Computer Science – Data Science (Elective)

Course # COMP 4002

Credits 6

Pre-requisites and Co-requisites: None

Course Description

This course offers a comprehensive exploration of key concepts and practical skills essential for leveraging Python in the field of data science. The course covers a wide array of topics, providing hands-on experience in dealing with diverse data sources and employing machine learning techniques. The course is focused on building proficiency in utilizing Python libraries such as Pandas, NumPy, and Scikit-Learn, while also covering essential statistical knowledge for effective data analysis. The course delves into ethical considerations, ensuring a holistic understanding of responsible data science practices. Additionally, it explores emerging trends in the future of data science, including artificial neural networks and deep learning models.

Course Learning Outcomes

Upon completion of this course, students will be able to:

- Utilize Python for data science tasks, including data manipulation, analysis, and visualization.
- Employ file handling techniques and SQL queries in Python to manage and process diverse data sources.
- Efficiently use Pandas and NumPy libraries for loading, cleaning, and manipulating datasets.
- Perform EDA to uncover patterns, trends, and insights within datasets, and visualize findings effectively.
- Implement machine learning classification models, regression techniques, and evaluate their performance.
- Optimize machine learning models, utilizing AutoML, implementing tree-based models, Support Vector Machines (SVM), and exploring deep learning models such as CNN, RNN, LSTM, and Transformers.

Course Assessments and Grading

Item	**Weight**
Class participation	10%
Quiz activities	15%
Assignments	15%

Item	Weight
Mid exam in two Parts 1. Objective Part: Online 2. Subjective Part: Paper based	30%
Final exam in two Parts 1. Objective Part: Online 2. Subjective Part: Paper based	30%