



## סילבוס קורס

שם הקורס: סמינר מתקדם בלמידה חישובית

שם קורס באנגלית: Advanced Seminar in Machine Learning

מס' קורס: 202-2-5701

סוג קורס: בחירה

נק"ז: 2.0

מרצה הקורס: פרופ' אריה קנטורוביץ

דרישות קדם: הסתברות (למדעי המחשב או מתמטיקה)

### מטרת ונושא הקורס

Consider a function  $f$  on some domain  $X$  with range in  $[0, 1]$ . Suppose that  $n$  points are drawn iid from  $X$  under some distribution  $P$ . This induces two quantities associated with  $f$  and  $P$ : the expectation of  $f$  under  $P$ , denoted by  $Pf$ , and the sample average over the  $n$  random points, denoted by  $P_n f$ . When  $f$  is a single function, the almost sure convergence of  $P_n f$  to  $Pf$  follows from Hoeffdings inequality. Suppose, however, that we have a collection of functions,  $F$ , and are interested in convergence uniformly over  $F$ . When the latter holds,  $F$  is said to be a Uniform Glivenko-Cantelli (UGC) class. These lie at the heart of probability theory, statistics, and machine learning. For binary function classes, UGC is characterized by a combinatorial parameter known as the Vapnik-Chervonenkis (VC) dimension; this quantity is also of key importance in computational geometry and set theory. For function classes with a finite VC-dimension  $d$ ,  $P_n f$  converges to  $Pf$  at a rate of  $\sqrt{d/n}$ , and this is tight. The seminar will be devoted to proving this fundamental fact.

עבור משתנה מקרי חסום, ניתן בקלות להסיק התכנסות כמעט ודאית של ממוצע אמפירי לתוחלת. הסמינר יעסוק בהתכנסות במידה שווה על פני אוסף של משתנים מקריים, המאנדקסים ע"י משפחת פונקציות. את קצב ההתכנסות ניתן לאפיין ע"י מדד קומבינטרי הנקרא מימד ופניק-צרבוונקיס. אנו נוכיח את קצב ההתכנסות ההדוק.

### נושאי ההרצאות

- a characterization of PAC learnability in terms of VC-dimension
- [margin-based](#) bounds for support-vector machines
- [Rademacher](#) complexity bounds
- [margin](#) analysis of boosting
- uniform Glivenko-Cantelli and Donsker classes

## דרישות הקורס

השתתפות בכיתה  
הגשת עבודה או הצגת מאמר

## מרכיבי ציון הקורס

השתתפות בכיתה, הצגת מאמר בכיתה או הגשת עבודה

## ספרות הקורס

David Pollard. *Empirical processes: theory and applications*

Richard Dudley. *A course on empirical processes*

M. Vidyasagar. *Learning and Generalization With Applications to Neural Networks*

Vladimir N. Vapnik. *Statistical Learning Theory*

Martin Anthony, Peter L. Bartlett. *Neural Network Learning: Theoretical Foundations*

Michel Ledoux and Michel Talagrand. *Probability in Banach Spaces*