# DSC 400 - Big Data Technology and Algorithms

**Course Description**This course provides an in-depth exploration of Big Data technologies and algorithms, focusing on their applications in data science. Using Python, students will learn about data collection, manipulation, visualization, and analysis. It covers both the theoretical foundations and practical implementations, including machine learning and artificial intelligence techniques for Big Data.

**Course Prerequisites**- Proficiency in Python programming.  
- Understanding of fundamental concepts in data structures and algorithms.

**Course Objectives**Upon successful completion of this course, students should be able to:  
1. Comprehend key concepts and the importance of Big Data in various domains.  
2. Use Python and its libraries for data collection, wrangling, and visualization.  
3. Implement and manage Big Data in both SQL and NoSQL databases.  
4. Understand and apply distributed computing concepts for Big Data processing.  
5. Develop machine learning and deep learning models for Big Data.  
6. Perform advanced data analysis with NLP and time series data.  
7. Apply AI algorithms to extract insights from Big Data.

**Grading Scale**- 93 – 100% = A  
- 90 – 92% = A-  
- 87 – 89% = B+  
- 83 – 86% = B  
- 80 – 82% = B-  
- 77 – 79% = C+  
- 73 – 76% = C  
- 70 – 72% = C-  
- 67 – 69% = D+  
- 63 – 66% = D  
- 60 – 62% = D-  
- 0 – 59% = F

**Course Outline**

**Week 1 - Introduction to Big Data and Python Fundamentals**

* Overview of Big Data and its significance.
* Introduction to Python for data analysis.

**Week 2 - Data Collection and Wrangling**

* Techniques for collecting and cleaning data.
* Preparing data for analysis.

**Week 3 - Data Visualization Techniques**

* Importance and methods of data visualization.
* Creating interactive dashboards and visualizations.

**Week 4 - Database Management for Big Data**

* Using SQL and NoSQL databases for Big Data.
* Storing and querying large datasets.

**Week 5 - Distributed Computing Foundations**

* Fundamentals of distributed computing.
* Utilizing frameworks like Spark**.**

**Week 6 - Machine Learning Fundamentals with Big Data**

* Machine learning techniques for predictive modeling.
* Building and evaluating models on large datasets.

**Week 7 - Deep Learning Applications in Big Data**

* Applying deep learning for pattern recognition.
* Designing deep neural networks for prediction tasks.

**Week 8 - Advanced Natural Language Processing (NLP)**

* Techniques for text data analysis at scale.
* Sentiment analysis and topic modeling.

**Week 9 - Time Series Analysis with Big Data**

* Analyzing time-series data for trends and forecasting.
* Identifying patterns and seasonal variations.

**Week 10 - AI for Big Data Analytics**

* AI techniques for efficient large dataset analysis.
* Large learning models for chat bots.

**Week 11 - Capstone Project in Big Data Analysis**

* Conceptualization and design of a comprehensive project.
* Implementation and presentation of project findings.