

This is a supervised learning task with nine features and a quantitative label. We have training and test sets available as csv files (with headers). The features are a mix of data types. Your goal is to train a Random Forest and compute MSE on the test set.

Use the `sparkHW.py` template. Read all of the comments and add your code under the appropriate comments. You may add comments, but you probably don't need to (except to enter your name). You probably want to develop your code in interactive pyspark, but be sure to test it using `spark-submit` before submitting on Canvas.

Comment out all test code including and `.show()` statements and any `print()` statements other than for the MSE.

Turn in **`sparkHW.py`**.