CIS 663 Biometrics

Assignment #5

Due: Before Live session 9

For this assignment it is expected you will write some code. You may use **any** language or even just an excel spreadsheet. But you must turn in your code and provide instructions on how to run.

Question 1) Bayes Classification You testing an biometric algorithm that warns security at the airport if someone has been flagged for a second screening. The test has a 99% sensitivity (true positive probability), 99% specificity (true negative probability), but only 0.5% of people have been flagged.

a) Using Bayes theorem where the test is P(+|Flagged) = 0.99, P(-|Not-Flagged), P(Flagged) = 0.005, and P(+) = P(+|Flagged)P(Flagged) + P(+|Not-flagged)P(Not-flagged). What is the probability that a randomly selected person that tested positive on the biometric test has been flagged, P(Flagged|+)?

See Excel Sheet for formula/working out

$P(\text{Flagged} \mid +) = 33.2\%$

b) The spec for your biometric system requires requires a P(Flagged | +) value of greater than 45% to be acceptable by the public, what do the sensitivity and specificity values need to be to meet that spec?

See Excel as well for work

We can hold $P(+ \mid \text{Flagged})$ / Sensitivity = 99%

If we bring $P(- \mid \text{Not Flagged})$ / Specificity = 99.39%

then $P(\text{Flagged} \mid +) = 45\%$.

Question 2) You need to clean up data from a continuous measurement for your classification system.

a) Your input appears to have a Gaussian distribution. You have determined there can be spikes in your input that are two standard deviations above the mean for the data where mu = 5.1 and sigma = 1.9.
From the below set,
{2.10, 4.20, 6.00, 4.50, 4.20, 4.30, 2.50, 6.90, 8.10, 7.00, 9.10, 5.30, 6.20, 4.80, 1.80, 5.80, 3.30, 3.30, 4.00, 6.4}
what are the outliers?

See Exeel for work.

Outlier is 9.10