SPE GCS ML Challenge 2025 — Team Polska

Problem statement

The 2025 SPE GCS ML Challenge aimed to accurately predict the Bottom Hole Circulating Temperature (BHCT) for five wells. The problem was divided into two phases:

- Phase 1: Develop a single, generalized model to predict BHCT.
- Phase 2: Fine-tune the model for improved accuracy on new wells.

Participants were restricted from using pre-trained models or external datasets, such as the Geothermal Data Repository. Model performance was evaluated using the R-squared (R^2) metric, measuring how well predictions matched actual values. The challenge reflected real-world constraints, including limited training data and variations in well conditions.

EDA and Data Pre-processing

The challenge dataset was divided into two phases:

- Phase 1: Data was collected from three wells, each within a similar depth range. The top section (7,000 to 8,500 ft) contained 88 input features along with known BHCT values for training. The bottom 500 ft included only input data, requiring BHCT predictions.
- Phase 2: Data came from two additional wells. The first well followed the same structure as in Phase 1, while the second well had a much shorter training section (2,000 ft) and was located at a greater depth than its test section.

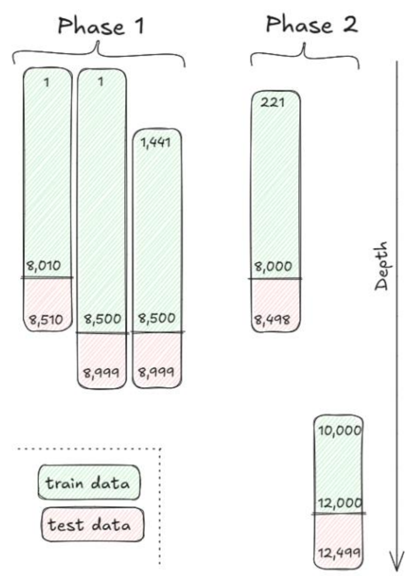


Figure 1 Spatial position of well data.

BHCT generally followed a linear trend consistent with the geothermal gradient. However, outliers were observed, primarily correlated with drill-pipe re-runs and interruptions in the drilling process.

Most existing BHCT prediction studies (Eppelbaum et al., 2014; Khaled et al., 2024) focus on time-domain approaches, utilizing either physical models (based on equivalent time) or machine learning techniques to estimate BHCT regression toward static formation temperature. However, since the challenge dataset was in the depth domain, with data recorded at fixed 1-ft intervals, time-based approaches could not be applied.

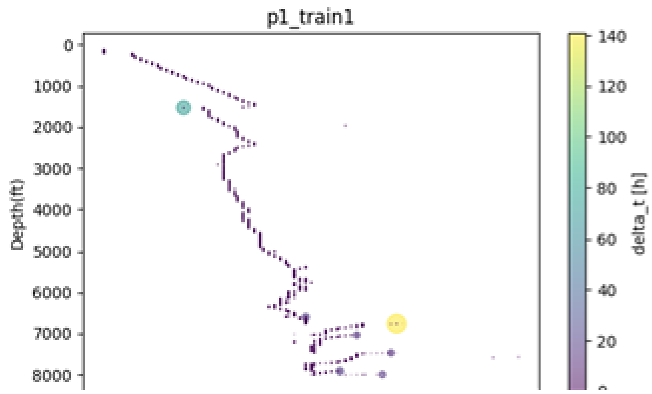


Figure 2. BHCT across well 1 from phase 1. Color and mark size stands for the time difference between consecutive measurements.

By studying correlation diagrams we observed that not all measurements in the training data are available in the test data. We used only the subset of features which is available in all training and test files.

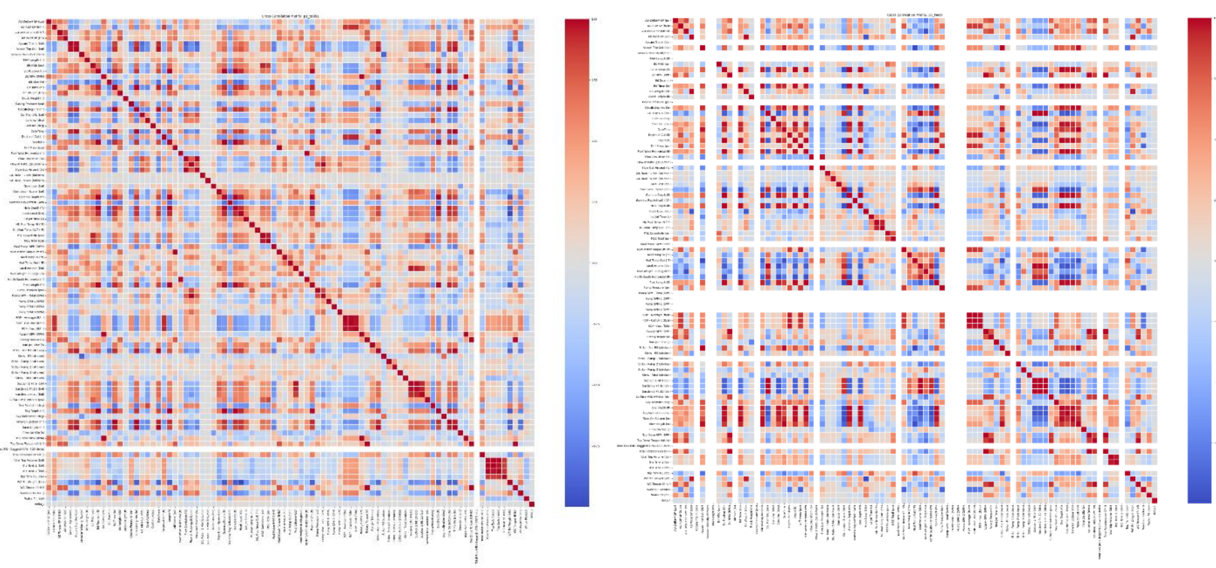


Figure 3 Cross-correlation between measurements in the well no. 1 from Phase 1 for training (left) and test data (right). Feature ordering is the same in both diagrams. Observation: Not all measurements in the training data are available in the test data. We used only the subset of features which is available in all training and test files.

To enhance model robustness and accuracy, we implemented a structured EDA and data pre-processing pipeline, described in the schematic below.

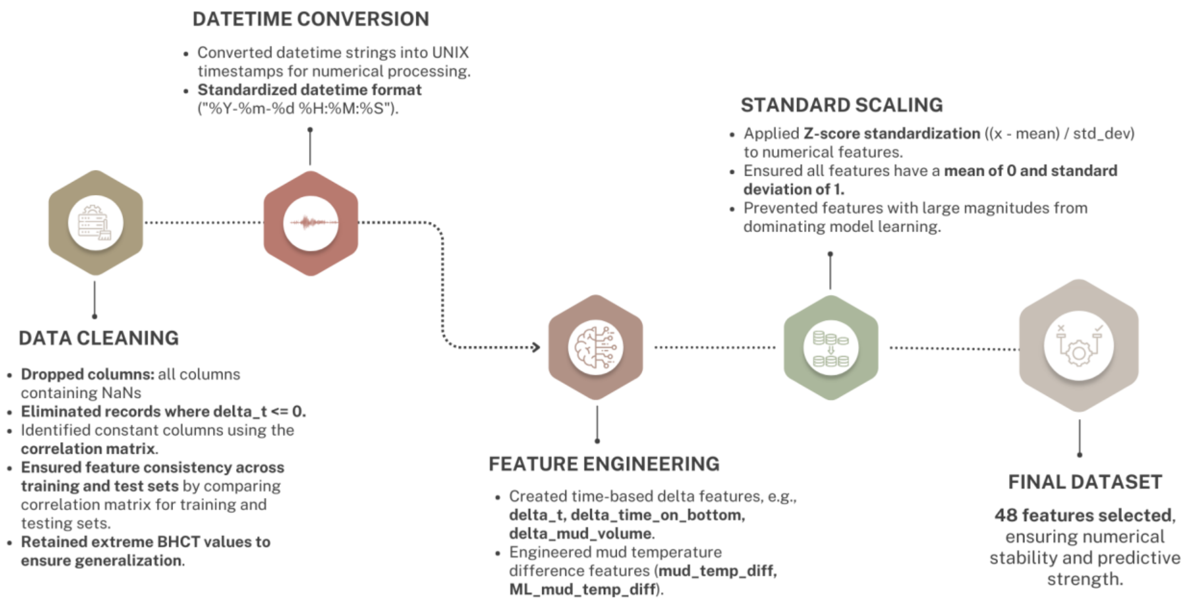


Figure 4 Data Pre-Processing Workflow of Cleaning, Datetime Conversion, Feature Engineering, and Standard Scaling for GRU Model Training.

Approach

To predict Bottom Hole Circulating Temperature (BHCT), we experimented with multiple model architectures, including both static and sequential models:

- Static models: Linear Regression, Support Vector Regression (SVR), and Feedforward Neural Networks (FFN).
- Sequential models: Long Short-Term Memory (LSTM), Gated Recurrent Unit (GRU), and Echo State Network (ESN).

Among these, GRU (Cho et al., 2014; Chung et al., 2014) emerged as the best-performing model, offering a strong balance between accuracy, efficiency, and adaptability to sequential geothermal drilling data.

To evaluate real-world performance, we excluded the last 500 rows from training and used them for testing. Given the limited dataset (only three wells in Phase 1), we applied synthetic data augmentation by generating 200+ trained models through:

- Noise injection in input data.
- Random weight initialization during training.

The simplified GRU architecture (compared to LSTM) enabled efficient training of multiple models. The final prediction was obtained by taking the median output across multiple models, which significantly improved generalization.

Model performance was assessed using the Mean Absolute Error (MAE), ensuring predictions closely aligned with actual BHCT values.

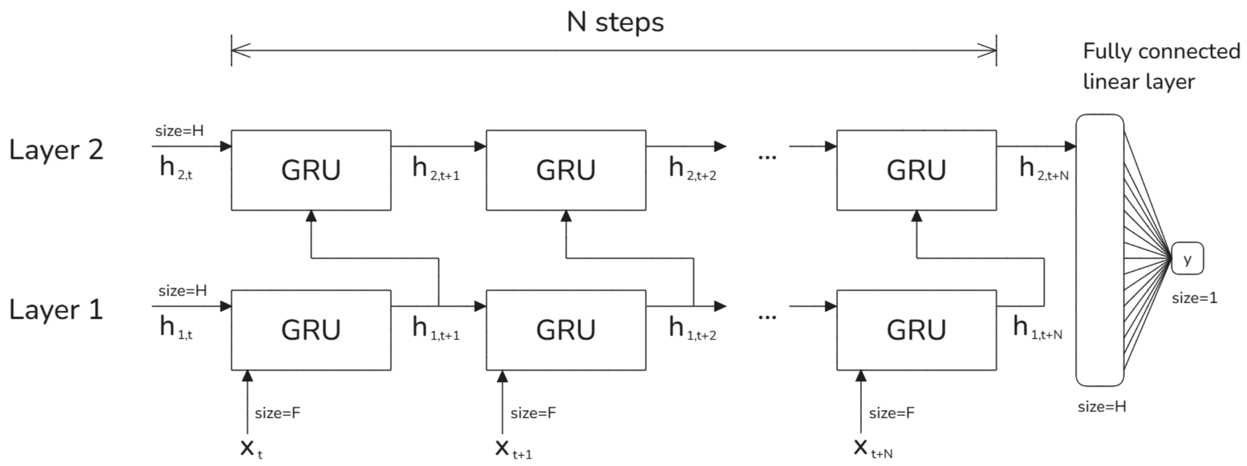


Figure 5 Final model architecture. The model consists of 2 layers of GRU units followed with a linear layer. The input x sequence length is N=16. The size of the hidden state h of each GRU is H=16. The hidden state is initialized with zeros. The input x feature size is F=48. The final state from the second GRU layer is passed to a fully connected linear layer to project it to a scalar output y.

Compiling Final Prediction

The final BHCT prediction was obtained by taking the median across multiple model predictions, ensuring robustness and reducing the impact of outliers.

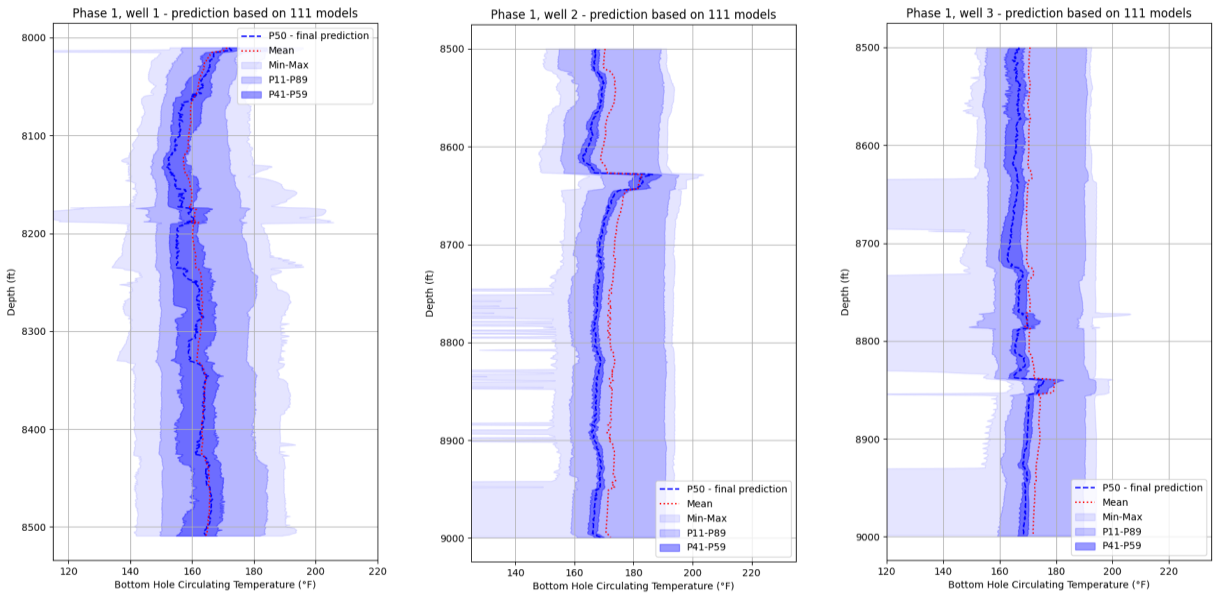
Phase 1:

- We used 111 models, with each model trained on one of the three wells from Phase 1 (37 models per well).
- Selection was agnostic to the dataset, meaning they were not actively selected based on any specific criteria.

Phase 2:

- We tested all models trained on data from all five wells across both phases.
- Models were evaluated by generating predictions on training data, and those with a Mean Absolute Error (MAE) ≤ 15 were selected for the final prediction.
- As in Phase 1, the median of selected model predictions was used as the final output.

This approach helped enhance prediction stability and generalization across different well conditions.



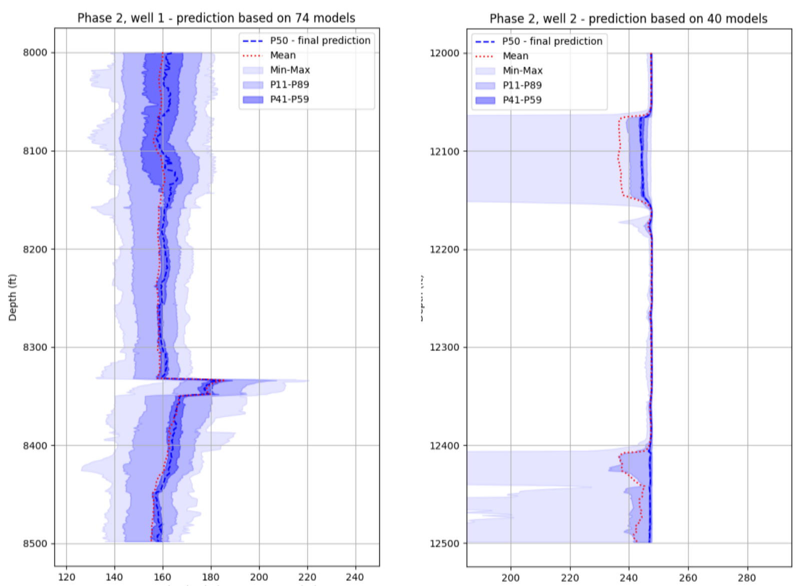


Figure 6 Final BHCT predictions with the range of predictions from multiple models. The median prediction represents the final output, while the shaded regions illustrate the variability across different models.

During testing, we observed that models generally performed better when trained and tested on data from the same well. This suggests that in a real-time drilling scenario, training a model on data from the current well could improve BHCT predictions for deeper sections.

However, this approach comes with a trade-off: while using only the current well's data helps capture its specific characteristics, it may reduce the model's generalization to other wells.

Balancing this trade-off is crucial for optimizing both accuracy and adaptability in practical applications.

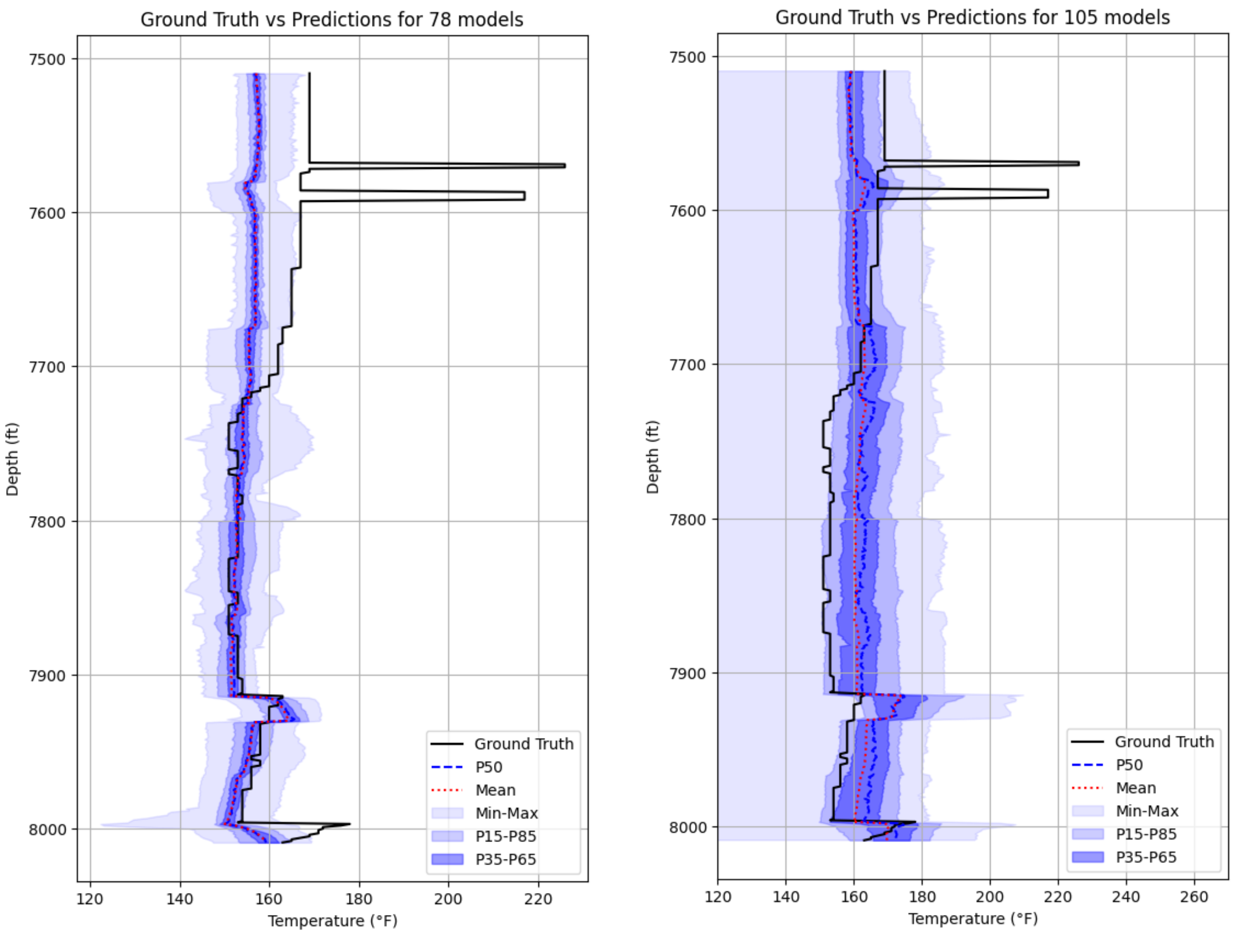


Figure 7 BHCT predictions for the last 500 rows of Well 1, Phase 1 (test data excluded from training). Left: Predictions from models trained only on Well 1 data. Right: Predictions from an ensemble of all models trained on data from all three wells in Phase 1.

Summary

Predicting BHCT is crucial for optimizing geothermal drilling performance, ensuring operational safety, and preventing costly failures of drilling tools. Traditional modelling techniques struggle to provide precise, adaptable predictions due to the complex and dynamic nature of downhole conditions, leading to inefficiencies, increased operational risks, and unplanned downtime. Additionally, the limited availability of high-quality labelled data poses a significant challenge in developing reliable predictive models.

We propose a specially designed Gated Recurrent Unit (GRU) architecture, optimized for BHCT prediction. Our approach is preceded by advanced Exploratory Data Analysis (EDA), feature engineering, data augmentation, and advanced regularization techniques, allowing the model to extract meaningful patterns and enhance performance despite data constraints. The GRU-based model is designed to generalize across multiple wells while adapting to new datasets, ensuring robustness and precision in varying drilling conditions. By integrating these advanced machine learning techniques, our solution aims to enhance decision-making, improve drilling efficiency, and mitigate risks associated with geothermal drilling operations.

Team Polska

Krzysztof Arendt
Tomasz Topór
Damian Gruszecki
Mirosław Wojnicki

References

Cho, Kyunghyun, et al. “On the Properties of Neural Machine Translation: Encoder-Decoder Approaches.” *arXiv.Org*, 3 Sept. 2014, https://arxiv.org/abs/1409.1259.

Chung, Junyoung, et al. “Empirical Evaluation of Gated Recurrent Neural Networks on Sequence Modeling.” *arXiv.Org*, 11 Dec. 2014, https://arxiv.org/abs/1412.3555.

Eppelbaum, Lev, et al. *Applied Geothermics*. Springer Science & Business, 2014.

Khaled, Mohamed Shafik, et al. “Real-Time Prediction of Bottom-Hole Circulating Temperature in Geothermal Wells Using Machine Learning Models.” *Geoenergy Science and Engineering*, vol. 238, July 2024, p. 212891, https://doi.org/10.1016/j.geoen.2024.212891.