

Practical Deep Learning

100 שעות לימוד אקדמיות

תיאור התפקיד:

בינה מלאכותית (**Artificial Intelligence** או **AI**) ולמידת מכונה (**Machine Learning**) הן כיום טכנולוגיות מוכרות שמשולבות בכל תחום בחיינו.

לדוגמא, המלצות על סרטים ומוצרים דומים באתרי קניות ותוכן או זיהוי סיכון של נטישת לקוחות על מנת להפעיל תוכניות שימור.

יחד עם זאת, אנו רואים גם בינה מלאכותית "משופרת" – למשל, מנגנוני שפה כמו **ChatGPT** שמנהלים שיחה כמעט אנושית או אפליקציות ציור כמו **Midjourney** שמייצרות תמונות ברמת ריאליזם מדהימה.

למרות שכל היישומים הללו הם אכן AI, ברור שיש הבדל מהותי בין שני הסוגים. את האחרונים מייצרים עם שימוש באלגוריתמים של "למידה עמוקה" (**Deep Learning** או **DL**) או בשמם הטכני – רשתות נוירונים מלאכותיות (**Artificial Neural Networks** או **ANN**).

התמחות בלמידה עמוקה היא היום השלב הבא עבור ה-Data Scientist שמחפש את האתגר הבא. מעבר לכל מה שכבר נלמד, ה-Data Scientist נדרש כעת להכיר כלים ואלגוריתמים נוספים ואפילו יכולות חומרה לעיבוד מקבילי (**Graphical Processing Unit** או **GPU**). על החוקר להתמודד עם כמויות גדולות של נתונים שבעת יכולים להיות תמונות, טקסט, קול ועוד סוגים של מידע שאינם ניתנים לייצוג בטבלה (כלומר, מידע שאינם מובנה). אימון מודל DL יכול לעיתים לקחת שבועות ואף חודשים, ולכן על החוקר לקחת בחשבון גם שיקולים של יעילות וניצולת של המחשב.

תיאור ההכשרה:

ברוכים הבאים לקורס **Deep Learning** עבור מדעני נתונים מנוסים!

Deep Learning התגלה כתחום מהפכני בעולם העצום הזה של הבינה המלאכותית, וגילה את הפוטנציאל האדיר לפתרון בעיות מורכבות והנעת חדשנות בתעשייה והפתחות של עולם ה-**Machine Learning**.

הקורס הזה נועד לספק הבנה מקיפה של עקרונות וטכניקות למידה עמוקה, להקנות מיומנויות וכישורים פרקטיים להתמודדות עם משימות מאתגרות בחזית מדעי הנתונים.

הקורס יספק צלילה מעמיקה לתוך העולם המרתק של רשתות נוירונים - **neural networks**. מודלים אלה, בהשראת המבנה והתפקוד של המוח האנושי, הוכחו כחזקים מאוד בזיהוי והפקת דפוסים התנהגותיים מכמויות אדירות של נתונים. תוך כדי לימוד של חקירת ארכיטקטורות **neural networks** שונות, הקורס ימנף את מלוא הפוטנציאל של למידה עמוקה כדי להתמודד עם אתגרים בעולם האמיתי.

עם התמקדות בלמידה מעשית, לסטודנטים בקורס **Deep Learning** תהיה הזדמנות ליישם את הידע והכישורים באמצעות תרגילים מעשיים והדגמות. הסטודנטים ירכשו ידע וניסיון מעשי באימון **neural networks**, כוונן עדין של פרמטרים של מודל ואופטימיזציה של ביצועים. בסוף קורס זה, יהיה הבוגרים ייצאו לשוק העבודה עם בסיס חזק בטכניקות למידה עמוקה, שיאפשרו להם ליישם מודלים של **neural networks** עבור מגוון רחב של יישומים.

מבנה ההכשרה:

בחלק הראשון של הקורס נתמקד בלהבין איך עובדת רשת נוירונים – מהם הרכיבים השונים וכיצד מודל חישובי פשוט יכול לייצר את התוצאות המדהימות שאנו רואים היום מ-ChatGPT ו-Midjourney. בפרט, נתמקד ביכולת של רשתות נוירונים לייצר פיצ'רים בצורה אוטומטית. נבין את הכלים הנדרשים (תיאורטיים ותכנותיים) ונבין כיצד מעבד גרפי (GPU) יכול להאיץ בצורה משמעותית את תהליך האימון. בסוף חלק זה נדע לקחת נתונים ולייצר מודל **Deep Learning** לנתח אותם.

בחלק השני של הקורס נעבור על ארכיטקטורות שונות של רשתות נוירונים – כיצד סידור שונה של נוירונים וקשרים ביניהם יכול מחד ליעל את תהליך האימון ומאידך להתאים למשימות כמו עיבוד תמונה, שפה וקול. נבין את היתרונות והחסרונות של כל ארכיטקטורה כזו ואיך לבחור ביניהן.

בחלק האחרון של הקורס נדון בנושא החם בעולם הלמידה העמוקה היום – AI גנרטיבי. אלה המודלים שמשוגלים לייצר תוכן חדש מתוך תכנים קיימים ולהראות מעין 'יצירתיות' אמנותית. נבין גם את העקרונות שעומדים בבסיסם של מודלי ה-State-of-the-art כמו ChatGPT.

הדגש בקורס הוא על יישום ופרקטיקה. לפחות 50% מהקורס מיועד לתרגול מעשי והתנסות בפועל בכלים השונים.

קורס Practical Deep Learning הוא ייחודי בשוק בכך שהוא מאפשר לסטודנטים – כבר בזמן הקורס – להתנסות באימון רשתות נוירונים על GPU-ים. GPU – ראשי תיבות של Graphical Processing Unit – הם רכיבי חומרה ייחודיים שהכרחיים לאימון של מודלים בעלי פוטנציאל ליישומים פרקטים. ללא GPU ניתן לאמן מודלים קטנים בעלי פוטנציאל לימודי בלבד, או לחכות זמן רב לכל אימון – מה שכמובן לא מתקבל בארגונים. כיוון שמדובר ברכיב חומרה נפרד עם אופן פעולה משלו, נדרשת הכרה של המאפיינים הייחודיים, הבעיות הפוטנציאליות וכיצד לפתור אותן והדרכים לייצר ביצועים מקסימליים. נושאים אלה משולבים כחלק אינטגרלי בתכני הקורס.

קהל יעד ודרישות קדם:

הקורס מיועד ל-Data Scientists מנוסים שמעוניינים להרחיב את הידע והיכולות שלהם מעבר לתחום Machine Learning ולהתמחות בדבר החם ביותר בתחום – Deep Learning.

דרישות קדם מקצועיות מתמקדות במספר נקודות הבאות:

- Splitting data sets into training, validation and test sets
- Applying cross-validation for hyper-parameter tuning
- Familiarity with classical machine learning topics such as regression and classification
- The Python data-science ecosystem (pandas, numpy, scikit-learn, etc.)

תנאים לקבלת תעודת סיום קורס:

- 80% נוכחות מינימום
- הגשת כל המשימות במהלך הקורס

תוכנית הלימודים:

Introduction to Neural Networks

- What is a neural network?
- 'Hello World' – The MNIST dataset
- Training a Neural Net: Forward and Backward Passes
- Components of a NN:
 - Layers & Weights
 - Activation Function
 - Loss Function
 - Optimizer & Learning rate
 - Gradients
- Demo: Training MNIST with PyTorch
 - Training and Test data
 - Batch sizes
 - Changing Learning Rate

Introduction to Graphical Processing Units (GPU)

- What is a GPU why is it required and what types are there
- Moving training from CPU to GPU – libraries and tools
- Issues to be aware of when using a GPU

Basic Neural Network Architectures 1: Feed-Forward

- Intro to fully connected networks
- Layer types:
 - Dense
 - Dropout
- Activation Functions:
 - Sigmoid
 - TanH
 - ReLU
 - Softmax
- Loss functions:
 - Mean Squared Error
 - Cross Entropy Loss
 - Binary Cross Entropy (BCELoss, BCEWithLogitsLoss)
- Application: Autoencoder on MNIST

Basic Neural Network Architectures 2: Convolutional

- Intro to convolutional networks
- Layer types:
 - Convolutional
 - Max Pooling
 - Avg Pooling
 - Flatten
- Activation Functions:
 - ReLU6
 - LeakyReLU
- Application: CIFAR-10

Basic Neural Network Architectures 3: Recurrent

- Intro to recurrent networks
- Layer types:
 - Embedding
 - GRU
 - LSTM
- Application: Part-of-Speech tagging

More Neural Network Architectures

- Layer types:
 - BatchNorm
- Loss Functions:
 - Triplet Loss
- Application: Siamese networks with MNIST
- Application: VGG16 with CIFAR10

Basic Neural Generative Architectures

- Activation Functions:
 - LeakyReLU
- Loss Functions:

-
- KL Divergence
 - Application: Variational auto-encoders on MNIST
 - Application: Generative adversarial networks on MNIST
 - Application: Attention, Transformers and Conformers

Advanced Topics in GPU Training

- Multi-GPU and multi-node training
- Optimization and Profiling
- GPU-accelerated Machine Learning