

מטלה 4

שם קוד הקבוצה- אריאלנפדןסקיאביפרדרמן

ת.ז חבריו הקבוצה-

. 1. 313306078 אריאל נפדןסק'

. 2. 316420132 אבי פרדמן

- קישור לפרויקט -

<https://github.com/ArielleNapa/LLMs-and-Multi-Agent-Orchestration---Assignment2.git>

ציון מוערך- 85 😊

טופס הגשת הערכה עצמית

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

שם הפרויקט: מורה דרב

תאריך הגשה: 04.12.2025

הציון העצמי שלי: 85/100

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 גבואה?

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

השערה: כמה זמן ומאזן השקעתם?

חדשנות: האם יש משהו ייחודי או מיוחד בעבודה?

למידה: מה למדת מהפרויקט?

נקודות חזק

1. **תיעוד וארქיטקטורה** מצוינים הפרויקט כולל תיעוד רחב היקף בrama אקדמית גבוהה: PRD, מקיף של 500+ שורות, תיעוד ארכיטקטורה עם טרשיימי C4 מלאים (4 רמות), README בرمת user manual, +10 מסמכי documentation נוספים. כל מסמך מבונה היטב עם diagrams, tabulars, ודוגמאות קוד.

- .2. קוד מקצועי עם modularity גבוהה המבוסס על ארכיטקטורת multi-agent מתוחכמת
עם 4 תוכנים מתמחים (YouTube, Music, Story, Judge), הפרדת אחריות מלאה, ו1,247 שורות קוד מתעדות. הקוד עוקב אחרי עקרונות SOLID, כולל type hints, ומשתמש בdesign patterns מתקדמים (Factory, Strategy, Singleton).
- .3. מחקר אמיתי ומדד מהחברת prompt engineering עם 18+ איטרציות מתעדות והשפעה מדודת (+ relevance, +20% consistency 15%+). ניתוח עלויות מקייף עם חישובים פיננסיים מפורטים (\$0.53 לכל journey), תחזית 3 שנים, ויזיהו אפשרות אופטימיזציה (72% הפקחת עלויות).
- .4. בדיקות מקיפות 82 fixtures עם 100% pass rate, CISCO 79.16% (קרוב מאוד ל-80%), ומערכת בדיקות מבנית עם 5- mocks. ביצועים concurrency שימוש מתוככם ב- ThreadPoolExecutor לביצוע מקביל של 3 תוכנים לכל waypoint, עם thread-safety מלא ו- x30 שיפור ביצועים לעומת סדרתית.

נקודות חולשה (כנות)

- .1. פער CISCO בבדיקות: Story Agent coverage 55.91% ו-Judge Agent רק 71.43%. מה שמריד את הicusio הכללי מתחת ל-80%. זה הפער המשמעותי ביותר.
- .2. העדר UI: אין ממשק משתמש גרפי - רק UI Web. CLI הוכר במהלך ניקיון הפרויקט.
- .3. סוגיות קטנות בהרחבה: למרות שהמערכת extensible, אין דוגמאות מעשיות להרחבת תוכנים חדשים.

השקעה

השענו כ-350 שעות בפרויקט (כמפורט בвитוח העליות), כולל:

- תכנון ארכיטקטורה: 40 שעות
- פיתוח ליבה: 120 שעות
- Prompt engineering: 50 שעות
- בדיקות: 60 שעות
- תיעוד: 80 שעות

חדשנות

- AI-curated content Multi-agent orchestration
- Judge algorithm ייחודי עם ציון 4 (multi-criteria)
- tier caching system-3 ל밋וב עלויות
- scalability מההתחלת עם דגש על Production-grade

למידה

למדנו עקרונות קרייטיים:

- תיאום סוכנים מ经纪人** •
- שיפור בן 18 איטרציות עם מדידות** •
- Concurrency patterns - ThreadPoolExecutor, thread-safety, queue management** •
- ניתוח פיננסי מעמיק והבנת trade-offs** •
- Production architecture - clean code, testing, documentation, security** •

הمسקנה: פרויקט בرمאה אקדמית גבוהה (M.Sc.) עם ביצוע מקצועני, תיעוד יוצא דופן, ומחקר ממשי.

הצהרת יושר אקדמי (Academic Integrity Declaration)

אני מצהיר/ה בזאת ש:

- העריכה העצמית שלי היא כנה ואמיתית
- בדקתי את העבודה מול כל הקרים/רשות לפניה/במשך הציג
- אני מודע/ת שציון עצמי גבוה יוביל לבדיקה דקדקנית יותר
- אני מקבל/ת את העבודה שהציג הוסף עשייה להזות שונה מהציון העצמי
- העבודה היא פרי עבודה/נו (של הקבוצה) ואני/ו אחרא/י מכל תוכנה

חתימה: אריאל נפנסקי ואבי פרדמן תאריך: 04.12.2025

