P1. 1) P(Y) = [P(spam), P(work), P(private)] = [0.4, 0.3, 0.3]
P(money|Y) = [P(money | spam), P(money | work), P(money | private)]
= [(4+alpha)/(8+3*alpha), (2+alpha)/(6+3*alpha), (1+alpha)/(7+3*alpha)]
= [5/11, 1/3, 1/5]

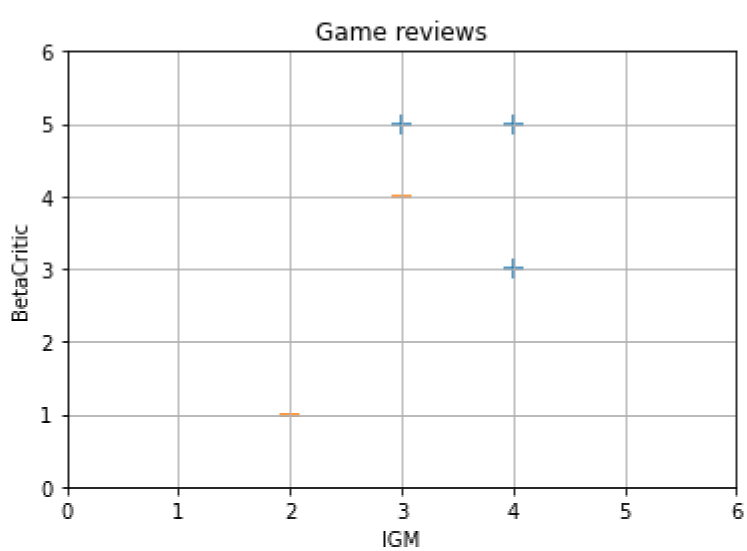
P(bank|Y) = [P(bank | spam), P(bank | work), P(bank | private)]
= [(1+alpha)/(8+3*alpha), (4+alpha)/(6+3*alpha), (2+alpha)/(7+3*alpha)]
= [2/11, 5/9, 3/10]

P(love|Y) = [P(love | spam), P(love | work), P(love | private)]
= [(3+alpha)/(8+3*alpha), (0+alpha)/(6+3*alpha), (4+alpha)/(7+3*alpha)]
= [4/11, 1/9, 1/2]

2) P(y, x1…xn) is proportional to P(y)*P(y|x1)... so y is argmax(P(y1, x1…), … P(yn, x1…))
P(spam) = 0.4 P(work) = 0.3 P(private) = 0.3
y1 = argmax([0.4*2/11*4/11*5/11, 0.3*5/9*1/9*1/3, 0.3*3/10*1/2*1/5]) = spam
y2 = argmax([0.4*4/11*5/11*2/11, 0.3*1/9*1/3*5/9, 0.3*1/2*1/5*3/10]) = spam
y3 = argmax([0.4*4/11*5/11*5/11*2/11, 0.3*1/9*1/3*1/3*5/9, 0.3*1/2*1/5*1/5*3/10]) = spam
y4 = argmax([0.4*4/11*5/11*2/11*2/11, 0.3*1/9*1/3*5/9*5/9, 0.3*1/2*1/5*3/10*3/10]) = work
y5 = argmax([0.4*4/11*4/11*4/11*5/11*2/11, 0.3*1/9*1/9*1/9*1/3*5/9, 0.3*1/2*1/2*1/2*1/5*3/10]) = private
y6 = argmax([0.4*2/11*2/11*2/11, 0.3*5/9*5/9*5/9, 0.3*3/10*3/10*3/10]) = work
y7 = argmax([0.4*4/11, 0.3*1/9, 0.3*1/2]) = private

P2.
denote the numbers of true positive, false positive, true negative, and false negative as TP, FP, TN, and FN; and non-spam as positive, spam as negative results.
Then TP = 4, FP = 0, TN = 1, FN = 2;
Precision = TP/(TP+FP) = 1
Recall = TP/(TP+FN) = 2/3
Accuracy = (TP+TN)/Total = 5/7
F1 = 2*(Pre*Rec)/(Pre+Rec) = 4/5

P3.
a)



It can be segregated linearly.

b)

step	w	activation	correct?
1	[-1, 0, 0]	-1	yes
2	[-1, 0, 0]	-1	no
3	[0, 4, 3]	27	yes
4	[0, 4, 3]	31	yes
5	[0, 4, 3]	24	no

After the 5th update, w is [-1, 1, -1].

c) $y1 = sign(-1*1 + 1*2 + (-1)*1) = 1$
$y2 = sign(-1*1 + 1*4 + (-1)*3) = 1$
$y3 = sign(-1*1 + 1*3 + (-1)*5) = -1$
$y4 = sign(-1*1 + 1*4 + (-1)*5) = -1$
$y5 = sign(-1*1 + 1*3 + (-1)*4) = -1$
Three of the results are wrong, so we cannot make the right predictions.

d)
Case 1: Let f1 = 5, f2 = 5, f(x) = [1, 5, 5], according to the problem, f1+f2>9, it will succeed. However, w^T*f(x) = −1 < 0 gives us an incorrect prediction. Therefore, we cannot classify games correctly in this case.

Case 2: Let f1 or f2 be 5, so suppose f1 = 3, f2 = 5, f(x) = [1, 3, 5]. As f2 = 5, the game will succeed. However, w^T*f(x) = −3 < 0 gives us an incorrect prediction. Therefore, we cannot classify games correctly in this case.

Case 3: If the scores sum to an even number, f1 + f2 = 2n for some positive integer n ≥ 1. Let f(x) = [1, f1, 2n−f1], and w^T*f(x) = −1 + f1 - (2n − f1) = 2*f1 - 2n − 1. Since n ≥ 1, 2n + 1 ≥ 3. If f1 = f2 =1, f1+f2 = 2, w^T*f(x) = -1, gives us a wrong prediction. Therefore, we cannot classify games correctly in this case.