

AIDA181, Big Data Analytics for Risk and Insurance, 2nd Edition

Instructor: Jamie Warner

Contact: JMWarnerData@gmail.com, 516-680-9136

Course page: <https://github.com/jmwarner/AIDA181>

Expectations

- **Class Structure:** Each class will meet for 2 hours with one 5-minute break. If you need to step away from your computer to use the restroom, please do so, but be aware the class will continue. CE credits will only be awarded to students who participate in the live webinars and who fulfill all the “check-in” requirements.
- **Dates:** Class will meet virtually, Mondays from 7-9pm on September 10 to October 15, 2018. October 22 will be an optional review session where we will complete additional practice problems and answer any outstanding questions.
- **Polling:** Questions will be asked during each topic to determine both understanding of the material and that each student is in attendance.
- **Homework:** Additional questions will be available as optional (but highly recommended!) homework. Homework quizzes will be available by Friday of each week.
- **Practice Exam:** The online AIDA 181 course material includes an online practice exam. I highly recommend taking the exam in similar circumstances to the ones you will have at the testing center. Please be aware of the following:
 - No food or drink is allowed in the testing center
 - If you need the restroom your test will keep running
 - You will be provided scratch paper, pencils, and a calculator. I prefer to always bring my own calculator so I know it works and I feel comfortable with it.
 - Other students will be coming in and out of the test center so be prepared to not have complete silence. Usually the centers have noise cancelling headphones you can wear.
- **Exam:** If you have not already, please schedule your certification exam and let me know the date.
- **Recommended Reading:** How to Prepare for Institute Exams (available on course webpage)

Course Outline

Week 1 – September 10

Assignment 1 — Exploring Big Data Analytics

1. Big Data and Technology—The Future of Insurance
 - a. Strategic Opportunities From Big Data and Technology
 - b. Big Data

- c. Technology
 - d. Data Science
 - e. Strategic Applications to Insurance
- 2. Evolution of Big Data
 - a. Big Data 1.0
 - b. Big Data 2.0
- 3. Big Data Characteristics and Sources
 - a. Data Characteristics
 - b. Internal and External Data
 - i. Internal Data
 - ii. External Data
 - c. Structured and Unstructured Data
- 4. Data Quality
 - a. Data Quality Defined
 - b. Metadata
 - c. Methods to Identify Inaccurate or Missing Data
 - i. Quantitative Data
 - ii. Categorical Data
 - d. Data Security
- 5. Data Mining
 - a. Data Mining Techniques
 - b. Developing a Predictive Model
 - c. The Data Mining Process

Break

- 6. Data Science
 - a. Data Science Concepts
 - b. The Data Scientist
 - c. The Data Science Team
- 7. Data-Driven Decision Making
 - a. Data Science and Data-Driven Decision Making

- i. The Descriptive Approach
 - ii. The Predictive Approach
- b. A Model for Data-Driven Decision Making in Risk Management and Insurance

Assignment 2 — Predictive Modeling Concepts

- 1. Basic Modeling Concepts
 - a. Supervised Versus Unsupervised Learning
 - b. Predictive Versus Descriptive Modeling
 - c. Algorithms
 - d. Entropy
 - e. Lift

Week 2 – September 17

- 2. Similarity and Distance
 - a. Measuring Similarity and Distance
 - b. Applications of Similarity and Distance
 - i. Nearest Neighbors
 - ii. Measuring Similarity in Networks
- 3. Training and Evaluating a Predictive Model
 - a. Training the Model
 - i. Training Data
 - ii. Holdout Data
 - iii. Cross-Validation
- 4. Evaluating the Model
 - a. Performance Metrics
 - b. Putting the Model into Production

Break

Assignment 3 — Traditional Analysis Techniques for Big Data

- 1. Traditional Data Analysis Techniques
 - a. Exploratory Data Analysis
 - b. Classification Trees
 - c. Regression Models

- d. Cluster Analysis
- 2. Classification Trees
 - a. Classification Tree Model
 - b. Classification Rules
 - c. Probability Estimation Tree
- 3. Linear Functions
 - a. Linear Discriminants
 - b. Linear Regression
 - i. Multiple Regression

Week 3 – September 24

- ii. Disadvantages of Linear Models
 - c. Generalized Linear Models
 - d. Logistic Regression
- 4. Cluster Analysis
 - a. Cluster Analysis
 - b. Hierarchical Clustering
 - c. K Nearest Neighbors Algorithm
 - d. K-Means Clustering Algorithm
 - e. Describing Clusters

Break

Assignment 4 — Text Mining, Social Network Analysis, and Neural Networks

- 1. Text Mining
 - a. Retrieve and Prepare Text
 - b. Create Structured Data from Unstructured Data
 - c. Create a Model Using Data Mining Techniques
 - d. Evaluate the Text Mining Model
- 2. Social Network Analysis
 - a. Introduction to Social Network Analysis
 - b. Link Analysis
 - c. Social Network Metrics

- d. Network Classification

Week 4 – October 1

- 3. Neural Networks
 - a. Mechanics of Neural Networks
 - b. Applications of Neural Networks

Assignment 5 — Technology and Smart Products

- 1. Emerging Technology for Underwriting, Claims, and Risk Management
 - a. Artificial Intelligence
 - b. Wireless Sensor Networks
 - c. Computer Vision
- 2. How Smart Products Apply to Risk Management
 - a. Property Management
 - b. Supply Chain Management
 - c. Transportation Management
 - d. Catastrophe Management
 - e. Workplace Safety Management
 - f. Construction and Engineering Management

Break

Assignment 6 — Underwriting Applications of Big Data Analytics

- 1. Automobile Ratemaking Using Vehicle Telematics
 - a. Traditional Automobile Rating Attributes
 - b. Vehicle Telematics
 - c. Telematics' Effect on Ratemaking
- 2. Segmenting Homeowners Policies Using Machine Learning
 - a. Need for Rate Adjustment
 - b. Identifying Loss Ratio Segments With Machine Learning
 - c. Evaluating the Effect of Machine Learning on Homeowners Policies

Week 5 – October 8

- 3. Underwriting Products Liability Risks Using Data Mining
 - a. Discovering Emerging Risks

- i. Cluster Analysis
- ii. Text Mining
- b. Predictive Modeling for Emerging Risks

Assignment 7 — Claims Applications of Big Data Analytics

1. Detecting Claims Fraud with Network Analysis and Clustering
 - a. Detecting Claims Fraud
 - b. Applying Network Analysis
 - c. Applying Cluster Analysis
2. Using Classification Tree Analysis in Claims Assignment
 - a. Overview of the Procedure
 - b. Information Gain by Complex Claims Attributes

Break

- c. Using a Classification Tree to Illustrate Attributes of Complex Claims
 - d. Validating the Complex Claim Model
 - e. Complex Claims Reporting
3. Improving Claims Processes with Business Process Analytics
 - a. Overview of the Procedure
 - b. Business Process Analytics and Claims
 - c. Process Mining Applied to Claims
 - d. Improving Claims Processes with Data Analytics

Week 6 – October 15

Assignment 8 — Risk Management Applications of Big Data Analytics

1. Preventing Employee Injuries With Sensor Data
 - a. Traditional Workplace Accident Analysis Techniques
 - b. Sensor-Generated-Data Analysis Techniques
2. Assessing Reputation Risk Through Text Mining and Social Network Analysis
 - a. Reputation Risk
 - b. Text Mining
 - c. Social Network Analysis
3. Using Clustering and Linear Modeling for Loss Development

- a. Overview of the Procedure
- b. Collect and Organize Past Data
- c. Apply Data Analysis Techniques
- d. Develop a More Accurate Method to Predict Losses

Break

Assignment 9 — Implementing a Data Analytics Strategy

1. Data Analytics Strategy
 - a. Reinforcing Strengths
 - i. Mitigating Weaknesses and Threats
 - ii. Exploiting Opportunities
2. Data Analytics Risk Management
 - a. Scan Environment
 - i. Internal Environment
 - ii. External Environment
 - b. Identify Risks
 - c. Analyze Risks
 - d. Treat Risks
 - e. Monitor and Assure
3. Data Analytics Change Management
 - a. Articulate the Need for Change
 - b. Appoint a Leadership Team
 - c. Develop a Written Statement of the Vision and Strategies
 - d. Communicate the Vision and Strategies
 - e. Eliminate Barriers to Change
 - f. Recognize Incremental Successes
 - g. Entrench the Change

Week 7 – October 22

Optional review session.