Pattern Project Report

-Ankit Nayan
-Ankita Balotia
-Pritesh Kumar

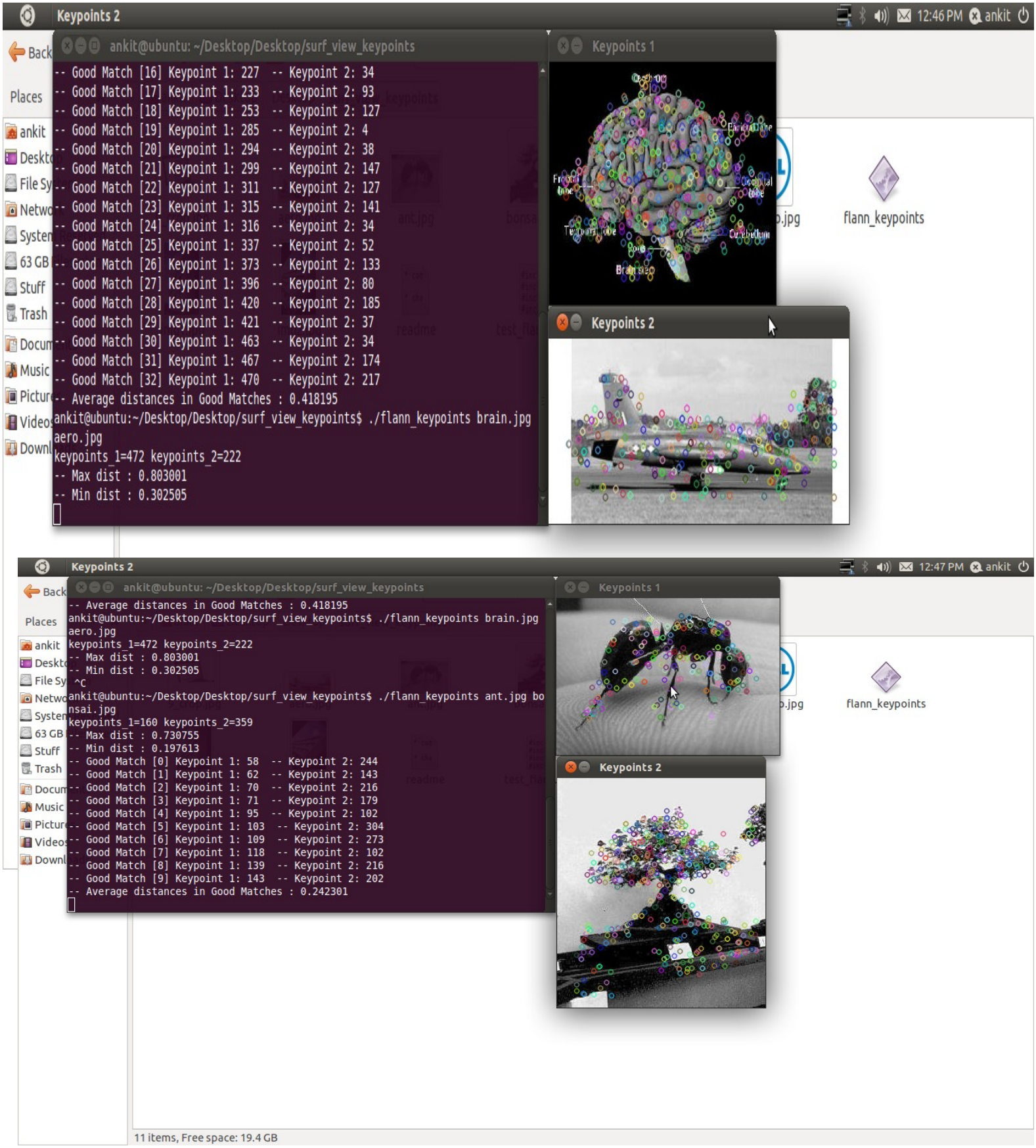
Opencv implementation of BOW model for classification of objects

Steps in Bow model :

1. Calculate SURF keypoints (a fast implementation of SIFT) for all training images
2. Cluster all the keypoints by K-means clustering
3. Make a Dictionary of those codewords (final cluster centres)
4. Match the keypoints of an image to codewords by Nearest Neighbour match
5. Calculate descriptors of the images by making histogram of the keypoints ,ie, frequency of keypoints of an image in different codewords.
6. Normalized histogram is written in csv file
7. Classify the objects by classifiers such as Naive Bayes, SMO, Logistic using **Weka**

- Keypoints are basically the points of extrema
- Keypoints are extracted with SURF algorithm with minHessian = 400
- SURF finds keypoints by finding laplacian, approximated using difference of gaussians
- SURF has feature descriptor (of each keypoint) of size 64 that contains gradients in dx,| dx|, dy, |dy| for 4x4 subregions each of size 5x5 samples
- SURF descriptors are scale and rotation invariant

Some surf keypoints

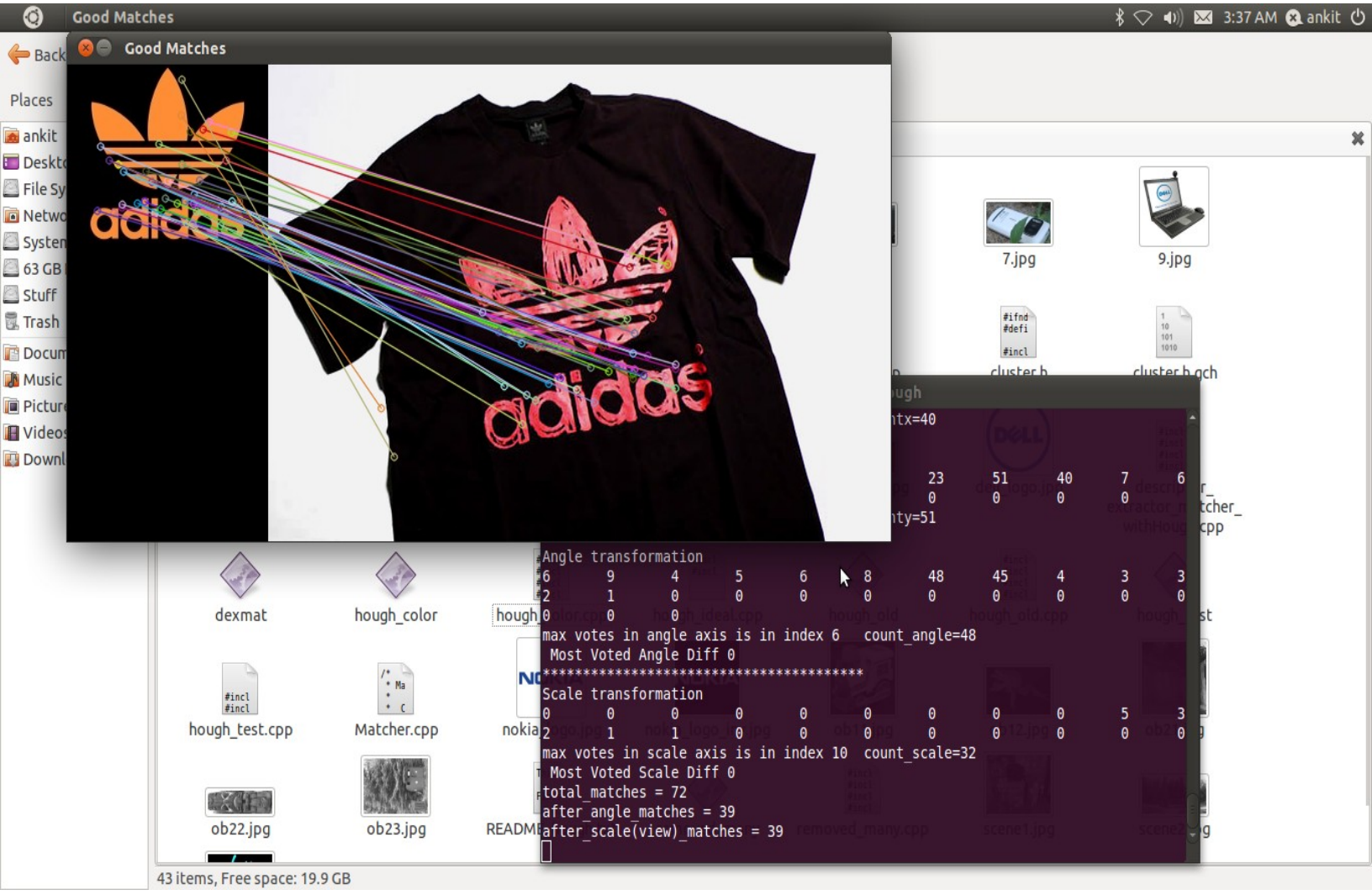


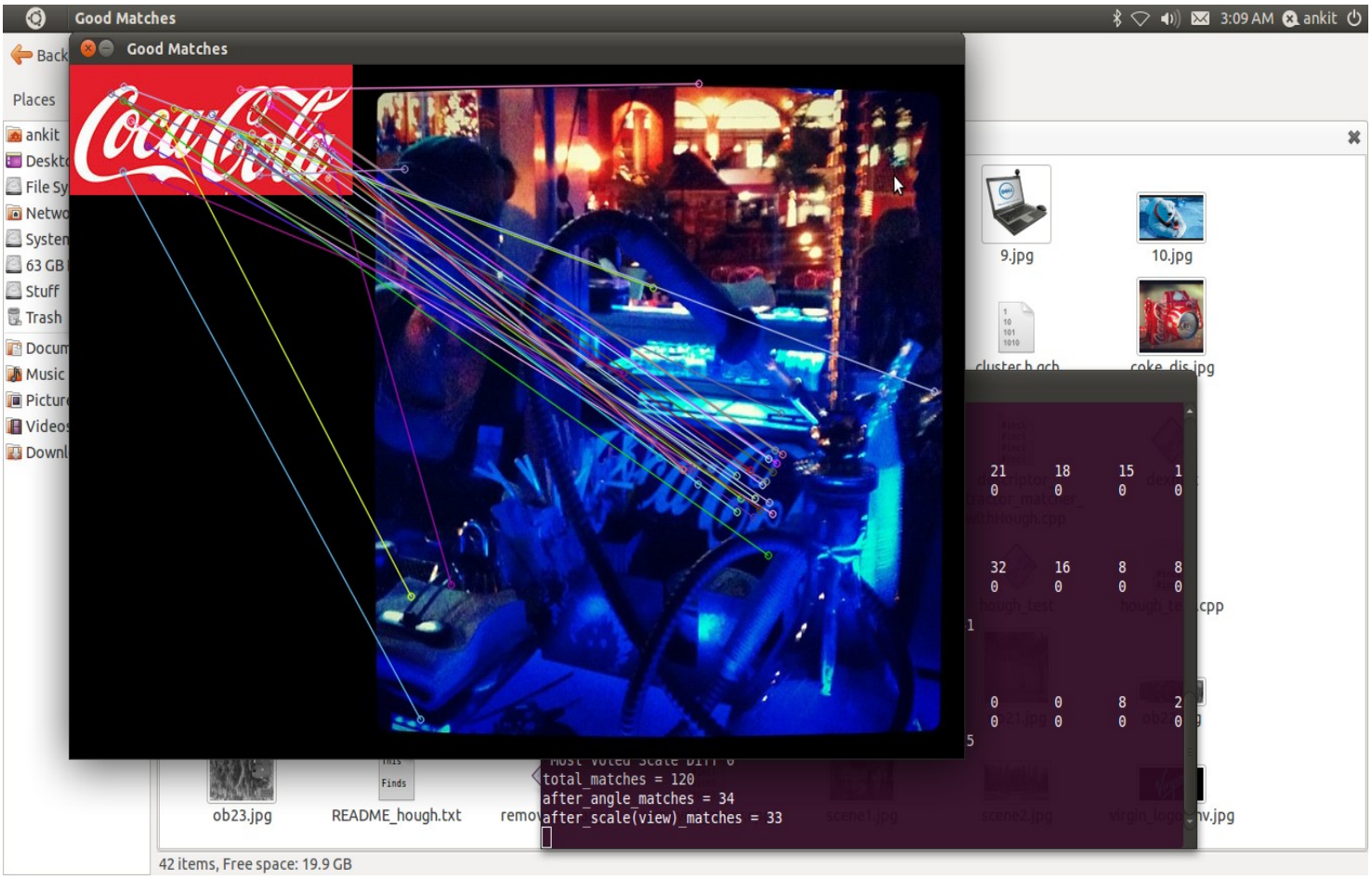
Simple image matching is done by calculating the distance between each pair of keypoints descriptors and then applying the formula :

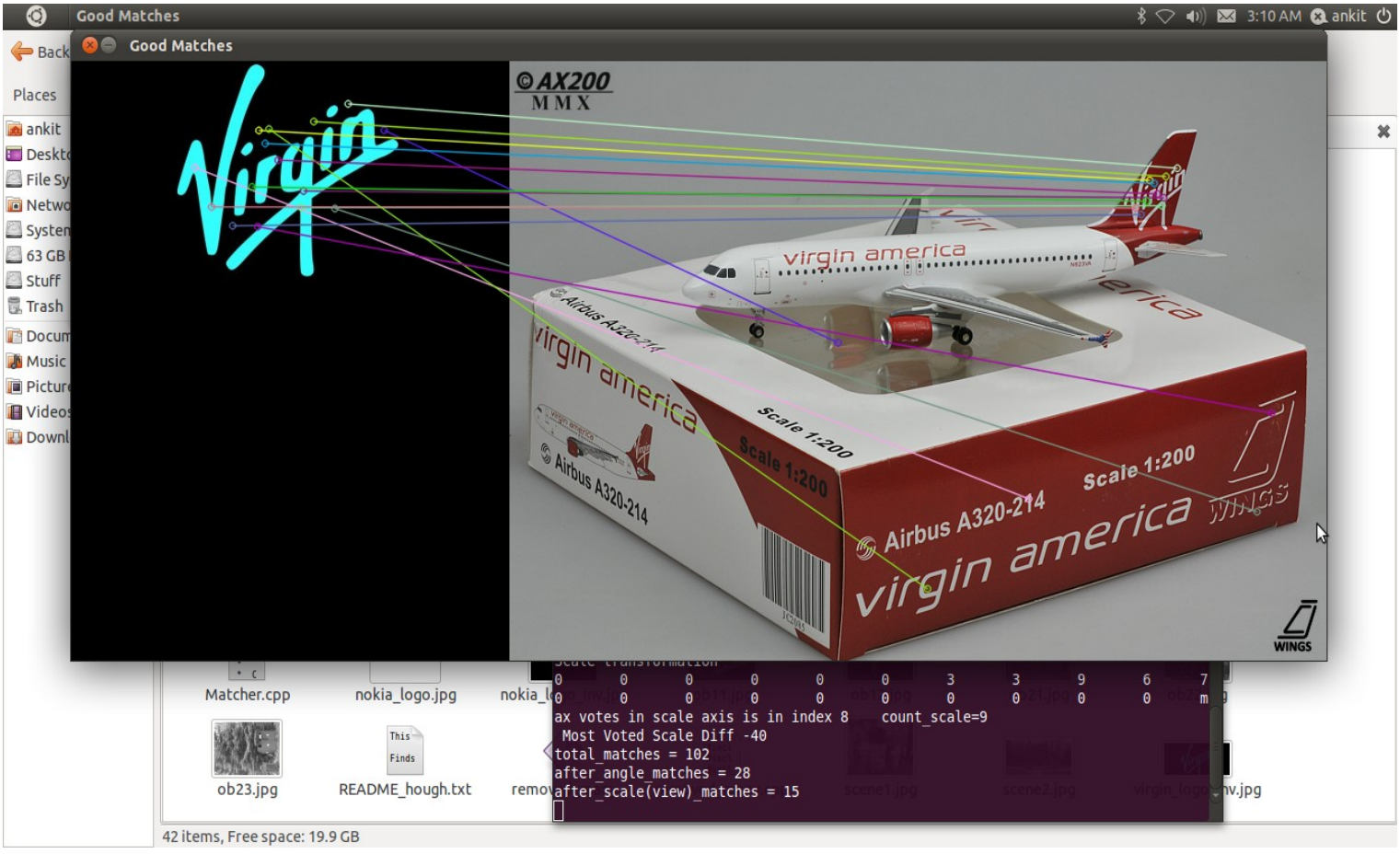
d1/d2 < 0.7

that is, ratio of nearest to 2nd nearest neighbour should be less than 0.7

- if match is not confident then d1/d2 ~ 1




Good Matches
Good Matches
Coca-Cola
total_matches = 120
after_angle_matches = 34
after_scale(view)_matches = 33
9.jpg
10.jpg
ob23.jpg
README_hough.txt
42 items, Free space: 19.9 GB


Good Matches
Good Matches
Virgin
©AX200
MMX
virgin america
Airbus A320-214
Scale 1:200
WINGS
Matcher.cpp
nokia_logo.jpg
ob23.jpg
README_hough.txt
ax votes in scale axis is in index 8 count_scale=9
Most Voted Scale Diff -40
total_matches = 102
after_angle_matches = 28
after_scale(view)_matches = 15
42 items, Free space: 19.9 GB

Results

Tested on 4 classes : nautilus, dolphin, crab & garfield each with 50 images
Results depend on vocabulary size here 100 and 200 are chosen.

K in K-means clustering equals vocabulary size.

• Vocab size = 100	Vocab size = 200
• SVM : 76%	79.9%
• Naive Bayes : 76.6%	75.5%
• Logistic : 69%	70.1%

vocabulary size = 200

svm : 79.9%

=== Confusion Matrix ===

```
 a  b  c  d   <-- classified as
42  1  6  1 |  a = nautilus
 5 35 10  0 |  b = dolphin
 4  5 41  0 |  c = crab
 1  1  3 29 |  d = garfield
```

naïve bayes : 75.5%

=== Confusion Matrix ====

```
 a  b  c  d   <-- classified as
38  5  4  3 |  a = nautilus
 8 34  8  0 |  b = dolphin
 3  5 40  2 |  c = crab
 1  1  5 27 |  d = garfield
```

logistic : 70.1%

=== Confusion Matrix ===

```
 a  b  c  d   <-- classified as
36  7   7   0 |  a = nautilus
12 24 14  0 |  b = dolphin
 4   4  41  1 |  c = crab
1  2  3 28 |  d = garfield
```

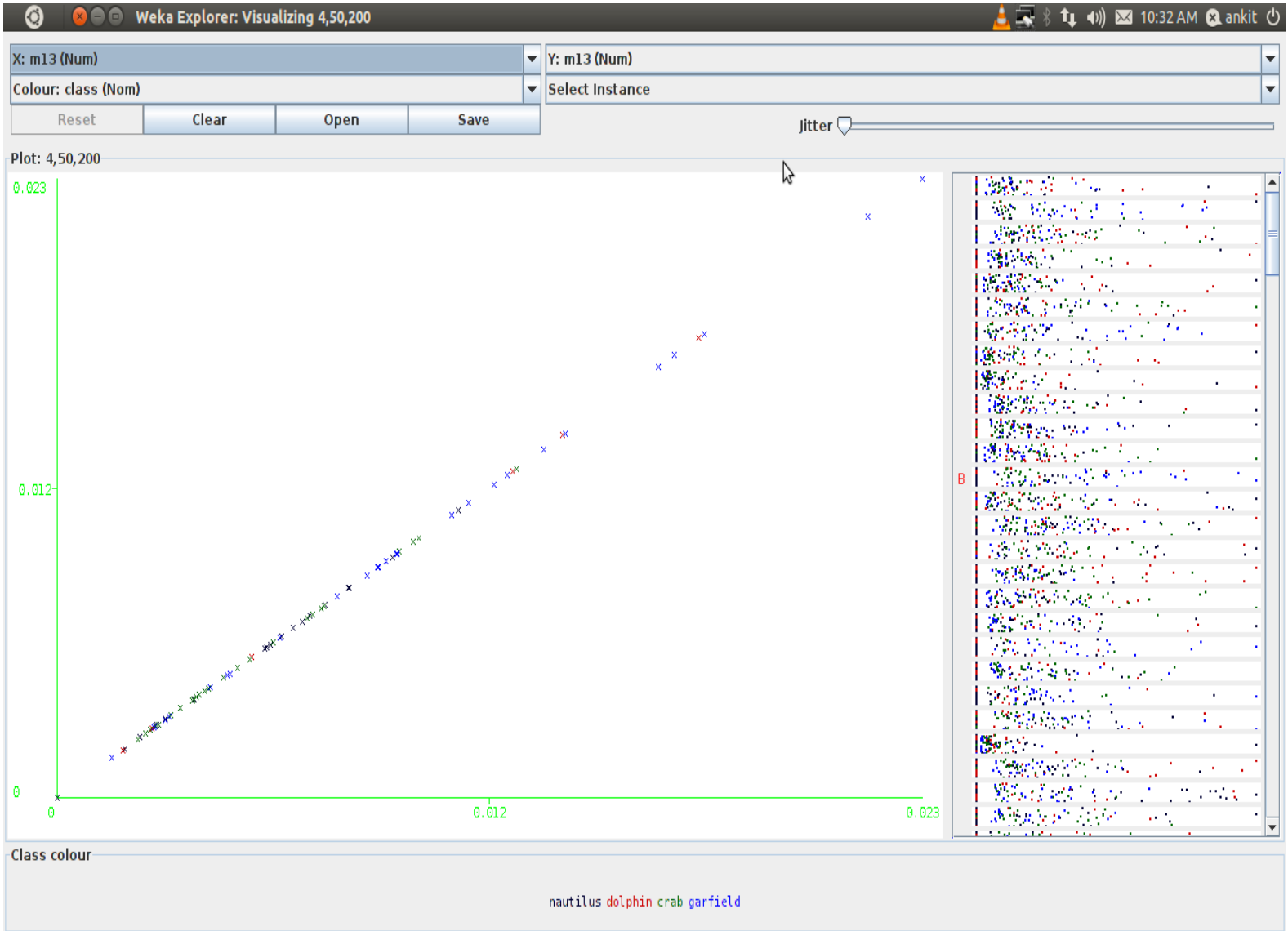
For 4 classes 25% accuracy is equivalent to no result since even if the objects were classified randomly there is ¼ probability of it being correct.

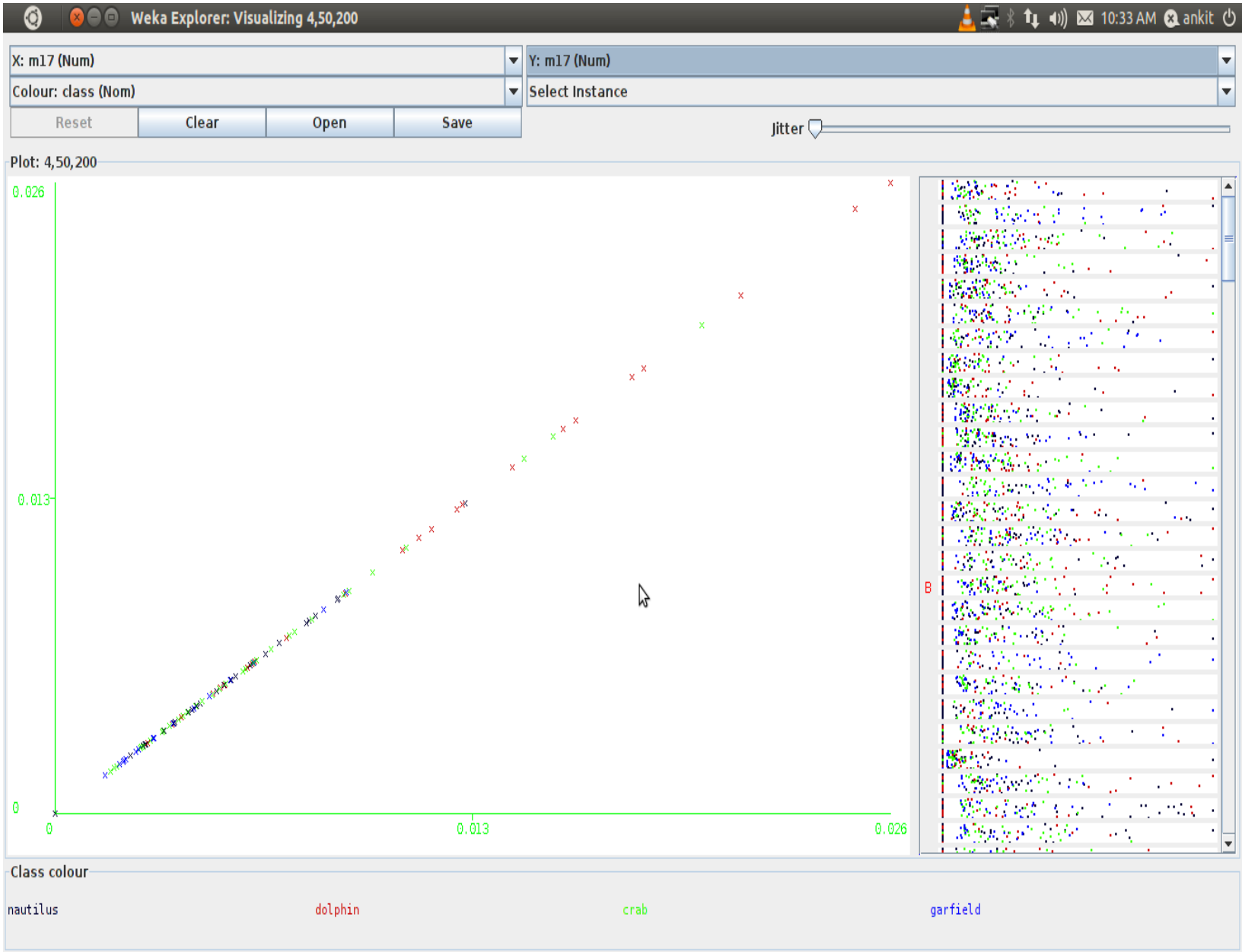
For the case of 4 classes the best two attributes that were found are m13 and m17

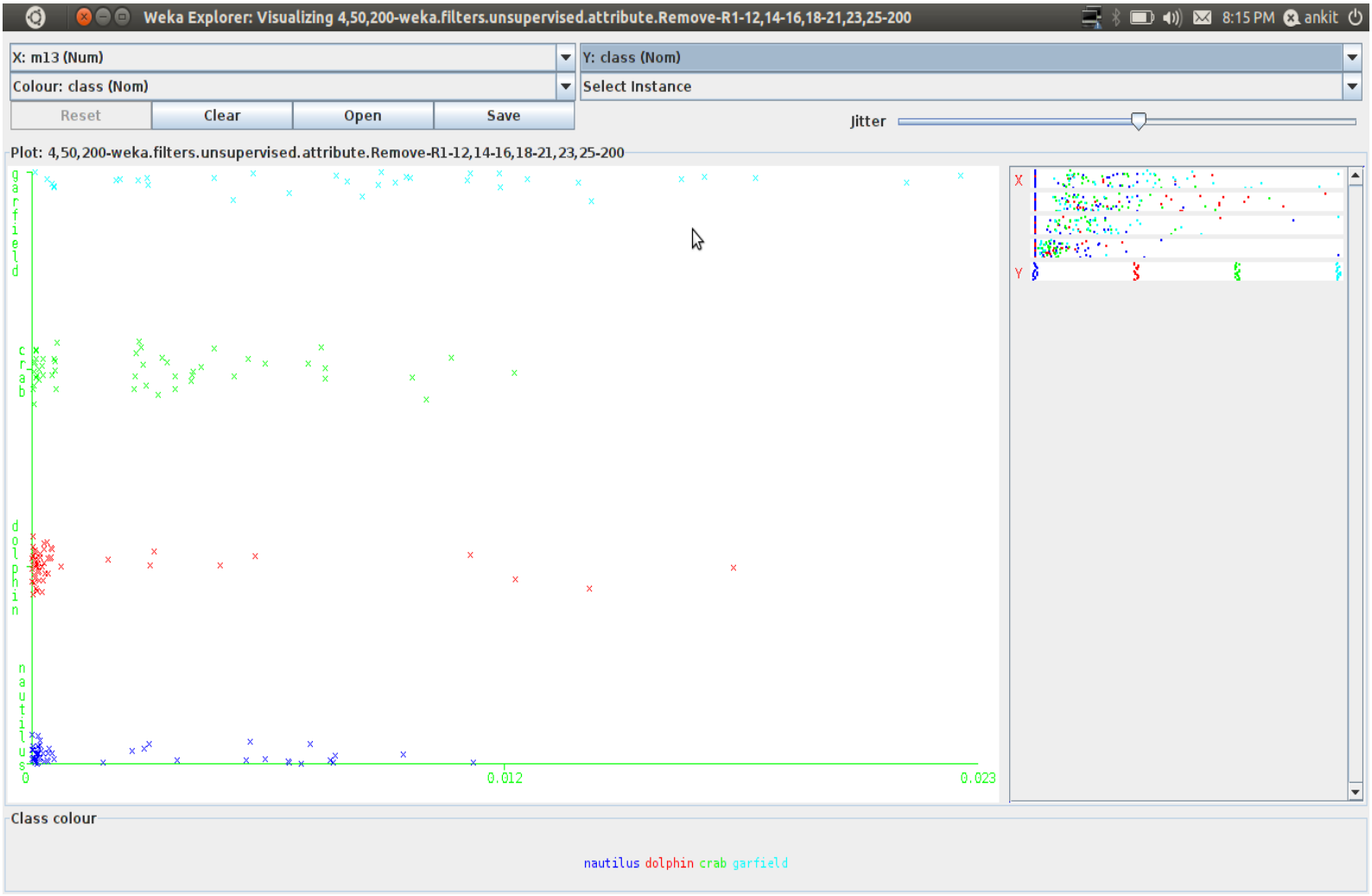
- m13 – 13th codeword in the dictionary of size 200
- m17 – 17th codeword in the dictionary of size 200

The figures below helps us visualize the distribution of 4 classes over the values of m13 and m17.

Not visible to human eyes but classifier models this distinctness of values for different classes to classify.




Weka Explorer: Visualizing 4,50,200
10:33 AM
ankit
X: m17 (Num)
Y: m17 (Num)
Colour: class (Nom)
Select Instance
Reset
Clear
Open
Save
Jitter
Plot: 4,50,200
0.026
0.013
0
0
0.013
0.026
B
Class colour
nautilus
dolphin
crab
garfield



When tested on 5 classes (aeroplane, ant, brain, nautilus, garfield) each with 50 images and vocab size = 200, SVM gave accuracy of 70.9%

for 10 classes ~60% accuracy

BOW limitations :

- It models only the basic appearance of the object in an image. Needs to be implemented with pyramid structure.
- It does not take into account shape of the object which alone is not efficient but when combined with appearance proves useful

THANK YOU