

מבוא לבינה מלאכותית

סמסטר חורף תשפ"ו

מטלה 4 - פרויקט

תאריך אחרון להגשה: 23:55 19.01.26

הנחיות

- שאלות בנושא מטלה זו יש לשאול דרך המודל, בפורום "מטלה 4".
- הוראות להגשת המטלה מופיעים בסוף מסמך זה.
- הקבצים הנדרשים להרצת הקוד הינם:
 - AgenticAI.ipynb
- העבודה ניתנת להגשה בזוגות או כיחידים.
- לפני שניגשים לממש את המטלה מומלץ לעיין בהסברים, בקוד המצורף ובתרגול מספר 7.
- פתרון המטלה שתגישו ייבדק מול שאר ההגשות על ידי תוכנת העתקות.
- **מי שימצא כי העתיק יכשל בקורס וכן יועבר לוועדת משמעת אוניברסיטאית.**
- הפרויקט נכתב וייבדק בשפת התכנות python.
- מסמך זה בנוי באופן הבא: תיאור המטלה, בעיית הסוכן המלאכותי, מרחב הבעיה, שאלות המטלה, הסבר על המימוש הקיים והסבר על הגשת המטלה.

תיאור המטלה

לצוות הקורס של בינה מלאכותית התגלה כי מטלת החיפוש, המשחקים והתכנון הפכו למפורסמות בכל העולם וכולם רוצים לנסות את האלגוריתמים שלכם! מאחר ורוב העולם לא יודע לכתוב קוד, המטרה שלכם היא לבנות רשת סוכנים שבראשה עומד סוכן מבוסס LLM שיפעיל שני סוכנים נוספים:

1. סוכן LLM שינסה להחזיר פתרון עבור הקלט ללא המימוש שלכם לאחת משלושת המטלות
2. סוכן LLM שישתמש במימוש המטלה שלכם כדי לפתור את הקלט של המשתמש.
3. ראש הסוכנים - סוכן LLM שישווה בין הפתרונות וינסה לתת הסבר והבדלים אם כאלו קיימים (LLM as a judge).

שאלות המטלה

- עליכם לבחור את אחת משלושת המטלות – חיפוש, משחקים או תכנון. וליישם את פונקציונליות הסוכנים רק עבור המטלה שבחרתם.
- מימשו בעזרת langgraph סוכן LLM המקבל קלט למטלה שבחרתם ומריץ את מימוש המטלה שלכם ומדפיס את הערך הנדרש:
 - עבור חיפוש – הסוכן יקבל מצב התחלתי וסופי, יריץ את המימוש שלכם של אלגוריתם החיפוש וידפיס את עלות הפתרון.
 - עבור משחקים – הסוכן יקבל את לוח המשחק והמצב ההתחלתי ויריץ את היוריסטיקה שמשתתם במטלה.
 - עבור תכנון – הסוכן יקבל מופע של הבעיה שתקבלו במטלת התכנון, יריץ את המימוש שלכם לאלגוריתם התכנון וידפיס תוכנית הפותרת את מופע הבעיה.
- מימשו בעזרת langgraph סוכן LLM המקבל קלט למטלה שבחרתם ומנסה לפתור את הבעיה בעצמו:
 - עבור חיפוש – הסוכן יקבל מצב התחלתי וסופי, ינסה לפתור את הבעיה בעצמו (ללא המימוש שלכם) וידפיס את עלות הפתרון.
 - עבור משחקים – הסוכן יקבל את לוח המשחק והמצב ההתחלתי ויכתוב פונקציה היוריסטית חדשה שתתחרה בפונקציה שמשתתם במטלה הראשונה.
 - עבור תכנון – הסוכן יקבל מופע של הבעיה שתקבלו במטלת התכנון ינסה לפתור את הבעיה בעצמו (ללא המימוש שלכם) ידפיס תוכנית הפותרת את מופע הבעיה.
- מימשו בעזרת langgraph סוכן LLM המקבל את הפלט של שני הסוכנים הקודמים, מסביר את ההבדלים בין הפתרונות ורושם סיכום קצר. אם בחרתם לממש את בעיית המשחקים, הריצו משחק בין 2 היוריסטיקות השונות.

הגשת המטלה

- יש להגיש מחברת פייתון מסוג `ipynb`, עם המימושים של המשימה שתבחרו ולציין איזה מ-3 המשימות בחרתם לממש.
- **שימו לב** שהפלטים הרלוונטים נשמרים במחברות הפייתון בזמן ההגשה.
- במקביל להגשת המטלה במודל יצורף סלוטים אליהם תצרכו להשתבץ ולבצע בדיקה פרונטלית מול בודקי העבודות. בפגישה זו תסביר על אופן המימוש שלכם של העבודה. ללא פגישה זו לא תוכלו לקבל ציון על ההגשה שלכם.

קובץ המטלה

שימו לב שבשונה מהקוד שהוצג בכיתה (תרגול 7), כאן יש לכם כבר מימוש בסיסי של סוכן שרץ לוקלית על המחשב שלכם, יכול להריץ את שני הסוכנים בתצורה המתבקשת ומשווה בין התוצר שלהם. בדוגמא שניתנה לכם כול סוכן מתבקש להחזיר תשובה לשאלה 5 בחזקת 3, כאשר לאחד הסוכנים יש גישה ל tool שיחשב לו את התוצאה ולשני לא. ברובריקה 2.3 האופן בו אנחנו מפעילים את הסוכנים השונים, ברובריקה 3 יש את ההרצה של LLM as a judge . אתם מוזמנים לשנות את הקוד, הפרומפטים וה LLM שרץ כפי שתמצאו.

שימו לב שאם תשתמשו ב Gemini הגרסא החינמית מוגבלת ל 5 בקשות בדקה ו 20 בקשות ביום ולכן יש לכם אפשרות להוריד LLM קטן לוקלית בעזרת ollama ולרוץ איתו. ברובריקה 2.1 ניתן לראות את הסוכן ibm/granite4:350m (סוכן שמומחה ב tools בגודל 350 מיליון פרמטרים) ממומש בצורה גלובלית ל-3 תתי הסוכנים שלנו (מוזמנים להגדיר כול סוכן עם LLM שונה, ממליץ להתנסות עם גדלים גדולים יותר של המודלים במידה והם רצים אצלכם בזמן סביר).