

Seminario: Tarea 1

Kevin García - Alejandro Vargas

6 de junio de 2019

Contenido

- 1 Antecedentes
 - Antecedente 1
 - Antecedentes 2
- 2 Efron & Hastie (2017)
 - Tema 1
 - Tema 2
- 3 Lantz (2013)
 - Tema 1
 - Tema 2

Planes de muestreo de aceptación de atributos mejorados basados en muestreo de nominaciones máximas

Este artículo, de los autores Jozani & Mirkamali (2010), demuestra el uso de la técnica de muestreo de nominación máxima (MNS) en diseño y evaluación de planes de muestreo de aceptación únicos para atributos.

Se analiza el efecto del tamaño de la muestra y el número de aceptación en el rendimiento de los planes MNS propuestos mediante la curva característica de operación (OC). Entre otros resultados, se muestra que los planes de muestreo de aceptación de MNS con un tamaño de muestra más pequeño y un número de aceptación mayor se desempeñan mejor que los planes de muestreo de aceptación comúnmente usados para atributos basados en la técnica de muestreo aleatorio simple (SRS). De hecho, los planes de muestreo de aceptación MNS dan como resultado curvas de OC que, en comparación con sus contrapartes de SRS, están mucho más cerca de la curva de OC ideal.

Redes neuronales y aprendizaje profundo(cap 18)

El aprendizaje profundo y las redes neuronales hacen parte de la inteligencia artificial o aprendizaje automatizado, esto es, que por medio de un conjunto de algoritmos se intente modelar abstracciones de alto nivel en datos usando arquitecturas compuestas de transformaciones no lineales múltiples.

Las redes neuronales como su nombre lo indica, trata de asemejarse a lo que vendría siendo un cerebro, conectando “neuronas artificiales” de manera que simulen el aprendizaje humano, como puede llegar a ser el reconocimiento de un rostro, la compresión de algún lenguaje o incluso el reconocimiento de voz.

Se puede aplicar en muchas áreas del conocimiento, por ejemplo, en la medicina, se podría usar este método para el reconocimiento de enfermedades en la piel por medio de una app para teléfonos inteligentes o en la agricultura, con la recolección de información por medio de drones para la toma de decisiones.

Bosques aleatorios e impulso(cap 17)

El concepto de bosques aleatorios surge con los árboles de decisión o diagramas de decisión, un árbol de decisión es un “mapa” de los posibles resultados de una serie de decisiones relacionadas, esto permite que una organización o individuo comparen diferentes acciones con el fin de maximizar la probabilidad de llegar al resultado esperado, esto puede ser por ejemplo conocer cuál es la ruta más óptima para llegar al trabajo o cuál va a ser la campaña y/o producto que me va a generar la mayor utilidad.

¿Qué sucede cuando tenemos gran cantidad de datos? Bueno, pasamos de tener un árbol de decisión a una infinidad de árboles, esto en función a la cantidad de datos y lo que queramos hacer con estos, así pues, pasamos a tener un bosque aleatorio.

Este método facilita en gran medida el proceso de toma de decisiones tanto en organizaciones como en proyectos personales, además, es muy utilizado en problemas de clasificación(clasificar en clientes buenos y malos, clasificar modelos de coches distintos, clasificación de muestras de ADN para determinar enfermedades,entre otros), arrojando muy buenos resultados.

Aprendizaje probabilístico - clasificación usando bayes ingenuo(cap 4)

En este tema, se utiliza la estadística bayesiana, más específicamente el teorema de Bayes, para tratar de predecir o clasificar, modelando la incertidumbre por medio de probabilidades y utilizando internamente pruebas de hipótesis.

Es un método muy utilizado en la actualidad en diversos contextos por su uso práctico y los buenos resultados en la mayoría de los casos. Por ejemplo, se puede utilizar para algo muy trivial, como clasificar una persona en hombre o mujer basándonos en las características de sus medidas: peso, altura y número de pie; pero, también se puede usar para problemas más complejos y muy útiles como por ejemplo, el que plantea el autor, filtrar el spam de los teléfonos celulares para eliminarlos automáticamente y que esto no se convierta en una molestia para los usuarios.

Con la proliferación de datos no estructurados, la clasificación de texto o la categorización de texto ha encontrado muchas aplicaciones en la clasificación de temas, análisis de sentimientos, identificación de autoría, detección de correo no deseado, etc.

Divide y vencerás - Clasificación usando reglas y arboles de decisión(cap 5)

La clasificación usando reglas y arboles de decisión, es un sistema de aprendizaje supervisado que aplica la estrategia “divide y vencerás” para hacer la clasificación, implementando métodos y técnicas para la realización de procesos inteligentes, representando así el conocimiento y el aprendizaje, con el propósito de automatizar tareas.

“Por su estructura son fáciles de comprender y analizar; su utilización cotidiana se puede dar en diagnósticos médicos, predicciones meteorológicas, controles de calidad, y otros problemas que necesiten de análisis de datos y toma de decisiones”, lo que quiere decir que los árboles de decisión pueden ser usados en cualquier ámbito sin importar que sea laboral o personal, siempre y cuando implique toma de decisiones con cierto grado de incertidumbre (Calancha Zuniga, Carrión Bárcena, Cori Vargas, & Villa Torres, 2010, pág. 2).

Es evidente que obtener un mecanismo como este sería muy útil debido a que podríamos predecir comportamientos futuros a partir de los comportamientos observados en el pasado.

Bibliografía



Bradley Efron and Trevor Hastie.

Computer age statistical inference. Algorithms, Evidence, and Data Science.
Cambridge, 2017.



Mohammad Jafari Jozani and Sayed Jamal Mirkamali.

Improved attribute acceptance sampling plans based on maxima nomination sampling.

Statistical Planning and Inference, 2010.



Brett Lantz.

Machine learning with R.

Packt Publishing, 2013.