

Aplicación de los árboles de decisión en la identificación de patrones de lesiones fatales por causa externa en el municipio de Pasto, Colombia

Resumen

Introducción: La Organización Panamericana de la Salud (OPS) desde el año 1993 y la Organización Mundial de la Salud (OMS) en 1996, aceptaron que la violencia es un problema de salud pública, situación que se corrobora en el Informe de Violencia y Salud, en el cual América Latina presentó una tasa de homicidios de 18 por cada 100.000 personas, y es considerada como una de las regiones más violentas del mundo.

Objetivo: Detectar patrones delictivos con técnicas de minería de datos en el Observatorio del Delito del municipio de Pasto (Colombia).

Materiales y métodos: Se aplicó Cross Industry Standard Process for Data Mining (CRISP-DM), una de las metodologías utilizadas en el desarrollo de proyectos de minería de datos en los ambientes académico e industrial. La fuente de información fue el Observatorio del Delito del municipio de Pasto, donde está almacenadas las cifras históricas, limpias y transformadas sobre las lesiones de causa externa (fatales y no fatales), registrados en 11 años.

Resultados: Se construyó un modelo de clasificación basado en árboles de decisión que permitió descubrir patrones de muertes por causa externa. Para el caso de homicidios, estos sucedieron en su mayoría en la Comuna 5 de Pasto, los fines de semana, en la madrugada, en el segundo semestre del año, en la vía pública y las víctimas fueron hombres adultos, de oficios varios, la causa de los homicidios fueron riñas y se produjeron con arma de fuego.

Conclusión: El conocimiento generado ayudará a los organismos gubernamentales y de seguridad a tomar decisiones eficaces en lo relacionado a la implementación de planes de prevención de delitos y seguridad ciudadana.

Palabras clave: Reconocimiento de normas patrones automatizadas; minería de datos; árboles de decisión; clasificación

Abstract

Introduction: The Pan American Health Organization (PHO) and the World Health Organization (WHO) accepted, since the year 1993 and 1996 respectively, that violence is a public health problem, a situation that is corroborated in the report on violence and health, in which Latin America presented a homicide rate of 18 per 100,000 people, and it is considered one of the most violent regions in the world.

Objective: To detect criminal patterns with data mining techniques in the Crime Observatory of the municipality of Pasto (Colombia).

Materials and methods: Cross Industry Standard Process for Data Mining (CRISP-DM) was applied, which is one of the methodologies used in the development of data mining projects in academic and industrial environments. The source of information was the Crime Observatory of the municipality of Pasto, where the historical clean and transformed figures on the injuries of external cause (fatal and nonfatal) recorded in 11 years are stored.

Results: A decision tree-based classification model was built that allowed the discovery of patterns of deaths from external causes. In the case of homicide, these happened mostly in the commune 5 in Pasto under the following circumstances: during the weekends, in the early morning, in the second semester of the year and in the public thoroughfare; besides, the victims were adult men of various professions; and the cause of the homicides were quarrels and they were produced with a fire gun.

Conclusion: The generated knowledge will help government and security agencies make effective decisions regarding the implementation of crime prevention and citizen security plans.

Keywords: Pattern recognition, automated, data mining, decision trees, classification