

Metodología basada en reportes de siniestros para la optimización de la gestión Municipal sobre seguridad vial.

Perez, Angueira Luciana; Marcos, Carlos; Gasselle, Gonzalo; Martinez Micakoski, Fernanda; Enrietti, Adhemar Raul;

Universidad Tecnológica Nacional Facultad Regional Trenque Lauquen; CP: 6400 Racedo 298
lucianaperezanguera@hotmail.com

RESUMEN.

El presente trabajo tiene como propósito compartir la metodología que se llevará a cabo en el proyecto de investigación que tiene como objetivo generar algoritmos que apoyen la gestión de los recursos para reducir la siniestralidad vial en municipios que posean entre 30.000 y 60.000 habitantes.

El caso de aplicación es la ciudad de Trenque Lauquen, pudiendo replicarse en otros municipios que posean estadísticas respecto de la siniestralidad urbana y la gravedad de las lesiones provocadas durante los últimos años.

Para el tratamiento de los datos se implementa el software libre R, que facilita la aplicación de técnicas estadísticas variadas como modelos lineales y no lineales, análisis de series de tiempo, clasificación, agrupamiento, análisis de correlación y visualización acordes al proyecto a desarrollar. Los modelos de predicción generados serán utilizados por el Municipio para apoyar la toma de decisiones sobre el lugar, el momento y a quiénes serán dirigidos los operativos de tránsito en el área urbana con la finalidad de reducir la gravedad de las lesiones causadas por la siniestralidad vial.

La asignación de recursos es algo crucial para la efectividad de los programas de seguridad vial. Por tanto, los procesos que la conforman deben estar vinculados a objetivos específicos para poder alcanzar el máximo rendimiento de los fondos asignados, para ello se utilizará la metodología SEMMA que permite obtener conclusiones que se utilicen en la proyección de asignación de recursos para el trabajo periódico de los agentes de tránsito.

La reducción del número de heridos y muertos por accidentes de tránsito mitigará el sufrimiento, desencadenará el crecimiento y liberará recursos para su utilización más productiva.

Palabras Claves: metodología, SEMMA, siniestros Viales, R, seguridad vial.

ABSTRACT (Resumen en Inglés)

The purpose of this paper is to share the methodology that will be carried out in the research project that aims to generate algorithms that support the management of resources to reduce road accidents in municipalities with between 30,000 and 60,000 inhabitants.

The case of application is the city of Trenque Lauquen, which can be replicated in other municipalities that have statistics regarding urban accidents and the severity of injuries caused in recent years.

For the treatment of the data, the free software R is implemented, which provides varied statistical techniques such as linear and non-linear models, time series analysis, classification, grouping, correlation analysis and graphs applicable to the project to be developed.

The prediction models generated will be used by the Municipality to support decision making about the location, timing and to whom traffic operations in the urban area will be directed in order to reduce the severity of injuries caused by road accidents.

The allocation of resources is crucial for the effectiveness of road safety programs.

Therefore, the processes that make it up must be linked to specific objectives in order to achieve maximum performance of the funds allocated, for this will use the SEMMA methodology that allows conclusions to be drawn that are used in the projection of resource allocation for the periodic work of transit agents.

Reducing the number of people injured and killed in traffic accidents will alleviate suffering, trigger growth and free up resources for more productive use.

1. INTRODUCCIÓN

En su resolución 64/255, de marzo de 2010, la Asamblea General de las Naciones Unidas proclamó el periodo 2011-2020 «Decenio de Acción para la Seguridad Vial», con el objetivo general de estabilizar y, posteriormente, reducir las cifras previstas de víctimas mortales en accidentes de tránsito en todo el mundo aumentando las actividades en los planos nacional, regional y mundial. En la resolución se solicita a la Organización Mundial de la Salud y a las comisiones regionales de las Naciones Unidas que, en cooperación con otros asociados del Grupo de colaboración de las Naciones Unidas para la seguridad vial y otros interesados, preparen un plan de acción del Decenio como documento orientativo que facilite la consecución de sus objetivos.

Este Plan mundial será útil para apoyar el desarrollo de planes de acción local y nacional y, al mismo tiempo, ofrecerá un marco para favorecer la realización de actividades coordinadas a nivel regional y mundial. Está destinado a un amplio público, en particular los gobiernos locales y nacionales, la sociedad civil y las empresas privadas que deseen ajustar sus actividades a la consecución del objetivo común, manteniendo una perspectiva genérica y flexible, conforme con las necesidades de los países.

Cada año, en el mundo, cerca de 1,3 millones de personas fallecen a raíz de un accidente de tránsito —más de 3.000 defunciones diarias— y más de la mitad de ellas no viajaban en automóvil. Entre 20 millones y 50 millones de personas más sufren traumatismos no mortales provocados por accidentes de tránsito, y tales traumatismos constituyen una causa importante de discapacidad en todo el mundo. El 90% de las defunciones por accidentes de tránsito tienen lugar en los países de ingresos bajos y medianos, donde se halla menos de la mitad de los vehículos matriculados en todo el mundo. Entre las tres causas principales de defunciones de personas de 5 a 44 años figuran los traumatismos causados por el tránsito. Según las previsiones, si no se adoptan medidas inmediatas y eficaces, dichos traumatismos se convertirán en la quinta causa mundial de muerte, con unos 2,4 millones de fallecimientos anuales. Ello se debe, en parte, al rápido aumento del mercado de vehículos de motor sin que haya mejoras suficientes en las estrategias sobre seguridad vial ni la planificación del uso del territorio. Se ha estimado que las colisiones de vehículos de motor tienen una repercusión económica del 1% al 3% en el producto bruto nacional (PNB) respectivo de cada país, lo que asciende a un total de más de \$ 500.000 millones. La reducción del número de heridos y muertos por accidentes de tránsito mitigará el sufrimiento, desencadenará el crecimiento y liberará recursos para una utilización más productiva.

En las Américas, la tasa de mortalidad por accidentes de tránsito cada 100.000 habitantes es en promedio de 16,1 mientras que en la Argentina se estima en 13,6. [1]

La tasa de siniestros por cada 100.000 automóviles en nuestro país asciende a 71,99 según datos del Observatorio Vial para el año 2014. Mientras que las lesiones culposas en siniestros viales cada 100 mil habitantes en el año 2014 totalizaron 225 [2].

En el plano nacional, se alienta a los países a que apliquen los cinco pilares siguientes, sobre la base de las recomendaciones del Informe mundial sobre prevención de los traumatismos causados por el tránsito, que propone la Comisión para la Seguridad Vial Mundial.

Los países deberían considerar estas cinco áreas en el marco de su propia estrategia nacional de seguridad vial, su capacidad en esta materia y sus sistemas de recopilación de datos.

- Pilar 1: Gestión de la seguridad vial
- Pilar 2: Vías de tránsito y movilidad más seguras
- Pilar 3: Vehículos más seguros
- Pilar 4: Usuarios de vías de tránsito más seguros
- Pilar 5: Respuesta tras los accidentes

Dentro de las actividades del *Pilar 1* se recomienda trabajar para garantizar que se disponga de fondos suficientes para la ejecución de actividades, mediante el fomento del establecimiento de procedimientos para la asignación eficiente y eficaz de recursos en todos los programas de seguridad [3].

Es esencial que cada país ponga en marcha un sistema científico y coherente de recolección, almacenamiento, análisis, difusión y aplicación de datos de accidentes de tránsito. [1]

Por ello, la comunidad internacional debe ahora asumir el liderazgo y promover buenas prácticas en la gestión de la seguridad vial y la adopción de esas intervenciones en otros países, adaptándolas a sus entornos particulares. [1]

Sin un diagnóstico y una gestión permanentes, y basados en datos, de los principales problemas relacionados con los traumatismos causados por el tránsito, no se lograrán reducciones significativas y sostenibles de la exposición al riesgo de accidentes ni de la gravedad de estos.

Como mínimo, un buen sistema de datos de accidentes de tránsito debería: capturar casi todos los accidentes con víctimas mortales y una proporción significativa de los causantes de traumatismos graves; ofrecer detalles suficientes sobre el vehículo, el usuario de la vía, y la vía y su entorno para ayudar a identificar las causas y a seleccionar las contramedidas; incluir información exacta sobre la ubicación del accidente; proporcionar oportunamente productos fiables para facilitar la toma de decisiones basadas en pruebas. [1]

Sobre éste último tópico se enfoca el actual proyecto. Con los procesos de adquisición de datos ya establecidos en los últimos cuatro años en la ciudad, la información generada y una concientización de la población y los funcionarios sobre la importancia de abordar este problema se requiere de herramientas de gestión que permitan asignar eficientemente los recursos que el Municipio dedica a la seguridad vial de manera que se transformen en una solución sostenible a largo plazo.

Dado que los recursos disponibles siempre resultan escasos en relación con el alto beneficio que se espera es necesario establecer criterios de priorización que aplicados sobre los datos relevados permitan gestionar eficientemente los recursos humanos y vehiculares de los agentes de tránsito.

En la ciudad de Trenque Lauquen existe una base de datos donde se registran los siniestros viales desde el año 2012 a la actualidad con información sobre el momento, lugar, condiciones del entorno y climatológicas, movilidad y perfil de los participantes de los siniestros, incluyendo la gravedad de las lesiones mediante la trazabilidad de las víctimas en los centros de salud.

Mediante la aplicación de técnicas de aprendizaje supervisado (machine learning) se apoyará la toma de decisiones del organismo de Contralor Municipal para ubicar los lugares, momentos, vehículos y perfil de los conductores que pertenezcan al segmento de mayor riesgo de participar en un siniestro vial en base a los datos históricos recabados en dicha localidad.

1.1. Objetivo general

Desarrollar una herramienta que permita proponer la asignación en espacio y tiempo de los recursos humanos y vehiculares de la Dirección de Contralor para reducir la siniestralidad vial en la ciudad.

1.2. Objetivos secundarios

Representar las asignaciones mediante un Sistema de Información Geográfica agrupado por cuadrículas.

Realizar el seguimiento, control y ajuste de la plataforma que sistematiza los datos de siniestralidad vial generados por Agentes de Tránsito Municipales, Bomberos Voluntarios y Agentes Sanitarios Municipales,

Establecer indicadores de performance de la herramienta de gestión que permitan evaluar la eficacia de los algoritmos desarrollados.

2. METODOLOGIA

El estudio actual se basa en los datos relevados a partir de proyectos de investigación desarrollados desde el año 2012 a la fecha en la ciudad de Trenque Lauquen, la cual se ubica en el noroeste de la provincia de Buenos Aires y posee aproximadamente 50.000 habitantes.

Se establecen una serie de datos a relevar, con un formato de categorización y descripción específica al momento de observar la ocurrencia de un siniestro vial. Los mismos se pueden observar en la Tabla 1.

Tabla 1 Datos relevados en los siniestros viales.

Variable	Categoría	Descripción
Fecha	Tiempo	Día/Mes/Año
Horario		HH:mm
Zona	Lugar	Ruta/Rural/Suburbana/Urbana
Calle		Nombre de la calle (texto)
Altura		Altura de la calle (número entero)
Entre calles		Calles próximas (texto)
Lugar	Vía Pública	Curva/Intersección/Recta
Tipo		Calle/Distribuidor/Rotonda/Paso FFCC/Otro
Superficie		Arena/Asfalto/Concreto/Ripio/Tierra/Otro
División		Cordón o bulevard/Guardarrail/Línea Divisoria/Ninguna/Otro
Estado		Bueno/Baches/Ahuellamiento/Mojado/Con escarcha/Banquina inutilizable/Resbaladizo
Tiempo		Bueno/Granizo/Llovizna/Lluvia/Nieve/Nublado
Luminosidad	Condiciones atmosféricas y luminosidad	Amanecer/Día/Atardecer/Noche
Luz artificial		Sí/No
Vientos fuertes		Sí/No
Niebla		Sí/No
Policía de Tránsito	Prioridades	Sí/No
Semáforo		Sí/No
Señal Pare		Sí/No

Señal ceda el paso		Sí/No	
Marca vial		Sí/No	
Senda Peatonal		Sí/No	
Rotonda		Sí/No	
Jerarquía de vía		Sí/No	
Otra señal		Sí/No	
Ninguna		Sí/No	
Alcohol o Droga	Causas aparentes	Sí/No	
Cansancio o Sueño		Sí/No	
Deficiencia en la vía		Sí/No	
Distracción		Sí/No	
Enfermedad		Sí/No	
Exceso de velocidad		Sí/No	
Evitó embestir un animal		Sí/No	
Evitó embestir otro vehículo		Sí/No	
Evitó embestir peatón		Sí/No	
Mal estado del vehículo		Sí/No	
Maniobra riesgosa		Sí/No	
Meteorología adversa		Sí/No	
Otros		Sí/No	
Restricción		Restricciones, Señalización y Semaforización	Accidente previo/Desvío provisorio/Obra en construcción/Reparación vía/Sin restricción/Otro
Señalización			Completa/Incompleta/Regular/Nula
Semaforización	Funcionaba/Intermitente/No funcionaba/Sin semáforo/Otro		
Siniestro	Tipo de siniestro	Atropello peatón/Caída sin colisión/Colisión frontal/Colisión frontolateral/Colisión lateral/Colisión por alcance/Otro/Vuelco en calzada	
Sexo	Participantes	Femenino/Masculino	
Edad		Edad (número entero)	
Movilidad		Peatón/Bicicleta/Motocicleta/Automovil/Camioneta/Transporte de carga/Transporte de pasajeros/Maquinaria/Tracción a sangre/Cuatriciclo/Vehículo desconocido (fuga)	
Ubicación		Conductor/Asiento delantero/Asiento trasero	
Uso del cinturón de seguridad		Sí/No/Nc	
Uso de Casco		Sí/No/Nc	
Hospitalizado		Sí/No	
Licencia conductor		Sí/No	
VTV	Licencias y seguros	Sí/No	
Seguro		Sí/No	
Notas	Notas	Información adicional (texto)	

EL procedimiento de la toma de datos en los formularios diseñados corresponde a personal de la Dirección de Contralor Municipal (Agentes de Tránsito), Bomberos Voluntarios y Hospital Municipal (Agentes de Salud). Una vez que se releva un siniestro vial, los datos del mismo se introducen en un sistema de registro que permite geolocalizar el hecho mediante un Sistema de Información Geográfica que brinda la latitud y longitud en base a los datos relevados y/o ajuste manual del operador (Figura 1).

La información obtenida por los profesionales permite clasificar la gravedad de las lesiones en Leves, Graves, Gravísimo y Fallecido, con un horizonte de 30 días de ocurrido el siniestro.

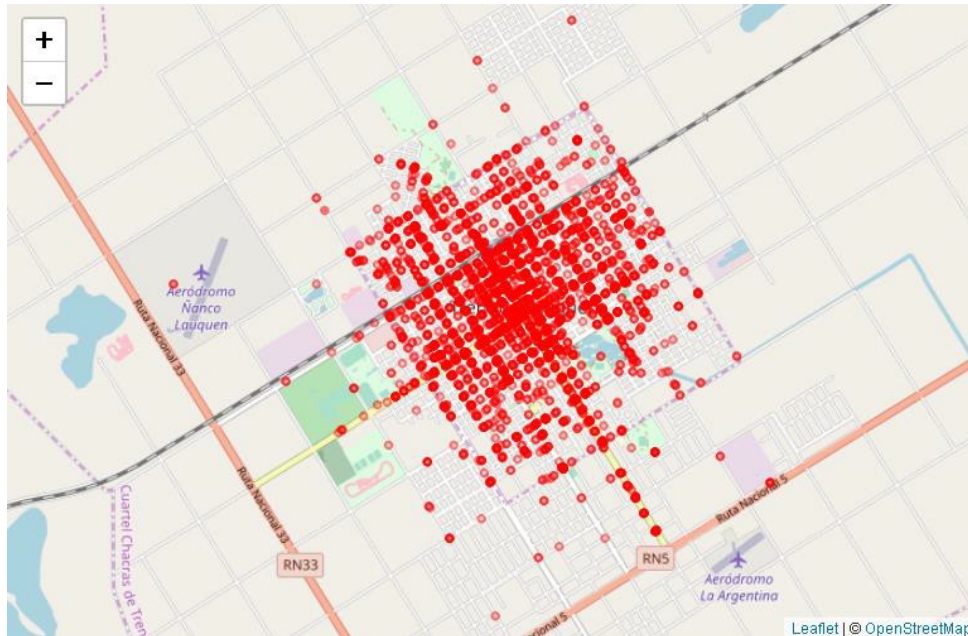


Figura 1 Geolocalización de los datos correspondientes a siniestros viales relevados.

Los operativos de la Dirección de Contralor dirigidos a acciones preventivas de seguridad vial se definen en base a tres criterios:

- ¿A quiénes va dirigido?: Tipo de movilidad, perfil de los individuos,
- ¿Dónde se realizan? Calle(s), Intersección, Cuadrícula o latitud y longitud,
- ¿Cuándo se realizan? Día de la semana y rango horario.

El algoritmo a desarrollar, como objetivo del presente trabajo, debe ser capaz de brindar información sobre uno o dos de esos criterios cuando se definan los restantes por parte del Director de tránsito. Para su implementación se adopta la metodología SEMMA [4] radicada en las siguientes etapas (Figura 2):

- Muestreo (Sample)
- Exploración (Explore)
- Selección e ingeniería de variables (Modify)
- Modelización (Model)
- Evaluación de los modelos (Assess)

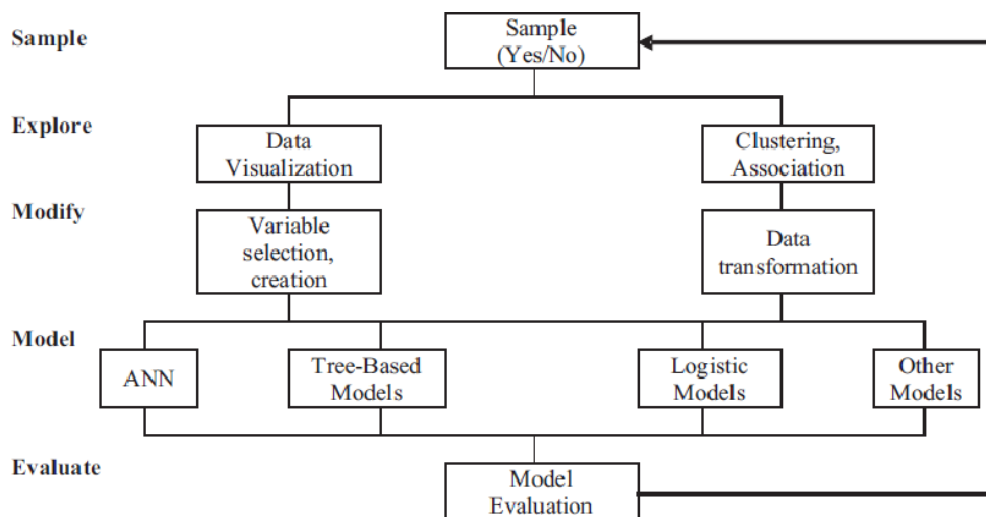


Figura 2 Etapas de la metodología SEMMA. [5]

Muestreo: La obtención de los datos se basará en el registro sistematizado de los siniestros viales y sus consecuencias en los participantes que ocurren en el área urbana de la ciudad. Esta restricción se debe a que la planificación de los operativos de seguridad vial se llevan a cabo en el tejido urbano; principalmente porque los agentes no tienen jurisdicción sobre las rutas nacionales que circundan

la ciudad y en áreas rurales y suburbanas la frecuencia vehicular es muy baja para realizar este tipo de operativos.

Se considerará como participante válido todo aquel individuo en el que se haya registrado al menos el sexo o la edad y el tipo de movilidad utilizada.

Los datos se ordenarán en una tabla tabulada por participante (Figura 3), por lo que la información relativa a las variables del siniestro aparecerá duplicada en cada uno de los individuos que participó en el mismo.

		Variables				
		idFormularioTránsito	idFormularioSalud	Fecha	...	Seguro
Observaciones (participantes)		00001	00001			
		00001	00002			
		00002	-			

Figura 3 Esquema de tabla tabulada.

Exploración: Se realizará un análisis exploratorio visual de las variables, tendiente a detectar datos ausentes, distribuciones y posibles relaciones entre las variables. Para ello se utilizará el software libre R [6], el cual proporciona una amplia variedad de técnicas estadísticas (modelos lineales y no lineales, pruebas estadísticas clásicas, análisis de series temporales, clasificación, clustering, etc.) y técnicas gráficas (Figura 4).



Figura 4 Clustering de datos georeferenciados.

Modificación: En esta etapa se seleccionarán las variables predictoras disponibles para aplicar en los modelos implementados, en base a las variables de salida que pueden fijar los agentes de tránsito. También se realizan imputaciones de valores ausentes e ingeniería de variables que permitan capturar mayor información sobre el evento (Figura 5 a 9) y por lo tanto aumentar el rendimiento de los modelos de la siguiente etapa.

Para llevar a cabo la imputación de valores ausentes en primer lugar se observará el patrón de datos faltantes y se verificará visualmente si se corresponde con un patrón MAR (missing at random) o MCAR (missing completely at random). En caso de ser así se llevará a cabo una imputación múltiple de los mismos.

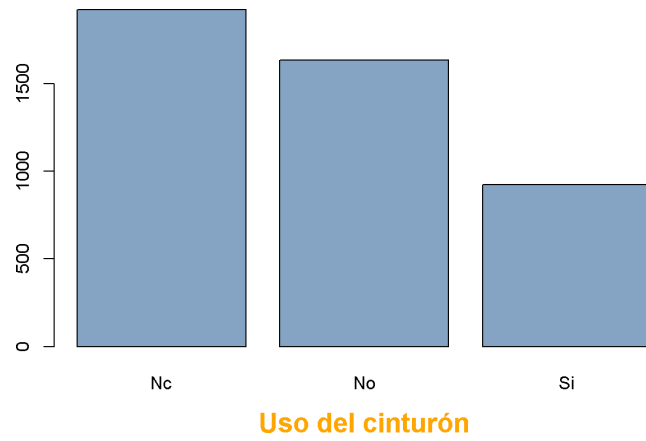


Figura 5 Observaciones sobre uso del cinturón.

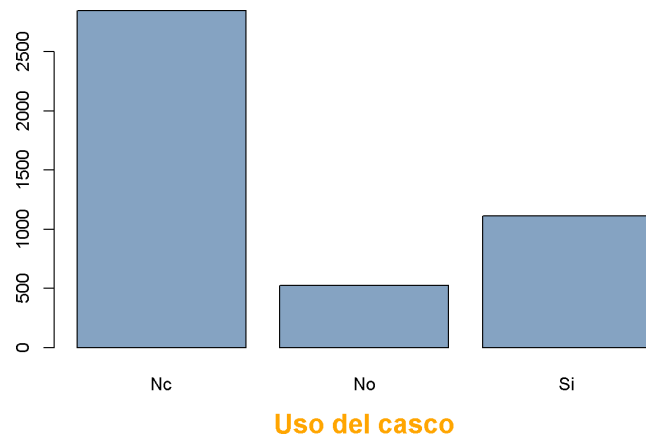


Figura 6 Observaciones sobre uso del cinturón.

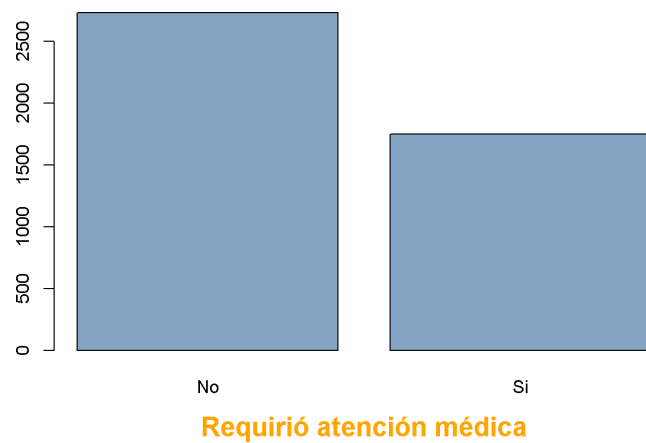


Figura 7 Observaciones sobre requerimiento de atención médica.

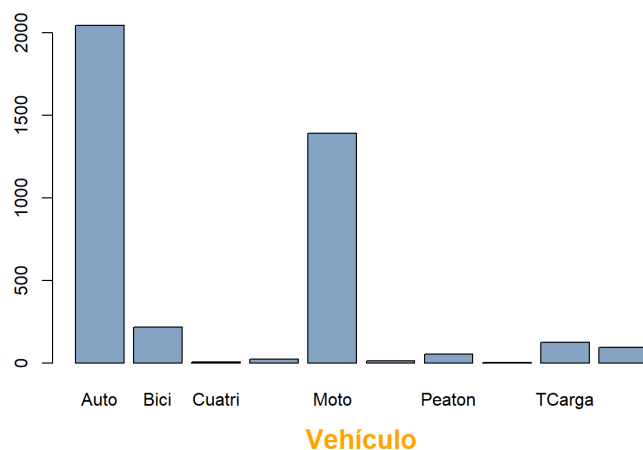


Figura 8 Observaciones sobre tipo de vehículo involucrado.

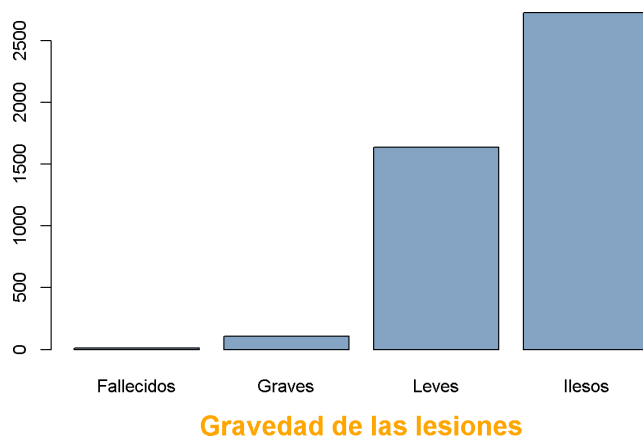


Figura 9 Observaciones sobre tipo el tipo de gravedad de las lesiones.

Se identificarán las variables con varianza nula o cercana a cero ya que estas últimas también podrían transformarse en variables de varianza nula al realizar submuestras de validación cruzada/bootsrap o al separar los datos de entrenamiento y testeo. El criterio para identificar estas variables se basará en dos métricas; la primera es la proporción del valor más frecuente de la variable respecto del segundo valor más frecuente, y la segunda es la proporción de valores únicos respecto del total de las observaciones. Cuando la primera obtenga un valor muy alto y la segunda un valor muy bajo, entonces, la variable deberá analizarse cuidadosamente para ver si se la mantiene o no dentro de los predictores que dispondrán los modelos. Una posible solución para evitar eliminar la variable es su reingeniería discretizando los valores y unificando aquellos con menor frecuencia en un mismo grupo.

Las variables cualitativas o categóricas que no posean una escala u orden se binarizarán mediante la técnica de one-hot-encoding [7] ya que ninguna de las variables posee alta cardinalidad por lo tanto no se generará un excesivo número de variables ficticias. Las variables cualitativas ordinales serán tratadas como factores ordenados.

A continuación se buscará el nivel de correlación entre los potenciales predictores y dependencias lineales, ya sea en las variables originales como las generadas por binarización con el objetivo de eliminarlas y obtener un mejor rendimiento de los modelos.

Por último se centrarán y escalarán las variables para evitar que los modelos que son sensibles a valores demasiados altos o bajos afecten en desempeño del modelo.

La etapa de modificación se utiliza en forma iterativa con las siguientes, ya que la necesidad de ingeniería de variables muchas veces es detectada en las etapas de generación y evaluación del modelo.

Modelización: Se particionarán los datos en dos grupos, uno de entrenamiento y otro de testeo. Los datos de entrenamiento se utilizarán para ajustar los modelos, utilizando técnicas de validación cruzada siempre que sea posible.

Se procederá a generar los modelos en función de las variables asociadas a los tres criterios utilizados por los agentes de tránsito. Dado que son más de una las posibles variables que responden a esos criterios se

conservarán los modelos con mejor performance.

A su vez se evaluarán los modelos en función de las variables predictoras significativas, ya que las mismas deberían incluir los dos criterios restantes. Por ejemplo: si el operativo está dirigido a un determinado tipo de movilidad entonces el lugar y/o el momento para realizarlo debe tener como variable predictora el tipo de vehículo.

Entre los modelos a utilizar se encuentran tanto modelos de clasificación como de regresión en función del tipo de variable de salida, aunque la mayoría de las variables de salida serán de tipo cualitativa (cuadrícula, tipo de movilidad, sexo, grupo etario (un agente de tránsito no puede estimar exactamente la edad, sino una franja etaria), día de la semana, etc.). Esto se realizará a través de algoritmos y técnicas de:

- Regresión Logística
- Análisis discriminante lineal
- K-nn (K-vecinos cercanos)
- Modelos aditivos generalizados
- Árboles de decisión simple
- Random Forests
- Máquinas de soporte vectorial
- Regresión lineal múltiple

Evaluación: Para los modelos, se utilizarán como herramienta de evaluación, las curvas ROC (Receiver Operating Characteristic), la curva de ganancia acumulada y métricas obtenidas a partir de la matriz de confusión como exactitud, sensibilidad y precisión.

El reporte que se generará una vez que el agente de tránsito parametrize las variables de interés de uno de los criterios, incluirá los valores de las variables correspondientes a los otros dos criterios, como así también el intervalo de confianza de la predicción o la probabilidad de pertenecer a la clase de la variable de salida.

La efectividad de la herramienta se verificará mediante su utilización por los agentes de tránsito a lo largo de 3 (tres) meses y evaluando los índices de siniestralidad urbana respecto de los períodos donde los operativos se realizaban en forma arbitraria o al azar.

3. CONCLUSIONES

El presente trabajo tiene como propósito compartir la metodología que se llevará a cabo en el proyecto de investigación que tiene como objetivo generar algoritmos que apoyen la gestión de los recursos para reducir la siniestralidad vial en municipios que posean entre 30.000 y 60.000 habitantes.

Mediante la aplicación de técnicas de aprendizaje supervisado se generarán modelos predictivos de manera que sobre una base científica y donde los criterios de asignación sean conocidos y compartidos por la gestión Municipal, se pueda utilizar la información de siniestralidad vial para ajustar proactivamente el despliegue de los agentes de tránsito en el territorio.

Adicionalmente se podrán conocer cuáles de las variables relevadas en la ciudad poseen valor predictivo para la adecuación de los operativos de seguridad vial.

Otro atractivo que tiene este tipo de aplicación de técnicas de aprendizaje supervisado es la utilización concreta de los datos relevados por personal municipal y bomberos voluntarios que le dan un sentido a continuar un registro que hasta el momento solo existía con finalidad descriptiva pero que no se intentaba utilizar en la propuesta de soluciones al problema relevado.

El proyecto es de interés para aquellas líneas de investigación o disciplinas que abarquen: Accidentología Vial; Gestión de Recursos Humanos; Estadística Aplicada; Modelización descriptiva de siniestros viales; Buenas prácticas en gestión Municipal.

Se transferirá al Municipio una herramienta metodológica de asignación de los recursos de la Dirección de Tránsito en función de la siniestralidad vial.

4. REFERENCIAS

[1] Organización Mundial de la Salud (2010). Sistemas de datos. *Manual de seguridad vial para*

decisores y profesionales.

- [2] Secretaría de Seguridad Interior. Ministerio de Seguridad. (2016). “Estadísticas Criminales en la República Argentina – Año 2015”. Argentina.
- [3] Naciones Unidas (2010). *Plan Mundial para el Decenio de Acción para la Seguridad Vial 2011–2020*.
- [4] Institute, S. A. S., Sas, U., & Software, E. M. (1998). Data Mining and the Case for Sampling - A SAS Institute Best Practices Paper Solving Business Problems Using SAS® Enterprise Miner™ Software.
- [5] Bulkley, J.; Gayle, S.; Hicks, B.; Stephens, R. (1999). *Adding the where to the who. In 24th SUGI - SAS Users Group International conference*. Paper 173, 3p. Miami, USA.
- [6] R Core Team (2013). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <http://www.R-project.org/>.
- [7] Garavaglia, Susan & Sharma, Asha. (2019). *A Smart Guide To Dummy Variables: Four Applications And A Macro*.