Тест 1 по курсу «Байесовский выбор моделей»

Время выполнения: 40 минут Максимальный балл: 45 баллов

Вариант 1

Задача 1 (15 баллов). а) Какое семейство распределений называется экспоненциальным? Что называется достаточными статистиками относительно его параметров? (4 балла)

- б) Получить представление правдоподобия НОР выборки из $\mathcal{N}(m, \sigma^2)$ в форме распределения из экспоненциального семейства с вектором параметров $\boldsymbol{\theta}$. Выписать достаточные статистики. (6 баллов)
- в) Пользуясь свойствами нормировочной константы экспоненциального семейства распределений, получить $\mathbb{D}\xi^2$ для случайной величины $\xi \sim \mathcal{N}(m, \sigma^2)$ (5 баллов).

Задача 2 (10 баллов). Вывести формулу прогноза целевой переменной на тестовой выборке $p(\mathbf{y}_{\text{test}}|\mathbf{X}_{\text{test}},\ \mathbf{X}_{\text{train}},\ \mathbf{y}_{\text{train}})$ по известному признаковому описанию \mathbf{X}_{test} и НОР обучающей выборке ($\mathbf{X}_{\text{train}},\ \mathbf{y}_{\text{train}}$) для вероятностной модели с совместным правдоподобием

$$p(\mathbf{y}, \mathbf{w}|\mathbf{X}) = p(\mathbf{w})p(\mathbf{y}|\mathbf{X}, \mathbf{w}).$$

Какие свойства совместного правдоподобия модели использованы при выводе результата?

Задача 3 (20 баллов). Пусть имеется две двухсторонние монеты, случайно и независимо выбранные из всех существующих монет достоинством в 2 рубля. Пусть было произведено $n_1=10$ бросаний первой монеты и $n_2=10000$ бросаний второй. Среди $n_1=10$ результатов бросания первой монеты было $k_1=2$ орла, а среди $n_2=10000$ бросаний второй – $k_2=5100$ орлов.

- а) Построить вероятностную модель эксперимента, записав правдоподобие и введя априорные распределения на вероятности p_1 и p_2 выпадания орлов для первой и второй монеты соответственно. Опишите, как и из каких соображений Вы выбрали априорные распределения $q(p_1)$ и $q(p_2).(4$ балла)
- б) Получить апостериорные распределения $q(p_1|k_1, n_1)$ и $q(p_2|k_2, n_2)$. (4 балла)
- в) Пусть теперь рассматривается две вероятностные модели: M_1 с $p_1=p_2=p$ и априорным распределением, которые было ранее выбрано Вами для p_1 и полная модель M_2 из пункта а), где p_1 и p_2 априорно выбраны независимо из $q(p_1)$ и $q(p_2)$. Сосчитать апостериорную вероятность обеих моделей, считая их априори равновероятными $(p(M_1)=p(M_2)=0.5)$. Какой вывод можно сделать из результата? (12 баллов)

Время выполнения: 40 минут Максимальный балл: 45 баллов

Вариант 2

Задача 1 (10 баллов). а) При каких условиях априорное распределение называется сопряженным к правдоподобию вероятностной модели? (3 балла)

б) Получить сопряженное априорное распределение на вероятность выпадения орла p у двухсторонней монеты для вероятностной модели n независимых ее бросаний (7 баллов).

Задача 2 (15 баллов). а) Что такое обоснованность вероятностной модели? Как она связана с устойчивостью качества аппроксимации обучающей выборки при малых вариациях параметров? (4 балла)

- б) Вывести формулу прогноза целевой переменной на тестовой выборке $p(\mathbf{y}_{\text{test}}|\mathbf{X}_{\text{test}},\mathbf{X}_{\text{train}},\mathbf{y}_{\text{train}})$ по известному признаковому описанию \mathbf{X}_{test} и НОР обучающей выборке $(\mathbf{X}_{\text{train}},\mathbf{y}_{\text{train}})$ для вероятностной смеси K моделей с известными априорными вероятностями $\mathbb{P}(M_i),\ i=\overline{1,K}$. При каком условии прогноз смеси моделей совпадает в точности с прогнозом одной из моделей в смеси? (6 баллов)
- в) Получить формулу апостериорной вероятности каждой модели из смеси $\mathbb{P}(M_i|\mathbf{X}_{\text{train}},\mathbf{y}_{\text{train}})$. Что такое принцип максимума обоснованности, и при каких условиях его использование оправданно? (5 баллов)

Задача 3 (20 баллов). Пусть имеется модель линейной регрессии с нормальным шумом

$$\mathbf{y} = \mathbf{X}\mathbf{w} + \boldsymbol{\varepsilon}, \ \boldsymbol{\varepsilon} \sim \mathcal{N}(0, \ \sigma^2 \mathbf{I}),$$

где σ^2 – известно, и априорным распределение на \mathbf{w} $p(\mathbf{w}) = \mathcal{N}(\mathbf{w}|\mathbf{m}, \operatorname{diag}(\mathbf{s}))$, где \mathbf{m} и $\operatorname{diag}(\mathbf{s})$ неизвестные гиперпараметры.

- а) Выписать совместное правдоподобие $p(\mathbf{y}, \mathbf{w}|\mathbf{X}, \mathbf{m}, \mathbf{s})$, задающее вероятностную модель. (4 балла)
- б) Получить апостериорное распределение на вектор ${\bf w}$, предполагая ${\bf m}$ и ${\bf s}$ известными. Что происходит, если $s_i=0$? (6 баллов)
- в) Решить задачу максимизации обоснованности

$$p(\mathbf{y}|\mathbf{X}, \mathbf{m}, \mathbf{s}) = \int p(\mathbf{y}|\mathbf{X}, \mathbf{w}) p(\mathbf{w}|\mathbf{m}, \operatorname{diag}(\mathbf{s})) d\mathbf{w}$$

по гиперпараметрам \mathbf{m} и \mathbf{s} . Какой вывод можно сделать из полученного результата? (10 баллов)