Web Mining System Architecture - v1

Raw Data → Local File System → Preprocess → Models → Buckets → UI

Container-1: Classification Component

class
ClassificationTask(TaskType)

Container-2: Regression Component

class
RegressionTask(TaskType)

Container-3: Clustering Component

class
RegressionTask(TaskType)

Kaggle
Body
Performance
Dataset

data/
raw/bodyPerformans.csv
preprocessed/preprocessed_data.csv

Preprocess Class
-p Classification
-p Regression
-p Clustering

Model Artifacts

models/
classification_models/
m1.pickle
m2.pickle
m3.pickle
regression_model/
m4.pickle
clustering_model/
m5.pickle

Cloud

app.py

Langs: Python
UI: StreamLit App
Platform: Local / Docker / UI (StreamLit App)