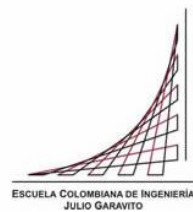


Maestría en Ingeniería Civil

**Modelo predictivo de accidentalidad para víctimas no fatales en la
concesión Devinorte**

Jairo Francisco Bravo Ramírez

Bogotá, D.C., 8 de octubre de 2020



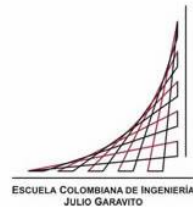
**Modelo predictivo de accidentalidad para víctimas no fatales en la
concesión Devinorte**

**Tesis para optar al título de magíster en Ingeniería Civil, con
énfasis en tránsito y transporte**

Santiago Henao Pérez

Director

Bogotá, D.C., 8 de octubre de 2020



La tesis de maestría titulada “Modelo predictivo de accidentalidad para víctimas no fatales en la concesión Devinorte”, presentada por Jairo Francisco Bravo Ramírez, cumple con los requisitos establecidos para optar al título de Magíster en Ingeniería Civil con énfasis en tránsito y transporte.

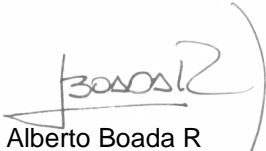
Director de la tesis


Santiago Henao Pérez

Jurado


Maritza Villamizar R.

Jurado


Alberto Boada R

Bogotá, D.C., 29 de octubre de 2020

Dedicatoria

A mamá Myriam y papá Jairo quienes con su apoyo y esfuerzo me han permitido lograr mis metas. A Dalma Sofia mi compañera de viajes por creer en mí y seguirme a todas partes; a mamá Tica, Diana, vane, tías y tíos, Jhon, Pablo, amigos y a todas aquellas personas que de una u otra forma contribuyeron para permitirme realizar este logro.

Agradecimientos

Ing. Santiago Henao Pérez

Ing. Lina López

Ing. Maritza Villamizar

Resumen

Identificando la necesidad de promover mecanismos que permitan caracterizar y plantear soluciones técnicas encaminadas a la mitigación de la siniestralidad, se busca incorporar avances tecnológicos para el análisis de datos y se observa la pertinencia de explotar modelos de predicción de siniestros mediante el uso de RNA (Redes Neuronales Artificiales). Aun cuando no han sido ampliamente masificados como alternativa para tratar de entender y buscar relaciones entre las variables que involucran la ocurrencia de un siniestro vial, el Machine Learning y las RNA ya se destacan como herramientas aplicadas en otras áreas, por ejemplo, identificación de patrones, análisis de imágenes o implementación de algoritmos predictivos con fines comerciales. Mostrando en todos los casos, su capacidad de procesar, predecir y resolver problemas que involucran un gran volumen de datos.

La aplicación de un modelo predictivo de siniestralidad mediante RNA, a través de modelos computacionales que incluyen el uso de algoritmos para el descenso del gradiente para optimizar los resultados, permiten desarrollar un modelo que involucra las variables de entrada obtenidas mediante formatos de recopilación para siniestros que involucran solo daños materiales en un periodo de operación de la Concesión DEVINORTE (2010-2016) y las salidas de datos reales obtenidos de dichos registros de siniestralidad, los cuales, posteriormente son comparados con las salidas proyectadas por el modelo de RNA, buscando obtener una aproximación cercana a las salidas reales. De modo que permite evaluar tanto la fidelidad del modelo, como identificar un potencial uso práctico en la resolución de problemas que involucran siniestralidad vial. Cabe resaltar que, si bien la aplicación de RNA equivalen a una solución específica para el conjunto de datos, se pretende verificar si es posible aplicar estos modelos, en otras concesiones o zonas de estudio dónde sea posible obtener información para diferentes líneas temporales con registros de siniestralidad vial.

Índice general

Introducción	11
Capítulo I	12
1.1. Siniestralidad del corredor	13
1.2. Análisis de puntos críticos	17
1.3. Caracterización de la Siniestralidad Con Solo Daños Materiales	20
Capítulo II	24
2.1. Antecedentes y Revisión Bibliográfica	24
2.2. Modelo de Redes Neuronales Artificiales (RNA).....	25
2.3. Construcción del modelo de RNA para siniestros que involucran solo daños	27
2.4. Aplicación de un modelo de RNA para predecir la Abscisa de ocurrencia de un siniestro con solo daños.....	29
2.5. Resultado de la aplicación de Modelos de RNA en la Concesión	34
2.6. Análisis de resultados para un modelo de RNA en la Concesión.....	36
Capítulo III	39
Capítulo IV.....	44
Referencias	46

Índice de tablas

Tabla 1 Variables registradas por Devinorte para siniestros ocurridos en la Concesión...	13
Tabla 2 Variables registradas por Devinorte para siniestros ocurridos en la Concesión...	14
Tabla 3 Variación de los modelos de RNA para la vía 45A04-Devinorte	29
Tabla 4 Variación de los modelos de RNA para la vía 45A04-Devinorte	34
Tabla 5 Variación de los modelos de RNA para la vía 45A04 Primeros 10 PR-Devinorte	34
Tabla 6 Variación de los modelos de RNA para la vía 5501 -Devinorte	35
Tabla 7 Variables incluidas en el modelo predictivo para determinar la clase de siniestro con implementación de RNA para el PR 19 de la vía 5501 -Devinorte	37
Tabla 8 Variables incluidas en el modelo predictivo para determinar la clase de siniestro con implementación de RNA para el PR 19 de la vía 5501 -Devinorte	37
Tabla 9 Puntos críticos por índice de Peligrosidad vía 45A04 -Devinorte.....	39
Tabla 10 Puntos críticos por índice de Peligrosidad vía 5501 -Devinorte	39

Índice de figuras

Figura 1 Desarrollo Vial del Norte de Bogotá-DEVINORTE	12
Figura 2 Salida de Bogotá DEVINORTE (SN/NS/KR7)-DEVINORTE	12
Figura 3 Vía Chía-Cajicá-DEVINORTE	13
Figura 4 Siniestralidad General Concesión Devinorte 2010-2016	15
Figura 5 Siniestralidad con Relación Morbilidad Total- Concesión Devinorte 2010-2016 .	16
Figura 6 TPD Anual Peaje Fusca vs Afectados Totales en Siniestros Viales Salida por Avenida Carrera 7-Concesión Devinorte 2010-2016	16
Figura 7 TPD Anual Salida Peaje Andes vs Afectados Totales en Siniestros Viales Salida por Autopista Norte-Concesión Devinorte 2010-2016	16
Figura 8 Siniestralidad por intervalo de Abscisa de ocurrencia-Concesión Devinorte 2010-2016	17
Figura 9 Abscisas-Concesión Devinorte	18
Figura 10 índice de peligrosidad por Abscisa de ocurrencia-Concesión Devinorte 2010-2016	19
Figura 11 Concentración de siniestros por índice de peligrosidad -Concesión Devinorte 2010-2016	19
Figura 12 Distribución por Clase de siniestros Salida por Carrera 7-Concesión Devinorte 2010-2016	21
Figura 13 Distribución por Clase de siniestros Salida por Autopista Norte-Concesión Devinorte 2010-2016	21
Figura 14 Causa posible en Siniestros Viales con solo daños Salida por Carrera 7-Concesión Devinorte 2010-2016.....	22
Figura 15 Causa posible en Siniestros Viales Salida por Autopista Norte-Concesión Devinorte 2010-2016	22
Figura 16 Interacción de siniestros por solo daños, con tipo de vehículos involucrados-Concesión Devinorte 2010-2016.....	23
Figura 17 Modelo de redes Neuronales Artificiales	26
Figura 18 Funcionamiento del modelo de RNA.....	26
Figura 19 Modelo general de RNA Para siniestros en la concesión Devinorte.....	28
Figura 20 Datos de entrada para modelo de RNA para predecir abscisa de ocurrencia de un siniestro	29
Figura 21 Distribución de los Datos de entrada para modelo de RNA para predicción de abscisas.....	30
Figura 22 Arquitectura del Modelo de RNA con 10 capas ocultas para predicción de abscisas en siniestros con solo daños	30
Figura 23 Resultados de un Modelo de RNA con 10 capas ocultas para predicción de abscisas en siniestros con solo daños	¡Error! Marcador no definido.
Figura 24 Resultados gráficos de un Modelo de RNA con 10 capas ocultas para predicción de abscisas en siniestros con solo daños	31
Figura 25 Arquitectura del Modelo de RNA con 15 y 20 capas ocultas para predicción de abscisas en siniestros con solo daños	32

Figura 26 Procedimiento de aprendizaje para predicción de abscisas en siniestros con solo daños.....	32
Figura 27 Resultados gráficos de un Modelo de RNA con 10 capas ocultas para predicción de abscisas en siniestros con solo daños	33
Figura 28 Resultados gráficos de un Modelo de RNA con 20 capas ocultas para predicción de la clase de accidente en siniestros con solo daños para la vía 5501 en la concesión Devinorte	36
Figura 29 Distribución horaria en siniestros con solo daños para el PR 19 de la vía 5501 en la concesión Devinorte.....	36
Figura 30 Ubicación PR 19 de la vía 5501 en la concesión Devinorte	38
Figura 31 Principales Puntos con mayor peligrosidad Concesión Devinorte. 2010-2016 .	40
Figura 32 Obras de mejoramiento Puente La Caro y ampliación Concesión Accenorte...	40
Figura 33 Obras de mejoramiento entre PR 9 y PR 10 Sector Fontanar-Devinorte.....	41
Figura 34 recomendaciones pacificación entre PR 11 y PR 12 Sector LA Caro Sopó-Devinorte	42
Figura 35 recomendaciones pacificación entre PR 14 y PR 15 Sector Hatogrande-Devinorte	43
Figura 36 PR 19 y PR 20 Sector Briceño-Sopó-Devinorte	43

Índice de anexos

- Anexo 1. Base de datos de siniestralidad Devinorte
- Anexo 2. Procesamiento de tablas y bases de datos
- Anexo 3. Archivos y procesamiento en Matlab
- Anexo 4. Registro fotográfico
- Anexo 5. Mapas y Archivos GIS

Introducción

En el año 2019, más de 6000 personas perdieron la vida como consecuencia de siniestros de tránsito (ONSV, 2020), según la OMS (Organización Mundial de la Salud), la problemática relacionada con los siniestros de tránsito se considera un tema de salud pública. Las medidas encaminadas al análisis, mitigación o supresión de las condiciones potenciales que desencadenan la aparición de siniestros, deben ser estudiadas, ajustadas y puestas en marcha, toda vez que su aplicación puede ser traducida en vidas salvadas.

El Desarrollo Vial del Norte, Devinorte, fue una concesión encargada del mejoramiento, operación y mantenimiento en los accesos por el norte de Bogotá, en la autopista norte por la Calle 245 y por la Carrera Séptima. Conectaba la Ciudad de Bogotá, con municipios como Chía, Cajicá, Zipaquirá, Sopó entre otros. Tenía una longitud aproximada de 53 Km, incluía el peaje Andes (Salida por Autopista Norte) y estuvo operativa desde 1994 hasta la finalización del contrato en el año 2017.

Como parte de la evaluación, se obtiene registros proporcionados por la Concesión, para información correspondiente del 2010 al 2016, sin embargo, teniendo en cuenta que se presenta información incongruente o incompleta en los dos primeros años de los registros, el análisis y modelos se realizan con la información de los años 2012 a 2016.

El presente documento plantea la utilización de modelos de Redes Neuronales Artificiales (RNA) buscando generar un modelo predictivo que sea desarrollado con las variables de entrada obtenidas de los registros de siniestros recolectados por la Concesión entre los años 2012 a 2016 y la variable de salida, la cual será comparada con los valores obtenidos de los registros de siniestralidad para evaluar la efectividad del modelo y posteriormente si se encuentran correlaciones significativas, confirmar si es posible realizar una predicción con un modelo para siniestros que presentan solo daños materiales.

Capítulo I

CARACTERIZACIÓN DEL CORREDOR

La concesión DEVINORTE, fue un proyecto que se encargaba del mantenimiento y operación de la conexión de los accesos hacia el norte de Bogotá, en las vías nacionales Ruta 45A-04 y Ruta 5501, cuya operación estuvo activa desde 1994 hasta 2017. En su reemplazo la sucedió la actual concesión ACCENORTE que incluye entre otras, las obras de ampliación de 3 a 5 carriles para el acceso por la Autopista Norte y la doble calzada para el acceso por la Carrera 7. El corredor se encuentra en el departamento de Cundinamarca, funciona como eje de conectividad de la capital, con el norte del país tal como se muestra en la siguiente imagen:

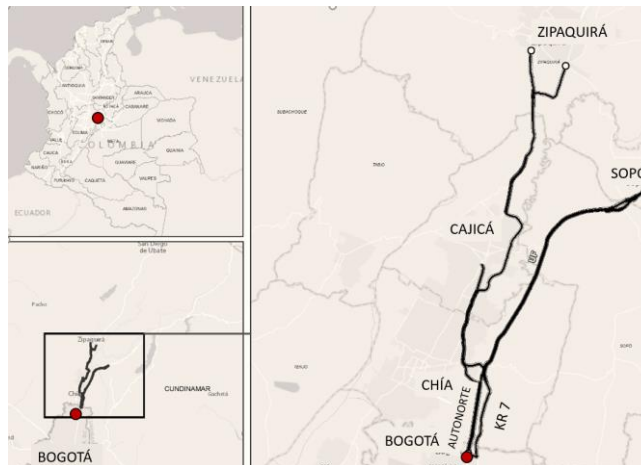


Figura 1 Desarrollo Vial del Norte de Bogotá-DEVINORTE

Fuente: Elaboración propia con base en Cuarto de datos ANI

Por ser una concesión con alto potencial de recaudo, se caracteriza por tener buenas condiciones de circulación, la mayoría de los tramos cuentan con bermas amplias, buena señalización horizontal y vertical; de igual forma sus parámetros geométricos y de operación como niveles de servicio y estado de la carpeta asfáltica.



Figura 2 Salida de Bogotá DEVINORTE (SN/NS/KR7)-DEVINORTE

Fuente: Elaboración propia con imagen de Google Earth



Figura 3 Vía Chía-Cajicá-DEVINORTE

Fuente: Elaboración propia

En general, la topografía del corredor se desarrolla en terreno plano y ondulado, cuenta en su mayoría con tramos unidireccionales con vías de 2 a 3 carriles por sentido y anchos de carril promedio de 3.6m con bermas entre 1.8m a 2 m. Salvo en la salida por la Carrera 7, que corresponde a una vía bidireccional con carriles de aproximadamente 3.6m y berma interrumpida o sin presencia de ella.

1.1. Siniestralidad del corredor

Con referencia a la siniestralidad histórica del corredor, se recolecta a través de la información obtenida de los reportes y levantamientos de siniestros registrados por la concesión dentro del corredor, los cuales eran diligenciados por personal del convenio entre la concesión y la policía de carreteras. Esta información se tomaba de forma manual en formatos propios, o a través de formatos para Informes Policiales de Accidentes de Tránsito (IPAT) y posteriormente eran digitalizados con la información relevante que varía en los primeros años de datos, sin embargo, en el año 2016 se incluyen 21 variables para cada registro, siendo determinado como el formato papara estandarizar las variables de modo que puedan ser comparadas en la temporalidad de los datos y listadas a continuación:

Tabla 1 Variables registradas por Devinorte para siniestros ocurridos en la Concesión.

VARIABLE IPAT	NOMBRE	VARIABLE IPAT	NOMBRE
1	FECHA	12	VICTIMAS (TOTAL AFECTADOS)
2	DIA	13	HERIDOS LEVES
3	HORA	14	HERIDOS GRAVES
4	CONDICIONES METEOROLOGICAS	15	MUERTOS
5	LOCALIZACION	16	CONDUCTOR VEHÍC. AUTOMOTOR MUERTO

VARIABLE IPAT	NOMBRE	VARIABLE IPAT	NOMBRE
6	ABSCISADO	17	PLACA VEHÍCULO CONDUCTOR FALLECIDO
7	CALZADA	18	INSPECTOR
8	GEOMETRIA	19	AGENTE QUE CONOCIÓ EL CASO
9	CLASE ACCIDENTE	20	CROQUIS
10	CAUSAS POSIBLES	21	DESESTIMIENTO
11	ELEMENTOS INVOLUCRADOS		

Fuente: Elaboración propia con datos de Devinorte

Teniendo en cuenta que la información recopilada varía entre los diferentes años, se homogeneiza con base en el último año de registros. Teniendo en cuenta que el modelo será realizado únicamente con siniestros que no reportan heridos de gravedad o muertes, se desestima aquellas variables que se relacionan con víctimas (variables 14,15,16,17,18,19,20 y 21) y, además se hace los siguientes tratamientos:

- Se elimina la precisión de las variables temporales con el fin de hacer anónima la base definitiva.
- Se elimina la variable 5 teniendo en cuenta la imprecisión y discrepancia en la ubicación del siniestro, tomando como referencia única de posición, la abscisa de ocurrencia del siniestro y el tramo de la vía.
- Para las variables 9,10 y 11 se toma únicamente los dos primeros valores del reporte, teniendo en cuenta la homogeneización de las variables y con el fin de evitar ruido por falta de datos en un gran porcentaje de los casos.

Finalmente, se depura y elimina los datos incompletos e incongruentes para realizar el diagnóstico del corredor, generando el siguiente listado de variables para su caracterización:

Tabla 2 Variables registradas por Devinorte para siniestros ocurridos en la Concesión.

NOMBRE VAR	DESCRIPCION	TIPO	LIM INF	LIM SUP
MES	MES [ENERO-DICIEMBRE]	CONTINUA	1	12
YEAR	AÑO DE OCURRENCIA	CONTINUA	2012	2016
DIA	DÍA DE OCURRENCIA [LUNES-DOMINGO]	DISCRETA	1	7
HNUM	HORA "HH" EN FORMATO MILITAR	DISCRETA	0	24
TPD	TRANSITO PROMEDIO DIARIO ESTACIONES DE PEAJE	DISCRETA	N/A	N/A
D_N	DIA/NOCHE	DISCRETA	1	2
COND_MET	CONDICIÓN METEREOLÓGICA [NORMAL-LLUVIA-NIEBLA]	DISCRETA	1	3
VIA	CODIGO DE LA VÍA	DISCRETA	45A04	5501
KM	ABSCISADO DÓNDE OCURRIÓ EL SINIESTRO	CONTINUA	0	46600
CALS	CALZADA [ESTE-OESTE- UNICA]	DISCRETA	1	3
GEOMETRIA	GEOMETRIA DE LA VÍA [RECTA-CURVA]	DISCRETA	1	5
CLASEAC1	CLASE DE ACCIDENTE PRIMER VALOR	DISCRETA	1	121
CLASEAC2	CLASE DE ACCIDENTE SEGUNDO VALOR	DISCRETA	2	132

CAUP1	CAUSA POSIBLE DEL SINIESTRO PRIMER VALOR	DISCRETA	1	503
CAUP2	CAUSA POSIBLE DEL SINIESTRO SEGUNDO VALOR	DISCRETA	0	999
ELEMENTO1	ELEMENTO IMPACTADO	DISCRETA	1	126
ELEMENTO2	SEGUNDO ELEMENTO IMPACTADO	DISCRETA	0	999
VICTIMAS	NÚMERO DE AFECTADOS QUE GENERA EL SINIESTRO	DISCRETA	0	20

Fuente: Elaboración propia con datos de Devinorte

Cabe resaltar que la vía Nacional 5501, hace referencia a la salida por la Carrera 7, dónde se ubica la estación de peaje Fusca; por su parte, la vía nacional 45A04, corresponde a la salida por la Autopista Norte Incluyendo el Peaje Andes.

Una vez realizado el procesamiento de la información, se observa que la siniestralidad en concesión presenta un comportamiento relativamente creciente respecto al número de casos ocurridos, tal como se muestra a continuación:

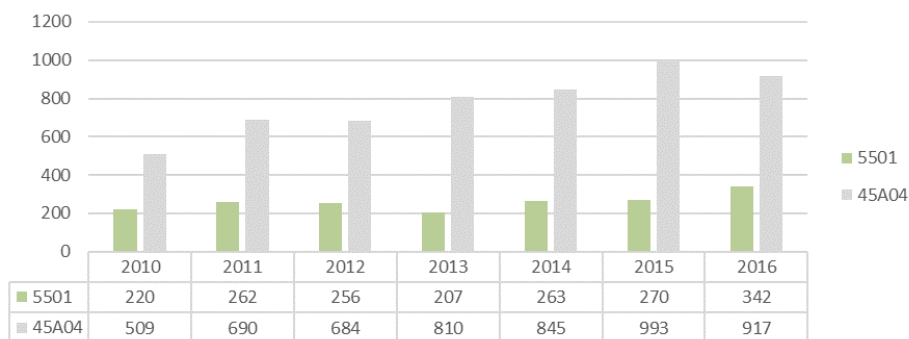


Figura 4 Siniestralidad General Concesión Devinorte 2010-2016

Fuente: Elaboración propia

Se observa que el año 2016, se presenta el mayor número de siniestros tanto para el acceso por Carrera 7, como para la salida por la Autopista Norte. De igual forma se observa que la ruta Nacional 45A04, presenta entre 3 a 4 veces el número de siniestros reportados.

Referente a la morbilidad de la concesión para el periodo de estudio, se presenta un promedio de 379 siniestros con solo daños, 55 siniestros con Víctimas no fatales y 19 víctimas fatales. Se resalta que el año 2016 fue particularmente crítico por el aumento de víctimas fatales.

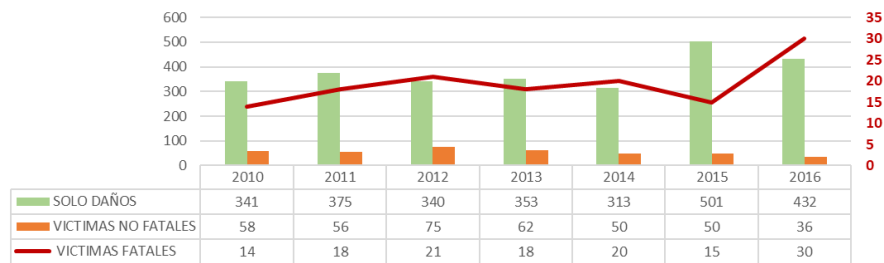


Figura 5 Siniestralidad con Relación Morbilidad Total- Concesión Devinorte 2010-2016

Fuente: Elaboración propia

Con respecto a la morbilidad de la concesión para el periodo de estudio, se presenta un promedio de 379 siniestros con solo daños, 55 siniestros con Víctimas no fatales y 19 víctimas fatales. Se resalta que el año 2016 fue particularmente crítico por el aumento de víctimas fatales; sin embargo, los siniestros con solo daños, así como las víctimas no fatales presentan un descenso.

Realizando una comparación entre el TPD obtenido del resumen anual de las estaciones de peaje (ANI, 2020); para los peajes Fusca y Andes , de igual forma las víctimas registradas en la concesión; se observa que el aumento en el TPD no necesariamente contempla un aumento en la siniestralidad, sino que intervienen otras variables que pueden ayudar a diagnosticar más asertivamente este fenómeno. De igual forma se encuentra que esto aplica tanto para los accesos por la Carrera 7, como por la autopista norte, A continuación, se presenta las gráficas que permiten observar la visualización de la información:



Figura 6 TPD Anual Peaje Fusca vs Afectados Totales en Siniestros Viales Salida por Avenida Carrera 7-Concesión Devinorte 2010-2016

Fuente: Elaboración propia



Figura 7 TPD Anual Salida Peaje Andes vs Afectados Totales en Siniestros Viales Salida por Autopista Norte-Concesión Devinorte 2010-2016

Fuente: Elaboración propia

1.2. Análisis de puntos críticos

Para analizar la ubicación de puntos críticos, se pretende realizar un análisis espacial por medio del procesamiento de las bases de datos proporcionadas; sin embargo, dado que la información recopilada por la concesión no se encuentra georreferenciada, se dificulta la implementación extensiva de herramientas GIS.

Con el fin de tener un mayor detalle en las zonas de concentración de siniestros presentados en la concesión, se busca realizar la identificación de puntos críticos apoyados en la elaboración de indicadores de peligrosidad y complementos con el software QGIS versión 3.10.2.

De manera inicial, se caracteriza el corredor únicamente por cantidad de eventos según la abscisa de ocurrencia tal como se presenta a continuación:

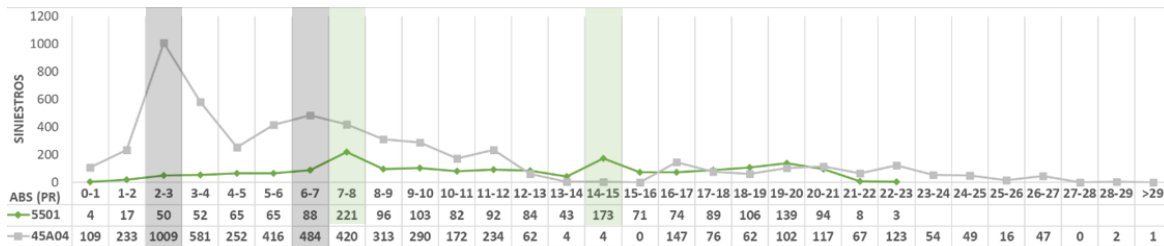


Figura 8 Siniestralidad por intervalo de Abscisa de ocurrencia-Concesión Devinorte 2010-2016

Fuente: Elaboración propia

Con el fin de identificar de mejor manera la ubicación de los siniestros, se presenta el mapa de la concesión con los Postes de Referencia (INVIAS, 2020).

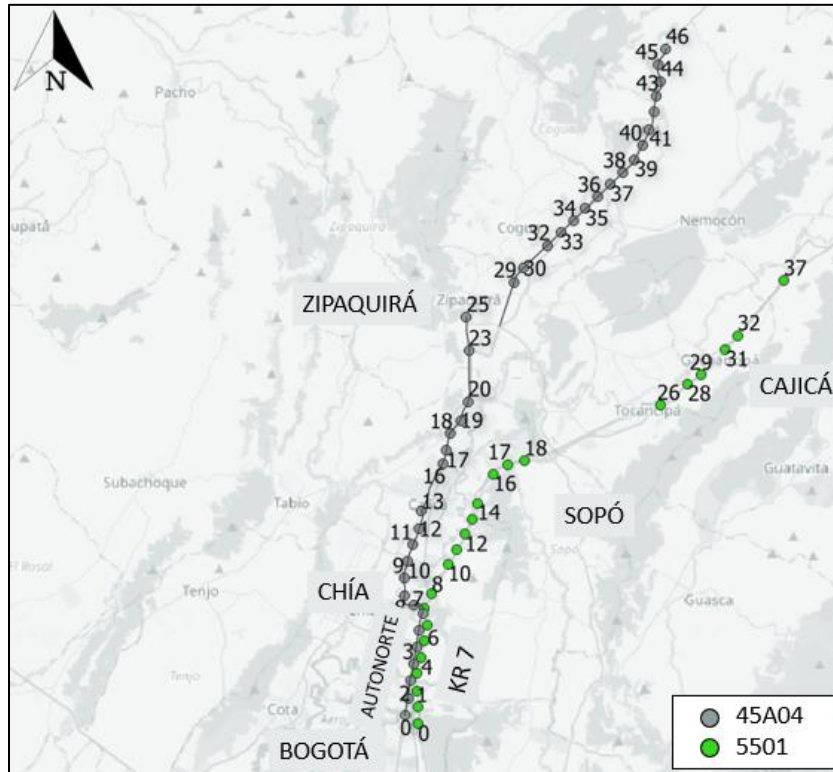


Figura 9 Abscisas-Concesión Devinoorte
Fuente: Elaboración propia con base en QGIS 3.10.2.

Para la definición del indicador de peligrosidad se adopta una equivalencia de siniestros con solo daños, de ampliación con factor de 2 veces para siniestros con heridos y 13,4 veces para siniestros que presentan víctimas fatales (ANSV, 2020). Posteriormente se relacionan con el TPD de las estaciones Andes y Fusca según corresponda y se adopta un valor único para longitud de 1000 metros a fin de estandarizar la medición en todos los PR.

:

$$IP = \frac{\#siniestros(SD + 2 * NF + 13.4 * F)}{TPD * 365 * L}$$

- SD= Número de siniestros con solo daños materiales
- NF= Número de siniestros con víctimas no fatales
- F= Número de Víctimas fatales
- TPD= Tránsito promedio diario (Estaciones Andes y Fusca)
- L= Longitud del tramo (homogeneizado a 1000m)

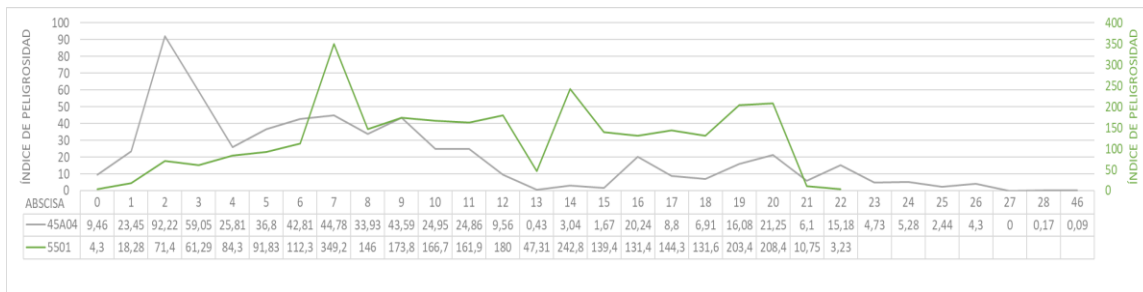


Figura 10 índice de peligrosidad por Abcisa de ocurrencia-Concesión Devinorte 2010-2016
Fuente: Elaboración propia

Con base en el indicador elaborado, se procede a incluir los datos de los 10 sectores más críticos para cada tramo de vía ya identificados, para su ubicación espacial y posterior análisis.

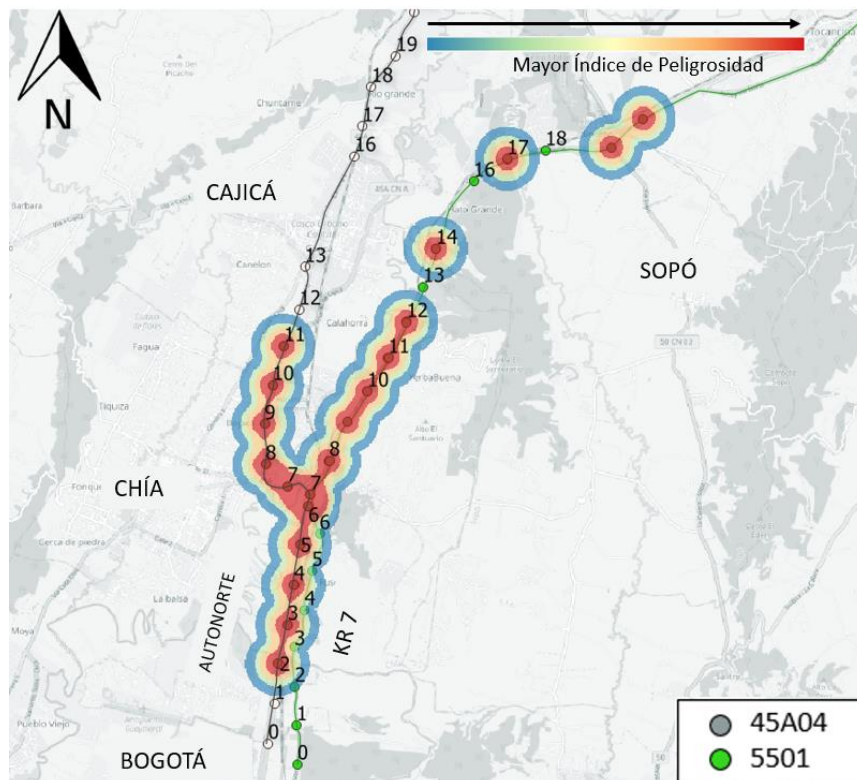


Figura 11 Concentración de siniestros por índice de peligrosidad -Concesión Devinorte 2010-2016
Fuente: Elaboración propia con QGIS 3.10.2

Las herramientas de GIS facilitan el análisis estadístico que permite encontrar, por ejemplo, la concentración de puntos por medio de un ráster, generado a través de un análisis por vecino más cercano (Silverman, 1986). De esta forma, se identifica el nudo de conexión en el sector denominado Puente de la Caro, como el sitio más crítico por peligrosidad. Particularmente se resalta su proximidad con sitios de

atracción y generación de viajes, tales como zonas escolares de colegios y universidades, así como el intercambiador vial que conecta Chía con Bogotá y la Vía Bogotá Tunja.

Con referencia a otros puntos críticos por ocurrencia, se observa que la tendencia presenta picos, que para el caso de la salida por la Carrera 7 (5501), se concentran en dos puntos, ubicados en entre el PR 7+000-PR 8+000 y PR 14+000-PR15+000. Para la Salida por autopista Norte (45A04); se observa un comportamiento bimodal entre los PR2+000 a PR 3+000 y un segundo pico entre el PR6+000 y PR 7+000. De igual forma, se observa una concentración en las primeras abscisas y una disminución en el número de casos a medida que se aleja de los centros poblados, que puede estar directamente relacionada con el número de viajes realizados entre Bogotá-Chía.

Si bien el análisis de datos apoyado en herramientas GIS no presenta mayor relevancia para la elaboración de los modelos, debido a no es un insumo que interfiera en la aplicación de RNA, si permite observar que la concentración de siniestros en la vía 45A04, corresponde principalmente a zonas predominantemente urbanas, lo cual toma sentido si se tiene en cuenta la cantidad de conflictos por ingreso y salida de vehículos, actividades comerciales y presencia recurrente de zonas de atracción y generación de viajes, como centros comerciales o zonas escolares.

1.3. Caracterización de la Siniestralidad Con Solo Daños Materiales

Una vez realizado el diagnóstico general del corredor, se presenta con mayor detalle, la descripción correspondiente a los siniestros que involucran únicamente daños materiales, en otras palabras, por el alcance del presente documento, se desprecia aquellos siniestros en los que se contempla víctimas fatales o no fatales.

Para el periodo comprendido entre los años 2012 a 2016, se tienen 1661 casos reportados para la vía 5501 y 5231 para la 45A04. Los siniestros sin víctimas graves o heridos representan aproximadamente el 95% de los datos registrados por la Concesión Devinorte en el periodo de estudio.

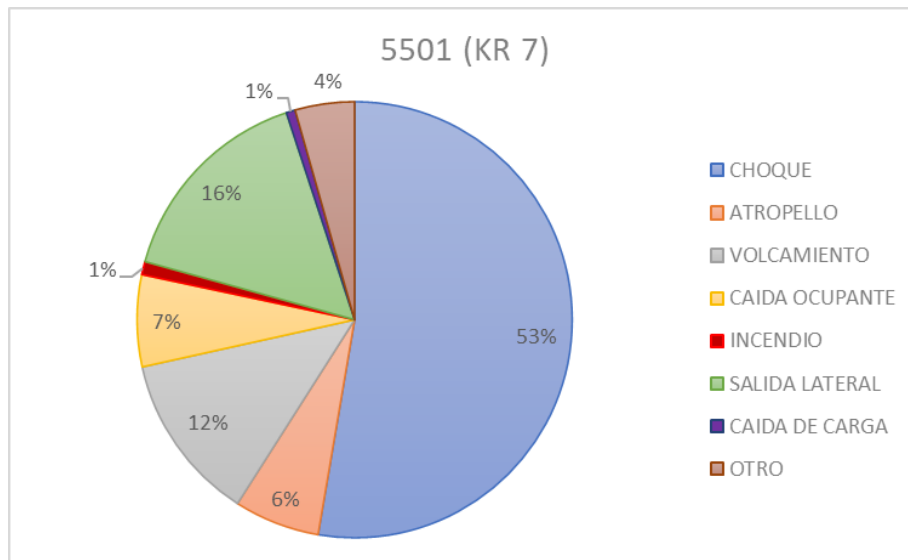


Figura 12 Distribución por Clase de siniestros Salida por Carrera 7-Concesión Devinorte 2010-2016

Fuente: Elaboración propia

La distribución porcentual de siniestros en la Vía 5501, denota los choques como la principal clase por ocurrencia, con 705 casos, seguido por las salidas laterales con 209 casos y los volcamientos con 167 casos ocurridos entre el 2010 y 2016.

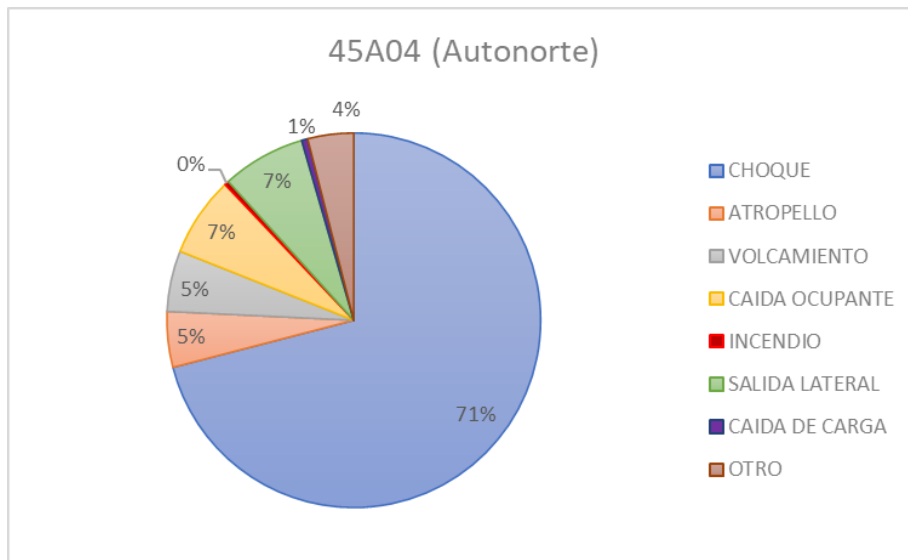


Figura 13 Distribución por Clase de siniestros Salida por Autopista Norte-Concesión Devinorte 2010-2016

Fuente: Elaboración propia

Para la vía 45A04, se tiene como evento más recurrente, aquellos que involucran choques con 3015 casos, seguido por la salida lateral con 307 casos y 223 casos en los que se presenta caída de ocupante.

Como parte de la caracterización de las hipótesis iniciales de los siniestros que involucran pérdidas materiales, se presenta las siguientes imágenes:

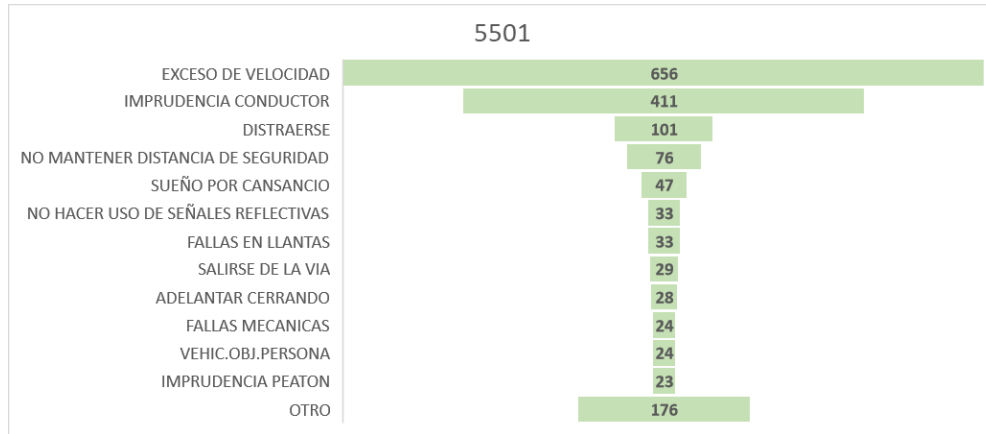


Figura 14 Causa posible en Siniestros Viales con solo daños Salida por Carrera 7-Concesión Devinorte 2010-2016

Fuente: Elaboración propia

Para la salida por la Carrera 7, cerca del 40% de los siniestros con solo daños, corresponden a un presunto exceso de velocidad, aproximadamente 25% a imprudencias del conductor y 6% a distracciones mientras conduce.

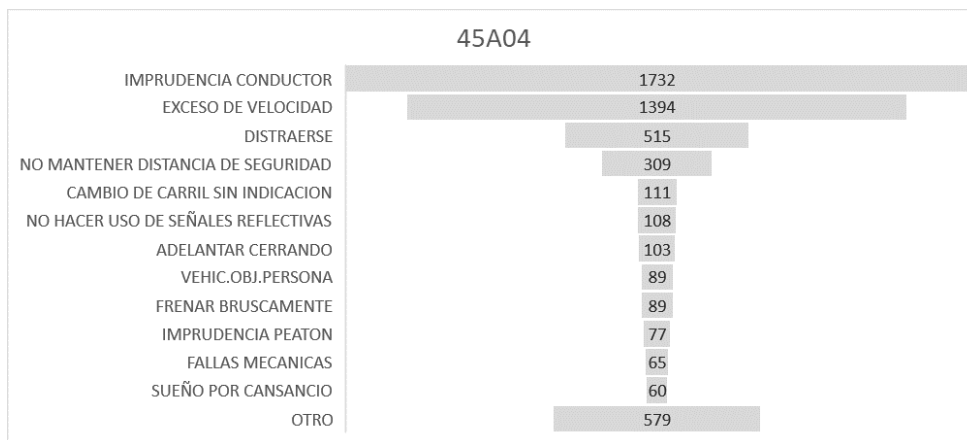


Figura 15 Causa posible en Siniestros Viales Salida por Autopista Norte-Concesión Devinorte 2010-2016

Fuente: Elaboración propia

En los siniestros con solo daños materiales que se presentaron en la concesión por la salida con la Autopista Norte, el 33% se clasifican como imprudencia del conductor, cerca del 27% por exceso de velocidad y 10% corresponden a distracción del conductor.

Con referencia a los vehículos involucrados en siniestros, se obtiene que el 32% de los casos no tienen información del vehículo u objeto al que impactaron; sin

embargo, en el 53% de los casos se puede extraer los resultados para la caracterización de interacciones entre vehículos siniestrados más relevantes, las cuales son presentadas a continuación:


















									
	12,01%	1,79%	5,41%	9,13%	17,39%	1,14%	2,50%	2,45%	12,80%
		0,14%	0,33%	0,46%	0,95%	0,05%	0,05%	0,14%	0,49%
			0,84%	1,47%	2,39%	0,11%	0,27%	0,14%	1,74%
				0,82%	3,86%	0,19%	0,95%	0,57%	2,09%
				0,03%	4,84%	0,46%	1,17%	0,65%	7,42%
					0,03%		0,03%	0,03%	0,43%
				0,03%				0,16%	0,52%
								0,05%	0,35%
									1,14%

Figura 16 Interacción de siniestros por solo daños, con tipo de vehículos involucrados- Concesión Devinorte 2010-2016
Fuente: Elaboración propia

Los anteriores datos se recopilan como vehículo primario y secundario involucrados en siniestro con solo daños. Se resalta que la interacción más frecuente se produce entre automóvil y camioneta (640 casos), seguido de automóvil-motocicleta (471 casos) y automóvil-automóvil (442 casos), lo cual permite identificar los 3 usuarios críticos para encaminar medidas que pueden ayudar a mitigar la ocurrencia de siniestros.

Capítulo II

Modelo de simulación

2.1. Antecedentes y Revisión Bibliográfica

En la recopilación de diversos documentos cuyo fin está relacionado con la implementación de Redes Neuronales Artificiales, se destaca que unos de los primeros análisis encontrados datan del 2008, cuando ya se empiezan a realizar modelos para estudiar la siniestralidad en vías rurales, cuyos resultados se perciben como exitosos teniendo en cuenta la capacidad de adaptación a series de tiempo con comportamientos no lineales (Ozgan, 2008); también sobresalen por ejemplo, aquellos que involucran estudios con países en vías de Desarrollo como Nigeria (Ogwueleka, Misra, Ogwueleka, & Fernandez-Sanz, 2014) resaltando la facilidad con la que se puede solventar situaciones difíciles de percibir con los análisis convencionales. Por otra parte dentro de la investigación que se desarrolla en os diferentes actores involucrados en la siniestralidad, se tiene casos ampliamente documentados como el aprendizaje autónomo para vehículos que no necesitan conductor, a través del reconocimiento de imágenes y aplicación de algoritmos para Deep Learning (Tian, Jana, Pei, & Ray, 2020), como herramientas de entrenamiento para mejorar la respuesta del vehículo autónomo que se perciben como uno de los grandes retos tanto de fabricantes de vehículos, como la adaptación de la normativa, dado que por una parte pueden ayudar a prevenir la ocurrencia de siniestros al evitar condiciones recurrentes como el cansancio por exceso de horas durante la conducción, pero en caso de ocurrencia de siniestros con víctimas, dificulta cuando menos, la reparación de víctimas en siniestros.

En el caso de Colombia, se encuentran, por ejemplo, documentos que relacionan modelos de seguridad vial para intersecciones con semáforos, realizando matrices de correlación entre las diferentes variables recolectadas para un evento de siniestralidad y la ampliación de modelos ya establecidos como el HSM (Manual Americano de Seguridad Vial) (Restrepo, 2018). Por otra parte, se encuentra el análisis de patrones y tendencias realizado por medio de modelos de gravedad (Urbano, Sánchez, & Quintero, 2018). Sin embargo, no se precisa información relacionada con análisis de siniestralidad mediante el uso de redes neuronales artificiales para zonas urbanas o rurales, en parte porque la recolección de datos concernientes a la siniestralidad son relativamente escasos salvo en las ciudades principales, o porque la información es limitada en cuanto al número de variables que se pueden acceder desde el ámbito académico. La razón puede estar ligada a que las morbilidades relacionadas con la seguridad vial, son temas sensibles en muchos casos y que tanto los operarios de las concesiones, como las entidades territoriales relacionadas con la movilidad, conservan gran cantidad de datos como reserva para análisis internos, siendo que finalmente, la información reportada o de fácil acceso sea considerablemente menor a la requerida para realizar análisis profundos o poder relacionar las variables involucradas en la siniestralidad, buscando comprender mejor los diversos componentes que interactúan para desencadenar la ocurrencia de un siniestro.

Conforme a la solicitud realizada a la concesión Devinorte para acceder a las bases de siniestralidad o reportes conservados y recopilados en la operación de la concesión, se logra acceder a un importante volumen de datos, el cual una vez organizado muestra un comportamiento no lineal, dificultando el ajuste con modelos matemáticos o estadísticos tradicionales, de forma que el comportamiento y distribución de los datos, no se asemeja a una distribución probabilística conocida (Normal, binomial, etc. (Walpole, Myers, & Myers, 2012)) haciendo que los análisis estadísticos mayormente conocidos se vean limitados en su aplicación.

Finalmente, al considerar la distribución no lineal y algunas tendencias identificadas en el preprocesamiento de los datos, se percibe los modelos de Redes Neuronales Artificiales, como una posible opción para desarrollar un modelo que permita predecir el comportamiento de la siniestralidad en el corredor, basado en los eventos ocurridos en una serie temporal que incluye los datos tomados diariamente entre enero y diciembre de los años 2010 a 2016.

2.2. Modelo de Redes Neuronales Artificiales (RNA)

Las redes neuronales son una variación de un modelo matemático que pretende asemejar el comportamiento de un cerebro biológico, lo que implica que un computador de alta complejidad sea capaz de resolver problemas cotidianos que involucran patrones, percepción del espacio y control motriz. Por ejemplo, la visión humana, realiza una representación del entorno permitiendo reconocer un rostro familiar o ayudar a predecir la trayectoria de un vehículo mientras se conduce, para poder sortear un obstáculo o curva (Haykin, 1999). Similar al proceso de sinapsis ocurrido en los ejemplos previos, una versión compleja de las combinaciones de atributos sintéticos (producto cartesiano entre variables de entrada) pueden tener la capacidad de predecir un evento particular. Así pues, los modelos computacionales de redes neuronales consisten en una aproximación arbitraria a un evento que puede ser expresado mediante una función, en dónde una serie de capas que van unidas por conectores (probando diferentes combinaciones), y cuyo desempeño pretenden emular el funcionamiento de las neuronas en un cerebro humano (Bravo Caicedo & López Sotelo, 2017), establecen un símil entre las neuronas biológicas y artificiales adoptando las siguientes características:

- Las neuronas tienen señales de entrada que son percibidas a través de estímulos o eventos que alteran su comportamiento y pueden influir en la conexión con otras neuronas o músculos.
- Las señales percibidas por la neurona pueden modificar la salida del estímulo, mediante pesos sinápticos, que corresponden a la sinapsis o forma en cómo se comunican las neuronas. Esta comunicación se realiza a través de los

neurotransmisores cuya labor radica en filtrar la señal que percibe la neurona para bloquear o transmitir dicha señal desde o hacia otras neuronas.

- Los elementos procesados suman las entradas que son afectadas por la sinapsis, esto significa que las señales se suman en la neurona receptora permitiendo tomar una decisión basada en toda la información de entrada.
- Bajo condiciones favorables, la neurona puede transmitir una salida, si la sumatoria de señales de entrada hacen activar la neurona; lo que significa que es lo suficientemente importante para activar la respuesta.
- La salida que emite la neurona puede conectarse con otra neurona y pasar a ser la señal de entrada en la neurona receptora.

De manera general, una RNA está compuesta por una capa de entrada (X), una capa oculta (a1...an) y una capa de salida (Y). Las capas de entrada corresponden a las variables de entrada en la red, las capas ocultas se relacionan con la zona dónde el proceso iterativo se realiza y la capa de salida corresponde al resultado de la interacción o corrida.

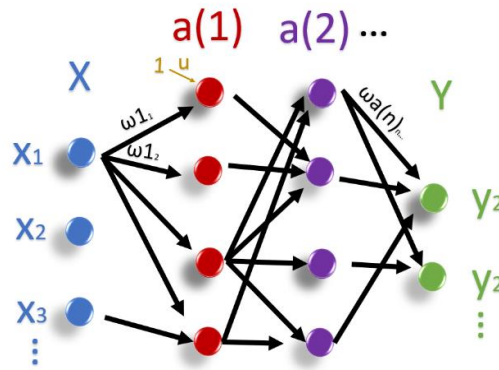


Figura 17 Modelo de redes Neuronales Artificiales
Fuente: Elaboración propia

A su vez, para cada neurona en la red, se tiene una propia capa de entrada que será denotada por la letra z, un peso sináptico ω y un parámetro conocido como umbral o función de activación (Giuseppe Ciaburro, 2017). La cual tiene como finalidad acotar la salida en la neurona y es muy importante que su derivada sea sencilla de tal modo que facilite el ajuste de los errores en las iteraciones del modelo.

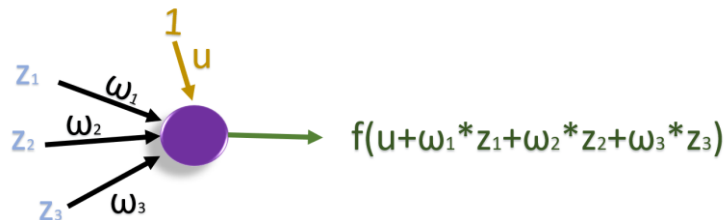


Figura 18 Funcionamiento del modelo de RNA
Fuente: Elaboración propia

Dado que el presente problema de investigación corresponde a una entrada de datos directa al modelo, se suponen como una distribución de probabilidad, adoptando una RNA con perceptrón multicapa. Esta consiste en modelo no paramétrico y no lineal que, en búsqueda de tener generalización y optimización de los pesos, se resuelve con el método bayesiano (Ceballos, Muñoz, & Moreno Cadavid, 2011).

Como parte de esa función de activación, existen varias opciones como la sigmoide (s), tangente hiperbólica, Rectified Lineal Unit, Softmax entre otras. En el caso particular, y como es de manera usual en modelos de predicción, se implementa la función sigmoide.

$$F(x)_{activacion} = (1 + e^{-x})^{-1}$$

Y su derivada como función de activación:

$$F'(x)_{activacion} = e^{-x} * (1 + e^{-x})^{-2}$$

El algoritmo de “aprendizaje” de la red se relaciona con la siguiente expresión:

$$\omega m + 1 = \omega m - \alpha f'(\omega m) = 0$$

Dónde el coeficiente α , se designa como razón de aprendizaje.

El error es una función de las omegas (ωi) dado que las entradas y las salidas son valores fijos. Se tiene que el descenso de gradiente en el perceptrón está condicionado por la alteración de cada una de las omegas, modificando el parámetro alfa y la derivada parcial de la función de error respecto a cada una de esas omegas.

En el proceso de aprendizaje y prueba de la red, se realiza la derivada parcial frente a la variación del error afectando solo una de las variables. Así pues, la expresión del error será determinada por:

$$\omega i \rightarrow \omega i - \alpha * \frac{\partial error}{\partial \omega i}$$

2.3. Construcción del modelo de RNA para siniestros que involucran solo daños

Teniendo en cuenta la caracterización del corredor y las diferencias marcadas tanto en geometría de la vía, como en los volúmenes vehiculares se considera necesario realizar dos modelos diferentes para reducir la posibilidad de generar sesgos o ruido en los datos de entrada. De esta manera realizando una analogía a los accesos que presenta el proyecto. Se plantea un modelo que es desarrollado con la información correspondiente a la Vía 5501 (Salida por Carrera séptima) y un segundo modelo se cargará a partir de los datos de la vía 45A04. Cabe resaltar la similitud en el procesamiento de los datos y las variables consideradas en ambos casos, hacen que la concepción lógica del modelo de

RNA sea la misma para el desarrollo de ambos tramos de vía. Sin embargo, se presentan las siguientes variaciones buscando encontrar los modelos que presenten un ajuste adecuado y cuyo desempeño involucre la mayor cantidad de variables.

Conforme a lo anterior, se plantea la variación de la variable dependiente conforme a los datos reportados de siniestralidad. A su vez, se plantea la modificación de las capas ocultas en el modelo de Redes Neuronales Artificiales, a fin de encontrar un mejor coeficiente de correlación.

El modelo simplificado incluye una función de activación sigmoide, método de disminución del gradiente y comprobación por diferencia de mínimos cuadrados entre el valor obtenido de los formatos de siniestralidad como variable dependiente y el resultado predicho por el modelo de RNA.

A continuación, se presenta la estructura básica del modelo realizado en el software Matlab R2020a, el cual cuenta con la caja de herramientas para Machine Learning y Deep Learning. Haciendo uso de la herramienta “Neural Fitting” (nftool); red tipo “feed forward” o prealimentada; con un perceptrón multicapa, la cual se aplica por su facilidad de aplicación y la posibilidad de variar de forma sencilla el número de capas ocultas, que la convierte en una alternativa viable para simplificar el proceso de identificación del ajuste en modelos con varias variables:

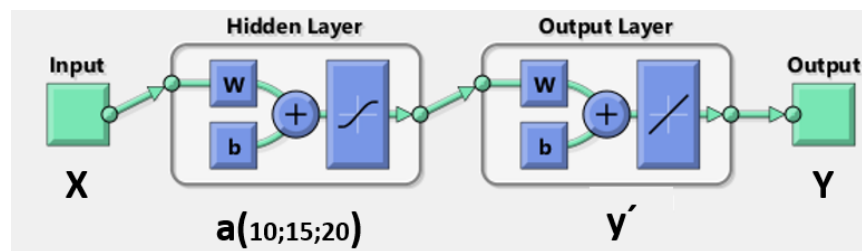


Figura 19 Modelo general de RNA Para siniestros en la concesión Devinorte
Fuente: Elaboración propia con base en Matlab R2020a

Dónde:

- X: Corresponde a las variables de siniestralidad incluidas como datos de entrada
- a: Corresponde a las capas ocultas de la RNA, que toma valores de 10, 15 y 20.
- y' : Corresponde a la variable dependiente o target (1 valor único)
- Y: Corresponde al valor de salida, predicho por el modelo (1 valor único)

2.4. Aplicación de un modelo de RNA para predecir la Abscisa de ocurrencia de un siniestro con solo daños

A continuación, se muestra el procedimiento realizado en la elaboración de un modelo que busca predecir la abscisa donde se presenta un siniestro en la concesión Devinorte para la vía 45A04:

Tabla 3 Variación de los modelos de RNA para la vía 45A04-Devinorte

	VARIABLE	DESCRIPCION	RANGO
1	MES	Mes de ocurrencia del siniestro	1-12
2	YEAR	Año de ocurrencia del siniestro	2010-2016
3	DIA	Día de la semana de ocurrencia del siniestro	1-7
4	HNUM	Hora en formato militar	0-24
5	TPD	Tránsito Promedio Diario Anual	20610-40241
6	D_N	Día o Noche	1-2
7	COND_MET	Condición meteorológica	1-3
8	CALS	Calzada E-W-Única	1-3
9	GEOMETRIA	Recta, Curva, etc.	1-4
10	CLASEAC1	Choque, atropello volcamiento Etc.	1-8
11	CLASEAC2	Choque, atropello volcamiento Etc.	1-8
12	CAUP1	Exceso de velocidad, Fallas mecánicas Etc.	1-304
13	CAUP2	Exceso de velocidad, Fallas mecánicas Etc.	1-304
14	ELEMENTO1	Automóvil, bus, buseta	1-20
15	ELEMENTO2	Automóvil, bus, buseta	1-20
16	HERIDOS	Se reporta herido, pero no requiere atención médica (0 NO; 1 SI)	1-2
17	ABSCISADO	Abscisa de ocurrencia del siniestro. (variable dependiente)	0-46

Fuente: Elaboración propia

Se ingresa como input las variables 1 a 16 y la variable 17 como Target

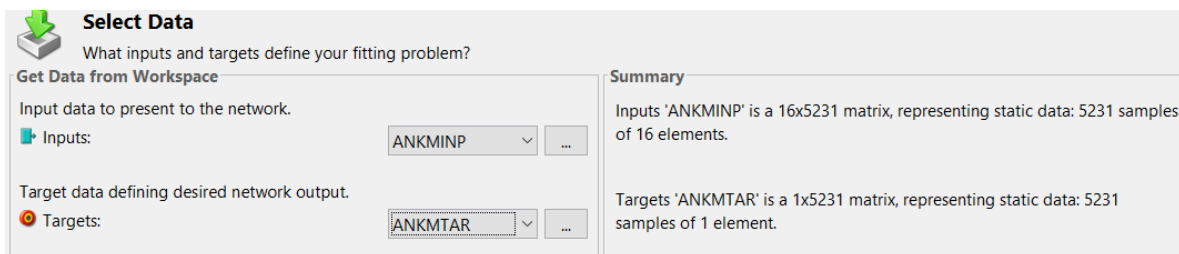


Figura 20 Datos de entrada para modelo de RNA para predecir abscisa de ocurrencia de un siniestro

Fuente: Elaboración propia con base en Matlab R2020a

Considerando que el software permite una variación en los datos de entrada para la realización del modelo; se toman el 70% de los datos en la etapa de “training” (correspondiente a 65% de training y 5% de validación) y 30% para “testing” (Koyel Chakraborty, 2020).

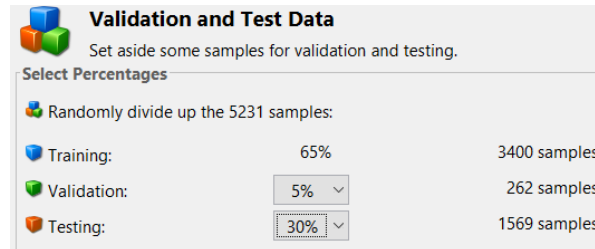


Figura 21 Distribución de los Datos de entrada para modelo de RNA para predicción de abscisas

Fuente: Elaboración propia con base en Matlab R2020a

Posteriormente, se ingresa los parámetros que definen la arquitectura del modelo, específicamente, la cantidad de capas ocultas:

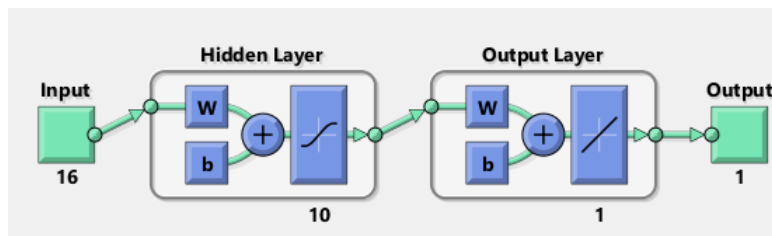


Figura 22 Arquitectura del Modelo de RNA con 10 capas ocultas para predicción de abscisas en siniestros con solo daños

Fuente: Elaboración propia con base en Matlab R2020a

Se procede a realizar el entrenamiento de la red con regulación bayesiana (Caudill, 1992) (grafos dirigidos) y se obtiene el resultado preliminar:

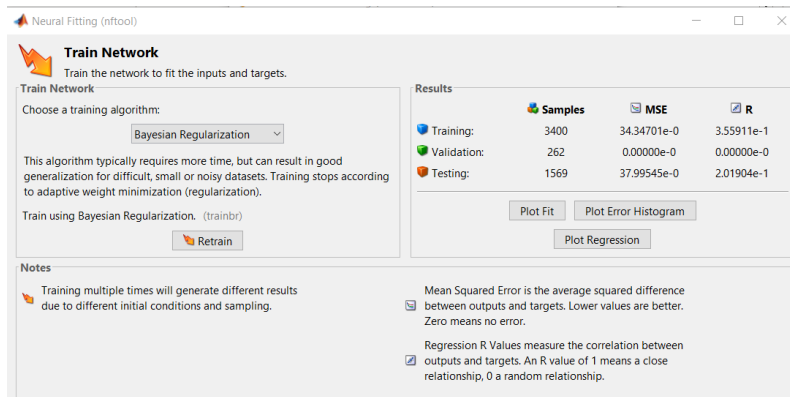


Figura 23 Resultados de un Modelo de RNA con 10 capas ocultas para predicción de abscisas en siniestros con solo daños

Fuente: Elaboración propia con base en Matlab R2020a

De los primeros entrenamientos se obtiene resultados de mínimos cuadrados y coeficiente de correlación, obtenidos de las diferencias entre el valor predicho por la red y el valor objetivo (Matlab, 2020).

Para este caso, se obtiene un coeficiente de correlación de 35.6, para la etapa de entrenamiento, un valor de 0.20 para la etapa de prueba y un combinado o total de 0.308; permitiendo inferir que, si bien hay relación, no es determinante para explicar la siniestralidad en el tramo 45A04. El resultado gráfico se presenta a continuación:

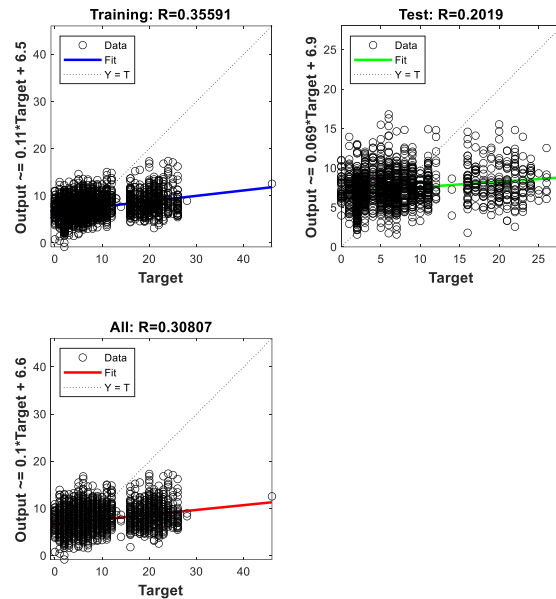


Figura 24 Resultados gráficos de un Modelo de RNA con 10 capas ocultas para predicción de abscisas en siniestros con solo daños

Fuente: Elaboración propia con base en Matlab R2020a

En el siguiente paso, se procede a modificar el número de capas ocultas para determinar si existe una mejora en el desempeño de la RNA:

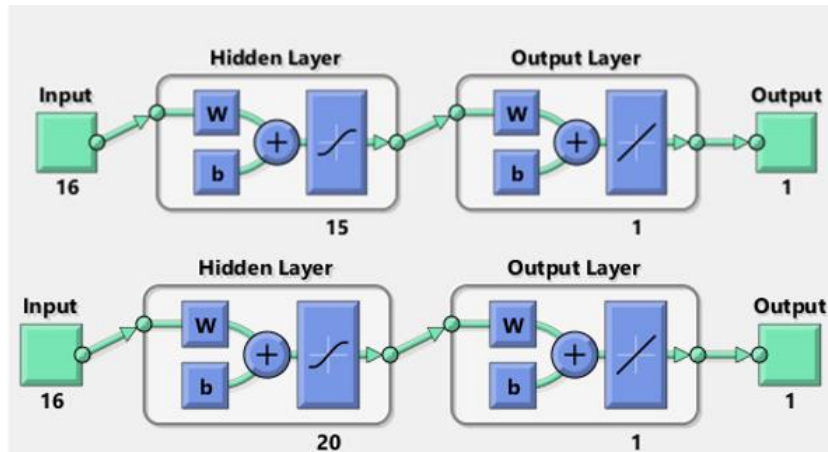


Figura 25 Arquitectura del Modelo de RNA con 15 y 20 capas ocultas para predicción de abscisas en siniestros con solo daños

Fuente: Elaboración propia con base en Matlab R2020a

A continuación, se muestra algunos de los resultados extraídos durante el proceso de entrenamiento de la red, donde se puede observar el descenso del gradiente y la validación, como pasos intermedios antes de la obtención de resultados:

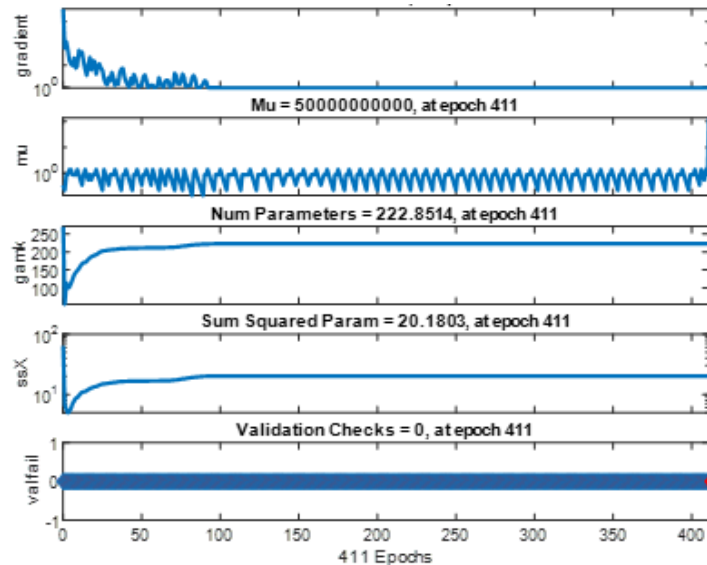
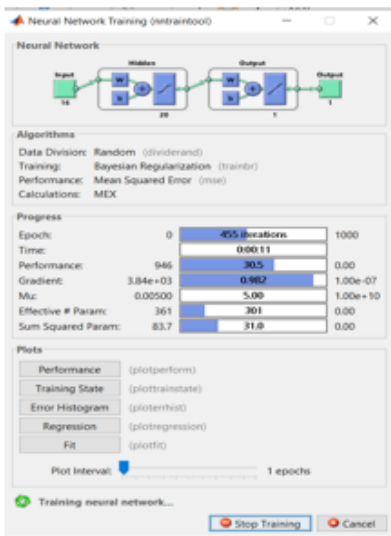


Figura 26 Procedimiento de aprendizaje para predicción de abscisas en siniestros con solo daños

Fuente: Elaboración propia con base en Matlab R2020a

De igual forma se realiza el procedimiento de cálculo para el modelo de RNA con 15 y 20 capas ocultas, reentrenado la red múltiples veces y obteniendo los siguientes resultados:

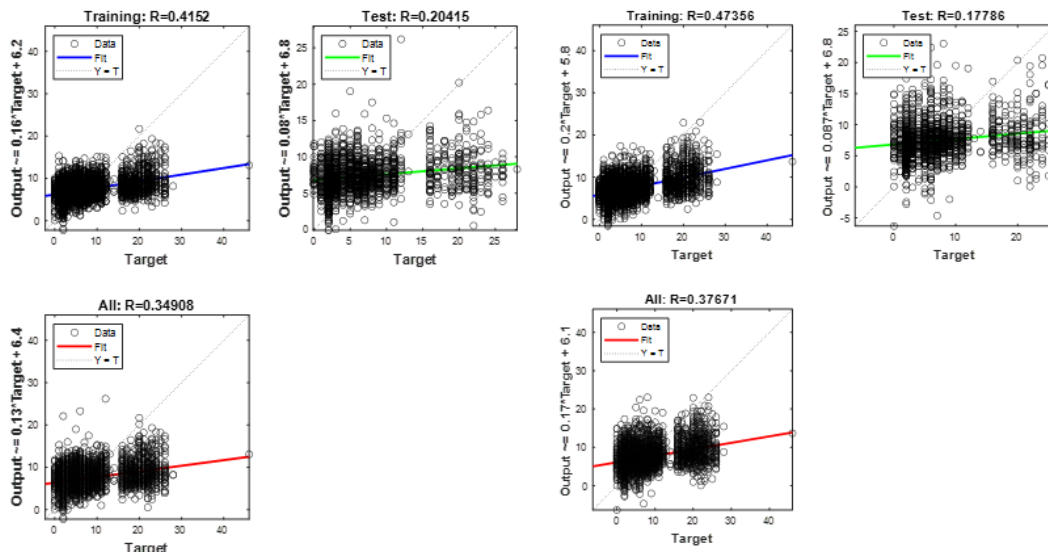


Figura 27 Resultados gráficos de un Modelo de RNA con 10 capas ocultas para predicción de abscisas en siniestros con solo daños

Fuente: Elaboración propia con base en Matlab R2020a

De lo anterior, se puede observar como en el proceso de aprendizaje de la red, se van obteniendo mejores resultados a medida que se incluye un mayor número de capas ocultas, sin embargo, esto altera de forma considerable el tiempo de procesamiento, y obtención de resultados, por tal razón se limita a 20 capas ocultas la ejecución del tercer modelos.

Si bien en este caso no se puede mostrar una correlación relativamente alta entre las variables escogidas y la abscisa de ocurrencia de siniestros con solo daños (coeficiente de correlación mixto e 0.37), si es posible interpretar como viable, una profundización, teniendo en cuenta que elementos como el salto en el comportamiento de los datos que se observa en las anteriores figuras, puede ser fraccionado buscando encontrar mayores correlaciones. Esto es posible, por ejemplo, si se realiza un proceso de sensibilidad mediante el cual se modifica la variable dependiente o variables independientes de tal forma que se obtenga un mejor desempeño del modelo.

2.5. Resultado de la aplicación de Modelos de RNA en la Concesión

Para determinar la pertinencia de escoger un modelo cuyo desempeño permita predecir con cierta precisión una situación o variable particular y cuya explicación matemática se pueda inferir a partir del coeficiente de correlación, se realizan varios ensayos donde se modifican los parámetros de entrada sin alterar el proceso descrito previamente.

Realizando los ajustes correspondientes, una vez Ingresados los valores de entrada y aplicando de manera similar para ambos segmentos viales (45A04 y 5501), se realizan variaciones de la cantidad de variables independientes, capas ocultas y variables dependientes obteniendo los siguientes resultados:

Tabla 4 Variación de los modelos de RNA para la vía 45A04-Devinorte

MODELO	VARIABLES INGRESADAS (X)	VARIABLE DEPENDIENTE (Y')	CAPAS OCULTAS (a)								
			10			15			20		
			TRAIN	TEST	TOTAL	TRAIN	TEST	TOTAL	TRAIN	TEST	TOTAL
1	16	# DE HERIDO LEVE	0,289	0,2783	0,29	0,27	0,21	0,26	0,33	0,261	0,319
2	12	ABS OCURRENCIA	0,307	0,132	0,291	0,337	0,19	0,321	0,32	0,12	0,301
3	10	ABS OCURRENCIA	0,307	0,164	0,292	0,34	0,159	0,321	0,32	0,119	0,301

Fuente: Elaboración propia

Como se puede observar, en la vía 45A04 se obtienen resultados inestables principalmente para la etapa de prueba, poniendo en consideración tanto el salto de datos mencionado en el anterior numeral, como el volumen de datos y la distribución de siniestros en este tramo, se pretende mejorar los valores del coeficiente de correlación, realizando un ajuste en los datos de entrada, que parte de la mayor concentración de siniestros en este tramo de vía, visible en los primeros 10 PR, y se procede a evaluar el comportamiento y la variación para estos valores:

Tabla 5 Variación de los modelos de RNA para la vía 45A04 Primeros 10 PR-Devinorte

MODELO	VARIABLES INGRESADAS (X)	VARIABLE DEPENDIENTE (Y')	CAPAS OCULTAS (a)								
			10			15			20		
			TRAIN	TEST	TOTAL	TRAIN	TEST	TOTAL	TRAIN	TEST	TOTAL
4	15	ABS OCURRENCIA	0,41	0,33	0,3986	0,45	0,273	0,432	0,487	0,278	0,463
5	13	ABS OCURRENCIA	0,385	0,339	0,381	0,432	0,337	0,422	0,461	0,227	0,437
7	13	CLASEAC1	0,436	0,325	0,402	0,497	0,2914	0,436	0,5373	0,317	0,47
8	13	HERIDO LEVE (SI NO)	0,451	0,495	0,453	0,518	0,4	0,5122	0,522	0,299	0,51
6	12	GEOMETRIA	0,509	0,344	0,458	0,508	0,297	0,442	0,572	0,296	0,49
9	10	HERIDO LEVE (SI NO)	0,521	0,336	0,468	0,517	0,335	0,463	0,57	0,316	0,492

Fuente: Elaboración propia

Se observa que al realizar la reducción de datos de entrada para la vía 45A04, pasando de 5231 casos a 4131, equivalente al 79% de los casos, se obtiene mejoras o aumentos cercanos al 18% para la etapa de aprendizaje, 14% para la etapa de prueba y 15% para el total. Lo que permite deducir, que el modelo de RNA responde mejor para los primeros 10 PR de la vía 45A04 en la concesión Devinorte.

De manera análoga, se procede a realizar las estimaciones con el segmento de la salida por la Carrera 7, vía 5501; para el cual se obtienen los siguientes resultados:

Tabla 6 Variación de los modelos de RNA para la vía 5501 -Devinorte

MODELO	VARIABLES INGRESADAS (X)	VARIABLE DEPENDIENTE (Y')	CAPAS OCULTAS (a)								
			10			15			20		
			TRAIN	TEST	TOTAL	TRAIN	TEST	TOTAL	TRAIN	TEST	TOTAL
1	16	# DE HERIDO LEVE	0,599	0,269	0,568	0,635	0,355	0,605	0,56	0,423	0,547
2	13	ABS OCURRENCIA	0,561	0,395	0,513	0,611	0,405	0,547	0,507	0,455	0,491
3	12	ABS OCURRENCIA	0,552	0,41	0,51	0,638	0,332	0,542	0,545	0,421	0,51
4	16	HERIDO LEVE (SI NO)	0,655	0,388	0,571	0,736	0,365	0,622	0,785	0,326	0,641
5	15	CLASE ACC	0,77	0,596	0,716	0,8	0,582	0,733	0,838	0,565	0,748

Fuente: Elaboración propia

De manera general, se observa un mejor desempeño en los modelos realizados para la vía 5501, teniendo en cuenta que la distribución de los casos es menos concentrada que en el tramo 45A04; lo cual puede ayudar a explicar la menor variabilidad de los modelos.

Se deduce que en el caso de la vía 45A04; los resultados más cercanos a la condición deseada se obtienen para la elaboración de un modelo predictivo con 15 capas ocultas, 13 variables de entrada, y con una certeza del 51%, que permite extraer la posibilidad de presencia de siniestros con solo daños, que además involucran el evento de presentar alguna morbilidad para el usuario que tiene el siniestro.

En el caso de la vía 5501; se tienen datos más homogéneos, con mejor desempeño y que permiten inferir que el modelo más confiable es un modelo predictivo con 15 capas ocultas y el cual, tras ingresadas 15 variables independientes, permite predecir la clase de siniestro que podría suceder con una confianza cercana al 73%.

2.6. Análisis de resultados para un modelo de RNA en la Concesión

Conforme a los resultados encontrados, se procede a profundizar el modelo cuya respuesta plantea un mejor desempeño, correspondiente al quinto modelo realizado para la Vía 5501, con 20 capas ocultas, con un R^2 combinado de 0.74 y cuyas variables se listan a continuación:

- Variable dependiente: “Clase de Accidente”
- Variables independientes: Mes, Año, Dia, Hora, TPD, D_N, Condición meteorológica, Abscisado, Calzada, Geometría, Causa Posible 1, Causa Posible 2, Elemento 1, Elemento 2, Presencia de herido leve.

El modelo presenta el siguiente diagrama:

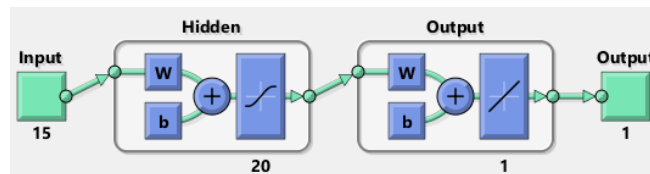


Figura 28 Resultados gráficos de un Modelo de RNA con 20 capas ocultas para predicción de la clase de accidente en siniestros con solo daños para la vía 5501 en la concesión Devinorte

Fuente: Elaboración propia con base en Matlab R2020a

En consecuencia, se plantea la variación de escenarios para evaluar las salidas del modelo. Esto ocurre con la inclusión de diversos escenarios planteados por el investigador, que permiten evaluar una condición particular.

Para esta evaluación se escoge un tramo de curva (PR 19) y se verifica la distribución horaria de los siniestros, la cual se presenta a continuación:



Figura 29 Distribución horaria en siniestros con solo daños para el PR 19 de la vía 5501 en la concesión Devinorte

Fuente: Elaboración propia con base en MS Excel

Sin embargo, teniendo en cuenta la cantidad de posibles combinaciones entre escenarios, se considera necesario establecer variables con valores fijos y realizar ajustes en algunas para analizar los resultados del modelo, los cuales se relacionan a continuación:

Tabla 7 Variables incluidas en el modelo predictivo para determinar la clase de siniestro con implementación de RNA para el PR 19 de la vía 5501 -Devinorte

Variable	Valor introducido
MES	AGOSTO
YEAR	2020
DIA	MIERCOLES
HNUM	16
TPD	**Variable ()
D_N	DIA
COND_MET	Normal
ABSCISADO	19
CALS	**Variable ()
GEOMETRIA	CURVA
CAUP1	EXCESO DE VELOCIDAD
CAUP2	IMPRUDENCIA DEL CONDUCTOR
ELEMENTO1	AUTOMOVIL
ELEMENTO2	N/A
HERYN	0

Fuente: Elaboración propia

Las variables que reciben modificaciones corresponden la calzada y el TPD, el cual se toma como base de 1287 (ANI, 2020).

Para los escenarios que se desea evaluar en la predicción, se incluye la variación de calzada según el sentido de circulación, y se ajusta un valor de TPD con incremento de 10% (1416); 20% (1544; 30% (1673); 40% (1802); y 50% (1931), generando 20 escenarios diferentes.

Mediante el uso de la herramienta nntool, se realiza la adaptación del modelo para los 20 escenarios obteniendo los siguientes resultados:

Tabla 8 Variables incluidas en el modelo predictivo para determinar la clase de siniestro con implementación de RNA para el PR 19 de la vía 5501 -Devinorte

TPD	INCENDIO (5)	SALIDA LATERAL (6)	CAIDA DE CARGA (7)	TOTAL
1287		2	2	4
1544	2	2		4
1673	2	2		4
1802	2	2		4
1931	2	2		4

TPD	INCENDIO (5)	SALIDA LATERAL (6)	CAIDA DE CARGA (7)	TOTAL
TOTAL	8	10	2	20

Fuente: Elaboración propia



Figura 30 Ubicación PR 19 de la vía 5501 en la concesión Devinorte

Fuente: Elaboración propia con base en MS Excel

Cabe resaltar que, si bien los valores no son exactos, se evidencia una clara tendencia a que el modelo relacione los resultados con un valor cercano a 6, el cual coincide con siniestros que involucran una salida lateral, siendo coherente con la ubicación del PR que está relacionado con un tramo de curva. Por otra parte, no presenta variación cuando se realiza el filtro por calzada, generando una distribución de resultados de 50% para cada costado de la vía

Capítulo III

Propuestas de atención para puntos críticos

Una vez recopilada la información de diagnóstico, el procesamiento de las bases de datos proporcionadas y la elaboración de modelos que involucran RNA; se tiene un amplio espectro de trabajo que permite tener diversas herramientas que facilitan la toma de decisiones para intervención en puntos específicos. Dado que estas zonas son identificadas en su mayoría como puntos de mayor concentración de siniestros, puntos de alta peligrosidad o puntos críticos con relación a todo el corredor; se percibe una necesidad de evaluar medidas específicas que promuevan la disminución de esos indicadores y permita una movilidad más segura para los usuarios de la vía.

A continuación, se resumen los 10 puntos críticos por nivel de peligrosidad conforme a la metodología expresada en el capítulo 1.

Tabla 9 Puntos críticos por índice de Peligrosidad vía 45A04 -Devinorte

45A04					
PR	Solo Daños	Heridos graves	Muertos	TPD	IP
2	997	11	4	31867	92,22
3	563	15	7	31867	59,05
6	476	11	0	31867	42,81
7	415	6	7	31867	44,78
9	272	17	15	31867	43,59

Fuente: Elaboración propia

Tabla 10 Puntos críticos por índice de Peligrosidad vía 5501 -Devinorte

5501					
PR	Solo Daños	Heridos graves	Muertos	TPD	IP
7	213	9	7	2548	349,24
12	77	5	6	2548	180
14	167	16	2	2548	242,79
19	125	12	3	2548	203,44
20	84	8	7	2548	208,38

Fuente: Elaboración propia

Con la finalidad de facilitar la ubicación espacial de los puntos, a continuación, se presentan de manera gráfica

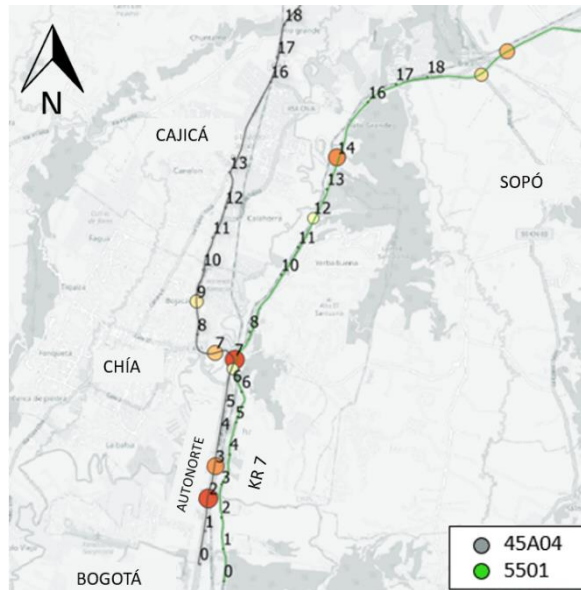


Figura 31 Principales Puntos con mayor peligrosidad Concesión Devinorte. 2010-2016
Fuente: Elaboración propia con base en QGIS

Como se puede observar, el 50% de los puntos críticos se ubican sobre la salida por la autopista norte y La Caro, sin embargo, teniendo en cuenta las mejoras que actualmente se llevan a cabo con la ampliación de los carriles enmarcados dentro de la concesión Accenorte, las obras de ampliación y señalización en el puente de la Caro y cuya operación llegó a reemplazar a Devinorte una vez fue finalizado el contrato de concesión. Para los demás puntos se resumen las siguientes recomendaciones:



Figura 32 Obras de mejoramiento Puente La Caro y ampliación Concesión Accenorte
Fuente: Elaboración propia con base en QGIS

Vía 45A04:

- **PR 9 a PR 10 sector Fontanar:**

En el punto se identifica diversos conflictos por ingreso y salida de vehículos, tanto de las zonas comerciales colindantes al centro comercial, como el retorno en sentido Zipaquirá-Zipaquirá cuya ubicación facilitaba las maniobras de incorporación riesgosas principalmente en sentido Bogotá-Zipaquirá.

Cabe resaltar que, en el año 2016, se efectuaron obras complementarias para mejorar los retornos y en la zona aledaña al Centro Comercial Fontanar; las cuales mejoraron la visibilidad, canalización de flujos y pretendían reducir las maniobras de invasión para dos o más carriles. Teniendo en cuenta lo anterior, la recomendación radica en mantenimiento, seguimiento y reemplazo de los dispositivos de segregación instalados a fin de que los dispositivos removidos por vida útil finalizada no representen facilidades para cometer maniobras con riesgo potencial alto para los usuarios que transitan por el corredor principal:



Figura 33 Obras de mejoramiento entre PR 9 y PR 10 Sector Fontanar-Devinorte

Fuente: Elaboración propia con imagen de Google Earth

Vía 5501:

- **PR 12 a PR 13 sector La Caro Sopó:**

Se identifican conflictos de ingreso y salida de vehículos por proximidad a zona estación de combustible en el sentido la Caro Sopó y zona comercial o bodegas. La visibilidad es buena considerando el alineamiento recto, pero también facilita el sobrepaso de la velocidad permitida; Se recomienda una posible forma de controlar la velocidad mediante dispositivos de limitación como bandas en agregado, las cuales pueden ser ubicadas de forma gradual, con un límite de velocidad de 40Km/h y sin ser implementadas con agregado en su totalidad tal como se muestra a continuación:



Figura 34 recomendaciones pacificación entre PR 11 y PR 12 Sector LA Caro Sopó-Devinorte

Fuente: Elaboración propia con imagen de Google Earth

- **PR 14 a PR 15 sector Hatogrande:**

En este sector se observan principalmente dos dificultades, la primera fue ajustada con el intercambiador de acceso a la Hacienda Hatogrande y la variante Cajicá-Hatogrande, la segunda tiene que ver con el asentamiento en el costado occidental en el sector Las Lomitas, dónde se presenta una concentración de vehículos por las oferta comercial y gastronómica del sector; causando entrecruzamientos y maniobras de acceso y salida de vehículos. De igual forma que en el caso anterior, se recomienda pacificar la zona con reductores de velocidad no tan agresivos, implementando límites de 40 Km/h y mejorando la visibilidad con iluminación artificial que sirva principalmente en horario nocturno. Finalmente se recomienda reforzar la señalización para advertir a los conductores sobre la posible presencia de peatones en la vía e incorporación de vehículos.



Figura 35 recomendaciones pacificación entre PR 14 y PR 15 Sector Hatogrande-Devinorte
Fuente: Elaboración propia con imagen de Google Earth

- **PR 19 a PR 20 Briceño-Sopó:**

Este sector tiene un comportamiento crítico principalmente en los fines de semana, debido a su proximidad con el municipio de Sopó; los diversos entrecruzamientos que se perciben por el intercambiador Briceño favorecen la aparición de siniestros con el aumento de volumen vehicular atraído de los días atípicos. Cabe resaltar que dentro del otrosí 61 del contrato de concesión 664 de 1994 y la APP Accenorte 2, se contempla la construcción de la Variante Sopó, cuya conexión se realiza entre el proyecto Perimetral Oriental de Bogotá (POB) y Accenorte Fase 1. Lo cual supone una mejora en las conexiones desde y hacia Sopó, disminuyendo también el volumen de vehículos que ingresan al intercambiador Sopó. Se recomienda tener en cuenta los indicadores presentados en este documento para los estudios definitivos de la Variante y generar alternativas de pacificación de la velocidad que permitan generar mayores brechas y, por tanto, facilitar las maniobras de incorporación con carriles de aceleración y desaceleración entre otros.



Figura 36 PR 19 y PR 20 Sector Briceño-Sopó-Devinorte
Fuente: Elaboración propia con imagen de Google Earth

Capítulo IV

Conclusiones

- Una vez verificados las distintas combinaciones entre variables dependientes e independientes, se confirma que es posible realizar un modelo predictivo con RNA que permita ayudar a la gestión de los indicadores de siniestralidad de la concesión. Sin embargo, teniendo en cuenta la gran variación de los datos, se recomienda la aplicación de estos modelos principalmente en el tramo 5501 de la concesión Devinorte.
- El limitado desempeño de los modelos realizados para el tramo 45A04 en la concesión Devinorte, puede obedecer a variables explicativas no contempladas y que no pueden ser explicadas por los datos de siniestralidad recopilados por la concesión. Para tal fin se recomienda la inclusión de parámetros como velocidades de operación, pendientes, parámetros específicos de diseño geométrico, e inclusive, mediciones de IRI o similares. De este modo es posible obtener un modelo robusto que pueda ayudar en la toma de decisiones.
- Teniendo en cuenta que el entrenamiento de la red obedece a aprendizaje autónomo, los resultados de los modelos varían cada vez que se realiza el entrenamiento, razón por la cual no se puede afirmar de forma definitiva que los resultados obtenidos obedecen al mejor modelo para cada variable estudiada.
- Los modelos de RNA se plantean como una alternativa viable y aun no masificada, que puede ser aplicados en diversos ámbitos y que pueden ser una herramienta para considerar en estudios de seguridad vial, ya que permiten recopilar información complementaria a las bases de siniestralidad y pueden profundizar los estudios tradicionales de seguridad vial, que en muchas ocasiones se ven estancados en las inferencias que se pueden obtener únicamente con estadística descriptiva.
- Se resalta que la aplicación de un modelo predictivo con RNA en la concesión arroja un espectro de resultados que si bien, ayudan al diagnóstico e identificación de potenciales variables o eventos que pueden desencadenar un siniestro. Estos datos no se deben extrapolar como un valor numérico específico.
- Los valores obtenidos deben ser interpretados de forma global, considerando que en el desarrollo del modelo se omiten diversas condiciones del terreno, geometría, externalidades y otras variables que no están contempladas como datos de entrada.
- El modelo predictivo con RNA, permite evaluar diversos escenarios facilitando la modificación de una o más variables de forma que sus resultados permiten inferir que debido a tratarse de eventos no aleatorios, sea posible determinar a menos en un diagnóstico inicial, las potenciales condiciones, variables o eventos que se pueden relacionar con la mitigación de la siniestralidad.
- Si bien es posible aplicar modelos de RNA en la base de datos de siniestralidad de la Concesión Devinorte, los procedimientos automatizados no reemplazan el

conocimiento y experiencia del investigador, debido a que se necesita tener una sensibilidad con los datos, requiriendo una alta preparación académica y laboral. Aun así, por la misma concepción teórica de las RNA, no es posible generar un alto grado de certeza para asegurar la escogencia del mejor u óptimo de los modelos.

- Durante los últimos años de operación de la concesión Devinorte (2016-2017), se dio atención a una serie de obras como el mejoramiento del Puente de La Caro, Mejoramiento de los accesos al Centro comercial Fontanar, pasos férreos y construcción e la ciclorruta Centro Chía, los cuales coinciden con varios de los puntos críticos de peligrosidad por siniestros ocurridos desde el año 2010; siendo un avance importante en la mitigación de los casos y morbilidades relacionadas con la siniestralidad.
- Cerca del 50% de los puntos críticos están relacionados con la salida por la Autopista norte hasta el puente de la Caro, siendo importante recalcar que la ampliación de carriles contemplados en el desarrollo de Accenorte Fase 1, ya contempla varias medidas de mitigación incluidas en los estudios de fase 3 para la construcción de las obras.
- El nuevo contrato de concesión adjudicado a Accenorte, tiene como uno de sus requisitos la implementación de sistemas de gestión de la seguridad vial, lo cual facilita que actualmente se continúen registrando los eventos de siniestralidad, que pueden permitir continuar alimentando las bases de datos incluidas en el presente documento, a fin de continuar alimentando el proceso de aprendizaje de las RNA que pueden contribuir a una mejora en el desempeño de los modelos predictivos.

Referencias

- ANI. (30 de 05 de 2020). *ESTADÍSTICA, TRÁFICO Y RECAUDO*. Obtenido de Agencia Nacional de Infraestructura: <https://www.ani.gov.co/estadistica-traffic-y-recaudo>
- ANSV. (01 de 05 de 2020). *Metodología para la identificación de sectores críticos de accidentalidad en zonas urbanas*. Obtenido de Agencia Nacional de Seguridad Vial: <https://ansv.gov.co/observatorio/index.html>
- Bravo Caicedo, E. F., & López Sotelo, J. A. (2017). *Una aproximación práctica a las redes neuronales artificiales*. Cali: Universidad del Valle.
- Caudill, M. a. (1992). *Understanding Neural Networks: Computer Explorations, Vols. 1 and 2*. Cambridge: The MIT Press.
- Ceballos, F., Muñoz, L. E., & Moreno Cadavid, J. (Diciembre de 2011). Selección de perceptrones multicapa usando aprendizaje bayesiano. *Scientia et Technica*, 110-115.
- Giuseppe Ciaburro, B. V. (2017). *Neural Networks with R: Smart models using CNN, RNN, deep learning, and artificial intelligence principles*. Birmingham: Packt.
- Haykin, S. O. (1999). *Neural Networks. A Comprehensive Foundation, second edition*. Ontario: Pearson Prentice.
- INVIAS. (15 de 06 de 2020). *INSTITUTO NACIONAL DE VIAS*. Obtenido de Poste de Referencia - PR: <https://inviasopendata-invias.opendata.arcgis.com/datasets/poste-de-referencia-pr>
- Koyel Chakraborty, A. A. (08 de 08 de 2020). *Neural Network Training*. Obtenido de <https://www.sciencedirect.com/topics/engineering/neural-network-training>
- Matlab. (2020). Version R2020a. Natick , Massachusetts: The MathWorks Inc.
- Ogwueleka, F. N., Misra, S., Ogwueleka, T. C., & Fernandez-Sanz, L. (2014). An Artificial Neural Network Model for Road Accident Prediction: A Case Study of a Developing Country. *Acta Polytechnica Hungarica*, 177-197.
- ONSV. (12 de julio de 2020). *AGENCIA NACIONAL DE SEGURIDAD VIAL*. Obtenido de OBSERVATORIO NACIONAL DE SEGURIDAD VIAL.
- Ozgan, E. a. (2008). Neural Networks-based Modeling of Traffic Accidents In Interurban Rural Highways. *Duzce Sampling. J. Applied Sci.*8, 146-151.

- Restrepo, B. (2018). *Aproximación a modelos de seguridad vial en intersecciones semaforizadas de Medellín*. Obtenido de Universidad Nacional de Colombia: <http://bdigital.unal.edu.co/71184/1/1035831424.2018.pdf>
- Silverman, B. W. (1986). *Estimación de densidad para las estadísticas y el análisis de datos*. New York: Chapman y Hall, .
- Tian, Y., Jana, S., Pei, K., & Ray, B. (22 de Abril de 2020). *DeepTest: Automated Testing of Deep-Neural-Network-driven Autonomous Cars*. Obtenido de s. <https://doi.org/10.1145/3180155.3180220>
- Urbano, E., Sánchez, J., & Quintero, I. (2018). *Modelo de predicción de gravedad de accidentes de tránsito: un*. Obtenido de https://vicisev.institutoivia.org/wp-content/uploads/2018/11/edwin-urbano-CISEV-Articulo_gravedad_accidentes-VRFNL.pdf
- Walpole, R. E., Myers, R., & Myers, S. L. (2012). *Probabilidad y estadística para ingeniería y ciencias*. ADDISON-WESLEY.