

## **דוקטורט ישיר בשיתוף עם המרכז למדעי הנתונים ובינה מלאכותית (TAD)**

המרכז למדעי הנתונים ובינה מלאכותית באוניברסיטת תל אביב מציע קורסים מתקדמים בתחום מדעי הנתונים ובינה מלאכותית ומלגות מוגדלות לסטודנטים מצטיינים. התוכנית מיועדת לסטודנטים/ות מצטיינים/ות בעלי תואר ראשון במדעים מדוייקים או הנדסה ותקנה לתלמידים ידע תיאורטי ומעשי בחזית המחקר בתחום מדעי הנתונים ובינה מלאכותית. התוכנית היא בשיתוף עם בתי הספר הבאים: מדעי-המחשב, מתמטיקה, הנדסת חשמל, מדעי המח ובתי הספר השונים בפקולטה למדעי החיים.

### **מטרות התוכנית:**

- להציע קורסים מתקדמים ועדכניים בבינה מלאכותית ומדעי הנתונים שינתנו על ידי חוקרים מובילים בתחומים אלה בשילוב עם תוכניות הלימודים של בתי הספר השותפים.
- לייצר קהילה של תלמידי מחקר מצטיינים שיעסקו במחקר בתחומי הבינה המלאכותית ומדעי הנתונים בבתי ספר שונים אשר שותפים לתוכנית.
- לקדם בינתחומיות במדעי הנתונים ובינה מלאכותית באמצעות:
  - חשיפת התלמידים/ות למחקר שנעשה בתחום הבינה המלאכותית ומדעי הנתונים בפקולטות השונות בקמפוס דרך סדנאות וסמינר שנתי.
  - תלמידים/ות הלומדים בתוכנית מגיעים מתוכניות לימודים שונות מרחבי הקמפוס (הנדסה, מדעי המחשב, מדעי המוח) ובכך יחשפו זה את זו לעולמות המחקר שלהם/ן
  - עידוד הנחיה משותפת של עבודות הדוקטורט
  - חיבור בין חוקרים/ות שעוסקים/ות בנושאי הליבה ופיתוח שיטות לאלה שעוסקים/ות ביישום של שיטות מתחום הבינה המלאכותית ומדעי הנתונים במגוון נושאים מדעיים.
  - עידוד תלמידים לקחת קורסים מחוץ לתחום המחקר שלהם (בתיאום עם בתי הספר השותפים)

### **הרשמה:**

התלמידים/ות ירשמו לתוכנית התואר השני/שלישי בבית הספר/חוג של חברת/ה הסגל שאיתו/ה הם מעוניינים/ות לבצע את עבודת התזה. במידה ועדיין לא נבחרו חברת/ה סגל להנחיה, ירשמו לתוכנית שתחום התוכן שלה תואם לרקע שלהם/ן או לתחומי העניין שלהם/ן.

### **דרישות קדם:**

התוכנית מיועדת לבוגרים/ות של הפקולטות למדעים מדוייקים והנדסה המעוניינים/ות לרכוש ידע ולעסוק בחזית המחקר בתחום מדעי הנתונים ובינה מלאכותית.

### **קבלה:**

הקבלה לתוכנית תקבע על ידי ועדה מטעם TAD לאחר שהסטודנטים/ות התקבלו לתוכנית של בתי הספר שאליו הם/ן נרשמו. התלמידים יתקבלו בשלב ראשון לבית הספר שאליו נרשמו ובית הספר/מנחה מתחייבים לתת לו 100% מלגה. TAD ישלים מלגה ל 100% נוספים. הסטודנטים יתחייבו ללמוד באופן מלא בתוכנית (לא ניתן לשלב עם עבודה).

### **קורסי השלמה:**

הועדה תקבע רשימה של קורסי השלמה בהתאם לרקע של התלמיד/ה שבאחריות התלמידים לעשות על מנת להצליח בקורסי התוכנית

### **מלגה:**

התלמידים/ות שיתקבלו לתוכנית יקבלו מלגה בגובה 200% שתמומן על ידי בתי הספר ו/או המנחה ומרכז TAD. התלמידים יידרשו לתרום לפעילות האקדמית של בתי הספר שאליהם ירשמו בהתאם לכללים המקובלים בבית הספר, על מנת להיות זכאים למלגה.

#### נקודות יציאה:

- ניתן יהיה לסיים עם תואר שני עם תזה במידה ותוכנית הדוקטורט לא תצא לפועל.
- במידה והתלמיד מעוניין להחליף מנחה ניתן יהיה לסיים את התואר השני באחד מבתי הספר ולעבור לבית הספר אחר ללימודי הדוקטורט. נציין שניתן יהיה להחליף מנחה (או אף להחליף בית ספר/חוג) באישור במהלך התוכנית ללא הפסקתה.
- לא ניתן לסיים את התוכנית ללא תזה

#### קורסים:

- תלמיד/ה שיתקבלו ידרשו לקחת את 3 קורסים הבאים (סילבוסים מצורפים למטה):  
חברי הסגל שילמדו את הקורסים יוכלו לקבוע את ההדגשים שינתנו בקורסים השונים.
- למידת מכונה ולמידה סטטיסטית (4 ש"ס סמסטריאלי)
  - נושאים ויישומים מתקדמים בלמידת מכונה ולמידה סטטיסטית (4 ש"ס סמסטריאלי)
  - סמינר לתלמידי התוכנית (2 ש"ס שנתי) שיכלול הרצאות אורח של חוקרים מרחבי הקמפוס / journal club /הצגת מחקרים של התלמידים בתוכנית אחת לשנה

את שאר הקורסים ייקחו התלמידים בהתאם לדרישות תוכנית הלימודים של בית הספר שאליו הם נרשמו להשלמה של מספר השעות הנדרשות.

#### סילבוסים של הקורסים

### **1: An intensive introduction to machine learning**

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 6-11: 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**

Weeks 1-6: deep learning - deep dive

1. Fundamentals: backprop, 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