



תוכנית לימודים אקדמית (2 סמסטרים) למפתחי בינה מלאכותית - GenAI

תקציר:

מסמך זה מציג מתווה לתוכנית לימודים ממוקדת ומתקדמת, המיועדת להכשרת מפתחי בינה מלאכותית (AI Engineers). התוכנית תוכננה במיוחד כדי להתאים לדרישות השוק הנוכחיות בתחום ה-GenAI, תוך התמקדות בלימוד תכנות ופיתוח של מערכות משולבות בינה, של סוכנים אוטונומיים (AI Agents), הנדסת פרומפטים, שימוש במודלי שפה גדולים (LLM) ופריסה של שירות לסביבת הענן (Production).

המודל החינוכי המוצע מקזז משאבים אקדמיים בתחום ה-Data Science ובניית מודל קלאסי (כדוגמת ספריות NumPy ו-Pandas עבור תכנות מדעי בשפת פייתון), לטובת התמחות בתחום הנדסת מערכת שלמה (מבנה ופעולה), הבנת הארכיטקטורה של צד שרת וצד לקוח, שימוש במשאבים בענן, ובניית אוטומציה של תהליכים.

תוכנית הלימודים האקדמית המוצעת מהווה הכשרה פורצת דרך שנותנת מענה לדרישות המיידיות של תעשיית ה-GenAI העולמית, תוך עמידה באילוצים ובמיקוד שהוגדרו.

מפתחים שיסיימו את המסלול, יהיו בעלי יכולת ייחודית לקחת פתרון בינה מלאכותית משלב הקונספט ועד הפריסה בענן (E2E), יתרון תחרותי מובהק בשוק הישראלי המחפש מפתחי מוצרי AI ממוקדי-Production.

התוכנית המוצעת היא בהיקף כולל של 234 שעות אקדמיות מרוכזות וממוקדות. המבנה, חלוקה לסמסטרים ופירוט הקורסים לפי שעות אקדמיות מצורף בהמשך למטה.

1. סיכום מנהלים והצדקה אסטרטגית

האבולוציה מ-Data Science ל-AI Engineering

המעבר הגורף בתעשיית הטכנולוגיה בשנים האחרונות אינו מתמקד רק בשיפור מודלי למידת מכונה (ML) קלאסיים, אלא בעיקר במינוף מודלי יסוד קיימים (Foundation Models), באמצעות שימוש בהם ואינטגרציה שלהם לתהליכים רחבים יותר.

המשמעות היא שהשוק עבר ממיקוד במהנדסי **אימון** של מודלים למפתחי בינה, שמבינים כיצד לחבר, להתאים ולפרוס את המודלים. תוכנית הלימודים מגיבה למעבר קריטי זה בכך שהיא מכשירה מפתחים בעלי יכולת לבנות צינורות (Pipelines) של מודלי LLM ולהפוך אותם למערכת ביצוע, או לסוכנים אוטונומיים המספקים ערך עסקי מהיר ומדיד.

מפתחי AI ב-2025 נדרשים להבין במגוון כלים רחב, ובעלי יכולת לבנות מערכות מבוצרות שמתקשרות בצורה יעילה ואפקטיבית עם APIs. דרישה זו שונה מהותית מהדרישה המסורתית למומחי Data Science שהתמקדו בניתוח נתונים טבלאיים ובסטטיסטיקה, או ממומחי למידת מכונה שמתמקדים בתכנון ובפיתוח של המודלים עצמם.

בהתאם, התוכנית שמה דגש מוחלט על הבנת מעגל הפיתוח, ארכיטקטורה של מערכות, ושימוש



בכלי בינה לכתובת קוד מהירה (Vibe Coding), ובפריסה E2E של מערכות בענן באמצעות FastAPI, Docker, ומשאבים בענן של GCP/AWS.

אסטרטגיית הוויתור: צמצום NumPy, Pandas & PyTorch

ההחלטה לצמצם באופן דרסטי את הלימוד על הספריות המדעיות NumPy ו-Pandas להיכרות בסיסית בלבד, היא ליבת האסטרטגיה של התוכנית. אלו הן ספריות מדעיות שמשתמשים בהן בעיקר במדעי המידע והמחשב, אך העמקה בהן היא הרבה מעבר להיקף הלימוד בסמסטר אחד, דורשת ידע מוקדם בתכנות וביעילות, ואינה בהכרח נדרשת. ניתוח המגמות בשוק מעיד כי האילוץ הזה הוא למעשה גורם מכריע, המאפשר מיקוד עמוק ומהיר יותר בתחום Gen-AI Engineering.

בעוד שתוכניות קיימות מכסות מגוון רחב של קורסי AI, תוכנית זו נועדה לייצר מפתחים רלוונטיים בתוך שני סמסטרים מרוכזים. אימון מודלי ML קלאסי, הדורש Data Science עמוק, הוא תהליך ארוך ומייגע, ומחייב ידע רחב יותר במתמטיקה. לעומת זאת, שימוש במודלי בינה לכתובת קוד מהירה וחיבור ממשקים, או פיתוח סוכני LLM באינטראקציה מול API, הם הרבה יותר אינטואיטיביים. הוויתור על קורסי Data Science מסורתיים מפנה משאבי זמן יקרים (כ-120 שעות אקדמיות) לטובת לימוד מתקדם של הנדסת פרומפטים ושימוש בכלי צ'אט ברמת גבוהה (גם לתכנות), הקמה והפעלה של סוכנים אוטונומיים, כתיבה ללא קוד (Vibe Coding), וכלים עבור פריסה לענן. מיומנויות אלו חשובות ומועילות יותר לפיתוח תוכנה מבוססת AI כיום.

בנוסף, הורדת החלק המדעי-מתמטי העמוק יותר, מאפשרת את הורדת תנאי הסף לקבלה למסלול, וגם מאפשרת את הגדלת קהל היעד העיקרי באופן ניכר.

2. סמסטר 1: יסודות התכנות, בינה מלאכותית, ולמידת

מכונה (117 שעות אקדמיות, 9 שעות בשבוע, ב-2 ימים * 13 שבועות)

הסמסטר הראשון נועד להקנות למשתתפים בסיס איתן בבינה מלאכותית, למידת מכונה ומודלים, שימוש בצ'אטים שונים ובכלי יצירה של בינה לטקסט ותמונה, כל זאת בנוסף ללימוד יסודות התכנות בשפת פייתון והבנת ממשקים, ארכיטקטורה ומבנה של מערכות מחשב מודרניות. **מיקוד:** תכנות באמצעות Python, פיתוח הבנה של מערכת, מודלים, תיאוריה ושימוש בכלי AI.

3. סמסטר 2: אוטומציה, ענן, סוכנים ותכנות ללא קוד

(Vibe-Coding, AI-Agents, Cloud, Automation)

117 שעות אקדמיות, 9 שעות בשבוע ב-2 ימים * 13 שבועות

סמסטר זה מהווה את ליבת המומחיות בתחום ה-GenAI, כאשר הדגש העיקרי הוא מבט מערכתי, פיתוח תהליכים מקצה לקצה (E2E) והבנת מחזור הפיתוח של מוצר תוכנה, שליטה עמוקה במודלים, שימוש ב-API וב-Frameworks מתקדמים לפריסה לענן ולבניית אוטומציות עסקיות.

סמסטר זה משלים את המסלול מהבנה ברמת הפיתוח המקומי למבט רחב ומקיף יותר. מפיתוח קוד ועד לפריסה בסביבת ענן, זה בונה את רשימת המלאי הדרושה למהנדס AI מודרני.

4. רשימת הקורסים

4.1 קורס תוכנה ופיתוח בשפת פייתון (60 שעות)

תכנות מונחה-עצמים (OOP) ותכנות מונחה-בדיקות (TDD)

הידע ב-OOP הוא הבסיס הקריטי לבניית Frameworks פנימיים ו-Agentic Systems. הלימוד מתמקד בארכיטקטורה של קוד מודולרי, עקרונות SOLID, ירושה וכתובת ממשקים, מעבר מפונקציה לתוכנה - לראייה רחבה של המערכת. מטרתו היא להכשיר מפתחים המסוגלים ליצור קוד ניתן להרחבה (Scalable) קריא וקל לתחזוקה, הדרוש במערכות AI מורכבות. נושאים נוספים שיכוסו הם תהליך הפיתוח, טיפול בחריגים, דיבאגינג, בדיקות יחידה (Unit Testing), ניהול גרסאות עם Git.

אינטגרציית AI לתהליך הפיתוח (AI-Assisted Coding)

הלמידה תשלב שימוש פרקטי ויומיומי במנועי קוד והשלמה גנרטיביים (כגון Gemini, Copilot, Claude). המפתחים ילמדו לא רק כיצד לכתוב קוד, אלא גם כיצד לבצע *Refactoring*, *Review*, *Debug* לקוד על ידי מנוע בינה. גישה זו משנה את הפרדיגמה הרגילה של לימוד שפות תכנות. במקום להתמקד רק בהכרת השפה והמבנה של פייתון, המפתח מתמקד בהבנה של כתיבת תוכנה פועלת על פי דרישות, במהירות הפיתוח, ובכלי פרסום לענן של המוצר המוגמר. זהו שינוי הכרחי בכדי להפוך את המפתחים למהירים ואפקטיביים יותר לדרישות השוק כיום.

4.2 קורס מסדי נתונים (DB), ממשקים ואינטגרציה (40 שעות)

חלק חשוב מכל מערכת היא מאגרי המידע, והחיבוריות לשירותים נוספים. מכיוון שהסטודנטים הם חסרי רקע קודם בתחום, המסלול חייב לתת כלים בסיסיים לשימוש במסדי נתונים. כמו כן, מכיוון ששילוב שירותים (כמו מודלים) היא מיומנות ליבה, קורס זה מעמיק בכלים סטנדרטיים של פייתון לבניית שירותי רשת ובשימוש ב-APIs ככלי תקשורת בין מערכות נפרדות, תוך שמירה על חוסר תלות במימוש הפנימי.

תכנים עיקריים:

- **יסודות בתקשורת, רשת, REST, HTTP, Json**
פרוטוקולים סטנדרטיים ברשת האינטרנט, כיצד REST APIs משמשים ככלי לתקשורת והעברת מידע, מהם Microservices, ניהול מפתחות, פרוטוקולי אימות והזדהות (Authentication).
- **FastAPI Mastery**
שימוש ב-FastAPI לבניית ממשקים אסינכרוניים. בניית נקודות קצה פונקציונליות ועבור בדיקות תקינות (Health Check, Keep Alive), שהן קריטיות לניטור המערכת בסביבת אמת (Production). יצירת תיעוד אוטומטי ל-API באמצעות Swagger/Redoc.



- **בדיקות תקינות מידע עם Pydantic**
בניית שירותים אמנים דורשת אכיפת חוזים עם סכמות מידע קבועות. שימוש בספריית Pydantic עבור הגדרות מבנה הנתונים (Data Validation) הוא חיוני ומאפשר פיתוח מקבילי של חלקים אחרים בקוד, תוך הסתמכות על חוזי התקשורת בין החלקים השונים וטיפול בכל אחד מהם בנפרד.

- **מסדי נתונים - בסיס**
SQL לעומת NoSQL, טבלאות ואינדקסים, שימוש בסיסי בשני מאגרי מידע מהסוגים השונים: MongoDB עבור NoSQL, PostgreSQL עבור SQL מבוזר.

- **כתיבת ממשקי צד לקוח באמצעות Streamlit.io**
פיתוח צד שרת עם פייתון בלבד וספריית קוד פתוח Streamlit, ושפת Markdown.

4.3 קורס תיאוריה - בינה מלאכותית, מודלים, למידת מכונה (20 שעות)

- כיצד מודלים של בינה מלאכותית לומדים, מאומנים ופועלים לחיזוי או ייצור מידע. תהליכי למידה מפוקחת, למידה לא מבוקרת, פידבק אנושי, ותיאוריה בסיסית.
- **הנדסת פרומפטים (Prompt Engineering)**: איך לכתוב נכון בקשות למודלים שונים, מהם חלקי הפרומפט, איך לקבל פלט בפורמט הנכון ועוד.
- **שליטה והשפעה על הפלט של המודל**: קביעת פרומפט מערכת, שינוי ערכי טמפרטורה של המודל להשפעה על היצירתיות שלו בהחזרת תשובות, קביעת מגבלות על אורך הפלט המקסימלי באמצעות Max Tokens ועוד.
- **שימוש מערכתי בפרומפטים** - בעיות בשימוש בשטח וטיפול במקרי קצה, יתרון בעצי מחשבה ופיצול לפרומפטים קצרים וממוקדים על פני פרומפט אחד ארוך שעושה את הכל, ניהול גרסאות של המערכת (Git basics), שימוש בתבניות עם משתנים (Placeholders).
- **שימוש בידע קיים על ידי RAG**
שימוש במודלים על בסיס ידע ארגוני קיים כדי לספק תשובות עובדתיות על בסיס הקשר נתון. זה מרכיב קריטי לשימוש מותאם חברה, בניגוד לשימוש הכללי. בקורס נכיר כיצד לתכנן ארכיטקטורת RAG Retrieval-Augmented Generation, מהו ייצוג וקטורי והבנת ה-Embeddings ואינדוקס נתונים ב-Vector Stores.

4.4 Vibe-Coding תכנות ללא קוד (30 שעות)

- שימוש בפלטפורמות חצי חנימיות (Loveable, Base44) בשילוב Github לפיתוח ממשקי משתמש (GUI) של צד הלקוח. הכרת היכולות ושימוש מתקדם. בנוסף נראה כיצד להשתמש בצ'אטים שהכרנו קודם לכן בצורה חדשה, על מנת לייצר ולנתח קוד.
- נכיר גם דרך לייצור ממשקי משתמש למפתחי Python ללא תלות בצד הלקוח, על ידי פלטפורמות כגון H2O Wave/Streamlit. פלטפורמות אלה מאפשרות פיתוח אפליקציות AI אינטראקטיביות באמצעות Python בלבד, תוך ביטול הצורך במיומנויות HTML, CSS, או JavaScript. לאחר בניית הסוכנים במהלך הסמסטר השני, הסטודנטים ילמדו כיצד לחבר את ממשק המשתמש ל-FastAPI Backend, על מנת ליצור אפליקציה שלמה.



4.5 פיתוח סוכנים אוטונומיים (45 שעות)

זהו המודל המתקדם ביותר בתוכנית, המכסה את הטרנד המרכזי של פיתוח סוכנים ומערכות אוטונומיות (Agentic AI, Frameworks). פיתוח סוכנים באמצעות Frameworks דוגמת LangChain, AutoGen ו-CrewAI הפך להיות הסטנדרט החדש לבניית מערכות AI ניתנות להרחבה.

- הכרת הפלטפורמות הקיימות כיום בשוק; Agentic Loop, Tool Calling, Crew-AI, LangChain, AutoGen
- הבנת מנגנון הפעולה של סוכני AI כולל תהליכים מרובי-שלבים. כיצד בונים סוכן אוטונומי רב-שלבי לפתרון בעיה. תכנון (Plan), ביצוע (Execute), תצפית (Observe), והשתקפות (Reflect/Refinement). הבנה של ניהול זיכרון ומצב (State).
- **LangChain**: ישמש כבסיס ללימוד עקרונות הניהול (Chains, Memory, Tools).
- **AutoGen ו-CrewAI** מאפשרות בניית מערכות רב-סוכנים (Multi-Agent Systems) מבוססות תפקידים. פיתוח אוטומציה עסקית מורכבת דורשת התמחות (למשל, סוכן אחד חוקר, סוכן שני מנתח, וסוכן שלישי מנסח דו"ח). CrewAI מספקת את המודל הטוב ביותר לניהול תהליכי עבודה שיתופיים אלו, ומבטיחה שהבוגרים יכולים לפתור בעיות ארגוניות רב-שלביות. AutoGen משמשת להדגמת תהליכי עבודה מבוססי שיחה ושיתוף פעולה בין סוכנים.
- **Tool Calling** ואינטראקציה חיצונית: מתן יכולת לסוכנים להפעיל APIs (כלים) חיצוניים, כגון גישה למערכת מזג אוויר או מערכת מלאי, חיוני לאוטומציה עסקית.

4.6 קורס אוטומציה עסקית (15 שעות)

היכרות ושימוש בפלטפורמות לניהול ויצירת תהליכי אוטומציה: **Make & N8N**. נבין מהו תהליך, טריגר או אירוע, שימוש ב-Webhooks לתקשורת בין תהליכים. נכיר אפשרויות לשלב בתהליכים מודלי בינה, שירותים אחרים ורשתות חברתיות.

4.7 קורס פריסה והטמעה לענן עם קונטיינרים (25 שעות)

פריסת המערכת היא החלק האחרון והקריטי ביותר של מוצר תוכנה שלם. נלמד על העקרונות הבסיסיים של הענן (Compute, Storage, Network) כפלטפורמה להרחבה (Scalability) ואבטחה. פלטפורמה דרכה ניתן לפרסם את המוצר שלנו לעולם הרחב. נלמד לבצע דחיפת קונטיינרים (Deployment) לסביבות ענן בסיסיות, למשל שימוש ב-**Google Cloud Run**.

בנוסף נכיר את **Docker** המאפשר בנייה, אריזה של כל רכיבי המערכת לקונטיינר מוכן להפצה, ופריסה של מערכת מלאה לענן. דוקר מקצר באופן משמעותי את מחזור הפיתוח המסורתי, ומספק תשתית אבטחה עם התנהגות עקבית שקל לדבג ולנטר. הפריסה בענן על גבי **GCP/AWS**. כמו כן, נכיר בצורה בסיסית את **Docker Compose** המיועד לאריזת כל רכיבי המערכת וניהול של סביבת פיתוח מלאה (יכולה לכלול מספר **Microservices** או צד שרת וצד לקוח על אותו קונטיינר).

5. קצת עליי - חגי און

אני חגי, בן 49 ואב לשני ילדים חמודים, בעל תואר בהנדסת חשמל ומחשבים בטכניון (2002), עם התמחות במסלולי למידת מכונה (ML) ורשתות עצביות. עוד במהלך התואר שלי, התחלתי לעבוד כמתכנת בחברת "רפאל - מערכות לחימה". בסיום התואר המשכתי לעבוד כמפתח זמן-אמת במספר חברות מוכרות כמו סאנדיסק, קוואלקום ועוד, ולאחר מכן בתפקידים של מתכנת בכיר וארכיטקט מערכת. בסך הכל, אני מפתח מומחה ב-C++, Java, JS ופייתון במשך 25 השנים האחרונות.

נקודות מעניינות:

- במהלך עבודתי תכננתי, עיצבתי ופיתחתי מודל בינה מלאכותית מסחרי מלא, המנתח נתוני ציר זמן אינסופיים של אירועים מקבילים בשנת 2016.
- בשנת 2022 פרשתי מעבודתי העיקרית כמפתח מערכות, והתחלתי להרצות על בינה מלאכותית, מודלים והטמעה של שימוש בבינה מלאכותית. בנוסף, אני יוצר ומוכר תמונות אמנות שאני יוצר עם כלי בינה מלאכותית, ומנחה 2 קהילות גדולות על יצירה עם בינה מלאכותית (כ-2000 איש).
- אני אוהב ללמד ולהעביר ידע פרקטי ומועיל, שנותן לאנשים עוד אפשרויות ליצור ולממש את עצמם, ועל הדרך לגעת בחדשנות ולהתמלא בהשראה!
- בנוסף לכל, אני מאמן קונג-פו מוסמך מוינגייט, ומדריך קרקס וג'אגלינג.

Love U