**דו"ח מסכם למידה עמוקה תרגיל בית 1**

מגישים: רועי זהר 209896174 רועי מזן 319042800

להלן התוצאות שהתקבלו:

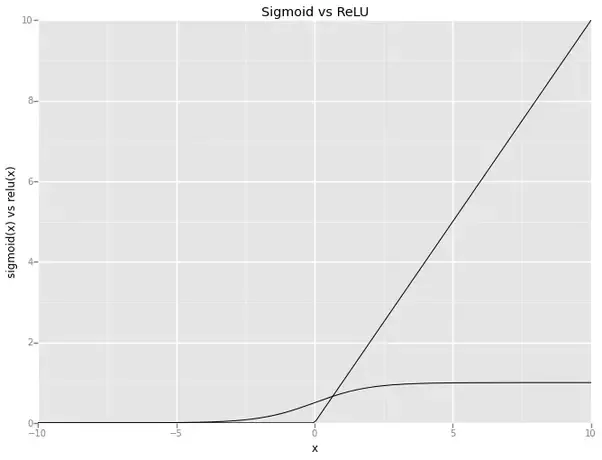
|  |  |
| --- | --- |
| **Accuracy:** | **Function:** |
| 0.897 | One\_layer\_no\_activation |
| 0.894 | Two\_layer\_no\_activation |
| 0.9406 | Two\_layer\_sigmoid\_activation |
| 0.9417 | Two\_layer\_RELU\_activation |

נשים לכמה תופעות מעניינות:

ראשית כל, ההרצות שהתבצעו בלי אקטיבציה הניבו את התוצאות הכי פחות טובות. הסיבה לכך היא שכפי שהוסבר בהרצאה, המודל שהתקבל הוא מודל **ליניארי**, שהוא מודל בעל סיבוכיות נמוכה יחסית. זאת מכיוון שהרכבה של פונקציות לינאריות היא לינארית, ולכן עבור אימון מספיק ארוך נצפה לקבל את אותה תוצאה. לכן, גם לא צפינו בשינוי כמעט בין שכבה אחת ושתי שכבות.

למעשה, הוספת השכבה השנייה ברשת ללא אקטיבציה, רק מוסיפה סיבוכיות מיותרת לשלב האימון, וייתכן שזו הסיבה לכך שהתוצאות שהרשת בעלת שתי השכבות הניבה תוצאות קצת פחות טובות

ניכר השיפור לאחר השימוש בפונקציות אקטיבציה, משום שהמודל שלנו נהיה מסובך יותר (במובן הטוב). נשים לב שאקטיבצית RELU פעלה במעט יותר טוב מאקטיבצית הsigmoid. ההצדקות לכך הן כפי שראינו בכיתה, שעבור אינפוטים גדולים הsigmoid כמעט לא מבדיל ביניהם. זאת לעומת הRELU שמבדיל גם בין אינפוטים גבוהים:



בנוסף, RELU נותן לנו יותר sparsity משום שהוא מאפס משקלים שליליים, שהיא תכונה שראינו כחיובית בשיעורים.