



המכללה האקדמית להנדסה ע"ש סמי שמעון



מדור בחינות ומערכת שעות

המחלקה להנדסת חשמל

30/01/25  
12:10-13:40

# למידה עמוקה

## בוחן

## דימה בחובסקי

תשפ"ה סמסטר א'

השאלון מכיל 4 עמודים (כולל עמוד זה)

בהצלחה!

=====

### שימוש במחשבוני

☐ ניתן להשתמש במחשבון

### חומר עזר

☐ ניתן להשתמש בחומר עזר/דף נוסחאות, כמפורט : 10 עמודים, שהם 5 דפים

### אחר / הערות

☐ הבוחן מכיל 11 שאלות, 10 נק' כל שאלה

1. מהן הסיבות לכך שלמידה עמוקה תפסה תאוצה בשנים האחרונות? (יש לבחור את התשובות הנכונות ביותר).

- ☐ זמינות של הרבה יותר כוח חישוב.
- ☐ רשתות עצביות הן תחום חדש לגמרי.
- ☐ זמינות של הרבה יותר נתונים.
- ☐ למידה עמוקה הביאה לשיפורים משמעותיים ביישומים חשובים כגון פרסום מקוון, זיהוי דיבור וזיהוי תמונות.

2. הסבר את המושג Logistic Loss.

נקרא גם cross-entropy loss :

- מדובר בסוג של פונקציית מחיר, כאשר במהלך האימון של המודל מנסים להגיע למינימום שלה.
- לתוצאה של פונקציית מחיר משמעות הסתברותית.

3. נניח שבנית רשת נוירונים (neural network). החלטת לאתחל את המשקלים (weights) וההטיות (biases) של הרשת ע"י אפסים. אילו מהמשפטים הבאים נכונים? (יש לבחור את התשובות הנכונות ביותר)

- ☐ כל נוירון (neuron) בשכבה החבוייה הראשונה יבצע את אותה החישוב. לכן, גם לאחר מספר איטרציות של gradient descent, כל נוירון בשכבה יבצע את אותו הדבר כמו שאר הנוירונים.
- ☐ כל נוירון בשכבה החבוייה הראשונה יבצע את אותה החישוב באיטרציה הראשונה. אבל לאחר איטרציה אחת של gradient descent, הם ילמדו לבצע דברים שונים כי "שברנו את הסימטריה" (broken symmetry).
- ☐ כל נוירון בשכבה החבוייה הראשונה יבצע את אותו הדבר, אך נוירונים בשכבות שונות יבצעו דברים שונים. לכן, השגנו "שבירת סימטריה" (symmetry breaking).
- ☐ הנוירונים בשכבה החבוייה הראשונה יבצעו חישובים שונים זה מזה כבר באיטרציה הראשונה, הפרמטרים שלהם ימשיכו להתפתח בדרכם.

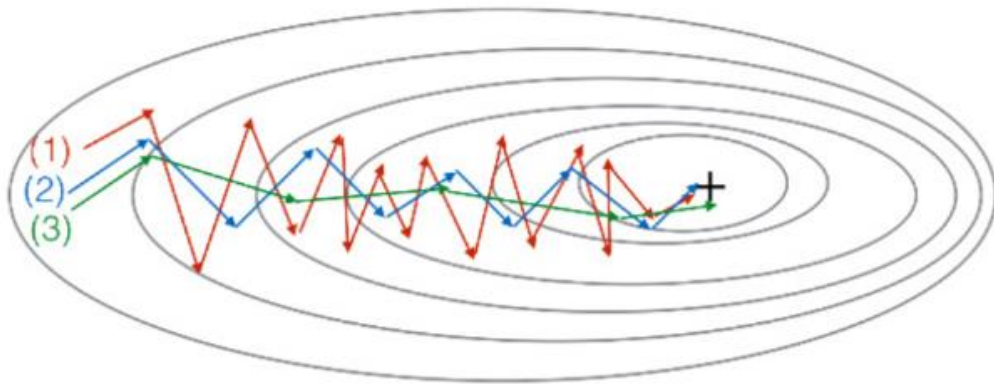
4. איזה מהמשפטים הבאים נכון? (יש לבחור תשובה אחת)

- a. השכבות העמוקות (deeper) יותר של רשת נוירונים בדרך כלל מחשבות מאפיינים (features) מורכבים יותר של הקלט, בהשוואה לשכבות המוקדמות (earlier).
- b. השכבות המוקדמות של רשת נוירונים בדרך כלל מחשבות מאפיינים מורכבים יותר של הקלט בהשוואה לשכבות העמוקות.
- c. השכבות העמוקות ברשת נוירונים אחראיות לחלץ תכונות פשוטות כגון קצוות (edges).
- d. רשתות נוירונים מחשבות תכונות באופן אחיד, ללא הבדלים בין השכבות.

5. אילו מהטכניקות הבאות מועילות להפחתת שונות (variance) ולהפחתת overfitting? (יש לבחור את התשובות הנכונות ביותר)

- Dropout** ☐
- L2 regularization** ☐
- Gradient Checking ☐
- Exploding gradient ☐
- Data augmentation** ☐
- Vanishing gradient ☐
- Xavier initialization ☐

6. התבונן בתרשים הבא :



הגרפים נוצרו באמצעות :

- a. Gradient descent
- b. Gradient descent עם  $\beta = 0.5$  momentum
- c. Gradient descent עם  $\beta = 0.9$  momentum

איזה עקומה מתאימה לכל שיטה? יש לרשום את התשובה הסופית בצורה של זוגות, כגון a1. יש לנמק בקצרה את התשובה.

**A1, b2, c3**

ללא momentum, הקפוצות משמעותיות ביותר.

תאורטית, ככל ש  $\beta$  גדול יותר כך קצב ההתקדמות גדול יותר כלומר קצב הלמידה מהיר יותר.

7. מה הבדל בין פרמטרים (parameters) של המודל לבין היפר-פרמטרים (hyper-parameters) (יש לבחור תשובה אחת) :

- a. הפרמטרים נבחרים על ידי המשתמש לפני האימון, בעוד שהיפר-פרמטרים נלמדים מתוך הנתונים במהלך האימון.
- b. הפרמטרים שולטים ישירות בתהליך האימון, בעוד שהיפר-פרמטרים הם המשקולות וההטיות (biases) הפנימיות של המודל.
- c. היפר-פרמטרים נבחרים על ידי המשתמש לפני האימון, בעוד שהפרמטרים נלמדים מתוך הנתונים במהלך האימון.



- d. אין הבדל משמעותי, שני המושגים מתייחסים לאותו הדבר.
8. הסבר אופן שימוש ב-Transfer Learning? (יש לבחור את התשובות הנכונות ביותר)
- ☐ יצירת מידע סינטטי עבור אימון המודלים גדולים.
  - ☐ הימנעות מ"דעות קודמות" של המודל.
  - ☐ הקטנת גודל המודל.
  - ☐ הקטנת זמן אימון.
  - ☐ הקטנת מספר משקלים שיש לאמן במודל.
  - ☐ שיטת אימון מודל מאפס.
  - ☐ שיטת אימון מודל על נתונים חדשים/נוספים.
9. יש להסביר את השיטה Dropout: מהו העיקרון ומהי המטרה המרכזית של השיטה עבור רשתות נוירונים. יש להמחיש את ההסבר בעזרת ציור מתאים.
- שיטת Dropout מונעת מרשת הנוירונים Overfitting. בכל epoch של המודל בזמן האימון מכבים באופן אקראי חלק מהנוירונים כך הרשת לומדת בצורה האוזנת יותר. בנוסף שיטה זאת עוזרת למודל ללמוד תבניות כלליות (שיפור generalization).**
10. מה ההבדל העיקרי בין Fine-Tuning ל-Transfer Learning? (יש לבחור תשובה אחת)
- a. Fine-Tuning משתמש בידע ממשימות קודמות, בעוד Transfer Learning מאמן מודל מאפס.
  - b. Fine-Tuning כולל עדכון "עדין" של המשקלים במודל, בעוד Transfer Learning שומר על הרוב המשקלים.
  - c. Transfer Learning מבצע אימון על כל שכבות המודל, בעוד Fine-Tuning משנה רק את השכבות האחרונות במודל, לרבות שכבת המוצא.
  - d. אין הבדל, שני המונחים מתארים את אותו תהליך בדיוק.
11. באילו מקרים עדיף להשתמש ב-CPU במקום להשתמש ב-GPU לצורך חישובים? (תנו לפחות שתי סיבות)
- 1) באלגוריתם שאינו ניתן לחישוב מקבלי כלומר לא ניתן לחלק למשימות מקבילות (דרך החישוב מתבצעת בצורה טורית).
  - 2) נפחי זיכרון קטנים, כאשר תקורת זמן בהעברת הנתונים ל-GPU וחזרה יכולה להיות משמעותית.
  - 3) בעיות בעלי סיבוכיות חישוב נמוכה יחסית.