1.) *Here Class 1 corresponds to “5”, and Class 2 corresponds to “6”.

Table 1: Mixture of Gaussians:

Accuracy	0.8676
w0	0.1057
w	−0.0004 0.0161 0.0304 0.0129 0.0531 0.1589 0.0227 0.0362 0.0819 −0.0245 0.1165 0.0121 0.0211 0.1565 0.0109 −0.0506 0.0723 0.0640 0.0098 −0.0201 −0.0118 −0.0565 −0.0338 0.0253 0.0524 0.0118 −0.0208 0.0180 0.0883 −0.0069 −0.0668 0.0022 0.0716 −0.0100 −0.0646 −0.0402 −0.0333 −0.0820 −0.0385 −0.0446 −0.0420 −0.0038 −0.2171 −0.0821 −0.0025 0.0272 0.0249 −0.0461 0.0329 −0.0994 −0.0726 −0.0532 0.0620 −0.0317 −0.0723 0.0098 −0.0259 −0.0731 0.0906 0.0586 −0.0185 −0.0489 0.0046 0.0158
Pi1	0.4964
Pi2	0.5036
u1	4.9201 5.0127 8.3103 10.0200 10.7532 9.5481 6.2976 4.6860 5.1180 6.0490 10.6661 9.8621 8.5844 7.7731 5.4211 4.5699 4.5753 6.5681 10.1506 7.0853 5.7260 4.8566 4.6515 4.5590 4.8657 6.4283 10.3031 9.7205 8.0617 6.4229 4.7114 4.7314 4.9546 5.4537 7.7187 7.9510 8.7060 7.9093 5.7768 4.3684 4.7387 4.8784 5.1016 5.7931 7.5045 8.6951 5.9347 4.5662 4.3884 5.0690 7.2123 7.8566 8.9365 8.0871 5.7768 4.7096 4.6479 4.9038 8.7514 10.4392 8.9147 6.5481 5.0998 5.0672
u2	4.9767 4.8605 5.5760 9.5188 8.5027 5.3900 5.3345 4.8050 4.6816 4.6333 7.9535 10.3148 6.8784 5.1145 4.6780 4.6011 4.7835 4.9821 9.8140 8.4293 4.7746 4.8426 4.6708 4.7692 4.5975 5.5653 10.5617 7.8837 6.1503 5.4061 5.0841 4.6172 4.7835 6.1449 10.6047 9.8318 9.5850 8.7048 5.6225 4.5206 4.7800 5.6261 10.4526 9.0519 7.6333 9.2826 8.4419 4.6261 4.5921 5.2057 9.2379 9.9714 7.2182 9.5528 9.0304 4.9106 4.8640 4.8533 5.7335 9.1377 10.9696 9.7531 6.4669 5.0268
diag(Sigma)	39.8587 39.5712 43.2563 42.4111 40.3930 41.9653 43.6573 35.7806 37.6420 39.6338 41.6589 41.4307 43.8446 39.8570 39.2544 37.2208 35.9671 40.9541 42.5452 46.7621 39.7362 38.2503 36.0834 36.7397 36.3302 40.3912 40.0378 46.0040 45.5309 43.4102 37.6236 37.1228 37.8276 42.6755 41.9397 44.2269 45.0235 45.2556 41.4459 33.0250 37.5696 39.7835 39.8710 44.4305 46.4573 43.3695 43.5041 35.6092 35.2377 40.2025 43.9491 44.2588 45.8758 44.7406 43.9388 38.8965 38.0409 37.4593 44.8849 41.9983 42.3220 44.0130 42.7988 40.4677

Table 2: Logistic Regression

Accuracy	0.8631
w	0.1347 −0.0295 0.0142 0.0349 0.0152 0.0442 0.1827 0.0378 0.0630 0.0940 −0.0047 0.1045 0.0168 0.0260 0.1929 0.0035 −0.0514 0.0529 0.0753 −0.0074 −0.0199 0.0069 −0.0969 −0.0220 0.0521 0.0608 0.0192 −0.0130 0.0317 0.0895 −0.0038 −0.1014 0.0271 0.1348 −0.0219 −0.0684 −0.0324 −0.0168 −0.0702 −0.0406 −0.0580 −0.0224 −0.0140 −0.2309 −0.0926 0.0062 −0.0001 0.0140 −0.0933 0.0120 −0.1506 −0.0892 −0.0742 0.0632 −0.0347 −0.0848 −0.0003 −0.0337 −0.0703 0.1127 0.1009 −0.0432 −0.0486 −0.0058 0.0350

The parameterizations/objectives for both Mixture of Gaussians and Logistic Regression successfully generate a linear separator of the data leading to an 86% accuracy rate (when tested with cross-validation). Thus we can conclude that a mostly accurate linear separator can be found for this data set. Also, due to the closeness of the accuracies of Mixture of Gaussians and Logistic Regression, it might be hypothesized that the objectives of these two methods represent very similar goals.

The K Nearest Neighbours algorithm in Assignment 1 only accomplished an 81% accuracy. Thus, we can conclude that a linear separator was more suited for this data because both methods that found a linear separator achieved a higher accuracy. We could say that the data was more linearly separable than “K Nearest Neighbour separable”.

2. (a) AND Possible - Take $w_0 = -1.5$ $w_1 = 1$ $w_2 = 1$

OR Possible - Take $w_0 = -0.5$ $w_1 = 1$ $w_2 = 1$

XOR Not Possible: Take $\phi(x) = \phi(x_1, x_2)$

$= ((x_1 - x_2)^2, (x_1 - x_2)^2)$

and $w_0 = -1$ $w_1 = 1$ $w_2 = 1$

EFF Not Possible: Take $\phi(x) = \phi(x_1, x_2)$

$= (-(x_1 - x_2)^2, -(x_1 - x_2)^2)$

and $w_0 = 1$ $w_1 = 1$ $w_2 = 1$

(b) The dataset is not separable. Our experiment (see linearSeparability.m) involved training and testing a logistic regression classifier on the entire data set. If a linear separator did exist, the tests would be a 100% accurate, because the linear separator found would divide the data into two classes perfectly.

However, this experiment only achieved an accuracy of 88.74% (not 100%). Thus, the data cannot be linearly separable.