

MACHINE LEARNING-BASED ESTIMATION OF OIL RECOVERY FACTOR USING XGBOOST: INSIGHTS FROM CLASSIFICATION AND DATA-DRIVEN ANALYSES

Alireza Roustazadeh¹ , Frank Male² , Behzad Ghanbarian^{3,4,5*} , Mohammad B. Shadmand⁶ , Vahid Taslimitehrani⁷ , Larry W. Lake⁸ 

¹Porous Media Research Lab, Department of Geology, Kansas State University, Manhattan, Kansas, United States; ²John and Willie Leone Family Department of Energy and Mineral Engineering, Pennsylvania State University, University Park, Pennsylvania, United States; ³Department of Earth and Environmental Sciences, University of Texas at Arlington, Arlington, Texas, United States; ⁴Department of Civil Engineering, University of Texas at Arlington, Arlington, Texas, United States; ⁵Division of Data Science, College of Science, University of Texas at Arlington, Arlington, Texas, United States; ⁶Department of Electrical and Computer Engineering, College of Engineering, University of Illinois at Chicago, Chicago, Illinois, United States; ⁷Machine Learning/Data Science Tech Lead at Meta, San Francisco, California, United States; ⁸Hildebrand Department of Petroleum and Geosystems Engineering, University of Texas at Austin, Austin, Texas, United States

1. APPENDIX A

This appendix includes the heat map matrices constructed based on the Spearman's rank correlation coefficient calculated for eleven input and one output features in different databases.



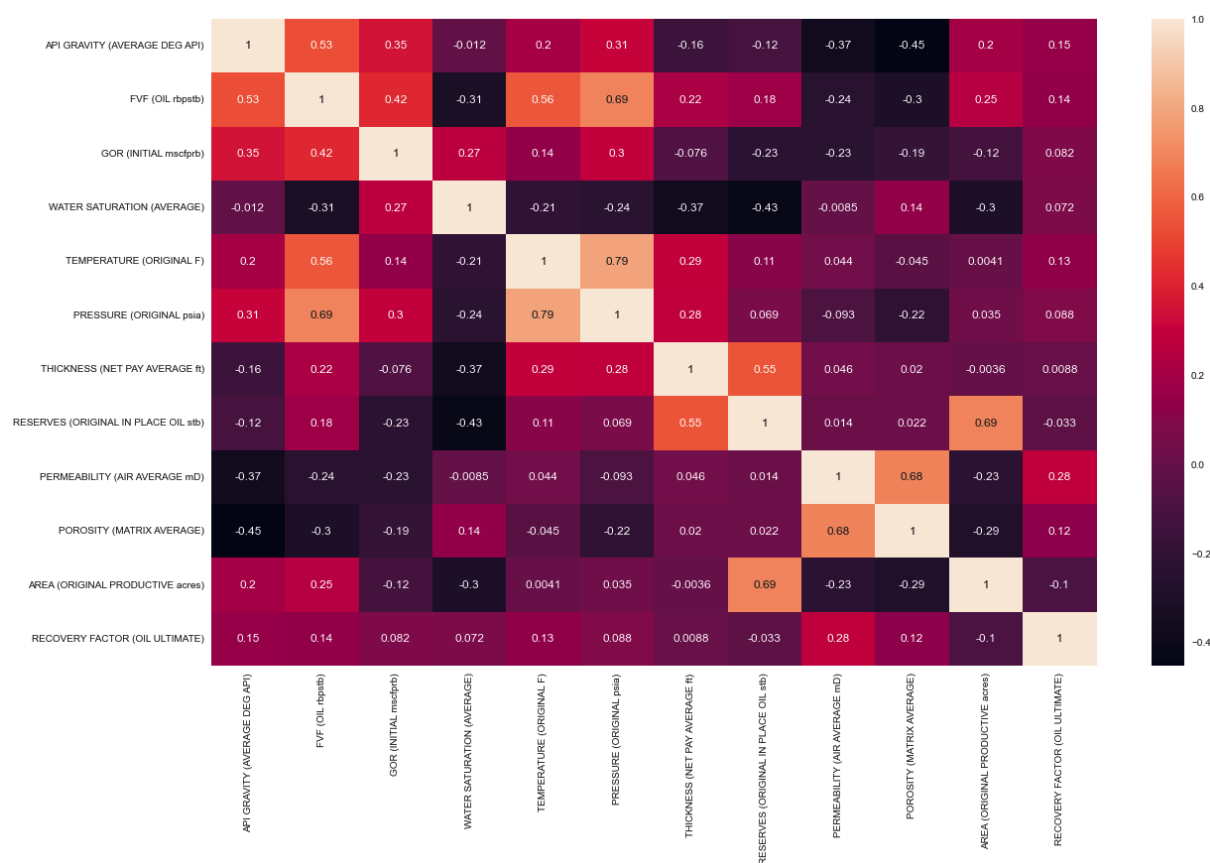


Figure S1: The heat map matrix based on Spearman's rank correlation coefficient calculated for eleven input and one output features in the TC database.

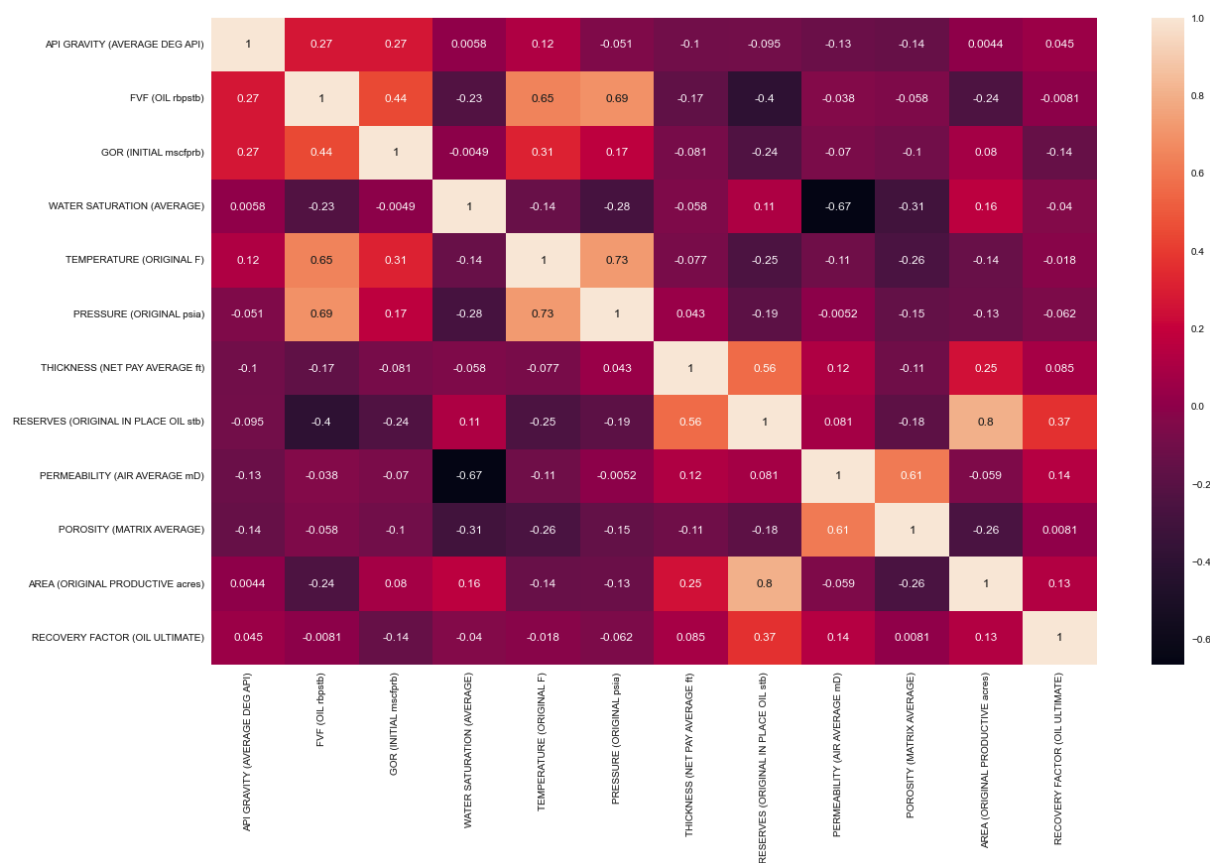


Figure S2: The heat map matrix based on Spearman's rank correlation coefficient calculated for eleven input and one output features in the TA database.

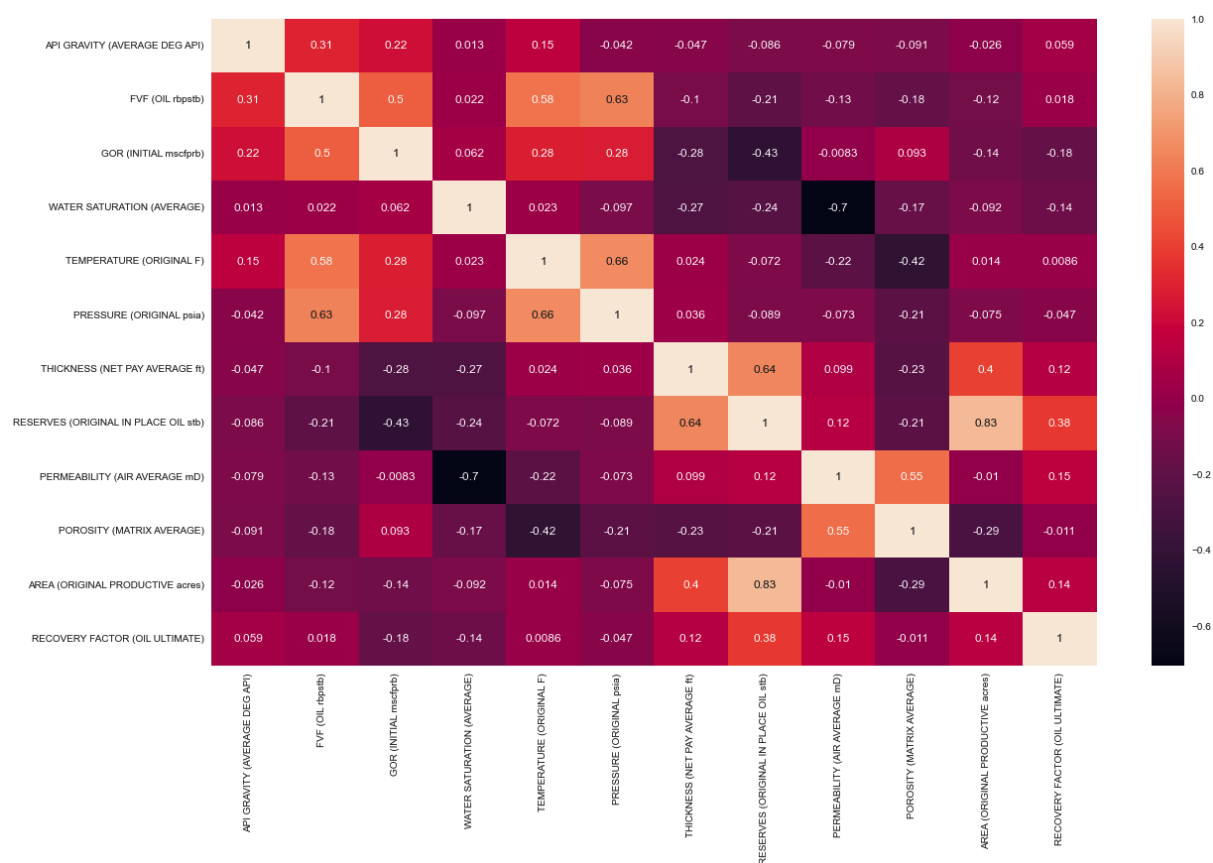


Figure S3: The heat map matrix based on Spearman's rank correlation coefficient calculated for eleven input and one output features in the CA database.

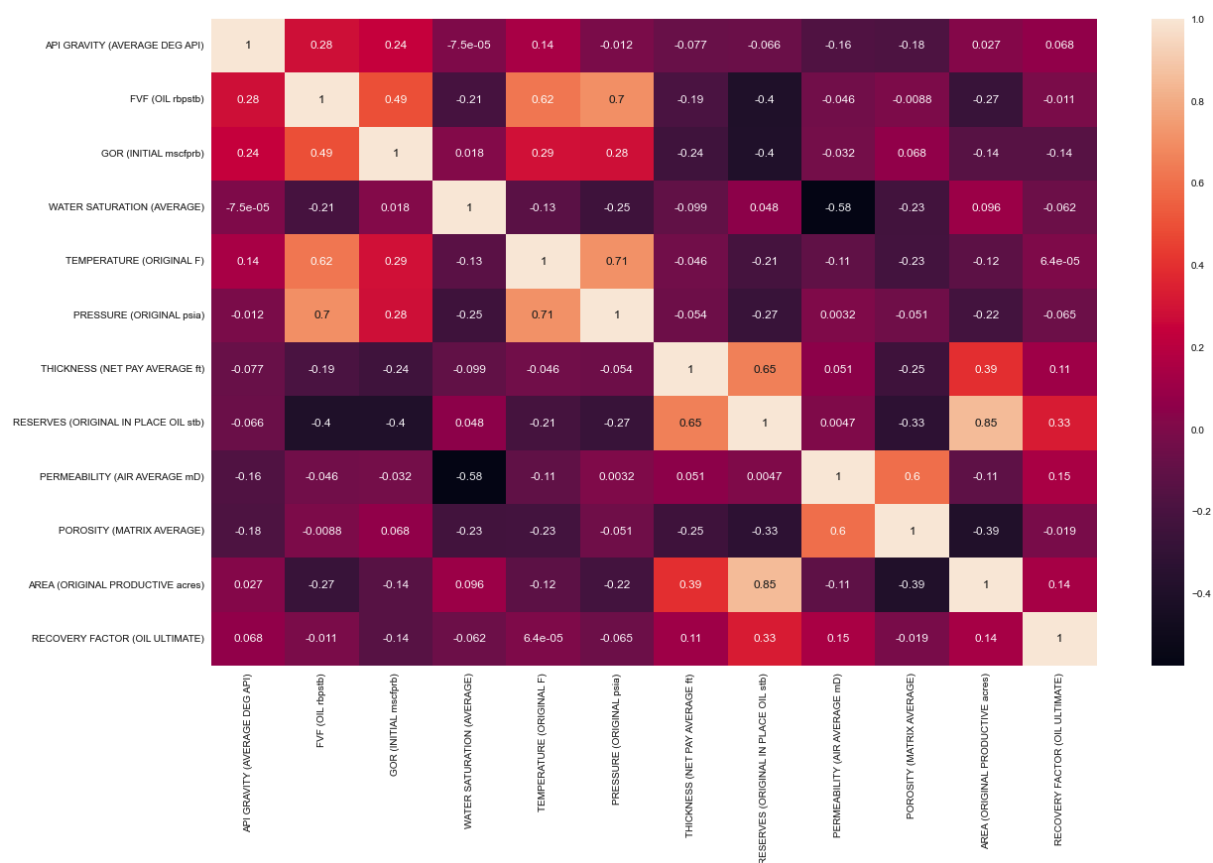


Figure S4: The heat map matrix based on Spearman's rank correlation coefficient calculated for eleven input and one output features in the TCA database.