

מטלה 3

שם קוד הקבוצה- אריאל נפדנסקי ואבי פרדמן

ת.ז חברי הקבוצה-

1. 313306078 אריאל נפדנסקי

2. 316420132 אבי פרדמן

קישור לפרויקט -

<https://github.com/aviferdman/LLMs-and-Multi-Agent-Orchestration---Assignment3.git>

ציון מוערך- 95 😊

טופס הגשת הערכה עצמית

שם הסטודנט/ים: אריאל נפדנסקי ואבי פרדמן

שם הפרויקט: Turing Machine

תאריך הגשה: 27.11.2025

הציון העצמי שלי: 95/100

הצדקה להערכה העצמית (חובה - 200-500 מילים)

בסעיף זה, הסבירו למה בחרתם בציון זה. כללו:

נקודות חוזק: מה עשיתם במיוחד טוב? אילו מרכיבים הם ברמה גבוהה?

נקודות חולשה: מה חסר או יכול היה להיות טוב יותר? (כנות מוערכת!)

השקעה: כמה זמן ומאמץ השקעתם?

חדשנות: האם יש משהו ייחודי או מיוחד בעבודה?

למידה: מה למדתם מהפרויקט?

נקודות חוזק:

הפרויקט שלנו מציג איכות יוצאת דופן בכל ממדי ההערכה. הגענו לניסוי אמפירי מקיף של 21 משפטים

עם 7 רמות שיבוש (20%-50%), כולל 63 פעולות תרגום ואימות 100% (21/21 משפטים עברו בדיקת

אימות עם סטייה ממוצעת של 1.5% בלבד). התיעוד שלנו הוא ברמת Production מלאה: יצרנו מסמך

PRD מלא הכולל יעדים ומדדי הצלחה; תיעוד ארכיטקטורה מפורט עם 3 מסמכי החלטות

ארכיטקטוניות (ADRs) המסבירים את בחירת המודל, עיצוב שרשרת התרגום, ואסטרטגיית

Multi-Agent; מתודולוגיית מחקר מפורטת הכוללת אימות שיטתי של שיעורי טעויות; וניתוח כמותי

ואיכותני מקיף. המערכת בנויה על ארכיטקטורת Multi-Agent מודולרית עם תקשורת מבוססת-קבצים, הפרדת אחריות ברורה (Claude לתרגום ויצירת משפטים, Python לחישובים מתמטיים), ומנגנוני הרחבה המאפשרים שינוי שפות ושרשראות תרגום בקלות. תיעדנו את כל ה-Prompts בספר הפרומפטים, ביצענו ניתוח עלויות מפורט, והכנו את הפרויקט כקוד פתוח עם רישיון Contribution-Guidelines.

נקודות חולשה:

למרות שהפרויקט עומד בכל הדרישות, אנו מזהים מספר נקודות לשיפור. הניסוי התמקד בשרשרת תרגום אחת (אנגלית→צרפתית→איטלקית→אנגלית) - הרחבה לשרשראות נוספות הייתה מחזקת את הממצאים. מערכת ה-SSL certificate verification חוסמת כעת הרצה מחדש ברשתות ארגוניות - ניתן היה להוסיף fallback mechanism או הורדה מקומית של המודל. בנוסף, אף שתיעדנו 21 משפטים, מדגם גדול יותר (100-50) היה מספק סטטיסטיקה חזקה יותר. הפרדת האחריות (Claude vs Python) אופטימלית מבחינה עיצובית, אך יוצרת תלות ב-API של Claude לתרגומים.

השקעה:

ההשקעה בפרויקט הייתה משמעותית וכללה תכנון מערכת Multi-Agent, מימוש שלושה סוכני תרגום אוטונומיים, יצירה ידנית של 21 משפטים מגוונים עם הזרקת טעויות בשיעורים מדויקים, ביצוע 63 פעולות תרגום, אימות שיטתי של כל משפט, חישוב 21 מרחקים סמנטיים, ויצירת ויזואליזציה 4-פאנל מקצועית. השקענו במתודולוגיה מחקרית קפדנית, תיעוד מקיף, וביקורות איכות מתמשכות.

חדשנות:

הפרויקט מציג חידושים משמעותיים. גילינו תבנית לא-ליניארית מפתיעה - שיא הסטייה הסמנטית ב-30% (שיבוש לא ב-150%), המעיד על יכולת תיקון שגיאות מובנית ב-LLMs. הוכחנו שתוכן/דומיין משפיע יותר משיעור השיבוש (תזמורת סימפונית ב-30% שיבוש: 0.824 סטייה; רומן היסטורי ב-40%: 0.295 סטייה). יישמנו ארכיטקטורת Multi-Agent אמיתית עם תקשורת מבוססת קבצים, המדגימה Orchestration אוטונומי. הפרדנו ביצירה אוטומטית של טעויות ממודלים סטטיסטיים - Claude יוצר טעויות קוגניטיביות (substitution, deletion, transposition) ולא רנדומליות.

למידה:

הפרויקט העניק לנו הבנה מעמיקה ב-Multi-Agent Orchestration, תקשורת בין-סוכנית מבוססת-קבצים, ומדידה סמנטית עם Embeddings. למדנו כיצד LLMs מתפקדים כ"מתקני שגיאות" סמויים, מה ההבדל בין Semantic Similarity ל-Lexical Similarity, וכיצד עיצוב ניסויים מבוקר משפיע על תקפות ממצאים. רכשנו ניסיון במתודולוגיות מחקר NLP, ניתוח סטטיסטי (Pearson correlation: -0.046!), וויזואליזציה מדעית. הלקח המשמעותי ביותר: תוצאות מחקר יכולות להפתיע - הציפיות שלנו

(ליניאריות) נסתרו, והממצאים (U-shaped curve) חשפו התנהגות מורכבת של מערכות AI. למדנו חשיבות של אימות קפדני, תיעוד מעמיק, והכנה לשכפול ניסויים.

רמת הדקדקנות המבוקשת בבדיקה

על פי הציון העצמי שנתתי, אני מבין/ה שרמת הבדיקה תהיה:
90-100: דקדקנית ביותר - חיפוש "פילים בקנה", הקפדה על כל פרט

מאפיינים:

- ☒ קוד ברמת production עם hooks, extensibility, וארכיטקטורת plugins
- ☒ תיעוד מושלם ומפורט בכל היבט: PRD מקיף, תיעוד ארכיטקטורה מלא, README מקצועי
- ☒ עמידה מלאה בתקן ISO/IEC 25010
- ☒ בדיקות מקיפות עם כיסוי 85%+, edge cases מתועדים ומטופלים
- ☒ מחקר מעמיק: ניתוח רגישות שיטתי, הוכחות מתמטיות, השוואה מבוססת נתונים
- ☒ ויזואליזציה ברמה גבוהה ביותר עם dashboard אינטראקטיבי
- ☒ ספר פרומפטים מפורט ומתועד
- ☒ ניתוח עלויות מקיף עם המלצות אופטימיזציה
- ☒ חדשנות וייחודיות: רעיונות מקוריים, פתרון לבעיה מורכבת
- ☒ תרומה לקהילה: קוד פתוח, תיעוד לשימוש חוזר

הצהרת יושר אקדמי (Academic Integrity Declaration)

אני מצהיר/ה בזאת ש:

- ☒ ההערכה העצמית שלי היא כנה ואמיתית
- ☒ בדקתי את העבודה מול כל הקריטריונים לפני קביעת הציון
- ☒ אני מודע/ת שציון עצמי גבוה יוביל לבדיקה דקדקנית יותר
- ☒ אני מקבל/ת את העובדה שהציון הסופי עשוי להיות שונה מהציון העצמי
- ☒ העבודה היא פרי עבודתי/נו (של הקבוצה) ואני/אחראי/ם לכל תוכנה

חתימה: אריאל נפדונסקי ואבי פרדמן, תאריך: 09.11.2025

