University of Rochester

DSCC 275/475: Time-Series Analysis and Forecasting for Data Science

Fall 2024

**Location and Time:**

Tuesdays and Thursdays 11:05 am -12:20 pm

Dewey Hall 2162

**Instructor**

Prof. [Ajay](http://www.cs.rochester.edu/u/jluo) Anand, Wegmans Hall Rm 1203, ajay.anand@rochester.edu  
Office hours: Posted on Blackboard under Announcements

**TAs/Office hours:**

Shah, Shyam [sshah77@UR.Rochester.edu](mailto:sshah77@UR.Rochester.edu)

Maisa, Aashrith <amaisa@UR.Rochester.edu>

*(Office hours will be posted on Blackboard)*

**Course description**

Time series analysis is a valuable data analysis technique in a variety of industrial (e.g., prognostics and health management), business (e.g., financial data analysis) and healthcare (e.g., disease progression modeling) applications. Moreover, forecasting in time series is an essential component of predictive analytics. The course will begin with an introduction to practical aspects relevant to time series data analysis such as data collection, characterization, and preprocessing. Topics covered will include smoothing methods (moving average, exponential smoothing), trend and seasonality in regression models, autocorrelation, AR and ARIMA models applied to time-series data. Deep learning models including feedforward, recurrent, gated and convolutional architectures will also be studied. Students shall work on projects with time-series data sets using modeling tools in Python/R.

Prerequisites: Introductory Statistics (DSC 262/STT212/STT213 or equivalent), Linear Algebra and Differential equations (MTH 165 or equivalent), and applied Python programming (CSC161 or equivalent)

**Required Text book:**

**(1st half of course) *Introduction to Time Series Analysis and Forecasting (Wiley Series in Probability and Statistics) 2nd Edition by*** [*Douglas C. Montgomery*](https://www.amazon.com/s/ref=dp_byline_sr_book_1?ie=UTF8&text=Douglas+C.+Montgomery&search-alias=books&field-author=Douglas+C.+Montgomery&sort=relevancerank)*(Author),*[*Cheryl L. Jennings*](https://www.amazon.com/s/ref=dp_byline_sr_book_2?ie=UTF8&text=Cheryl+L.+Jennings&search-alias=books&field-author=Cheryl+L.+Jennings&sort=relevancerank)*(Author),*[*Murat Kulahci*](https://www.amazon.com/s/ref=dp_byline_sr_book_3?ie=UTF8&text=Murat+Kulahci&search-alias=books&field-author=Murat+Kulahci&sort=relevancerank)*(Author)*

* Series: Wiley Series in Probability and Statistics
* Hardcover: 672 pages
* Publisher: Wiley-Interscience; 2nd edition (April 27, 2015)
* Language: English
* ISBN-10: 1118745116
* ISBN-13: 978-1118745113

**Additional Reading:**

Sanchez, J. (2023). *Time Series for Data Scientists: Data Management, Description, Modeling and Forecasting*. Cambridge: Cambridge University Press.

Woodward, W.A., Sadler, B.P., & Robertson, S. (2022). Time Series for Data Science: Analysis and Forecasting (1st ed.). Chapman and Hall/CRC. https://doi.org/10.1201/9781003089070

Practical Time Series Forecasting with R: A Hands-On Guide [2nd Edition] (Practical Analytics) 2nd Edition by Galit Shmueli (Author), Kenneth C. Lichtendahl Jr (Author)

*Deep Learning (Adaptive Computation and Machine Learning series) by*[*Ian Goodfellow*](https://www.amazon.com/Ian-Goodfellow/e/B01MQGN8N0/ref=dp_byline_cont_book_1)*(Author),*[*Yoshua Bengio*](https://www.amazon.com/Yoshua-Bengio/e/B00IWC47MU/ref=dp_byline_cont_book_2)*(Author),*[*Aaron Courville*](https://www.amazon.com/Aaron-Courville/e/B01N8XGWRL/ref=dp_byline_cont_book_3)*(Author)*

* Series: Adaptive Computation and Machine Learning series
* Hardcover: 800 pages
* Publisher: The MIT Press (November 18, 2016)
* Language: English
* ISBN-10: 0262035618
* ISBN-13: 978-0262035613
* Available <https://www.deeplearningbook.org/>

*Pattern Recognition and Machine Learning (Information Science and Statistics) by*[*Christopher M. Bishop*](https://www.amazon.com/Christopher-M.-Bishop/e/B001IGLMNY/ref=dp_byline_cont_book_1)*(Author) (available* <https://www.microsoft.com/en-us/research/uploads/prod/2006/01/Bishop-Pattern-Recognition-and-Machine-Learning-2006.pdf>*)*

**Course schedule (chapters refer to the textbook)**

|  |  |
| --- | --- |
| **Topic** | **Textbook Chapter Reference** |
| Introduction to Time-series Analysis and Forecasting | (Chap. 1, Montgomery) |
| Statistics introduction to Time-series Analysis | (Chap. 2, Montgomery) |
| Smoothing Methods (MA, Exponential Smoothing) | (Chap. 4, Montgomery) |
| Auto-regressive Techniques (AR, ARMA, ARIMA) | (Chap. 5, Montgomery) |
| Time-series Forecasting using Machine Learning Techniques | (Chap. 8, Shmueli) and Notes/slides |
| Additional Topics (e.g., Spectral analysis) | (Chap. 7, Montgomery) and Notes |
| Deep Learning for Time-series Analysis | (Notes; Chap. 5, Goodfellow and Chap. 3-4 Bishop) |
| Introduction to Neural Networks (optimization, backpropagation, regularization) | (Notes; Chap. 6-8, Goodfellow and Chap. 5 Bishop) |
| Recurrent Neural Networks | (Notes; Chap. 10, Goodfellow) |
| Gated Recurrent Neural Networks | (Notes; Chap. 10, Goodfellow) |
| AutoEncoders for Classification with Time-series Data | Notes |
| Attention Mechanisms and Transformers (\*NEW\*) | Notes |
| Temporal Convolutional Neural Networks | Notes |

### Course Schedule and Outline:

(Will be posted on a Google Doc as a live link with updates)

**Grading**

Grading (total 100%) [Graduate students will have additional problems in select assignments]

* Homework assignments: 20% (3 HWs) – 1st half of course
* Midterm: 20%
* Project #1: 20% (based on material covered until mid-term)
* Project #2 and #3 [3.1 (RNNs/LSTMs) and 3.2 (AutoEncoders)] (Deep Learning): 40%

Project #1 involves applying classical statistical models to time-series data and make predictions over different time intervals, and quantitatively evaluate the performance to determine the best performing model.

Project examples for #2 and #3 include: 1) Electric consumption forecasting using hourly and daily consumption data. 2) Anomaly detection using data from wearable sensors. These projects apply Deep Learning architectures including LSTMs, Auto-encoders. Detailed project description and additional guidelines will be provided when project is assigned.

*Homeworks and Projects* are due at 11:59 pm on the due date.**Late** **points** will be subtracted (2% per hour). Must submit on **Blackboard** online. Only the latest submission is considered.

Graduate students will have additional questions in homeworks and projects to receive graduate level credit.

Re-grading deadline is 1 week from the date graded assignment is returned.

Exam policy: No make-up exams will be given. If a student has to miss an exam for an unavoidable circumstance, please discuss with the instructor as soon as possible.

**Letter Grades**

Note: The below grading thresholds are nominal values. They are subject to change based on circumstances in the semester at discretion of the instructor.

A: >90  
B: 80-89  
C: 70-79

D: 60-69  
E: <60.  
  
Within each letter grade, sub grades such as B+, B, B- will be assigned at equal intervals.

**Programming Experience**

The HWs and the projects require implementing time series algorithms based on variety of techniques you will learn in this course. You are welcome to choose a programming language of your choice for the HWs and Project #1 (e.g. Python, R, Matlab). Project #2 and #3 is based on using Deep-learning libraries in Python. The goal of the projects will be to develop predictive models on sequential data by applying different deep learning strategies and hyper-parameter tuning.

**Academic Honesty Policy**

*PLEASE COMPLETE THE SURVEY ON BLACKBOARD ASAP (Under “Course HomePage”)!!*

The University of Rochester academic honesty policy applies to all assignments and exams for this class. The full-text of the academic honesty policy can be found at: http://www.rochester.edu/College/CCAS/AdviserHandbook/AcadHonesty.html

In examinations, you must work individually with no communication with others and use only materials/tools that have been explicitly allowed. For homework, you may discuss problems with your colleagues, but final solutions need to be worked out, written and submitted individually. Any external material used should be clearly cited. In your own writings (example term papers, homework solutions, proposals etc), no more than one or two sentences may be used verbatim from any source. READ THESE INSTRUCTIONS CAREFULLY! If any aspect of the academic honesty policy and guidelines for this course are unclear, please ask the instructor. Lack of awareness or understanding of this policy will not be an acceptable excuse or defense against disciplinary action. You may also be in violation if you excessively rely on AI “Co-Pilot” systems to assist you with writing your code.

Notes on Plagiarism: Plagiarism is a serious offense and is in violation of university policy. If you are unsure of what constitutes plagiarism in written documents, a good description can be found here: <https://rochester.edu/college/gradstudies/assets/pdf/Plagiarism_Misconduct.pdf> Plagiarism does not just occur in written documents; it also occurs in code. It is unacceptable (and it is considered plagiarism) to copy code developed by others and submit it as your own. If you do consult any online sources of code, you must properly attribute the corresponding sections in your code to their original source, as you would add quotations, footnotes, or references in a written document.

### Credit Hour Policy:

This course follows the College credit hour policy for four-credit courses. This course meets two times weekly for three academic hours per week. The course requires students to complete homework, projects, and a mid-term exam. The students are expected to complete supplementary work averaging 8 hours per week. The student activities on the homework and projects would include developing algorithms, writing code, explaining concepts learned in class, and preparing documentation.