

1. Machine learning is the ability to make predictions based on previous data and continue to roll with the punches as data changes.
2. Classifying things like Spam/Ham; Being able to work around and complete tasks repetitively like Data Mining; Working through a lot of data like Finding Fastest Routes; And multiple other complex problems.
3. A set of training data with predetermined labels to work with.
4. Regression and Classification
5. KNN, Clustering, Rule Learning, Anomaly Detection
6. Would need to start small with a basic map and work its way up so it should be semi supervised at first and continue with more rigorous training with reinforcement learning as a reward system for success.
7. KNN clustering because it uses the information around a certain point to fit them into categories.
8. Supervised because it uses examples to judge what spam is.
9. It is feeding information sequentially either individually or in batches to basically learn on the go.
10. This type of learning is used with big data where it will take segments at a time, train it, and do another segment.
11. Instance-Based Learning
12. A model uses model parameters to predict while the hyperparameters are part of the learning algorithm that found the most optimal model. The hyperparameters aren't part of the model just the algorithm.
13. Model-based learning makes models based on examples to make predictions with. It makes predictions on training data and optimizes based on the correct predictions made.
14. Finding the right algorithm, Bad data, Not enough data, Overfitting data
15. The data is overfitting the training data and cannot work with new information. A way to correct it is to use less features, get more data, do a different model.
16. A test set is something you use to test the functionality of the model in use. It will help to see how efficient or inefficient the model is before putting it to use on something more extreme.
17. When training multiple models with different hyperparameters, a validation set compares the models to find which is more optimal.
18. Hyperparameters are a part of the algorithm so tuning them to the test set will cause more wrong predictions when put through other tests due to overfitting to that particular set the hyperparameters were tuned to.
19. Cross-validation is comparing models without multiple validation sets. So instead of using multiple validation sets it is possible to just cross-validate and get the same result.