**Essential Question #3 - NK**

*JD’s AI/ML Laboratory*

**Due: 07/11/2025, 11:59 PM EST**

**General Instruction**: Please **answer all questions** and provide relevant support from literature using **APA references**. You can answer the questions using bullet points with citations (preferably from peer-review journals). Keep in mind that most science competitions focus on participants having a very strong grasp of the literature. I would highly recommend printing out the relevant journal articles and do a deep dive. Using a highlighter and writing notes on the journal article can help you narrow down important content for future reference. I will be thorough in my review of your response and expect all the answers to be well thought-out and clearly explained with supporting literature.

1. How is Random Forest Classification model different from Random Forest Regression? Explain using a decision tree based on you project features/variables
2. Add to your table and learn about these models Elastic Net Regression, Gradient Boosting Regressor, Support vector regressor, Multilayer Perceptron Regressor, Gaussian Process Regression (GPR) and Bayesian Regression.
3. Review article, focus on ML models, SHAP analysis, data curation for ML analysis and ML analysis workflow https://www.sciencedirect.com/science/article/pii/S259012302400183X#ab0010
4. Select 4 most relevant model for your project
5. Learn about multi objective optimization using ML, in your case you are trying to optimize both flowrate and filter life span
6. Train a machine learning model that would help you answer your scientific question
7. Draft PowerPoint with Background, Research Question, Hypothesis, research plan, EDA, Analysis Flowchart, ML model analysis, results, conclusions.
8. Prepare for oral presentation on Friday
9. Draft research paper outline