

## דף תרגיל 2

1. (20 נק') שטוחלנדיה (flatland) היא ארץ המתקיימת בשני ממדים בלבד. ליצורים בשטוחלנדיה רשתית חד ממדית בעלת 16 ניוונים. תבנית הפעילות של הרשתית היא פונקציה חד ממדית המייצגת את עוצמת האור בכל נקודה על-פני הרשתית. תבנית פעילות זו עוברת עיבוד ברשת שכבתית שניתן למדל כרשת קונבולוציה עם שכבת קונבולוציה יחידה בעלת פילטר יחיד באורך 3, שכבת max-pooling ושכבה שמהווה פרספטרון בינארי (כולל איבר bias).  
בשטוחלנדיה חיים ארנבונים ודביבונים. לכל הארנבונים בשטוחלנדיה יש צורה עגולה בעלת קוטר קבוע של 4 ולכל הדביבונים צורה מלבנית בעלת אורך 8 ורוחב 6.  
בשכבת הקונבולוציה מתפתח במהלך הלמידה פילטר יחיד כאשר כל המשקלים בו הם 1.

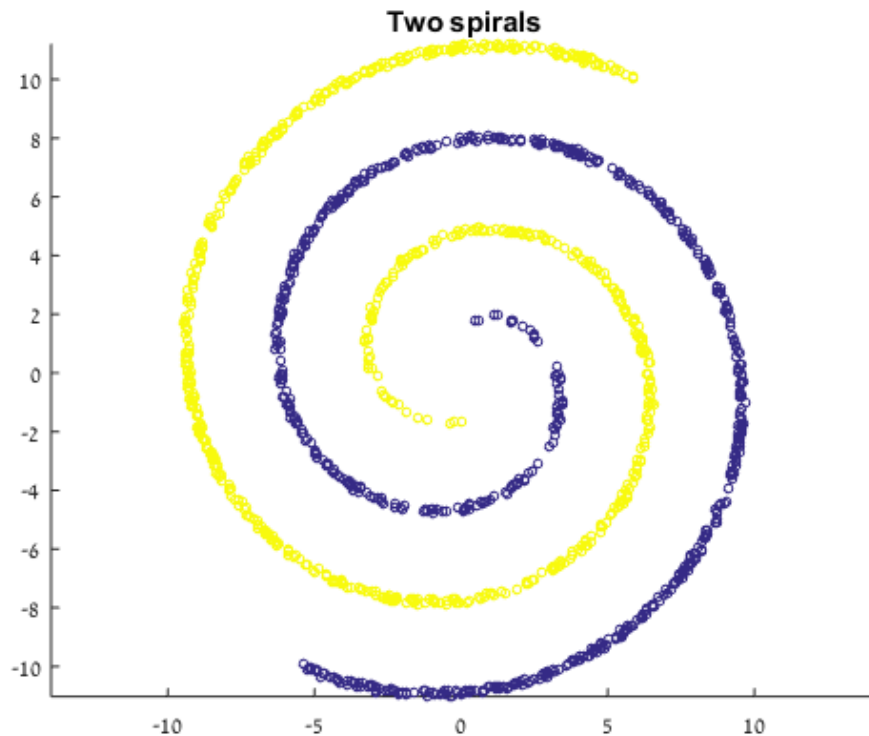
- א. אחת ההפרעות הנירולוגיות שמתרחשות בשטוחלנדיה היא מחלת הדדפולינג (deadpooling), המתבטאת בהרס של שכבת ה-max pooling ומעבר ישיר של המידע משכבת הקונבולוציה לשכבת הפרספטרון הבינארי. כיצד תתבטא הפרעה זו ביכולת זיהוי האובייקטים של החולים במחלה?
- ב. ת'ור, חולה בדדפולינג, הגיע לבדיקה נירולוגית. בבדיקה מציגים לת'ור שתי תמונות: באחת רואים ארנבון במרכז שדה הראייה ובשנייה רואים דביבון במרכז שדה הראייה. כיצד תיראה הפעילות בשכבת הקונבולוציה בכל אחת מהתמונות? הניחו כי יש 14 ניוונים בשכבת הקונבולוציה. התייחסו למצבים שבהם הדביבון מפנה בתמונה את צידו הצר ביותר או הרחב ביותר (האלכסון של המלבן).
- ג. האם על-ידי בחירת משקלים מתאימה בשכבת הפלט יכול ת'ור לעבור את הבדיקה ולהבחין בין ארנבון לדביבון במרכז שדה הראייה? אם כן, הציעו תבנית קשרים מתאימה עבור הפרספטרון הבינארי, אם לא – נמקו.

2. (20 נק') נתונה רשת עם כלל למידה שמבוסס על ירידה במורד הגרדיינט בתוספת איבר מומנטום:

$$\Delta W(n) = -\eta \frac{\partial \varepsilon(n)}{\partial W(n)} + \alpha \Delta W(n-1), \quad \alpha > 0$$

- א. פתרו את משוואת ההפרש שלעיל, וקבלו ביטוי עבור  $\Delta W(n)$  כפונקציה של  $\alpha$  ושל סדרת הערכים  $\frac{\partial \varepsilon(t)}{\partial W(t)}$  בזמנים  $t = 1$  עד  $t = n$ .
- ב. מהו התנאי על  $\alpha$  (תנאי הכרחי) כדי שגודל העדכון יתכנס? (הניחו שסדרת הערכים  $\left(\frac{\partial \varepsilon(t)}{\partial W(t)}\right)_{t=1}^{\infty}$  חסומה.)
- ג. מה קורה לביטוי שקיבלתם בסעיף א' כאשר ל- $\frac{\partial \varepsilon(t)}{\partial W(t)}$  יש אותו סימן באיטרציות עוקבות? מה קורה אם הסימן מתהפך באיטרציות עוקבות? כיצד תשפיע תשובתכם על קצב ההתכנסות למינימום המקומי?

3. (60 נק') ישמו רשת נוירונים הכוללת מספר שכבות, לצורך בעיית two spirals (ראו קבצי נתונים באתר הקורס).



בעיה זו עליכם לחזות האם נקודה שייכת לספירלה אחת או לספירלה השנייה. נתונות זוג קוארדינטות  $(x, y)$  ומשתנה בינארי (0 עבור ספירלה אחת, ו-1 עבור הספירלה השנייה). עליכם לאמן את הרשת בעזרת סט האימון ולבחון ארכיטקטורות ופרמטרים בעזרת סט הולידציה. עליכם להגיש קוד אשר יבחן על סט נתונים שלא הוצג בפניכם (test set). הפלט הנדרש עבור נקודה נתונה הוא הספירלה שאליה היא שייכת (0 או 1).

המלצה: תחילה בנו רשת לפתרון בעיה פשוטה יותר, כגון חיזוי פונקציית  $\sin$  או  $\cos$ . בבעיות מסוג זה קצת יותר פשוט לאתר בעיות בקוד.

- א. ממשו רשת נוירונים לפתרון הבעיה:
  - בחרו פונקציית מחיר מתאימה.
  - ישמו כלל למידה מסוג stochastic gradient descent.
  - נסו ארכיטקטורות שונות (מומלץ להשתמש ביותר משכבת ביניים אחת).
  - נסו מספר פונקציות אקטיבציה (sigmoid, tanh, ReLU, וכו').
  - בחנו מספר דרכים לאתחול המשקולות וקצב הלימוד (למשל Xavier's initialization).
- ב. נסו להחליף את כלל הלמידה בזה שמתואר בשאלה הקודמת (איבר מומנטום). האם יש הבדל בקצב ההתכנסות?
- ג. בנוסף: נסו כללי למידה נוספים, כגון adagrad או RMSprop.
- הוסיפו לכלל הלמידה רגולריזציות מסוגים שונים  $(L_1, L_2)$ . האם יש הבדל בתוצאות? מדוע?

הבהרה: לבסוף יש להגיש רק את הרשת שהגיעה לתוצאות הכי טובות, אך כל האלגוריתמים שבחנתם צריכים להופיע בקוד.