
Proyecto ICD

Wilmer González

2 de agosto de 2015

Fecha de entrega11 de agosto de 2015

1. Objetivo

La clusterización de personas ha sido una de las tareas mas comunes cuando los datos a caracterizar son generados en redes sociales. Los post de facebook, por ejemplo, pueden darnos información sobre la semejanza de los amigos de diferentes individuos, en esta oportunidad le será provisto un conjunto de post de los amigos de sus compañeros con el que se le solicita agruparlos de acuerdo a la similitud de los post de sus amigos.

2. Data

Nombre del archivo	Formato	Uso
CI1_CI2_facebook_posts	.csv	Posts tomados del API de facebook Sin preprocesar.
CI1_CI2_facebook_usuarios-grupos	.csv	Usuarios y las palabras que los relacionan con los otros usuarios

3. Actividades

1. Realice las tareas de pre-procesamiento necesarias para mejorar e identificar las palabras contenidas en cada post.
2. para cada individuo identifique las palabras que los relacionan con los otros.
3. Aplique **algun método de clasificación** para agrupar a los usuarios según las palabras que elija.(**justifique en el informe**).
4. Genere un nuevo data frame en el que almacene una matriz individuos* individuos y cada entrada son las palabras que los relacionan, luego guardelo en un nuevo archivo con el nombre CI1_CI2_facebook_usuarios_grupos.csv
5. Defina las relaciones que considere mas importantes.

4. Formato de entrega

Adjunto del correo:

- T1_CI1_CI2_CI3:
 - CI1_CI2_facebook_posts.csv
 - CI1_CI2_facebook_usuarios_grupos.csv
 - source.R
 - inform [.Rmd — .pdf]