

vclsi-01-Duitama-Lacerda

April 15, 2018

Visual Computing
Prof. Dr.-Ing. Thomas Schultz

Assignment 1
Students
Mauricio Lacerda
Camila Duitama

Exercise 1) a)

```
In [152]: import pandas as pd
          data = pd.read_excel('breast-cancer-wisconsin.xlsx', 0, \
                               index_row=0, na_values=[''])
          print("Number of columns:", len(data.columns))
          print("Number of instances:", len(data.index))
          print("Column names:", data.columns.get_values())
```

Number of columns: 11

Number of instances: 699

Column names: ['code' 'thickness' 'uniCelS' 'uniCelShape' 'marAdh' 'epiCelSize' 'bareNuc'
'blaChroma' 'normNuc' 'mitoses' 'class']

Exercise 1) b)

```
In [153]: data = data.interpolate(method='linear')
```

We used the linear method for the interpolation because, since the data seems to be linearly distributed, it is less likely that the results could be affected by this interpolation.

Exercise 1) c)

```
In [154]: benign = data[data['class']==2]
          malign = data[data['class']==4]
          print("Number of benign records:", len(benign.index))
          print("Number of malignant records:", len(malign.index))
```

Number of benign records: 458

Number of malignant records: 241

Exercise 1) d)

```
In [155]: import math
def f_score(x1, x2, x):
    mu_x1 = x1.mean()
    mu_x2 = x2.mean()
    mu_x = x.mean()
    # numerator
    num = ((mu_x1-mu_x)**2) + ((mu_x2-mu_x)**2)
    # denominator
    den=0
    acc1=0
    for i in x1:
        acc1+=(i-mu_x1)**2
    den+=acc1/(len(x1.index)-1)
    acc2=0
    for i in x2:
        acc2+=(i-mu_x2)**2
    den+=acc2/(len(x2.index)-1)
    if den==0:
        return math.inf
    return num/den
for i in data.columns:
    print(i, f_score(benig[i], malig[i], data[i]))
```

```
code 0.00946920773691
thickness 1.13172472256
uniCelS 1.83630641175
uniCelShape 1.89863511519
marAdh 0.848820438714
epiCelSize 0.808413912364
bareNuc 1.80596723137
blaChroma 1.30138955286
normNuc 0.92825503621
mitoses 0.187838518258
class inf
```

The best classes to describe the data are: Clump Thickness, Uniformity of Cell Size, Uniformity of Cell Shape, Bare Nuclei and Bland Chromatin (larger values). In the case of the class, since it describes each of the subgroups, the f-score tends to infinity.

Exercise 1) e) Since it wasn't mentioned to add the code, we left this column out.

```
In [156]: try:
    data.loc[:,['thickness','uniCelS','uniCelShape', 'bareNuc', \
                'blaChroma']].to_excel('reduced_dataset.xlsx', sheet_name='Sheet1')
except:
    print("Couldn't write to file")
```

Exercise 2)

a) What is the Helmholtz-Kohlrausch effect? (2P)

Phenomenon where one of two objects with the same luminance is perceived to be brighter. This effect increases with the saturation of the colour (the more saturated it is the brighter it is perceived)

b) Why are the authors proposing to use images of faces? (2P)

Because humans have a region in our brains to recognise faces (mostly when the light comes from above), and luminance plays a major role in such process. Therefore, subjects in the study would be more likely to correctly select the “positive” (the image that stands out more) or “negative” image, when faced with the picture of two faces with a gray and a coloured region of different luminance levels.

c) To what alternative method do the authors compare their newly proposed one in the user study? (2P)

To the minimally distinct boundary method (MDB), where two adjacent fields are compared several times by varying the brightness. The procedure is done until the observer decides that the border between the fields is barely recognisable

d) Based on the result of the user study, what is the advantage of the newly proposed method? (2P)

The newly proposed method is easily reproducible (because it relies on our natural ability to recognise faces), simple, more precise and equally fast.

e) Why do the authors have to know the monitor gamma while creating a colormap based on the result of the user study? (2P)

With the monitor gamma parameter, one could convert two colours of equal luminance from RGB coordinates to intensity levels

In []: