

## Guía de lectura

### Machine Learning with R, Brett Lantz, 2<sup>nd</sup> Ed. Capítulo 4: Probabilistic Learning – Classification Using Naïve Bayes

Desarrolle los siguientes puntos basándose en la lectura. Se puede apoyar en fuentes adicionales (artículos, libros, etc).

- Explique la diferencia entre probabilidad de un evento, probabilidad conjunta de dos eventos, y probabilidad condicional de un evento dado otro evento.
- ¿Qué significa que dos eventos sean independientes?
- Explique las suposiciones en las que se basa el algoritmo de Naïve Bayes. ¿Cuál es el propósito de hacer esas suposiciones? ¿Por qué se le llama “Naïve” (ingenuo) a este método?
- Analizar el significado del teorema de Bayes, y qué se necesita para estimar la probabilidad  $P(\text{Clase}=X \mid \text{Salario}=Y)$ .
- Aplicar el algoritmo de Naïve Bayes según el ejemplo del capítulo.
- ¿Cuáles son las implicaciones de tener una variable predictiva categórica (por ejemplo  $x_i \in \{A, B, C\}$ ) cuando en el dataset no se tiene ninguna instancia para una categoría en particular (por ejemplo, no hay ningún registro con  $x_i = A$ )? ¿Cómo se pueden abordar esas implicaciones)?