

MACHINE LEARNING IN EMOJI 🧐

- Supervised
- Unsupervised
- Reinforcement

- **Supervised:** Humans build models based on input/output.
- **Unsupervised:** Human input, machine output, human utilizes if satisfactory.
- **Reinforcement:** Human input, machine output, human reward/punish, cycle continues.

CLUSTER ANALYSIS

- ● **K-MEANS** `cluster.KMeans()`
Similar datum into groups based on centroids 

- ● **Anomaly detection** `covariance.EllipticalEnvelope()`
Finding outliers through grouping 

BASIC REGRESSION

- **Linear** `linear_model.LinearRegression()`
Lots of numerical data 
- **Logistic** `linear_model.LogisticRegression()`
Target variable is categorical 

FEATURE REDUCTION

- T-Distrib. stochastic** `manifold.TSNE()`
Visualize high dimensional data. Convert similarity to joint probabilities 

- Principle component analysis** `decomposition.PCA()`
Distill feature space into components that describe greatest variance 

- Canonical correlation analysis** `decomposition.CCA()`
Making sense of cross-correlation matrices 

- Linear discriminant analysis** `lda.LDA()`
Linear combination of features that separate classes 

CLASSIFICATION

- ● ● **Neural net** `neural_network.MLPClassifier()`
Complex relationships, prone to overfitting. Basically magic. 

- **K-NN** `neighbors.KNeighborsClassifier()`
Group membership based on proximity 

- **Decision tree** `tree.DecisionTreeClassifier()`
if/then/else Non - contiguous data
Can also be regression 

- ● **Random forest** `ensemble.RandomForestClassifier()`
Find best split randomly, can also be regression 

- ● **SVM** `svm.SVC()` `sv.LinearSVC()`
Maximum margin classifier. Fundamental Data Science algorithm 

- **Naive Bayes** 
Updating knowledge step by step with new info
`GaussianNB()` `MultinomialNB()` `BernoulliNB()`

OTHER IMPORTANT CONCEPTS

BIAS VARIANCE TRADEOFF

UNDERFITTING / OVERFITTING

INERTIA

ACCURACY FUNCTION $(TP + TN) / (P + N)$ PRECISION FUNCTION $TP / (TP + FP)$ SPECIFICITY FUNCTION $TN / (FP + TN)$ SENSITIVITY FUNCTION $TP / (TP + FN)$