

USO DE LA INTELIGENCIA ARTIFICIAL PARA EL ANÁLISIS DEL COMPORTAMIENTO DE CONDUCTORES Y LA GENERACIÓN DE CHOQUES DE TRÁNSITO¹

Wilson Arias Rojas², Jorge Eliecer Córdoba Maquilón³, Germán Jairo Hernández Pérez⁴

RESUMEN: Esta investigación, es el resultado del análisis del comportamiento de conductores en un escenario controlado, en un simulador de conducción, en el cual, mediante la medición de ondas cerebrales, se midió el grado de concentración al conducir y por medio del uso de redes neuronales y de inteligencia artificial, se planteó un modelo de comportamiento de los conductores al someterse a un efecto distractor mientras se conduce, el cual permite analizar los factores más relevantes que se reflejan en errores y malas prácticas al momento de conducir. En esta investigación se determinó una muestra poblacional desde los 16 años hasta los 90 años, compuesta de hombres y mujeres, a partir de un universo obtenido de una base de datos de fatalidades durante 7 años, se construyó un simulador de conducción con un software para la simulación que permite diferentes escenarios de conducción. Finalmente se planteó un modelo de comportamiento de conductores ante los efectos distractores.

Palabras clave: comportamiento humano, conductas riesgosas, Neurosky, redes neuronales, simulación

USE OF THE ARTIFICIAL INTELLIGENCE FOR DRIVERS BEHAVIOR ANALYSIS AND VEHICULAR CRASHES GENERATION

ABSTRACT: This research is the result of the analysis of drivers' behavior in a controlled scenario, using a driving simulator, in which, by measuring brain waves, the degree of concentration was (measured when driving and through the use of networks neuronal and artificial intelligence, a model of behavior of drivers was proposed to be subjected to a distracting effect while driving, which allows analyzing the most relevant factors that are reflected in errors and bad practices at the time of driving. In this research it was determined a population sample of men and woman whose ages oscillate between 16 to 90 years, from a universe obtained from a database of fatalities for 7 years. A driving simulator was built, and it was using a software for the simulation that allows different driving scenarios. Finally, risk behaviors were classified to be a factor of distraction.

Keywords: human behavior, risky behaviors, Neurosky, neural networks, simulation

INTRODUCCIÓN A LA GENERACIÓN DE CHOQUES DE TRÁNSITO

El estudio de la frecuencia y la causalidad de las colisiones de tránsito en las vías públicas es un tema prominente debido a su impacto y costo a la sociedad. Cada año, de acuerdo con las estadísticas del extinto Fondo de Prevención Vial de Colombia, en un promedio de 2 minutos y medio ocurre un choque de tránsito, cada 10 minutos hay un herido y cada hora hay una víctima fatal, debido a la pobre seguridad vial que se presenta en el país. Los grupos de la población especialmente vulnerables son: los jóvenes de 15 a 30 años, los peatones, los ciclistas y en los últimos cinco (5) años, los motociclistas. Para Ferrer et al., (2013) no coinciden las cifras de choques y fatalidades que se manejan en Colombia por parte de la Policía Nacional, el "Dane", y el Instituto de Medicina Legal.

¹ Artículo recibido el 5 de noviembre de 2020 y aceptado para publicación el 8 de diciembre de 2020.

² Candidato a Doctor, Universidad Nacional de Colombia; Profesor, Fundación Universitaria Agraria de Colombia, Uniagraria, Cl. 170 ##54a-10, Bogotá, Colombia. Email: wiarasro@unal.edu.co/arias.wilson@uniagraria.edu.co

³ Profesor, Universidad Nacional de Colombia sede Medellín, Cl. 65 ## 59A - 110, Medellín, Antioquia, Colombia Email: jecordob@unal.edu.co

⁴ Profesor, Universidad Nacional de Colombia: sede Bogotá. Email: gjhernandezp@unal.edu.co

El exceso de velocidad es la principal causa en las estadísticas de choques de tránsito. El 89% de los choques de tránsito ocurren por factores humanos, el 11% restante, ocurren por problemas mecánicos, factores climáticos y defectos en la carretera entre otros (Arias y Colucci, 2006). La mayoría de los choques de conductores, ocurren en un promedio de 5 minutos antes de llegar a su destino, ya que el conductor se relaja y por ende, disminuye la atención en la conducción del vehículo (Arias y Colucci, 2006). Sin embargo, los autores no tuvieron en cuenta que el factor de las condiciones de la carretera contribuye en el 28 al 34 por ciento y el factor de los dispositivos de seguridad del vehículo contribuye en el 8 al 12 por ciento (Rumar, 1985 y Johnston, 2006).

La severidad de un choque puede estar relacionada con cambios repentinos en la velocidad límite de un segmento de carretera que podrían pasar desapercibidos por los conductores si la geometría o las características generales del segmento no varían considerablemente. Esta situación puede ser más definitiva en un segmento de carretera donde la velocidad límite no es acorde con las expectativas de los conductores considerando las características del segmento. Además, se deben tener en cuenta factores como la educación, la ingeniería, el manejo de incidentes en las carreteras y la ejecución de planes en las mismas para analizar las causas por las cuales ocurren los choques sobre el corredor vial seleccionado.

Se puede considerar que un conductor que esté en la zona urbana, donde la velocidad máxima permitida es de 60 kilómetros por hora (37.3 mph), y cambia a una zona rural donde aumenta la velocidad a 100 kilómetros por hora (62.1 mph), y se encuentra con una zona de construcción temporal, donde se debe disminuir la velocidad a 30 kilómetros por hora (18.6 mph), puede no reaccionar adecuadamente a este cambio y perder el control del vehículo ocasionando un choque, ya sean de índole severo o fatal.

Las fatalidades que ocurren en las carreteras por la ausencia o falta de mantenimiento de los dispositivos de control del tránsito y de seguridad instalados, como la demarcación inadecuada y la señalización en las vías, posiblemente sean una de las causas en el aumento de los choques fatales que tienen un muy alto costo para la sociedad y para la economía de un país.

El aumento en el precio de la gasolina puede tener un efecto reductor en los viajes durante ciertos períodos de tiempo, aumentando o disminuyendo la probabilidad de choques. El efecto de una vigilancia policíaca intensiva resulta en una modificación fugaz en el comportamiento de los conductores en la carretera debido al aumento en la percepción del riesgo de ser multado. Shinar y Stiebel (1986) y Benekohal et al. (1992) demostraron el efecto reductor fugaz en las velocidades de los conductores debido a un aumento en la presencia policíaca en la carretera. Vaa (1997) demostró que el efecto de una vigilancia policíaca intensiva en un segmento de carretera tiene una duración máxima de 8 meses en las velocidades de los conductores.

En términos de las muertes asociadas a la velocidad, el 54 por ciento de estos choques ocurrieron en segmentos de carretera con una velocidad límite rotulada de 56 Km/h o menos (NCSA, 2006). La "American Association of State Highway and Transportation Officials" (AASHTO, 2004) recomienda que la velocidad de diseño de una carretera sea consistente con la velocidad que los usuarios esperan con base en las características geométricas de la carretera. Cabe preguntarse cuál es la relación existente entre la velocidad de diseño, los componentes geométricos y la velocidad límite rotulada en estos segmentos que pueda ser un factor contribuyente a estos choques fatales. Es necesario analizar cuál es la relación entre el diseño geométrico de la carretera y la percepción del riesgo de choque de los conductores en estos segmentos.

Una investigación realizada en el Recinto Universitario de Mayagüez de la Universidad de Puerto Rico demostró que el 70 por ciento de los choques con peatones ocurren en zonas urbanas, el 98.5 por ciento ocurren en lugares fuera de las intersecciones y el 73 por ciento ocurre en segmentos de carretera rectos y llanos (Alicia, 2004). Este trabajo de investigación presentó varias recomendaciones a corto, mediano y largo plazo para mitigar el problema de seguridad de los peatones en Puerto Rico. Algunas de estas medidas son el establecimiento de estrategias de educación, instalación de barreras a peatones, instalación de señales de tránsito para peatones en selectos segmentos de carretera, enmiendas a la ley de tránsito y el uso de técnicas de ingeniería ("traffic calming") para reducir las velocidades de los vehículos.

“En tasas comparativas por población, la Unión Europea alcanzó en 2014 la cifra más baja hasta el momento en cualquier región del mundo de 51 muertes por millón de habitantes, 12 menos que en 2010. Los países con tasas más bajas son Suecia, Holanda, Reino Unido y Malta, por debajo de los 30 muertos por millón de habitantes. Mientras otros como Lituania, Bulgaria, Rumanía y Latvia superan los 90 muertos por millón de habitantes” (Alonso 2016). En Colombia esta tasa es de 121 muertos por millón de habitantes (WHO, 2015).

La tasa de choques y la tasa de mortalidad son indicadores útiles para conocer la escala del problema y para realizar comparaciones en seguridad vial. La frecuencia de choques está asociada al volumen de tránsito en la carretera por lo que cualquier evaluación estaría indebidamente influenciada por la cantidad de vehículos en la misma. La tasa de choques o de mortalidad por vehículo-Km recorridos (VKR) normaliza la frecuencia de los eventos por su exposición al riesgo de choque o muerte, respectivamente.

En varios países, se han establecido Sistemas de Gerencia en Seguridad Vial (“Safety Management Systems”) que implantan procesos sistemáticos e integrados en la planificación, el diseño y la construcción de carreteras con el propósito de reducir la cantidad de muertes. Además, estos sistemas incorporan procesos integrados en el manejo de datos de choques, la evaluación y análisis de los problemas de seguridad y la administración de un sistema de carreteras, entre otros procesos, que están enfocados en cuatro áreas esenciales: la educación y entrenamiento al conductor y otros usuarios del sistema de carretera; la ingeniería; el cumplimiento de las leyes (“enforcement”), y los servicios médicos de emergencia (en inglés este concepto se conoce las Cuatro E’s: “Engineering”, “Education”, “Enforcement” y “Emergency Services”).

Para la ejecución de este experimento se utilizó un simulador de conducción, apoyado del uso de inteligencia artificial para el análisis de los datos recopilados.

De acuerdo a Regan et al., (2008) se enumeran las ventajas y desventajas conocidas sobre los simuladores de conducción:

Ventajas

- Tienen la capacidad de crear escenarios a los conductores en situaciones probables de un choque sin ocasionar daños, como cuando hay consumo de drogas, cansados, en condiciones climáticas extremas, utilizando nuevas tecnologías, entre otras actividades peligrosas.
- Permite controlar muchas variables de confusión que ocurren mientras se conduce, por ejemplo, cuando se utiliza la simulación de conducción (p. ej., clima, tránsito, iluminación, frecuencia de usuarios vulnerables en la vía, viento, hoyos en la vía, proporción de tipos de vehículos, comportamiento inesperado o irracional de otros conductores, entre otros).
- Todos los detalles sensoriales del mundo real no son utilizados por los conductores de todos modos. La información perceptiva (Gibson, 1986) para conducir es conocible y puede ser reproducida de manera fidedigna utilizando simuladores.
- Los escenarios se pueden repetir idénticamente para cada participante del experimento.
- Los simuladores ofrecen ahorros de dinero ya que se puede hacer una configuración flexible para que se planteen una diversidad de preguntas de investigación (Jamson, 2001).
- Los simuladores de bajo costo y baja fidelidad en las manos adecuadas pueden abordar una amplia variedad de interesantes preguntas de investigación.
- La simulación de conducción es convincente y provoca reacciones emocionales de los conductores que son similares en la conducción real.
- Los simuladores son buenos para evaluar el desempeño del conductor o lo que puede hacer un conductor (Evans, 2004).

- Se puede configurar y ejecutar un currículo estructurado para la capacitación de nuevos conductores y para algunas habilidades, transferencias a carreteras reales (Pollatsek et al., 2006).

Desventajas

- Los choques simulados no tienen las mismas consecuencias que un choque real y puede afectar el comportamiento posterior. Los choques en un simulador pueden tener un impacto psicológico desconocido en los participantes.
- Las variables de confusión o interacción que ocurren en el mundo real también deben entenderse y, dado que no pueden recrearse completamente en simuladores, no son necesariamente susceptibles de prueba (todavía). Dicho de otra manera, comprender el comportamiento del conductor está en los detalles que interactúan en la simulación.
- Los escenarios reales nunca se pueden reproducir perfectamente (por ahora).
- Cada exposición de conducción afecta las respuestas a exposiciones posteriores.
- Los simuladores de alta gama requieren un hardware de altas especificaciones y el desarrollo de software para abordar un número limitado de preguntas de investigación.
- Los simuladores de bajo costo pueden ser imprecisos e inflexibles y, por lo tanto, abordar todas las preguntas de investigación que puedan surgir.
- Los conductores no creen en la autenticidad de la simulación.
- Los simuladores no pueden abordar preguntas sobre el comportamiento del conductor, ya que el conductor no conduce como lo hace en su propio vehículo (Evans, 2004).
- Se desconoce el grado de la experiencia y habilidades del conductor transferidas en la conducción en el simulador por lo tanto afecta la relación costo-efectividad de dichos programas (Jamson, 2001).

Materiales y Métodos

El material utilizado en este experimento se describe en el siguiente listado:

- Diadema “Neurosky MindWave mobile 2” utilizada para captar las ondas cerebrales.
- “Matlab” utilizada para grabar y registrar las señales captadas por la diadema “Neurosky”.
- “Forest Simulation”, programa para hacer la simulación de conducción.
- Equipo de simulación compuesto de: Simulador de autos estático: Procesador “Intel Core” I7 7700ATX – “Board” MSI-H110M - 16GB DDR4 - 1 TERA HDD-Unidad DVD-RW, Teclado y Mouse, Disco de estado sólido: Unidad SSD ADATA SPP550 - 240GB, Tarjeta gráfica: PNY GTX 1070 TI 8GB DDR5, “Bluetooth”, Timón y pedales Logitech modelo G29 con adaptador para timón de vehículo de calle, Chasis en tubería estructural, hecho en tubo cuadrado, con soporte de silla y corredera, timón pedales y caja con ajuste gradual y soporte de pantalla, sonido y acabado en pintura electro estática.
- Televisor LED SAMSUNG 32" - UN32J4000DKXZL, Bocinas Logitech Z506 - 5.1 CH. Este simulador fue construido por el autor del proyecto para la realización de esta investigación doctoral.
- Celular marca Alcatel modelo U5.

Las figuras 1, 2 y 3 muestran la diadema utilizada para la captura de la información, una imagen obtenida del simulador de conducción y el simulador construido para esta investigación.



Figura 1: Diadema “Neurosky”.



Figura 2: Escenario urbano en el simulador de conducción.



Figura 3: Simulador construido para esta investigación.

La muestra poblacional que participó en esta investigación cumplió ciertos criterios básicos, los cuales se presentan a continuación:

- Tener una licencia válida de conducción.
- Haber conducido al menos por más de 1 año.
- Haber conducido al menos durante 1 semana en el último año de la prueba.
- No presentar desordenes físicos ni cognitivos.
- No tener enfermedades del sistema nervioso central.
- No estar embarazada.
- No haber consumido bebidas alcohólicas ni sustancias alucinógenas antes de la prueba.

La enfermedad del simulador es un fenómeno que se ve afectado por las especificaciones del simulador y características de los participantes. Produce síntomas similares, pero típicamente lo que presentan los participantes que conducen en el simulador de conducción es mareo, náuseas, incomodidad ocular y desorientación (Kennedy et al., 1993).

Condiciones del área de conducción

El escenario para ejecutar la prueba de conducción fue el mismo para los 167 participantes del experimento. Se condujo en una zona urbana, con tránsito de hora pico, entre las 6:00 PM y las 7:00 PM, en un ambiente dentro y fuera de zonas urbanas, con peatones y vehículos circulando en el escenario. La zona urbana presenta andenes, zonas comerciales, zonas de parqueo a la orilla de la vía, gasolineras, señales de tránsito y glorietas a lo largo del desarrollo de la ruta.

Escenario de tránsito

El efecto del flujo de tránsito y su efecto distractor mientras se conduce, es importante en esta investigación. La simulación del ambiente (por ejemplo, el comportamiento de otros vehículos en la simulación de la red vial del experimento) puede ser muy compleja en algunos casos. Es importante observar la simulación del comportamiento de no más de 1 o 2 vehículos en relación con el vehículo simulado en este experimento. La simulación del ambiente del tránsito es mucho más demandante que el modelo clásico de microsimulación de tránsito, por tal razón, el software tiene implementado un marco de movimiento similar que cuando se conduce de manera natural, de esta manera, el ambiente no es estático.

La distribución de los participantes del experimento, por edad, género, estrato socioeconómico y grado de escolaridad. 39 participantes fueron mujeres y 128 fueron hombres, para un total de 167 personas.

Distribución de los participantes por grupo y por edad

La Tabla 1 muestra la distribución de los participantes por rango y edad, concentrándose la mayor población femenina en el rango de 31 a 40 años y la masculina de 21 a 30 años, seguida por el rango de 31 hasta los 60 años.

Tabla 1: Distribución de participantes por grupo y rango de edad.

Rango de edad	Femenino	Masculino	Gran Total
0 a 20	2	15	17
21 a 30	8	28	36
31 a 40	15	23	38
41 a 50	7	23	30
51 a 60	3	23	26
61 a 70	2	7	9
71 a 80	1	6	7
81 a 90	1	3	4
Gran Total	39	128	167

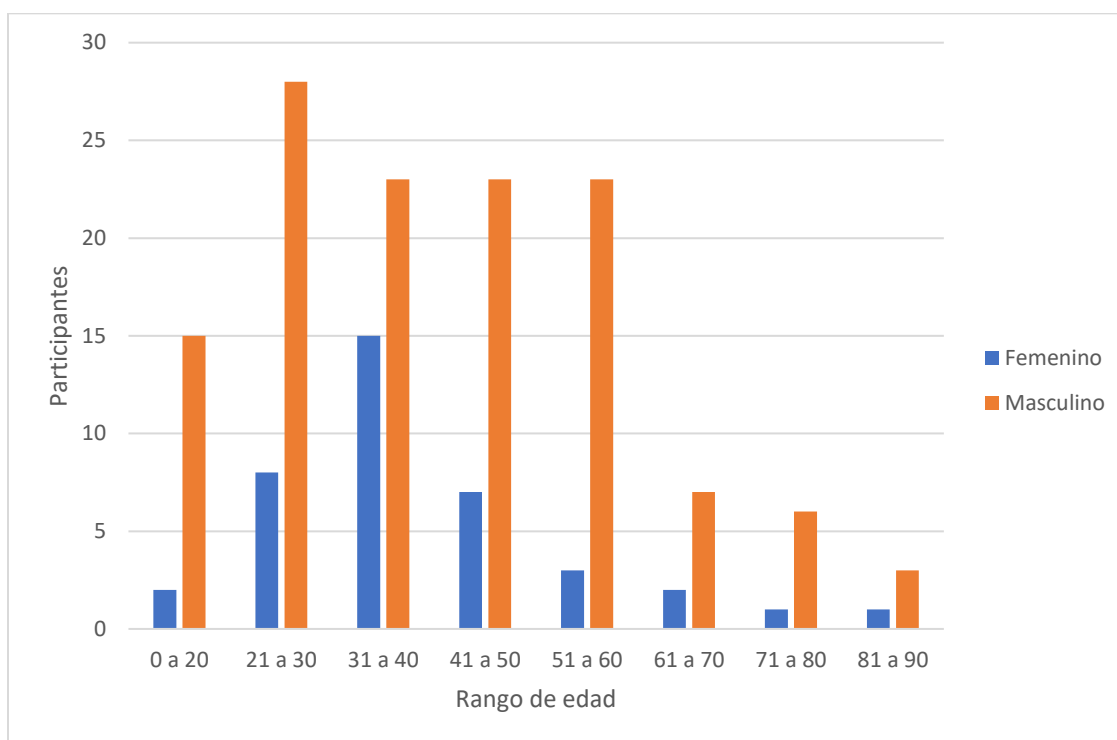


Figura 4: Distribución de participantes por grupo y rango de edad.

El estrato socioeconómico es una clasificación de los inmuebles residenciales en Colombia para cobrar de manera diferencial por estratos los servicios públicos domiciliarios, subsidiando a los estratos más bajos (1 y 2) y cobrar impuestos a los estratos más altos (5 y 6). Por lo tanto, quienes reciben mayores ingresos económicos, deben pagar más por los servicios públicos domiciliarios y se subsidian a los estratos más bajos.

Tabla 2: Distribución de participantes por estrato socioeconómico.

Estrato socioeconómico	Femenino	Masculino	Gran Total
1		5	5
2	7	30	37
3	28	71	99
4	3	17	20
5		5	5
6	1		1
Gran Total	39	128	167

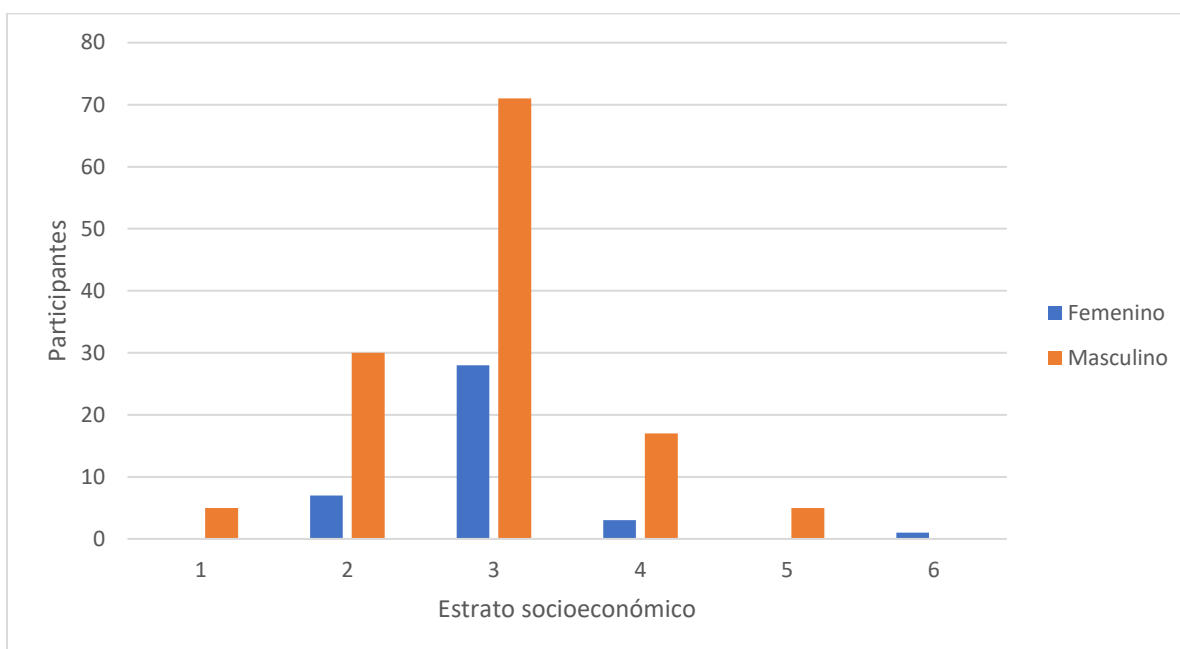


Figura 5: Distribución de participantes por estrato socioeconómico.

Tabla 3: Distribución de participantes por nivel de escolaridad.

Nivel educativo	Femenino	Masculino	Gran Total
Primaria	1	9	10
Bachillerato	10	33	43
Tecnológico	12	26	38
Pregrado	12	52	64
Maestría	4	5	9
Doctorado		3	3
Gran Total	39	128	167

Se escogió como escenario la penumbra (de 6:00 PM a 7:00 PM) en un entorno urbano, con un lapso de tiempo de 10 minutos de conducción. Condiciones climáticas secas.

Al conductor se le preguntó si desea conducir el vehículo con transmisión mecánica o transmisión automática. Esta información es necesaria para configurar el software del simulador.

Al conductor se le instala la diadema “Neurosky” en la cabeza y la información que se obtenga de la misma será registrada vía “bluetooth” al computador, en el cual con el programa SDK de “Neurosky” y el desarrollo del programa en “Matlab”, se va grabando la información en tiempo real del grado de concentración del conductor durante la prueba.

Además el programa del simulador va grabando los errores cometidos por el conductor durante la prueba.

Al terminar el escenario, los datos del sujeto son grabados tanto los obtenidos con la diadema, como los procesados por el simulador, entre otros la hora de inicio y final de la prueba y el tiempo en segundos de cada uno de los errores cometidos en la conducción.



Figura 6: Participante utilizando el simulador de conducción.

RESULTADOS

La información de estudios de transporte generalmente es modelada utilizando dos tipos de aproximaciones: la estadística clásica y la inteligencia artificial. En la estadística, se recolecta, organiza e interpreta la información numérica con herramientas matemáticas, particularmente cuando esta información está relacionada con características de la población como la inferencia de una muestra (Glymour et al., 1997). Los modelos estadísticos tienen bases y fundamentos matemáticos muy robustos ampliamente aceptados y suministran ideas acerca de los mecanismos de la creación de datos. Sin embargo, con frecuencia fallan cuando se trata de datos complejos y altamente no lineales. La inteligencia artificial combina conceptos de aprendizaje, adaptación, evolución y lógica difusa para crear modelos que son “inteligentes en el sentido que la estructura surge de un comienzo no estructurado (Engelbrecht, 2007; Sadek et al., 2003).

Las redes neuronales son una clase extremadamente popular de la inteligencia artificial y han sido ampliamente utilizados en diferentes problemas de transporte, particularmente porque son modelos matemáticos muy genéricos, precisos y convenientes, capaces de simular fácilmente componentes numéricos de un modelo. Las redes neuronales presentan una propensión inherente a almacenar conocimiento empírico y pueden usarse en cualquiera de las tres maneras básicas (Haykin, 1999): i. Como modelos de sistemas nerviosos biológicos; ii. Como procesadores/controladores de señal adaptativos en tiempo real; iii. Como métodos analíticos de datos. En las investigaciones de transporte, las redes neuronales se han utilizado principalmente como métodos de análisis de datos debido a su capacidad de trabajar con grandes volúmenes de datos multidimensionales, su flexibilidad de modelado, su aprendizaje, su capacidad de adaptación y predicción que es buena en general. (Karlaftis y Vlahogianni, 2010).

Con bastante frecuencia, los investigadores que dominan uno de los dos enfoques, discuten fervientemente en apoyo de su método elegido. La selección del enfoque de análisis es uno de los temas más debatidos en las reuniones y publicaciones de investigación y aunque estos argumentos proporcionan interesantes debates científicos, logran

confundir completamente a los investigadores más jóvenes, en particular a profesionales que están más interesados en el modelo que deberían usar en lugar de concentrarse en los fundamentos filosóficos o matemáticos de los enfoques. (Karlaftis y Vlahogianni, 2010).

En la etapa experimental del proyecto, se registraron datos de 167 sujetos distribuidos en 39 mujeres y 128 hombres, cumpliendo los rangos de edad de la muestra poblacional y con diferentes tipos de escolaridad y estrato socioeconómico.

El volumen de información total de todo el experimento se resume con la siguiente operación matemática:

$167 \text{ sujetos} \times 30 \text{ errores de conducción} \times 28000 \text{ registros} = 140.280.000 \text{ datos promedio.}$

Se observa que la cantidad de datos es muy alta, sin embargo, este tipo de datos no se clasifican como Big. Data, ya que para serlo se deben cumplir los siguientes criterios

- Volumen de la información.
- Frecuencia de la información.
- Variabilidad de la información

Se obtienen muchos datos de múltiples fuentes pero la frecuencia es muy corta por la duración del experimento de cada sujeto que solo fueron 10 minutos. Si fuera “Big Data”, la información que se recibiría constantemente en el tiempo.

El simulador entregó después de la prueba, 30 parámetros, los cuales se encuentran codificados de acuerdo a un número y que se puede ver en el archivo de errores de conducción o “types” que suministra el desarrollador del “software”.

Estos parámetros son errores o situaciones que se presentan mientras el sujeto está conduciendo en el lapso de tiempo de la prueba.

Con los datos de los 167 sujetos y los 30 parámetros, se procede a hacer una matriz de 167 filas por 30 columnas y se parametrizan los valores obtenidos entre cero (0) y uno (1), para ser procesado por la red neuronal. Estos datos parametrizados entre cero (0) y uno (1) son los grados de concentración de los sujetos durante la ejecución de la prueba.

- Se calculó la media aritmética de cada uno de los eventos y se multiplicó por el tiempo de ocurrencia de cada uno de los eventos. Se usó este tipo de medida por ser más sensible a los cambios de las variables.
- Si el error de conducción ocurre al inicio del experimento, puede ser un error del sujeto.
- Si el error ocurre al final del experimento, el sujeto no está concentrado en la conducción del simulador.
- Se calcula el promedio del tiempo por el evento, para convertir toda la información en un solo dato, con el fin de procesarlo más adelante con la inteligencia artificial.
- Se calculó la derivada a la variable concentración para determinar la variación de la misma, para cada uno de los sujetos. Si es positivo el valor, el sujeto se encontró concentrado y si es negativo, se encontró desconcentrado.
- Se normalizaron los valores obtenidos, en un rango de 0 a 1, para poder procesar la información en la red neuronal.
- Se obtienen los errores más comunes cometidos por los sujetos de la prueba. Estos errores son suministrados por el programa del simulador y tienen el nombre de “types”. Durante la ejecución de este experimento, se presentaron 30 errores de conducción recurrentes en los 167 sujetos del experimento.

- El peso se obtiene introduciendo la matriz de 167 x 30 en “Matlab” y luego es recalculado con el entrenamiento de la red neuronal.
- Los pesos son multiplicados por los errores de conducción más recurrentes y como resultado se obtiene la concentración de cada sujeto que ejecutó la prueba.

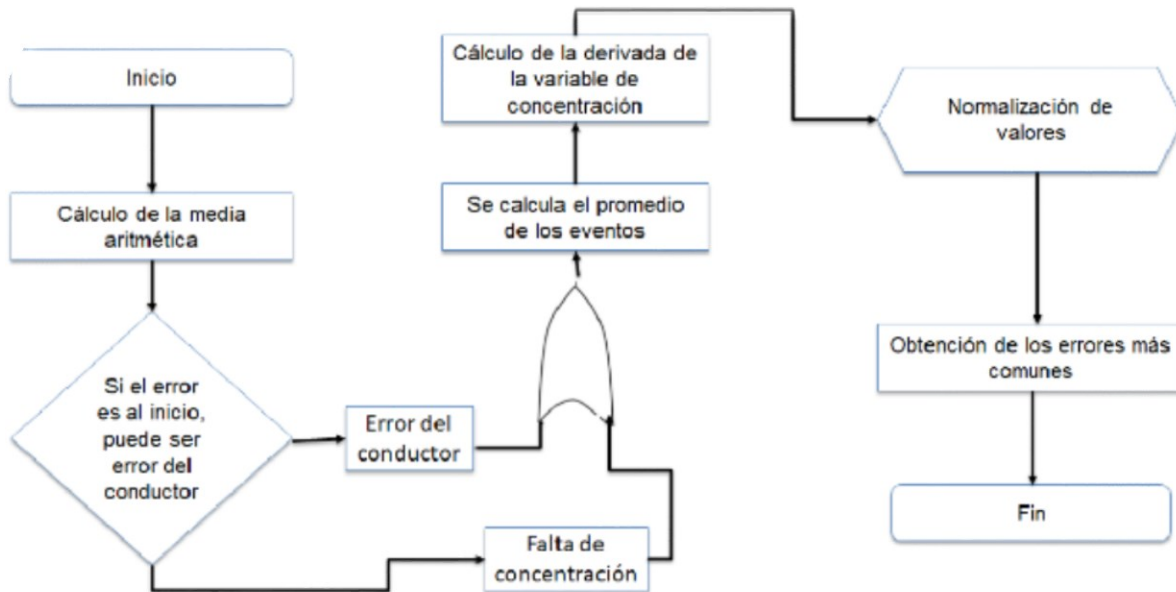


Figura 7: Diagrama de flujo el proceso de análisis de datos en la red neuronal.

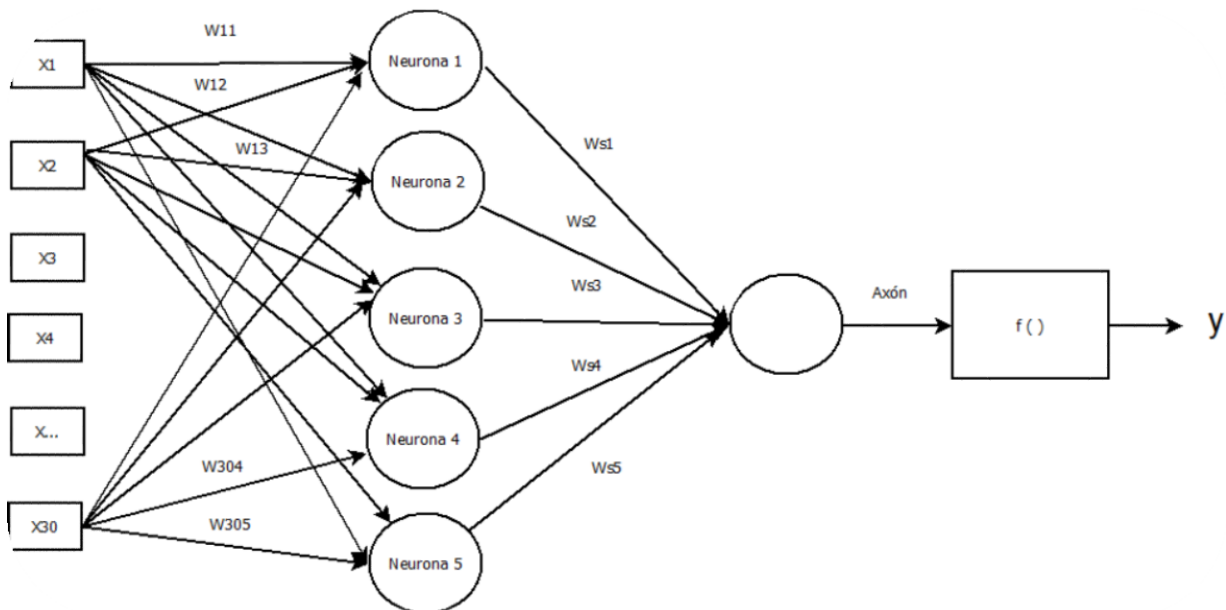


Figura 8: Modelo de la red neuronal utilizada.

Una vez se ha analizado toda la información, se procede a observar cual es la función de activación la cual representa de manera simultánea la salida de datos de la neurona y el estado de activación.

De igual manera, al ejecutar la red neuronal, la misma calcula el valor del umbral en el cual se discriminan los valores relevantes tanto del eje positivo como el eje negativo del modelo matemático, presentando una ecuación como se expone a continuación:

$$y = \frac{W_s x [W_i X_i]}{b} \quad (1)$$

Donde:

W_i matriz de pesos de la capa de entrada, ponderación que asigna la red neuronal a las entradas.

X_i son los datos de entrada.

W_s es la matriz de pesos o ponderación que asigna la red neuronal a la capa de salida o a las entradas procesadas. Como el umbral es una función escalonada, el modelo de comportamiento planteado en esta investigación doctoral está dado por la siguiente ecuación:

$$y = f(x) = \begin{cases} \text{Concentrado} & \forall [X \cdot W] \cdot W_s \geq b \\ \text{Desconcentrado} & \forall [X \cdot W] \cdot W_s < b \end{cases} \quad (2)$$

Siendo la pendiente, los pesos W , de los errores de conducción del experimento y b el umbral calculado por la red neuronal.

El modelo que permite predecir si la persona está concentrada o no y en consecuencia si existe probabilidad de accidente está descrito por la ecuación (1).

Uso de “Auto Machine Learning”

Al aplicar un modelo de regresión “Stepwise Forward”, utilizando “Auto Machine Learning”, ejecutando regresiones logísticas, se obtuvieron los siguientes valores, donde se hizo una relación entre el intercepto y la concentración, encontrando una asociación significativa o “p-value” de 0.0211, encontrando que la concentración es un factor que explica la accidentalidad. De la misma manera, se muestra la significancia del modelo y se observa con el “p-value” Para ello se utilizó el “software Real Statistics” V.7.0 de Marzo 2020 el cual es el software R V.3.2.6 de Marzo 2020 para Excel.

La Tabla 4 contiene los valores explicados anterior y adicionalmente el intervalo de confianza (“lower y upper”).

Tabla 4: “Stepwise Forward”: probabilidad de accidentalidad.

	“coeff b”	“s.e.”	“wald”	“p-value”	OR	“lower”	“upper”
“Intercept”	1.18	0.26	21.25	0.000	3.25		
“Conc. Norm.”	1.05	0.45	5.37	0.021	2.85	1.17	6.90

En la Tabla 5 se observan los valores obtenidos de, χ^2 , “p-value”, el área bajo la curva y la precisión del modelo.

Tabla 5: Valores de para el modelo de concentración.

Chi-Sq	5.84
p-value	0.016
AUC	0.87
Accuracy	0.83

La ecuación obtenida para este modelo se presenta a continuación al igual que la Figura 9 muestra el área bajo la curva para este modelo obtenido. Esta ecuación representa la curva del modelo planteado en esta investigación, donde se obtiene una precisión el modelo del 83% con el 87% de los datos totales obtenidos.

$$p = \frac{1}{1 + e^{-1.18 - 1.05x}} \quad (3)$$

Y la ecuación correspondiente al modelo de concentración es determinada por

$$Y = 1.18x_1 + 1.05x_2 \quad (4)$$

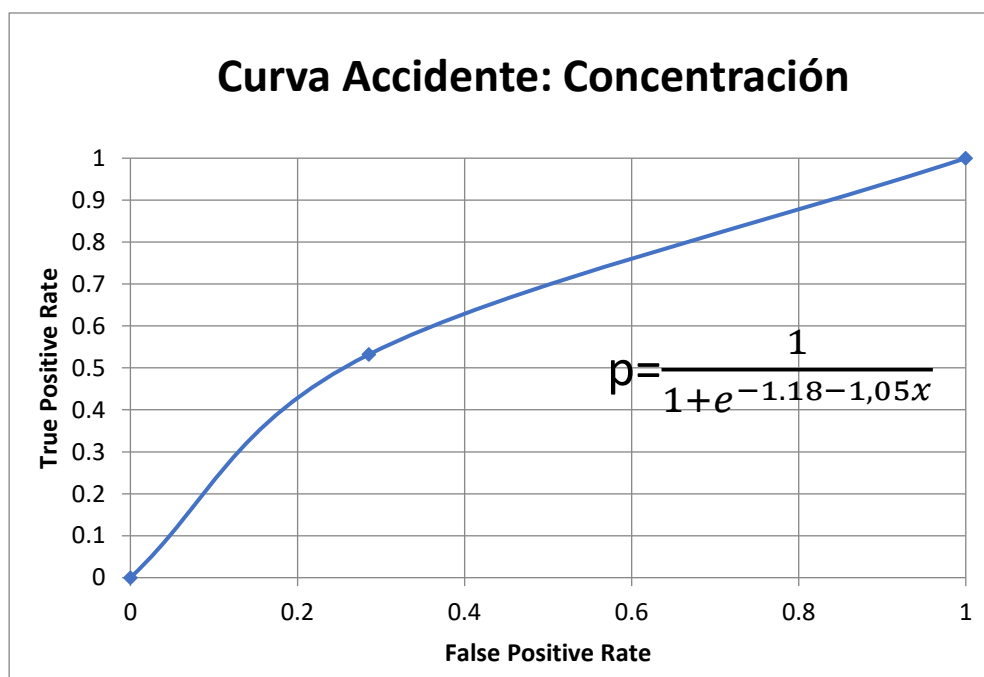


Figura 9: Área bajo la curva del modelo obtenido.

CONCLUSIONES

El objetivo principal de esta investigación consiste en el desarrollo de un modelo de comportamiento por medio del uso de inteligencia artificial, en el cual se evaluaron las conductas riesgosas mientras los sujetos de la prueba conducían en el simulador de conducción, evaluando los errores y creando un modelo que determina a partir de un umbral, si el conductor en la prueba se encuentra concentrado o desconcentrado al momento de conducir en el simulador. Adicionalmente, se evaluó si el incremento del riesgo choques incide por la distracción del conductor.

La ecuación de este modelo de comportamiento está definida por la siguiente expresión:

$$Y = 1.18x_1 + 1.05x_2 \quad (5)$$

Esta ecuación es el resultado de haber podido medir el factor más importante del comportamiento humano y es la conducta, a través de la atención y concentración de los conductores que participaron en el experimento.

Se clasificaron y caracterizaron las conductas riesgosas de conductores al ser sometidos a un factor distractor, sin importar el género ni la edad y que de alguna manera influyeron en el comportamiento al momento de conducir el escenario recreado.

Se observó que los conductores que hicieron el experimento con vehículo con transmisión mecánica presentaron mayores dificultades y por ende, se distraen más fácilmente, al momento de recibir mensajes de "Whatsapp" por el celular, cometiendo equivocaciones más recurrentes que los conductores que prefieren la transmisión automática.

Se observó que las personas mayores de 50 años, tienden a no revisar el celular mientras están conduciendo, mientras que las personas entre el rango de 18 a 45 años revisan el celular y en varias ocasiones, contestan los mensajes mientras están conduciendo, afectando el grado de concentración al momento de la prueba y presentando errores de conducción y choques de manera más recurrente.

El modelo matemático obtenido, depende del grado de concentración de los conductores y varía, dependiendo del escenario, condiciones climáticas, tipo de carretera, hora en la que se conduce, tipo de vehículo y tipo de transmisión del vehículo.

REFERENCIAS BIBLIOGRÁFICAS

- American Association of State Highway and Transportation Officials (AASHTO, 2004). *A Policy on Geometric Design of Highways and Streets*, Washington, D.C., USA.
- Af Wählberg, A.E. (2012). "Changes in driver acceleration behavior over time: Do drivers learn from collisions?" *Transportation Research Part F: Traffic Psychology and Behaviour*, 15(5), pp.471–479.
- Alonso G. et al. (2004). Entrenamiento de una red neuronal artificial usando el algoritmo simulated annealing, *Revista Scientia et Technica*, Año X, No 24, Mayo 2004. UTP. ISSN 0122-1701.
- Alicea, L. (2004). "Analysis and Evaluation of Crashes Involving Pedestrians in Puerto Rico", Tesis de Maestría, Departamento de Ingeniería Civil, Universidad de Puerto Rico, Recinto Universitario de Mayagüez.
- Alonso, M. (2016). La integración del factor humano en el ámbito técnico de la gestión de las carreteras y la seguridad vial: Un enfoque investigativo. Disponible en: <http://roderic.uv.es/handle/10550/51943>.
- Arias, W., Colucci, B. (2006). "Road Safety Audit", 19(3), p.28.
- Benekohal, R.F., Hashmi, A.M. (1992). "Procedures for estimating accident reductions on two-lane highways", *Journal of Transportation Engineering*, 118 (1), pp. 111-129.
- Engelbrecht, A.P. (2007). *Computational Intelligence. An Introduction*, Second Edition, Wiley, Nueva York, Estados Unidos.
- Evans, L. (2004). *Traffic safety*. Bloomfield Hills, MI: Science Serving Society.

- Ferrer, A., Smith, R. & Cuellar, M. (2013). Análisis de la Capacidad de Gestión de la Seguridad Vial. *Banco Mundial*, pp.92–93.
- Figuroa, A. (2005). "Speed factors on four lane highways in free flow conditions", Tesis doctoral, Purdue University, Indiana, Estados Unidos.
- Figuroa, A., Kong, S., Tarko A. (2005). "Roadway and Driver Factors of Risk Perception on Four-Lane Highways, International: Road Safety on Four Continents", Warsaw, Poland.
- Figuroa, A., Colucci, B. Arias, W. (2006). Sistema de Gerencia en Seguridad Vial: Integrando la Planificación, el Diseño Geométrico y la Auditoria de las Carreteras Primera cumbre puertorriqueña de seguridad vial, San Juan, Puerto Rico, 2006.
- Gibson, J.J. (1986). *The ecological approach to visual perception*, Hillsdale, NJ: Lawrence Erlbaum.
- Glymour, C., Madigan, D., Pregibon, D., Smyth, P. (1997). *Statistical themes and lessons for data mining*, Data Mining and Knowledge Discovery 1 (1), 11–28.
- Haykin, S. (1999). *Neural Networks: A Comprehensive Foundation*, Macmillan, Nueva York.
- Horberry, T., Anderson, J., Regan, M.A., Triggs, T.J., Brown, J. (2006). "Driver distraction: The effects of concurrent in-vehicle tasks, road environment complexity and age on driving performance", *Accident Analysis and Prevention*, Número 38, pp.185-191.
- Jamson, A.H. (2001). "Image characteristics and their effect on driving simulator validity", *Proceedings of the first international driving symposium on human factors in driver assessment, training and vehicle design*, Aspen, CO. (pp.190-195).
- Karlaftis, M.G, Vlahogianni E.I. (2010). "Statistical methods versus neural networks in transportation research: Differences, similarities and some insights", *Transportation Research Part C*, Volumen 19, Número 3, pp. 387-399.
- Kesting, A., Trieber, M. Helbing, D. (2009). *Agents for traffic Simulation*, Uhrmacher A. & Weyns, D. (Ed.), Multi-agent systems, Simulation and applications, pp. 325-356.
- MATLAB, *User's Manual*, Mathworks Inc, 2017.
- National Center for Statistics and Analysis (NCSA). 2006a. "Race and Ethnicity in Fatal Motor Vehicle Traffic Crashes 1999-2004", National Highway Traffic Safety Administration, Washington, D.C.
- National Center for Statistics and Analysis (NCSA). 2006. "Puerto Rico Toll of Motor Vehicle Crashes, 2005", National Highway Traffic Safety Administration, Washington, D.C.
- National Center for Statistics and Analysis (NCSA). 2005. "Traffic Safety Facts 2004", National Highway Traffic Safety Administration, Washington, D.C.
- Pollatsek, A., Fisher, D.L., Pradhan, A.K. (2006). "Identifying and remediating failures of selective attention in younger drivers", *Current directions in Psychological Science*, pp. 255-259.
- Rumar, K. (1985). *The Role of Perceptual and Cognitive Filters in Observed Behavior. En: Human Behavior in Traffic Safety*, Evans and Schwing, Plenum Press.
- Sadek, A.W., Spring, G., Smith, B.L. (2003). "Toward more effective transportation applications of computational intelligence paradigms", *Transportation Research Record*, Journal of the Transportation Research Board, 1836, 57–63
- Vaa, T. (1997). "Increased police enforcement: effects on speed", *Accident Analysis and Prevention*, Volumen 29, Número 3, pp. 373-85.