

## Las deficiencias actuales en el tratamiento del *bias* en *Machine learning*

Ing. Yadira Lizama-Mué

### Resumen

El uso de sistemas predictivos se ha ampliado con el desarrollo de métodos computacionales relacionados y la evolución de las ciencias en las que se aplican estos métodos (Solon y Selbst, Calif L REV 104: 671–732, 2016; Pedreschi et al, 2007). Estos métodos, clasificados en el dominio de la inteligencia artificial, incluyen técnicas de aprendizaje automático, reconocimiento de rostro y / o voz, mapeo de temperatura, procesamiento de lenguaje natural, visión computacional, entre otros. Además, se están aplicando actualmente para resolver problemas en áreas sensibles desde el punto de vista social y político como la prevención del delito, la gestión de la justicia, la gestión de multitudes y el análisis de emociones, solo por mencionar unos pocos. Sin embargo, hoy en día se pueden encontrar varias predicciones como resultado de la aplicación de estos métodos que resultan en clasificación errónea, por ejemplo, para el caso de evaluación de riesgo de condena (Oficina de Libertad Condicional y Servicios Prejuicios, 2011) o el proceso de toma de decisiones al diseñar políticas públicas (Lange, 2015).

El objetivo de este documento es identificar las brechas actuales en logro de la equidad en el contexto de los sistemas predictivos en inteligencia artificial mediante el análisis de literatura académica y científica hasta 2020. Para lograr este objetivo, se han reunido materiales indexados en la Web of Science y Scopus de los últimos 5 años, y se analizaron los diferentes métodos propuestos en dichas publicaciones y sus resultados en relación con el sesgo (*bias*) como un problema emergente en el campo de estudio de la Inteligencia Artificial. Las conclusiones preliminares indican que el aprendizaje automático tiene algunas limitaciones intrínsecas que conducen a automatizar el sesgo al diseñar algoritmos predictivos. En consecuencia, otros métodos deben ser explorados o deberíamos redefinir la forma en que se utilizan los enfoques actuales de aprendizaje automático al construir sistemas de apoyo a la toma de decisiones para instituciones cruciales de nuestra sociedad, como el sistema judicial, solo por mencionar uno.

Ver el video de la conferencia:

<https://www.youtube.com/watch?v=29d3RwF7JIY>

Ing. Yadira Lizama-Mué



**Yadira Lizama-Mué** es graduada de Ingeniería en Ciencias Informáticas, Universidad de las Ciencias Informáticas, Cuba. Miembro del Grupo de Investigaciones de Sociocibernética y LegalTech de la Sociedad Cubana de Derecho e Informática. Actualmente es aspirante a doctorado en Western University, Ontario, Canadá.