*תרגיל 2 – אייל באן ונדב גת*

*הערה לתרגיל: מסמן מכפלה סקלרית או וקטורית רגילה (לפי ההקשר), מסמן מכפלה חיצונית.*

***שאלה 1***

1. *במקרה בו :*

*במקרה בו :*

סה"כ נקבל:

כעת נגזור כל איבר בסכום בנפרד עם חלוקה למקרים.

כאשר :

כאשר :

ובסה"כ קיבלנו:

מכך ש-y הוא one-hot vector. ולכן:

1. *נגדיר ונחשב:*

*כאשר מסמל את הערך בתא ה- במטריצה .*

*כעת נגדיר ומסעיף א' נקבל:*

*בנוסף נחשב בעזרת כלל השרשרת:*

*כאשר הנגזרת הראשונה היא באופן זהה לחישוב של , ואת הנגזרת של סיגמואיד חישבנו בתרגיל הקודם. עכשיו נוכל להיעזר בכלל השרשרת ולחשב את הנגזרת הרצויה:*

1. *נחשב את הנגזרות הנדרשות למימוש:*

**שאלה 2**

1. *נגדיר ואז נחשב לפי כלל השרשרת:*

*כאשר השתמשנו בנגזרת של סיגמואיד, אותה חישבנו בתרגיל הקודם.*

*כאשר חישובי הביניים זהים לשאלה 1.*

***שאלה 3***

*בשפות כמו עברית וערבית, שבהן יש לעתים מספר מיליות המרכיבות את אותה מילה, עשוי להיות יתרון למודל שפה מבוסס אותיות. זאת משום שלכל מילית יש משמעות של מילה נפרדת ולכן למעשה נתייחס לאותה המילה עם וריאציות שונות כמילים נפרדות, לדוגמה, המילה "אני" - ההטיות "ואני", "כשאני", יהיו מיוצגות כל אחת ע"י טוקן אחד ולכן "יאבדו" את המשמעות הסמנטית לעומת חלוקת המילים ל- "ו", "כש", "אני".*

*כמו כן, נאבד גם קשרים בין מילים כמו "כאשר" ו-"כש", "בתוך" ו-"ב".*

*חלוקה של המודל לאותיות מבטלת גם את הסיכון להיתקל במילה לא ידועה משום שכל תו שנקרא הוא בוודאות במילון שהוגדר.*

*מהצד השני, חלוקה של המודל למילים מאפשרות יותר אקספרסיביות במובן שהמילון הרבה יותר גדול ועל כן יש יותר "קשרים" בין מילים. מחלוקה כזו נקבל קשר סמנטי וסינטקטי בין מילים לעומת אותיות שלרוב אין משמעות סינטקטית או סינטקטית לשכיחות שהן מופיעות יחדיו במשפט. במשפט "הלכתי לבנק ומשכתי כסף" נקבל ככה"נ תלות יותר גדולה בין המופעים של "לבנק" ו"כסף" מאשר בין המופעים של האות ב' והאות ס'.*

***שאלה 4***

*נראה כי:*

*וקיבלנו את הדרוש. המעבר השני הוא מחוקי חזקות, השלישי הוא מחוקי לוגריתם והאחרון הוא מכך ש-.*