< Previous   ■◀✓   ▤✓   ■◀✓   ■◀✓   ☑✓   ☑✓   ☑✓   Next >

5. A Concrete Example of a Supervised Learning Task

🔖 Bookmark this page

Movie Recommender Problem

▶ | 0:00 / 0:00| | | ▸ 1.25x | ◀» | ✖ | CC | 66

Now, let's take a very concrete example

of how to specify a machine learning task

and how to solve it.

Let's consider a movie recommender problem.

I have a set of movies I've already seen,

Video
Download video file

Transcripts
Download SubRip (.srt) file
Download Text (.txt) file

Feature Vector Demystified 1

1/1 point (graded)

We have a movie recommending system that reads description of each movie and determines some important characteristics of the movie. In particular, it examines whether each of the criterion below is true for that movie:

1. Is it a comedy movie?

2. Is it an action movie?

3. Was the movie directed by Spielberg?

4. Do dinosaurs appear in the movie?

5. Is it a Disney film?

For example, when the recommending system reads descriptions of "Jurassic Park", the answers for the five questions above will be "no, yes, yes, yes, no." On the other hand if the recommending system reads descriptions of "High School Musical", the answers will be "no, no, no, no, yes"

The system converts "yes" into 1, "no" into 0, and makes a feature vector X for each movie. So $X_{JurrasicPark}$ will be $[0, 1, 1, 1, 0]$, while $X_{HighSchoolMusical}$ will be $[0, 0, 0, 0, 1]$

Question 1: Now we have a comedy movie that is not an action movie, that was not directed by Spielberg, that does not have dinosaurs in it, but was produced by Disney. What is this movie's feature vector?

- 🔘 $[1, 0, 0, 0, 1]$

- ⚪ $[0, 0, 1, 0, 0]$

- ⚪ $[1, 0, 0, 1, 0]$

○ $[1, 1, 0, 0, 0]$

✔

Submit   You have used 1 of 3 attempts

Feature Vector Demystified 2

1/1 point (graded)
Question 2: What is the dimension of the feature vector of this movie?

5   ✔

Submit   You have used 2 of 3 attempts

Training Set vs Test Set 1

1/1 point (graded)
The ultimate goal of our recommending system is to predict whether John will like this movie. Now suppose our movie recommending system knows whether John likes or dislikes the following movies:

	comedy	action	Spielberg	Dinosaur Appearance	Disney	**Liked by John?**
movie 1	0	1	0	0	1	1
movie 2	1	1	1	0	0	-1
movie 3	0	1	0	1	1	1
movie 4	1	1	0	1	0	1

(Like is denoted as 1 and dislike as -1 in the above table) On the other hand, the movie recommender does not know whether John likes the following movies when building the model, but will know them after the model is built:

	comedy	action	Spielberg	Dinosaur Appearance	Disney	**Liked by John?**
movie 5	1	0	0	0	0	Don't know yet
movie 6	0	0	0	0	1	Don't know yet
movie 7	0	0	0	1	1	Don't know yet

Assume that, when John evaluates movies, he only does so based on the five criteria.

Question 1: What is the **label** of movie 1, based on the fact that John likes the movie?

◉ 1

○ -1

✔

Training Set vs Test Set 2

1/1 point (graded)
Question 2: What movies are in the **training set**? Select all those apply.

- [x] movie 1
- [x] movie 2
- [x] movie 3
- [x] movie 4
- [ ] movie 5
- [ ] movie 6
- [ ] movie 7

✔

Training Set vs Test Set 3

1/1 point (graded)
Question 3: What movies are in the **test set**? Select all those apply.

- [ ] movie 1
- [ ] movie 2
- [ ] movie 3
- [ ] movie 4
- [x] movie 5
- [x] movie 6
- [x] movie 7

✔

Discussion

Hide Discussion

Topic: Unit 1 Linear Classifiers and Generalizations (2 weeks):Lecture 1. Introduction to

Add a Post

Show all posts ⌄		by recent activity ⌄
?	**Facing issue in submitting the answers.** I am not able to submit my answers. When I click on submit button I am not getting any feedback or answers and also showing as 0 atte...	3
💬	**Is it me or the questions are a bit too obvious?** They are asking us which movies are in the training set?!! Please improve the level of the exercises. Note: I do not mean to be disrespect...	3
💬	**Ambiguous Question Statement (Question 3)** *Question 1: What is the label of movie 1, based on the fact that John likes the movie?* Movie 1 is in training set. It already has a label. D...	2
?	**How can we have the slides?** Hello :) I would like to annotate the slides with my own notes. Are the slides available anywhere for download? Thanks in advance!	3
?	**Is the test set completely unlabelled ?** Is the test set completely unlabelled? Or is it just unlabelled for the ML Developers, but its actually labelled and will be used to check the...	2
?	**is this standard 1 st class question as compare to Homework 0** @ Staff, can we compare Homework 0 question with these question, in Homework u didn't shown or put any documents, but these que...	5
☑	**Subjective feature vectors** Since the feature vectors are created manually, there would be some subjectivity involved. Wouldn't some standard criteria need to be e...	3
?	**What do I enter in the feature vector dimension question?** I don't understand if it should be a list a tuple or just a simple number? Like how do I enter the answer if the dimensions are j,k. I have al...	6

‹ Previous   **Next ›**

edX

edX

About

Affiliates

edX for Business

Open edX

Careers

News

Legal