PRD: Indonesian Oil & Gas Fields Indicative Economic Value Predictions

I. PROBLEM ALIGNMENT

A. The Problem

When oil companies discover a field, they have to make an early economic calculation to determine the future of the discovered field, either to be developed or abandoned. Thus, they need a model that can calculate the preliminary economic (economic scale) in a short time before handing over the field to the development team to gain significant information before designing the development concept.

Indonesia has 100 years of oil and gas field data and with the fast computing engine and big data analytics algorithm, we can utilize the data to make predictions on the future commercial of the field.

B. High Level Approach

1. Using reservoir properties data such as porosity, permeability, gas or oil saturation, and the size of the field (hydrocarbon pore volume) a.k.a hydrocarbon initial in place (HCIP) to calculate the possible recoverable oil or gas as representation of field subsurface data
2. Using field environment setting (onshore or offshore)
3. Using cost and NPV data of developed field as economic data
4. Make model from field subsurface, field setting, and economic data (train the model)
5. Make recoverable oil or gas prediction and the economics for the newly discovered field

C. Goal & Success

Providing benchmarking tools for Oil & Gas participants to get early indicative economic value such as profitability index (PI) of newly discovered oil and gas fields. This information is crucial to enhance the field development concept from a technical and economic point of view.

II. SOLUTION ALIGNMENT

A. Key Solution

We will use eSDC dataset, consist of 17 features, we will build the model using Linear regression, Logistic Regression, Trees Method, SVM, clustering and dimensionality reduction as needed

B. Key Flows

We will use python ML libraries such as sklearn and tensorflow and isolate them by using venv to preserve the version and compatibility. eESDC dataset are ready to use subsurface dataset so during this work we could minimize data cleaning jobs and focus on feature engineering and modeling to obtain the best

optimized model. On the modeling part, we will generate the least effort model as a baseline such as linear regression and continue the experiment on top of that. The iteration/ experiment process will be conducted until the model could deliver sufficient results represented as an optimum tradeoff between accuracy and utilization.

III. LAUNCH READINESS

A. Key Milestones

Date	**Milestones**	**Description**
Sun, 12 June 2022	Data Understanding	EDA & Pre-Processing
Sun, 19 June 2022	Feature Update	Feature Selection & Data Ready for Modeling
Sun, 26 June 2022	Experiment with Models	Baseline & Promising Candidates
Sun, 03 July 2022	Modeling	Expected Matrics Results: ~93%
Sun, 24 July 2022	Serving Ready	Backend and model image is ready to deploy

B. Artifact

Artifacts	**Where to Check?**
Dataset Final	eSDC
Project Milestone	Github & Google Docs

References

1. https://github.com/aprds/field_profitability_index
2. https://docs.google.com/document/d/1P7m6gwpmw3LaD39ss0641h_H6m3UjYW2KkZCm_GpUu0/edit