

למידת מכונה – תרגיל מספר 3

מבנה הרשת:

השכבה הראשונה ברשת היא בגודל 784 ניורונים כי מקבלים כקלט תמונה בגודל 28×28 שמכילה בעצם 784 פיקסלים. הערך של כל ניורונים הוא צבע בטווח בין 0 לבין 255. בשכבה השנייה יש 34 ניורונים. החיבור בין שתי השכבות הוא fully connected ועל השכבה השנייה מפעילת את פונקציית האקטיבציה סיגמואיד. השכבה השלישית והאחרונה, שהיא למעשה שכבת ה-output שמחוברת ב fully connected לשכבה השנייה והיא מונה 10 ניורונים. מפעילים את פונקציית האקטיבציה softmax על השכבה האחרונה כי המודל מתאמן ונבחן על סט דוגמאות שהוא multiclass. לאחר פונקציית האקטיבציה על שכבת ה-output ניתן לצפות את class התמונה שהוא בטווח בין 0 לבין 9.

בתחילת כל איטרציה נבצע Forward כדי לקבוע את הפרמטרים $W1, W2, B1, B2$ ובסוף ה Forward נפעיל softmax על השכבה האחרונה שנוצרה כפי שנלמד בתרגול.

את ה- backward נבצע גם כפי שנלמד בתרגול רק שההבדל היחיד הוא שנעדכן

$dz2 = (h2 - y)$ במקום: $dz2 = (h2 - \text{one_hot_label}(y))$. כיוון שהמודל שלנו מטפל בדוגמאות שהתיוג שלהם הוא multiclass נעדכן את הנגזרת של $z2$ להיות ערכי שכבת ה output פחות וקטור בגודל 10 שכולו אפסים פרט לאינדקס של y . כלומר אם $y=5$ אז $\text{one_hot_label}(y) = (0,0,0,0,0,1,0,0,0,0)$. סה"כ המודל הגיע לדיוק של 0.91 בעזרת kfold.

נרמול לסט האימון והמבחן: על פי מאמר שמצאתי באינטרנט נרמלתי את הפרמטרים הבאים:

```
W1 = np.random.rand(neurons, 784) / np.sqrt(784)
b1 = np.random.rand(neurons, 1) / np.sqrt(784)
W2 = np.random.rand(10, neurons) / np.sqrt(neurons)
b2 = np.random.rand(10, 1) / np.sqrt(neurons)
```

בנוסף, נרמלתי את כל ערכי x לפי הנוסחה $x = x/255$ וזה על פי מקורות באינטרנט שהבסיס הוא שטווח ערכי x הוא בין 0 לבין 255.

בחירת מקדם הלמידה:

בדקתי דיוק על מקדמי למידה שונים בין 0.02 לבין 0.09 בסט הולידציה. דיוק מקסימלי שהתקבל:

מקדם הלמידה	דיוק
0.05	0.857023088955862

בחירת מספר האיטרציות:

בדקתי מספר איטרציות שונה בטווח בין 5 איטרציות לבין 60 איטרציות ללמידה. דיוק מקסימלי שהתקבל:

מספר האיטרציות	דיוק
50	0.858613095587885

בחירת מספר ניורונים ברשת הפנימית:

בדקתי מספר נורונים שונה ברשת הפנימית בטווח בין 16 לבין 784.

דיוק מקסימלי שהתקבל:

מספר הנורונים	דיוק
34	0.914224