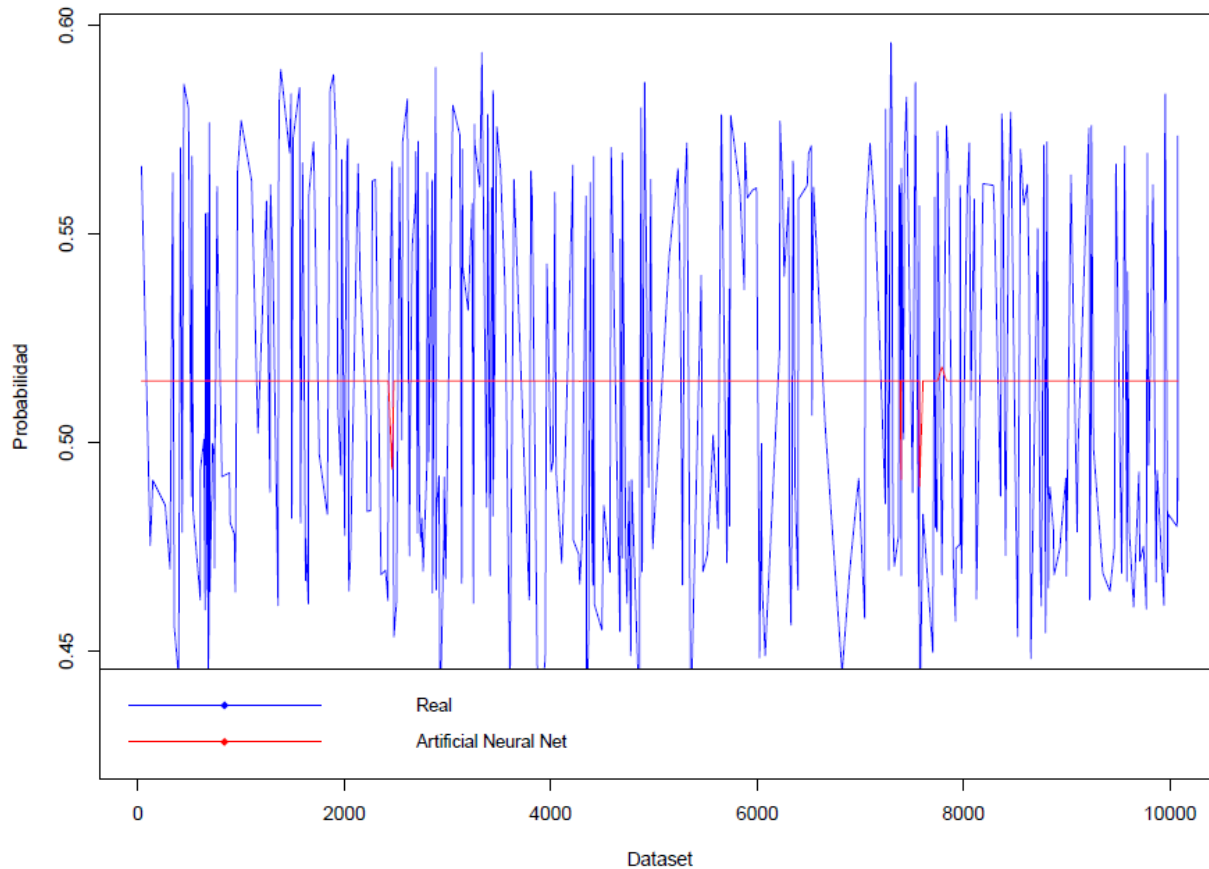


Figura 42: Desempeño de configuración 12. Dos capas (6n, 3n) con umbral 0.01.

Fuente: Elaboración propia.

Si bien es cierto la selección realizada cumple con los lineamientos establecidos y satisface adecuadamente todas las premisas hasta el momento, pero no existe garantía de que en el futuro y derivado de la mecánica de comportamiento de la delincuencia varíe por lo cual será necesario realizar auditorías cada cierto tiempo del modelo para garantizar el funcionamiento óptimo con la configuración elegida, recordando que la selección de configuraciones es sólo una parte del mismo.

Capítulo 5: Conclusiones

5.1 Conclusiones generales

En la hipótesis planteada al inicio de este trabajo, se estableció que *“La implementación de un modelo basado en redes neuronales permitirá definir un pronóstico efectivo del comportamiento de la incidencia delictiva superior al 80 %”*.

El desarrollo de la metodología utilizada en el proyecto, permitió observar serias fallas en la estandarización de los datos capturados respecto a la incidencia delictiva, para lo cual la concreción del modelo de trabajo utilizado permite proporcionar una guía estandarizada para la captura de la información.

Por otra parte, con los resultados obtenidos se establece que, bajo los parámetros planteados, se pudo obtener al menos una configuración estable que emula el comportamiento de los datos reales, generando datos satisfactorios y consistentes, suficientemente apropiados para generar líneas de acción sólidas con expectativas de éxito reales.

Una vez establecido el método para la implementación del modelo de procesamiento de incidencia delictiva a través de redes neuronales artificiales y después de probar su factibilidad, tenemos un mecanismo procedimental adecuado para el procesamiento de la información. De igual manera, el método proporciona una guía sólida que permite al analista basar su trabajo en un método lógico-racional, en lugar de sólo documentar los eventos delictivos y dejar la interpretación de los datos de manera empírica o intuitiva a los mandos superiores.

Las variables elegidas de los tres entornos considerados (delictivo, socioeconómico y ambiental), una vez probadas dentro del vector de evaluación procesado por la red, muestran un comportamiento consolidado, lo que al momento de presentar el proyecto actual se pueden calificar como estables, bajo la consideración de que la incidencia es un bloque de datos que debe actualizarse constantemente, los datos socioeconómicos se actualizan en intervalos de 5 a 10 años de la mano de los censos y consultas de los órganos facultados para tal efecto, y conforme el paso del tiempo las variables ambientales también van mostrando ligeras variaciones conforme a años previos que a su vez también deben ser consideradas.

El propósito nuclear del modelo planteado es de origen generar pronósticos del comportamiento delincuencia con base en las tendencias y patrones que se van presentando, y los resultados generados en la etapa de experimentación muestran a sobremanera que el modelo es efectivo en generar pronósticos válidos para la toma de decisiones.

Por todo lo anterior, se concluye de manera ineludible que el objetivo del trabajo realizado cumple satisfactoriamente con la hipótesis establecida al comienzo del mismo, dando incluso cabida a su implementación práctica, bajo las respectivas consideraciones para su actualización y perfeccionamiento.

5.2 Futuras líneas de investigación

En primer lugar, para el refinamiento del desempeño del modelo es preciso ahondar en la revisión del desempeño del vector de entrada, probando y descartando o manteniendo, por el peso y la aportación, a cada una de las variables elegidas, considerando que el

bloque socioeconómico es prácticamente un conjunto constante con consistencia válida por prácticamente 10 años, al momento de obtenerse una actualización de esos valores debe plantearse la vigencia de los mismos y la manera de su implementación en el modelo.

Por otra parte, es factible la consideración de otras herramientas, además del mapa de interpolación por distancia inversa ponderada, tales como diagramas de barras, matrices de datos y relojes delictivos de pronóstico.

Es viable también, la integración con otros métodos numéricos y estadísticos, aunque el proyecto se enfocó al uso de redes neuronales artificiales no excluye de manera tácita la incorporación de otros elementos lo que desde otras perspectivas permitiría agudizar su desempeño o simplemente experimentar la influencia en el modelo o bien comparar la efectividad con métodos más tradicionales.

Finalmente, la automatización total del proceso, que incorporaría la obtención automática del vector de entrada a través de un sistema de información o generador de reportes, hasta la obtención instantánea de otros productos de inteligencia que aprovechen los datos generados del pronóstico.

5.3 Reflexiones finales

Las características del modelo presentado permiten que pueda adaptarse a otros métodos de aprendizaje, incluso a otras herramientas distintas a el algoritmo incorporado del lenguaje R. Su aplicación es de uso general y no presenta restricciones específicas.

Las aplicaciones del modelo son extensas y el mismo es un punto de partida para que quienes se interesan en el comportamiento de la delincuencia puedan desarrollar

herramientas y procedimientos más sólidos y efectivos. El trabajo aquí presentado está para ser aplicado y mejorado con nuevas y mejores ideas, que permitan ayudar al establecimiento de una seguridad pública confiable y por ende una administración de justicia más efectiva y menos requerida.

Por último, y a título personal del autor, más allá de las cuestiones técnicas, organizacionales y procedimentales, se debe tener en cuenta que el factor más importante y probablemente el más decisivo para contener la violencia y la delincuencia que de años atrás nos aqueja con cada vez mayor ímpetu, yace, no en las herramientas, ni en las policías, ni en los equipos, ni en los operativos, ni en la efectividad del sistema de procuración de justicia, ni siquiera en que un equipo de cómputo que sea capaz de mostrarnos con base en cálculos la ocurrencia de un delito, el factor más importante somos cada uno de nosotros, la conciencia del bien común en cada persona, los valores personales y una ética social empática con el prójimo, la seguridad y bienestar de los demás radica en lo que tú y yo hacemos para que prevalezca.

Referencias

- Aldabas-Rubira, E. (2002). Introducción al reconocimiento de patrones mediante redes neuronales. *XIX Jornades de Conferències d'Enginyeria Electrònica del Campus de Terrassa*. Tarrasa: Universitat Politècnica de Catalunya.
- Aleksander, I., & Morton, H. (1990). *An introduction to neural computing*. Chapman and Hall.
- Anderson, J. (1995). *An Introduction to Neural Networks*. London: The MIT Press.
- Auditoria Superior de la Federación. (2013). *Evaluación número 1207 "Política Pública de Seguridad Pública"*. CDMX: Cámara de Diputados.
- Ballesteros González, A. (2000). *Neural Networks Framework*. Obtenido de <http://www.redes-neuronales.com.es>
- Bishop, C. (1995). *Neural Networks for Pattern Recognition*. Birmingham: Clarendon Press.
- Brahan, J., Lam, K., Hilton, C., & Leung, W. (1998). AICAMS: artificial intelligence crime analysis and management system. *Knowledge-Based Systems*, 355-361.
- Bursik, R. J., & Grasmik, H. G. (1993). *Neighborhoods and crime: the dimensions of effective community control*. San Francisco: Lexington Books.
- Cámara de Diputados del H. Congreso de la Unión. (2016). *Ley General del Sistema Nacional de Seguridad Pública*. CDMX: Diario Oficial de la Federación.
- Chainey, S., & Ratcliff, J. (2005). *GIS and Crime Mapping*. Wiley.

- Cohn, E. (1990). Weather and crime. *British Journal of Criminology* 30, 51-64.
- Debes, C., Koenig, A., & Gross, H.-M. (2005). Transfer Functions in Artificial Neural Networks A Simulation-Based Tutorial . *Brains, minds media*, 1-11.
- Flatley, J., & Bradley, J. (2013). Analysis of variation in crime trends. *Office for National Statistics*, 1-15.
- Flórez López, R., & Fernández Fernández, J. M. (2008). *Las Redes Neuronales Artificiales, Fundamentos Teóricos y Aplicaciones Prácticas*. Oleiros: Netbiblio, S.L.
- Friedman, M., & Kandel, A. (2000). *Introduction to Pattern Recognition Statistical, Structural, Neural and Fuzzy Logic Approaches*. London: Imperial College Press.
- Fritsch, S., & Guenther, F. (16 de Agosto de 2016). *Repositorio CRAN*. Obtenido de neuralnet: Training of Neural Networks: <https://cran.r-project.org/web/packages/neuralnet/index.html>
- Fukushima, K. (1980). Neocognitron: A Self-Organizing Neural Network Model for a Mechanism of Pattern Recognition Unaffected by Shift in Position. *Biological Cybernetics*, 192-202.
- Galushkin, A. (2007). *Neural Networks Theory*. Moskva: Springer.
- García, J. (1 de julio de 2015). *Redes Neuronales Fácil y desde cero*. Obtenido de Youtube:
https://www.youtube.com/playlist?list=PLAnA8FVrBI8AWkZmbswwWiF8a_52dQ3JQ

- Groff, E. R., Weisburd, D. L., & Yang, S.-M. (2010). It Is It Important to Examine Crime Trends at a Local 'Micro' Level? A Longitudinal Analysis of Street to Street Variability in Crime Trajectories. *Journal of Quantitative Criminology*, 7-32.
- Haykin, S. (1999). *Neural Networks: A Comprehensive Foundation, second edition*. New Jersey: Prentice-Hall.
- Heaton, J. (2008). *Introduction to Neural Networks for C#*. Saint Louis: Heaton Research.
- Hebb, D. O. (1949). *The Organization of Behavior: A A neuropsychological theory*. New York: John Wiley and Sons.
- Herrman, C. (2015). The dynamics of robbery and violence hot spots. *Crime Science An Interdisciplinary Journal*, 4-33.
- INEGI, I. N. (s.f.). *INEGI*. Obtenido de www.inegi.org.mx
- Isasi Viñuela , P., & Galván León, I. (2004). *Redes de Neuronas Artificiales Un enfoque práctico*. Madrid: Pearson Educación.
- Jackson, P. (1998). *Introduction to Expert Systems*. Addison Wesley.
- Jayaweera, I., Sajeewa, C., Liyanage, S., Wijewardane, T., & Perera, I. (2015). Crime Analytics: Analysis of Crimes Through. *Moratuwa Engineering Research Conference (MERCon), 2015* (págs. 277-288). Moratuwa: IEEE.
- Kang, H.-W., & Kang, H.-B. (24 de Abril de 2017). *National Center for Biotechnology Information*. Obtenido de PLOS One A Peer-Reviewed, Open Access Journal: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5402948/>

- Kawaguchi, K. (2000). *A Multithreaded Software Model for Backpropagation Neural Network Applications*. El Paso: The University of Texas at El Paso.
- Koch, R. (1999). *The 80/20 Principle: The Secret to Success by Achieving More with Less*. Broadway Business.
- Krawczak, M. (2013). *Multilayer Neural Networks A Generalized Perspective*. Warsaw: Springer International Publishing.
- Krenker, A., Bester, J., & Kos, A. (2011). *Artificial Neural Networks - Methodological Advances and Biomedical Applications*. Rijeka: University of Ljubljana.
- Kvasnicka, V. (1998). *How to Process Structured Data by Neural Networks?*. Bratislava: Slovak Technical University.
- Minsky, M. (1961). Steps Toward Artificial Intelligence. *Proceedings of the IRE*, 8-30.
- Montaño Moreno, J. (2002). *Redes Neuronales Artificiales aplicadas al Análisis de Datos*. Palma de Mallorca: Universitat de Les Illes Balears.
- Morenoff, J. D., & Sampson, R. J. (1997). Violent Crime and the Spatial Dynamics of Neighborhood Transition: Chicago, 1970–1990. *Social Forces*, 31-64.
- Navarro, D. (2004). El ciclo de inteligencia y sus límites. *Cuadernos constitucionales de la Cátedra Fadrique Furió Ceriol*(48), 51-66.
- Perversi, I. (2007). *Aplicación de Minería de Datos para la Exploración y Detección de Patrones Delictivos en Argentina*. Buenos Aires: Instituto Tecnológico de Buenos Aires.

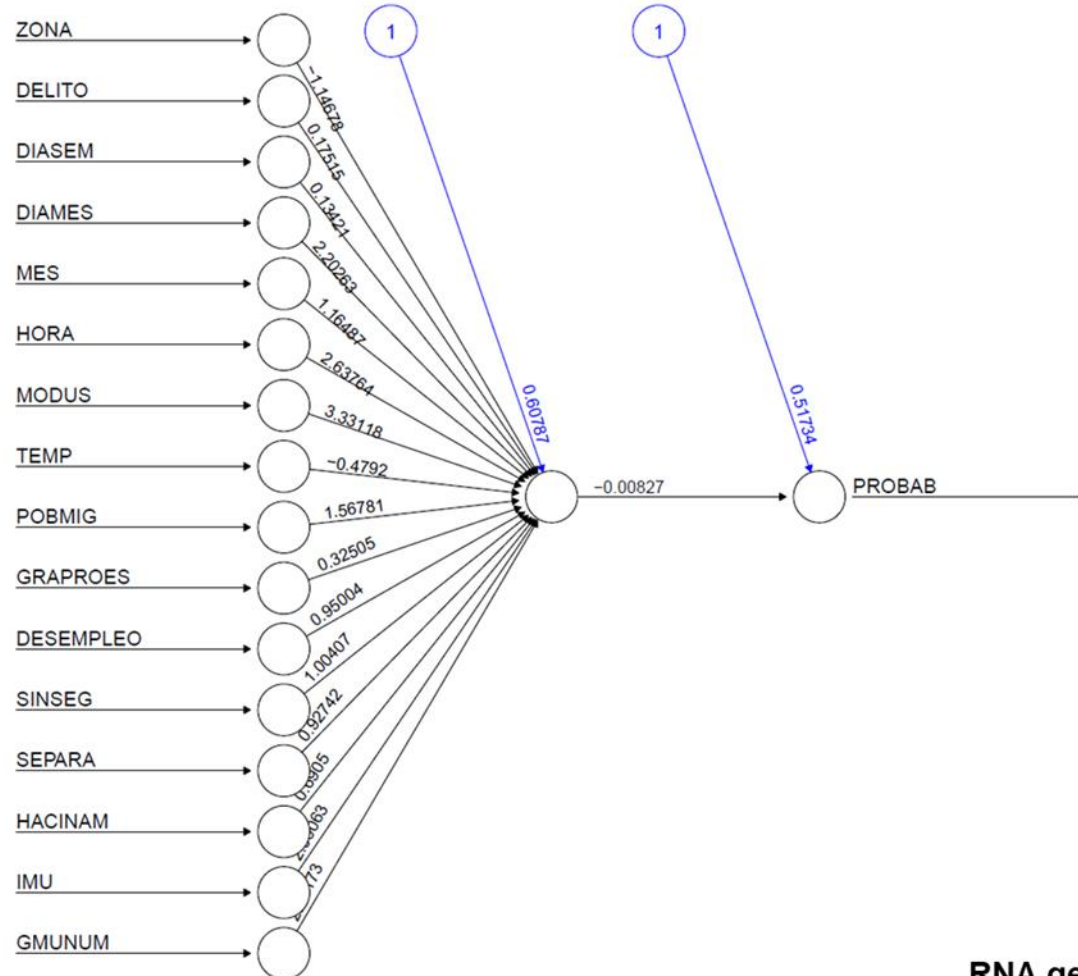
- Ponce Cruz, P. (2010). *Inteligencia Artificial con Aplicaciones a la Ingeniería*. Ciudad de México: Alfaomega.
- Ratcliffe, J. (2011). *Intelligence-Led Policing*. New York: Willan Publishing.
- RumelHart, D. E., & McClelland, J. L. (1986). *Parallel Distributed Processing: Exploration in the*. Cambridge: MIT Press.
- Salas, R. (2004). *Redes neuronales artificiales*. Valparaíso: Universidad de Valparaíso.
- Santana, J., & Mateos Farfán, E. (2014). *El arte de programar en R: un lenguaje para la estadística*. Jiutepec: Instituto Mexicano de Tecnología del Agua.
- Secretariado Ejecutivo del SNSP. (2015). *Iniciativa Delictiva del Fuero Común 2015*. CDMX: Secretaría de Gobernación.
- Secretariado Ejecutivo del SNSP. (2016). *Incidencia Delictiva del Fuero Común 2016*. CDMX: Secretaría de Gobernación.
- Serralde, A. (2011). Implementación de un Sistema de Inteligencia Empresarial. *Seminario de profundización* (pág. 1). Guadalajara: Universidad de Guadalajara.
- Shiffman, D. (2012). *The Nature of Code*. Mountain View: Theoklesia.
- Shiju, S., & Surya, G. (2014). Crime Analysis and Prediction Using Data Mining. *Networks & Soft Computing (ICNSC), 2014 First International Conference on* (págs. 406-412). Guntur: IEEE.

- Short, M., D'Orsogna, M., Pasour, V., Tita, G., Brantingham, P., Bertozzi, A., & Chayes, L. (2008). A Statistical Model Of Criminal Behavior. *Mathematical Models and Methods in Applied Sciences*, 1249-1267.
- Siddique, N., & Adeli, H. (2013). *Computational Intelligence Synergies of Fuzzy Logic, Neural Networks and Evolutionary Computing*. New Delhi: John Wiley & Sons.
- Taylor, R. B. (1997). Social Order and Disorder of Street Blocks and Neighborhoods: Ecology, Microecology, and the Systemic Model of Social Disorganization. *Journal of Research in Crime and Delinquency*, 133-155.
- Torres Estrada, P. R. (2015). *La inteligencia en el nuevo sistema de justicia penal*. México, D.F.: Instituto Nacional de Ciencias Penales.
- Tudela, P. (2016). Estándares de trabajo en una Unidad de Análisis del Criminal: asegurando la eficacia y la calidad para la persecución penal. *Buenas prácticas para el análisis delictual en América Latina*. Santiago: Fundación Paz Ciudadana.
- Werbos, P. J. (1974). *Beyond Regression: New Tools for Prediction and Analysis in the Behavioral Sciences*. Cambridge: Harvard University.

Anexos

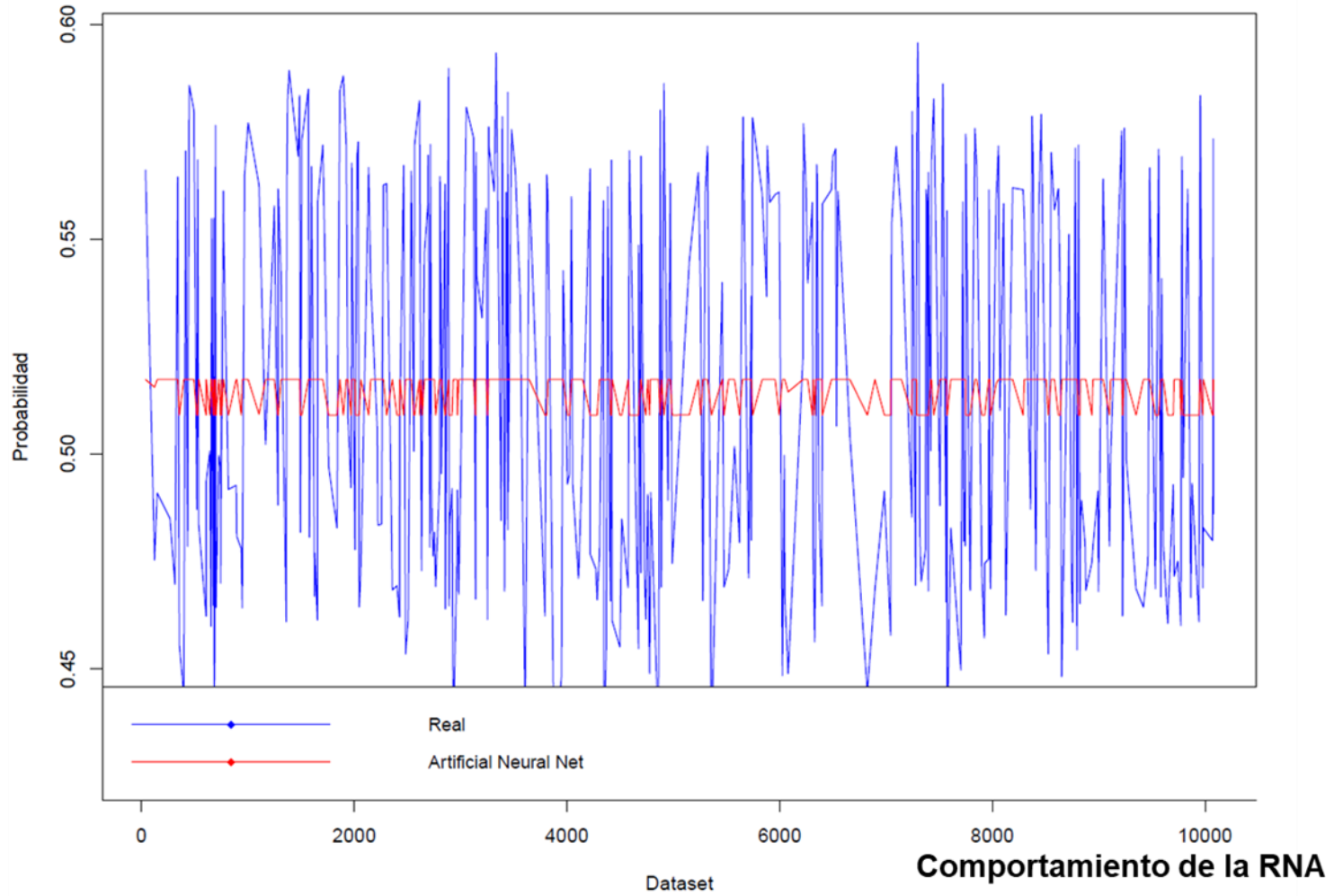
Anexo 1: Conjunto de corridas de configuraciones de red neuronal.

1. Configuración: 1 capa, 1 neurona, umbral 0.1

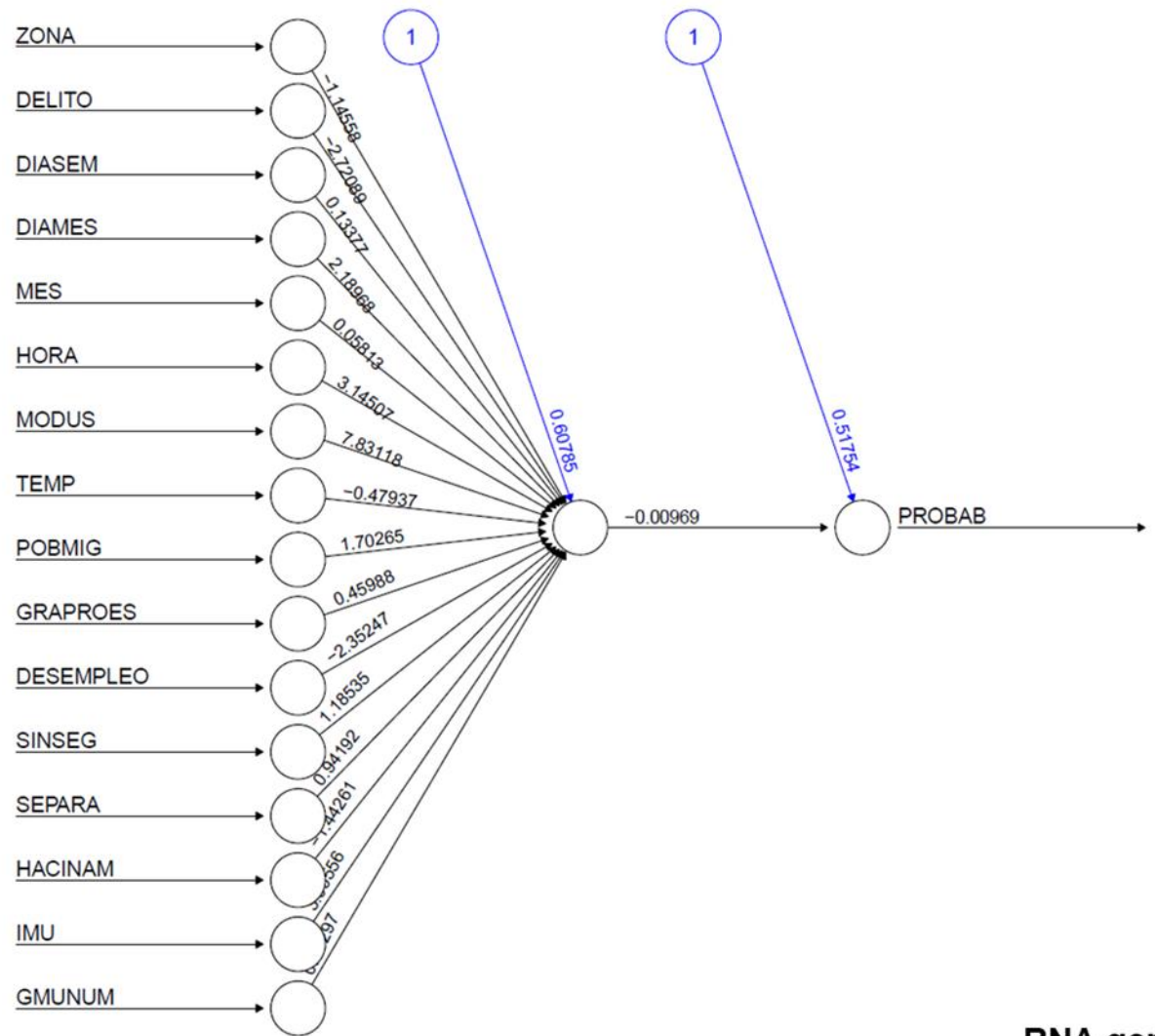


RNA generada

1. Configuración: 1 capa, 1 neurona, umbral 0.1

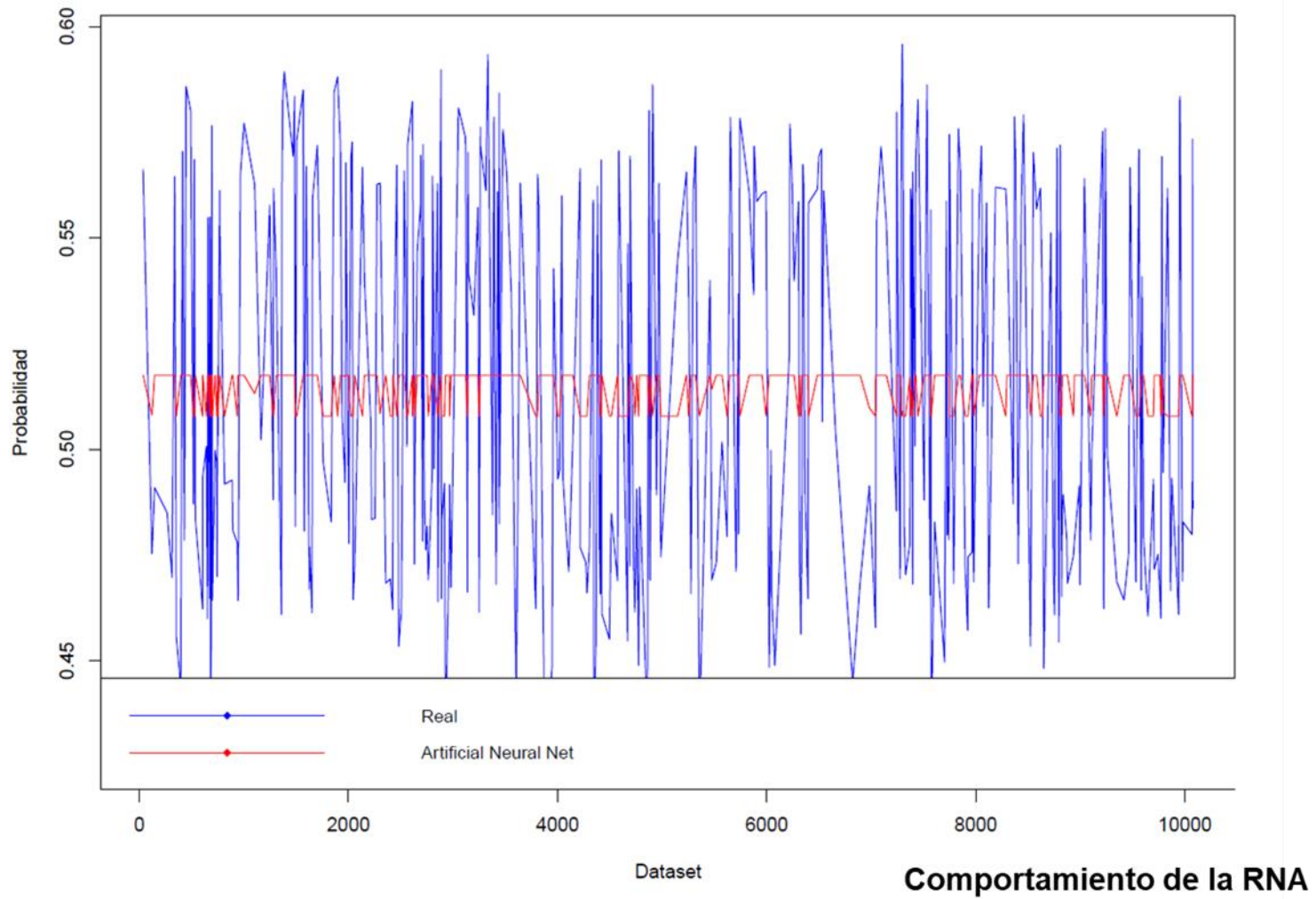


2. Configuración: 1 capa, 1 neurona, umbral 0.05.

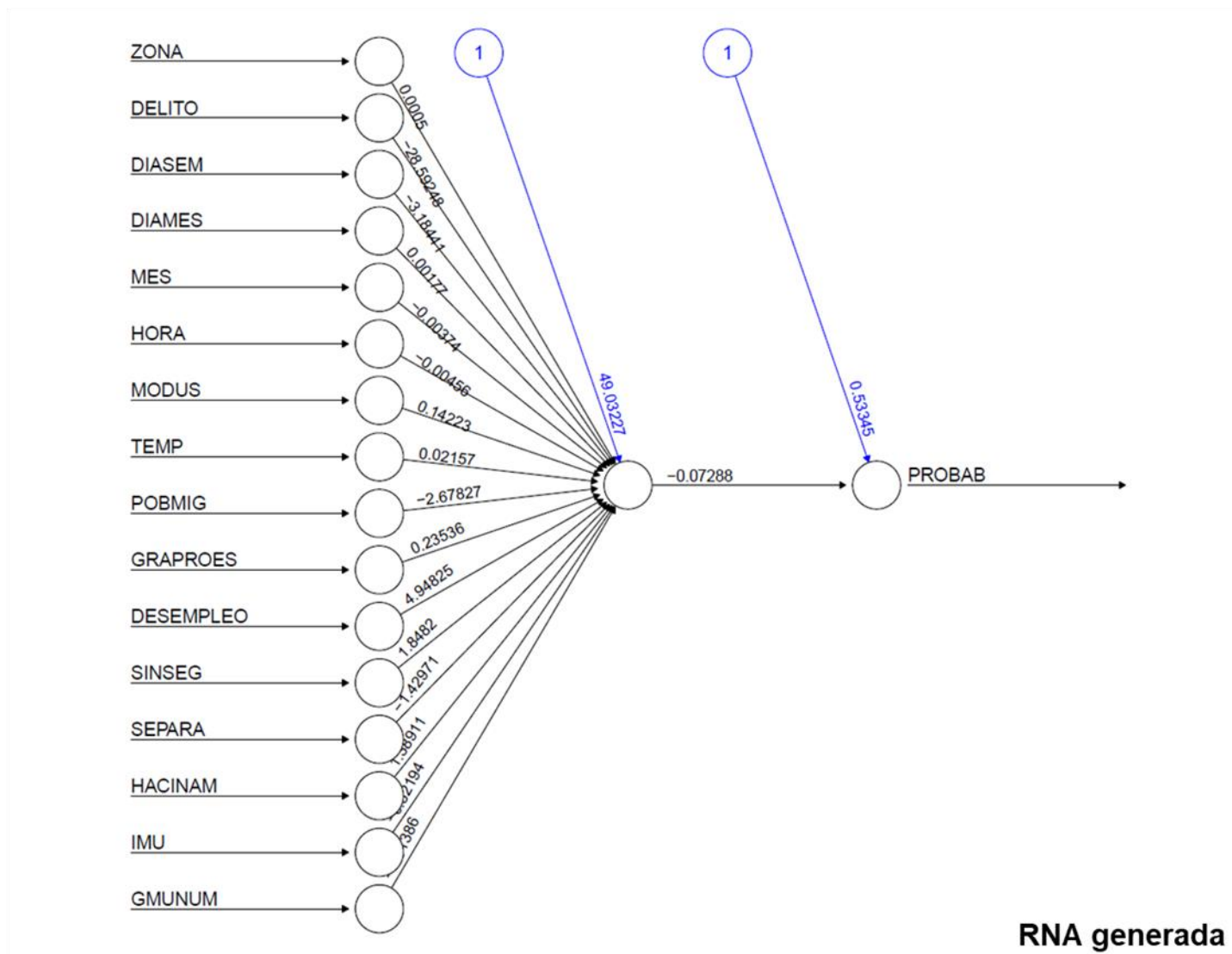


RNA generada

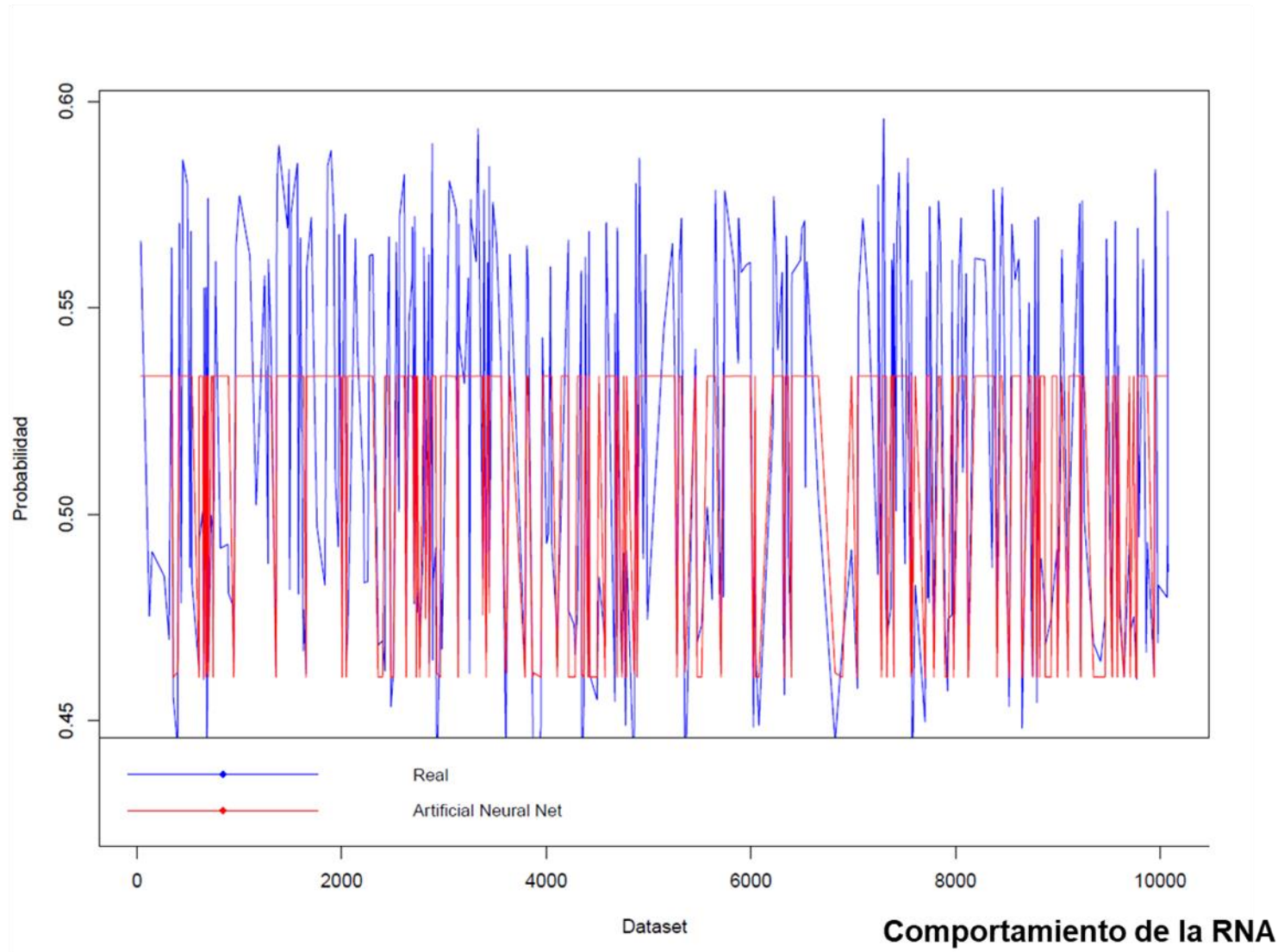
2. Configuración: 1 capa, 1 neurona, umbral 0.05.



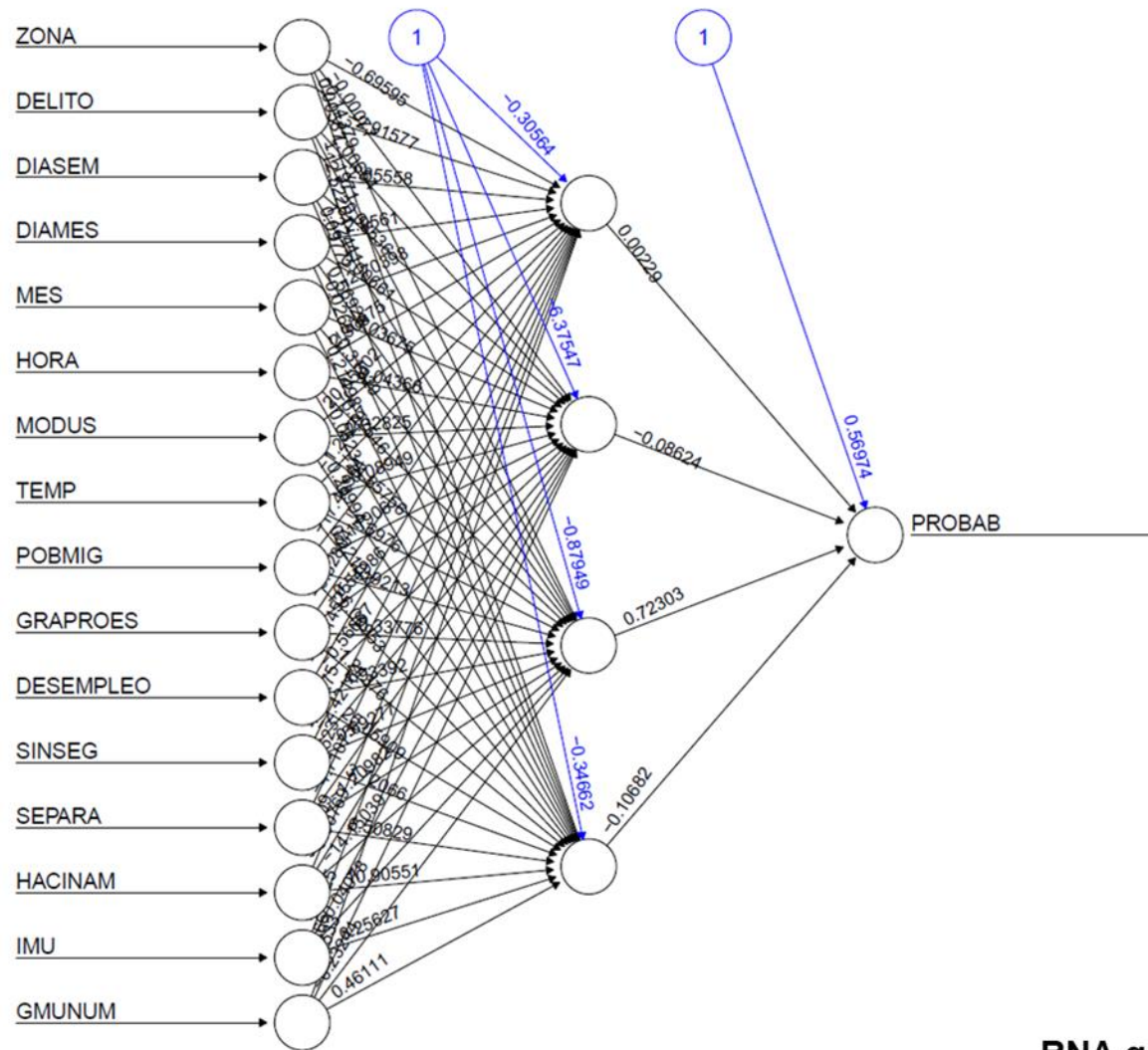
3. Configuración: 1 capa, 1 neurona, umbral 0.01.



3. Configuración: 1 capa, 1 neurona, umbral 0.01.

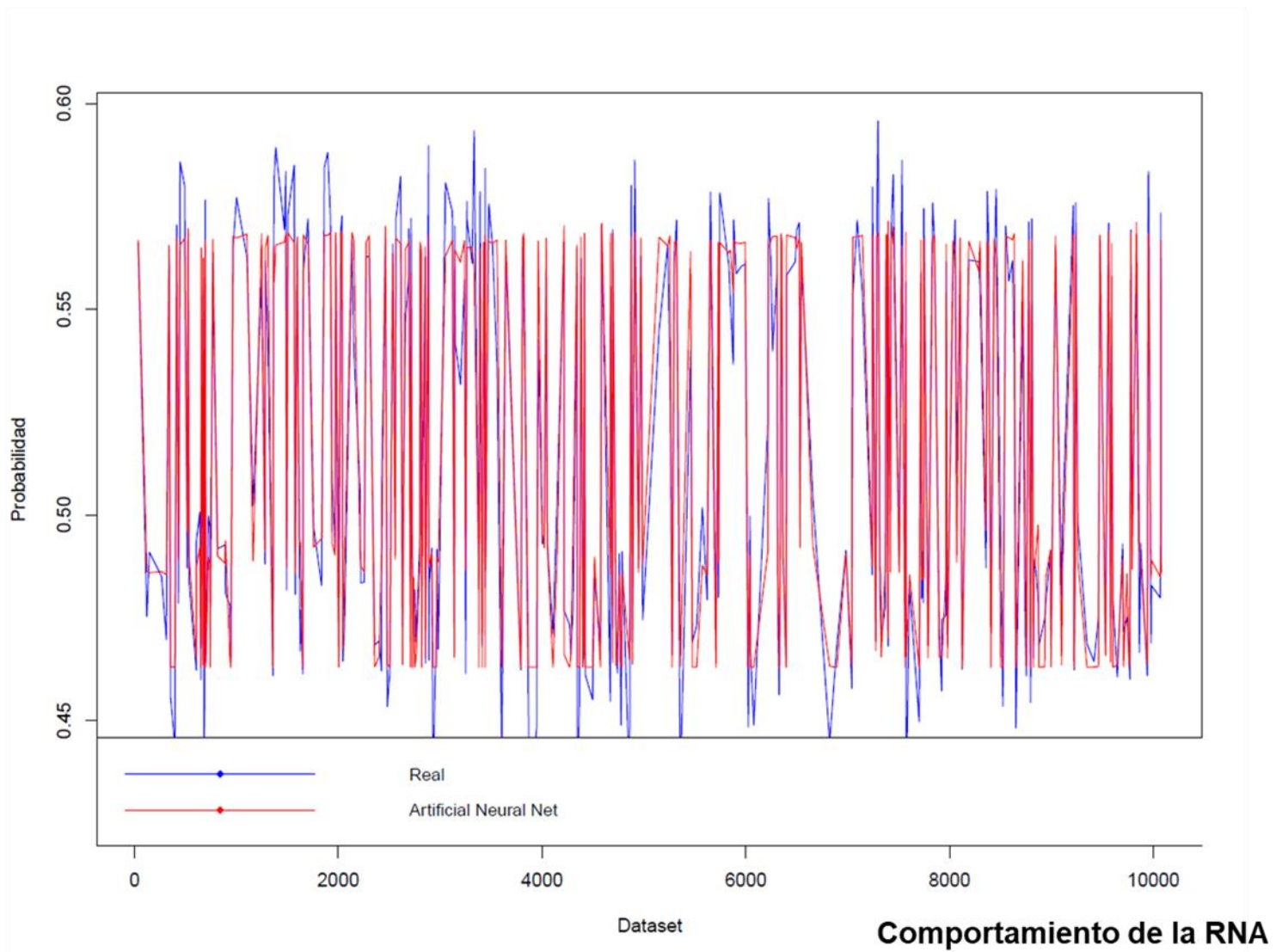


4. Configuración: 1 capa, 4n, umbral 0.1.

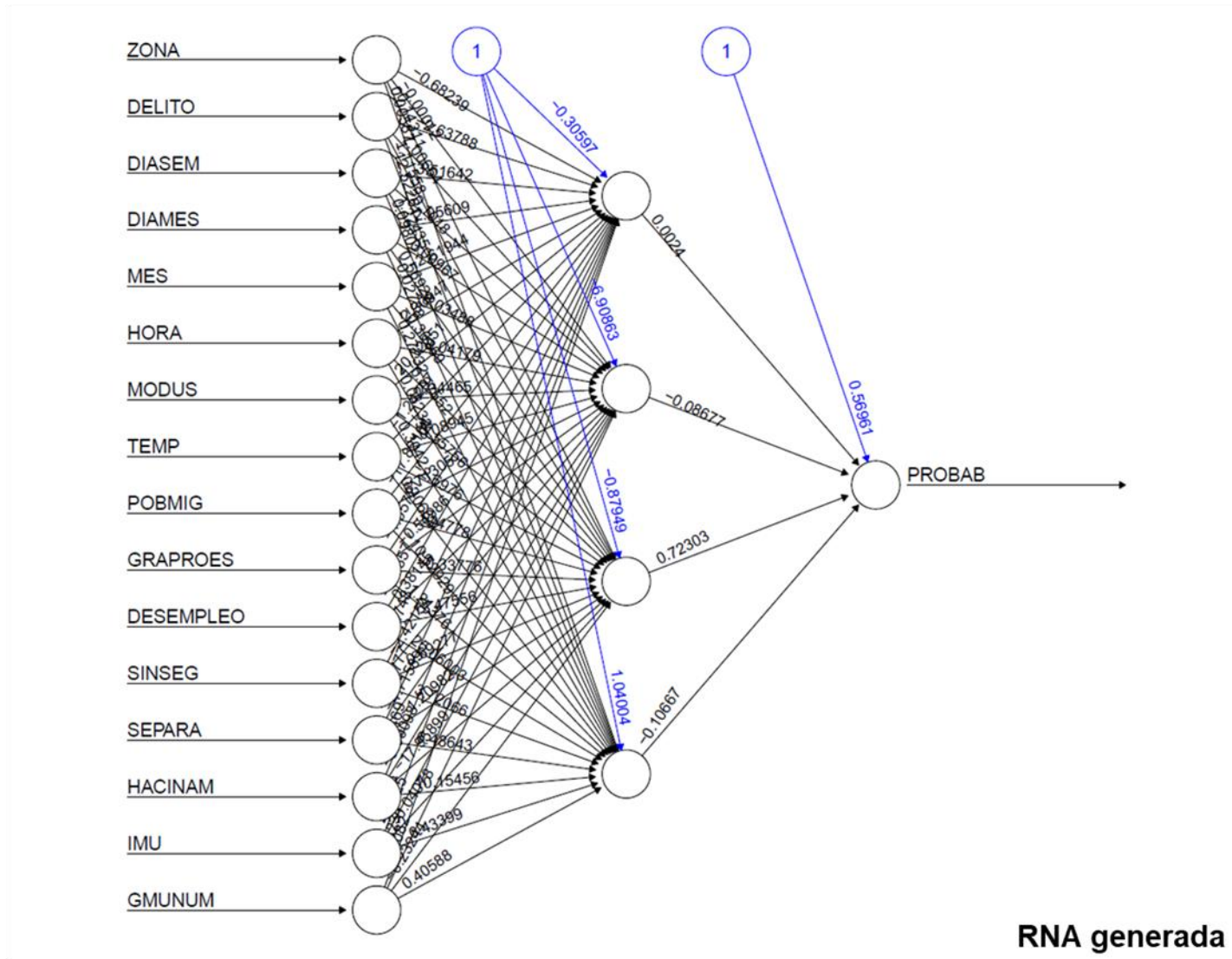


RNA generada

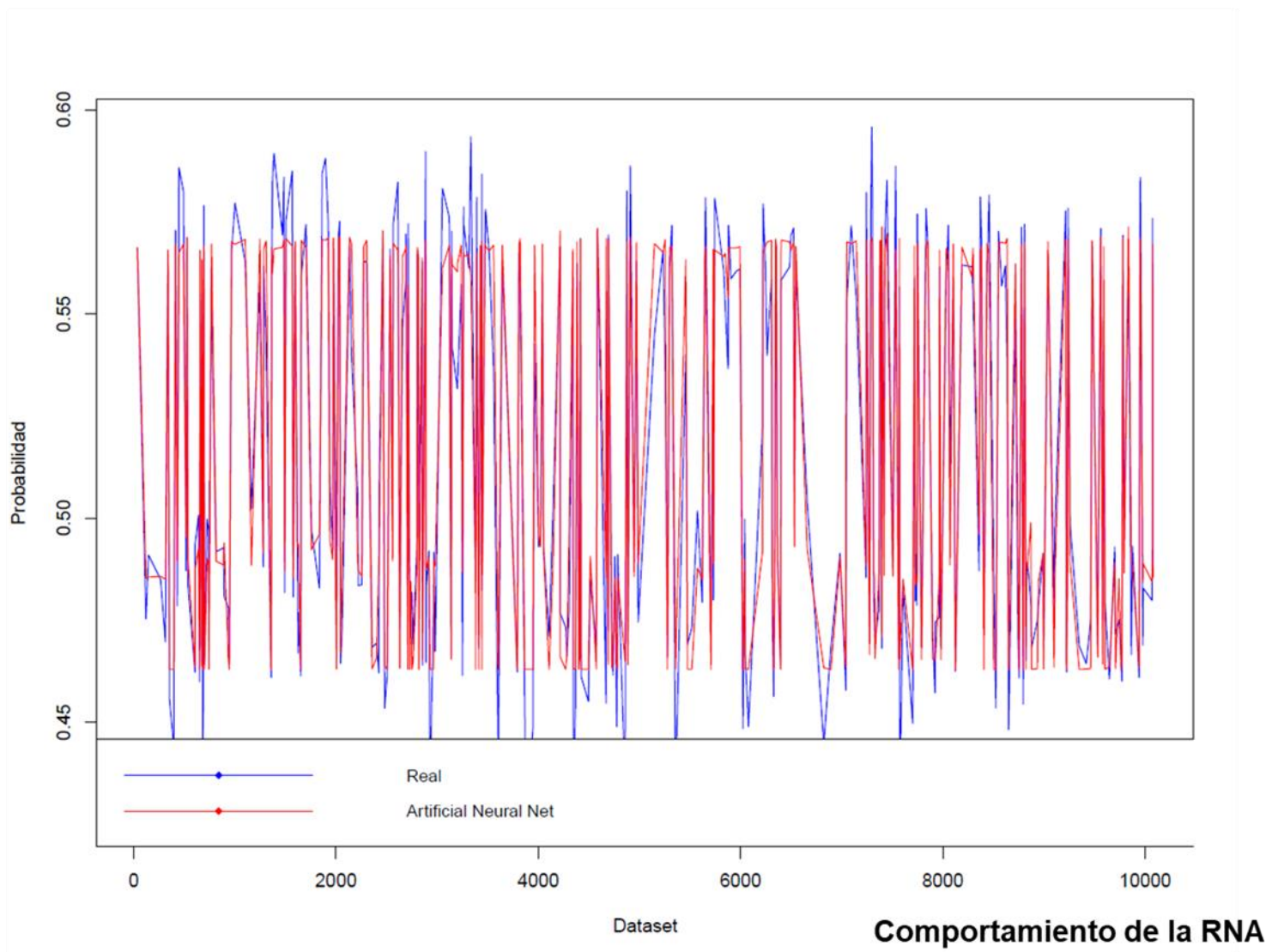
4. Configuración: 1 capa, 4n, umbral 0.1.



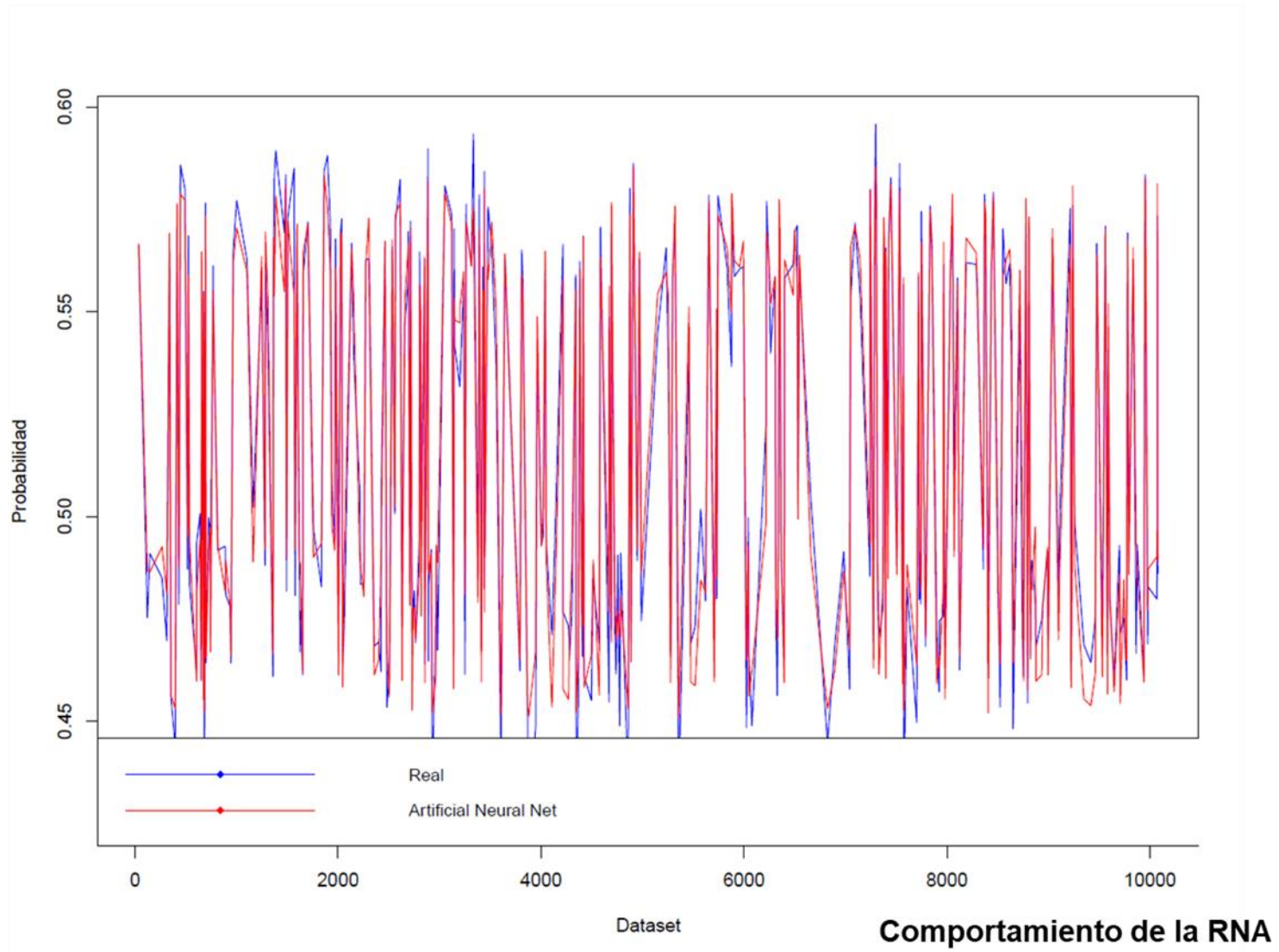
5. Configuración: 1 capa, 4n, umbral 0.05.



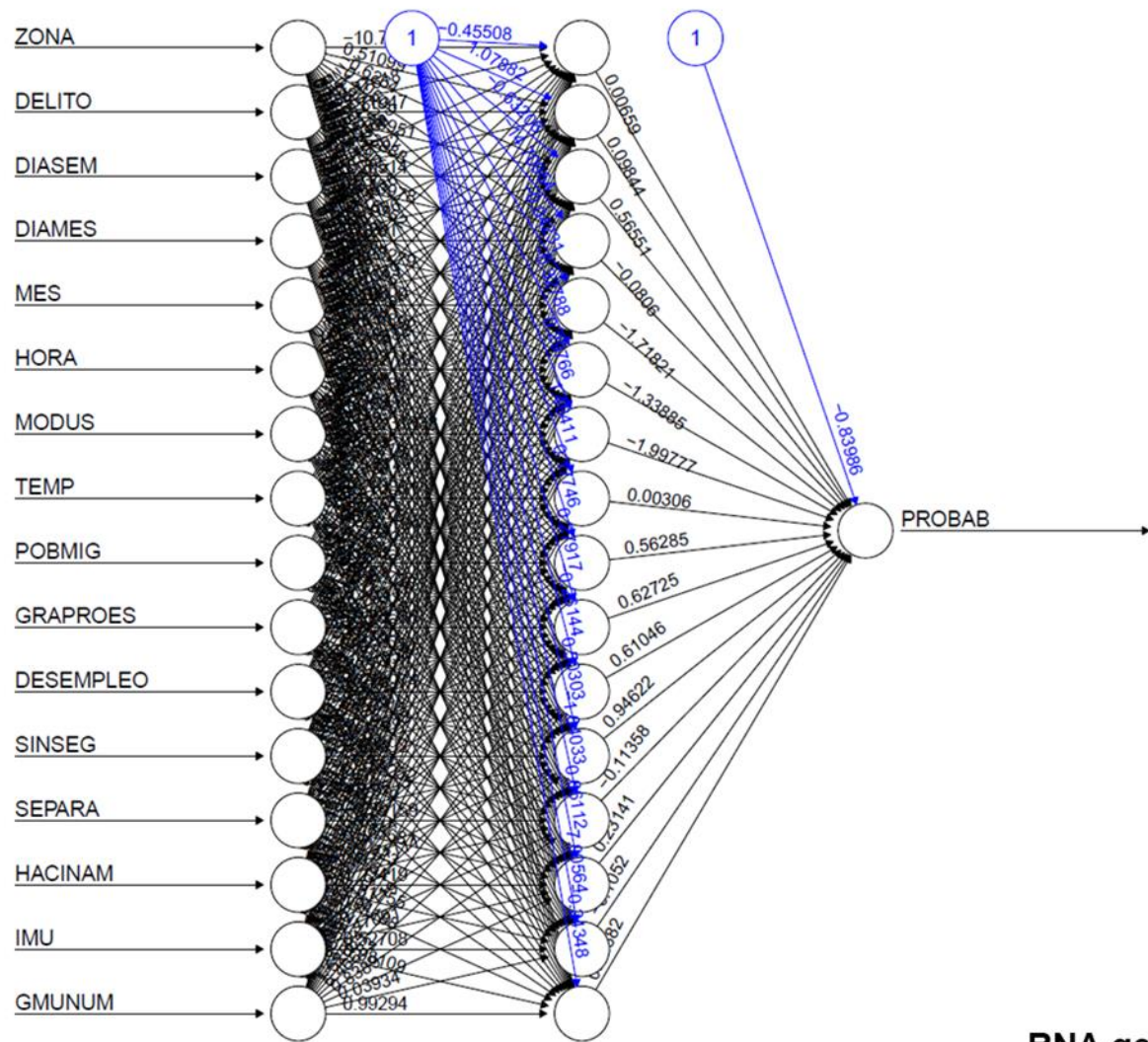
5. Configuración: 1 capa, 4n, umbral 0.05.



6. Configuración: 1 capa, 4n, umbral 0.01.

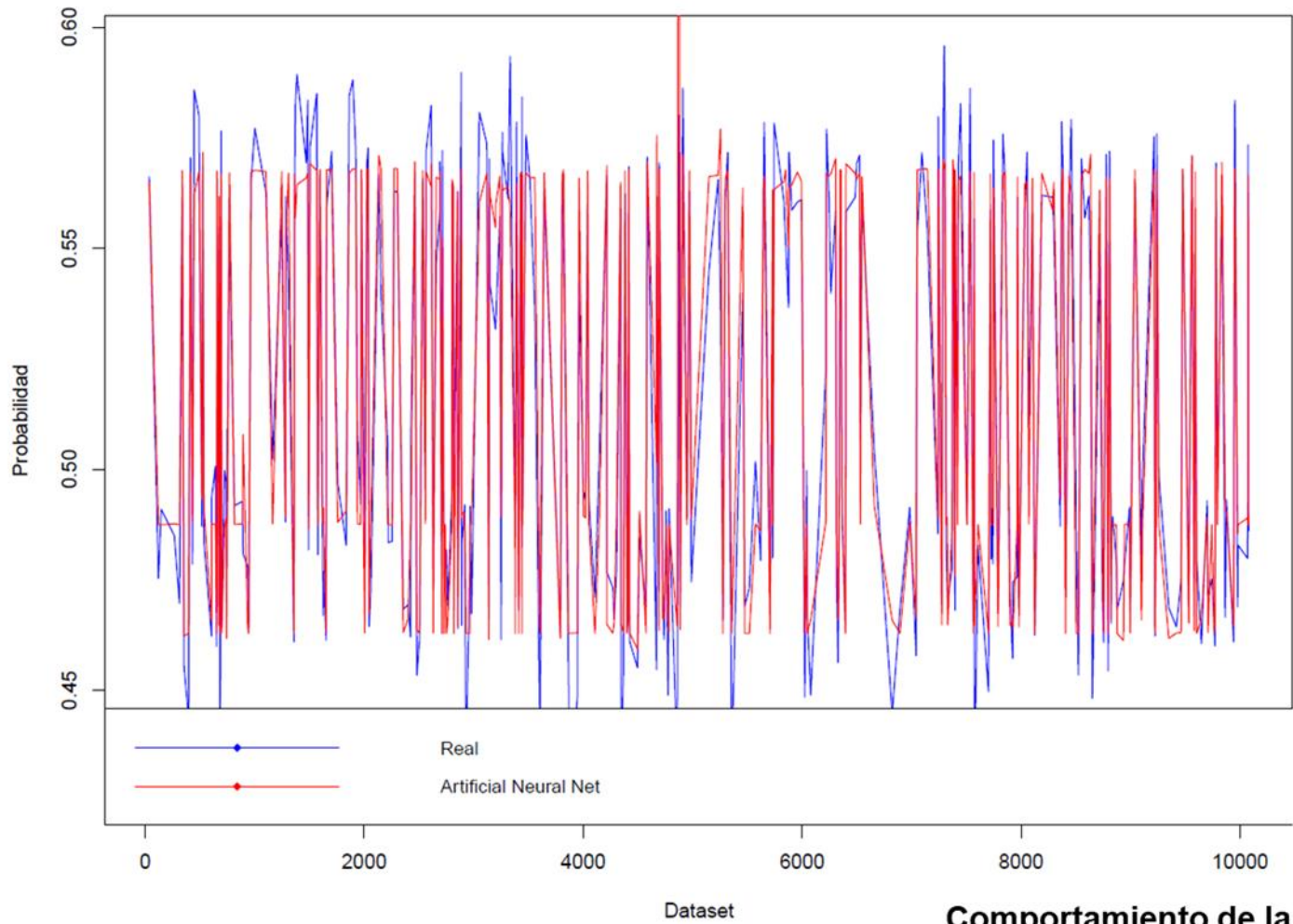


7. Configuración: 1 capa, 16n, umbral 0.1.

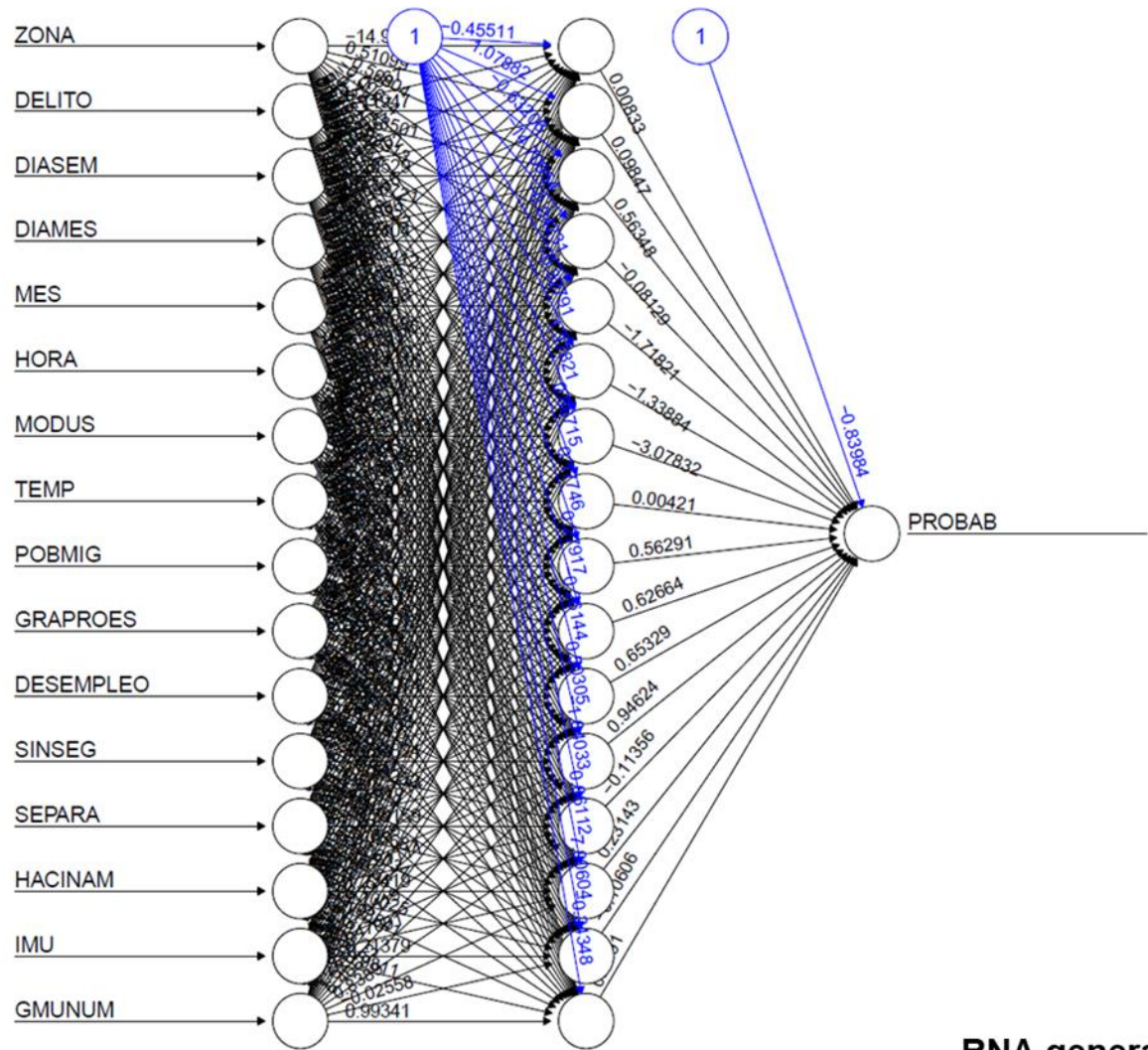


RNA generada

7. Configuración: 1 capa, 16n, umbral 0.1.

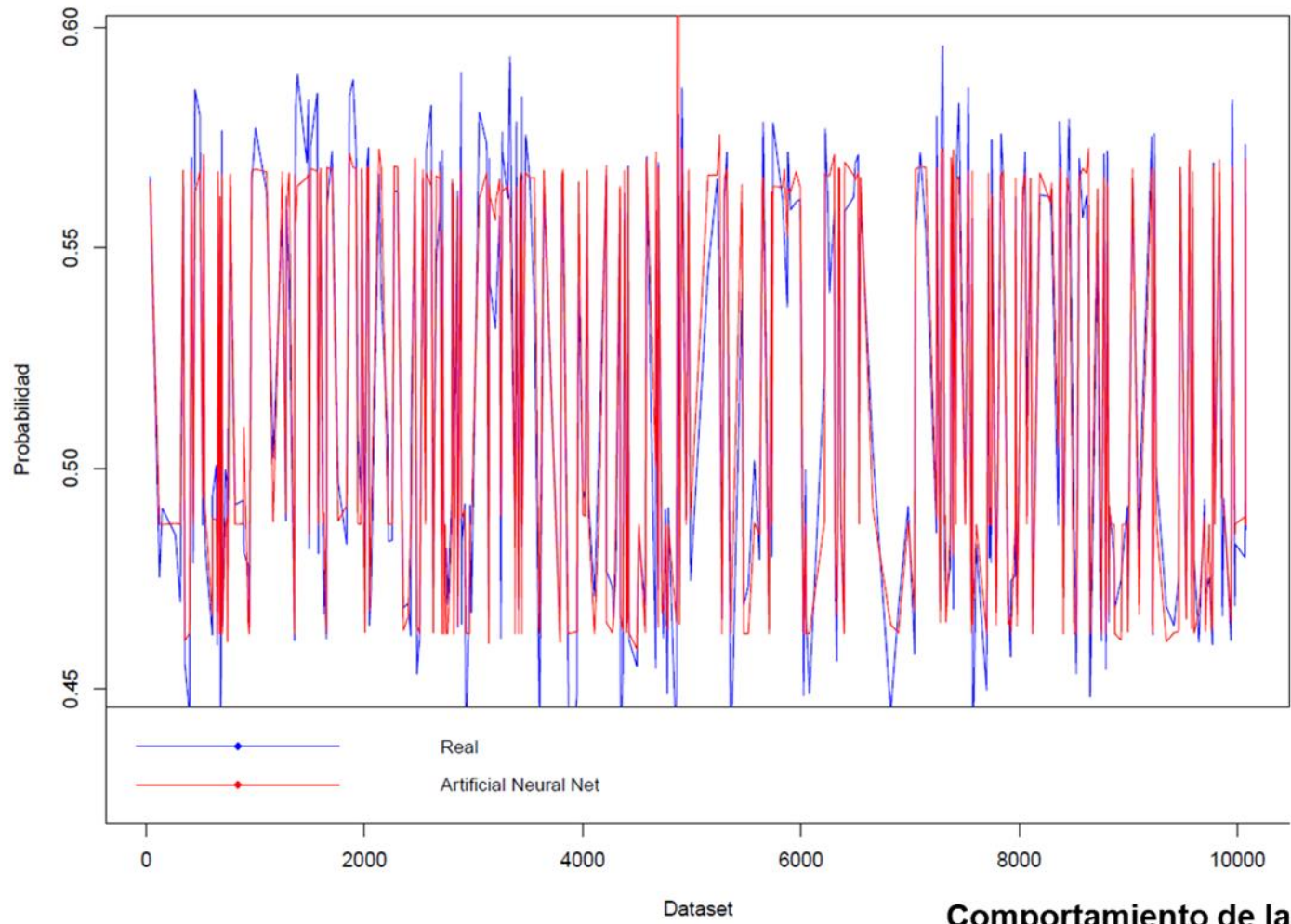


8. Configuración: 1 capa, 16n, umbral 0.05.



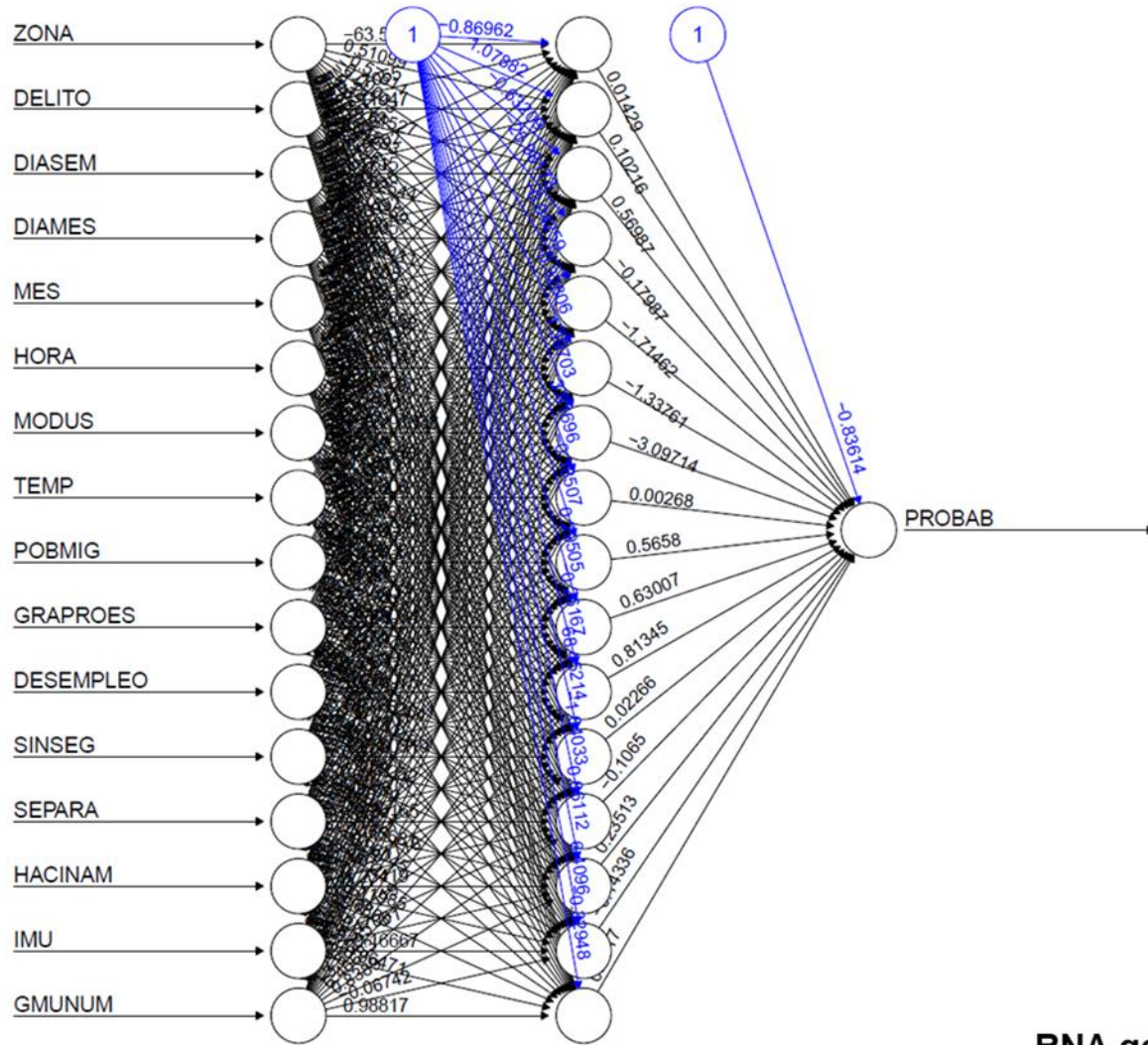
RNA generada

8. Configuración: 1 capa, 16n, umbral 0.05.



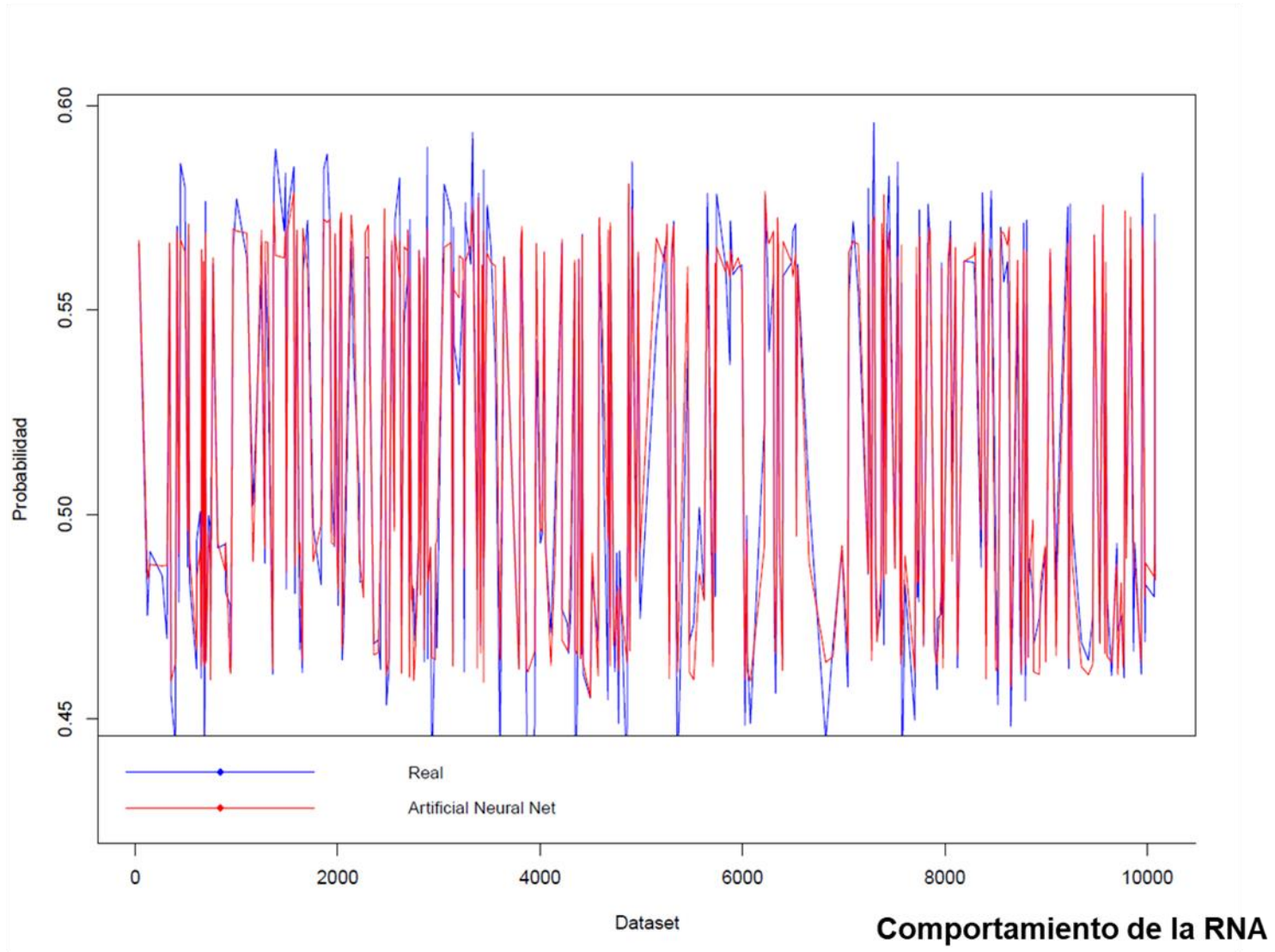
Comportamiento de la RNA

9. Configuración: 1 capa, 16n, umbral 0.01.

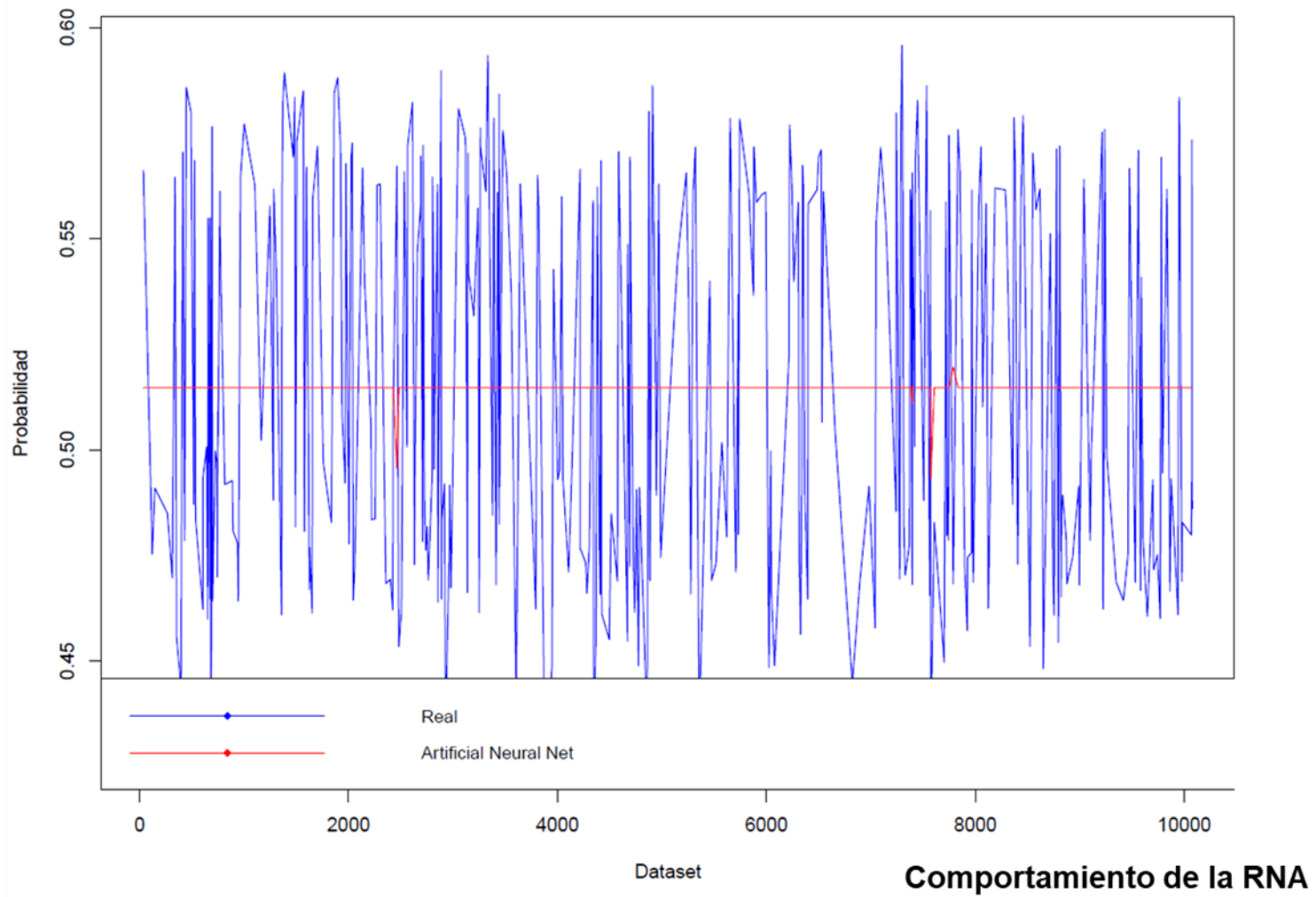


RNA generada

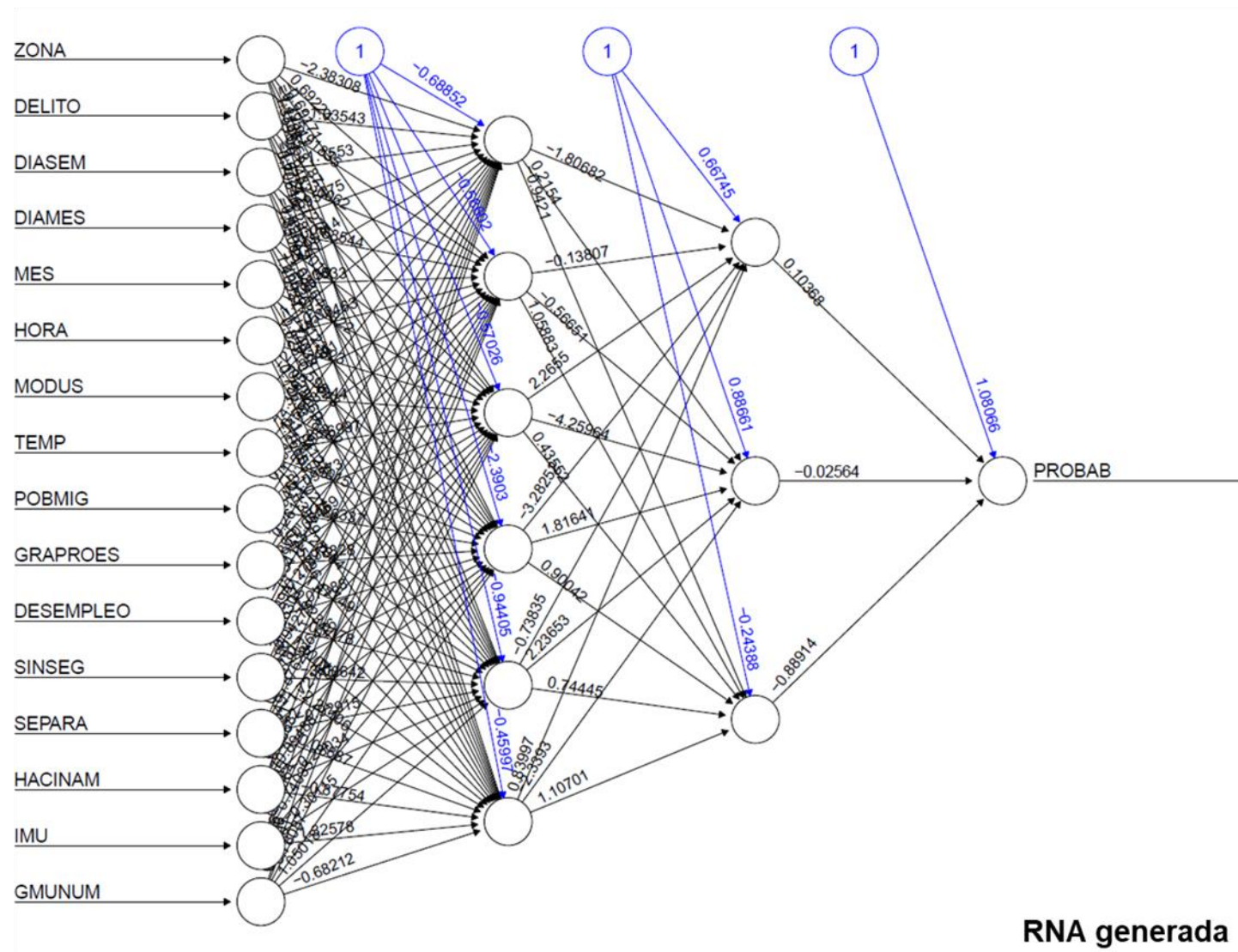
9. Configuración: 1 capa, 16n, umbral 0.01.



10. Configuración: 2 capas, (6n,3n), umbral 0.1.

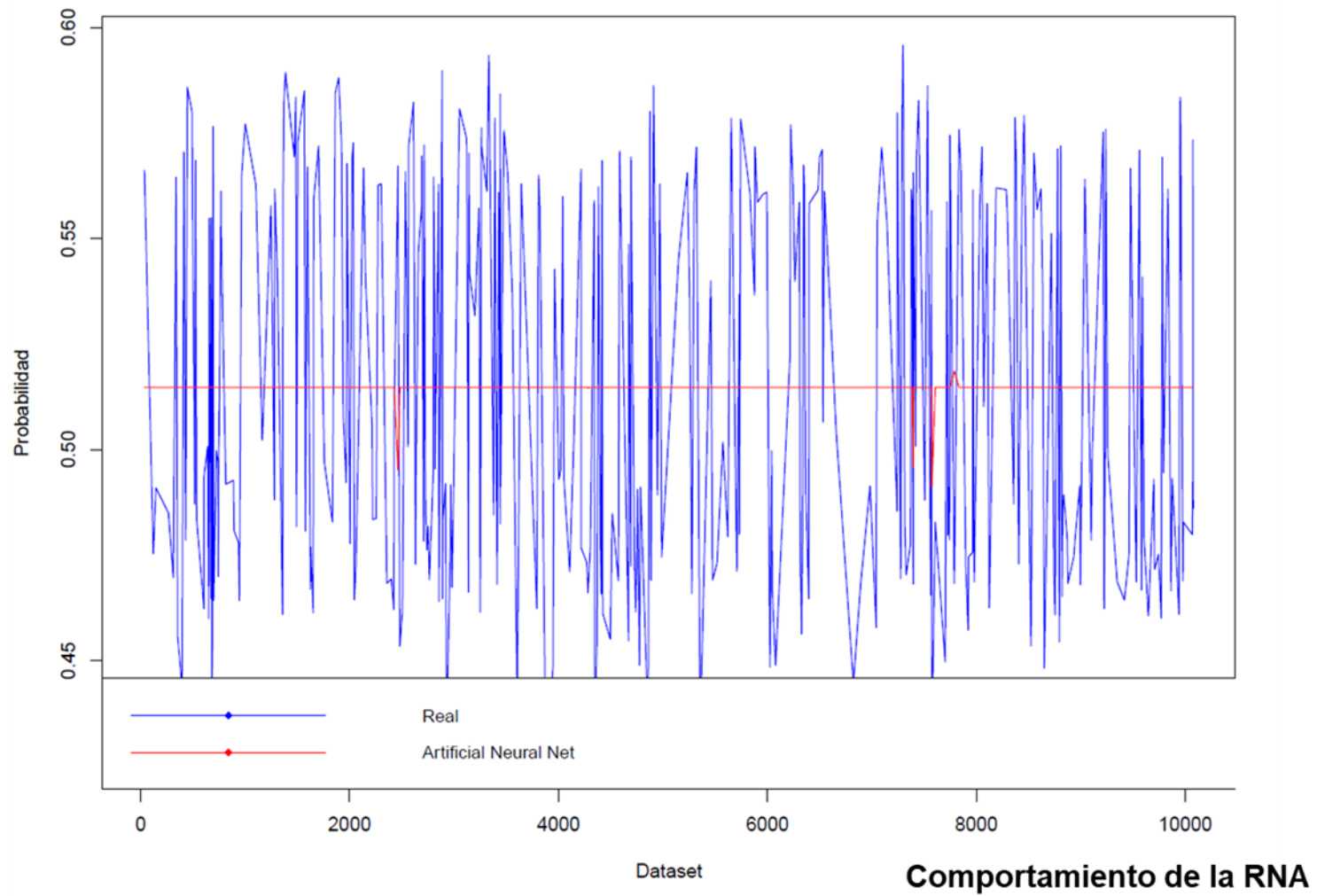


11. Configuración: 2 capas, (6n,3n), umbral 0.05.

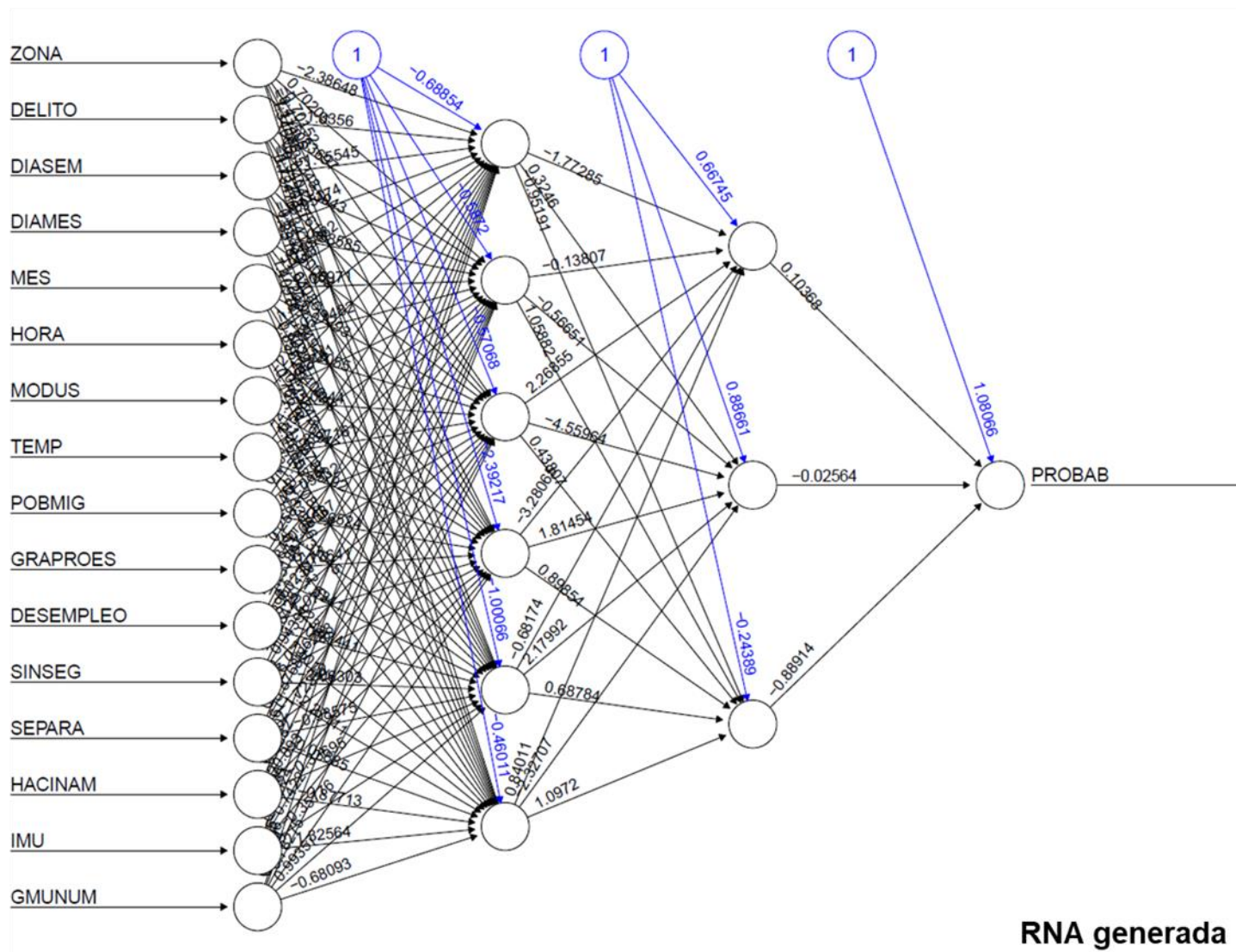


RNA generada

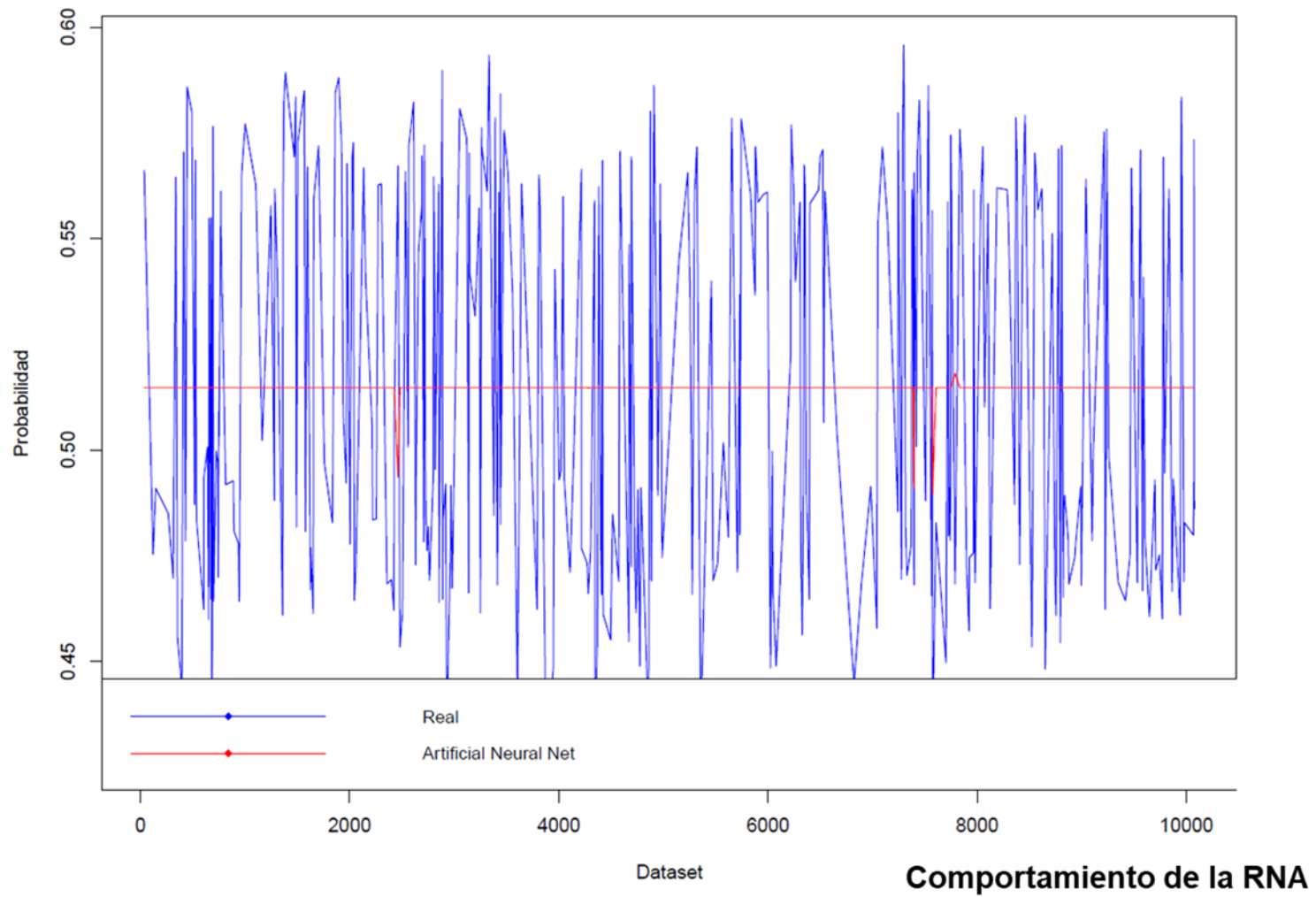
11. Configuración: 2 capas, (6n,3n), umbral 0.05.



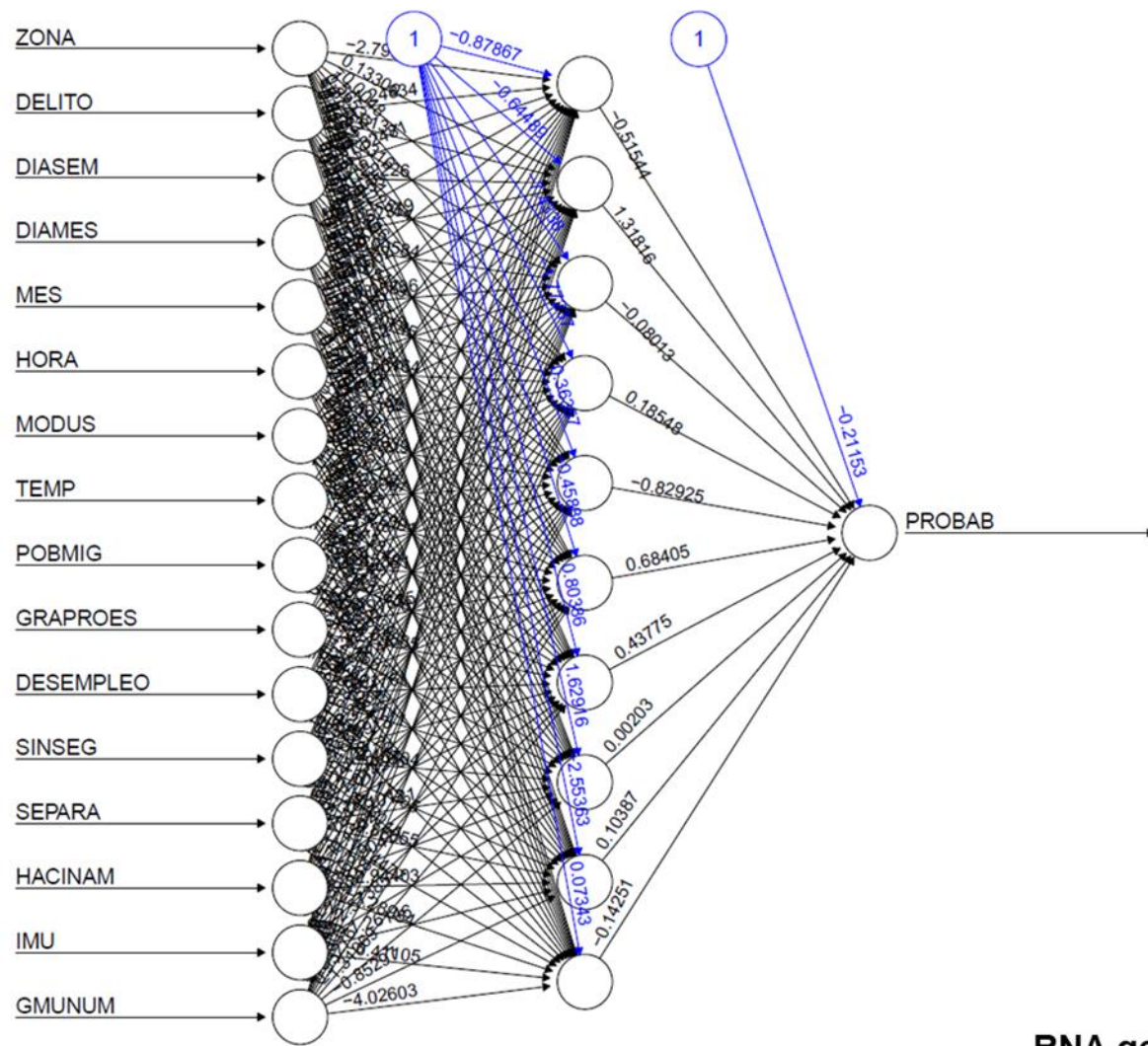
12. Configuración: 2 capas, (6n,3n), umbral 0.01.



12. Configuración: 2 capas, (6n,3n), umbral 0.01.

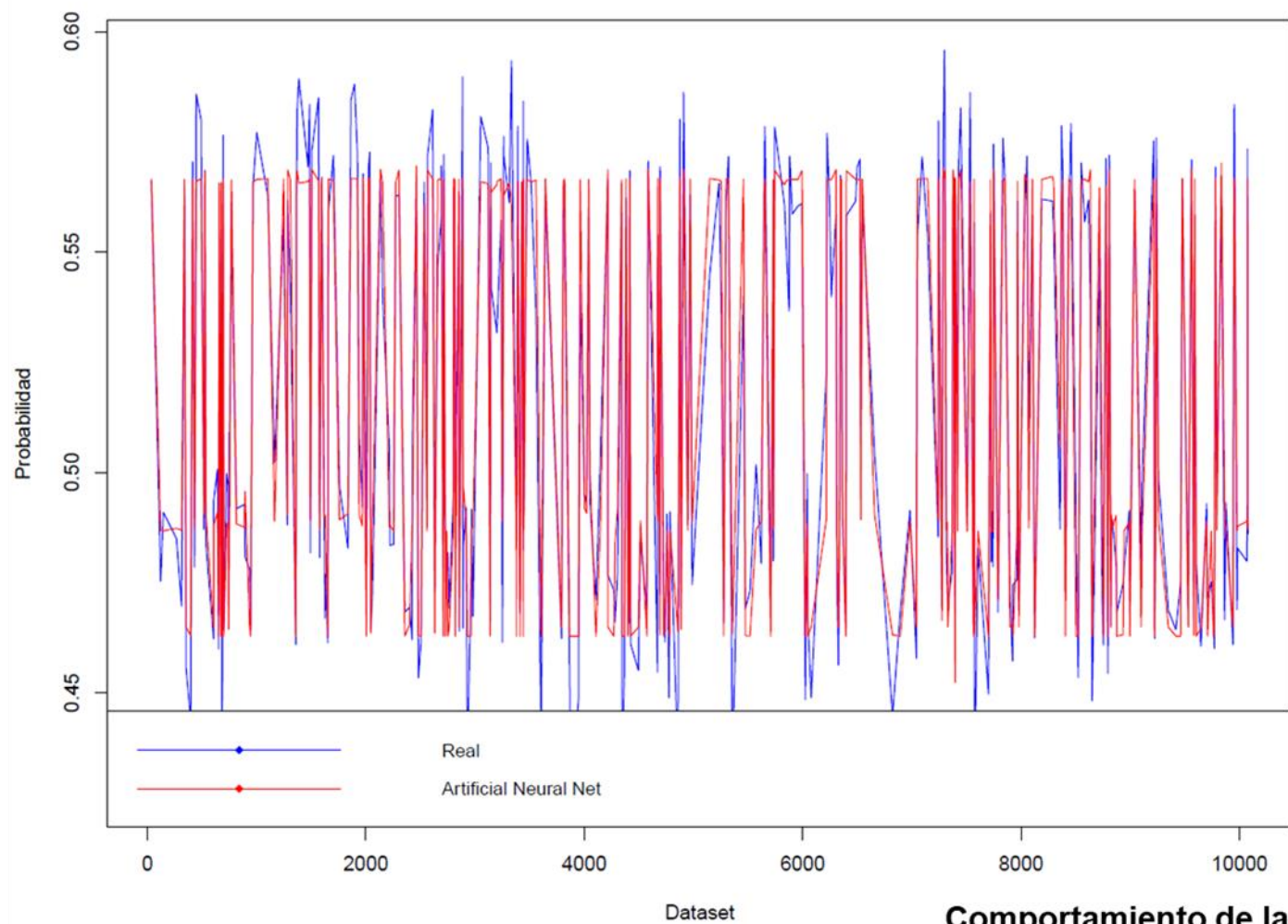


13. Configuración: 1 capa, 10n, umbral 0.1.



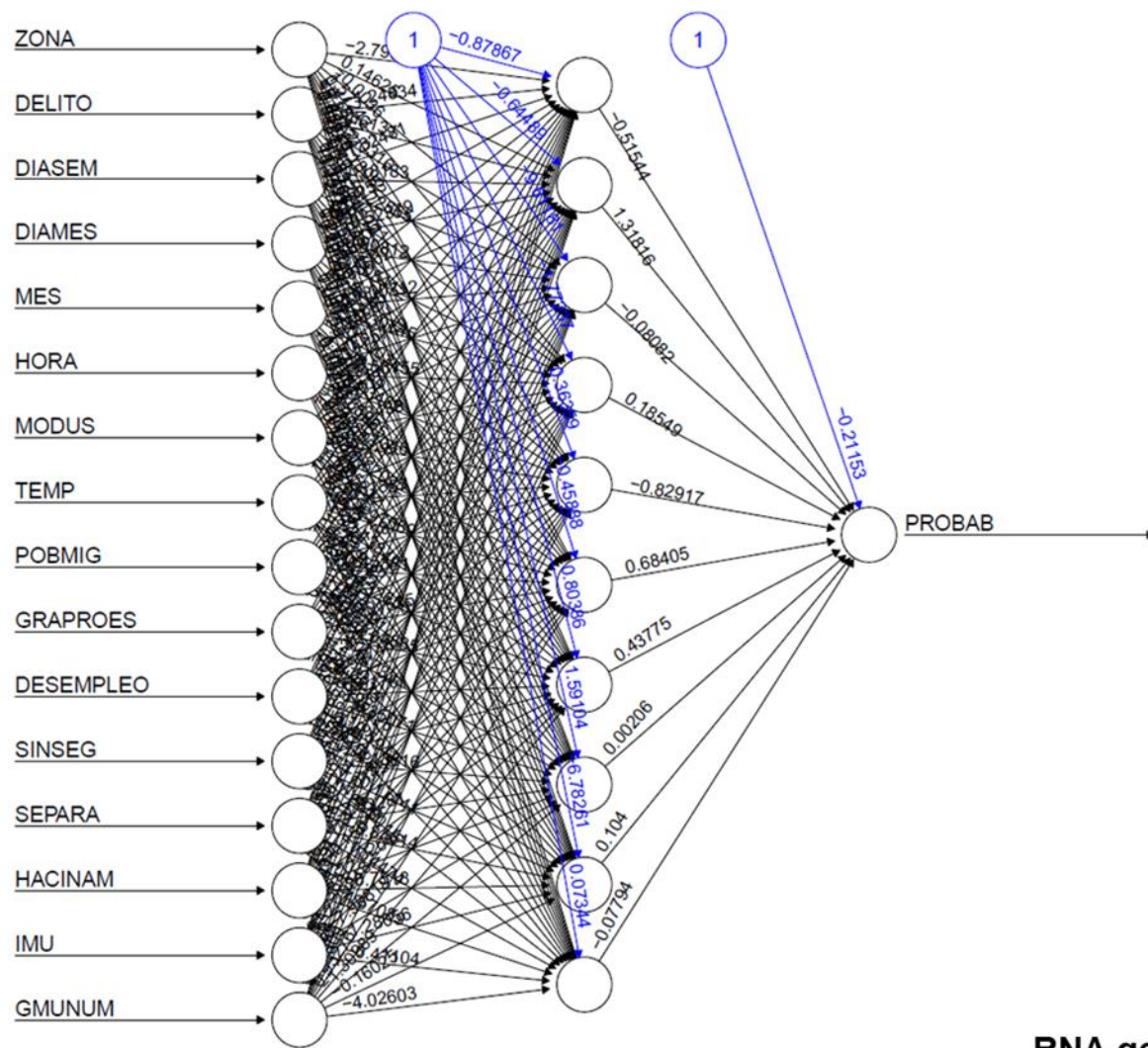
RNA generada

13. Configuración: 1 capa, 10n, umbral 0.1.



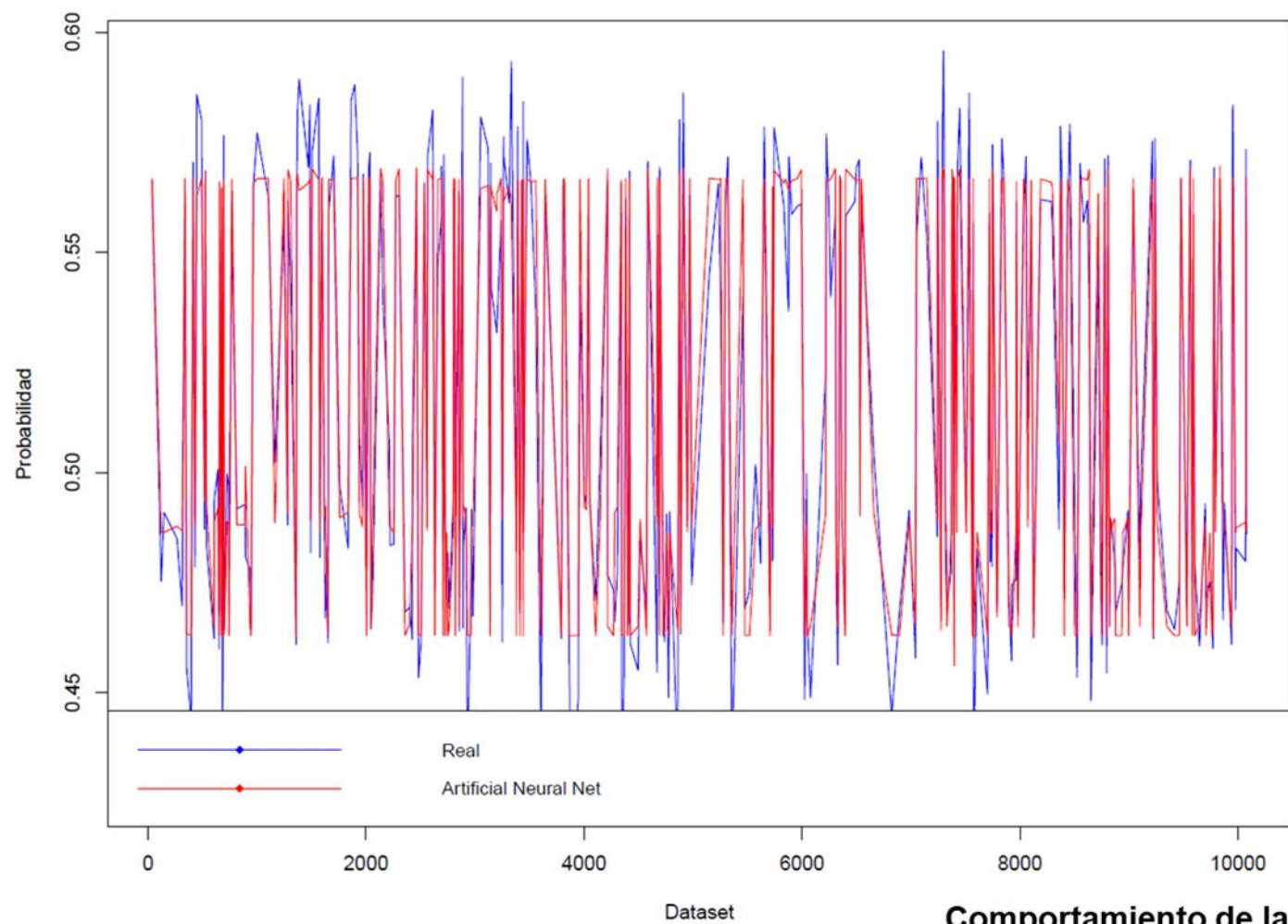
Comportamiento de la RNA

14. Configuración: 1 capa, 10n, umbral 0.05.

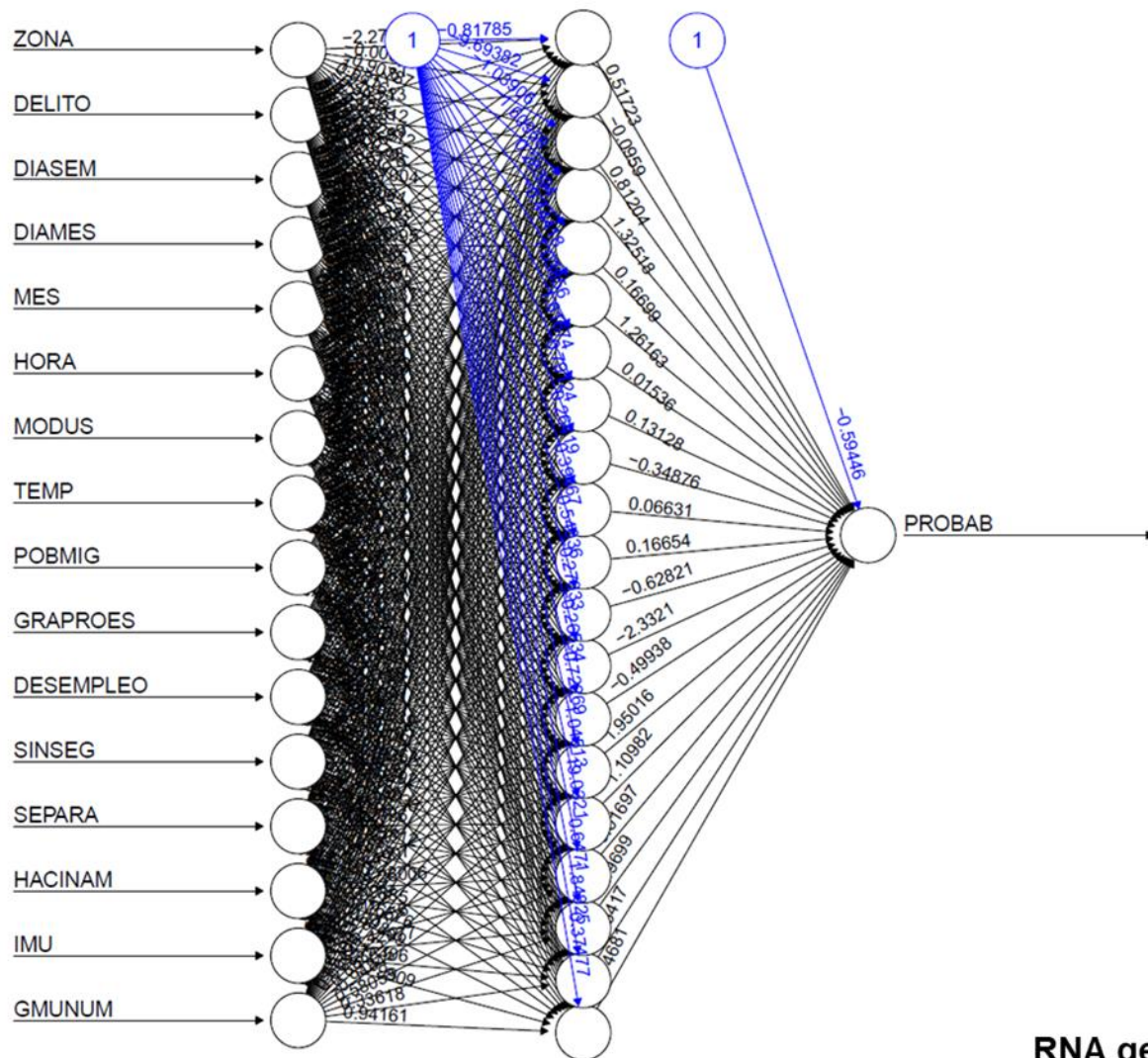


RNA generada

14. Configuración: 1 capa, 10n, umbral 0.05.

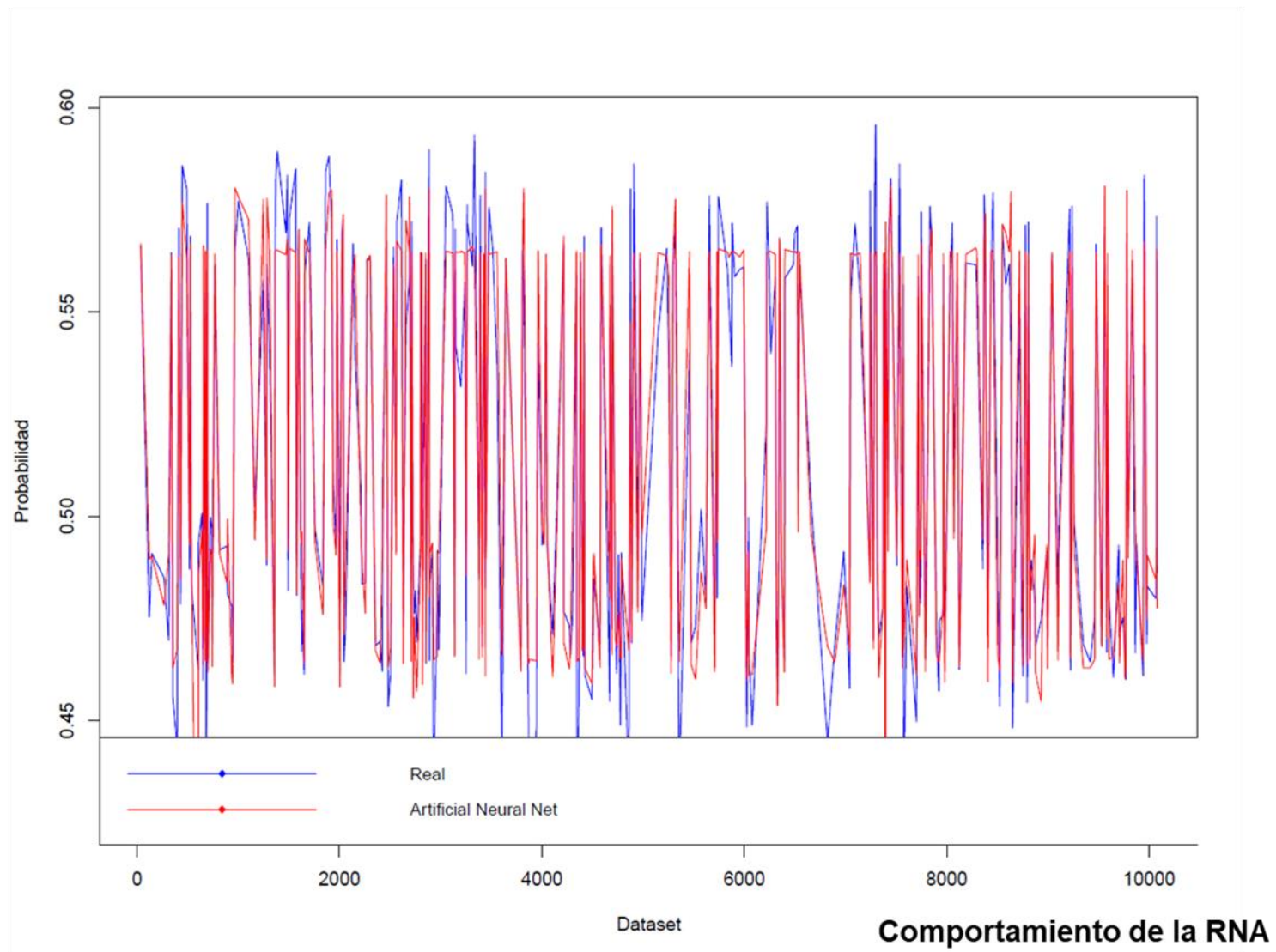


15. Configuración: 1 capa, 20n, umbral 0.1.

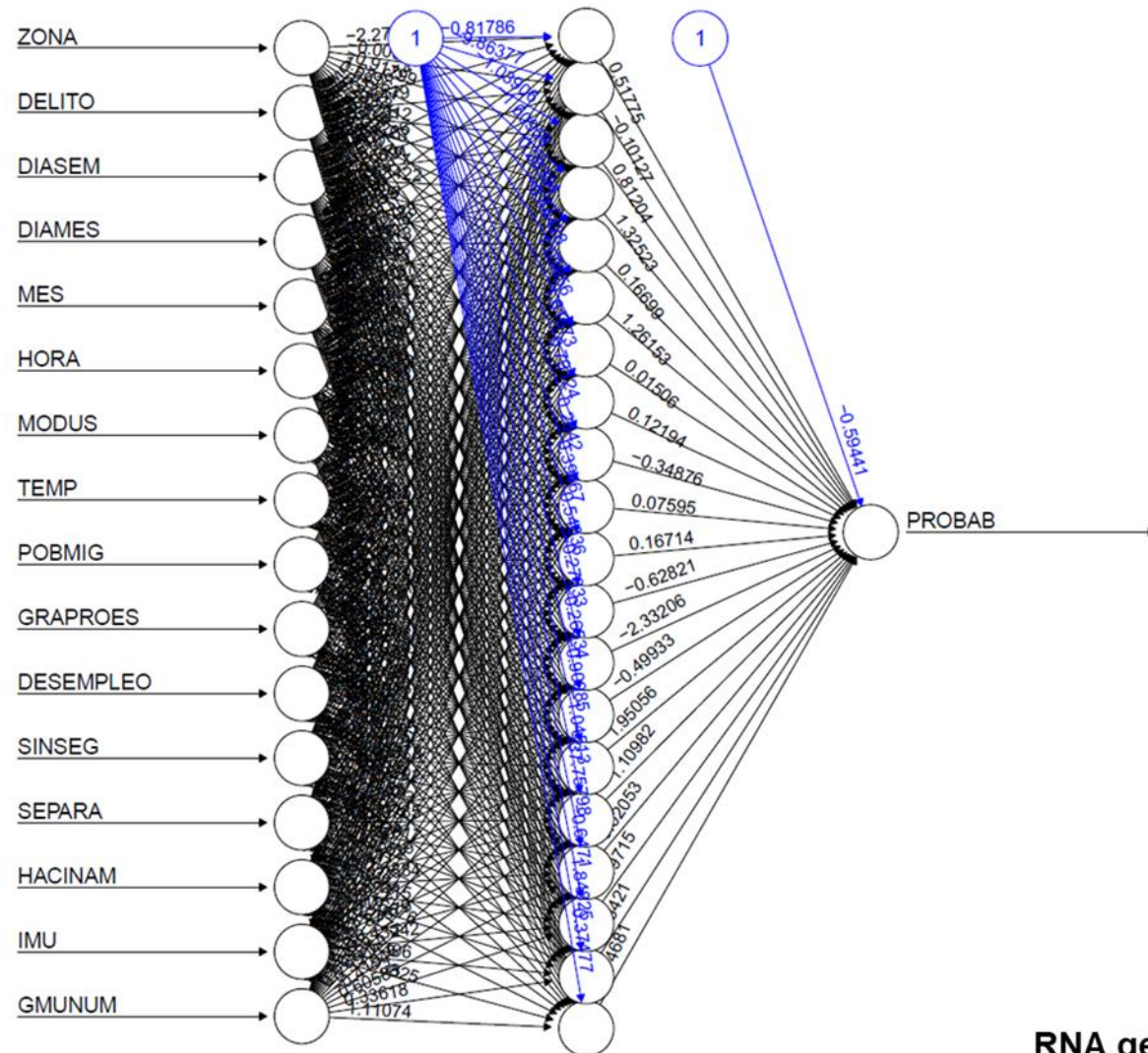


RNA generada

15. Configuración: 1 capa, 20n, umbral 0.1.

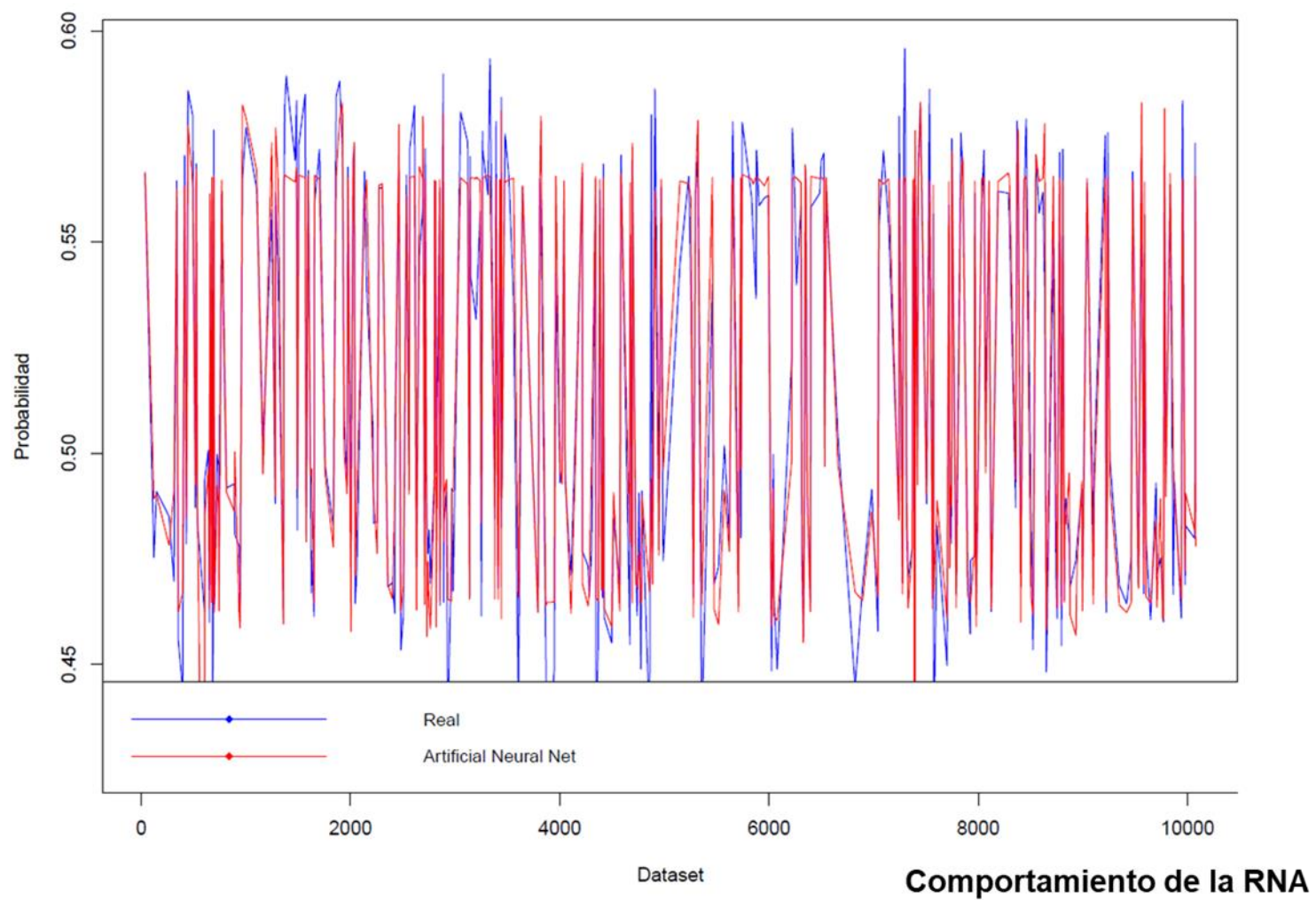


16. Configuración: 1 capa, 20n, umbral 0.05.

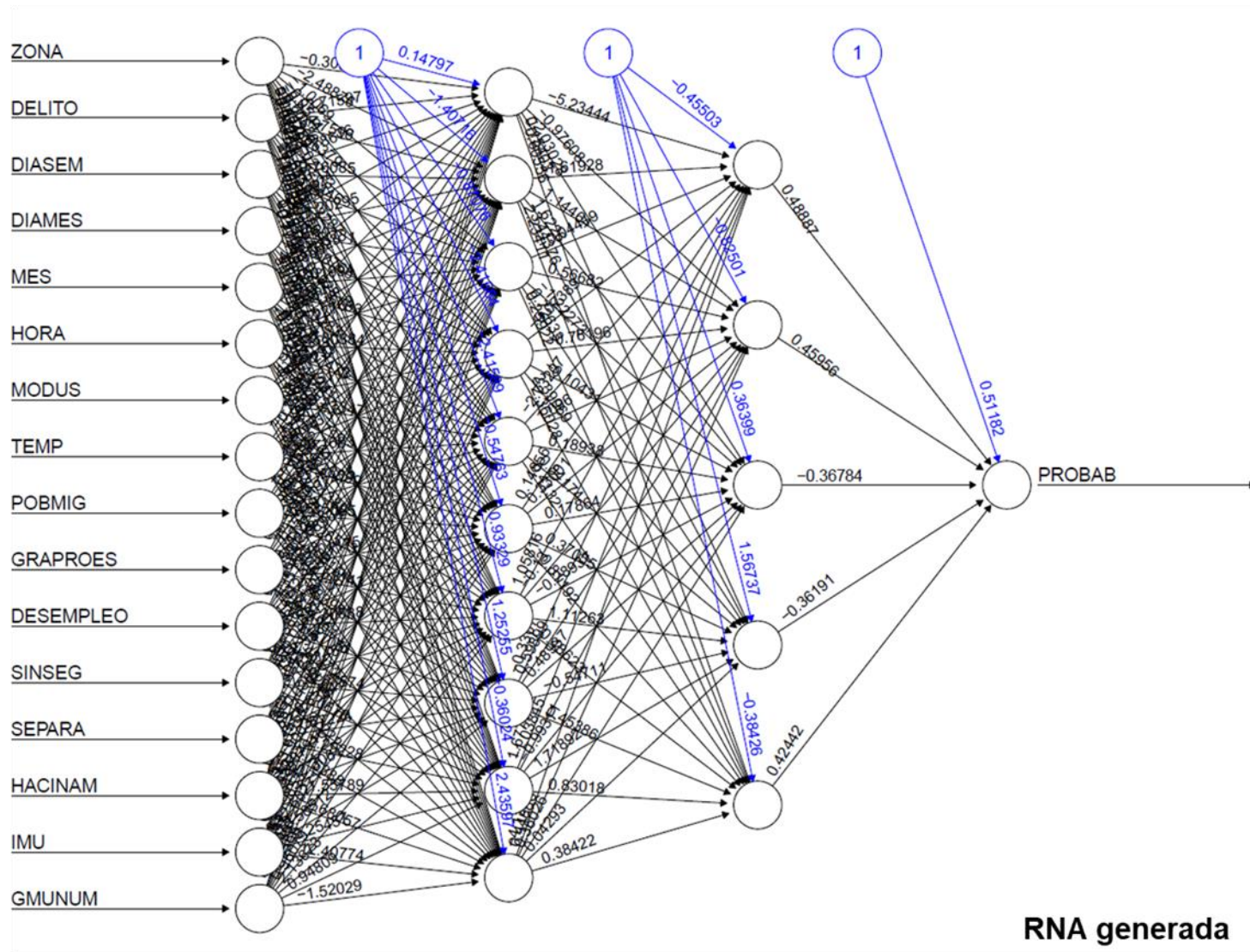


RNA generada

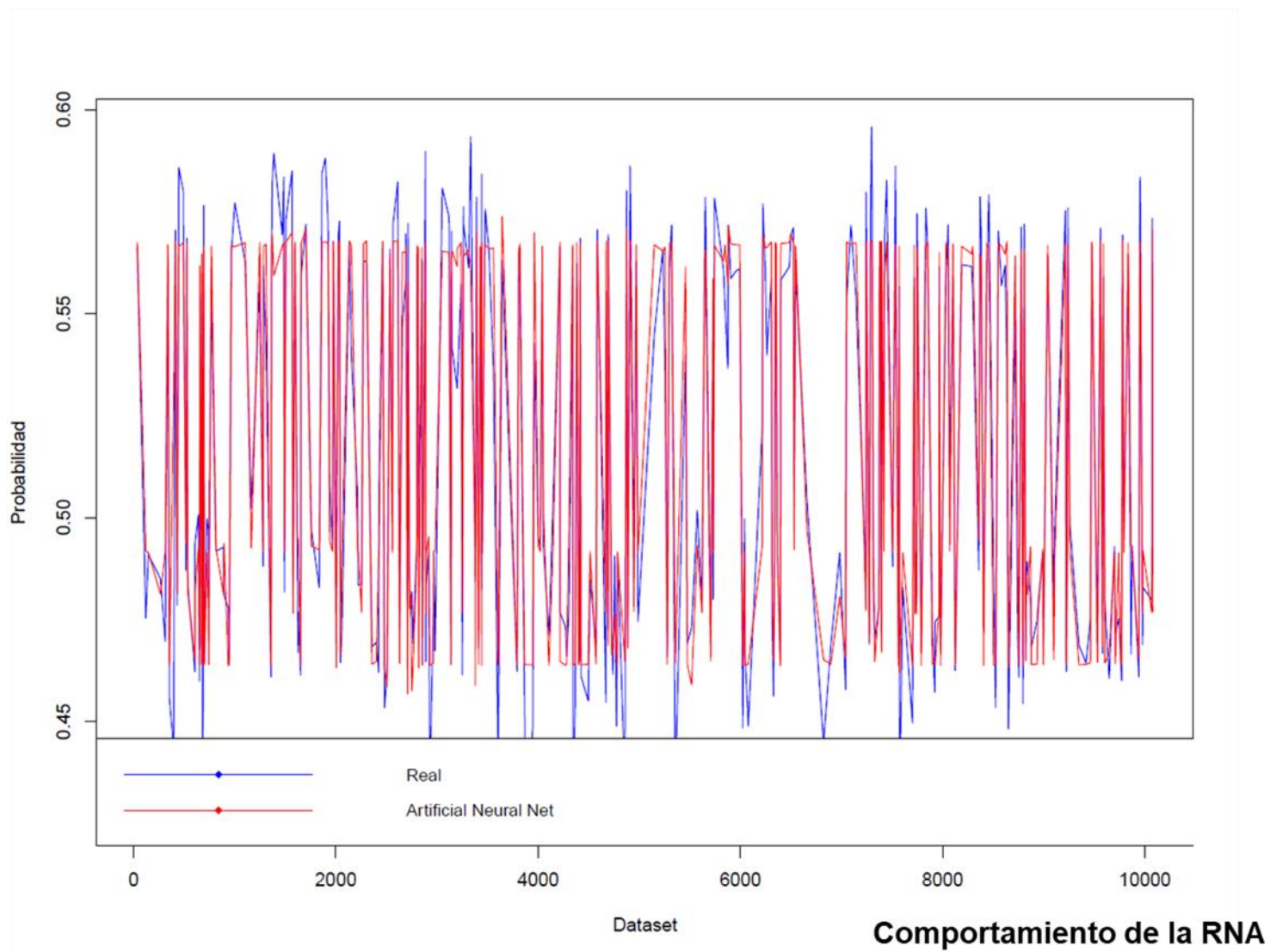
16. Configuración: 1 capa, 20n, umbral 0.05.



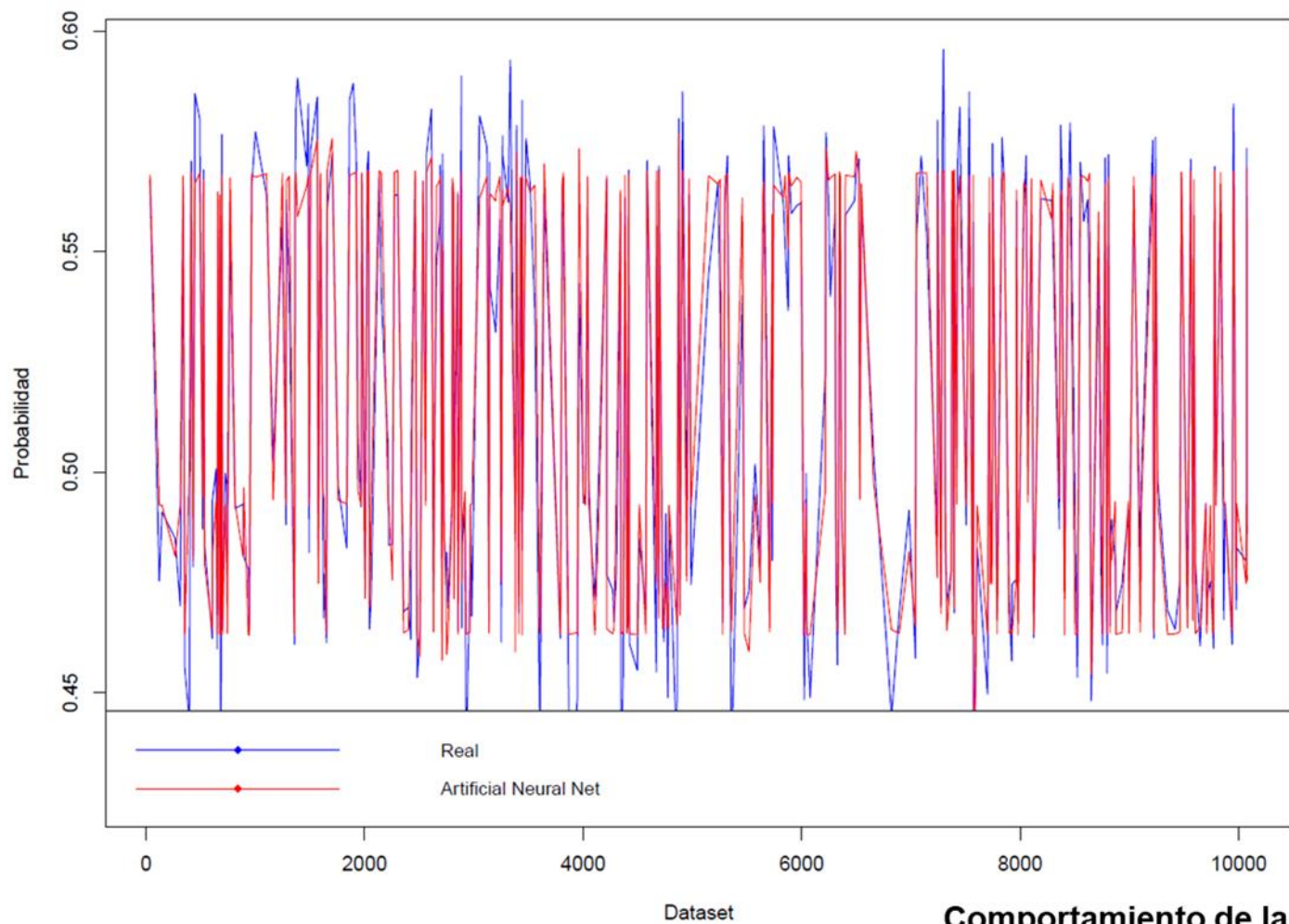
17. Configuración: 2 capas, (10n,5n), umbral 0.1.



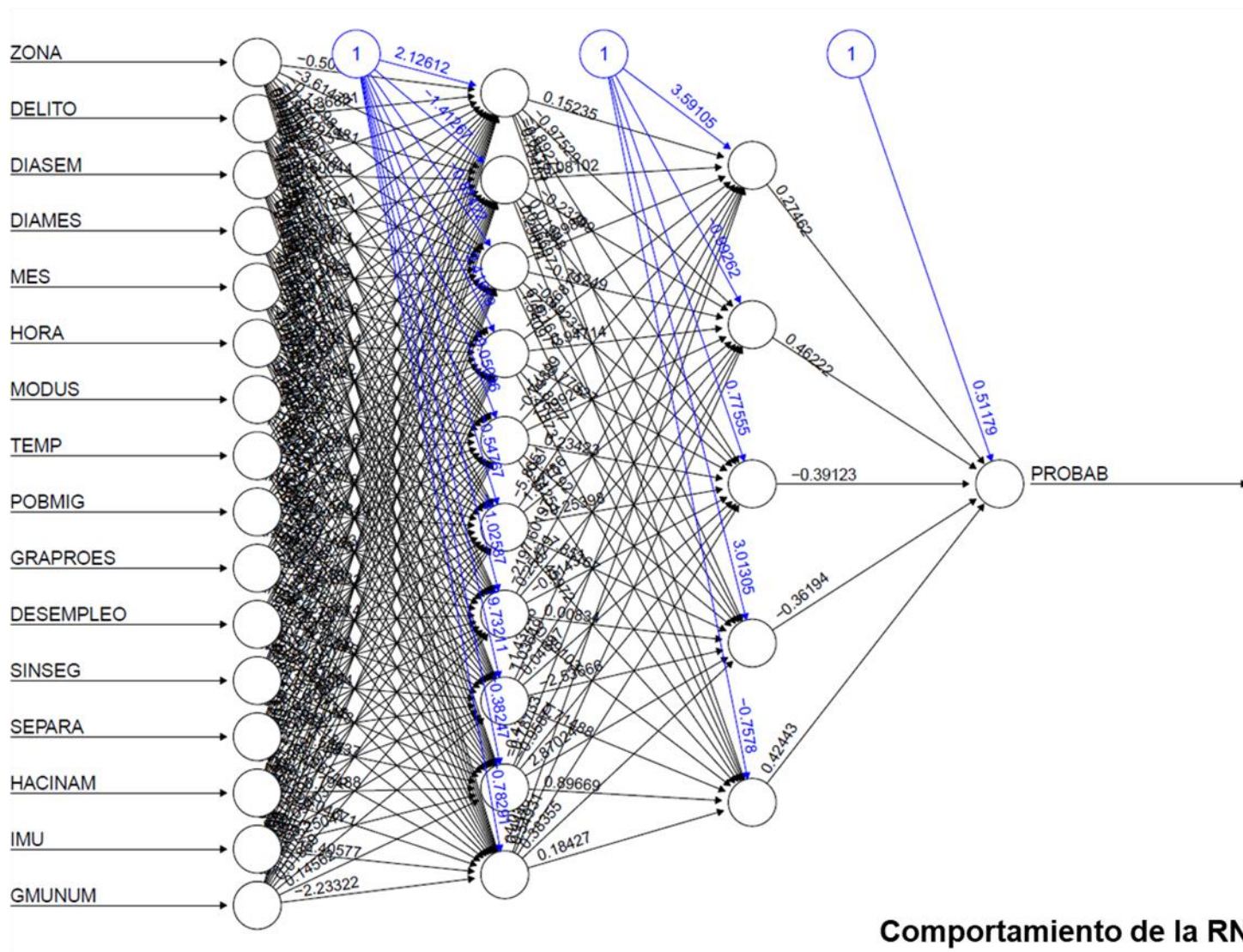
17. Configuración: 2 capas, (10n,5n), umbral 0.1.



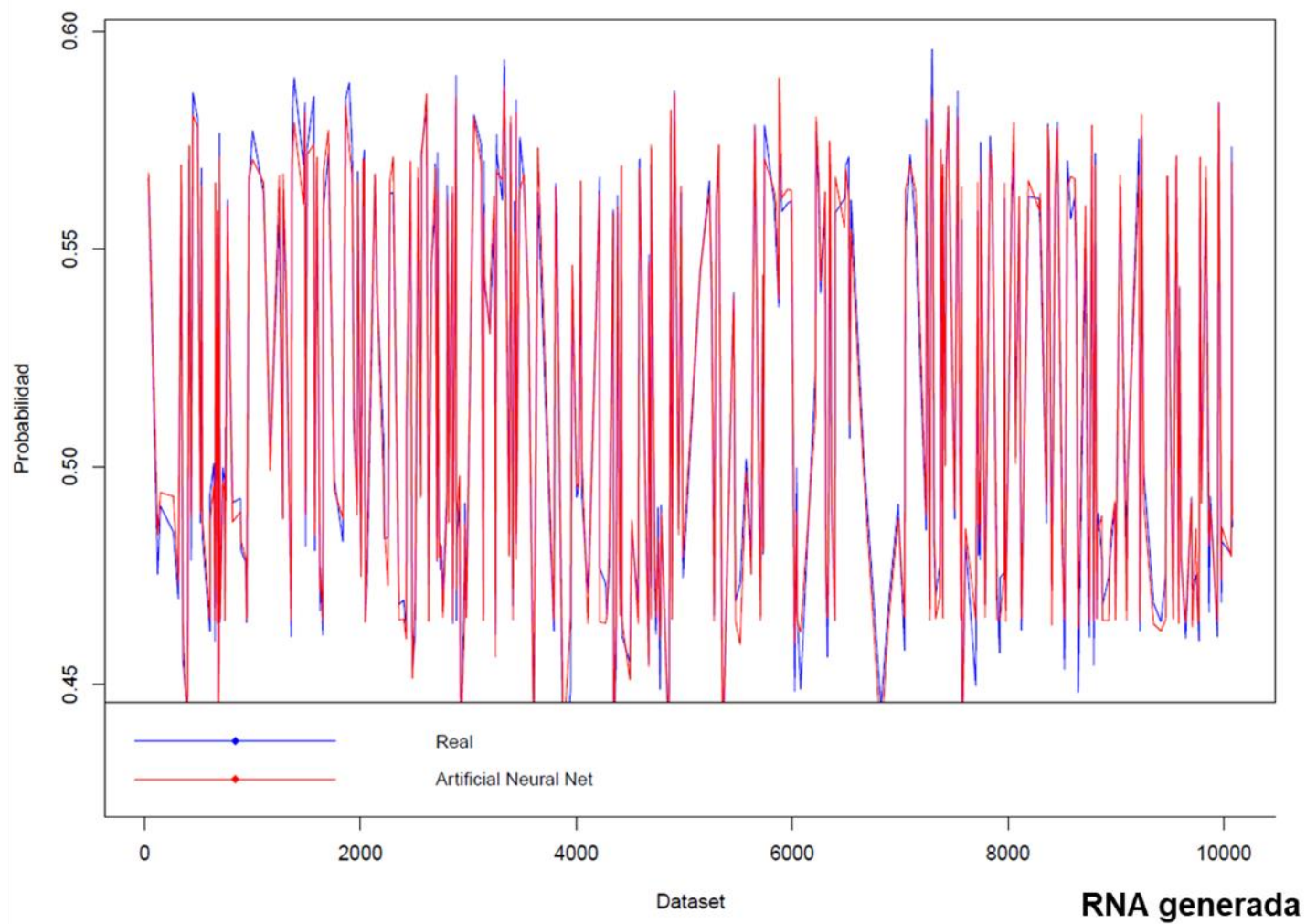
18. Configuración: 2 capas, (10n,5n), umbral 0.05.



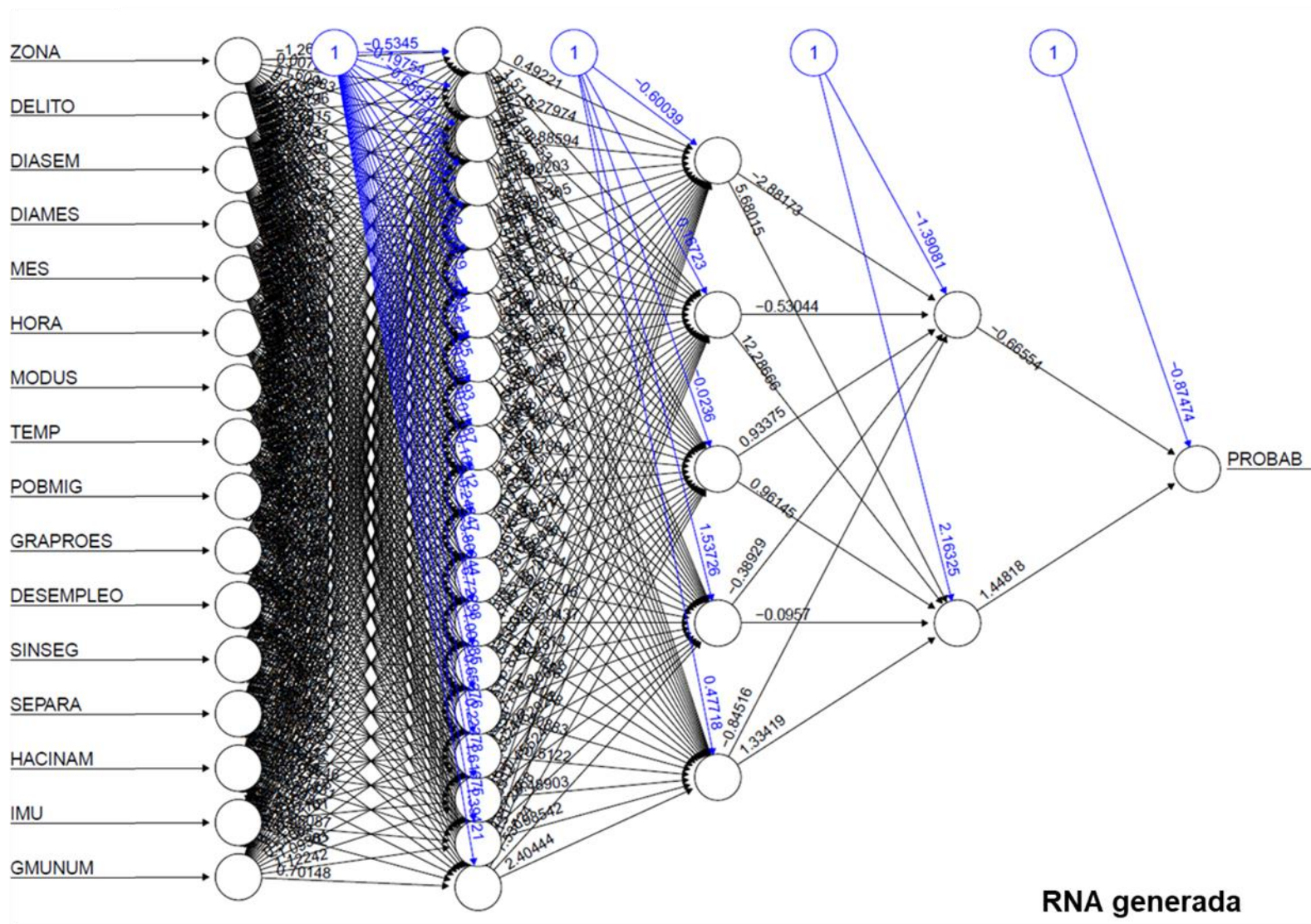
19. Configuración: 2 capas, (10n,5n), umbral 0.01.



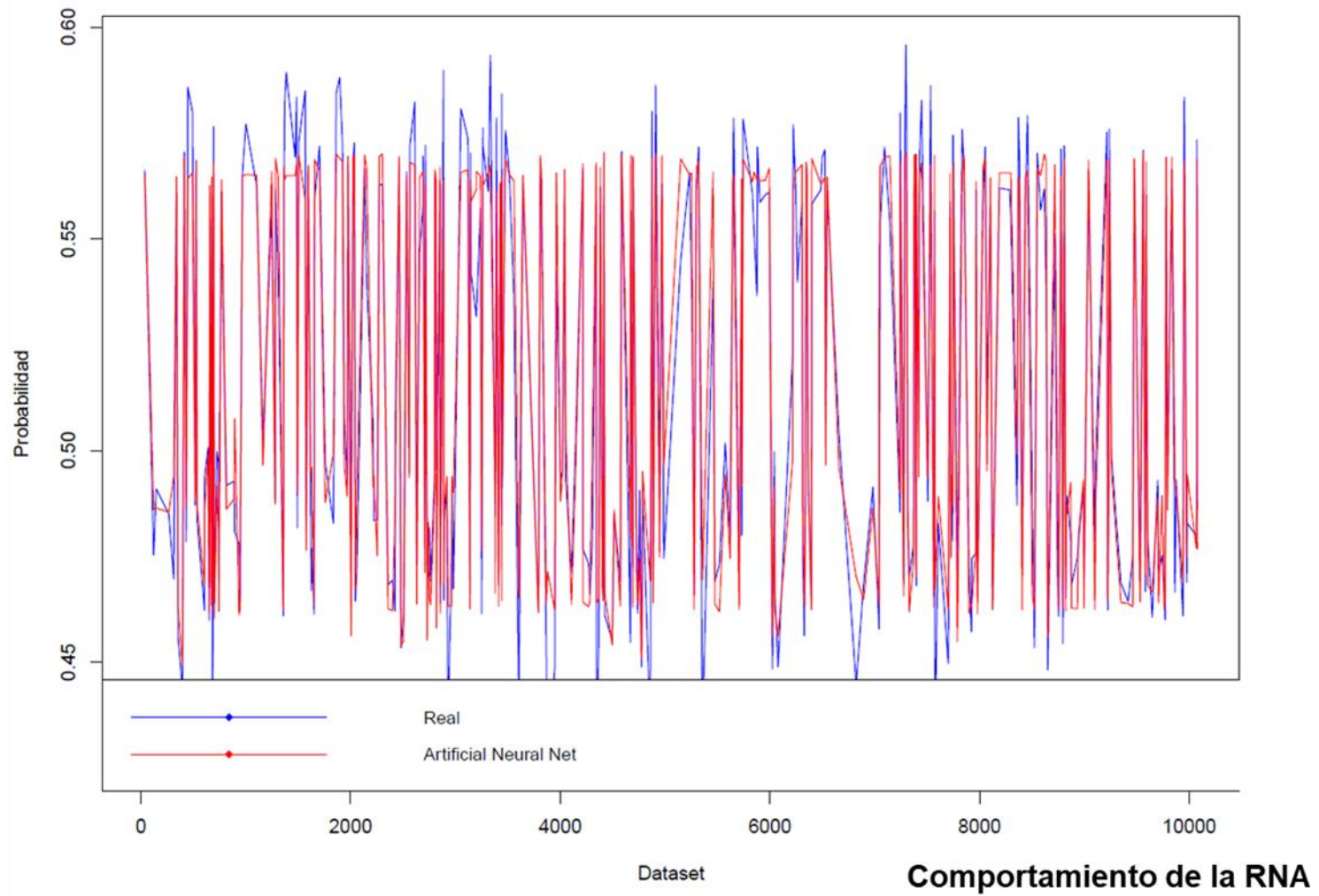
19. Configuración: 2 capas, (10n,5n), umbral 0.01.



20. Configuración: 3 capas, (20n,5n,2n), umbral 0.05.



20. Configuración: 3 capas, (20n,5n,2n), umbral 0.05.



Anexo 2: Relación de correspondencia entre Zonas y AGEB.

RELACIÓN DE CORRESPONDENCIA ZONAS/AGEB									
ZONA	AGEB	ZONA	AGEB	ZONA	AGEB	ZONA	AGEB	ZONA	AGEB
1	2402800010172	29	2402800010844	57	2402800011842	85	2402800012234	113	2402800012662
2	2402800010187	30	2402800010859	58	2402800011880	86	2402800012249	114	2402800012677
3	2402800010191	31	2402800010878	59	2402800011895	87	2402800012253	115	2402800012696
4	2402800010223	32	2402800010929	60	2402800011908	88	2402800012268	116	2402800012709
5	2402800010238	33	2402800010952	61	2402800011927	89	2402800012272	117	240280001279A
6	2402800010242	34	2402800011077	62	2402800011946	90	2402800012287	118	2402800012817
7	2402800010257	35	2402800011096	63	2402800011950	91	2402800012291	119	2402800012821
8	2402800010261	36	2402800011122	64	2402800011965	92	2402800012304	120	2402800012855
9	2402800010276	37	240280001126A	65	240280001197A	93	2402800012319	121	240280001286A
10	2402800010312	38	2402800011293	66	2402800011984	94	2402800012323	122	2402800012889
11	2402800010350	39	2402800011344	67	2402800011999	95	2402800012342	123	2402800012893
12	2402800010365	40	2402800011359	68	2402800012003	96	2402800012357	124	2402800012906
13	240280001037A	41	2402800011433	69	2402800012037	97	2402800012380	125	2402800012910
14	2402800010399	42	2402800011471	70	2402800012041	98	2402800012395	126	2402800012925
15	2402800010420	43	2402800011486	71	2402800012060	99	2402800012408	127	240280001293A
16	240280001044A	44	2402800011518	72	2402800012075	100	2402800012412	128	2402800012959
17	2402800010454	45	2402800011541	73	240280001208A	101	2402800012427	129	2402800012963
18	2402800010469	46	2402800011556	74	2402800012094	102	2402800012431	130	2402800012982
19	2402800010473	47	240280001158A	75	2402800012107	103	2402800012450	131	2402800013001
20	2402800010488	48	2402800011594	76	2402800012111	104	2402800012484	132	2402800013035
21	2402800010539	49	2402800011645	77	240280001215A	105	2402800012535	133	240280001304A
22	2402800010558	50	2402800011715	78	2402800012164	106	240280001254A	134	2402800013054
23	2402800010562	51	240280001172A	79	2402800012179	107	2402800012588	135	2402800013069
24	2402800010581	52	2402800011749	80	2402800012183	108	2402800012592	136	2402800013073
25	2402800010609	53	2402800011753	81	2402800012198	109	2402800012605	137	2402800013088
26	2402800010717	54	2402800011772	82	2402800012200	110	2402800012624	138	2402800013092
27	2402800010755	55	2402800011787	83	2402800012215	111	2402800012643	139	2402800013124
28	2402800010793	56	2402800011804	84	240280001222A	112	2402800012658	140	2402800013139

RELACIÓN DE CORRESPONDENCIA ZONAS/AGEB

ZONA	AGEB	ZONA	AGEB	ZONA	AGEB	ZONA	AGEB	ZONA	AGEB
141	2402800013143	169	2402800013603	197	2402800014298	225	2403500010552	253	2403500010995
142	2402800013158	170	2402800013622	198	2402800014334	226	2403500010567	254	240350001100A
143	2402800013177	171	2402800013637	199	2402800014372	227	2403500010571	255	2403500011014
144	2402800013196	172	2402800013641	200	2402800014419	228	2403500010590	256	2403500011029
145	2402800013228	173	2402800013656	201	2402800014461	229	2403500010603	257	2403500011033
146	2402800013232	174	2402800013660	202	2402800014480	230	2403500010622	258	2403500011048
147	2402800013247	175	2402800013675	203	2402800014550	231	2403500010637	259	2403500011052
148	2402800013251	176	2402800013694	204	2402800014601	232	2403500010641	260	2403500011067
149	2402800013266	177	2402800013707	205	2402800014781	233	2403500010660	261	2403500011071
150	240280001329A	178	2402800013711	206	240280218005A	234	2403500010675	262	2403500011086
151	2402800013302	179	2402800013726	207	240280239425A	235	2403500010694	263	2403500011090
152	2402800013336	180	240280001375A	208	2402802522766	236	2403500010707	264	2403500011103
153	2402800013355	181	2402800013798	209	2402802522770	237	2403500010779	265	2403500011118
154	240280001336A	182	2402800013800	210	2402802642732	238	2403500010783	266	2403500011122
155	2402800013374	183	240280001382A	211	2403500010247	239	2403500010800	267	2403500011137
156	240280001343A	184	2402800013834	212	2403500010321	240	2403500010815	268	2403500011141
157	2402800013459	185	2402800013849	213	2403500010340	241	2403500010853	269	2403500011156
158	2402800013463	186	2402800013853	214	2403500010355	242	2403500010868	270	2403500011160
159	2402800013482	187	2402800013868	215	2403500010425	243	2403500010872	271	240350001118A
160	2402800013497	188	2402800013887	216	240350001043A	244	2403500010887	272	2403500011226
161	240280001350A	189	2402800013923	217	2403500010444	245	2403500010891	273	2403500011283
162	2402800013514	190	2402800013942	218	2403500010459	246	2403500010904	274	2403500011315
163	2402800013533	191	2402800014090	219	2403500010463	247	2403500010919	275	2403500011334
164	2402800013548	192	2402800014141	220	2403500010478	248	2403500010923	276	2403500011368
165	2402800013552	193	2402800014156	221	2403500010482	249	2403500010942	277	2403500261211
166	2402800013567	194	2402800014160	222	240350001050A	250	2403500010957		
167	2402800013571	195	2402800014264	223	2403500010529	251	2403500010961		