

- امتحان شامل ۲ صفحه و ۴ پرسش است.
- استفاده از مداد مشکی اشکالی ندارد. از رنگ قرمز استفاده نکنید.
- قوانین آزمون رعایت شوند.
- به برنامه‌های نوشته شده دقیقاً به همان صورتی که هستند پاسخ دهید و حتی یک کما هم اضافه یا کم نکنید.
- شماره خط‌های ذکر شده در کنار برخی برنامه‌ها، جزو برنامه‌ها نیستند و صرفاً برای آن است که شما بتوانید در صورت نیاز به خط موردنظرتان در پاسخنامه اشاره کنید.

پرسش	۱	۲	۳	۴	جمع نمرات
بارم	۳۰	۳۰	۳۰	۳۰	۱۲۰
نمره					

۱. (۳۰ نمره) اگر  $\frac{\partial \mathbf{y}}{\partial \mathbf{x}}$  به صورت زیر تعریف شده باشد:

$$\frac{\partial \mathbf{y}}{\partial \mathbf{x}} = \begin{bmatrix} \frac{\partial y_1}{\partial x_1} & \frac{\partial y_1}{\partial x_2} & \dots & \frac{\partial y_1}{\partial x_n} \\ \frac{\partial y_2}{\partial x_1} & \frac{\partial y_2}{\partial x_2} & \dots & \frac{\partial y_2}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial y_m}{\partial x_1} & \frac{\partial y_m}{\partial x_2} & \dots & \frac{\partial y_m}{\partial x_n} \end{bmatrix}$$

where  $\mathbf{y}$  is an  $m \times 1$  vector and  $\mathbf{x}$  is an  $n \times 1$  vector

و  $\mathbf{y} = \mathbf{Ax}$  نشان دهید:  $\frac{\partial \mathbf{y}}{\partial \mathbf{x}} = \mathbf{A}$ . به فرض  $\mathbf{A}$  یک ماتریس  $m \times n$  است که وابسته به  $\mathbf{x}$  نیست و عناصر آن به صورت  $a_{ij}$  هستند.

۲. (۳۰ نمره) جدول زیر را کامل کرده، Accuracy، Binary cross entropy loss را محاسبه کنید.

actual	Predicted probabilities	Corrected probabilities	log
0	0.2		
0	0.7		
1	0.3		
1	0.8		

The formula of binary cross entropy loss for a single data point is:

$$L(y, \hat{y}) = -[y \log \hat{y} + (1 - y) \log(1 - \hat{y})]$$

where  $y$  is the true label (0 or 1), and  $\hat{y}$  is the predicted probability (between 0 and 1). The formula of binary cross entropy loss for a batch of data points is the average of the individual losses:

$$L(Y, \hat{Y}) = -\frac{1}{N} \sum_{i=1}^N [y_i \log \hat{y}_i + (1 - y_i) \log(1 - \hat{y}_i)]$$

$$\log(0.2) = -1.6, \log(0.3) = -1.2, \log(0.7) = -0.35, \log(0.8) = -.22$$

۳. (۳۰ نمره) کار برنامه زیر چیست؟ برای هر بخش آن یک توضیح مختصر بنویسید. (اگر در پاسخنامه پاسخ می‌دهید، فقط شماره خط برنامه ذکر شود، کفایت می‌کند).

```

1 import tensorflow as tf
2 import random
3
4 A = tf.Variable(random.random())
5 B = tf.Variable(random.random())
6 C = tf.Variable(random.random())
7 D = tf.Variable(random.random())
8
9 EPOCHS = 1000
10 optimizer = tf.keras.optimizers.Nadam(lr=0.1)
11 for _ in range(EPOCHS):
12     with tf.GradientTape() as tape:
13         y1 = A + B - 9
14         y2 = C - D - 1
15         y3 = A + C - 8
16         y4 = B - D - 2
17         sqerr = y1*y1 + y2*y2 + y3*y3 + y4*y4
18         gradA, gradB, gradC, gradD = tape.gradient(sqerr, [A,B,C,D])
19         optimizer.apply_gradients([(gradA,A), (gradB,B), (gradC,C),
                                     (gradD,D)])

```

۴. (۳۰ نمره) تابع هزینه در رگرسیون به صورت رابطه‌ی (۱) تعریف شده بود. در الگوریتم گرادیان کاهشی پارامترهای مدل مبتنی بر مشتق تابع هزینه بر روزرسانی شدند. مدلی که در کلاس و برنامه‌ی رگرسیون دیدیم به صورت  $\hat{y} = b + wx$  بود. در این سؤال مشتق تابع هزینه بر اساس پارامترهای مجهول را بر اساس مدل  $\hat{y} = b + wx^2$  به دست آورید. آیا الان هم رگرسیون خطی انجام می‌شود؟

$$Loss = MSE = \frac{1}{n} \sum_{i=1}^n error_i^2 = \frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i)^2 \quad (1)$$