



CentroSur
e-ISSN: 2600-5743
Social Science Journal

UTILIZACIÓN DE ALGORITMOS DE CLASIFICACIÓN PARA LA PREDICCIÓN DE LOS DELITOS QUE AFECTAN LA SEGURIDAD CIUDADANA

USE OF CLASSIFICATION ALGORITHMS FOR THE PREDICTION OF CRIMES THAT AFFECT CITIZEN SECURITY

Centro Sur.
Social Science Journal
2022 – Número continuo
<http://centrosureditorial.com/index.php/revista>
eISSN: 2600-5743
revistacentrosur@gmail.com

Atribución/Reconocimiento-
NoComercial-CompartirIgual 4.0
Licencia Pública Internacional —
CC BY-NC-SA 4.0
<https://creativecommons.org/licenses/by-nc-sa/4.0/legalcode.es>

Byron Oviedo Bayas
Jorge Guanín Fajardo
Jorge Gómez-Gómez

Resumen

El presente trabajo está orientado para la obtención de conocimiento mediante la aplicación de minería de datos usando algoritmos de clasificación para diseñar y aplicar políticas de seguridad que adapten las necesidades de la población a un diagnóstico fiable de los delitos más comunes por parte del personal especializado en seguridad ciudadana para conocer los estratos de la población de estudio y analizar los factores que disparan los índices delictivos en áreas determinadas

Palabras Clave: algoritmos de minería de datos, crimen, seguridad pública

Universidad Técnica Estatal de Quevedo,
<https://orcid.org/0000-0002-5366-5917>,
boviedo@uteq.edu.ec, 120508, Quevedo-Ecuador

Universidad Técnica Estatal de Quevedo,
<https://orcid.org/0000-0001-9150-4009>,
jorgeguanin@uteq.edu.ec, 120508, Quevedo-Ecuador

Universidad de Córdoba, <https://orcid.org/0000-0001-8746-9386>,
jeliecergomez@correo.unicordoba.edu.co,
Montería-Colombia

Abstract

This work is aimed to obtain knowledge by applying data mining using classification algorithms to design and implement security policies that suit the needs of the population to a reliable diagnosis of the most common crimes by specialized personnel public safety

to meet the strata of the study population and analyzing the factors that trigger the crime rates in certain areas.

Key words: data mining algorithms, crime, public safety

Introducción

Actualmente vivimos en una sociedad muy variable, estos comportamientos sean positivos o negativos denotan lo vinculante de las acciones en nuestro entorno, es así que a medida que la sociedad evoluciona, la delincuencia va desarrollando nuevas formas de organización, logrando infundir pavor en el entorno en que se desenvuelve un individuo convirtiéndose esto en un problema de carácter social, llevando a que el Estado Ecuatoriano a través de su Función Legislativa incorpore nuevos delitos y sanciones al Código Orgánico Integral Penal, con el propósito de frenar el auge del crimen organizado.

Es así que con el fin de mantener la seguridad en el Ecuador y evitar que la delincuencia organizada llegue a corromper a los adolescentes, jóvenes y niños que son los más vulnerables a esta actividad criminal; la Policía Nacional del Ecuador basada en su talento humano, profesionalismo y capacidad logística, ha logrado entablar una coordinación e intercambio de información oportuna con la ciudadanía en general, la misma que ha sido el eje fundamental para combatir de manera estratégica y fuerte a la delincuencia en cada una de las provincias del Ecuador.

El índice de delincuencia que posee la provincia del Guayas ha sido un referente para llevar a cabo esta investigación, con el propósito de indagar las causales que han ocasionado que estadísticamente Guayas sea la provincia más conflictiva en el Ecuador.

La inseguridad que se vive actualmente en la provincia del Guayas es alarmante, más que todo por el incremento que ha tenido en estos tres años, y las nuevas formas de delinquir que han buscado los bandidos que comprende, desde asesinar por un teléfono móvil hasta el sicariato.

Problema: Si bien es cierto en muchas ciudades uno de los problemas que aqueja siempre es el económico y falta de trabajo que tienen los ciudadanos, citando estás como una de las principales causas para que la delincuencia se incremente. Sin desmerecer el trabajo constante que realiza la Policía Nacional para mantener la seguridad ciudadana en cada una de las ciudades del Ecuador, en especial la ciudad en estudio que es Guayaquil.

Otra de las causas que ha generado el incremento de la delincuencia es la fragilidad en la que se encuentran actualmente las familias, siendo la base para formar niños, jóvenes y adolescentes de bien; actualmente el trabajo y la tecnología adsorben el tiempo de los padres y el mal uso que generalmente dan los niños, jóvenes y

adolescentes a la tecnología sumado a las malas amistades y la no existencia de supervisión de los padres, ha llevado a involucrarse a los jóvenes en actos delictivos.

Propósito del Trabajo: Obtener mediante algoritmos de minería de datos el conocimiento a través de reglas de clasificación, que permitan a los gestores de la seguridad ciudadana mejorar sus estrategias para combatir a la delincuencia.

Trabajos Relacionados.

Ruiz en su trabajo señala que resulta paradójico que cuando se diseña una política de seguridad local en muchas ocasiones no se establezca un diagnóstico que igualmente interprete las necesidades de los sectores de esa localidad. La intervención inicial de la policía contempla reuniones con los vecinos, las cuales resultan en muchas ocasiones poco productivas. En el mejor de los casos estas reuniones proveen un listado de quejas de la comunidad y reflejan cierta animosidad hacia la policía. En el peor de los casos, estas reuniones son simplemente preventivas y para invitar a los vecinos a estar vigilantes frente al delito, sugerirles la instalación de alarmas, la pronta denuncia en caso de incidentes, etc. La mayoría de las veces no se logra construir un diagnóstico riguroso de la situación de cada sector. La información recolectada por la policía queda supeditada a percepciones e intuiciones [1].

En efecto, pese a la inexistencia y/o a la poca confiabilidad de las estadísticas criminales, a las diferencias en la forma de recopilar y clasificar los delitos descubiertos por la policía o denunciados ante ellas o las autoridades judiciales, el panorama se caracteriza por el continuo incremento de los índices delictuales y la victimización (particularmente significativos a partir de la segunda mitad de los años ochenta); el acelerado crecimiento de la delincuencia violenta, de los delitos contra la vida y la integridad física de las personas; la alta participación de la población masculina joven en la criminalidad (tanto en calidad de víctimas y de victimarios) y la diversidad de factores criminógenos coadyuvantes, tales como: las persistentes condiciones de desigualdad social, las secuelas de enfrentamientos armados, las extensión del consumo de alcohol y drogas psicoactivas, el acceso a armas y los significativos índices de la violencia intrafamiliar, entre otros [2].

La evolución del conflicto, el perfil del crimen, así como las condiciones sociales demandan un método especializado. Por ello, el modelo que implementarán recoge enseñanzas dejadas por administraciones pasadas y le imprime un estilo de vigencia, respeto y promoción de la dignidad humana para garantizar la vida, libertad, integridad, patrimonio y tranquilidad ciudadana. Sin embargo, dentro de las recomendaciones a corto plazo se tiene en principio, la evaluación del Sistema Unificado de Información de Violencia y Delincuencia de Bogotá SUIVD, además de definir los factores de riesgo que atentan contra la dignidad humana en Bogotá [3].

Metodología

Para el desarrollo de este trabajo se consideró como herramienta de análisis de los datos el programa de minería de datos KNIME (Konstanz Information Miner) [7].

Preparación de los datos: Los datos que se obtuvieron es información de los delitos denunciados en la Policía Judicial en Guayaquil que es la ciudad de estudio para este trabajo. Información que se filtró y depuró de las hojas electrónicas de Excel Microsoft Office.

De acuerdo como se muestra en la **Figura 1** de forma descriptiva a través de un gráfico de burbuja se tuvo los delitos más comunes efectuados en la ciudad de Guayaquil.

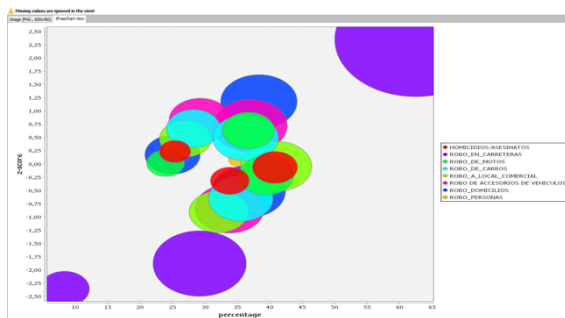


Fig. 1 Visualización de variables de estudio.

Además del gráfico anterior en la **Figura 2** se analizó el comportamiento de las variables semana y mes de la infracción y cuartil_edad, ésta última presenta menos variación en relación a las otras dos primeras.

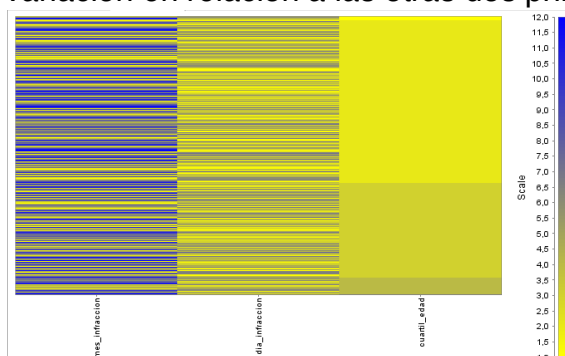


Fig. 2 visualización de variables de estudio.

Descripción de los datos: De entre la información de estudio se abordaron las variables de estudio tanto numéricas como categóricas. Se seleccionó 13 variables y una clase.

Tabla 1: Variables usadas en el trabajo

Variables	Descripción
distrito	Distrito donde se encuentra ubicada la policía UPC
circuito	Circuito del distrito al que pertenece el UPC
zona	Zona donde se comete el delito
semana	Semana donde se comete el delito
mes_infraccion	Mes_Infraccion donde se comete el delito
dia_infraccion	Dia_Infraccion donde se comete el delito
delito_pj	Delitos que se denuncian en la Policía Judicial
modus_operandi	Modus_Operandi donde se comete el delito
vd_sexo	Sexo de la víctima denunciante
vd_edad	Edad de la víctima denunciante
cuartil_edad	Cuartil_Edad
vd_nacionalidad	Nacionalidad de la víctima denunciante
cmi	Tipo de crimen efectuado
class	Clase

Extracción del conocimiento: Previo a la ejecución de los algoritmos de clasificación en la herramienta de minería de datos KNIME.

Técnicas de clasificación supervisadas: Para este trabajo se utilizó como técnicas de clasificación tres algoritmos. Random Forest [4], Decision Tree [5], NetBayes [6].

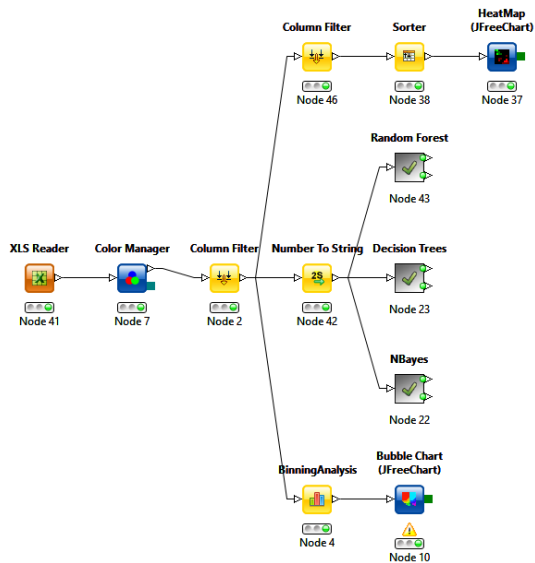


Fig. 3 Flujo de Nodos Knime

En la **Figura 3** se efectuó el flujo de nodos con el programa KNIME para obtener las reglas de clasificación a través de los algoritmos seleccionados, tal como se muestra a continuación en la **Tabla 2** los resultados de los algoritmos.

Tabla 2: Resultados de ejecución de algoritmos.

Algoritmo	Precisión
Net Bayes	97.74
Decision Tree	99.89
Random Forest	99.99

El algoritmo NetBayes a pesar de tener un resultado menor, se atribuyó que estos resultados son más fiables, ya que en comparación de los otros dos tienen sobreentrenamiento y resulta más conveniente para este trabajo el algoritmo de NetBayes.

Conclusiones

Es necesario aclarar, que la combinación de otras técnicas de minería de datos lograría mejorar la interpretabilidad de las reglas, sin embargo, este trabajo logró obtener reglas concretas a fin de que sean analizadas por expertos en seguridad ciudadana.

La mayoría de los delitos que se cometieron de acuerdo a los resultados obtenidos corresponde a los meses de marzo, abril, mayo y agosto, en sectores urbanos, siendo las víctimas más comunes los ciudadanos comprendidos entre 20 y 39 años de edad.

Referencias

- [1] J. C. Ruiz, «Los Mitos acerca de la Seguridad Local y la Policía Comunitaria,» *Ciencia Política y Gobierno. Grupo de Estudios Regionales, Borradores de Investigación*, pp. 3-26, 2006.
- [2] P. Tudela, Reforma policial, COP y POP: Perspectivas latinoamericanas de aplicación y paradigmas de gestión policial dirigida a la comunidad, Chile: Policía de investigaciones de Chile, 2007.
- [3] J. Libreros, POLÍTICA DE SEGURIDAD CIUDADANA EN BOGOT, Bogotá: FESCOL, 2008.
- [4] B. L., «Random Forest,» *Machine Learning*, vol. 45, n° 1, pp. 5-32, 1999.
- [5] Q. J.R., «C4.5,» *Programs for machine learning*, pp. 544-555, 1999.
- [6] G. J. P., «Estimating Continuous Distributions in Bayesian Classifiers,» *In proceedings of the Eleventh Conference on Uncertainty in Artificial Intelligence*, pp. 1-8, 1995.
- [7] KNIME, «Konstanz Information Miner,» *KNIME*, 2004.