


Applying the Unsupervised Isolation Forest (IF) technique to detect spikes in petrophysical well log data

Background

In Petrophysical Data Quality Control (QC), "spikes" refer to abrupt, sharp changes in data values that stand out from the surrounding dataset. These anomalies can arise due to measurement errors, logging tool malfunctions, data transmission issues, or genuine but abrupt changes in subsurface properties such as porosity, permeability, or fluid saturation. Detecting these spikes is critical, as they can distort the interpretation of subsurface properties if left unaddressed.

For instance, in well logging data like Gamma Ray (GR) or Resistivity Logs, spikes manifest as sudden jumps in readings that deviate from the general formation trend or are inconsistent with neighboring data points. Identifying and flagging these anomalies is a key step during QC before any corrections are applied.

By leveraging the Unsupervised Isolation Forest (IF) technique, we can automate spike detection in a structured, efficient, and accurate manner. This approach not only accelerates the QC process but also reduces manual errors, providing a reliable solution for anomaly detection.

To explore how this technique works in practice, check out my sample spike detection utility available on GitHub  [LogSpikeDetection_IsoForest](#)

It is freely available under the MIT license. Feel free to give it a try, and let's improve the reliability of petrophysical data together!

How to use the LogSpikeDetection Script & Requirements

1. LAS File Reading: The `load_las_file` function uses `lasio` to extract depth and gamma-ray curves.
2. Interactive Input: Prompts the user to specify LAS file paths interactively.
3. Error Handling: Provides meaningful error messages if a file cannot be processed or lacks required data.

You can install the `lasio` library if it's not already installed using the following command:

```
> pip install lasio
```

This script will handle single or multiple LAS files and plot the Gamma Ray logs with detected spikes for each file.

