

למידת מכונה תרגיל 3 אלעד ישראל 313448888

פרטי המימוש:

פיצלתי את הדוגמאות שקיבלתי ל-75% אימון (training samples) ו-25% בדיקת האימון (testing samples), מכיוון שאחוזים קטנים/גדולים יותר נתנו הצלחה נמוכה יותר עבור הדוגמאות הנתונות.

התחלתי מאתחול אקראי של הפרמטרים על מנת למנוע הטיה של התוצאות.

הפעלתי את רשת הניורונים על training samples.

השכבה הראשונה של רשת הניורונים היא שכבת input- ובה 784 ניורונים, מכיוון שכל תמונה קטנה מורכבת מ-28x28 פיקסלים.

האלגוריתם ניסה לחזות את הסיווגים של הדוגמאות- מעבירים כל דוגמא ב forward_prop ומשנים (מנסים לתקן) את הפרמטרים (המשקולות בין כל שכבה וה bias) ב back_prop בהתאם ע"י חישוב הגרדיאנטים.

שכבת הניורונים האמצעית היא שכבת ניורונים שבה מספר הניורונים נתון לבחירתי בהתאם לתוצאות ההרצה (מפורט בהמשך).

Labels. The possible labels are:

0. T-shirt/top

1. Trouser

2. Pullover

3. Dress

4. Coat

5. Sandal

6. Shirt

7. Sneaker

8. Bag

9. Ankle boot

לבסוף (לאחר 15 epochs – איטרציות על הקלט) קיבלתי פרמטרים מאומנים- w_1, b_1, w_2, b_2 .

העברתי את שאר הדוגמאות (testing samples) ב forward_prop עם הפרמטרים המאומנים, ועבור כל דוגמא לקחתי את האינדקס של הוקטור γ_{hat} שנתן את התוצאה הכי גבוהה- כלומר את הסיווג החזוי.

לאחר מכן השווייתי את הסיווג האמיתי שלהם (הנתון) עם הסיווג שחזיתי ע"י המסווג. הסיווג הוא מספר בין 0 ל 9 הממפה לסוג אביזר הלבוש המתאים, כמפורט משמאל.

לכן רשת הניורונים האחרונה (output layer) היא בת 10 ניורונים.

ספרתי את כמות ההצלחות (המרתי לאחוזים ע"י חלוקת מס' ההצלחות במס' הדוגמאות) וכך קיבלתי את אחוזי ההצלחה של רשת הניורונים שאימנתי.

Hyper-parameters:

hidden_layer_neuron_count: כמות הניורונים בשכבה האמצעית- 2 הערכים שהניבו את אחוזי ההצלחה הגבוהים ביותר היו 150 ו-250 (עם הפרש שולי ביניהם), בחרתי ב-250.

eta: ניסיתי ערכי eta רבים ושונים מכיוון שלבחירת ה eta הייתה השפעה גדולה מאוד על אחוזי ההצלחה.

עבור ערכי eta גבוהים קיבלתי אחוזי הצלחה נמוכים יחסית. ככל שהקטנתי את eta התחלתי לקבל אחוזים גבוהים יותר, עד לגבול מסויים.

לבסוף בחרתי ב-0.003 מכיוון שזו הייתה ה eta שמיקסמה את אחוזי ההצלחה.

epochs: בכל האלגוריתמים מספר ה epochs הוא 15 מכיוון שזאת כמות האיטרציות שעבורם קיבלתי אחוז הצלחה גבוה (15-20) איטרציות הניבו פחות או יותר אותה תוצאה, בחרתי ב-15 לשיפור מהירות התוכנית.

עבור כמות איטרציות גדולה יותר רמת השיפור הייתה זניחה, ולכן העדפתי להשאיר אלגוריתם מהיר יחסית (15 epochs) עם אחוזי הצלחה גבוהים, מאשר אלגוריתם עם עוד עשרות איטרציות שיהיה הרבה יותר איטי, כאשר ההפרש בין אחוזי ההצלחה זניח.

אקטיבציה:

לפי דרישות התרגיל השתמשתי בReLU כפונקציות אקטיבציה.

נורמליזציה:

על מנת להימנע מoverflow שעלול לקרות בכפל המטריצות הענקיות שאיתן אנו עובדים(מה שמוביל לערכים שגויים – NaN), ניסיתי מספר אלג' נרמול- Z-Score normalization מהתרגיל הקודם עבד כמצופה ולא פגע באחוזי ההצלחה.

לאחר ניסויים רבים עם פרמטרי האלגוריתמים וכו'(מוצג לעיל) הגעתי לאחוזי ההצלחה הבאים:

Success rate: **84.03%**

כאמור, בעזרת הפרמטרים הבאים:

eta= 0.003

epochs=15

hidden_layer_neuron_count= 250