

## דגימת נתונים סינטיים והגדרת מודל הסיווג

רשתות הנירונים אשר נשתמש בהן בהמשך למימוש מטרות מורכבות כגון תרגום בין שפות או זיהוי אובייקטים בתמונה דורשות מורכבות חישובית הבאה לידי ביטוי בין השאר בכך שהן מורכבות ממיליוני נירונים המחוברים זה לזה בדרכים מתוחכמות. למזלנו גם היחידה הבסיסית של מורכבות זו, הנירון הבודד, מאפשרת בפני עצמה מימוש של מטרות כנ"ל, אם כי פשוטות יותר. למשל נירון יחיד יוכל להתמודד עם סיווג בינארי של קבוצות נק' פשוטות במרחב דו מימדי.

בעזרת הספרייה scikit-learn נוכל לייצור אוספי נתונים כאלו בנקל. בקטע הקוד הבא, בפקודה אחת נייצר דגימות רנדומליות משתי מחלקות שונות.

```
import sklearn.datasets as skds
X, Y = skds.make_blobs(n_samples=100, n_features=2,
                       centers=2, random_state=1)
print(X[:5, :], type(X))
print(Y[:5], type(Y))
```

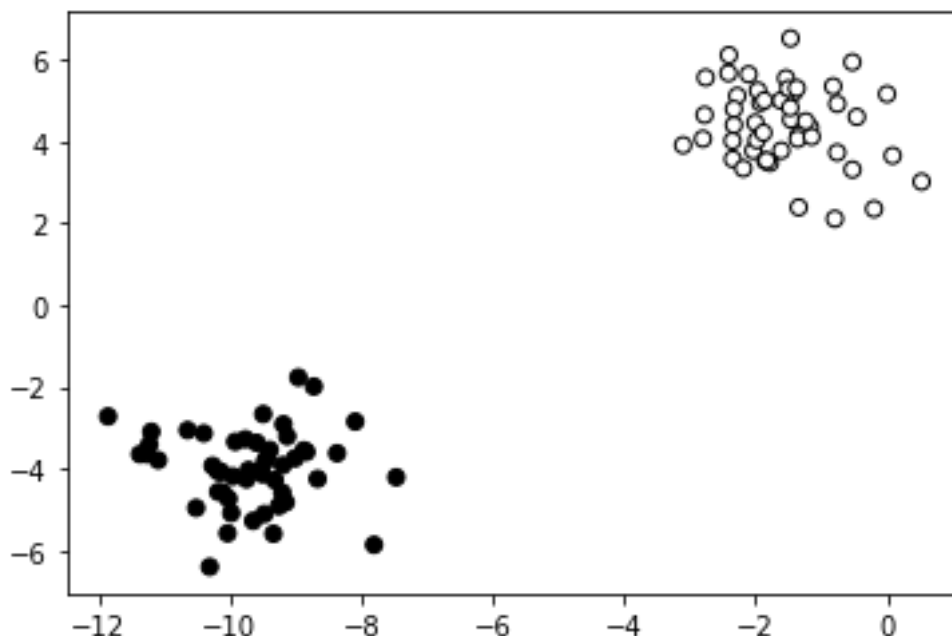
פלט:

```
[[ -0.79415228  2.10495117]
 [-9.15155186 -4.81286449]
 [-3.10367371  3.90202401]
 [-1.42946517  5.16850105]
 [-7.4693868  -4.20198333]] <class 'numpy.ndarray'>
[0 1 0 0 1] <class 'numpy.ndarray'>
```

ניתן לראות שקואורדינטות הנתונים והסיווג התקבלו במשתנים מסוג ndarray של NumPy ולכן בנקל נוכל להמיר אותם בעת הצורך לטנזורים של Pytorch בעזרת הפונקציה `torch.tensor()`, בנוסף בעזרת הספרייה Matplotlib נוכל לצייר את הנתונים במישור:

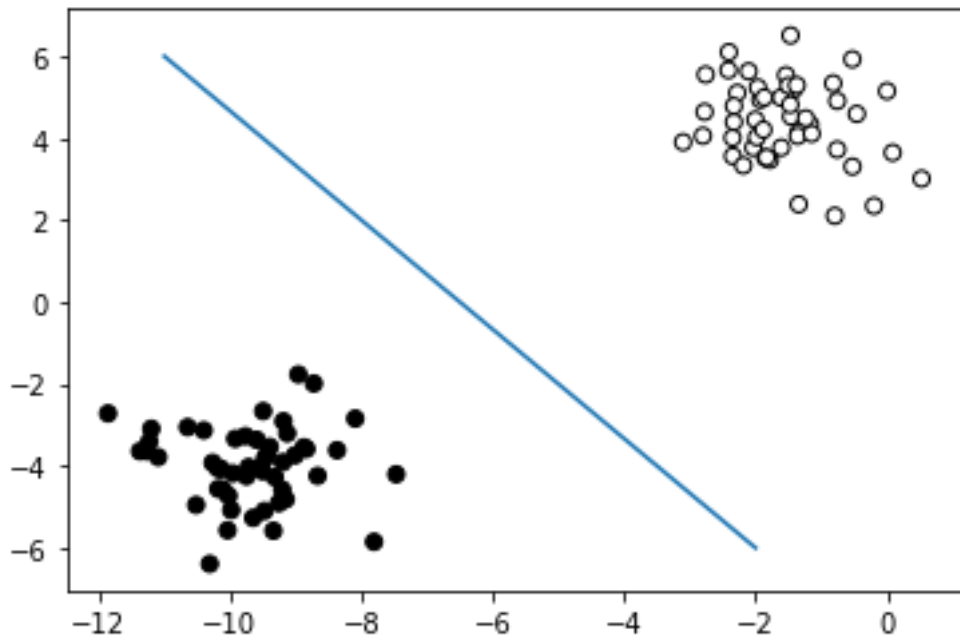
```
import matplotlib.pyplot as plt
plt.scatter(X[:, 0], X[:, 1],
           c=Y, cmap="Greys", edgecolor="black");
```

פלט:



שימו לב לשורה השנייה בקריאה לפונקציה `scatter()`, שם אנו קובעים שהנקודות ייצבעו בצבעים שונים לפי הערכים במשתנה  $y$ , שהצבעים השונים יהיו שחור/לבן וששפת העיגולים תהיה צבועה בשחור (כדי שנוכל לראות את העיגולים הלבנים), וזאת על ידי שינוי שלושת הפרמטרים המתאימים. ניתן לשנות פרמטרים רבים נוספים של הגרף, כגון החלפת העיגולים בריבועים או כוכביות, הפיכת העיגולים לחצי שקופים ועוד. אין צורך לזכור בעל פה פקודות אלו, אלא לחפש את הדרוש בתייעוד של Matplotlib.

אוסף נתונים זה מספיק פשוט בכדי נירון הבודד יוכל להבדיל בין נקודות שחורות ללבנות, וזאת מפני שהנקודות ניתנות להפרדה על ידי פונקציה ליניארית, כפי שנבין בהמשך. לעת עתה רק נשים לב שבבירור ניתן לצייר קו במישור המפריד בין שתי המחלקות, למשל:



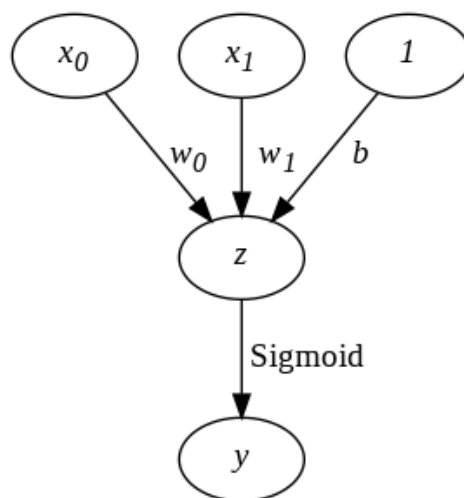
כעת כשאוסף הנתונים בידינו, נפנה להגדרת מודל הסיווג, המורכב מנירון יחיד:

1. הנירון יקבל כקלט את קואורדינטות הנקודות,  $(x_0, x_1)$ .
2. על הקלט תופעל פונקציה ליניארית  $z = w_0 x_0 + w_1 x_1 + b$ .
3. תוצאת חישוב זה תועבר לפונקציית אקטיבציה (activation function),  

$$y = S(z) = \frac{1}{1 + e^{-z}}$$

4. פלט הנירון,  $y$ , הוא ההסתברות שהנקודה הנתונה שייכת למחלקת הנקודות השחורות.

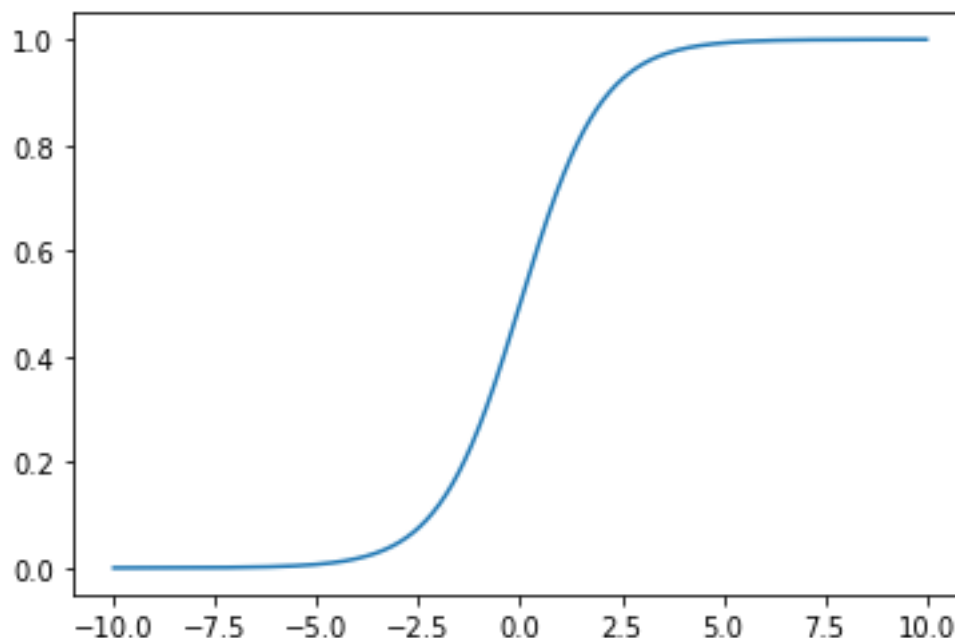
ובאיוור,



שימו לב שפונקציית האקטיבציה אותה בחרנו עבור נירון זה, נקראת סיגמואיד שכן גרף הפונקציה  $y = S(z)$  דמוי האות S, אשר מתחילה ב-0 ונגמרת ב-1. בהתאם, אפשר לפרש תוצאת אקטיבציה זו כהסתברות ולכן היא שימושית למשימות סיווג בינאריות. בקטע הקוד הבא אנו מציירים את גרף הפונקציה לצורך המחשה.

```
import numpy as np
z=np.linspace(-10,10,1000)
y=1/(1+np.exp(-z))
plt.plot(z,y);
```

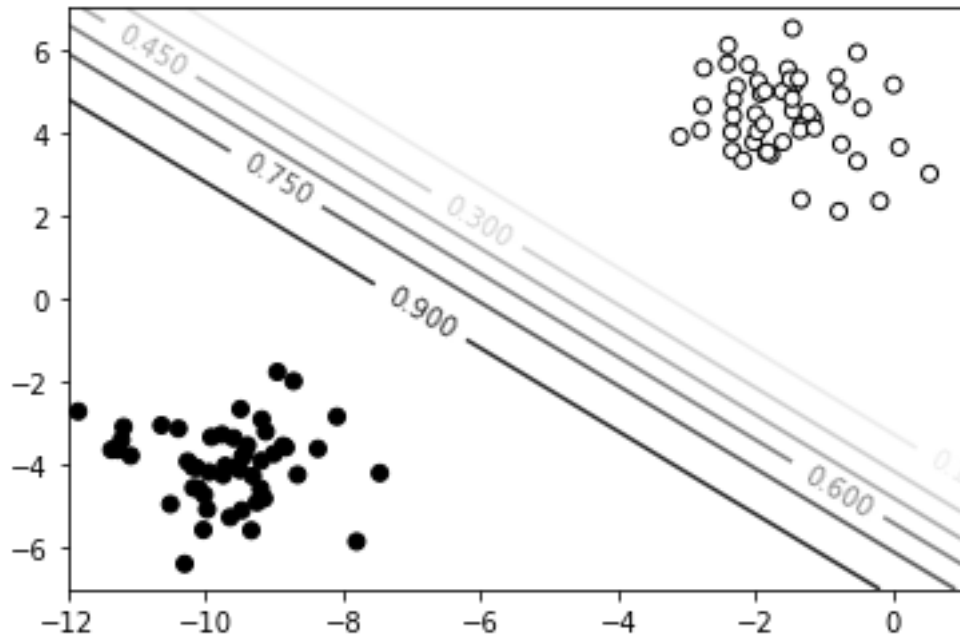
פלט:



מטרתנו כעת היא למצוא את ערכי הפרמטרים של המודל, הלא הם  $w_0, w_1, b$ , כך שהמודל יבצע עבודתו נאמנה: כאשר קלט הניירון יהיה נקודה לבנה, נרצה שהמודל יחזיר כפלט הסתברות קרובה ל-0, ועבור נקודה שחורה – הסתברות קרובה ל-1. במקרה הנוכחי משימה זו קלה במיוחד, שכן פרמטרים אלו קובעים את מיקומו של הקו המפריד בין המחלקות, כמצוייר לעיל וככל שנקודות הקלט

מרוחקת מהקו, כך ההסתברות שהמודל יחזיר תהיה קרובה יותר ל-1 (עבור מתחת לקו) או ל-0 (עבור נקודות מעל הקו). לכן, אם נבחר קו העובר "באמצע", נקבל מודל מעולה.

למשל עבור בחירה מסויימת של פרמטרים, נקבל את המודל המסווג המופיע בציור הבא. שימו לב שעל כל קו כתובה ההסתברות שהמודל מייחס לנק' הנמצאת עליו להיות שחורה. כדורש, הקווים הקרובים למחלקה הלבנה מצביעים על הסתברות נמוכה, ואלו הקרובים למחלקה השחורה – על הסתברות גבוהה.



בעתיד, כאשר הבעיות איתה נתמודד תהיה מורכבת יותר, ובהתאם גם המודל, לא תהיה לנו אפשרות לבחור את הפרמטרים המתאימים באופן ידני, ועל כן דרוש לנו אלגוריתם שיבצע זאת באופן אוטומטי, אותו נשלב בדוגמה זו בהמשך.

### שאלות לתרגול

1. הסבירו למה קווי הגובה באיור האחרון הם קווים ישרים.
2.
  - א. בהנתן בחירה מסויימת של פרמטרים עבור הנוירון, מצאו את משוואת הישר אשר כל נקודה מתחתיו מסווגת בהסתברות גבוהה מ-0.5 כנקודה שחורה.
  - ב. האם יש משוואה נוספת המתארת את אותו קו?
  - ג. האם יש הבדל בין המודלים המתקבלים משני סטים של פרמטרים אשר עבורם הקו המפריד זהה? אם כן, הסבירו מהו.
3. החליפו את פונקציית האקטיבציה של הנוירון ל-ReLU:
 
$$\text{ReLU}(x) = \begin{cases} 0 & x < 0 \\ x & x \geq 0 \end{cases}$$
  - א. ציירו גרף של פונקציית האקטיבציה.
  - ב. נסו להשתמש באקטיבציה זו במקום הסיגמואיד לצורך סיווג הנקודות, והסבירו את הקשיים בהם אתם נתקלים.
  - ג. נסו לפתור את הקשיים על ידי שינוי קל של המודל.