Laboratory 2 - Binary Classification for fingerprint spoofing detection
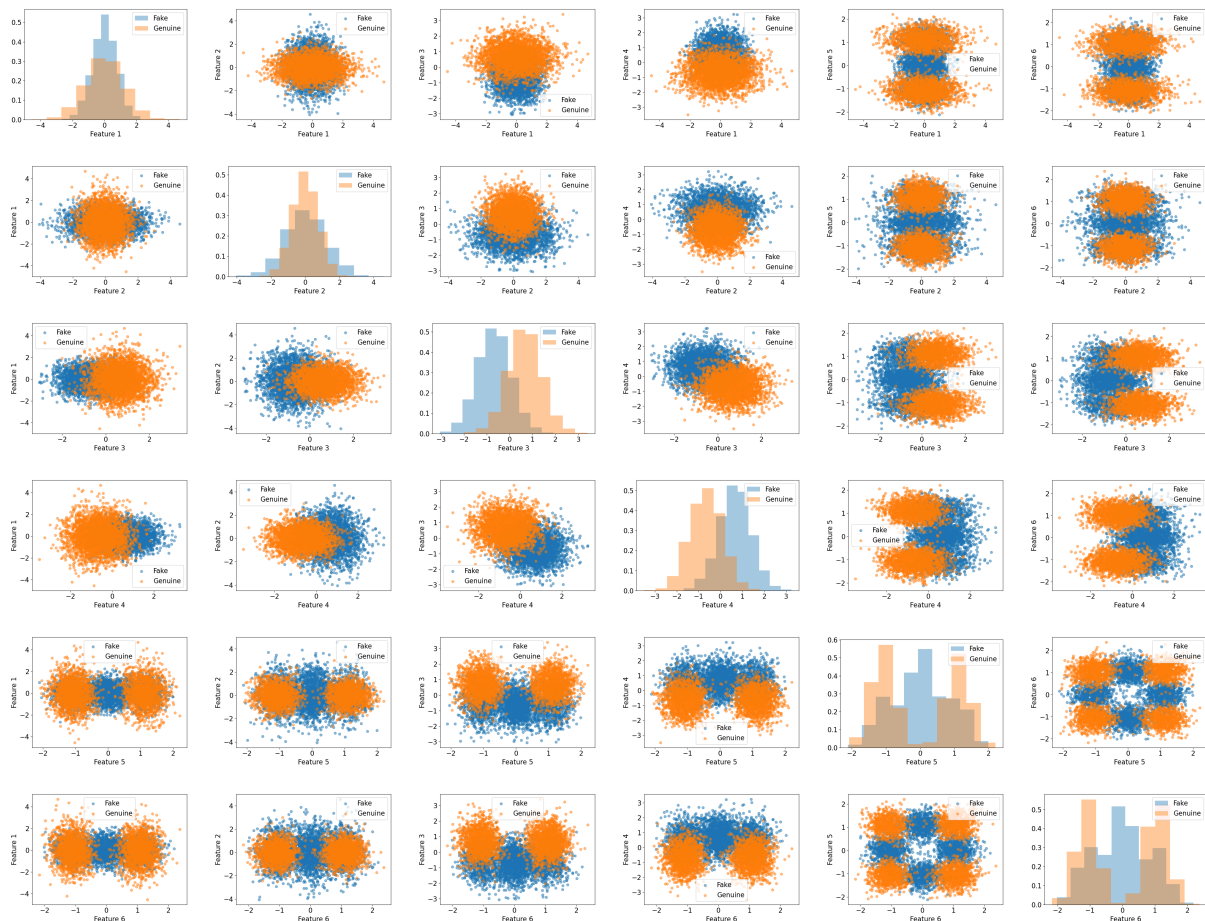
Exploit all the data

Figure 1: All features

First question

Analyze the first two features. What do you observe? Do the classes overlap? If so, where? Do the classes show similar mean for the first two features? Are the variances similar for the two classes? How many modes are evident from the histograms (i.e., how many "peaks" can be observed)?
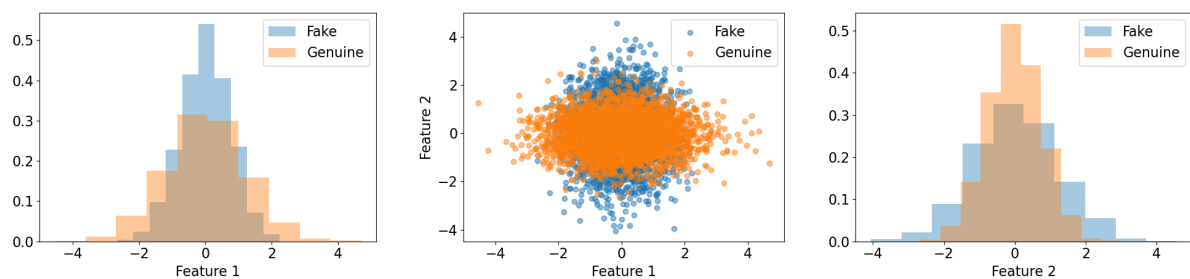

Figure 2: Analyse Feature 1 and Feature 2

Looking at the histograms, we can notice in both features classes overlap for each value assumed by the feature. For class Fake, mean of Feature 1 and Feature 2 are respectively 0.00287744 (0.003) and 0.01869316 (0.019); variance is 0.56958105 and 1.42086571 For classe Genuine, mean of Feature 1 and

Feature 2 are respectively 5.44547838e-04 (0.0005) and −8.52437392e-03 (-0.008); variance is 1.43023345 and 0.57827792

Histograms have few peaks: in both features, peak is near 0, causing a class overlap around this value in the scatterplot.

Looking at the scatterplot, we can see that classes overlap for Feature 1 and Feature 2, in particular in the interval [−1.5, 1.5] of Feature 2.

Second question

Analyze the third and fourth features. What do you observe? Do the classes overlap? If so, where? Do the classes show similar mean for these two features? Are the variances similar for the two classes? How many modes are evident from the histograms?
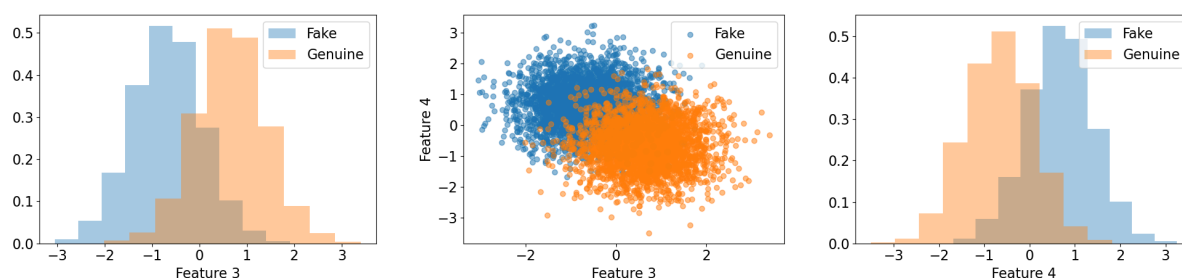

Figure 3: Analyse Feature 3 and Feature 4

Third question

Analyze the last two features. What do you observe? Do the classes overlap? If so, where? How many modes are evident from the histograms? How many clusters can you notice from the scatter plots for each class?
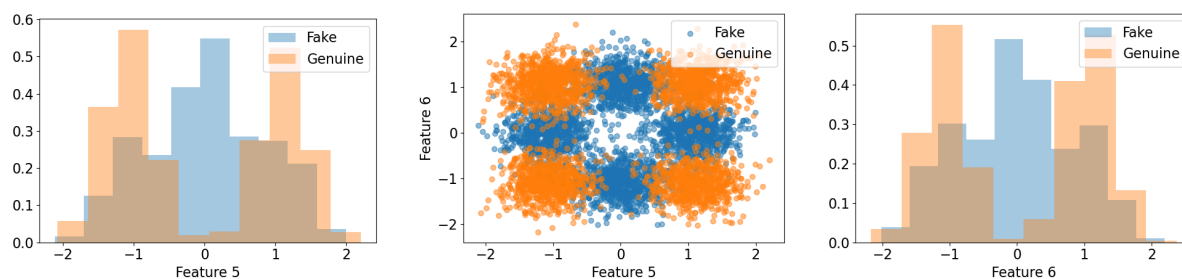

Figure 4: Analyse Feature 5 and Feature 6