

<u>Unit 2 Nonlinear Classification</u>, <u>Linear regression, Collaborative</u>

<u>Course</u> > <u>Filtering (2 weeks)</u>
7. Classification Using Manually
Crafted Features

> Project 2: Digit recognition (Part 1) >

## **Audit Access Expires May 11, 2020**

You lose all access to this course, including your progress, on May 11, 2020. Upgrade by Apr 1, 2020 to get unlimited access to the course as long as it exists on the site. **Upgrade now** 

## 7. Classification Using Manually Crafted Features

The performance of most learning algorithms depends heavily on the representation of the training data. In this section, we will try representing each image using different features in place of the raw pixel values. Subsequently, we will investigate how well our regression model from the previous section performs when fed different representations of the data.

## **Dimensionality Reduction via PCA**

Principal Components Analysis (PCA) is the most popular method for linear dimension reduction of data and is widely used in data analysis. For an in-depth exposition see: <a href="http://snobear.colorado.edu/Markw/BioMath/Otis/PCA/principal\_components.ps">http://snobear.colorado.edu/Markw/BioMath/Otis/PCA/principal\_components.ps</a>.

Briefly, this method finds (orthogonal) directions of maximal variation in the data. By projecting an  $n \times d$  dataset X onto  $k \leq d$  of these directions, we get a new dataset of lower dimension that reflects more variation in the original data than any other k-dimensional linear projection of X. By going through some linear algebra, it

can be proven that these directions are equal to the k eigenvectors corresponding to the k largest eigenvalues of the covariance matrix  $\widetilde{X}^T\widetilde{X}$ , where  $\widetilde{X}$  is a centered version of our original data.

**Remark:** The best implementations of PCA actually use the Singular Value Decomposition of  $\widetilde{X}$  rather than the more straightforward approach outlined here, but these concepts are beyond the scope of this course.

## **Cubic Features**

In this section, we will also work with a **cubic feature** mapping which maps an input vector  $x=[x_1,\ldots,x_d]$  into a new feature vector  $\phi(x)$ , defined so that for any  $x,x'\in\mathbb{R}^d$ :

$$\phi(x)^T\phi\left(x'\right) = \left(x^Tx'+1\right)^3$$

Discussion

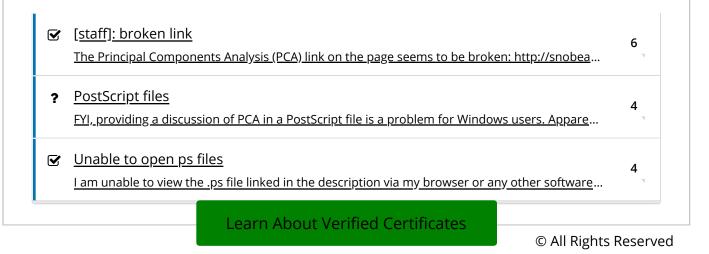
**Hide Discussion** 

**Topic:** Unit 2 Nonlinear Classification, Linear regression, Collaborative Filtering (2 weeks):Project 2: Digit recognition (Part 1) / 7. Classification Using Manually Crafted Features

Add a Post

Show all posts by recent	activity
<u>PCA explanation</u> <u>In Pinned</u>	1
Singular Value Decomposition I have found the article provided [above about Principal Components Analysis (PCA)][1] most  ▼ Pinned ▲ Community TA	3
? What is x'? maybe im missing something obvious, but can anyone please tell me what is x'?	4

7. Classification Using Manually Crafted Features ... https://courses.edx.org/courses/course-v1:MITx+...



3 of 3