Homework 4

Theory Part

Q1 Consider a layer in CNN that takes in a single channel input of 64 × 64, and has 96 filters. In each of the following cases, compute the number of parameters that are learned in this layer. We assume that bias is present for each weight.

[1] A convolution layer with filters of same size as the input.

$$Param = (ksize^2 + 1) * channel$$
$$Param = (64^2 + 1) * 96$$
$$Param = 393312$$

In [1]:
```
(64**2 + 1)* 96
```

Out[1]: 393312

[2] A convolution layer with 8 × 8 filters with stride of 4

$$Param = (ksize)^2 * channel + bias$$
$$Param = (8^2 + 1) * 96$$
$$Param = 6240$$

In [2]:
```
(8**2 + 1) * 96
```

Out[2]: 6240

[3]. A convolution layer with 1 × 1 filter and a stride of 1

$$Param = (ksize)^2 * channel + bias$$
$$Param = (1^2 + 1) * 96$$
$$Param = 192$$

Q2 Suppose you would have a neuron which has an RBF kernel as activation function (remember the evil wolf? Drop your linear style of thoughts. Circumferential thoughts can be nice too.)

$$y = exp(-(x_1^2 + x_2^2)) + b$$

with parameter b. *What would be the shapes realized by the set of points {$(x1, x2) : y((x1, x2)) = 0$} as a function of b ? Explain in at most 2 sentences and/or a little math.*

$$0 = exp(-(x_1^2 + x_2^2)) + b$$

$$-b = exp(-(x_1^2 + x_2^2))$$

$$-ln(-b) = x_1^2 + x_2^2$$

Therefore, it is a circle centered around the origin with radius $\sqrt{-ln(-b)}$. Obviously this is only valid when $-1 < b < 0$.

Supposed now we add weight,

$$y = exp(-(w_1 x_1^2 + w_2 x_2^2)) + b$$

what shapes can we realize now? Explain in at most 5 sentences and/or a little math. You can make references to publicly available in the internet materials to explain.

$$-ln(-b) = w_1 x_1^2 + w_2 x_2^2$$

$$-ln(-b) = \frac{x_1^2}{w_1^{-1}} + \frac{x_2^2}{w_2^{-1}}$$

$$1 = \frac{x_1^2}{-ln(-b) \cdot w_1^{-1}} + \frac{x_2^2}{-ln(-b) \cdot w_2^{-1}}$$

$$1 = \left(\frac{x_1}{\sqrt{-ln(-b) \cdot w_1^{-1}}}\right)^2 + \left(\frac{x_2}{\sqrt{-ln(-b) \cdot w_2^{-1}}}\right)^2$$

hence it is an ellipse, centered around the origin with radius $\sqrt{-ln(-b) \cdot w_1^{-1}}$ along the x_1 axis and radius $\sqrt{-ln(-b) \cdot w_2^{-1}}$ along the x_2 axis.

Q3 Suppose you have five linear neurons neurons n1, . . . , n5, realizing above decision boundaries as shown in Figure 1. That is: for every decision boundary we have outputs are = 0.5 in the zones marked with red plusses, and = 0.2 in the zones marked with the blue minuses.

figure1

As you know, each neuron is realized by:

$$n_i = 0.3H(w_1^{(i)} x_1 + w_2^{(i)} x_2 + b^{(i)}) + 0.2, \quad H(z) \in 0, 1$$

where H is the threshold activation function. You want to predict positive values in a shape marked in green in Figure 1. You want to achieve this prediction by combining these neurons using a threshold neuron H:

$$y = H(\sum_i v_i^* n_i + b^*)$$

[1] _what do you have to do with the weights of n_5 so that you can move the decision boundary of n_5 so that you can realize the shape in green shown above (in the sense of having positive values inside and negative values outside.)? Give a qualitative description. Note: Give a qualitative description in 3 sentences at most. Note that there is an x- and an y-axis, which helps you to express vectors qualitatively._

The position of the decision boundary of n_5 depends on its weight and biases. Particularly, the ratio between w_1 and w_2 determines the slant of the boundary, while the ratio between the b and w_2 determines its offset from origin. As the desired position is a shift upwards (given that the boundary continues infinitely), we want to decrease b so that the boundaries shift upwards.

[2] _after moving the decision boundary of n5 appropriately, the green shape looks a bit like an logical AND-combination of the +-zones for every neuron. How to choose the weights v_i^* and the bias b^* in $y = H(\sum_i v_i^* n_i + b^*)$ so that you can realize the green shape (in the sense of having positive values inside and negative values outside that shape)? Note: n_i gives out values either 0.5 or 0.2_

Lets say that function H has a threshold $h = 0$, such that

$$H(z) = 1[z > h] = 1[z > 0]$$

If we were to take green area as {+1}, for the threshold neuron to fire $+1$, we will need all neurons n_i to fire 0.5,

$$1 = H(\sum_i 0.5v_i^* + b^*)$$

$$1 = 1[(\sum_i 0.5v_i^* + b^*) > 0]$$

$$(\sum_i 0.5v_i^* + b^*) > 0 \quad -- \quad (1)$$

If we were to take non-green area as {-1}, for the threshold neuron to fire -1, we will need at least one neuron n_j to fire 0.2,

$$0 = H(\sum_{i-1} v_i^* n_i + v_j^* n_j + b^*)$$

$$0 = H(\sum_{i-1} 0.5v_i^* + 0.2v_j^* + b^*)$$

$$0 = 1[\sum_{i-1} 0.5v_i^* + 0.2v_j^* + b^* > 0]$$

$$\sum_{i-1} 0.5v_i^* + 0.2v_j^* + b^* <= 0 \quad -- \quad (2)$$

For simplicity, we set $v_i^* = 1$ for all i. Finding b,

$$0.5 \cdot 5 + b^* > 0 \quad -- \quad (1)$$

$$2.5 + b^* > 0$$

$$b^* > -2.5$$

$$0.5 \cdot 4 + 0.2 + b^* <= 0 \quad -- \quad (2)$$

$$2.2 + b^* <= 0$$

$$b^* <= -2.2$$

$$therefore \quad -2.5 < b <= -2.2$$

we can pick any b within this range, e.g. $b = -2.4$ with our $v_i^* = 1$, such that

$$H(\sum_i 0.5 - 2.4) = H(2.5 - 2.4) \quad for\ all\ n_i = 0.5$$

$$= H(0.1) = 1[0.1 > 0] = 1 \quad (green)$$

$$H(\sum_{i-1} 0.5 + 0.2 - 2.4) = H(2.2 - 2.4) \quad for\ four\ n_i = 0.5\ and\ one\ n_j = 0.2$$

$$= H(-0.2) = 1[-0.2 > 0] = 0 \quad (outside)$$

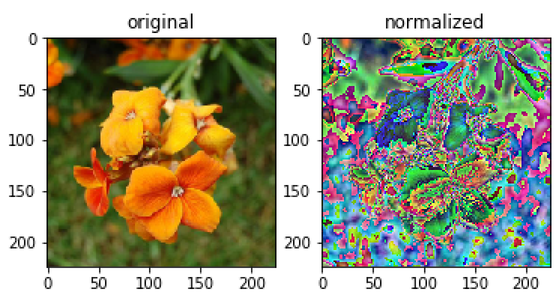
Coding - Part 1

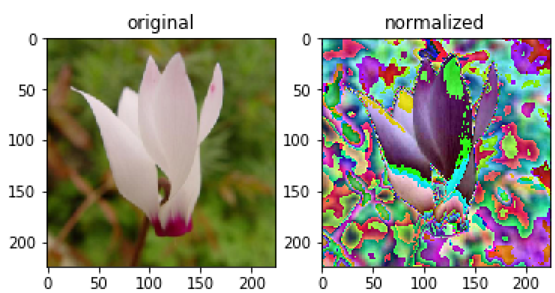
For this section, there is a script called `hw4_code.py` that will hold some definitions e.g. the `FlowerDataset` class and `train_model` function.

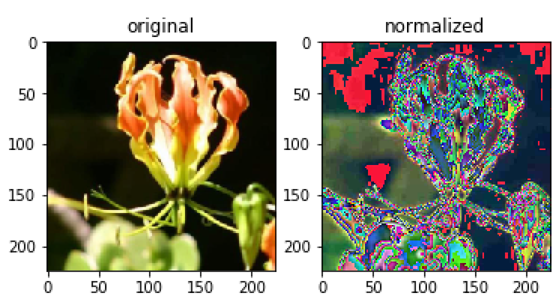
Below we will do some sanity check on the dataset.

In [7]:
```
from hw4_code import *
import matplotlib.pyplot as plt
%matplotlib inline
from torchvision import transforms

# testing dataset
flower_dataset = FlowerDataset('..\\datasets\\flowersstuff\\102flowers\\flowers_data', mode='train')
flower_dataset_val = FlowerDataset('..\\datasets\\flowersstuff\\102flowers\\flowers_data', mode='val')

for i in range(3):
    flower1 = flower_dataset[i]
    print('label:',flower1['label'])
    image = transforms.ToPILImage()(flower1['image'])

    invTrans = transforms.Compose([ transforms.Normalize(mean = [ 0., 0., 0. ],
                                                         std = [ 1/0.229, 1/0.224, 1/0.225 ]),
                                    transforms.Normalize(mean = [ -0.485, -0.456, -0.406 ],
                                                         std = [ 1., 1., 1. ]),
                                    transforms.ToPILImage(),
                                   ])
    clear_image = invTrans(flower1['image'])

    plt.figure()
    plt.subplot(121)
    plt.imshow(clear_image)
    plt.title('original')
    plt.subplot(122)
    plt.imshow(image)
    plt.title('normalized')
```

```
label: 45
label: 87
label: 20
```







We will then use resnet models to classify our flowers.

In [8]:

```
# Getting pretrained resnet
from torchvision import models

def get_pretrained_resnet(use_gpu=True):
    model = models.resnet18(pretrained=True)
    if use_gpu:
        model = model.cuda(0)
    return model

resnetmodel = get_pretrained_resnet()
resnet_dict = resnetmodel.state_dict()
```

In [9]:

```
# training with various learn rate
for lr in [0.1, 1, 10]:
    print("\n\nWith learning rate {},\n".format(lr))
    learnrate = lr
    optimizer = optim.SGD(resnetmodel.parameters(), lr=learnrate, momentum=0.9)

    trainedmodel = train_model(flower_dataset, resnetmodel, optimizer,
                               epoch=5, mode='train', use_gpu=True, print_every=1)
    trainedmodel = train_model(flower_dataset_val, trainedmodel, optimizer,
                               epoch=5, mode='val', use_gpu=True, print_every=1)
```

With learning rate 0.1,

```
[train] - Epoch 0..
      >> Epoch loss 1.84182 accuracy 0.034                      in 442.9377s
[train] - Epoch 1..
      >> Epoch loss 1.19127 accuracy 0.195                      in 448.1263s
[train] - Epoch 2..
      >> Epoch loss 0.84649 accuracy 0.355                      in 434.1745s
[train] - Epoch 3..
      >> Epoch loss 0.63054 accuracy 0.498                      in 434.5283s
[train] - Epoch 4..
      >> Epoch loss 0.48427 accuracy 0.627                      in 435.3186s
[val] - Epoch 0..
      >> Epoch loss 0.50609 accuracy 0.561                      in 28.6890s
[val] - Epoch 1..
      >> Epoch loss 0.50609 accuracy 0.561                      in 29.2429s
[val] - Epoch 2..
      >> Epoch loss 0.50609 accuracy 0.561                      in 28.0560s
[val] - Epoch 3..
      >> Epoch loss 0.50609 accuracy 0.561                      in 27.9266s
[val] - Epoch 4..
      >> Epoch loss 0.50609 accuracy 0.561                      in 28.2003s
```

With learning rate 1,

```
[train] - Epoch 0..
      >> Epoch loss 0.32680 accuracy 0.739                      in 435.5767s
[train] - Epoch 1..
      >> Epoch loss 0.06936 accuracy 0.977                      in 438.4414s
[train] - Epoch 2..
      >> Epoch loss 0.02621 accuracy 0.999                      in 434.8344s
[train] - Epoch 3..
      >> Epoch loss 0.01594 accuracy 1.000                      in 432.8960s
[train] - Epoch 4..
      >> Epoch loss 0.01160 accuracy 1.000                      in 435.0222s
[val] - Epoch 0..
      >> Epoch loss 0.11504 accuracy 0.879                      in 27.7213s
[val] - Epoch 1..
      >> Epoch loss 0.11504 accuracy 0.879                      in 27.9815s
[val] - Epoch 2..
      >> Epoch loss 0.11504 accuracy 0.879                      in 27.9220s
[val] - Epoch 3..
      >> Epoch loss 0.11504 accuracy 0.879                      in 27.9824s
[val] - Epoch 4..
      >> Epoch loss 0.11504 accuracy 0.879                      in 28.0526s
```

With learning rate 10,

```
[train] - Epoch 0..
      >> Epoch loss 0.14960 accuracy 0.876                      in 435.1083s
[train] - Epoch 1..
      >> Epoch loss 0.28600 accuracy 0.707                      in 435.4502s
[train] - Epoch 2..
      >> Epoch loss 0.04029 accuracy 0.961                      in 438.7462s
[train] - Epoch 3..
      >> Epoch loss 0.00938 accuracy 0.994                      in 440.7829s
[train] - Epoch 4..
      >> Epoch loss 0.00379 accuracy 0.998                      in 437.6852s
[val] - Epoch 0..
      >> Epoch loss 0.04910 accuracy 0.945                      in 27.5148s
[val] - Epoch 1..
      >> Epoch loss 0.04910 accuracy 0.945                      in 27.5783s
[val] - Epoch 2..
      >> Epoch loss 0.04910 accuracy 0.945                      in 27.5548s
[val] - Epoch 3..
      >> Epoch loss 0.04910 accuracy 0.945                      in 27.5676s
[val] - Epoch 4..
      >> Epoch loss 0.04910 accuracy 0.945                      in 27.4915s
```

In [11]:
```python
# now lets do the same thing, but train with an empty resnet

def get_empty_resnet(use_gpu=True):
    model = models.resnet18(pretrained=False)
    if use_gpu:
        model = model.cuda(0)
    return model

emptymodel = get_pretrained_resnet()
empty_dict = resnetmodel.state_dict()
```

In [13]:
```python
# training with various learn rate
for lr in [0.1, 1, 10]:
    print("\n\nWith learning rate {},\n".format(lr))
    learnrate = lr
    optimizer = optim.SGD(emptymodel.parameters(), lr=learnrate, momentum=0.9)

    trainedmodel = train_model(flower_dataset, emptymodel, optimizer,
                               epoch=5, mode='train', use_gpu=True, print_every=1)
    trainedmodel = train_model(flower_dataset_val, trainedmodel, optimizer,
                               epoch=1, mode='val', use_gpu=True, print_every=1)
```

```
With learning rate 0.1,

[train] - Epoch 0..
      >> Epoch loss 0.82801 accuracy 0.366                      in 442.9514s
[train] - Epoch 1..
      >> Epoch loss 0.61871 accuracy 0.508                      in 445.8555s
[train] - Epoch 2..
      >> Epoch loss 0.47620 accuracy 0.635                      in 433.9480s
[train] - Epoch 3..
      >> Epoch loss 0.37526 accuracy 0.734                      in 435.6844s
[train] - Epoch 4..
      >> Epoch loss 0.30134 accuracy 0.807                      in 436.9786s
[val] - Epoch 0..
      >> Epoch loss 0.37327 accuracy 0.649                      in 28.3487s


With learning rate 1,

[train] - Epoch 0..
      >> Epoch loss 0.23467 accuracy 0.837                      in 437.0729s
[train] - Epoch 1..
      >> Epoch loss 0.05134 accuracy 0.989                      in 435.1459s
[train] - Epoch 2..
      >> Epoch loss 0.02194 accuracy 1.000                      in 437.3835s
[train] - Epoch 3..
      >> Epoch loss 0.01423 accuracy 1.000                      in 433.7015s
[train] - Epoch 4..
      >> Epoch loss 0.01061 accuracy 1.000                      in 433.5076s
[val] - Epoch 0..
      >> Epoch loss 0.12527 accuracy 0.875                      in 27.4427s


With learning rate 10,

[train] - Epoch 0..
      >> Epoch loss 0.08040 accuracy 0.939                      in 435.1173s
[train] - Epoch 1..
      >> Epoch loss 0.35823 accuracy 0.642                      in 438.3859s
[train] - Epoch 2..
      >> Epoch loss 0.05526 accuracy 0.942                      in 435.5303s
[train] - Epoch 3..
      >> Epoch loss 0.01359 accuracy 0.988                      in 436.0038s
[train] - Epoch 4..
      >> Epoch loss 0.00421 accuracy 0.998                      in 436.5753s
[val] - Epoch 0..
      >> Epoch loss 0.05358 accuracy 0.944                      in 27.9151s
```

In [16]:
```python
# now lets do the same thing, but train with an unfrozen resnet
import torch.nn as nn

def get_unfrozen_resnet(numcl, use_gpu=True):
    model = models.resnet18(pretrained=False)
    num_ftrs = model.fc.in_features
    model.fc = nn.Linear(num_ftrs, numcl)
    if use_gpu:
        model = model.cuda(0)
    return model

unfrozenmodel = get_pretrained_resnet()
unfrozen_dict = resnetmodel.state_dict()
```

In [17]:
```python
# training with various learn rate
for lr in [0.1, 1, 10]:
    print("\n\nWith learning rate {},\n".format(lr))
    learnrate = lr
    optimizer = optim.SGD(unfrozenmodel.parameters(), lr=learnrate, momentum=0.9)

    trainedmodel = train_model(flower_dataset, unfrozenmodel, optimizer,
                               epoch=5, mode='train', use_gpu=True, print_every=1)
    trainedmodel = train_model(flower_dataset_val, trainedmodel, optimizer,
                               epoch=1, mode='val', use_gpu=True, print_every=1)
```

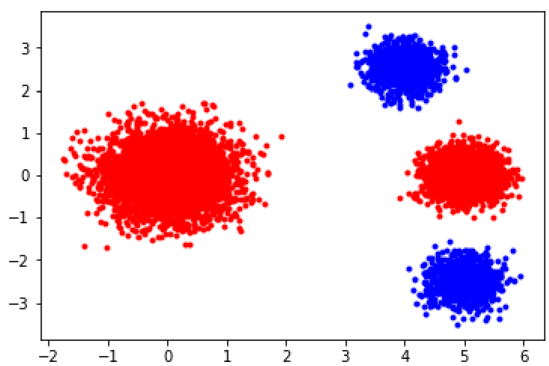
```
With learning rate 0.1,

[train] - Epoch 0..
      >> Epoch loss 1.84156 accuracy 0.033                      in 436.6851s
[train] - Epoch 1..
      >> Epoch loss 1.19135 accuracy 0.194                      in 438.1763s
[train] - Epoch 2..
      >> Epoch loss 0.84679 accuracy 0.352                      in 439.2018s
[train] - Epoch 3..
      >> Epoch loss 0.63122 accuracy 0.499                      in 439.1671s
[train] - Epoch 4..
      >> Epoch loss 0.48525 accuracy 0.624                      in 438.4431s
[val] - Epoch 0..
      >> Epoch loss 0.50701 accuracy 0.557                      in 28.5880s


With learning rate 1,

[train] - Epoch 0..
      >> Epoch loss 0.32747 accuracy 0.740                      in 438.2457s
[train] - Epoch 1..
      >> Epoch loss 0.06945 accuracy 0.977                      in 437.4092s
[train] - Epoch 2..
      >> Epoch loss 0.02642 accuracy 0.999                      in 435.1354s
[train] - Epoch 3..
      >> Epoch loss 0.01607 accuracy 1.000                      in 436.9965s
[train] - Epoch 4..
      >> Epoch loss 0.01164 accuracy 1.000                      in 437.0905s
[val] - Epoch 0..
      >> Epoch loss 0.11233 accuracy 0.877                      in 27.9271s


With learning rate 10,

[train] - Epoch 0..
      >> Epoch loss 0.21452 accuracy 0.814                      in 437.2194s
[train] - Epoch 1..
      >> Epoch loss 0.25315 accuracy 0.732                      in 436.8206s
[train] - Epoch 2..
      >> Epoch loss 0.03555 accuracy 0.967                      in 436.9103s
[train] - Epoch 3..
      >> Epoch loss 0.00800 accuracy 0.996                      in 435.5079s
[train] - Epoch 4..
      >> Epoch loss 0.00257 accuracy 0.999                      in 435.8060s
[val] - Epoch 0..
      >> Epoch loss 0.05316 accuracy 0.948                      in 28.1246s
```

Coding Part 2

In [101]:
```
from hw4_code2 import *
import numpy as np

def plot_dataset(data, label):
    true_mask = label.reshape((-1)).astype(bool)
    false_mask = (1-label).reshape((-1)).astype(bool)

    data_true = np.transpose(data[true_mask])
    data_false = np.transpose(data[false_mask])

    print('samples with positive class:', len(data_true[0]),
          '\nsamples with negative class:', len(data_false[0]))

    plt.figure()
    plt.plot(data_true[0], data_true[1], '.b')
    plt.plot(data_false[0], data_false[1], '.r')
```

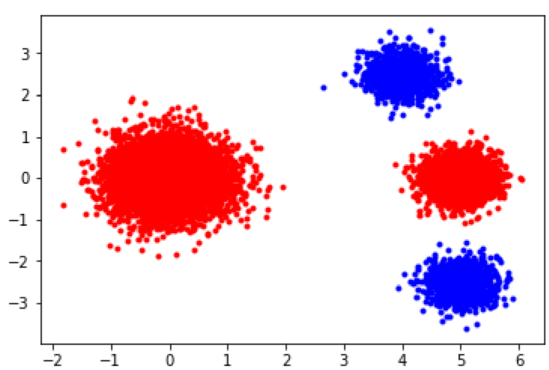
In [102]:
```
print("from samplestr.txt:")
data_tr, label_tr = from_text('samplestr.txt')
plot_dataset(data_tr, label_tr)
```

```
from samplestr.txt:
samples with positive class: 2000
samples with negative class: 8000
```



In [103]:
```
print("from sampleste.txt:")
data_te, label_te = from_text('sampleste.txt')
plot_dataset(data_te, label_te)
```

```
from sampleste.txt:
samples with positive class: 2000
samples with negative class: 8000
```



In [104]:

```python
'''
Creating a new Neural Network
'''
from torch import nn
import torch.nn.functional as F

class OurNet(nn.Module):
    def __init__(self):
        super(OurNet, self).__init__()
        self.fc = nn.Linear(2, 1)

    def forward(self, x):
        x = F.relu(self.fc(x))
        return x

our_net = OurNet()
print(our_net)
```

```
OurNet(
  (fc): Linear(in_features=2, out_features=1, bias=True)
)
```

In [105]:

```python
# testing the Dataset subclass from hw4_code2

train_set = ImbaDataset('.', mode='train')
print('train:',train_set[2])
test_set = ImbaDataset('.', mode='test')
print('test:', test_set[2])
```

```
train: {'data': tensor([ 0.7959,  0.3896], dtype=torch.float64), 'label': tensor([ 0.], dtype=torch.float64)}
test: {'data': tensor([-0.5324, -0.7683], dtype=torch.float64), 'label': tensor([ 0.], dtype=torch.float64)}
```

In [106]:

```
# simple training
import torch
from torch.utils.data import DataLoader
import torch.optim as optim
from torch.autograd import Variable
from torch.utils.data.sampler import WeightedRandomSampler

our_net = OurNet().double()
optimizer = optim.SGD(our_net.parameters(), lr=0.1)
criterion = nn.BCEWithLogitsLoss()

def train_ournet(dataset, model, optimizer, criterion, batch_sampler=None,
                 mode='train', balanced_acc=True):
    print_truecount = False#True
    if batch_sampler:
        loader = DataLoader(dataset, batch_sampler=batch_sampler)
    else:
        loader = DataLoader(dataset, batch_size=128)
    running_loss = 0
    running_corrects = 0
    total_data = 0

    # for balanced sampling
    true_pos = 0
    true_neg = 0
    total_positive = 0
    total_negative = 0

    if mode == 'train':
        model.train()
    else:
        model.eval()

    for inputdata in loader:
        data = Variable(inputdata['data'])
        labels = Variable(inputdata['label']).view(-1,1)
        if print_truecount:
            print('First batch true count:', labels.sum().item(), 'out of', len(labels),
                  "({}%)".format(labels.sum().item() * 100 / len(labels)))
            print_truecount = False

        outputs = model.forward(data)
        predictions = (outputs > 0)

        # balanced stuff
        true_pos += (predictions.double() * labels.double()).sum().item()
        true_neg += ((1-predictions).double() * (1-labels).double()).sum().item()
        total_positive += labels.sum().item()
        total_negative += (1-labels).sum().item()

        loss = criterion(outputs, labels)
        if mode == 'train':
            loss.backward()
            optimizer.step()

        # balanced stuff
#         total_data += (labels * (positive_weight-1) + 1).sum().item()
        total_data += len(labels)
        corrects = (predictions.double().cpu() == labels.cpu()).double()
#         weights = (positive_weight - 1) * labels.cpu() + 1

        running_corrects += (corrects).sum().item()
        running_loss += loss.item()


#     print("total data:", total_data)
    running_corrects /= (total_data)#float(128*100)
    if mode == 'train':
        print("Trained with accuracy {} and loss {}".format(running_corrects, running_loss))
    elif mode == 'test':
        print("Tested with accuracy {} and loss {}".format(running_corrects, running_loss))

    if balanced_acc:
        true_pos_rate = true_pos/total_positive
        true_neg_rate = true_neg/(total_negative)
        b_acc = (true_pos_rate + true_neg_rate) / 2
```

```
#         print("   True Positive:", true_pos)
#         print("   True Negative:", true_neg)
        print("   True Positive Rate:", true_pos_rate)
        print("   True Negative Rate:", true_neg_rate)
        print("   Balanced Accuracy:", b_acc)
        print('-----')
    return model
```

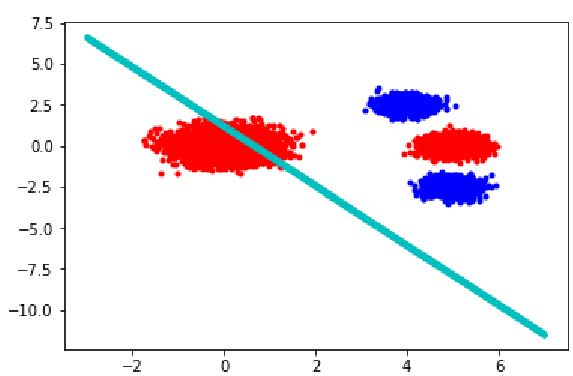
In [107]:
```
our_net = OurNet().double()

our_net_trained = train_ournet(train_set, our_net, optimizer, criterion)
our_net_trained = train_ournet(train_set, our_net_trained, optimizer, criterion, mode='test')
our_net_trained = train_ournet(test_set, our_net_trained, optimizer, criterion, mode='test')
print(our_net_trained.state_dict())
```

```
Trained with accuracy 0.6371 and loss 65.16015314912875
   True Positive Rate: 1.0
   True Negative Rate: 0.546375
   Balanced Accuracy: 0.7731875
-----
Tested with accuracy 0.6371 and loss 65.16015314912875
   True Positive Rate: 1.0
   True Negative Rate: 0.546375
   Balanced Accuracy: 0.7731875
-----
Tested with accuracy 0.6395 and loss 65.14335670596058
   True Positive Rate: 1.0
   True Negative Rate: 0.549375
   Balanced Accuracy: 0.7746875
-----
OrderedDict([('fc.weight', tensor([[ 0.2524,  0.1391]], dtype=torch.float64)), ('fc.bias', tensor([-0.1658],
dtype=torch.float64))])
```

In [108]:
```
w = our_net_trained.state_dict()['fc.weight']
b = our_net_trained.state_dict()['fc.bias']
w1 = w[0,0].item()
w2 = w[0,1].item()
b1 = b[0].item()

m = -w1/w2
c = -b/w2

line_x = [ i/100 for i in range(-300, 700)]
line_y = [ m*i + c for i in line_x]

plot_dataset(data_tr, label_tr)
plt.plot(line_x, line_y, '.c')
```

```
samples with positive class: 2000
samples with negative class: 8000
```

Out[108]: `[<matplotlib.lines.Line2D at 0x203efea45c0>]`



To achieve 50-50 on minibatches, we need a Sampler subclass.

In [109]:
```python
class OurBatchSampler(object):
    r"""
    Special batch sampler class that ensures a 50-50
    dataset classes.
    """

    def __init__(self, dataset, batch_size, iteration, random=False):
        self.dataset = dataset
        self.batch_size = batch_size
        self.iteration = iteration
        self.random = random

        self.idx_plus = [idx for idx in range(len(dataset)) if dataset[idx]['label'].item() > 0]
        self.idx_minus = [idx for idx in range(len(dataset)) if idx not in self.idx_plus]

    def __iter__(self):
        batch = []
        for i in range(self.iteration):
            if self.random:
                idplus = np.random.choice(self.idx_plus, self.batch_size // 2).tolist()
                idminus = np.random.choice(self.idx_minus, self.batch_size // 2).tolist()
            else:
                idplus = [self.get_positive_det() for i in range(self.batch_size // 2)]
                idminus = [self.get_negative_det() for i in range(self.batch_size // 2)]
            batch += idplus + idminus

            yield batch
            batch = []

    def get_positive_det(self):
        idx = self.idx_plus[0]
        self.idx_plus = self.idx_plus[1:] + [idx]
        return idx

    def get_negative_det(self):
        idx = self.idx_minus[0]
        self.idx_minus = self.idx_minus[1:] + [idx]
        return idx

    def __len__(self):
        return self.batch_size * self.iteration

train_sampler = OurBatchSampler(train_set, 128, 100)
print('train sampler built')
test_sampler = OurBatchSampler(test_set, 128, 100)
print('test sampler built')
```

```
train sampler built
test sampler built
```

In [110]:
```
our_net = OurNet().double()
print('\nWithout 50-50 sampling:')
our_net_trained = train_ournet(train_set, our_net, optimizer, criterion)
# print('Train: accuracy')
# our_net_trained = train_ournet(train_set, our_net_trained, optimizer,
#                                criterion, mode='test')
print('Test: accuracy')
our_net_trained = train_ournet(test_set, our_net_trained, optimizer,
                               criterion, mode='test')

our_net = OurNet().double()
print('\n\n\n')
print('With 50-50 sampling:')
our_net_trained = train_ournet(train_set, our_net, optimizer, criterion,
                               batch_sampler=train_sampler)
# print('Train: accuracy')
# our_net_trained = train_ournet(train_set, our_net_trained, optimizer,
#                                criterion, mode='test')
print('Test: accuracy')
our_net_trained = train_ournet(test_set, our_net_trained, optimizer,
                               criterion, mode='test')
# print(our_net_trained.state_dict())
```

```
Without 50-50 sampling:
Trained with accuracy 0.6981 and loss 55.94245333436154
   True Positive Rate: 0.0
   True Negative Rate: 0.872625
   Balanced Accuracy: 0.4363125
-----
Test: accuracy
Tested with accuracy 0.6986 and loss 55.89674747519486
   True Positive Rate: 0.0
   True Negative Rate: 0.87325
   Balanced Accuracy: 0.436625
-----




With 50-50 sampling:
Trained with accuracy 0.47140625 and loss 62.13626624408322
   True Positive Rate: 0.46875
   True Negative Rate: 0.4740625
   Balanced Accuracy: 0.47140625
-----
Test: accuracy
Tested with accuracy 0.4954 and loss 54.08188945714033
   True Positive Rate: 0.5
   True Negative Rate: 0.49425
   Balanced Accuracy: 0.49712500000000004
-----
```

In [111]:
```
w = our_net_trained.state_dict()['fc.weight']
b = our_net_trained.state_dict()['fc.bias']
w1 = w[0,0].item()
w2 = w[0,1].item()
b1 = b[0].item()

m = -w1/w2
c = -b/w2

line_x = [ i/100 for i in range(-300, 700)]
line_y = [ m*i + c for i in line_x]

plot_dataset(data_tr, label_tr)
plt.plot(line_x, line_y, '.c')
```

```
samples with positive class: 2000
samples with negative class: 8000
```

Out[111]:
```
[<matplotlib.lines.Line2D at 0x203efe9dc18>]
```