**שאלה 2:**

*הרשת שבניתי מורכבת משכבת קלט ופלט ושתי שכבות ביניים.*

*שכבת הקלט: 4 נוירונים שמתאימים לקלט של בדיקה/אימון*

*שכבת הפלט: 3 נוירונים שמתאימים לשלושת הסיווגים האפשריים של שלושת זני האירוסים.*

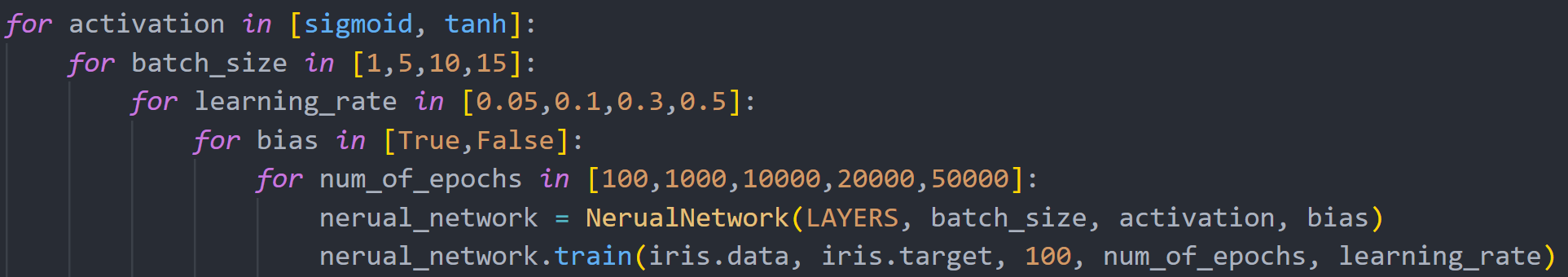
*שכבות הביניים: מורכות מ-7 נוירונים.*

*כל אחד מהנוירונים בשכבה מסוימת מחובר לכל אחד מהנוירונים בשכבה שאחריו. ההשראה למודל הגיעה מחיפוש באינטרנט ומציאת מספר רב של דוגמות לרשתות נוירונים מלאכותיות שמבצעות סיווג ועושות שימוש במודל זהה.*

*את אימון המודל עשיתי על המאה דוגמאות הראשונות והבדיקה שלו על חמישים הדוגמאות האחרות.*

*למודל קונפיגורבילי לכן יכולתי בקלות לבחון את השפעת הפרמטרים הבאים:*

* *סוג פונקציות האקטיבציה: sigmoid / tanh*
* *גודל הbatch: 1/5/10/15 דוגמות*
* *קצת הלמידה: 0.05/0.1/0.3/0.5*
* *הפעלת bias: כן/לא*
* *מספר epochs: 100/1,000/10,000/20,000/50,000*

*על מנת לבדוק קונפיגורציה עדיפה נעזרתי בקוד הבא:*

*לאורך הבדיקות הבחנתי במספר דברים:*

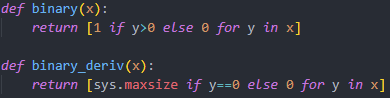
1. *פונקציית sigmoid מביאה לטעויות קטנות יותר בבדיקה ובאימון מאשר פונקציית tanh*
2. *באזור ה-10,000 epochs (לרוב) נגיע לטעות מינימלית ולאחר מכן היא עלולה לעלות מעט (overfitting)*
3. *ככל שקצב הלמידה גבוה יותר, הרשת תלמד יותר טוב (במסדרת הepochs שהגדרתי) את פונקציית הסיווג.*
4. *ככל שה-batches קטנים יותר, הרשת לומדת יותר טוב והטעות קטנה יותר.*
5. *הגרף שבחרתי להציג הוא של הריצה הטובה ביותר של הרשת שעבורה:*

* *פונקציית האקטיבציה היא sigmoid*
* *הרשת מבצעת 10,000 epochs על מידע האימון*
* *קצב הלמידה הוא 0.5*
* *גודל ה-batch הוא 1*

1. *ניתן לראות מהגרף שהצגתי שאכן הרשת למדה כיצד לסווג נכון סוגי אירוסים שונים ולא סתם שיננה את קלט הבדיקה כי גרף הטעות של הלמידה ושל הבדיקה לאורך epochs תואם במגמה שלו, לכן אין סימן של overfitting ברשת.*
2. *ישנן מספר דוגמאות שהרת לא מסווגת נכון:*
   * *5.9, 3.2, 4.8, 1.8 – מסווג כ- virginica למרות שהוא בפועל versicolor*
   * *6.3, 2.5, 4.9, 1.5 – מסווג כ- virginica למרות שהוא בפועל versicolor*
   * *6, 2.7, 5.1, 1.6 – מסווג כ- virginica למרות שהוא בפועל versicolor*

*סביר להניח שהם מסווגים לא נכון בכלל כיוון שגודל עלה הגביע שלהם גדול יחסית לזן שלהם. ניתן לראות מבחינה מהירה של נתוני הזנים שעלי גביע גדולים יותר מזוהים לרוב עם virginica ולכן סביר שהמודל מסווג לא נכון.*

1. *כאשר שיניתי את פונקציית האקטיבציה שלי להיות בינארית, התוצאות היו שהרשת פשוט לא הצליחה ללמוד ולשנות את משקלי הסינפסות ולכן ונתקעה על התוצאה הראשונית שלה:*

*פונקציית האקטיבציה הבינארית היא:*

1. *שימוש בפונקציית אקטיבציה בינארית יחד עם שימוש באלגוריתם backpropagation בעייתי כיוון שפונקציות בינאריות לא מאפשרות מידע על מגמה כאשר מסתכלים על הנגזרת שלהם. באלגוריתם backpropagation עושים שימוש בנגזרת של פונקציית האקטיבציה על מנת לפעפע לאחור את השינוי הנדרש בקשתות הרשת שיגרור שגיאה קטנה יותר בשכבת הפלט, אך אם נעזרים בפונקציית אקטיבציה בינארית המידע הנגזרת שלה תהיה 0 בכל מקום שהוא לא 0 ושם הנגזרת לא מוגדרת (או אינסוף) לכן אין לרשת דרך ללמוד את מגמת הבעיה.*

*לינק לסרטון הסברה:*

[*https://drive.google.com/drive/folders/1KPikJNz8u8V8e86dnzbG7TuJ7uDq56uQ?usp=sharing*](https://drive.google.com/drive/folders/1KPikJNz8u8V8e86dnzbG7TuJ7uDq56uQ?usp=sharing)