Problem Set 6

Problem 1

1

$P(Y) =< P(spam), P(work), P(private) >=< 0.4, 0.3, 0.3 >$

$P(money|Y) =< \frac{4+1}{8+3}, \frac{2+1}{6+3}, \frac{1+1}{7+3} >=< \frac{5}{11}, \frac{3}{9}, \frac{2}{10} >$

$P(bank|Y) =< \frac{1+1}{8+3}, \frac{4+1}{6+3}, \frac{2+1}{7+3} >=< \frac{2}{11}, \frac{5}{9}, \frac{3}{10} >$

$P(love|Y) =< \frac{3+1}{8+3}, \frac{1}{6+3}, \frac{4+1}{7+3} >=< \frac{4}{11}, \frac{1}{9}, \frac{5}{10} >$

2

1. It Spam. $P(Y|S) = \frac{P(S|Y)P(Y)}{P(S)}$.
   $P(S|Y)P(Y) =< 0.4, 0.3, 0.3 > * < \frac{2}{11}, \frac{5}{9}, \frac{3}{10} > * < \frac{4}{11}, \frac{1}{9}, \frac{5}{10} > * < \frac{5}{11}, \frac{3}{9}, \frac{2}{10} >=< \frac{40}{1331}, \frac{5}{243}, \frac{3}{100} > * < 0.4, 0.3, 0.3 >= \langle \frac{16}{1331}, \frac{1}{16}$
2. Its Spam.
   $P(S|Y)P(Y) =< 0.4, 0.3, 0.3 > * < \frac{2}{11}, \frac{5}{9}, \frac{3}{10} > * < \frac{4}{11}, \frac{1}{9}, \frac{5}{10} > * < \frac{5}{11}, \frac{3}{9}, \frac{2}{10} >=< \frac{40}{1331}, \frac{5}{243}, \frac{3}{100} > * < 0.4, 0.3, 0.3 >= \langle \frac{16}{1331}, \frac{1}{16}$
3. Its Spam.
   $P(S|Y)P(Y) =< \frac{5}{11}, \frac{3}{9}, \frac{2}{10} > * < 0.4, 0.3, 0.3 > * < \frac{2}{11}, \frac{5}{9}, \frac{3}{10} > * < \frac{4}{11}, \frac{1}{9}, \frac{5}{10} > * < \frac{5}{11}, \frac{3}{9}, \frac{2}{10} >= \langle \frac{80}{14641}, \frac{1}{486}, \frac{9}{5000} \rangle$
4. Its Work.
   $P(S|Y)P(Y) =< 0.4, 0.3, 0.3 > * < \frac{2}{11}, \frac{5}{9}, \frac{3}{10} > * < \frac{4}{11}, \frac{1}{9}, \frac{5}{10} > * < \frac{5}{11}, \frac{3}{9}, \frac{2}{10} > * < \frac{2}{11}, \frac{5}{9}, \frac{3}{10} >= \langle \frac{32}{14641}, \frac{5}{1458}, \frac{27}{10000} \rangle$
5. It's Private.
   $P(S|Y)P(Y) =< 0.4, 0.3, 0.3 > * < \frac{2}{11}, \frac{5}{9}, \frac{3}{10} > * < \frac{4}{11}, \frac{1}{9}, \frac{5}{10} > * < \frac{5}{11}, \frac{3}{9}, \frac{2}{10} > * < \frac{4}{11}, \frac{1}{9}, \frac{5}{10} > * < \frac{4}{11}, \frac{1}{9}, \frac{5}{10} >= \langle \frac{256}{161051}, \frac{1}{13122},$
6. It's Work. $P(S|Y)P(Y) =< 0.4, 0.3, 0.3 > * < \frac{2}{11}, \frac{5}{9}, \frac{3}{10} > * < \frac{2}{11}, \frac{5}{9}, \frac{3}{10} > * < \frac{2}{11}, \frac{5}{9}, \frac{3}{10} >= \langle \frac{16}{6655}, \frac{25}{486}, \frac{81}{10000} \rangle$
7. It's Private. $P(S|Y)P(Y) =< \frac{4}{11}, \frac{1}{9}, \frac{5}{10} > * < 0.4, 0.3, 0.3 >= \langle \frac{8}{55}, \frac{1}{30}, \frac{3}{20} \rangle$

Problem 2

Spam Class

Precision: $\frac{1}{3}$

Recall: $\frac{1}{1}$

Accuracy: $\frac{5}{7}$

F_1 Score: $2 * \frac{\frac{1}{3}}{\frac{4}{3}} = \frac{1}{2}$

Non-Spam Class

Precision: $\frac{4}{4} = \frac{1}{1}$

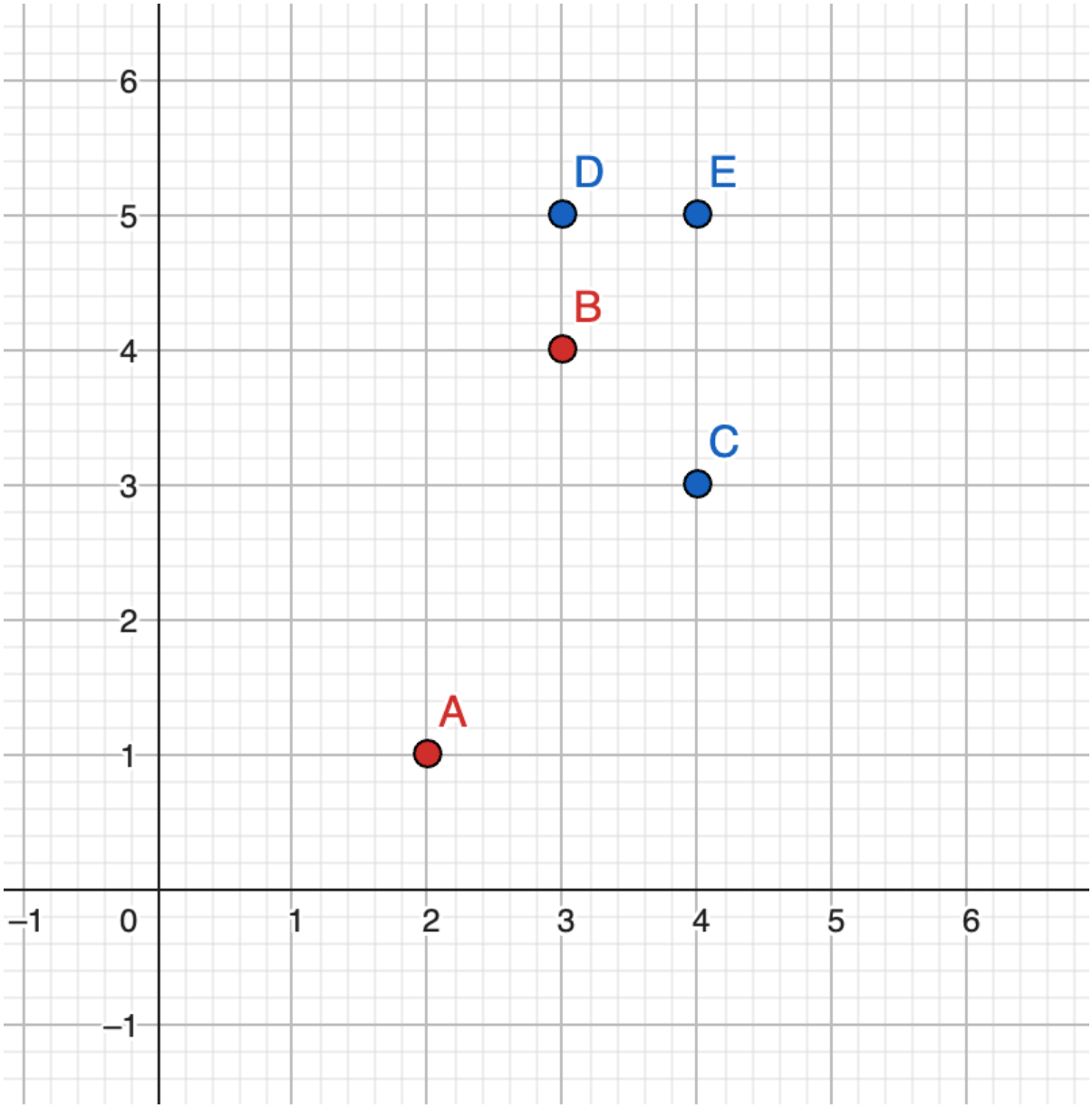
Recall: $\frac{4}{6} = \frac{2}{3}$

Accuracy: $\frac{5}{7}$

F_1 Score: $2 * \frac{\frac{2}{3}}{\frac{5}{3}} = \frac{4}{5}$

Problem 3

(a)



It is possible considering there are lots of lines (like DC) that can separate two points groups.

(b)

Step	w	Activation	Correct?
1	[-1,0,0]	-1	Yes
2	[-1,0,0]	-1	No
3	[0,4,3]	25	Yes
4	[0,4,3]	27	Yes
5	[0,4,3]	24	No
6	[-1,1,-1]		

So W is $< 1, -1, 1 >$

(c)

No we can't. It's not correct. Because for some data point, when you multiply f to the weight, the result is not correct.

(d)

1. Yes we can because the separation can be represented as a line $x + y = 9$
2. No, we can't. Because it is not linearly separable.
3. No, we can't. Because it is also not linearly separable.