The Glowing Python

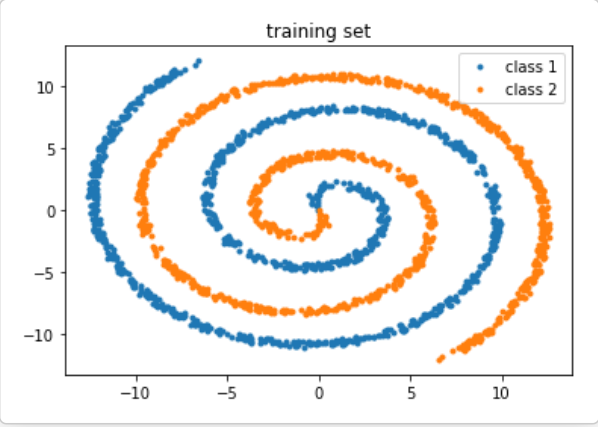
A collection of sloppy snippets for scientific computing and data visualization in Python.

Thursday, April 27, 2017

Solving the Two Spirals problem with Keras

In this post we will see how to create a Multi Layer Perceptron (MLP), one of the most common Neural Network architectures, with Keras. Then, we'll train the MLP to tell apart points from two different spirals in the same space.
To have a sense of the problem, let's first generate the data to train the network:

```
import numpy as np
import matplotlib.pyplot as plt

def twospirals(n_points, noise=.5):
    """
     Returns the two spirals dataset.
    """
    n = np.sqrt(np.random.rand(n_points,1)) * 780 * (2*np.pi)/360
    d1x = -np.cos(n)*n + np.random.rand(n_points,1) * noise
    d1y = np.sin(n)*n + np.random.rand(n_points,1) * noise
    return (np.vstack((np.hstack((d1x,d1y)),np.hstack((-d1x,-d1y)))),
            np.hstack((np.zeros(n_points),np.ones(n_points))))

X, y = twospirals(1000)

plt.title('training set')
plt.plot(X[y==0,0], X[y==0,1], '.', label='class 1')
plt.plot(X[y==1,0], X[y==1,1], '.', label='class 2')
plt.legend()
plt.show()
```



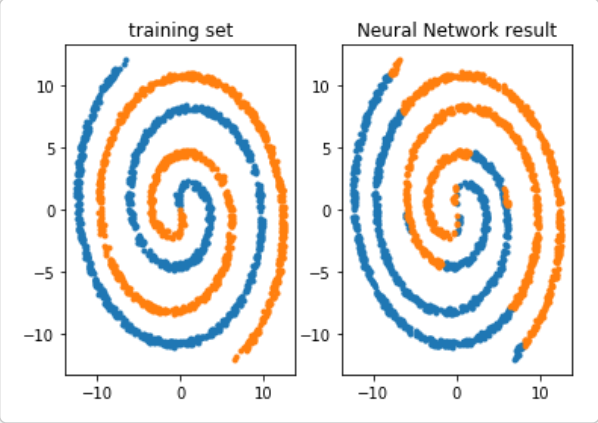
As we can see, this dataset contains two different spirals. This kind of dataset has been named as Worst Dataset Ever!, indeed telling apart the points from the two spirals is not an easy part if your MLP is not sophisticated enough. Let's build a simple MLP with Keras and see what we can achieve:

```
from keras.models import Sequential
from keras.layers import Dense

mymlp = Sequential()
mymlp.add(Dense(12, input_dim=2, activation='tanh'))
mymlp.add(Dense(1, activation='sigmoid'))

mymlp.compile(loss='binary_crossentropy',
              optimizer='rmsprop',
```

```
              metrics=['accuracy'])

# trains the model
mymlp.fit(X, y, epochs=150, batch_size=10,  verbose=0)
```

Here we created a Neural Network with the following structure: 2 inputs (the data is in a 2D space) fully connected to 12 hidden neurons and 1 output. Let's generate some test data and see if our model is able to classify them:

```
X_test, y_test = twospirals(1000)

yy = np.round(mymlp.predict(X_test).T[0])

plt.subplot(1,2,1)
plt.title('training set')
plt.plot(X[y==0,0], X[y==0,1], '.')
plt.plot(X[y==1,0], X[y==1,1], '.')
plt.subplot(1,2,2)
plt.title('Neural Network result')
plt.plot(X_test[yy==0,0], X_test[yy==0,1], '.')
plt.plot(X_test[yy==1,0], X_test[yy==1,1], '.')
plt.show()
```



We have the original train set on the left and the results of the Neural Network on the right. It's easy to note that the model misclassified most of the points on the test data. Let's add two hidden layers to our model and see what happens:

```
mymlp = Sequential()
mymlp.add(Dense(12, input_dim=2, activation='tanh'))
mymlp.add(Dense(12, activation='tanh'))
mymlp.add(Dense(12, activation='tanh'))
mymlp.add(Dense(1, activation='sigmoid'))

mymlp.compile(loss='binary_crossentropy',
              optimizer='rmsprop',
              metrics=['accuracy'])

# Fit the model
mymlp.fit(X, y, epochs=150, batch_size=10,  verbose=0)

yy = np.round(mymlp.predict(X_test).T[0])

plt.subplot(1,2,1)
plt.title('training set')
plt.plot(X[y==0,0], X[y==0,1], '.')
plt.plot(X[y==1,0], X[y==1,1], '.')
plt.subplot(1,2,2)
plt.title('Neural Network result')
plt.plot(X_test[yy==0,0], X_test[yy==0,1], '.')
plt.plot(X_test[yy==1,0], X_test[yy==1,1], '.')
plt.show()
```

scikit-learn sorting statistics visualization wav

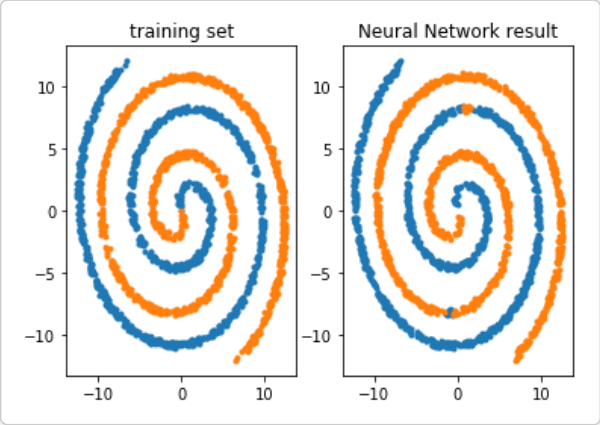
Followers

Seguidores (173) Siguiente

Seguir

Archive

► 2019 (9)
► 2018 (2)
▼ 2017 (4)
► Oct (1)
► Jul (1)
► Jun (1)
▼ Apr (1)
Solving the Two Spirals problem with Keras
► 2016 (1)
► 2015 (3)
► 2014 (9)
► 2013 (20)
► 2012 (23)
► 2011 (62)



The structure of our Network is now more suited to solve the problem and we see that most of the points used for the test were correctly classified.

Labels: keras, machine learning, neural network, spirals

3 comments:

ducatindia September 3, 2017 at 1:51 PM

This comment has been removed by the author.

Reply

Upender June 21, 2018 at 10:42 AM

This comment has been removed by the author.

Reply

Unknown July 5, 2018 at 12:18 AM

FYI https://www.researchgate.net/publication/287229974_Learning_to_Tell_Two_Spirals_Apart

Reply

Enter your comment...

Comment as: dbanshee@gm ▾ Sign out

Publish Preview Notify me

Newer Post Home Older Post

Subscribe to: Post Comments (Atom)

Quote

The greatest value of a picture is that it forces us to notice what we never expected to see

John Tukey