

פרויקט סיום קורס- מבוא לנוירו-חישוביות
מדעי המחשב, מסלול מדעי הנתונים, שנה ב', סמסטר א'
מרצה- מירב שקרון

הנחיות כלליות:

בפרויקט זה תדרשו ללמידה עצמית ולפיתוח קוד בפייטון עבור רשת נוירונים שאותה תבחרו. בשיעור האחרון של הסמסטר תדרשו להציג את עבודתכם בכיתה לכלל הסטודנטים.

- הגשה בזוגות
- קריטריונים לציון – הבנת עקרונות הרשת שנבחרה, דיוק המימוש, תפקוד הרשת, איכות הקוד, איכות המצגת ועוד.
- תכולת המצגת- מבנה הרשת, איך היא עובדת, אתגרי המימוש, תוצאות, מסקנות.
- אורך המצגת (כ- 10 דקות)

שלבי הפרויקט:

שלב א'- בחירת מודל הרשת

בשלב זה אתם נדרשים לבחור סוג של רשת וסוג של מודל שאותו תרצו לממש מבין רשימת האפשרויות:

1. רשתות סיווג – Hebb net
2. רשתות ללמידה אסוציאטיבית - BAM
3. רשתות מבוססות תחרות - Hamming Net

שלב ב'- חקר תיאורטי

בשלב זה אתם נדרשים ללמוד על הרשת אותה בחרתם:

1. לימדו את העקרונות התאורטיים של הרשת שבחרתם
2. הבינו את האלגוריתם ללמידה/אימון של הרשת.
3. סכמו יתרונות וחסרונות של הרשת.

שלב ג' – מימוש הרשת בפייטון

בשלב זה עליכם לממש את הרשת שבחרתם:

1. כתבו קוד מאפס שמממש את האלגוריתם של הרשת
2. אין להשתמש בספריות מוכנות ללמידה עמוקה כמו TensorFlow, PyTorch וכו'
3. ניתן להשתמש ב NumPy-לביצוע חישובים מטריציוניים.
4. ניתן להתשמש בספריות נוספות לצורך טיפול בדאטה ולפעולות נוספות אם יש צורך בכל מקרה לא להשתמש בספריות של למידת מכונה/ למידה עמוקה וכו'

שלב ד' – יצירת dataset ומשימות על הנתונים-

בשלב זה עליהם להזין dataset לתוכנית שלכם ובעזרתה לבחון את ביצועי הרשת. סוג הנתונים משתנה מעט לפי סוג הרשת שבחרתם:

1. רשתות סיווג –

- a. מטרת הרשת - ברשת Hebb net תצטרכו לממש רשת שמקבלת כקלט ציורים של תווי אותיות גדולות באנגלית בייצוג בינארי ותחזיר כפלט את הקבוצה אליה משתייכת האות.
 - b. יצירת הדאטה –
 - i. קלט (Input Vector): וקטור באורך 64 שמייצג ציור של אות באנגלית בפורמט בינארי. כל אות מיוצגת על ידי תמונה מוקטנת לגודל, כאשר כל פיקסל הוא ערך בינארי (0 או 1).
 - ii. פלט (Output Vector): וקטור באורך 3 שמייצג את הקבוצה שאליה שייכת האות:
 - $[1, 0, 0]$ עבור הקבוצה הראשונה (A-I).
 - $[0, 1, 0]$ עבור הקבוצה השנייה (J-R).
 - $[0, 0, 1]$ עבור הקבוצה השלישית (S-Z).
 - iii. לדוגמא: האות B שייכת לקבוצה הראשונה ולכן תחזיר $[1, 0, 0]$
 - c. משימות על הדאטה-
 - i. למידה - הרשת תלמד את הזוגות של תמונות האותיות וקבוצות הקטגוריה המתאימות להן.
 - ii. זיהוי קטגוריה - לאחר תהליך הלמידה, תנו לרשת קלט (וקטור בינארי שמייצג אות) ובדקו אם הרשת מזהה את הקטגוריה הנכונה.
 - iii. זיהוי עם רעש - צרו וקטורי קלט רועשים (הוספת שינויים קטנים לערכי הפיקסלים הבינאריים). בדקו אם הרשת מצליחה לזהות את הקטגוריה הנכונה גם בתנאים של רעש. (5% רעש)
 - iv. בדיקת עמידות לרעש ברמות שונות - הגדירו רמות רעש שונות בקלט (5%, 10%, 20%) ובדקו כיצד הן משפיעות על ביצועי הרשת. סכמו את תוצאות הניסוי והשוו בין רמות הרעש.
 - v. הכללת קטגוריות – תנו לרשת דגימות חדשות מתוך אותה קטגוריה שהיא לא ראתה קודם (למשל, אותיות חדשות מאותה קבוצה). בדקו אם הרשת מצליחה לזהות את הקטגוריה הנכונה על בסיס התבניות הכלליות שהיא למדה.

2. רשתות ללמידה אסוציאטיבית-

- a. מטרת הרשת - ברשת BAM תצטרכו לממש רשת שמקבלת כקלט ציורים של תווי אותיות גדולות באנגלית בייצוג בינארי ותחזיר כפלט את הקוד האסקי של האות הזאת.
 - b. יצירת הדאטה-

מבוא לנוירו חישוביות תשפ"ה

- i. קלט: (Input Vector) - וקטור באורך 64 שמייצג ציור של אות באנגלית בפורמט בינארי. כל אות מיוצגת על ידי תמונה מוקטנת לגודל, כאשר כל פיקסל הוא ערך בינארי (0 או 1).
- ii. פלט: (Output Vector) - וקטור באורך 7 שמייצג את קוד ה-ASCII של האות בפורמט בינארי.
- iii. לדוגמה: 'A' -> קוד 65 ASCII -> בינארי: [1, 0, 0, 0, 0, 0, 1].
- c. משימות על הדאטה-
 - i. למידה- הרשת תלמד את הזוגות של אות גדולה באגלית וקוד ASCII הבינארי שלו.
 - ii. שחזור- אם ניתן הקלט כקוד ASCII רועש, הרשת תצטרך לשחזר את התו המקורי. שנו את הדאטה כך שיכנס ערך רועש ובידקו האם הרשת ממשיכה בתהליך הזיהוי בצורה נכונה (5% רעש).
 - iii. בדיקת עמידות לרעש ברמות שונות - הגדירו רמות רעש שונות בקלט (5%, 10%, 20%) ובדקו כיצד הן משפיעות על ביצועי הרשת. סכמו את תוצאות הניסוי והשוו בין רמות הרעש.
 - iv. משימת התאמה דו-כיוונית - ניתן קלט של תו או קוד ASCII, והרשת תוכל להשלים את הפלט המתאים.

3. רשתות מבוססות תחרות-

- a. מטרת הרשת - ברשת Hamming Net תצטרכו לבצע משימות זיהוי תבניות עם רעש.
- b. יצירת הדאטה -
 - i. קלט (Input Vector): וקטור באורך 64 שמייצג ציור של אות באנגלית בפורמט בינארי. כל אות מיוצגת על ידי תמונה מוקטנת לגודל, כאשר כל פיקסל הוא ערך בינארי (0 או 1).
 - ii. פלט: (Output Vector) - וקטור יחידה באורך N כאשר N זה מס' האותיות שאוחסנו.
 - iii. לדוגמה: אם הזנו רק 3 אותיות לרשת- [1, 0, 0] עבור האות A, [0, 1, 0] עבור האות B, [0, 0, 1] עבור האות C.
- c. משימות על הדאטה-
 - i. למידה - הרשת תלמד את התבניות הידועות של האותיות, המיוצגות כוקטורים בינאריים, ותשמור אותן בזיכרון.
 - ii. זיהוי עם רעש- אם תינתן לרשת גרסה רועשת של אחת התבניות הידועות, הרשת תצטרך לזהות את התבנית הקרובה ביותר. שנו את הדאטה כך שיכנס ערך רועש ובידקו האם הרשת ממשיכה בתהליך הזיהוי בצורה נכונה (5% רעש)
 - iii. בדיקת עמידות לרעש ברמות שונות - הגדירו רמות רעש שונות בקלט (5%, 10%, 20%) ובדקו כיצד הן משפיעות על ביצועי הרשת. סכמו את תוצאות הניסוי והשוו בין רמות הרעש.

מבוא לנוירו חישוביות תשפ"ה

iv. זיהוי תבנית קרובה ביותר - הרשת תנסה לזהות את התבנית הקרובה ביותר מבין התבניות המאוחסנות, גם כאשר הקלט אינו תואם באופן מלא לאף תבנית.

בהצלחה!

מירב