

# מטלת מנחה (ממ"ן) 14

הקורס: 22961 למידה עמוקה

חומר הלימוד למטלה: יחידות 6-7

מספר השאלות: 4

סמסטר: 2022ב

משקל המטלה: 10 נקודות

מועד אחרון להגשה: 10.06.2022

## קיימות שתי חלופות להגשת מטלות:

- שליחת מטלות באמצעות מערכת המטלות המקוונת באתר הבית של הקורס
  - שליחת מטלות באמצעות הדואר או הגשה ישירה למנחה במפגשי ההנחיה
- הסבר מפורט ב"נוהל הגשת מטלות מנחה"

## חלק א' (20 נק')

ענו על אחת מהשאלות הבאות.

### שאלה 1

בשאלה זו עליכם לאמן רשת RNN לסיווג אוסף הנתונים SST-2 ולהתמודד בעצמכם עם בעיית הגרדיאנט המתפוצץ על ידי קיצוץ חישוב הגרדיאנט לאורך הזמן: עטפו את הרשת בפונקציה המאמנת אותה, אשר תקבל פרמטר  $t$  - זהו מספר צעדי הזמן המקסימלי שיש לפעפע את גרדיאנט השגיאה לאחר דרך התאים הנשנים.

לאחר כתיבת הפונקציה, אמנו את הרשת והראו שעבור קלט באורך  $T$  טוקנים הגרדיאנט אינו מפועפע מעבר למצב החבוי המחושב בזמן  $T-t$  וזאת על ידי איור גדלי הגרדיאנטים כפונקציה של מיקום הטוקן במשפט.

### שאלה 2

בשאלה זו עליכם לממש תא RNN בעל בקרת זרימה בסיסית. הנחיות:

- לתא יהיה מצב חבוי מועמד  $\hat{h}_t$  המחושב ראשית כמו בתא אלמן.
- בנוסף יהיה לתא "כפתור אתחול מחדש": מצב חבוי נוסף,  $R_t$  התלוי בקלט התא בזמן  $t$  בלבד.
- מצב התא המועבר ליחידת הזמן הבאה יהיה  $h_t = \hat{h}_t \cdot R_t$ .
- עליכם לתכנן את המצב החבוי הנוסף כך שבעזרתו התא יוכל "לשכוח" את המצב החבוי  $h_t$  בהנתן קלט מתאים. החליטו לבדכם מהי האקטיבציה המתאימה למשימה זו, וכיצד בדיוק  $R_t$  יחושב. נמקו שיקוליםכם בפירוט.
- אמנו רשת RNN המשתמשת בתא זה על מספר מועט של משפטים מאוסף נתונים לבחירתכם עד להתאמת יתר, בכדי לוודא שהוא מסוגל ללמוד.

## חלק ב' (80 נק')

ענו על אחת מהשאלות הבאות.

### שאלה 3

בשאלה זו עליכם לתכנן ולאמן מקודד עצמי עמוק כך שמספר השכבות במקודד זהה למספר השכבות במפרש, וכל שכבה במפרש קרובה ככל האפשר להיות הפונקציה ההפוכה של השכבה המקבילה לה במקודד. באיור:

אם המקודד הוא בעל הצורה הסכמטית הבאה,



אז המפרש המאומן צריך להיות קרוב ככל האפשר לצורה הבאה,



מובן שדבר זה לא תמיד אפשרי, אם למשל הפונקציות בהן אתם משתמשים במימוש הרשת אינן הפיכות, אך יש לתמרץ את הרשת להתקרב למבנה זה ככל האפשר.

- ממשו מקודד-מפרש מתאים בעזרת שכבות בעלות קישוריות מלאה.
- בחרו פונקציית מחיר מתאימה למשימה ואמנו את המודל על אוסף הנתונים MNIST.
- ציירו מספר תוצאות של הזנת הנתנים לרשת יחד עם הקלט המקורי.
- בדקו את התוצאות המתקבלות בכל שכבה: כמה קרובה שכבה נתונה במפרש להיות ההפכית של השכבה המקבילה לה במקודד?
- חזרו על הסעיפים הקודמים עבור ארכיטקטורה של רשת קונבולוציונית באופן מלא.
- נסו להסביר את ההבדלים המתקבלים בין שתי הרשתות שאימנתם.

### שאלה 4

בשאלו זו עליכם לתכנן ולאמן מקודד עצמי מנקה רעש לנתוני טקסט.

- הסבירו בפירוט מהי ארכיטקטורת הרשת הנבחרת, מהם הקלט והפלט של הרשת ומהי פונקציית המחיר המתאימה.
- בחרו סט נתונים מתאים לאימון הרשת למשימה זו, טענו אותו לזכרון ובצעו עיבוד מקדים מתאים.
- תארו את תהליך השחתת הנתונים, והסבירו למה בחרתם דווקא בו.
- אמנו את הרשת על batch קטן על מנת להראות שהמודל מסוגל ללמוד.
- בדקו את ביצועי הרשת על נתונים שהמודל לא ראה בזמן האימון והסבירו את התוצאות המתקבלות.
- השתמשו במודל הנלמד למטרת סיווג רגש על ידי העברת למידה מתאימה ודווחו על התוצאות המתקבלות. הסבירו שיקולים בעת תכנון העברת הלמידה.

יש להגיש בקובץ ZIP אחד :

1. מחברת ג'ופיטר המכילה את הקוד והפלט הדרוש לכל השאלות.
  2. **קישור לסרטון קצר בו אתם מסבירים את הקוד, מריצים אותו ומנתחים/מסבירים את הפלט** (את הסרטון יש להכין בעזרת תוכנת Zoom ושימוש בפונקציה Share screen).
  3. תשובות בכתב ככל שנדרש.
- מומלץ להשתמש במערכת הגשה האינטרנטית.**