

מטלה 6

שם קוד הקבוצה- אריאלנפדןסקיאביפרדמן

ת.ז חברה הקבוצה-

. 1. 313306078 אריאל נפדןסק'

. 2. 316420132 אבי פרדמן

קישור לפרויקט -

<https://github.com/ArielleNapa/LLMs-and-Multi-Agent-Orchestration---Assignment6.git>

ציון מוערך- 85 😊

טופס הגשת הערכה עצמית

שם הסטודנט/ים: אריאל נפדןסק' ואבי פרדמן

שם הפרויקט: Prompt Engineering for mass production

תאריך הגשה: 16.12.2025

הציון העצמי שלי: 85/100

הצדקה להערכת העצמית (חוובה - 200-500 מילים)

בסעיף זה, הסבירו למה בחרתם בציון זה. כללו:

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

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 למדתם מהפרויקט?

לאחר בחינה ביקורתית, שיטתיות ועמוקה של הפרויקט ביחס לקריטריונים שנקבעו, אנו מעריכים כי

הוא ראוי לציון 100/85. ציון זה משקף עבודה אקדמית אינטלקטואלית, רצינית ומושקעת, העומדת בrama

גבואה ברוב היבטי הערכה, תוך מודעות מלאה לנקודות בהן קיימים מקום לשיפור ולהעמקה נוספת.

נקודות חזק

הפרויקט מציג מסגרת מחקר מקיפה לאופטימיזציה של prompts למודלי שפה גדולים, עם עמידה מלאה בעקרונות של מחקר אקדמי שיטתי. היקף ואיכות התיעוד בולטים במיוחד: למעלה מ-16 מומכים מפורטים (מעל 12,000 מילימ'), הכוללים תיעוד ארכיטקטוני מלא, מסמך יסודות מתמטיים (MATHEMATICAL FOUNDATIONS.md), ו-24 אסמכתאות אקדמיות בפורמט APA.

איכות הקוד עומדת בסטנדרטים תעשייתיים גבוהים: 21 מודולים ממוקדים (כולם קיצרים מ-150 שורות), 264 בדיקות אוטומטיות עם כיסוי של 88%, שימוש עקבי ב-type hints, docstrings מלאים בסגנון NumPy/Google לכל פונקציה ציבורית. התוכן הניסויי קפדי וינוון לרבייה: שבע רמות אופטימיזציה, שלוש ריצות לכל שיטה (סה"כ 100,510 פריטים), שימוש ב-random seeds קבועים, וניתות סטטיסטי מקיף הכלול t-tests עם תיקון Bonferroni, Cohen's ANOVA, ורוחוי סמרק של 95% הוויזואלייזציות ברמה מקצועית, עם גרפים ברזולוציה גבוהה (300 DPI), פלטת צבעים נגישה (Okabe-Ito) וסימון מובהקות סטטיסטיות.

נקודות חולשה

לצד החזקות, קיימים מספר תחומים בהם ניתן היה להעמיק. שלב 2 Phase, הכלל עבודה מול API אמיתי, הושלם אך בהיקף מוגבל בשל מגבלות תקציב (10-\$–15), דבר שמצמצם את היקף האימונות האמפירית של השערות שנבנו בסימולטור. כמו כן, חלק מהמסמכים המרכזים (למשל ARCHITECTURE.md) ארוכים במיוחד, דבר המקשה על ניוט מהיר. כיסוי הבדיקות, על אף היותו גבוה, אינו מלא בכל המודולים – רכיבים מסוימים כגון (CLI) 82% ו-(ExperimentRunner) 83% נמצאים מעט מתחת ליעד האידיאלי של 85%. בנוסף, המחקר מתמקד בעיקר במידדי accuracy, בעוד שמדדים مثلיים כגון factual consistency, calibration או robustness, נבדות אלן משקפות בחירות מודעות במסגרת מגבלות זמן ומשאבים, ולא ליקויים מהותיים באיכות העבודה.

השquaה ומאמץ

הפרויקט דרש השקעת זמן ומאזן ממשמעותיים – כ-80 שעות עבודה אפקטיבית לאורך תשעה שבועות. העבודה כללה שלב תכנון מפורט (PRD-PLANNING.md), ייצור datasets מאומתים (650 פריטים), פיתוח simulator מבוסס ספרות מחקרית, ביצוע ניסויים שיטתיים, ניתוח סטטיסטי عميق ו-refactoring מקיף לumedה מלאה בדרישות המבניות. התמודדנו עם אTEGRIM טכניים מורכבים, בהם ניהול refactoring מקיף לumedה מלאה בדרישות המבניות. התמודדנו עם אTEGRIM טכניים מורכבים, בהם ניהול caching וניהול מנגנון ההרצה. הטענת שיטות סטטיסטיות מתקדמות, ופיתוח מנגנון caching והרצה מקובלית שהניבו שיפור ביצועים של פי 2–4. התהילה היה מסודר, מתווך ומבוקר, עם אבני דרך ברורות או אינטראקצייתם בחלק שלב.

חדשנות וייחודיות

הפרויקט כולל מספר אלמנטים ייחודיים: **גש 1 Mock-first (Phase 1 בעלות אפסית)** המאפשרת איטרציות מהירות לפני שימוש ב-API אמיתי; ארכיטקטורה מבוססת ADRs עם תיעוד שיטתי של החלטות עיצוב; simulator מכיל למחקר המבוסס על ספורות אקדמית (למשל Wei et al., 2022; Brown et al., 2020) ועמידה מלאה בטkn 0 IEC 25010 ISO לאיכות תוכנה, כולל היבטי אבטחה וביצועים ברמה ארגונית. אף שאין מדובר בפריצת דרך מחקרית, אלמנטים אלו מעידים על חשיבה מתודולוגית עצמאית ושאייפה לעמידה בסטנדרטים מתקיימים גבוהים מעבר לדרישות הבסיסיות של הקורס.

למידה והפנמה

הפרויקט תרם למידה עמוקה הן בrama הטכנית והן ברמה המחברית. העמקנו בהבנת מתודולוגיות מחקר אמפיריות, חשיבות קבוצות ביקורת, ניתוח מובהקות סטטיסטי, תיקון לשוואות רבות ופרשנות של גDAL אפקט. במקביל, יישמו עקרונות מתקדמים בהנדסת תוכנה כגון Single Responsibility Principle, תכנון מודולרי-Test-Driven Development, והפנימו את ערכו של תיעוד חלק אינטגרלי מהתהילך הפיתוח. לבסוף, פיתחנו הרצה מעשית של האיזון הנדרש בין שאייפה לשלמות ובין מגבלות ריאליות של זמן ותקציב, לצד יכולת הערכה עצמית ביקורתית – מיומנות מרכזית בהפתחות אקדמית ומתקצועית.

הצהרת יושר אקדמי (Academic Integrity Declaration)

אני מצהיר/ה בזאת ש:

- הערכתה העצמית של היא כנה ואמיתית**
 - בדקה את העובדה מול כל החריטריזמים לפני קביעת ההחלטה**
 - אני מודע/ת שציון עצמי גבוה יוביל לבדיקה הדקדקנית יותר**
 - אני מקבל/ת את העובדה שהציון הסופי עשוי להיות שונה מהציון העצמי**
 - העובדת היא פרט בעובדתי/נו (של הקבוצה) ואני/נו אחראים/יים לכל תוצאה**

חתימה: אריאל נפנסקי ואבי פרדמן תאריך: 16.12.2025

