

67822 - 78102 גן נון נס ציון

$$: 1 \frac{\sqrt{2x+1}}{2}$$

נתנו:

$$208489732$$

$$\sqrt{3k^3}$$

לפנינו.

$$314997388$$

$$\sqrt{p}$$

לפנינו.

נמצא $\sqrt{g(x)}$ $\sqrt{f(g(x))}$ $\sqrt{f(x)}$ $\sqrt{f(g(x))} = \sqrt{f(Bx)} = \sqrt{Bx}$ $\sqrt{f(g(x))} = \sqrt{f(Ax)} = \sqrt{Ax}$ $\sqrt{f(g(x))} = \sqrt{f(x)} = \sqrt{x}$

$$\forall x \quad f(x) = Ax, \quad g(x) = Bx \quad \text{ונר.}$$

$$- \text{נайдו } A, B \quad \text{ונר.}$$

$$\forall x \quad f(g(x)) = f(Bx) = ABx$$

$$- \text{נайдו } C = AB \quad \text{ונר.}$$

$$f \circ g(x) = Cx \Rightarrow \text{נайдו } f, g \quad \text{ונר.}$$

$$\sqrt{f(g(x))} = \sqrt{Cx} = \sqrt{C}$$

$f(x) = Ax + a$ $\sqrt{f(x)} = \sqrt{Ax + a}$ $\sqrt{f(x)} = \sqrt{Ax} + \sqrt{a}$ $\sqrt{f(x)} = \sqrt{Ax} + \sqrt{a}$

$$\forall x \quad f(x) = Ax + a, \quad g(x) = Bx + b$$

$$- \text{נайдו } a, b \quad \text{ונר.}$$

$$\forall x \quad f(g(x)) = f(Bx + b) = A(Bx + b) + a =$$

$$= ABx + Ab + a = Cx + C'$$

$$f \circ g \quad \text{נайдו } C = AB, \quad C' = Ab + a \quad \text{ונר.}$$

$$\sqrt{f(g(x))} = \sqrt{Cx + C'} = \sqrt{C} \sqrt{x + \frac{C'}{C}}$$

.10 (2)

$$\theta^{n+1} = \theta^n - \alpha \nabla F_{\theta^n}(x)$$

$$\theta^{n+1} = \theta^n \quad \text{אם } \nabla F_{\theta^n}(x) = 0$$

אם לא, (θ^n הוא נקודתstationary) נסמן $\nabla F_{\theta^n}(x) \neq 0$ (הארה)

$-\alpha \nabla F_{\theta^n}(x) \neq 0$ ($\alpha > 0$)

$$\alpha \nabla F_{\theta^n}(x) = 0 \Rightarrow \nabla F_{\theta^{n+1}}(x) = 0$$

לעתה נוכיח θ^{n+1} מינימום מקומי של $F_{\theta^n}(x)$.

הארה. נוכיח כי $F_{\theta^{n+1}}(x) \leq F_{\theta^n}(x)$ (אך לא מינימום).

נוכיח כי $F_{\theta^{n+1}}(x) \leq F_{\theta^n}(x)$.

ונוכיח כי $F_{\theta^{n+1}}(x) < F_{\theta^n}(x)$.

ונוכיח כי $F_{\theta^{n+1}}(x) < F_{\theta^n}(x)$.

ונוכיח כי $F_{\theta^{n+1}}(x) < F_{\theta^n}(x)$.

הארה. נוכיח כי $F_{\theta^{n+1}}(x) \leq F_{\theta^n}(x)$.

נוכיח כי $F_{\theta^{n+1}}(x) \leq F_{\theta^n}(x)$.

$$f(x + dx) = f(x) + \nabla f(x)dx + dx^T H(x)dx + O(\|dx\|^3)$$

נוכיח כי $\nabla f(x) = 0$.

$$f(x + dx) - f(x) = dx^T H(x)dx$$

נוכיח כי $f(x + dx) - f(x) \leq 0$.

$f(x + dx) - f(x) \leq -\lambda \|dx\|^2$.

$-\lambda \|dx\|^2 \leq 0$.

$$dx^T H(x)dx = \sum_{i,j} \frac{\partial^2 f}{\partial x_i \partial x_j} dx_i dx_j$$

לעתה נוכיח כי $\frac{\partial^2 f}{\partial x_i \partial x_j} \geq 0$.

$$y = \text{הערך הימני} \quad \text{ריבוע} \quad \cdot 10^3 \quad (3)$$

$$p = \text{הערך המינימלי}$$

loss loss loss loss

def loss(y, p):

$$d = \text{torch. abs}(y - p)$$

$$d[d > 180] = 360 + / \quad 360 \quad \text{הערך}$$

return torch.square(d)

הסבר.

השאלה שאלת הדרישה שוקrk שלם
השאלה שוקrk שלם שוקrk שלם שוקrk שלם
השאלה שוקrk שלם שוקrk שלם שוקrk שלם
השאלה שוקrk שלם שוקrk שלם שוקrk שלם

$$\frac{\partial}{\partial x} f(x+y, 2x, z) = \left(\frac{\partial(x+y)}{\partial x}, \frac{\partial(2x)}{\partial x}, \left(\frac{\partial(z)}{\partial x} \right) \right) = (1, 2, 0) \quad (4)$$

$$\frac{\partial}{\partial x} f_1(f_2(\dots f_n(x))) = \text{הערך של}$$

$$= \frac{\partial f_1}{\partial x} \cdot f_2(\dots f_n(x)) \cdot \frac{\partial}{\partial x} f_2(f_3(\dots f_n(x))) =$$

:

$$= \frac{\partial f_1}{\partial x} \cdot f_2(\dots f_n(x)) \cdot \dots \cdot \frac{\partial f_n}{\partial x}.$$

$$\frac{\partial}{\partial x} f_1(x, f_2(x, f_3(\dots f_{n-1}(x, f_n(x)))) = \text{הערך של}$$

$$= \left(\frac{\partial(x)}{\partial x}, \frac{\partial}{\partial x} f_2(x, f_3(\dots f_{n-1}(x, f_n(x)))) \right) =$$

$$= (1, (1, \dots, (1, \frac{\partial f_1}{\partial x} f_2(\dots f_n(x)) \dots \frac{\partial f_n}{\partial x})))$$

$$\frac{\partial}{\partial x} f(x + g(x + h(x))) =$$

$$= \text{def } f'(x + g(x + h(x)) \cdot (g' + g'(x + h(x)) \cdot (1 + h'(x)))$$

- גוף הוכיח . $D_{KL} \geq 0$ כ. (5)

$$D_{KL}(P||Q) = \sum_{i=1}^n p_i \log\left(\frac{y_i}{p_i}\right)$$

- גוף הוכיח סכום של כפונקציית שיפוט

$$\sum_{i=1}^n a_i \log\left(\frac{a_i}{b_i}\right) \geq \left(\sum_{i=1}^n a_i\right) \log\left(\frac{\sum_{i=1}^n a_i}{\sum_{i=1}^n b_i}\right)$$

לוג. כ.

$$\sum_{i=1}^n p_i \log\left(\frac{y_i}{p_i}\right) \geq \left(\sum_{i=1}^n p_i\right) \cdot \log\left(\frac{\sum_{i=1}^n y_i}{\sum_{i=1}^n p_i}\right)$$

$$\sum_{i=1}^n p_i = \sum_{i=1}^n y_i = 1$$

$$D_{KL}(P||Q) \geq \sum_{i=1}^n p_i \cdot \log\left(\frac{1}{1}\right) = 0$$

$$D_{KL}(P||Q) = 0 \iff P = Q$$

$$D_{KL}(P||Q) = \sum_{i=1}^n p_i \log\left(\frac{y_i}{p_i}\right) = \sum_{i=1}^n p_i \cdot \log(1) = 0$$

$$p_i = y_i \quad D_{KL}(P||Q) = 0 \quad \text{ולכן}$$

$$\sum_{i=1}^n p_i \log\left(\frac{y_i}{p_i}\right) = 0 \Rightarrow \log\left(\frac{y_i}{p_i}\right) = 0 \Rightarrow y_i = p_i \quad \text{ולכן}$$

ר' מון נסיך דק' 6.

- גודל גודל

$$D_{KL}(\lambda p_1 + (1-\lambda) p_2 \parallel \lambda q_1 + (1-\lambda) q_2) =$$

$$= \sum_{x \in X} (\lambda p_1(x) + (1-\lambda) p_2(x)) \cdot \log \left(\frac{\lambda p_1(x) + (1-\lambda) p_2(x)}{\lambda q_1(x) + (1-\lambda) q_2(x)} \right)$$

$$\leq \sum_{x \in X} \left(\lambda p_1(x) \cdot \log \left(\frac{\lambda p_1(x)}{\lambda q_1(x)} \right) + (1-\lambda) p_2(x) \log \left(\frac{(1-\lambda) p_2(x)}{(1-\lambda) q_2(x)} \right) \right)$$

$$= \lambda \sum_{x \in X} p_1(x) \cdot \log \left(\frac{p_1(x)}{q_1(x)} \right) + (1-\lambda) \sum_{x \in X} p_2(x) \cdot \log \left(\frac{p_2(x)}{q_2(x)} \right) =$$

$$= \lambda D_{KL}(p_1 \parallel q_1) + (1-\lambda) D_{KL}(p_2 \parallel q_2)$$

7

ReLU פונקציית מילוט גראDED בינה מילוט (7)

הצגה, $C[0,1] \rightarrow$ מילוט מילוט

$$\forall x \in \mathbb{R} \quad f(x) = \text{ReLU}(ax+b)$$

הצגה, a, b קבועים

וניגן, f גראDED מילוט גראDED גראDED כורן (7)

$$g(x) = \text{ReLU}(ax+b) - \text{ReLU}(ax+b-1)$$

$$\lim_{x \rightarrow -\infty} g(x) = 0 \quad , \quad \lim_{x \rightarrow \infty} g(x) = 1$$

ר' מון

בנוסף נשים \Rightarrow מילוט גראDED גראDED גראDED גראDED (7)

$O(\omega)$ ו α גורם שה μ מינימלי נתקה (8)

$O(\omega)$ ו α גורם שפונקיה פולינומית (לפחות).

$$f(x) = \sum_{i=1}^n \alpha_i \sigma(\omega_i x + b_i) \quad \underline{\alpha_i \in \mathbb{R}}$$

בנוסף, $(x \rightarrow \infty)$ הוליך ל- L - (8)

אם $\alpha_i > 0$ כי σ גורמת לכך α_i מוגברת. אם $\alpha_i \leq 0$ מוגברת α_i .

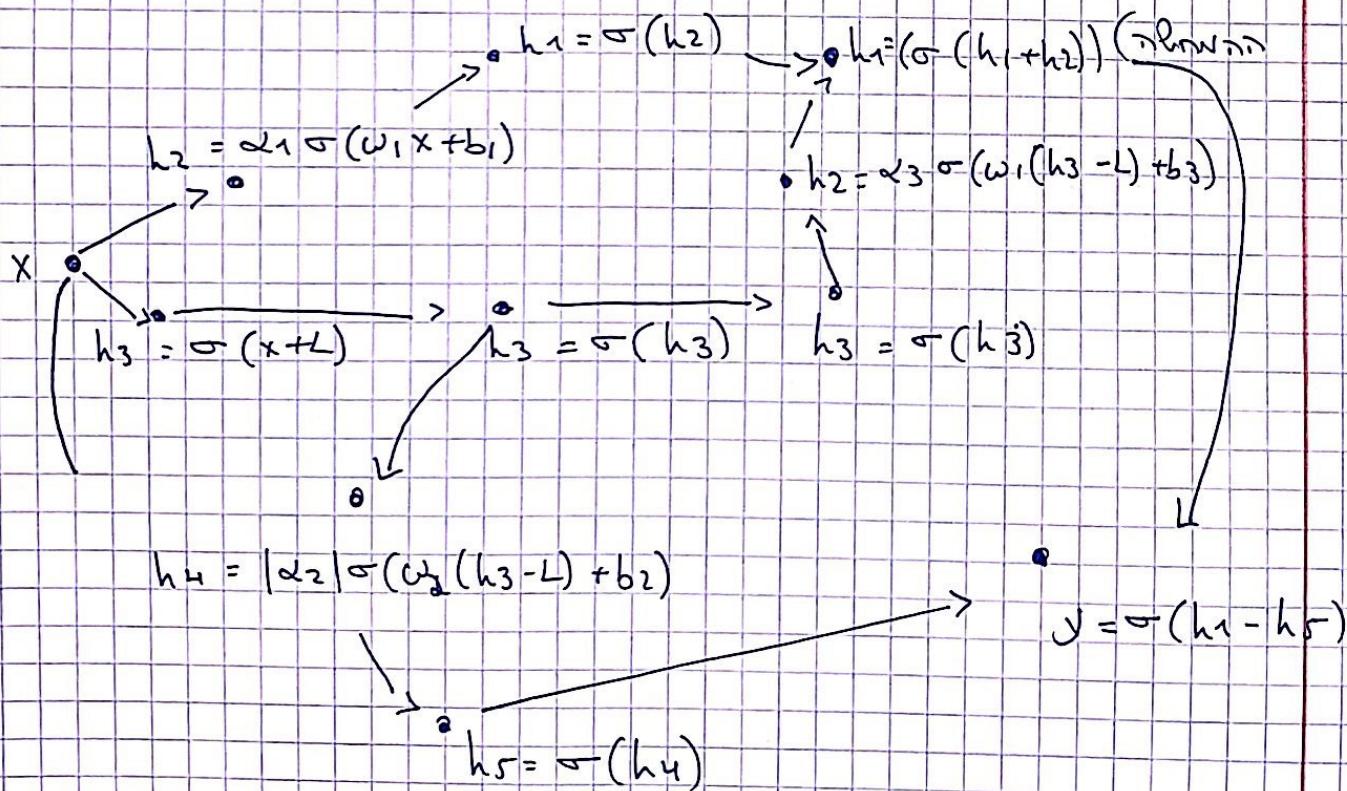
לפיכך $|\alpha_i|$ מוגברת. מוגברת α_i מוגברת. מוגברת α_i מוגברת.

לפיכך מוגברת. מוגברת. מוגברת. מוגברת. מוגברת.

בנוסף, α_i מוגברת. מוגברת. מוגברת. מוגברת. מוגברת.

לפיכך מוגברת. מוגברת.

ר-36 $\alpha_2 \leq 0$ - 1 $\alpha_1, \alpha_3 > 0$, $n=3$ כ. (8)



$$h_4 = |\alpha_2| \sigma(\omega_2 (h_3 - L) + b_2)$$

$$y = \sigma(h_1 - h_5)$$

$$h_5 = \sigma(h_4)$$

ר-37 מילוי 2 הוליך כ. מילוי 2 (8)

מילוי גורם גורם מילוי מילוי מילוי (8)

מילוי גורם גורם מילוי מילוי מילוי (8)

$\alpha_2), O(\omega) -$

מבוא ללמידה عمוקה – תרגיל 1:

בדוח זה נתאר את תהליך העבודה שלנו בחלק המעשי של התרגיל. נכלול גרפים ונסביר את תהליך החשיבה שboveע מאחריו הקוד. נתקלנו במספר קשיים במהלך העבודה, כל אחד מהם פתרנו בשיטה אותה למדנו בקורס ובקורסים קודמים (IMI).

הकושי הראשון בו נתקלנו היה לחשב על אופן נכון לייצוג המידע. לאחר ולא ידענו להעיר שני / מרחק בין חומצות אמינים בחרכנו לייצג כל חומצה על ידי One Hot Encoding. באופן זה יכולנו לאחר מכן לחשב שני אופן שקול בין החומצות כמו כן, שמרנו על יחס הסדר הראשוני.

הוספנו Labels לכל קובץ מידע אותו טענו ואיחדנו את המידע. למעשה, ייצגנו כל רצף של חומצות אמינים בן 9 אותיות, כוקטור בעל 180 איברים.

לאחר מכן, פיצלנו את הדadata באופן רנדומלי בצורה הבאה:
% 90 מהdataset שימוש לאימון הרשות | -% 10 שימר לבחינה.

כמצין בתרגיל, הקושי העיקרי בניתוח המידע מופיע לאחר והdata שקיבלנו לא היהamazon. לאחר קריאה ובחינת מספר שיטות להטמוד עם זה, בחרנו לככלול באימון המודל שלנו שתי שיטות.

הראשונה, הינה להכפיל את הדוגימות החיוביות (Oversampling) וכן בעצם לצמצם את הפער בין הבדיקות החיוביות לשיליות. נציג כי בדקנו מספר פרמטרים להכפל בהם את מספר הדוגימות החיוביות אך הגענו איתן ל-Overfit.

שיטה שנייה בה נקטו, הייתה לדוגם את הדוגימות באימון בצורה הסתברותית. הוספנו משקלות (WeightedRandomSampler) לטוען האימון שלנו וכן בעצם, למדנו את הרשות באופן שקול לצדוק גם על דוגימות חיוביות.

לאחר מכן, הינו מוכנים להתחיל באימון הרשות. ניסינו מספר רב של קונסטרוקציות שונות לבניית הרשות עד שהגענו לאחת שמקסמת את טוויותיה על הדdata המיעוד לבחינה.

נציג מספר גרפים המתארים את תהליך החשיבה ובחרית הפרמטרים לרשות:

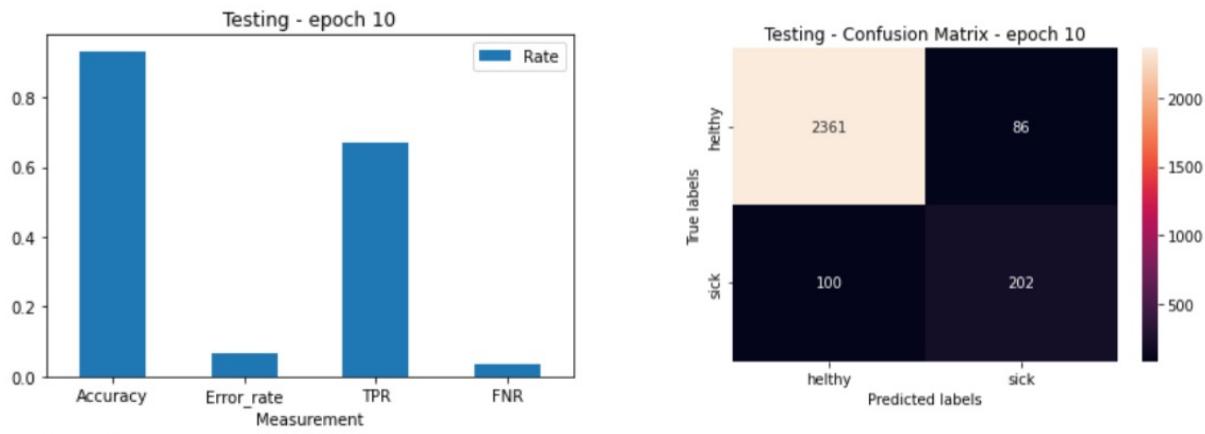
תחילה נציג כי המדריך העיקרי בדקנו למזרע היה מספר הטעויות של הרשות על חולים מאומתים. שכן געדיף לטעות ולשלוח אנשים בריאים לבידוד מאשר לתת לאנשים חולים להסתובב חופשי ולהדביק אחרים.

הרשות הראשונה אותה בדקנו, הייתה רשות פשוטה, עם מספר יחסית קטן של נוירונים שכבה פנימית אחת. חלוקת הדdata לא כללה את ה-oversampling ודגםת הדdata עם משקלות. צפוי הרשות סיוגה את רוב הקלטים כבראים.

```
class Net(nn.Module):
    def __init__(self):
        super(Net, self).__init__()
        self.fc1 = nn.Linear(180, 20)
        self.fc2 = nn.Linear(20, 1)

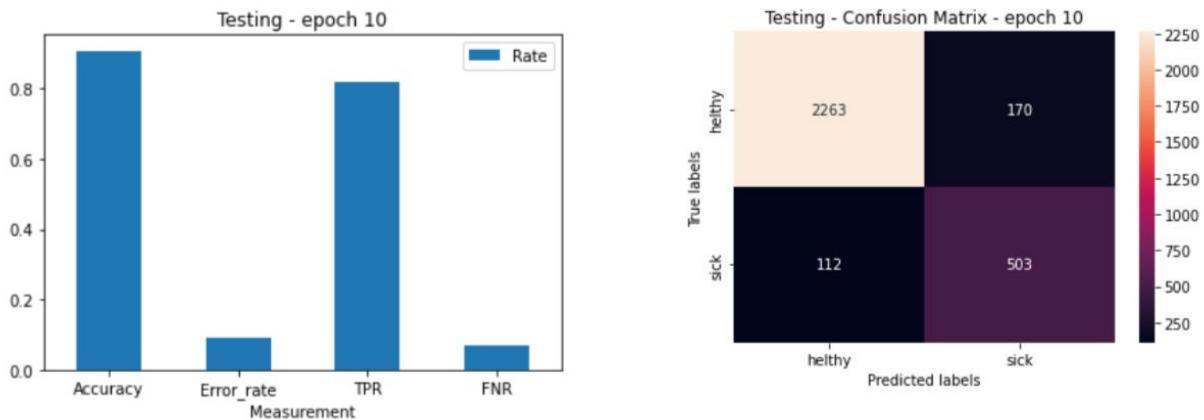
    def forward(self, x):
        x = F.relu(self.fc1(x))
        x = torch.sigmoid(self.fc2(x))
        return x
```

-Net



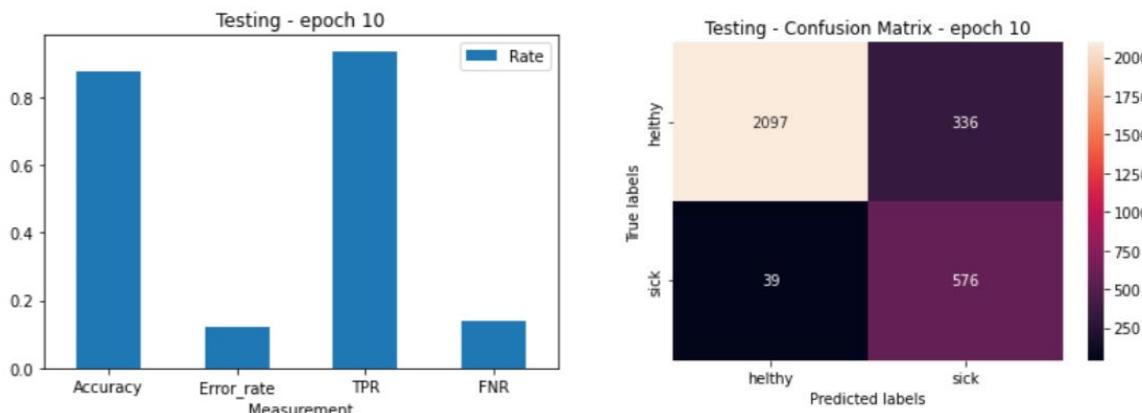
ניתן לראות כי אמונם הדיווק של הרשות גבוהה, אך הרשות אכן טואה בבדיקה החיה/בוגר (TPR) נמוך יחסית ~ 0.66) וכן נכשלת בבדיקה שלנו.

- Oversampling את הנקודות לפתרות הדאטה הלא מאוזן, ראשית ביצענו



ניתן לראות כי ה- TPR עלה ממשמעותית לכיוון ה-0.8, אולם ה- FNR עלה גם הוא (מספר הטעויות על הבדיקות השליליות), אך שוב זהה מכך אותו אנו מוכנים לשלים.

לאחר מכן הוספנו את המשקלות –



שוב, ה- TPR עלה לכיוון ה-0.9 במחיר עליית הטעויות על הבדיקות הבריאות.

השינוי הבא אותו הכננו לרשת היה הגדלת מספר הנוירונים והשכבות, ובנוסף הוספה שכבות BatchNorm. בנוסף בחרנו משתנים כמו גודל התקדמות, גודל Batch, ומספר epochs לבעץ, לא נתאר בחירת כל פרמטר בנפרד שכן ההבדלים הם קטנים.

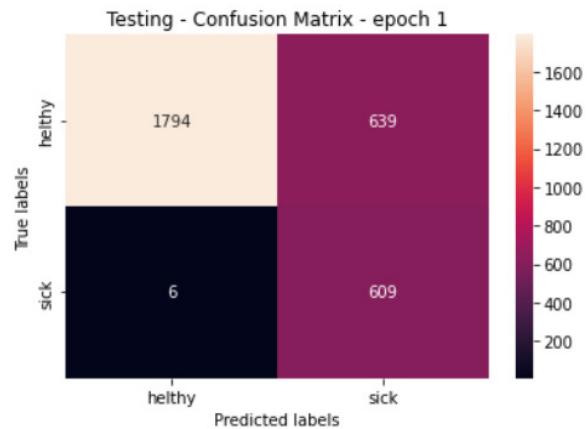
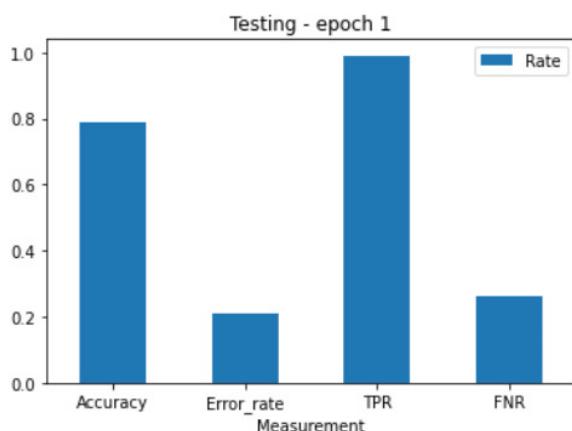
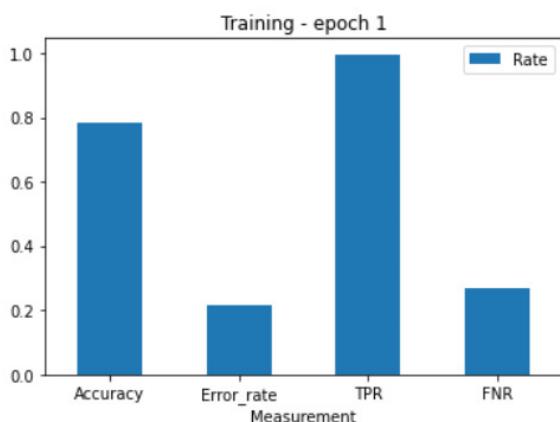
נציג כעת את התוצאה הסופית ל-epoch, כאשר נציג את תוצאות האימון מול תוצאות המבחן. ניתן לראות כי היחס בין הבדיקה נשמר מה שמעיד על כך שהרשת לא מגיעה ל-Overfit:

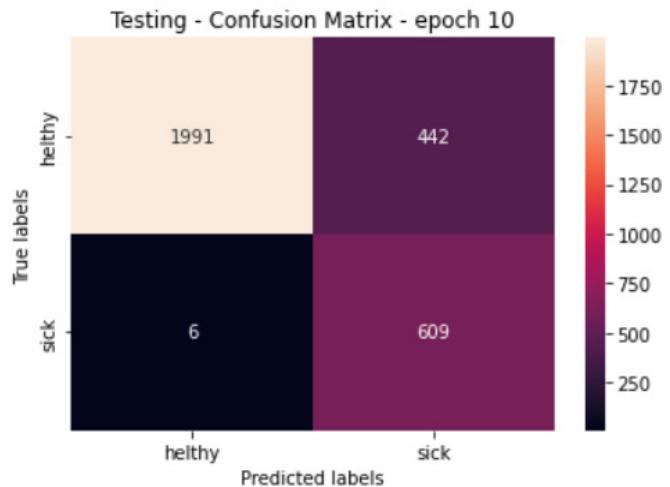
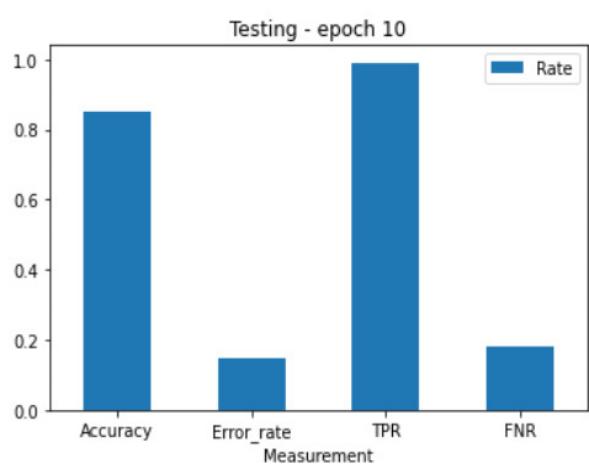
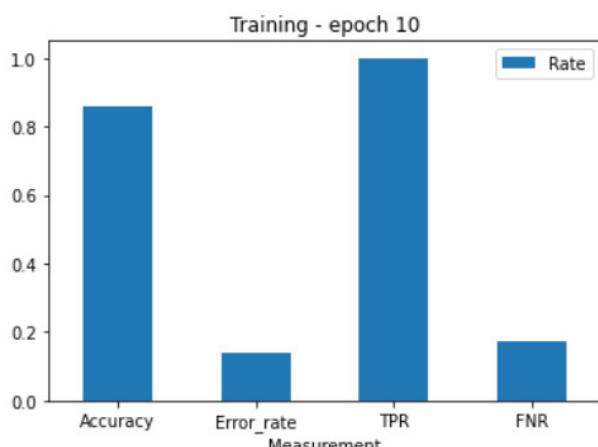
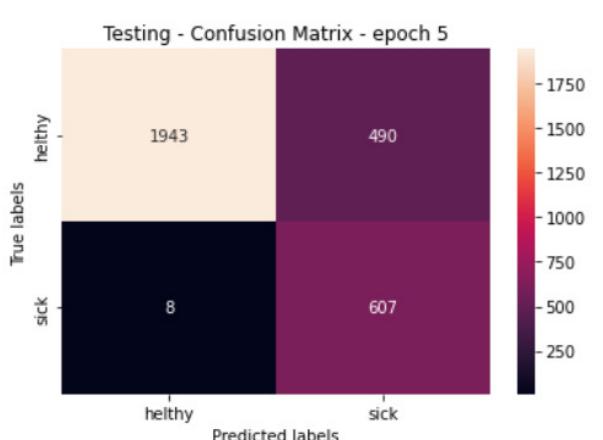
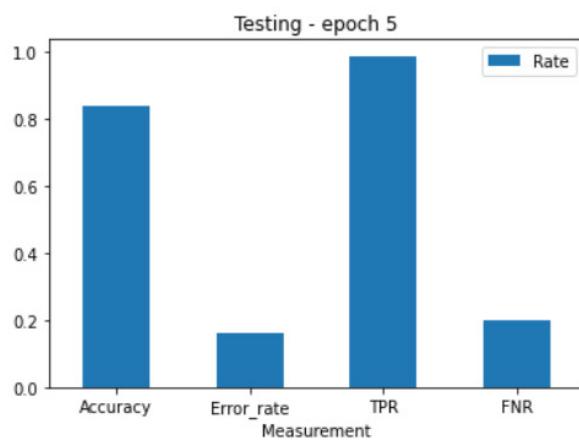
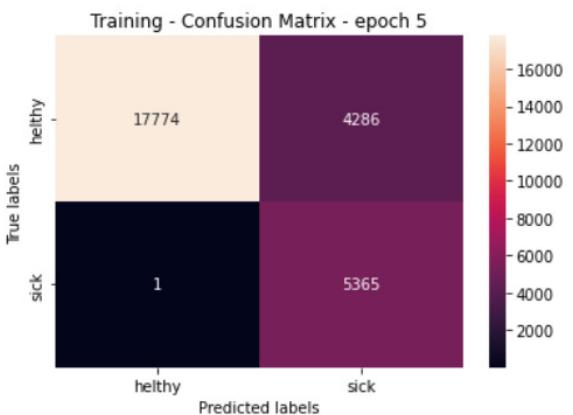
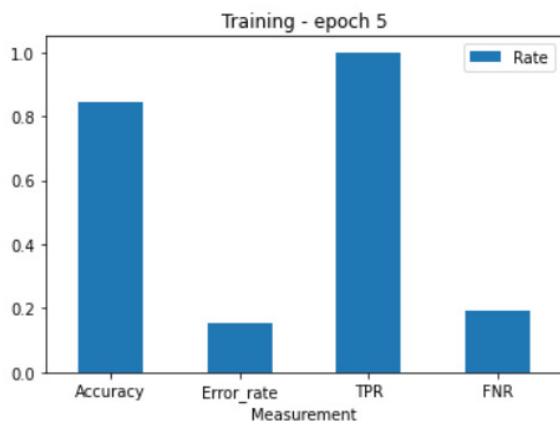
הרשת המוצגת הינה בעלת קצת התקדמות של 0.001, גודל Batch הוא 64 דוגמאות, אנו מבצעים 15 Epochs, קריטריון ה-loss הוא BCE והאופטימיזר הוא ADAM.

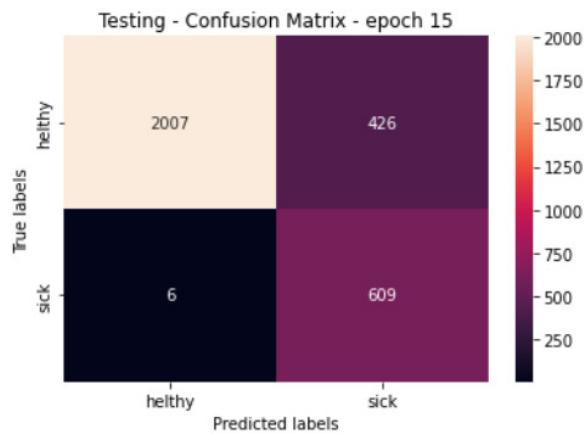
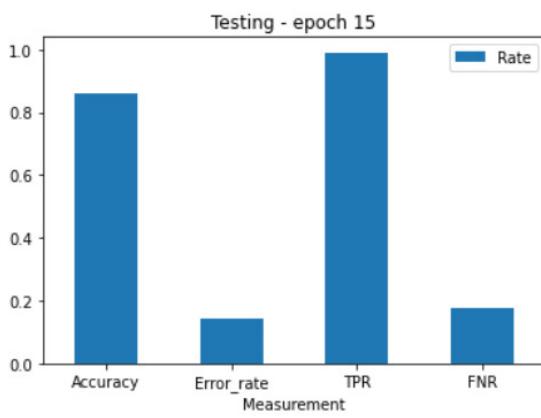
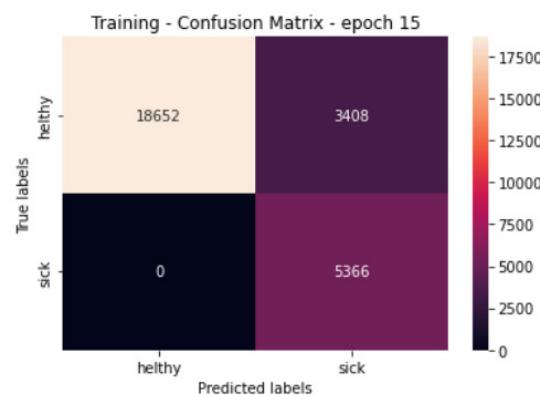
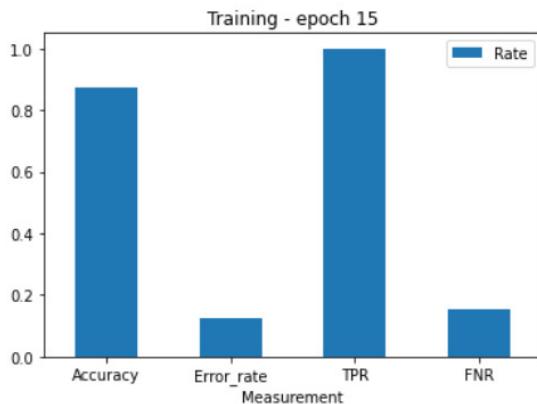
```
class Net(nn.Module):
    def __init__(self):
        super(Net, self).__init__()
        self.fc1 = nn.Linear(180, 60)
        self.fc2 = nn.BatchNorm1d(60)
        self.fc3 = nn.Linear(60, 20)
        self.fc4 = nn.BatchNorm1d(20)
        self.fc5 = nn.Linear(20, 1)

    def forward(self, x):
        x = F.relu(self.fc2(self.fc1(x)))
        x = F.relu(self.fc4(self.fc3(x)))
        x = torch.sigmoid(self.fc5(x))
        return x
```

-Net







אכן ניתן לראות, כי במהלך כל epoch היחס נשמר, אנו מגיעים לכמעט 100% דיוק באנשים החולים ומשלמים על כך במחיר של 400 טעויות ל-2400 אנשים בריאים.

שאלה 6,7:

השתמשנו במודל המאומן כדי למצוא את התחזית הטובה ביותר ביותר לרצפי חומצות האמינו הנתונים, נציג את תוצאות 5 החומצות עם ההסתברות הגבוהה ביותר על פי המודל. נציג Ci התשובה לשאלה 6 היא הרცף הראשון שנציג:

Pepides- DCALDPLSE : probability according to model- 0.9999998807907104
 Pepides- KCTLKSFTV : probability according to model- 0.9999996423721313
 Pepides- VIRGDEVRQ : probability according to model- 0.9999995231628418
 Pepides- DLLFNKVTL : probability according to model- 0.9999991655349731
 Pepides- NIIRGWIFG : probability according to model- 0.9999974966049194

שאלה 8:

בשאלה זו, חיפשנו את רצף החומצות אשר המודל שלנו יופעל על ידו בצורה הגבוהה ביותר על ידו. אימנו את הרשות למקסם את ה שלו כדי להגיע לפט הרצוי. נظر פלט זה ואת ההסתברות לקבלו:

Pepides- ['S' 'L' 'W' 'R' 'L' 'I' 'P' 'I' 'A'] : probability according to model- 0.9906257390975952