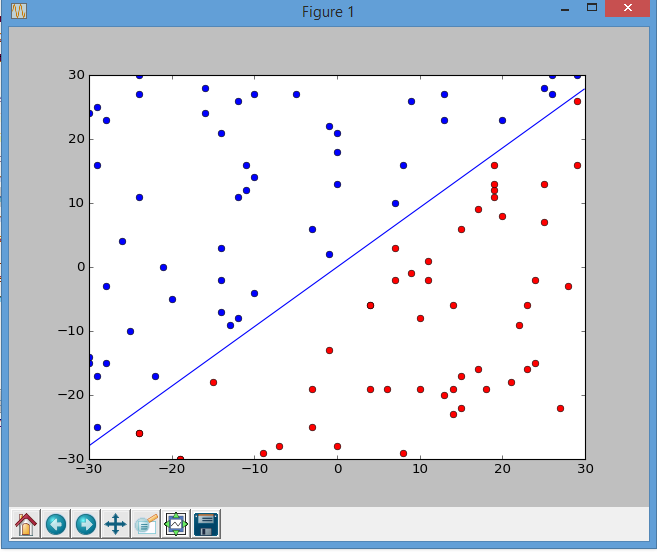
גיא כהן 304840283

למידת מכונה – דו"ח תרגיל 2

f.

h.

מספר השגיאות לאחר ערבוב ה-data השתנה, מאחר והתכנסות ה-w לוקטור האופטימלי תלויה בסדר המעבר על ה-data. כיוון שהסדר השתנה, ישנה התכנסות שונה לוקטור w האופטימלי – ולכן גם וקטור w שקיבלנו בתוצאה הינו שונה. ישנם מספר וקטורים (מספר היפותזות) שמצליחים להפריד את הדוגמאות, אך העובדה כי עברנו בסדר שונה על הדוגמאות מביאה להתכנסות לוקטורים שונים וכן למספר שגיאות שונה בכל הרצה.

i.

לפי ה-data והקו המפריד, נין לראות שקיים w אופטימלי שנורמה 1 שלו היא 1. זהו הוקטור:

W = [0 -0.5 0.5]

שנובע מהישר y = x המפריד בין הדוגמאות.

נוכל לחשב את חסם השגיאות לפי וקטור זה ונקבל שהחסם הינו 13924.0.

החסם אכן מחזיק היות ואנו קיבלנו תוצאות יותר נמוכות בהרבה במספר השגיאות (והצלחנו להגיע לוקטור המפריד).

j. נריץ עם eta = 0.5 ונקבל:

עבור eta = 1:

W: [1.0 -40.0 43.0]

Iterations: 3

Mistakes: 13

עבור eta = 0.5:

W: [0.5 -20.0 21.5]

Iterations: 3

Mistakes: 13

מס' השגיאות ומס' האיטרציות לא השתנו.

הוקטור עצמו כן השתנה, בהתאם ל-eta שבחרנו. מכיוון שבחרנו eta = 0.5 אז כל וקטור w קטן פי 2.

הסיבה שהוקטור השתנה בצורה הזו היא שבכל טעות העדכון של w הוא בדיוק אותו עדכון (כי התנאי לא השתנה) שמוכפל ב-eta. בתנאי עצמו w אכן משתנה, אך מאחר וכל האיברים מוכפלים ב-eta, ניתן לחלק ב-eta ובעצם לשמור על אי השוויון כמו שהוא. אי לכך, בכל מקום בו נעדכן בוקטור המקורי נעדכן גם את הוקטור החדש, פשוט העדכון מוכפל ב-eta.