

MITx 6.86x

Machine Learning with Python-From Linear Models to Deep Learning

Course **Progress** Discussion Dates Resources

A Course / Unit 2. Nonlinear Classification, Linear regression, ... / Project 2: Dig



10. Kernel Methods

 \square Bookmark this page

Project due Mar 15, 2023 08:59 -03 Past due

As you can see, implementing a direct mapping to the high-dimensional features is a lo an even higher dimensional feature mapping.) This is where the kernel trick becomes u

Recall the kernel perceptron algorithm we learned in the lecture. The weights $m{ heta}$ can be combination of features:

$$heta = \sum_{i=1}^n lpha^{(i)} y^{(i)} \phi\left(x^{(i)}
ight)$$

In the softmax regression fomulation, we can also apply this representation of the weig

$$heta_j = \sum_{i=1}^n lpha_j^{(i)} \phi\left(x^{(i)}
ight).$$

$$h\left(x
ight) = rac{1}{\sum_{j=1}^k e^{\left[heta_j\cdot\phi(x)/ au
ight]-c}} egin{bmatrix} e^{\left[heta_1\cdot\phi(x)/ au
ight]-c} \ e^{\left[heta_2\cdot\phi(x)/ au
ight]-c} \ dots \ e^{\left[heta_k\cdot\phi(x)/ au
ight]-c} \end{bmatrix}$$

$$h\left(x
ight) = rac{1}{\sum_{j=1}^{k} e^{\left[\sum_{i=1}^{n} lpha_{j}^{(i)} y^{(i)} \phi(x^{(i)}) \cdot \phi(x)/ au
ight] - c}} egin{bmatrix} e^{\left[\sum_{i=1}^{n} lpha_{1}^{(i)} y^{(i)} \phi(x^{(i)}) \cdot \phi(x)/ au
ight] - c} \ & e^{\left[\sum_{i=1}^{n} lpha_{2}^{(i)} y^{(i)} \phi(x^{(i)}) \cdot \phi(x)/ au
ight] - c} \ & dots \ e^{\left[\sum_{i=1}^{n} lpha_{k}^{(i)} y^{(i)} \phi(x^{(i)}) \cdot \phi(x)/ au
ight] - c} \end{bmatrix}$$

We actually do not need the real mapping $\phi(x)$, but the inner product between two few $\phi(x_i)\cdot\phi(x)$, where x_i is a point in the training set and x is the new data point for what the probability. If we can create a kernel function $K(x,y)=\phi(x)\cdot\phi(y)$, for any two then kernelize our softmax regression algorithm.

You will be working in the files part1/main.py and part1/kernel.py in this problem

```
compute the polynomial kernel between two matrices x and Y::
 5
               K(x, y) = (\langle x, y \rangle + c)^p
 6
           for each pair of rows x in X and y in Y.
 7
 8
           Args:
 9
               X - (n, d) NumPy array (n datapoints each with d features)
               Y - (m, d) NumPy array (m datapoints each with d features)
10
               c - a coefficient to trade off high-order and low-order terms
11
12
               p - the degree of the polynomial kernel
13
14
           Returns:
15
               kernel_matrix - (n, m) Numpy array containing the kernel matri
```

Press ESC then TAB or click outside of the code editor to exit

Unanswered

Submit

You have used 0 of 25 attempts

Gaussian RBF Kernel

1 point possible (graded)

Another commonly used kernel is the Gaussian RBF kenel. Similarly, write a function retwo matrices and and computes the RBF kernel for every pair of rows

Available Functions: You have access to the NumPy python library as np

```
1
 2 def rbf_kernel(X, Y, gamma):
 3
 4
          Compute the Gaussian RBF kernel between two matrices X and Y::
 5
               K(x, y) = \exp(-gamma ||x-y||^2)
 6
          for each pair of rows x in X and y in Y.
 7
 8
          Args:
9
               X - (n, d) NumPy array (n datapoints each with d features)
10
              Y - (m, d) NumPy array (m datapoints each with d features)
               gamma - the gamma parameter of gaussian function (scalar)
11
12
13
          Returns:
14
               kernel_matrix - (n, m) Numpy array containing the kernel matri
15
```

Press ESC then TAB or click outside of the code editor to exit

Unanswered

Submit

You have used 0 of 25 attempts

kernelized features.

In the next project, you will apply neural networks to this task.

Previous

Next >

Discussion



edX

About

Affiliates

edX for Business

Open edX

Careers

News

Legal

Terms of Service & Honor Code

Privacy Policy

Accessibility Policy

Trademark Policy

Sitemap

Cookie Policy

Do Not Sell My Personal Information

Connect

Blog

Contact Us

Help Center

Security

Media Kit









