**פרויקט גמר -‘Semantle’ AI**

חברי צוות- יובל היטר, בר דמרי, נסים חי אפריאט.

המחלקה להנדסת תוכנה, אוניברסיטת בן גוריון.

**המטרה-**

Table

Description automatically generatedליצור אלגוריתמם חכם, אשר מביס את משחק ניחושי המילים "סמנתעל" ומוצא את המילה הנכונה במספר ניחושים מינימלי ככל הניתן.

**שיטת עבודה-**

שימוש במודלים של מילים ואלגוריתמים מתמטיים לחיזוי מילה על פי מרחק סמנטי באמצעות ניחושים. \

**שלבי עבודה:**

1. יצירת סביבה מלאה ועובדת שאפשר להריץ בה ניסויים ולקבל דוחות csv וגרפים.
2. ניצור [3 וריאנטים](#_שלושת_הvariants:) שונים ועבור כל אחד מהם נגבש פיתרון טריוויאלי אחד ולפחות פיתרון אינטלגנטי אחד, עבור כל וריאנט יש מנחה משחק.
3. נריץ בדיקות דומות עבור כל אחד מן הוריאנטים, נגבש תוצאות סביב איכות פעולותיהם, קצב ההתקרבות, מספר הניחושים, זמן הריצה וכדומה בתצורה גרפית וטבלאית (csv).
4. נגבש מסקנות מן תוצאות הניסויים, נקדם את הפרויקט בתחרות ונגבש מאמר מחקרי.

**התנהלות,דגשים ועקרונות מנחים לעבודה:**

* פגישות אחת לשבועיים/חודשיות לחיתוכים והצגת התקדמות.
* יצירת תוכנית עבודה שנתית- בוצעה.
* כתיבת קוד בפייתון.
* שימוש ב-GitHub/GitLAB
* לשחק במשחק ברשת- ללמוד כללים.
* Word embedding - ללמוד עצמאית.
* קוד ברור קריא ניתן להרחבה והרצה, שימושי, אחרי זה למחקר.
* להעמיד את הסביבה כמה שיותר מהר.
* להשאיר זמן לפיתוח הAI ולעשות כל מיני ניסויים עד שנצליח להשיג תוצאות טובות.
* בשלב ראשון משחקים מול ה-offline לאחר יצירת כל וריאנט. uאחרי שעושים את כל המשחק לוריאנט הראשון\שני לשחק כנגד המנחה אונליין כדי לראות עד כמה מצליחים לצמצם את מספר הניחושים.
* דוחות גרפים \*\* טרם ידוע בשלב זה מה נדרש.

**תוצר ראשוני רצוי-**

Pipeline ראשוני מלא מהתחלה ועד הסוף לאחד הוריאנטים כולל תוצאות כולל הכל.

**דוחות ומעקבים -**

הדוחות – 5 אלגוריתמים בסך הכל. מריצים עבור כל אחד מהם מאתיים פעם לדוגמא עם מילה סודית שונה. תוך כמה זמן, קצב התקרבות וכדומה.

Log עבור כל משחק – אופציה רלוונטית.

**נספחים:**

# שלושת הvariants:

* **הוריאנט הראשון –** למנחה ולשחקן יש אותו מודל בדיוק (word embedding) והשחקן לא יודע מה המילה הסודית שהמנחה חושב עליה, צריך למצוא אותה. הבעיה זו היא בעיה טרילטלציה\ טרינגלציה. מספר הניחושים כמספר המימדים במודל ועוד 1. המטרה של האלגוריתם האינטלגנטי זה לתת פחות ניחושים.
* **הוריאנט השני-** למנחה ולשחקן יש מודלים שונים, עם מספר מימדים ייתכן שונה, אבל שניהם מודלים סמנטים.

אז המרחקים שהמנחה אומר הם בדיוק אותם מרחקים כפי שהשחקן רואה במודל שלו. (עדיין צריך להתמצא).

\*\* נגבש טריוואלי בהמשך.

* **הוריאנט השלישי –** השחקן לא מחזיק מודל שפה בזיכרון שלו. אבל, המנחה מספק לו עוד שאילתא אחת חוץ ממרחק (השחקן יכול לשאול את המנחה " תן לי בבקשה את k המילים הקרובות ביותר למילה כלשהי נניח בית)

ואז המנחה ומסתכל במודל שלו ומחזיר לו אותן.

\*\* המודל הטריוויאלי pure euristic search - ללכת תמיד לקודקוד שנראה הכי קרוב ליעד ולהסתכל מי חברים שלו.