# Project Proposal

CSCI 470: Introduction to Machine Learning

**OilDigger**

Nadima Dwihusna

Xiaoyu (Rosie) Zhu

Mohamed Ibrahim Mohamed

**Preference:** High

**Topic Area:** Geophysics

**Project Name:** Mapping Geological Formation Using Well Logs.

**Problem Statement:** There are numerous well logs data that are collected but are not being used to map geologic formation. Instead of relying on one source of data (seismic), geophysicists can make use and take advantage of these acquired well log data to map the formation layers. Currently interpreting formation layers from numerous well logs data is a tedious method. Using a Machine Learning algorithm will fasten the process, improve geologic formation layer interpretation, and improve data-driven decision.

**Proposed Solution:** We will use several Machine Learning Algorithms such as Random Forest, Support Vector Machines, and K-Nearest Neighbors, then pick the best model for this problem. First, utilize part of the collected well logs in one region to train the model, test the model with the other wells. Then, predict the formations based on the model and map the formation layers. This is followed by comparing the mapped formation layer results with the Seismic Data. Subsequently, weight which logs and data sets are more accurate and provide a close match with the seismic, using feature_importance from Random Forest. Our solution also considers the uncertainties and error within each well log. The process is then repeated using the top picked or highly scored well logs. The final predicted map will then be compared with the seismic data.

**Data:** Well log data at measured depth including: Formation Density, Gamma Log, Resistivity, Caliper, Neutron Porosity, Spontaneous potential, and Seismic data.

**Timeline**

| Week Starting | Nadima Dwihusna | Xiaoyu (Rosie) Zhu | Mohamed Mohamed |
|---|---|---|---|
| 9/17 | Obtain Data Sets | Obtain Data Sets and Research | Research on ML Methods |
| 9/24 | Data Processing | Data Visualization | Research on ML Methods |
| 10/1 | Data Processing | Data Visualization | Research on ML Methods |
| 10/8 | Data Processing | Data Visualization | ML Method Selection |
| 10/15 | | All: Progress Report (Due Oct18) | |
| 10/22 | Data Training | Data Prediction | Help Team Members |
| 10/29 | Data Training | Data Prediction | Help Team Members |
| 11/5 | ML Method 1 | ML Method 2 | ML Method 3 |
| 11/12 | ML Method 1 | ML Method 2 | ML Method 3 |
| 11/19 | ML Method 1 | ML Method 2 | ML Method 3 |
| 11/26 | | All: Visualize Results with Seismic | |
| 12/3 | | All: Prepare Final Presentation | |

**Preference:** Medium

**Topic Area:** Petroleum

**Project Name:** Identifying Sweet Spots Using Well Logs.

**Problem Statement:** Utilizing well logs to identify the formation bearing fluids (Oil, Gas, and Water) has been common in the oil and gas industry. However, this has been done manually for the past decades. Manual identification of sweet spots can result in many errors as human or user errors and moreover manual identification overlooks the instrumental errors, environmental disturbance and units inconsistency. .

**Proposed Solution:** In this project we are proposing several machine learning classification algorithms that can identify the different formation fluid using well logs. This project will use classification algorithms (Random Forest and Knn) to predict the fluid type (Water, Oil, and gas) from the well log data. It returns the most promising depth/layers the geophysicists should later focus on. Then we will choose the best algorithm based on their performance. We also analyze which well log is more important to be used in the random forest.

**Data:** Well log data at measured depth including: Formation Density, Gamma Log, Resistivity, Caliper, Neutron Porosity, Spontaneous potential, and Seismic data.

**Timeline**

| Week Starting | Nadima Dwihusna | Xiaoyu (Rosie) Zhu | Mohamed Mohamed |
|---|---|---|---|
| 9/17 | Obtain Data Sets | Obtain Data Sets and Research | Research on ML Methods |
| 9/24 | Data Processing | Data Visualization | Research on ML Methods |
| 10/1 | Data Processing | Data Visualization | Research on ML Methods |
| 10/8 | Data Processing | Data Visualization | ML Method Selection |
| 10/15 | | All: Progress Report (Due Oct18) | |
| 10/22 | Data Training | Data Prediction | Help Team Members |
| 10/29 | Data Training | Data Prediction | Help Team Members |
| 11/5 | ML Method 1 | ML Method 2 | ML Method 3 |
| 11/12 | ML Method 1 | ML Method 2 | ML Method 3 |
| 11/19 | ML Method 1 | ML Method 2 | ML Method 3 |
| 11/26 | | All: Visualize Results with Seismic | |
| 12/3 | | All: Prepare Final Presentation | |

**Preference:** Low

**Topic Area:** Petroleum and Geophysics

**Project Name:** Multidisciplinary Approach for Mapping Geological Formation and Detecting Sweet Spots Using Well Logs.

**Problem Statement:** The most productive areas of a reservoir can be accurately identify by combining a multidisciplinary approach of identifying the geological formation and formation bearing fluids or sweet spots (Oil, Gas, Water). Currently well logs are manually interpreted and the programming process to identify sweet spots is time consuming. .

**Proposed Solution:** We will identify the geological formation and formation bearing fluids using well logs data by implementing machine learning. The models of machine learning that will be utilized are Random Forest, Support Vector Machines, and K-Nearest Neighbors. We will choose the best algorithm based on the performance. The current programming process for formation evaluation is time consuming. Stratigraphers manually interpret geological formation and formation bearing fluids to identify sweet spots. A machine learning approach allows us to improve efficiency. The algorithm will input well logs (weighted with uncertainties) and outputs geological formation and sweet spots characterizing the reservoir.

**Data:** Well log data at measured depth including: Formation Density, Gamma Log, Resistivity, Caliper, Neutron Porosity, Spontaneous potential, and Seismic data.

**Timeline**

| Week Starting | Nadima Dwihusna | Xiaoyu (Rosie) Zhu | Mohamed Mohamed |
|---|---|---|---|
| 9/17 | Obtain Data Sets | Obtain Data Sets and Research | Research on ML Methods |
| 9/24 | Data Processing | Data Visualization | Research on ML Methods |
| 10/1 | Data Processing | Data Visualization | Research on ML Methods |
| 10/8 | Data Processing | Data Visualization | ML Method Selection |
| 10/15 | All: Progress Report (Due Oct18) | | |
| 10/22 | Data Training | Data Prediction | Help Team Members |
| 10/29 | Data Training | Data Prediction | Help Team Members |
| 11/5 | ML Method 1 | ML Method 2 | ML Method 3 |
| 11/12 | ML Method 1 | ML Method 2 | ML Method 3 |
| 11/19 | ML Method 1 | ML Method 2 | ML Method 3 |
| 11/26 | All: Visualize Results with Seismic | | |
| 12/3 | All: Prepare Final Presentation | | |