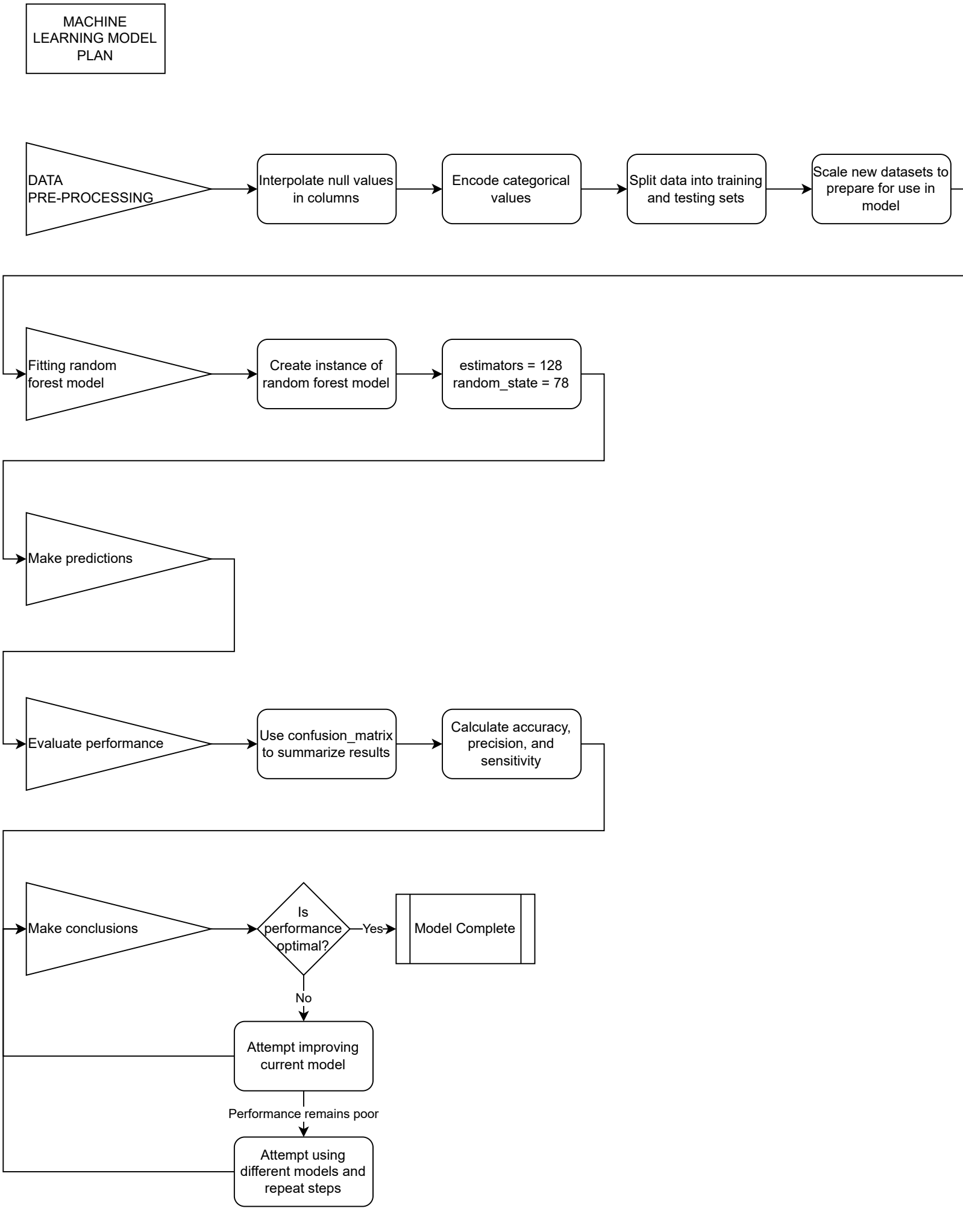
MACHINE LEARNING MODEL PLAN
DATA PRE-PROCESSING
Interpolate null values in columns
Encode categorical values
Split data into training and testing sets
Scale new datasets to prepare for use in model
Fitting random forest model
Create instance of random forest model
estimators = 128
random_state = 78
Make predictions
Evaluate performance
Use confusion_matrix to summarize results
Calculate accuracy, precision, and sensitivity
Make conclusions
Is performance optimal?
Yes
Model Complete
No
Attempt improving current model
Performance remains poor
Attempt using different models and repeat steps