Level: Graduate

Focus on Programming: ★★★★☆

Focus on Data Analytics/Statistics: ★★★★☆

Focus on Business Application: ★★★☆☆

|  |  |
| --- | --- |
| Course name | **ACCY 577: Machine Learning for Accounting** |
| Objectives | * Understand various types of machine learning algorithms, classification, regression, clustering, text analysis, time series analysis; * Be able to apply proper machine learning algorithms with python on a specific business problem; * Be able to optimize algorithms to achieve best result; * Be able to complete the whole data analytics process with Python. |
| Key knowledge and skill areas | * How to apply a machine learning algorithm on a dataset? * How to choose a model that fit a business problem best? * How to optimize a machine learning model to achieve best result? |
| Class activities | Videos, notebook lessons, quizzes, assignments, group project |
| Software | Jupyter Notebook(Anaconda) |
| Textbook | Jupyter Notebooks |

**Course Description**

ACCY577, as for now, is named as "Accounting Machine Learning with Python".

This course introduces machine learning algorithms and their applications in accounting problems with Python. It teaches the fundamentals of different machine learning algorithms, including classification, regression, clustering, text analysis, time series analysis. It also discusses model evaluation and model optimization for machine learning models.

This course provides an entry point for students to be able to apply proper machine learning algorithms on business related datasets to solve various problems. By the end of the course, combined with the skills learned from Accounting Data Analytics with Python course, students should be able to complete an entire data analytics process with Python.

**Course Content**

* Module 1: Introduction to Machine Learning
  + Lesson 1: Introduction to Machine Learning
  + Lesson 2: Introduction to Data Pre-Processing
  + Lesson 3: Introduction to Machine Learning Algorithms
* Module 2: Fundamental Algorithms I
  + Lesson 1: Linear Regression
  + Lesson 2: Logistic Regression
  + Lesson 3: Decision Tree
* Module 3: Fundamental Algorithms II
  + Lesson 1: K-Nearest Neighbors
  + Lesson 2: Support Vector Machine
  + Lesson 3: Random Forest
* Module 4: Model Evaluation
  + Lesson 1: Regression Evaluation
  + Lesson 2: Classification Evaluation I
  + Lesson 3: Classification Evaluation II
* Module 5: Model Optimization
  + Lesson 1: Feature Selection
  + Lesson 2: Cross Validation
  + Lesson 3: Model Selection
* Module 6: Introduction to Text Analysis
  + Lesson 1: Introduction to Text Analytics
  + Lesson 2: Introduction to Text Classification I
  + Lesson 3: Introduction to Text Classification I
* Module7: Introduction to Clustering
  + Lesson 1: Introduction to K-means Clustering
  + Lesson 2: K-means Case Study
  + Lesson 3: Introduction to Density-based Clustering
* Module 8: Introduction to Time Series Analysis
  + Lesson 1: Working with Date and Time
  + Lesson 2: Analyzing Time Series Data

This list is tentative. But the format of the course is fixed, we will have:

1. 8 Modules
2. Each module has 3 lessons
3. Each module has 4 major videos, an overview video and one for each lesson
4. Each module has one fundamental assignment, one high engagement assignment, one quiz
5. For high engagement students, there's one Live-Session-Prep quiz which is due one day within live session.
6. There's one final group project
7. There's no final exam