

Tesis Montserrat

Juan Carlos Martinez Ovando

5 Junio 2019

Comentarios Generales

Capitulo 1

- El **capitulo 1** debe ser de introduccion a tu tesis, i.e. resumen del problema que estudiaras, el problema que motivo tal estudio, asi como una descripcion del contenido del documento.

Todo esto considerando un espacio de 4-8 paginas.

Capitulo 2

- El **capitulo 2** debe incluir una descripcion del problema de conglomeracion/clustering (basado en distancias o basado en modelos de probabilidad). En ambos casos, ahondar sobre el enfoque tradicional en que las **mediciones/atributos** son escalares, y el reto que representa realizar la **segmentacion/conglomeracion** combinando otras escalas de medicion: **conteos, categorias**, etc.

Capitulo 3

- El **capitulo 3** debe incluir una descripcion del paradigma bayesiano de inferencia y prediccion (como lo tienes actualmente en el capitulo 2 actual).
- La ecuacion 2.2 actual, enunciala en su forma parametrica, i.e.

$$p(x_1, \dots, x_n | \theta) = \prod_{i=1}^n p(x_i | \theta),$$

comentando que $p(x_i | \theta)$ puede ser una **funcion de masas de probabilidades** (caso discreto) o una **funcion de densidad de probabilidades** (caso absolutamente continuo).

- La especificacion de $P(X_i = x_i) = F(X_i = x_i | \theta)$ que describes en la siguiente pagina (pp.13) tendra validez solo considerando los casos que comento en el punto anterior.
- El **teorema de representacion** debe presentarse antes de la **verosimilitud** y antes del **teorema de Bayes**. Con el teorema de representacion establecido, puedes ahondar en los dos componentes que lo componen: a) **verosimilitud** y b) **prior**. con esos dos ingredientes puedes plantear la **actualizacion de informacion**, a.k.a. **aprendizaje estadistico** via el teorema de Bayes.
- La seccion actual 2.8.1. **Métodos de Aproximación** debe ser una seccion que siga a la de **prediccion**.

Capitulo 4

- El **capitulo 4** debe incluir la presentacion del modelo de clasificacion via mezclas para variables **continuas/discretas** que propones.
- El mismo capitulo debe describir el algoritmo MCMC implementado, particularmente las distribuciones condicionales completas

Capítulo 5

- El capítulo 5 debe reportar los resultados del análisis para tu base de datos.

Capítulo 6

- El capítulo 6 debe incluir conclusiones de la tesis.

Apendice A

- En el apéndice A describe el análisis conjugado bayesiano, necesario para el cálculo de las distribuciones condicionales completas del capítulo 4/

Apendice B

- El apéndice B puede ser una descripción de las funciones en R que desarrollaste y que distribuyes en Github.