**STAT 3348.001 Course Syllabus**  
**2023 / Fall**  
**Predictive Modeling**

**Instructor:** Jacob Turner, Ph.D.  
**Department:** Mathematics and Statistics  
**Email:** turnerja2@sfasu.edu  
**Phone:** 936-468-1692  
**Office:** 342 NM  
**Office Hours:**

<table>
<thead>
<tr>
<th>Monday</th>
<th>Tuesday</th>
<th>Wednesday</th>
<th>Thursday</th>
<th>Friday</th>
</tr>
</thead>
<tbody>
<tr>
<td>11am-12:30pm</td>
<td></td>
<td>11am-12:30pm</td>
<td></td>
<td>Or by appointment</td>
</tr>
</tbody>
</table>

**Class meeting time and place:** MW 2:30pm - 3:45pm, Mathematics 202

**Text and Materials:**

<table>
<thead>
<tr>
<th>Introduction to Statistical Learning</th>
<th>James, Witten, Hastic, &amp; Tibshirani</th>
<th>Free Online</th>
<th>Springer</th>
</tr>
</thead>
<tbody>
<tr>
<td>R statistical software</td>
<td>Free Download Online</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Course Details:**  
**Course Objectives:**  
A broad overview of various statistical methods with an emphasis in predictive modeling and R programming. Topics will include statistical learning concepts and terminology, classification, advanced feature selection methods, bootstrap and cross validation procedures, tree algorithms, unsupervised learning. In addition to the general skill set, a great emphasis will be placed on understanding how to conduct an analysis workflow from start to finish, from exploratory work, model development, and validation.

**Course outline (Approximate time spent)**

**Course Topics:**

- **What is statistical learning?**  
  - Exploratory Analysis  
  - Key Definitions  
  - Need for Validation  
  - The Bias/Variance Trade Off  
  Approximate time spent: 15%

- **Multiple Linear Regression**  
  - Estimation and assumptions  
  - Pitfalls and Shortcomings  
  - K-Nearest Neighbors  
  Approximate time spent: 15%

- **Classification**  
  - Logistic Regression  
  - Linear Discriminate Analysis  
  - Intro to ROC Curves  
  Approximate time spent: 25%

- **Resampling Methods**  
  Approximate time spent: 15%

*Updated: August 2023*
• Cross Validation
• Bootstrap
• Feature Selection Using LASSO
• Tree Based Methods
  o CART models
  o CART vs Regression
  o Random Forrest
• Unsupervised Learning and Data Visualization
  o Clustering methods
  o Intro to PCA
  o Visualizations for Unsupervised learning

Student Learning Outcomes (SLO): At the end of STAT 3348, a student who has studied and learned the material should be able to:
1. Exhibit a firm understanding of how to enter, manipulate, analyze, and visualize data within R statistical software.
2. Identify and apply correct modeling procedures and strategies to different data types depending on the study design
3. Assess model and prediction performance through resampling techniques and correct interpretation of ROC curves
4. Explain the general concept of the trade offs between model complexity and interpretability
5. Explain the difference between how resampling can play a role in feature selection and how that potentially differs from estimating prediction performance
6. Apply appropriate unsupervised analysis techniques to aid in data descriptions, reduction, and visualization
7. Use machine learning algorithms such as random forest or support vector machines as an alternative to traditional predictive modeling techniques.

Grading: The final average will be computed using the following weights:

- Homework and additional to dos 30%
- 2 projects 40%
- Midterm 15%
- Cheat Sheet Final Exam 15%

Homework: Homework assignments will be given out of the text and potentially some standalone work that I will provide. It is very important that you stay up on trying to master some of your R capabilities and the HW will help you to practice. I highly recommend you take a look at the end of each chapter in the text book. The authors provide very nice R examples and tutorials for many of the things that we will be doing.

Midterm and Final: The midterm and final will focus on definitions, key statistical learning concepts, small mathematical derivations, and the general information about each method covered. Depending on how the class go, some material may be based on a few assigned readings.

Projects: There will be two projects during the course. These will be group projects of 2-3 students. The first project will be more technical in nature to help facilitate the early concepts and techniques of the course. The second project will be a contest among the groups on a big data modeling and prediction. Dr. Turner will have a hold out data set to assess each groups predictive capabilities. Bonus points will be allotted for the winning group! For each project, the students will present their work and approach to the class via a ppt presentation. R code must be provided as a secondary document.
Grading Scale:

<table>
<thead>
<tr>
<th>Letter Grade</th>
<th>% Scale</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>90-100</td>
</tr>
<tr>
<td>B</td>
<td>80-89</td>
</tr>
<tr>
<td>C</td>
<td>70-79</td>
</tr>
<tr>
<td>D</td>
<td>60-69</td>
</tr>
<tr>
<td>F</td>
<td>below 59</td>
</tr>
</tbody>
</table>

Expectations and Class Rules:
ALL cell phones are to be turned OFF and put away. If someone violates this rule (texting, playing games, etc.), then the person must leave the class room for that class period without any further delay according to the professor’s instructions. No food in the classroom. Don’t leave the classroom in the middle of the lecture.

Attendance Policy:
Be on time. There is no need for an attendance policy in advanced mathematics courses. They are self-regulating.

Formalities Which Must Be Included in This Document:

Attendance
This is a graduate level class, and I do not expect attendance issues. If you know you are going to have to miss a specific class, please let me know via e-mail or phone prior to the class.

Academic Integrity
The Code of Student Conduct and Academic Integrity outlines the prohibited conduct by any student enrolled in a course at SFA. It is the responsibility of all members of all faculty, staff, and students to adhere to and uphold this policy.

Articles IV, VI, and VII of the new Code of Student Conduct and Academic Integrity outline the violations and procedures concerning academic conduct, including cheating, plagiarism, collusion, and misrepresentation. Cheating includes, but is not limited to: (1) Copying from the test paper (or other assignment) of another student, (2) Possession and/or use during a test of materials that are not authorized by the person giving the test, (3) Using, obtaining, or attempting to obtain by any means the whole or any part of a non-administered test, test key, homework solution, or computer program, or using a test that has been administered in prior classes or semesters without permission of the Faculty member, (4) Substituting for another person, or permitting another person to substitute for one’s self, to take a test, (5) Falsifying research data, laboratory reports, and/or other records or academic work offered for credit, (6) Using any sort of unauthorized resources or technology in completion of educational activities.

Plagiarism is the appropriation of material that is attributable in whole or in part to another source or the use of one’s own previous work in another context without citing that it was used previously, without any indication of the original source, including words, ideas, illustrations, structure, computer code, and other expression or media, and presenting that material as one’s own academic work being offered for credit or in conjunction with a program course or degree requirements.

Collusion is the unauthorized collaboration with another person in preparing academic assignments offered for credit or collaboration with another person to commit a violation of any provision of the rules on academic dishonesty, including disclosing and/or distributing the contents of an exam.

Misrepresentation is providing false grades or résumés; providing false or misleading information in an effort to receive a postponement or an extension on a test, quiz, or other assignment for the purpose of obtaining an academic or financial benefit for oneself or another individual or to injure another student academically or financially.

Withheld Grades Semester Grades Policy (5.5)
Ordinarily, at the discretion of the instructor of record and with the approval of the academic chair/director, a grade of WH will be assigned only if the student cannot complete the coursework because of unavoidable circumstances.
Students must complete the work within one calendar year from the end of the semester in which they receive a WH, or the grade automatically becomes an F. If students register for the same course in future terms the WH will automatically become an F and will be counted as a repeated course to compute the grade point average. For additional information, go to https://www.sfasu.edu/policies/course-grades-5.5.pdf.

**Students with Disabilities**
To obtain disability-related accommodations, alternate formats, and/or auxiliary aids, students with disabilities must contact the Office of Disability Services (ODS), Human Services Building, and Room 325, 468-3004 / 468-1004 (TDD) as early as possible in the semester. Once verified, ODS will notify the course instructor and outline the accommodation and/or auxiliary aids to be provided. Failure to request services promptly may delay your accommodations. For additional information, go to http://www.sfasu.edu/disabilityservices/.

**Student Wellness and Well-Being**
SFA values students’ overall well-being, mental health and the role it plays in academic and overall student success. Students may experience stressors that can impact both their academic experience and their personal well-being. These may include academic pressure and challenges associated with relationships, emotional well-being, alcohol and other drugs, identities, finances, etc.

If you are experiencing concerns, seeking help, SFA provides a variety of resources to support students’ mental health and wellness. Many of these resources are free, and all of them are confidential.

**On-campus Resources:**

**The Dean of Students Office** (Rusk Building, 3rd floor lobby)
www.sfasu.edu/deanofstudents
936.468.7249
dos@sfasu.edu

**SFA Human Services Counseling Clinic** Human Services, Room 202
www.sfasu.edu/humanservices/139.asp
936.468.1041

**The Health and Wellness Hub** “The Hub”
Location: corner of E. College and Raguet St.

To support the health and well-being of every Lumberjack, the Health and Wellness Hub offers comprehensive services that treat the whole person – mind, body and spirit. Services include:
- Health Services
- Counseling Services
- Student Outreach and Support
- Food Pantry
- Wellness Coaching
- Alcohol and Other Drug Education

www.sfasu.edu/thehub
936.468.4008
thehub@sfasu.edu

**Crisis Resources:**
- Burke 24-hour crisis line: 1.800.392.8343
- National Suicide Crisis Prevention: 9-8-8
- Suicide Prevention Lifeline: 1.800.273.TALK (8255)
- johCrisis Text Line: Text HELLO to 741-741
STAT 3346–Experimental Design Analysis
Course Syllabus

Course description: Analysis of variance, single factor completely randomized designs, blocking and Latin square designs. Multifactor experiments including factorial experiments, nested, blocked and split-plot designs, analysis of covariance. Quality control, acceptance sampling, reliability issues. SAS or other statistical software used throughout. Report writing, data driven problems and/or case studies incorporated throughout.

Credit hours: 3

The following is an excerpt from SFA Policy 5.4:
The federal definition of a credit hour is an amount of work represented in intended learning outcomes and verified by evidence of student achievement that is an institutionally established equivalency that reasonably approximates:

1. Not less than one hour of classroom or direct faculty instruction and a minimum of two hours out-of-class student work each week for approximately fifteen weeks for one semester or trimester hour of credit, or 10 to 12 weeks for one quarter hour of credit, or the equivalent amount of work over a different amount of time, or;

2. At least an equivalent amount of work as outlined in item 1 above for other academic activities as established by the institution including laboratory work, internships, practica, studio work, and other academic work leading to the award of credit hours.

To this end, all students in courses offered by the Department of Mathematics and Statistics that wish to be successful should plan to spend a minimum of two hours outside of class for every credit hour associated with this course. Expected activities to be completed in the time outside of class include reviewing notes from previous class meetings, reading assigned course resources, completing all assigned exercises and projects, and performing periodic assessment preparation.

Course Prerequisites and Corequisites: STA 3342

Course outline: Approximate time spent

- Single Factor Experiments 35%
  - Completely Randomized Designs (CR)
  - Multiple Comparison Procedures
  - Randomized Complete Block Designs (RCB)
  - Latin Square Designs
  - Unbalanced and Incomplete Structures in Designs

- Multi-Factor Experiments 35%
  - Factorial Treatment Structure in a CR Design
    - Interaction Assessment
  - Factorial Treatment Structure in a RCB Design
    - Interaction Assessment
  - Nested Treatment Structures
  - Repeated Measures Designs
  - Split-Plot Designs
  - Covariance (Analysis of)

- Special Topics 10%
  - Quality Control
    - Control Charts for Attributes
Control Charts for Variables
  o Acceptance Sampling and General Sampling Theory
  o Reliability in Industrial Settings

- **Statistical Consulting/Report Writing**  20%
  o The Consultant/Client Relationship
  o Writing a Statistical Report
  o Statistical Computing and Software Usage
  o Presenting (Written and Oral) Statistical Reports

*Incorporated throughout the previous three "bullets" as opposed to a stand alone subject or unit.

**Student Learning Outcomes (SLO):** At the end of STAT 3348, a student who has studied and learned the material should be able to:

1. Distinguish between CR and RCB designs.  [EEO: 1, 6]
2. Plan and analyze CR, RCB and other statistical experiments.  [EEO: 1, 2, 4, 5]
3. Compare and contrast popular multiple comparison techniques for experiments.  [EEO: 1, 5, 6]
4. Explain the nature of factorial treatment structure and know how to properly assess interaction effects in statistical models.  [EEO: 1, 2, 5]
5. Distinguish nested from crossed factors and analyze data resulting from experiments that contain each type of factor.  [EEO: 1, 2, 5, 6]
6. Plan and analyze repeated measures and split-plot experiments.  [EEO: 1, 2, 4, 5]
7. Distinguish analysis of variance from analysis of covariance situations in nature.  [EEO: 1, 2, 4, 5, 6]
8. Recognize physical scenarios which are matched to the experiments discussed throughout the course.  [EEO: 1, 7]
9. Write both statistical reports and brief summaries for client usage.  [EEO: 3, 7]
10. Demonstrate a basic understanding of the client/consultant relationship during consulting meetings.  [EEO: 3, 7]

*There are no specific program learning outcomes for this major addressed in this course. It is a general education core curriculum course and/or a service course.*

**Exemplary Educational Objectives (EEO):**

1. To apply arithmetic, algebraic, geometric, higher-order thinking, and statistical methods to modeling and solving real-world situations.
2. To represent and evaluate basic mathematical information verbally, numerically, graphically, and symbolically.
3. To expand mathematical reasoning skills and formal logic to develop convincing mathematical arguments.
4. To use appropriate technology to enhance mathematical thinking and understanding and to solve mathematical problems and judge the reasonableness of the results.
5. To interpret mathematical models such as formulas, graphs, tables and schematics, and draw inferences from them.
6. To recognize the limitations of mathematical and statistical models.
7. To develop the view that mathematics is an evolving discipline, interrelated with human culture, and understand its connections to other disciplines.

**Academic Integrity**

The Code of Student Conduct and Academic Integrity outlines the prohibited conduct by any student enrolled in a course at SFA. It is the responsibility of all members of all faculty, staff, and students to adhere to and uphold this policy.

Articles IV, VI, and VII of the new Code of Student Conduct and Academic Integrity outline the violations and procedures concerning academic conduct, including cheating, plagiarism, collusion, and misrepresentation. Cheating includes, but is not limited to: (1) Copying from the test paper (or other assignment) of another student,
(2) Possession and/or use during a test of materials that are not authorized by the person giving the test, (3) Using, obtaining, or attempting to obtain by any means the whole or any part of a non-administered test, test key, homework solution, or computer program, or using a test that has been administered in prior classes or semesters without permission of the Faculty member, (4) Substituting for another person, or permitting another person to substitute for one's self, to take a test, (5) Falsifying research data, laboratory reports, and/or other records or academic work offered for credit, (6) Using any sort of unauthorized resources or technology in completion of educational activities.

Plagiarism is the appropriation of material that is attributable in whole or in part to another source or the use of one's own previous work in another context without citing that it was used previously, without any indication of the original source, including words, ideas, illustrations, structure, computer code, and other expression or media, and presenting that material as one's own academic work being offered for credit or in conjunction with a program course or degree requirements.

Collusion is the unauthorized collaboration with another person in preparing academic assignments offered for credit or collaboration with another person to commit a violation of any provision of the rules on academic dishonesty, including disclosing and/or distributing the contents of an exam.

Misrepresentation is providing false grades or résumés; providing false or misleading information in an effort to receive a postponement or an extension on a test, quiz, or other assignment for the purpose of obtaining an academic or financial benefit for oneself or another individual or to injure another student academically or financially.

**Withheld Grades Semester Grades (SFA Policy 5.5)**
Ordinarily, at the discretion of the instructor of record and with the approval of the academic chair/director, a grade of WH will be assigned only if the student cannot complete the course work because of unavoidable circumstances. Students must complete the work within one calendar year from the end of the semester in which they receive a WH, or the grade automatically becomes an F. If students register for the same course in future terms the WH will automatically become an F and will be counted as a repeated course for the purpose of computing the grade point average. The circumstances precipitating the request must have occurred after the last day in which a student could withdraw from a course. Students requesting a WH must be passing the course with a minimum projected grade of C. For additional information, go to https://www.sfasu.edu/policies/course-grades-5.5.pdf.

**Students with Disabilities**
To obtain disability related accommodations, alternate formats and/or auxiliary aids, students with disabilities must contact the Office of Disability Services (ODS), Human Services Building, and Room 325, 468-3004 / 468-1004 (TDD) as early as possible in the semester. Once verified, ODS will notify the course instructor and outline the accommodation and/or auxiliary aids to be provided. Failure to request services in a timely manner may delay your accommodations. For additional information, go to http://www.sfasu.edu/disabilityservices.

**Student Wellness and Well-Being**
SFA values students’ overall well-being, mental health and the role it plays in academic and overall student success. Students may experience stressors that can impact both their academic experience and their personal well-being. These may include academic pressure and challenges associated with relationships, emotional well-being, alcohol and other drugs, identities, finances, etc.

If you are experiencing concerns, seeking help, SFA provides a variety of resources to support students’ mental health and wellness. Many of these resources are free, and all of them are confidential.

**On-campus Resources:**
The Dean of Students Office (Rusk Building, 3rd floor lobby)
www.sfasu.edu/deanofstudents
936.468.7249
dos@sfasu.edu

www.sfasu.edu
The Health and Wellness Hub “The Hub”
Location: corner of E. College and Raguet St.

To support the health and well-being of every Lumberjack, the Health and Wellness Hub offers comprehensive services that treat the whole person – mind, body and spirit. Services include:

- Health Services
- Counseling Services
- Student Outreach and Support
- Food Pantry
- Wellness Coaching
- Alcohol and Other Drug Education

Crisis Resources:

- Burke 24-hour crisis line: 1.800.392.8343
- National Suicide Crisis Prevention: 9-8-8
- Suicide Prevention Lifeline: 1.800.273.TALK (8255)
- Crisis Text Line: Text HELLO to 741-741

Acceptable Student Behavior
Classroom behavior should not interfere with the instructor’s ability to conduct the class or the ability of other students to learn from the instructional program (see the Student Conduct Code, policy 10.4). Unacceptable or disruptive behavior will not be tolerated. Students who disrupt the learning environment may be asked to leave class and may be subject to judicial, academic or other penalties. This prohibition applies to all instructional forums, including electronic, classroom, labs, discussion groups, field trips, etc. The instructor shall have full discretion over what behavior is appropriate/inappropriate in the classroom.

Date of document: 08/23/2023