SeismoScan

NAU NORTHERN ARIZONA UNIVERSITY
School of Informatics, Computing, and Cyber Systems

Team: Noah Carges, Alyssa Sombrero, Thomas Rotchford, and Brady Wisniewski
Client: Donna Shillington, James Gaherty, and Christine Regalla
NAU School of Earth and Sustainability

Team Mentor: Scott LaRocca

Motivation

Mapping faults on the ocean floor is vital for geologists to better understand Earth processes.

Fault locations allow geologists and seismologists research the following:

- Deep Water Cycle
- Earthquakes
- Tsunamis

However, the traditional method of hand mapping faults took **80** hours to cover an area of ~40,000 square kilometers.

We have developed a machine learning model that dramatically reduces the time geologists spend manually mapping faults.

Key Features

- **Clustering:** Grouping lat long points with based on seafloor slope
- **Merging:** A process which combines connected points from Clustering into an object
- **Random Forest:** Making a decision on weather or not an object is a fault.

Challenges

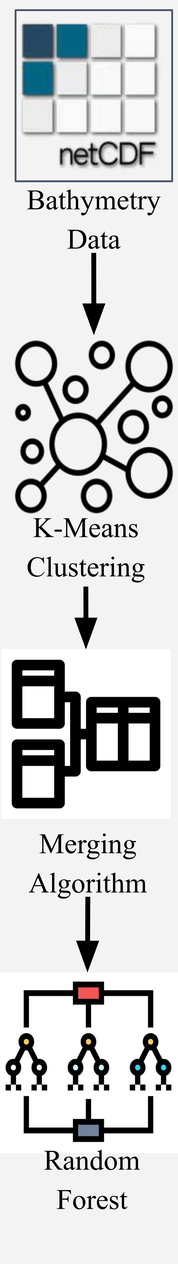
- Training a machine learning model with limited labeled geological data.
- Differentiating natural seafloor variants (sonar issues, sea vents) from true faults.
- Managing & processing large bathymetry datasets.
- Building a model that can analyze diverse regions across the world

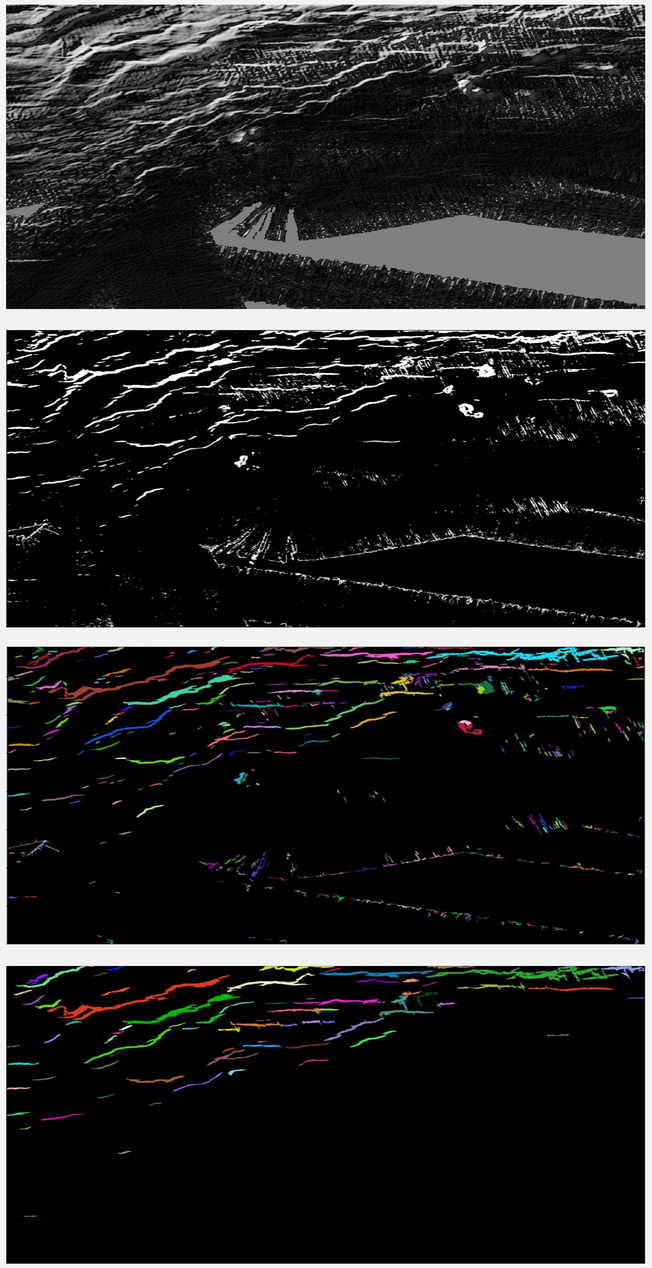
Overview

Developing machine learning tools to map faults within bathymetry data

- **Bathymetry Data:** Topography data specific to the sea floor.
- **Faults:** Bends in the sea floor due to tectonic plate movement.

Solution





Outcomes

We have:

- Constructed a complete data processing system to extract all necessary features
- Built a Machine Learning Pipeline that analyzes a dataset to detect faults
- Implemented expansive configuration options to allow users to tune the process to any dataset

Technologies Used

Programming Language: Python
Data Format: NetCDF 4
Primary Libraries/Frameworks:

- PyGMT & Matplotlib
- NumPy & Xarray
- Scikit-learn

Environment: Linux Windows Windows 11
Version Control: Github

Future Work

- Optimize performance for large-scale datasets.
- Expand model training with more diverse datasets and preprocessing pipeline in order to improve accuracy.
- Testing with various configuration values to obtain the best output that has less false positives.

We feel confident that our product that will improve geological research related to faults in a significant way