

GESTION DEL CONOCIMIENTO E INVESTIGACION AMBIENTAL

Entidad: CORPORACIÓN AUTONOMA REGIONAL para la defensa de la meseta de Bucaramanga.

Departamento: Santander

Municipio: Bucaramanga

Orden : Territorial

Sector: Ambiente y Desarrollo Sostenible, Estaciones Hidrometeorológicas durante el año 2021.

AUTOR: HUGO FORERO GUERRA

RESUMEN

PROBLEMA: La Corporación Autónoma regional de Santander, en aras de aportar información para la investigación sobre el medio ambiente, busca la manera de procesar y obtener resultados fiables de los datos recolectados durante el año 2021 en las estaciones Hidrometeorológicas de la meseta de Bucaramanga.

DESCRIPCION: Los datos son suministrados por la página Datos Abiertos Colombia- www.datos.go.co. Aplicaremos el modelo Machine Learning con regresión lineal múltiple y Random Forest, la variable a predecir es "Index" y las variables predictoras son, const, out, Hum, Speed, Bar, Rain, Rad.

RESULTADOS:

La clasificación de las variables predictoras por orden de importancia

```
[66] important.sort_values(by="importance")
```

	variable	importance
1	Hum	-0.002996
4	Rain	-0.002469
3	Bar	0.000444
5	Rad,	0.005339
0	Out	0.008898
2	Speed	0.762835

Los resultados

score_LR is 0.4432404790534472

score_LR2 is 0.44179345637846545

RMS2_SM is 1.6999526929794797

Predicción: score_LR is 0.4432404790534472

RMS2_SM is 1.6999526929794797

RMS2_LR is 1.6999755252906628

RMS2_RF is 0.503786313770233

Los valores no presentaron variaciones con los dos métodos, dando una mejor claridad con Random Forest.