From the equations presented below, express the probability of a tweet being positive given that it contains the word happy in terms of the probability of a tweet containing the word happy given that it is positive

$$P(\text{Positive}|\text{“happy”}) = \frac{P(\text{Positive} \cap \text{“happy”})}{P(\text{“happy”})} \qquad P(\text{“happy”}|\text{Positive}) = \frac{P(\text{“happy”} \cap \text{Positive})}{P(\text{Positive})}$$

○ $$P(\text{Positive}|\text{“happy”}) = P(\text{“happy”}|\text{Positive}) \times \frac{P(\text{“happy”})}{P(\text{Positive})}$$

○ $$P(\text{Positive}|\text{“happy”}) = P(\text{“happy”} \cap \text{Positive}) \times \frac{P(\text{Positive})}{P(\text{“happy”})}$$

○ $$P(\text{Positive}|\text{“happy”}) = P(\text{“happy”} \cap \text{Positive}) \times \frac{P(\text{“happy”})}{P(\text{Positive})}$$

◉ $$P(\text{Positive}|\text{“happy”}) = P(\text{“happy”}|\text{Positive}) \times \frac{P(\text{Positive})}{P(\text{“happy”})}$$

✓ **Correct**

That's right. You just derived Bayes' rule.

Question

Here, again, is Bayes' rule:

$$P(X|Y) = P(Y|X) \times \frac{P(X)}{P(Y)}$$

Suppose that in your dataset, 25% of the positive tweets contain the word 'happy'. You also know that a total of 13% of the tweets in your dataset contain the word 'happy', and that 40% of the total number of tweets are positive. You observe the tweet: ''happy to learn NLP'. What is the probability that this tweet is positive?

- (selected) P(Positive | "happy") = 0.77
- P(Positive | "happy") = 0.08
- P(Positive | "happy") = 0.10
- P(Positive | "happy") = 1.92

✓ **Correct**

That's right. You just applied Bayes' rule.

1. Assume that there are 2 happy people and 2 unhappy people in a room. Concretely, persons A and B are happy and persons C and D are unhappy. If you were to randomly pick a person from the room, what is the probability that the person is happy.

1 / 1 point

- (selected) 1/2
- 1/4
- 3/4
- 0

Correct

2. Assume that there are 2 happy people and 2 unhappy people in a room. Concretely, persons A and B are happy and persons C and D are unhappy. If a friend showed you the part of the room where the two happy people are, what is the probability that you choose person B?

1 / 1 point

- (selected) 1/2
- 1/4
- 3/4
- 1

Correct

3. From the equations presented below, express the probability of a tweet being positive given that it contains the word happy in terms of the probability of a tweet containing the word happy given that it is positive

1 / 1 point

$$P(\text{ Positive } \mid \text{ "happy" }) = \frac{P(\text{ Positive} \cap \text{"happy" })}{P(\text{ "happy" })}$$

$$P(\text{ "happy" } \ \text{ Positive }) = \frac{P(\text{ "happy"} \cap \text{Positive })}{P(\text{ Positive })}$$

- (selected) $P(\text{ Positive } \mid \text{ happy }) = P(\text{ happy } \mid \text{ Positive }) \times \frac{P(\text{ Positive })}{P(\text{ happy})}$
- $P(\text{ Positive } \mid \text{ happy }) = P(\text{ "happy" } \mid \text{ Positive }) \times \frac{P(\text{ happy })}{P(\text{ Positive })}$
- $P(\text{ Positive } \bigcap \text{ happy }) = P(\text{ happy } \mid \text{ Positive }) \times \frac{P(\text{ Positive })}{P(\text{ happy})}$
- $P(\text{ Positive } \bigcap \text{ happy }) = P(\text{ "happy" } \mid \text{ Positive }) \times \frac{P(\text{ happy })}{P(\text{ Positive })}$

Correct

Yes, that is the correct answer.

4. Bayes rule is defined as

1 / 1 point

- (selected) $P(X \mid Y) = P(Y \mid X) \times \frac{P(X)}{P(Y)}$
- $P(X \mid Y) = P(Y \mid X) \times \frac{P(Y)}{P(X)}$
- $P(X \mid Y) = P(X \mid Y) \times \frac{P(X)}{P(Y)}$
- $P(X \mid Y) = P(Y \mid X) \times \frac{P(X)}{P(Y|X)}$

Correct

Yes.

5. Suppose that in your dataset, 25% of the positive tweets contain the word ‘happy’. You also know that a total of 13% of the tweets in your dataset contain the word 'happy', and that 40% of the total number of tweets are positive. You observe the tweet: ''happy to learn NLP'. What is the probability that this tweet is positive?

1 / 1 point

0.77

Correct

**** That’s right. You just applied Bayes’ rule.

6. The log likelihood for a certain word w_i is defined as:

1 / 1 point

$\log(\frac{P(w_i|pos)}{P(w_i|neg)})$.

- [x] Positive numbers imply that the word is positive.

Correct

- [ ] Positive numbers imply that the word is negative.
- [x] Negative numbers imply that the word is negative.

Correct

- [ ] Negative numbers imply that the word is positive.

7. The log likelihood mentioned in lecture, which is the log of the ratio between two probabilities is bounded between

1 / 1 point

- ○ -1 and 1
- ◉ $-\infty$ and ∞
- ○ 0 and ∞
- ○ 0 and 1

Correct

Yes!

6. When implementing naive Bayes, in which order should the following steps be implemented. 1 / 1 point

- (selected)
  1. Get or annotate a dataset with positive and negative tweets
  2. Preprocess the tweets: process_tweet(tweet) ➞
  3. Compute freq(w, class)
  4. Get P(w | pos), P(w | neg)
  5. Get λ(w)
  6. Compute logprior = log(P(pos) / P(neg))

-
  1. Get or annotate a dataset with positive and negative tweets
  2. Preprocess the tweets: process_tweet(tweet) ➞
  3. Compute freq(w, class)
  4. Get λ(w)
  5. Get P(w | pos), P(w | neg)
  6. Compute logprior = log(P(pos) / P(neg))

-
  1. Get or annotate a dataset with positive and negative tweets
  2. Compute freq(w, class)
  3. Preprocess the tweets: process_tweet(tweet) ➞
  4. Get P(w | pos), P(w | neg)
  5. Get λ(w)
  6. Compute logprior = log(P(pos) / P(neg))

-
  1. Get or annotate a dataset with positive and negative tweets
  2. Compute freq(w, class)
  3. Preprocess the tweets: process_tweet(tweet) ➞
  4. Compute logprior = log(P(pos) / P(neg)
  5. Get P(w | pos), P(w | neg)
  6. Get λ(w)

Correct

Yes, that is correct.

9. To test naive bayes model, which of the following are required?

1 / 1 point

- (selected) $X_{\text{val}}, Y_{\text{val}}, \lambda, logprior$
- $X_{\text{val}}, Y_{\text{val}}, logprior$
- $X_{\text{val}}, \lambda, logprior$
- $Y_{\text{val}}, \lambda, logprior$

Correct

This is correct.

10. Which of the following is NOT an application of naive Bayes?

1 / 1 point

- ○ Sentiment Analysis
- ○ Author identification
- ○ Information retrieval
- ○ Word disambiguation
- ◉ Numerical predictions

Correct
This is correct.