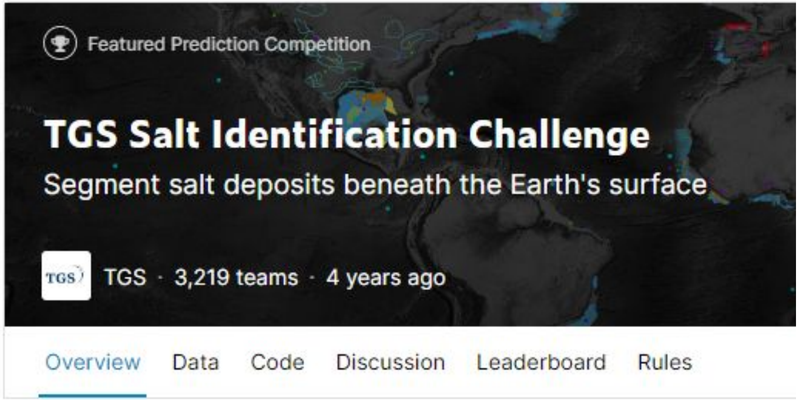
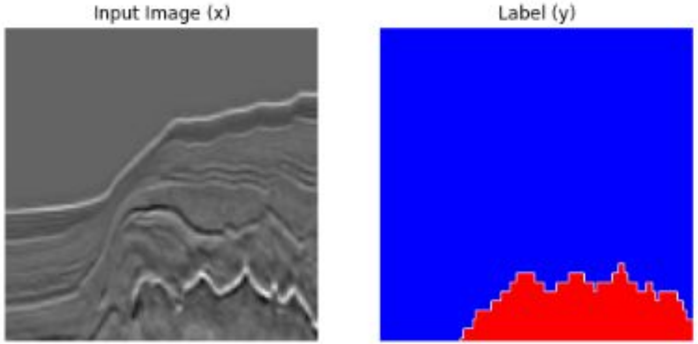
Explainable-Christmas

Marcos Jacinto - Madrid - Geowellex
Edwin Brown - UK - Optic Earth

Intro to Idea



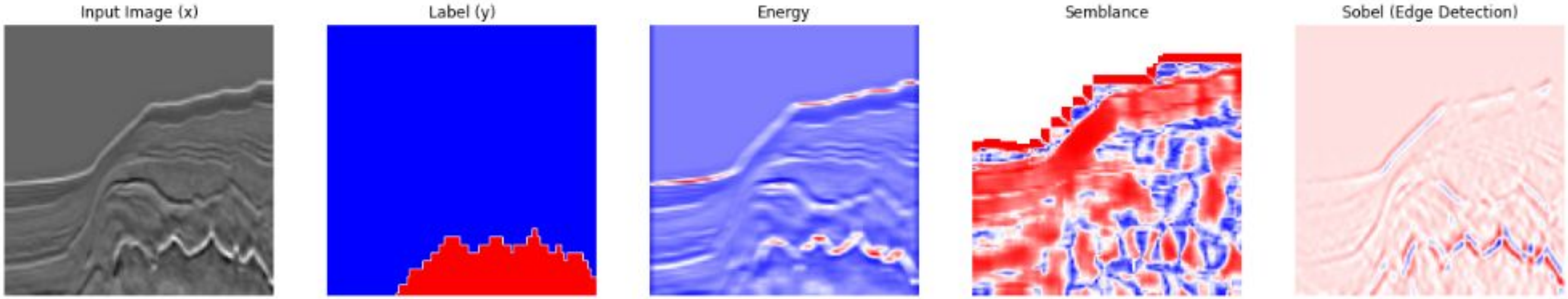


We used the **TGS Salt Identification Challenge data** along with a **U-Net model** trained from scratch using Tensorflow's library.

Seismic attributes were created from the seismic data.

Finally, we used **Shap library** to **calculate SHAP values** and obtain a measure of how important each feature (seismic data and attributes) is.

Attribute Generation



A set of **seismic attributes** were calculated from the original seismic section using the **Bruges library.**

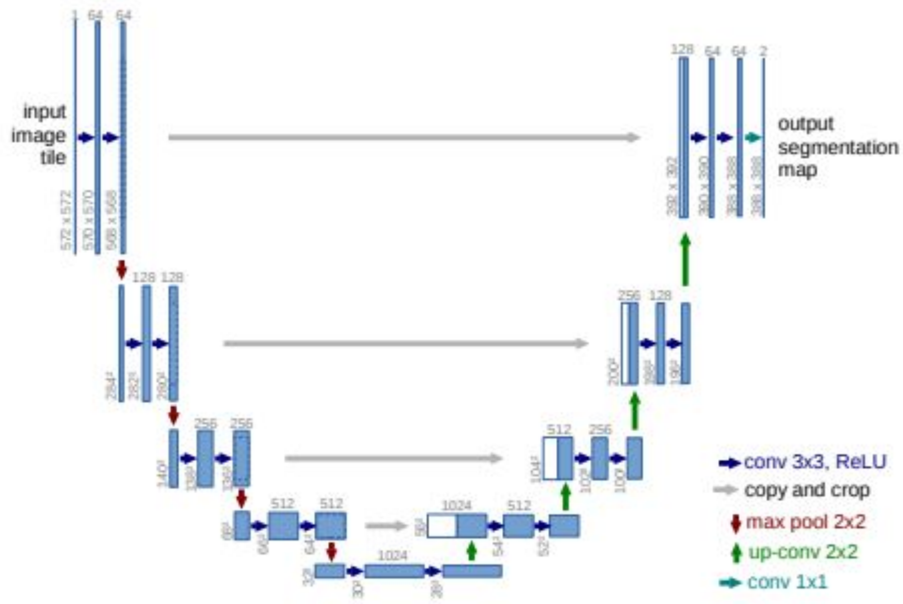
Salt often has a different seismic character compared with other type of geological structures. Therefore, seismic attributes could offer some useful input features.

Model Summary

We used a **U-Net model** trained from scratch using Tensorflow's library.

During training, we used two NVIDIA GPUs* from the provided cluster.

*Thanks NVIDIA!



Ronnerberger et al. (2015)

Model Training

mlflow 1.26.1 Experiments Models GitHub Docs

Default > Run 19cc549a15b343b2bf7d8eb2068782bb > val_auc

val_auc

Validation AUC: 92.4%
Validation Acc.: 88.8%
Test AUC: 92.5%
Test Acc.: 87.7%

Completed Runs ⓘ

1/1

Points: Off

Line Smoothness ⓘ

1

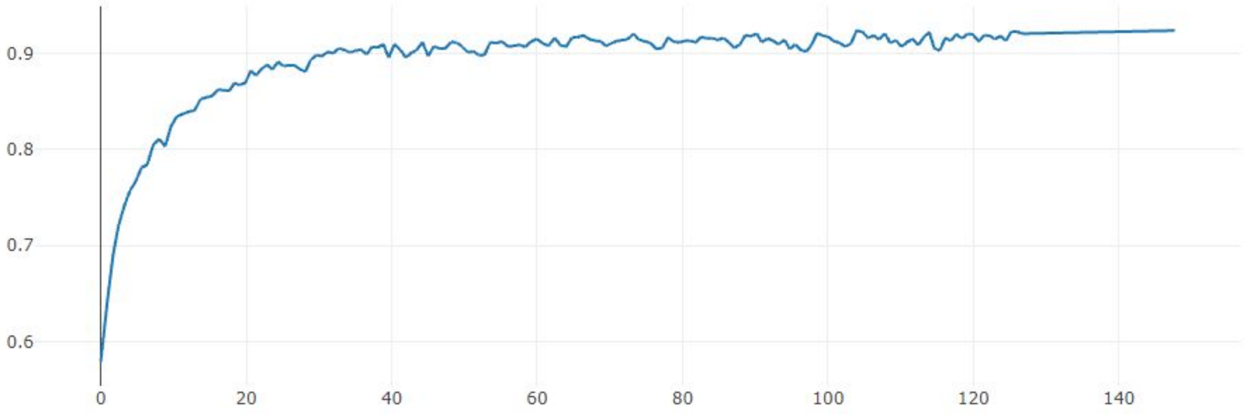
X-axis:

- Step
- Time (Wall)
- Time (Relative)

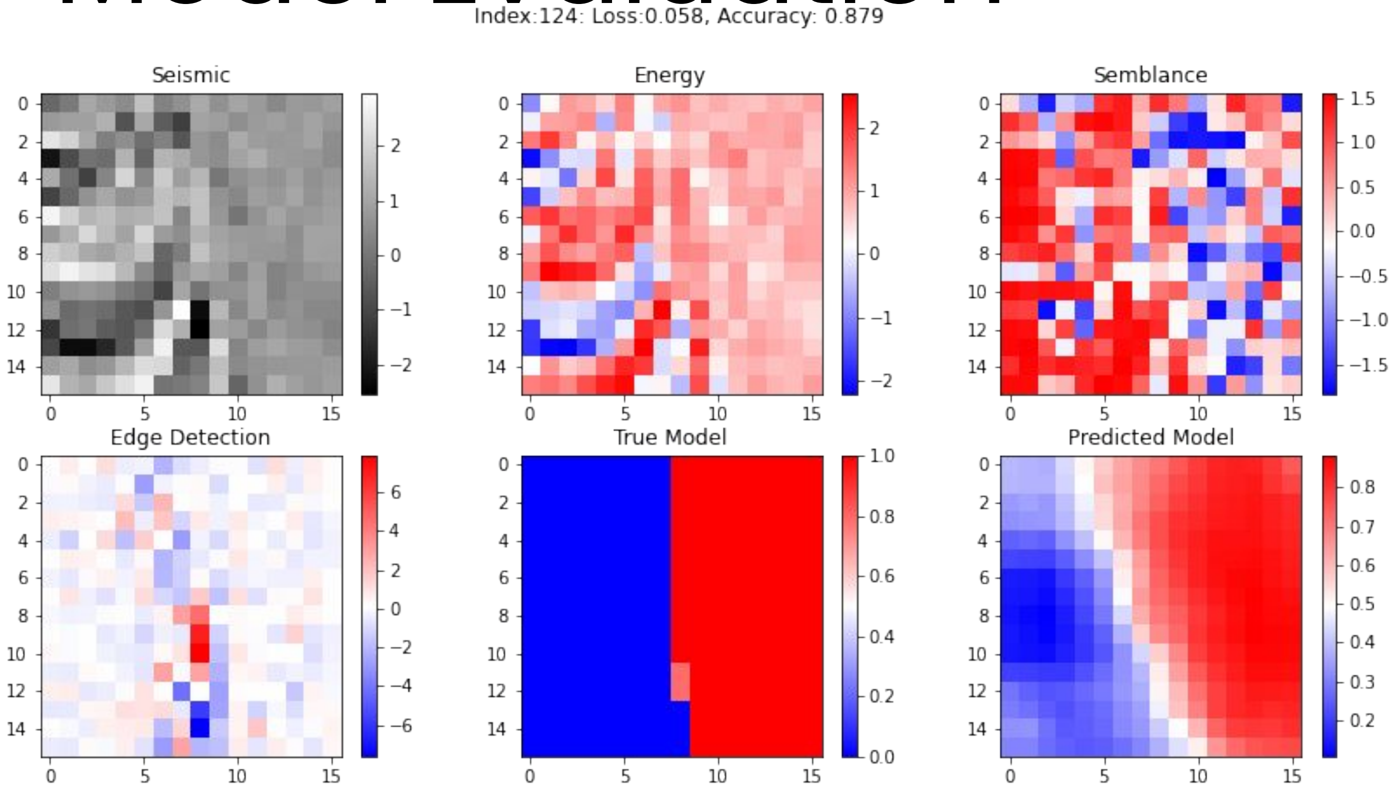
Y-axis:

val_auc ×

Y-axis Log Scale: Off

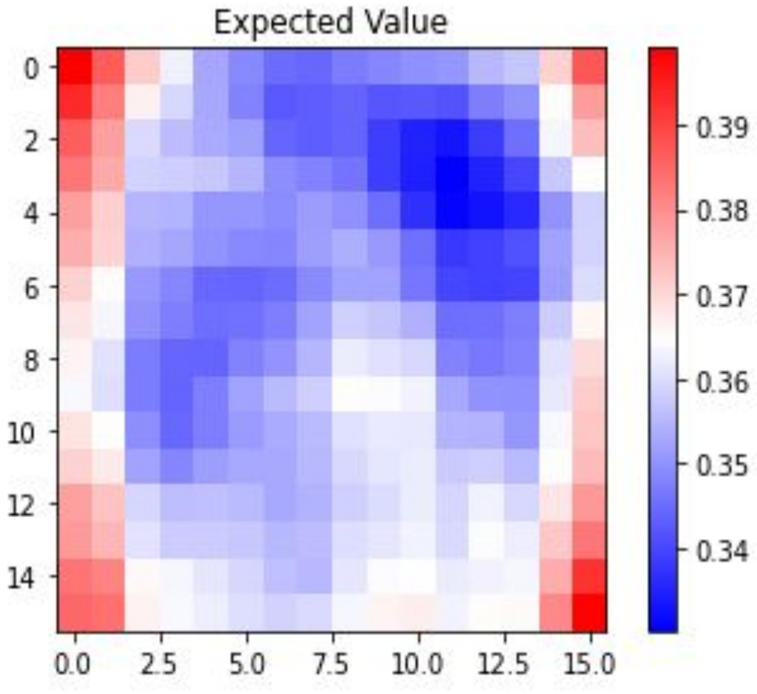


Model Evaluation

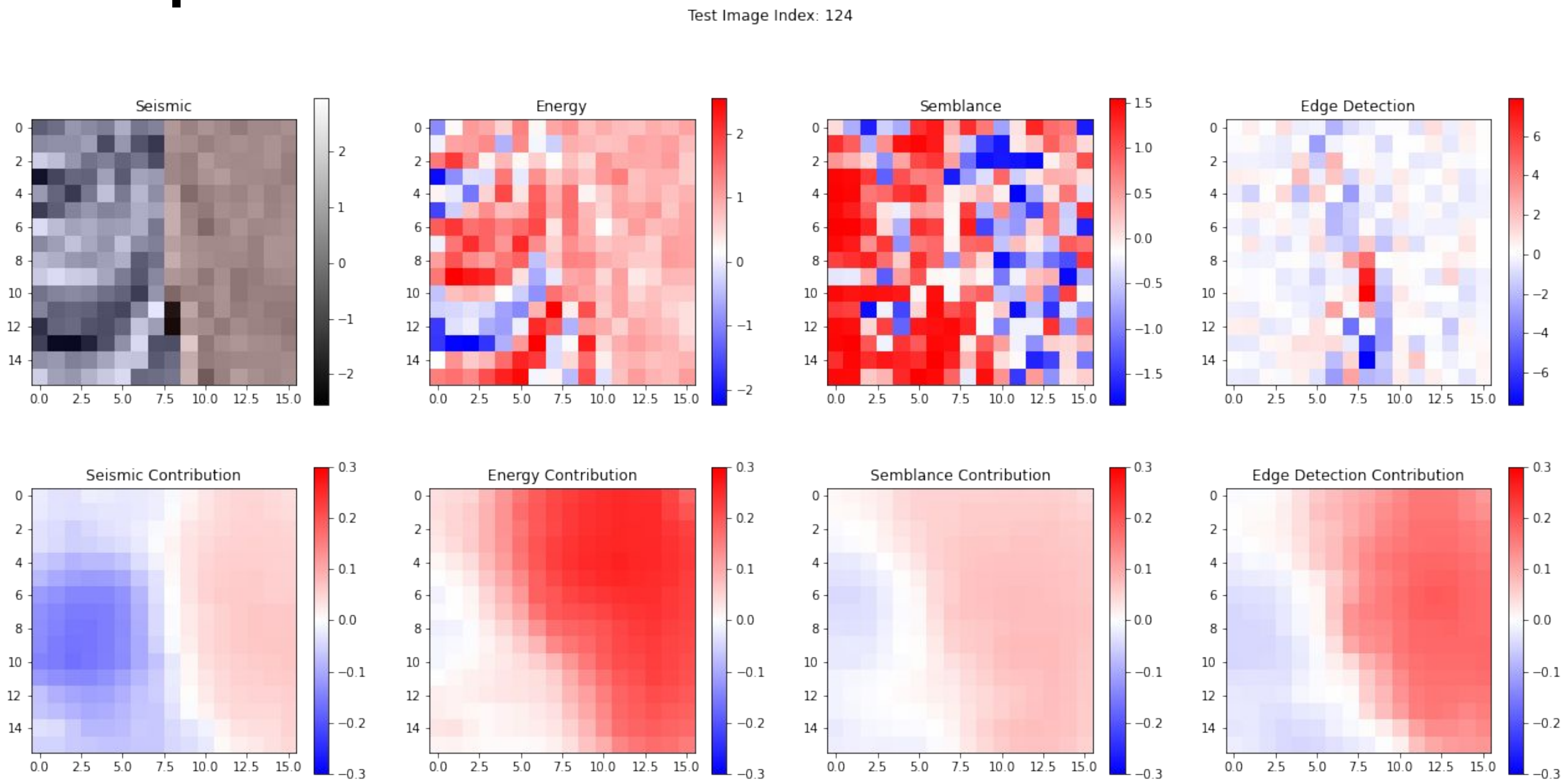


We had to reduce samples to 16x16

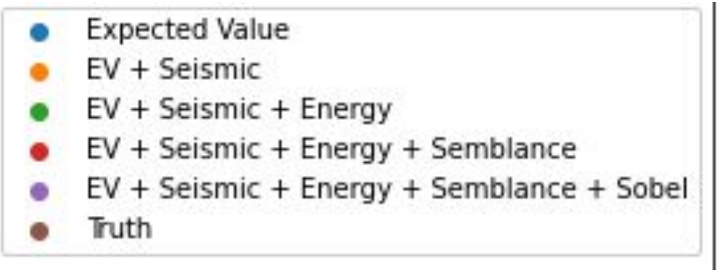
Shap Evaluation

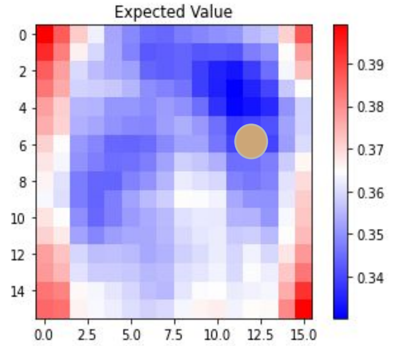


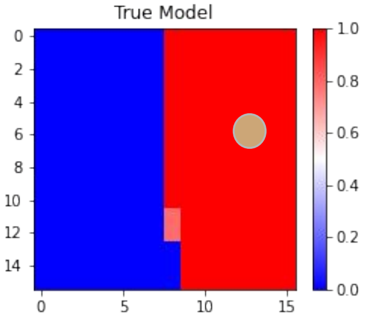
Shap Evaluation

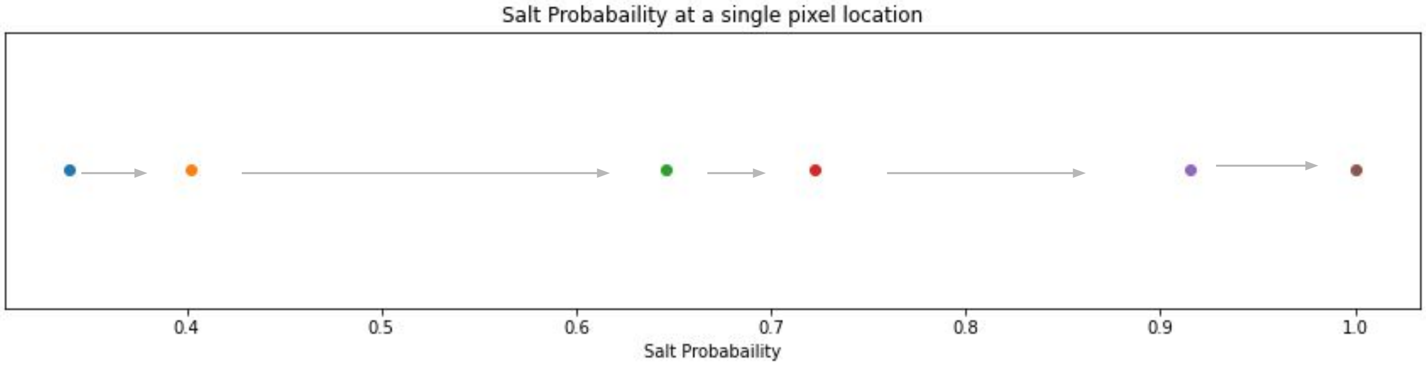


Shap Evaluation

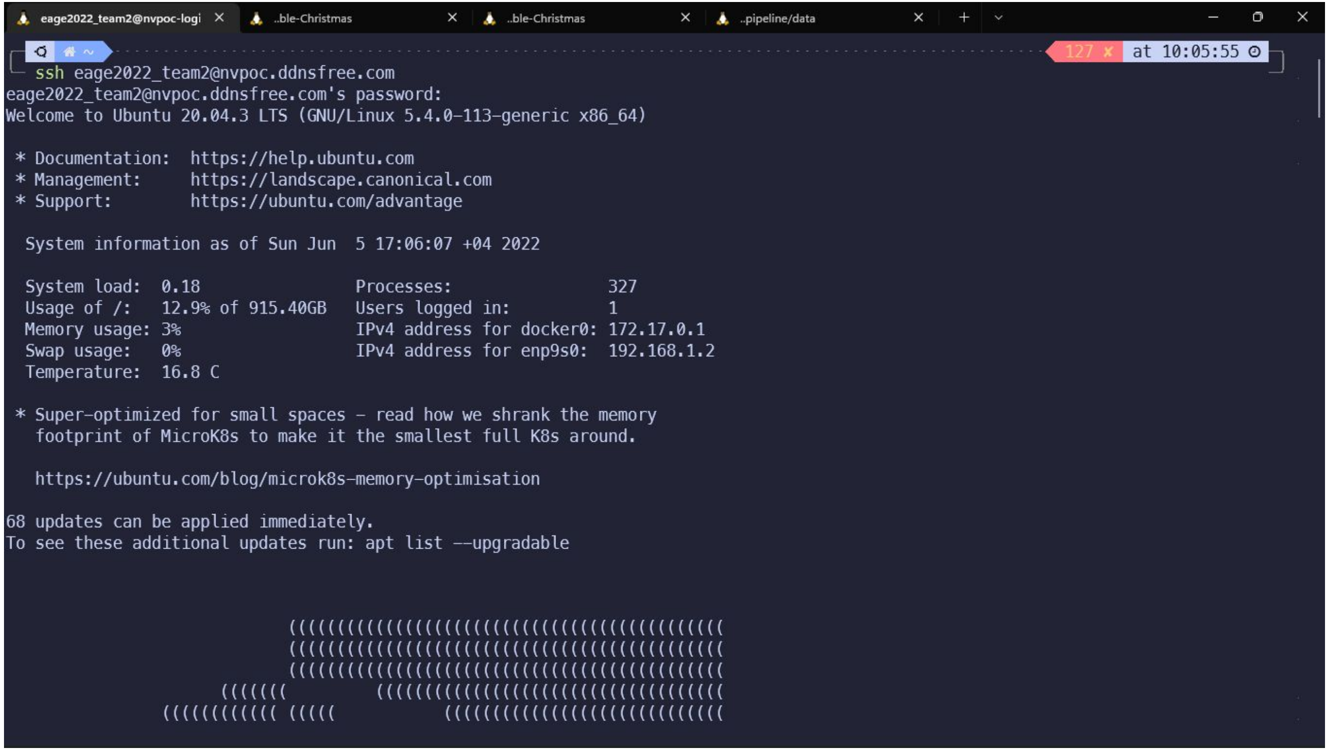








Challenges



- Computing the SHAP values **isn't optimized yet to GPUs**: had to reduce samples to 16x16

- Using NVIDIA's cluster

Conclusions & Further Work

Using SHAP to explain allowed us to:

- Evaluate possible biases: all the attributes and the seismic data contribute to the predictions;
- Complete data-driven approach;

Next steps:

- Optimize a few parts of the code so it runs faster;
- Reproducible pipeline and experiment tracking.

Food for thought. .

“Deep learning also makes problem-solving much easier, because it completely automates what used to be the most crucial step in a machine learning workflow: feature engineering. “

Francois Chollet, Deep Learning with Python, 2021