Lec Mon

Wednesday, September 26, 2018 11:06 PM

Introduction to Machine Learning
Kai-Wei Chang
kw+cm146@kwchang.net
OH Mon 3-4 EVI 374

Machine Learning is the study of algorithms that improve their performance P at some task T with experience E.
Ex:
T: Playing Checkers
P: Percentage of games won against arbitrary opponent
E: Playing practice games against itself

Ex:
Spam Detection
- A **binary classification** task
- Assign one of two **labels** (yes/no) to the **input** (an email)
- Classification requires a **model** (a **classifier**) to determine which label to assign

In this class, we study **algorithms** to **learn** (**train**) models from data

We'll cover:
- Supervised Learning - given **labeled** examples, goal: learn mapping that predicts label for test instances
    - Decision Tree, Perceptron, Linear Models, SVM's, Kernel Methods
    - Learning Theory
- Unsupervised Learning - given **unlabeled** input, goal: learn some intrinsic structure in inputs
    - Clustering, Hidden Markov Models
    - EM Algorithms
- Practical Issues
    - Experimental Evaluation, Implementing ML Models

Won't Cover:
- Reinforcement Learning - given sequence of states and actions with rewards
    - goal: learn policy that maximizes agent's reward
    - like how people train a mouse - it goes through a maze and looks for cheese

<u>Framing a Learning Problem</u>
- need to define an **instance** and represent it by machine
    - and what **features** we'll use for making predictions
- train a **classifier** that **learns** (**generalizes**) based on features that can make predictions

<u>Challenges in Machine Learning</u>
- Representation - how to represent input/output?
- What is the right model?
    - depends on size of data, type of problem, prior knowledge, annotation quality…
    - depends on goal: model size, test-time budget, accuracy vs. speed
- Debugging - bugs can come from your implementation or your design and within each, lots of possible causes
- Structured Inference
- Robustness - lots of hard edge cases you might not be prepared for

- Adversarial Attack - can make targeted changes to an image that messes with the prediction