ENHANCED ARTIFICIAL INTELLIGENCE WORKFLOW FOR PREDICTING PRODUCTION WITHIN THE BAKKEN FORMATION

Jackson R. Tomski, Mrinal K Sen, and Michael J. Pyrcz

Department of Geological Sciences
The University of Texas at Austin

ABSTRACT

The research reported in this paper focuses on a production problem, where a relationship is established between completion parameters and production patterns in the field given a dataset from the Bakken Formation in the northwest North Dakota region. A random forest model is employed alongside a Bayesian Neural Network (BNN) model to predict production given a set of predictive features found through a series of feature selection methods. The BNN model utilizes Markov Chain Monte Carlo (MCMC) via Langevin Dynamics in order to sample from the probability distribution and to estimate uncertainity. Various training and testing dataset scenarios are created through random sampling and clustering via K-means Clustering and Gaussian Mixture Models (GMMs). This is done in order to reduce the sampling bias within the generated training and testing datasets and to ensure that the machine learning models are being trained and tested on data coming from similar geological regions with similar production rate values. With the integration of these techniques, a better understanding of the parameters useful for optimizing oil production is possible with a degree of uncertainity when using the BNN model.

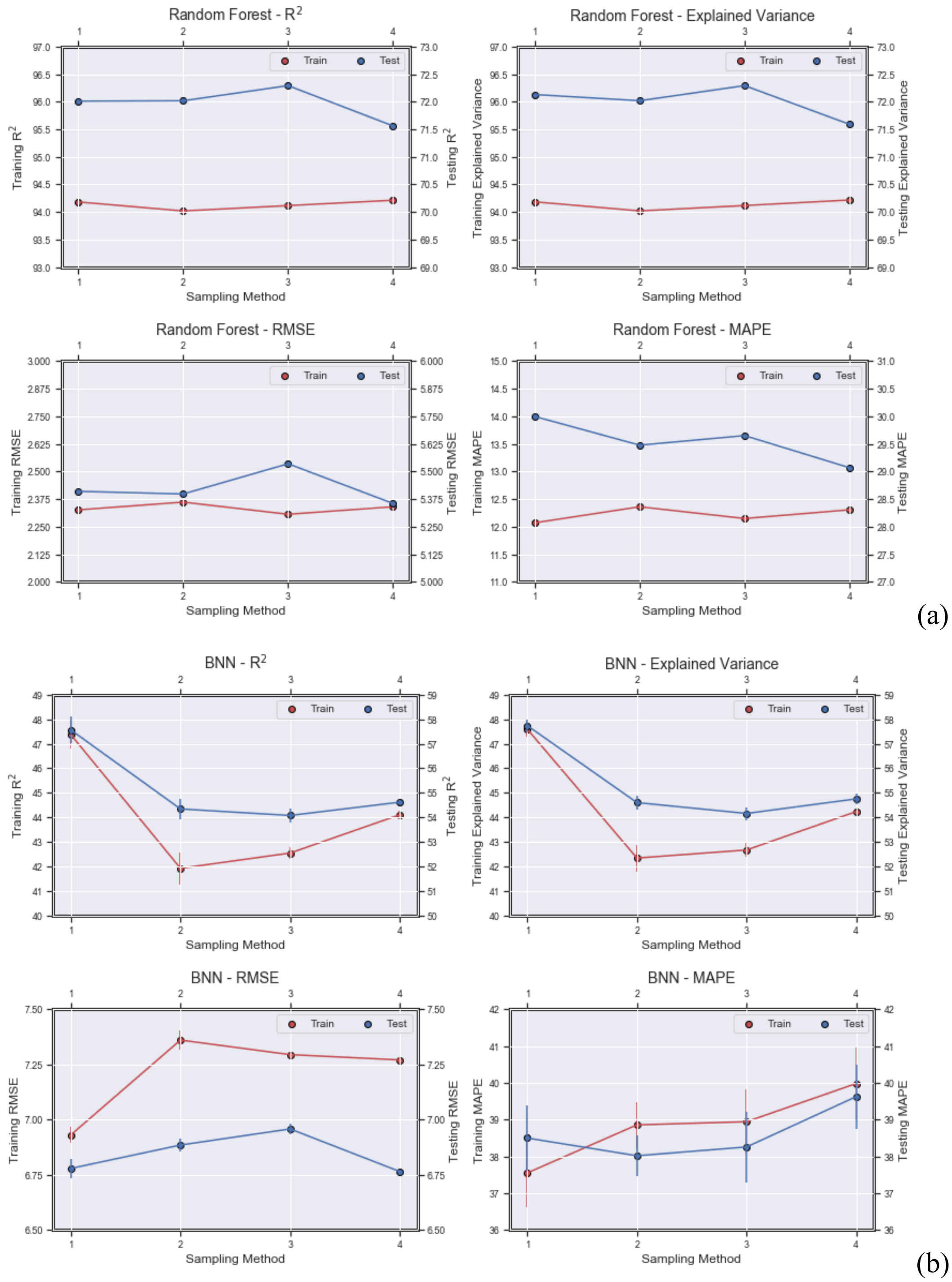

FIGURE: Model performance metrics for (a) random forest model and BNN model (b) with red curves and dots indicating training data, and blue is the testing data. The vertical bars in (b) at each sampling method represent the error bars/STD associated with the computed metric value.