



בינה מלאכותית בסביבות עבודה סגורות

להריץ AI בתנאים שלכם

תיאור הקורס:

האם הייתם רוצים שהמידע הכי רגיש שלכם יעזוב את קירות המשרד? קורס זה מלמד אתכם איך להחזיר את השליטה לידיים שלכם. אנחנו נלמד איך להריץ מודלי שפה עוצמתיים על שרתים מקומיים או על המחשב האישי - ללא תלות באינטרנט וללא תשלום דמי מנוי חודשיים לחברות הענק.

זהו הקורס המושלם עבור ארגונים ביטחוניים, רפואיים או פיננסיים שצריכים בינה מלאכותית חזקה אך מחויבים לפרטיות מוחלטת. נבין את עולם הקוד הפתוח, נלמד על חומרה ונבין איך להפוך מחשב חזק למנוע אינטליגנציה פרטי. הידע הזה יעניק לכם עצמאות טכנולוגית ויכולת לבנות פתרונות שאיש לא יכול לחסום או לרגל אחריהם.

קהל יעד:

אנשי IT, מנהלי אבטחת מידע, חובבי טכנולוגיה וכל מי שמאמין שהמידע שלו צריך להישאר שלו.

סילבוס נושאים

פרק 1: מהפכת הקוד הפתוח

- היכרות עם עולם המודלים החופשיים: הבנת המגמות המובילות בתחום ה-Open Source.
- סקירת מודלים מובילים: העמקה ביכולות של מודלים דוגמת Llama ו-Mistral.
- עצמאות טכנולוגית: המשמעות של עבודה עם מודלים שאינם תלויים בספקי ענן.

פרק 2: תשתית חומרה ל-AI

- בחירת רכיבי עיבוד: קריטריונים לבחירת מעבדים (CPUs) וכרטיסי מסך (GPUs) ייעודיים.
- התאמת חומרה לביצועים: הבנת היחס בין נפח זיכרון (VRAM) למהירות הרצת המודל.
- תכנון מקומי: הקמת עמדת עבודה או שרת המותאמים לעומסי עבודה של אינטליגנציה מלאכותית.

פרק 3: כלי התקנה וניהול

- פלטפורמות הרצה מהירות: עבודה עם כלי פריסה כמו Ollama ו-LM Studio.
- נוחות תפעולית: הגדרה והרצת מודלים מורכבים בשיטת "לחיצת כפתור".
- ניהול סביבות עבודה: שליטה בממשקי הניהול של המודלים המקומיים.

פרק 4: אופטימיזציה

- טכניקות דחיסה: איך לגרום למודלי ענק לרוץ על חומרה סטנדרטית ונגישה.
- שיפור מהירות (Inference): איזון בין דיוק המודל למהירות התגובה שלו.
- מיקסוס משאבים: ניצול מרבי של כוח המחשוב הקיים בארגון.

פרק 5: אבטחת שרתים מקומיים

- הגנת המידע הפרטי: בניית חומות הגנה סביב תשתית ה-AI הארגונית.
- עבודה בסביבה מבודדת: הבטחה שהמידע אינו זולג אל מחוץ לכותלי המשרד.
- ניהול סיכונים: זיהוי נקודות תורפה בתקשורת מול השרת המקומי.

פרק 6: חיבוריות פנימית וריבוי משתמשים

- נגישות משרדית: הגדרת הרשת כך שכל צוות המשרד יוכל להשתמש ב-AI.
- עבודה ללא אינטרנט: יצירת קישוריות פנימית מאובטחת ללא תלות בחיבור חיצוני.
- שיתוף משאבים: ניהול גישה של מספר משתמשים בו-זמנית לשרת ה-AI המרכזי.

פרק 7: עדכונים ותחזוקה שוטפת

- ניהול גרסאות: מעקב ושדרוג המודלים בהתאם לחידושים האחרונים בשוק.
- תחזוקת מערכת: הבטחת יציבות השרת לאורך זמן.
- עדכוני אבטחה ותוכנה: הטמעת שיפורים טכנולוגיים מבלי לפגוע ברציפות העבודה.

פרק 8: בניית ממשק משתמש (UI) מקומי

- חוויית משתמש (UX): יצירת סביבת עבודה נוחה ומזמינה לשימוש יומיומי.
- התאמה אישית: בניית ממשק המותאם לצרכים הספציפיים של הארגון.
- הנגשת הטכנולוגיה: הפיכת המנוע המורכב לכלי עבודה פשוט וידידותי לכל עובד.