הפקולטה למדעים מדויקים מועד א'

החוג לסטטיסטיקה וחקר ביצועים סמסטר א', תשפ"ב

אוניברסיטת תל אביב 25.1.2022

**למידה עמוקה**

מרצה: דר' שמעון שחר

כל חומר עזר דומם מותר בשימוש, אסור להתייעץ עם אנשים אחרים.

את הפתרונות יש לדחוס לקובץ zip יחיד, ולשלוח במייל אל monishahar@tauex.tau.ac.il.

שורת הכותרת של המייל צריכה להיות: Deep Learning Course Test <id number>.

לדוגמה: Deep Learning Course Test 024484214

שאלה 1 (50)

כאשר כותבים chatbot אחת המטלות העיקריות נקראת intent classification. במטלה זאת יש לבנות מודל שיקבל משפט או פסקה קצרה ויחזיר אחת מ-k אפשרויות מוגדרות מראש. למשל במערכת מענה קולי של רופא, האפשרויות אותן יש לזהות הן: "הזמנת תור", "קבלת אישור" או "אחר". אם הקלט הוא: "אני רוצה לקבוע תור לרופא ביום ראשון" ה- class שהמודל יחזיר הוא "הזמנת תור".

בלינק ל- Google Drive שמצא כאן, נתון dataset לאימון ובחינה של מודל ל- intent classification.

**שימו לב:** הקבצים בסיומת out שהועלו לדרייב אינם רלוונטיים לבעיה. בדרך כלל פותרים בנוסף ל- intent classification בעיה נוספת שקשורה להבנת המשפטים, וקבצים אליה רלוונטיים אליה.

<https://drive.google.com/drive/folders/1vtjtClS22qSQVuhSd4VoAb9_xfRPXs9x?usp=sharing>

1. כתוב מודל ל- intent classification אשר אינו משתמש ב- language model
2. כתוב מודל ל- intent classification אשר משתמש ב language model. מומלץ להשתמש ב- BERT אך לא חובה.

הקוד לשני המודלים שכתבתם צריך להריץ אימון, פרדיקציה, ולהדפיס את תוצאות ה- evaluation על ה- test data שנתון. במידה ואתם משתמשים בספריות (למשל ספריית transformers שהיא ספרייה מקובלת ל- language models) ציינו את זה, או כתבו קובץ requirements.txt.

שאלה 2 (30)

Conditional Random Field הוא מודל גרפי שנמצא בשימוש בבעיות machine vision. המודל מוגדר על ידי גרף שבו כל קודקוד מייצג משתנה מקרי Y שמאונדקס לפי אינדקס הקודקוד v, ואוסף תצפיות X. התכונה העיקרית של המודל היא שמתקיימת מרקוביות בשכנים שבגרף. כלומר:

P[Yv |X, Y\{Yv}] = P[Yv | X, {Yw; w neighbor of v}]. הגדרה מפורטת של המודל נמצאת למשל בוויקיפדיה: <https://en.wikipedia.org/wiki/Conditional_random_field>

בעיית ה- semantic segmentation מוגרת כבעיה שהקלט אליה הוא תמונה והפלט הוא label לכל פיקסל, כאשר ה- labels נתונים מתוך קבוצה דיסקרטית ידועה מראש (לדוגמה: אנשים, עצים ורקע).

1. הסבר עבור בעיית semantic segmentation מה מייצג כל משתנה Y? מה מייצגות התצפיות X?

המשתנה Y מייצג את המחלקות השונות אליהן ניתן לסווג את הפיקסלים. התצפיות X מייצגות את הפיצ'רים של כל פיקסל (המאפיינים של הפיקסל שבעזרתם נסווג אותו למחלקה ב-Y).

1. קרא את המאמר: Efficient Inference in Fully Connected CRFs with Gaussian Edge Potentials והסבר את המודל להסתברות המותנה ואת הקירוב שמשתמשים בו ללמידה (sections 2,3).

במאמר, המודל שתואר הינו Fully Connected CRF, כלומר, הוא מייצג כל פיקסל על ידי קודקוד בגרף. כל קודקוד, מחובר לכלל הקודקודים האחרים (לכל הפיקסלים האחרים), זאת על מנת לבטא קשרים יותר "ארוכי טווח" מאשר תלות אך ורק בפיצ'רים של הפיקסל הספציפי או רק בשכנים שלו.

1. קרא את המאמר: Conditional Random Fields as Recurrent Neural Networksוהסבר כיצד מקרבים את הבעיה כרשת נוירונים. בפרט: (1) תאר את ההבדל בין שכבת קונבולוציה רגילה לשכבה המתוארת במאמר ו- (2) תאר כיצד מתבצעת הגזירה (sections 3-5).

שאלה 3 (20)

בשאלה זאת מספר סעיפים אשר אינם קשורים בינהם.

1. הסבר כיצד ניתן להפוך autoencoder נתון ל- variational autoencoder? מהם השינויים בארכיטקטורה ובפונקצית המטרה. הנח שהרשת המקורית מיושמת באמצעות מספר שכבות fully connected.
2. האם אפשרי להשתמש בשכבות convolution ב- autoencoder, ואם כן מה נרצה להניח שהדאטה בקירוב מקיים על מנת שיהיה כדאי להשתמש בהן?
3. אותה שאלה כמו בסעיף ב' עבור LSTM.
4. נתון dataset שבו התצפיות מגיעות משתי התפלגויות רב-מימדיות שונות. כמו-כן נתון כי רק 0.001 מתוך התצפיות באות מההתפלגות השניה, אך לא ידוע אילו תצפיות מקורן בכל אחת מההתפלגויות. הצע דרך שתשתמש ב- autoencoders כדי לזהות את המקור לתצפיות והסבר מדוע אתה חושב שהיא תעבוד.

**רמז:** כדוגמה קונקרטית הנח שכל אחת מההתפלגויות מרוכזת מאוד סביב יריעה שונה במימד 2 והדאטה הנתון הוא במימד גבוה בהרבה. מה יהיה המימד הרצוי של ה- encoding?