

מסלול ישיר לדוקטורט בשילוב עם המרכז לבינה מלאכותית ומדעי הנתונים

המרכז למדעי הנתונים ובינה מלאכותית מציע לסטודנטים.ות מצטיינים.ות המתחילים.ות את השנה הראשונה או השנייה של התואר השני לשלב לימודים בתחומי הבינה המלאכותית ומדעי הנתונים. המטרה היא להכשיר סטודנטים בתחומי דעת שונים למחקר המשלב כלים ושיטות חדשניות במדעי הנתונים והבינה המלאכותית.

קהל היעד: בוגרי ובוגרות תואר ראשון במדעים מדויקים, הנדסה או מדעי החיים (עם רקע חישובי מתאים) שהתקבלו לתואר שני או שסיימו שנה ראשונה של התואר השני בבתי הספר המשתתפים בתוכנית*, ומעוניינים להמשיך למסלול ישיר לדוקטורט.

הערות ספציפיות לבתי הספר השונים:

- במדעי החיים, תלמידים שמתחילים את השנה הראשונה יהיו במסלול "ישיר מהיר", ואלו שיתקבלו לאחר שנה של לימודי מוסמך יהיו שקולים למסלול ישיר לדוקטורט.

מה מקנים הלימודים:

- הכשרה מעמיקה בתחומי הבינה המלאכותית, למידת מכונה ומדעי הנתונים.
- מלגה של 200% לתקופה של לכל היותר 4 שנות לימודים או עד סיום הדוקטורט. המלגה מתחלקת ל- 100% מלגה מהמרכז לבינה מלאכותית והיתרה מבית הספר או המנחה לפי חוקי בית הספר.
- התואר שיתקבל הוא במסגרת ביה"ס שאליו שייך.ת הסטודנט.ית. בגיליון הציונים יופיעו הקורסים הרלוונטים.

מה נדרש מהסטודנטים.יות:

- קבלה כסטודנט.ית מן המניין לתואר שני באחד מבתי הספר שבהם התוכנית מאושרת, כלומר לא בלימודי השלמה, צבירה או תואר שני ללא תזה.
- לקבל את הסכמת המנחה או בית הספר להשלמת המלגה בהיקף של 100%.
- לימודים במסגרת הקדשת זמן מלא – לפי חוקי בית הספר.
- מציאת מנחה עד תום השנה הראשונה של התואר השני.
- מועמדים.ות בסיום השנה הראשונה של התואר השני צריכים לעבור לסטטוס של דוקטורט ישיר.
- עמידה בדרישות הקבלה של בית הספר לקבלה לדוקטורט.
- במסגרת השנה הראשונה של התכנית יש לסיים את קורסי החובה של התוכנית בציון ממוצע של 80 לפחות. אי עמידה בדרישות תבטל את הזכאות למלגה במסגרת המרכז.

הערות ספציפיות לבתי הספר השונים:

- במדעי החיים, התלמידים חייבים למצוא מנחה כבר מהשנה הראשונה.

סטודנטים. יות שלא ימשיכו לאחר השנה הראשונה לדוקטורט יוכלו לסיים תואר שני לפי חוקי בית הספר, ולא יהיו זכאים. יות למלגה במסגרת המרכז.

תנאי קבלה:

- סינון על-פי גיליון ציונים מכתבי המלצה וקו"ח.
- ראיון קבלה שיתקיים בשני מועדים: יולי וספטמבר 2024.

הקורסים:

שני קורסים מתקדמים בלמידה חישובית וסמינר מחקר שנתי אינטרדיסציפלינרי. הסילבוס המפורט של הקורסים מופיע בהמשך. הקורסים יהיו פתוחים רק לתלמידי התוכנית, או באישור מיוחד של המרצה על בסיס מקום פנוי. הקורסים מאושרים על ידי וועדות ההוראה של בתי הספר כקורסים לתואר שני, הסמינר מוכר כסמינר מחקר על ידי בתי הספר.

כיצד נרשמים:

סטודנטים. ות המעוניינים. ות להגיש מועמדות לתוכנית ישלחו מייל לכתובת: datascience@tauex.tau.ac.il בצירוף המסמכים הבאים:

- גיליון ציונים של התואר הראשון + של השנה הראשונה של התואר שני במידה וקיים.
- קורות חיים.
- מכתב המלצה ממנחה פרויקט בתואר הראשון/מהמנחה לתואר המתקדם במידה ויש.
- כל חומר רלוונטי אחר.

ניתן להגיש מועמדות עד לתאריך **15.6.24**

במידה ויתקיים סבב שני של ראיונות בספטמבר על בסיס מקום פנוי ניתן יהיה להגיש מועמדות עד 15.8.2024

*בתי הספר המשתתפים בתוכנית בשנת תשפ"ה: הנדסת חשמל, הנדסת תעשייה וניהול, בתי הספר בפקולטה למדעי החיים, מדעי המח, מדעי המחשב, מתמטיקה במסלול סטטיסטיקה.

פירוט הקורסים:

בתכנית יש להשלים שני הקורסים והסמינר הבאים:

- **An intensive introduction to machine learning** (3 hours) **1886.5100**
- **Deep learning and advanced topics** (3 hours) **1886.5101**
- **Seminar in AI and Data Science** (2 hours) **1886.5102**

Presentations by AI/DS scientists from the different departments on campus/industry/
students presentations.

את שאר הקורסים יש ללמוד בהתאם לדרישות תוכנית הלימודים של בית הספר שאליו רשומים.

סילבוסים של הקורסים:

1: An intensive introduction to machine learning 1886.5100

Weeks 1-5: A unified theoretical framework:

2 weeks: the statistical learning approach: decision theory, parametric vs local learning, curse of dimensionality, error decompositions (bias-variance), demonstration on nearest neighbors and linear regression, optimism and model selection

2 weeks: the PAC learning approach: ERM, Generalization bounds, VC dimension, Rademacher complexity

1 week: Summarizing the model selection problem.

Weeks 6-13: methods

0.5 weeks: quick introduction to convex optimization and duality, to be used for the rest of the course

1 week: statistical vs algorithmic/optimization view: logistic regression and linear SVM (likelihood vs optimization, unified view as loss-based learning, similarity in results)

1 week: regularization approaches: ridge/lasso, statistical properties, algorithms as time allows

1.5 weeks: Trees, Random Forest, Boosting.

1 week: Neural Networks. Deep Neural Networks: Introduction, Examples, Basic Experimenting

1 week: the generic view of embedding based learning: embedding in high dimension, regularization, computational tricks: boosting (trees, "L1", coordinate descent) and kernel machines (RKHS, "L2", representer)

2 weeks: intro to unsupervised learning: PCA, factor analysis, clustering: k-means, GMM=>EM, beyond linear: t-SNE, Auto Encoders, metric learning, random projections (if time permits)

2: Deep learning and advanced topics 1886.5101

Weeks 1-6: deep learning - deep dive

1. Fundamentals: backpropagation, SGD, computation, expressivity=>universality
2. Optimization / regularization: SGD=>L2, momentum, dropout
3. Architectures (2 weeks): CNN, RNN, transformers, algorithmic implications: back-propagation through time
4. Diffusion and generative modeling (2 weeks)

Weeks 7-13: Selected topics. Suggested examples:

1. Reinforcement learning
2. Statistics of Big Data: high dimensional modeling
3. Multiple testing and selective inference
4. Causality
5. Big data in science: genomics, particle physics, brain imaging
6. Modeling graphs and networks
7. Explainability / fairness / privacy
8. Information Theory approach