Proposal for a new course

Title: Special topics in CS: Mathematics for Machine Learning

Course No: CS698AB

Units: 3-0-0-9

Proposer: Harish Karnick, Prateek Jain, Adjunct Faculty (Microsoft Research, Bangalore)

Others interested in teaching course: -

Pre-requisites: CS201, CS203, (or CS602), CS210 (or equivalent). MSO201 (or equivalent knowledge) is desirable.

About the course:
The course will cover the important mathematical concepts, methods and results that are widely used and needed in machine learning. It has three major sub-parts.

Matrix analysis and introductory functional analysis:

- Vector spaces, metrics, norms, inner products, linear transformations and properties.
- Banach spaces, inner product spaces, reproducing kernel Hilbert spaces, kernels, PSD matrices and properties, elementary properties of linear transformations.
- Basic spectral theory, eigenvalue decompositions, SVD and applications, Lanczos method.
- Matrix norms, spectral norms, elementwise and mixed norms, induced norms, matrix inversion, least squares and pseudo-inverse.
- Basic matrix analysis.

Probability theory:

- Probability as a measure, sample space, $\sigma-$algebra, pmf, pdf, elementary properties, conditional probabily, independence.
- Random variables, expectation, moments, law of large numbers, central limit theorem.
- Concentration inequalities - Markov, Tchebyshev, Bennett, Chernoff, Azuma-Höffding, McDiarmid etc. Johnson-Lindenstraus lemma.
- Matrix concentration bounds.
- Rademacher complexity, generalization bounds, uniform convergence.
- Design of experiments and basic hypothesis testing.

Convex analysis and programming:

- Convex sets and functions, lower and upper semi-continuous sets, Jensen's inequality.
- Lagrange multipliers, duality theory, Fenchel's duality, dual norms, KKT conditions.
- LP, Farkas' lemma, QP and S-lemma, cone programming, SDP.
- Gradient descent, Newton's method, conjugate gradient descent.

References:

1. Lloyd N Trefethen, David Bau III, Numerical Linear Algebra, SIAM, 1997.
2. Gene H Golub, Charles F Van Loan, Matrix Computations, 3rd Ed., John Hopkins Univ. Press, 1996.
3. David A Harville, Matrix Algebra From a Statistician's Perspective, Springer, 1997. (Reference for results from Matrix Algebra).
4. Erwin Kreyszig, Introductory Functional Analysis with Applications, Wiley, 1978.
5. William Feller, An Introduction to Probability Theory and its Applications, 3rd Ed., Wiley, 1968.
6. David Stirzaker, Probability and Random Variables: A Beginner's Guide, Cambridge Univ. Press, 2003.
7. Geoffrey R Grimmet, David Stirzaker, Probability and Random Processes, 3Ed., Oxford Univ. Press, 2001.
8. Stephen Boyd, Lieven Vandenberghe, Convex Optimization, Cambridge Univ. Press, 2004.
9. A Ben-Tal, Arkadi Nemirovski, Lectures on Modern Convex Optimization: Analysis, Algorithms, Engineering Applications, SIAM, 2001.
10. Various resources on the internet - lecture notes, videos, papers etc.

Proposers signature: Harish Karnick, Prateek Jain (Adjunct Faculty).

Convenor, DPGC **Chairman, SPGC**