

Practice 16: Reviews of my classmates.

Mayra Cristina Berrones Reyes 6291

December 15, 2020

1 Alberto M.

Inferencia y estadísticas bayesianas para la imputación de datos en datasets Los datos faltantes son problemas muy comunes encontrados en los datasets de la vida real, estos pueden perturbar el análisis de datos dado que disminuyen el tamaño de las muestras y en consecuencia la potencia de las pruebas de contraste de hipótesis, además hace que no se puedan utilizar directamente técnicas y modelos de machine learning, deep learning. Los anteriores nos lleva a la necesidad de rellenar o imputar datos en datasets, existen diversas técnicas para lograr este objetivo como la sustitución por la media, la sustitución por constante, imputación por regresión, entre otras. Dichas técnicas tienen ciertas deficiencias, por lo que en este trabajo se utilizará la estadística inferencial y bayesiana para la imputación flexible de datos faltantes en algunos datasets.

(Mayra) Este trabajo también me parece bastante interesante. Como mencionas, la parte de rellenar datos faltantes es muy importante cuando estas trabajando temas de minería de datos, sobre todo cuando la indicación general para tratar con datos faltantes es trabajar alrededor del problema o de plano eliminarlos de tu investigación. Me gustaría bastante ver que técnicas encuentras para poder subsanar este problema. Un pequeño detalle es nada más el uso que le das a la palabra de Imputación. La palabra rellenar creo que se entiende perfectamente. Imputación se refiere a algo más como dar la culpa de algo. No sé si querías usar más o menos ese significado.

2 Palafox.

Networks arise in many scientific and technological fields [Newman, 2018]. The internet, social networks, electrical networks, are among many available examples. To study network processes, sometimes it is convenient to have a model which preserves the essential characteristics of the network. A random graph is a model network in which the values of certain properties are fixed, but the network is in other respects random [Newman, 2018]. For example, a number n nodes and m edges could be fixed, but edges between any two nodes placed at random. The aim of this project is to do a theoretical and computational study of random graphs, and analyze how closely some of these resemble real world networks [Leskovec and Krevl, 2014].

(Mayra) La teoría de grafos es un tema que en lo personal me parece bastante interesante. Entiendo el enfoque que le quieres dar al final a tu tema, que es explicar de manera teórica el comportamiento de los grafos. Pero si estaría super interesante si, en caso de hacer este es el trabajo que quieras desarrollar, intentes con datos reales. Si no me equivoco, escuchamos algo de esto en una conferencia acerca de las redes eléctricas. También hay algunas aplicaciones en el

area de transporte.

3 Gabriela.

Comparación de soluciones: Como parte del trabajo de tesis se tienen datos sobre las soluciones obtenidas con dos formulaciones diferentes, se desea analizar dichas soluciones para verificar si hay diferencias significativas entre ambas. Como primera instancia se pretende usar estadística descriptiva y después verificar con pruebas de hipótesis que sean aplicables a los datos.

(Mayra) Esto es porque conozco tu tema, y si me gustaría bastante ver la diferencia que hay entre los dos modelos que tienes. Lo que no me queda muy claro es la parte en donde describes la primera instancia que vas a utilizar en tu experimento. Entiendo la parte en que vas a realizar estadística descriptiva. Lo demás me gustaría escucharte explicarlo.