

Proyecto FINAL

Descripción

1. Elegir alguno de los corpus localizados en Teams en la carpeta de Proyecto. Encontrarán ahí documentos para entrenamiento y pruebas o en su defecto tome el 80 por ciento de datos para entrenar y el 20 para probar.
2. Preprocesen los documentos para obtener la representación vectorial de los mismos.
3. Utilicen Weka o Python para clasificar los textos, experimenten clasificando con tf con tf-idf y reduciendo el vocabulario
4. Generen su reporte en el mismo formato de las prácticas reportando la exactitud, precisión, recuerdo y medida F obtenidas.