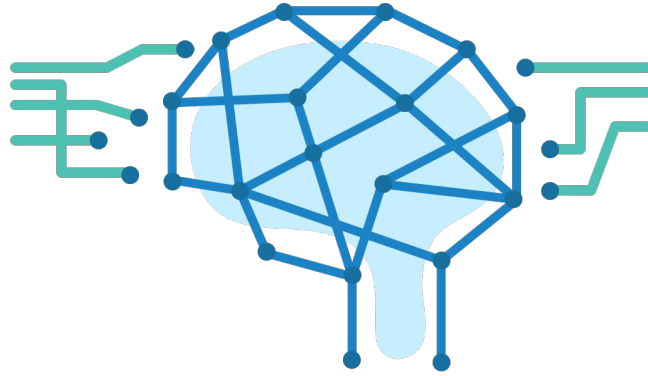




# **Machine Learning for Non Data Scientist (Key Terminology)**

# Machine Learning



A branch of science that “trains” computers to be **able to learn on their own**, without the need to be specially programmed



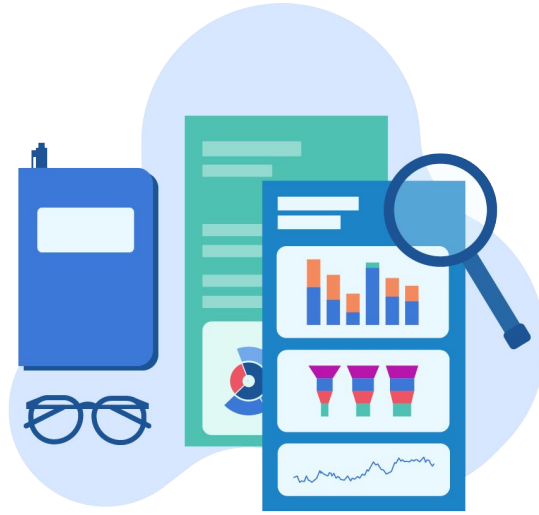
# Machine Learning Process



- Train computer by creating machine learning models using different algorithms, the test it
- The choice of algorithm depends on problem and the the type of data obtained and what task will be automated



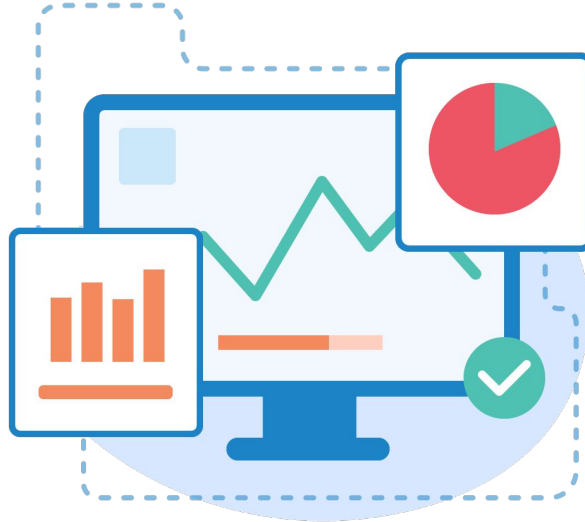
# Training Data



The sample of data used to fit the model. The actual dataset that we use to train the model. The model sees and learns from this data.



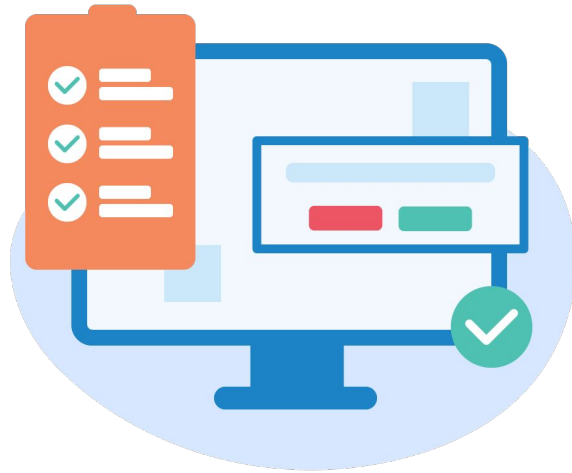
# Validation Data



The sample of data used to provide an unbiased evaluation of a model fit on the training dataset while tuning the model.



# Testing Data



*The sample of data used to provide an unbiased evaluation of a final model fit on the training dataset.*

