```
1)
a)False Positive Rate = FP / (FP + TN) = 10 / (10 + 25) = .286
False Negative Rate = FN / (FN + TP) = 9 / (9 + 36) = .2
b)If we increase t, our threshhold for classifying positives, then we will wont classify as many points as
positive and classify more as negative.
c)
```

CS148 HW #4

2)

$P(fire) = .01$

$P(alarm \mid fire) = .99$

$P(\neg alarm \mid \neg fire) = .91$

a) Prob of alarm

$$P(fire \mid alarm) = \frac{P(alarm \mid fire)\, P(fire)}{P(alarm)}$$

$$P(alarm) = \frac{P(alarm \mid fire)\, P(fire)}{P(fire \mid alarm)} = \frac{.99 \cdot .01}{P(fire \mid alarm)}$$

$$P(fire \mid alarm) = \frac{P(alarm \mid fire)\, P(fire)}{P(alarm \mid fire) P(fire) + P(alarm \mid \neg fire) P(\neg fire)}$$

$$= \frac{\cancel{.99} \cdot .01}{\cancel{.99} \cdot .01 + .1 \cdot \cancel{.99}} = \frac{.01}{.01 + .1} = \frac{.01}{.11} = 1/11 = .091$$

$$P(alarm) = \frac{.99 \cdot .01}{.091} = .109$$

b) See $P(fire \mid alarm) = .091$ above

c) I would say that the fire alarm is still useful as it sounds the alarm when there is a fire 99% of the time and only gives false alarms about 10% of the time.

$$\log((1+e^{-\beta^T x_i})^{-1})$$

3) $L(\beta) = -\sum y_i \log \frac{1}{(1+e^{-(\beta^T x_i)})} + (1-y_i)\log\left(1-\frac{1}{1+e^{-(\beta^T x_i)}}\right)$

Prove:

$$\frac{\partial L(\beta)}{\partial \beta_j} = \sum\left(\frac{1}{1+e^{-\beta^T x_i}} - y_i\right)x_i^j$$

assuming log base e

$$\frac{\partial L(\beta)}{\partial \beta_j} = -\sum y_i \frac{(+1)(1+e^{-\beta^T x_i})^{-2}\cdot(+x_i \cdot e^{-\beta_j x_i})}{\cancel{\ln(e)}\cdot(1+e^{-\beta^T x_i})^{-1}}$$

$$+(1-y_i)\cdot\frac{0+(+1)(1+e^{-\beta^T x_i})^{-2}(-x_i \cdot e^{-\beta_i x_j})}{\cancel{\ln(e)}\cdot\left(1-\frac{1}{1+e^{-\beta x_i}}\right)}$$

$$1-\frac{1}{1+e^{-\beta^T x_i}}$$

$0-$

$$\frac{1+e^{-\beta x_i}}{1+e^{-\beta x_i}} - \frac{1}{1+e^{-\beta x_i}}$$

$$= \frac{e^{-\beta x_i}}{1+e^{-\beta x_i}}$$

$$\frac{\partial L(\beta)}{\partial \beta_j} = -\sum y_i \frac{(+x_i e^{-\beta_j x_i})\cancel{(1+e^{-\beta^T x_i})}}{(1+e^{-\beta^T x_i})^{\cancel{2}}} + (1-y_i)\cdot\frac{(-x_j \cancel{e^{-\beta x}})\cancel{(1+e^{-\beta x})}}{\cancel{(e^{-\beta^T x_i})}(1+e^{-\beta^T x_i})^{\cancel{2}}}$$

$$\frac{\partial L(\beta)}{\partial \beta_j} = -\sum +y_i \frac{(+x_i e^{-\beta_j x_i})}{(1+e^{-\beta_j^T x_i})} + \frac{-x_j}{(1+e^{-\beta x_i})} + \frac{x_j y_i}{(1+e^{-\beta x_i})}$$

$$= \sum \frac{-y_i x_i e^{-\beta_j x_i} + x_i - x_i y_i}{(1+e^{-\beta^T x_i})}$$

$$= \sum x_i\left(\frac{-y_i e^{-\beta x_i} + 1 - y_i}{(1+e^{-\beta x})}\right) = \sum x_i \frac{(1 - y_i(1+e^{-\beta x}))}{(1+e^{-\beta x_i})}$$

$$= \sum x_i^j\left(\frac{1}{(1+e^{-\beta x_i})} - y_i\right)$$

4) a)One vs All: One vs all trains n models to recognize each potential class in the data. If it is that class then 1, else if it is anything else it ouputs 0. We then classify based on which classifier identifies the data point. Easy to implement for each class but number of data points, and thus accuracy, may be different for each model. May also have issues with ties and cases where no clasifier accepts the data point. 3)All vs All: Create a classifier between each possible combination of possible classes. We train each classifier only on the data it is able to identify. We then run all classifiers and choose the class that it is identified as the most. Once again we run into the issue where the training set for each is small and unbalanced and may have ties.

5) a)False, it should be the other way around. The positive predictive value is the probability that the true label of a sample is positive given the prediction was positive. b)True as we create a classifier for each possible class (4), each with two betas. c)False, the decision boundary for logistic regression is linear. d)False, we have to take into account how often fraud actually occurs. If fraud only occurs 30 times out of a thousand, and we miss it those thirty times but predict correctly that there was no fraud the other 970 times, then we have a model 97% percent accuracy, but is actually useless. e)True, as those are the cases where we have true-positives and true-negatives.