Clarity on dataset

Data cleaning /preprocessing

Correlation analysis

Clustering

Log transformation

Feature engineering

Feature scaling

Model xgboost evaluation why

Anomaly stats

Lof isolation evaluation then hybrid approach risk score why hybrid evaluation

Literature review

Process flow for ML model