



UNIVERSIDAD AUTÓNOMA DEL
ESTADO DE MORELOS



UNIVERSIDAD AUTÓNOMA DEL ESTADO DE MORELOS
FACULTAD DE CIENCIAS QUÍMICAS E INGENIERÍA
MAESTRÍA EN INGENIERÍA ELÉCTRICA Y ELECTRÓNICA

Modelación, Simulación Y Predicción De Los Accidentes Viales En Una Intersección
Mediante Automatas Celulares E Inteligencia Artificial

TESIS

QUE PARA OBTENER EL GRADO DE:

MAESTRO EN INGENIERÍA ELÉCTRICA Y ELECTRÓNICA

PRESENTA:

PABLO HUMBERTO HERNANDEZ PERALTA

ASESORES:

DR. RACHID MARZOUG

DR. OUTMANE OUBRAM

CUERNAVACA MOR.

NOVIEMBRE, 2023

Agradecimientos

A CONAHCYT por el apoyo económico para continuar con mis estudios.

A la Universidad Autónoma Del Estado De Morelos.

A la Facultad De Ciencias Químicas e Ingeniería, por las instalaciones y facilidades de ingreso a laboratorios y oficinas.

Al comité de la Maestría De Ingeniería Eléctrica Electrónica, por su acompañamiento y atención en los procesos administrativos.

Al Dr. Outmane por compartir sus conocimientos y por su enorme paciencia, lo admiro y lo estimo.

Al Dr. Rachid por su gran perseverancia, paciencia y dedicación en la realización del proyecto.

Agradezco todos los docentes que me brindaron acompañamiento durante este proceso.

Dedicatoria

A Dios por colmar de bendiciones a mi familia y seres queridos.

A mis padres J. Belem y Carolina que han estado conmigo apoyándome en todo momento de mi vida, sepan que los amo.

A mis hermanas María de Jesús y Belem Natali por todo su cariño hacia mí, siempre estaré para ustedes.

A todas las personas, amigos y familiares de los cuales he recibido un gran apoyo y han estado cerca de mí.

Resumen

En la actualidad, en diferentes ciudades del mundo, existe una gran variedad de problemas por el uso de los diferentes sistemas de transporte, principalmente el terrestre, los cuales son generados por el aumento del parque vehicular y deficiencias en las vías de tránsito que son rebasadas en capacidad y diseño, esto propicia mayor tiempo de traslado de un punto a otro y al mismo tiempo una alta emisión de gases que contribuyen al efecto invernadero.

En la última década, los científicos del mundo han puesto un mayor interés en la busca de soluciones a esta problemática, destacando los modelos que implementan inteligencia artificial y ciencia de datos para realizar modelos de simulación de distintas condiciones de tráfico vial.

Como resultado del desarrollo de estos modelos, podemos obtener soluciones a las nuevas problemáticas que se manifiestan en los sistemas de transporte, mejorando los tiempos de traslado, disminuyendo las emisiones de gases que causan el efecto invernadero y reduciendo la probabilidad de accidentes viales.

En el presente estudio, se realizó una simulación de los accidentes viales en una intersección utilizando las reglas de Nagel & Schreckenberg llamadas autómatas celulares. Posteriormente se analizaron los datos resultantes para hacer una predicción de los accidentes viales con inteligencia artificial, específicamente las herramientas de aprendizaje automático.

Abstract

Currently, in different cities around the world, there is a wide variety of problems due to the use of different transportation systems, mainly land transportation, which are generated by the increase in the vehicle fleet and deficiencies in traffic routes that are exceeded. In capacity and design, this leads to a longer transfer time from one point to another and at the same time a high emission of gases that contribute to the greenhouse effect.

In the last decade, scientists around the world have placed greater interest in finding solutions to this problem, highlighting models that implement artificial intelligence and data science to create simulation models of different road traffic conditions.

As a result of the development of these models, we can obtain solutions to the new problems that arise in transportation systems, improving travel times, reducing gas emissions that cause the greenhouse effect and reducing the probability of road accidents.

In the present study, a simulation of road accidents at an intersection was carried out using the Nagel & Schreckenberg rules called cellular automata. The resulting data was then analyzed to make a prediction of road accidents with artificial intelligence, specifically machine learning tools.

Contenido

Agradecimientos	II
Dedicatoria	III
Resumen	IV
Abstract	V
Contenido	VI
Lista de figuras	IX
1. Introducción	1
1.1. Planteamiento del problema	4
1.2. Justificación.....	5
1.3. Objetivo general	5
1.5. Hipótesis.....	5
2. Marco teórico	6
2.1. Tráfico Vial	6
2.2. Accidentes viales	6
2.3. Clases de accidentes:	6
2.4. Parámetros y mediciones del tráfico vehicular	8
2.5. Congestión o atasco.....	8
2.6. Sensores	9
2.7. Diagrama Fundamental	12
2.8. Diagrama espacio-tiempo	13
2.9. Clasificación de modelos.....	15
2.10. Seguimiento de automóviles y cambio de carril	17
2.11. Dinámica determinista o estocástica	18
3. Estado del arte	19

3.1. Los modelos de autómatas celulares (AC).....	19
3.2. Análisis de actividad de cambio de carril	20
3.3. Análisis de límites abiertos y cerrados.....	21
3.4. Análisis de accidentes viales	21
3.5. Análisis en intersecciones y rotondas	22
3.6. Análisis de modelos de seguimiento de automóviles.....	23
3.7. Los modelos de Redes Neuronales Artificiales (RNA).....	24
4. Autómatas celulares e inteligencia artificial.....	27
4.1 Autómatas celulares	27
4.2 Límite de frontera abierta y cerrada.....	27
4.3 Modelo de Nagel-Schreckenberg.....	28
4.4. Reglas de actualización de vehículos con autómatas celulares de Nagel y Schreckenberg	29
4.5. Inteligencia artificial.....	31
4.6. ¿Qué es inteligencia artificial?.....	31
4.7. Clasificación de la IA	32
4.8. Aplicaciones de la inteligencia artificial	33
4.9. Aprendizaje Automático (<i>Machine Learning</i>).....	34
4.10. Aprendizaje supervisado.....	34
4.11. Aprendizaje no supervisado.....	35
4.12. Sobreajuste	35
4.13. Evitar el sobreajuste	35
4.14. Aprendizaje por refuerzo	36
4.15. Aprendizaje profundo (<i>Deep learning</i>).....	36
4.16. Regresión Lineal.....	37
4.17. Concepto de interacción	37
4.18. Diagrama de dispersión	37

4.19. Redes Neuronales Artificiales	38
4.20. Modelo biológico red neuronal	39
4.21. Red neuronal artificial	40
4.22. Estructura artificial	41
4.23. Clasificación.....	41
4.24. Clasificación en aprendizaje automático (<i>classification learner</i>)	42
4.25. Interpretación geométrica de la clasificación.....	43
4.26. Algoritmos de clasificación	44
4.27. Elección de número de vecinos k	46
4.28. Métricas de distancia del algoritmo KNN.....	48
4.29. Aplicaciones de KNN en <i>Machine Learning</i>	50
4.30. Evaluación del rendimiento de los modelos de clasificación de <i>machine learning</i>	50
4.31. Matriz de confusión.....	52
5. Metodología	54
5.1 Etapas de generación de una intersección vial.	54
5.2. Modelo de estudio	55
5.3. Base de datos de la simulación con autómatas celulares.....	60
6.1. Proceso de entrenamiento y modelo de red neuronal artificial	62
6.2. Conclusión de entrenamiento con RNA	64
6.3. Entrenamiento con <i>classification learner</i> KNN (k vecinos más cercanos)	64
6.4. Datos de entrenamiento de modelo.....	65
6.5. Evaluación de rendimiento de algoritmo de clasificación.....	66
6.6. Matriz de confusión.....	67
7. Conclusión	69
ANEXO.....	70
Bibliografía	71

Lista de figuras

Figura 1. Intersección vial en zona urbana	7
Figura 2. Gap.	9
Figura 3. Sensor de bucles inductivos.	10
Figura 4. Tubos neumáticos.	10
Figura 5. Sensor infrarrojo.	11
Figura 6. Sensor microondas.	12
Figura 7. Sensor ultrasónico.	12
Figura 8. Diagrama fundamental.	13
Figura 9. Diagrama espacio tiempo.	14
Figura 10. Diagrama espacio tiempo en estado de flujo saturado. (Zeng 2021).	14
Figura 11. Contorno de frontera abierta.	28
Figura 12. Contorno de frontera cerrada.	28
Figura 13. Discretización espacial con un tamaño de 7.5 metros por celda y ocupación de partículas en el modelo de NaSch. El número en la esquina superior derecha es la velocidad del vehículo.	29
Figura 14. Ejemplo paso a paso de la aplicación de las reglas de actualización con $v_{max} = 2$ y $p = 1/3$	31
Figura 15. Modelo de neurona biológica. (Ramírez-Agundis, 2023).	39
Figura 16. Modelo de una neurona artificial. (Rodríguez, 2023).	40
Figura 17. Ejemplo de una red neuronal totalmente conectada. (Matich 2001).	41
Figura 18. Definición de clases en el espacio de variables (Murcia, 2023).	43
Figura 19. Descripción del algoritmo KNN.	46
Figura 20. Comparación de valores de k minoritario y mayoritario (Toth, 2020).	47
Figura 21. Curva ROC con bajo nivel de clasificación.	51
Figura 22. Típica curva ROC.	52
Figura 23. Curva ROC con perfecta clasificación.	52
Figura 24. Matriz de confusión.	53
Figura 25. Etapas de generación de una intersección.	55
Figura 26. Intersección vial en estudio.	56
Figura 27. Relación velocidad $V - \text{Alpha}$	57
Figura 28 Relación $J - \text{Alpha}$	58

Figura 29 Relación densidad $d - Alpha$	59
Figura 30 Relación cambio de carril $Chg - Alpha$	59
Figura 31 Relación probabilidad de accidentes $Pacc - Alpha$	60
Figura 32 Relación de ocurrencia de accidentes contra no accidentes.....	63
Figura 33 Definición de clases en el espacio de variables.	65
Figura 34 Curva ROC del modelo entrenado.	66
Figura 35 Matriz de confusión del modelo KNN.	67

1. Introducción

En los últimos años se ha observado un enorme incremento en el uso de vehículos de transporte desde entornos rurales hasta grandes ciudades, lo que facilita el desplazarse de un lugar a otro en menos tiempo y con mayor comodidad. El transporte terrestre es uno de los medios predilectos de traslado de bienes y pasajeros en el mundo entero, que incorporado a las ciudades, permite la existencia del transporte urbano: autobuses, automóviles, motocicletas y en algunos casos tranvías, trenes o subterráneos, que trasladan a personas de un punto a otro (Editorial, 2021).

Ahora bien, el sistema de transporte terrestre se caracteriza por ser complejo. Este se compone por carreteras, puentes, intersecciones, señalizaciones, semáforos, vehículos, peatones, conductores, entre otros elementos más. Tales elementos de infraestructura vial deben de estar contruidos bajo un análisis profundo de los beneficios y limitaciones que puedan tener a corto, mediano y largo plazo. En cambio, al desplazarse por una carretera urbana es común encontrar escenarios de tráfico que generan un caos y limitaciones al libre tránsito, mencionamos algunos problemas a continuación:

- Intersecciones rebasadas en capacidad de mantener un flujo constante.
- Vialidades que ya no son suficientes para gestionar el tráfico.
- Reductores de velocidad colocados en lugares no necesarios.
- Paradas de transporte publico mal ubicadas.
- Escasas señalizaciones para trayectos adecuados y cortos.
- Semáforos no sincronizados.

La falta de una buena administración vial puede producir consecuencias graves como:

- Accidentes viales.
- Mayores emisiones de gases efecto invernadero.
- Fenómenos de congestión del tráfico.
- Mayor consumo de combustible.
- Retrasos en el tiempo de traslado.
- Contaminación auditiva.

Lo anterior plantea un gran desafío para la construcción, diseño y mantenimiento de infraestructura apta para el tráfico urbano. Actualmente emergen alternativas relacionadas a la transición energética, que al usar combustibles de origen fósil nos alienta cada vez más al uso de los combustibles generados a partir de energía renovable.

Asimismo, surgen limitantes con respecto al consumo energético ya que a medida que la población mundial crece y se urbaniza, incrementa la demanda de energía y transporte, lo que constituye la fuerza motriz para la emisión de sustancias a la atmósfera. También se trata del efecto económico en los usuarios y beneficiarios del transporte, ya sean individuos, instituciones o gobiernos, que requieren realizar un gasto para cubrir sus necesidades de desplazamiento o aprovisionamiento de insumos. De hecho, en el mundo hay 16.5 millones de autos eléctricos (Agency, 2022), no obstante, ha sido insuficiente para disminuir la dependencia de las energías no renovables y esto ha colocado al sector transporte en la cima de la lista con relación al consumo de energía.

En México el sector de transporte ocupa el primer lugar en consumo energético con el 38.44% lo cual revela el balance nacional de energía de 2020 de la secretaria de energía (Energía, 2020). Un factor importante en relación con el gran consumo de energía es su impacto negativo en el medio ambiente, al emitir una gran cantidad de toneladas de dióxido de carbono (CO₂) por los tipos de combustibles utilizados. Con relación al efecto que ocasiona, ser el sector que más energía consume en México da como resultante que el 25% de emisiones de gases efecto invernadero provenga del sector transporte (SEMARNAT, 2018), colocándolo en primer lugar como emisora contaminante.

Sin embargo, en la búsqueda para resolver las constantes deficiencias ocasionadas por el sistema de transporte terrestre, se opta por incrementar el presupuesto público y así combatir el problema en el transporte por tráfico vial, buscando solución a esto, se realizan obras de infraestructura tales como: puentes, autopistas, extensión de carriles, distribuidores viales, asfaltado de caminos rurales, rutas de transporte público, mantenimiento constante de las vías, etc. Lo anterior se traduce directamente a un alto costo de inversión pública y que en ocasiones no son una verdadera solución a la problemática. Además, en algunas ciudades o países en desarrollo no pueden realizar construcciones por escasez de recursos económicos. Entre los países que cuentan con una buena calidad de red vial se encuentran: Singapur,

Países Bajos, Suiza, Hong Kong, Japón. Por otro lado, los países que tienen una baja calidad de infraestructura vial son: Chad, Madagascar, Haití, Angola. Según el indicador de calidad vial, estudio realizado por el foro económico mundial (Forum, 2019).

Una alternativa para hacer más eficientes las vías de transporte terrestre es el modelado y simulación de transporte vehicular, donde se tiene la capacidad de simular parámetros específicos de los vehículos como: tamaño, velocidad, aceleración, frenado, entre otros, tratando a cada vehículo como una partícula independiente que interactúa con las demás partículas (autos), de tal forma se logra analizar algunos fenómenos como: aceleración, velocidad, cambio de carril, flujo, frenado, densidad, además de accidentes, congestionamiento y emisiones contaminantes. En vías como carreteras, autopistas, intersecciones, rotondas, etc. esto es posible lograrlo mediante modelación con autómatas celulares antes mencionadas.

La implementación de modelos que usan inteligencia artificial para predecir fenómenos de tráfico también ha desarrollado muchas aplicaciones entre ellas en el transporte vehicular, como lo han observado en varios modelos como el realizado por R. Yasdi llamado “Predicción del tráfico vial utilizando un enfoque de red neuronal” (Yasdi, 1999), donde tuvieron como objetivo presentar la efectividad de un sistema de redes neuronales para la predicción basada en datos de series de tiempo. Usando un parámetro, a saber, que es el volumen de tráfico para el pronóstico. Emplearon redes neuronales artificiales para la previsión de tráfico aplicada en un tramo de carretera.

Por lo tanto, es de vital importancia, el estudio y análisis de fenómenos relacionados con el tráfico vehicular como el consumo energético, el flujo vial, la problemática de una excesiva emisión de CO₂ y ser de los países con más accidentes viales en Latinoamérica que nos lleve a buscar y generar soluciones para prevenir los conflictos viales y sus derivados.

En la presente investigación, se propone un modelo teórico basado en autómatas celulares y las redes neuronales artificiales para la predicción de accidentes vehiculares en una intersección señalizada. Sin embargo, en el proceso de entrenamiento, no se obtuvieron los resultados esperados, por lo tanto se implementó un algoritmo de aprendizaje automático.

1.1. Planteamiento del problema

Uno de los problemas más graves relacionados con los sistemas de transporte son los accidentes.

La mayoría de los accidentes viales suceden en las intersecciones de las vialidades, esto se deriva a causa de las rupturas a la ley y las faltas al orden vehicular tales como cruzar la calle sin respetar el semáforo, exceder los límites de velocidad, conducir en estado de intoxicación, utilizar el teléfono celular e ignorar las señales de tránsito (Guadarrama, 2019).

Las intersecciones pertenecen a las zonas viales con mayor número de choques en México ya que más de 300 mil accidentes ocurrieron en ellas durante el 2016, lo que equivale al 87 por ciento de los siniestros que se registraron en las áreas urbanas (Mexico, 2017).

México ocupa el séptimo lugar a nivel mundial y el tercero en la región de Latinoamérica en muertes por siniestros viales, con 22 decesos de jóvenes de entre 15 y 29 años al día, convirtiéndolo en la primera causa de muerte de jóvenes en ese rango de edad, y la quinta para la población en general con 24 mil decesos en promedio al año (INSP, 2020).

Si se pone sobre la mesa el tema de la seguridad vial encontramos cifras importantes que nos revela la Organización Mundial de la Salud, donde cada año se pierden aproximadamente 1.3 millones de vidas a consecuencia de los accidentes de tránsito en todo el mundo y más de 20 millones de personas lesionadas, para una población del orden de 7 mil millones de habitantes (OMS, 2022). Las colisiones debidas al tránsito cuestan a la mayoría de los países el 3% de su PIB. La mayoría de esas víctimas ocurre en los países en desarrollo, y las cifras aumentan en la medida en que su parque automotor crece (OMS, 2022).

En 2019, los accidentes viales ocuparon el séptimo lugar dentro de las pandemias que más muertes generan en los países de ingresos bajos y la décima en los países de ingresos medios (OMS, 2020). Superados por las enfermedades del corazón, infecciones de las vías respiratorias inferiores etc. De manera que esta problemática cada vez irá ganando prioridad en relación con las otras causas, por lo que resulta un problema que debe atenderse con mayor interés.

La real tragedia de los accidentes viales es que, en gran medida, ellos y sus consecuencias, pueden evitarse mediante una inversión adecuada y creciente para generar mejores

conductores y usuarios de las vialidades, superiores estándares de diseño, fabricación y mantenimiento vehicular, y mejores estándares de diseño, construcción y mantenimiento carretero.

1.2. Justificación

El análisis de los sistemas viales siempre será de gran importancia, pues con el aumento en la densidad de tráfico también tienden al alza los percances viales, los cuales son un tema serio ya que está en riesgo la vida de las personas involucradas. Para facilitar esta labor en la busca de métodos que eviten este tipo de eventos catastróficos, ha sido de gran ayuda la utilización de diversos programas computacionales y tecnológicos, los cuales implementan en la presente investigación.

1.3. Objetivo general

Desarrollar mediante autómatas celulares e inteligencia artificial el modelo de un programa que simule y prediga los accidentes automovilísticos en una intersección vial.

1.4. Objetivos específicos

- Simular el tráfico vehicular en una intersección señalizada utilizando autómatas celulares.
- Calcular algunos parámetros del tráfico como la velocidad promedio, flujo y densidad vehicular y la probabilidad de accidentes.
- Evaluar el modelo en conjunto de tal forma que permita conocer su eficiencia y los alcances de este.
- Analizar el impacto de algunos parámetros sobre las características del tráfico vehicular.

1.5. Hipótesis

Realizar un modelo que simule y prediga el tráfico y accidentes viales en una intersección es posible mediante autómatas celulares e inteligencia artificial. Para cumplir con el objetivo se propondrá un modelo que permitirá utilizar las dos herramientas para simular y predecir el tráfico y accidentes en una intersección vial.

2. Marco teórico

2.1. Tráfico Vial

El tráfico vial se relaciona con la congestión vehicular, es decir, abundante flujo vehicular en una sola zona de la ciudad y es consecuencia de múltiples factores sociales, culturales, económicos y políticos que se presentan en las principales ciudades del mundo.

Tránsito es la acción de transitar, ir de un lugar a otro por vías o parajes públicos. El concepto suele utilizarse para nombrar al movimiento de los vehículos y las personas que pasan por una calle, una carretera u otro tipo de camino (Pérez Porto, 2022).

2.2. Accidentes viales

Los accidentes viales son aquellos que ocurren sobre la vía y se presenta súbita e inesperadamente, determinado por condiciones y actos irresponsables potencialmente previsibles, atribuidos a factores humanos, vehículos preponderantemente automotores, condiciones climatológicas, señalización y caminos, los cuales ocasionan pérdidas prematuras de vidas humanas y/o lesiones, así como secuelas físicas o psicológicas, perjuicios materiales y daños a terceros.

2.3. Clases de accidentes:

- Colisión: Comprende el choque de uno o más vehículos en movimiento.
- Atropello: Es la acción en la que uno o varios peatones son arrollados por un vehículo en movimiento.
- Los términos colisión, atropello y vuelco son usados para definir una serie de accidentes relacionados entre sí, considerándose para la elaboración estadística, como un solo accidente, de acuerdo al orden de ocurrencia.
- Caída de persona o cosa del vehículo en marcha: Se refiere al caso en que una persona o cosa cae de un vehículo en marcha y esa caída ocasiona daños personales o a la propiedad.
- Accidentes de tránsito fatales: Es todo aquel en el cual una o más personas pierden la vida.

Principales causas de accidentes viales:

1. El uso de teléfonos celulares al conducir.
2. Estar bajo sustancias intoxicantes tales como alcohol o droga.
3. Sobrepasar los límites de velocidad.
4. Agentes climatológicos (lluvia y neblina).
5. No detener en señales de tránsito o luz roja.
6. Escuchar el sistema de audio en alto volumen.
7. Adelantamiento de forma indebida.
8. No mantener distancia de seguridad.
9. Carreteras en mal estado.
10. Fatiga y somnolencia.
11. Calor en la cabina.
12. Estado físico y emocional del conductor.

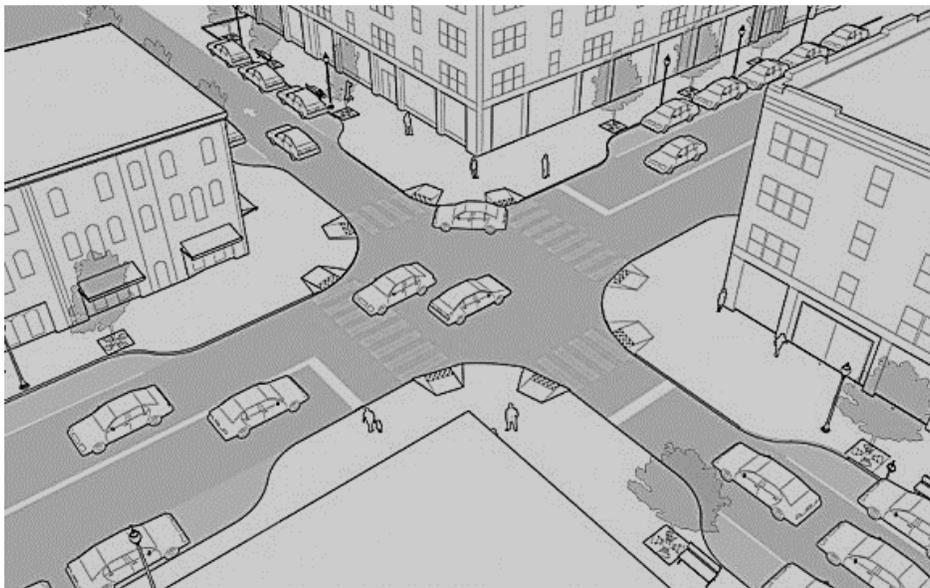


Figura 1. Intersección vial en zona urbana.

Vía: Es toda calle, avenida, camino o carretera (inclusive banquetas o aceras) destinadas para el tránsito de vehículos. Así como también, sitios para el estacionamiento de vehículos.

Vehículo: Es cualquier artefacto en el cual pueden ser transportadas personas o cosas.

Intersección: Es el área en donde se cruzan dos o más vías, como ejemplifica en la figura 1.

2.4. Parámetros y mediciones del tráfico vehicular

Flujo promedio J : durante el intervalo de tiempo T viene dado por:

$$\langle J \rangle = \frac{N(T)}{T} \quad (1)$$

Dónde: $N(T)$ = número de vehículos que pasan por el detector durante ese intervalo de tiempo.

T = intervalo de tiempo

Velocidad v : es la velocidad del vehículo n que pasa de un punto determinado t_{D_1} a un punto t_{D_2} , está dada por la siguiente ecuación:

$$v_n = \frac{d_D}{t_{D_2} - t_{D_1}} \quad (2)$$

d_D = distancia entre los detectores

t_{D_1} = tiempo de llegada al primer inductor

t_{D_2} = tiempo de salida del segundo inductor

Densidad ρ : es el número de vehículos en un tramo fijo de la carretera.

Se determina generalmente a partir del flujo y la velocidad promedio de los vehículos que pasan por el intervalo en el tiempo T .

$$\langle \rho \rangle = \frac{\langle J \rangle}{\langle v \rangle} \quad (2)$$

2.5. Congestión o atasco

La congestión es la condición que prevalece si la introducción de un vehículo en un flujo de tránsito aumenta el tiempo de circulación de los demás. Hasta un cierto nivel de tránsito, los vehículos pueden circular a una velocidad relativamente libre, determinada por los límites de velocidad, la frecuencia de las intersecciones, etc. Sin embargo, a volúmenes mayores, cada

vehículo adicional estorba el desplazamiento de los demás, es decir, comienza el fenómeno de la congestión. A medida que aumenta el tránsito, se reducen cada vez más las velocidades de circulación.

Además, es un estado en el que la densidad es casi máxima y la velocidad es cercana a cero. Se distinguen dos tipos principales de atascos.

Atascos en cuellos de botella: área donde localmente se reduce la capacidad de transporte, esto se produce si el flujo de entrada es mayor que la capacidad de dicha área.

Atascos espontáneos: ocurren sin ninguna razón externa (obvia) normalmente el origen son fluctuaciones locales.

Gap: distancia de separación en celdas de un vehículo a otro (fig. 2).

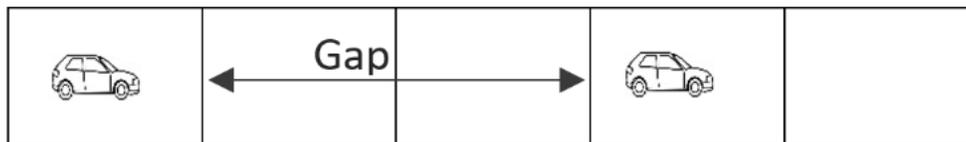


Figura 2. Gap.

2.6. Sensores

Los sistemas de control de tráfico funcionan alimentándose de datos tales como: mediciones de velocidad, tiempo entre vehículos, flujo de vehículos en cada carril y en cada sentido de una vía, densidad de vehículos, tipos de vehículos circulando en la vía etc. Para obtener estos datos se necesitan de sensores que hagan mediciones precisas y confiables. Existen sensores invasivos y no invasivos, los **sensores invasivos** se instalan en el asfalto o dentro del mismo, como lo son: bucles inductivos, sensores magnéticos, tubos neumáticos, cables piezoeléctricos etc.

Sensor de bucles inductivos

Se pueden utilizar en forma temporal o permanente, siendo ésta última la más normal, detectan el paso del vehículo por variación de la masa magnética sobre el lazo, permiten clasificar vehículos, son económicos, son los más comunes en todo tipo de mediciones de tráfico (fig. 3). La instalación permanente requiere cierre de la vía y daña la misma, es necesario varios bucles para obtener información pertinente y son sensibles a la temperatura,

no detectan con detalle el eje, de modo que no pueden diferenciar el eje doble o el eje triple de un eje sencillo. (PIARC, 2023).

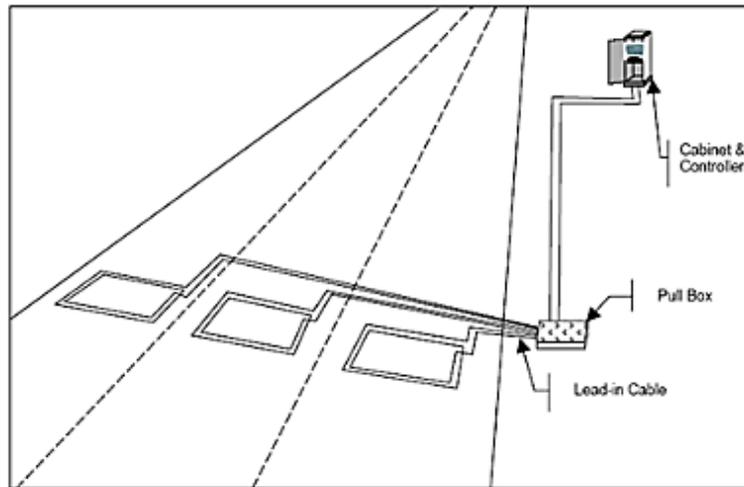


Figura 3. Sensor de bucles inductivos.

Tubos neumáticos

Son sensores de ejes que detectan el paso del vehículo en base al impulso de presión que se genera (fig. 4). Pueden contar y clasificar, pero solo para utilización temporal y en tránsitos fluidos. Cuando la manguera atraviesa varios carriles debe estar protegida en los carriles en los que no esté efectuando la medida. Hay que cuidar también la fijación de la manguera a la calzada para evitar que los vehículos la arrastren o la rompan.

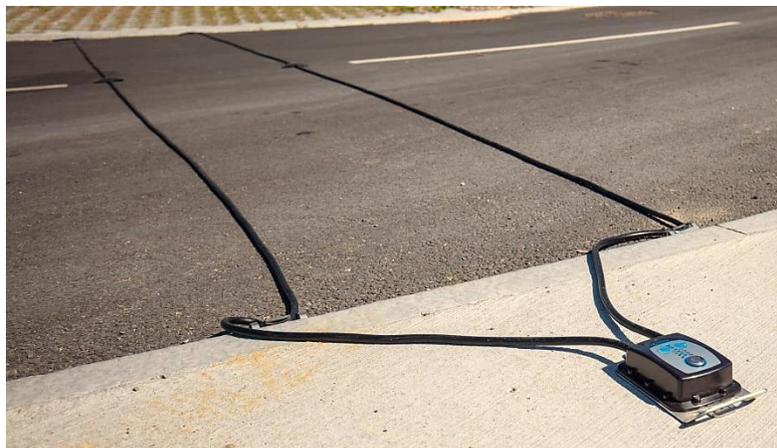


Figura 4. Tubos neumáticos.

Los **sensores no invasivos** cambian en la ubicación, predomina la altura este tipo de sensores. Su instalación es más sencilla y flexible, se pueden cubrir varios carriles con el mismo sensor

y son más económicos con respecto a los sensores invasivos, los que trabajan con propagación de ondas son más susceptibles a condiciones climáticas, como se ve en los siguientes ejemplos:

Sensor infrarrojo

Es sensible a la luz infrarroja ya sea producida por la reflexión de un haz viniendo de una fuente (sensores activos) o producida por el vehículo mismo (sensores pasivos). El sensor proporciona la siguiente información: longitud, alto y ancho del vehículo, posición respecto al acotamiento, velocidad (ver fig. 5). La desventaja de estos sensores es que son sensibles a la luz ambiente y por lo tanto a condiciones climáticas (PIARC, 2023).

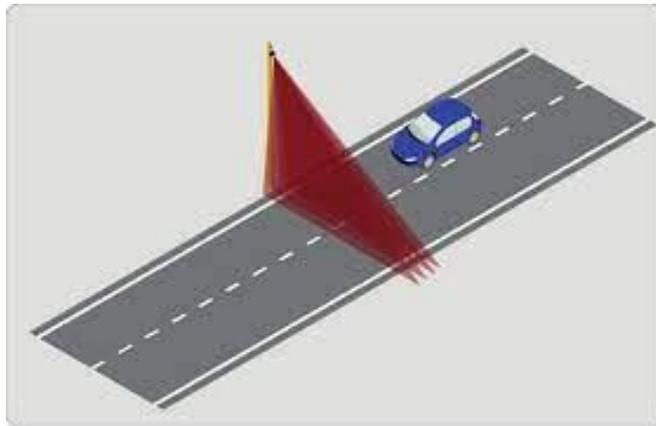


Figura 5. Sensor infrarrojo.

Sensor tipo cámara

Este tipo de sensores proporciona una mayor cantidad de información, su instalación es muy flexible y cuenta con facilidad de uso, además de medir amplias zonas de la vía, son características que han hecho que este tipo de sensores sean utilizados en la actualidad. Al final, el uso de cámaras representa la solución económicamente más eficiente.

Sensor tipo microondas

Basado en tecnología de microondas convencional determina la distancia a la que se encuentra un objeto, en movimiento o estacionado, en función del tiempo entre la emisión y la recepción de la onda reflejada (fig. 6).

Este tipo de sensores son sensibles al ruido ambiental, a la velocidad de los vehículos y a condiciones climatológicas.

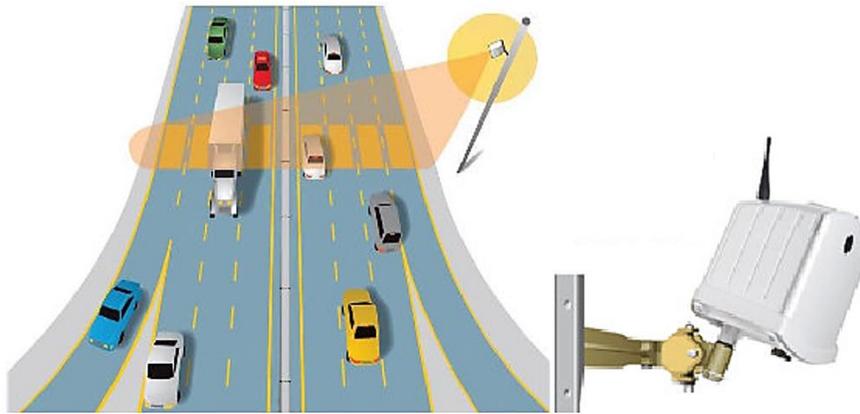


Figura 6. Sensor microondas.

Sensor de tipo ultrasonido

Consiste de la emisión de onda ultrasonido y medición de la recepción. Para la medición de velocidad por el efecto Doppler, se realiza comparando la frecuencia de la onda emitida con la frecuencia de la onda reflejada. Este tipo de sensores son sensibles a condiciones climáticas. (fig. 7).

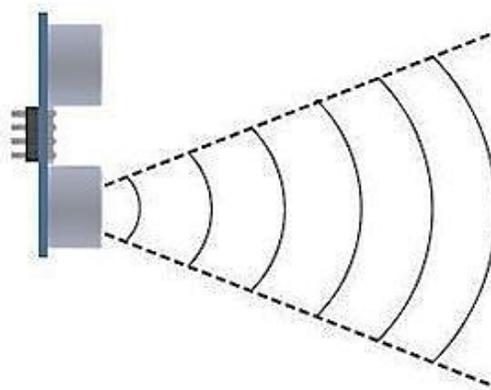


Figura 7. Sensor ultrasónico.

2.7. Diagrama Fundamental

Es la característica cuantitativa más importante del flujo de tráfico, es la relación $J(\rho)$ entre la densidad de tráfico ρ y el flujo de tráfico J .

El diagrama fundamental captura las características básicas, lo cual consta de dos ramas lineales y tiene forma triangular. La forma del diagrama muestra que, a mayor cantidad de vehículos en el camino, menor velocidad promedio de movimiento de estos.

El producto de la densidad de vehículos y la velocidad promedio es la cantidad de flujo promedio. El diagrama fundamental comienza en el origen, (cuando la densidad es cero, lo es también el flujo) y tiene al menos un máximo como se observa en la figura 8.

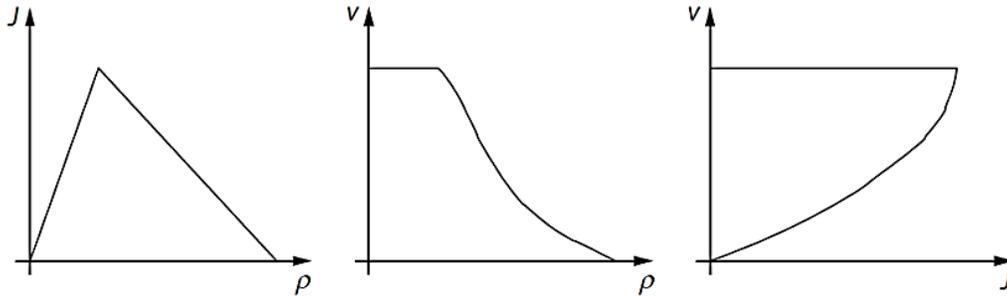


Figura 8. Diagrama fundamental.

- a) Relación entre flujo y densidad $J(\rho)$
- b) Relación entre velocidad y densidad $v(\rho)$
- c) Relación entre velocidad y flujo $v(J)$

El diagrama fundamental consiste en dos curvas que son representadas por intervalos. La región del diagrama fundamental correspondiente a $0 \leq \rho < \rho_{max}$ es conocida como la región estable o de flujo laminar, debido a que la densidad de autos es lo suficientemente baja que prácticamente no existe interacción entre ellos y por tanto se da un “flujo libre”. La región correspondiente a $\rho_{max} \leq \rho \leq 1$ (ó 100%) es llamada región inestable o de congestión. En ésta, la cantidad de autos en el camino provoca un flujo generalmente más lento. El punto con densidad ρ_{max} indica cuando un camino se encuentra cerca de un congestionamiento (Maerivoet, 2005).

2.8. Diagrama espacio-tiempo

En la figura 9, se observan las trayectorias de tres vehiculos en el tiempo t a $t + 6$. Se puede apreciar que el auto azul se detiene en un paso de tiempo, afectando la trayectoria libre de coche verde y amarillo que vienen atras, generando un pequeño atasco.

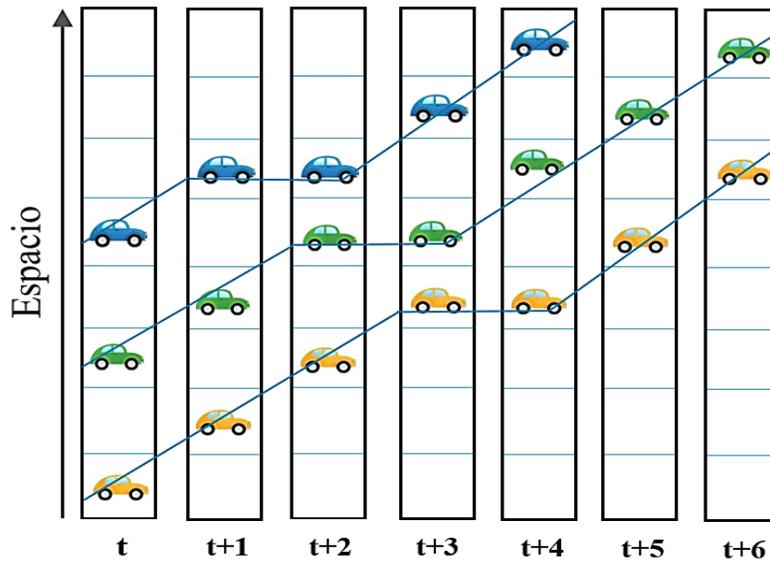


Figura 9. Diagrama espacio tiempo.

Se observa en la figura 10, que cada línea blanca punteada representa la trayectoria de un vehículo en un carril de la carretera. El gráfico de espacio-tiempo (es decir, las trayectorias $x(t)$ de los vehículos) muestra la formación y propagación de un embotellamiento.

Se observa que al inicio del diagrama las líneas punteadas están separadas unas de otras, conforme van avanzando en el espacio aumenta la densidad hasta que en la celda 800 disminuye el movimiento casi totalmente, que puede ser causado por algún obstáculo para los vehículos.

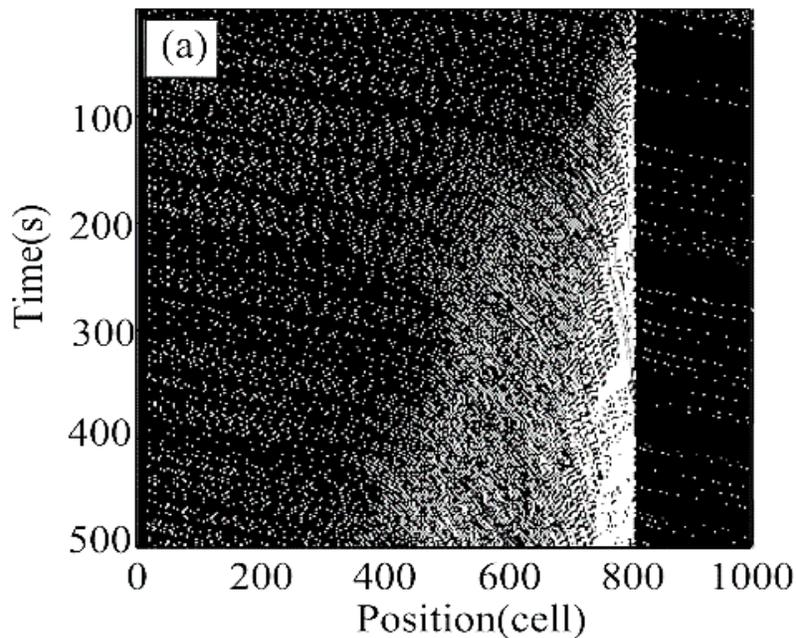


Figura 10. Diagrama espacio tiempo en estado de flujo saturado. (Zeng 2021).

Tal embotellamiento puede propagarse hacia atrás, en dirección opuesta a la circulación de los vehículos. Además, es posible que dos o más atascos coexistan en una carretera.

2.9. Clasificación de modelos

Hay diferentes formas de clasificar los modelos de tráfico vehicular, como son los modelos microscópico, macroscópico y mesoscópico. Los modelos microscópicos describen el comportamiento espaciotemporal de cada partícula o vehículo y sus interacciones, mientras que los modelos macroscópicos describen propiedades colectivas sin distinguir partículas individuales, por otra parte los modelos mesoscópicos tienen características de ambos enfoques. Estos no distinguen ni rastrean partículas individuales, pero especifican el comportamiento individual, por ejemplo, en términos de distribuciones de probabilidad (Andreas Schadschneider, 2011).

2.9.1. Modelo Macroscópico

Utiliza el método de la mecánica de fluidos para estudiar el comportamiento promedio de los vehículos, ignorando las características de un solo vehículo, por lo que estos modelos no son muy convenientes para describir dinámicas discretas, tales como, la interacción individual entre vehículos. La principal característica de este enfoque es el flujo global, en donde los vehículos son vistos como partículas a través de un ducto (Morales Fernandez, 2010).

Un estudio de análisis de tráfico ha sido realizado por Payne, donde muestra en que analiza un modelo en la categoría macroscópica que es particularmente útil para evaluar las operaciones en las autopistas. El modelo se describe con detalles matemáticos para la simulación básica de flujo, medición y desvío de rampa, vigilancia y representación de incidentes en autopistas. También se detalla el cálculo de las medidas de desempeño (Payne, 1979).

Otro modelo donde se analiza el flujo de tráfico a nivel macroscópico es el realizado por Kai Nagel, el cual es llamado modelo fluido-dinámico para flujo de tráfico, en él se muestra cómo los modelos de salto de partículas encajan en el contexto de la teoría del flujo de tráfico. Se muestran las conexiones entre los modelos de flujo de tráfico fluidodinámico, que se derivan de las ecuaciones de Navier-Stokes, y los modelos de salto de partículas (Nagel, Particle hopping models and traffic flow theory, 1996).

Dentro de los modelos macroscópicos encontramos el realizado por Fernando Pérez, donde introduce un modelo simple para describir el flujo vehicular en un congestionamiento de tráfico y así obtener la ecuación que relaciona el comportamiento característico entre flujo vehicular y densidad de tráfico, conocida en la literatura como relación fundamental (Fernando Perez, 2014).

2.9.2. Modelo microscópico

Este modelo estudia las características del flujo de tráfico analizando la interrelación entre vehículos individuales, el flujo es tratado de manera individual a los vehículos. De manera que la dinámica del tráfico y su desempeño estarán determinados por la manera en que cada vehículo influye sobre el movimiento de los demás.

El modelo básico de autómatas celulares que marco un punto de partida para posteriores investigaciones en el flujo de tráfico microscópico, es el de Nagel-Schreckenberg que, a través de una serie de pasos, reflejan las propiedades muy generales del tráfico de un solo carril que se modelan sobre la base de reglas de autómatas celulares con valores probabilísticos de valor entero. Este modelo simple ya muestra un comportamiento no trivial y realista. El tercer paso es esencial para simular un flujo de tráfico realista ya que, de lo contrario, la dinámica es completamente determinista. Tiene en cuenta las fluctuaciones naturales de la velocidad debidas al comportamiento humano o a las variaciones de las condiciones externas. Sin esta aleatoriedad, cada configuración inicial de vehículos y velocidades correspondientes alcanza muy rápidamente un patrón estacionario que se desplaza hacia atrás (es decir, opuesto al movimiento del vehículo) un sitio por paso de tiempo (Nagel & Schreckenberg, 1992).

Encontramos también un modelo realizado por J. Esser et al., donde se presenta una herramienta de simulación de tráfico para redes de carreteras urbanas que se basa en el modelo Nagel-Schreckenberg. Los tipos arbitrarios de caminos y cruces se modelan como combinaciones de solo unos pocos elementos básicos. Además, se consideran las capacidades de estacionamiento y las circulaciones de los transportes públicos (Esser, 1997).

Un modelo microscópico más que utiliza autómatas celulares es de Das, AK y Chattaraj, U. donde pretenden explicar con precisión la actividad de cambio de carril, que se trata como un proceso continuo en lugar de un evento discreto, como sugerían los modelos anteriores. El

modelo propuesto podría explicar varias propiedades importantes del flujo de tráfico, y los resultados simulados son bastante razonables. Tal modelo puede dar una explicación a las propiedades microscópicas, como la estabilidad local, la estabilidad asintótica, el acercamiento/rechazo y la insensibilidad del avance de la distancia segura al patrón de perturbación, el avance de la distancia y la velocidad iniciales (AK Das, Cellular Automata Model for Lane Changing Activity, 2022).

2.9.3. Modelo mesoscópico

Un modelo mesoscópico no distingue ni rastrea vehículos individuales, sino que especifica el comportamiento de los individuos, es decir en términos probabilísticos. Con este fin, el tráfico está representado por (pequeños) grupos de entidades de tráfico, cuyas actividades e interacciones se describen con un bajo nivel de detalle. Por ejemplo, una maniobra de cambio de carril, podría representarse para un vehículo individual como un evento instantáneo, donde la decisión de realizar un cambio de carril se basa en densidades de carril relativas y diferenciales de velocidad. Algunos modelos mesoscópicos se derivan en analogía con la teoría cinética de gases. Estos llamados modelos cinéticos de gas describen la dinámica de las distribuciones de velocidad (Hoogendoorn, 2001).

2.10. Seguimiento de automóviles y cambio de carril

El movimiento del vehículo también se divide en dos partes: seguimiento de automóviles y cambio de carril. En cuanto al modelo de seguimiento de automóviles, se puede dividir en modelo discreto y modelo continuo. El modelo discreto típico es el modelo de autómatas celulares, Nagel y Schreckenberg (NaSch) desarrollado en 1992 (Nagel & Schreckenberg, A cellular automaton for freeway traffic, 1992) donde establecieron reglas simples de actualización de vehículos como una velocidad máxima, una probabilidad de frenado, una aleatorización a través de autómatas celulares, por lo tanto cada vehículo se describe por medio de una ecuación de movimiento propia para todos los vehículos por lo que las características individuales de estos producen grandes efectos en el comportamiento global (Fernandez, 2010).

2.11. Dinámica determinista o estocástica

La dinámica de las partículas puede ser **determinista** o **estocástica**. En el primer caso, el comportamiento en un momento determinado está completamente determinado por el estado actual. En los modelos estocásticos, el comportamiento está controlado por ciertas probabilidades, de modo que las partículas pueden reaccionar de manera diferente en la misma situación.

3. Estado del arte

3.1. Los modelos de autómatas celulares (AC)

Estos modelos son sistemas dinámicos abstractos, cuyos estados se describen en términos de variables discretas y viven en un espacio-tiempo discreto. La dinámica de un autómata celular se constituye a través de un conjunto de reglas de actualización, locales y homogéneas, que son aplicadas en forma paralela y simultánea a los sitios de la retícula espacial. En cada paso temporal, el nuevo estado del autómata en cada sitio de la retícula se determina mediante la aplicación de las reglas de actualización a los estados presentes del sitio mismo y de los sitios dentro de alguna vecindad pequeña de éste (Salcido, 2015).

En la década de 1930 comenzó el estudio de la teoría del flujo de tráfico con Bruce D. Greenshields (Greenshields, 1935) quien utiliza la teoría de la probabilidad y las estadísticas matemáticas para describir la relación entre la velocidad, el flujo de tráfico y la densidad. Cercanos a 1950 se desarrolló la planificación y gestión del tráfico por el aumento de este.

Debido a las características discretas del modelo de autómatas celulares, se puede simular efectivamente el estado de movimiento microscópico de los vehículos en el flujo de tráfico. En los últimos años, se ha convertido en un centro de investigación en el campo del transporte. Los primeros modelos de flujo de tráfico establecidos por autómatas celulares incluyen el modelo ya antes mencionado de NaSch (Nagel & Schreckenberg, A celular automaton for freeway traffic, 1992), el modelo Fukui-Ishibashi (M Fukui, 1996) entre otros.

En 1992, Kai Nagel y Michael Schreckenberg (Nagel & Schreckenberg, 1992), presentan un modelo para el tráfico de vehículos en una autopista de un solo carril. En el modelo, divide el tiempo en pequeños intervalos iguales y en cada uno de ellos, se actualizan las posiciones y las velocidades de los vehículos simultáneamente (Salat Figols, 2014). Cada vehículo puede estar en reposo o moviéndose, saltando de un sitio a otro sobre una retícula de una dimensión, con una rapidez de valores enteros positivos hasta un valor máximo dado. Las reglas de actualización del autómata son probabilísticas y controlan los eventos de propagación, aceleración y frenado de los vehículos, conservando el número de éstos y evitando las colisiones y que los vehículos se rebasen unos a otros (Salcido, 2015).

En 1997 D Chowdhury , De Wolf, M Schreckenberg desarrollaron modelos de tráfico de dos carriles con salto de partícula (adelantamiento) con dos tipos de vehículo grande y compacto, con dos valores diferentes de velocidad máxima, tomado como referencia el modelo básico de NaSch. Mencionan que el modelo más simple es el simétrico con respecto a los dos carriles donde gobiernan diferentes reglas de cambio del carril “rápido” al “lento” y viceversa. En el modelo asimétrico los conductores de vehículos rápidos pueden anticipar a menudo con mucha posibilidad de quedar atascados detrás de un vehículo lento (D Chowdhury, 1997).

En 2003, K. Daoudia y N. Moussa, estudiaron un modelo de AC de tres carriles basado sobre el modelo de NaSch (D-M) y analizaron el efecto de los vehículos lentos en el sistema. Su objetivo fue señalar los parámetros que definen la forma del diagrama fundamental y compararlo con el de dos carriles (AK Daoudia, 2003).

En 2007 A Aponte, JA Moreno, realizaron un estudio de un modelo donde incluyen un solo tipo de vehículos y las velocidades de los mismos se asignan a través de una distribución gaussiana, de manera de tomar en cuenta que no todos los conductores acatan los límites de velocidad. Validan el modelo mediante la comparación entre los valores de flujo vehicular simulados y medidos empíricamente por medio de contadores mecánicos (A Aponte, 2007).

3.2. Análisis de actividad de cambio de carril

W. Knospe L. Santen, generalizaron el modelo NaSch, a dos carriles adyacentes e independientes, y a este le agregaron reglas que permiten la simulación del cambio de carril, su objetivo es señalar los parámetros que definen la forma del diagrama fundamental (W Knospe, 2002).

En 2021 Wang Jinghui et al. Realizaron un modelo de autómatas celulares basado en múltiples agentes para la simulación de control en intersecciones, proponen reglas de seguimiento de automóviles y cambio de carril donde las aplican en diferentes escenarios de tráfico. Muestran resultados donde el efecto de control de tráfico con AC es mejor que la estrategia tradicional de control de tiempo fijo (J Wang, 2021).

En 2022 AK Das, U Chattaraj, en su modelo pretenden explicar con precisión la actividad de cambio de carril, donde lo tratan como un proceso continuo en lugar de un evento discreto

como sugieren otros modelos. Los resultados de la simulación de un solo carril coinciden con los valores sugeridos por el manual de capacidad de carreteras (AK Das, 2022).

3.3. Análisis de límites abiertos y cerrados

En 2000 S Cheybani, J Kertesz, M Schreckenberg, estudiaron las fases del modelo de tráfico de Nagel-Schreckenberg con condiciones de frontera abiertas en función de las probabilidades de aleatorización $p_{ags} > 0$ y la velocidad máxima $v_{max} > 1$. Debido a la existencia de "sitios de amortiguamiento" que mejoran la región de flujo libre (S. Cheybani, 2000).

En 2022 MK Singh, K Ramachandra Rao, intentan evaluar la diferencia en los resultados de simulaciones de límites abiertos y cerrados en tráfico heterogéneo no basado en carriles. La metodología usada en su estudio se relaciona con las condiciones de campo. En su estudio incluyen algunas de las características comunes, como la filtración de vehículos pequeños. Además, incluyen algunas de las características desapercibidas al modelar el tráfico no basado en carriles en las intersecciones. El modelado de la simulación de límites abiertos puede ser mejor y más fácil en la mayoría de las situaciones en comparación con los límites cerrados (MK Singh, 2022).

3.4. Análisis de accidentes viales

En 1997 Boccara et al., estudian un modelo simple de circulación de automóviles en una autopista de un solo carril, analizan la probabilidad de ocurrencia de accidentes de tránsito cuando los conductores no respetan la distancia de seguridad entre automóviles, y, como consecuencia del bloqueo durante el tiempo T necesario para despejar la carretera, determinaron el número de automóviles detenidos en función de la densidad de automóviles. Dieron una teoría simple en buen acuerdo con sus simulaciones numéricas (N Boccara, 1997).

En 2004 Yang et al., estudiaron numéricamente el impacto del desorden extinguido inducido por accidentes automovilísticos en el flujo de tráfico en el modelo de Nagel-Schreckenberg (NS). Explican que los accidentes automovilísticos ocurren cuando se dan las condiciones necesarias propuestas por Boccara *et al.* Se cumplen. Consideraron dos situaciones realistas de automóviles involucrados en accidentes automovilísticos. Modelo A presenta a considerar

que los autos accidentados quedan atascados temporalmente. Sus estudios exhiben la "inversa- λ form" o el estado metaestable para el flujo de tráfico en el diagrama fundamental y ondas de atascos de amplio movimiento en el patrón de espacio-tiempo. Modelo B propone tener en cuenta que los autos "siniestrados" se quedan ahí para siempre y los autos de atrás pasarán por los sitios ocupados por los autos "siniestrados" con una tasa de transmisión (Yang, 2004).

En 2015 Lakouari et al., estudiaron el sistema de rampa de entrada bajo la condición de frontera abierta expandida. Calcularon el diagrama de fase del sistema de rampa de entrada de dos carriles. Encontraron que la estrategia de inserción del límite izquierdo expandido mejora el flujo en el carril de rampa de acceso. Además, estudiaron la probabilidad de ocurrencia de accidentes automovilísticos. Distinguieron dos tipos de accidentes automovilísticos: el accidente en la vía de acceso (P_{rc}) y el accidente por alcance en la vía principal (P_{ac}). Se muestra que es más probable que ocurran accidentes automovilísticos en el sitio de la vía de acceso cuando el tráfico es libre en la carretera A. Sin embargo, los accidentes en la parte trasera comienzan a ocurrir por encima de una tasa de inyección crítica α_{c1} (Echab, 2015).

En 2017 Pang et. al., investigaron la influencia del clima lluvioso en los accidentes de tránsito de una autopista. Analizaron los comportamientos de conducción a micro escala en clima lluvioso y los posibles accidentes de colisión trasera y lateral del vehículo. Presentaron un modelo AC mejorado de autopista de un solo sentido de dos carriles, donde ocurrirán algunos accidentes vehiculares cuando se cumplan simultáneamente las condiciones necesarias. Discutieron las características del flujo de tráfico bajo diferentes intensidades de lluvia y analizaron las probabilidades de accidente a través de los experimentos de simulación utilizando el límite de velocidad variable (VSL) y el control de flujo entrante. Los resultados indican que las medidas son efectivas especialmente durante fuertes tormentas o lluvias intensas de corta duración (Pang, 2017).

3.5. Análisis en intersecciones y rotondas

En 2015 C. Chai, Y.D. Wong, realizaron un estudio sobre los impactos de seguridad en intersecciones con cámaras de luz roja basado en modelos de autómatas celulares, su objetivo fue evaluar el impacto de las cámaras de luz roja en varios tipos de conflicto en diversas

condiciones de tráfico en intersecciones, creados por simulación con autómatas celulares (C Chai, 2015).

En 2017 Han-Tao Zhao et al. Realizaron un modelo donde tomaron en cuenta el efecto de la velocidad del vehículo delantero, la influencia de la luz de freno y el conflicto del flujo de tráfico, establecieron un modelo de autómatas celulares llamado CE-NS para el flujo de tráfico en la intersección en el entorno de red sin vehículos (Han Tao Zhao, 2017).

En 2018 Younes Regragui, Najem Moussa proponen un modelo bidimensional de autómatas celulares (CA), donde todas las calles son de dos sentidos, con un carril en cada dirección. Diseñan una rotonda para cada intersección donde confluyen cuatro caminos. El comportamiento de giro de los conductores lo modelan con un parámetro γ . Los resultados de la simulación muestran que el tráfico urbano está en estado de flujo libre cuando la densidad del vehículo es lo suficientemente baja. Sin embargo, cuando la densidad excede una densidad crítica ρ_c , el tráfico urbano estará en estado de embotellamiento.

3.6. Análisis de modelos de seguimiento de automóviles

Los primeros modelos matemáticos de seguimiento de automóviles que se desarrollaron se basaron en una descripción de la interacción entre dos vehículos vecinos en un flujo de tráfico, es decir, un seguidor y su líder.

En 2020 Bingmei Jia et al. Proponen un nuevo modelo de seguimiento de automóviles considerando el efecto de prevención del cambio de carril. La velocidad final del siguiente vehículo es una combinación de una velocidad segura y una velocidad de prevención de cambio de carril. La condición de estabilidad del modelo se deriva y verifica mediante simulación numérica, y analizan los impactos de varios factores en la estabilidad (Jia, 2020).

En 2019 Lee et al. Proponen un procedimiento estocástico novedoso para modelar comportamientos de seguimiento de automóviles en una autopista de varios carriles. Desarrollan un modelo integrado de seguimiento continuo estocástico de varios carriles en el que se utiliza una arquitectura de aprendizaje profundo para estimar la probabilidad de maniobras de cambio de carril (LC) (Lee, 2019).

3.7. Los modelos de Redes Neuronales Artificiales (RNA)

Los modelos RNA utilizan un enfoque estadístico no lineal para modelar relaciones o patrones complejos entre variables independientes y dependientes. De manera similar a cómo funciona el cerebro humano, las RNA adopta un enfoque no algorítmico para procesar las señales de entrada y salida en lugar de seguir un conjunto predefinido de instrucciones (Bailer-Jones, 2002).

Las RNA normalmente constan de tres capas:

- Capa de entrada
- Capa oculta
- Capa de salida

Primero, la capa de entrada procesa los vectores de entrada normalizando cada uno de los vectores de entrada antes de que se introduzcan en la capa oculta. Luego, la capa oculta aplica pesos y sesgos a la señal recibida de la capa de entrada. La RNA sigue ajustando esos pesos y sesgos durante el entrenamiento para mejorar la función de rendimiento.

El proceso de entrenamiento se puede detener debido a diferentes criterios, como no observar ninguna mejora en los objetivos de entrenamiento (es decir, la función de rendimiento) o alcanzar el número máximo de épocas (es decir, iteraciones de entrenamiento). Finalmente, la capa de salida finalmente transfiere la señal que recibe de la capa oculta al resultado del modelo ANN (Mohammad Shareef Ghanim, 2019).

En 2021 Olayode et al. Estudiaron la aplicación de una red neuronal artificial entrenada por optimización de enjambre de partículas (ANN-PSO) para analizar el problema de la congestión del tráfico. Las variables de flujo de tráfico, como la velocidad de los vehículos en la carretera, el número de diferentes categorías de vehículos, la densidad del tráfico, el tiempo y los volúmenes de tráfico, fueron variables de entrada y salida para modelar el flujo de tráfico de vehículos no autónomos en una intersección vial señalizada. Los resultados de su investigación mostraron un rendimiento de entrenamiento y prueba de 0.98356 y 0.98220 (Olayode, 2021).

En 2019 Mohammad Sharif Ghanim et al. Entrenan un modelo de redes neuronales artificiales (RNA) para analizar la relación entre los volúmenes de aproximación y los

movimientos de giro correspondientes. Los resultados muestran que el modelo RNA desarrollado se puede utilizar para predecir movimientos de giro con un alto nivel de precisión (Mohammad Shareef Ghanim, 2019).

En 2019 A Ata et al. Utilizaron Redes Neuronales Artificiales (RNA) para predecir la congestión del tráfico y así controlar o minimizar los atascos y suavizar el tráfico vial. La propuesta de modelado inteligente para el control de la congestión del tráfico vial utilizando redes neuronales de retropropagación artificial (MSR2C-ABPNN) disminuye la congestión vial. En este artículo, la predicción de la congestión se operacionaliza mediante el uso del algoritmo de retropropagación para entrenar la red neuronal (Ata, 2019).

En 2018 García et al. Realizaron un modelo de RNA para predecir los accidentes en las carreteras nacionales de Suiza con resultados confiables. Examinan los datos disponibles de los accidentes previos de 2009 a 2012, para desarrollar modelos utilizando redes neuronales artificiales (RNA) y establecen un proceso sistemático de optimización de RNA para determinar la arquitectura óptima del modelo. Se ha descubierto que las RNA pueden utilizarse como un método viable para predecir la frecuencia de los accidentes de tráfico. Dado que los accidentes son eventos relativamente raros, los datos se caracterizan por una gran parte de cero observaciones (García de Soto, 2018).

En 2016 Sharaf Aljeder et al. Utilizaron una red neuronal artificial (RNA) para predecir la gravedad de las lesiones de los accidentes de tránsito en base a 5973 registros de accidentes de tránsito ocurridos en Abu Dhabi durante un periodo de 6 años (2008 a 2013). Los resultados experimentales revelaron que los clasificadores RNA desarrollados pueden predecir la gravedad de los accidentes con una precisión razonable (Sharaf Aljeder, 2016).

En 2010 Rocha et al. Diseñaron un sistema de detección basado en redes neuronales y momentos invariantes capaz de adaptarse y entrenarse con ciertas señales de tránsito con la meta de asistir al conductor de no cometer una infracción o en el peor de los casos un accidente, la red neuronal será capaz de reconocer una señal de tránsito a cierta distancia para que así el conductor a priori tenga el conocimiento de esta (Rocha & Escorcía G., 2010).

En 2010 Akin et al. Desarrollaron un modelo de red neuronal (RN) para predecir choques en intersecciones, la capacidad predictiva del modelo RN la determinaron agrupando los

choques en estos tipos: accidentes fatales, con lesiones y daños a la propiedad solamente. El enfoque de RN se usó para desarrollar y probar RN de avance de múltiples capas entrenados con el algoritmo de retropropagación para modelar la relación no lineal entre los tipos de choque y las propiedades del choque, como el tiempo, el clima, la luz y las condiciones de la superficie, el conductor y el vehículo. Demostraron que el enfoque adaptado en este estudio es capaz de proporcionar una predicción muy precisa (90,9 %) de los tipos de accidentes (Darcin Akin, 2010).

En 1999 R. Yasdi et al. Presento la efectividad de un sistema de redes neuronales para la predicción basada en datos de series de tiempo. Solo se dedicó a usar un parámetro, a saber, el volumen de tráfico para el pronóstico. Emplearon redes neuronales artificiales para la previsión de tráfico aplicada en un tramo de carretera (Yasdi, 1999).

4. Autómatas celulares e inteligencia artificial

4.1 Autómatas celulares

Los autómatas celulares son modelos discretos en todas las variables. Son modelos microscópicos y genéricamente la dinámica es reglada y estocástica. En las simulaciones por computadora, generalmente se realiza un tiempo discreto a través de la actualización paralela o síncrona donde todas las partículas se mueven al mismo tiempo. Esto introduce una escala de tiempo. Los autómatas celulares (CA) son idealizaciones de sistemas físicos en los que se supone que tanto el espacio como el tiempo son discretos y cada una de las unidades que interactúan puede tener solo un número finito de estados discretos. Debido a que, por lo general, los modelos CA tienen una dinámica basada en reglas, son adecuados para modelar problemas de sistemas de tráfico. Hasta donde sabemos, el primer modelo CA para tráfico vehicular fue introducido por Cremer y Ludwig (Cremer, 1986), pero el modelo ahora fundamental fue introducido en 1992 por Nagel y Schreckenberg (Nagel & Schreckenberg, 1992).

En los modelos CA de tráfico, la posición, la velocidad, la aceleración y el tiempo se tratan como variables discretas. En este enfoque, un carril se representa mediante una red unidimensional. Cada uno de los sitios de la red representa una celda, que puede estar vacía u ocupada como máximo por un vehículo en un instante dado (ver Fig. 4). En cada paso de tiempo discreto $t \rightarrow t + 1$, el estado del sistema se actualiza siguiendo una prescripción bien definida. La eficiencia computacional de los modelos CA discretos es la principal ventaja de este enfoque sobre los enfoques de celosía de mapas acoplados y de seguimiento de automóviles.

4.2 Límite de frontera abierta y cerrada

4.2.1. Límite de frontera abierto

Las partículas que entran en el circuito salen y no vuelven a entrar. En los sistemas abiertos los vehículos van ingresando y saliendo del arreglo según una regla dada (fig. 11). Los vehículos entran con una probabilidad de ingreso (o de inyección) α cuando la primera celda del vector está vacía y salen con una probabilidad de salida (o extracción) β . Para que los vehículos salgan de la vialidad, estos deben de tener la velocidad suficiente y la probabilidad de salida debe de ser menor a β (Canul, 2022).



Figura 11. Contorno de frontera abierta.

4.2.2. Límite de frontera cerrada

El inicio y fin del circuito están conectados, las partículas que salen vuelven a entrar, en los sistemas cerrados (fig. 12) los vehículos que se encuentren cerca del final del arreglo y tengan velocidad suficiente para salir de este son regresados al inicio del arreglo o a una celda que emule la distancia recorrida por el vehículo. La densidad de vehículos se mantenga estable.

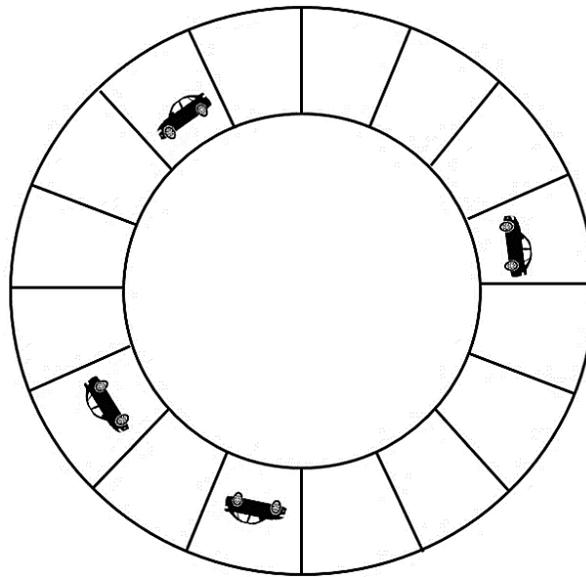


Figura 12. Contorno de frontera cerrada.

4.3 Modelo de Nagel-Schreckenberg

Debido a las características discretas del modelo AC, puede simular efectivamente el estado de movimiento microscópico de los vehículos en el flujo de tráfico. En los últimos años, se ha convertido en un centro de investigación en el campo del transporte. Los primeros modelos de flujo de tráfico establecidos por autómatas celulares incluyen el modelo NaSch, el modelo Fukui-Ishibashi (M Fukui, 1996).

En los modelos de tráfico de CA, la posición, la velocidad, la aceleración, y el tiempo se tratan como variables discretas. Donde un carril está representado por una celosía unidimensional. Cada sitio de la celosía (fig. 13) representa una celda que puede estar vacía

u ocupada por un vehículo como máximo. En cada paso de tiempo discreto $t + 1$, el estado del sistema se actualiza siguiendo una prescripción bien definida.

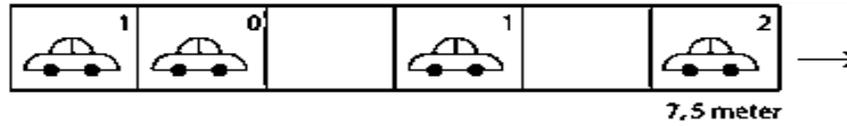


Figura 13. Discretización espacial con un tamaño de 7.5 metros por celda y ocupación de partículas en el modelo de NaSch. El número en la esquina superior derecha es la velocidad del vehículo.

Un movimiento de 3 sitios/iteración representa la velocidad $3 \cdot 7.5 \text{m/s} = 75 \text{km/h}$ en la realidad.

4.4. Reglas de actualización de vehículos con autómatas celulares de Nagel y Schreckenberg

En la vía, N vehículos se mueven de modo que su densidad viene dada por $\rho = N / L$. La velocidad v de cada vehículo puede tomar uno de los valores enteros permitidos $v_{max} + 1$ $v = 0, 1, \dots, v_{max}$. Suponga que x_n y x_n denotan la posición y la velocidad, respectivamente, del n -ésimo vehículo. Entonces, $d_n = x_n + 1 - x_n - 1$ es el avance (espacial) del n -ésimo vehículo en el tiempo t , es decir, el número de celdas vacías frente a este automóvil.

En cada paso de tiempo $t \rightarrow t + 1$, la disposición de los N vehículos en una red finita de longitud L se actualiza de acuerdo con las siguientes reglas:

NaSch1: Aceleración.

Si $v_n < v_{max}$, la velocidad del n -ésimo vehículo aumenta en uno, pero v_n permanece inalterado si $v_n = v_{max}$, es decir,

$$v_n \rightarrow \min (v_n + 1, v_{max}) \quad (4)$$

- La regla (NaSch1) refleja la tendencia general de los conductores a conducir lo más rápido posible, si se les permite hacerlo, sin cruzar el límite de velocidad máxima.

NaSch2: Deceleración.

(Debido a otros vehículos). Si $v_n > d_n$, la velocidad del n-ésimo vehículo es reducido a d_n , es decir,

$$v_n \rightarrow \min (v_n, d_n) \quad (5)$$

- La regla (NaSch2) está destinada a evitar colisiones entre los vehículos.

NaSch3: Aleatorización.

Si $v_n > 0$, la velocidad del n-ésimo vehículo se reduce aleatoriamente por unidad con probabilidad p pero v_n no cambia si $v_n = 0$, es decir,

$$v_n \rightarrow \max (v_n - 1, 0) \quad (6)$$

Con una probabilidad p

- La aleatorización en (NaSch3) tiene en cuenta los diferentes patrones de comportamiento de los conductores individuales, especialmente, la aceleración no determinista, así como la reacción exagerada al reducir la velocidad; esto es de vital importancia para la formación espontánea de atascos.

NaSch4: Movimiento del vehículo. (Actualización)

Cada vehículo se mueve hacia adelante de acuerdo con su nueva velocidad determinada en las reglas anteriores, es decir,

$$x_n \rightarrow x_n + v_n. \quad (7)$$

El modelo NaSch es un modelo mínimo en el sentido de que los cuatro pasos son necesarios para reproducir las características básicas del tráfico real (ver fig. 14). Sin embargo, es necesario formular reglas adicionales para capturar situaciones más complejas. Incluso cambiar el orden preciso de los pasos de las reglas de actualización indicadas anteriormente cambiaría las propiedades del modelo (Schadschneider, 2002).

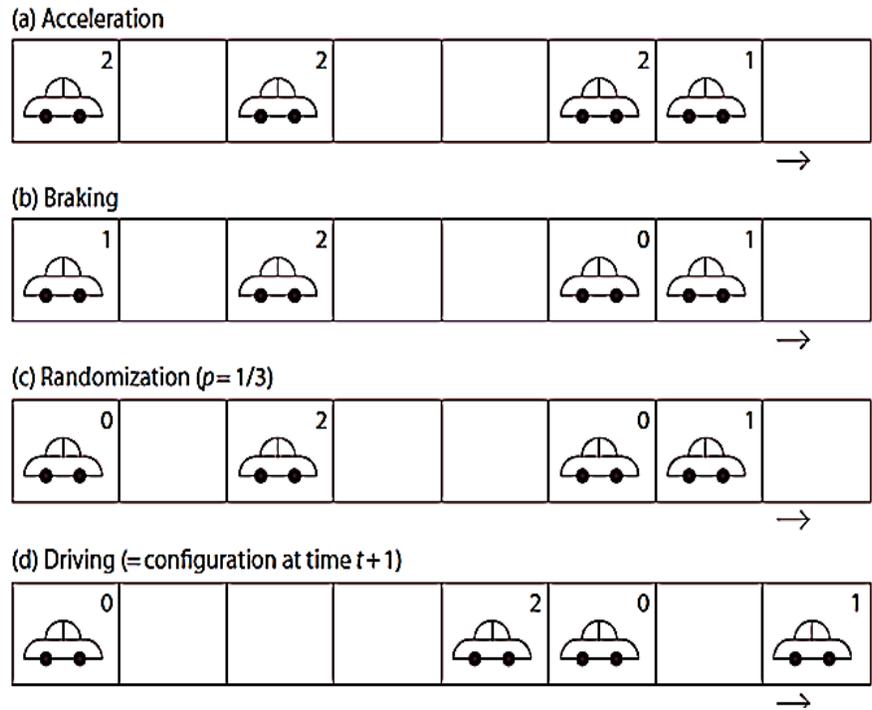


Figura 14. Ejemplo paso a paso de la aplicación de las reglas de actualización con $v_{max} = 2$ y $p = 1/3$.

4.5. Inteligencia artificial

4.6. ¿Qué es inteligencia artificial?

La inteligencia artificial (IA) es una rama de las ciencias de la computación que incluye conceptos muy transversales relacionados con la lógica y el aprendizaje. Se trata, por lo tanto, de diseñar herramientas informáticas que simulen procesos de inteligencia humana que incluyen el aprendizaje, el razonamiento y la autocorrección. Mediante diversos algoritmos las máquinas «aprende» y son capaces de «tomar decisiones». No se trata de ciencia ficción, sino de una realidad que está presente actualmente y que poco a poco va ocupando espacios cotidianos en nuestra casa, trabajo y también relacionados con nuestra movilidad (Avila-Tomás, 2020).

Sin embargo, a diferencia de las personas, los dispositivos basados en IA no necesitan descansar y pueden analizar grandes volúmenes de información a la vez. Asimismo, la proporción de errores es significativamente menor en las máquinas que realizan las mismas tareas que sus contrapartes humanas.

4.7. Clasificación de la IA

Los expertos en ciencias de la computación Stuart Russell y Peter Norvig (Stuart Russell, 2004) diferencian varios tipos de inteligencia artificial:

- Sistemas que piensan como humanos

Automatizan actividades como la toma de decisiones, la resolución de problemas y el aprendizaje. Un ejemplo son las redes neuronales artificiales.

- Sistemas que actúan como humanos

Se trata de computadoras que realizan tareas de forma similar a como lo hacen las personas. Es el caso de los robots.

- Sistemas que piensan racionalmente

Intentan emular el pensamiento lógico racional de los humanos, es decir, se investiga cómo lograr que las máquinas puedan percibir, razonar y actuar en consecuencia. Los sistemas expertos se engloban en este grupo.

- Sistemas que actúan racionalmente

Idealmente, son aquellos que tratan de imitar de manera racional el comportamiento humano, como los agentes inteligentes.

Las tecnologías basadas en la IA ya están siendo utilizadas para ayudar a los humanos a beneficiarse de mejoras significativas y disfrutar de una mayor eficiencia en casi todos los ámbitos de la vida. Pero el gran crecimiento de la IA también nos obliga a estar atentos para prevenir y analizar las posibles desventajas directas o indirectas que pueda generar la proliferación de la IA.

Chatsbots

Un *chatbot* es un programa informático que utiliza inteligencia artificial (IA) y procesamiento natural del lenguaje (PNL) para comprender las preguntas de los clientes y automatizar las respuestas, simulando una conversación humana (Machines, 2023).

4.8. Aplicaciones de la inteligencia artificial

- **Asistentes personales virtuales**

Conviviremos con *chatbots* interactivos que podrán sugerirnos productos, restaurantes, hoteles, servicios, espectáculos, según nuestro historial de búsquedas.

- **Finanzas**

Las tecnologías inteligentes pueden ayudar a los bancos a detectar el fraude, predecir patrones del mercado y aconsejar operaciones a sus clientes.

- **Logística y transporte**

Será útil a la hora de evitar colisiones o atascos y también para optimizar el tráfico. Tesla ha desarrollado un sistema gracias al cual, cuando uno de sus coches transita una ruta por primera vez, comparte la información con el resto.

- **Educación**

Permite saber si un estudiante está a punto de cancelar su registro, sugerir nuevos cursos o crear ofertas personalizadas para optimizar el aprendizaje.

- **Comercial**

Posibilita hacer pronósticos de ventas y elegir el producto adecuado para recomendárselo al cliente. Empresas como Amazon utilizan robots para identificar si un libro tendrá o no éxito, incluso antes de su lanzamiento.

- **Climáticas**

Flotas de drones capaces de plantar mil millones de árboles al año para combatir la deforestación, vehículos submarinos no tripulados para detectar fugas en oleoductos, edificios inteligentes diseñados para reducir el consumo energético, etc.

- **Agrícolas**

Plataformas específicas que, por medio de análisis predictivos, mejoran los rendimientos agrícolas y advierten de impactos ambientales adversos.

- **Sanidad**

Existen *chatbots* que nos preguntan por nuestros síntomas para realizar un diagnóstico. La recolección de datos genera patrones que ayudan a identificar factores genéticos susceptibles de desarrollar una enfermedad.

4.9. Aprendizaje Automático (*Machine Learning*)

El *machine learning*, o aprendizaje automático, es una rama de la inteligencia artificial que ha revolucionado la forma en que las computadoras procesan y utilizan datos. A través de algoritmos avanzados y modelos matemáticos, las máquinas pueden aprender de experiencias pasadas y mejorar su rendimiento en tareas específicas. Este avance tecnológico ha tenido un impacto significativo en nuestra sociedad actual y está transformando diversas industrias.

En esta forma de la IA permite a un sistema aprender de los datos en lugar de aprender mediante la programación explícita. Sin embargo, *machine learning* no es un proceso sencillo. Conforme el algoritmo ingiere datos de entrenamiento, es posible producir modelos más precisos basados en datos. Un modelo de *machine learning* es la salida de información que se genera cuando entrena su algoritmo de *machine learning* con datos. Después del entrenamiento, al proporcionar un modelo con una entrada, se le dará una salida.

Existen diferentes tipos de aprendizaje automático, que incluyen:

4.10. Aprendizaje supervisado

El aprendizaje supervisado comienza típicamente con un conjunto establecido de datos y una cierta comprensión de cómo se clasifican estos datos. El aprendizaje supervisado tiene la intención de encontrar patrones en datos que se pueden aplicar a un proceso de analítica. Estos datos tienen características etiquetadas que definen el significado de los datos (Machines, 2023). Por ejemplo, se puede crear una aplicación de *machine learning* con base en imágenes y descripciones escritas que distinga entre millones de animales.

En este tipo de aprendizaje hay dos algoritmos (entrenamientos): el de clasificación y el de regresión.

Algoritmo de clasificación: se espera que el algoritmo diga a qué grupo pertenece el elemento en estudio. El algoritmo encuentra patrones en los datos que le damos y los clasifica en grupos. Luego compara los nuevos datos y los ubica en uno de los grupos y es así como puede predecir de qué se trata.

Algoritmo de regresión: en este método lo que se espera es un número. No lo ubica en un grupo, sino que devuelve un valor específico (Sandoval, 2018).

4.11. Aprendizaje no supervisado

El aprendizaje no supervisado se utiliza cuando el problema requiere una cantidad masiva de datos sin etiquetar. Por ejemplo, las aplicaciones de redes sociales, tales como Twitter, Instagram y Snapchat, tienen grandes cantidades de datos sin etiquetar. La comprensión del significado detrás de estos datos requiere algoritmos que clasifican los datos con base en los patrones o clústeres que encuentra. El aprendizaje no supervisado lleva a cabo un proceso iterativo, analizando los datos sin intervención humana. Se utiliza con la tecnología de detección de spam en e-mails. Existen demasiadas variables en los correos electrónicos legítimos y de spam para que un analista etiquete una cantidad masiva de correos electrónicos no solicitado. En su lugar, los clasificadores de *machine learning*, basados en *clustering* y asociación, se aplican para identificar correos electrónicos no deseados.

4.12. Sobreajuste

El sobreajuste es la tendencia que tienen la mayoría de los algoritmos de *machine learning* a ajustarse a unas características muy específicas de los datos de entrenamiento que no tienen relación causal con la función objetivo que estamos buscando para generalizar. El ejemplo más extremo de un modelo sobreajustado es un modelo que solo memoriza las respuestas correctas; este modelo al ser utilizado con datos que nunca antes ha visto va a tener un rendimiento azaroso, ya que nunca logró generalizar un patrón para predecir.

4.13. Evitar el sobreajuste

Todos los modelos de *machine learning* tienen tendencia al sobreajuste, Existen dos principales estrategias para lidiar con el sobreajuste son:

Retención de datos

Consiste en dividir el conjunto de datos, en uno o varios conjuntos de entrenamiento y otro/s conjuntos de evaluación. Es decir, que no van a pasar todos los datos al algoritmo durante el entrenamiento, sino que se retiene una parte de los datos de entrenamiento para realizar una evaluación de la efectividad del modelo. Con esto se busca evitar que los mismos datos usados para entrenar sean los mismos que se utilizan para evaluar. De esta forma se analiza con más precisión como el modelo se va comportando a medida que se entrena y se puede

detectar el punto crítico en el que el modelo deja de generalizar y comienza a sobreajustarse a los datos de entrenamiento.

Validación cruzada

En lugar de solo obtener una simple estimación de la efectividad de la generalización; la idea es realizar un análisis estadístico para obtener otras medidas del rendimiento estimado, como la media y la varianza, y así poder entender cómo se espera que el rendimiento varíe a través de los distintos conjuntos de datos. Esta variación es fundamental para la evaluación de la confianza en la estimación del rendimiento. La validación cruzada también hace un mejor uso de un conjunto de datos limitado; ya que a diferencia de la simple división de los datos en uno el entrenamiento y otro de evaluación; la validación cruzada calcula sus estimaciones sobre todo el conjunto de datos mediante la realización de múltiples divisiones e intercambios sistemáticos entre datos de entrenamiento y datos de evaluación (Briega, 2023).

Clustering

Consiste en formar grupos cerrados y homogéneos a partir de un conjunto de elementos que tienen diferentes características o propiedades, pero que comparten ciertas similitudes, ejemplo patrones climáticos, agrupar artículos por temas o para segmentar clientes (Data, 2021).

4.14. Aprendizaje por refuerzo

El aprendizaje de refuerzo es un modelo de aprendizaje conductual. El algoritmo recibe retroalimentación del análisis de datos, conduciendo al usuario hacia el mejor resultado. El aprendizaje de refuerzo difiere de otros tipos de aprendizaje supervisado, porque el sistema no está entrenado con el conjunto de datos de ejemplo. Más bien, el sistema aprende a través de la prueba y el error. Por lo tanto, una secuencia de decisiones exitosas conduce al fortalecimiento del proceso, porque es el que resuelve el problema de manera más efectiva.

4.15. Aprendizaje profundo (*Deep learning*)

El *deep learning* es un método específico de *machine learning* que incorpora las redes neuronales en capas sucesivas para aprender de los datos de manera iterativa. El *deep learning* es especialmente útil cuando se trata de aprender patrones de datos no estructurados. Las redes neuronales complejas de *deep learning* están diseñadas para emular cómo funciona

el cerebro humano, así que las computadoras pueden ser entrenadas para lidiar con abstracciones y problemas mal definidos. Las redes neuronales y el *deep learning* se utilizan a menudo en el reconocimiento de imágenes, voz y aplicaciones de visión de computadora (Machines, 2023).

4.16. Regresión Lineal

La regresión lineal o *linear regression* es una técnica paramétrica utilizada para predecir variables continuas, dependientes, dado un conjunto de variables independientes. De un modo general se dice que existe regresión de los valores de una variable con respecto a los de la otra cuando hay alguna línea, denominada línea de regresión, que se ajusta más o menos claramente a los valores observados. La regresión se usa para la identificación de relaciones potencialmente causales o bien, cuando no existen dudas sobre su relación causal, para predecir una variable a partir de la otra. Cuando dos variables tienen una relación de tipo determinista, el valor de una define exactamente el valor de la otra; un ejemplo puede ser la relación entre la presión y el volumen de un gas a temperatura constante (Abuín, 2007).

4.17. Concepto de interacción

Otro concepto importante que debe ser tenido en cuenta al construir un modelo de regresión es que pueden introducirse términos independientes únicos (una sola variable, por ejemplo efecto del tabaco) y además las interacciones entre variables de cualquier orden (efecto del tabaco según género), si consideramos que pueden ser de interés o afectar a los resultados. Al introducir los términos de interacción en un modelo de regresión es importante para la correcta estimación del modelo respetar un orden jerárquico, es decir siempre que se introduzca un término de interacción de orden superior ($x \cdot y \cdot z$), deben introducirse en el modelo los términos de interacción de orden inferior ($x \cdot y, x \cdot z, y \cdot z$) y por supuesto los términos independientes de las variables que participan en la interacción (x, y, z) (Mata, 2020).

4.18. Diagrama de dispersión

Previo a todo análisis, resulta conveniente una primera inspección visual de los datos al objeto de comprobar la conveniencia o no de utilizar el modelo de regresión simple. Se recurre a este respecto, a la representación conjunta de los datos mediante el diagrama de dispersión o nube de puntos. Una simple ojeada nos permitirá determinar (se entiende grosso

modo): a) si existe relación o no entre las variables y b) si ésta es o no lineal. Además pueden extraerse otras informaciones de interés, como son: c) el grado de estrechez de la nube de puntos, indicadora de la intensidad de la relación, d) si existen valores anómalos que distorsionan la posible relación, o e), si la dispersión de los datos a lo largo de la nube de puntos es uniforme etc. (Camacho, 2023).

4.19. Redes Neuronales Artificiales

Los desarrollos actuales de los científicos se dirigen al estudio de las capacidades humanas como una fuente de nuevas ideas para el diseño de las nuevas máquinas. Así, la inteligencia artificial es un intento por descubrir y describir aspectos de la inteligencia humana que pueden ser simulados mediante máquinas. Esta disciplina se ha desarrollado fuertemente en los últimos años teniendo aplicación en algunos campos como visión artificial, demostración de teoremas, procesamiento de información expresada mediante lenguajes humanos... etc.

Las redes neuronales son más que otra forma de emular ciertas características propias de los humanos, como la capacidad de memorizar y de asociar hechos. En definitiva, las redes neuronales no son más que un modelo artificial y simplificado del cerebro humano. Una red neuronal es “un nuevo sistema para el tratamiento de la información, cuya unidad básica de procesamiento está inspirada en la célula fundamental del sistema nervioso humano: la neurona”.

Una red neuronal artificial es una estructura compuesta de un número de unidades interconectadas (neuronas artificiales). Cada unidad posee una característica de entrada/salida e implementa una computación local o función. La salida de cualquier unidad está determinada por su característica de entrada/salida, su interconexión con otras unidades, y (posiblemente) de sus entradas externas. La red neuronal artificial desarrolla usualmente una funcionalidad general a través de una o más formas de entrenamiento (Schalkoff, 1997).

Debido a su constitución y a sus fundamentos, las redes neuronales artificiales presentan un gran número de características semejantes a las del cerebro. Por ejemplo, son capaces de aprender de la experiencia, de generalizar de casos anteriores a nuevos casos, de abstraer características esenciales a partir de entradas que representan información irrelevante, etc. Esto hace que ofrezcan numerosas ventajas y que este tipo de tecnología se esté aplicando en múltiples áreas.

Aprendizaje adaptativo. Capacidad de aprender a realizar tareas basadas en un entrenamiento o en una experiencia inicial.

Auto-organización. Una red neuronal puede crear su propia organización o representación de la información que recibe mediante una etapa de aprendizaje.

Tolerancia a fallos. La destrucción parcial de una red conduce a una degradación de su estructura; sin embargo, algunas capacidades de la red se pueden retener, incluso sufriendo un gran daño. Operación en tiempo real, los cálculos neuronales pueden ser realizados en paralelo; para esto se diseñan y fabrican máquinas con hardware especial para obtener esta capacidad. Fácil inserción dentro de la tecnología existente, se pueden obtener chips especializados para redes neuronales que mejoran su capacidad en ciertas tareas. Ello facilitará la integración modular en los sistemas existentes.

4.20. Modelo biológico red neuronal

El cerebro humano contiene más de 100 billones de elementos de procesos llamados neuronas, como se observa en la fig. 15, que se comunican a través de conexiones llamadas sinapsis.

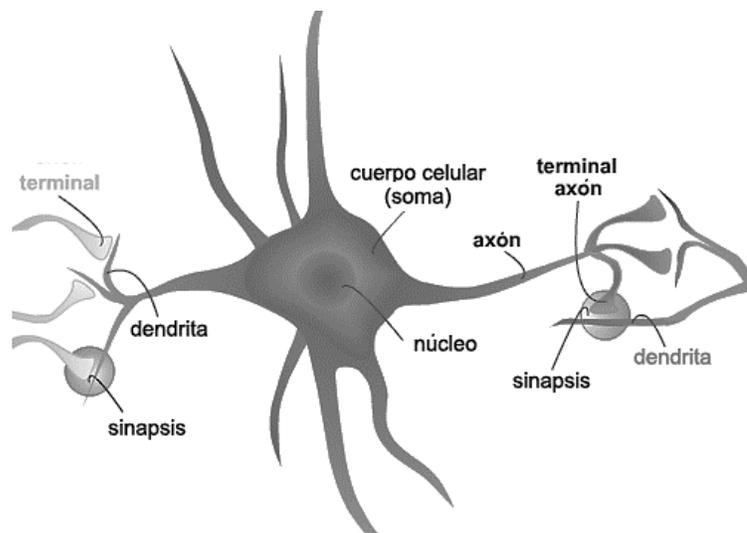


Figura 15. Modelo de neurona biológica. (Ramírez-Agundis, 2023).

Las dendritas que son como las ramas que salen del cuerpo, poseen algunas conexiones sinápticas en donde se reciben señales que generalmente vienen de otros axones.

El axón se encarga de activar o inhibir otras neuronas las cuales a su vez son activadas por cientos o miles de otras neuronas.

4.21. Red neuronal artificial

Una red neuronal artificial (fig. 16) es una estructura compuesta de un número de unidades interconectadas (neuronas artificiales). Cada unidad posee una característica de entrada/salida e implementa una computación local o función. La salida de cualquier unidad está determinada por su característica de entrada/salida, su interconexión con otras unidades, y (posiblemente) de sus entradas externas. La RNA desarrolla usualmente una funcionalidad general a través de una o más formas de entrenamiento (Schalkoff, 1997).

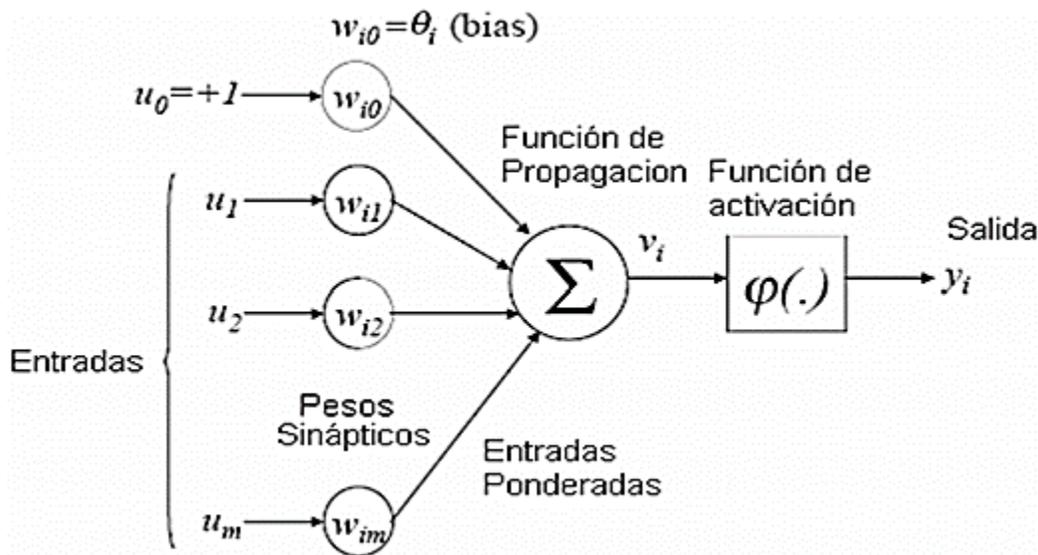


Figura 16. Modelo de una neurona artificial. (Rodríguez, 2023)

Básicamente, un modelo de una neurona simple consiste en aplicar un conjunto de entradas, cada una representando la salida de otra neurona, o una entrada del medio externo, realiza una suma ponderada con estos valores y “filtra” este valor con una función.

Cada neurona artificial recibe un vector In de entrada que corresponde a todas aquellas señales $p_1, p_2, p_3 \dots p_R$ que llegan a la sinapsis. Cada una de estas señales se multiplica por un peso asociado $W_1, W_2, W_3 \dots W_R$. Al conjunto de pesos se le denomina vector W . Cada peso representa la “intensidad” o fuerza de conexión de una sinapsis en una neurona biológica. Los resultados de estas multiplicaciones se suman y simulan vagamente el cuerpo de una neurona biológica.

4.22. Estructura artificial

En la figura 17 podemos observar el ejemplo de una red neuronal totalmente conectada. Se interconectan neuronas en tres tipos de capas:

- De entrada: reciben estímulos externos.
- Oculta: elementos internos de procesamiento (se pueden estructurar en varias capas).
- De salida: reciben la información procesada y retornan la respuesta del sistema al exterior.

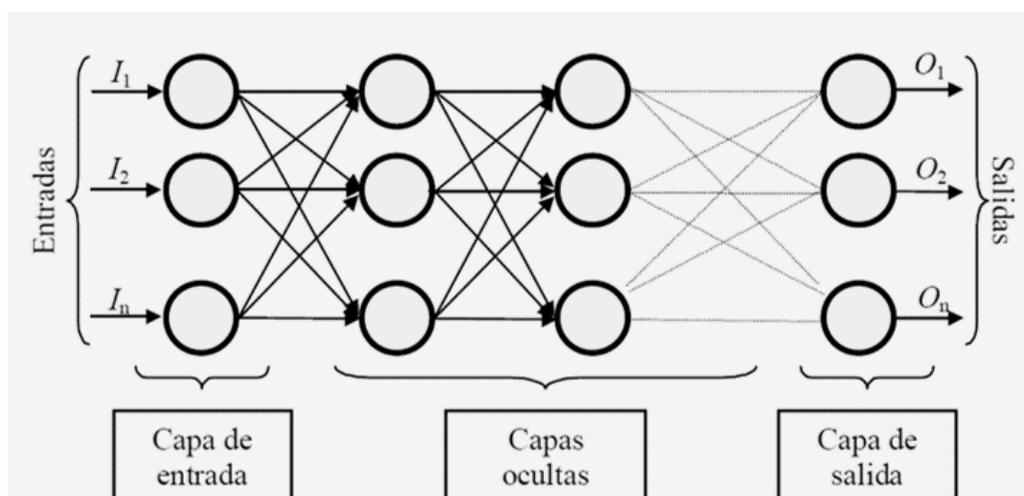


Figura 17. Ejemplo de una red neuronal totalmente conectada. (Matich 2001).

La misma está constituida por neuronas interconectadas y arregladas en tres capas (esto último puede variar). Los datos ingresan por medio de la “capa de entrada”, pasan a través de la “capa oculta” y salen por la “capa de salida”. Cabe mencionar que la capa oculta puede estar constituida por varias capas.

4.23. Clasificación

La clasificación se define como el proceso de reconocimiento, comprensión y agrupación de objetos e ideas en categorías preestablecidas, también conocidas como "subpoblaciones". Con la ayuda de estos conjuntos de datos de entrenamiento pre categorizados, la clasificación en los programas de aprendizaje automático aprovecha una amplia gama de algoritmos para clasificar conjuntos de datos futuros en categorías respectivas y relevantes (Banoula, 2023).

4.24. Clasificación en aprendizaje automático (*classification learner*)

La "clasificación por aprendizaje" es un término que puede referirse a un proceso en el campo del aprendizaje automático (*machine learning*) donde un algoritmo se entrena para clasificar datos en diferentes categorías o clases en función de las características de esos datos. La clasificación es una tarea fundamental en el aprendizaje automático y se utiliza en una amplia variedad de aplicaciones, como reconocimiento de imágenes, detección de spam, diagnóstico médico, análisis de sentimientos, entre otros.

Descripción de cómo funciona la clasificación por aprendizaje:

Recopilación de datos: En primer lugar, se reúnen y preparan datos de entrenamiento que contienen ejemplos etiquetados. Cada ejemplo consta de un conjunto de características y una etiqueta que indica la categoría a la que pertenece.

Selección de algoritmo: Luego, se elige un algoritmo de clasificación adecuado. Hay varios algoritmos de clasificación disponibles, como regresión logística, máquinas de vectores de soporte (SVM), árboles de decisión, redes neuronales, vecinos más cercanos entre otros. La elección del algoritmo depende del tipo de datos y del problema que se esté abordando.

Entrenamiento del modelo: El algoritmo se entrena utilizando los datos de entrenamiento. Durante el entrenamiento, el modelo ajusta sus parámetros para aprender a hacer coincidir las características de entrada con las etiquetas de salida. El objetivo es minimizar la discrepancia entre las predicciones del modelo y las etiquetas reales en los datos de entrenamiento.

Validación y ajuste: Después del entrenamiento, se evalúa el rendimiento del modelo utilizando datos de validación o pruebas que no se utilizaron durante el proceso de entrenamiento. Se ajustan los hiperparámetros del modelo y se realizan mejoras para lograr un rendimiento óptimo.

Predicción y clasificación: Una vez que el modelo está entrenado y validado, se puede utilizar para clasificar nuevos datos sin etiquetas en las categorías deseadas. El modelo toma las características de entrada y predice la etiqueta o clase correspondiente.

La clasificación por aprendizaje se utiliza en una amplia gama de aplicaciones en la vida real. Por ejemplo, en la clasificación de correos electrónicos, un algoritmo de aprendizaje automático puede decidir si un correo electrónico es spam o no en función de su contenido y características. En el diagnóstico médico, puede ayudar a identificar enfermedades a partir de datos de pacientes. En el reconocimiento de imágenes, puede categorizar objetos en una imagen en función de sus características visuales. El predecir la probabilidad de un accidente a partir de valores suministrados por sensores.

La clasificación por aprendizaje es un proceso fundamental en el aprendizaje automático que implica entrenar modelos para asignar datos a categorías o clases específicas basadas en características (Banoula, 2023).

4.25. Interpretación geométrica de la clasificación

Uno de los modos más sencillos de entender la clasificación cuando se basa en medidas cuantitativas es pensar en el conjunto de variables que caracterizan a individuos y clases como conjunto de ejes que definen un espacio de variables multidimensional. En la figura 18 se muestra un ejemplo de cuatro clases de variables representadas en un plano tridimensional.

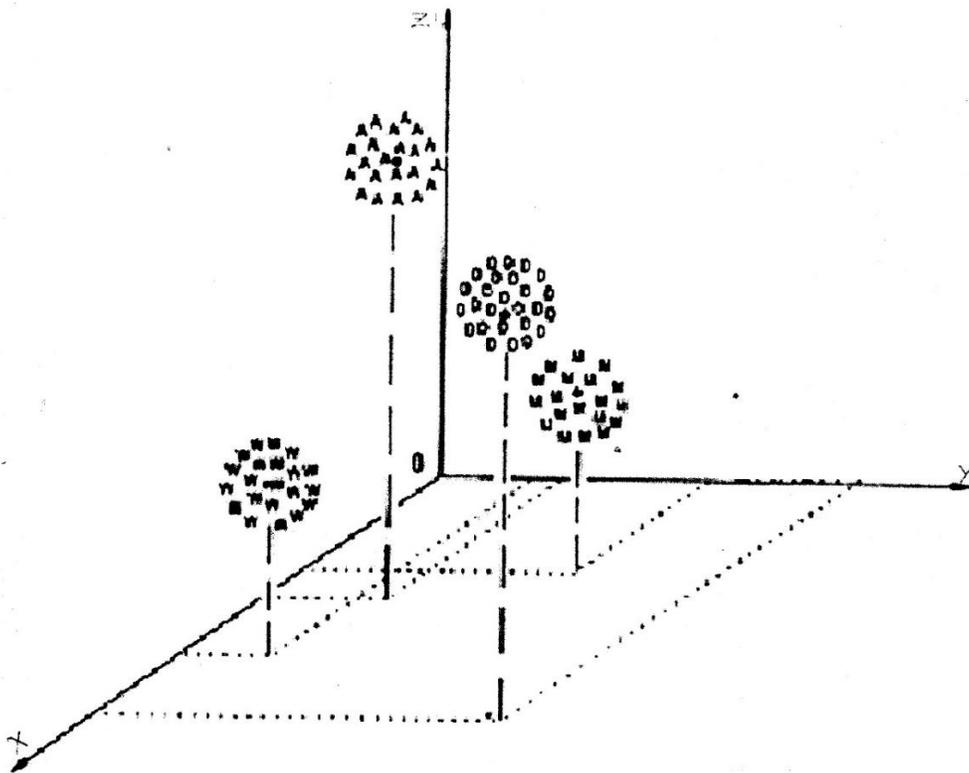


Figura 18. Definición de clases en el espacio de variables (Murcia, 2023).

4.26. Algoritmos de clasificación

Existen varios algoritmos de clasificación en el campo del aprendizaje automático, y la elección del algoritmo adecuado depende de la naturaleza de los datos y el problema en cuestión. La elección del algoritmo de clasificación adecuado depende de la naturaleza de los datos, la cantidad de datos disponibles, la complejidad del problema y otros factores.

A menudo, es una buena práctica probar varios algoritmos y ajustar sus hiperparámetros para encontrar el que mejor se adapte a una tarea específica (Machines, 2023).

Algoritmos de clasificación más comunes:

Regresión Logística: A pesar de su nombre, la regresión logística se utiliza para problemas de clasificación binaria, donde se deben predecir dos clases distintas. Utiliza una función logística para modelar la probabilidad de que un ejemplo pertenezca a una de las dos clases. Es simple, rápido y eficaz en muchas aplicaciones.

Máquinas de Vectores de Soporte (SVM): Las SVM son excelentes para la clasificación binaria y pueden extenderse a problemas de clasificación multiclase. Intentan encontrar un hiperplano que mejor separe las clases en un espacio de alta dimensión. Son especialmente útiles cuando hay una separación clara entre las clases.

Árboles de Decisión: Los árboles de decisión dividen el conjunto de datos en subconjuntos más pequeños en función de las características, con el objetivo de clasificar los ejemplos en clases. Son fáciles de entender e interpretar, pero pueden ser propensos a sobreajuste (CEUPE, 2023).

Bosques Aleatorios (Random Forest): Un bosque aleatorio es un conjunto de múltiples árboles de decisión. Combina las predicciones de varios árboles para obtener resultados más robustos y reducir el sobreajuste. Es adecuado para una amplia variedad de problemas de clasificación.

Naïve Bayes: Basado en el teorema de Bayes, el algoritmo Naïve Bayes es efectivo para problemas de clasificación de texto y minería de datos. Supone independencia condicional entre las características, lo que puede ser una simplificación útil.

Redes Neuronales Artificiales (RNA): Las redes neuronales son modelos de aprendizaje profundo que se utilizan en una variedad de tareas de clasificación. Son especialmente efectivas en problemas complejos, como el procesamiento de imágenes y el procesamiento del lenguaje natural, pero a menudo requieren grandes cantidades de datos y recursos computacionales.

Máquinas de Soporte Vectorial (SVM): Las máquinas de soporte vectorial se utilizan para problemas de clasificación binaria y multiclase. Intentan encontrar un hiperplano que mejor separe las clases en un espacio de alta dimensión. Son especialmente útiles cuando hay una separación clara entre las clases.

Gradient Boosting: Este enfoque combina múltiples modelos más simples para construir un modelo más robusto y preciso. XGBoost y LightGBM son ejemplos populares de implementaciones de gradient boosting.

Clasificadores Bayes Ingenuos (Naïve Bayes): Estos algoritmos se basan en el teorema de Bayes y asumen independencia condicional entre las características. Son particularmente efectivos en problemas de clasificación de texto, como la clasificación de spam o análisis de sentimientos (CEUPE, 2023).

K Vecinos Más Cercanos (KNN): El algoritmo de vecinos más cercanos (k-Nearest Neighbor) también conocido como KNN, es un clasificador de aprendizaje supervisado no paramétrico, que utiliza la proximidad para hacer clasificaciones o predicciones sobre la agrupación de un punto de datos individual. La letra “k” es un término variable que implica que se podría usar cualquier número de vecinos más cercanos.

Si bien se puede usar para problemas de regresión o clasificación, generalmente se usa como un algoritmo de clasificación, partiendo de la suposición de que se pueden encontrar puntos similares cerca uno del otro.

Para los problemas de clasificación, se asigna una etiqueta de clase sobre la base de un voto mayoritario, es decir, se utiliza la etiqueta que se representa con más frecuencia alrededor de un punto de datos determinado, se observa en la figura 19 esa asignación. Si bien esto técnicamente se considera "voto por mayoría", el término "voto por mayoría" se usa más comúnmente en la literatura. La distinción entre estas terminologías es que "voto

mayoritario" técnicamente requiere una mayoría superior al 50 %, lo que funciona principalmente cuando solo hay dos categorías. Cuando tiene varias clases, por ejemplo, cuatro categorías, no necesita necesariamente el 50 % de los votos para llegar a una conclusión sobre una clase; puede asignar una etiqueta de clase con un voto superior al 25 % (Raschka, 2018).

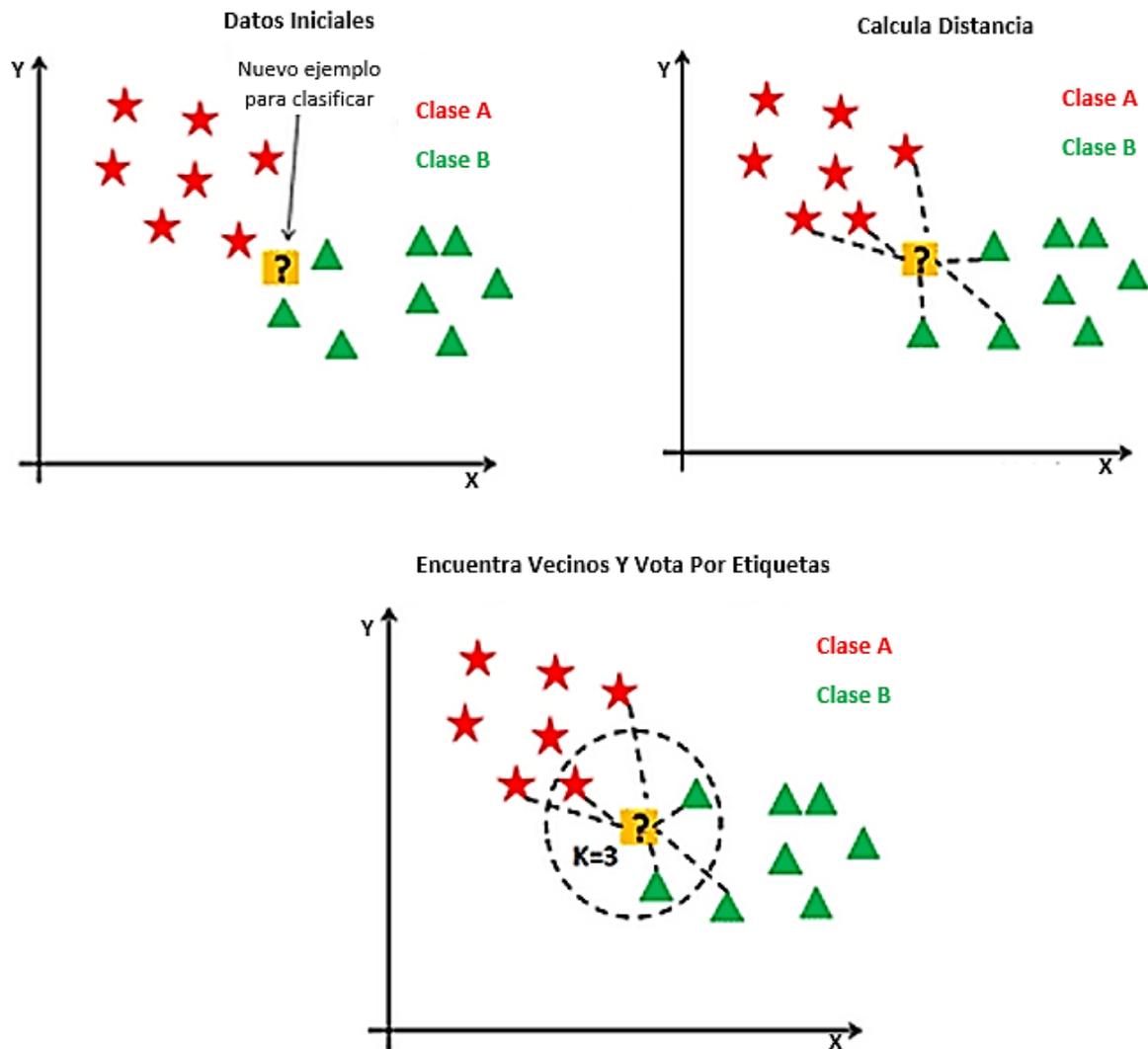


Figura 19. Descripción del algoritmo KNN.

4.27. Elección de número de vecinos k

La decisión de cuántos vecinos usar para KNN determina qué tan bien el modelo generalizará para futuros datos.

La elección de un k grande reduce el impacto o la varianza causada por la data con ruido, pero puede sesgar el aprendizaje con el riesgo de ignorar patrones pequeños pero importantes.

Suponiendo que se elija un k tan grande como el número total de observaciones en los datos de entrenamiento. Con cada instancia de entrenamiento representado en la votación final, la clase más común siempre tiene la mayoría de votos. Por lo tanto, el modelo siempre predeciría la clase mayoritaria, independientemente de los vecinos más cercanos.

En el extremo opuesto, el uso de un $k=1$ permite la data con ruido u *outliers* que influyen indebidamente en la clasificación de ejemplos, ver figura 20.

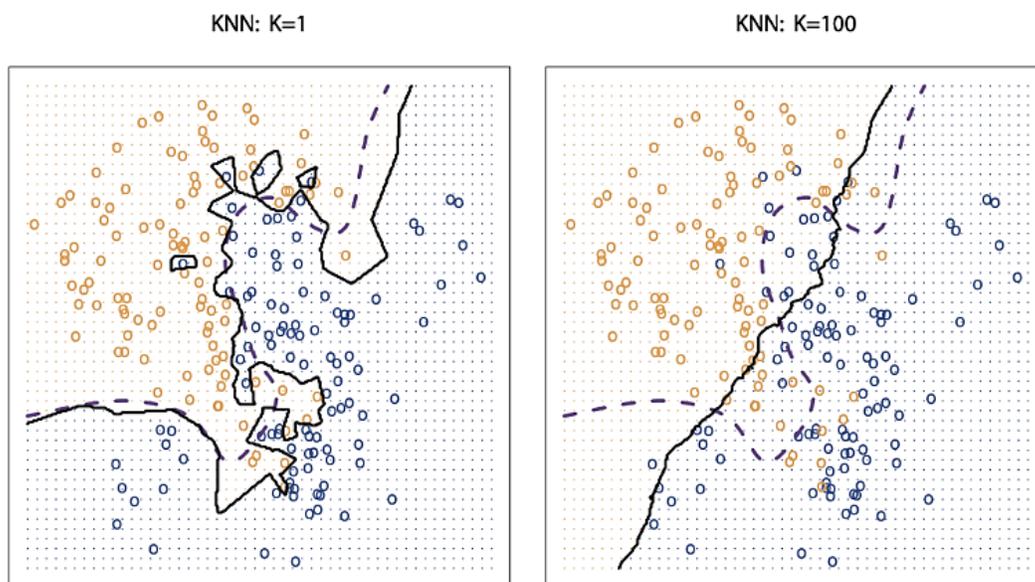


Figura 20. Comparación de valores de k minoritario y mayoritario (Toth, 2020).

Suponiendo que algunos de los ejemplos de entrenamiento fueron mal etiquetados accidentalmente. Cualquier ejemplo sin etiqueta que sea el más cercano al vecino etiquetado incorrectamente, se predice que tendrá la clase incorrecta, incluso si otros nueve vecinos más cercanos hubieran votado de manera diferente. El mejor valor k está en algún punto entre estos dos extremos (Servan, 2022).

Después de elegir k , el algoritmo requiere un conjunto de datos de entrenamiento compuesto por ejemplos que ya están clasificados en varias categorías, según la etiqueta de una variable categórica (Y). Para cada registro no etiquetado, KNN identifica los registros k en la data de

entrenamiento que son los “más cercanos” en similitud. Al registro no etiquetado se le asigna la clase de la mayoría de los k vecinos más cercanos.

4.28. Métricas de distancia del algoritmo KNN

El objetivo del algoritmo del vecino más cercano es identificar los vecinos más cercanos de un punto de consulta dado, de modo que podamos asignar una etiqueta de clase a ese punto.

Para determinar qué puntos de datos están más cerca de un punto de consulta determinado será necesario calcular la distancia entre el punto de consulta y los otros puntos de datos. Estas métricas de distancia ayudan a formar límites de decisión, que dividen los puntos de consulta en diferentes regiones (Machines, 2023).

Se mencionan las siguientes métricas de distancia, donde:

p = Exponente de la distancia de Minkowski.

Para el caso especial de $p = 1$, la distancia de Minkowski ofrece la distancia Manhattan. Para el caso especial de $p = 2$, la distancia de Minkowski ofrece la distancia euclidiana. Para el caso especial de $p = \infty$, la distancia de Minkowski ofrece la distancia de Chebyshev.

Distancia Minkowski: Esta medida de distancia es la forma generalizada de las métricas de distancia Euclidiana y Manhattan. El parámetro, p , en la fórmula a continuación, permite la creación de otras métricas de distancia. La distancia euclidiana se representa mediante esta fórmula cuando p es igual a dos, y la distancia de Manhattan se denota con p igual a uno.

$$Distancia\ Minkowski = \left(\sum_{i=1}^n |x_i - y_i| \right)^{\frac{1}{p}} \quad (8)$$

Formula de la distancia de Minkowski.

Distancia Manhattan ($p = 1$): Esta es también otra métrica de distancia popular, que mide el valor absoluto entre dos puntos. También se conoce como distancia de taxi o distancia de cuadra de la ciudad, ya que comúnmente se visualiza con una cuadrícula, que ilustra cómo se puede navegar de una dirección a otra a través de las calles de la ciudad.

$$\text{Distancia Manhattan} = d(x, y) = \left(\sum_{i=1}^m |x_i - y_i| \right) \quad (9)$$

Formula de la distancia de Manhattan.

Distancia Euclidiana ($p = 2$): Esta es la medida de distancia más utilizada y está limitada a vectores de valor real. Usando la fórmula a continuación, mide una línea recta entre el punto de consulta y el otro punto que se mide.

$$d(x, y) = \sqrt{\sum_{i=1}^n (y_i - x_i)^2} \quad (10)$$

Formula de distancia Euclidiana.

Distancia de Chebyshev: Es una métrica definida en un espacio vectorial donde la distancia entre dos puntos (representados por sus vectores) es la mayor de sus diferencias a lo largo de cualquiera de sus dimensiones coordenadas.

$$\text{Distancia Chebyshev} = \lim_{p \rightarrow \infty} \left(\sum_{i=1}^n |x_i - y_i| \right)^{1/p} \quad (11)$$

Formula de distancia de Chebyshev.

La distancia de Chebyshev es un caso especial de la distancia de Minkowski, donde $p = \infty$.

Distancia de Hamming: Esta técnica se usa típicamente con vectores booleanos o de cadena, identificando los puntos donde los vectores no coinciden. Como resultado, también se la conoce como la métrica de superposición. Esto se puede representar con la siguiente fórmula:

$$\text{Distancia Hamming} = D_H = \left(\sum_{i=1}^k |x_i - y_i| \right) \quad (12)$$

$$x = y \quad D = 0$$

$$x \neq y \quad D \neq 1$$

Fórmula de la distancia de Hamming.

4.29. Aplicaciones de KNN en *Machine Learning*

El algoritmo KNN se ha utilizado en una variedad de aplicaciones, principalmente dentro de la clasificación. Algunos de estos paquetes incluyen:

-Pre-procesamiento de datos: Los conjuntos de datos suelen tener valores faltantes, pero el algoritmo KNN puede estimar esos valores en un proceso conocido como imputación de datos faltantes.

-Motores de recomendación: utilizando datos de flujo de clics de sitios web, el algoritmo KNN se ha utilizado para proporcionar recomendaciones automáticas a los usuarios sobre contenido adicional.

-Reconocimiento de patrones: KNN también ha ayudado a identificar patrones, como en texto y clasificación de dígitos. Esto ha sido particularmente útil para identificar números escritos a mano que puede encontrar en formularios o sobres de correo.

-Sistemas de transporte: se realizan estudios de implementación del algoritmo KNN para realizar una clasificación y predicción de los accidentes viales en zonas de alto riesgo, como intersecciones viales.

4.30. Evaluación del rendimiento de los modelos de clasificación de *machine learning*

Las curvas ROC

Curva ROC o característica operativa del receptor (figuras 21), constituye una herramienta importante para evaluar el rendimiento de un modelo de *machine learning*. Por lo general, se utilizan en problemas de clasificación binaria, concretamente, problemas con dos clases de salidas distintas. La curva ROC muestra la relación entre la tasa de verdaderos positivos (true positive rate, TPR) y la tasa de falsos positivos (false positive rate, FPR) del modelo. TPR describe la tasa a la que el clasificador predice como “positivo” a las observaciones que son “positivas”. FPR describe la tasa a la que el clasificador predice como “positivo” a las observaciones que en realidad son “negativas”. Los clasificadores perfectos tienen una TPR de 1 y una FPR de 0.

La mayoría de los modelos de *machine learning* para clasificación binaria no generan solo 1 o 0 cuando hacen una predicción. En realidad, generan un valor continuo en un intervalo [0,1]. Los valores iguales o superiores a un determinado umbral (por ejemplo, 0,5) se clasifican como 1 y los valores por debajo de ese umbral se clasifican como 0. Los puntos de la curva ROC representan la FPR y la TPR para diferentes umbrales.

El umbral seleccionado puede encontrarse en cualquier punto del intervalo [0,1], y las clasificaciones resultantes cambiarán según el valor del umbral. Por ejemplo, si el umbral se establece como 0 (cualquier valor igual o superior a 0 se clasifica como 1), el modelo siempre predecirá 1, dando como resultado una TPR de 1 y una FPR de 1. En el otro extremo de la curva ROC, si el umbral se establece como 1 (cualquier valor por debajo de 1 se clasifica como 0), el modelo siempre predecirá 0, dando como resultado una TPR de 0 y una FPR de 0 (Deland, 2023).

Al evaluar el rendimiento de un modelo de clasificación, el enfoque reside en el comportamiento entre extremos. En general, cuanto más “arriba y a la izquierda” del diagrama se encuentre la curva ROC, mejor será el clasificador.

Las curvas ROC se utilizan normalmente con validación cruzada para evaluar el rendimiento del modelo con datos de validación o prueba.

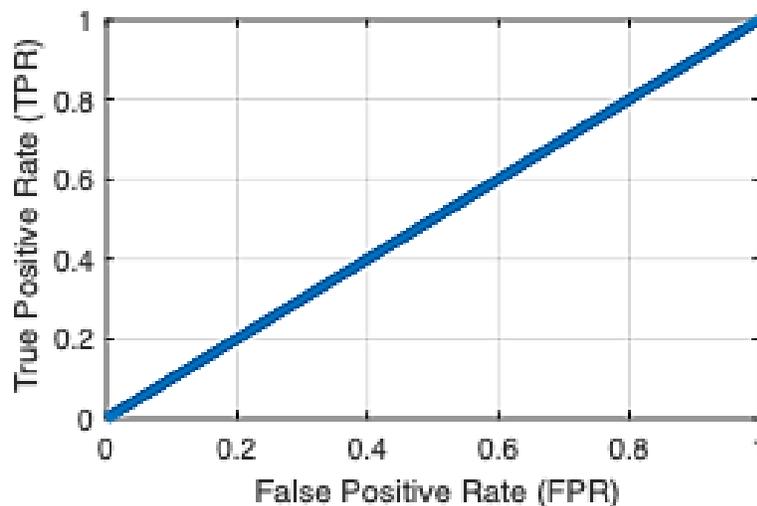


Figura 21. Curva ROC con bajo nivel de clasificación.

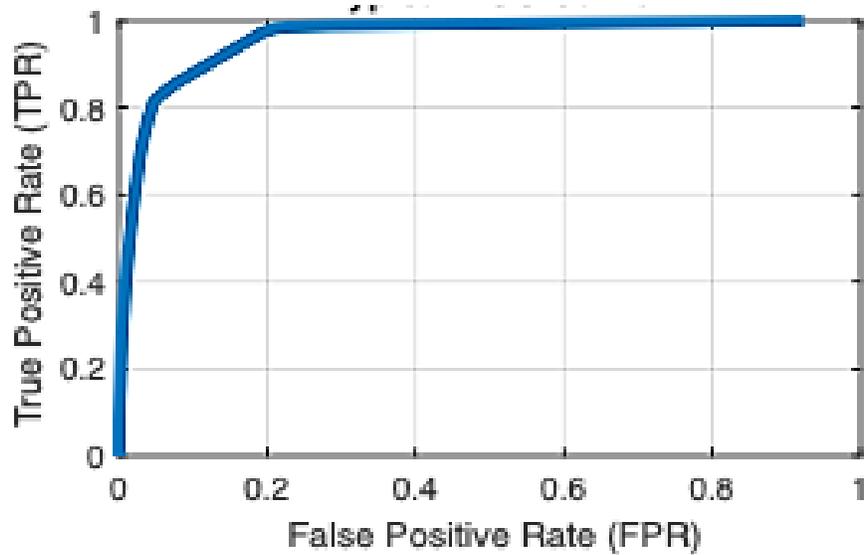


Figura 22. Típica curva ROC.

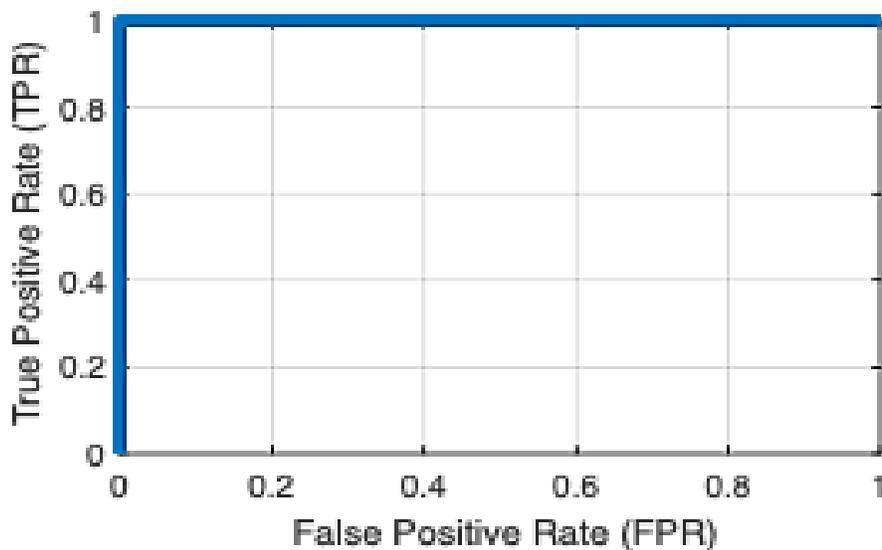


Figura 23. Curva ROC con perfecta clasificación.

4.31. Matriz de confusión

Una matriz de confusión (fig. 24), también conocida como matriz de error, es una tabla resumida que se utiliza para evaluar el rendimiento de un modelo de clasificación. El número de predicciones correctas e incorrectas se resumen con los valores de conteo y se desglosan por cada clase.

A continuación se muestra una imagen de la estructura de una matriz de confusión de 2x2. Para dar un ejemplo, digamos que hubo diez casos en los que un modelo de clasificación

predijo "Sí" en los que el valor real fue "Sí". Entonces el número diez iría en la esquina superior izquierda en el cuadrante verdadero positivo. Esto nos lleva a algunos términos clave:

		Valores actuales	
		Si	No
Valores predichos	Si	Verdaderos positivos	Falsos positivos
	No	Falsos negativos	Verdaderos negativos

Figura 24. Matriz de confusión.

Positivo: La observación es positiva (por ejemplo, es un perro).

Negativo: La observación no es positiva (por ejemplo, no es un perro).

Verdadero Positivo: Resultado en el que el modelo predice correctamente la clase positiva.

Verdadero Negativo: Resultado donde el modelo predice correctamente la clase negativa.

Falso Positivo: También llamado error de tipo 1, resultado donde el modelo predice incorrectamente la clase positiva cuando en realidad es negativa.

Falso Negativo: También llamado error de tipo 2, un resultado en el que el modelo predice incorrectamente la clase negativa cuando en realidad es positiva.

5. Metodología

El proyecto a desarrollar se divide principalmente en dos partes:

1. En la primera etapa se aplica el modelo de Autómatas Celulares para el estudio de flujo vial y el estudio de los posibles choques en una intersección señalizada.
2. La segunda etapa se recaba la información y los resultados obtenidos de la primera etapa. Estos resultados serán el punto de partida para el análisis y la predicción con inteligencia artificial el flujo y los accidentes viales.

5.1 Etapas de generación de una intersección vial.

- a) Se recreó el sistema con un solo carril para dominar el diagrama básico aplicando el modelo de NaSch calculando el flujo, la densidad y la velocidad.
- b) Más adelante se incrementó la carretera a dos carriles donde se toma en cuenta el cambio de carril P_{chg} con una probabilidad dada de 0.8, esto para aumentar su velocidad y evitar congestionamiento. Para realizar este cambio de carril debe de cumplirse la regla de vacancia trasera:

Todo vehículo que pretenda un cambio de trayectoria con un cambio en el ángulo de dirección debe poseer un espacio libre trasero, que debe ser la mitad del ancho y largo del vehículo. Si se viola cualquiera de las dos reglas, el vehículo prefiere un camino recto.

En todo momento existe un riesgo de accidentes por el comportamiento de los conductores que no respetan la señalización, se induce una probabilidad de 0.1 de que se produzca un accidente de coche cuando los conductores no respetan la distancia de seguridad. Esto produce que 0.1% del total de las iteraciones sea un accidente, puesto que los accidentes viales son fenómenos raros, ya que la mayoría no son provocados. Las condiciones para que exista un accidente son: Que un vehículo cambie de carril cuando hay otro ocupando la celda de destino. Que un vehículo lento cambie de carril al tiempo que se acerca un vehículo rápido en el carril de destino.

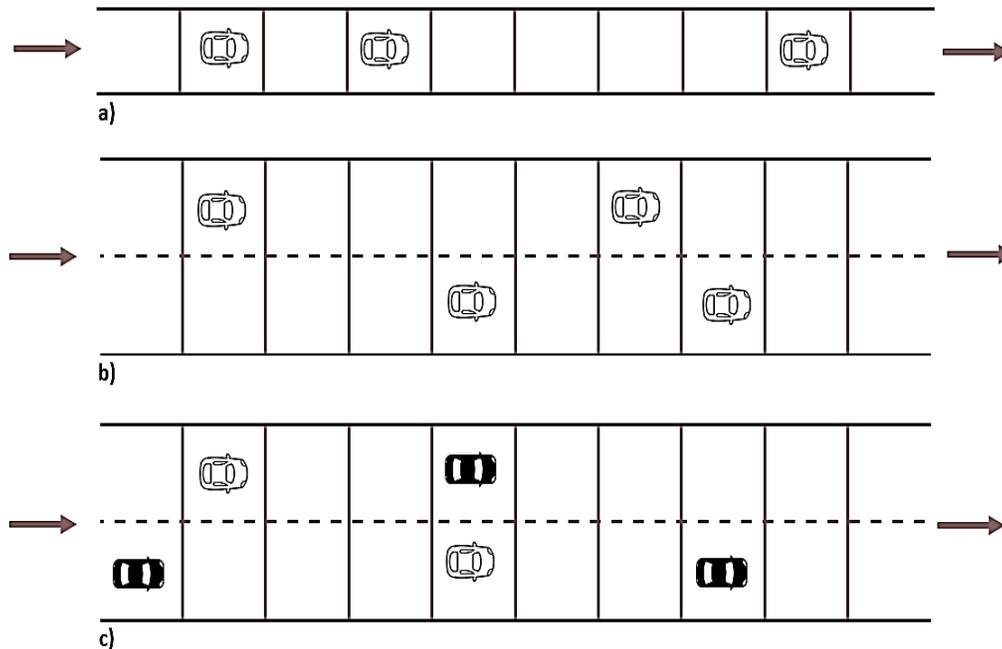


Figura 25. Etapas de generación de una intersección.

- c) Posteriormente se introdujo dos tipos de vehículos: el primero es lento con una velocidad máxima (V_{max}) de dos celdas por unidad de tiempo (1 segundo) y rápidos con una velocidad máxima (V_{max}) de tres celdas por unidad de tiempo (1 segundo) (fig. 25).

Por último se realizó una combinación de las etapas b y c para plantear una intersección de dos carreteras con dos tipos de vehículos y con semáforos, donde se recrea la dinámica básica del modelo de NaSch cuyas reglas son: aceleración, deceleración, aleatorización, movimiento del vehículo (actualización). De esa manera planteamos el programa y obtuvimos los siguientes resultados.

5.2. Modelo de estudio

Se considera una intersección vial señalizada de dos carreteras (fig. 26), donde circulan dos tipos de vehículos donde cada uno entra con una probabilidad Alpha y sale con una probabilidad beta. Cada carretera está dividida en sitio idénticos de 7.5 metros de longitud. La velocidad puede tomar solamente los valores enteros.

Del modelo de análisis obtenemos los valores de las variables para realizar la predicción de los accidentes viales.

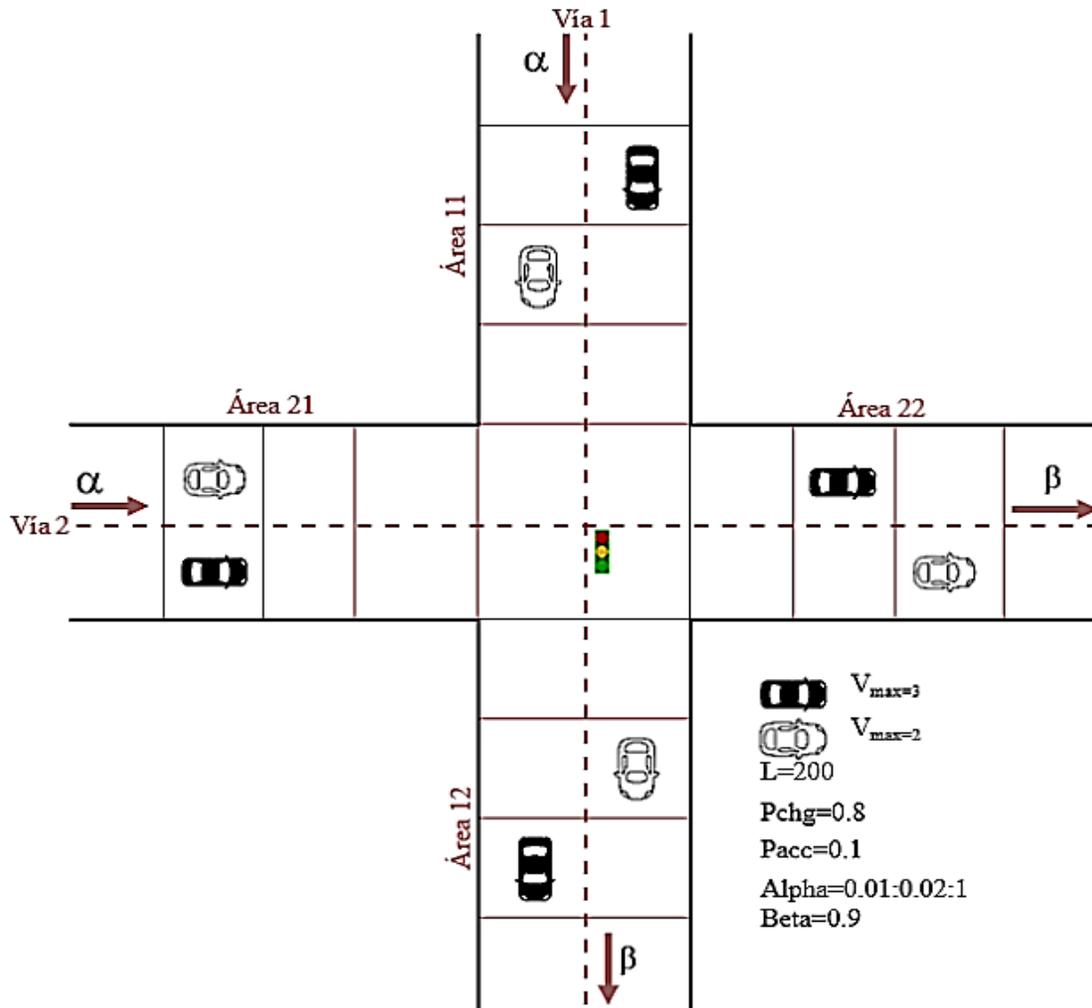


Figura 26. Intersección vial en estudio.

- Una unidad de velocidad máxima (V_{max}) corresponde a 25 km/hr.
- La velocidad máxima del vehículo en color negro es $V_{max}: 3$ es equivalente a 75 km/hr.
- La velocidad máxima del vehículo en color blanco es $V_{max}: 2$ es equivalente a 50 km/hr.
- Longitud de vía en celdas (L): 200
- Cada celda tiene una longitud de 7.5 metros lo cual indica que el modelo a en estudio tiene una longitud total de 1500 metros de norte a sur y de este a oeste.
- Probabilidad dada de cambio de carril (P_{chg}): 0.8
- Probabilidad dada de accidentes viales (P_{acc}): 0.1

- Probabilidad dada de ingreso de autos a la intersección Alpha (α): 0.01:0.02:1 (la inyección de automóviles aumenta de 0.01 a 1 con incrementos de 0.02).
- Probabilidad dada de salida de autos de la intersección Beta (β): 0.9.

Se aprecia en la figura 23, la intersección vial a estudiar, donde ya se encuentran dos tipos de autos, rápido con $V_{max}=3$ y lento con $V_{max}=2$, en dos carriles.

En la simulación de esta intersección podemos observar distintos fenómenos que vivimos en el tráfico vial real como lo es el flujo, la velocidad, la densidad, cambios de carril y accidentes de los cuales obtenemos datos que como resultado graficamos para un mayor análisis.

Relación velocidad V – $Alpha$

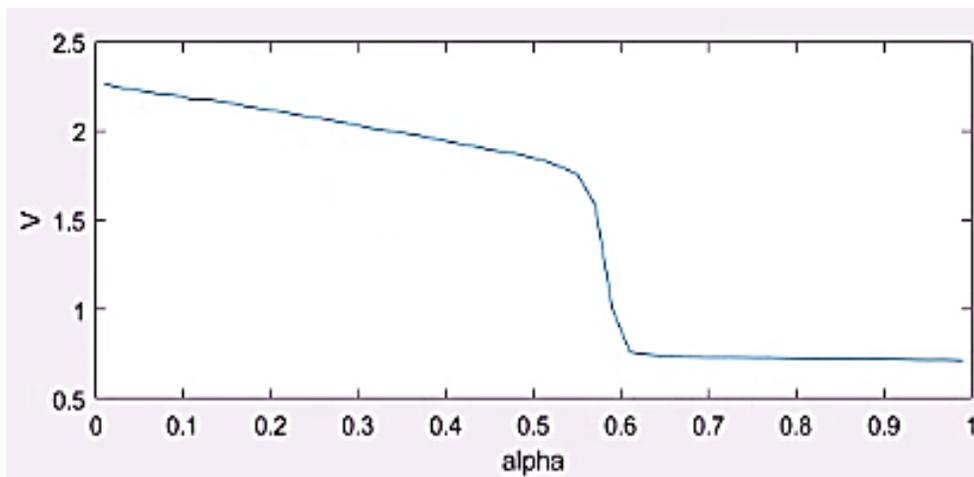


Figura 27. Relación velocidad V – $Alpha$.

En la figura 27 se observa la velocidad promedio en función de la probabilidad de inyección. La velocidad inicia cerca de 2.4 celdas/segundo (c/s) para una baja cantidad de automóviles, esto representa una vía libre con muy poco tráfico y por consiguiente una velocidad mayor. Disminuye cerca de $alpha = 0.6$ ya que la vía empieza a congestionarse puesto que los automóviles ya se encuentran ocupando un mayor número de celdas dentro de la vía, esto representa una disminución de velocidad cercana a $v = 0.5$ c/s y se mantiene constante hasta $alpha = 0.99$.

El parámetro alpha, representa la probabilidad de inyección de automóviles a la intersección $alpha=0.01:0.02:1$, en este caso la inyección de automóviles aumenta de 0.01 a 1 con incrementos de 0.02.

Relación flujo J – $Alpha$

Se observa en la figura 28, que el flujo vehicular J aumenta de una manera lineal hasta $\alpha = 0.28$ donde consigue un punto de inflexión en $alpha = 0.6$ donde conserva un valor fijo de 0.27 aproximadamente. Esto nos dice que el flujo en la intersección variara según la cantidad de autos que entran, hasta llegar a un punto donde ya no aumentará, debido a la disminución del espacio en las vías.

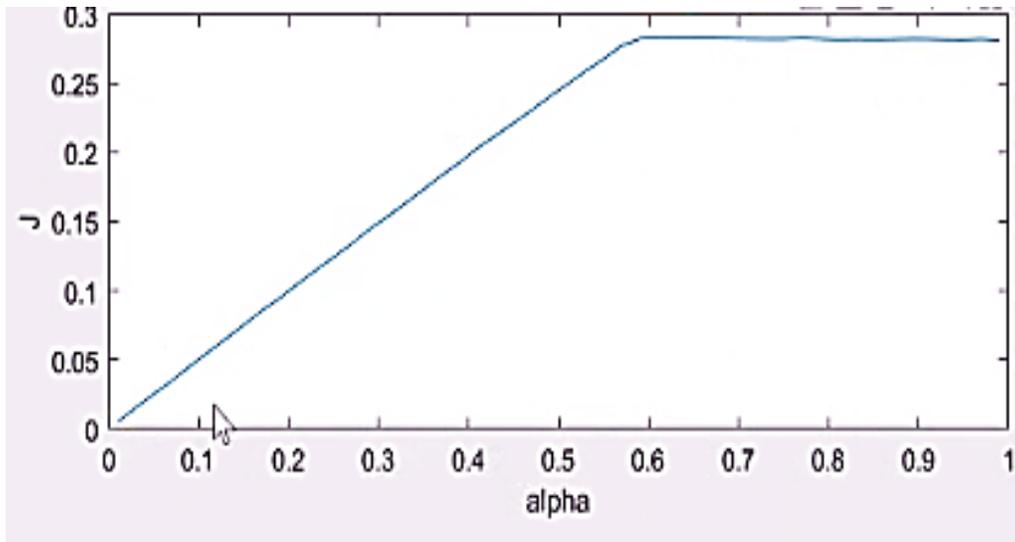


Figura 28. Relación J – $Alpha$.

Relación densidad d – $Alpha$

Se observa en la figura 29, que la densidad aumenta lentamente hasta el punto $alpha=0.55$ aproximadamente y en el punto $alpha=0.6$ se eleva rápidamente. Esto se debe a que existe una alta ocupación de las celdas en la intersección hasta llegar a un punto antes de $alpha=1$ donde la movilidad de los vehículos será prácticamente nula.

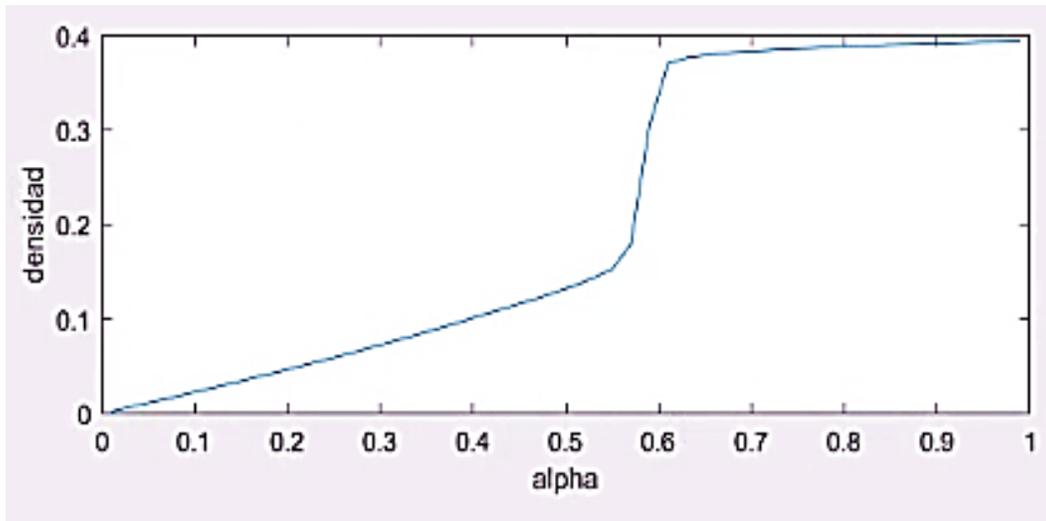


Figura 29. Relación densidad d – Alpha.

Relación cambio de carril Chg – Alpha

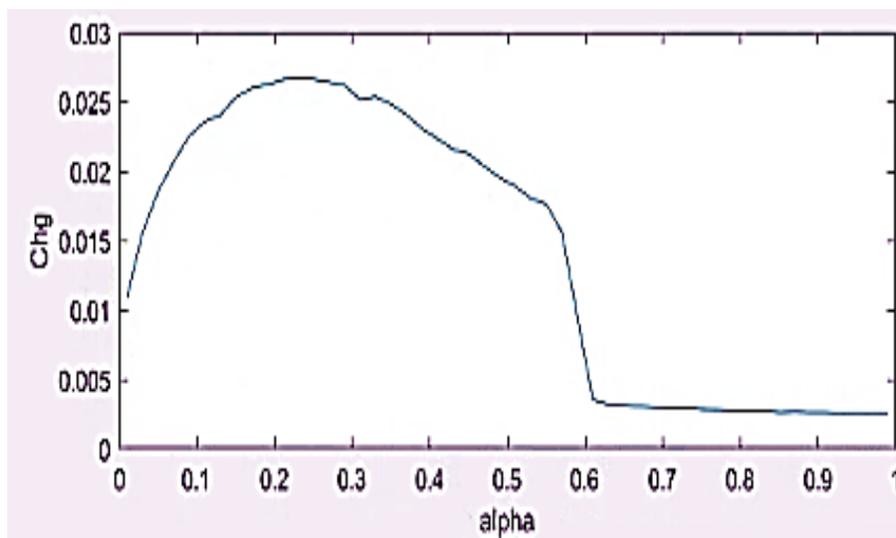


Figura 30 Relación cambio de carril Chg – Alpha

Se observa en la figura 30, que la probabilidad de cambio de carril inicia con un índice bajo y aumenta creando una media luna debido a que en el punto más alto los autos con mayor velocidad intentan adelantar a los de menor velocidad, hasta llegar al punto de $\alpha=0.6$ donde existen menores condiciones de un cambio de carril, pues la densidad vehicular es más alta, después de ese punto la probabilidad de cambiar de carril disminuye considerablemente y se mantiene así ya que las celdas disponibles ya han sido ocupadas.

Relación probabilidad de accidentes $Pacc$ – Alpha:

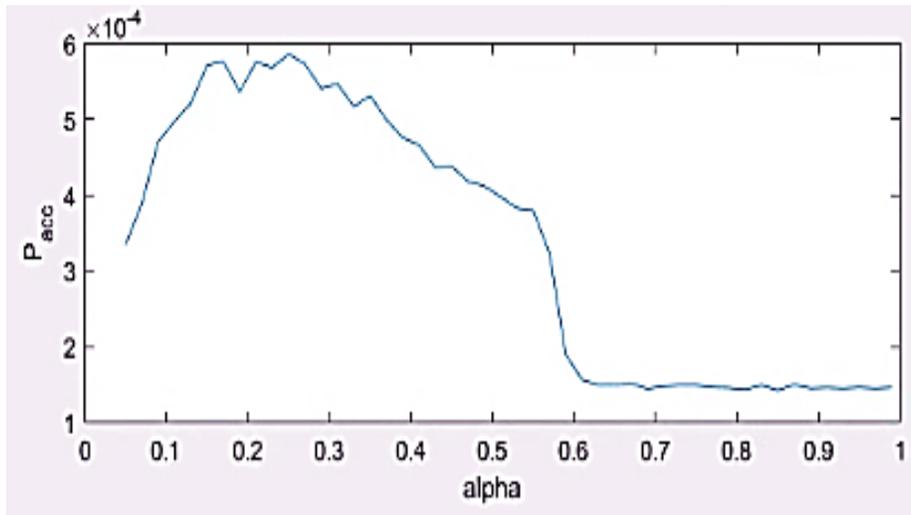


Figura 31. Relación probabilidad de accidentes P_{acc} – Alpha.

Se observa en la figura 31, que la probabilidad de accidentes aumenta desde el punto 0.1 hasta 0.5 puesto que es donde existe un mayor flujo (J) de automóviles al mismo tiempo de una velocidad (V) alta y una mayor probabilidad de cambio de carril (p_{chg}) mayor ocasionando una alta probabilidad de colisión. Esto disminuye en el punto de $alpha = 0.6$ dado que las condiciones anteriormente habladas hacen que disminuyan las variables que condicionan una alta probabilidad de accidente, lo que nos demuestra que el planteamiento del problema donde se simula una intersección señalizada y que existe una probabilidad de accidentes dada.

5.3. Base de datos de la simulación con autómatas celulares

Al realizar la simulación de la intersección con autómatas celulares, logramos obtener una base de datos concisa y suficientemente grande para realizar una predicción de los accidentes viales con redes neuronales artificiales con algoritmos de *machine learning*.

Esta base de datos consiste en 101,501 filas y 14 variables (ver anexo).

Variables de entrada	
1.	Velocidad promedio de los vehículos en el tramo antes de la intersección 1
2.	Velocidad promedio de los vehículos en el tramo antes de la intersección 2
3.	Número de vehículos que entran en vía 1

4.	Número de vehículos que entran en vía 2
5.	Número de vehículos que cambian de carril de izquierda a derecha en la vía 1 (LR1)
6.	Número de vehículos que cambian de carril de derecha a izquierda en la vía 2 (RL2)
7.	Número de vehículos que cambian de carril de izquierda a derecha en la vía 2 (LR2)
8.	Número de vehículos que cambian de carril de derecha a izquierda en la vía 1 (RL1)
9.	Núm. de vehículos que se encuentran en el área 11/tiempo(1s)
10.	Núm. de vehículos que se encuentran en el área 12/tiempo(1s)
11.	Núm. de vehículos que se encuentran en el área 21/tiempo(1s)
12.	Núm. de vehículos que se encuentran en el área 22/tiempo(1s)
13.	Distancia promedio entre vehículos en área 11
14.	Distancia promedio entre vehículos en área 21

Variable de salida	
1.	Accidentes viales

Tabla 1. Variables de entrenamiento.

6. Resultados

6.1. Proceso de entrenamiento y modelo de red neuronal artificial

El proceso de selección de la arquitectura *RNA* consiste en tres partes: selección del algoritmo de propagación hacia atrás, selección de las funciones de transferencia con el mejor efecto en la red y establecimiento del número de neuronas en la capa oculta con el mejor desempeño.

Procedimos a entrenar un modelo de red neuronal artificial obteniendo los siguientes resultados.

- a) Se utilizó la herramienta de *neural net fitting* para realizar un modelo de red neuronal para la predicción de los accidentes.
- b) Se introdujeron los valores de x_n que son las entradas normalizadas de los valores obtenidos de las variaciones de los vectores β y α , corresponde a una única salida que son los accidentes.
- c) Se introducen los valores de validación como se recomienda al hacer modelos de *RNA* para un entrenamiento del 70% validación en 15% y un *testing* de 15%, lo que representa 71,051 valores para entrenamiento, 15,225 valores para validación y 15,225 valores para *testing*.
- d) Posteriormente se introduce el número de neuronas deseada para realizar el entrenamiento de la red neuronal artificial, en este caso se realizaron varias pruebas desde 2 en 2 hasta 20 neuronas.
- e) Se elige el algoritmo para trabajar la *RNA* con mejores referencias correspondiente a Levenberg-Marquardt (LM) y se procede a entrenar la red.

Resultado del entrenamiento con *RNA*

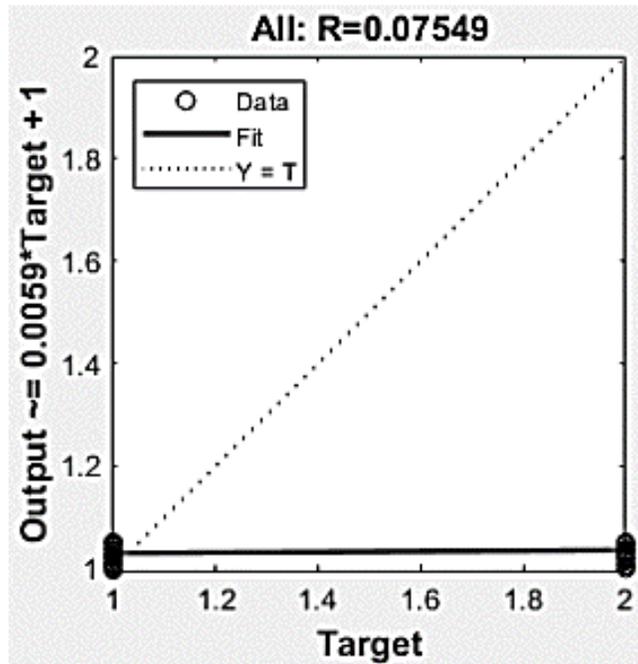


Figura 32. Relación de ocurrencia de accidentes contra no accidentes.

En la figura 32 se observa el coeficiente de determinación con el valor más alto de R , resultado de una arquitectura de red con 10 neuronas en la capa oculta. Las funciones de transferencia de la capa oculta y de salida se establecieron en *tansig* y *purelin* respectivamente. Se estableció el algoritmo LM.

El valor de $R_{all}=0.07549$ significa que no es un modelo aceptable para la predicción de accidentes. Un modelo se considera eficiente cuando se obtienen valores mayores a $R_{all}=0.9$.

El coeficiente de determinación, mide en una escala de 0 a 1. Un valor de 1 indica un modelo que predice perfectamente los valores en el campo de destino. Un valor de 0 indica un modelo que no tiene ningún valor predictivo (IBM, 2023).

Posteriormente se realizaron cambios en la base de datos para mejorar el modelo, como los descritos a continuación:

- En la base de datos se cambió las salidas 1 y 2, donde 1 corresponde a que no hay accidente y 2 a que si lo hay.
- Se modificaron las salidas a 0.01 y 0.09, donde 0.01 corresponde a que no hay accidente y 0.09 que si lo hay.
- Se trabajó con las variables de entrada sin normalizar.

- Se dejó la variable de salida sin normalizar.
- Se realizó el entrenamiento con muestras aleatorias de la base de datos.
- Se omitieron algunas variables de entrada.

Realizando los cambios descritos en los puntos anteriores, no se logró obtener resultados aceptables para un modelo eficiente.

Cabe recordar que las entradas están normalizadas como se muestra en el anexo A.1 adjunto.

6.2. Conclusión de entrenamiento con RNA

Los modelos de RNA utilizan regresión lineal para trabajar con información en el entrenamiento, estos modelos son muy útiles cuando se trata de predecir datos cuya información en sus datos de entrada contienen una relación causal, para predecir una variable a partir de la otra. Por ejemplo la temperatura con la humedad o la velocidad con el tiempo; lo cual es opuesto a nuestro objetivo que consiste en validar si habrá un accidente o no, esto limita solo a dos opciones a un “sí” o un “no” lo cual se resume a un problema de clasificación, el cual se abordara con algoritmos de aprendizaje por clasificación, para entrenar un nuevo modelo y así obtener los resultados necesarios, para la predicción de accidentes en una intersección vial.

6.3. Entrenamiento con classification learner KNN (k vecinos más cercanos)

Se realizó el entrenamiento con el algoritmo de *classification learner* obteniendo los siguientes resultados.

Modelo actual

En la figura 33 se observan las clases en el espacio para la variable 1 con relación a la variable 2 de la base de datos utilizada para entrenar el modelo de clasificación, esta imagen se puede lograr solo con dos tipos de variables ya que resultaría muy difícil el lograr graficar en conjunto las 15 variables utilizadas en el modelo, esto no quiere decir que solo se puedan entrenar dos variables, pues el modelo es más preciso si se utiliza un mayor número de variables con una estrecha relación entre ellas.

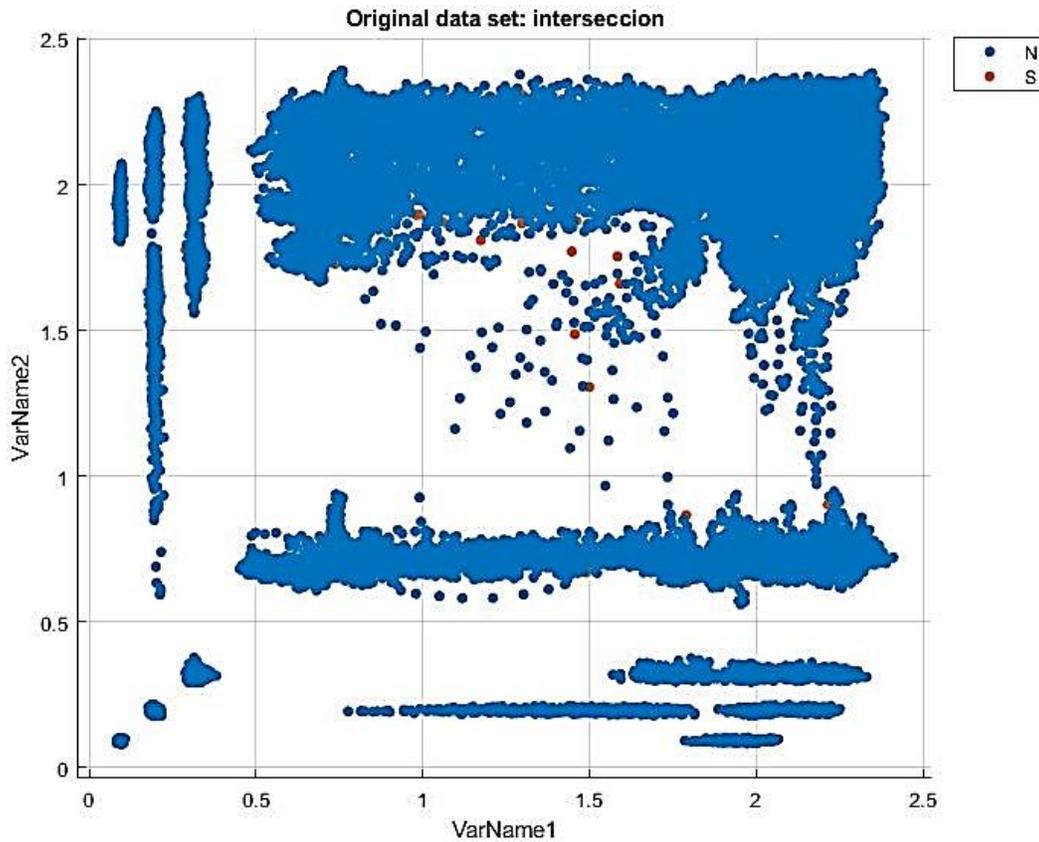


Figura 33. Definición de clases en el espacio de variables.

6.4. Datos de entrenamiento de modelo

- Tipo de modelo preestablecido
- Mejor modelo de KNN optimizado

Número de vecinos	21
Métrica de distancia	Chebyshev
Distancia peso	Inverso al cuadrado.
Base de datos	Datos normalizados
Opción de optimizador	Bayesiano
Función de adquisición	Mejora esperada por segundo (<i>expected improvement per second</i>).
Numero de iteraciones	30

Tabla 2. Datos de entrenamiento para el modelo de kNN.

6.5. Evaluación de rendimiento de algoritmo de clasificación

A partir de la base de datos obtenida resultante de la simulación de la intersección con el modelo de autómatas celulares. Se realiza el entrenamiento con el algoritmo de clasificación k vecinos más cercanos (k nearest neighbors), arrojándonos un modelo capaz de realizar una labor importante de clasificación de datos objetivamente el de los accidentes viales en la intersección señalizada, con valores de $AUC=0.97$, esto quiere decir que cuenta con una eficiencia del 97% para realizar la predicción de cuándo habrá un accidente vial en una intersección señalizada, esto nos lo corrobora la figura 34 con la Curva ROC-AUC.

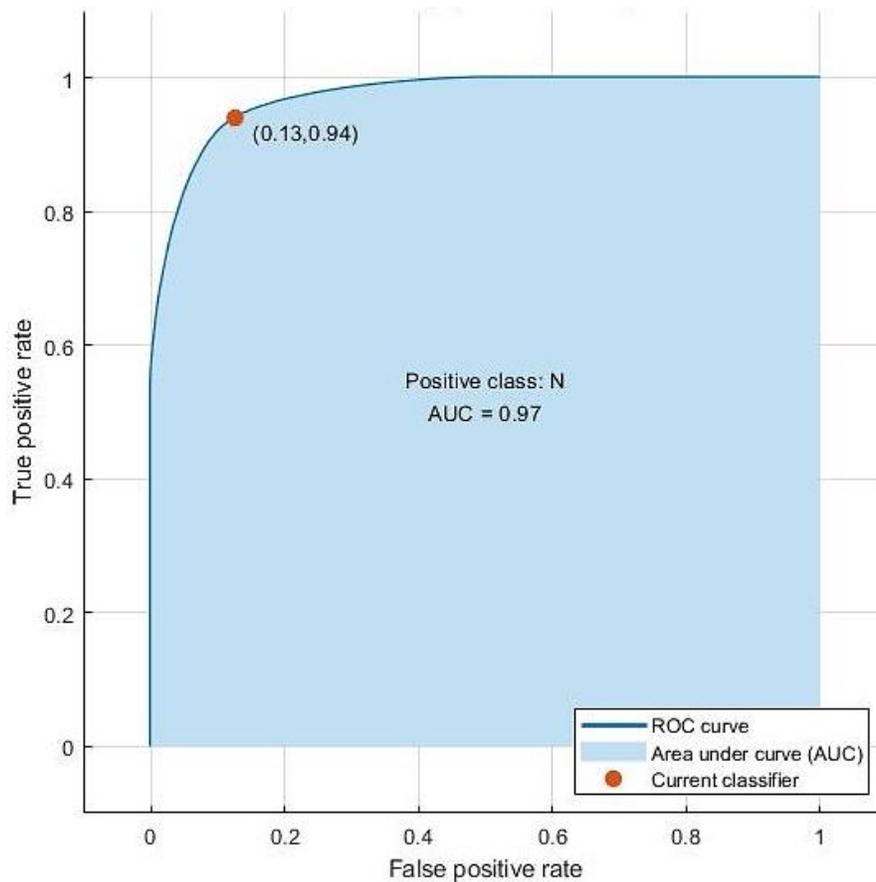


Figura 34. Curva ROC del modelo entrenado.

La curva ROC nos demuestra que nuestro modelo entrenado de KNN es eficiente en la predicción de clasificación de los accidentes en una intersección vial.

6.6. Matriz de confusión

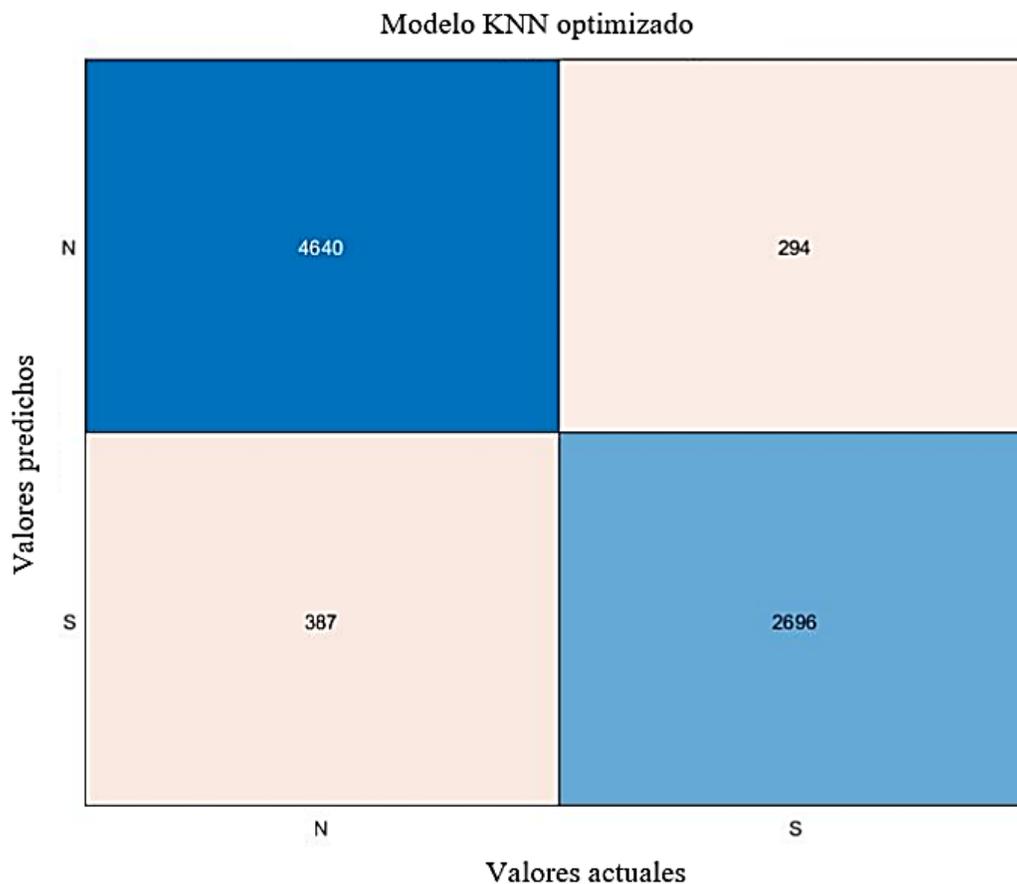


Figura 35. Matriz de confusión del modelo KNN.

En la matriz de confusión fig. 35, demuestra que la clasificación fue muy precisa, con valores mayoritariamente correctos en el entrenamiento de este modelo. Lo cual nos confirma su eficiencia.

Verdadero Positivo: Resultado en el que el modelo predice correctamente la clase positiva con 4,640 valores acertados.

Verdadero Negativo: Resultado donde el modelo predice correctamente la clase negativa con 2,696 valores acertados.

Falso Positivo: El modelo predice incorrectamente la clase positiva cuando en realidad es negativa, en este caso de 294 valores.

Falso Negativo: El modelo predice incorrectamente la clase negativa cuando en realidad es positiva, en este caso fue de 387 valores.

Con la matriz de confusión se ratifica la precisión del modelo de predicción por clasificación, lo cual comprueba la eficiencia de este modelo para predecir los accidentes viales en una intersección señalizada.

7. Conclusión

Se modeló, simuló y predijo una intersección vial señalizada y los accidentes que en ella ocurrían con las herramientas computacionales de autómatas celulares en conjunto con la inteligencia artificial concretamente en el aprendizaje supervisado con el algoritmo de calificación de k vecinos más cercanos (KNN) con excelentes resultados. Se describen a continuación algunas recomendaciones para lograr realizar modelos similares utilizando estas herramientas:

- Trabajar con un equipo de cómputo adecuado con una considerable capacidad de procesamiento como ejemplo un icore5, para reducir el tiempo de simulación de autómatas celulares al igual que el entrenamiento con los algoritmos de aprendizaje automático.
- Realizar un tratamiento de la base de datos que contenga más de cinco variables con una relación cercana; si es necesario realizar una normalización de valores para un mejor entrenamiento; verificar que no tenga valores NaN o faltantes.
- Para seleccionar el algoritmo adecuado para la predicción es recomendable realizar comparaciones entre los algoritmos disponibles según el propósito del modelo a construir, y realizar variaciones en los valores de entrada, realizando una labor heurística.

ANEXO

Se observa en la imagen la base de datos normalizada utilizada en el entrenamiento del modelo de algoritmo de clasificación KNN para predecir los accidentes viales.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1	0.28264421	0.27563944	0.94909862	0.96312849	0.39902784	0.38205499	0.41089772	0.39678676	0.70225606	0.13344234	0.70854668	0.12753507	0.1012037	0.09885124	N
2	0.95926457	0.92271217	0.05037116	0.26648045	0.19929297	0.13603473	0.48414471	0.41626095	0.00663117	0.00544647	0.03628524	0.03032988	0.66133919	0.53791791	N
3	0.05152018	0.55912856	0.4602333	0.44413408	0.05656209	0.05016884	0.68155427	0.66163583	0.87587747	0.87999571	0.0639015	0.15552985	0.02049268	0.2949497	N
4	0.87600554	0.87900699	0.24390244	0.42793296	0.45072912	0.40569223	0.67753461	0.63485881	0.03629585	0.0345643	0.05946926	0.05548275	0.59052071	0.45333026	N
5	0.81615249	0.89736322	0.78419936	0.39441341	0.92311091	0.89966233	0.62393926	0.57254138	0.13057266	0.10883911	0.05581587	0.04725703	0.34569722	0.47150027	N
6	0.27668769	0.81567836	0.9485684	0.65642458	0.38179408	0.39604438	0.80393033	0.7677702	0.70462668	0.14769711	0.09971129	0.09233754	0.10275545	0.37971674	N
7	0.33198236	0.25276535	0.93160127	0.94972067	0.6495802	0.63000482	0.39794551	0.4352483	0.46955188	0.27237003	0.7179038	0.17822479	0.13784488	0.09604385	N
8	0.10467007	0.81910724	0.71792153	0.4301676	0.10384445	0.11143271	0.67396159	0.63437196	0.7858815	0.77404551	0.06138979	0.06583212	0.03513195	0.46138058	N
9	0.25426276	0.72181676	0.93955461	0.82905028	0.37295625	0.38977328	0.87583743	0.87000974	0.70609992	0.20960576	0.13500655	0.14024598	0.1053558	0.30482872	N
10	0.27917761	0.87694439	0.95493107	0.41620112	0.36942112	0.38446696	0.63510496	0.57789679	0.70081091	0.15140432	0.05941452	0.05237342	0.10118288	0.48295346	N
11	0.8639533	0.29196678	0.58165429	0.94916201	0.80114892	0.75397974	0.47565878	0.42307692	0.09009117	0.07659901	0.65850962	0.13026039	0.39888526	0.10516959	N
12	0.85998311	0.90283538	0.57741251	0.23407821	0.73840035	0.68547998	0.42697633	0.3563778	0.09028157	0.076375	0.03420049	0.02691896	0.4113203	0.55037779	N
13	0.00622028	0.00744953	0.17232238	0.18826816	0.01060539	0.01013025	0.01473872	0.00925024	0.97620333	0.98468891	0.98269175	0.98177004	0.00446593	0.00515244	N
14	0.84022889	0.93926475	0.62089077	0.03631285	0.79319487	0.75542692	0.20053595	0.14800389	0.09845153	0.08571069	0.00430399	0.00387111	0.37733969	0.59973699	N
15	0.48217611	0.84697099	0.91781548	0.62513966	0.70393283	0.67390256	0.78517195	0.74196689	0.38426904	0.13373169	0.09559341	0.08012412	0.18204575	0.38746581	N
16	0.75953018	0.80241745	0.7767625	0.61117318	0.91162174	0.90641582	0.80571684	0.78481013	0.12731716	0.12983625	0.09123624	0.0917101	0.32333844	0.38874178	N
17	0.95692197	0.27408536	0.02120891	0.96759777	0.16968626	0.11095031	0.39794551	0.37682571	0.00310568	0.00157591	0.71567987	0.12775739	0.5726981	0.09684729	N
18	0.9315718	0.94014221	0.03022269	0.02905028	0.19133893	0.14664737	0.18758374	0.14410906	0.00390629	0.00422992	0.00296838	0.00375762	0.84956971	0.58261448	N
19	0.10112841	0.88218063	0.70519618	0.23910615	0.11798498	0.11095031	0.45377401	0.41139241	0.78390416	0.78271069	0.03228622	0.03234939	0.02983464	0.54675019	N
20	0.92530414	0.26861223	0.03711559	0.96312849	0.19619973	0.14230584	0.37918714	0.38899708	0.00568386	0.00475107	0.71026003	0.14332887	0.65298746	0.10033079	N
21	0.91624061	0.27534551	0.26002333	0.97039106	0.47370747	0.40328027	0.36802144	0.35248296	0.03572309	0.02891404	0.71167302	0.13210888	0.58910936	0.09540632	S
22	0.28487926	0.26820097	0.94644751	0.96089385	0.42819266	0.41051616	0.38231353	0.37974684	0.68879241	0.14178394	0.71861853	0.13629714	0.10207003	0.09588646	N
23	0.05303916	0.04990241	0.47030753	0.46815642	0.04949183	0.05885191	0.05627512	0.05014606	0.87659693	0.87661053	0.88609428	0.88233703	0.02007424	0.017757	N

Tabla 3. Parte inicial de la base de datos normalizada.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
101481	0.94103991	0.105903	0.02704136	0.75307263	0.1727795	0.12301013	0.20723537	0.16455696	0.00315718	0.00393589	0.75058883	0.77589249	0.63792636	0.03288844	N
101482	0.00592525	0.0061267	0.16808059	0.17932961	0.00618648	0.0130246	0.01205895	0.00827653	0.98403307	0.98591946	0.98185347	0.98449692	0.00531686	0.00481972	N
101483	0.48341981	0.27077127	0.96818664	0.95810056	0.77463544	0.75976845	0.41491737	0.40652386	0.36859708	0.14825561	0.70577067	0.14158765	0.18530133	0.09753933	N
101484	0.90332677	0.10616707	0.04559915	0.75977654	0.21387539	0.17317897	0.10585083	0.10564752	0.00704162	0.0070286	0.78141118	0.77131091	0.64370218	0.03363764	N
101485	0.10625773	0.11068019	0.72269353	0.7849162	0.21166593	0.17221418	0.11299687	0.10954236	0.74677962	0.77773716	0.77428895	0.76486839	0.03388464	0.03511833	N
101486	0.28049314	0.84319627	0.94697773	0.62067039	0.36279275	0.35793536	0.82626172	0.77896787	0.70523533	0.13808762	0.09472073	0.08077708	0.09900379	0.37501556	N
101487	0.93786286	0.27352184	0.02704136	0.96927374	0.17057004	0.11770381	0.39883877	0.39191821	0.00429489	0.0027769	0.71624602	0.12970693	0.58563433	0.09607898	N
101488	0.39189844	0.27305361	0.96818664	0.97039106	0.69730446	0.67824409	0.36623493	0.37779942	0.48021418	0.13880791	0.71193889	0.13665938	0.15193377	0.09766157	N
101489	0.00499494	0.78804511	0.16648993	0.0424581	0.01148917	0.01398939	0.20366235	0.17770204	0.97947756	0.98799164	0.00568653	0.01448633	0.00328521	0.71118366	N
101490	0.7940105	0.27831578	0.78950159	0.97541899	0.91427309	0.90593343	0.40107191	0.38851022	0.1329542	0.11676379	0.70475254	0.1361339	0.33589795	0.09786395	N
101491	0.88404699	0.82732032	0.4135737	0.63743017	0.67565179	0.61312108	0.84814649	0.81791626	0.06164379	0.05501694	0.09567161	0.08826278	0.47063645	0.38238684	N
101492	0.24481072	0.26081834	0.94326617	0.9547486	0.36323464	0.38977328	0.40509156	0.40165531	0.70668829	0.24622512	0.70350919	0.17016232	0.09189683	0.10139263	N
101493	0.10291279	0.10022077	0.71633086	0.74469274	0.12549713	0.11963338	0.0973649	0.09931184	0.77474788	0.78148637	0.79385868	0.7832165	0.02935022	0.03130968	N
101494	0.21192183	0.93012532	0.94326617	0.02793296	0.29695095	0.30824891	0.19338991	0.14410906	0.70663991	0.37826734	0.00303563	0.00345601	0.07715922	0.62126607	N
101495	0.28633298	0.89503225	0.94326617	0.41564246	0.39991162	0.36565364	0.61679321	0.55014606	0.69070107	0.13183531	0.05895941	0.04893451	0.10166637	0.47731418	N
101496	0.29126694	0.27212949	0.95811241	0.95977654	0.47238179	0.45296671	0.3930326	0.39143135	0.68318034	0.13857922	0.70914411	0.13702938	0.10491007	0.09812909	N
101497	0.93399846	0.90352035	0.23382821	0.42178771	0.43393725	0.35890014	0.65073694	0.59347614	0.03351478	0.02716544	0.05920964	0.05013471	0.59665137	0.47349475	N
101498	0.28301667	0.28580529	0.94803818	0.96536313	0.41405214	0.42016401	0.49218401	0.43816943	0.69602598	0.14065607	0.67477002	0.13869598	0.10213255	0.10318785	N
101499	0.26453444	0.25881725	0.9485684	0.95810056	0.43791427	0.47033285	0.42876284	0.44644596	0.69510832	0.19406131	0.71106777	0.17229532	0.11499799	0.10184492	N
101500	0.27563327	0.85168504	0.94697773	0.44413408	0.53247901	0.50506512	0.69673962	0.65384615	0.64198564	0.21963995	0.0635371	0.06114949	0.10443983	0.48008253	N
101501	0.53157276	0.8501762	0.91781548	0.62960894	0.75298277	0.72503618	0.77445288	0.71859786	0.36085161	0.13380326	0.09610639	0.08031068	0.19263706	0.38099234	N
101502															
101503															

Tabla 4. Parte final de la base de datos normalizada.

Bibliografía

- A Aponte, J. M. (2007). Autómatas celulares y su aplicación a la modelación del tráfico automotor en la ciudad de Caracas. *Tekhné*.
- Abuín, J. M. (2007). Regresión lineal múltiple. Madrid, España: Instituto de Economía y Geografía.
- Agency, I. E. (2022). *International Energy Agency*. Obtenido de <https://www.iea.org/>
- AK Daoudia, N. M. (2003). Numerical simulations of a three-lane traffic model using cellular automata. *Chinese journal of physics*.
- AK Das, U. C. (2022). Cellular Automata Model for Lane Changing Activity. *International Journal of Intelligent Transportation*, 446-455.
- AK Das, U. C. (2022). Cellular Automata Model for Lane Changing Activity. *International Journal of Intelligent Transportation*.
- Andreas Schadschneider, D. C. (2011). *STOCHASTIC TRANSPORT IN COMPLEX SYSTEMS*. Amsterdam, The Netherlands: Elsevier BV.
- Ata, A. a. (2019). Modelling smart road traffic congestion control system using machine learning techniques. *Neural Network World*, 99-110.
- Avila-Tomás, J. F.-P.-V. (2020). La inteligencia artificial y sus aplicaciones en medicina I: introducción antecedentes a la IA y robótica. *atención primaria*, 778-784.
- Bailer-Jones, D. M.-J. (2002). Modeling data: Analogies in neural networks, simulated annealing and genetic algorithms. *Model-based reasoning: Science, technology, values*, 147-165.
- Banoula, M. (2023). Classification in Machine Learning: What it is & Classification Models. *Simplilearn* .
- Briega, R. E. (05 de 10 de 2023). *github*. Obtenido de <https://iaarbook.github.io/ML/>
- C Chai, Y. W. (2015). Safety Impacts of Red Light Cameras at Signalized Intersections Based on Cellular Automata Models. *Traffic injury prevention*.

- Camacho, C. (2023). Regresión lineal simple. En U. d. Sevilla. Sevilla.
- Canul, E. J. (2022). *Análisis y Estudio del Tránsito Vehicular Usando Autómatas Celulares y Redes Neuronales*. Cuernavaca.
- Carpio Peláez , K. P., & Oñate Valdivieso, F. (2020). *Redes Neuronales Artificiales Aplicadas En Sistemas De Prediccion Para La Seguridad Vial*. Ecuador.
- CEUPE. (05 de 06 de 2023). *Centro Europeo de Posgrado*. Obtenido de <https://www.ceupe.mx/blog/cuales-son-los-algoritmos-de-clasificacion.html#:~:text=Los%20algoritmos%20de%20clasificaci%C3%B3n%20sobre%20un%20conjunto%20de%20t%C3%A9cnicas%20utilizadas,la%20investigaci%C3%B3n%20y%20la%20academia>.
- Chen, X. a. (2014). A traffic breakdown model based on queueing theory. *Networks and Spatial Economics*, 485-504.
- Cremer, M. a. (1986). A fast simulation model for traffic flow on the basis of Boolean operations. *Mathematics and computers in simulation*, 297-303.
- D Chowdhury, D. W. (1997). Particle hopping models for two-lane traffic with two kinds of vehicles: Effects of lane-changing rules. *Physica A: Statistical*.
- Darcin Akin, B. A. (2010). A neural network (NN) model to predict intersection crashes based upon driver, vehicle and roadway surface characteristics. *Sci. Res. ensayos*, 2837--2847.
- Data, C. q. (22 de 04 de 2021). *UNIR*. Obtenido de Universidad Internacional de La Rioja: <https://www.unir.net/ingenieria/revista/clustering/>
- Deland, S. (05 de 10 de 2023). *MathWorks*. Obtenido de <https://la.mathworks.com/discovery/roc-curve.html>
- Demuth, H., & Beale, M. (2005). *Neural Network Toolbox for use with Matlab, User's Guide Version 4*. Massachusetts, USA: The Math Works, Inc.
- Echab, H. y.-Z. (2015). Cellular automata model simulating traffic car accidents in the on-ramp system. *International Journal of Modern Physics C*.

- Editorial, E. D. (16 de 07 de 2021). *Transporte terrestre*. Obtenido de concepto: <https://concepto.de/transporte-terrestre/>
- Energía, S. d. (2020). *Balance Nacional De Energia de 2020*. Ciudad de México.
- Esser, J. y. (1997). Microscopic simulation of urban traffic based on cellular automata. *International Journal of Modern Physics C*, 1025-1036.
- Fernandez, A. M. (2010). *Modelación y Simulación De Una Via Rápida Con Rampas Usando Autómatas Celulares*. Ciudad de México.
- Fernando Perez, A. B. (2014). Analysis of vehicular traffic flow using a macroscopic model. *DYNA*, 36-40.
- Forum, W. E. (2019). *Calidad de las carreteras-Clasificaciones*. Obtenido de https://es.theglobaleconomy.com/rankings/roads_quality/
- García de Soto, B. a. (2018). Predicting road traffic accidents using artificial neural network models. *Infrastructure Asset Management*, 132--144.
- Greenshields, B. D. (1935). A study of traffic capacity. *Proceedings of the Highway Research*, 448-477.
- Grether, D. a. (2012). Simulation of urban traffic control: A queue model approach. *Procedia Computer Science*, 808-814.
- Guadarrama, G. (17 de Diciembre de 2019). *Portal*.
- Han Tao Zhao, X. R. (2017). Cellular automata model for traffic flow at intersections in. *Physica A*, 40-51.
- Helbing, D. (2001). Traffic and related self-driven many-particle systems. *Reviews of modern physics*, 1067.
- Hoogendoorn, S. y. (2001). State-of-the-art of vehicular traffic flow modelling. *Journal of Systems and Control Engineering*, 283-303.
- IBM. (28 de 11 de 2023). *Cognos Analytics*. Obtenido de <https://www.ibm.com/docs/es/cognos-analytics/11.1.0?topic=terms-r2>

- INSP. (2020). *Instituto Nacional De Salud Publica*. Obtenido de Gobierno de México: <https://www.insp.mx/avisos/4761-seguridad-vial-accidentes-transito.html>
- J Wang, W. L. (2021). cellular automata model based on multiple agents for the simulation of control in intersections. *Physica A: Statistical Mechanics and its Applications*.
- Jia, B. a. (2020). Car-following model considering the lane-changing prevention effect and its stability analysis. *The European Physical Journal B*, 1-9.
- Jin, Y. a. (2020). Dynamical analysis of an optimal velocity model with time-delayed feedback control. *Communications in Nonlinear Science and Numerical Simulation*.
- Kloska, S. P. (2021). Queueing theory model of Krebs cycle. *Bioinformatics*, 2912-2919.
- Lee, S. a.-E. (2019). Integrated deep learning and stochastic car-following model for traffic dynamics on multi-lane freeways. *Transportation research part C: emerging technologies*, 360-377.
- M Fukui, Y. I. (1996). Traffic Flow in 1D Cellular Automaton Model Including Cars Moving with High Speed. *Journal of the Physical Society of Japan*.
- Machines, I. B. (02 de 10 de 2023). *IBM*. Obtenido de <https://www.ibm.com/mx-es/topics/chatbots>
- Maerivoet, S. a. (2005). Transportation planning and traffic flow models. *arXiv preprint physics*.
- Mata, F. B. (2020). Macrodatos, inteligencia artificial y proceso: luces y sombras. *Revista General de Derecho Procesal*, 51.
- Matich, D. J. (2001). *Redes Neuronales: Conceptos Básicos y Aplicaciones*. Rosario, Argentina: Universidad Tecnológica Nacional – Facultad Regional Rosario.
- Mexico, C. (06 de 2017). Obtenido de <http://cesvimexico.com.mx/index.html>
- MK Singh, K. R. (2022). Selection of open or closed boundaries in a cellular automata model for heterogeneous non-lane-based traffic. *SIMULACIÓN*.

- Mohammad Shareef Ghanim, K. S. (2019). Estimating Turning Movements at Signalized. *IEEE TRANSACTIONS ON INTELLIGENT TRANSPORTATION SYSTEMS*, 1828-1836.
- Montt, C., Rodriguez, N., Valencia, A., & Contreras, Á. (2015). *Predicción De Accidentes De Transito En La Region Metropolitana De Chile Basado En Redes Neuronales*. Pontificia Universidad Católica de Valparaíso , Chile.
- Morales Fernandez, A. (2010). *Modelación y Simulación De Una Via Rápida Con Rampas Usando Autómatas Celulares*. Ciudad de México.
- Murcia, U. d. (05 de 10 de 2023). *um.es*. Obtenido de <https://www.um.es/geograf/sigmur/teledet/tema09.pdf>
- N Boccara, H. F. (1997). Accidentes de tráfico y número de coches detenidos por bloqueo de vía en una vía de un solo carril. *Journal of Physics A: Matemática y General*.
- Nagel, K. (1996). Particle hopping models and traffic flow theory. *Phys. Rev. E, American Physical Society*, 4655--4672.
- Nagel, K., & Schreckenberg, M. (1992). A celular automaton for freeway traffic. *Journal de Physique I*, 2221-2229.
- Nagel, K., & Schreckenberg, M. N. (1992). *A celular automaton for freeway traffic*. France.
- Olayode, I. O. (2021). Development of a hybrid artificial neural network-particle swarm optimization model for the modelling of traffic flow of vehicles at signalized road intersections. *Applied Sciences*, 8387.
- OMS. (9 de Diciembre de 2020). *Organización Mundial de la Salud*. Obtenido de <https://www.who.int/es/news-room/fact-sheets/detail/the-top-10-causes-of-death>
- OMS. (20 de junio de 2022). *Organización Mundial de la Salud*. Obtenido de <https://www.who.int/es/news-room/fact-sheets/detail/road-traffic-injuries>
- Pang, M.-B. a.-N. (2017). Effects of rainy weather on traffic accidents of a freeway using cellular automata model. *Chinese Physics B*.

- Payne, H. J. (1979). FREFLO: A macroscopic simulation model of freeway traffic. *Transportation Research Record, National Academy of Sciences*, 68-77.
- Pérez Porto, J. G. (15 de 7 de 2022). Obtenido de definicion.de: <https://definicion.de/transito/>
- PIARC. (22 de 10 de 2023). *MANUAL EXPLOTACIÓN DE LA RED VIAL*. Obtenido de <https://rno-its.piarc.org/es/conceptos-basicos-its-tecnologias-its-datos-e-informacion/sensores-viales>
- Ramírez-Agundis, A. (1 de 12 de 2023). La neurona biológica. *Instituto Tecnológico de Celaya*. México: https://www.researchgate.net/publication/50837616_Diseño_y_experimentación_de_un_cuantizador_vectorial_hardware_basado_en_redes_neuronales_para_un_sistema_de_codificación_de_video/figures?lo=1.
- Raschka, S. (2018). Machine Learning Lecture Notes. *University of Wisconsin*.
- Rocha, C., & Escorcía G., J. (2010). *Sistema de Vision Artificial para la Detección y el Reconocimiento de Señales de Tráfico basadas en Redes Neuronales*. Peru.
- Rodríguez, J. L. (01 de 12 de 2023). Modelo de una neurona artificial. Colombia.
- Rui Jiang, Q.-S. (2003). Cellular automata models for synchronized traffic flow. *JOURNAL OF PHYSICS A: MATHEMATICAL AND GENERAL*, 381-390.
- S. Cheybani, J. K. (2000). Nondeterministic Nagel-Schreckenberg traffic model with open boundary conditions. *PHYSICAL REVIEW E*.
- Salat Figols, R. S. (2014). *Tráfico de vehículos con autómatas celulares*. Instituto Politécnico Nacional, Departamento de Matemáticas, Escuela Superior de Física Y Matemáticas, CDMX.
- Salcido, A. (2015). *Tres iniciativas de Innovación Aplicando Modelos de Autómatas Celulares*. Instituto de Investigaciones Eléctricas, División de Energías Alternas, Cuernavaca, Morelos, México.
- Sandoval, L. J. (2018). Machine Learning Algorithms For Data Analysis And Prediction. *ITCA FEPADE*.

- Schadschneider, A. (2002). Traffic flow: a statistical physics point of view. *Physica A: Statistical Mechanics and its Applications*, 153-187.
- Schalkoff, R. (1997). *Artificial neural networks*. McGraw-Hill Higher Education.
- SEMARNAT, S. d. (2018). *Inventario Nacional de Emisiones de Gases Y Compuestos De Efecto Invernadero 1990-2015*. Ciudad de México.
- Servan, W. J. (02 de 05 de 2022). K-Vecinos mas cercanos(KNN).
- Sharaf Aljeder, M. M. (2016). Severity Prediction of Traffic Accident Using an Artificial Neural Network: Traffic Accident Severity Prediction Using Artificial Neural Network. . *Journal of Forecasting*.
- Soh, A. C. (2007). Modelling and Optimisation of a Traffic Intersection Based on Queue Theory and Markov Decision Control Methods. *First Asia International Conference on Modelling & Simulation*, 478-483.
- Stuart Russell, P. N. (2004). *Inteligencia Artificial Un Enfoque Moderno*. Madrid: Pearson educacion.
- Toth, G. (16 de 02 de 2020). K-Nearest Neighbors. *datasklr*. <https://www.datasklr.com/select-classification-methods/k-nearest-neighbors>.
- W Knospe, L. S. (2002). A realistic two-lane traffic model for highway traffic. *Journal of Physics A: Mathematical and General*.
- Yang, X.-q. a.-q.-m. (2004). Effects of quenched randomness induced by car accidents on traffic flow in a cellular automata model. *Physical Review E*.
- Yasdi, R. (1999). Prediction of Road Traffic using a Neural Network. *Informática neuronal y aplicaciones*, 135-142.
- Younes Regragui, N. M. (2018). A Cellular Automata Model for Urban Traffic with. *HAL open science*, 18.
- Zinat Matin, H. N. (2020). Nonlinear Optimal Velocity Car Following Dynamics (II): Rate of Convergence In the Presence of Fast Perturbation. *National Science Foundation*.



UNIVERSIDAD AUTÓNOMA DEL
ESTADO DE MORELOS

FACULTAD DE CIENCIAS QUÍMICAS e INGENIERÍA

Programas educativos de calidad reconocidos por CIEES, CACEI y CONACYT
SGI Certificada en la norma ISO 9001:2015 e ISO 21001:2018



Facultad de Ciencias
Químicas e Ingeniería

FORMATO T-2
NOMBRAMIENTO DE COMISIÓN REVISORA
Y DICTAMEN DE REVISIÓN

Cuernavaca, Mor., a 7 de diciembre de 2023.

Dr. Outmane Oubram
Dr. Rachid Marzoug
Dr. Mario Acosta Flores
Dr. Roy Lopez Sesenes
Dr. Pedro Moreno Bernal
PROFESORES DE LA FCQeI
P R E S E N T E

Por este conducto, me permito informarle que ha sido asignado como integrante de la Comisión Revisora de la tesis que presenta el Ing. Pablo Humberto Hernández Peralta titulada: Modelación, Simulación Y Predicción De Los Accidentes Viales En Una Intersección Mediante Autómatas Celulares E Inteligencia Artificial realizada bajo la dirección del Dr. Outmane Oubram en el Programa Educativo de Maestría en Ingeniería Eléctrica y Electrónica. Agradezco su valiosa participación en esta Comisión y quedo a sus órdenes para cualquier aclaración o duda al respecto.

A T E N T A M E N T E
Por una humanidad culta

MTRA. ANGÉLICA GALINDO FLORES
DIRECTORA INTERINA DE LA FCQeI

D I C T A M E N

MTRA. ANGÉLICA GALINDO FLORES
DIRECTORA INTERINA DE LA FCQeI
P R E S E N T E

En respuesta a su amable solicitud para formar parte de la Comisión Revisora de la tesis mencionada y una vez realizada la revisión correspondiente, me permito informarle que mi VOTO es:.

D I C T A M E N

NOMBRE	VOTO	FIRMA
Dr. Outmane Oubram	Aprobatorio	*Se anexa firma electrónica
Dr. Rachid Marzoug	Aprobatorio	*Se anexa firma electrónica
Dr. Mario Acosta Flores	Aprobatorio	*Se anexa firma electrónica
Dr. Roy Lopez Sesenes	Aprobatorio	*Se anexa firma electrónica
Dr. Pedro Moreno Bernal	Aprobatorio	*Se anexa firma electrónica



UNIVERSIDAD AUTÓNOMA DEL
ESTADO DE MORELOS

Se expide el presente documento firmado electrónicamente de conformidad con el ACUERDO GENERAL PARA LA CONTINUIDAD DEL FUNCIONAMIENTO DE LA UNIVERSIDAD AUTÓNOMA DEL ESTADO DE MORELOS DURANTE LA EMERGENCIA SANITARIA PROVOCADA POR EL VIRUS SARS-COV2 (COVID-19) emitido el 27 de abril del 2020.

El presente documento cuenta con la firma electrónica UAEM del funcionario universitario competente, amparada por un certificado vigente a la fecha de su elaboración y es válido de conformidad con los LINEAMIENTOS EN MATERIA DE FIRMA ELECTRÓNICA PARA LA UNIVERSIDAD AUTÓNOMA DE ESTADO DE MORELOS emitidos el 13 de noviembre del 2019 mediante circular No. 32.

Sello electrónico

ANGELICA GALINDO FLORES | Fecha:2023-12-07 10:17:58 | Firmante

fx3HQzu/OMqm4rzgA3+3Oi4fQC+mb141R0WZdzmgCygpnpo93rTOerCx0GuDxwhzxGUbgBKa1JkwhXGJ7EtMzjr4iUHD4t72r5xer/UBjn4L/S5Xt+Klo7fSgH7vm5tSo7Dsk69YAb4qQbfHGndNAoTqnN6OP/KKLMf+T2ObNswR/L5eXOh2zxk4rKmvn7gbDUZQdx0dM/NmOXMKUp5AubM1ZUIIKxIBczO99r0dsyV0KrO4xRnfKu99bYsYXYK/4Q4wObIQ778D/VogMcgGJwoyaSggzt4DGSBhgYU/jWFIINUET4i+yhw4GhGMEk9NkULpLVIAWMDljHhK7NKtg==

Puede verificar la autenticidad del documento en la siguiente dirección electrónica o escaneando el código QR ingresando la siguiente clave:



[XS0mo4WhT](#)

<https://efirma.uaem.mx/noRepudio/PstJ3gHRDt3Ps08sVsiMzWNhGfsI7Hkx>





UNIVERSIDAD AUTÓNOMA DEL
ESTADO DE MORELOS

Se expide el presente documento firmado electrónicamente de conformidad con el ACUERDO GENERAL PARA LA CONTINUIDAD DEL FUNCIONAMIENTO DE LA UNIVERSIDAD AUTÓNOMA DEL ESTADO DE MORELOS DURANTE LA EMERGENCIA SANITARIA PROVOCADA POR EL VIRUS SARS-COV2 (COVID-19) emitido el 27 de abril del 2020.

El presente documento cuenta con la firma electrónica UAEM del funcionario universitario competente, amparada por un certificado vigente a la fecha de su elaboración y es válido de conformidad con los LINEAMIENTOS EN MATERIA DE FIRMA ELECTRÓNICA PARA LA UNIVERSIDAD AUTÓNOMA DE ESTADO DE MORELOS emitidos el 13 de noviembre del 2019 mediante circular No. 32.

Sello electrónico

OUTMANE OUBRAM | Fecha:2023-12-07 12:56:34 | Firmante

Z501XDdimGnQTKd4zx5TXyOqFYFr5m+hbCgn1RIJQ/ILobVY9L9GyzHvmokxk6pfHfFRd/eTnODDc9ezSRfwtfBQouc6zZinmVlWU9nOdXmNmKPePzBBaLV8G833z3gWwPs5yld/tNTXgolNRwkfTQtB+p1mzR6mQuJY9O6O6LwY/0hYf3wnvAFQHIPdVof4P8MiOjTczLcuxCjSY6Xmj6P9eDQn5WGSspjZXFIfAe0j41D5fZxA9r0X948WBNPwtmAl9cD H3uia2wnCs0eM6mKKbNmDNLczOhTE+PBnMPDjvvlrcWLABwQFfkg7EilHBe/QBkPRIH3A==

MARIO ACOSTA FLORES | Fecha:2023-12-07 13:12:25 | Firmante

GAVd+SmFCDDMZx5vYWWkZqxSYh7TJneNXC4QG1hxOZKvAzquJYELFGnfjlz3gBf6uUbAUI+2lBytdhkh7AGKdmm0iM0G4ajle7s5p419m5xfb3iTJbQj8Ja6N1s/s776yAYFxoxx8jvl/7jhOsdM+OpnzHhPLsZxhvVyn5bi83D8PxoHN07vZ532igogAUj6q4ZrRdDDSG/j/v/mpQhx8hmUkH5XW3V8sKx8LElxh1Te+RUcxy73NjXUeMgy42qxpEUcNyl1NhEviVaWxP5ycAYA50QbRdnnyGxSzxwGB/2MMW6Mu1vC2HrsToE5dS+3Kft3k2aaXHoeF8/uyEA==

ROY LOPEZ SESENES | Fecha:2023-12-07 13:21:50 | Firmante

qQqJYIQF7JHcmCL1UgJ/90/D83N3LggoxSn/tN2ZgeQV8neyHX5/5klywTFDotboAslX11fS7KooOltPMkr5skSmcy0Q3VtQ+VK0TMJL1cQlsXLxq75WxTdOGxs2SLmg3RyXisLMZu/jPGIEdz+cl4f/IV+dp/7cJl58Oge8pw6Uos7HuqO5bRFSdy0yqYa1LY9fLPlaob+Ut1C+qn2qODuT+oyo9CrKnXgtWga+fy3rayBS7VznmjLJ8vGZHsoAdlYe8rVQASQcYyteUwzKbBU6laZWP6fLKTMeDciiVrCsnvLzb+K2kh8vM+pwB7qiepsi1EJ6IC1TQPK7A==

RACHID MARZOUG | Fecha:2023-12-07 14:29:45 | Firmante

WQD71NFr4FrqhdY0vs1+HiHcG260lcyMR+OMFU+wDIWtF/S5IMQRnpYJb5wzglscJFpaUhcEk+mnT2fMLZFWG8VPhzehp7MYB6BgubHoAtqE/iRjoKnyoEHswEtLGilzGN54OpOCiAbEmAtHZBtSD+udlSNiS0lQpMAormvNuHFCGlgvCkdSolq/V306eugWw7suluPLphgJ0Um6bkM4c/YljFsyh657pRlloYkdyhmEFiHRxxMf5+DFsI82fVX3Wl1760LYerZRLds2SLEZ2BWJPQXOVk2SYiZ5PaAgUZkkLBBs9FaucoFsRe8TfjBvxu4z/bMCNzH4k6+2Csw==

PEDRO MORENO BERNAL | Fecha:2023-12-12 10:49:26 | Firmante

hd9MqhTBLA6FCk6SjHi0aR+Qf/5z23xxzX9+ykeDlhAxQtpygf1F0oCcVT9tEkj6STLboLw5K7AZAP/otatOaZPY0vzrdLVIoehGEJyREz52hzf6vwGCg2xV9AvNYZPCmUkwSxOg37rVfweudR1Ti46x2vQj+AFA55KbB/6viEkUeStQuiBYl6jSL9NkiCylGnD//OnLof8oQKMXd5qam6T/Ki7+O9wfnzG/orQcp4v2pbfmBcGvjieiGWedW/8OSCjmw/6ihQb+CRXEETfKU5TFXfoOHh+/e8jTJADnuD8k6EstEn6YwMCwNzjz8MLIBlcwKsYYIoKr0wSrELA==

Puede verificar la autenticidad del documento en la siguiente dirección electrónica o escaneando el código QR ingresando la siguiente clave:



sjgYxwXGV

<https://efirma.uaem.mx/noRepudio/ZHBw4PDfijE6PCOAiwjCvYZMGxX76t>

