



STA 3241 Statistical Learning Syllabus

Spring Semester 2022

Course Information

- **Course Number and Title:** STA 3241 Statistical Learning
- **Credit Hours:** 3 credits
- **Current Academic Term:** Spring 2022

Instructor Information

- **Instructor:** Sravani Vadlamani, PhD
- **Office:** IST 2040
- **Office Hours:** TR: 12:30 – 1:30 PM, W: 1:30 – 2 :30PM (In-person or Teams), or by appointment
- **Office Phone:** 863-874-8806
- **E-mail:** svadlamani@floridapoly.edu
- **Class Meeting Day, Time & Location:** TR, 2 – 3:15 PM EST, IST 1065

Course Details

- **Class Delivery Mode:** The class will be delivered in a face-to-face format where the students are expected to attend all their scheduled university classes to satisfy all academic objectives as defined by the instructor.
- **Course Website:** <https://floridapolytechnic.instructure.com/courses/6098>
- **Official Catalog Course Description:** This is an introductory-level course in supervised learning. Topics include classification and regression, cross-validation and bootstrap, model selection, dimension reduction, tree-based methods, random forests and boosting, support-vector machines, principal components, and cluster analysis. Students will have hands-on experience in model building, machine learning, and implementation.
- **Course Pre and/or Co-Requisites:** QMB 3200 - Advanced Quantitative Methods or (MAS 3114 - Computational Linear Algebra AND STA 2023 - Statistics 1)
- **Communication/Computation Skills Requirement (6A-10.030):** No
- **Required Texts:**
 - *(ISLR) Gareth James, Daniela Witten, Trevor Hastie, and Robert Tibshirani. "An Introduction to Statistical Learning - with Applications in R" Second Edition.* Available at: <https://www.statlearning.com/> **ISBN-13:** 978-1071614174
 - Supplementary materials: "R for Data Science" by Garrett Grolemund and Hadley Wickham accessible at <https://r4ds.had.co.nz/>

Equipment and Materials: This course is an application-driven introduction to statistical learning. We will use the R programming language and RStudio, both are free. The course covers fundamental and popular R packages for statistical learning, introduced as working examples. The format of the course will include lectures by the instructor, class discussions, directed readings, and students' presentations.

- **Suggested:** You can simply create an account in <https://rstudio.cloud> to access a cloud-based version of RStudio to work on your assignments/projects
- **Alternative:** a local installation of R can be completed by downloading R from the [R Project web site](#). After installing R, a free and open-source Integrated Development Environment (IDE) for R can be downloaded [from the RStudio web site](#).
- **Course Objectives:** The specific objective is for students to understand and apply basic *concepts and methods in statistical learning*. The broader objective is for students hone their ability to *apply modern techniques of data analysis* to provide insights that help in making real world decisions.
- **Course Learning Outcomes:**
 1. Explain statistical learning methodology.
 2. Implement the techniques covered, interpret and understand results, and validate models.
 3. Monitor performance of ongoing implementations where appropriate.
 4. Effectively communicate the results of model implementation orally and in writing.
- **Alignment with DS Program Outcomes**

Data Science Program Student Outcomes	Course Learning Outcome and Learning Level*			
	1	2	3	4
(1) Apply current data science concepts, techniques, and practices to solve complex problems.	Comprehension	Application	Evaluation	
(2) Analyze a given data science problem and formulate a solution in terms of the datasets needed, the techniques required or the technologies to be utilized.		Application	Evaluation	
(3) Communicate effectively insights, analysis, conclusions, or solutions to a diverse audience.				Create

*: learning level as described in Bloom's taxonomy and Anderson and Krathwohl's taxonomy.

Academic Support Resources

- **Library:** Students can access the Florida Polytechnic University Library through the University website and [Canvas](#), on and off campus. Students may direct questions to Academic Success Center success@floridapoly.edu or by email, library@floridapoly.edu.
- **ASC:** The Academic Success Center provides essential services that directly support the student experience at Florida Polytechnic University. Located on the first floor of the Innovation, Science & Technology Building in room 1019 and at ASC East in Phase 2 Dorms, Academic Support Services is a hub that connects the community with the resources needed to succeed academically. The desk is staffed by success coaches who provide academic coaching on a variety of topics, including time management, test preparation, and test-taking skills. Success coaches also provide academic guidance and help students manage their schedule and academic progress. Additionally, coaches lead career development initiatives on campus and are available

to review resumes and conduct mock interviews. Students may direct questions to success@floridapoly.edu.

Course Policies:

- **Attendance:** Students in face-to-face courses are expected to attend all their scheduled University classes and to satisfy all academic objectives as defined by the instructor. Attendance is highly correlated with student motivation and performance and is highly encouraged. Students whose absences exceed those allowed by the course syllabus may see a reduction in their final letter grade. If you know that you will miss a class for any reason discuss the situation with your instructor BEFORE the class is missed. Only valid and documented excuses will be considered. Any documentation must be submitted within 1 week of your absence. Class attendance will be taken at every class period. Students are responsible for arranging to make up work missed due to a valid, excused class absence. You will not be able to make up any pop quizzes. The university considers attendance a vital part of the learning process and as a result, there will be a penalty for unexcused missed classes. Also, due to the disruption caused to your classmate's learning process late arrivals will likewise incur a penalty. As a result, this course is implementing the following final grade reduction policy for absences or late arrivals:

Number of unexcused missed meetings	Effect on final grade
0 classes	3 points extra credit
1 class	1 points extra credit
2 or more classes	0 points extra credit

- **Students Feeling Sick: I am a student; what should I do if I think I may have COVID-19?**
Students who are showing symptoms or who have been exposed to COVID-19 are expected to stay in their residences (at home or in their dorm rooms) and immediately notify the FL Poly CARE manager at care@floridapoly.edu. The CARE manager will work with each student to triage their individual situation and will notify faculty of students who are not attending courses due to COVID-19 symptoms.
- **Class Participation:** Asking and answering questions and solving problems in class is strongly encouraged.
- **In-Class Quizzes:** In-class pop quizzes will be given occasionally to assist students in practicing problem-solving and conceptual understanding. These will be graded and account for a portion of your final grade.
- **Grading Scale:** The following grading scale will be used for this class.

A	93% – 100%	B	83% – <86%	C	73% – <76%	D	63% – <66%
A–	90% – <93%	B–	80% – <83%	C–	70% – <73%	D–	60% – <63%
B+	86% – <90%	C+	76% – <80%	D+	66% – <70%	F	0% – <60%
- **Assignment/Evaluation Methods:**
 - Exams: There will be one midterm exam.
 - Assignments: There will be several homework assignments that may involve the use of the statistical program Minitab, SAS, R or Excel.
 - Final Project: This is a data analysis assignment that involves application of the concepts

- on a data set and reporting the inferences from the data analysis.
- Final Exam: There will be a cumulative final exam.

Assignment	Percentage
Attendance & Participation	10%
Assignments	20%
Quizzes	10%
Data Analysis Project	20%
Midterm Exam	20%
Comprehensive Final Exam	20%
Total	100%

Participation in all course activities is a very important element of this course and is a basic expectation. Course participation consists of active and respectful involvement in class discussions, peer feedback, postings, replies, projects, and other interactions. The participation grade considers quality, quantity, and timeliness of student participation.

- **Late Work/Make-up work:** All class assignments will have due dates communicated at the time of assignment. **It is the student's responsibility to know the deadlines and turn work in ON TIME.** The make-up policy applies only to missed midterm exams. If you miss an exam with an excused absence, you must meet with your professor as soon as possible to deal with the issue. In most cases your final score will be calculated with that exam, you missed, so it is in your best interest to not miss any exams. In-class quizzes and the Final exam cannot be made up. If you have a schedule conflict for the final exam let your professor know as soon as possible.
- **Lecture Expectations:** Lecture meets for seventy-five minutes, twice per week in person. The intent of lecture time is for you to develop your conceptual understanding and practice problem-solving. The lecture will be interactive - you are expected, at appropriate times, to work with your neighbor, express your thoughts, ask, and answer questions, discuss ideas, patiently listen to and respect other's ideas.
- **CANVAS Policy:** Assignments, announcements, and information will be posted on CANVAS. **Students are responsible for checking CANVAS regularly to be aware of their assignments** and other class information.
- **Email Policy:** All students are required to use **studentuserID@floridapoly.edu** email system (most preferable) OR the CANVAS e-mail system to communicate with the instructor. On occasion, email may be used to disseminate important class-related assignments, announcements, and information. Students are responsible for any information or assignments given in e-mail.

University Policies

Basic rules for in the classroom, IST, and Campus:

1. We highly recommend, until further notice, that you always wear your face-covering during class and throughout the building.
2. Absolutely **no eating or drinking** during class.
3. Study hard and engage in all your courses!

Academic Integrity: All students must commit to the highest ethical standards in completion of all academic pursuits and endeavors, whether in classroom or online environments: [Academic Integrity](#).

“Behaviors of academic dishonesty in violation of this policy are listed below and are not intended to be all inclusive. Violations may result in the imposition of academic sanctions under this regulation and/or disciplinary sanctions under the Student Code of Conduct.

- Intentionally using or attempting to use unauthorized materials, information, or study aids in any type of academic exercise.
- Intentionally or knowingly representing the words or ideas of another as one’s own in any academic exercise.
- Intentional and unauthorized falsification or invention of any information or citation in an academic exercise.
- Multiple Submission. Submission of the same or substantially the same work for credit in two or more courses. Multiple submissions shall not include those situations where the instructor gives the student prior written approval to use such prior academic work or endeavor.
- Facilitating Academic Dishonesty. Intentionally or knowingly assisting or attempting to assist another in violating any provision of this regulation.
- Misconduct in Research and Creative Endeavors. Serious deviation from the accepted professional practices within a discipline or from the policies of the University in carrying out, reporting, or exhibiting the results of research or in publishing, exhibiting, or performing creative endeavors. This does not include honest error or honest disagreement about the interpretation of data.
- Misuse of Intellectual Property. Illegal use of copyright materials, trademarks, trade secrets, or intellectual properties.”

Reasonable Accommodations: Florida Polytechnic University is committed to assisting students with disabilities and offering reasonable accommodations to those with documented eligibility. The Office of Disability Services (ODS) coordinates accommodations for students with disabilities in accordance with the ADA Amendments Act of 2008 (ADAAA), the Americans with Disabilities Act of 1990 (ADA), and Section 504 of the Rehabilitation Act of 1973. Reasonable accommodations are determined on an individual basis through an interactive process between you, ODS, and your instructor(s). If you have already registered with ODS, please ensure that you have requested an accommodation letter for this course and communicate with your instructor about your approved accommodations at your earliest convenience. If you are not registered with ODS but believe you have a temporary health condition or permanent disability requiring an accommodation, please contact ODS as soon as possible at DisabilityServices@floridapoly.edu or (863)874-8770 or ASC East building or [ODS website](#): www.floridapoly.edu > Student Affairs > Health Wellness > Disability Services

Accommodations for Religious Observances, Practices and Beliefs: The University will reasonably accommodate the religious observances, practices, and beliefs of individuals in regard to admissions, class attendance, and the scheduling of examinations and work assignments. (See [University Policy](#).)

Title IX: Florida Polytechnic University is committed to ensuring a safe, productive learning environment on our campus that prohibits sex discrimination and sexual misconduct, including sexual harassment, sexual assault, dating violence, domestic violence and stalking. It is important for you to know that there are resources available if you or someone you know needs assistance. You may speak to your professor, but your professors have an obligation to report the incident to the Title IX Coordinator. It is an educational goal that you feel able to share information related to your life experiences in classroom discussions and in one-on-one meetings. However, it is requirement for university employees to share information with the Title IX Coordinator regarding disclosure. However, please know that your information will be kept private to the greatest extent possible. You will not be required to share your experience. If you want to speak to someone who is permitted to keep your disclosure confidential, please seek assistance from the Florida Polytechnic University [Ombuds Office](#), BayCare's Student Assistance Program, 1-800-878-5470 and locally within the community at [Peace River Center](#), 863-413-2707 (24-hour hotline) or 863-413-2708 to schedule an appointment.

Statement of Academic Continuity: In the event of an emergency (such as a hurricane), it may be necessary for Florida Poly to suspend normal operations. During this time, Florida Poly may opt to continue delivery of instruction through methods that include but are not limited to the Learning Management System (Canvas), online conferencing, email messaging, and/or an alternate schedule. It is the responsibility of the student to monitor the Learning Management System for each class for course-specific communication, and the Florida Poly website and emails for important general information.

Student Record of Lectures: Students may, without prior notice, record video or audio of a class lecture for a class in which the student is enrolled for their own personal educational use. Recordings may not be used as a substitute for class participation or class attendance. Recordings may not be published or shared in any way, either intentionally or accidentally, without the written consent of the faculty member. Failure to adhere to these requirements is a violation of state law (subject to civil penalty) and the student code of conduct (subject to disciplinary action). *Recording class activities other than class lectures, including but not limited to lab sessions, student presentations (whether individually or part of a group), class discussion (except when incidental to and incorporated within a class lecture), and invited guest speakers is prohibited.* For further information, go to [the Registrar's webpage](#) and click on [HB233 Guidance](#).

Course Schedule

Date	Topics and Readings	Suggested Textbook Problems
01/11	Overview of statistical learning, RStudio setup and intro	Ch. 2: 2, 4, 8, 10
01/13	Canvas resources, Fundamentals of linear algebra overview	
01/18	Review of Linear Regression, Algorithms overview	Ch.3: 1, 3, 8, 9, 15
01/20	Other considerations in regression models, Examples, and applications	

Date	Topics and Readings	Suggested Textbook Problems
01/25	Overview of classification problem, Review of logistic regression	Ch. 4: 1 and lab
01/27	Multiple logistic regression, Examples, and applications	
02/01	Linear Discriminant Analysis (LDA), Examples and Applications	Ch 4: 5, 12, 13 (a) – (f), 14
02/03	Quadratic Discriminant Analysis (QDA), Comparison of methods	
02/10	Cross-validation, k-fold, leave-one-out, Bootstrapping	Ch. 5: 3, 5, 8, 9
02/15	Linear Model Selection, Best subset selection, stepwise selection	Ch. 6: 1, 8 (a) – (d)
02/17	Ridge regression formulation and examples, The LASSO formulation and examples	Ch. 6: 2, 9, 11
02/22	Selecting the tuning parameters, Examples, and applications	
02/24	Principal Component Regression, Partial Least Squares	Ch. 6: 9, 11
03/01	Considerations in high dimensions, Examples, and applications	
03/04	MIDTERM EXAM	Ch. 7: 6, 7, 9, 10
03/08 & 03/10	SPRING BREAK – NO CLASSES	
03/15	Polynomial regression, Generalized additive models	
03/17	Basics of decision trees, Fitting regression and classification trees	Ch. 8: 8, 9
03/22	Bagging, Random Forests, Boosting, Examples and applications	
03/24	Explainable models with decision trees, More on random forests and tree-based models	Ch. 8: 10, 11
03/29	Maximal Margin Classifier, Support Vector Classifiers	Ch. 9: 5, 6
03/31	Multiclass Support Vector Machines (SVMs), Relationship with logistic regression	Ch. 9: 7, 8
04/05	Kernels and SVMs, Case studies	
04/07	Principal Component Analysis, Examples, and applications	Ch. 12: 8, 10

Date	Topics and Readings	Suggested Textbook Problems
04/12	Other uses for principal component methods, Examples, and applications	
04/14	Clustering methods overview, K-means algorithm Examples and applications, Problems in unsupervised learning	Ch. 12: 9, 13
04/19	Hierarchical clustering, Examples and applications, Practical issues in clustering, Dissimilarity measures	
04/21	Introduction to Deep Learning	Ch 10
04/26	Final project presentations	
04/30 & 05/01 – 05/05	Final Exams week (04/30 & 05/01 – 05/05)	

****This is a tentative schedule and I reserve the right to modify this schedule as required by the progression of the class.*

Quizzes

The course will have at most ten in class quizzes that are either announced before or impromptu. The objective of these quizzes is to ensure students are keeping up with the course material.

Final Project

In the final project you will show your knowledge and skills in statistical learning, using any combination of the different tools and topics discussed throughout the semester applied to an area/field of your interest.

- *Final Project Report*
Your goal is to submit a cohesive project report that conveys that you have mastered the techniques discussed during the semester.
- *Final Project Presentation*
An important aspect of doing research is taking time to share your findings with others. We will give everyone 10 minutes to share their final project and summarize their findings. *The final project presentation accounts for 15% of your final project grade.*

Your instructor will provide you with specific guidelines for the final project report and final project presentation shortly after the first few weeks of classes (format and length, call for proposals, reference materials, presentation guidelines and logistics, rubric, etc.)

Sample final project topics from previous years include:

- Unsupervised learning of IMDB movie ratings
- Regression analysis of nutritional indicators from fast-food restaurant menus
- Classification model for competitive eBay auctions
- Predictive analytics using eSports games data
- Trend analysis and characterization of the Tour de France winners
- Predicting early career earning after graduation

Rubric for Project Report and Presentation

The final presentations and reports will be evaluated using the rubrics included below.

Report Rubric

Objective	Category	Below Expectations	Weak	Average	Good	Excellent
	Score	1	2	3	4	5
Students can write professional quality documents	Introduction	Opening is off-topic and inappropriate to the purpose, not concise and no clarity	Opening is somewhat related to the topic and appropriate to the purpose but is not concise and clear	Opening is related to the topic and appropriate to the purpose. Somewhat clear and concise	Opening is related to the topic and appropriate to the purpose. Clear and concise	Strong opening that is clear and concise
	Organization	Disorganized; incorrect format; unclear direction	Somewhat organized; incorrect format; unclear direction	Organized; correct format; unclear direction	Organized; correct format; clear direction	Correct formatting, strong clarity and organization in the development of main points
	Literature Review	Does not present information from any source	Presents information from irrelevant sources representing limited points of view/approaches	Presents information from relevant sources representing limited points of view/approaches	Presents in-depth information from relevant sources representing limited points of view/approaches	Synthesizes in-depth information from relevant sources representing limited points of view/approaches
	Research Design (weighted twice)	Does not provide information on research design	Inquiry design demonstrates misunderstanding of the methodology or theoretical framework	Critical elements of the methodology or theoretical framework are missing, incorrectly developed or unfocused	Critical elements of the methodology or theoretical framework are appropriately developed however, more subtle elements are ignored or unaccounted for	All elements of the methodology or theoretical framework are skillfully developed and may be synthesized from across disciplines or relevant subdisciplines
	Analysis (weighted twice)	Incorrect, Irrelevant, no supporting evidence	Correct, irrelevant, no supporting evidence	Correct, relevant, no supporting evidence	Relevant and correct with supporting evidence	Relevant, correct, complete, incorporates innovative insights
	Next Steps	Missing or content does not support conclusion	Conclusion irrelevant to the findings	Conclusion somewhat relevant to the findings	Conclusion relevant to the findings	Strong conclusion that is clear, complete and compelling
	Grammar & Spelling	Uses language that often impedes meaning due to errors	Uses language that often sometimes meaning due to errors	Uses language that generally conveys meaning to readers with clarity, although writing includes some errors	Uses straightforward language that conveys meaning to readers. Language has few errors	Uses graceful language that communicates meaning to readers with clarity and fluency and is virtually error free
	Reference Style (APA)	Did not follow APA style	Numerous errors in APA style, did not cite sources correctly, formatting issues	Some errors in APA style, cited correctly but formatting issues persist	Minimum errors in style and formatting but does not detract from readability	No errors in APA style
Total points for Report = 50						

Presentation Rubric

Objective	Category	Below Expectations	Weak	Average	Good	Excellent
	Score	1	2	3	4	5
Students can demonstrate mastery of communication technology	Use of Media	Lack of media detracts from the presentation objective	Misuse of media that detracts from the presentation objective	Use of media barely supports and contributes to the presentation objective	Use of media supports and contributes to the presentation objective	Use of media supports, clarifies and reinforces the presentation objective
	Quality of Slides	Very poor quality. Not enough or too much colors, fonts and animations that detract from project objective	Poor quality. Not enough or too much colors, fonts and animations that detract from project objective	Fonts, colors and animations barely support the presentation objective	Fonts, colors and animations support the presentation objective	Fonts, colors and animations support, clarify and reinforce the presentation objective
Students can develop and deliver a compelling oral talk with relevant facts and information	Opening statement	Opening is off-topic and inappropriate to the purpose, not concise and no clarity	Opening is somewhat related to the topic and appropriate to the purpose but is not concise and clear	Opening is related to the topic and appropriate to the purpose. Somewhat clear and concise	Opening is related to the topic and appropriate to the purpose. Clear and concise	Strong opening that is clear and concise
	Organization	Disorganized; incorrect format; unclear direction	Somewhat organized; incorrect format; unclear direction	Organized; correct format; unclear direction	Organized; correct format; clear direction	Correct formatting, strong clarity and organization in the development of main points
	Literature Review	Does not present information from any source	Presents information from irrelevant sources representing limited points of view/approaches	Presents information from relevant sources representing limited points of view/approaches	Presents in-depth information from relevant sources representing limited points of view/approaches	Synthesizes in-depth information from relevant sources representing limited points of view/approaches
	Analysis	Incorrect, Irrelevant, no supporting evidence	Correct, irrelevant, no supporting evidence	Correct, relevant, no supporting evidence	Relevant and correct with supporting evidence	Relevant, correct, complete, incorporates innovative insights
	Next Steps	Missing or content does not support conclusion	Conclusion irrelevant to the findings	Conclusion somewhat relevant to the findings	Conclusion relevant to the findings	Strong conclusion that is clear, complete and compelling
	Timing	Presentation is too short, insufficient coverage of material	Presentation is too long. Unable to cover all the material	Able to cover all the material within five extra minutes	Utilizes allotted time to provide sufficient coverage of material	Well-paced coverage of material within the allotted time
Students can deliver an oral talk with clarity and appropriate poise	Delivery Techniques	Does not participate in the oral presentation	Delivery techniques (posture, gesture, eye contact, and vocal expressiveness) detract from the understandability of the presentation, and speaker appears uncomfortable.	Delivery techniques (posture, gesture, eye contact, and vocal expressiveness) make the presentation understandable, and speaker appears tentative.	Delivery techniques (posture, gesture, eye contact, and vocal expressiveness) make the presentation interesting, and speaker appears comfortable.	Delivery techniques (posture, gesture, eye contact, and vocal expressiveness) make the presentation compelling, and speaker appears polished and confident.
	Peer Evaluation	5 points				
Total Points = 50						

Important Dates

January 10	M	First Day of Classes
January 10 -14	M-F	Drop/Add Week
January 14	F	Withdrawal Deadline - No Academic or Fee Liability
January 17	M	Martin Luther King Jr. Holiday - No Classes
January 31	M	Spring/Summer 2022 Graduation Application Deadline
February 8	T	Day Break (Career Day) – No Classes
March 5 - 13	Sat-Sun	Spring Break – No Classes
March 7	M	Mid-term Grades Due
April 15	F	Withdrawal without Academic Penalty Deadline (W grade assigned)
April 27	W	Last Day of Classes
April 28-29	Th & F	Reading Days - No Classes
April 30, May 1-5	Sat, M-Th	Final Exams
May 8	Sun	Commencement
May 11	W	Final Grades Available Online