

Data Science & AI



המרכז הבינלאומי
ללימודי הייטק וחדשנות

בחסות
האוניברסיטה הפתוחה
מערך לימודי החוץ



Data Science & AI

תיאור המסלול

בעידן המודרני ההשפעות של מהפכת הבינה המלאכותית הופכות להיות יותר דומיננטיות, החל מתחום יחסית קטן של מודלים בעולם למידת המכונה, דרך מודלים מורכבים יותר של למידה עמוקה בעולם העיבוד תמונה ועיבוד דיבור ועד לפסגה האחרונה של ההתפתחות הטכנולוגית שבאה לכדי ביטוי ב-ChatGPT. מומחים כלכלים מנתחים את המצב ומצביעים על השינויים האלו כלא פחות מאשר המהפכה התעשייתית הרביעית.

כדי לעמוד בביקוש הגובר לתחום זה, הקמנו את התוכנית החדשה להכשרת מדעני נתונים. במסגרת התוכנית, אנו מכשירים בוגרי תואר ראשון מרקעים שונים להיות מדעני נתונים באמצעות כל הכלים הנדרשים לבניית פרויקטים של בינה מלאכותית.

צוות המומחים שלנו פיתח שיטה בה הלמידה התיאורטית משולבת בפרויקטים מעשיים ועשיה משמעותית בנושאים שמעניינים את הסטודנטים עצמם. תוך כדי ההכשרה נתמודד עם בעיות אמתיות, עם כתיבה נכונה של קוד עבור יישומי פרודקשן ושימוש בכלים מתקדמים לפתרון אתגרים רלוונטיים בעולם ההיי-טק הנוכחי.

מפרט טכני מינימלי למחשב

- כונן אחסון SSD 512
- מסך 14" לפחות
- מעבד i5 דור 8 ומעלה
- זיכרון GB16

Data Science & AI

חשוב לדעת

היקף שעות

392 שעות לימוד אקדמיות הכוללות עבודה עצמית על פרויקטים ופרויקט בתעשייה.

קהל יעד

הקורס לא מסתמך על ידע מוקדם אבל דורש היכרות מוקדמת עם לפחות אחד משלושת האספקטים המרכזיים הדרושים למדעני נתונים:

- יכולות מתמטיות כמותיות עבור סטודנטים שלמדו בעבר תואר אקדמי ממקצועות המדעים המדויקים, הנדסה או מתמטיקה (STEM)
- סטודנטים בעלי רקע תעסוקתי בעולם התוכנה כגון מתכנתים, אנשי QA ואוטומציה.
- בעלי רקע מקצועי או אקדמי ביכולות ניתוח נתונים כגון Data Analyst, כלכלנים, אנשי ניתוח נתונים בעולם הפרסום או מקצועות דומים שמתמקדים ביכולת ניתוח הנתונים

דרישות קדם

בנוסף לאחד משלושת הדרישות הבסיסיות:

- נדרשת אנגלית ברמה גבוהה וחשיבה ריאלית
- דרשת מחויבות ויכולת למידה עצמית

זכאות לתעודת גמר מטעם מכללת INT

תעודת גמר מטעם מכללת INT תוענק לעומדים בתקנון הלימודים לרבות נוכחות ב-85% מהשיעורים לפחות, הגשת המטלות פרויקט הגמר בהתאם לדרישות המסלול.

למה ללמוד ב-INT?

- מתודולוגית הלמידה במסלול יהיה במודל PBL (Project-Based Learning), מודל המבוסס על למידה אקטיבית של הסטודנט, המשלבת בין יישום פרקטי (Hands-on) ורכישת ידע. מודל זה מקנה יכולות חשיבה ביקורתית המאפשר יישום מעשי של משימות מאתגרות, ומערבות את הסטודנט בפתרון בעיות וקבלת החלטות.
- רכישת הידע נעשית ע"י הרצאות תאורטיות, בשילוב פעילויות אינטראקטיביות המסייעות בהבנה ובהטמעת החומר הנלמד, תרגול וסימולציות, ובאמצעות למידה עצמאית.
- INT מספקת לסטודנטים את המעטפת המתאימה לחוויית לימודים מיטבית ופרקטית

מסלול זה מכין את בוגריו להשתלב בשוק העבודה עם ניסיון ממשי ויכולות אוטו-דיקטיות, הנדרשות מכל מי שמעוניין להשתלב בעולם ההייטק.

המכללה שומרת לעצמה את הזכות לערוך מעת לעת, לפי שיקול דעתה, שינויים בתכנית הלימודים, היקף שעות הלימוד, גל המדריכים וכד', ולא יראו בכל מידע המפורט בדפי מידע של המכללה כהתחייבות כלשהי מצד המכללה.

Data Science & AI

חשוב לדעת

פרויקט לתעשייה

במסגרת תוכנית הלימודים תוכלו להצטרף ל"פרוייקט בתעשייה" שנעשה בשיתוף פעולה עם סטארט-אפים, חברות ומיזמים מובילים בתעשיית ההיי-טק בו תגישו תוצר טכנולוגי מוגמר - משלב הרעיון ועד הפיתוח בפועל בחסות מרצים מנוסים שילוו את כל התהליך. מיזם זה יעניק לכם ניסיון מוכח בתכנון והבנה של פרוייקט, התמודדות עם אתגרים ומצבים מורכבים שעולים מן השטח, וחשיפה לגיוס והעסקה פוטנציאלית עבור סטודנטים המפגינים יכולות גבוהות תוך כדי עבודה פעילה בפרוייקט.

*כפוף להמלצת המרצה ובגיוס חברה מתאימה לשת"פ

Data Science & AI

מבנה התוכנית

תוכנית ההכשרה בנויה משישה פרקים נפרדים, במהלך כל פרק תעברו הכשרה תיאורטית ותדרשו ליישם את החומר הנלמד באמצעות פרויקטים מעשיים.

Module 1

Basic Machine Learning

60 hours

בחלק זה נכיר את המודלים הבסיסיים של למידת מכונה - כיצד לכתוב קוד לצורך בנית המודלים ואת הרקע התיאורטי שעומד מאחוריהם, בנוסף נכיר שיטות בסיסיות להערכת מודלים ושימוש בטכניקות בסיסיות של חלוקת המידע כדי ליצור מודלים עמידים יותר.

- Basic Python & Introduction to Machine Learning
- Linear models
 - Simple and multiple regression
 - Cost functions
 - Least square method
 - Gradient Descent
 - Analytical vs Numerical solutions
- K-nearest Neighbors Classifier
 - Distance metrics: Euclidean, Manhattan other alternatives
 - Bias-Variance Tradeoff
 - Implementations of the model and sci-kit learn
- Decision tree
 - Entropy, Information Gain, Gini Impurity
 - Pruning and Regularization
 - Visualization of DT

המכללה שומרת לעצמה את הזכות לערוך מעת לעת, לפי שיקול דעתה, שינויים בתכנית הלימודים, היקף שעות הלימוד, סגל המדריכים וכד', ולא יראו בכל מידע המפורט בדפי מידע של המכללה כהתחייבות כלשהי מצד המכללה.

Data Science & AI

- Random forest
 - Ensemble method and bootstrapping
 - Bagging
 - Feature importance and visualization
 - Hyperparameter tuning and optimization
- XGboost
 - XGboost Vs other boosting methods
 - Tree boosting, regularization
 - Learning rate and tree depth
 - Handling missing value

Module 2

Mathematics

100 hours

המתמטיקה מהווה את הבסיס ליכולות של מדען הנתונים, היא הלוגיקה והגיון מאחורי המודלים האלגוריתמיים. בחלק זה נלמד באופן פרקטי את הנושאים המרכזיים הדרושים להבנת האלגוריתמים והשימושים בהם שאותם נלמד בחלקים המתקדמים יותר.

- Probability
 - Basic Probability
 - Bayes theorem and conditional probability
 - Discrete and continuous probability distribution
 - Moments: expectation, variance and higher moments
- Statistics
 - Descriptive statistics: Mean, median, mode variance and standard deviation
 - Hypothesis testing, Confidence interval
 - Correlation and regression analysis
 - Central Limit Theorem

המכללה שומרת לעצמה את הזכות לערוך מעת לעת, לפי שיקול דעתה, שינויים בתכנית הלימודים, היקף שעות הלימוד, סגל המדריכים וכד', ולא יראו בכל מידע המפורט בדפי מידע של המכללה כהתחייבות כלשהי מצד המכללה.

Data Science & AI

- Linear algebra
 - Matrix operations: addition, subtraction, multiplication and inversion
 - Vector space
 - (Eigenvalues, eigenvectors and Singular Value Decomposition (SVD
- Calculus
 - Limit, continuity, differentiation, integration
 - Multivariable calculus
 - Fundamental theory and application in optimization problems

Module 3

Advanced Python and Machine learning concepts

44 hours

במסגרת הפרק, נכיר ספריות כגון Pandas וספריות דומות כדי לעבד את המידע לצורך ניקוי והכנת המידע ללמידת מכונה.
בנוסף, נשתמש בכלים שלמדנו לפני כדי לבדוק ולשפר את המודלים שלנו באמצעות כלים כגון:
Feature Engeneering, Data Cleaning

- Advance python techniques
 - Data Visualization with seaborn
 - Handling missing data with missingno library
 - Advanced imputation with IterativeImputer and KNNImputer
- Advance Data Cleaning Techniques
 - Outlier detection and treatment
 - Data transformation: Log and Box-Cox
- Working with SQL database
 - Basic SQL and non SQL database design
 - Using the SQLAlchemy

המכללה שומרת לעצמה את הזכות לערוך מעת לעת, לפי שיקול דעתה, שינויים בתכנית הלימודים, היקף שעות הלימוד, סגל המדריכים וכד', ולא יראו בכל מידע המפורט בדפי מידע של המכללה כהתחייבות כלשהי מצד המכללה.

Data Science & AI

- Feature Engineering
 - Categorical encoding: one hot, ordinal, binary
 - Feature scaling and normalization
 - Dimensionality reduction using Principal Component Analysis (PCA) and t-SNE
- Advance Overfitting Concepts
 - Bias variance trade-off
 - Regularization techniques
 - Advance overfitting techniques: data augmentation, early stoping

Module 4

Basic Neural Network and Deep Neural Network

60 hours

בחלק זה נלמד על אבני הבסיס של המודלים המתקדמים ללמידת מכונה עמוקה שעל בסיסה מתבססים המודלים המורכבים יותר.

- Introduction to Neural Network - NN
 - Neurons and activation functions
 - Feedforward and Backpropagation
 - Using the libraries: PyTorch and TensorFlow
- Convolutional Neural Networks - CNN
 - Convolutional layers, Pooling, fully connected layers
 - Application for image processing
 - CNN architecture: VGG, ResNet, Inception
- Recurrent Neural Networks - RNN
 - RNN and the vanishing gradient problem
 - (Long Short Term Memory (LSTM) and Gated Recurrent Unit (GRU
 - Sequence-to-sequence models

המכללה שומרת לעצמה את הזכות לערוך מעת לעת, לפי שיקול דעתה, שינויים בתכנית הלימודים, היקף שעות הלימוד, סגל המדריכים וכד', ולא יראו בכל מידע המפורט בדפי מידע של המכללה כהתחייבות כלשהי מצד המכללה.

Data Science & AI

- Current Applications of NN in the topics
 - NLP
 - Voice Recognition
 - Time Series Analysis

Module 5

Advanced ML & Deep Neural Network

32 hours

במסגרת הפרק, נלמד על ארכיטקטורות של מודלים עמוקים שמשמשות לשימושים שונים כגון NLP, Time Series Analysis, Image Processing. ואספקטים מתקדמים ופרקטיים יותר בלמידת מכונה שמטרתם שיפור המודלים והתוצאות בשטח.

- Languages and Chatbots
 - Basics Languages models
 - Word2Vec
 - Transofrmer architectures: attention mechanisim
- Generative Pre-Trained Transformer GPT
 - Languege Generative models
 - The GPT architecture
 - Application of GPT
- LLM, Chat GPT and open source alternatives
 - Overview and current status
 - Major Models: BLOOM, Bard, LLaMA2, Falcon
 - Finetuning applications
- Lang chain and Autonomies agents
 - Autonomies Agents concepts
 - Plugins and Langchain alternatives
 - Codeless solutions such as flowise

המכללה שומרת לעצמה את הזכות לערוך מעת לעת, לפי שיקול דעתה, שינויים בתכנית הלימודים, היקף שעות הלימוד, סגל המדריכים וכד', ולא יראו בכל מידע המפורט בדפי מידע של המכללה כהתחייבות כלשהי מצד המכללה.

Data Science & AI

במסגרת הפרק, נלמד על כלים ורעיונות כמו Python, Jupyter, Pandas, Numpy, SQL, Scikit-learn, Matplotlib
Git & Github, AWS Cloud platform, Docker, Streamlit

Final Project

24 hours

חלק זה הינו גולת הכותרת של ההכשרה!
נלמד כיצד להתמודד עם אתגרים מציאותיים מהעולם וכיצד לחבר נכון את הכלים שרכשנו כדי ליצור פתרונות פרקטיים.

Module 6

Advanced Topics

32 hours

לאחר סיום הפרויקט וההתמודדות עם האתגרים שנתקל בהם בעולם האמתי, נעמיק לשיטות מודרניות לפתרון האתגרים הללו באמצעות מודלים מתקדמים המשמשים את אנשי המקצוע השונים בתעשיית ההיי-טק.

- Model Interoperability
- Transfer learning
- generative adversarial networks (GANs)
- Deep reinforcement learning

Industry Project

40 hours

החלק המסכם של ההכשרה הוא פרויקט שבו נתלווה לצוות קיים מהתעשייה, נלמד על האתגרים שלו ונייצר מענה לאתגרים ממשיים העולים מן השטח.



המרכז הבינלאומי
ללימודי הייטק וחדשנות

בחסות
האוניברסיטה הפתוחה
מערך לימודי החוץ



***6377** | int-college.co.il