GENERAL ASSEMBLY

NAÏVE BAYES

Matt Brems

Data Science Immersive, GA DC

DATA SCIENCE PROCESS

1. Define problem.
2. Gather data.
3. Explore data.
4. Model with data.
5. Evaluate model.
6. Answer problem.

LEARNING OBJECTIVES

- By the end of this lesson, students should be able to:
    - **Intuitively explain** how Bayes' Theorem can be used as a modeling tactic.
    - **Implement** Naive Bayes in scikit-learn.
    - **Discuss** assumptions, advantages, and disadvantages of Naive Bayes as a classifier.

CONDITIONAL PROBABILITY

- Recall that we use $P(A)$ to refer to the probability that A occurs, where A is some event.

- If we want to describe the probability that A occurs given that we know something else to be true, we use $P(A|B)$.

- Note that $P(A|B)$ is usually not the same as $P(B|A)$!

BAYES' THEOREM

- Bayes' Theorem (Bayes' Rule) relates $P(A|B)$ to $P(B|A)$.

BREAKING DOWN BAYES' THEOREM

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

- $P(A)$ is the probability that A occurs given no supplemental information.
- $P(B|A)$ is the likelihood of seeing evidence (data) B assuming that A is true.
- $P(B)$ is the probability that B occurs given no supplemental information.
  - $P(B)$ what we scale $P(B|A)P(A)$ by to ensure we are only looking at A within the context of B occurring.

APPLYING BAYES' THEOREM TO SPAM CLASSIFICATION

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

- Bayes' Theorem is really neatly set up as a classification model.
- We can estimate the probability – `.predict_proba()` – that an observation falls into a specific class, then classify that observation – `.predict()` – accordingly!

APPLYING BAYES' THEOREM TO SPAM CLASSIFICATION

$$P(\text{spam}|\text{words in email}) = \frac{P(\text{words in email}|\text{spam})P(\text{spam})}{P(\text{words in email})}$$

APPLYING BAYES' THEOREM TO SPAM CLASSIFICATION

$$P(\text{spam}|\text{words}) = \frac{P(w_1|\text{spam})P(w_2|w_1 \cap \text{spam})P(w_3|w_2 \cap w_1 \cap \text{spam}) \cdots P(\text{spam})}{P(w_1)P(w_2|w_1)P(w_3|w_2 \cap w_1) \cdots}$$

- This gets **really** complicated. Can we simplify this?

NAÏVE BAYES

- The Naïve Bayes classification algorithm is a:
  - classification modeling technique
  - that relies on Bayes Theorem
  - that makes one simplifying assumption.

- **<u>We assume that our features are independent of one another.</u>**

APPLYING BAYES' THEOREM TO SPAM CLASSIFICATION

$$P(\text{spam}|\text{words}) = \frac{P(w_1|\text{spam})P(w_2|w_1 \cap \text{spam})P(w_3|w_2 \cap w_1 \cap \text{spam}) \cdots P(\text{spam})}{P(w_1)P(w_2|w_1)P(w_3|w_2 \cap w_1) \cdots}$$

$$P(\text{spam}|\text{words}) = \frac{P(w_1|\text{spam})P(w_2|\text{spam})P(w_3|\text{spam}) \cdots P(\text{spam})}{P(w_1)P(w_2)P(w_3) \cdots}$$

NAÏVE BAYES

- **Advantages** of making this assumption of feature independence:
  - Easier to calculate probabilities.
  - Empirically, our classifications are surprisingly accurate.

- **Disadvantages** of making this assumption of feature independence:
  - It's incredibly unrealistic, especially in the case of text data.
  - While our classifications are accurate, our predicted probabilities are usually quite bad.

PROCESS OF NAÏVE BAYES

1. Decide which Naïve Bayes model to use.
    - BernoulliNB
    - MultinomialNB
    - GaussianNB

2. Decide what your priors will be.
    - Based on your data. *(default)*
    - Manually set.

3. `.fit()`, `.predict()`!

WHICH NAÏVE BAYES MODEL SHOULD I USE?

- BernoulliNB

- MultinomialNB

- GaussianNB

WHAT SHOULD MY PRIORS SHOULD BE?

$$P(\text{spam}|\text{words in email}) = \frac{P(\text{words in email}|\text{spam})P(\text{spam})}{P(\text{words in email})}$$

- Estimated from data.

- Manually set.

PROCESS OF NAÏVE BAYES

1. Decide which Naïve Bayes model to use.
    - BernoulliNB
    - MultinomialNB
    - GaussianNB

2. Decide what your priors will be.
    - Based on your data. *(default)*
    - Manually set.

3. `.fit()`, `.predict()`!

WHERE DOES NAÏVE BAYES FIT IN OUR MENTAL SET OF MODELS?

INTERVIEW QUESTION

- Suppose we want to detect whether Amazon reviews are spam or ham. How would you do this?