



FACULTAD DE MATEMÁTICAS
PONTIFICIA UNIVERSIDAD CATÓLICA DE CHILE

Tarea 1 - EYP3407

Profesor: Mauricio Castro
Ayudante: Bladimir Valerio

1. (30%) Sea Z una variable aleatoria con función de densidad dada por

$$f(z | \mu, \sigma^2, \alpha) = \frac{2}{\sigma} \phi\left(\frac{z - \mu}{\sigma}\right) \Phi\left(\alpha \left(\frac{z - \mu}{\sigma}\right)\right), z \in \mathbb{R},$$

donde $\phi(\cdot)$ y $\Phi(\cdot)$ representan la función de densidad y la función distribución, respectivamente, de una variable aleatoria normal estándar. La notación utilizada para referirnos a la distribución de Z será $Z \sim SN(\mu, \sigma^2, \alpha)$. Note que cuando $\alpha = 0$, $Z \sim N(\mu, \sigma^2)$. Muestre que la entropía $H_{SN(\mu, \sigma^2, \alpha)}$ de Z , está dada por

$$H_{SN(\mu, \sigma^2, \alpha)} = H_{N(\mu, \sigma^2)} - E_{X_0} [\ln(2\Phi(\alpha X_0))],$$

donde $X_0 \sim SN(\alpha)$ y $H_{N(\mu, \sigma^2)}$ es la entropía cuando $Z \sim N(\mu, \sigma^2)$. Grafique la entropía $H_{SN(\mu=0, \sigma^2=1, \alpha)}$ para diferentes una grilla de valores de α entre -20 y 20 . ¿Qué puede concluir de lo anterior?

2. (70%) En el GitHub <https://github.com/caserec/Datasets-for-Recommender-Systems>. git creado por Arthur F. da Costa Ph.D. del Instituto de Matemática y Ciencias de la Computación de la Universidad de Sao Paulo se encuentran interesantes set de datos utilizados para generar sistemas de recomendación. Cada equipo deberá elegir uno (o más) conjunto(s) de datos en los cuales pueda implementar al menos dos métodos de recomendación (pueden ser vistos o no vistos en clase, y pueden ser implementados con set de datos diferentes). Para cada caso (si corresponde) se debe considerar:
- (a) (10%) Descripción y objetivo del problema.
 - (b) (10%) Análisis exploratorio de los datos
 - (c) (10%) Descripción del método utilizado.
 - (d) (30%) Implementación computacional.
 - (e) (30%) Resultados obtenidos.
 - (f) (10%) Medida de performance.

Responda además las siguientes preguntas:

- (a) ¿Cuál es en su opinión el mejor sistema de recomendación? Discutir.
- (b) ¿Los métodos considerados son escalables? Comente

Aspectos a evaluar en la presentación oral (si corresponde): Cada grupo deberá presentar sus principales hallazgos en una presentación de 10 minutos. Solo se evaluará la presentación (calidad, duración, claridad de la exposición, contenido) y si el código funciona o no en vivo (programación).

Grupos

	Tarea 1	
Grupo 1	Alonso Antonio Campos López	Isidora Vanessa Cariola Wernar
Grupo 2	Andrés David Díaz Carrasco	Camila Andrea Echeverría Cabeza
Grupo 3	Axel Alfred González Espinoza	Lucas Benjamín Gutiérrez Azócar
Grupo 4	Camila Zahira Villagrán Pinto	Tomás Gabriel Valdenegro Cid
Grupo 5	Diego Adolfo Aravena Morales	Felipe Ignacio Vera Ojeda
Grupo 6	Diego Ignacio Gallo Romero	Sebastián Nicolás Celaya Vergara
Grupo 7	Fabián Esteban Sepúlveda Rivas	Josefa Silva Muñoz
Grupo 8	Francisca Fernanda Vilca Sánchez	Javiera Alejandra Valenzuela Soto
Grupo 9	Francisco Andrés Vega Mac-Namara	Pedro Varela Reyes
Grupo 10	Iván Alberto Toledo Andrade	Sebastián Andrés Figueroa Espina
Grupo 11	Javiera Victoria Lagos Turenne	Paolo Lorenzo Ciani Rozas
Grupo 12	María José Gutiérrez Rivera	Mariella Catalina Tognarelli Buono-Core

Figure 1: Grupos para la T1