

סיכום שיעור 72

למידה عمוקה והקדמה לפרשptron

Deep Learning and Perceptron Introduction

מרצה: ד"ר יורם סgal

Dr. Yoram Segal

מילות מפתח

בינה מלאכותית (AI) – **למידת מכונה** (Machine Learning) – **למידה عمוקה** (Deep Learning) – **פרשptron** (Perceptron) – **רשתות נוירוניים** (Neural Networks) – **רשות כונבולוציה** (CNN) – **רשות נשנות פרופגציה לאחור** (Back Propagation) – **רשתות טרנספורמרים** (Transformers) – **מנגנון תשומת לב** (Attention) – GPU – Hyperparameters

1 מבוא: היררכיית הבינה המלאכותית

בינה מלאכותית (AI - Artificial Intelligence) היא התחום הרחב ביותר, הכלול את כל הטכנולוגיות המאפשרות למחשבים לבצע משימות הדורשות אינטיגנציה. בתחום זה נמצאת **למידת מכונה** (Machine Learning), ובתוכה נמצאת **למידה عمוקה** (Deep Learning).

1.1 הקשר בין התחומיים

- **בינה מלאכותית (AI)**: התחום הכללי - כולל רובוטיקה, ראייה ממוחשבת, עיבוד שפה טבעית, חיפוש, תכנון ועוד
- **למידת מכונה (ML)**: תת-תחום של AI - אלגוריתמים הלומדים מנתונים, כולל סטטיסטיקה, גרפסיה, סיוג ועוד
- **למידה عمוקה (DL)**: תת-תחום של ML - מבוססת על רשות כונבולוציה (Convolutional Neural Networks) עם שכבות רבות

2 ארכיטקטורת הלמידה

2.1 משתנים בלמידה

בכל בעית למידה יש שלושה סוגי משתנים:

1. **משתנה הכניסה (X)**: הנתונים שאנו מזינים למודל
2. **משתנה היציאה (Y)**: התוצאה הרצואה (the Ground Truth)
3. **משתנה הפרמטרים (W)**: המשקولات שהמודל לומד

2.2 היפר-פרמטרים (Hyperparameters)

היפר-פרמטרים הם פרמטרים שאינם נלמדים אלא קבועים מראש על ידי המתכנת. הם קובעים את המבנה הכללי של המודל, כגון:

- מספר השכבות בראשת
- מספר הנוירונים בכל שכבה
- קצב הלמידה (Learning Rate)
- גודל האצווה (Batch Size)

3 היסטוריה של הפרספטرون והלמידה העמוקה

3.1 שנת 1958 – לידת הפרספטרון

פרנק רוזנבלט (Frank Rosenblatt) המציא את הפרספטרון, שהיה הצעד הראשון לקריאת הרשותות הנוירונליות. הפרספטרון היה מודל פשוט של נוירון מלאכותי שיכל לבצע סיווג ביןארי.

3.2 שנת 1969 – המשבר הראשון

מינסקי ופארט (Minsky and Papert) הראו שהפרספטרון לא יכול לפתור בעיות שאינן ניתנות להפרדה לינארית (כמו XOR). זה גרם לירידה חדה במימון ובמחקר בתחום.

3.3 שנת 1989 – התחייה הראשונה

יאן לכוון (Yann LeCun) הציג את רשותות הקונבולוציה (CNN) לזיהוי ספרות בכתב יד. זו הייתה פריצת דרך ממשמעותית.

3.4 שנת 2006 – הפריצה

ג'פרי הינטון (Geoffrey Hinton) הציג את רשות האמונה העמוקה (Deep Belief Network) והוכיח שניתן לאמן רשותות עמוקות בהצלחה באמצעות טכניקות חדשות.

3.5 שנת 2012 ואילך – עידן הלמידה העמוקה

מאז 2012, עם הזמינות של GPU חזקים וכמוויות גדולות של נתונים, הלמידה העמוקה הפכה לכלי המרכזי בבינה מלאכותית.

4 אלגוריתם הפרופנציה לאחר (Back Propagation)

4.1 עקרון הפעולה

הפרופנציה לאחר היא האלגוריתם המאפשר לאמן רשותות נוירוניים. העיקרונות:

1. מעבר קדימה (Forward Pass): חישוב הפלט בהתבסס על הכניסות והמשקילות
2. חישוב השגיאה: ההפרש בין הפלט בפועל לפלט הרצוי
3. מעבר לאחר (Backward Pass): חישוב הגרדיינטס לכל משקלות
4. עדכון המשקלות: שינוי המשקלות בכיוון שמקטין את השגיאה

4.2 גרדיאנט דיסנט (Gradient Descent)

שיטת ירידת הגרדיינט משמשת לעדכון המשקלות:

$$(1) \quad W_{new} = W_{old} - \eta \cdot \frac{\partial E}{\partial W}$$

כאשר η הוא קצב הלמידה ו- E היא פונקציית השגיאה.

5 סוגים רשותות נוירוניים

5.1 רשותות מחוברות לחלווטין (Fully Connected / MLP)

- כל נוירון בשכבה מחובר לכל הנוירונים בשכבה הבאה
- מתאימות לנ נתונים טבלאים
- מספר רב של פרמטרים

5.2 רשותות קונבולוציה (CNN - Convolutional Neural Networks)

- מותאמות לעיבוד תמונות
- משתמשות בפילטרים (גרעינים) המזהים דפוסים מקומיים
- מספר פרמטרים קטן יחסית
- יכולת לזהות אובייקטים ללא תלות במיקוםם בתמונה

5.3 רשותות נשנות (RNN - Recurrent Neural Networks)

- מותאמות לנ נתונים סדרתיים (טקסט, זמן)
- יש להן "זיכרון" המאפשר לזכור מידע מצעדים קודמים
- LSTM הוא וריאנט משופר עם יכולת לזכור מידע לטוח אורך

5.4 טרנספוררים (Transformers)

- ארכיטקטורה מודרנית המבוססת על מנגן תשומת לב (Attention)
- מאפשרים עיבוד מקבילייעיל
- הבסיס למודלים כמו GPT ו-BERT
- משמשים ב-NLP, ראייה ממוחשבת ועוד

6 מנגן תשומת הלב (Attention)

מנגן ה-Attention מאפשר למודל "לשדים לב" לחלקים רלוונטיים בקלט. במקומות העבודה את כל הקלט באופן שווה, המודל לומד לאילו חלקים לתת יותר משקל.

יתרונות:

- יכולת להתמקד במידע הרלוונטי
- עיבוד מקבילי (בנייה RNN-L)
- ביצועים מעולים במשימות NLP

7 חישבות ה-GPU בלמידה عمוקה

GPU (יחידת עיבוד גרפי) חיוני ללמידה عمוקה מכיוון ש:

- אפשר חישובים מקבילים מסיביים
 - מנצח את תהליך האימון פי עשרות עד מאות
 - ניתן להשתמש בשירותי ענן כמו Google Colab
- דוגמה:** אימון שלוקח שבועיים על CPU יכול לדרוש יומם אחד על GPU.

8 סיכום

הລמידה העמוקה היא כלי רב-עוצמה המאפשר פתרון בעיות מורכבות שלא ניתן היה לפתור בעבר. מהפרנספטרון הפשטוט של 1958, דרך רשותות הקונבולוציה, הרשותות הנשנות ועד הטרנספורמרים של היום – התחום התפתח באופן דרמטי. הבנת העקרונות הבסיסיים היא המפתח להבנת הטכנולוגיות המתקדמות ביותר כיום.