DORITALAND.SL

Salary Prediction

Judit Roca

COMPANY GROUP

Table of Content

	work_year	experience_level	employment_type	job_title	salary	salary_currency	salary_in_usd	employee_residence	remote_ratio	company_location	company_size
0	2023	SE	FT	Principal Data Scientist	80000	EUR	85847	ES	100	ES	L
1	2023	MI	CT	ML Engineer	30000	USD	30000	US	100	US	S
2	2023	MI	CT	ML Engineer	25500	USD	25500	US	100	US	S
3	2023	SE	FT	Data Scientist	175000	USD	175000	CA	100	CA	M
4	2023	SE	FT	Data Scientist	120000	USD	120000	CA	100	CA	M

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 3755 entries, 0 to 3754
Data columns (total 11 columns):
 #   Column              Non-Null Count  Dtype
---  ------              --------------  -----
 0   work_year           3755 non-null   int64
 1   experience_level    3755 non-null   object
 2   employment_type     3755 non-null   object
 3   job_title           3755 non-null   object
 4   salary              3755 non-null   int64
 5   salary_currency     3755 non-null   object
 6   salary_in_usd       3755 non-null   int64
 7   employee_residence  3755 non-null   object
 8   remote_ratio        3755 non-null   int64
 9   company_location    3755 non-null   object
 10  company_size        3755 non-null   object
dtypes: int64(4), object(7)
memory usage: 322.8+ KB
```

EUR

valor
$-€

SIZE

S
M
L

PAIS

Salario
medio

EXP

EN: 1
MI: 2
SE: 3
EX: 4

JOB

Scientist -Data Science

cloud

Data Analyst/Analytics
Data Engineer/Strategist

Machine Learning/ ML
AI / MLOps

Head

DORITALAND.SL

Regresión lineal

Polynomial Regression

Random Forest Regressor (GS)

XG_Boost

Pipeline

Linear Regresion

Random Forest

XG_Boost

Scaler

GS

PCA

	Linear	Poly	RForest	GXBoost	Pipe	PCA
Score	0.33	10696613916437	0.35	0.32	0.37	0.35
MAE	35590	15558563751	34166	35473	34132	34547

ESTIMATOR: XGBREGRESSOR

PARAM_GRID:

BASE_SCORE=0.5
BOOSTER='GBTREE'
LEARNING_RATE=0.1
MAX_DEPTH=3
N_ESTIMATORS=50
REG_LAMBDA=1
SCORING: NEG_MEAN_ABSOLUTE_ERROR

NEXT STEPS

Streamlit

Nuevos data engineering

Nueva recolección de data

GRACIAS
POR SU ATENCION