



UNIVERSIDAD TÉCNICA DEL NORTE

INSTITUTO DE POSTGRADO

MAESTRÍA EN TELECOMUNICACIONES

Detección de patrones de tiempo y espacio en el Sistema Integrado SIS ECU 911 Ibarra utilizando técnicas de minería de datos para la toma de decisiones informadas y reducción de costos técnicos-económicos.

Trabajo de Investigación previo a la obtención del Título de Magíster en
TELECOMUNICACIONES

AUTOR: César Gustavo Chacón Encalada

TUTOR: PhD. Iván García Santillán

IBARRA - ECUADOR

2021

APROBACIÓN DEL TUTOR

En calidad de Director del Trabajo de Investigación con el tema “DETECCIÓN DE PATRONES DE TIEMPO Y ESPACIO EN EL SISTEMA INTEGRADO SIS ECU 911 IBARRA UTILIZANDO TÉCNICAS DE MINERÍA DE DATOS PARA LA TOMA DE DECISIONES INFORMADAS Y REDUCCIÓN DE COSTOS TÉCNICOS-ECONÓMICOS.”, de autoría de César Gustavo Chacón Encalada, para obtener el Título de Magíster en Telecomunicaciones, doy fe que dicho trabajo reúne los requisitos y méritos suficientes para ser sometidos a presentación y evaluación por parte del jurado examinador que se designe

En la ciudad de Ibarra, a los 5 días del mes de febrero de 2021

Lo certifico:

PhD. Iván García Santillán

CI: 1002292603

TUTOR DE TESIS



UNIVERSIDAD TÉCNICA DEL NORTE

BIBLIOTECA UNIVERSITARIA

AUTORIZACIÓN DE USO Y PUBLICACIÓN A FAVOR DE LA

UNIVERSIDAD TÉCNICA DEL NORTE

1. IDENTIFICACIÓN DE LA OBRA

En cumplimiento al Art. 144 de la Ley de Educación Superior, hago entrega del presente trabajo a la Universidad Técnica del Norte para que sea publicado en el Repositorio Digital Institucional, para lo cual pongo a disposición la siguiente información:

Datos de contacto	
Cédula de identidad:	1002698445
Apellidos y nombres:	Chacón Encalada César Gustavo
Dirección:	Ibarra, Río Chota 4-54 y Río Cenepa
Email:	cgchacone@utn.edu.ec
Teléfono fijo:	
Teléfono celular:	0995867371

Datos de la Obra	
Título:	Detección de Patrones de tiempo y espacio en el Sistema Integrado SIS ECU 911 Ibarra utilizando técnicas de minería de datos para la toma de decisiones informadas y reducción de costos técnicos-económicos
Autor:	César Gustavo Chacón Encalada
Fecha:	05 de febrero de 2021
SOLO PARA TRABAJOS DE GRADO	
PROGRAMA	PREGRADO () POSTGRADO (X)

TÍTULO POR EL QUE OPTA:	Magíster en Telecomunicaciones
TUTOR:	PhD. Iván García Santillán

2. CONSTANCIA

El autor manifiesta que la obra objeto de la presente autorización es original y se la desarrolló, sin violar derechos de autor de terceros, por lo tanto, la obra es original y que es titular de los derechos patrimoniales, por lo que asumen la responsabilidad sobre el contenido de esta y saldrán en defensa de la Universidad en caso de reclamación por parte de terceros.

Ibarra, a los 20 días del mes de julio de 2021

EL AUTOR

Nombre: César Gustavo Chacón Encalada

C.I.: 1002698445

DEDICATORIA

La dedicación, el esfuerzo lo pongo YO, la inspiración es mi familia

Gustavo Chacon 2021

Este trabajo lo dedico primero a Dios que con su bendición todo es posible, a la Virgen María por ser mi abogada e interceder a mis súplicas.

A mi familia con todo el cariño y amor que mi cuerpo pueda dar, en especial a mi amada esposa, mi amor: Lorena, que por su ayuda y empuje he logrado completar muchos sueños, a mis Hijos Alexis Agustín y Gustavito Rafael para los cuales he querido ser un ejemplo de superación y dedicación; y, a mis Hijas: Sofía Elizabeth y María Susana que con sus sonrisas iluminan mis días y llenan de alegría mi corazón.

Gustavo Ch.

AGRADECIMIENTO

Un sincero reconocimiento a la Universidad Técnica del Norte y al Instituto de Postgrado, ya que me han brindado la oportunidad de crecer profesionalmente, con una formación integral con valores y principios éticos y humanistas.

A la Dirección Zonal de Tecnología del SIS ECU 911 Ibarra quienes con su apoyo y entusiasmo se logró superar muchas barreras y lograr este objetivo planteado hace años atrás.

Un especial reconocimiento al PhD Iván García Santillán, Tutor de esta investigación, por su apoyo profesional y humano en el desarrollo de este proyecto.

A mi madre que con sus consejos y ayuda me han permitido estar donde estoy y ser mi ejemplo de lucha para lograr muchos objetivos y sueños, gracias por estar siempre a mi lado.

Finalmente, a toda mi familia que de una u otra manera han compartido conmigo palabras de aliento y motivación para llegar a cumplir cada una de las metas que me he propuesto, espero ser ejemplo para mis primos y sobrinos. Va por usted papá Gonzalo.

Gustavo Ch.

ÍNDICE DE CONTENIDOS

ÍNDICE DE CONTENIDOS	7
ÍNDICE DE FIGURAS	8
ÍNDICE DE GRÁFICOS	9
ÍNDICE DE TABLAS	10
RESUMEN	11
ABSTRACT	13
CAPÍTULO I: INTRODUCCIÓN	15
1.1. Problema de investigación	17
1.2. Objetivos de la investigación	17
1.2.1. Objetivo general	17
1.2.2. Objetivos específicos	18
1.3. Justificación	18
CAPÍTULO II: MARCO REFERENCIAL	22
2.1. Antecedentes	22
2.2. Referentes teóricos	26
2.2.1. Telecomunicaciones	26
2.2.2. Minería de Datos	30
2.2.3. Big Data	36
2.2.4. Base de Datos Espacial (Geodatabase)	43
2.2.5. Modelo de Predicción “Regresión Geográficamente Ponderada (GWR)	46
2.2.6. Mapas de Calor	48
2.2.7. Reloj de Datos	49

2.3.	Servicio Integrado de Seguridad (SIS) ECU 911	50
2.4.	Marco legal	52
CAPÍTULO III: MARCO METODOLÓGICO		55
3.1.	Descripción del área de estudio	55
3.2.	Enfoque y tipo de investigación.....	58
3.3.	Procedimiento de investigación	59
3.3.1.	Primera Fase:	59
3.3.2.	Segunda Fase	64
3.3.3.	Tercera fase.....	66
3.3.4.	Cuarta fase	69
CAPÍTULO IV: RESULTADOS Y DISCUSIÓN		71
CAPÍTULO V: CONCLUSIONES Y RECOMENDACIONES		88
CONCLUSIONES		88
RECOMENDACIONES.....		89
BIBLIOGRAFÍA		91

ÍNDICE DE FIGURAS

Figura 1.	Proceso de Minería de Datos	19
Figura 2.	Modelo de Gestión SIS ECU 911	20
Figura 3.	Etapas de un sistema de telecomunicación	28
Figura 4.	Mapa de calor.....	48
Figura 5.	Reloj de Datos.....	50
Figura 6.	Mapa de la provincia de Imbabura.....	56

Figura 7. Niveles administrativos de planificación	57
Figura 8. Fases del Procedimiento de investigación	59
Figura 9. Triangulación	60
Figura 10. Proceso de geolocalización de llamadas	63
Figura 11. Sistema de Información Geográfica y sus partes	65
Figura 12. Fórmula que representa la estadística G_i^*	66
Figura 13. Esquema metodológico del Software Crime Analyst	67
Figura 14. Tasa de llamadas mal intencionadas por parroquia 2018-2019	74
Figura 15. Llamadas mal intencionadas por cantón 2018-2019.....	76
Figura 16. Hot Spot de llamadas mal intencionadas por cantones de Imbabura.....	77
Figura 17. Mapa Caliente (hot spot) de llamadas mal intencionadas divididas por cantones año 2018	78
Figura 18. Hot Spot llamadas mal intencionadas divididas por cantones año 2019	79
Figura 19. Llamadas mal intencionadas 2018	80
Figura 20. Reloj rectangular de llamadas mal intencionadas 2018	81
Figura 21. Reloj de Datos de llamadas mal intencionadas 2019.....	82
Figura 22. Reloj de datos rectangular de llamadas mal intencionadas 2019.....	82
Figura 23. Comparativa de llamadas mal intencionadas 2018-2019.....	84
Figura 24. Comparativa de llamadas mal intencionadas en reloj de datos años 2018-2019.....	85
Figura 25. GWR de llamadas mal intencionadas en Imbabura por Cantones al 2030	86
Figura 26. GWR de llamadas mal intencionadas en Imbabura por Cantones al 2030	86

ÍNDICE DE GRÁFICOS

Gráfico 1. Llamadas ingresadas al SIS ECU 911 Ibarra - Año 2017	23
Gráfico 2. Llamadas ingresadas al SIS ECU 911 Ibarra - Año 2018	23
Gráfico 3. Llamadas ingresadas al SIS ECU 911 Ibarra - Año 2019	24
Gráfico 4. Comparativa de Llamadas ingresadas 2017-2019.....	25

ÍNDICE DE TABLAS

Tabla 1. Llamadas ingresadas al SIS ECU 911 Ibarra 2017-2019	24
Tabla 2. Descripción de distritos y circuitos de Imbabura	55
Tabla 3. Descripción del proceso de ingreso de llamadas	62
Tabla 4. Información para medir grado de confianza de los resultados	66
Tabla 5. Criterios de aceptabilidad del coeficiente R2 para el análisis de regresión.....	69
Tabla 6. Resumen de llamadas mal intencionadas 2018-2019	71
Tabla 7. Porcentaje de llamadas por cantón	72
Tabla 8. Tasa de llamada por cantón 2018-2019.....	72
Tabla 9. Tasa de llamadas mal intencionadas por parroquia	73
Tabla 10. Porcentaje de llamadas por cantón	75



UNIVERSIDAD TÉCNICA DEL NORTE

INSTITUTO DE POSTGRADO



MAESTRÍA EN TELECOMUNICACIONES

Detección de patrones de tiempo y espacio en el Sistema Integrado SIS ECU 911 Ibarra utilizando técnicas de minería de datos para la toma de decisiones informadas y reducción de costos técnicos-económicos.

Autor: César Gustavo Chacón Encalada

Tutor: Phd. Iván García Santillán

Año: 2021

RESUMEN

La presente investigación aborda la detección de patrones de tiempo y espacio en el Sistema Integrado SIS ECU 911 Ibarra utilizando técnicas de minería de datos para la toma de decisiones informadas y reducción de costos técnicos-económicos, este estudio parte de la información en la cual indican que el Servicio Integrado de Seguridad ECU 911 durante el año 2018 recibió del total de llamadas entrantes, aproximadamente, el 47.4% de llamadas catalogadas como mal intencionadas, esto generó costos innecesarios. El enfoque de la investigación es cuantitativo ya que se partió de los datos generados en el SIS ECU 911 Ibarra de llamadas mal intencionadas en el período 2018 - 2019. Esta investigación encontró patrones espacio temporales que ayudaron a entender la problemática. La investigación fue realizada en fases: la primera fue la obtención de los datos (ingreso de llamadas y geolocalización), la segunda fase generó la Geodatabase y los mapas calientes o hotspots, mientras que la tercera fase se encargó de generar los relojes de datos y finalmente en la

cuarta fase se usó el modelo de predicción GWR aplicando la regresión ponderada geográficamente. En conclusión, revisando el modelo GWR se encontró una varianza baja del 43% entre las llamadas mal intencionadas y la variable relacionado a la población en parroquias como Pimampiro, Ambuquí, Ibarra, Salinas, Pablo Arenas y Lita, estas parroquias se ajustaron a este modelo. En el resto de las parroquias, a pesar de tener una tasa de llamadas más altas de la provincia como González Suárez, Pablo Arenas, y Selva Alegre todas rurales el modelo no ha permitido focalizar geográficamente una dependencia espacial que permita determinar un patrón predictivo significativo. Se confirmó parcialmente que las llamadas mal intencionadas responden a un patrón espaciotemporal.

Palabras clave: minería de datos, patrones de tiempo, hotspots, relojes de datos, modelo predictivo.



UNIVERSIDAD TÉCNICA DEL NORTE

INSTITUTO DE POSTGRADO



MAESTRÍA EN TELECOMUNICACIONES

Detección de patrones de tiempo y espacio en el Sistema Integrado SIS ECU 911 Ibarra utilizando técnicas de minería de datos para la toma de decisiones informadas y reducción de costos técnicos-económicos.

Autor: César Gustavo Chacón Encalada

Tutor: Phd. Iván García Santillán

ABSTRACT

This research addresses the detection of time and space patterns in the SIS ECU 911 Ibarra Integrated System using data mining techniques for making informed decisions and reducing technical-economic costs, this study starts from the information in which they indicate The ECU 911 Integrated Security Service during 2018 received approximately 47.4% of the total incoming calls of calls classified as malicious, this generated unnecessary costs. The focus of the research is quantitative since it was based on the data generated in the SIS ECU 911 Ibarra of malicious calls in the period 2018 - 2019. This research found spatial-temporal patterns that helped to understand the problem. The research was carried out in phases: the first was to obtain the data (incoming calls and geolocation), the second phase generated the Geodatabase and the hot maps or hotspots, while the third phase was in charge of generating the data clocks and finally, in the fourth phase, the GWR prediction model was used applying the geographically weighted regression. In conclusion, reviewing the GWR model, a low variance of 43% was found between malicious calls and the variable related to the population in parishes such as Pimampiro, Ambuquí, Ibarra, Salinas, Pablo Arenas and Lita, these parishes were adjusted to this model . In the rest of the parishes, despite having a higher rate

of calls in the province such as González Suárez, Pablo Arenas, and Selva Alegre, all rural, the model has not made it possible to geographically focus a spatial dependence that allows determining a significant predictive pattern. It was partially confirmed that malicious calls respond to a spatio-temporal pattern.

Keywords: data mining, time patterns, hotspots, data clocks, predictive model.

CAPÍTULO I: INTRODUCCIÓN

El Estado ecuatoriano garantizará la seguridad humana a través de políticas y acciones integradas, de tal manera que aseguren la convivencia pacífica de las personas, a fin de promover un ambiente de paz y prevenir las formas de violencia y discriminación, así como el cometimiento de infracciones y delitos. De igual manera se debe garantizar a sus habitantes el derecho a una cultura de paz, a la seguridad integral y a vivir en una sociedad democrática y libre de corrupción tal como indica la Constitución del Ecuador (2008), artículo 3, numeral 8.

El enfoque relativo a la Seguridad Integral se analiza desde dos dimensiones: la seguridad ciudadana y la seguridad del Estado las cuales buscan garantizar y proteger los derechos humanos y las libertades de las personas, la reducción de vulnerabilidades a través de la prevención de riesgos, así como la protección, respuesta y remediación ante emergencias y desastres. Para ello se desarrolló una resolución política con profundas transformaciones en la Policía Nacional, Fuerzas Armadas y en todas aquellas instituciones relacionadas con la seguridad en el ámbito interno y externo de tal manera que se pueda construir un país seguro, incluyente y pacífico. El nuevo enfoque de seguridad que orienta este empeño rompe con los paradigmas reactivos y represivos señalando un nuevo camino enfocado en la atención de los factores estructurales, protectores y de riesgo de la inseguridad.

El presidente de la República emitió el Decreto Ejecutivo No. 988 del 29 de diciembre de 2011, publicado en el Registro Oficial No. 618 del 13 de enero de 2012, en donde se regula la implementación del Servicio Integrado de Seguridad ECU – 911 (SIS ECU 911), como herramienta integradora de los servicios de emergencia que prestan el Cuerpo de Bomberos, las Fuerzas Armadas, la Policía Nacional e instituciones que conforman el Sistema Nacional de Salud.

Mediante Decreto Ejecutivo No. 31 de 24 de junio de 2013, se emite reformas al anterior decreto y se sustituye el artículo 8 por el siguiente texto: “El Servicio Integrado de Seguridad ECU-911 estará dirigido y administrado por el Director General, ...”; así mismo, en la Disposición General del referido Decreto se establece: “Concédase al Servicio Integrado de Seguridad ECU-911 la calidad de “Servicio” en los términos de la letra h) del Artículo 10.1 del Estatuto del Régimen Jurídico y Administrativo de la Función Ejecutiva y, por tanto, personalidad jurídica como organismo público con autonomía administrativa, operativa y financiera, y jurisdicción nacional, con sede principal en la ciudad de Quito, conformado por centros operativos a nivel nacional” (Plan Nacional de Seguridad Integral, 2014).

El SIS ECU 911 se encarga de gestionar, en todo el territorio ecuatoriano, la atención de las situaciones de emergencia de la ciudadanía, reportadas a través del número 911, y las que se generen por video vigilancia y monitoreo de alarmas, mediante el despacho de recursos de respuesta especializados pertenecientes a organismos públicos y privados articulados al sistema, con la finalidad de contribuir, de manera permanente, a la consecución y mantenimiento de la seguridad integral ciudadana.

Según (ECU 911, 2018) en su informe de gestión, ingresaron al SIS ECU 911 12,6 millones de llamadas de emergencias, de las cuales 6 millones que representa el 47,4% fueron alertas falsas. En el primer cuatrimestre del año 2019, el SIS ECU 911 a nivel nacional recibió alrededor de 5'300.000 llamadas de emergencia de las cuales 2'056.000 (40%) representan el mal uso que se realiza a la línea única de emergencia lo que ocasiona un grave problema tanto en lo económico como en el tiempo que dedica el operador de llamadas a responder y recolectar datos que no son verídicos. Se debe tomar en cuenta que al hacer mal uso de la línea única de emergencia se ven inmersos distintos tipos de problemas ya que se pueden despachar innecesariamente recursos como: ambulancias, Policía Nacional, Bomberos, etc.; todo dependerá del contexto que presente la llamada malintencionada, provocando demora

en la atención de emergencias reales. Referente al aspecto económico, en el primer cuatrimestre del 2019, se generó un gasto innecesario de \$616.800 dólares por avisos falsos de emergencia. Estas cifras, por su elevado costo técnico-económico para el país motivan a proponer la presente investigación.

1.1. Problema de investigación

El Servicio Integrado de Seguridad ECU 911, en el año 2018 recibió del total de llamadas entrantes, aproximadamente, el 47.4% de llamadas catalogadas como mal intencionadas, dichas llamadas generaron costos innecesarios al despachar recursos que podrían ser usados para emergencias reales y no generar una lentitud en el servicio, y en el peor de los casos la falta de cumplimiento a una emergencia real que requiere uno o varios recursos de los organismos de socorro, los costos económicos en el 2018 bordean los USD. 12 455 169.

Por lo anteriormente indicado, el SIS ECU 911 a través de la detección de patrones de tiempo y espacio basado en minería de datos establecerá las estrategias de aplicación, que permitan tener con mayor certeza las zonas sensibles y mitigar el riesgo de que se sigan generando llamadas mal intencionadas y por ende pérdidas económicas para el Estado.

1.2. Objetivos de la investigación

1.2.1. Objetivo general

- Detectar patrones de tiempo y espacio en el Sistema Integrado SIS ECU 911 Ibarra utilizando técnicas de minería de datos para la toma de decisiones informadas y reducción de costos técnicos-económicos.

1.2.2. *Objetivos específicos*

- Elaborar un marco teórico que sustente las técnicas de minería de datos y el proceso de descubrimiento de conocimiento en base de datos.
- Obtener una vista minable a partir de los datos del SIS ECU 911 de los últimos 2 años utilizando la herramienta Crime Analyst de ArcGIS.
- Elaborar un Reloj de datos y Mapas de Calor (hotspot) para obtener un patrón de tiempo e identificar los puntos de concentración geográfica de una manera visual.
- Obtener resultados de un análisis aplicando la regresión ponderada geográficamente (GWR) para obtener una predicción a futuro.
- Validar y documentar los resultados de los años 2018 y 2019 mediante métodos estadísticos o métricas cuantitativas.

1.3. Justificación

La Minería de Datos es el descubrimiento eficiente de información valiosa, no-obvia de una gran colección de datos y constituye una herramienta indispensable para conseguir un conocimiento más profundo del sistema y descubrir esas claves ocultas que pueden marcar la diferencia y solucionar problemas o riesgos a los que está expuesta una Institución o empresa.

Según (Álvarez Menéndez, 2008), uno de los rasgos que caracterizan al sector de las telecomunicaciones es la cantidad de información que generan y almacenan las organizaciones o empresas. Esta enorme fuente de datos puede resultar imposible de abordar desde un punto de vista manual, por lo que es necesaria la utilización de algún tipo de análisis automatizado para obtener algún rendimiento. De manera general, el proceso de Minería de Datos se lo puede observar en la Figura 1.

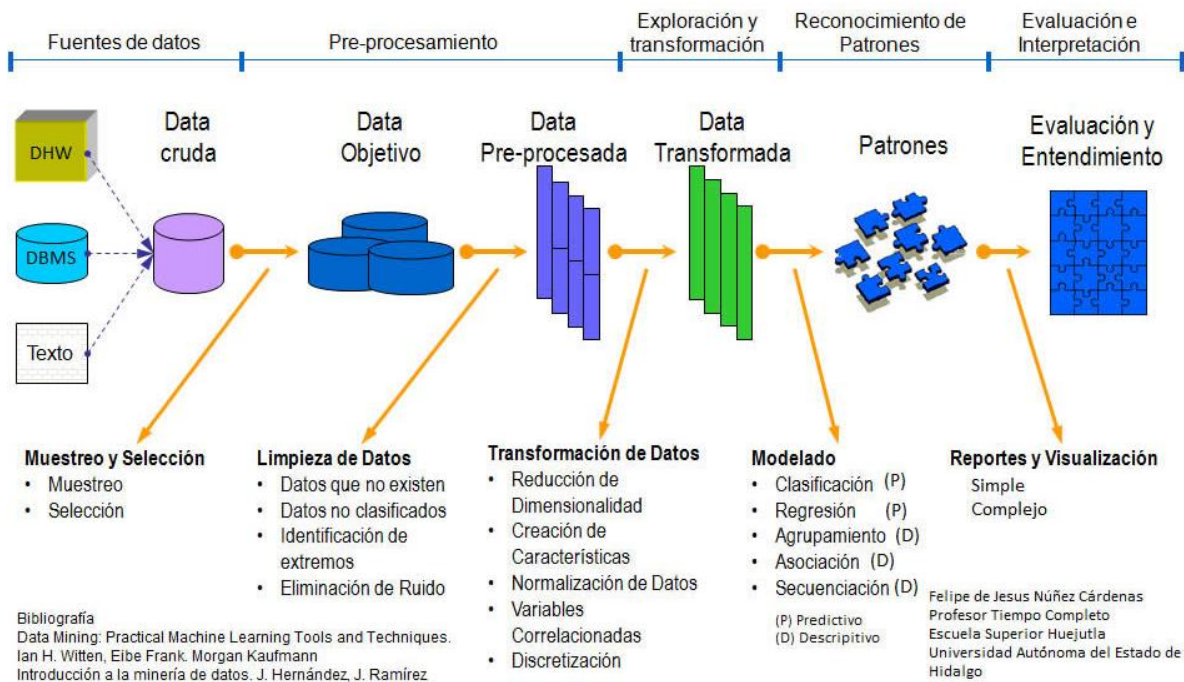


Figura 1. Proceso de Minería de Datos

Fuente: (Núñez Cárdenas, 2018)

Dentro de la clasificación de información almacenada se puede nombrar a: Registro de llamadas, Información de la red y Datos de los clientes; en este caso se hace referencia al Registro de llamadas, que se generan cada vez que se produce una llamada en una red de telecomunicaciones, almacenándose en una base de datos asociada. Cada uno de estos registros debe guardar al menos el origen, destino, fecha, hora y duración de la comunicación.

Según (SENPLADES, 2017), el enfoque relativo a la Seguridad Integral se analiza desde dos dimensiones: la seguridad ciudadana y la seguridad del Estado donde se busca garantizar y proteger los derechos humanos y las libertades de las personas, la reducción de vulnerabilidades a través de la prevención de riesgos, así como la protección, respuesta y remediación ante emergencias y desastres.

El Servicio de Seguridad Integrado ECU 911 es un servicio de respuesta inmediata e integral a una determinada emergencia. Se encarga de coordinar la atención de los organismos de respuesta articulados, para casos de accidentes, desastres y emergencias movilizandolos recursos disponibles para brindar atención rápida a la ciudadanía. Policía Nacional, Fuerzas Armadas, Cuerpo de Bomberos, Comisión Nacional de Tránsito, Ministerio de Salud Pública, Instituto Ecuatoriano de Seguridad Social, Secretaría de Gestión de Riesgos, Cruz Roja Ecuatoriana y otros organismos locales encargados de la atención de emergencias, han unido esfuerzos para brindar la mejor atención a través de un número único: 9-1-1, tal como lo indica (ECU 911, 2017). La Figura 2 muestra el modelo de gestión al cual responde este Servicio.



Figura 2. Modelo de Gestión SIS ECU 911

Fuente: (ECU 911, 2017)

Las llamadas catalogadas como mal intencionadas se registran luego que interactúa el operador de llamadas con el alertante, en ese instante se registra la geolocalización, y los datos necesarios para proveer los servicios sin embargo al ser una llamada falsa se genera

costos innecesarios los cuales se podrían disminuir mediante la aplicación de estrategias que permitirán obtener patrones como resultado de un proceso de minería de datos.

Con el desarrollo de un marco metodológico se debe establecer las estrategias de aplicación, así como los responsables y medios de verificación que permitan tener con certeza las zonas sensibles y posibles patrones para mitigar el riesgo de seguir generando llamadas falsas y por ende pérdidas económicas para el Estado. Además, se busca minimizar el número de sancionados por este problema ya que la normativa legal existente determina que la persona que de manera indebida realice uso del número único de atención de emergencias 911 para dar un aviso falso de emergencia y que implique desplazamiento, movilización o activación innecesaria de recursos de las instituciones de emergencia, será sancionada con pena privativa de libertad de quince a treinta días (COIP, 2008).

El presente trabajo de investigación aporta a los objetivos 8 y 9 del Plan Nacional del Buen Vivir, enmarcados en “Promover la transparencia y la corresponsabilidad para una nueva ética social” y “Garantizar la soberanía y la paz, y posicionar estratégicamente el país en la región y el mundo”.

Finalmente, es necesario identificar los beneficiarios directos que según el problema planteado es directamente el Estado ecuatoriano al no incurrir en gastos innecesarios por llamadas mal intencionadas, y como beneficiarios indirectos están las Instituciones articuladas al SIS ECU 911 Ibarra y ciudadanía en general, porque se mejorará la eficiencia en despacho de recursos y su disponibilidad.

CAPÍTULO II: MARCO REFERENCIAL

2.1. Antecedentes

Una emergencia es cualquier situación que perturba el funcionamiento de la comunidad o sociedad con afectación a las personas, la salud, los bienes o al medio ambiente. Una emergencia puede ser generada por eventos naturales o propios de la actividad humana, ya sea de forma repentina o como resultando de un proceso (ECU 911, 2019).

Mediante Decreto Ejecutivo No. 988, publicado en el Registro Oficial No. 618 de 13 de enero de 2012, se emiten disposiciones para la implementación del Servicio Integrado de Seguridad ECU 911, como herramienta integradora de los servicios de emergencia que prestan los Cuerpos de Bomberos, las Fuerzas Armadas, la Policía Nacional e instituciones que conforman el Sistema Nacional de Salud para la atención de las situaciones de emergencia de la ciudadanía reportadas a través del número 911 (Plan Estratégico 2014-2017 ECU 911, 2014).

La utilización de técnicas de minería de datos, en el momento tecnológico actual, está adquiriendo una gran importancia, debido a que las empresas tanto públicas como privadas están implementando estas técnicas con la finalidad de procesar y obtener la información no visible y convertirla en resultados efectivos para la toma de decisiones (Moreno Muñoz & Lara Torralbo, 2016).

Con la creación del SIS ECU 911, la comunicación que realizan al número 9-1-1 para el servicio de emergencias, tendrá lugar a través de una línea telefónica. En esta entidad se maneja un gran volumen de información, refiriéndose específicamente a la provincia de

Imbabura, en el año 2017 se reciben 439.614 llamadas, en el 2018 se recibieron 434.576 y en el 2019 un total de 402.698 de acuerdo con el detalle que consta en los Gráficos 1, 2 y 3.

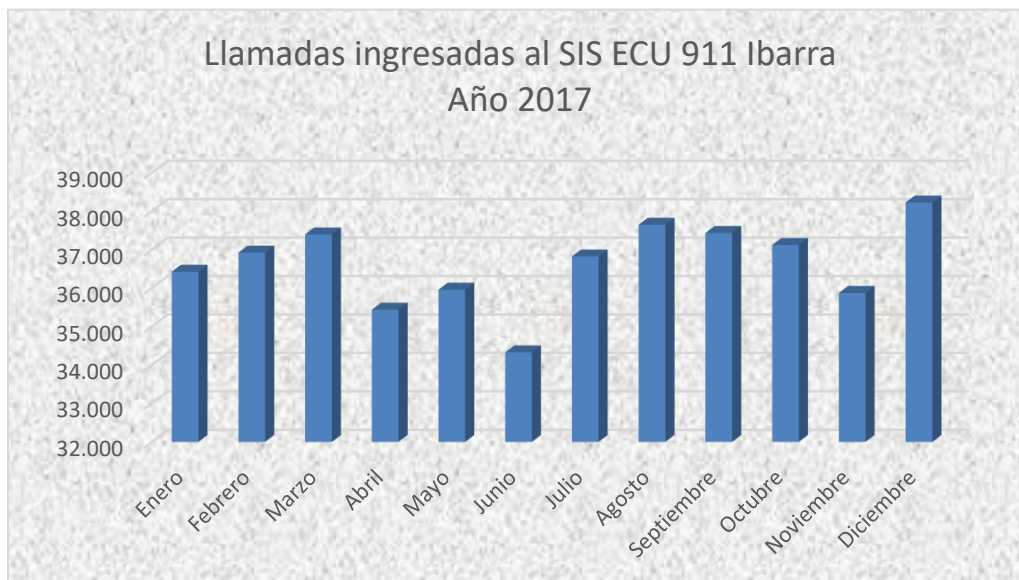


Gráfico 1. Llamadas ingresadas al SIS ECU 911 Ibarra - Año 2017

Fuente: Ecu 911 – Dirección Zonal Estadística

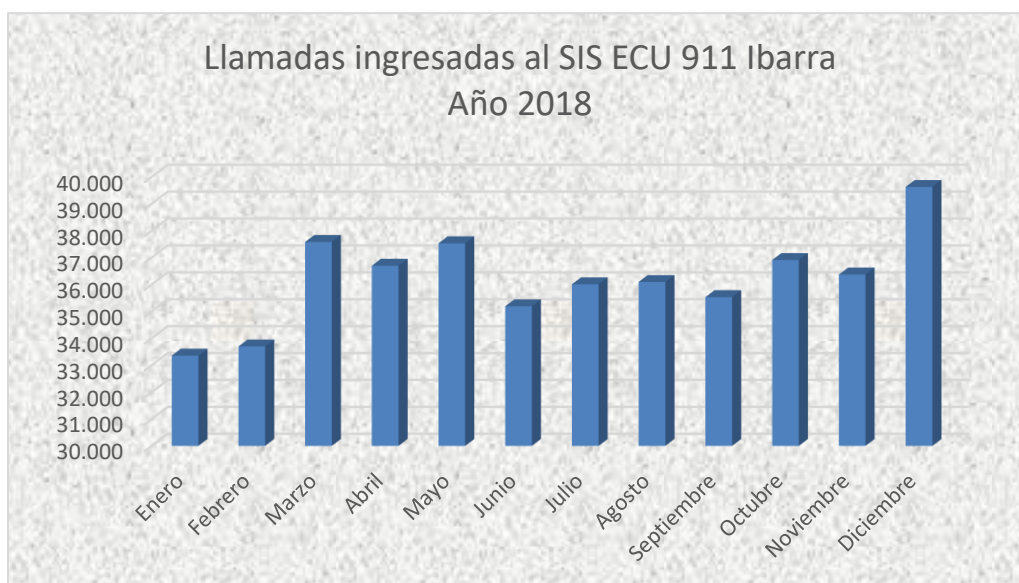


Gráfico 2. Llamadas ingresadas al SIS ECU 911 Ibarra - Año 2018

Fuente: Ecu 911 – Dirección Zonal Estadística

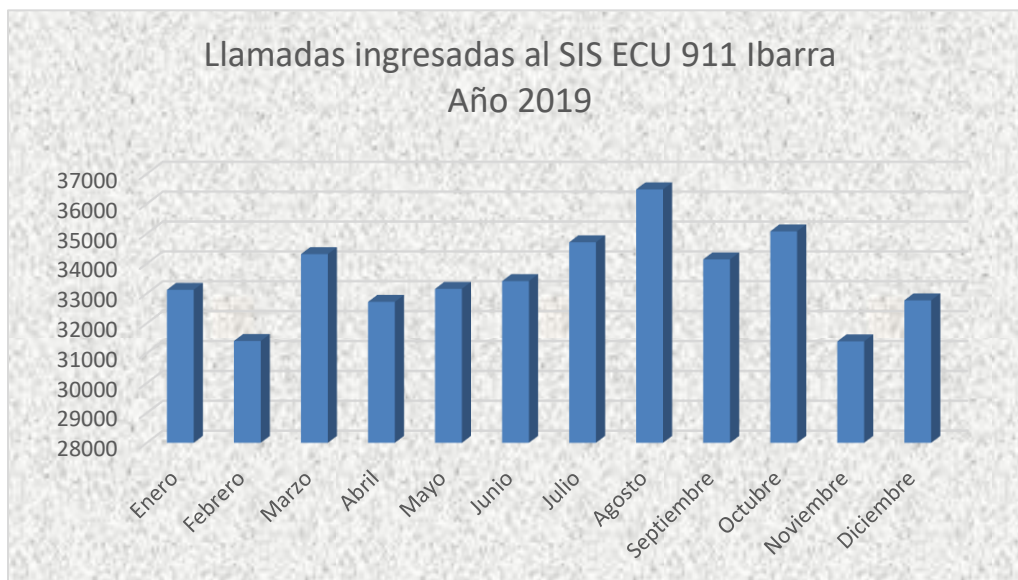


Gráfico 3. Llamadas ingresadas al SIS ECU 911 Ibarra - Año 2019

Fuente: Ecu 911 – Dirección Zonal Estadística

La Tabla 1 y el Gráfico 4, contiene la información resumida con los datos de los 3 años, para tener una referencia del total de llamadas que ingresaron mensualmente al SIS ECU 911 Ibarra, durante estos años cada uno con su particularidad.

Tabla 1. Llamadas ingresadas al SIS ECU 911 Ibarra 2017-2019

Mes	2017	2018	2019
Enero	36.427	33.354	33119
Febrero	36.924	33.696	31409
Marzo	37.399	37.581	34314
Abril	35.441	36.696	32718
Mayo	35.960	37.531	33151

Junio	34.331	35.188	33412
Julio	36.823	36.007	34712
Agosto	37.649	36.091	36482
Septiembre	37.434	35.530	34139
Octubre	37.124	36.908	35078
Noviembre	35.876	36.369	31395
Diciembre	38.226	39.625	32769
Total	439.614	434.576	402698

Fuente: Ecu 911 – Dirección Zonal Estadística

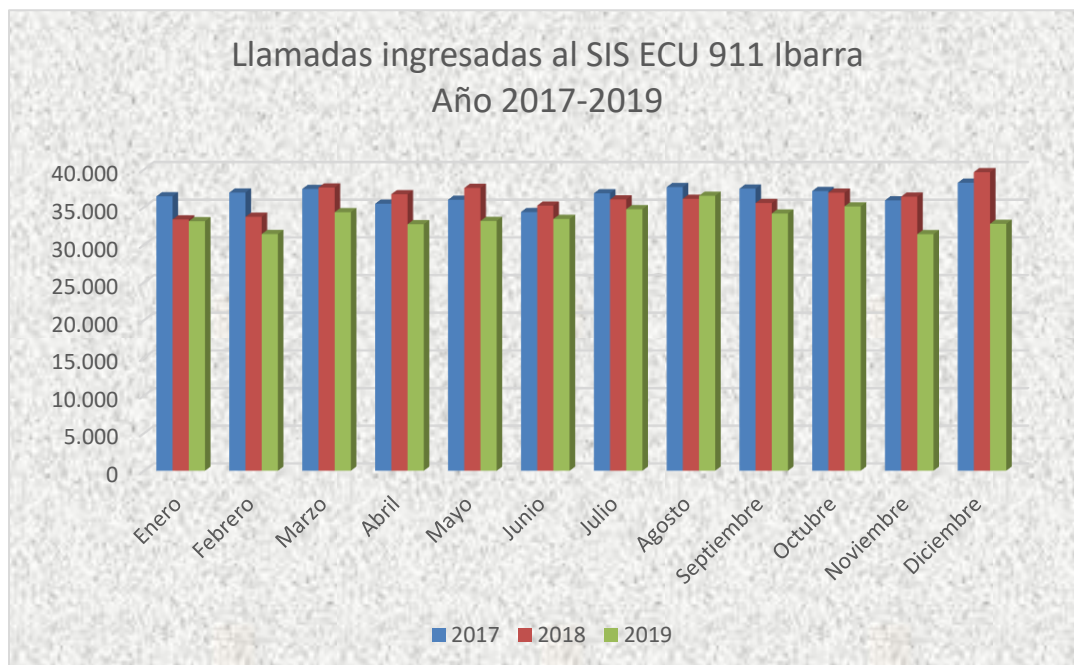


Gráfico 4. Comparativa de Llamadas ingresadas 2017-2019

Fuente: Ecu 911 – Dirección Zonal Estadística

El mal uso de la línea de emergencias 9-1-1 en la provincia de Imbabura, ha ocasionado que durante el año 2019 se registren un total de 196.898 avisos falsos y en el primer semestre del año 2020 un total de 92.207 avisos de este tipo, dando como resultado retrasos en la atención de emergencias reales, por lo que la entidad a nivel nacional diseñó e implementó la campaña denominada ‘Usa Bien el 911’ con la finalidad de concienciar a la población para que haga uso adecuado de la línea de emergencias; además, para que se entienda lo que implica realizar una llamada falsa o mal intencionada y los recursos que se mal usan en esas atenciones.

2.2.Referentes teóricos

2.2.1. Telecomunicaciones

El origen de las telecomunicaciones se remonta a muchos siglos atrás, sin embargo, inicia su desarrollo de forma acelerada el momento que se empiezan a aplicar las tecnologías emergentes a finales del siglo XIX. En los últimos años las telecomunicaciones han experimentado una gran transformación, se ha revolucionado totalmente el sector, sobre todo si se analiza la evolución de la telecomunicación desde su comienzo se encuentra que se han producido grandes avances, precisamente, en periodos de veinticinco años, sí es cierto que la última etapa ha sido excepcional y por tanto ha despuntado de tal manera que está permitiendo el estudio de grandes proyectos e investigaciones usando las nuevas tecnologías (Romeo López & Romero Frías, 2016).

Los primeros indicios se descubrieron en África, donde se hacía una transmisión de información mediante de uso de señales de humo y tambores, luego de muchos años en 1830 se inventaron sistemas eléctricos de telecomunicaciones. Samuel Morse desarrolló en 1837 el primer sistema electrónico de comunicaciones, lo hizo mediante el uso de la inducción electromagnética para transferir información en forma de puntos, rayas y espacios entre un transmisor y un receptor sencillos, usando una línea de transmisión que consistía en un tramo de conductor metálico, se le conoce como telégrafo, así mismo en 1876, Alexander Graham

Bell y Thomas A. Watson fueron los primeros en transferir en forma exitosa la conversación humana a través de un sistema sencillo de comunicaciones con hilo metálico, al que llamaron teléfono (Telectrónica, 2018).

Desde el punto de vista etimológico de acuerdo con el diccionario etimológico en línea, telecomunicaciones proviene del prefijo griego “tele”, que significa distancia, y del latín “communicare”, que significa compartir, su uso moderno proviene principalmente de una adaptación del francés, dado que, fue el ingeniero y novelista francés Édouard Estaunié quien la utilizó por primera vez en sus escritos en el año 1904.

De acuerdo con el diccionario Inglés Oxford, las telecomunicaciones se definen como una comunicación sobre un medio que puede ser un cable, telégrafo, teléfono o radiodifusión. Por otra parte, según la RAE la describe como un sistema de transmisión y recepción a distancia de señales de diversa naturaleza por medios electromagnéticos. Al margen del medio por donde se transmita, entonces, las telecomunicaciones se definen como la comunicación que se realiza entre, por lo menos, dos puntos en lugares distintos.

Actualmente, las telecomunicaciones se han convertido en una necesidad tanto para las personas como para las empresas ya sean estas públicas o privadas, su importancia se nota claramente en los usuarios ya que cuanto más usuario haya conectado a los sistemas de telecomunicaciones mayores son las posibilidades y las necesidades de comunicación.

La telecomunicación se refiere a la comunicación a distancia y es aquella tecnología que se usa para transferir información desde un transmisor hacia un receptor, por medio de un canal de comunicación, esta información es transportada por una señal que puede ser un voltaje proporcional a la amplitud de la voz, como en un teléfono simple, una secuencia de pulsos de luz en una fibra óptica o una onda de radio irradiada por una antena.

En (Galileo Universidad, 2019) se manifiesta que la Unión Internacional de Telecomunicaciones (ITU) define a la telecomunicación como: “Toda emisión, transmisión y recepción de signos, señales, escritos e imágenes, sonidos e informaciones de cualquier naturaleza, por hilo, radioelectricidad, medios ópticos u otros sistemas electromagnéticos”, concluyendo así que las telecomunicaciones son efectivas cuando existe un emisor, un receptor y un canal de comunicación.

La figura muestra las etapas de las telecomunicaciones en donde el transmisor y el receptor en cualquier estación se pueden combinar en un solo dispositivo llamado transceptor. El medio de transmisión de la señal puede ser a través de un cable o cable eléctrico (también conocido como «cobre»), fibra óptica, campos electromagnéticos o luz. La transmisión de espacio libre y la recepción de datos por medio de campos electromagnéticos se llama comunicaciones inalámbricas, así lo manifiesta (Telectrónica, 2018).

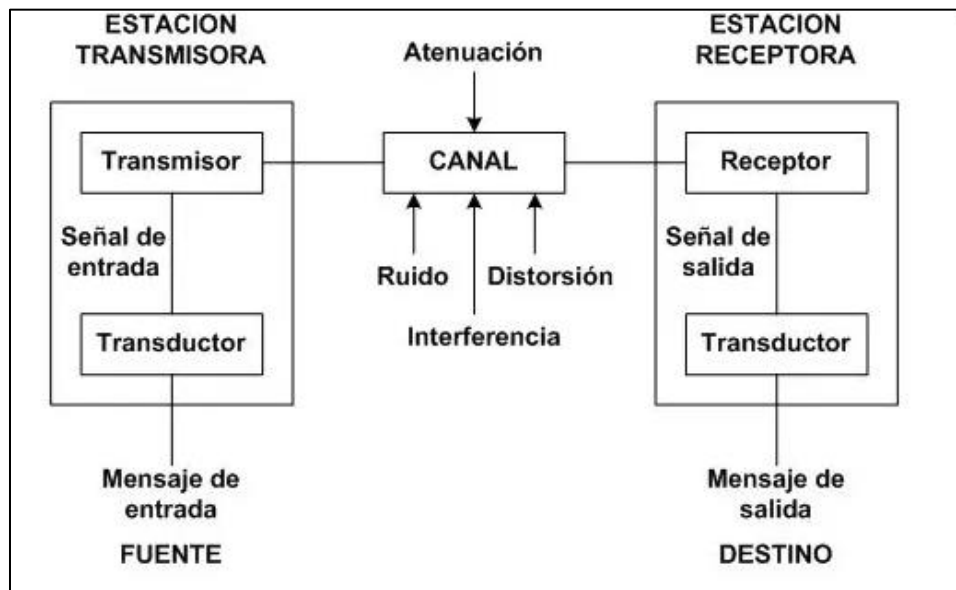


Figura 3. Etapas de un sistema de telecomunicación

Fuente: (Telectrónica, 2018)

Las telecomunicaciones sirven para transmitir información, la misma que puede adquirir infinitas formas o empaquetarse de múltiples maneras, y se los conoce como contenidos. Las redes y servicios de telecomunicación manejan los contenidos que pueden ser de cualquier naturaleza: películas, música, cursos de formación, páginas web, documentos, fotografías, vídeos o simple voz. Con las posibilidades tecnológicas actuales esos contenidos pueden estar almacenados en un servidor situado en cualquier lugar y ser accesibles desde todos los lugares del planeta. Es decir, están almacenados en la “nube”, lo que permite disponer de ellos con todo tipo de dispositivos y desde cualquier lugar donde el usuario se encuentre.

Con la creciente difusión y disponibilidad de equipos móviles tabletas y smartphones, aparecen unos programas sencillos llamados aplicaciones (apps), estos permiten realizar acciones concretas, jugar, editar fotografías o simplemente acceder a informaciones sobre temas específicos, tanto de ocio y entretenimiento como profesionales. Estos programas pueden ser gratuitos o pagados, son fácilmente descargables y quedan instaladas en los terminales inteligentes, de forma que su uso es rápido e inmediato.

El internet es la Red que ejemplifica de la manera más clara lo que significan las telecomunicaciones del siglo XXI, ya que ha presentado un sinnúmero de posibilidades al usuario, tales como disponibilidad de contenidos e interés de acceso universal. Las redes sociales, la banda ancha y la movilidad, son los pilares de un elemento que está transformando la sociedad y llevándola a una sociedad de la información y del conocimiento y a un nuevo mundo digital. (Roca Chillada, 2019) resume a las telecomunicaciones como una herramienta que está cambiando el mundo y todo lo que sucede en él y haciendo posibles cosas que hace poco no eran más que un sueño digital.

2.2.2. Minería de Datos

La minería de datos conocida también como Data Mining se encarga de buscar los patrones ocultos en los datos que permitirán predecir un comportamiento futuro. Los gobiernos, las empresas, y los científicos han usado este enfoque por años para transformar los datos en conocimientos proactivos y en la toma de decisiones, empleando una amplia variedad de técnicas, la información resultante la utilizan para incrementar sus ingresos, recortar costos, mejorar sus relaciones con clientes, reducir riesgos entre otras.

El objetivo de la minería es el análisis de grandes volúmenes de datos para la obtención de modelos y patrones predictivos o descriptivos, por medio del descubrimiento de conocimiento en las bases de datos, para esto se emplean técnicas de aprendizaje automático y estadísticas. La minería de datos (MD) apunta a la necesidad de metodologías de análisis inteligente de datos, las cuales puedan descubrir conocimiento útil de los mismos, así lo afirman (Riquelme, Ruiz, & Gilbert, 2006).

La minería de datos o Knowledge Discovery in Databases (KDD) es un campo de la estadística y las ciencias de la computación referido al proceso que intenta descubrir patrones en grandes volúmenes de conjuntos de datos. El Análisis predictivo y aprendizaje automático también son términos conocidos como minería de datos y usa principios estadísticos, los mismos que se documentan para detectar patrones en los datos obtenidos. En la minería de datos se aplican algoritmos que permiten predecir ciertas tendencias, identificar patrones, crear reglas, así como recomendaciones que son fruto del análisis de una secuencia de eventos de datos complejos y con esto se logra nuevos puntos de vista del comportamiento de los datos.

Este proceso permite entre otras cosas filtrar todo el ruido caótico y repetitivo en los datos, entender los datos que son relevantes y luego hacer un buen uso de esa información para

evaluar resultados probables, acelerar el ritmo de la toma de decisiones informadas en base a las predicciones realizadas. De forma general, los datos son la materia prima bruta, cuando el usuario les atribuye algún significado especial pasan a convertirse en información. Los especialistas son los que se encargan de elaborar o encontrar un modelo, logrando que la interpretación que resulta entre la información y ese modelo demuestre un valor agregado, sólo ahí se puede decir que existe un nuevo conocimiento.

Según (Bruner, 2018), la minería de datos es el conjunto de técnicas y tecnologías que permiten explorar grandes bases de datos con el objetivo de encontrar patrones que aporten información valiosa en la toma de futuras decisiones, este proceso tiene normalmente cuatro etapas:

- **Determinación de los objetivos:** el usuario determina qué objetivos quiere conseguir usando minería de datos, y el tipo de información que se desea extraer de la base de datos.

- **Procesamiento de los datos:** esta etapa requiere más trabajo ya que se tiene que seleccionar, limpiar, enriquecer, reducir y transformar la base de datos obtenida ya sea del cliente o de la institución que requiere usar minería de datos.

- **Determinación del modelo:** primero se debe realizar análisis estadísticos de los datos, para posteriormente llevar a cabo una visualización gráfica de los mismos y así conseguir una primera aproximación, cabe indicar que, esto depende de los objetivos planteados y la tarea que debe llevarse a cabo, para determinar los algoritmos idóneos desarrollados en diferentes áreas de la Inteligencia Artificial representando esto un desafío para el investigador ya que de esto depende la resolución o no del problema planteado.

- **Análisis de los resultados:** se deberán verificar si los resultados obtenidos son coherentes y se los coteja con los obtenidos por los análisis estadísticos y de visualización gráfica. El cliente determina si son novedosos y si en realidad le aportan un nuevo conocimiento para considerarlos en sus decisiones.

La minería de datos, como una disciplina compuesta, representa diversos métodos o técnicas que se utilizan en diferentes capacidades analíticas que abordan una gama de necesidades organizacionales, hacen diferentes tipos de preguntas y utilizan diferentes niveles de aportación humana o reglas para llegar a una decisión. Valcárcel (2004) manifiesta que para usar algunas técnicas y/o algoritmos de Minería de datos, es necesario tener la vista minable, esto quiere decir, contar con los datos preparados, conocer su descripción, y para seleccionar la técnica, se requiere saber el tipo de dato que se tiene, por otro lado, el método a utilizar depende del problema que se desea solucionar. Dentro de este marco se debe considerar que se debe realizar diferentes análisis para encontrar el algoritmo indicado.

Los problemas planteados, generalmente tienen dos tipos de enfoque: descriptivos y predictivos, para cada uno de ellos se tiene un grupo de técnicas que pueden ser usadas de acuerdo con el problema a resolver, las mismas que se describen a continuación:

- **Técnicas de Data Mining para Problemas Descriptivos:** El objetivo de resolver problemas descriptivos consiste en encontrar una descripción de los datos. En muchos de los casos se suele agrupar clientes con ciertas características semejantes para enviar notificaciones que sean más personalizadas. Según (Romero, 2019), las técnicas para este tipo de problemas son las siguientes:
 - ✓ **Análisis de Segmentación:** la intención de esta técnica es encontrar grupos similares, homogéneos en los datos. Para su solución se construyen modelos de aprendizaje no supervisados o clustering.

- ✓ **Análisis de Asociación:** el objetivo es obtener relaciones entre los datos, un ejemplo de esta técnica es cuando se requiere analizar el carrito de compras en un supermercado.

- **Técnicas de Data Mining para Problemas Predictivos:** se aplica para obtener modelos que se usarán para aplicarlo en datos futuros, esencialmente para predecir comportamientos. En inteligencia artificial se llaman modelos de aprendizaje supervisado. Las variables utilizadas pueden ser categóricas y numéricas.

- ✓ **Clasificación:** referente a modelos en los cuales la variable a predecir tiene valores definidos y son contables. Se usan las variables categóricas. Un ejemplo, es la predicción de si un cliente va o no a comprar cierto producto.

- ✓ **Predicción de Valores:** intenta predecir valores, por lo que usa en su mayoría variables numéricas, frecuentemente se obtiene la probabilidad de un evento: la probabilidad de que un cliente siga con su tarjeta de crédito, o de que adquiera una tarjeta adicional.

- ✓ **Otras técnicas de minería de datos:** como se sabe las técnicas de minería de datos provienen de la inteligencia artificial y de la propia estadística, se trata de algoritmos que se aplican sobre un conjunto de datos con el objetivo de obtener resultados. Las más representativas que se utilizan son:
 - ✓ **Las redes neuronales:** se refieren a un paradigma de aprendizaje y de un procesamiento automatizado. Es una de las técnicas más utilizadas para la formación de relaciones entre datos, es un sistema rápido y eficaz pero que a su vez cuenta con insuficiencias en los espacios multidisciplinares donde puedan relacionarse más de 2 variables. En este sistema se puede interconectar las neuronas en una red (red neuronal) que presta colaboración para la producción de estímulos de salida. La regresión lineal y los mapas autoorganizados son ejemplos de este tipo de técnica.

- ✓ **Los árboles de decisión:** se refiere a un modelo de predicción que se utiliza en el campo de la inteligencia artificial a partir de una base de datos en donde se construyen los diagramas de construcciones lógicas. Es un sistema similar a la predicción que está basado en reglas que sirven para la representación de una serie de condiciones que se presentan de manera sucesiva en la resolución de problemas.
- ✓ **Los modelos estadísticos:** hace referencia a una expresión simbólica en forma de igualdad que es empleada en los diseños experimentales y en la regresión, con la intención de identificar los factores que modifican la variable de respuesta.
- ✓ **El agrupamiento:** consiste en la agrupación de una serie de vectores según determinados criterios que habitualmente son a distancia. Se refiere a la disposición de los vectores de entrada de manera que estén más cercanos a los que tengan características comunes.
- ✓ **Algoritmos de Minería de datos:** según el objetivo que tenga la realización del análisis los algoritmos se pueden clasificar como algoritmos supervisados, que predicen un dato desconocido inicialmente a partir de otros datos que son de conocimiento previo. Y los algoritmos no supervisados, los cuales descubren patrones y tendencias que presentan los datos, dentro de los algoritmos más usados están los siguientes:
 - ✓ **Árboles de decisión:** es una secuencia de decisiones que se organizan de manera jerárquica, a manera de ramas de un árbol. Estos algoritmos aceptan tanto datos numéricos como categóricos. Este algoritmo se aplica frecuentemente para tareas de clasificación, agrupación y pronóstico, cuando predicen categorías se suele llamar árboles de clasificación, sin embargo, si son numéricos y se pretende predecir, se llaman árboles de regresión.
 - ✓ **Redes neuronales artificiales:** estos algoritmos son muy potentes, y permiten modelar prácticamente con cualquier tipo de problemas, además realizan tareas de clasificación, predicción y agrupación, sin embargo, presenta desventajas y una de

ellas es que trabajan con datos numéricos. Se debe tomar en cuenta que las variables categóricas se suelen discretizar para aplicar estos algoritmos. Las redes se pueden concebir como grafos con nodos y enlaces, se organizan por capas, la primera es de entrada, las siguientes son denominadas ocultas, y finalmente la capa de salida.

- ✓ **Análisis Factorial:** el objetivo del análisis factorial es disminuir la dimensión de los datos, es una especie de combinación que condensan la información completa. El algoritmo está diseñado para trabajar con variables cuantitativas.
- ✓ **Análisis de Correspondencia:** este algoritmo es el indicado para resolver problemas de dimensionalidad con variables categóricas, se toma en cuenta el análisis de correspondencia simple, que evalúa dos variables; toma como base la tabla de contingencia. y el análisis de correspondencias múltiples, que considera más de dos variables.
- ✓ **Escalamiento Multidimensional:** es un algoritmo usado para representar de forma gráfica a través de un mapa perceptual las similitudes que tienen los objetos de una nube de datos, considerando el posicionamiento entre ellos. Es muy parecido al análisis clúster con la diferencia que en este modelo no se conocen las variables para determinar la semejanza, mientras que en el clúster sí.
- ✓ **El algoritmo K-means:** el agrupamiento de los medios K es uno de los algoritmos de aprendizaje de máquinas no supervisadas más simples y populares. Típicamente, los algoritmos no supervisados hacen inferencias a partir de conjuntos de datos usando sólo vectores de entrada sin referirse a resultados conocidos o etiquetados.
- ✓ **Algoritmo del vecino k más cercano:** K-Nearest Neighbors es uno de los algoritmos de clasificación más conocido y básico pero esencial en el aprendizaje por máquina, pertenece al dominio del aprendizaje supervisado y encuentra una intensa aplicación en el reconocimiento de patrones, la minería de datos y la detección de intrusos.

- ✓ **Naive Bayes:** Naive Bayes es un algoritmo probabilístico de aprendizaje automático basado en el Teorema de Bayes, utilizado en una amplia variedad de tareas de clasificación. En este post, obtendrá una comprensión clara y completa del algoritmo Bayes Ingenuo y todos los conceptos necesarios para que no haya lugar a dudas ni a lagunas en la comprensión.

- ✓ **Algoritmo CART:** este algoritmo está estructurado como una secuencia de preguntas, cuyas respuestas determinan cuál será la siguiente pregunta, si es que la hay. El resultado de estas preguntas es una estructura en forma de árbol donde los extremos son nodos terminales en cuyo punto no hay más preguntas. La utilización de árboles de decisión, y entre ellos los de tipo CART (Classification and Regression Trees), aporta una aproximación que intenta solucionar problemas y muestra, en algunas aplicaciones, ciertas ventajas respecto a otros modelos (Trujillano, y otros, 2008).

En función de lo planteado la minería de datos se está convirtiendo en uno de los trabajos con mayor proyección para el futuro, ya que ahorra grandes cantidades de dinero para las empresas y abre nuevas oportunidades de negocio.

2.2.3. *Big Data*

El concepto de "big data" no es nuevo ni recientemente inventado, los inicios de grandes volúmenes de datos se remontan a las décadas de 1960 y 1970, cuando el mundo de los datos acababa de empezar con los primeros centros de datos y el desarrollo de las bases de datos relacionales. A mediados del 2005, los ingenieros que manejan Datos empezaron a preocuparse por la cantidad de datos que generaban los usuarios a través de Facebook, YouTube y otros servicios online. En ese mismo año, se plasmaría la idea de desarrollo de Hadoop, un marco de código abierto, diseñado con la finalidad de almacenar y analizar conjuntos de datos extremadamente grandes (Oracle.com, 2020).

Con la creación de software con código abierto tales como Hadoop, y unos de los más recientes Spark, son la base para el crecimiento y el manejo del big data, pues estas tecnologías ayudaban a que el big data resultase más fácil de usar y más barato de almacenar. En el transcurso de los años el volumen de big data se ha disparado. Los usuarios continúan generando grandes cantidades de datos.

Las siete V de Big data

Por su terminología las tres “V” referidas son: volumen, variedad y velocidad, las cuales representan las tres características principales que lo definen, además de sus principales retos. El Big Data hace referencia a volúmenes de datos tan grandes que resulta imposible que pueden ser accedidos, almacenarse ni procesarse en un único lugar, por lo que requieren de sistemas específicos para ello (UNIR, 2020).

A continuación, se detalla cada una de las características nombradas:

- ✓ **Volumen:** hace referencia a la cantidad de datos del Big Data, los cuales se generan y recopilan constantemente. Se estima que en el año 2020 se hayan superado ampliamente los 25 mil millones de dispositivos remotos como: ordenadores de escritorio, portátiles, tablets, servidores, conectados a Internet. La cantidad de datos importa, con big data, tendrá que procesar grandes volúmenes de datos no estructurados de baja densidad. Puede tratarse de datos de valor desconocido, como feeds de datos de Twitter, cantidad de clics de una página web o aplicación para móviles. Para algunas empresas u organizaciones, esto puede generar decenas de terabytes de datos. Para otras, incluso cientos de petabytes (Oracle.com, 2020).
- ✓ **Velocidad:** la velocidad se refiere a la rapidez con que los datos son creados, accedidos, almacenados y procesados. Continuamente se están generando datos por segundo, es esta la razón de que las plataformas de Big Data deben ser capaces de acceder a dichos datos,

almacenarlos y tratarlos para que las organizaciones puedan obtener los resultados y de esta manera tomar mejores decisiones basadas en dichos resultados (UNIR, 2020).

La velocidad se refiere al ritmo con el que se reciben los datos, la mayor velocidad de los datos normalmente se transmite directamente a la memoria. Algunos productos inteligentes habilitados para Internet tienen su funcionamiento prácticamente en tiempo real y requieren una evaluación y actuación en tiempo real es por eso la importancia de la velocidad al momento de procesar datos (Oracle.com, 2020).

- ✓ **Variedad:** los datos generados para Big Data no son homogéneos, esto significa, que tienen heterogéneas y diferentes características entre todos los datos, lo cual indica que no se puede procesar los datos de la igual manera y en una única iteración. La razón de que los datos no son homogéneos ya que provienen de distintos orígenes o fuentes (archivos, bases de datos, sensores, webs...), adicional son de distintos tipos (texto, numéricos, imágenes...) y tienen distintos formatos (timestamp, fecha...). Los tipos de datos convencionales eran estructurados (estructura definida perfectamente por su formato, tipo y tamaño como por ejemplo los de las BBDD relacionales), semiestructurados (su estructura no está estrictamente definida pero sí están organizados, como por ejemplo XML, JSON, HTML...) y no estructurados (sin formato específico, como documentos de Word, PDFs...), con todo lo anteriormente explicado para procesar estos datos se requieren preprocesamientos adicionales con la finalidad de obtener significado y habilitar los metadatos (Oracle.com, 2020).

- ✓ **Viabilidad:** es la capacidad para el manejo eficaz de el gran volumen de los datos obtenidos de esta manera transformarlos en información que le aporte a la toma de decisiones acertadas y de esta manera diseñar estrategias eficaces (UNIR, 2020).

- ✓ **Veracidad:** big data permite la aplicación de métodos y soluciones que garantizan la fiabilidad de los datos generados y almacenados, eliminando cualquier inexactitud o

incertidumbre en la recopilación de estos, vinculándose en este punto el tratamiento de los datos o limpieza de ellos; esto es un gran desafío a medida que los datos aumentan y crecen significativamente las bases de datos (Datamall, 2018). Se necesita información útil y de calidad, el grado de fiabilidad se lo debe mantener controlado y el índice de calidad debe ser muy alto, ya que de esta información dependerá la toma de decisiones (Prometeusgs, 2019).

- ✓ **Visualización:** se trata de la forma en que los datos son presentados con big data, luego de que se han recopilado, procesado y modelado. Se pueden representar de forma legible y accesible, a fin de encontrar patrones y claves ocultas en la data que se analizará con el fin de obtener conclusiones, (Techoarget, 2020). Cuando se habla de visualización se refiere al modo en el que los datos son presentados, una vez que los datos son procesados (los datos están en tablas y hojas de cálculo), se requiere representarlos visualmente de manera que sean legibles y accesibles, para encontrar patrones y claves ocultas en el tema a investigar (IIC, 2019).

Algunos de los beneficios de la visualización de datos son:

- La capacidad de absorber información rápidamente, mejorar los conocimientos y tomar decisiones más rápidas.
- Una mayor comprensión de los próximos pasos que deben tomarse para mejorar y tomar decisiones.

Según (Techoarget, 2020), la visualización de una gran cantidad de volúmenes de datos va más allá de las técnicas comunes utilizadas en la visualización normal, como lo son gráficos circulares, histogramas y gráficos corporativos. En su lugar, utiliza representaciones más complejas, como mapas de calor (HotSpot) y relojes de tiempo.

- ✓ **Valor:** el valor se obtiene de datos que se transforman en información, disponer de los datos puede resultar fácil, lo arduo y verdaderamente diferencial es el uso que se da de estos para comprender. De acuerdo con (IIC, 2019), en el caso de la tecnología Big Data, se puede decir que la información es valor, ya que después de haber obtenido los datos, haberlos almacenado, clasificado por tipología y analizado su veracidad, llega el momento de aplicar los resultados a la estrategia, el valor de los datos está en que sean accionables, es decir, que con la información de los datos se puedan tomar mejores decisiones.

El hecho de almacenar datos no es de mucha ayuda, es necesario obtener el valor de ellos, y para esto es preciso transformarlos en conocimiento y posteriormente en acciones. El Big Data hace uso de tecnologías aplicadas para lograr que los datos te aporten información de calidad y, así, se pueda llegar efectivamente al público adecuado y de esta manera gestionar bien el negocio (Cognodata, 2019).

La meta de Big Data es convertir la información en conocimiento y para esto se usa la analítica la cual nos ayuda en la toma de decisiones. En conclusión, se usa los datos obtenidos de un modo rentable y su vez eficiente. la parte medular del Big Data es subir el siguiente nivel: usar y analizar los datos para ser más inteligentes.

Adicionalmente, se podrían incluir muchas otras propiedades como los cambios en el tiempo (variabilidad), la confiabilidad de las fuentes (veracidad) y qué tan útiles son (valor). Estas propiedades pueden estar dentro del mismo conjunto, pero antes de analizarlos es necesario depurarlos, estandarizarlos y conectarlos o catalogarlos de manera adecuada. A este proceso se le denomina gestión de datos, así lo afirman (Rodríguez, Palomino, & Mondaca, 2017).

Para la gestión de datos se debe tener en cuenta los metadatos, información complementaria que puede indicar su origen, calidad o su confiabilidad, fechas de creación y modificación.

Por ejemplo, si se tiene una colección de fotografías, éstas serían los datos y los metadatos pueden ser la fecha de creación, si ha sido editada o no, el autor, la cámara utilizada, etc. La información útil o sensible de la cual se puede obtener información primordial pueden estar tanto en los datos como en los metadatos. Para extraer sus cualidades se utilizan métodos avanzados de procesamiento que manejan el vasto volumen y su complejidad. A todo este conjunto de técnicas se le conoce como Big Data Analytics o simplemente Analytics.

Algunos ejemplos de estas técnicas son:

- ✓ **Minería de datos (Data Mining):** se usa para encontrar patrones que no puedan localizarse con métodos tradicionales, ya sea por la complejidad o por el volumen del conjunto de datos (Microsoft.com, 2018) . Este concepto es frecuentemente mal usado ya que su nombre pareciera indicar que consiste en extraer los datos, pero realmente consiste en obtener conocimiento de ellos, como identificar patrones de relación.

- ✓ **Aprendizaje computacional o automático (Machine learning):** Es una aplicación de la inteligencia artificial donde una máquina o computadora “aprende” a partir de un gran número casos, este aprendizaje consiste en automatizar, mediante distintos algoritmos, la identificación de patrones o tendencias que se “esconden” en los datos. Por ello, resulta muy importante no sólo la elección del algoritmo más adecuado, sino también el hecho de disponer de un gran volumen de datos de suficiente calidad. (Recuero de los Santos, 2018).

El big data puede ayudarle a abordar una serie de actividades empresariales, desde la experiencia de cliente hasta la analítica. Además, le aporta nuevas perspectivas que abren paso a nuevas oportunidades y modelos de negocio. Para iniciar en esto requiere de tres acciones clave:

1. Integrar

El big data concentra datos de numerosas fuentes y aplicaciones distintas. Los mecanismos de integración de datos convencionales, tales como ETL (extract, transform, load [extraer, transformar, cargar]), generalmente no están a la altura de dicha tarea. Analizar conjuntos de big data de uno o más terabytes, o incluso petabytes, de tamaño requiere de nuevas estrategias y tecnologías. Durante la integración, es necesario incorporar los datos, procesarlos y asegurarse de que estén formateados y disponibles de tal forma que los analistas empresariales puedan empezar a utilizarlos.

2. Gestionar

El big data requiere almacenamiento. Su solución de almacenamiento puede residir en la nube, on premises o en ambos. Puede almacenar sus datos de cualquier forma que desee e incorporar los requisitos de procesamiento de su preferencia y los motores de procesamiento necesarios a dichos conjuntos de datos on-demand. Muchas personas eligen su solución de almacenamiento en función de dónde residan sus datos en cada momento. La nube está aumentando progresivamente su popularidad porque es compatible con sus requisitos tecnológicos actuales y porque le permite incorporar recursos a medida que los necesita.

3. Analizar

La inversión en big data se rentabiliza en cuanto se analizan y utilizan los datos. Adquiera una nueva claridad con un análisis visual de sus diversos conjuntos de datos. Continúe explorando los datos para realizar nuevos descubrimientos. Comparta sus hallazgos con otras personas. Construya modelos de datos con aprendizaje automático e inteligencia artificial. Ponga sus datos a trabajar.

Según (PowerData, 2020), el big data es útil para muchas empresas ya que proporciona respuestas a muchas preguntas que las empresas ni siquiera sabían que tenían y que necesitaban, proporciona un punto de referencia para la toma de decisiones. Con una cantidad tan grande de información, los datos pueden ser moldeados o probados de cualquier manera

que la empresa considere adecuada. Al hacerlo, las organizaciones son capaces de identificar los problemas de una forma más comprensible.

Empresas como Netflix y Procter & Gamble usan big data para prever la demanda de los clientes. Construyen modelos predictivos para nuevos productos y servicios clasificando atributos clave de productos anteriores y actuales, y modelando la relación entre dichos atributos y el éxito comercial de las ofertas. Además, P&G utiliza los datos y la analítica de grupos de interés, redes sociales, mercados de prueba y avances de salida en tiendas para planificar, producir y lanzar nuevos productos (Oracle.com, 2020).

El machine learning o aprendizaje automático es un tema que en la actualidad es muy usado. Los datos, concretamente big data, es uno de los motivos de que así sea. En lugar de programarse, las máquinas pueden aprender. Esto es posible gracias a la disponibilidad de big data para crear modelos de machine learning.

De acuerdo con lo que afirma (Oracle.com, 2020), la eficiencia operativa no es un tema que se lo mire muy a menudo, pero es el área en que big data tiene un mayor impacto. El big data le permite analizar y evaluar la producción, la opinión de los clientes, las devoluciones y otros factores para reducir las situaciones de falta de stock y anticipar la demanda futura. El big data también puede utilizarse para mejorar la toma de decisiones en función de la demanda de mercado en cada momento.

2.2.4. Base de Datos Espacial (Geodatabase)

La Geodatabase es un modelo que permite el almacenamiento físico de la información geográfica, ya sea en archivos dentro de un sistema de ficheros o en una colección de tablas en un Sistema Gestor de Base de Datos sean estos Oracle, Microsoft SQL Server, IBM DB2 e Informix (Jimpako, 2008).

De la misma forma se puede decir que es la estructura de datos nativa para ArcGIS y es el formato de datos principal que se utiliza para la edición y administración de datos. Las geodatabases cuentan con un modelo de información integral para representar y administrar información geográfica. Este modelo de información integral se implementa como una serie de tablas que almacenan clases de entidad, datasets ráster y atributos (www.esri.com, 2019).

Una geodatabase es una estructura de datos, un contenedor físico para la nuestra información geográfica, para lograr acceder y trabajar con esta instancia física del conjunto de datasets se lo hace a través de ArcGIS o mediante un sistema de administración de bases de datos utilizando SQL, las geodatabases poseen un modelo de transacción para administrar flujos de trabajo de datos SIG. En otras palabras, se trata de un “contenedor” para los datos. (Alonso, 2016).

Una de las principales ventajas de usar información espacial en una geodatabase y no en archivos del sistema es que se aprovechan las ventajas del SGBDR (Sistema Gestor de Base de Datos Relacional), esta incluye:

- ✓ Restricciones de acceso y seguridad de la información.
- ✓ Soporte para SQL (StructuredQueryLanguage) para realizar consultas espaciales complejas.
- ✓ La arquitectura cliente servidor de la base de datos permite que múltiples usuarios realicen consulta y edición simultánea

Según (Fernández, 2014), las Geodatabases garantizan una alta interoperabilidad, información cartográfica ordenada y jerarquizada, alto control del tipo de dato a introducir, así como de su formato (cifras decimales y enteros), no acepta duplicidad, permite restringir

los datos utilizando dominios, para así disminuir errores en atributos, permite utilizar y almacenar reglas de topología, e incluso programar reglas propias.

La geodatabase es un "contenedor" utilizado para alojar un conjunto de datasets. Se pueden clasificar entres tipos:

1. **Geodatabases de archivos:** almacenados como carpetas en un sistema de archivos. Cada dataset se aloja como un archivo que puede escalar hasta 1 TB de tamaño. Las geodatabases de archivos se recomiendan por sobre las geodatabases personales. Este tipo de geodatabase proporciona una solución de geodatabase ampliamente disponible, sencilla y escalable para todos los usuarios tiene una cualidad de que es de proporcionar una fácil portabilidad que funcione en todos los sistemas operativos.
2. **Geodatabases personales:** todos los datasets se almacenan dentro de un archivo de datos de Microsoft Access con un límite de tamaño de 2 GB. Este tipo de geodatabase tiene una peculiaridad la cual es que solo son en Microsoft Windows.
3. **Geodatabases corporativas:** también conocidas como geodatabases multiusuario, pueden no tener límite de tamaño y cantidad de usuarios. Se almacenan en una base de datos relacional con Oracle, Microsoft SQL Server, IBM DB2, IBM Informix o PostgreSQL. Las geodatabases corporativas se utilizan principalmente en una amplia variedad de opciones de configuración individuales, de grupos de trabajo, de departamentos y corporativas. Aprovechan al máximo las arquitecturas DBMS el cual es sistema manejador de bases de datos.

La geodatabase corporativa admite administración de datos en un entorno multiusuario utilizando un método llamado versionado. El versionado es un mecanismo en el que todos los cambios de la base de datos se registran como filas en tablas. Por ejemplo, cada vez que

actualiza un valor en una fila, la fila antigua se "retira" y se agrega una fila nueva actualizada (ARCGIS, 2019).

2.2.5. Modelo de Predicción “Regresión Geográficamente Ponderada (GWR)

La regresión geográficamente ponderada (GWR), tiene sus inicios a mediados de los años noventa del siglo pasado, en la estadística, en particular la estadística espacial, puede ayudar en la interpretación de datos en las pruebas estándar, por una muy sencilla razón: el lugar sí importa. La GWR es un tipo novedoso de regresión que permite una mejor comprensión de los modelos, siempre y cuando se cuente con una variable espacial. La GWR permite abordar los modelos localmente y evaluarlos de cara al modelo global. Además, contribuye a desarrollar un análisis de regresión apropiado sobre un conjunto de datos, partiendo esencialmente de la dependencia espacial y asumiendo el carácter local de los datos. No contemplar la espacialidad de los datos sin duda lleva a sacar conclusiones erradas de un análisis de regresión y emergen problemas de especificación en los modelos cuando la dependencia espacial está presente en los datos (Ruiz-Linares, Guzmán-Patiño, & Castro-Ávila, 2017).

El análisis de predicción se encuentra relacionado con las técnicas de regresión espacial. El objetivo de este tipo de análisis es descubrir la relación entre variables ya sean independientes o dependientes (Dueñas-Reyes, 2009).

El término regresión geográficamente ponderada (GWR), fue introducido con el fin de aludir a una familia de modelos de regresión “ajustados al espacio” donde es posible observar las variaciones espaciales de los parámetros estimados y con ello saber exactamente dónde y cuánto es el efecto de una variable explicativa sobre la dependiente. Este modelo GWR tiene la capacidad de capturar la estacionalidad espacial, es decir, las variaciones espaciales, al

permitir que los parámetros de la regresión varíen con la ubicación. Para lograrlo en el modelo de regresión geográficamente ponderada se incorporan coordenadas espaciales.

La regresión ponderada geográficamente (GWR), es una de las técnicas de regresión espacial que se utiliza en geografía y otras disciplinas. La GWR evalúa un modelo local en base de variables, o el proceso que trata de comprender o predecir ajustando una ecuación de regresión a todas las variables del dataset. GWR construye estas ecuaciones individuales mediante la incorporación de las variables dependiente y explicativa de las entidades que caen dentro de la vecindad de cada entidad de destino.

La forma y extensión de todas las vecindades que se analizan se basan en los parámetros: Tipo de vecindad y Método de selección de vecindad. GWR se debe aplicar a los datasets con varios cientos de entidades. No es un método adecuado para datasets pequeños y no funciona con datos de multipunto (www.esri.com, 2019).

Esta técnica hace posible explorar si la asociación entre llamadas mal intencionadas y sus variables explicativas es constante en toda la provincia o si es posible identificar variaciones por cantones.

Existe un modelo de regresión que se ha creado para realizar predicciones para otras entidades (puntos o polígonos) en la misma área de estudio. Crear dichas predicciones requiere que cada una de las ubicaciones de predicción tenga valores para cada una de las variables explicativas proporcionadas. Si los nombres de campo de los parámetros Entidades de entrada y ubicaciones de predicción no coinciden, se proporciona un parámetro variable coincidente. Al comparar las variables explicativas, los campos de los parámetros entidades de entrada y ubicaciones de predicción deben ser del mismo tipo (los campos dobles deben corresponderse con los campos dobles, por ejemplo) (ARCGIS, 2019).

2.2.6. Mapas de Calor

Los mapas de calor muestran las áreas con más entidades de puntos calientes, los cuales ayudan a responder a preguntas sobre los datos, por ejemplo: ¿cuál es la distribución geográfica de los puntos calientes? Estos mapas son más una ayuda visual que una forma precisa de mostrar la densidad de puntos y es mejor utilizarlos junto con otro tipo de perspectiva (www.esri.com, 2019).

Estos mapas constituyen una ayuda visual más que una forma precisa de mostrar la densidad de puntos, existen variadas formas en las que se puede representar visualmente los datos. Los mapas de calor son fácilmente interpretables siempre y cuando se conozca lo que se desea medir. Esto permite que las personas que los vean no necesiten saber de alguna otra ciencia para saber lo que se está representando. De esta forma, de un vistazo se puede sacar una conclusión bastante acertada y clara, lo que conllevaría a una toma de decisión mucho más rápida y clara. Los mapas de calor se usan en muchos ámbitos debido a su capacidad de representar cantidad importante de información de una manera rápida y sencilla para cualquier usuario, independientemente de sus conocimientos, los mapas de calor ayudan a interpretar los resultados de los estudios realizados.

Un ejemplo de mapa de calor se puede observar en la Figura 4.

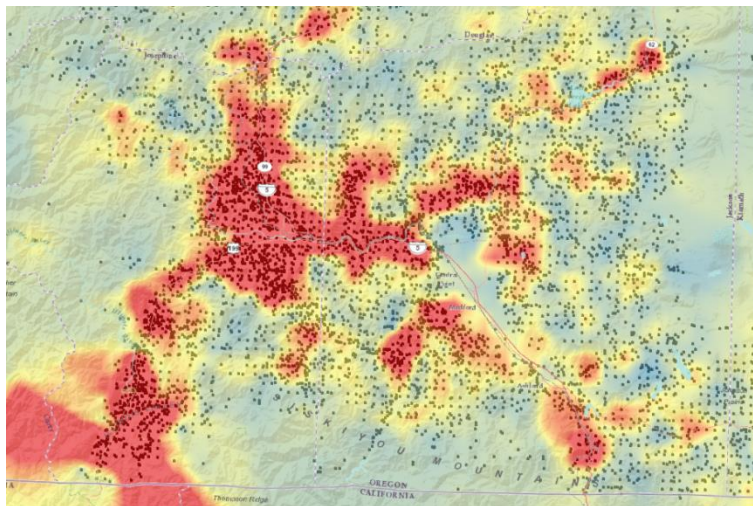


Figura 4. Mapa de calor

Fuente: (Oregon State University, 2016)/

2.2.7. Reloj de Datos

El reloj de datos es un diagrama bidimensional que permite analizar patrones en la distribución temporal de sus datos en dos frecuencias diferentes. Al momento de trabajar con datos temporales, es muy importante que se entienda la distribución temporal de los datos para que sea útil entender cuántos datos existen en períodos diferentes de tiempo.

El Gráfico de reloj de datos es un gráfico circular. Está dividido en celdas mediante una combinación de círculos concéntricos y líneas radiales, similar a los radios en una rueda de bicicleta. Los círculos concéntricos dividen los datos de una manera como, por ejemplo, mes del año y las líneas radiales dividen los datos de una manera diferente.

Con el fin de tener claro estos conceptos, se va a imaginar que dispone de una capa de rastreo las incidencias de llamadas mal intencionadas. El histograma en Administrador de reproducción podría indicar que las llamadas son constantes durante un período de 12 meses. Sin embargo, podría haber un patrón oculto en los datos que no está viendo. Después de crear un reloj de datos para trazarlos por mes y también por hora del día, en este instante se observará un patrón, el mismo que sería difícil observar utilizando otros métodos de análisis de datos.

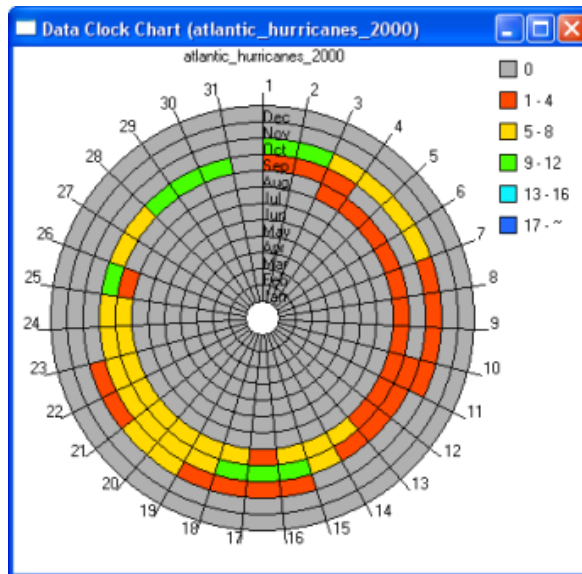


Figura 5. Reloj de Datos

Fuente: (ARCGIS, 2020)

El reloj de datos es un gráfico circular como se muestra en la Figura 5. Está dividido en celdas mediante una combinación de círculos concéntricos y líneas radiales, similar a los radios en una rueda de bicicleta. Los círculos concéntricos dividen los datos, por ejemplo, mes del año y las líneas radiales dividen los datos de una manera diferente, por ejemplo, día del mes (www.esri.com, 2019).

2.3.Servicio Integrado de Seguridad (SIS) ECU 911

La Zona de Planificación 1 (Norte), cuya sede administrativa se encuentra ubicada en la ciudad de Ibarra, corresponde a la frontera norte del país, la cual está conformada por las provincias de Esmeraldas, Carchi, Imbabura y Sucumbíos (Senplades, 2018). Con este antecedente cabe indicar que el Servicio Integrado de Seguridad (SIS) ECU-911 Zonal 1 en Ibarra se inauguró el lunes 5 de mayo del 2014, cuenta con una plataforma tecnológica que articula Cámaras de Videovigilancia, Dispositivos de Atención Inmediata (DAI), Despacho de recursos los cuales tienen como propósito atender las emergencias en los cantones de Ibarra, Otavalo, Antonio Ante, Cotacachi, Pimampiro y Urcuquí, mismos que conforman la

provincia de Imbabura. En Ibarra el SIS ECU 911 empieza su funcionamiento con 80 cámaras a nivel de la provincia, a las cuales también se suman 19 cámaras denominadas ojos de águila que pertenecen al GAD Ibarra, 13 cámaras de Antonio Ante. El SIS ECU 911 Ibarra, tiene dos áreas agregadores de valor: Video vigilancia y despacho de emergencia en la cual se contesta las llamadas y se despacha recursos; los cuales funcionan las 24 horas del día, los 365 días del año, y atienden un promedio mensual de emergencias reales de 36.215 según fuentes oficiales (ECU 911, 2015).

El mal uso de la línea ha provocado un decremento en los tiempos de respuesta en cuanto a la coordinación de recursos para atención de emergencias y a su vez gastos innecesarios. Según (ECU 911, 2019), el SIS ECU 911 Ibarra cuenta con 21 consolas de despacho desde donde se enviarán las unidades de respuesta necesarias para atender las diferentes emergencias dentro de la provincia de Imbabura, 14 consolas de recepción de llamadas, 15 consolas de video vigilancia. A este centro de atención se articulan diferentes instituciones, tales como: Policía Nacional, Cuerpo de Bomberos, Fuerzas Armadas, Gestión de Riesgos, Agentes de Tránsito, Ministerio de Salud, Instituto Ecuatoriano de Seguridad Social (IESS), Agencia Nacional de Tránsito (ANT), entre otras.

✓ *Mal uso de línea de emergencia*

Cuando un alertante realiza el uso de la línea única de emergencias las cuales NO están relacionadas con incidentes o emergencias, existen algunas razones de este tipo de llamadas como: El descuido en el uso del teléfono o marcaciones accidentales, llamadas sin interlocutor o silenciosas, y llamadas mal intencionadas que atenten contra la integridad del Servicio, entre otros (ECU 911, 2015).

Este Tipo de llamadas afectan al Servicio Integrado de Seguridad ECU 911, debido a que congestionan el canal de acceso de la ciudadanía y al servicio de emergencias, ocasionando saturación en las líneas telefónicas y se dejen de atender emergencias reales.

✓ *Llamada mal intencionada*

Las llamadas malintencionadas son las que no se relacionan con la atención o gestión de situaciones de emergencia, y que no implican desplazamiento de recursos de las entidades de atención de servicios de emergencia, tales como:

- Llamadas sin interlocutor
- Insultos o burlas
- Agresiones o acoso sexual
- Insinuaciones que atentan contra la integridad del servicio y del servidor

✓ *Llamada Falsa*

Según (ECU 911, 2015), son aquellas que hacen uso indebido del 9-1-1, reportando eventos falsos que implican el desplazamiento innecesario de recursos para atención de emergencias. Este tipo de llamadas derivan en desperdicio de tiempo, recursos económicos y técnicos, humanos, equipamiento e infraestructura, necesarios para la atención de situaciones de emergencia real, limitando el accionar operativo en caso de que hubiera una emergencia real.

2.4.Marco legal

La Agencia de Regulación y Control de las Telecomunicaciones (ARCOTEL), con la información reportada por el Servicio Integrado de Seguridad ECU 911 entrega a las empresas operadoras de telefonía el listado de números telefónicos desde los que se

realizaron llamadas mal intencionadas, el Artículo 10 de la Resolución Tel-756-21-CONATEL-2011 detalla la clasificación de las afectaciones (CONATEL, 2011).

Sobre las medidas y aplicación, a fin de evitar que se afecten derechos constitucionalmente protegidos para la ciudadanía, la ARCOTEL dispondrá a los prestadores de servicios finales de telecomunicaciones, la implementación de las siguientes medidas, que consta en la Resolución mencionada:

- **Suspensión de corto plazo.** - por realizar más de tres llamadas no relacionadas con la atención de emergencias dentro de un mes. Las líneas son suspendidas parcialmente por 30 días calendario; únicamente podrá recibir llamadas y SMS en su terminal. Esta medida se aplicará en caso de afectaciones leves (ARCOTEL, 2019).
- **Suspensión de largo plazo.** - por reincidencia de llamadas de mal uso del servicio en periodo de un año o la movilización innecesaria de recursos (ambulancias, patrulleros, entre otros). Las líneas son suspendidas totalmente y bloqueado el terminal por un periodo de 180 días calendario. En el caso del Servicio Móvil Avanzado (SMA), adicionalmente bloqueo del terminal asociado, cuyo IMEI del dispositivo se incluirá en la base de datos terminales suspendidos por realización de llamadas que afectan a servicios de emergencia, de las operadoras del SMA (ARCOTEL, 2019).
- **Definitiva.** - Por reincidencia en afectaciones graves se suspende definitivamente el servicio y se bloquea permanentemente el terminal móvil. En el caso del Servicio Móvil Avanzado (SMA), adicionalmente se generará el bloqueo definitivo del terminal asociado, cuyo IMEI se incluirá en la base de datos de terminales suspendidos por realización de llamadas que afectan a servicios de emergencia (ARCOTEL, 2019).

Existe además sanciones estipuladas en el Código Orgánico Integral Penal (COIP), en su Art. 396.- Contravenciones de cuarta clase. - Será sancionada con pena privativa de libertad de quince a treinta días: literal 3. La persona que de manera indebida realice uso del número único de atención de emergencias para dar un aviso falso de emergencia y que implique desplazamiento, movilización o activación innecesaria de recursos de las instituciones de emergencia (COIP, 2008).

Prisión Según el COIP en el **Art. 396.- Contravenciones de cuarta clase.** - Será sancionada con pena privativa de libertad de quince a treinta días: y en su literal indica:

La persona que de manera indebida realice uso del número único de atención de emergencias para dar un aviso falso de emergencia y que implique desplazamiento, movilización o activación innecesaria de recursos de las instituciones de emergencia (COIP, 2008).

CAPÍTULO III: MARCO METODOLÓGICO

3.1.Descripción del área de estudio

El Ecuador está dividido en 9 zonas de planificación, 140 distritos y 1,134 circuitos administrativos según los datos del SIN (2015). Según el proceso de transformación del Estado que plantea niveles administrativos de planificación pública. Imbabura está considerada dentro de la Zona 1 Norte, junto con las provincias de Carchi, Esmeraldas, Imbabura y Sucumbíos.

La provincia de Imbabura está dividida en 3 distritos y 44 circuitos administrativos que se detallan en la Tabla 2.

Tabla 2. Descripción de distritos y circuitos de Imbabura

Nro.	Provincia	Nombre	Nro. Circuitos	Población
10D01	Imbabura	Ibarra- Pimampiro - Urcuquí	22	209,816
10D02		Antonio Ante, Otavalo	16	148,392
10D03		Cotacachi	6	40,036

Fuente: Decreto Ejecutivo 988 (2012).

Imbabura se conforma de 6 cantones Ibarra, Antonio Ante, Cotacachi, Otavalo, Pimampiro y Urcuquí, su capital es Ibarra, la ciudad más grande y poblada. Tiene una superficie total de 4,587.51 km² representando el 1.79 % de la superficie nacional, cuenta con una población de 398,244 habitantes según el Censo 2010 mostrados por el INEC (2010b). Limita al norte con Carchi, al sur con Pichincha, por el occidente con Esmeraldas y al este con Sucumbíos, como se muestra en la Figura 6.

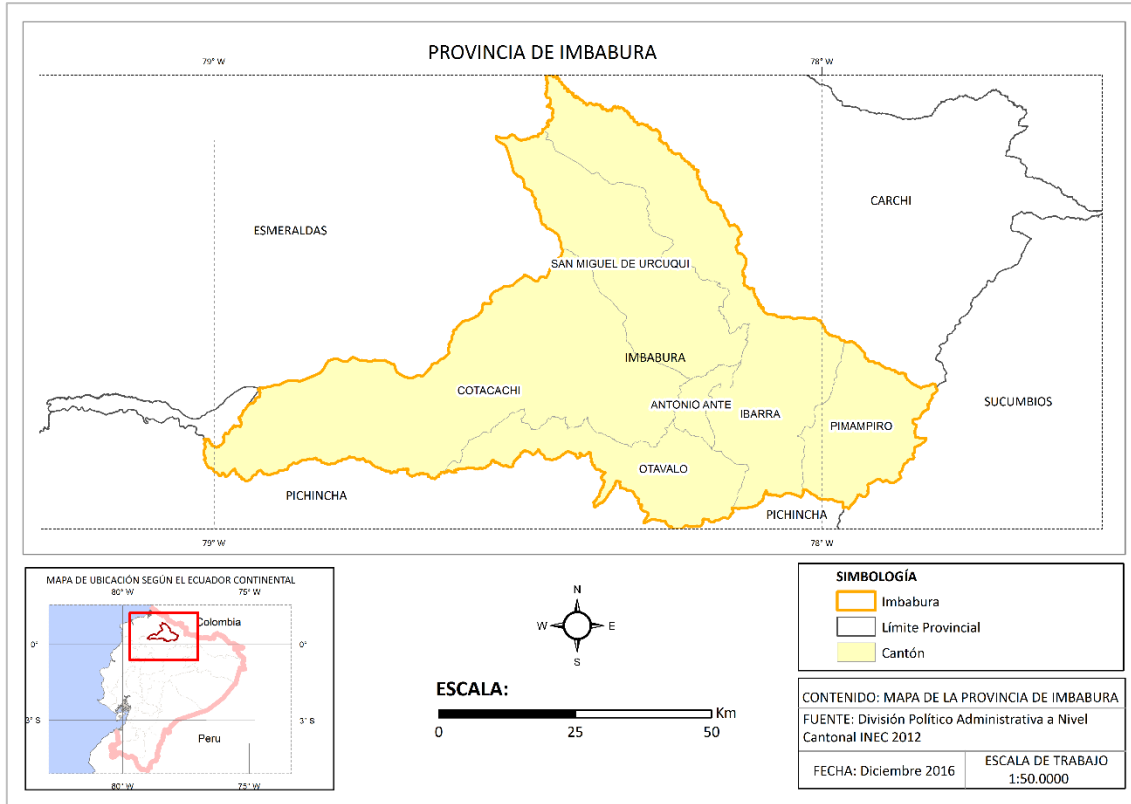


Figura 6. Mapa de la provincia de Imbabura
Fuente: INEC, 2010

En este contexto, el centro zonal ECU 1911 Ibarra como organismo público con autonomía administrativa, operativa y financiera (SIS ECU 911, 2017a), mantiene una estructura de servicio territorial para la prestación de servicios en la provincia de Imbabura que permita garantizar el derecho de los ciudadanos a la seguridad integral establecido en la Constitución de la República en los niveles administrativos de planificación indicados en la Figura 7.

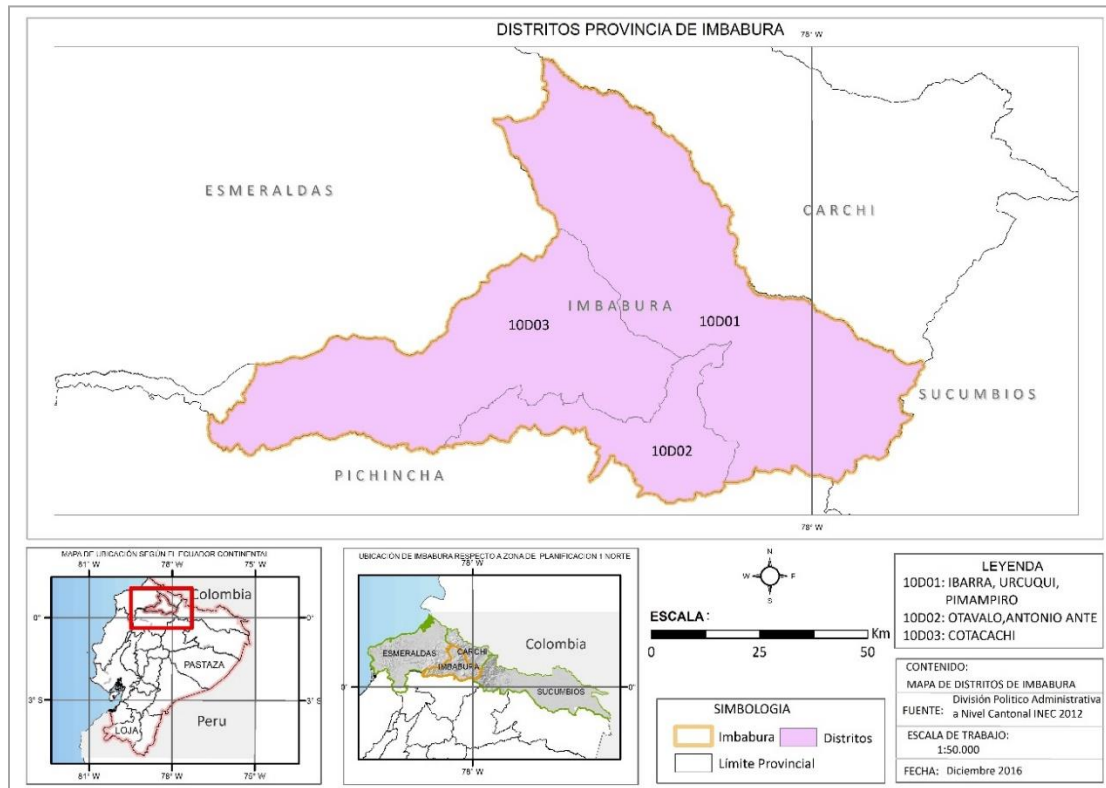


Figura 7. Niveles administrativos de planificación.
Fuente: Decreto Ejecutivo 557 (2012).

El SIS ECU 911 fue implementado mediante el Decreto Ejecutivo No. 988, publicado en el Registro Oficial No. 618 de 13 de enero de 2012, este sistema trabaja como una herramienta integradora de los servicios de emergencia que prestan las Instituciones Articuladas, tales como Cuerpo de Bomberos, Fuerzas Armadas, Policía Nacional e instituciones que conforman el Sistema Nacional de Salud y de esta manera se atiende las situaciones de emergencia reportadas mediante el número 911 (Plan Estratégico 2014-2017 ECU 911, 2014).

La Investigación se realizó en el SIS ECU 911 Ibarra, Institución que posee la información de las llamadas de emergencia generadas por los seis cantones de la Provincia de Imbabura: Ibarra, Otavalo, Cotacachi, Antonio Ante, Pimampiro y Urcuquí. El mal uso de la línea de emergencia ocasiona un grave problema, tanto en lo económico como en tiempos de

respuesta, ya que el tiempo que dedica el operador de llamadas a responder y recolectar datos que no son verídicos perjudica a la resolución de emergencias reales, debido que al hacer mal uso de la línea 911 se ven inmersos distintos tipos de recursos que se despachan innecesariamente, tales como ambulancias, Policía Nacional, Bomberos, etc.; todo dependerá del contexto que presente la llamada malintencionada.

3.2.Enfoque y tipo de investigación

El enfoque de la investigación es cuantitativo ya que se partió de los datos generados en el SIS ECU 911 Ibarra de llamadas mal intencionadas que se realizaron al número 911 en el período 2018 - 2019. Esta investigación propuso una herramienta de análisis de las llamadas mal intencionadas que permita encontrar un patrón temporal-espacial para entender la problemática y el origen de estas. El análisis se realizó a partir de una investigación cuantitativa y una perspectiva conductual.

Por el aumento de la tasa de llamadas mal intencionadas, se maneja una gran cantidad de datos sobre llamadas mal intencionadas, que sería muy difícil analizar manualmente, en este contexto usando el concepto de minería de datos, la herramienta Crime Analyst de Esri permite identificar y analizar patrones y tendencias en relación con las llamadas mal intenciones en la provincia de Imbabura. Esta técnica analítica y predictiva identifica de forma temporal y espacial los sectores y zonas que hacen mal uso del servicio de emergencia.

Con esta referencia la investigación partió de datos generados y almacenados en el SIS ECU 911 Ibarra, con las llamadas definidas como mal intencionadas que se realizaron al número único de emergencia 911 en el periodo: 2018 - 2019.

3.3.Procedimiento de investigación

La investigación fue realizada de acuerdo con las siguientes fases:

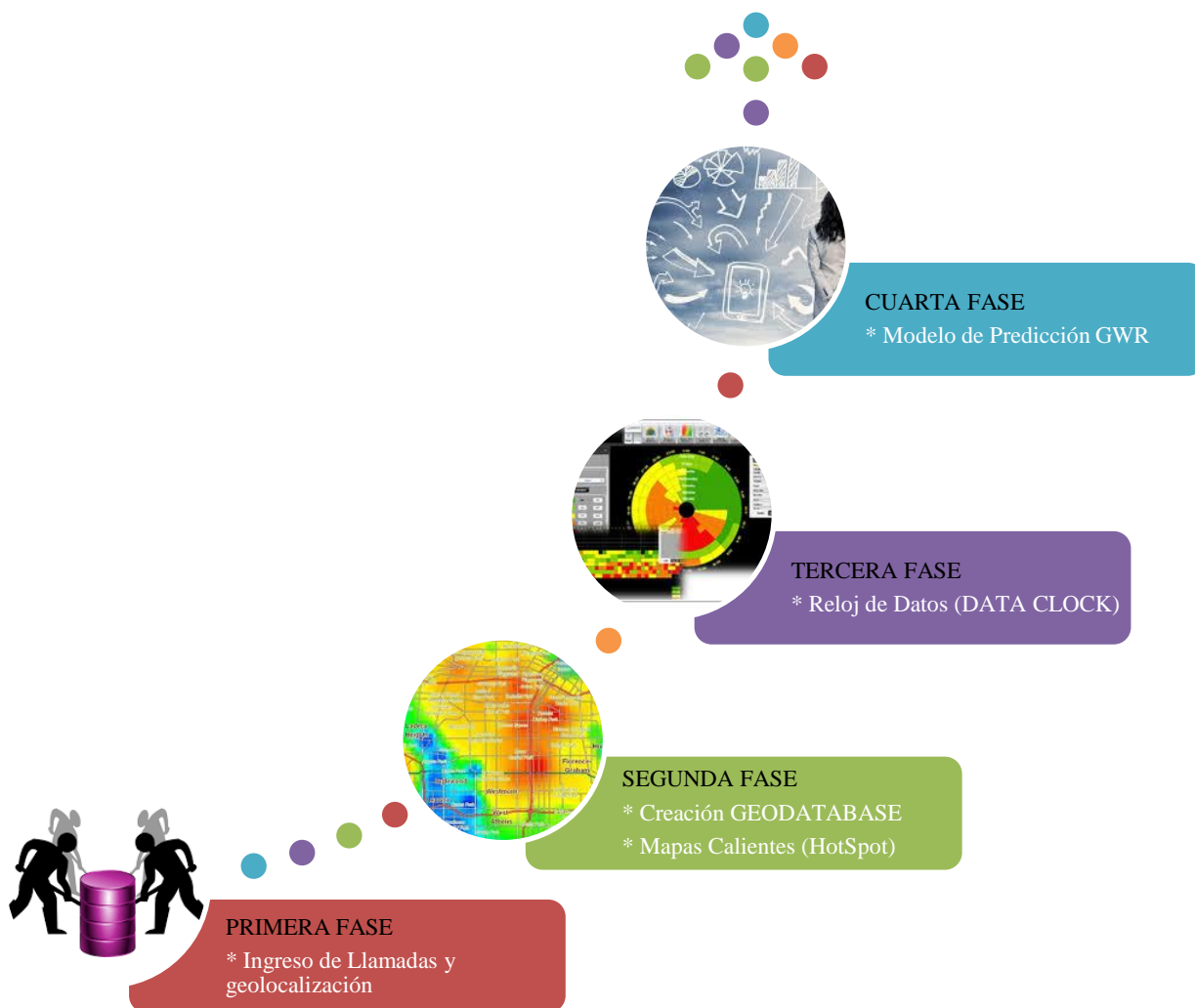


Figura 8. Fases del Procedimiento de investigación

Fuente: Gustavo Chacón

3.3.1. Primera Fase:

Se realizó la obtención de datos necesarios para la investigación, estos datos se encuentran en formato alfanumérico y reposan en las bases de datos de los servidores del SIS ECU 911 Ibarra. En esta fase se definió la forma en la cual ingresan las llamas al SIS ECU 911 Ibarra,

el almacenamiento de estos datos, y la forma como se obtiene la información del geoposicionamiento de la llamada.

Es importante tener en cuenta que el Ecuador en el año 2010, resuelve crear una norma la cual regula el procedimiento para la entrega de información relacionada con la ubicación geográfica aproximada de una llamada realizada por un usuario del SMA (Servicio de Móvil Avanzado). En el 2012 las tres operadoras que tienen cobertura de servicio en el Ecuador se comprometen a entregar la ubicación geográfica aproximada de una llamada de un número celular a la central de emergencias ECU 911 por medio de un único número el cual es 911 (Resolución 464-16-CONATEL, 2010).

En Ecuador, las comunicaciones móviles se encuentran bajo la tecnología GSM (Sistema global para las comunicaciones móviles), lo cual permite ubicar un teléfono móvil sin que se haya usado este equipo para realizar la llamada, basta con que se encuentre encendido. Para obtener la localización las operadoras realizan mediciones de la señal en varios puntos de la red, luego envían los datos a un centro de procesamiento y finalmente determinan la posición de la estación móvil, este proceso se denomina triangulación, el cual se lo puede visualizar en la Figura 9. Una vez geolocalizado el teléfono móvil se envía dicha información al SIS ECU 911 Ibarra.

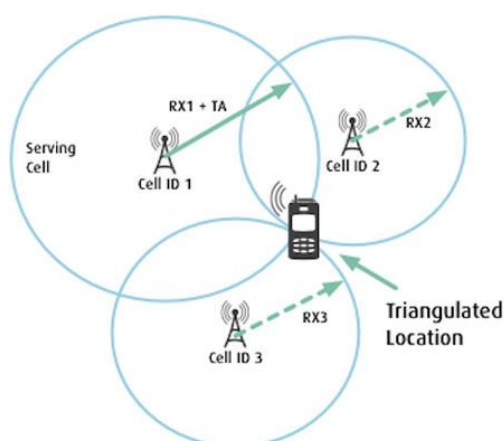


Figura 9. Triangulación

Fuente: (Informador Chile, 2020)

El proceso de triangulación se resume en tres aspectos:

- La antena más cercana capta la señal de la llamada del teléfono móvil.
- Transmite los datos de su ubicación a su proveedor de telefonía móvil.
- El proveedor entrega los datos al SIS ECU 911 Ibarra.

La precisión de la geolocalización depende de muchos factores entre ellos:

- La cantidad de radio bases celulares existentes en el área geográfica donde se encuentra el teléfono móvil. Mientras más radio bases o antenas existan, mejor será la precisión de la información de ubicación del teléfono (SIS ECU 911, 2010)..
- Se realiza el posicionamiento del teléfono móvil en base a la información del sistema de posicionamiento global (GPS) interno, dicha función se encuentra activada en la mayoría de los teléfonos móviles inteligentes. Con este método, existe mayor precisión en la información de localización (SIS ECU 911, 2010).

Una vez que la operadora envía la información de la localización del teléfono móvil, la plataforma tecnológica de atención de emergencias del SIS ECU 911 Ibarra, permite visualizar de manera automática en el sistema de información geográfica (GIS) la ubicación aproximada de la llamada realizada al número único de emergencias 911, dicha ubicación dependerá del proceso de triangulación que realicen las radio bases que se tenga de la operadora (SMA) ubicadas en el sector.

El SIS ECU 911 Ibarra, cuenta con un servidor que tiene como nombre *MobileLocator*, el mismo que mediante un aplicativo recepta todas las peticiones de geolocalización de las llamadas entrantes al SIS ECU 911 Ibarra, tal como lo establece el acuerdo con las diferentes operadoras, para esto existe un procedimiento interno que interactúa con los webservices de

las prestadoras SMA. Para un mejor entendimiento del proceso se lo puede apreciar en la Tabla 3.

Tabla 3. Descripción del proceso de ingreso de llamadas

Proceso	Descripción
Sala Operativa	Las llamadas ingresan al SIS ECU 911.
Infraestructura Tecnológica	<p>El número celular es identificado en el sistema mediante el aplicativo diseñado.</p> <p>En ese instante se consulta a la operadora mediante los <i>webservices</i> proporcionadas por dichas operadoras móviles. La búsqueda tiene un orden específico el cual es:</p> <ul style="list-style-type: none"> • CLARO • MOVISTAR • CNT
Triangulación de Operadoras (Localización)	Una vez identificado el número celular, cada operadora entrega la información por el método de triangulación y es el <i>web service</i> quien luego de hacer la consulta, devuelve la información necesaria para la ubicación geográfica como (latitud y longitud) de la base celular más cercana.
Proceso de captura	<p>La información es almacenada en una base de datos <i>Oracle</i>, y visualizada en una interfaz <i>web</i> usando el <i>api</i> ArcGis de ESRI, compatible con toda la plataforma GIS del SIS ECU911 Ibarra.</p> <p>Se almacenan 6 variables:</p> <ul style="list-style-type: none"> • OPERADOR • REFERENCIA • TELEFONO • LATITUD • LONGUITUD • AZIMUT • DIRECCION • TIPO DE INCIDENTE
Tiempo del proceso	<p>Inicio: 1 de enero de 2018.</p> <p>Fin: 30 de diciembre de 2019.</p>

Los datos de la Tabla 2 se pueden visualizar de manera gráfica en la Figura 10.

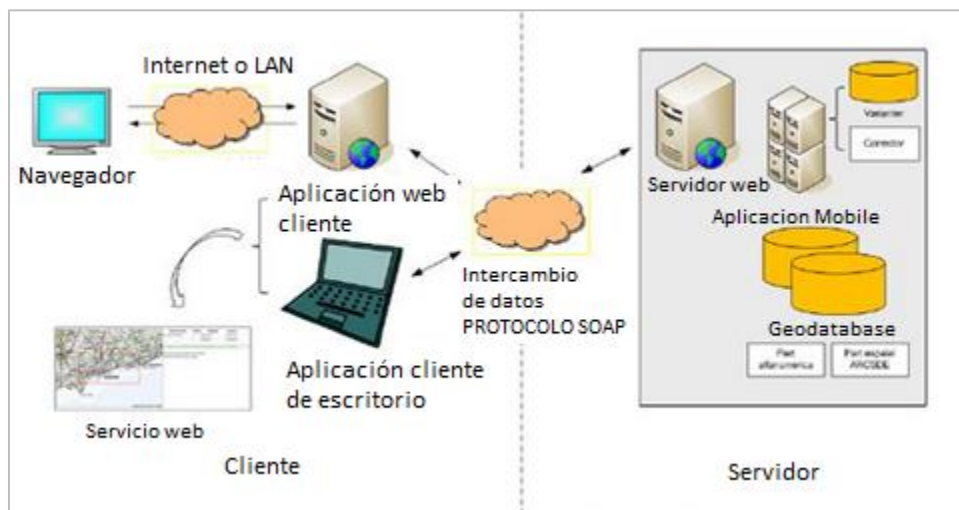


Figura 10. Proceso de geolocalización de llamadas

Fuente: (Güell & Torre, 2010)

Una vez realizado el proceso con los webservices de las operadoras móviles, se obtiene las coordenadas de localización tanto latitud como longitud de la llamada, ésta es almacenada en una base de datos alfanumérica en un servidor de base de datos, el cual usa como gestor de base de datos la plataforma Oracle.

Una vez que la información se encuentre en la base de datos se realiza scripts personalizados (consulta) a la base de datos Oracle con el fin de obtener información como: llamadas por día, por mes, así como latitud y longitud de las llamadas, dicha información se extrae en un formato de texto plano txt. o en un formato separado por comas cvs. Esta información es necesaria para la creación de la geodatabase la cual permite almacenar la información geográfica que será la base para el análisis espacial de esta investigación.

3.3.2. Segunda Fase

Esta fase se encargó de transformar los datos obtenidos en información espacial y con esto se creó una base datos espacial (GEODATABASE). Seguidamente se aplicó la técnica descriptiva de minería de datos, ya que con esta técnica no se supone variables dependientes ni independientes así como tampoco existe un modelo previo para los datos, los modelos se crean automáticamente partiendo del reconocimiento de patrones, para esta técnica se usó la herramienta Crime Analyst (Weiss, 2015) para ARCGIS con la finalidad de crear los mapas de calor (HOTSPOT) y así detectar los puntos calientes existentes con la finalidad de determinar un patrón geográfico. .

En base al proceso de extracción del conocimiento, KDD (Knowledge Discovery in Databases), se procedió con la selección de la información, es decir, el filtrado desde enero del 2018 a diciembre del 2019, en esta etapa se realiza la limpieza y filtrado de la información a fin de tener datos válidos para el objetivo de esta investigación. Con los resultados del script ejecutado en el paso anterior se obtiene un archivo en formato csv, como primer paso se elimina los campos en los cuales no está completa toda la información necesaria para este estudio tal como, latitud y longitud de la llamada, es decir, dicha llamada no fue geolocalizada en el sistema del SIS ECU911 Ibarra. Una vez obtenido el archivo en formato cvs final se usó herramientas de ARGIS de ESRI (Environmental Systems Research Institute), con la finalidad de crear la geodatabase. La Figura 11 muestra las partes de un SIG (Sistema de Información Geográfica).

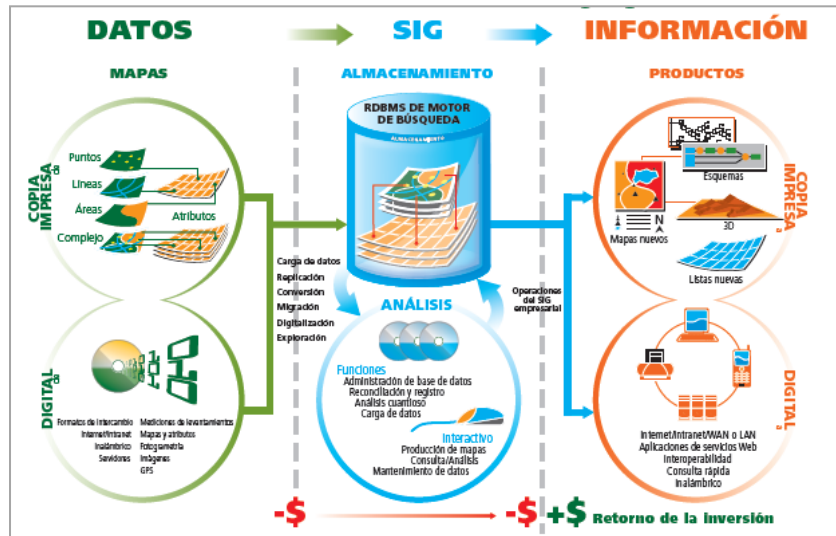


Figura 11. Sistema de Información Geográfica y sus partes
Fuente: (Lozada, 2017)

El proceso para realizar los HotSpot, inicia realizando un cálculo estadístico G_i^* de Getis-Ord para cada entidad en un dataset ¹. Las puntuaciones z las cuales son desviaciones estándar y los valores P los cuales son una probabilidad, cuando el valor P es muy pequeño significa que es muy poco probable, es así que se van obteniendo resultados que indican dónde se agrupan espacialmente las entidades con valores altos o bajos, de esta manera se forman las agrupaciones y se empieza a crear asociaciones y relaciones. Esta herramienta funciona por medio de la búsqueda de cada entidad dentro del contexto de entidades adyacentes.

Cabe indicar también que una entidad con un valor alto es importante, pero es posible que no sea un punto caliente estadísticamente significativo. Para ser un punto caliente estadísticamente significativo, una entidad debe tener un valor alto y también estar rodeada por otras entidades con valores altos. La suma local para una entidad y sus vecinos se compara proporcionalmente con la suma de todas las entidades; cuando la suma local es muy

¹ Un dataset de entidad es una colección de clases de entidad relacionadas que comparten un sistema de coordenadas común. Los datasets de entidades se utilizan para integrar espacial o temáticamente clases de entidad relacionadas, cuyo propósito primario es organizar clases de entidad relacionadas en un dataset común para generar una topología.

diferente de la esperada, y esa diferencia es demasiado grande como para ser el resultado de una opción aleatoria, se obtiene como consecuencia una puntuación z estadísticamente significativa. En la Tabla 4 se muestran los valores p y las puntuaciones z críticas sin corregir para distintos niveles de confianza.

Tabla 4. Información para medir grado de confianza de los resultados

puntuación z (Desviaciones estándar)	valor P (Probabilidad)	Nivel de confianza
< -1,65 o > +1,65	< 0,10	90%
< -1,96 o > +1,96	< 0,05	95%
< -2,58 o > +2,58	< 0,01	99%

Fuente: (ARCGIS, 2017)

La estadística G_i^* devuelta para cada entidad en el dataset es una puntuación z. Para las puntuaciones z positivas que son estadísticamente significativas, mientras más grande es la puntuación z, más intenso es el clustering de valores altos (punto caliente). Para las puntuaciones z negativas que son estadísticamente significativas, mientras más pequeña es la puntuación z, más intenso es el clustering de valores bajos (punto frío) (ARCGIS, 2017). Matemáticamente la estadística G_i^* se representa en la fórmula de la Figura 12:

$$G_i^* = \frac{\sum_{j=1}^n w_{i,j} x_j - \bar{X} \sum_{j=1}^n w_{i,j}}{S \sqrt{\frac{n \sum_{j=1}^n w_{i,j}^2 - \left(\sum_{j=1}^n w_{i,j} \right)^2}{n-1}}}$$

*Figura 12. Fórmula que representa la estadística G_i^**

Fuente: (ARCGIS, 2017)

3.3.3. Tercera fase

Esta fase permitió la creación de los relojes de datos los cuales muestran una distribución temporal la cual que se puede utilizar para analizar patrones de datos que pueden no ser evidentes al ver los datos, de esta manera se pudo interpretar de una mejor forma los datos

minados. Se usó la herramienta Crime Analyst (Weiss, 2015) para ARCGIS. Esta herramienta es muy usada para las entidades que se dedican al análisis de patrones de delincuencia, gestionando operaciones especiales o patrullando, los mapas inteligentes pueden ayudar a las entidades de Seguridad a mantener a su comunidad segura y protegida. Los mapas del sistema de información geográfica (SIG) hacen que la información sea procesable, y se pueda visualizar, analizar y compartir datos con mapas interactivos, además, con esta información se puede realizar capacitaciones al personal involucrado sobre ubicación en tiempo real y geolocalización de tal manera que contribuya a una mejor toma de decisiones.

Cada decisión que se tome debe comenzar con el análisis, con la finalidad de reducir riesgos, mejorar las operaciones; y tomar decisiones más inteligentes basadas en datos reales. Las personas que se encargan de estos análisis aprovechan el software moderno de mapeo de delitos y las herramientas de análisis espacial para entregar la información que la entidad de seguridad necesita para tomar decisiones informadas.

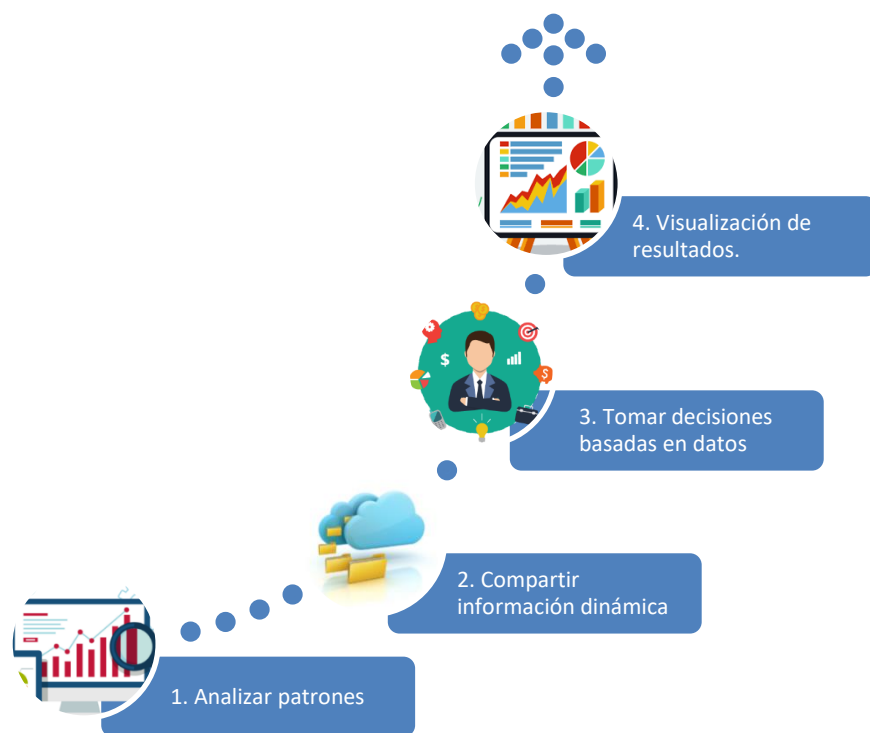


Figura 13. Esquema metodológico del Software Crime Analyst

Fuente: Gustavo Chacón

La Figura 13, muestra de manera general el comportamiento del software Crime Analyst, en cuanto a los tres procesos esenciales que realiza. Cada proceso se describe a continuación:

- ✓ **Analizar patrones:** ArcGIS permite realizar muchos trabajos de análisis de problemas más comunes referentes al tema de seguridad (delitos, uso indebido de la línea 911), analiza los datos sobre ellos, identifica de qué tipo son los problemas y dónde se encuentran los puntos críticos, esto con la finalidad de realizar pronósticos para ayudar a las entidades de seguridad a planificar iniciativas de prevención del problema presentado. El software de mapeo de delitos ESRI admite análisis táctico, estratégico e investigativo y facilita compartir y colaborar con mapas web y aplicaciones (ARCGIS, 2020).

- ✓ **Compartir información dinámica:** se encarga de traducir el análisis en productos de información operacionalmente útiles. Crea y comparte aplicaciones dinámicas para el análisis de autoservicio y la conciencia situacional en el campo, así como también genera informes convincentes de los problemas presentados, pero lo muestra con mapas interactivos.

- ✓ **Tomar decisiones basadas en datos:** con la información obtenida se puede proporcionar las herramientas adecuadas para todos los miembros de la entidad de seguridad. De acuerdo con el ámbito de su competencia cada uno de ellos podrán aprovechar los mapas y las aplicaciones que permiten la toma de decisiones basada en datos.

En el caso de estudio esta herramienta toma en cuenta las 24 horas por día y 7 días a la semana, para la creación de los relojes de datos, estos se dividen en celdas mediante círculos concéntricos donde se visualiza los días de la semana y en líneas radiales que son las que representan las horas del día, cada color de la celda indica el número de eventos ocurridos en ese tiempo. Los relojes de datos permitieron profundizar con más detalle las llamadas mal

intencionadas ocurridas en el año 2018 y 2019, por cantones y a su vez internamente dentro de cada cantón se determinó cómo fue el comportamiento por parroquia. De igual manera se generó un reloj de datos rectangular en donde se tomó en cuenta los días y los meses, en este caso el uso de los colores definirá y resaltarán los valores más altos de llamadas mal intencionadas realizadas durante los meses y días de los años objeto de este estudio.

3.3.4. Cuarta fase

En esta fase se usó la herramienta de software GWR para obtener las predicciones futuras lo que permitió determinar las posibles estrategias para la mitigación del problema y la toma de decisiones. El modelo de regresión geográficamente ponderada GWR por sus siglas en inglés se representa una variable dependiente (tasa de llamadas MAL INTENCIONADAS) y la población como variable independiente. Este modelo delimita la zona geográfica de las llamadas mal intencionadas, es un tipo de estadística local que puede producir un conjunto de estimaciones de parámetros locales que muestran la forma cómo varían las relaciones en el espacio para una mejor comprensión de las posibles causas ocultas de este patrón como una estimación de parámetros locales (Fotheringham, Brunson, & Charlton, 2002).

Este proceso fue realizado automáticamente por la herramienta GWR del software ArcGIS 10.2.2. La interpretación de estos valores está basada en el coeficiente R^2 local, mismo que presentó ciertos valores, para esta investigación se usó los rangos de credibilidad descritos en la Tabla 5 para el análisis de regresión ponderada geográficamente.

Tabla 5. Criterios de aceptabilidad del coeficiente R^2 para el análisis de regresión

Rango R^2	Aceptabilidad
0 – 0.25	Muy baja
0.25 – 0.50	Baja
0.50 – 0.75	Alta
0.75 – 0.96	Muy alta

Fuente: (Nash & Sutcliffe, 1970)

Para este proceso se aplicó la técnica de Regresión ponderada geográficamente (GWR). Con el fin de validar el comportamiento se tomó una variable de tiempo de 20 años, estos datos se obtuvieron de las Proyecciones Referenciales de la Población a nivel cantonal 2010-2030, (SNI, Sistema Nacional de Información) donde se encuentra el tamaño de la población a 20 años.

CAPÍTULO IV: RESULTADOS Y DISCUSIÓN

El periodo de tiempo que se usó para esta investigación fue desde 01 de enero del año 2018 al 31 de diciembre del año 2019, si bien es cierto el SIS ECU 911 Ibarra inicia sus actividades en el año 2014, a partir del año 2018 se implementa el proceso de geolocalización de las llamadas mal intencionadas, ya que en un inicio únicamente se almacenaba la georreferenciación de las llamadas normales. En la Tabla 6 se muestra un porcentaje de las llamadas que ingresaron al número único de emergencias 911 en la provincia de Imbabura desde los teléfonos móviles.

Tabla 6. Resumen de llamadas mal intencionadas 2018-2019

Tipo de llamada	Número	Porcentaje (%)
Llamadas mal intencionadas registradas	34036	100%
Registros no válidos	3728	10.95%
Registros Válidos	30308	89.05%

Fuente: Gustavo Chacón

Los datos obtenidos se encuentran en la Tabla 7, con estos datos se puede realizar un análisis cuantitativo obteniéndose que en la provincia de Imbabura de las 30308 llamadas mal intencionadas válidas se las divide por cantón y se tiene en cuenta el número de llamadas, así como la población de cada cantón, este dato es tomado según el SNI (2015).

Tabla 7. Porcentaje de llamadas por cantón

CANTÓN	Nro. Llamadas	Porcentaje de llamadas
IBARRA	12932	43%
ANTONIO ANTE	3158	10%
COTACACHI	2707	9%
OTAVALO	9172	30%
PIMAMPIRO	1321	4%
SAN MIGUEL DE		
URCUQUI	1018	3%
TOTAL	30308	100%

Fuente: Gustavo Chacón

En la Tabla 8 se puede observar que el cantón Urcuquí según el número de llamadas y su población es el que mayor porcentaje de llamadas mal intencionadas genera. Para realizar el cálculo del porcentaje se usó la siguiente fórmula:

$$\% \text{ tasa de llamadas} = \frac{\text{nro llamadas}}{\text{población}} \times 100$$

Tabla 8. Tasa de llamada por cantón 2018-2019

CANTON	Nro. de Llamadas Población		Tasa de llamadas %
IBARRA	12932	139721	9.26
ANTONIO ANTE	3158	23299	13.55
COTACACHI	2707	17139	15.79
OTAVALO	9172	52753	17.39
PIMAMPIRO	1321	9077	14.55
SAN MIGUEL DE			
URCUQUI	1018	5205	19.56
TOTAL	30308	247194	100%

Fuente: Gustavo Chacón

Mientras tanto, a nivel de parroquias, la Tabla 9 muestra el porcentaje de las llamadas mal intencionadas, de donde se puede analizar que en Ibarra la parroquia La Esperanza es la que tiene la tasa más alta de llamadas mal intencionadas, ya que al realizar el cálculo del porcentaje se tomó en cuenta el número de llamadas y el número de habitantes de cada parroquia y en este caso el resultado fue un 7.93% para esta parroquia, seguidamente con un margen un poco menor se encuentran los cantones, Antonio Ante la parroquia de Natabuela con 17.25%, Cotacachi la parroquia de Cotacachi con un porcentaje del 12.21%, Otavalo la parroquia Selva Alegre presenta un porcentaje alto de 24.81%, en Pimampiro la parroquia de Pimampiro el 12.60% y finalmente en Urcuquí la parroquia de Pablo Arenas el 17.19%.

Tabla 9. Tasa de llamadas mal intencionadas por parroquia

CANTON	PARROQUIA	NRO. LLAMADAS	POBLACION	TASA DE LLAMADAS
IBARRA	SAN MIGUEL DE IBARRA	10981	139721	7.86
	AMBUQUI	286	5477	5.22
	ANGOCHAGUA	2	3263	0.06
	LA ESPERANZA	584	7363	7.93
	LITA	2	3349	0.06
	SALINAS	9	1741	0.52
	SAN ANTONIO	1068	17522	6.10
ANTONIO	ATUNTAQUI	1840	23299	7.90
ANTE	IMBAYA (SAN LUIS DE COBUENDO)	7	1279	0.55
	SAN JOSE DE CHALTURA	80	3147	2.54
	SAN ROQUE	256	10142	2.52
	SAN FCO. DE NATABUELA	975	5651	17.25
	COTACACHI	2092	17139	12.21
COTACACHI	GARCIA MORENO (LLURIMAGUA)	75	5060	1.48
	IMANTAG	528	4941	10.69
	QUIROGA	12	6454	0.19
	OTAVALO	4489	52753	8.51

	DOCTOR MIGUEL EGAS	480	4883	9.83
	CABEZAS (PEGUCHE)			
	EUGENIO ESPEJO (CALPAQUI)	52	7357	0.71
	GONZALEZ SUAREZ	1325	5630	23.53
	SAN JOSE DE QUICHINCHE	672	8476	7.93
	SAN JUAN DE ILUMAN	625	8584	7.28
	SAN PABLO	1121	9901	11.32
	SAN RAFAEL	11	5421	0.20
	SELVA ALEGRE (CAB. EN SAN MIGUEL DE PAMPLONA)	397	1600	24.81
PIMAMPIRO	PIMAMPIRO	1144	9077	12.60
	MARIANO ACOSTA	177	1544	11.46
URCUQUI	URCUQUI	654	5205	12.56
	PABLO ARENAS	364	2118	17.19

Fuente: Gustavo Chacón

La Figura 14 muestra de forma gráfica los valores determinados en la Tabla 9.

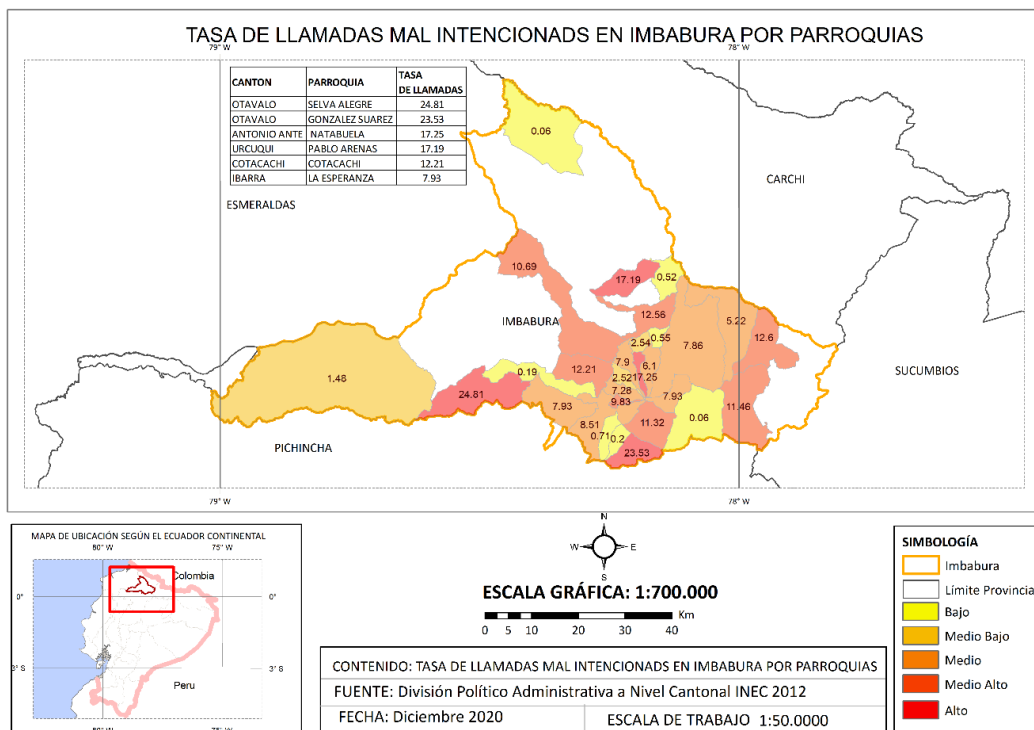


Figura 14. Tasa de llamadas mal intencionadas por parroquia 2018-2019

Fuente: Gustavo Chacón

A continuación la Tabla 10 contiene información procesada en base a la siguiente fórmula:

$$\% \text{ de llamadas por cantón} = \frac{\text{Nro de Llamadas} \times 100\%}{\text{Total}}$$

Esta tabla contiene el porcentaje de llamadas mal intencionadas por cantones en Imbabura, resaltando al cantón Ibarra con mayor porcentaje de llamadas.

Tabla 10. Porcentaje de llamadas por cantón

CANTÓN	Nro Llamadas	Porcentaje de llamadas
IBARRA	12932	43%
ANTONIO ANTE	3158	10%
COTACACHI	2707	9%
OTAVALO	9172	30%
PIMAMPIRO	1321	4%
SAN MIGUEL DE URCUQUI	1018	3%
TOTAL	30308	100%

Fuente: Gustavo Chacón

Estos datos se les puede observar también de forma gráfica en la Figura 15 la misma que muestra el mapa de las llamadas mal intencionadas por cantones en Imbabura, resaltando al cantón Ibarra con mayor porcentaje de llamadas, los datos mostrados en la figura son resultado de la Tabla 10 en donde se describe el número de llamadas y su porcentaje, el cual fue calculado en base a la siguiente fórmula:

$$\% \text{ de llamadas por cantón} = \frac{\text{Nro de Llamadas} \times 100\%}{\text{Total}}$$

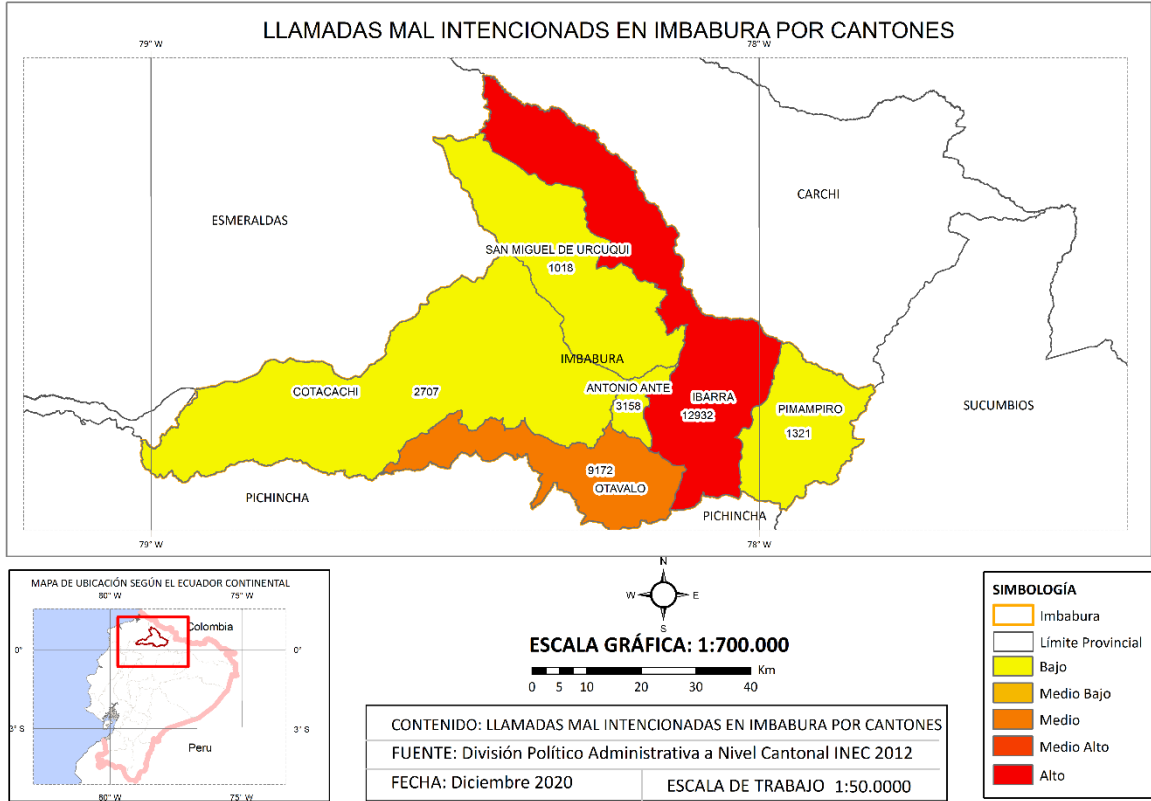


Figura 15. Llamadas mal intencionadas por cantón 2018-2019

Fuente: Gustavo Chacón

4.1. Análisis Espacial (hot spot) de los años 2018 y 2019

Una vez procesada la información en referencia a las llamadas mal intencionadas en la provincia de Imbabura, la determinación de los HotSpots o puntos calientes es importante para focalizar en el territorio la concentración de estas llamadas. A continuación, el mapa que indica los puntos calientes de las llamadas mal intencionadas por cantones se puede observar en la Figura 16.

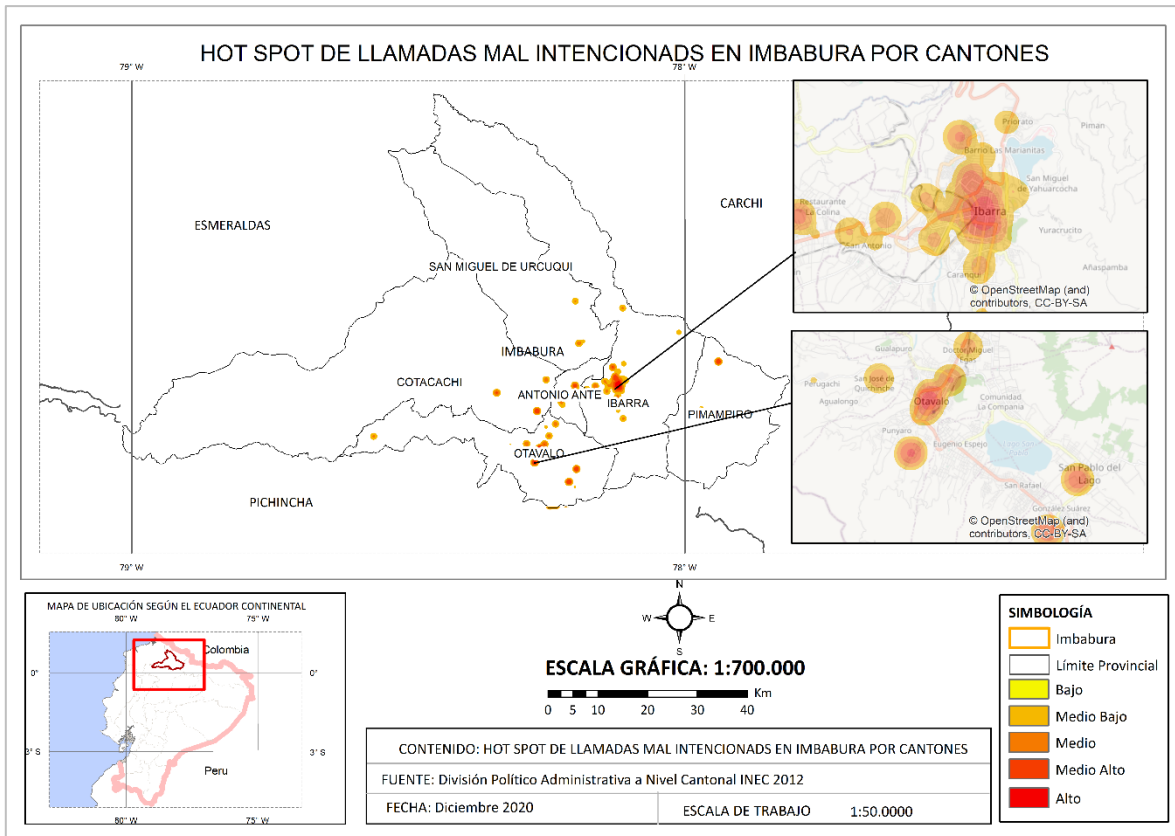


Figura 16. Hot Spot de llamadas mal intencionadas por cantones de Imbabura

Fuente: Gustavo Chacón

Las zonas en rojo de la figura anterior son agrupaciones estadísticamente importantes con valores altos (puntos calientes - hotspots) de llamadas “mal intencionadas”. En este caso el mayor número de llamadas definidas, provienen del cantón Ibarra y Otavalo, por lo que se puede apreciar mayor concentración en las categorías: Medio-Alto, Alto. Mientras que en cantones como: Atuntaqui, Cotacachi, Urcuquí y Pimampiro están caracterizados como: Medio Bajo y Medio, las llamadas de este tipo no contienen valores significativos.

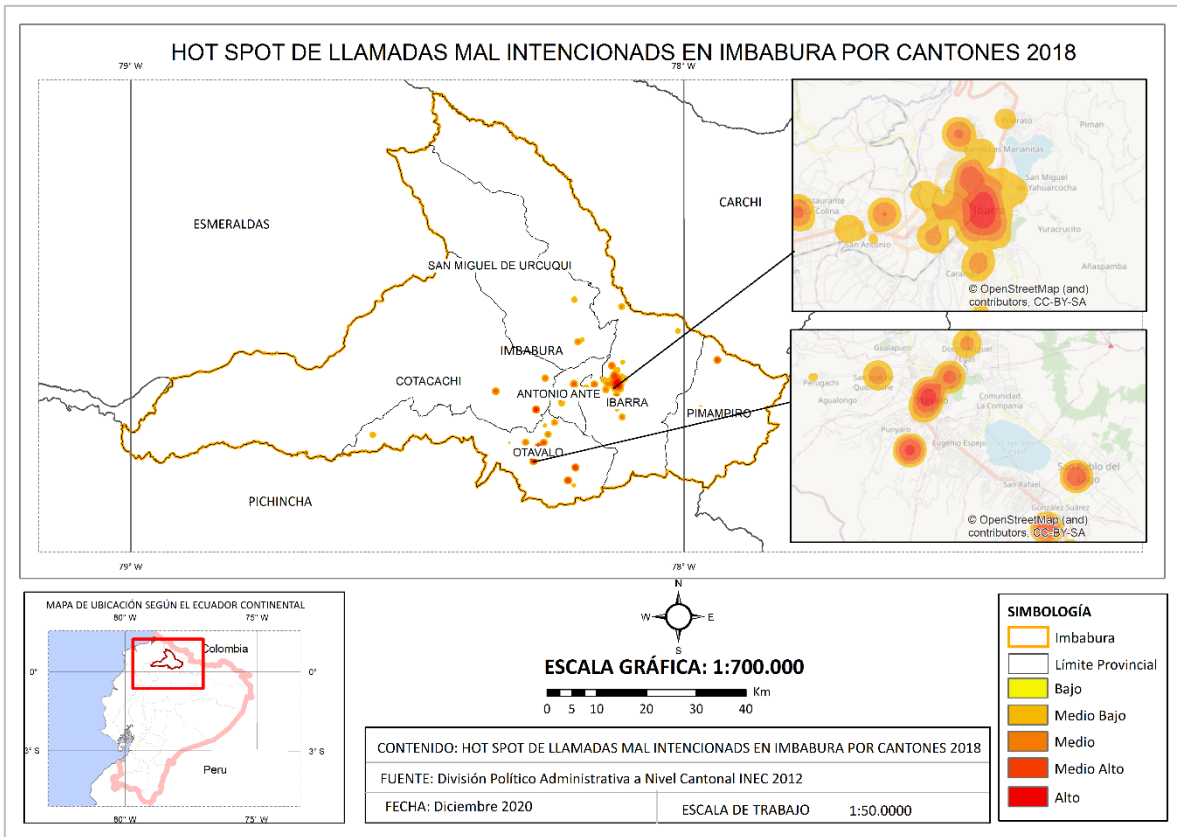


Figura 17. Mapa Caliente (hot spot) de llamadas mal intencionadas divididas por cantones año 2018

Fuente: Gustavo Chacón

En el 2018 las zonas en rojo de la Figura 17, muestra cuales son agrupaciones estadísticamente importantes con valores altos de llamadas “mal intencionada” (puntos calientes). En este caso el mayor número de llamadas definidas, provienen del cantón Ibarra y Otavalo, por lo que se puede apreciar mayor concentración en las categorías: Medio-Alto, Alto. Mientras que en cantones como: Atuntaqui, Cotacachi, Urcuquí y Pimampiro están caracterizados como: Medio Bajo y Medio, concluyendo que las llamadas de este tipo no tienen valores significativos.

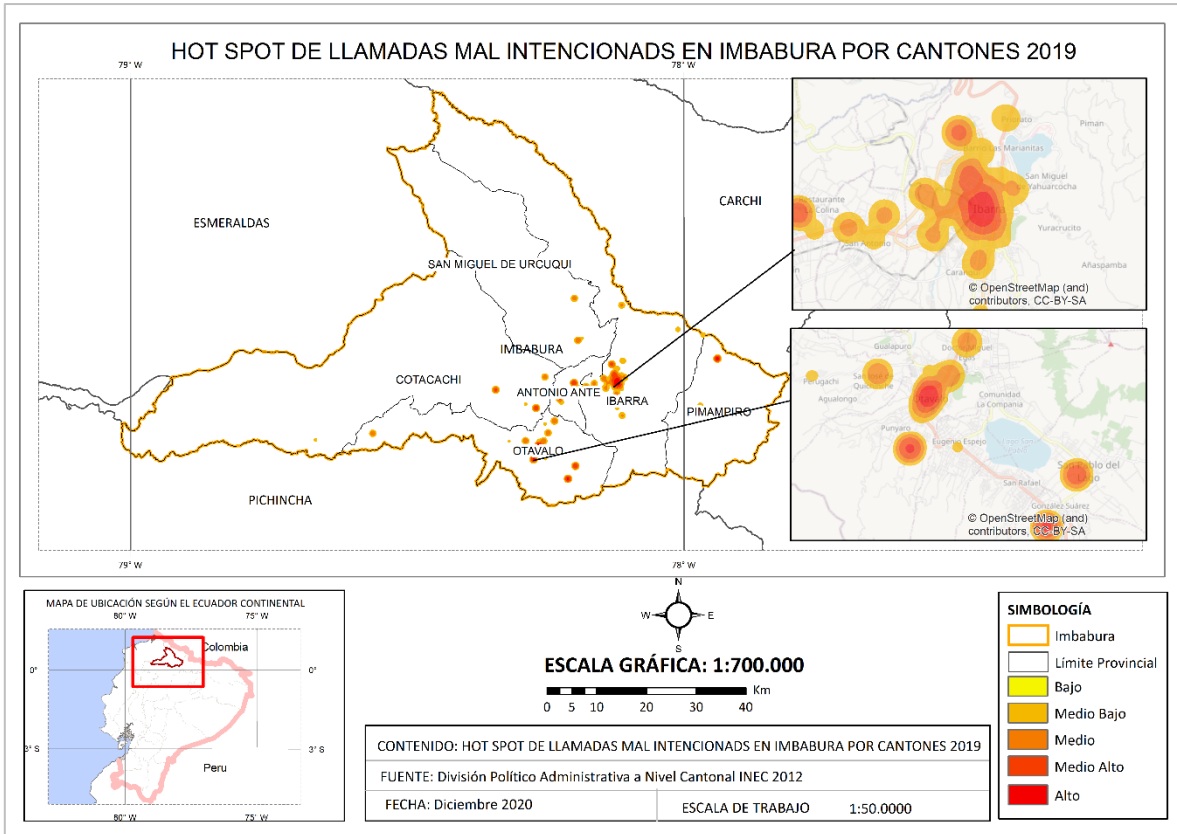


Figura 18. Hot Spot llamadas mal intencionadas divididas por cantones año 2019

Fuente: Gustavo Chacón

En el 2019 las zonas en rojo mostradas en la Figura 18, indican cuales son agrupaciones estadísticamente importantes con valores altos de llamadas “mal intencionada” (puntos calientes). En este caso el mayor número de llamadas definidas, se encuentran en los cantones Ibarra y Otavalo, notándose claramente que existe una mayor concentración en las categorías: Medio-Alto, Alto, cabe indicar además que en cantones como: Atuntaqui, Cotacachi, Urucuquí y Pimampiro las categorías son: Medio Bajo y Medio, mostrando que las llamadas de este tipo no poseen valores significativos como para considerarlos dentro del grupo problema.

4.2. Análisis Temporal de llamadas mal intencionadas en los años 2018 y 2019

La Figura 19 muestra un reloj de datos el cual se divide en celdas mediante círculos concéntricos que representan los días de la semana, mientras que las líneas radiales representan las horas del día, cada color de la celda indica el número de eventos ocurridos en ese tiempo, este reloj de datos permite profundizar con más detalle las llamadas mal intencionadas ocurridas en el año 2018, en la gráfica se puede distinguir claramente el área de color rojo que en este caso representa al rango de horas de entre las 22:00 y las 6:00 del siguiente día, esto ya va generando una idea específica de los horarios que se debería tomar en cuenta para realizar alguna acción preventiva y mitigar el problema.

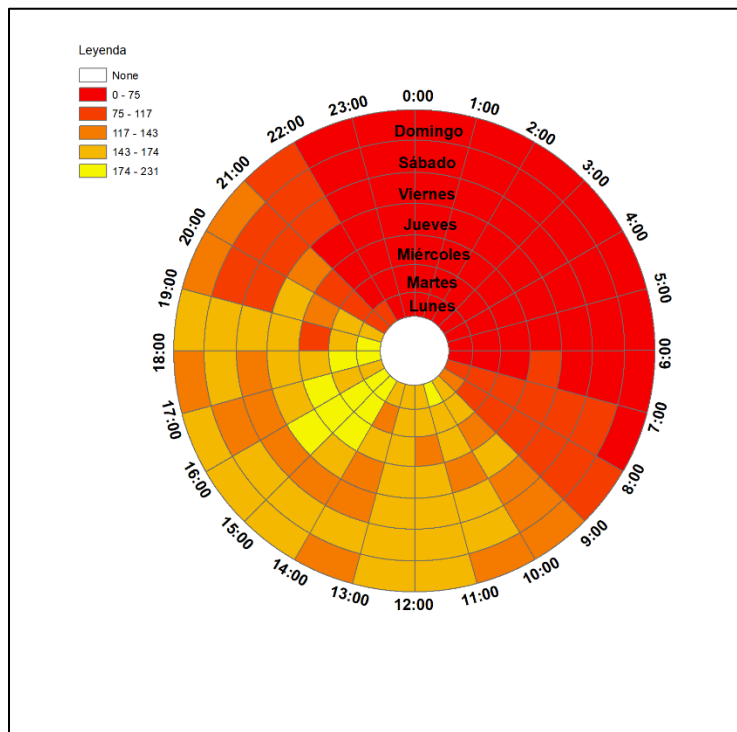


Figura 19. Llamadas mal intencionadas 2018

Fuente: Gustavo Chacón

El reloj de datos rectangular de la Figura 20 muestra como variables los días y los meses, de igual manera resalta el color rojo correspondiente a los valores más altos de llamadas mal intencionadas realizadas durante esos meses y días, durante este año los meses que sobresalen son junio y julio, así se cruza la información con los días es notorio que los días 2, 6, 13, 20, 26, 30 presentan el indicio de alerta y se lo puede asociar a alguna situación en particular que haya sucedido esos días.

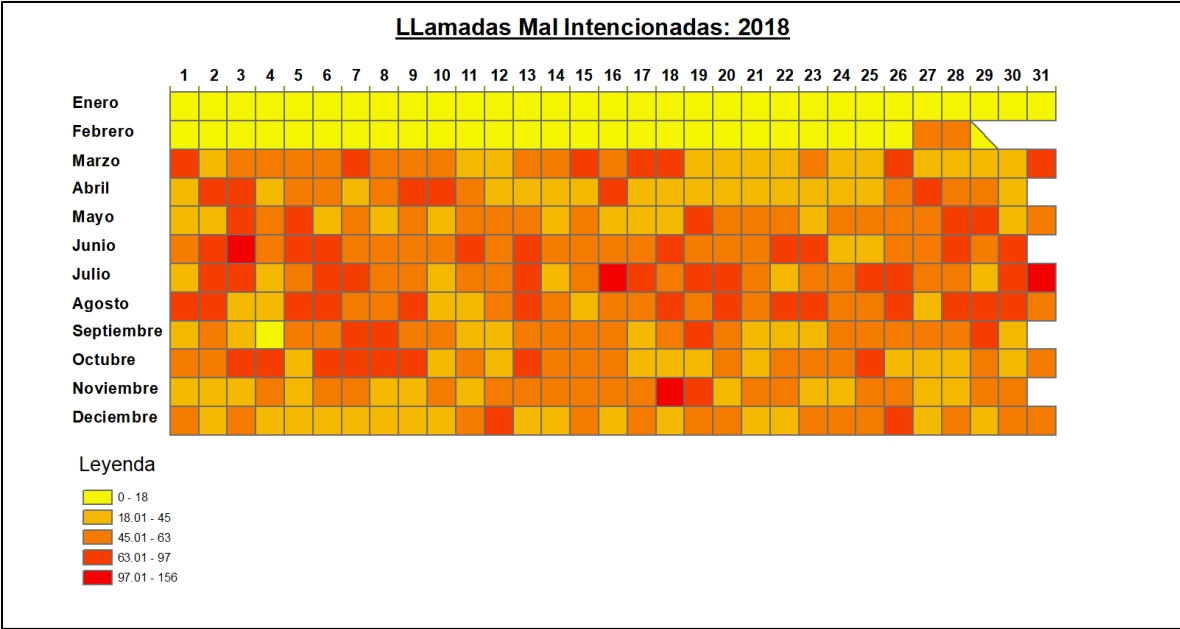


Figura 20. Reloj rectangular de llamadas mal intencionadas 2018

Fuente: Gustavo Chacón

En la Figura 21 se observa un reloj de datos el cual se divide en celdas mediante círculos concéntricos en el que se visualizan los días de la semana y las líneas radiales corresponden a las horas del día, cada color de la celda indica el número de eventos ocurridos en ese tiempo, este reloj de datos permite profundizar con más detalle las llamadas mal intencionadas ocurridas en el año 2019. Se puede notar claramente que los eventos en este año corresponden al mismo horario que se determinó en el año 2018, estos se encuentran en color rojo y es el horario de 22:00 a 06:00.

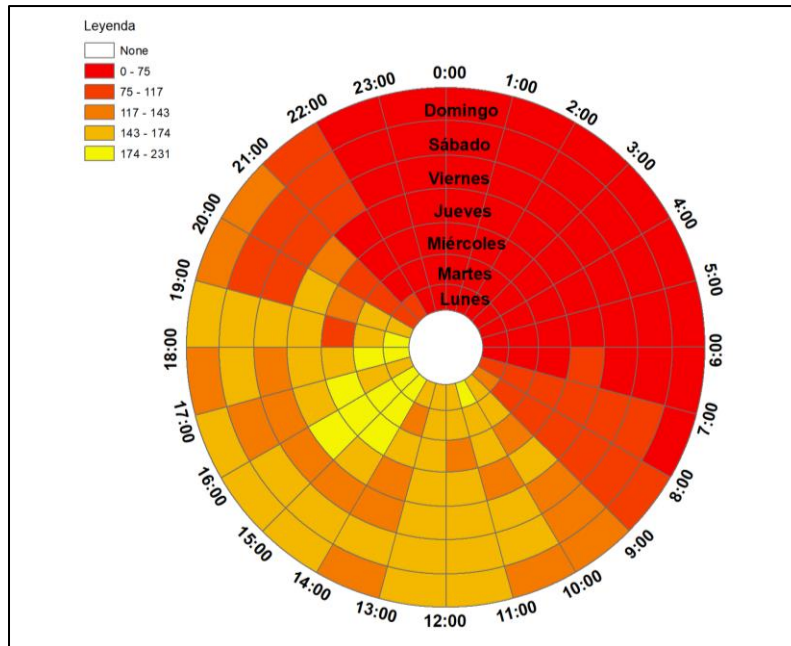


Figura 21. Reloj de Datos de llamadas mal intencionadas 2019

Fuente: Gustavo Chacón

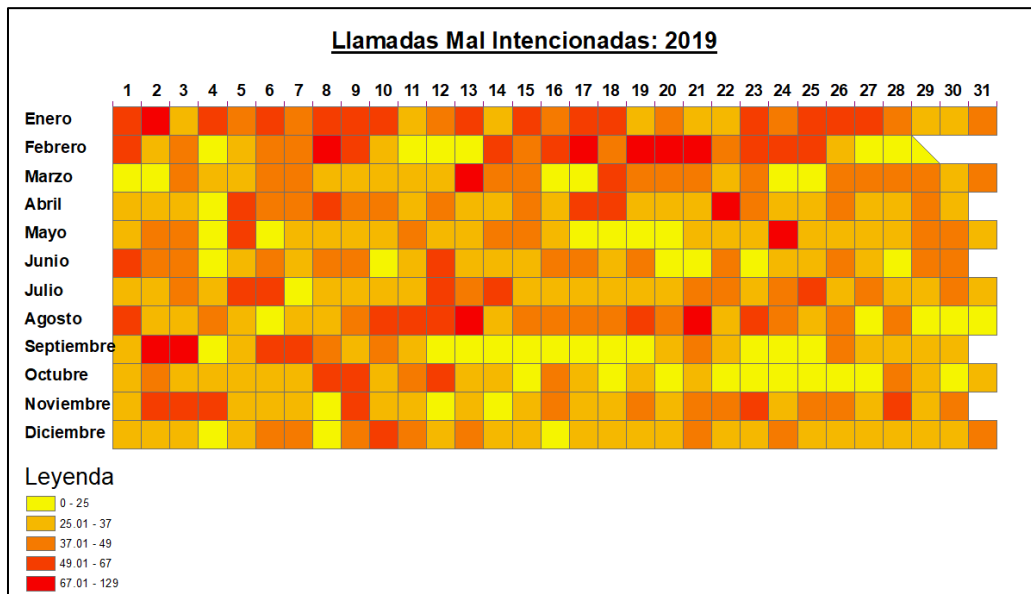


Figura 22. Reloj de datos rectangular de llamadas mal intencionadas 2019

Fuente: Gustavo Chacón

El reloj de datos rectangular muestra los días y los meses, el color rojo resalta los valores más altos de llamadas mal intencionadas realizadas durante esos meses y días, tal como se aprecia en la Figura 22, en ella sobresalen los meses de febrero, agosto y septiembre con mayor cantidad de eventos en rojo, y de los tres meses quien más sobresale es febrero, de la misma manera se puede observar los días que tienen mayor concentración de rojo en cada mes. En febrero se nota claramente los días 8, 17, 19, 20, 21, agosto muestra los días 13 y 21 y finalmente septiembre los días 2 y 3.

4.3.Comparativa de llamadas mal intencionadas años 2018 y 2019

Se detectó patrones temporales en determinados periodos relacionados con el cruce de información entre los años 2018 y 2019, así como, desde el lunes al viernes los horarios en los cuales se realizan el mayor número de llamadas mal intencionadas son desde la 13:00 hasta las 21:00. De la misma manera se encuentra un patrón en los meses los cuales dan como resultado los siguientes datos; los meses de julio y agosto son los más críticos en lo referente a llamadas mal intencionadas, de igual manera se detectan patrones críticos en los días lunes, martes, viernes y sábado de los meses de julio y agosto.

En las Figuras 23 y 24 se muestran unas comparativas entre los años 2018 y 2019 en el reloj circular de datos que muestra los días y las horas y el reloj rectangular el cual muestra los días y los meses. Se realiza el análisis y se detecta algunos patrones los cuales son:

- De lunes a viernes los horarios en los cuales se realizan el mayor número de llamadas mal intencionadas entre los años 2018 y 2019 son desde las 13:00 hasta las 21:00.

- Realizando el mismo análisis por meses da como resultado los siguientes patrones, los meses de julio y agosto son los meses más críticos en lo referente a llamadas mal intencionadas, así también se detectan patrones críticos en los días 2, 5, 6, 12, 13, 30, de estos meses.

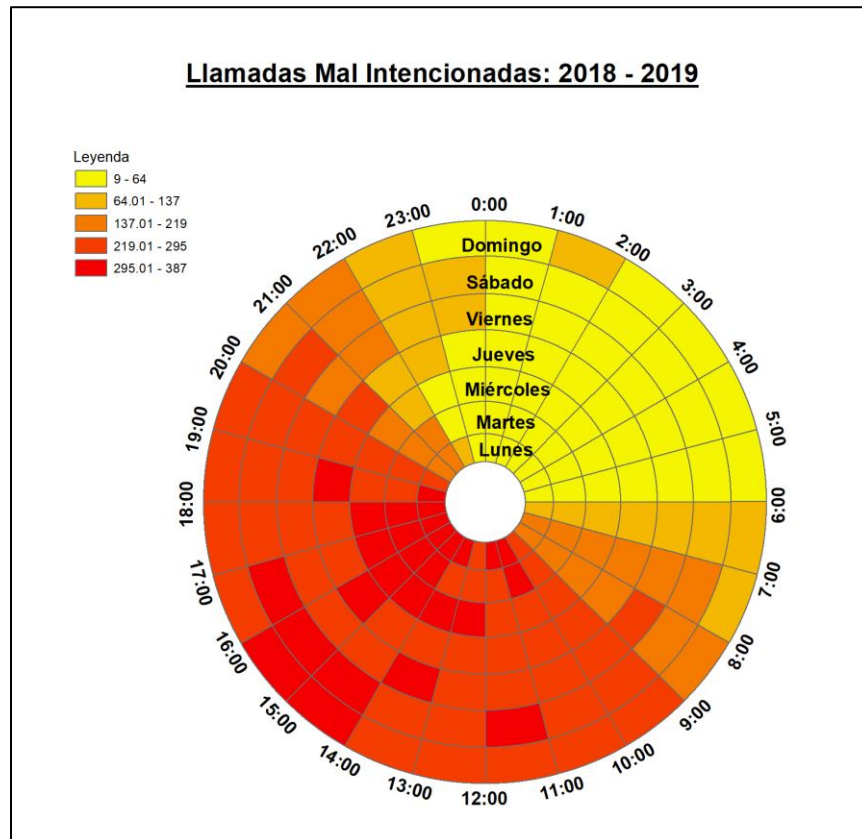


Figura 23. Comparativa de llamadas mal intencionadas 2018-2019.

Fuente: Gustavo Chacón

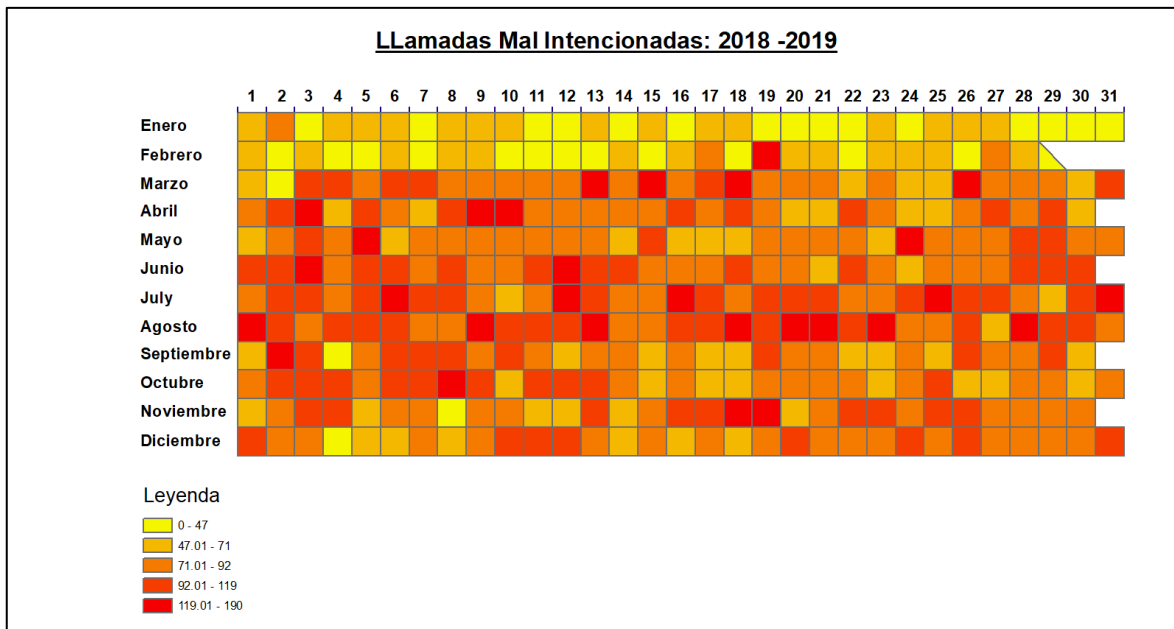


Figura 24. Comparativa de llamadas mal intencionadas en reloj de datos años 2018-2019

Fuente: Gustavo Chacón

4.4.Resultado Modelo de regresión GWR

En la Figura 25 se aprecia los resultados del modelo GWR, el R^2 local para el modelo local de GWR osciló entre 0.43248 y 0.43404. El modelo de GWR explicó hasta el 43 % de la varianza de las llamadas Mal Intencionadas relacionado con la población de la provincia de Imbabura. Parroquias como Pimampiro, Ambuquí, Ibarra, Salinas, Pablo Arenas y Lita sobresalen en este modelo. Sin embargo, en las parroquias de cantones de Cotacachi y Otavalo, el modelo no explicó gran parte de esta variabilidad a pesar de que la variabilidad del coeficiente R^2 local es mínima. Es decir, la dependencia espacial no es muy alta, ya que los coeficientes son mínimos.

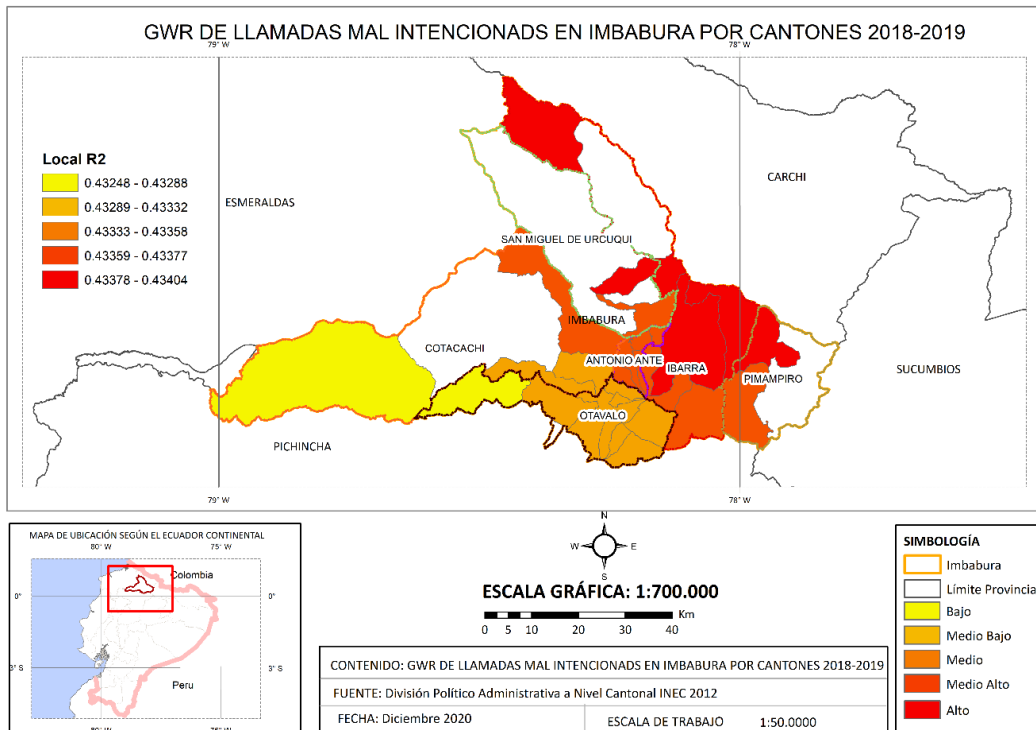


Figura 25. GWR de llamadas mal intencionadas en Imbabura por Cantones al 2030

Fuente: Gustavo Chacón

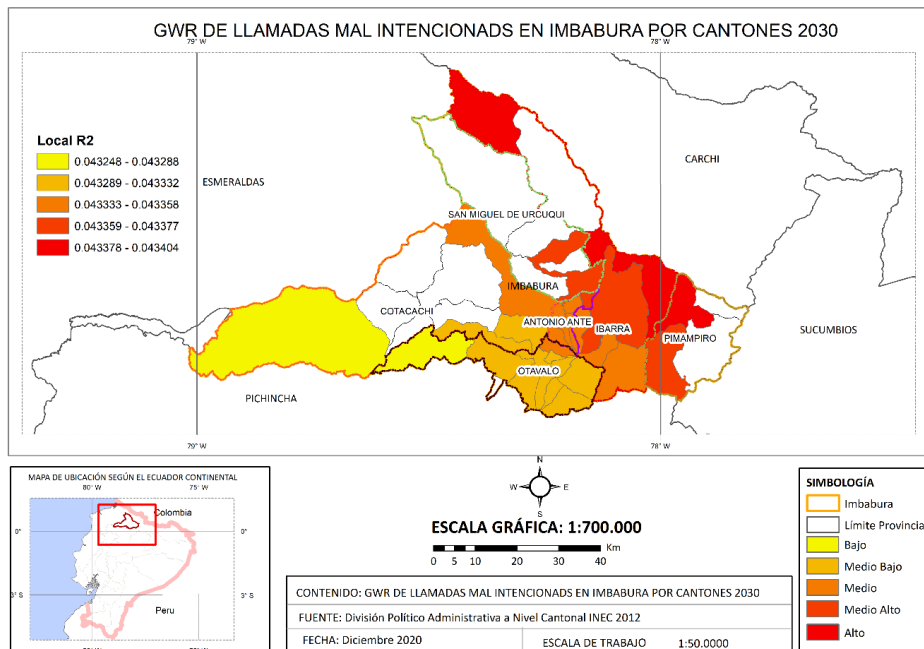


Figura 26. GWR de llamadas mal intencionadas en Imbabura por Cantones al 2030

Fuente: Gustavo Chacón

En la Figura 26 el modelo GWR tuvo modificación en relación con la variable independiente (población). En base a la proyección población del 2030 tomada del SNI (2021) se tiene que la zona urbana del cantón Pimampiro y las parroquias rurales de Ambuquí, Salinas y Lita tienen un 43% de la varianza en relación con las llamadas mal intencionadas.

4.5. Análisis de este resultado (GWR)

En este contexto, la tasa de llamadas mal intencionadas obedece a un patrón espacial especialmente en el sector rural del cantón Ibarra y Pimampiro. El modelo mostró un 43% de la varianza en la tasa de llamadas mal intencionadas y la población, el modelo GWR no pudo predecir en el resto de las parroquias y poder focalizar geográficamente la incidencia en el servicio de emergencia.

Para el año 2030 el modelo GWR mostró que la mayoría de las parroquias son rurales (Ambuquí, Salinas y Lita) y tienen un 43% de la varianza en relación con las llamadas mal intencionadas a diferencia de Pimampiro (cabecera cantonal). La dependencia espacial para este modelo con esta varianza le excluye a Ibarra a pesar de que el modelo representó un coeficiente R^2 bajo.

CAPÍTULO V: CONCLUSIONES Y RECOMENDACIONES

CONCLUSIONES

- ✓ Se detecta patrones de tiempo y espacio en el Sistema Integrado SIS ECU 911 Ibarra utilizando técnicas de minería de datos para la toma de decisiones informadas y reducción de costos técnicos-económicos. Revisando el moldeo GWR en la Figura 25 se encuentra una varianza baja del 43% entre las llamadas mal intencionadas y la variable relacionado a la población en parroquias como Pimampiro, Ambuquí, Ibarra, Salinas, Pablo Arenas y Lita, estas parroquias se ajustaron a este modelo. En el resto de las parroquias, a pesar de tener una tasa de llamadas más altas de la provincia como González Suárez, Pablo Arenas, y Selva Alegre todas rurales el modelo no ha permitido focalizar geográficamente una dependencia espacial que permita determinar un patrón predictivo significativo. Se confirma parcialmente que las llamadas mal intencionadas responden a un patrón espaciotemporal.
- ✓ Esta investigación traza su metodología con relación a un marco teórico basado en minería de datos con la herramienta de ESRI Criminal Analyst, una estructura de ingeniería de software desde el origen de la llamada mal intencionada (triangulación celular) hasta el procesamiento y análisis de esta. El aspecto teórico se fundamenta en organizar las herramientas técnicas en el proceso de detectar la llamada mal intencionada, en sectores o puntos interés (mapas calientes - hotspot), así como un análisis temporal (reloj de datos) en el proceso de encontrar una justificativo del alto número de este tipo de llamadas.
- ✓ El reloj de datos determina que todos los días y a toda hora existe llamadas mal intencionadas. Sin embargo, el comportamiento temporal se enfatiza en horarios entre las 14:00 hasta las 18:00 según la Figura 23. El mayor número de llamadas mal intencionadas entre el 2018 y 2019 los meses de julio y agosto son críticos, es decir, las llamadas mal intencionadas responden a un patrón temporal. Mientras que a nivel territorial las parroquias que tienen una mayor concentración de este tipo de llamadas se muestran en

las parroquias González Suarez y Selva Alegre y en sus respectivos sitios de interés como mercados y terminales, es decir, se aprecia una relación espacial mediante los mapas de calor.

- ✓ Para el año 2030, el modelo GWR en la Figura 26 muestra una varianza baja del 43% en parroquias rurales de: Ambuquí, Salinas y Lita en relación con las llamadas mal intencionadas con la población. Estas parroquias son colindantes con la provincia del Carchi, así como la densidad de antenas y bases celulares en ese sector contextualizan la problemática de estas llamadas. A pesar de que la varianza del modelo GWR para los años (2018-2019) y 2030 es parecida, refleja cierta parcialidad la relación de las llamadas mal intencionadas que responden a un patrón sobre todo espacial.

RECOMENDACIONES

- ✓ Con la carga de llamadas mal intencionadas, previo a los análisis temporales (reloj de datos) los cronogramas de asistencia técnica a la plataforma del ECU 911 Ibarra se podrían planificar sin perjudicar el rol del servicio de emergencia. A esto se añadiría un acompañamiento en territorio especialmente a las parroquias de González Suárez (Otavalo), Pablo Arenas (Urcuquí), Selva Alegre (Otavalo), en conjunto con instituciones articuladas permitirá causar un efecto de consciencia, donde se informe sobre las sanciones que implica el mal uso de la línea 911 y de esta manera reforzar la campaña “Buen uso de la línea 911”.
- ✓ Esta investigación usó ArcGis con la extensión Criminal Analyst de ESRI como herramienta de software licenciado que permitió identificar patrones espacio-temporales de llamadas mal intencionadas en bases de datos espaciales a través de la minería de datos. Como trabajo a futuro se puede crear un mecanismo que permita analizar el modelo de predicción GWR, puntos calientes y temporalidad de las llamadas con la información espacial almacenada definida como más precisa, es decir cuyas llamadas mal intencionadas entregaron información a una distancia no mayor de 500 metros. Esto implica actualizar la información geográfica a nivel de sectores censales, así como una

mejora continua de los webservices de las diferentes prestadoras de servicio móvil en el Ecuador.

- ✓ El empleo de herramientas de software libre (software WEKA) brinda la posibilidad de aplicar algoritmos de minería de datos para extraer información relevante, de una gran cantidad de información geográfica. En este sentido la metodología propuesta en el presente trabajo puede ser utilizada sobre todo para sugerir predicciones en el territorio de sectores o zonas que hacen mal uso de la línea única de emergencias 911 en referencia a las llamadas mal intencionadas.

BIBLIOGRAFÍA

- Admin. (18 de Abril de 2010). *www.enseguridad.com*. Obtenido de <http://www.enseguridad.com/la-historia-del-numero-de-emergencias->
- Alonso, D. (2016). *mappinggis.com*. Obtenido de <https://mappinggis.com/2016/06/trabajar-geodatabases-esri>
- Álvarez Menéndez, J. (2008). Minería de Datos: Aplicaciones en el sector de las telecomunicaciones. 8.
- ARCGIS. (2017). Obtenido de <https://desktop.arcgis.com/es/arcmap/10.3/tools/spatial-statistics-toolbox/h-how-hot-spot-analysis-getis-ord-gi-spatial-stati.htm>
- ARCGIS. (2019). Obtenido de <https://desktop.arcgis.com/es/arcmap/10.3/manage-data/geodatabases/what-is-a-geodatabase.htm>
- ARCGIS. (2020). *DESKTOP ARCGIS*. Obtenido de <https://desktop.arcgis.com/es/arcmap/latest/extensions/tracking-analyst/094-the-data-clock.htm>
- ARCOTEL. (2019). *ARCOTEL*. Obtenido de <http://www.arcotel.gob.ec/arcotel-solicita-suspension-de-lineas-telefonicas-por-llamadas-falsas-al-sis-ecu-911/>
- Bruner, R. (2018). *Muyinteresante.com*. Obtenido de <https://www.muyinteresante.es/tecnologia/preguntas-respuestas/que-es-la-mineria-de-datos-311477406441>
- Capdevila Montes, E., & Mínguez García, M. (2019). *Introducción a los Sistemas de Información Geográfica*.

Cognodata. (19 de Abril de 2019). *Cognodata*. Obtenido de <https://www.cognodata.com/blog/tecnologia-big-data/>

COIP. (5 de Febrero de 2008). Obtenido de https://www.defensa.gob.ec/wp-content/uploads/downloads/2018/03/COIP_feb2018.pdf

COIP. (2018). *Código Orgánico Integral Penal*.

CONATEL. (20 de Octubre de 2011). RESOLUCION TEL -756-21 -CONATEL-2011. Quito, Pichincha , Ecuador.

Constitución del Ecuador. (2008). *Constitución del Ecuador*. Montecristi.

Datamall. (25 de Octubre de 2018). *DatamallEC*. Obtenido de <http://datamall.com.ec/estason-las-7-v-que-caracterizan-el-big-data/>

Dueñas-Reyes, M. (2009). Minería de datos espaciales en búsqueda de la verdadera informacion. *Ingeniería y Universidad*.

ECU 911. (2015). Obtenido de <https://www.ecu911.gob.ec/preguntas-frecuentes/>

ECU 911. (2015). Obtenido de <https://www.ecu911.gob.ec/servicio-integrado-de-seguridad-ecu-911/>

ECU 911. (2017). Obtenido de <https://www.ecu911.gob.ec/servicio-integrado-de-seguridad-ecu-911/>

ECU 911. (2018). Obtenido de <https://www.ecu911.gob.ec/wp-content/uploads/2019/02/Informe-Preliminar-Rendicion-cuentas-Planta-Central.pdf>

ECU 911. (2019). Obtenido de <https://www.ecu911.gob.ec/como-reportar-al-9-1-1/>

ECU911. (2017). Obtenido de <https://www.ecu911.gob.ec/servicio-integrado-de-seguridad-ecu-911/>

El Comercio. (2014). *El Comercio*. Obtenido de https://www.elcomercio.com/app_public.php/actualidad/ecuador/ibarra-se-inauguro-nuevo-ecu.html

Estoy full ebrio’, ‘se me bloqueó el celular’, ‘me trae una salchipapa’; 12 millones cuestan las 'bromas' al ECU 911. (24 de julio de 2019). *El Comercio*. Obtenido de <https://www.elcomercio.com/actualidad/llamadas-falsas-ecu911-gasto-emergencias.html>

Fernández, M. (2014). Integración de Información Georeferenciada para optimizar el.

Fotheringham, A., Brunson, M., & Charlton, M. (2002). Geographically weighted regression: The analysis of spatially variable relationships. 284.

Galileo Universidad. (09 de 10 de 2019). Obtenido de www.galileo.edu/trends-innovation/que-son-las-telecomunicaciones/

Güell, A., & Torre, M. (1 de Julio de 2010). *GENCAT*. Obtenido de https://territori.gencat.cat/es/detalls/Article/GeoServeis_ICC_01

IIC. (2019). *Instituto de Ingeniería del Conocimiento*. Obtenido de <https://www.iic.uam.es/innovacion/big-data-caracteristicas-mas-importantes-7-v/#valor-dato>

INEC, I. (2010b). *Manual de Actualización Cartográfica*. Quito.

Informador Chile. (21 de Junio de 2020). Obtenido de <https://www.elinformadorchile.cl/2020/06/21/noticias-chile-gobierno-rastrear-antenas-de-celular-para-poder-ver-la-movilidad-de-los-chilenos-en-cuarentena/>

Jimpako. (10 de Agosto de 2008). *Servidores Geograficos*. Obtenido de <http://servidoresgeograficos.blogspot.com/2008/07/geodatabase.html>

Lozada, E. (6 de diciembre de 2017). *GEOPAISA*. Obtenido de <https://geopaisa.blog/2017/12/06/funciones-de-un-sig>

Microsoft.com. (09 de 03 de 2018). Obtenido de <https://docs.microsoft.com/es-es/sql/analysis-services/data-mining/data-mining-concepts>

Moreno Muñoz, A., & Lara Torralbo, J. (25 de Noviembre de 2016). Análisis de actividad de un servicio de teleasistencia social mediante Big Data y Data Mining. Madrid, España.

Nash, J., & Sutcliffe, J. (1970). River flow forecasting through conceptual models. Part 1. A discussion of principles. *Journal of Hydrology*.

NETEC. (27 de mayo de 2019). *netec.com*. Obtenido de <https://bit.ly/3fyETQe>

Núñez Cárdenas, F. (2018). *Universidad Autónoma del Estado de Hidalgo*. Obtenido de El proceso de minería de datos: <https://www.uaeh.edu.mx/scige/boletin/huejutla/n1/m2.html>

Oracle.com. (2020). Obtenido de <https://www.oracle.com/es/big-data/what-is-big-data.html>

Oregon State University. (13 de 04 de 2016). *Oregon State*. Obtenido de <https://blogs.oregonstate.edu/geo599spatialstatistics/2016/04/13/arcgis-hot-spot-analysis-with-a-fire-occurrence-dataset/>

(2014). *Plan Estratégico 2014-2017 ECU 911*. QUITO.

Plan Nacional de Seguridad Integral. (2014).

PowerData. (2020). Obtenido de <https://www.powerdata.es/big-data>

Prometeusgs. (19 de Febrero de 2019). *Prometeusgs*. Obtenido de <https://prometeusgs.com/volumen-variedad-velocidad-veracidad-y-valor-las-5-dimensiones-del-big-data-la/>

- Recuero de los Santos, P. (09 de 03 de 2018). *Think Big*. Obtenido de Machine Learning: supervisado y no supervisado: <https://empresas.blogthinkbig.com/que-algoritmo-elegir-en-ml-aprendizaje/>
- Riquelme, J., Ruiz, R., & Gilbert, K. (2006). Minería de Datos: Conceptos y Tendencias. *Inteligencia Artificial Volumen 10. Nro 29*, 18.
- Rivas, E. (08 de 01 de 2018). *iebschool.com*. Obtenido de <https://www.iebschool.com/blog/data-mining-mineria-datos-big-data/>
- Roca Chillada, J. (2019). Obtenido de www.informeticplus.com/que-son-las-telecomunicaciones
- Rodriguez, P., Palomino, N., & Mondaca, J. (2017). El uso de datos masivos y sus técnicas analíticas para el diseño e implementación de políticas públicas en Latinoamérica y el Caribe.
- Romeo López, J., & Romero Frías, R. (2016). *El Big Data y el futuro de las Telecomunicaciones*. Obtenido de http://cdn2.hubspot.net/hub/239039/file-1959511423-pdf/docs/%5BIC%5D_OFFER_-_EBOOK-_Big_data_y_sector_telecomunicacionesv2.pdf
- Romero, J. (11 de junio de 2019). *jorgeromero.net*. Obtenido de <https://jorgeromero.net/tecnicas-y-algoritmos-de-mineria-de-datos/>
- Ruiz-Linares, J., Guzmán-Patiño, F., & Castro-Ávila, M. (2017). Comparing the National Saber 11 and Saber Pro Exams in Antioquia, Colombia: An Approach from Geographically Weighted Regression (GWR).
- Schrage, M. (2019). *sas.com*. Obtenido de https://www.sas.com/es_mx/insights/analytics/data-mining.html
- SENPLADES. (2017). *Plan Nacional de Desarrollo Toda una Vida*. Quito-Ecuador.

- Senplades. (2018). *Senplades*. Obtenido de <https://www.planificacion.gob.ec/zona-de-planificacion-1-norte/>
- SIS ECU 911. (2010). *Servicio Integrado de Seguridad ECU911*. Obtenido de <https://www.ecu911.gob.ec/localizador-mobil/>
- SIS ECU 911, S. (2017a). *Plan estratégico*. Quito.
- SNI, S. (Agosto de 2015). *SNI Consultas Interactivas*. Obtenido de <http://menucloud.sni.gob.ec/web/menu>
- SNI, S. (Enero de 2021). *SNI*. Obtenido de <https://sni.gob.ec/proyecciones-y-estudios-demograficos>
- SNI, S. (s.f.). *Sistema Nacional de Información*. Obtenido de <https://sni.gob.ec/proyecciones-y-estudios-demograficos>
- Techtarget. (Marzo de 2020). *SearchdataCenter*. Obtenido de <https://searchdatacenter.techtarget.com/es/definicion/Visualizacion-de-datos>
- Telectronika*. (23 de 04 de 2018). Obtenido de <https://www.telectronika.com/articulos/que-son-las-telecomunicaciones/>
- Telectronika. (23 de abril de 2018). *telectronika*. Obtenido de <https://www.telectronika.com/articulos/que-son-las-telecomunicaciones/>
- Trujillano, J., Sarria-Santamera, A., Esquerda, A., Badia, M., Palma, M., & March, J. (2008). Aproximación a la metodología basada en árboles de decisión (CART). Mortalidad hospitalaria del infarto agudo de miocardio. *Gaceta Sanitaria*.
- UNIR. (26 de 06 de 2020). *Unir.com*. Obtenido de <https://www.unir.net/ingenieria/revista/noticias/3-v-big-data/549205067239/>
- Valcárcel, V. (2004). Data Mining y el descubrimiento del conocimiento. *Revista de la Facultad de Ingeniería Industrial - UNMSM*.

Weiss, R. (2015). Crime Analyst ARCGIS.

www.esri.com. (2019). Obtenido de <https://pro.arcgis.com/es/pro-app/tool-reference/spatial-statistics/how-geographicallyweightedregression-works.htm>

www.esri.com. (2019). Obtenido de <https://doc.arcgis.com/es/insights/create/heat-maps.htm>

www.esri.com. (2019). Obtenido de <https://desktop.arcgis.com/es/arcmap/latest/extensions/tracking-analyst/094-the-data-clock.htm>

www.esri.com. (2019). Obtenido de <http://desktop.arcgis.com/es/arcmap/10.3/manage-data/geodatabases/what-is-a-geodatabase.htm>