Фамилия:..

Имя:...

Группа:..

Задача №1

Вы прогнозируете покупку клиентом VIP подписки в зависимости от времени, проводимого им еженедельно в вашем приложении, а также от расходов на покупки внутри приложения.

Индекс	1	2	3	4	5	6	7	8	9
Время	20	30	41	21	31	32	18	45	50
Расходы	400								
Покупка	1	1	0	0	1	1	1	0	1
Удовлетворенность	5	8	16	10	24	32	40	48	25
Часть	1			2			3		
Выборка	Обучающая			Тестовая					

Для прогнозирования вы применяете метод ближайших соседей с 2 соседями и расстоянием Минковского с параметром $\lambda = 1$. Также, осуществляется стандартизация признаков. Прогнозируется 1, если оценка условной вероятности не меньше 0.6.

С каждой покупки VIP подписки фирма зарабатывает 10 рублей. Если фирма ожидает, что клиент не совершит покупку, то она предоставляет ему скидку в размере 2 рубля. Фирма предполагает, что в таком случае тот, кто хотел совершить покупку, также совершит ее, а тот, кто не хотел – совершит с вероятностью 0.5.

1. Известно, что независимо от того, используется для прогнозирования покупки лишь переменная времени, либо сразу обе переменные (время и расходы), прогнозы оказываются одинаковыми. Также, имеются не менее двух индивидов с различающимися расходами. Используя данную информацию заполните в таблице выше пропущенные значения переменной на расходы. **(5 баллов)**

2. Посчитайте точность (accuracy) прогнозов на обучающей выборке. **(5 баллов)**

3. Рассчитайте прибыль прогнозов на тестовой выборке. **(10 баллов)**

4. Вычислите полноту (recall) с использованием 3-х частной кросс-валидации. **(10 баллов)**

5. Используя взвешивание на обратные условные вероятности (IPW) оцените, на всей выборке, средний эффект воздействия покупки VIP подписки на удовлетворенность клиента. При этом условные вероятности, оцененные как 0 и 1, заменяются на 0.2 и 0.8 соответственно. **(10 баллов)**

Решение

Для удобства аггрегируем результаты промежуточных расчетов в таблице:

Индекс	1	2	3	4	5	6	7	8	9
Время	20	30	41	21	31	32	18	45	50
Расходы	400	600	820	420	620	640	360	900	1000
Покупка	1	1	0	0	1	1	1	0	1
$\hat{\mathbb{P}}$ (обучающая)	1	1	0.5	1	0.5	0.5	1	0.5	0.5
$\widehat{\text{Покупк}}$ (обучающая)	1	1	0	1	0	0	1	0	0
Тип (обучающая)	TP	TP	TN	FP	FN	FN	TP	TN	FN
$\hat{\mathbb{P}}$ (CV)	0.5	1	0.5	1	0.5	0.5	0.5	0.5	0.5
$\widehat{\text{Покупк}}$ (CV)	0	1	0	1	0	0	0	0	0
Тип (CV)	FN	TP	TN	FP	FN	FN	FN	TN	FN
$\hat{\mathbb{P}}$ (вся)	0.5	0.8	0.2	0.5	0.8	0.8	0.8	0.2	0.5
Часть	1			2			3		
Выборка	Обучающая			Тестовая					

1. Задача имеет несколько решений. Однако, проще всего заметить, что прогнозы окажутся идентичными в случае, если векторы расходов и времени будут коллинеарны, то есть $\text{Расходы}_i = 20\text{Время}_i$.

2. Точность на обучающей выборке составила $\text{ACC} = 1$.

3. Согласно условию цены прогнозов являются следующими:

$$
\begin{gathered}
\text{P}_{\text{TP}} = 10 \\
\text{P}_{\text{TN}} = (10 - 2) \times 0.5 = 4 \\
\text{P}_{\text{FN}} = 10 - 2 = 8 \\
\text{P}_{\text{FP}} = 0
\end{gathered}
$$

   В результате получаем прибыль:

$$\pi = 1 \times 10 + 1 \times 4 + 3 \times 8 + 1 \times 0 = 38$$

4. Рассчитаем полноту для каждой из частей:

$$\text{recall}_1 = \frac{1}{1+1} = 0.5 \qquad \text{recall}_2 = \frac{0}{0+1} = 0 \qquad \text{recall}_3 = \frac{0}{0+1} = 0$$

   Вычислим полноту по кросс-валидации:

$$\text{recall}_{\text{CV}} = \frac{0.5 + 0 + 0}{3} = \frac{1}{6}$$

5. Рассчитаем необходимую оценку:

$$\widehat{\text{ATE}} = \frac{1}{9}\left(\frac{5}{0.5} + \frac{8}{0.8} - \frac{16}{1-0.2} - \frac{10}{1-0.5} + \frac{24}{0.8} + \frac{32}{0.8} + \frac{40}{0.8} - \frac{48}{1-0.2} + \frac{25}{0.5}\right) = 10$$

Задача №2

Вы оцениваете эффект воздействия показа рекламы индивиду на факт покупки им лодки.

Индекс	1	2	3	4	5	6	7	8	9	10
Город$_i$	1	0	0	1	0	0	1	0	1	0
Водоем$_i$	1	1	1	0	1	1	0	0	1	1
Реклама$_i$	1	0	1	0	1	0	1	0	1	0
Лодка$_i$	1	1	1	1	1	0	0	0	0	0

В данных применяются следующие обозначения:

- **Город** – бинарная контрольная переменная, принимающая значение 1, если индивид живет в городе и 0 – в противном случае.
- **Водоем** – бинарная контрольная переменная, принимающая значение 1, если индивид проживает достаточно близко к реке или озеру и 0 – в противном случае.
- **Реклама** – переменная воздействия, принимающая значение 1, если вы показали рекламу индивиду и 0 – в противном случае.
- **Лодка** – бинарная целевая переменная, принимающая значение 1, если индивид купил лодку и 0 – в противном случае.

1. С помощью T-learner оцените условный средний эффект воздействия рекламы на покупку товара горожанином, проживающим рядом с рекой. Для оценивания условных математических ожиданий используйте наивный Байесовский классификатор **(20 баллов)**.
2. Нарисуйте такую Байесовскую сеть (включающую все 4 переменные), использование которой для расчета условных математических ожиданий (вместо наивного Байесовсокого классификатора) в предыдущем пункте с заменой T-learner на S-learner даст точно такую же (как в предыдущем пункте) оценку условного среднего эффекта воздействия. Ответ подробно обоснуйте с точки зрения используемых для расчетов факторов **(15 баллов)**.

Подсказка: у бернуллиевских случайных величин математическое ожидание и вероятность принять значение 1 – совпадают.

Решение:

1. Сперва оценим значения необходимых для итоговых расчетов факторов:

$$\hat{P}(\text{Л}_i = 1|\text{Р}_i = 1) = \frac{3}{5} \qquad \hat{P}_0(\text{Л}_i = 1|\text{Р}_i = 0) = \frac{2}{5}$$

$$\hat{P}(\text{Г}_i = 1|\text{Л}_i = 1, \text{Р}_i = 1) = \frac{1}{3} \qquad \hat{P}(\text{Г}_i = 1|\text{Л}_i = 1, \text{Р}_i = 0) = \frac{1}{2}$$

$$\hat{P}(\text{Г}_i = 1|\text{Л}_i = 0, \text{Р}_i = 1) = 1 \qquad \hat{P}(\text{Г}_i = 1|\text{Л}_i = 0, \text{Р}_i = 0) = 0$$

$$\hat{P}(\text{В}_i = 1|\text{Л}_i = 1, \text{Р}_i = 1) = 1 \qquad \hat{P}(\text{В}_i = 1|\text{Л}_i = 1, \text{Р}_i = 0) = \frac{1}{2}$$

$$\hat{P}(\text{В}_i = 1|\text{Л}_i = 0, \text{Р}_i = 1) = \frac{1}{2} \qquad \hat{P}(\text{В}_i = 1|\text{Л}_i = 0, \text{Р}_i = 0) = \frac{2}{3}$$

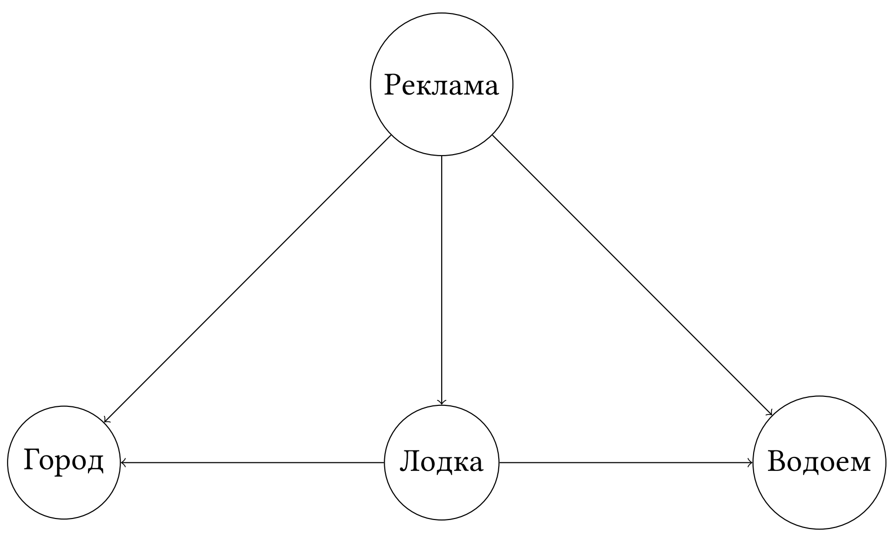
Рассчитаем условные математические ожидания:

$$\hat{\mathrm{E}}\left(\text{Л}_i|\Gamma_i=1,\text{В}_i=1,\text{Р}_i=1\right)=\hat{\mathrm{P}}\left(\text{Л}_i=1|\Gamma_i=1,\text{В}_i=1,\text{Р}_i=1\right)=$$
$$=\frac{\hat{P}\left(\text{Л}_i=1|\text{Р}_i=1\right)\hat{P}(\Gamma_i=1|\text{Л}_i=1,\text{Р}_i=1)\hat{P}(\text{В}_i=1|\text{Л}_i=1,\text{Р}_i=1)}{...+\hat{P}\left(\text{Л}_i=0|\text{Р}_i=1\right)\hat{P}(\Gamma_i=1|\text{Л}_i=0,\text{Р}_i=1)\hat{P}(\text{В}_i=1|\text{Л}_i=0,\text{Р}_i=1)}=$$
$$=\frac{\frac{3}{5}\times\frac{1}{3}\times 1}{\frac{3}{5}\times\frac{1}{3}\times 1+\frac{2}{5}\times\frac{1}{2}\times 1}=\frac{1}{2}$$
$$\hat{\mathrm{E}}\left(\text{Л}_i|\Gamma_i=1,\text{В}_i=1,\text{Р}_i=0\right)=\hat{\mathrm{P}}\left(\text{Л}_i=1|\Gamma_i=1,\text{В}_i=1,\text{Р}_i=0\right)=$$
$$=\frac{\hat{P}\left(\text{Л}_i=1|\text{Р}_i=0\right)\hat{P}(\Gamma_i=1|\text{Л}_i=1,\text{Р}_i=0)\hat{P}(\text{В}_i=1|\text{Л}_i=1,\text{Р}_i=0)}{...+\hat{P}\left(\text{Л}_i=0|\text{Р}_i=0\right)\hat{P}(\Gamma_i=1|\text{Л}_i=0,\text{Р}_i=0)\hat{P}(\text{В}_i=1|\text{Л}_i=0,\text{Р}_i=0)}=$$
$$=\frac{\frac{2}{5}\times\frac{1}{2}\times\frac{1}{2}}{\frac{2}{5}\times\frac{1}{2}\times\frac{1}{2}+\frac{3}{5}\times 0\times\frac{2}{3}}=1$$

Оценим условный средий эффект воздействия:

$$\widehat{\text{CATE}}=\hat{\mathrm{E}}\left(\text{Л}_i|\Gamma_i=1,\text{В}_i=1,\text{Р}_i=1\right)-\hat{\mathrm{E}}\left(\text{Л}_i|\Gamma_i=1,\text{В}_i=1,\text{Р}_i=0\right)=$$
$$=\frac{1}{2}-1=-\frac{1}{2}$$

2. Исходя из факторизации, описанной в решении предыдущего пункта, следует, что для получения аналогичных результатов достаточно рассмотреть следующую Байесовскую сеть:



Задача №3

Имеется нейросеть, включающая всего $n = 1$ наблюдение по 2 признакам: $x_1 = 1$ (первый признак) и $x_2 = 2$ (второй признак). Значение целевой переменной равняется $y = 2240$. Имеется 2 скрытых слоя, в первом из которых содержится 2 нейрона, а во втором находится 3 нейрона. В качестве функции активации в скрытом и выходном слоях используется ReLU. Применяется квадратичная функция потерь. В нейросети нет смещений (констант) и все ее параметры равняются 5.

1. Изобразите графически описанную нейросеть. **(5 баллов)**
2. Рассчитайте значение функции потерь данной нейросети при заданных значениях весов. **(10 баллов)**
3. Для обучения нейронной сети используется градиентный спуск со скоростью обучения 0.001. Найдите, чему после одной итерации данного алгоритма будет равен вес, с которым второй признак входит в первый нейрон первого скрытого слоя. **(10 баллов)**

Решение

1. Очевидно.
2. Последовательно рассчитаем значения в различных нейронах, а затем найдем величину функции потерь:

$$
\begin{gathered}
h_{11} = h_{12} = \max(0, 5 \times 1 + 5 \times 2) = 15 \\
h_{21} = h_{22} = h_{23} = \max(0, 5 \times 15 + 5 \times 15) = 150 \\
o = \max(0, 5 \times 150 + 5 \times 150 + 5 \times 150) = 2250 \\
l = (2250 - 2240)^2 = 100
\end{gathered}
$$

3. Используя метод обратного распространения ошибки получаем значение производной:

$$
\begin{aligned}
\frac{\partial l}{\partial \omega_{12}^{(1)}} &= \frac{\partial l}{\partial o}\frac{\partial o}{\partial q}\left(\frac{\partial q}{\partial h_{21}}\frac{\partial h_{21}}{\partial s_{21}}\frac{\partial s_{22}}{\partial h_{11}} + \frac{\partial q}{\partial h_{22}}\frac{\partial h_{22}}{\partial s_{22}}\frac{\partial s_{21}}{\partial h_{11}} + \frac{\partial q}{\partial h_{23}}\frac{\partial h_{21}}{\partial s_{23}}\frac{\partial s_{23}}{\partial h_{11}}\right)\frac{\partial h_{11}}{\partial s_{11}}\frac{\partial s_{11}}{\partial \omega_{12}^{(1)}} = \\
&= \frac{\partial l}{\partial o} \times 1 \times \left(\omega_{11}^{(3)} \times 1 \times \omega_{11}^{(2)} + \omega_{21}^{(3)} \times 1 \times \omega_{12}^{(2)} + \omega_{31}^{(3)} \times 1 \times \omega_{13}^{(2)}\right) \times 1 \times x_2 = \\
&= 2 \times (2250 - 2240) \times 1 \times (5 \times 1 \times 5 + 5 \times 1 \times 5 + 5 \times 1 \times 5) \times 1 \times 2 = 3000
\end{aligned}
$$

Применяя градиентный спуск с заданной скоростью обучения рассчитаем обновленный вес:

$$
\omega_{11,\text{new}}^{(1)} = \omega_{11,\text{old}}^{(1)} - \alpha \times \frac{\partial l}{\partial \omega_{12}^{(1)}} = 5 - 0.001 \times 3000 = 2
$$