VHydro - Hydrocarbon Potential Prediction App: User Tutorial

This tutorial will guide you through using the VHydro application to analyze well log data and predict hydrocarbon potential zones using Graph Convolutional Networks.

Getting Started

Installation

1. First, install the required dependencies:

bash

Copy

```bash
pip install streamlit pandas numpy matplotlib seaborn scikit-learn lasio plotly networkx
```

2. Download the application files:
   - `app.py` (Main Streamlit application)
   - Any test LAS files you want to use
3. Run the application:

bash

Copy

```bash
streamlit run app.py
```

Application Workflow

The VHydro application follows a step-by-step workflow to analyze well log data and predict hydrocarbon potential:

1. Home Page

When you first launch the application, you'll see the home page with an overview of the VHydro methodology and the application workflow. Use the navigation sidebar to proceed to the first step.

2. Data Upload & Processing

On this page, you can upload your LAS file containing well log data. The application will process the file and calculate petrophysical properties such as:

- Shale volume
- Porosity

- Water/oil saturation
- Permeability

Required curves in your LAS file:

- GR (Gamma Ray)
- RHOB (Bulk Density)
- ILD (Deep Resistivity)
- Additional curves like CGR (Computed Gamma Ray) are beneficial but not required

After uploading, the application will:

1. Display the processed data
2. Show a visualization of the well log curves
3. Allow you to proceed to the next step

3. Graph Dataset Creation

This page converts the well log data into a graph dataset for the GCN model:

1. **Set parameters:**
   - Number of clusters (5-10): Controls the facies identification
   - Train/Test Split: Determines the proportion of data for training
2. **Generate the graph dataset:**
   - The application will perform K-means clustering to identify facies
   - Create nodes representing depth points
   - Define edges connecting related depth points
   - Assign PET (Petrophysical Entity) labels based on the calculated properties
3. **Review the results:**
   - View the clustered well log visualization
   - Examine samples of the node and edge data
   - Proceed to graph visualization

4. Graph Visualization

This page provides visualizations of the graph dataset:

1. **Interactive graph visualization:**
   - Nodes represent depth points

- Colors represent hydrocarbon potential classifications
- Connections show relationships between depth points

2. **Additional visualizations:**
   - PET label distribution
   - Feature importance for classification
   - Correlation between petrophysical properties and hydrocarbon potential

5. Model Prediction

This page uses a Graph Convolutional Network to predict hydrocarbon potential:

1. **Configure the model:**
   - Number of GCN layers
   - Learning rate
   - Training epochs
2. **Run the prediction:**
   - The application will train the GCN model on the graph dataset
   - Generate predictions for hydrocarbon potential
   - Display the prediction results and model accuracy

6. Results & Interpretation

The final page provides detailed visualizations and interpretation of the results:

1. **Potential Distribution tab:**
   - Shows the distribution of hydrocarbon potential classifications
   - Compares true and predicted distributions
2. **Depth Analysis tab:**
   - Visualizes hydrocarbon potential vs. depth
   - Identifies zones with high potential
3. **Well Log with Predictions tab:**
   - Combines well log data with prediction results
   - Shows the relationship between log curves and hydrocarbon potential
4. **Summary and Recommendations:**
   - Overall assessment of hydrocarbon potential

- Recommendations for further exploration

5. **Export Results:**
   - Download prediction results as CSV
   - Export visualization report

Example Analysis

Here's an example of how to analyze a well using the VHydro application:

1. **Upload a LAS file** containing well log data for the well of interest.
2. **Create a graph dataset** with 7 clusters and an 80% train/test split.
3. **Visualize the graph** to understand the relationships between different depth points and their potential hydrocarbon classifications.
4. **Run the GCN model** with 2 GCN layers, a learning rate of 0.01, and 100 training epochs.
5. **Interpret the results** to identify zones with high hydrocarbon potential and make recommendations for further exploration.

Troubleshooting

- **Missing curves error:** Ensure your LAS file contains the required curves (GR, RHOB, ILD).
- **Graph visualization not loading:** Try reducing the number of nodes by setting a higher sampling rate.
- **Low model accuracy:** Experiment with different numbers of clusters and model parameters.

Additional Resources

- Research Paper: Novel Graph Dataset for Hydrocarbon Potential Prediction
- Streamlit Documentation
- Graph Convolutional Networks Documentation
- LASio Documentation

Understanding the VHydro Graph Dataset

The VHydro approach uses a graph-based representation of well log data, which is a key innovation compared to traditional methods:

1. **Nodes**: Each node represents a depth point in the well.
2. **Edges**: Edges represent connections between depth points that share similar petrophysical properties.

3. **Node Features**: Each node has features derived from well log data (e.g., porosity, permeability, saturation).
4. **Petrophysical Entity (PET) Labels**: Binary features that represent the presence or absence of specific petrophysical properties at each depth point.

This graph representation captures the spatial relationships and property distributions throughout the well, enabling the GCN model to identify patterns that indicate hydrocarbon potential.

Understanding Prediction Results

The application classifies hydrocarbon potential into five categories:

1. **Very Low**: Areas with poor reservoir properties and low hydrocarbon indicators.
2. **Low**: Areas with marginal reservoir properties.
3. **Moderate**: Areas with acceptable reservoir properties.
4. **High**: Areas with good reservoir properties and hydrocarbon indicators.
5. **Very High**: Areas with excellent reservoir properties and strong hydrocarbon indicators.

When interpreting results, consider:

- **Vertical Distribution**: Look for continuous zones of high potential rather than isolated points.
- **Correlation with Lithology**: High potential zones should align with appropriate lithology.
- **Consistency with Other Indicators**: Results should be consistent with other hydrocarbon indicators.

Example Case Study

A case study from the McKee Field showed that:

- Zones with high hydrocarbon potential correlated strongly with known productive intervals.
- The GCN model achieved 87% accuracy in predicting hydrocarbon potential.
- The graph-based approach outperformed traditional machine learning methods by 15%.

This demonstrates the effectiveness of the VHydro approach for identifying hydrocarbon potential zones in well log data.