**Brandeis University**



**Division of Graduate Professional Studies Rabb School of Continuing Studies**

**Course Syllabus**

1. **COURSE INFORMATION**

**Predictive Analytics:** January 12th, 2022—March 22nd, 2022 RSAN-160: Predictive Analytics & Machine Learning, Spring1 2022 Distance Learning Course Week: Wednesday through Tuesday Office Hours: Monday & Thursdays from 7:00 PM to 8:00 PM EST

**Instructor: Stephanie Wallace**

You can contact me through discussion forums on our course web site: by replying to any of my posted messages, posting a new topic on the Questions and Comments forum, or posting to one of the various forums established for each class assignment. To reach me privately, please use the Private Forum, which is also the method I will use to contact you. The Private Forum is used instead of email in the Strategic Analytics degree program.

**Syllabus Overview**

This syllabus contains all relevant information about the course: its objectives & outcomes, grading criteria, texts, other materials of instruction, weekly objectives, outcomes, readings, assignments, and due dates. Consider this your roadmap for the course. Please read through the syllabus carefully and feel free to share any questions that you may have.

**Course Description**

This course will focus on the fundamentals of predictive analytics as it relates to improving business performance, and will cover predictive models, key modeling techniques, scoring, non-parametric regression and classification, principal components analysis and dimension reduction, time series, quality control methods, multiple predictor variables, and decision trees. The course will utilize best practices and case studies to illustrate how predictive analytics can facilitate educated decision-making to reduce costs, increase revenues, and provide competitive advantage across a variety of industries.

**Relevant Programs:** Graduate elective course for the M.S. degree in Strategic Analytics

**Prerequisites (either one of the following):**

* RSAN101: Foundations of Data Science and Analytics
* RSAN110: Business Intelligence, Analytics and Decision Making

**Welcome to Predictive Analytics!**

This course is one of the elective courses offered to complete the M.S. degree in Strategic Analytics. This course will utilize case studies, trends, techniques, and best practices as it examines the various components and applications of predictive analytics. The course procedures and policies are clearly detailed throughout this syllabus with materials posted on the LATTE web site. Please familiarize yourself with these materials and feel free to ask me any questions.

**Required Texts**

Applied Predictive Analytics: Principles and Techniques for the Professional Data Analyst, Dean Abbott, 2014. Wiley ISBN: 978-1118727966

Predictive Analytics: The Power to Predict Who Will Click, Buy, Lie, or Die, Eric Siegel, 2016. Wiley ISBN: 978-1119145677.

**Topic Notes and Assignments**

* Weekly required and Topic Notes & R Coding Notes available on the course site (in Latte)
* 5 R Coding Assignments, available on the course site (in Latte)
* Final Project

**Online Course Content**

This section of the course will be conducted completely online using the Brandeis LATTE site. The site contains the course syllabus, assignments, discussion forums, links/resources to course-related professional organizations and sites, and weekly checklists, objectives, outcomes, topic notes, self-tests, and discussion questions. Access information is emailed to enrolled students before the start of the course.

**Overall Course Objectives**

The course is intended to provide students with an understanding of:

* The definitions of key terms and concepts in predictive analytics and text analytics
* The importance of understanding and preparing data for predictive analytics
* The differences between descriptive modeling vs. predictive modeling, as well as the key concepts of model ensembles and model deployment
* The ways in which predictive analytics are utilized in different fields and industries
* The business issues that predictive analytics addresses and resolves
* The most common challenges in developing and utilizing predictive analytics
* The strategic purpose and impact predictive analytics can serve within an organization
* The future direction predictive analytics will likely take, including application and careers

**Overall Course Outcomes**

At the end of the course, students will be able to:

* Explain what predictive analytics are, and related key concepts including model ensembles, model deployment, and text analytics
* Identify the differences between descriptive modeling vs. predictive modeling
* Describe how predictive analytics are utilized in different fields and industries
* Identify and describe the business issues that predictive analytics addresses and resolves
* Identify and describe the ways in predictive analytics can have a strategic impact within an organization
* Describe the likely path predictive analytics will take in the future, as well as possible career choices

**Description of Grading Components**

|  |  |
| --- | --- |
| **Percent** | **Component** |
| 30% | Weekly Discussions/Online Participation (3% each week) |
| 30% | R Coding Assignments (5 at 6% each) |
| 30% | Final Project |
| 10% | Quiz (Week 10) |

**Weekly Discussions / Online Participation (30%, 3% per week)**

All student participation will be done online via LATTE. Each weekly block has a page that includes "Discussion Questions" which describes the topics for discussion related to the course materials posted for that week.

Each topic description includes a series of discussion questions or other directions for providing a response. To earn full credit for the Participation component of the grade, students will be expected to complete the following during weeks 1 through 10 of the course:

* Respond to at least 2 discussion topics each week. Post an original response to one topic by end of day Saturday, midnight EST, and post another by end of day Monday, midnight EST.
* Post at least 2 other substantive replies to the discussions each week (on different days) by end of day Tuesday, midnight EST. These messages are replies to the original response messages of others or replies to someone else’s reply message. The assumption is that you will read through the posts of your classmates to enhance your learning; reply to those of your choice, based upon your own experiences and insights.
* Post messages on three different days of the course week. While you may post all the required original responses and replies before the due dates, it is important for you to be involved in the discussions throughout the week.

During some weeks, responses to specific topics are due on certain dates; in other weeks, students may choose from among the available topics. Please review the discussion topic requirements carefully. These discussion requirements are described within the Discussion Questions page within each weekly block on the course home page; they are also listed in the Checklist page for each week.

Each of the two required original response messages contributes 30% of the weekly participation grade. Maximum grade is given for each of these if the posted message:

* Answers all questions asked and follows all directions specified in the topic description.
* Includes shared industry experiences and/or relates concepts to the topic notes and readings as appropriate.
* Provides at least two recent citations beyond the course textbooks or required reading. Note, all sources must be recent, high-quality, academic, peer-reviewed sources, denoted with in-text-citations and properly referenced using APA format (refer to the Research Help > Citing Sources” link in LATTE Resources block).
* Uses good scholarship, cite sources, and follows proper grammar, spelling, and format.
* Contains original content, states in the student’s own words, rather than the words of the source. Minimizes the use of (or contains no) quoted material, images, graphs, charts, or diagrams.
* Provides sufficient detail; original responses should include a minimum of 200-300 words. Some topics require lengthier responses to answer all questions.
* The original response must also be included in a text document (such as Microsoft Word document) and attached to the post. This will allow for individual feedback from the instructor. The response should not include the text of the question. Instead, the question must be addressed in the response.

Each of the two required substantive reply messages contributes 15% of the weekly participation grade. Maximum grade is given for each of these if the posted message:

* Provides substantive comments (beyond an "I agree" post) with follow-on points or questions to extend the conversation. Substantive replies should include a minimum of 100-200 words.
* Uses good scholarship, original content, cites sources and follows proper grammar, spelling, & format.
* Note, as with the original response, all sources must be recent, high-quality, academic, peer-reviewed sources, denoted with in-text-citations and properly referenced using APA format (refer to the Research Help > Citing Sources” link in the LATTE Resources block).

All discussion forum postings are subject to review for originality and adherence to academic honesty principles. See the section “Academic Honesty & Student Integrity” for more information. Posting of discussion messages needs to be done in a timely manner to allow peers the opportunity to review and reply.

***Late Policy:***

* *Half credit is deducted for an original response that is one day late.*
* *No credit is earned for original responses that are posted more than one day late.*
* *No credit is earned for substantive replies that are posted late.*

10% of weekly participation grade is based on your participation in discussions throughout the week.

* Maximum grade: is given for those that post messages to the weekly discussion forum on three (or more) days during the course week.
* Partial credit: is given for those that post their messages to the weekly discussion forum on only one or two days of the course week.

Online participation grade for each week is based on your contribution to the weekly discussion forum, for example, “Week 1 Discussions”. Posts to the forums set up for discussion of general questions and comments, exercises, or assignments are not considered in the weekly participation grade.

To summarize, the online participation grade for each week is based on the following requirements:

|  |  |
| --- | --- |
| **Weekly Requirement** | **Portion of Weekly Participation Grade** |
| Post original response by Saturday Night | 50% |
| Post substantive reply #1 by Tuesday Night | 25% |
| Post substantive reply #2 by Tuesday Night | 25% |

Each week, the online participation in these discussions contributes 5% to the overall course grade. Over ten weeks, this amounts to 50% of the overall course grade.

**Assignments (30%, 6% per week)**

There will be 5 assignments throughout the semester with each worth 6% of the course grade. Assignment submissions will be due Tuesday at midnight in the week it has been assigned. Late Policy: Half credit is deducted for an assignment submitted one day late. No credit is earned for assignments submitted more than one day late.

**Final Project (30%)**

Your task is to apply your understanding of the CRISP-DM process to either an industry or company of your choosing. Following some research and exploration, select either an industry or company and identify a potential business problem that would benefit from leveraging a predictive model. Based on your defined goals and objectives, you will then source a publicly available structured dataset and use the sourced data to work through the six stages of the CRISP-DM process with the ultimate goal of building a predictive model. Once you have identified your target variable as part of defining the business problem you will both profile, prepare your data, and then ultimately develop a set of five predictive models that will be trained to your training dataset. The best three performing models will then be assessed using your test dataset, at which point you will make a final model selection. A final project write-up has been posted, as well as a detailed grading rubric. For the most detailed requirements of the final project, please reference both the final project write-up and corresponding grading rubric.

**Quiz – Week 10 (10%)**

In the final week of the course, Week 10, in addition to your final project being due, you will also have a 20- question quiz that is a combination of multiple choice and true/false questions. The quiz will be untimed, open book, and you will be allowed unlimited attempts. The scope of the quiz will be the major themes of the course, with questions most directly coming from the readings and topic notes throughout each week of the semester.

1. **WEEKLEY INFORMATION**

On the course site, the home page contains 10 weekly blocks, one for each week of the course. Within each weekly block on the home page, you will find information and resources about the activities for each week:

* Overview: Checklist, Objectives and Outcomes
* Discussions
* Topic Notes & R Coding Notes & Other Required Readings/ Videos
* Assignments / Assessments

Initially some of these items (related to discussions, assignments, or assessments) will be hidden on the course home page. As we come to the appropriate point in the course, they will become visible and available. A schedule for availability is included within this syllabus. Most of the items listed in the checklists are required for this course, but some are highlighted as "[optional]" for this course. As your schedule permits, you are encouraged to complete the optional work, as it will benefit your learning.

The following pages of this syllabus present a summary of the weekly objectives, outcomes, readings, assignments, and assessments. Chapter readings for both books are planned to generally follow the sequence of the weekly topic notes. Some of the references to PMBOK Guide readings include mention of the weekly topic that is highlighted within the chapter.

**Week #1 – Introduction to Predictive Analytics**

Objectives:

* The definitions of key terms and concepts in predictive analytics & text analytics.
* The business issues that predictive analytics addresses and resolves.

Outcomes:

* Explain what data science is, including its purpose, history, and applications in business.
* Explain what analytics are, including the various types and their application in business.
* Describe the business issues that data science and analytics can address and resolve.

Readings:

* Applied Predictive Analytics – Chapter #1
* Predictive Analytics – Chapter #1
* Week 1 Topic Notes & Readings

Videos: “Introduction to Predictive Analytics”

Deliverables:

* Complete Academic Integrity Report
* Watch the “Welcome” session video
* Introduce yourself in the “Introduce Yourself” forum
* Week 1 Discussion Topic (1 original post and 2 substantive replies to your peer’s original posts)
* Install R Studio environment.

**Week #2 – Setting Up the Problem**

Objectives:

* The definitions of key terms and concepts in predictive analytics & text analytics.
* The ways in which predictive analytics are utilized in different fields and industries.
* The business issues that predictive analytics addresses and resolves.

Outcomes:

* Explain what predictive analytics are, and related key concepts including: model ensembles, model deployment, and text analytics.
* Describe how predictive analytics are utilized in different fields and industries.
* Identify and describe the business issues that predictive analytics addresses and resolves.

Readings:

* Applied Predictive Analytics – Chapter #2
* Predictive Analytics – Chapter #2
* Week 2 Topic Notes , R Coding Notes, & Readings

Videos: “When Big Data & Predictive Analytics Collide”

Deliverables:

* R Coding Assignment #1
* Week 2 Discussion Topic (1 original post and 2 substantive replies to your peer’s original posts)

**Week #3 – Understanding the Data**

Objectives:

* The definitions of key terms and concepts in predictive analytics & text analytics.
* The importance of understanding and preparing data in predictive analytics.
* The business issues that predictive analytics addresses and resolves.
* The most common challenges in developing and utilizing predictive analytics.

Outcomes:

* Explain what predictive analytics are, and related key concepts including: model ensembles, model deployment, and text analytics.
* Identify and describe the business issues that predictive analytics addresses and resolves.

Readings:

* Applied Predictive Analytics – Chapter #3
* Predictive Analytics – Chapter #3
* Week 3 Topic Notes, R Coding Notes, & Readings

Deliverables:

* R Coding Assignment #2
* Week 3 Discussion Topic (1 original post and 2 substantive replies to your peer’s original posts)
* Submit final project proposal & publicly available dataset to be used.

**Week #4 – Preparing the Data**

Objectives:

* The definitions of key terms and concepts in predictive analytics & text analytics.
* The importance of understanding and preparing data in predictive analytics.
* The business issues that predictive analytics addresses and resolves.
* The most common challenges in developing and utilizing predictive analytics.

Outcomes:

* Explain what predictive analytics are, and related key concepts including: model ensembles, model deployment, and text analytics.
* Identify and describe the business issues that predictive analytics addresses and resolves.

Readings:

* Applied Predictive Analytics – Chapters #4 & #5
* Predictive Analytics – None
* Week 3 Topic Notes, R Coding Notes, & Readings

Deliverables:

* R Coding Assignment #3
* Week 4 Discussion Topic (1 original post and 2 substantive replies to your peer’s original posts)

**Week #5 – Descriptive Modeling**

Objectives:

* The definitions of key terms and concepts in predictive analytics & text analytics.
* The differences between descriptive modeling vs. predictive modeling, as well as key concepts of model ensembles and model deployment.
* The ways in which predictive analytics are utilized in different fields and industries.
* The business issues that predictive analytics addresses and resolves.
* The most common challenges in developing and utilizing predictive analytics.

Outcomes:

* Explain what predictive analytics are, and related key concepts including: model ensembles, model deployment, and text analytics.
* Identify the differences between descriptive modeling vs. predictive modeling.
* Describe how predictive analytics are utilized in different fields and industries.
* Identify and describe the business issues that predictive analytics addresses and resolves.

Readings:

* Applied Predictive Analytics – Chapters #6 & #7
* Predictive Analytics – Chapter #5
* Week 5 Topic Notes, R Coding Notes, & Readings

Deliverables:

* R Coding Assignment #4

**This document was truncated here because it was created in the Evaluation Mode.**