



دانشگاه صنعتی اصفهان  
دانشکده مهندسی برق و کامپیوتر

درس مبانی هوش محاسباتی

تکلیف تئوری اول

تاریخ تحویل: ۱۴ فروردین

## سؤال ۱ (۶ نمره)

رگرسیون خطی<sup>۱</sup> و رگرسیون لجستیک<sup>۲</sup> را در هر مورد مقایسه کنید.

- نتیجه (خروجی)<sup>۳</sup>
- ارتباط بین متغیرها<sup>۴</sup>
- خطا<sup>۵</sup>
- روش‌های تخمین<sup>۶</sup> (برآورد)

## سؤال ۲ (۱۰ نمره)

چرا نمی‌توانیم از تابع هزینه<sup>۷</sup> میانگین مربعات خطا<sup>۸</sup> استفاده شده در رگرسیون خطی<sup>۱</sup> برای رگرسیون لجستیک<sup>۲</sup> استفاده کنیم؟

## سؤال ۳ (۱۰ نمره)

با استفاده از کمی جبر ثابت کنید که دو معادله زیر معادل هستند. به عبارت دیگر، نمایش تابع لجستیک<sup>۹</sup> و نمایش لاجیت<sup>۱۰</sup> برای مدل رگرسیون لجستیک<sup>۲</sup> معادل هستند.

$$p(X) = \frac{e^{\beta_0 + \beta_1 X}}{1 + e^{\beta_0 + \beta_1 X}}$$

$$\log_e \frac{p(X)}{1 - p(X)} = \ln \frac{p(X)}{1 - p(X)} = \beta_0 + \beta_1 X$$

## سؤال ۴ (۱۰ نمره)

چگونه با استفاده از رگرسیون لجستیک<sup>۲</sup> با مسئله طبقه‌بندی چند کلاسه<sup>۱۱</sup> برخورد خواهید کرد؟

---

<sup>1</sup>Linear regression

<sup>2</sup>Logistic regression

<sup>3</sup>Outcome

<sup>4</sup>Relationship

<sup>5</sup>Error

<sup>6</sup>Estimation Methods

<sup>7</sup>Cost Function

<sup>8</sup>Mean Square Error

<sup>9</sup>logistic

<sup>10</sup>logit

<sup>11</sup>multiclass classification problem

## سؤال ۵ (۶ نمره)

فرض کنید یک طبقه‌بندی کننده رگرسیون لجستیک<sup>۱</sup> را آموزش می‌دهید و تابع فرضیه شما  $H$  است:

$$h(\theta) = g(\theta_0 + \theta_1 x_1 + \theta_2 x_2)$$

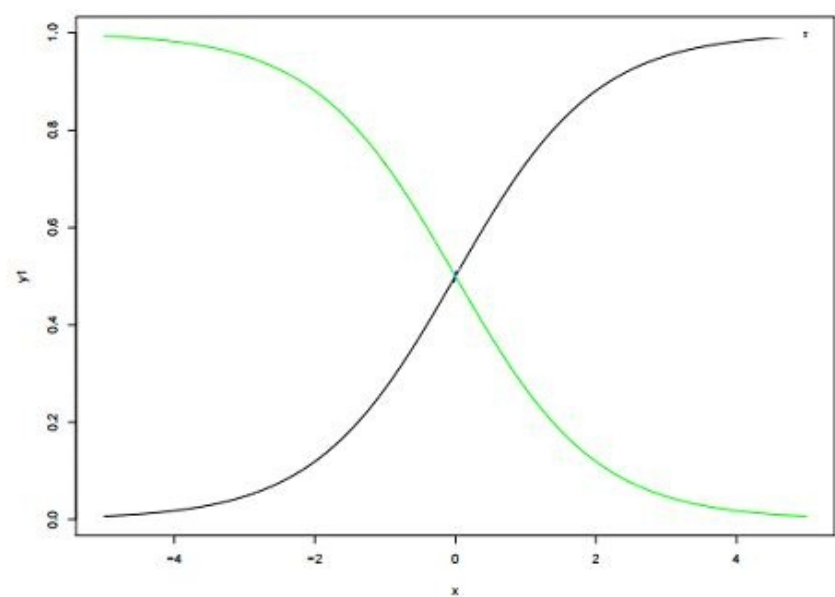
در هر مورد مرز تصمیم را رسم کنید.

۱.  $\theta_2 = 0, \theta_1 = -3, \theta_0 = 6$

۲.  $\theta_2 = 6, \theta_1 = 0, \theta_0 = 18$

## سؤال ۶ (۸ نمره)

در شکل زیر دو مدل لجستیک<sup>۱</sup> مختلف با مقادیر متفاوت برای  $\beta_0$  و  $\beta_1$  آورده شده است:



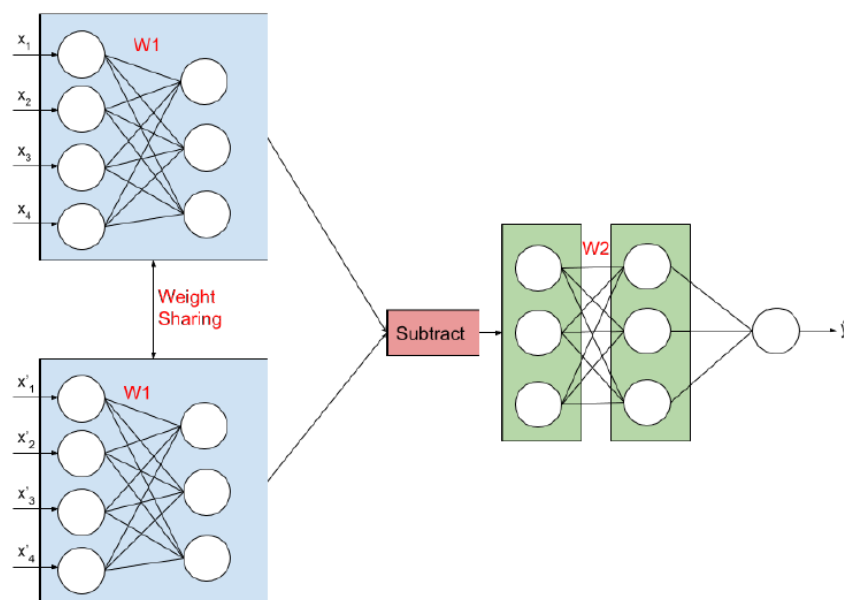
شکل ۱: منحنی‌های دو مدل رگرسیون لجستیک<sup>۱</sup> متفاوت

در مورد مقادیر ضرایب  $(\beta_0, \beta_1)$  برای هر یک از نمودارها چه می‌توان گفت؟ ( $Y = \beta_0 + \beta_1 X$ )

<sup>1</sup> Logistic regression

## سؤال ۷ (۵۰ نمره)

شبکه عصبی زیر را در نظر بگیرید. این شبکه، شبکه عصبی Siamese نام دارد که شامل یک شبکه دو قلو است که ورودی‌های مجزایی دارند اما وزن‌ها در این دو شبکه مشترک هستند. خروجی این شبکه دوقلو در یک یا چند لایه به هم متصل می‌شوند. فرض کنید یک شبکه عصبی Siamese به صورت زیر داریم:



شکل ۲: شبکه عصبی Siamese

روابط زیر در این شبکه عصبی برقرار است: (توابع فعالسازی نسبت به توابع فعالسازی خود شبکه عصبی اصلی تغییر یافته‌اند).

$$z_1 = W_1 x^{(i)} + b_1$$

$$a_1 = \tanh(z_1)$$

$$z_2 = W_1 x'^{(i)} + b_1$$

$$a_2 = \tanh(z_2)$$

$$a = a_1 - a_2$$

$$z_3 = W_2 a + b_2$$

$$\hat{y}^{(i)} = \sigma(z_3)$$

$$L^{(i)} = y^{(i)} \log(\hat{y}^{(i)}) + (1 - y^{(i)}) \log(1 - \hat{y}^{(i)})$$

$$J = -\frac{1}{m} \sum_{i=1}^m L^{(i)}$$

توجه کنید که  $x^{(i)}$  و  $x'^{(i)}$  ورودی‌های شبکه‌ی دوقلو هستند که ابعاد هر کدام از آن‌ها  $1 \times D_x$  است. همچنین  $y^{(i)}$  خروجی مربوط به جفت ورودی  $i$ ام بوده و یک عدد می‌باشد. در مجموعه داده مورد نظر،  $m$  داده وجود دارد. تعداد گره‌ها در لایه پنهان اول،  $D_{a_1}$  بوده بنابراین ابعاد  $z_1$  و  $z_2$ ،  $1 \times D_{a_1}$  است. توجه کنید که در دو لایه اول وزن‌ها مشترک هستند. با توجه به اطلاعات داده شده، به سؤالات زیر پاسخ دهید:

الف) ابعاد پارامترهای  $W_1$ ،  $b_1$ ،  $W_2$  و  $b_2$  به چه صورت است؟ اگر از تکنیک برداری‌سازی<sup>۱</sup> برای مجموعه داده استفاده کنیم، ابعاد  $X$  و  $Y$  به چه شکل می‌شود؟

ب) مشتقات زیر را حساب کنید:

$$\bullet \partial J / \partial z_3$$

$$\bullet \partial a / \partial z_2$$

$$\bullet \partial J / \partial W_1$$

ج) فرمول‌های به روزرسانی پارامترهای  $W_1$ ،  $b_1$ ،  $W_2$  و  $b_2$  با روش gradient descent و با نرخ یادگیری  $\alpha$  را بنویسید. توجه کنید که مشتقات باید محاسبه شوند. می‌توانید از مقادیر مشتق بدست آمده در قسمت ((ب)) استفاده نمایید.

د) برای دو نورون در هر بخش و مقادیر مشخص شده، شبکه داده شده را به روش SGD یک mini batch آموزش دهید و مقادیر جدید وزن‌ها را بدست آورید. (برای ورودی‌های  $x'$  در مورد عملکرد شبکه تحقیق کنید.)

$$x_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, x_2 = \begin{bmatrix} 2 \\ 2 \end{bmatrix}, y = 1$$

$$W_1 = \begin{bmatrix} 1 & 1 \end{bmatrix}, b_1 = \begin{bmatrix} 0 \end{bmatrix}$$

$$W_2 = \begin{bmatrix} 1 \end{bmatrix}, b_2 = \begin{bmatrix} 0 \end{bmatrix}$$

$$\alpha = 1$$

---

<sup>1</sup>vectorization

## نکات تکمیلی

۱. لزومی به تایپ کردن سوالات تئوری نیست؛ ولی در صورتیکه پاسخ آنها به صورت تایپ شده تحویل داده شود، ۱۰ درصد نمره اضافه به شما تعلق میگیرد. در صورتیکه پاسخهای شما تایپ شده نیست، باید پاسخها خوانا و باکیفیت در قالب فایل pdf ارسال شوند.
۲. فرمت نامگذاری تکلیف ارسالی باید به صورت زیر باشد: HWX\_Theory\_LastName\_StudentID که X شماره تکلیف LastName نام خانوادگی شما و StudentID شماره دانشجویی شما است.
۳. انجام این تکلیف به صورت تک نفره است. در صورت مشاهده تقلب، نمرات هم مبدا کپی و هم مقصد آن صفر لحاظ می شود.
۴. برای تکالیف تئوری امکان ارسال با تاخیر وجود ندارد.
۵. در صورت وجود هر گونه ابهام و یا سوال می توانید سوالات خود را در گروه تلگرام بپرسید. هم چنین می توانید برای رفع ابهامات با دستیاران آموزشی از طریق تلگرام در تماس باشید.  
آیدی ها:

@Fatemeh\_1241

@alirez anum1

@alireza20010226