# MATH 7339 - Machine Learning and Statistical Learning Theory 2

**Section 0  Introduction**

**Instructor: He Wang**

**Goal of the course:**

1. Further study of the theory and methods of advanced machine learning models and frameworks

2. More on the sequence data: time series and natural language processing

3. Project/presentation/independent study/lab/paper reading: skills

## ➢ Course topics:

1. **Advanced machine learning and statistical learning theory**
   - **Overview of machine learning models, Bias-variance trade-off.**
   - **Model Complexity/Hyperparameters/degrees of freedom/VC dimension**
   - **Exponential families and generalized linear methods.**
   - **Splines**
   - **Kernel Smoothing Methods**
   - **Bayesian Methods, Bayesian inference/ Bayesian Statistics/ methods**
   - **Kernel Methods**
   - Factor Analysis and PCA and ICA
   - All Markov: Markov chain, Markov Monte Carlo, Markov decision processes

2. **Time Series and Forecasting**
   - **Overview, intro and examples,**
   - **AR ; MA ; ARIMA**
   - **RNN for time series data**
   - **FFT**
   - **Facebook: Forecasting at scale**

3. **Natural Language Processing**
   - **Vectorization (Word Embeddings)**
   - **RNN,NN for sequences**
   - **Transformer**
   - **Bidirectional Encoder Representations from Transformers (BERT)**

**Tentative Schedule:**

Week 1-Week 7 (Advanced machine learning)

Week 8 (Feb22) Midterm Presentation

Week 8-Week 12 Times Series  (Week 9 is Spring Break)

Week 12-14 NLP

April 12, 19 Final project Presentation.

## ➢ Grades distribution:

1. **Homework.** (20%) There will be a few homework questions on the theory of machine learning
2. **Computer labs** (25%) (There will be a few computer labs focusing on the implementation of algorithms on real world data sets.)
3. **Attendance and Class discussion participation (**10%)
4. **Midterm: Theorical topics report/presentation/Paper review/paper presentation (**15%**)** For the paper presentation, (1) summarize the paper, (2) discuss the paper's strengths and weaknesses, and (3) discuss the paper's impact. (1-2 students)
5. **Final project**. (30%) The final project be a computational analysis of a data set using sufficiently complicated or novel techniques from this course. It consists of a proposal, middle stage progress report, project report and presentation (with poster or slides). Project groups should contain 2-5 students.

- **Midterm report/presentation**

Suggested theoretical topics:

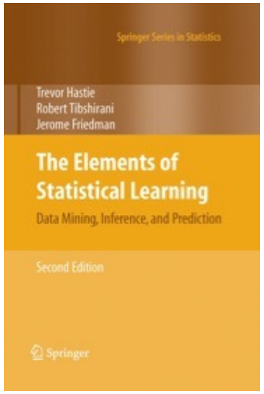
i. Casual Inference.
ii. A/B testing.
iii. Metric, Hilbert Space. (RKHS_ Reproducing kernel Hilbert space,)
iv. Sampling: Monte Carlo/ Markov Monte Carlo (Bishop's book)
v. EM Algorithm
vi. Gaussian Processing
vii. Graphical model
viii. A topic in Murphy1 or Murphy2 or Bishop not covered in class.
ix. Study and report a few papers in one topic.

- **Final project**.

Time series data (ARIMA, Spectral**, etc. Prophet** )
NLP data (RNN, Bert, Transformer, …)
Others

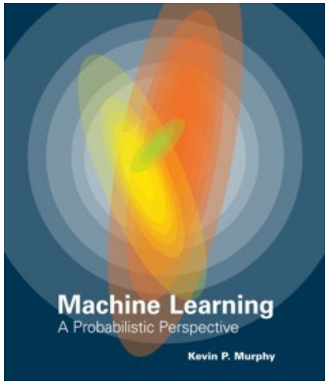
Here are some references used along the class.(In particular, we refer the abbreviation)
More references will be provided along each section.

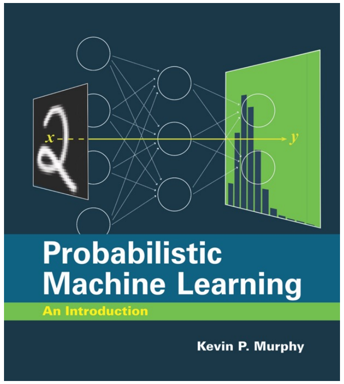
## ❑ Advanced Machine Learning References:



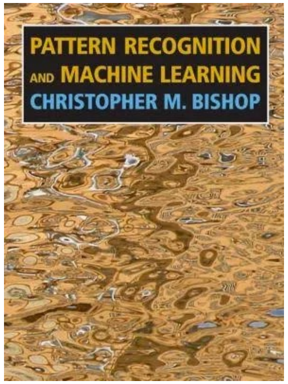

- **[Hastie] or [HTF]** *The Elements of Statistical Learning: Data Mining, Inference, and Prediction* by Trevor Hastie, Robert Tibshirani, Jerome Friedman
  https://web.stanford.edu/~hastie/ElemStatLearn/

- [**Murphy 0**] *Machine Learning: A Probabilistic Perspective* by Kevin P. Murphy
  https://probml.github.io/pml-book/book0.html




- [**Murphy 1**] Probabilistic Machine Learning: An Introduction by Kevin P. Murphy
  https://probml.github.io/pml-book/book1.html




- **[Bishop]** *Pattern Recognition and Machine Learning* by Christopher Bishop

  https://www.microsoft.com/en-us/research/uploads/prod/2006/01/Bishop-Pattern-Recognition-and-Machine-Learning-2006.pdf

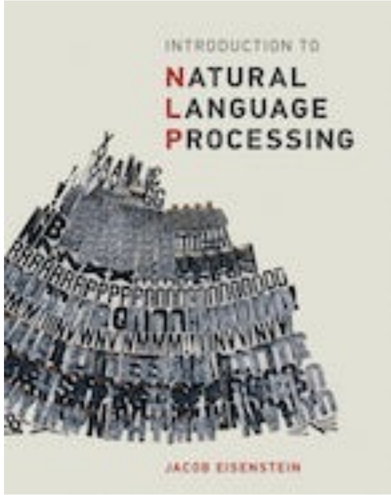
## ❑ **Time Series References**:

- "Time Series Analysis and Its Applications", 4th ed. 2017, by Shumway and Stoffer.
- "Introduction to Time Series and Forecasting", 3rd ed. 2016, by Brockwell and Davis.

**Facebook: Forecasting at scale: https://facebook.github.io/prophet/**

## ❑ **NLP References**:

Jacob Eisenstein: *Introduction to Natural Language Processing* (MIT Press, 2019)




Jurafsky and Martin: *Speech and Language Processing*

https://web.stanford.edu/~jurafsky/slp3/

## More online sources:

**Famous Deep Learning Papers:**
https://papers.baulab.info/00_README.html

**Paper With Code:**
https://paperswithcode.com/sota

Berkeley-Statistical Learning Theory Readings
https://www.stat.berkeley.edu/~bartlett/courses/2014spring-cs281bstat241b/readings.html