

Report

בתרגיל 3 בנינו רשת נוירונים לחיזוי ספרות מ-0 עד 9 בתמונות.

ברשת הנוירונים שבניתי קיימת שכבה חבויה אחת, ובה 170 נוירונים.

אתחלתי את וקטורי המשקולות להיות מספרים שקרובים ל-0, כך שההטיה ההתחלתית תהיה נמוכה ככל האפשר ועם זאת לא תהיה אפסים. את הווקטורים אתחלתי להיות מספרים בין 0.001 ל-0.001-, באופן רנדומלי בהתפלגות אחידה. את הביאסים אתחלתי להיות אפסים.

את דוגמאות האימון והמבחן נרמלתי למספרים בין 0 ל-1 על ידי חלוקה ב-255 את כל הערכים.

עברתי על כל הקלט 25 פעמים (=מספר ה-epochs), והעברתי כל דוגמה דרך הרשת.

בפעפוע קדימה:

השתמשתי בפונקציית סיגמואיד ע"מ להפוך את הdata ללא לינארי, ובשכבה האחרונה השתמשתי על הפלט ב-softmax על מנת לקבל הסתברויות.

את פונקציית ההפסד חישבתי באמצעות ביצוע לוג על ההסתברות של התיוג המקורי, והכפלתו ב-1.

בפעפוע אחורה:

חישבתי את הנגזרות המתאימות כפי שראינו בתרגול ובהרצאה.

את הערכים החדשים עדכנתי בעזרת הנוסחה שראנו בתרגול ובהרצאה, כאשר את אטא אתחלתי להיות 0.01.

פרדיקציה:

עשיתי פעפוע קדימה עבור כל דוגמה, ולקחתי את האינדקס של הערך הגבוה ביותר מתוך מערך ההסתברויות שקיבלתי כתוצאה מ-softmax.

אופן אתחול הפרמטרים:

את הפרמטרים אתחלתי בערכים שונים בצורה שרירותית. חילקתי את הקלט ל-80% אימון ו-20% ואלידציה על מנת שאוכל למדוד את ההתקדמות שלי בכל שלב. בדקתי את ההפסד בכל epoch, וידאתי התכנסות ובדקתי את אחוזי ההצלחה בכל שלב כתלות בשינוי בהיפר-פרמטרים השונים.