

מטלה 5 - אפשרות מס' 1

שם קוד הקבוצה- אריאלנפדןסקיאביפרדרמן

ת.ז חברה הקבוצה-

. 1. 313306078 אריאל נפדןסק'

. 2. 316420132 אבי פרדמן

קישור לפרויקט -

<https://github.com/aviferdman/LLMs-and-Multi-Agent-Orchestration---Assignment5.git>

ציון מוערך- 89 😊

טופס הגשת הערכה עצמית

שם הסטודנט/**ים**: אריאל נפדןסק' ואבי פרדמן

שם הפרויקט: מעבדה ראשונה - context window

תאריך הגשה: 10.12.2025

הציון העצמי שלי: 100/89

הצדקה להערכת העצמית (חובה - 200-500 מילים)

בסעיף זה, הסבירו למה בחרתם בציון זה. כללו:

נקודות חזק: מה עשיתם במיוחד טוב? אילו מרכיבים הם בrama גבוהה?

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

השערה: כמה זמן ומאזן השקעתם?

חדשנות: האם יש משהו ייחודי או מיוחד בעבודה?

למידה: מה למדת מהפרויקט?

נקודות חזק

הפרויקט מציג יכולות גבוהות במספר ממדים מרכזיים. מבחינה טכנית, בניית ארכיטקטורה מודולרית המבוססת על Strategy Pattern ו-Abstract Base Classes, המאפשרת הרחבה קלה. הקוד כולל 181 בדיקות עם כיסוי של 93.69% – גבוהה משמעותית מהיעד של .85%.

מבחינות תייעוד, נוצרו שבעה מסמכי תייעוד עיקריים עם למעלה מ-5,000 שורות, כולל PRD מקיף (על 2,268 שורות), תייעוד ארכיטקטורה מלא (1,077 שורות), וניתוח סטטיסטי מפורט (על 650 שורות)

הכול נסוחאות LaTeX. המחבר מבוסס על 220 שאלות ו-15 אמירות שבוצעו בארכעה ניסויים מתוכננים היטב, עם ניתוח סטטיסטי שיטתי (ANOVA, רגסיה, effect sizes) ומודלים מתמטיים.

הויזואלייזציות כוללות 13 גרפים באיכות פרטום (DPI 300) עם אנוטציות סטטיסטיות.

הפרויקט עומד בתקני IEC 25010 ISO/BRC בכל הקטגוריות (Performance, Reliability, Usability, Maintainability, Security).

נקודות חולשה

לצד האיכות הגבוהה, ישנו מספר תחומי הדורשים שיופיעו. ראשית, Dashboard אינטראקטיבי – הנדרש ב-PRD – לא יושם. שנית, ספר פרומפטים לא תועד במסמך נפרד, אף שהפרומפטים קיימים בקוד. שלישית, ניתוח העליות מופיע ב-PRD אך לא בוצע בפועל עם נתונים אמיתיים.

בנוסף, הפרויקט מוגבל למודל אחד (2 llama). אף שהארקיטקטורה תומכת במספר ספקים, לא בוצעו ניסויים עם מודלים נוספים. לבסוף, מספר יכולות מתקדמות לא יושמו, כגון Kaplan-Meier, ויזואלייזציות תלת-ממדיות וアニメציות.

השקעה

השעתי בין 80 ל-100 שעות בפרויקט, כולל:

תכנון ותיעוד (25 שעות), פיתוח קוד (35 שעות), ביצוע 220 שאלות (15 שעות), ניתוח סטטיסטי (10 שעות), ויזואלייזציה (8 שעות), בדיקות (12 שעות) ו-5 polish שעות).

ה השקעה הייתה משמעותית ורחבת בכל היבטי המערכת, עם תשומת לב גבוהה לפרטים ולאיכות המחבר.

חדשנות

הפרויקט כולל מספר תגליות מקוריות: זיהוי Performance Cliff אמפירי ב-2,500 tokens עם מודל מתמטי ($R^2 = 0.42$), הפרקת "Lost in the Middle" בקנה מידה קטן, השוואת שיטות של Context Management Strategies (יתרונות של 7%), והערכת כמותית של Semantic Strategies.

הפתרון מהווה מערכת מחקר אמפירית מלאה, reproducible, עם שימוש חוזר framework.

למידה

העמקתי בהבנת windows context MLL, מתודולוגית מחקר הכללת ניתוח השערות וסטטיסטייה,
ישום RAG pipeline עקרונות הנדסת תוכנה, יצירת ייזואלייזציות באיכות פרטום. התובנה המרכזית:
מחקר אמפירי איות דרוש תכנון קפדי, ניתוח סטטיסטי מדויק ותיעוד מקיף לצורכי reproducibility.

סיכום

הפרויקט עומד בכל הדרישות העיקריות ואף מעבר לכך, עם תיעוד יוצא דופן, קוד איות ומחבר
מעמיך. הצין 89 משקף איות גבוהה מאוד לצד מספר תחומים לשיפור (Dashboard, 59 פרוgettiים,
ישום ניתוח עלויות, השוואה בין מודלים).

הצהרת יושר אקדמי (Academic Integrity Declaration)

אני מצהיר/ה בזאת ש:

- ההערכה העצמית שלי היא כנה ואמיתית
- בדקתי את העבודה מול כל הクリיטריונים לפני קביעת הצין
- אני מודע/ת שציון עצמי גבוה יוביל לבדיקה נוספת
- אני מקבל/ת את העבודה שהצין הסופי עשי להזות שזו מהציון העצמי
- העבודה היא פרי עבודת/נו (של הקבוצה) ואני/ אנחנו/ם לכל תופעה

חתימה: אריאל גולדנסקי ואבי פרידמן תאריך: 10.12.2025

