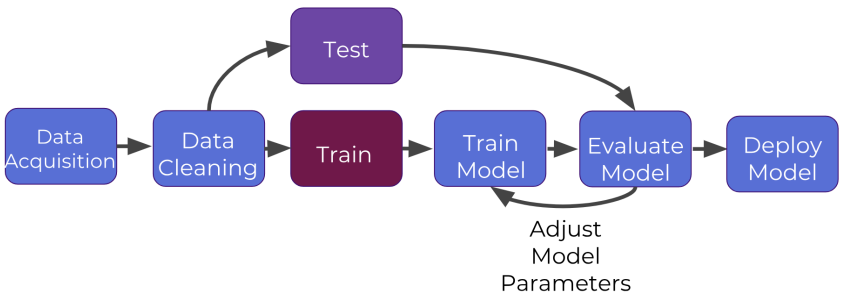
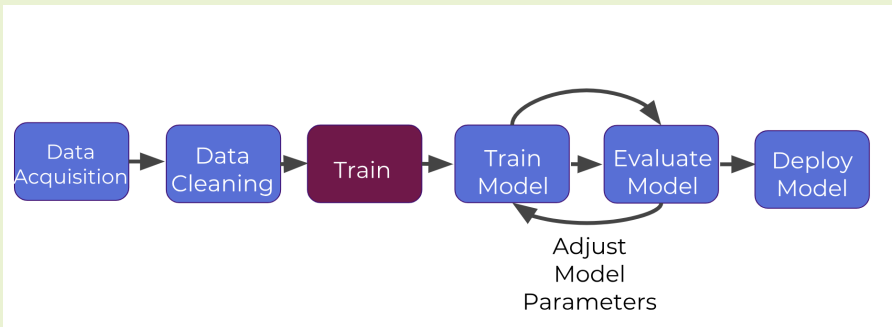
Machine Learning Algorithms

Supervised	Uses labelled data to predict a label given some features. If the label is continuous we have a regression problem (house value based on house features, if categorical we have a classification problem (gender based on physical measures). 
Unsupervised	Uses unlabelled data to look for patterns in and categorize data (clustering). ie. grouping breeds of dogs when we don't know what breeds the dogs are 
Reinforcement	Uses trial-and-error to maximize a specified reward metric.

Note: Another version/diagram of supervised learning uses holdout sets (even more labelled data) that tests the model in-between the evaluate and deploy stages to estimate a "true" accuracy for the model. Kind of like a final exam that can't be resat.

Evaluating Models (seeing how good they are)

Supervised - regression	Mean absolute error (+ has original scale) Mean squared error (+ punishes larger errors) Root mean squared error (+ has original scale **and** punishes larger errors)
Supervised - categorical	Accuracy (correctly classified / total samples) Recall (?) Precision (?)
Unsupervised	(harder to evaluate due to no labels, method depends on goal of task) Cluster homogeneity (?) Rand Index (?)
Reinforcement	More obvious/intuitive - how well the model performs its task