

למידת מכונה תרגיל 3 – דו"ח

בתרגיל זה נדרשנו לממש רשת נוירונים אשר מסווגת 10 קלאסים של פרטי לבוש שונים. בדו"ח זה אתאר על ארכיטקטורת הרשת ועל אופן בחירת ההיפר פרמטרים.

הרשת מכילה שיכבה אחת חבויה כאשר המשקלים בה מאותחלים רנדומלית (ניסיתי בהתחלה לאתחלה ב-0 אך קראתי שזה גורם לבעיות עם מציאת מינימום לוקאלי). מספר הנוירונים בשכבה הינו 97 בחרתי במספר זה לאחר ביצוע grid search על הפרמטר והגעה לאחוז הדיוק המקסימלי.

את הeta אתחלתי בהתחלה ל-0.001 אולם קיבלתי יחסי דיוק נמוכים יחסית על כן ניסיתי ערכים אחרים. שמתתי לב שהעלת הeta גורמת לירידה בדיוק ועל כן ניסיתי להקטינה עד שלבסוף מצאתי את הערך האופטימלי של 0.00075.

פרמטר נוסף הינו מספר epochs בדקתי בטווח של (1-100) את אחוז הדיוק. מה-epoch ה-14 לא הייתה עלייה ניכרת באחוזים והייתה אף ירידה, כ"כ הייתה עלייה גדולה בזמן הריצה ועל כן בחרתי במספר זה.

בעדכון הפרמטרים השתמשתי ב-SGD כאשר batch size של 1105 הביא לתוצאות הטובות ביותר.

את אחוזי הדיוק בדקתי על validation set של כ-20% מהדאטא והמשכתי עם כיול ההיפר פרמטרים עד לקבלת ממוצע של כ-85%-86% דיוק.