HOMEWORK 6: LEARNING THEORY, PROBABILISTIC LEARNING, FAIRNESS METRICS, AND SOCIETAL IMPACT

10-301/10-601 Introduction to Machine Learning (Fall 2024)
https://www.cs.cmu.edu/~mgormley/courses/10601/
OUT: Sunday, October 27th
DUE: Saturday, November 2nd
TAs: Mihir, Rohini, Hailey, Kushagra, Joaquin

Homework 6 covers topics on Learning Theory, Fairness Metrics, and Societal Impacts. The homework includes multiple choice, True/False, and short answer questions. There will be no consistency points in general, so please make sure to double check your answers to all parts of the questions!

START HERE: Instructions

- **Collaboration Policy**: Please read the collaboration policy here: http://www.cs.cmu.edu/~mgormley/courses/10601/syllabus.html
- **Late Submission Policy: For this homework, you will only have 2 late days instead of the usual 3.** This allows us to provide feedback before the exam. See the late submission policy here: http://www.cs.cmu.edu/~mgormley/courses/10601/syllabus.html
- **Submitting your work:** You will use Gradescope to submit answers to all questions and code. Please follow instructions at the end of this PDF to correctly submit all your code to Gradescope.
  - **Written:** For written problems such as short answer, multiple choice, derivations, proofs, or plots, please use the provided template. Submissions can be handwritten onto the template, but should be labeled and clearly legible. If your writing is not legible, you will not be awarded marks. If your scanned submission misaligns the template, there will be a 5% penalty. Alternatively, submissions can be written in LaTeX. Each derivation/proof should be completed in the boxes provided. If you do not follow the template, your assignment may not be graded correctly by our AI assisted grader.

Instructions for Specific Problem Types

For "Select One" questions, please fill in the appropriate bubble completely:

Select One: Who taught this course?

- ● Matt Gormley
- ○ Marie Curie
- ○ Noam Chomsky

If you need to change your answer, you may cross out the previous answer and bubble in the new answer:

Select One: Who taught this course?

- ● Henry Chai
- ○ Marie Curie
- ~~●~~ Noam Chomsky

For "Select all that apply" questions, please fill in all appropriate squares completely:

Select all that apply: Which are instructors for this course?

- ■ Matt Gormley
- ■ Henry Chai
- □ Isaac Newton
- □ I don't know

Again, if you need to change your answer, you may cross out the previous answer(s) and bubble in the new answer(s):

Select all that apply: Which are the instructors for this course?

- ■ Matt Gormley
- ■ Henry Chai
- ~~■~~ Isaac Newton
- ~~■~~ I don't know

For questions where you must fill in a blank, please make sure your final answer is fully included in the given space. You may cross out answers or parts of answers, but the final answer must still be within the given space.

Fill in the blank: What is the course number?

10-601	10-~~6~~301

Written Questions (75 points)

1 LaTeX Point and Template Alignment (1 points)

1. (1 point) **Select one:** Did you use LaTeX for the entire written portion of this homework?

   ● Yes

   ○ No

2. (0 points) **Select one:** I have ensured that my final submission is aligned with the original template given to me in the handout file and that I haven't deleted or resized any items or made any other modifications which will result in a misaligned template. I understand that incorrectly responding yes to this question will result in a penalty equivalent to 2% of the points on this assignment.
   Note: Failing to answer this question will not exempt you from the 2% misalignment penalty.

   ● Yes

2 Learning Theory (19 points)

1. Neural the Narwhal is given a classification task to solve, which he decides to use a decision tree learner with 2 binary features X_1 and X_2. On the other hand, you think that Neural should not have used a decision tree. Instead, you think it would be best to use logistic regression with 16 real-valued features in addition to a bias term. You want to use PAC learning to check whether you are correct. You first train your logistic regression model on N examples to obtain a training error $\hat{R}$.

   (a) (1 point) Which of the following case of PAC learning should you use for your logistic regression model?

   ○ Finite and realizable

   ○ Finite and agnostic

   ○ Infinite and realizable

   ● Infinite and agnostic

   (b) (2 points) What is the upper bound on the true error R in terms of $\hat{R}$, δ, and N? You may use big-$\mathcal{O}$ notation if necessary. Write only the final answer. Your work will *not* be graded.
   Note: Your answer may not contain any other symbols.

   Your Answer

   $\hat{R} + O\left(\sqrt{\frac{1}{N}\log\frac{1}{\delta}}\right)$

(c) (3 points) **Select one:** You want to argue your method has a lower bound on the true error as compared to the Neural's true error bound. Assume that you have obtained enough data points to satisfy the PAC criterion with the same ϵ and δ as Neural. Which of the following is true?

- ○ Neural's model will always classify unseen data more accurately because it only needs 2 binary features and therefore is simpler.
- ○ You must first regularize your model by removing 14 features to make any comparison at all.
- ○ It is sufficient to show that the VC dimension of your classifier is higher than that of Neural's, therefore having a lower bound for the true error.
- ● It is necessary to show that the training error you achieve is lower than the training error Neural achieves.

2. In lecture, we saw that we can use our sample complexity bounds to derive bounds on the true error for a particular algorithm. Consider the sample complexity bound for the infinite, agnostic case:

$$N = O\left(\frac{1}{\epsilon^2}\left[\text{VC}(\mathcal{H}) + \log\frac{1}{\delta}\right]\right).$$

(a) (2 points) What is the big-$\mathcal{O}$ bound of ϵ in terms of N, δ, and VC($\mathcal{H}$)?
Note: $A = \mathcal{O}(B)$ (for some value B) $\Leftrightarrow$ there exists a constant $c \in \mathbb{R}$ such that $A \leq cB$.

Your Answer

$$N = O\left(\frac{1}{\epsilon^2}\left[\text{VC}(\mathcal{H}) + \log\frac{1}{\delta}\right]\right) \tag{1}$$

$$\Rightarrow N \leq c\left(\frac{1}{\epsilon^2}\left[\text{VC}(\mathcal{H}) + \log\frac{1}{\delta}\right]\right), \text{where } c \in \mathbb{R} \tag{2}$$

$$\Rightarrow \epsilon^2 \leq c\left(\frac{1}{N}\left[\text{VC}(\mathcal{H}) + \log\frac{1}{\delta}\right]\right) \tag{3}$$

$$\Rightarrow \epsilon \leq c'\sqrt{\left(\frac{1}{N}\left[\text{VC}(\mathcal{H}) + \log\frac{1}{\delta}\right]\right)}, \text{where } c' \in \mathbb{R} \tag{4}$$

$$\Rightarrow \epsilon = O\sqrt{\left(\frac{1}{N}\left[\text{VC}(\mathcal{H}) + \log\frac{1}{\delta}\right]\right)} \tag{5}$$

(b) (2 points) Now, using the definition of ϵ (i.e. $|R(h) - \hat{R}(h)| \leq \epsilon$) and your answer to part a, prove that with probability at least $(1 - \delta)$:

$$R(h) \leq \hat{R}(h) + O\left(\sqrt{\frac{1}{N}\left[\text{VC}(\mathcal{H}) + \log\frac{1}{\delta}\right]}\right).$$

Your Answer

Since we have $|R(h) - \hat{R}(h)| \leq \epsilon$ with probability at least (1-δ), and $\epsilon = O\sqrt{\left(\frac{1}{N}\left[\text{VC}(\mathcal{H}) + \log\frac{1}{\delta}\right]\right)}$, then we can apply the ϵ to the upper bound of $|R(h) - \hat{R}(h)|$ and get the following equation: $R(h) \leq \hat{R}(h) + O\sqrt{\left(\frac{1}{N}\left[\text{VC}(\mathcal{H}) + \log\frac{1}{\delta}\right]\right)}$.

3. (3 points) Consider the hypothesis space of functions that map M binary attributes to a binary label. A function f in this space can be characterized as $f : \{0,1\}^M \rightarrow \{0,1\}$. Neural the Narwhal says that regardless of the value of M, a hypothesis class containing all possible functions in this space can always shatter 2^M points. Is Neural wrong? If so, provide a counterexample. If Neural is right, briefly explain why in 1-2 *concise* sentences.

Your Answer

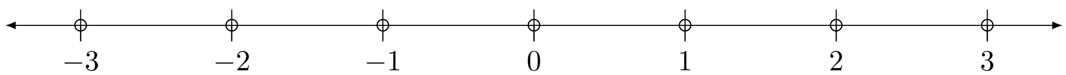
Neural is right, since the number of all possible functions that maps M binary attributes to a binary label is 2^{2^M}, which is indeed sufficient to classify all the labeling of 2^M points, which is also 2^{2^M}. Hence the hypothesis class containing all these functions always shatter 2^M points.

4. Consider an instance space $\mathcal{X}$ which is the set of real numbers.

(a) (3 points) **Select one:** What is the VC dimension of hypothesis class H, where each hypothesis h in H is of the form "if $a < x < b$ or $c < x < d$ then $y = 1$; otherwise $y = 0$"? (i.e., H is an infinite hypothesis class where a, b, c, and d are arbitrary real numbers).

- ○ 2
- ○ 3
- ● 4
- ○ 5
- ○ 6

(b) (3 points) Given the set of points in $\mathcal{X}$ below, construct a labeling of some subset of the points to show that any dimension larger than the VC dimension of H by *exactly* 1 is incorrect (e.g. if the VC dimension of H is 3, only fill in the answers for 4 of the points). Fill in the boxes such that for each point in your example, the corresponding label is either 0 or 1. For points you are not using in your example, write N/A (do *not* leave the answer box blank).



Answer for -3	Answer for -2	Answer for -1
NA	1	0

Answer for 0	Answer for 1	Answer for 2	Answer for 3
1	0	1	NA

3 MLE/MAP (7 points)

1. (1 point) **True or False:** Suppose you place a Beta prior over the Bernoulli distribution, and attempt to learn the parameter θ of the Bernoulli distribution from data. Further suppose an adversary chooses "bad" but finite hyperparameters for your Beta prior in order to confuse your learning algorithm. As the number of training examples grows to infinity, the MAP estimate of θ can still converge to the MLE estimate of θ.

   ● True

   ○ False

2. (2 points) **Select one:** Let Γ be a random variable with the following probability density function (pdf):

$$f(\gamma) = \begin{cases} 2\gamma & \text{if } 0 \leq \gamma \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

   Suppose another random variable Y, which is conditioning on Γ, follows an exponential distribution with $\lambda = 3\gamma$. Recall that the exponential distribution with parameter λ has the following pdf:

$$f_{exp}(y) = \begin{cases} \lambda e^{-\lambda y} & \text{if } y \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

   What is the MAP estimate of γ given $Y = \frac{2}{3}$ is observed?

Your Answer
$\gamma = 1$

3. (4 points) Neural the Narwhal found a mystery coin and wants to know the probability of landing on heads by flipping this coin. He models the coin toss as sampling a value from Bernoulli(θ) where θ is the probability of heads. He flips the coin three times and the flips turned out to be heads, tails, and heads. An oracle tells him that $\theta \in \{0, 0.25, 0.5, 0.75, 1\}$, and *no other values of θ should be considered.*

   Find the MLE and MAP estimates of θ. Use the following prior distribution for the MAP estimate:

$$p(\theta) = \begin{cases} 0.9 & \text{if } \theta = 0 \\ 0.04 & \text{if } \theta = 0.25 \\ 0.03 & \text{if } \theta = 0.5 \\ 0.02 & \text{if } \theta = 0.75 \\ 0.01 & \text{if } \theta = 1 \end{cases}.$$

   Again, remember that $\theta \in \{0, 0.25, 0.5, 0.75, 1\}$, so the MLE and MAP should also be one of them.

MLE of θ	MAP of θ
0.75	0.5

4 Fairness Metrics (21 points)

Neural works for the Bank of ML and is given the following dataset from another bank on whether or not to issue a loan to individuals. Each row in this dataset represents one individual's data, which includes their FICO credit score, their savings rate (percentage of their income that goes into their savings), and credit history in months. The data was collected in two different cities, city A and city B, as denoted in the first column. The "Label" column refers to the true label, where "1" refers to loan issued, and "0" refers to no loan issued. A csv file of this dataset could be found in the handout folder.

Region	FICO Score	Savings Rate (%)	Credit History (months)	Label
A	544.0625	28.0	21	1
A	489.0625	33.9	40	0
A	433.125	62.3	100	0
A	429.0625	56.7	203	1
A	417.8125	56.5	5	0
A	506.5625	32.7	75	1
A	400.625	60.7	216	0
A	836.875	10.7	86	1
A	471.875	36.2	92	1
A	402.8125	62.0	199	0
B	809.4285714	5.6	213	1
B	480.9375	40.2	72	1
B	505.0	31.1	20	0
B	438.4375	51.3	122	0
B	385.9375	76.2	89	0
B	505.625	34.7	39	1
B	514.0625	31.0	41	1
B	385.9375	76.2	89	0
B	446.25	44.5	51	0
B	428.75	55.6	215	1

1. Neural took the average value of the features (for example, the average value for the first data point is 197.69), and developed the following observation. In general, for all three features in this dataset, a high value indicates better credibility. Hence Neural trained the following decision stump on this dataset: if the average feature value is above the median (198.09), then we determine that the individual will receive the loan (prediction = 1). Otherwise, we decide that the individual will not receive the loan. **For parts (a), (b), (c) below, please round your answer to three decimal places**.

   (a) (1 point) Using the model that Neural proposed, what is the training error rate on the entire dataset?

Your Answer

0.4

(b) (1 point) What is the training error rate for region A?

Your Answer
0.4

(c) (1 point) What is the training error rate for region B?

Your Answer
0.4

(d) (1 point) How many false positives were there in region A?

Your Answer
3

(e) (1 point) How many false negatives were there in region A?

Your Answer
1

(f) (1 point) How many false positives were there in region B?

Your Answer
1

(g) (1 point) How many false negatives were there in region B?

Your Answer
3

2. (2 points) **True or False**: Using your responses to the previous question, we achieve statistical parity between regions A and B. Justify your answer.

   ○ True

   ● False

Your Answer

The selection rate (portion of positive prediction instance) of region A is 70%, while the selection rate of region B is only 30%. So, we do not achieve statistical parity between region A and B.

3. (2 points) **True or False**: We achieve equality of accuracy between regions A and B. Justify your answer.

   ● True

   ○ False

Your Answer

The accuracy of region A is 60%, and the accuracy of region B is also 60%. So, we do achieve the equality of accuracy between region A and B.

4. (2 points) **True or False**: We achieve equality of FPR/FNR between regions A and B. Justify your answer.

   ○ True

   ● False

Your Answer

The FPR of region A is 60%, and the FNR of region A is 20%. On the other hand, the FPR of region B is 20%, and the FNR of region B is 60%. As the result, we do not achiece equality of FPR/FNR between regions A and B.

5. (2 points) **True or False**: We achieve equality of PPV/NPV between regions A and B. Justify your answer.

   ◯ True

   ● False

Your Answer

The PPV of region A is $\frac{4}{7}$, and the NPV of region A is $\frac{2}{3}$. On the other hand, the PPV of region B is $\frac{2}{3}$, and the NPV of region B is $\frac{4}{7}$. So, we do not achieve equality of PPV/NPV between regions A and B.

6. (3 points) Using your responses from the previous questions, comment on the fairness of this model between cities A and B.

Your Answer

The model does not achieve Independence(statistical parity), Separation (equality of FPR/FNR), and Sufficiency (equality of PPV/NPV). Accordingly, the model is not fair between cities A and B.

7. (3 points) A Type I error occurs when you erroneously predict a positive label (false positive), and a Type II error is when you erroneously predict a negative label (false negative). Compare and contrast the consequences of making a Type I error and Type II error in this setting. Which would cause more significant consequences?

Your Answer

Type I error (false positive) means you **wrongly** give a person the loan. The consequence is that there would be a huge risk that the bank would lose its money, as the person may not be able to repay it. On the other hand, the Type II error (false negative) means that you reject to give loan to an applicant who is actually qualified. The consequence is that the bank may loss the profit from interest payments. I think the false positive one would have more significant consequence because the bank would directly loss its capitals due to unpaid loans, while the false negative would only cause the bank to lose its opportunity to earn more money. However, to the perspective of the applicants, the false negative one would have more significant consequence, as he/she would loss the opportunity to buy a house, start a business or do something they want.

5 Societal Impacts (27 points)

The fictional country, Xtopia, is in the midst of an epidemic. The Xtopian healthcare system has been under a great deal of strain in the past year due to a regional epidemic caused by an airborne virus called Xvid. The number of hospital beds is limited and as the result, healthcare professionals have to frequently make very difficult choices about which subset of Xvid patients can be hospitalized. Hospital care greatly increases the chance of recovering from the illness with no subsequent long-term health complications.

To save time and make these decisions more efficient and consistent, a team of ML practitioners have been brought in to automate the decision-making process. They have been given access to a data set consisting of the information about prior Xvid patients who sought hospital care along with the binary decision made about them by the hospital doctors ('+' indicates hospitalization and '-' indicates no hospitalization). The ML team has determined that the decision about each patient is highly correlated with his/her age as well as his/her prior utilization of medical insurance. This observation reflects the fact that Xtopian doctors are on average more likely to allocate scarce medical resources to the young and the vulnerable (i.e., those with prior medical conditions and comorbidities). Here, the insurance utilization serves as a proxy for severity of the patient's health conditions.

Figure 1 provides a snapshot of the Xvid training data and the predictive model that the ML team has come up with. Each instance corresponds to an individual patient, and each patient belongs to one of the two socially salient groups in Xtopia, indicated by blue and red.

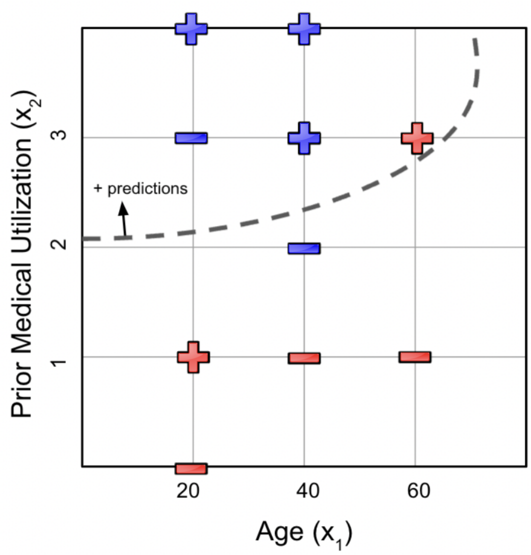


Figure 1: Xvid Training Data

Answer the following questions with respect to the above hypothetical context and data set.

1. (2 points) **Select all that apply:** Does the predictive model above satisfy the following notions of fairness across blue and red groups?

   - ☐ False Negative Rate (FNR) parity
   - ☐ False Positive Rate (FPR) parity
   - ☐ Negative Predictive Value (NPV) parity
   - ☐ Positive Predictive Value (PPV) parity
   - ■ Error parity
   - ☐ Statistical parity (or Selection rate parity)
   - ☐ None of the above

2. (2 points) **Select all that apply:** Which of the above notions of fairness would be satisfied if the ML team could train a model with 0 true error (i.e., a model that always predicts the correct label for every patient)?

   - ■ False Negative Rate (FNR) parity
   - ■ False Positive Rate (FPR) parity
   - ■ Negative Predictive Value (NPV) parity
   - ■ Positive Predictive Value (PPV) parity
   - ■ Error parity
   - ☐ Statistical parity (or Selection rate parity)
   - ☐ None of the above

3. (2 points) **Select all that apply:** Which of the above notions of fairness would be satisfied in expectation by a random classifier (i.e., a model that makes a randomized prediction for every patient: with probability 0.5 the patient is hospitalized regardless of their attributes)?

   - ■ False Negative Rate (FNR) parity
   - ■ False Positive Rate (FPR) parity
   - ☐ Negative Predictive Value (NPV) parity
   - ☐ Positive Predictive Value (PPV) parity
   - ☐ Error parity
   - ■ Statistical parity (or Selection rate parity)
   - ☐ None of the above

4. (2 points) From the perspective of a patient subject to the predictions made by this model, the violation of which of the parity conditions below would be most problematic? Justify your answer.

   ● False Negative Rate (FNR) parity
   ○ False Positive Rate (FPR) parity
   ○ Negative Predictive Value (NPV) parity
   ○ Positive Predictive Value (PPV) parity
   ○ Error parity
   ○ Statistical parity (or Selection rate parity)
   ○ None of the above

Your Answer

The False Negative Rate (FNR) means the portion of patients that truly need hospital care but be predicted as negative, which may lead to serious consequence on the health of these patients. The disparity of FNR means some groups have higher FNR, meaning that the group may have more people suffer from the FNR disparity bias of the model, causing their risk on not being able to recover from illness due to no hospitalization.

5. (2 points) **Causes of unfairness:** Name one potential cause of disparity in false negative rates across the two groups in the above context.

Your Answer

I think one potential cause is the feature that the model use, like age and prior utilization of medical insurance, is insufficient to identify all patients in different groups who need hospitalization. For example, assuming that a model considers a person needs more care if he uses more health insurance, and a young and poor person may not have enough money to have medical insurance, then the model may decide that the person does not need hospitalization, while the person may already be infected and indeed needs health care. On the other hand, a rich family may have constant check-up that misleads the model to decide they need more cares. As the model does not have the ability to discover the truly high-risk group, such as old or lower-income patients, they may be neglected and cause the high False Negative Rate and thus the disparity over groups.

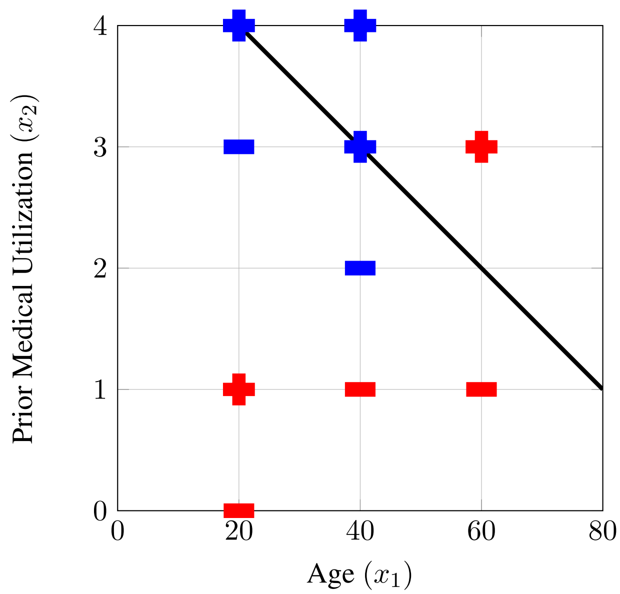
6. **Fairness interventions:** consider the following pre-processing method to improve statistical parity:

```
While the selection rate is unequal across the two groups:

(a) Pick the group with lowest selection rate.
(b) From this group in the training data, pick the data point
    closest to the decision boundary predicted as negative.
(c) Change the label of this instance to positive.
(d) Retrain the model on the modified training data by finding the
    highest accuracy classifier in the hypothesis class.
```

Suppose our hypothesis class is the class of all linear separators defined over $\mathbb{R}^2$.

(a) (1 point) The highest accuracy linear separator is shown in the figure below (assume that points on the decision boundary are characterized as '+'):

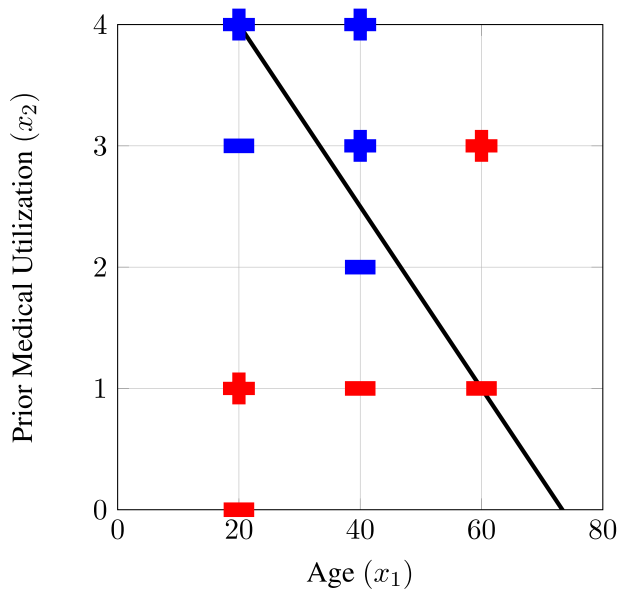


What are the coordinates of the data point whose label would be flipped first by this pre-processing method?

Your Answer

(60, 2)

(b) (2 points) After flipping the label of the point you identified in the previous question, plot the linear separator with the highest accuracy. For your convenience, we have provided a mechanism for you to input your answer by specifying the coordinates of two points on the decision boundary.
Note: Look at the LaTeX comments for instructions on how to redraw the decision boundry.



(c) (1 point) Using the linear decision boundary your plotted in the previous question, what are the coordinates of the data point whose label would be flipped next by this pre-processing method?

Your Answer

(40, 1)

(d) (1 point) **True or False:** The algorithm terminates at this point.

○ True

● False

7. **The fairness impossibility theorem:** Prove by contradiction that if prevalence rate, $r_s = P[Y = 1|S = s]$ across the two groups $s \in \{\text{blue}, \text{red}\}$ is different, then there does not exists a classifier that can satisfy PPV parity, FPR parity, and FNR parity simultaneously.

   (a) (4 points) Verify that the following identify holds for any $s \in \{\text{blue}, \text{red}\}$:

$$FPR_s = \frac{r_s}{1 - r_s} \times (1 - FNR_s) \times \frac{(1 - PPV_s)}{PPV_s}$$

.

Your Answer

We can verify the identity as the following:

$$\frac{r_s}{1-r_s} = \frac{(TP+FN)}{(FP+TN)} \tag{6}$$

$$(1-FNR_s) = 1 - \frac{FN}{(TP+FN)} = \frac{TP}{(TP+FN)} \tag{7}$$

$$\frac{(1-PPV_s)}{PPV_s} = \frac{(1-TP/(TP+FP))}{TP/(TP+FP)} = \frac{FP/(TP+FP)}{TP/(TP+FP)} = \frac{FP}{TP} \tag{8}$$

$$\frac{r_s}{1-r_s} \times (1-FNR_s) \times \frac{(1-PPV_s)}{PPV_s} = \frac{(TP+FN)}{(FP+TN)} \times \frac{TP}{(TP+FN)} \times \frac{FP}{TP} \tag{9}$$

$$= \frac{FP}{FP+TN} \tag{10}$$

$$= FPR_s \tag{11}$$

As the result, the identity holds for any s $\in$ {blue, red}.

(b) (4 points) Show that the expression from part (a) can be rewritten as

$$1/r_s = 1 + \frac{(1-FNR_s)}{FPR_s} \times \frac{(1-PPV_s)}{PPV_s}$$

Your Answer

We can show that expression in part(a) can be rewritten as the following:

$$(1-r_s)FPR_s = r_s \times (1-FNR_s) \times \frac{1-PPV_s}{PPV_s} \tag{12}$$

$$\frac{1-r_s}{r_s} = \frac{(1-FNR_s)}{FPR_s} \times \frac{1-PPV_s}{PPV_s} \tag{13}$$

$$\frac{1}{r_s} = 1 + \frac{(1-FNR_s)}{FPR_s} \times \frac{1-PPV_s}{PPV_s} \tag{14}$$

(c) (4 points) Finally, using results from parts (a) and (b), show that if FPR_s, FNR_s, and PPV_s are equal for $s \in \{\text{blue}, \text{red}\}$, then r_s must be equal for $s \in \{\text{blue}, \text{red}\}$, which is a contradiction.

Your Answer

From part(a) and part(b), we have the following equations:

$$\frac{1}{r_s} = 1 + \frac{(1 - FNR_s)}{FPR_s} \times \frac{1 - PPV_s}{PPV_s} \tag{15}$$

$$\rightarrow r_s = \frac{1}{1 + \frac{(1-FNR_s)}{FPR_s} \times \frac{1-PPV_s}{PPV_s}} \tag{16}$$

As the result, if we have same FPR_s, FNR_s and PPV_s, then we would have the same r_s, for s $\in$ $\{$blue, red$\}$, which is a contradiction.

6 Collaboration Questions

After you have completed all other components of this assignment, report your answers to these questions regarding the collaboration policy. Details of the policy can be found here.

1. Did you receive any help whatsoever from anyone in solving this assignment? If so, include full details.
2. Did you give any help whatsoever to anyone in solving this assignment? If so, include full details.
3. Did you find or come across code that implements any part of this assignment? If so, include full details.

Your Answer

1. No.
2. No.
3. No.