

# Machine Learning Model Tracking Document

## 1. Dataset Information

Dataset Name:	Well 160
Number of Samples:	160
Number of Features:	4 (Tf, Rs, Gg, Api)
Target Variable:	Pb
Outlier Handling:	None
Feature Engineering Applied:	None
Scaling/Normalization Applied:	Only for Neural Networks
Encoding Applied:	None

## 2. Preprocessing Steps

Step	Description
Train-Test Split	70% - 30%
Shuffling	Yes, using random_state=42
Handling Missing Data	None
Feature Scaling	Only for Neural Networks
Feature Selection	None

## 3. Models Used & Hyperparameters

Model	Hyperparameters	Training Time
XGBoost	n_estimators=200, learning_rate=0.1, max_depth=4	0.0252
CatBoost	iterations=1000, learning_rate=0.1, depth=4, l2_leaf_reg=7	0.1036
Neural Network	[128, 64, 32], epochs=150, batch_size=32	3.4472
Stacking Ensemble	Default base models + CatBoost final estimator	0.4703
Extra Trees	n_estimators=200	0.0569
Deep Neural Network	[256, 128, 64], epochs=150, batch_size=16	3.4314

## 4. Evaluation Metrics

Model	MSE	RMSE	MAE	R <sup>2</sup> Score	Adjusted R <sup>2</sup>
XGBoost	13412.360	115.812	83.490	0.9902	0.9892
CatBoost	22057.603	148.518	86.771	0.9838	0.9823
Neural Network	58507.281	241.883	176.604	0.9571	0.9531
Stacking Ensemble	15595.798	124.883	79.742	0.9886	0.9875
Extra Trees	21312.234	145.987	93.944	0.9844	0.9829
Deep Neural Network	18709.108	136.781	90.925	0.9863	0.9850

## 5. Cross-Validation Summary (5-Fold)

Model	RMSE Mean	RMSE Std	MAE Mean	MAE Std	R <sup>2</sup> Mean	R <sup>2</sup> Std
Stacking Ensemble	119.651	24.637	81.829	15.582	0.9867	0.0053
CatBoost	122.980	35.282	76.138	14.875	0.9854	0.0090
Extra Trees Regressor	136.914	23.415	97.075	9.053	0.9826	0.0076
XGBoost	142.822	26.733	97.016	13.295	0.9813	0.0070
Deep Neural Network	531.527	117.639	409.308	89.067	0.7317	0.1421
Neural Network	593.133	128.666	464.182	100.915	0.6587	0.1796

## 6. Visualizations

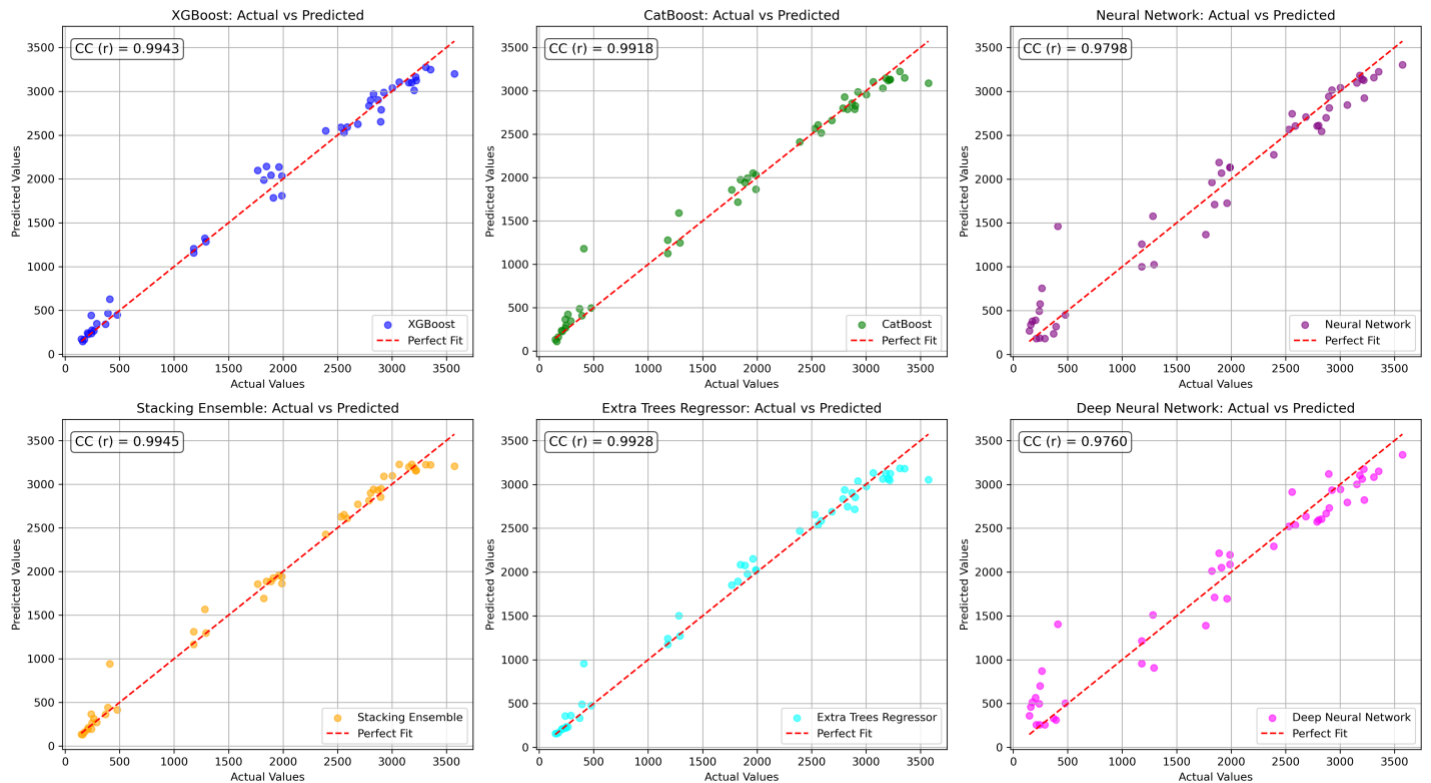


Figure 1: Actual vs. Predicted Values for Pb (Well 160)

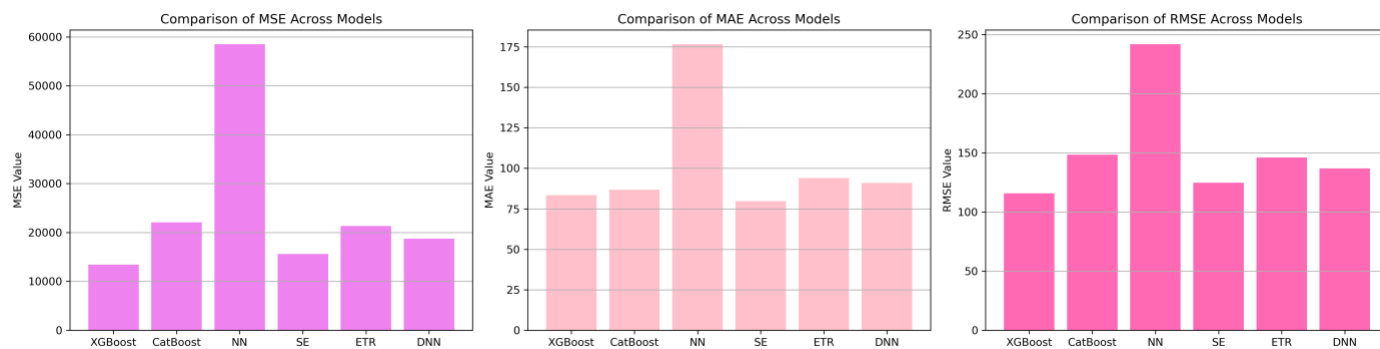


Figure 2: Bar Charts of MSE, RMSE, MAE for Pb (Well 160)

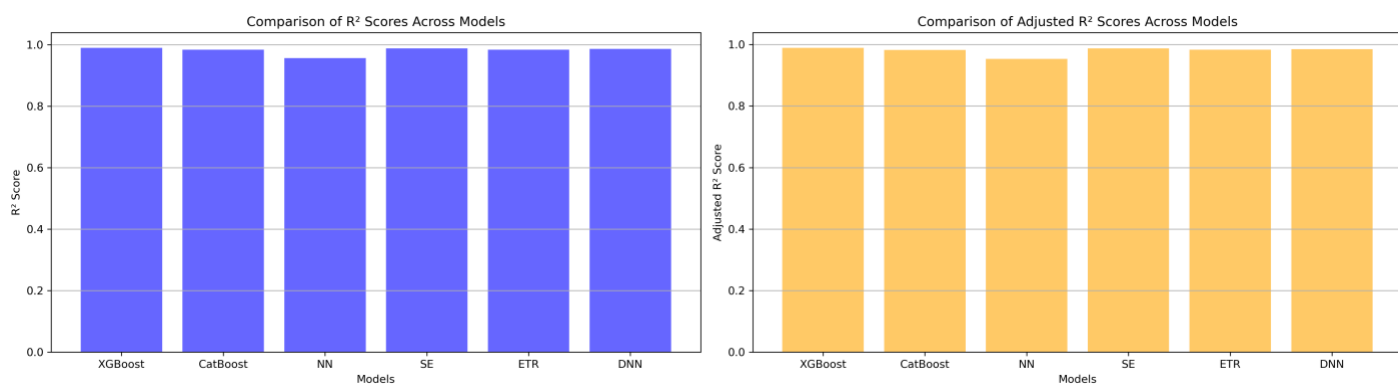


Figure 3: Bar Charts of  $R^2$  and Adjusted  $R^2$  for Pb (Well 160)

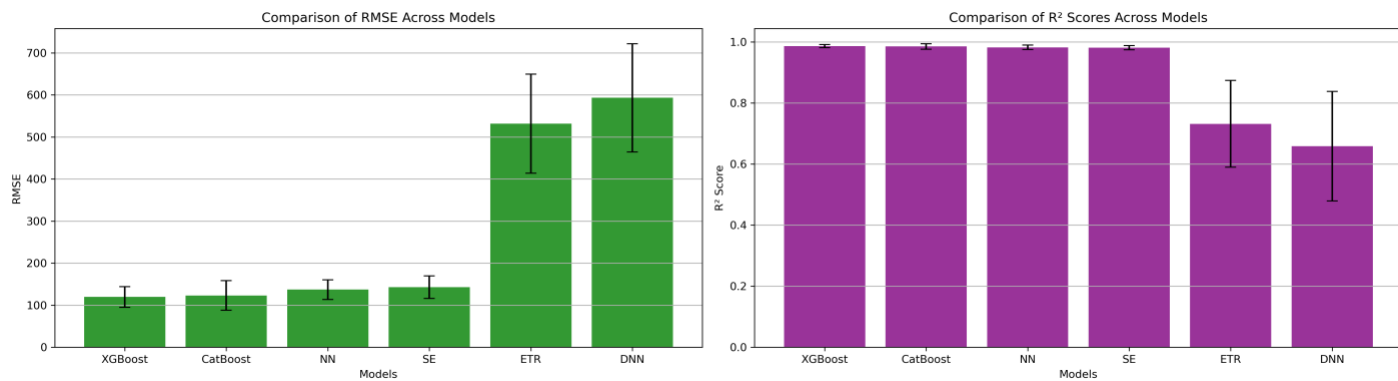


Figure 4: Error Bars for RMSE and  $R^2$  from CV for Pb (Well 160)

## 7. Observations & Next Steps

**Best Performing Model: Stacking Ensemble** (RMSE  $\sim 124.88$ ,  $R^2 \sim 0.9886$ )

- XGBoost and CatBoost also performed strongly but slightly behind the stacking model.
- **Neural Networks (NN and DNN)** performed poorly compared to ensemble methods.
  - Deep learning models showed much higher RMSE and lower  $R^2$ .
  - Particularly bad generalization across folds in cross-validation (high RMSE, unstable  $R^2$ ).
- Small sample size ( $\sim 160$ ) limits the effectiveness of neural networks.

## 8. Code Access

The complete source code for data preprocessing, model training, evaluation, and visualization is [available here](#). The repository includes organized Jupyter notebooks for each phase, dataset, and target, as well as requirements for reproducibility.