

סילבוס - תוכנית הוראה לקורס מערכות בינה מלאכותית מתקדמות

מבוא ל-LLMs, סוכנים וריבוי סוכנים

| מחלקה לניהול

55-023

Advanced AI Systems

Introduction to LLMs, Agents and Multi-Agents

סוג הקורס: שיעור

היקף נ"ז: 3

שנת לימודים:

סמסטר:

יום ושעה

שעת קבלה:

מייל מרצה:

קישור לאתר למדה: _____



תיאור הקורס ומטרות למידה

תקציר הקורס

הקורס עוסק בהיכרות מעמיקה עם מודלי שפה גדולים (LLMs), פיתוח סוכני בינה מלאכותית ומערכות רב-סוכנים, בדגש על שילוב תיאוריה עם יישום מעשי בסביבות מתקדמות. הסטודנטים יפתחו מיומנויות בהנדסת הנחיות, פתרון בעיות מבוסס חשיבה בצעדים, ותכנון פתרונות AI בארגונים. הקורס משתלב במיומנויות הנדרשות בניהול טכנולוגיה ומכין את הבוגרים לעבודה מקצועית עם טכנולוגיות AI עדכניות.

מטרות/תוצרי הלמידה של הקורס

א. ידע

1. הלומדים יתארו את סוגי מודלי השפה הגדולים (LLMs), סוכני AI, ומערכות רב-סוכנים תוך הבחנה בין רמות מורכבות ויכולות.
2. הלומדים יגדירו מושגים מרכזיים כגון Prompt Engineering, Chain-of-Thought Reasoning, MCP ו-Orchestration בין סוכנים.
3. הלומדים יכתבו תכניות עבודה הכוללות עיצוב סוכן AI או מערכת רב-סוכנים לפתרון בעיה עסקית יישומית.

ב. מיומנויות

1. הלומדים ינתחו בעיות עסקיות מורכבות ויבחרו עבורן פתרונות בינה מלאכותית מתקדמים המתאימים לתרחיש.
2. הלומדים יעריכו את איכות הביצועים של סוכני AI או מערכות מרובות סוכנים, תוך שימוש במדדים כמותיים ובקריטריונים מוסכמים.

ג. ערכים (רלוונטי לקורס זה)

1. הלומדים יפתחו תובנות ביקורתיות לגבי מגבלות, סיכונים ואתגרים אתיים בשילוב טכנולוגיות AI בארגונים.
2. הלומדים יאמצו גישה אחראית ומושכלת בהטמעת פתרונות AI, תוך שיקול דעת מוסרי, רגולטורי וחברתי.



למידה פעילה - תכנון מהלך השיעורים:

מס' השיעור	נושא השיעור	למידה פעילה	קריאה/ צפיה נדרשת	הערכה תהליכית/מעצבת
1	מבוא למודלי שפה גדולים: (LLM) יסודות והשוואת מודלים	למידה בזוגות – השוואה בין GPT, Claude, Gemini	פרק 1 בספר Large Language Models, סרטון השוואה מ-YouTube	הגשת טבלה השוואתית קצרה בזוגות
2	אתיקה ואחריות בשימוש במערכות AI	דיון כיתתי מבוסס מקרים אמיתיים ("דילמה מוסרית")	מאמר מ ACL-על אתיקה ב-AI	שאלון רפלקטיבי אישי לאחר הדיון
3	Prompt Engineering מתקדם	תרגול מעשי – כתיבת prompt דרישות שונות	מדריך prompt של OpenAI	הגשת prompt לבעיה עסקית + הסבר הבחירות
4	פתרון בעיות בעזרת חשיבה בצעדים-Chain of Thought)	פתרון תרגיל בקבוצות – שלב אחר שלב	פרק 3 ב Chain of Thought, מתוך Grok	הגשה קבוצתית – תיעוד שרשרת חשיבה למקרה
5	שילוב LLM עם מקורות מידע חיצוניים	סימולציית סוכן מחפש מידע	תיעוד AutoGPT/Gemini Docs	דו"ח קצר: חיפוש ותגובה מבוססת מקור
6	מבוא לסוכנים חכמים (Agents)	ניתוח תרחיש – איזה סוכן מתאים למשימה מסוימת	פרק רלוונטי מתוך Artificial Intelligence: Agents	שאלון הבחנה בין סוגי סוכנים
7	פיתוח סוכן עצמאי בפיתוח	תרגול מעשי במחשב – בניית סוכן בסיסי באמצעות LangChain	דוקומנטציה של LangChain	הגשת קוד + תיאור מילולי קצר
8	מערכות רב-סוכנים ופרוטוקול MCP	סימולציית תקשורת בין סוכנים בקבוצות	סרטון MCP מהתעשייה, פרק מתוך Multi-Agent Architecture	תיעוד אינטראקציה בין סוכנים
9	תכנון מערכת רב-סוכנים לארגון	עבודה קבוצתית על תרשים מערכתי	סקירה של פלטפורמות כמו CrewAI / LangGraph	הגשה: דיאגרמה + הסבר

10	שילוב LLM עם עיבוד תמונה וקול	פרויקטון: "סוכן מולטימודלי" לניתוח תוכן	Canva + ElevenLabs דוקומנטציה	הדגמה קצרה בכיתה
11	שימוש בסוכנים למחקר שוק	בניית תהליך מחקר סמי-אוטונומי בקבוצות	קריאה מתוך JAIR + פרויקט דוגמה מ-Manus	הצגת תהליך המחקר לכיתה
12	הטמעת פתרונות AI בארגון	סימולציית יעוץ – הצגת פתרון למנכ"ל מדומה	מאמר מתוך Nature Machine Intelligence על Implementation	משוב עמיתים לאחר הסימולציה
13	הצגת פרויקטי הגמר	פיץ' קבוצתי + הדגמה	חזרות עצמאיות, שימוש ב-Google Slides - NotebookLM	הערכה לפי רובריק הצגה, כולל משוב מהכיתה ומהמרצה

*ייתכנו שינויים בסילבוס ובאופן הלמידה הפעילה בהתאם לקצב ההתקדמות ואפקטיביות הלמידה



ציון סופי

מסקל בציון הסופי	תיאור התוצר
50%	מטלות: השלמה והגשת כל הערכה תהליכית/מעצבת שהתבקשו לבצע בשיעור ובבית
50%	בחינה



דרישות הקורס

הערכה מעצבת

במהלך הסמסטר יוגשו כל המטלות, שכל אחת מהן תאפשר לתרגל מיומנויות רלוונטיות לתוכן השיעור ולהדגים הבנה יישומית של הנושאים הנלמדים.

תוצר מסכם

מבחן מסכם, שבו יידרשו הסטודנטים לענות על שאלות וגם לתכנן מערכת AI הכוללת מערכת מרובת סוכנים לפתרון בעיה ארגונית, תוך הסבר של ההחלטות התכנוניות, הגדרת תפקידים, בחירת כלים והערכת התוצאה.

- חובה להגיש את כל המטלות במהלך הקורס.
- חובה לקבל ציון של לפחות 60 בכל מטלה בנפרד ובמבחן, לצורך מעבר הקורס.
- ציון הקורס הוא מספרי.
- נוכחות בשיעורים היא חובה (עד שתי היעדרויות מותרות).
- באחריות הסטודנטים לקבל רישיונות חנימיים לפלטפורמות AI שישמשו במהלך הקורס, כגון: Grok, Gemini, Perplexity, Claude, Qwen, DeepSeek, ChatGpt, Manus, Cursor, Google AI Studio, ElevenLab, Canva, Monica, You, Notebooklm, Grammarly, QuilBot, Google Translate, Napkin, Storm, Scispace, N8N, Activepieces, CrewAI, LangChain



שם הקורס	מס' הקורס
שימושי מחשב בעולם הפיתוח או מבוא תכנות בסביבת פייתון	55011 או 55096
מתמטיקה 2	55005
סטטיסטיקה א' וסטטיסטיקה ב'	55003, 55002

ביבליוגרפיה: תכנים לקריאה, צפיה והאזנה

ביבליוגרפיה (ספרים מרכזיים):

1. J. Atkinson-Abutridy, *Large Language Models: Concepts, Techniques and Applications*. CRC Press, 2024. doi: 10.1201/9781003517245
2. C. M. Bishop, *Pattern Recognition and Machine Learning*. Springer, 2006.
3. A. Ng, *Machine Learning Yearning*. deeplearning.ai, 2018. (מהדורת טיוטה מקוונת)
4. M. K. Jeffrey, *Chain of Thought: Unlocking the Power of Step-by-Step Reasoning*. Independently Published, 2024.
5. D. L. Poole and A. K. Mackworth, *Artificial Intelligence: Foundations of Computational Agents*. Cambridge University Press, 2023. doi: 10.1017/9781009258227
6. A. Vemula, *Designing Multi-Agent Architecture for Advanced Generative AI Applications*. Self-published, 2024.

מקורות נוספים (מאמרים וכתבי עת נבחרים):

1. *Journal on Data Semantics*, Springer, vol. [various], 2003–present. [Online]. Available: <https://link.springer.com/journal/13740>
2. *Transactions of the Association for Computational Linguistics (TACL)*, ACL, vol. [various], 2013–present. [Online]. Available: <https://transacl.org>
3. *Proceedings of the Annual Meeting of the Association for Computational Linguistics (ACL)*, ACL, vol. [various], 1963–present. [Online]. Available: <https://aclanthology.org>
4. *Journal of Artificial Intelligence Research (JAIR)*, AI Access Foundation, vol. [various], 1993–present. [Online]. Available: <https://jair.org>
5. *Nature Machine Intelligence*, Springer Nature, vol. [various], 2019–present. [Online]. Available: <https://www.nature.com/natmachintell>

משאבים נוספים:

- מדריכי Prompt Engineering למשל, מאתר OpenAI וממקורות מובילים נוספים
- דוקומנטציה רשמית של פלטפורמות וכלי LLM מובילים
- מקרי בוחן מהתעשייה לשילוב מערכות LLM ו AI-מתקדמות בארגונים.

טבלת סרטוני יוטיוב מתקדמים להעשרה לקורס מערכות בינה מלאכותית מתקדמות

סרט	נושאי לימוד	נושא הסרט	איך הסרט קשור להרצאה	5 דברים שניתן ללמוד מהסרט	קישור לסרט
1	מבוא למודלי שפה גדולים (LLM): יסודות והשוואת מודלים	Stanford CS229 - Building Large Language Models (LLMs)	הסרטון מספק מבוא מקיף למודלי שפה גדולים, מכסה את היסודות הטכניים ומשווה בין מודלים שונים, בדיוק כפי שנדרש בהרצאה הראשונה	1. הבנת המושגים הבסיסיים של מודלי שפה גדולים ואיך הם עובדים 2. תהליך ה-tokenization וחשיבותו בעיבוד שפה טבעית 3. שיטות הערכה של מודלי שפה, כולל Perplexity ו-MMLU 4. ההבדל בין pretraining ל-post-training במודלי שפה 5. הבנת הארכיטקטורה של Transformer והשפעתה על ביצועי המודל	https://www.youtube.com/watch?v=9vM4p9NN0Ts
2	אתיקה ואחריות בשימוש במערכות AI	Yoshua Bengio - The catastrophic risks of AI — and a safer path (TED Talk)	הסרטון מתמקד בדיוק באתיקה ואחריות בפיתוח ושימוש במערכות AI, מציג את הסיכונים והאתגרים האתיים, ומציע דרכים לפיתוח אחראי של טכנולוגיית AI	1. הבנת הסיכונים הקטסטרופליים הפוטנציאליים של מערכות AI מתקדמות 2. זיהוי התנהגויות של בעייתיות של AI כמו רמאות ושימור עצמי 3. חשיבות הפיתוח	https://www.ted.com/talks/yoshua_bengio_the_catastrophic_risks_of_ai_and_a_safer_path?language=en

סרט	נושאי לימוד	נושא הסרט	איך הסרט קשור להרצאה	5 דברים שניתן ללמוד מהסרט	קישור לסרט
				האחראי של AI ושמירה על שליטה אנושית 4. הצורך בתכנון מדיניות ורגולציה למניעת נזקים מ-AI 5. איזון בין חדשנות טכנולוגית לבין אחריות חברתית ואתית	
3	Prompt Engineering מתקדם	Anthropic - AI prompt engineering: A deep dive	הסרטון מספק מבט מעמיק על prompt engineering מתקדם ממומחים מובילים בתחום, מכסה טכניקות מתקדמות וטיפים מעשיים שמתאימים בדיוק לנושא ההרצאה	1. טכניקות מתקדמות לכתובת prompts יעילים ומדויקים 2. שיטות עבודה מקצועיות של מומחי prompt engineering בתעשייה 3. הבנת האתגרים והפתרונות בפיתוח prompts מורכבים 4. מגמות עתידיות ופיתוחים חדשים בתחום prompt engineering 5. טיפים מעשיים לשיפור ביצועי מודלי שפה	https://www.youtube.com/watch?v=T9aRN5JkmL8

סרט	נושאי לימוד	נושא הסרט	איך הסרט קשור להרצאה	5 דברים שניתן ללמוד מהסרט	קישור לסרט
				באמצעות prompting	
4	פתרון בעיות בעזרת חשיבה בצעדים (Chain-of-Thought)	Stanford CS25 - Large Language Model Reasoning (Denny Zhou, Google DeepMind)	הסרטון מתמקד בדיוק בנושא החשיבה בצעדים (Chain-of-Thought) ומספק הסבר מעמיק על הטכניקות והיסודות המתמטיים של reasoning במודלי שפה, מומחה מוביל מ-Google DeepMind באוניברסיטת Stanford	1. הבנת העקרונות הבסיסיים של Chain-of-Thought prompting ויישומו 2. למידה על טכניקת Self-consistency לשיפור דיוק הרזוניטיבי 3. הבנת היסודות המתמטיים של in-context learning במודלי שפה 4. זיהוי המוטיבציות והיתרונות של פיתוח יכולות reasoning ב-LLMs 5. הכרת טכניקות מתקדמות לשיפור יכולות פתרון בעיות של מודלי AI	https://www.youtube.com/watch?v=enX5Ur1hBk&ab_channel=StanfordOnline
5	שילוב LLM עם מקורות מידע חיצוניים	Stanford Webinar - Agentic AI: A Progression of Language Model Usage	הסרטון מכסה בפירוט את נושא RAG (Retrieval Augmented Generation) ושילוב מודלי שפה עם מקורות מידע חיצוניים, בדיוק כפי שנדרש	1. הבנת מושג RAG ואיך לשלב מקורות מידע חיצוניים עם LLMs 2. למידה על טכניקות tool usage function calling במודלי שפה	https://www.youtube.com/watch?v=kJLiOGle3Lw

סרט	נושאי לימוד	נושא הסרט	איך הסרט קשור להרצאה	5 דברים שניתן ללמוד מהסרט	קישור לסרט
			בהרצאה החמישית	3. הכרת דפוסי עיצוב agentic ויישומם המעשי 4. הבנת מגבלות מודלי שפה ואיך לפתור אותן עם שילוב מידע חיצוני 5. דוגמאות מעשיות לבניית AI agents עם יכולות חיפוש ואחזור מידע	
6	מבוא לסוכנים חכמים (Agents)	Stanford Webinar - Agentic AI: A Progression of Language Model Usage	הסרטון מספק מבוא מקיף לסוכנים חכמים (Agentic AI), מסביר את המושגים הבסיסיים, דפוסי העיצוב והיישומים המעשיים, בדיוק כפי שנדרש בהרצאה השישית	1. הבנת ההגדרה והמאפיינים של סוכנים חכמים (Agentic AI) 2. למידה על דפוסי reasoning action-1 בסוכנים אוטונומיים 3. הכרת דפוסי עיצוב מרכזיים לפיתוח סוכנים יעילים 4. הבנת יישומים מעשיים כמו customer support agents 5. טכניקות לשיפור ביצועי סוכנים באמצעות	https://www.youtube.com/watch?v=kJLiOGle3Lw

סרט	נושאי לימוד	נושא הסרט	איך הסרט קשור להרצאה	5 דברים שניתן ללמוד מהסרט	קישור לסרט
				-I reflection iterative improvement	
7	פיתוח סוכן בפייתון - מעשי	Andrew Ng Explores The Rise Of AI Agents And Agentic Reasoning	הסרטון מציג את עליית הסוכנים החכמים ו-agentic reasoning מנקודת מבט מעשית, כולל דוגמאות קוד ויישומים בפייתון	1. הבנת העקרונות של המעשיים של פיתוח AI agents 2. למידה על agentic reasoning patterns ויישומם 3. הכרת כלים וספריות לפיתוח סוכנים בפייתון 4. דוגמאות מעשיות לבניית סוכנים אוטונומיים 5. טכניקות debugging של AI testing agents	https://www.youtube.com/watch?v=KrRD7r7y7NY
8	בדיקות ואימות מערכות AI	Stanford AA228V - Validation of Safety Critical Systems	הסרטון מתמקד בוולידציה ובדיקות של מערכות קריטיות לבטיחות, כולל מערכות AI, ומציג שיטות מתקדמות לאימות ובדיקה	1. הבנת עקרונות הוולידציה של מערכות AI קריטיות 2. למידה על שיטות בדיקה ואימות מתקדמות 3. הכרת סטנדרטים ופרוטוקולים לבדיקת מערכות AI 4. טכניקות לזיהוי ומניעת כשלים במערכות AI	https://www.youtube.com/watch?v=hE9iWwMZANE

סרט	נושאי לימוד	נושא הסרט	איך הסרט קשור להרצאה	5 דברים שניתן ללמוד מהסרט	קישור לסרט
				5. שיטות הערכת ביצועים ואמינות של מודלי AI	
9	ביצועים ואופטימיזציה של מודלי שפה	Fine-Tuning LLM Made Easy	הסרטון מתמקד בטכניקות אופטימיזציה ושיפור ביצועים של מודלי שפה גדולים, כולל fine-tuning ושיטות מתקדמות	1. הבנת טכניקות fine-tuning למודלי שפה גדולים 2. למידה על שיטות אופטימיזציה לשיפור ביצועים 3. הכרת טכניקות לחיסכון במשאבים 4. שיטות הערכה ומדידת ביצועי מודלים 5. טכניקות מתקדמות לשיפור דיוק ומהירות מודלים	https://www.youtube.com/watch?v=MT7DmvpPngU
10	מודלים מולטימודליים (טקסט + תמונה)	Multimodal AI Agents with Ruslan Salakhutdinov	הסרטון מתמקד בדיוק במודלים מולטימודליים המשלבים טקסט ותמונה, מציג את Visual Web Arena ומסביר איך סוכנים יכולים להבין ולפעול על בסיס מידע ויזואלי וטקסטואלי	1. הבנת העקרונות של multimodal AI agents המשלבים טקסט ותמונה 2. למידה על Visual Web Arena כ-benchmark לבדיקת יכולות סוכנים ויזואליים 3. הכרת טכניקות visual	https://www.youtube.com/watch?v=wK0Tpl3gu28

סרט	נושאי לימוד	נושא הסרט	איך הסרט קשור להרצאה	5 דברים שניתן ללמוד מהסרט	קישור לסרט
				grounding set of marks representation 4. הבנת אלגוריתמי חיפוש לשיפור ביצועי סוכנים מולטימודליים 5. זיהוי אתגרים ובעיות בטיחות בפיתוח סוכנים אוטונומיים	
11	ראייה ממוחשבת ומודלי שפה	Introduction to Vision Language Models - OpenCV Live!	הסרטון מציג מבוא למודלי vision-language כולל ארכיטקטורת CLIP ויישומים מעשיים של שילוב ראייה ממוחשבת עם עיבוד שפה טבעית	1. הבנת העקרונות של vision-language models 2. למידה על ארכיטקטורת CLIP ויישומיה 3. הכרת טכניקות לשילוב מידע ויזואלי וטקסטואלי 4. יישומים מעשיים של VLMs בתעשייה 5. טכניקות הערכה ובדיקה של מודלי vision-language	https://www.youtube.com/watch?v=trYjGmlouk
12	בטיחות ויישור מודלי AI (AI Safety & Alignment)	Yann LeCun - How Could Machines Reach Human-Level Intelligence?	הרצאה מעמיקה על איך מכונות יכולות להגיע לרמת אינטליגנציה אנושית, כולל	1. הבנת האתגרים בפיתוח AI ברמת אינטליגנציה אנושית	https://www.youtube.com/watch?v=xL6Y0dpXEwc

סרט	נושאי לימוד	נושא הסרט	איך הסרט קשור להרצאה	5 דברים שניתן ללמוד מהסרט	קישור לסרט
			דיון על בטיחות AI ויישור מודלים עם ערכים אנושיים	2. למידה על עקרונות AI safety alignment - הכרת ארכיטקטורות קוגניטיביות למודלי AI מתקדמים 4. הבנת JEPA (Joint Embedding Predictive Architecture) 5. דיון על בטיחות ושליטה במערכות AI אוטונומיות	
13	מגמות עתידיות ב-AI	Andrew Ng Explores The Rise Of AI Agents And Agentic Reasoning	הסרטון מציג את המגמות העתידיות ב-AI, התמקדות בעליית הסוכנים החכמים ו-agentic reasoning כמגמה מובילה בתחום	1. הבנת המגמות העתידיות המובילות בתחום AI 2. למידה על התפתחות הסוכנים החכמים והשפעתם 3. הכרת טכנולוגיות מתפתחות ויישומיהן העתידיים 4. הבנת ההשפעה הכלכלית והחברתית של AI מתקדם 5. תחזיות לפיתוחים עתידיים	https://www.youtube.com/watch?v=KrRD7r7y7NY

סרט	נושאי לימוד	נושא הסרט	איך הסרט קשור להרצאה	5 דברים שניתן ללמוד מהסרט	קישור לסרט
				בתחום הבינה המלאכותית	

קריטריונים לבחירת הסרטים:

- שפה: כל הסרטונים באנגלית ברמה גבוהה
- מקור: הרצאות מאוניברסיטאות מובילות (Stanford, Harvard, CMU) או מומחים מפורסמים
- תאריך: כל הסרטונים מ-2023 ואילך
- צפיות: רוב הסרטונים עם מעל 50,000 צפיות (למעט מקרים חריגים של מומחים מובילים)
- תוכן: סרטוני לימוד אקדמיים, לא סרטוני שיווק

מרצים מובילים שנכללו:

- **Yann LeCun** - זוכה פרס טיורינג, מומחה מוביל ב-AI
- **Yoshua Bengio** - "Godfather of AI", זוכה פרס טיורינג
- **Andrew Ng** - מומחה מוביל ב-AI, מייסד Coursera
- **Ruslan Salakhutdinov** - פרופסור ב-CMU, מומחה ב-deep learning
- **Denny Zhou** - חוקר מוביל ב-Google DeepMind