פרויקט במעבדה – סיכום פגישת התנעה, לשם ואופיר, 7.11.18 :

משימות:

1. להמשיך לקרוא מאמרים ותכנים רלוונטים (לשם כבר שלח לי)
2. לחלץ כ-20 משפטים מNUCLE
   1. לקבל את המאגר מלשם
   2. האם קיים איזשהו כלי פירסור נוח? או לפרסר בעצמי
   3. להחליט איך לסנן את המשפטים לאחר הגרלה רנדומאלית שלהם
   4. עם רפרנס או לא?
3. לפתוח משימת crowed source ב - Amazon Turk:
   1. לחפש פורמט DA (direct assessment) מתאים.
   2. להחליט כמה אנשים למשפט (3?).
   3. פילטורים רלוונטים: מעריכים מארה"ב, בעלי קרדיט רלוונטי לנושא
4. להריץ את הפלטפורמה על הפיילוט
   1. כמה צריך לשלם לאנשים (כמה זמן אנחנו מוכנים לחכות לתוצאות?)

בירוקרטיה:

1. לתאם ציפיות מול עמרי לגבי משך הפרויקט והיקפו (כמה זמן וכמה שעות בשבוע בערך)

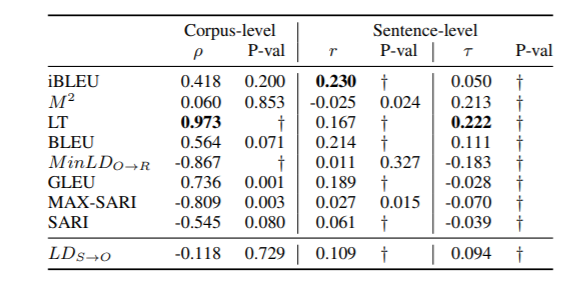
בתכל'ס – להמשיך לקרוא וללמוד, בתחילת שבוע הבא אתחיל להתעסק בחילוץ המשפטים, וברקע ללמוד על Amazon Turk כדי להבין מה נדרש מהחילוץ (אם או בלי רפרנס? אם עם, איזה? עוד שאלות?)

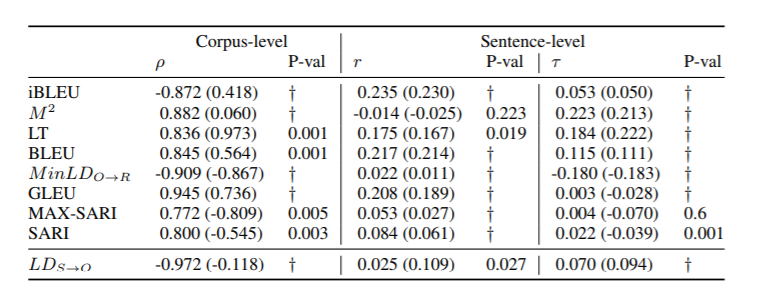
שאלות:

1. במאמר שלכם מוזכרות שיטות הערכה רבות לטיב התרגום. באילו מהן כדאי להתמקד ולהרחיב? ההסברים היו קצרים וכללו מושגים שלא הבנתי, אשמח להרחבה בנוגע לשיטות הרלוונטיות.

מושגים בסיסיים שחסרים לי:

* 1. lattice (סריג) שעליו מתבססת ההערכה.

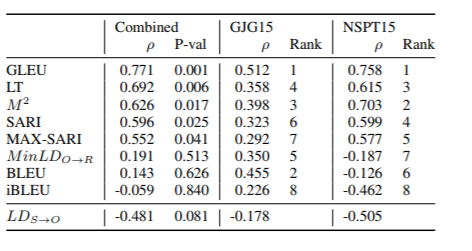
1. בניסיון להבין איזה שיטה יותר רלוונטית (כלומר המשך המאמר), נתקלתי בפערים במונחים סטטיסטים. 



קראתי קצת על מתאם ספירמן ועל P-VAL אבל יצאתי מבולבל, כי חשבתי שP-VAL נמוך מעיד על מהימנות נמוכה לתוצאות, אבל ניכר שהP-VAL פה נמוך במקרים רבים.

גם פה אשמח לסיוע בקריאת התוצאות, המדדים השונים וחשיבותם.

1. MAEGE



- איך בדיוק חישבתם את העמודה השמאלית?

לשם השוואה - לקחתם רק משפטים שדורגו על ידי אנשים בשתי שיטות המדידה (GJG15, NSPT15)? אם כן, יש מספיק כאלה, שניהם עבדו על מאגרים גדולים (חלק מNUCLE?)?

1. Corpus-level Analysis - הבנתי שיצרתם עץ של כל התיקונים האפשריים למשפט, אך לא ממש הבנתי איך אח"כ אתם משווים את המערכות השונות ומדרגים אותן בעזרת העץ הזה. אשמח לעזרתך בניסיון להבין זאת טוב יותר.