

Вопрос 1

Выполнен

Баллов: 1,00 из 1,00

Проанализируем формулу кросс-энтропии для бинарной классификации. Для единственного примера она вычисляется следующим образом:

$$BCE(\hat{y}, y) = -y \log \hat{y} - (1 - y) \log(1 - \hat{y})$$

где $y \in \{0, 1\}$ - настоящая метка класса для объекта, а $0 < \hat{y} < 1$ - вероятность класса 1, предсказанная моделью.

В процессе обучения классификатора мы минимизируем кросс-энтропию на всех обучающих примерах (сравните с описанием [BCELoss в PyTorch](#)):

$$BCE_{1..n} = \sum_{i=1}^n (-y_i \log \hat{y}_i - (1 - y_i) \log(1 - \hat{y}_i)) \rightarrow \min$$

Например, на двух обучающих примерах $y_1 = 1, y_2 = 1$ формула примет вид:

$$BCE_{1,2} = -\log \hat{y}_1 - \log \hat{y}_2$$

Мы предлагаем вам проанализировать полученную формулу на следующих предсказаниях модели \hat{y}_1, \hat{y}_2 :

- $\hat{y} = [0.99, 0.01]$ - первый объект классифицируется уверенно правильно, а второй уверенно неправильно
- $\hat{y} = [0.5, 0.5]$ - модель не может принять решения, абсолютно неуверенное предсказание
- $\hat{y} = [0.99, 0.45]$ - первый объект классифицируется уверенно правильно, а второй неуверенно неправильно
- $\hat{y} = [0.65, 0.65]$ - оба объекта классифицируются правильно, но классификатор не очень уверен в принятом решении

Какие виды ошибок с точки зрения кросс-энтропии более критичны и насколько это согласуется с Вашими ожиданиями как человека?



Фраза "**Модели выгоднее предсказывать А, чем Б**" означает, что **суммарное значение функции потерь будет ниже**, если модель предскажет набор ответов А, по сравнению с ситуацией, в которой она для тех же объектов предскажет набор ответов Б.

P.S. В этом задании Вам может пригодиться Python, чтобы было удобнее сравнивать значения BCE для разных вариантов предсказаний.

Выберите один или несколько ответов:

- ☒ а. С точки зрения BCE модели выгоднее предсказывать для этих объектов $[0.99, 0.45]$, чем $[0.99, 0.01]$
- ☐ б. С точки зрения BCE модели выгоднее предсказывать для этих объектов $[0.99, 0.01]$, чем $[0.99, 0.45]$
- ☐ в. С точки зрения BCE модели выгоднее предсказывать для этих объектов $[0.65, 0.65]$, чем $[0.99, 0.45]$
- ☒ г. С точки зрения BCE модели выгоднее предсказывать для этих объектов $[0.99, 0.45]$, чем $[0.65, 0.65]$
- ☒ д. С точки зрения BCE модели выгоднее предсказывать для этих объектов $[0.5, 0.5]$, чем $[0.99, 0.01]$
- ☐ е. С точки зрения BCE модели выгоднее предсказывать для этих объектов $[0.99, 0.01]$, чем $[0.5, 0.5]$

Ваш ответ верный.

Вопрос **2**

Выполнен

Баллов: 1,00 из 1,00

Оцените количество настраиваемых параметров в логистической регрессии для 5 классов и 10 признаков.

Не забудьте про bias!

Выберите один ответ:

- ☐ a. 50
- ☐ b. 5
- ☒ c. 55
- ☐ d. 500
- ☐ e. 10

Ваш ответ верный.

[◀ 2.1 Векторная модель текста и TF-IDF](#)

Перейти на...

[2.3 Задачи по теме: Векторная модель текста ▶](#)