

למידת מכונה דו־ח תרגיל 2

בתרגיל בניתי רשת נוירונים עם שכבה הקלט, שכבה נסתרת ושכבת הפלט

שכבת הקלט בגודל הפיקסלים של התמונה ולכן 784 נוירונים

בגלל שמדובר ב multiclass classification אזי תוצאת שכבת הפלט צריכה להיות כגודל class האפשריים ובמקרה שלנו מדובר ב10.

את שכבת הפלט בחרתי להיות 110 נוירונים(אחרי ניסוי וטעיה של תוצאות טובות ביותר)

בהתאם לכך הגדרתי 2 מטריצות w_1, w_2 להיות המשקולות בין השכבות. בנוסף הגדרתי גם 2 וקטורים b_1, b_2 להיות bias בכל שכבה.

התחלתי את המטריצות והווקטורים עם ערכים נמוכים בשביל התכנסות מהירה וטובה יותר (לאחר ניסוי וטעיה של אחוזי הצלחה על validation set)

אחרי כל דוגמא עדכנתי את המשקולות וה-bias בפונקציה update כמו כן בפונקציה זאת השתמשתי ב learning rate ובכל 4 epochs הקטנתי אותו פי 2

היפר פרמטר נוסף הוא כמות epochs וגם כאן לאחר ניסוי וטעיה הגעתי למקסימום הצלחה תחת 15 epochs

בין כל epochs עשיתי shuffle לדוגמאות בשביל שלא ילמד לפי סדר מסוים של הקלט בכל סיבוב. בנוסף עשיתי shuffle בהתחלה בשביל לא ללמוד מהקלט הממוין.

בנוסף ביצעתי נורמליזציה לקלט x בשביל להשתמש בערכים בין 0-1.

בין השכבות השתמשתי בפונקציית sigmoid בשביל להפוך את הלמידה ללא לינארית ובנוסף השתמשתי בפונקציית softmax בשכבה האחרונה במקום sigmoid כי מדובר בmulticlass.

היפר פרמטרים:

כמות הנוירונים בשכבה אמצעית : 110

אתחול מטריצות w : בערכים בין 0-0.0001

אתחול וקטורים b : בערכים בין 0-1

אתחול ערך הלמידה ϵ : 0.05 ובכל 4 סיבובים מקטין פי 2

כמות epochs: 15