Gaussian Processes for the Social Sciences

Koç University

Summer 2021

Syllabus

Instructor
David Carlson
dcarlson@ku.edu.tr
Office: CASE 140

Class Schedule
TBA

Office Hours
TBA

Online Access
https://ku.blackboard.com (for grades and readings)
https://github.com/carlson9/GPindep (for in-class material)

Introduction

This course is designed to primarily familiarize the student with the use of Gaussian processes (GPs) for social science research. GPs are adept as both a machine learning tool and a statistical inference tool, and both will be covered in the course. Topics include but are not limited to measurement problems, classification, time-series analyses, cross-sectional analyses, prediction, estimating counter-factuals, imputation, and causal GPs. The course includes an individual final project, and co-authoring a project with the professor. These are the only two components of the grade, but regular progress is expected. Most instruction will be done in Python, but some Stan and `R` may be utilized. Previous programming experience is not assumed.
Prerequisite: INTL 450 (or equivalent).

Required Book

This class requires extensive reading to be adequately prepared for class. We will go over more recent advances, but the primary text is fundamental and freely available:

Rasmussen, C.E. & C.K.I. Williams. *Gaussian Processes for Machine Learning*, the MIT Press, 2006, ISBN 026218253X. Massachusetts Institute of Technology. www.GaussianProcess.org/gpml.

Requirements and Grading

Grades will not be rounded, these represent strict cut-offs. In the rare event of, for example, exactly a 90, the higher grade will be assigned. Pluses and minuses will be applied at the instructor's discretion and will only be used if there are clear separations within a given grade. **Note that the Koç suggested grades are not followed in this course.**

A	90–100
B	80–90
C	70–80
D	60–70
F	<60

1) *Final Project: 50%*

   A major component of the course is to develop methodological and presentation skills for the development of your research. Rather than work on unrelated research questions, the knowledge obtained in the course should be applied to your topic of interest. The methodological rigor and presentation style will be the key determinants of your grade for this section. In order to track your progress and keep you on-track, we will discuss your goals throughout the semester. If, at the end of the semester, you do not have a working thesis, you will be required to write a report detailing your progress and submit this for a final grade.

2) *Co-Authored Project: 50%*

   The other component of the grade is to work with the professor on a project, with the final goal a publishable article. This may include conference or other presentations. Because

this will be co-authored, extensive contribution is expected, and progress must be made throughout the semester.

Course Schedule

Please note this schedule is subject to change.

Week 1: Introduction to Background, GP Regression

Readings: R&W Introduction, Appendix A, Appendix B, Chapter 2

Week 2: Classification

Readings: R&W Chapter 3

Week 3: Covariance Functions, Model Selection and Adaptation of Hyperparameters

Readings: R&W Chapter 4 and 5

Week 4: Relationships between GPs and Other Models, Theoretical Perspectives

Readings: R&W Chapter 6 and 7

Week 5: Approximation Methods for Large Datasets

Readings: R&W Chapter 8 and 9

Week 6: Flexibly Building a GP in Stan and Our Own Sampler

Readings: Selected articles on Blackboard

Week 7: Causal GPs, Other Extensions

Readings: Selected articles on Blackboard

Week 8: Touching Base on Projects and Review

Readings: Selected articles on Blackboard