Spring 2018 **STAT 597: Nonparametric Methods and Learning Theory**

Time and Place: 9:05am-10:20am Tu, Th in 215 Thomas Building

Instructor: Dr. Bharath K. Sriperumbudur, 314 Thomas Building, `bks18@psu.edu`, 867-5948
Office Hours: 1.30pm-2.30pm (F) or by appointment.

Syllabus: The course aims to provide a graduate level mathematical introduction to nonparametric statistics with a major focus on estimation. The following are some of the topics that will be covered in the course.

- **Function estimation:** Estimation of cumulative distribution function (cdf), probability density function (pdf) and regression function using empirical, kernel and penalized estimators.

- **Consistency and convergence rates:** Theoretical gurantees for nonparametric function estimators will be established in terms of consistency and convergence rates. While consistency provides a basic gurantee that an estimator is reasonable, the convergence rate allows to compare different estimators.

- **Minimax theory and lower bounds:** General theory of minimax estimation will be covered using which information theoretic lower bounds for function estimation problems will be established. These bounds will establish the minimax optimality of the aforementioned estimators.

- **Introduction to statistical learning theory:** We will introduce the formalism of risk minimization and develop/study empirical risk minimzation and structural risk minimization (aka penalized/regularized risk minimization). Some examples of this framework include support vector machines for classification and regression, kernel ridge regression, etc. An introduction to reproducing kernel Hilbert spaces will be provided and some kernel algorithms such as kernel PCA, kernel independence measures will be introduced. Time pertaining, more discussion on kernel methods will be made in addition to topics in nonparametric hypothesis testing.

The mathematical tools required for this course span various areas such as functional analysis, empirical process theory, approximation theory and high-dimensional probability (particularly concentration inequalities). The required basics will be covered to make the course self-contained.

Textbook: There is no SINGLE textbook for the course. Given the vastness of the subject area, in fact there is no good textbook available. The following are some good references for the course.

- **Concentration Inequalities:**
  - (C1) *Concentration Inequalities: A Nonasymptotic Theory of Independence*, S. Boucheron, G. Lugosi and P. Massart, Oxford University Press.
  - (C2) *Concentration-of-measure inequalities.* Lecture notes by G. Lugosi.

- **Empirical Process Theory:**

(E1) *Weak Convergence and Empirical Processes*, A. van der Vaart and J. Wellner, Springer.

(E2) (C1).

- **Nonparametric Estimation:**

  (N1) *Introduction to Nonparametric Estimation*, A. Tsybakov, Springer.

  (N2) *A Probabilistic Theory of Pattern Recognition*, L. Devroye, L. Gyröfi and G. Lugosi, Springer.

  (N3) *A Distribution-Free Theory of Nonparametric Regression*, L. Györfi, M. Kohler, A. Krzyzak and H. Walk, Springer.

- **Learning Theory:**

  (S1) *Introduction to Statistical Learning Theory.* Lecture notes by O. Bousquet, S. Boucheron and G. Lugosi.

  (S2) *On the Mathematical Foundations of Learning*, F. Cuker and S. Smale, Bulletin of the AMS.

- **Kernel Methods:**

  (K1) *Support Vector Machines*, I. Steinwart and A. Christmann, Springer.

  (K2) *Reproducing Kernel Hilbert Spaces in Probability and Statistics*, A. Berlinet and C. Thomas-Agnan, Kluwer-Academic.

  (K3) *Kernel mean embedding of distributions: A review and beyond*, K. Muandet, K. Fukumizu, B. K. Sriperumbudur and B. Scholkopf, Foundations and Trends in Machine Learning.

Homework: Overall, 5-6 homeworks will be assigned for the entire course and the details will be announced on Canvas. **Late assignments will not be accepted, regardless of reason.** You are encouraged to work in groups, but you are expected to write up the solutions in your own words.

Project: A major component of the course is a project and presentation. It is expected that the project will have some theoretical component. Literature survey in a particular area is also considered as a valid project. Final project report should be LaTeXed in 11pt font. It should be at least 10 pages long and should not exceed 20 pages. The project can be carried out in groups of two and will include a presentation, the details of which will be announced on Canvas.

Grading: Class participation (10%), Homework (30%), Project (40%), Presentation (20%)

Integrity: All Penn State and Eberly College of Science policies regarding academic integrity apply to this course. See `http://www.science.psu.edu/academic/Integrity/` for details.

Climate: The Eberly College of Science Code of Mutual Respect and Cooperation (`www.science.psu.edu/climate/Code-of-Mutual-Respectfinal.pdf`) embodies the values that we hope our faculty, staff, and students possess and will endorse to make the Eberly College of Science a place where every individual feels respected and valued, as well as challenged and rewarded.

Disabilities: Penn State welcomes students with disabilities into the Universitys educational programs. If you have a disability-related need for reasonable academic adjustments in this course, contact the Office for Disability Services (ODS) at 814-863-1807 (V/TTY). For further information regarding ODS, please visit the Office for Disability Services web site at `http://equity.psu.edu/ods/`.

Emergencies: Campus emergencies, including weather delays, are announced on Penn State Live (`http://psutxt.psu.edu/`) and communicated to cellphones, email, the Penn State Facebook page, and Twitter via PSUTXT (to sign up, please see `http://psutxt.psu.edu`)