

Conclusion

En este TP pudimos usar herramientas mas avanzadas para construir modelos a diferencia del TP anterior.

Hicimos un analisis mas avanzado del dataset usando de base el que obtuvimos en el TP1, nos centramos mas que nada en la columna de descripcion de la propiedad ya que con las nuevas herramientas de procesamiento de lenguaje natural, era el analisis que nos faltaba.

Cabe destacar que la columna de descripciones de las propiedades era bastante fea y necesito normalizacion y limpieza para poder ser usada.

Se vio un aumento de accuracy en el caso de las regresiones en comparacion al TP anterior, se lo atribuimos en mayor parte al metodo de Voting y otro poco a las nuevas columnas que salieron del analisis de lenguaje natural.

Y en el caso de la regresion, la mejora fue bastante grande, con un RMSE mucho menor en el caso de Stacking, tambien atribuimos la mejora al nuevo modelo y a su enfoque para generar los resultados.

Nos gustaria haber podido aprovechado mas la extraccion de informacion para generar mas columnas de atributos de cada propiedad, tambien quedaron totalmente fuera del TP2 los algoritmos geneticos, lexicones y analisis de sentimientos.