186

Activity 16: Support Vector Machines

Mary Chris Go / 2014 11122

Goal

To find the best decision line between two classes

Mary Chris Go / 2014 11122

SVM Algorithm

$$g(x)=w_o+x^T w=0$$

The same with the perceptron algorithm, an SVM algorithm calculates the decision line between two linearly separable classes clustered in feature space. But SVM assumes the best decision line as it offers the best generalization when testing using the novel data. The following formula above is the best decision line. It has excellent generalization properties.

The perpendicular distance z is given by this formula:

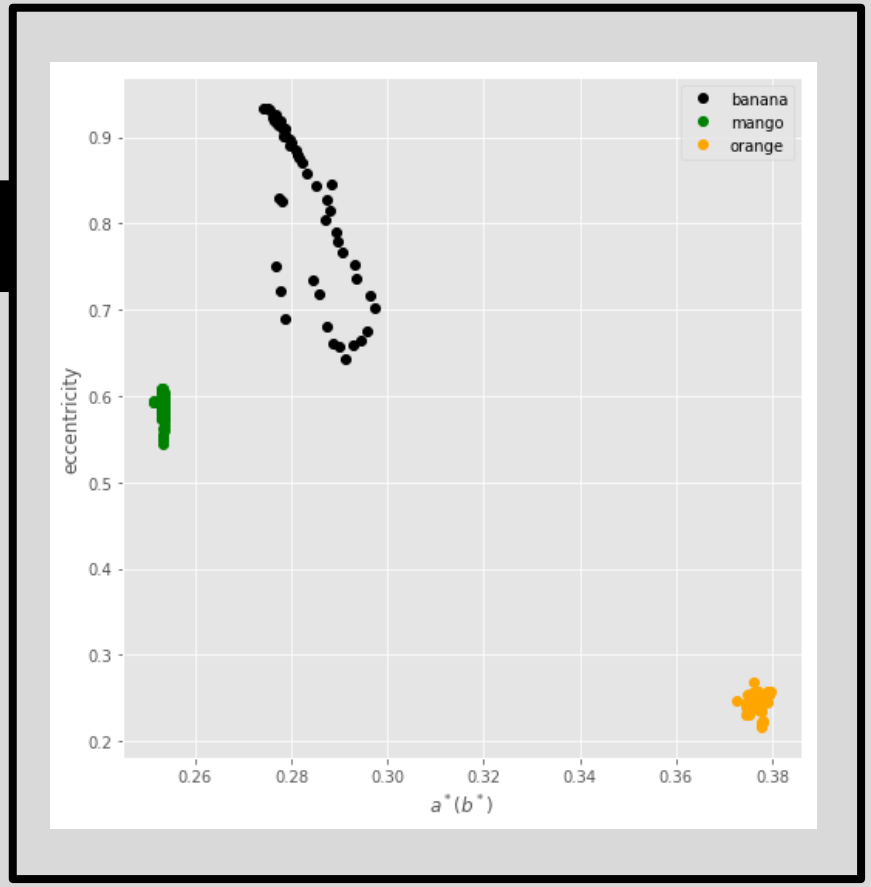
$$z=\frac{|g(x)|}{\|w\|}$$

While the margin is given by this formula (which is maximized:

$$\frac{1}{\|w\|}+\frac{1}{\|w\|}=\frac{2}{\|w\|}$$

Mary Chris Go / 2014 11122

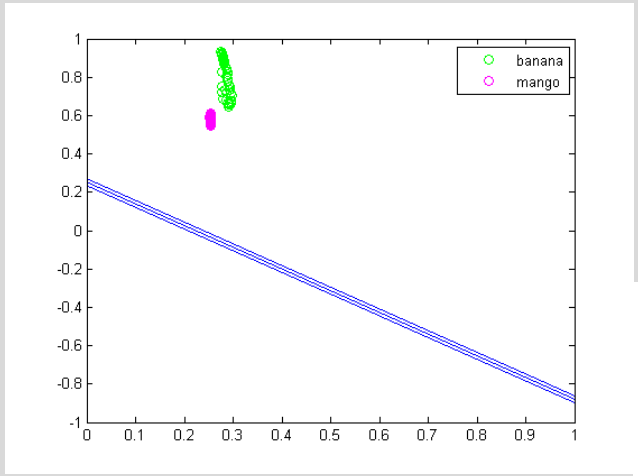
Feature Space

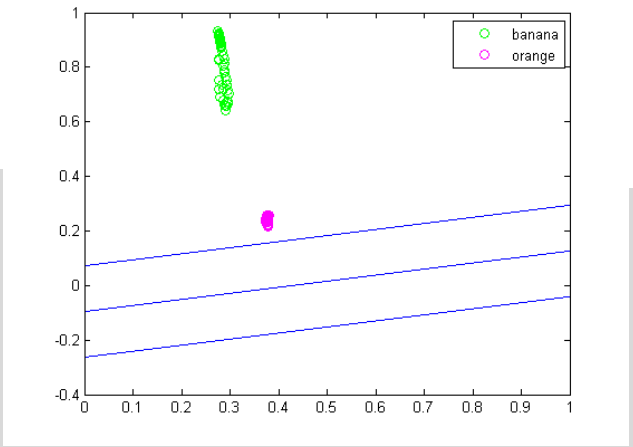


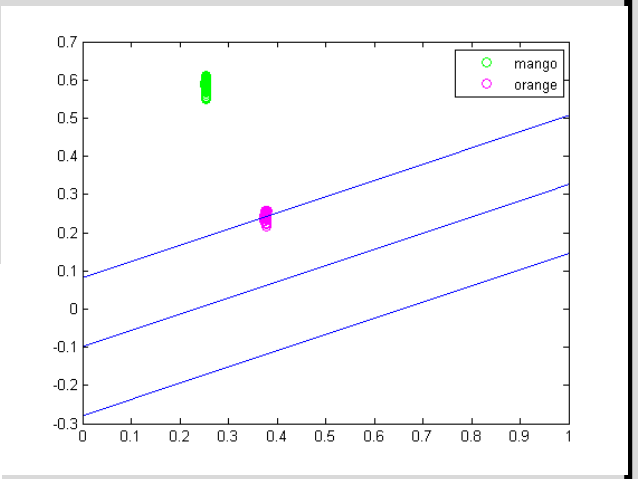

```
for j in range(3):
    filenames = os.listdir(dirs[j])
    for i,f in enumerate(filenames):
        if i == 50:
            break
        #eccentricity
        img = cv.imread(dirs[j] + f)
        img_gray = cv.cvtColor(img, cv.COLOR_BGR2GRAY)
        thres, out = cv.threshold(img_gray, 127, 255, cv.THRESH_OTSU)
        out = (img_gray < thres).astype(float)
        img_label = meas.label(out)
        props = meas.regionprops(img_label)
        ecc = props[0]['eccentricity']

        # a* b*
        img_Lab = cv.cvtColor(img, cv.COLOR_BGR2Lab).astype(float)
        img_Lab /= img_Lab[:,:,0].max()
        img_L, img_a, img_b = cv.split(img_Lab)

        ass[j].append(img_a.mean())
        bss[j].append(img_b.mean())
        ecs[j].append(ecc)
```

This snippet of code was already used several times in clustering three kinds of fruit: orange, banana, and mango. These values were saved as csv files.

Results







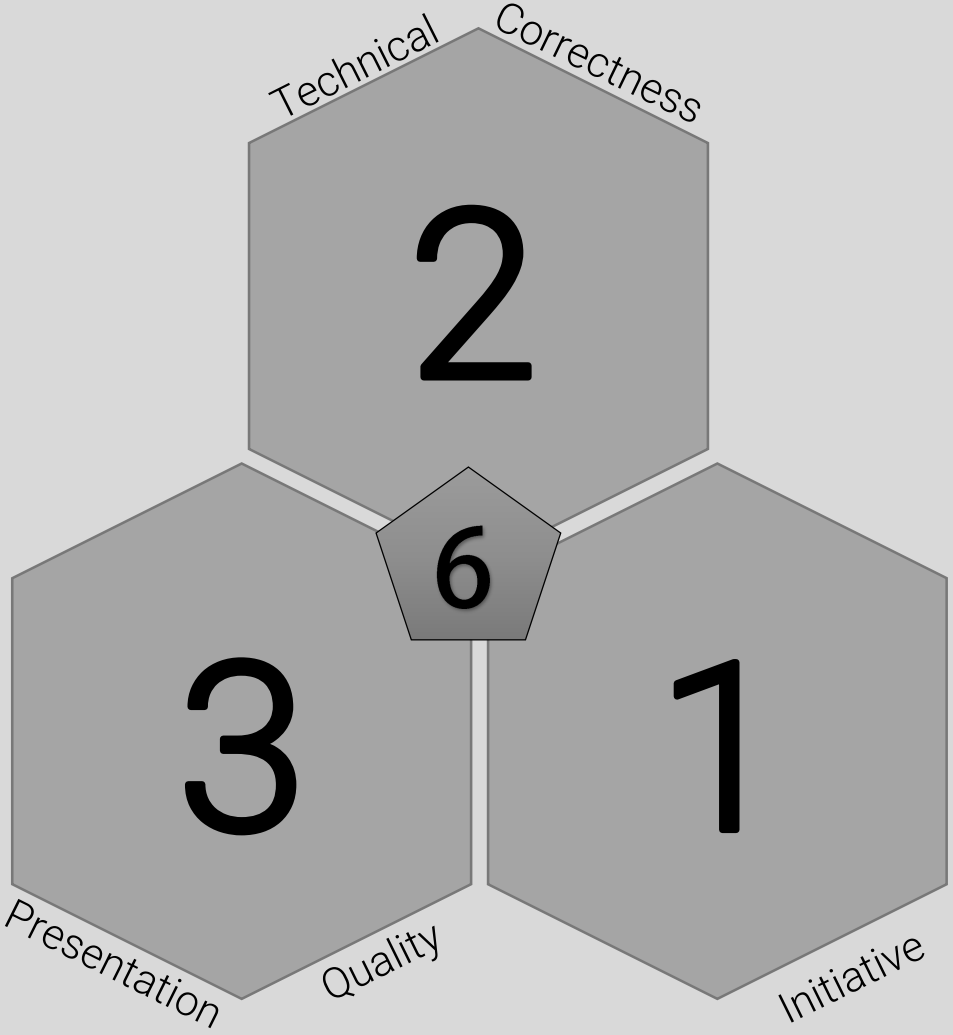
Following the instructions from the powerpoint, svm can be implemented through different programming languages. For this activity, I used MATLAB and the csv files were used as the values were needed to run. The lines also showed the margin and was able to appear. Although it showed the line, it wasn't able to classify different classes. I tried manipulating the code into something that would make it a good classifier but I'm already out of time. The angle of the line actually made sense but the placement is something to think about.

Summary

Support vector machine algorithm has good generalization properties. Compared to perceptron algorithm, it is slower as quadratic programming is computationally heavy.

All in all, it was interesting to check another method in differentiating various classes. This was supposed to be the most efficient and effective method of them all. I was a bit sad that I wasn't able to classify the classes effectively as I was hoping this method would be easier. But with beating the deadline and everything we're cramming (haha) now, I think I'll spend time to work this out after the semester.

Self-Evaluation



References

- Soriano, M., “Support Vector Machines”. 2019
- Theodoridis, Chapter 3, Pattern Recognition
- Veksler, O. , CS 434a/541a:Pattern Recognition Lecture 11 slides