

Segundo Trabajo Técnicas de Aprendizaje Estadístico

Camila Acosta Ramírez¹, Nicolas Amado Aristizábal², Sergio Andrés Arboleda Zuluaga³,
Esteban Palacio Londoño⁴, Natalia Ramírez Ossa⁵ & Carlos Alejandro Torres Albarracín⁶

1 Universidad Nacional de Colombia - Sede Medellín. caacostara@unal.edu.co

2 Universidad Nacional de Colombia - Sede Medellín namadoa@unal.edu.co

3 Universidad Nacional de Colombia - Sede Medellín. saarboledaz@unal.edu.co

4 Universidad Nacional de Colombia - Sede Medellín. epalaciol@unal.edu.co

5 Universidad Nacional de Colombia - Sede Medellín. nramirezo@unal.edu.co

6 Universidad Nacional de Colombia - Sede Medellín. caatorresal@unal.edu.co

Elabore un ensayo de una página argumentando cómo desde el aprendizaje estadístico y los datos abiertos se puede contribuir a el problema de la inseguridad en Medellín

El aprendizaje estadístico y las diferentes técnicas que incorpora pueden contribuir a mitigar el problema de inseguridad en Medellín, ya que se pueden usar los datos abiertos para interpretar y modelar la problemática. Al analizar los datos se pueden observar los delitos más cometidos, identificar zonas con alta concurrencia de los delitos, armas que fueron empleadas al cometer el delito, entre muchas otras variables. Es posible realizar un mapeo para revelar la relación que hay entre los delitos y la ubicación geográfica en los cuales son cometidos.

El tener acceso a los datos de la Secretaria de Movilidad de Medellín hace posible la participación de la comunidad académica, empresarial e iniciativas personales lo cual aporta distintos puntos de vista del problema desde el entendimiento de este hasta las soluciones y recomendaciones presentadas.

Existen diversos métodos de análisis que se pueden realizar para obtener información sobre asociación, interpretación y modelado; por ejemplo, con **métodos de análisis multivariado** se puede hacer un **análisis de coeficientes de correlación** para medir las relaciones entre los delitos y armas usadas, delitos y zonas/barrios entre otras variables, **análisis de componentes principales** para reducir la dimensionalidad creando un conjunto de variables independientes, **clustering** para agrupar e identificar por zonas/barrios de la ciudad se están presentando más delitos e inclusive a que horas se presentan; se pueden realizar por medio de **Machine Learning** modelos capaces de hacer predicciones y prever con base en los datos cuales zonas pueden presentar un aumento en el corto y mediano plazo.

Con los resultados obtenidos se puede visualizar en mapas la distribución espacial de las zonas más afectadas revelando información acerca de los delitos que ocurren y con su frecuencia. De esta manera, se pueden crear planes personalizados por zona que permiten disminuir la cantidad de delitos cometidos, aumentar los índices de seguridad y mejorar la percepción de seguridad general, generando así un impacto positivo en la comunidad.