Homework 1 of CS 165A (Spring 2022)

University of California, Santa Barbara

To be discussed on Apr 15, 2022 (Friday)

Notes:

- The homework is optional. You do not need to submit your solutions anywhere and you will not be evaluated by these.
- To maximize your learning, you should try understanding the problems and try solving them as much as you can before the discussion class.
- Feel free to discuss with your peers / form small groups to solve these problems.
- Feel free to discuss any questions with the instructor and the TA in office hours or on Piazza.

---

1 Why should I do this homework?

This homework is about intelligent agent in general, classifier agents and machine learning. In Problem 1, you will practice organizing your thoughts, coming up with descriptions of agents. In Problem 2, you will zoom into classifier agents and develop an understanding of the idea of a "decision boundary". In Problem 3, you will do a simple theoretical exercise to see the gist of statistical learning theory, and to understand how data splitting works. Understanding these questions may help you in the midterm.

2 Homework problems

Problem 1. Intelligent agents, rationality and descriptions of the environment.

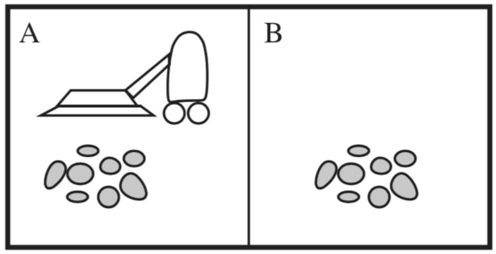


Figure 2.2 A vacuum-cleaner world with just two locations.

Percept sequence	Action
$[A, Clean]$	*Right*
$[A, Dirty]$	*Suck*
$[B, Clean]$	*Left*
$[B, Dirty]$	*Suck*
$[A, Clean], [A, Clean]$	*Right*
$[A, Clean], [A, Dirty]$	*Suck*
⋮	⋮
$[A, Clean], [A, Clean], [A, Clean]$	*Right*
$[A, Clean], [A, Clean], [A, Dirty]$	*Suck*
⋮	⋮

Figure 2.3 Partial tabulation of a simple agent function for the vacuum-cleaner world shown in Figure 2.2.

(a) Let us inspect the vacuum-cleaner example again from the textbook. Under the assumption listed on textbook page 38:

(i) Explain that why this vacuum-cleaner agent function described in Figure 2.3 is rational.

(ii) If each movement of the cleaner generate a unit cost, explain that now a rational agent needs to maintain an internal state.

(iii) Back to the original problem, now there is a naughty pet dog in the environment. At each time step, each clean square has a 50% chance of becoming dirty. Briefly explain how can you modified the rational agent in this case.

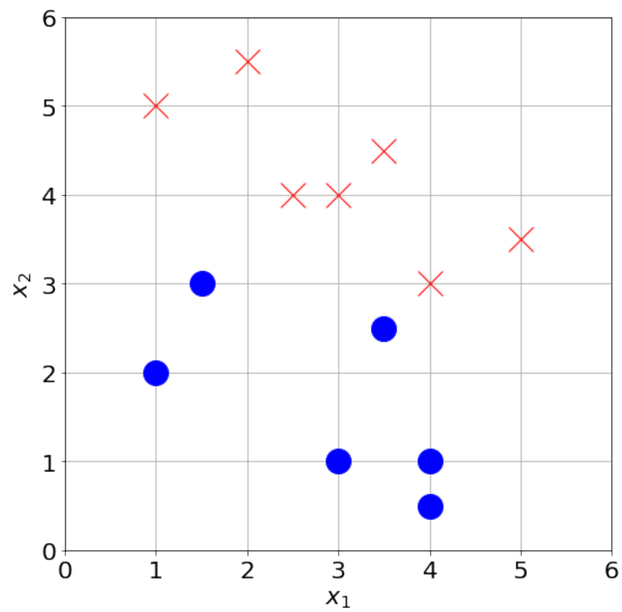
(b) For each of the following activities, give a **PEAS** description of the task environment and characterize it in terms of the properties listed in textbook Section 2.3.2.

(i) Playing soccer as a robot.

(ii) Internet laptop-shopping agent. (an intelligent agent that helps a customer to shop for a laptop on a website.)

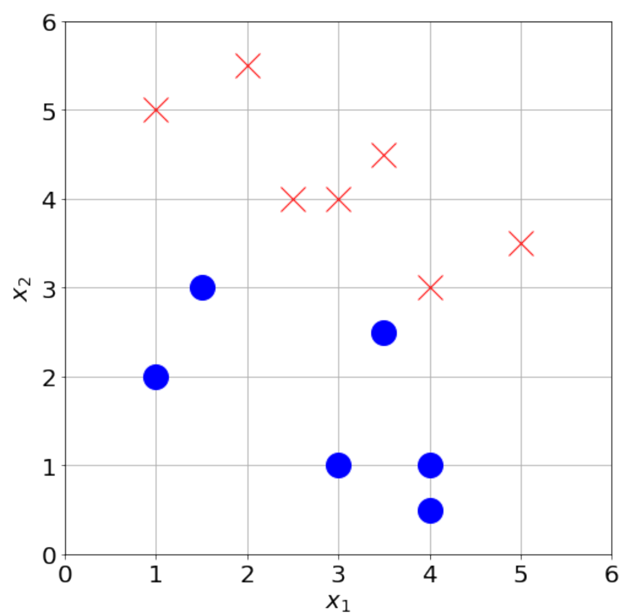
Problem 2. Classifiers and Machine Learning.

(a) What are the P,E,A,S description of a classifier agent?



(b) Let the feature space be $\mathbb{R}^2$, i.e., two continuous features. Consider a decision tree classifier with depth = 2 (branching on one variable by thresholding on each level, the variables to branch on can repeat). What are the possible boundaries representable by this classifier on a 2D feature plane? Is there a depth 2 decision tree that gives perfect classification in the example above?

(c) Draw the decision boundary of a 1-nearest neighbor classifier in the example above.



(d) (Challenge question) For a problem with a feature vector $x \in \mathcal{X} = \{0,1\}^d$ (Notice that all features are binary valued). How many unique binary classifiers are there in total in the entire universe (not necessarily linear classifiers or decision trees)?

Problem 3. (Training error, test error, Hold-out data and cross validation) The next step is to build some tool for evaluating the performance of a classifier learned by a machine learning algorithm.

Let h be a classifier. More formally, $h : \mathcal{X} \rightarrow \mathcal{Y}$, where $\mathcal{X}$ is the feature space and $\mathcal{Y}$ is the label space. And we denote the space of all classifiers to be $\mathcal{H}$. Let $(x_1, y_1), ..., (x_n, y_n)$ be the training dataset used for training, and assume that they are drawn i.i.d. (independently and identically distributed) from some unknown distributions $\mathcal{D}$ defined on $\mathcal{X} \times \mathcal{Y}$.

The most natural performance metric is the classification error, which measures the expected error rate.

$$\mathrm{Err}(h) = \mathbb{E}_{(x,y)\sim\mathcal{D}}[\mathbf{1}(h(x) \neq y)]$$

where $\mathbf{1}(\cdot)$ is the indicator function that outputs 1 is the condition is true and 0 otherwise.

We can also define the error that we calculate on a dataset. We denote the empirical error on this data set as

$$\hat{\mathrm{Err}}(h) = \frac{1}{n}\sum_{i=1}^{n} \mathbf{1}(h(x) \neq y).$$

Furthermore, we define the generalization error to be

$$\mathrm{GenErr} = \max_{h\in\mathcal{H}} |Err(h) - \hat{Err}(h)|$$

— the difference between the training error and the *expected* error on a **new data point** from the same distribution $\mathcal{D}$ on **all** classifier $h \in \mathcal{H}$.

(a) Let $\hat{h}$ be the learned classifier and h^* be the optimal classifier., i.e.

$$\hat{h} = \arg\min_{h} \hat{\mathrm{Err}}(h), \quad h^* = \arg\min_{h} \mathrm{Err}(h)$$

Prove that:

$$\mathrm{Err}(\hat{h}) \leq \mathrm{Err}(h^*) + 2\mathrm{GenErr}.$$

(b) In practice, we can evaluate a classifier with data splitting. We randomly partition the data into a training data set and a holdout dataset.

Show that

$$\mathbb{E}[\hat{\mathrm{Err}}_{\mathrm{Holdout}}(\hat{h})] = \mathbb{E}[\mathrm{Err}(\hat{h})],$$

where $\hat{\mathrm{Err}}_{Holdout}$ denotes the empirical error calculated on the holdout set. Notice that $\hat{h}$ is random (because its training data is random), but here let us fix $\hat{h}$ and take the expectation over only the distribution of the *relevant* part of the dataset that is assumed to be drawn iid.