

מטלה 6

שם קוד הקבוצה - אריאל נפדנסקי ואבי פרדמן

ת.ז חברי הקבוצה -

1. 313306078 אריאל נפדנסקי

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 נפדנסקי ואבי פרדמן

שם הפרויקט: Prompt Engineering for mass production

תאריך הגשה: 16.12.2025

הציון העצמי שלי: 85/100

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 "יחודי" או מיוחד בעבודה?

למידה: מה למדתם מהפרויקט?

לאחר בחינה ביקורתית, שיטתית ומעמיקה של הפרויקט ביחס לקריטריונים שנקבעו, אנו מעריכים כי

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

גבוהה ברוב היבטי ההערכה, תוך מודעות מלאה לנקודות בהן קיימים מקום לשיפור ולהעמקה נוספת.

נקודות חוזק

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 מסמכים מפורטים (מעל 12,000 מילים), הכוללים תיעוד ארכיטקטוני מלא, מסמך יסודות מתמטיים (MATHEMATICAL_FOUNDATIONS.md), שמונה רשומות החלטות ארכיטקטוניות (ADRs), ו-24 אסמכתאות אקדמיות בפורמט APA.

איכות הקוד עומדת בסטנדרטים תעשייתיים גבוהים: 21 מודולים ממוקדים (כולם קצרים מ-150 שורות), 264 בדיקות אוטומטיות עם כיסוי של 88%, שימוש עקבי ב-type hints, docstrings מלאים בסגנון Google/NumPy לכל פונקציה ציבורית. התכנון הניסויי קפדני וניתן לרבייה: שבע רמות אופטימיזציה, שלוש ריצות לכל שיטה (סה"כ 5,100 פריטים), שימוש ב-random seeds קבועים, וניתוח סטטיסטי מקיף הכולל t-tests עם תיקון Bonferroni, ANOVA, Cohen's d ורווחי סמך של 95%. הוויזואליזציות ברמה מקצועית, עם גרפים ברזולוציה גבוהה (DPI 300), פלטת צבעים נגישה (Okabe-Ito) וסימון מובהקות סטטיסטית.

נקודות חולשה

לצד החוזקות, קיימים מספר תחומים בהם ניתן היה להעמיק. שלב Phase 2, הכולל עבודה מול API אמיתי, הושלם אך בהיקף מוגבל בשל מגבלות תקציב (\$10-\$15), דבר שמצמצם את היקף האימות האמפירי של השערות שנבנו בסימולטור. כמו כן, חלק מהמסמכים המרכזיים (למשל ARCHITECTURE.md) ארוכים במיוחד, דבר המקשה על ניווט 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 במדד accuracy, בעוד שמדדים משלימים כגון robustness, calibration או factual consistency לא נבחנו לעומק. נקודות אלו משקפות בחירות מודעות במסגרת מגבלות זמן ומשאבים, ולא ליקויים מהותיים באיכות העבודה.

השקעה ומאמץ

הפרויקט דרש השקעת זמן ומאמץ משמעותיים – כ-80 שעות עבודה אפקטיבית לאורך תשעה שבועות. העבודה כללה שלב תכנון מפורט (PRD ו-PLANNING.md), יצירת datasets מאומתים (650 פריטים), פיתוח simulator מבוסס ספרות מחקרית, ביצוע ניסויים שיטתיים, ניתוח סטטיסטי מעמיק ו-refactoring מקיף לעמידה מלאה בדרישות המבניות. התמודדנו עם אתגרים טכניים מורכבים, בהם עיצוב ארכיטקטורה מודולרית, הטמעת שיטות סטטיסטיות מתקדמות, ופיתוח מנגנוני caching והרצה מקבילית שהניבו שיפור ביצועים של פי 2-4. התהליך היה מסודר, מתועד ומבוקר, עם אבני דרך ברורות ואימות בכל שלב.

חדשנות וייחודיות

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

למידה והפנמה

הפרויקט תרם ללמידה עמוקה הן ברמה הטכנית והן ברמה המחקרית. העמקנו בהבנת מתודולוגיות מחקר אמפיריות, חשיבות קבוצות ביקורת, ניתוח מובהקות סטטיסטית, תיקון להשוואות מרובות ופרשנות של גדלי אפקט. במקביל, יישמנו עקרונות מתקדמים בהנדסת תוכנה כגון Single Responsibility Principle, תכנון מודולרי ו-Test-Driven Development, והפנמנו את ערכו של תיעוד כחלק אינטגרלי מתהליך הפיתוח. לבסוף, פיתחנו הבנה מעשית של האיזון הנדרש בין שאיפה לשלמות לבין מגבלות ריאליות של זמן ותקציב, לצד יכולת הערכה עצמית ביקורתית – מיומנות מרכזית בהתפתחות אקדמית ומקצועית.

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 הציין
- ☒ אני מודע/ת שציין עצמי גבוה יוביל לבדיקה דקדקנית יותר
- ☒ אני מקבל/ת את העובדה שהציין הסופי עשוי להיות שונה מהציין העצמי
- ☒ העבודה היא פרי עבודתי/נו (של הקבוצה) ואני/ז אחראי/ם לכל תוכנה

חתימה: אריאל נפדנסקי ואבי פרדמן תאריך: 16.12.2025

