

תרגיל 5 - יותם גרדוש - 208541334

שאלות הקדמה

1. מה נכון לומר על מקורות הסטוכסטיות (אקראיות) בעיות למידת חיזוק כפי שהגדנו בכיתה? תיתכן יותר מتسובה אחת נcona

(א) הקלט למערכת חייב להיות סטוכסטי על מנת שהמערכת תוכל ללמידה משזו

(ב) הפלט של המערכת חייב להיות פונקציה דטרמיניסטיבית של הקלט

(ג) פונקציית הגמול, בהנתן קלט ופלט, יכולה להיות דטרמיניסטיבית

(ד) הפלט של המערכת יכול להיות סטוכסטי, אבל זה לא מומלץ כי הלמידה תהיה איטית יותר

(ה) בהנתן קלט, פלט וגורם מסוים - חשוב שהייה מרכיב סטוכסטי בכלל הלמידה של הפרמטרים של המערכת

בנוסף לכך, מטרת ה-MLP היא לearn את היחסים בין ה- $x$  ו- $y$ , כלומר, לearn את היחסים בין ה- $x$  ו- $\hat{y}$ . מטרת ה-MLP היא לearn את היחסים בין ה- $x$  ו- $\hat{y}$ .

ל) רכון - נקי גביה כ' כל הדרים נחה ביאר יוניסטרו נקי רם

הנורווגי יונסן מילר, מנהל תיאטרון גשר, מציין כי "במשך שנים רבות לא נתקל במחזה שזכה לתשומת כל כך הרבה מבקרים ומבקרים".

2. נניח כי מערכת הלומדות בעזרת מידת חיזוק קיבלה קלט, והוצאה פלט הסטברוטי בהינתן הקלט ופרמטר פנימי  $s$ , ולאחר מכן קיבלה גמול שלילי. מה מהבאים אפשר?

תיקון יותר מושבה אחת נכונה. עבור התשובות האפשריות, כתבו מדוע הן מסודרות עם התוצאה הרצiosa.

- (א) ערך הזכאות  $e$  הוא חיובי, וכך כל הלמידה עבור  $s$  יהיה חיובי
- (ב) ערך הזכאות  $e$  הוא חיובי, וכך כל הלמידה עבור  $s$  יהיה שלילי
- (ג) ערך הזכאות  $e$  הוא שלילי, וכך כל הלמידה עבור  $s$  יהיה חיובי
- (ד) ערך הזכאות  $e$  הוא שלילי, וכך כל הלמידה עבור  $s$  יהיה שלילי

$$\Delta w_i \leftarrow \eta e_i \text{sgn}(w_i - b)$$

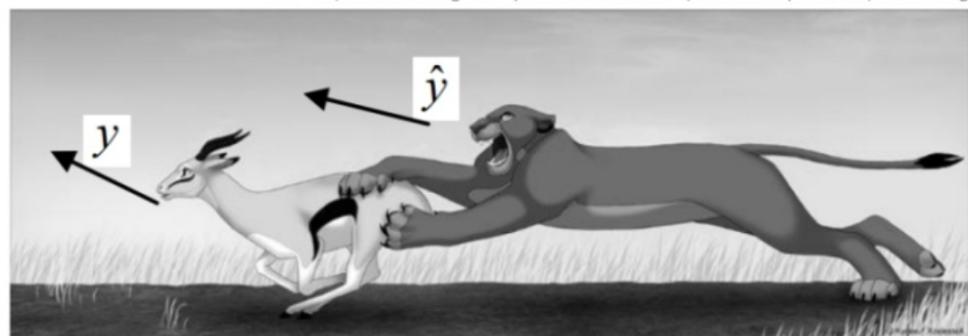
$$\Delta w_i \leftarrow \eta e_i \text{sgn}(w_i - b) \quad \text{אם } e_i > 0 \quad \text{ולא}$$

$$\Delta w_i \leftarrow \eta e_i \text{sgn}(w_i - b) \quad \text{אם } e_i < 0 \quad \text{ולא}$$

$$\Delta w_i \leftarrow \eta e_i \text{sgn}(w_i - b) \quad \text{אם } e_i = 0 \quad \text{ולא}$$

## שאלה 1

ע' ואט מיקום הטרף המשוערך במוחו של הטרוף ב-<sup>ב</sup> (שניהם סקלרים). נסמן את המיקום האמתי של הטרוף ב-



- עבור כל  $trial$ , הטרוף מקבל גמול  $r$ , שערךו:
 
$$r = -(\hat{y} - y)^2$$
  - נניח שהפרמטרים  $\mu$  ו- $\sigma^2$  נלמדים ע"י הטרוף באמצעות אלגוריתם *REINFORCE*.
  - השערוך  $\hat{y}$ , המיצר במוח הטרוף, הוא משתנה מקרי נורמלי המופיע על ידי שני פרמטרים  $\sigma$ ,  $\mu$ , ומתפלג כך:
 
$$\hat{y} \sim \mathcal{N}(\mu, \sigma^2)$$
  - מיקום הטרוף,  $y$ , הוא משתנה מקרי כלשהו עם ממוצע  $m$  ושונות  $s^2$ . המשתנים  $m$  ו- $s^2$  הם קבועים.

1. בכיתה, כאשר פרטנו את הנושא של לימודי חיזוק, הגדנו כי ישנו: מערכת לומדת, פרטורים פנימיים, קלט, פלט וגמול. הסבירו בקצרה מהם כל אחד מן הרכיבים האלה בשאלת הזאת (אם קיימים, ראיינו בתרגול שלא בכלל קיימים כל הרכיבים), ואיזה סוגים של תלות ישנים בעיה (אילו רכיבים תלויים באילו רכיבים?)

1. כבז' תר' הכהנים הילג זאג הילג הילג:

הו מושג בפערת ה- $\lambda$ , גודל ה- $\lambda$  מוגדר כ- $\lambda = \frac{1}{\tau}$ , פערת ה- $\lambda$  מוגדרת כ- $\Delta\lambda = \lambda - \lambda_0$ .

הנתקה מהתפקידים הדרושים בהסמכה והמוניטין בהסמכה.

g)  $\sin x$  ist auf  $[0, \pi]$  konkav, r. K. f.  $\sin x$  ist j.a. -  $\sin x$  -

ב- trial מסויים, מודיע הכללים שקיבלו את הגיוניות?)

$$\ell_i = \frac{\partial}{\partial w_i} \ln(P(y|x,w)) \quad \Delta w_i = \eta \ell_i (y, x, w) \cdot r(y, x) : \text{new rule} \quad \text{repeat until } \Delta w_i < 10^{-6}$$

$$Y \sim N(\mu, \sigma^2) \quad | \sim$$

$$\ln(P(\hat{y}|x, \omega)) \Rightarrow \ln(P(\hat{y}|\mu, \sigma^2)) = \ln\left(\frac{1}{\sqrt{2\pi\sigma^2}} \cdot e^{-\frac{(\hat{y}-\mu)^2}{2\sigma^2}}\right) = \ln\left(\frac{1}{\sqrt{2\pi\sigma^2}}\right) + \ln\left(e^{-\frac{(\hat{y}-\mu)^2}{2\sigma^2}}\right)$$

$\downarrow$   
 $\ln$   
 $\uparrow$

$$= -\ln\left(\frac{1}{\sqrt{2\pi\sigma^2}}\right) - \frac{(\hat{y}-\mu)^2}{2\sigma^2}$$

• M Fe 11298 55 > fm > 0.75

$$\Delta\mu = p_e \cdot r = \eta \cdot r \cdot \frac{\partial}{\partial \mu} \left( -\ln \left( \sqrt{2\pi\sigma^2} \right) - \frac{(\hat{y} - \mu)^2}{2\sigma^2} \right) = \eta \cdot r \cdot 0 + \frac{\cancel{\eta}(\hat{y} - \mu) \cdot \cancel{(-1)}}{\cancel{\sigma^2}} = \eta \cdot r \cdot \left( \frac{\hat{y} - \mu}{\sigma^2} \right)$$

: 9 1128 12 2011

$$\Delta \sigma = \eta' r \cdot \frac{\partial}{\partial \sigma} \left( -\ln(\sqrt{2\pi\sigma^2}) - \frac{(\hat{y}-\mu)^2}{2\sigma^2} \right) = \eta r \cdot \frac{\partial}{\partial \sigma} \left( -\frac{\sqrt{2}}{\sigma\sqrt{\pi}} \cdot (\hat{y}-\mu) + \frac{(\hat{y}-\mu)^2}{2\sigma^3} \right) = \eta r \cdot \left( \frac{-1}{\sigma} + \frac{(\hat{y}-\mu)^2}{\sigma^3} \right)$$

$\ln(p(\tilde{y}|x, \theta))$  is a log-likelihood function for  $\theta$  given  $x$  and  $y$ .

נול מילר  $r=0$  גורם  $\Delta M < 0$  למספר  $|M|$  גורם  $r \neq 0$  גורם  $\Delta M > 0$  למספר  $|M|$ .

$$\left( \frac{(y - \mu)^2}{r^3} \right) = \left( \frac{(y - \mu)^2}{r^3} - \frac{1}{r} \right) + \frac{1}{r}$$

$e_i < 0$  נסמן  $\downarrow$  ו-  $e_i > 0$  נסמן  $\uparrow$ . מכאן ש- $\sum e_i = \sum \downarrow - \sum \uparrow$  ו- $\sum e_i < 0$  מציין ש- $\sum \downarrow > \sum \uparrow$ .

3. חשבו את השינוי הממוצע בפרמטרים של הטורף (מומוצע על גבי כל מקורות הסטטיסטיות בעיה!!)

$w_i = \gamma \frac{\partial}{\partial w_i} E(\gamma)$  կույտագրություն է կամ լինելու վեհականությունը է.

$$\Rightarrow E(r) = E(-(g - \hat{y})^2) = -1 \cdot E(g^2 - 2g\hat{y} + \hat{y}^2) = -E(\hat{y}^2) + 2E(g\hat{y}) - E(g^2)$$

$$A: E(\hat{y}^2) = E(y^2) - E(\hat{y})^2 + E(\hat{y})^2 = \text{Var}(\hat{y}) + E(\hat{y})^2 = \sigma^2 + \mu^2$$

↗ Erwartungswert  
 ↗ Varianz  
 ↗ Erwartungswert  
 ↗ Varianz  
 ↗ Varianz  
 ↗ Erwartungswert  
 ↗ Varianz

$$\textcircled{B}: E(y^2) = E(y^2) - E(y)^2 + E(y)^2 = \text{Var}(y) + E(y)^2 = S^2 + m^2$$

$$\Leftrightarrow \mathbb{E}(\hat{y}|y) = \mathbb{E}(y)\mathbb{E}(\hat{y}) = m \cdot \mu \Rightarrow \mathbb{E}(r) = -\mu^2 - \sigma^2 + fm\mu - s^2 - m^2$$

לעתה נסמן  $\mathcal{P}(N)$  כ $\mathcal{P}$ , ו $\mathcal{P}^*$  כ $\mathcal{P}^N$ .

$$E(\Delta\mu) = \eta \frac{\partial}{\partial\mu} E(I) = \eta \frac{\partial}{\partial\mu} (-\mu^2 - \sigma^2 + Jm\mu - S^2 - m^2) = \eta (Jm - J\mu) = J\eta(m - \mu)$$

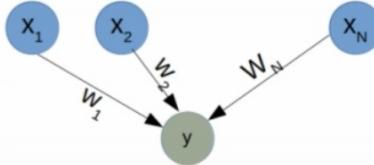
$$E(\Delta r) = \eta \frac{\partial}{\partial r} E(r) = \eta \frac{\partial}{\partial r} (-\mu^2 - \sigma^2 + f m \mu - S^2 - m^2) = \eta \cdot (-2r)$$

4. לאיזה ערך הפרמטרים מתכנסים בפועל? הסבירו בקצרה את ההיגיון בתשובה שקיבלתם.

שאלה 2

חלק אנליטי

נתון פרספטורון ביןארי כמתואר בצייר, שבו הפלט הוא הסתברותי.



בhinintn kllt  $\bar{x} = [x_1 \dots x_N]^\top$ , hplt y matpulg baofen ha'av:

$$\mathbb{P}[y=1|\bar{x}; \bar{w}] = \frac{1}{1+e^{-\bar{w}^\top \bar{x}}}$$

$$\mathbb{P}[y=0|\bar{x}; \bar{w}] = \frac{e^{-\bar{w}^\top \bar{x}}}{1+e^{-\bar{w}^\top \bar{x}}}$$

1. חשבו את 'ערך הזכאות'  $e_i$  ביחס לפרמטר  $w_i$ , כלומר את  $\frac{\partial \log P(y|\bar{x};\bar{w})}{\partial w_i}$ , עבור שני המקרים האפשריים:  $y = 0$  ו-  $y = 1$ .

$$\text{. פ'יר } y \text{ 'אכ' } 0 \text{ ו } 1 \text{ און } e_i \text{ ו } w_i \text{ .}$$

$$y=1$$

$$\ell_i = \frac{\partial}{\partial w_i} \ln(P(y=1|x, w)) = \frac{\partial}{\partial w_i} \ln\left(\frac{1}{1 + e^{-w^T x}}\right) = \frac{\partial}{\partial w_i} \left( \ln(1) - \ln(1 + e^{w^T x}) \right)$$

$$= \frac{\partial}{\partial w_i} \left( -\ln(1 + e^{-w^T x}) \right) = + \left( \frac{e^{w^T x} \cdot x_i}{1 + e^{-w^T x}} \right) = x_i \cdot \frac{e^{w^T x}}{1 + e^{w^T x}} = x_i P(y=1|x, w)$$

$$\begin{aligned}
 \text{J}(\theta) &= -\frac{1}{n} \sum_{i=1}^n \left[ y_i \ln(P(Y=1|X_i, w)) + (1-y_i) \ln(P(Y=0|X_i, w)) \right] \\
 &= -\frac{1}{n} \sum_{i=1}^n \left[ y_i \ln \left( \frac{e^{w^T x_i}}{1+e^{w^T x_i}} \right) + (1-y_i) \ln \left( \frac{1}{1+e^{w^T x_i}} \right) \right] \\
 &= -\frac{1}{n} \sum_{i=1}^n \left[ y_i \ln \left( \frac{e^{w^T x_i}}{1+e^{w^T x_i}} \right) + (1-y_i) \ln \left( \frac{1}{1+e^{w^T x_i}} \right) \right] \\
 &= -\frac{1}{n} \sum_{i=1}^n \left[ y_i \ln \left( \frac{e^{w^T x_i}}{1+e^{w^T x_i}} \right) + (1-y_i) \ln \left( \frac{1}{1+e^{w^T x_i}} \right) \right]
 \end{aligned}$$

2. רשמו ביטוי אחד שմבטא את  $e_i$  עבור שני המקרים.

$$\begin{aligned} P(y=0|x,w) &= 1 - P(y=1|x,w) \quad : \text{이 } \sum_{y \in \{0,1\}} P(y=1|x,w) \text{ 는 } y \text{의 확률 } \text{이므로 } P(y=0|x,w) \text{ 를 } 1 \text{에서 뺀 값} \\ y - P &= P(y=0|x,w) \Leftrightarrow y = 1 : \text{인 경우 } \rightarrow \text{이 } P \text{ 를 } 1-P(y=1|x,w) \text{ } \mid \text{인 경우} \\ y - P &= -P(y=1|x,w) \Leftrightarrow y = 0 \end{aligned}$$

$e_i = x_i(y - P)$  : the rank "def" of  $e_i$  as a linear map

3. בהינתן שב trial מסוים הנמול היה  $R$ , רשמו את כלל העדכון של הסינפסה  $w_i$ . האם כלל העדכון הוא לوكלי? הסבירו.

$$\Delta w_i = \eta \cdot r \cdot e_i \quad r = R \text{ ו } e_i = \int_{\Omega} \Delta u_j \cdot \delta_{ij} dx$$

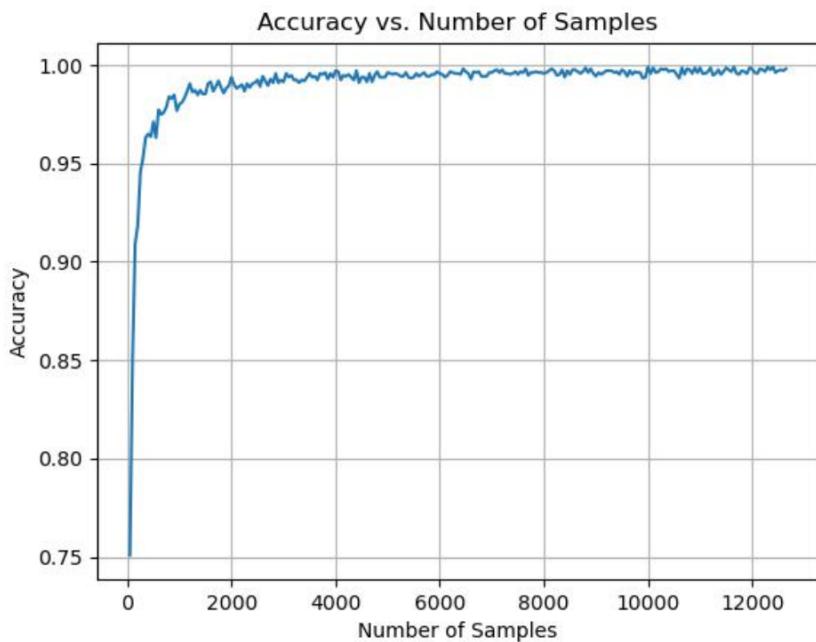
$$\Delta w_i = \eta R x_i (y - P)$$

הסבירו הראתך אם בהשאלה ההשאלה היא נכונה!  
 $\int_{\Omega} \Delta u_j \cdot \delta_{ij} dx$  איזה מושג שפירושו  $(w_j \text{ s.t. } j \neq i) \rightarrow \text{השאלה}$ ?

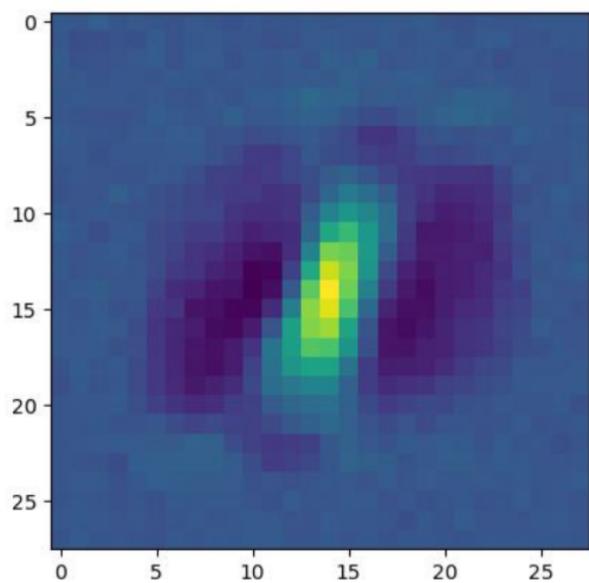
### חלק תכניות

ב חלק זו תממשו פרספטורון Bayesian סטטיסטי (כפי שהוגדר ב חלק הקודם) שילמד לסוג תמונות של הספרות 0 ו 1 הכתובות בכתב יד (מתוך הדאטאסט שנקרא MNIST).

3



לישר  $\rightarrow$  כ- 1000 ניסויים נזקק כדי לאריך כ- 100 שיטות נורמאליזציה ופונקציות נזקק כדי לאריך כ- 100 שיטות נורמאליזציה.



.4

## חישוביות וקוגניציה - תרגיל 5

להגשה עד: 29/02/2024

שימו לב: בתחילת התרגיל מופיעה כמה שאלות הקדמה אמריקאיות. יש להגיש את התשובות אליהם, עם משפט נימוק קצר לכל שאלה. לאחר מכן, שאלה 1 היא שאלה אנליטית ו שאלה 2 משלבת חלק אנליטי וחלק תכוני.

### שאלות הקדמה

1. מה נכון לומר על מקורות הסטטיסטיות (אקראיות) בעיות למידת חיזוק כפי שהגדנו בכיתה? תיתכן יותר מتسובה אחת נכון

- (א) הקלט למערכת חייב להיות סטטיסטי על מנת שהמערכת תוכל ללמוד משוחו
- (ב) הפלט של המערכת חייב להיות פונקציית דטרמיניסטי של הקלט
- (ג) פונקציית הגמול, בהנתן קלט ופלט, יכולה להיות דטרמיניסטי
- (ד) הפלט של המערכת יכול להיות סטטיסטי, אבל זה לא מומלץ כי הלמידה תהיה איטית יותר
- (ה) בהנתן קלט, פלט וGamal מסויימים - חשוב שיהיה מרכיב סטטיסטי בכל הלמידה של הפרמטרים של המערכת

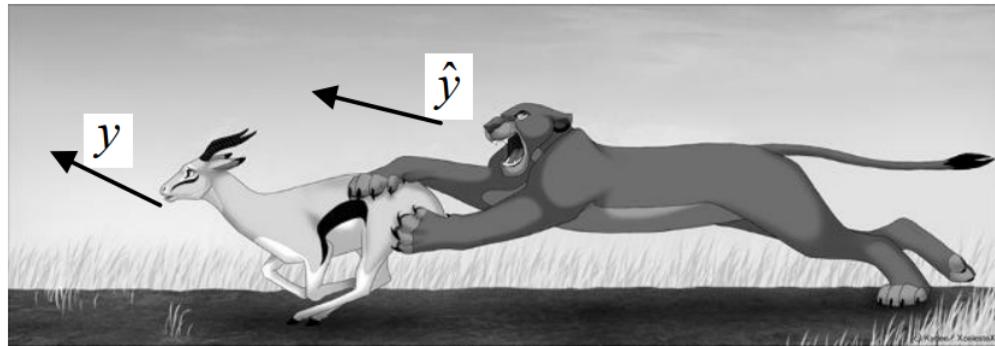
2. נניח כי מערכת הלומדת באמצעות למידת חיזוק קיבלה קלט, הוצאה פלט הסתברותי בהינתן הקלט ופרמטר פנימי  $w$ , ולאחר מכן קיבלה גמול שלילי. מה מהבאים אפשרי? תיתכן יותר מتسובה אחת נכון. עברו התשובות האפשריות, כתבו מודע הן מסתדרות עם התוצאה הרצiosa.

- (א) ערך הזכאות  $e$  הוא חיובי, וכך כל הלמידה עבור  $w$  יהיה חיובי
- (ב) ערך הזכאות  $e$  הוא חיובי, וכך כל הלמידה עבור  $w$  יהיה שלילי
- (ג) ערך הזכאות  $e$  הוא שלילי, וכך כל הלמידה עבור  $w$  יהיה חיובי
- (ד) ערך הזכאות  $e$  הוא שלילי, וכך כל הלמידה עבור  $w$  יהיה שלילי

זכרו לנמק את שתי השאלות.

## שאלה 1

נדון בחישוב שמתבצע במוחו של טורף המנסה לחזות את מיקום הטורף שלו. נסמן את המיקום האמתי של הטורף ב- $y$  ואת מיקום הטורף המשוערך במוחו של הטורף ב- $\hat{y}$  (שניים סקלרים).

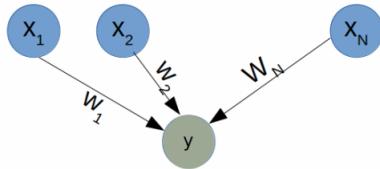


- מיקום הטורף,  $y$ , הוא משתנה מקרי כלשהו עם ממוצע  $m$  ושונות  $s^2$ . המשתנים  $m$  ו- $s^2$  הם קבועים.
  - השערוך  $\hat{y}$ , המיצר במוח הטורף, הוא משתנה מקרי נורמלי המופיע על ידי שני פרמטרים  $\mu$ ,  $\sigma$ , ומתפלג כך:  $\hat{y} \sim \mathcal{N}(\mu, \sigma^2)$ .
  - נניח שהפרמטרים  $\mu$  ו- $\sigma$  נלמדים ע"י הטורף באמצעות אלגוריתם *REINFORCE*.
  - עבור כל  $trial$ , הטורף מקבל גמול  $r$ , שערךו:  $r = -(y - \hat{y})^2$
1. בכיתה, כאשר פרלנו את הנושא של למידת חיזוק, הגדרנו כי ישנו: מערכת למדת, פרמטרים פנימיים, קלט, פלט וגמול. הסבירו בקורס מהם כל אחד מן הרכיבים האלה בשאלת הזאת (אם קיימים, ראיינו בתרגול שלא בכל בעיה קיימים כל הרכיבים), ואיזה סוגים של תלות ישנים בבעיה (אילו רכיבים תלויים באילו רכיבים?)
2. רשמו את כללי העדכון המפורטים של שני הפרמטרים המאפיינים את הטורף,  $\mu$  ו- $\sigma$ . בתשובהכם בסעיף זה הניחו כי הגמול הנקחי הינו  $r$  וקבעו העדכון הינו  $\hat{\eta}$  (כלומר בסעיף זה אין צורך להציב את הפונקציה של  $r$ ). הסבירו את התוצאות שקיבלתם ואיך הן מסתדרות עם למידה ע"י פרסים ועונשים (כלומר - בהנתן  $r$  כלשהו ב- $trial$  מסוים, מדוע הכללים שקיבלתם הגיוניים?)
3. חשבו את השינוי המוצע בפרמטרים של הטורף (ממוצע על גבי כל מקורות הסטוכסטיות בעיה!)
4. לאיזה ערך הפרמטרים מותכנים בממוצע? הסבירו בקורס את ההיגיון בתשובה שקיבלתם.

## שאלה 2

### חלק אנליטי

נתון פרס派טרון בינהרι כמתואר בציור, שבו הפלט הוא הסתברותי.



בהתנאי קלט  $\bar{x}$ , הפלט  $y$  מתפלג באופן הבא:

$$\begin{aligned}\mathbb{P}[y = 1|\bar{x}; \bar{w}] &= \frac{1}{1 + e^{-\bar{w}^\top \bar{x}}} \\ \mathbb{P}[y = 0|\bar{x}; \bar{w}] &= \frac{e^{-\bar{w}^\top \bar{x}}}{1 + e^{-\bar{w}^\top \bar{x}}}\end{aligned}$$

1. חשבו את 'ערך האזאות'  $e_i$  ביחס לפרמטר  $w_i$ , כלומר את  $\frac{\partial \log \mathbb{P}(y|\bar{x}; \bar{w})}{\partial w_i}$ , עבור שני המקרים האפשריים:  $y = 0$  ו- $y = 1$ .
2. רשמו ביטוי אחד שمبטא את  $e_i$  עבור שני המקרים.
3. בהינתן שב- trial מסוים הגמול היה  $R$ , רשמו את כלל העדכון של הסינפסה  $w_i$ . האם כלל העדכון הוא לוקלי?סבירו.

## חלק תבניות

ב חלק זו ת认真学习ו פרגספטורון בינהרִי סטטיסטי (כפי שהוגדר ב חלק הקודם) שילמד לסwoג תמונות של הספרות 0 ו- 1 הכתובות בכתב יד (מתוך הדאטאסט שנקרא MNIST).

ל שאלה מצורפים קבצי הנתונים הבאים: מוגם אימון (data, labels) ודוגמ בדיקה (data, labels). הנטונים מצורפים כארבעה קבצי csv או כקובץ מטלאב אחד הכלול בתוכו את הנתונים.

- למדו את הפסטטורון דוגמא מוגם האימון באופן הבא: הציגו לרשות את הדוגמא הנוכחית, והגרילו את הפלט לפי התפלגות המתאימה (שהוגדרה ב חלק האנליטי). נסמן ב-  $y$  את הפלט של הרשות וב-  $c$  את התיאוג הנכון (0 או 1). הגמול יוגדר באופן הבא:

$$R = \begin{cases} 1 & y = c \\ 0 & y \neq c \end{cases}$$

עדכנו את המשקولات לפי כלל הלמידה של REINFORCE (שמצאותם בשאלה 2) לאחר הצגת כל דוגמא. לאחר כל 50 צעדים (כלומר, הצגה של 50 דוגמאות), בדקו את הדיוק של הרשות על מוגם ה-  $\text{test}$  (ambilי למורי מהדוגמאות הללו): הציגו לרשות את כל הדוגמאות מה-  $\text{test}$ , לכל דוגמא הגרילו את הפלט של הרשות והשו אותו לתיאוג הרצוי. חשבו את הדיוק על פני כל הדוגמאות ב-  $\text{test}$  (כלומר, אחוז התמונות שהרשota מתויגת נכון).

3. הציגו גרפ שמתאר את התפתחות הדיוק על מוגם ה-  $\text{test}$  כפונקציה של הלמידה (מספר הדוגמאות שהרשota נכונה עד כה).

4. הציגו את וקטור המשקولات של הרשות שהתקבל לאחר הלמידה בתור תמונה (הפכו את הווקטור באורך 784 לתמונה בגודל  $28 \times 28$ ). הסבירו מדוע התמונה נראה כפי שהיא נראה.

הערות:

- השתמשו בקצב לימוד של  $0.01 = \eta$ .
- אתחלו את וקטור המשקولات באקראי, למשל מהתפלגות נורמלית עם ממוצע אפס וסטיית תקן  $\sigma = 0.01$ .
- מומלץ מאד לכתוב פונקציה עבור סעיף 3, כלומר פונקציה שמקבלת את וקטור המשקولات הנוכחיים ואוסף דוגמאות מתייגנות, ומחזירה את הדיוק הממוצע על פני הדוגמאות שהתקבלו (ambilי לשנות את המשקولات ברשות).
- מימדים:
  - שימו לב שבפייטון, יתכן שהמערכיים של התיאוגים ייטנו בתור מטריצת numpy מגודל  $.1 \times \#examples$  מומלץ להפוך אותם לוקטורים עי' ( $-1$ ),  $\text{labels} = \text{labels.reshape}(-1, \text{labels})$ , או שימוש בפונקציה squeeze.
  - כאשר מציגים את התמונות, שימו לב שבמטלאב יש לבצע reshape של כל דוגמא (או של וקטור המשקولات) למימדים  $28 \times 28$ .