MÉTODOS ESTADÍSTICOS DE APRENDIZAJE

ERIKA TANGER Y WILMAN VALDES

Universidad Antonio Nariño

15 de mayo del 2014

INTRODUCCIÓN

- En este capítulo aprenderemos como los agentes pueden tratar con la incertidumbre utilizando métodos que hagan uso de las probabilidades y de la teoria de la decisión pero primero deben aprender las teorias probabilisticas sobre el mundo apartir de la experiencia.

Aprendizaje Estadistico

- En el aprendizaje estadistico las variable claves son los datos y las hipótesis.
- Los datos son evidencias, es decir instancias de todas o algunas de las variables aleatorias.
- Las hipótesis son teorias proabilisticas.

Razonamiento Bayesiano

- Nos da un enfoque probabilístico de la inferencia.
- Está basado en asumir que las incógnitas de interés siguen distribuciones probabilísticas.
- Se puede conseguir una solución óptima por medio de estas distribuciones y datos observados.
- Nos da la posibilidad de realizar una ponderación de la posibilidad de ocurrencia de una hipótesis de manera cuantitativa.

Importancia del Razonamiento Bayesiano

- Los algoritmos de aprendizaje bayesiano pueden calcular probabilidades explícitas para cada hipótesis.
- También nos proporcionan un marco para estudiar otros algoritmos de aprendizaje.

Primera Caracteristica

- Cada ejemplo de entrenamiento afecta a la probabilidad de las hipótesis. Esto es más efectivo que descartar directamente las hipótesis incompatibles.
- Se puede incluir conocimiento a priori: probabilidad de cada hipótesis; y la distribución de probabilidades de los ejemplos.
- Es sencillo asociar un porcentaje de confianza a las predicciones, y combinar predicciones en base a su confianza

Segunda Caracteristica

- Una nueva instancia es clasificada como función de la predicción de múltiples hipótesis, ponderadas por sus probabilidades.
- Incluso en algunos casos en los que el uso de estos métodos se ha mostrado imposible, pueden darnos una aproximación de la solución óptima.

Teorema de Bayes

- P(h) es la probabilidad a priori de la hipótesis h.
- P(D) es la probabilidad de observar el conjunto de entrenamiento D.
- P(D—h) es la probabilidad de observar el conjunto de entrenamiento D en un universo donde se verifica la hipótesis h.

Readable Mathematics

Let $X_1, X_2, \ldots, X_n$ be a sequence of independent and identically distributed random variables with $\mathrm{E}[X_i] = \mu$ and $\mathrm{Var}[X_i] = \sigma^2 < \infty$, and let

$$S_n = \frac{X_1 + X_2 + \cdots + X_n}{n} = \frac{1}{n} \sum_{i}^{n} X_i$$

denote their mean. Then as n approaches infinity, the random variables $\sqrt{n}(S_n - \mu)$ converge in distribution to a normal $\mathcal{N}(0, \sigma^2)$.