

Ahora vamos a estudiar el siguiente tema en mi programa, que es Minería de opinión y detección de polaridad y sentimientos.

Primero estudien por favor las páginas 410-421 del PDF del libro “Text Data Management and Analysis” de ChangXiang Zhai y Sean Massung.

Luego descarguen de <http://www.lsi.us.es/~fermin/index.php/Datasets> el primer dataset que se llama Spanish Movie Reviews y investiguen los archivos del corpus para comprender su estructura y el contenido.

En las páginas mencionadas se explican los conceptos básicos de la tarea de minería de opinión y también el método que se llama Regresión logística ordinal. Este método vamos a usar implementado en el paquete de Python que se llama “mord”. Por favor busquen la documentación que corresponde a Regresión logística ordinal y estúdienla.

Por lo pronto es todo, ahora estamos estudiando la teoría. Luego vamos a aplicar la Regresión logística ordinal al corpus de críticas de cine que será la tarea para el próximo martes 19 de mayo. La tarea explicaré mañana.