|  |  |
| --- | --- |
| AIUB-Logo | **AMERICAN INTERNATIONAL UNIVERSITY-BANGLADESH (AIUB)**  Faculty of Science and Technology (FST)  Department of Computer Science (CS)  Undergraduate Program |



I. Course Core and Title

**CSC 4180: Introduction to Data Science**

II. Credit

**3 credit hours (2 hours of theory and 3 hours Lab per week)**

III. Nature

**Elective Course for CSE**

IV. Prerequisite

**CSC4121 Artificial Intelligence & Expert System**

**CSC2107 Introduction to Database**

**V. Vision:**

Our vision is to be the preeminent Department of Computer Science through creating recognized professionals who will provide innovative solutions by leveraging contemporary research methods and development techniques of computing that is in line with the national and global context.

**VI. Mission:**

The mission of the Department of Computer Science of AIUB is to educate students in a student-centric dynamic learning environment; to provide advanced facilities for conducting innovative research and development to meet the challenges of the modern era of computing, and to motivate them towards a life-long learning process.

|  |  |
| --- | --- |
| **COURSE PLAN** | **SEMESTER: Fall 2024-2025** |



## **VII - Course Description:**

* Comprehend introduction to the modern study of data science.
* Introduce R programming language for data science.
* Discuss various process for data collection and data preprocessing.
* Discuss the basic data management for data science.
* Describe basic statistics for data science.
* Discuss data visualization and basic graphs.
* Create graphs/plots using R for data visualization.
* Discuss web scrapping using R for data collection from websites.
* Implement interactive dashboard using R for data analysis.

****

## **VIII – Course outcomes (CO) Matrix:**

By the end of this course, students should be able to:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **COs** \* | **Description** | **Domain  Level** \*\*\* | | | **PO  Assessed** \*\*\*\* |
| C | P | A |
| **CO1** \*\* | ***Analyze* the data to draw a valid conclusion following the data science principles and considering the limitations.** | 4 |  |  | PO-d-2 |
| CO2 | *Break down* the processing of collected data according to the data science principles. | 4 |  |  | PO-d-2 |
| CO3 | *Evaluate* the solution of data science problem by considering all the requirements. | 5 |  |  | PO-d-3 |
| CO4 | *Justify* the solution of data science problem to provide a valid conclusion. | 5 |  |  | PO-d-3 |
| *C: Cognitive; P: Psychomotor; A: Affective Domain*  *\* CO assessment method and rubric of COs assessment is provided in later section*  *\*\* COs will be mapped with the Program Outcome (POs) \*\*\* The numbers under the ‘Level of Domain’ columns represent the level of Bloom’s Taxonomy each   CO corresponds to.*  *\*\*\*\* The numbers under ‘PO Assessed’ column represent the POs each CO corresponds to.* | | | | | |

## **IX – Topics to be covered in the class:\***

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Time Frame** | **CO**  **Mapped** | **Topics** | **Teaching**  **Activities** | **Assessment Strategy(s)** |
| Week 1 | CO1, CO2 | Introduction to Data Science (what is data science, steps in data science, datafication, and data science profile).  Introduction to R language (why use R, installing R and RStudio, and installing and loading R packages). | Lecture, Question-answer, Lab Practice | Quiz/Exam, Project |
| Week 2 | CO1, CO2 | Introduction to data and dataset (what is data, structured vs unstructured, what is dataset and variable, variable types, data collection sources).  R language basics (variable, data types, and input and output). | Lecture, Question-answer, Lab Practice | Quiz/Exam, Project |
| Week 3 | CO1, CO2 | Data structures (vector, matrix, array, data frame, list, and factor).  Conditional statements in R.  Loop statements in R. | Lecture, Question-answer, Lab Practice | Quiz/Exam, Project |
| Week 4 | CO3, CO4 | Data preprocessing (data cleaning, data integration, data transformation, data reduction, and data discretization). | Lecture, Question-answer, Lab Practice | Quiz/Exam, Project |
| Week 5 | CO2, CO3 | Basic data management (missing data, data type conversion, sorting data, merging datasets, subsetting dataset, and manipulating data). | Lecture, Question-answer, Lab Practice | Quiz/Exam, Project |
| Week 6 | CO1,CO4 | Sorting data, Rearrange Data, Ranking Data. Discussion on midterm project (Apply the knowledge of data science given through the lectures so far and complete the midterm project.  Use R language for the midterm project implementation.). | Lecture, Question-answer, Lab Practice | Quiz/Exam, Project |
| Midterm (Week 7) | | | | |
| Week 8 | CO1, CO4 | Data visualization (what is data visualization, getting started with graphs, creating a graph using R). | Lecture, Question-answer, Lab Practice | Quiz/Exam, Project |
| Week 9 | CO1, CO4 | Basic graphs/plotting (such as scatter plot, bar plot, histogram, etc.) | Lecture, Question-answer, Lab Practice | Quiz/Exam, Project |
| Week 10 | CO2, CO3 | Advanced data management (reshaping data, and aggregating data).  Functions in R. | Lecture, Question-answer, Lab Practice | Quiz/Exam, Project |
| Week 11 | CO1, CO4 | Introduction to statistics (gathering data, describing data, population and sample, parameter and statistics, study types, and sampling types). | Lecture, Question-answer, Lab Practice | Quiz/Exam, Project |
| Week 12 | CO1, CO4 | Descriptive statistics (what is descriptive statistics, mean, mode, median, range, variance, standard deviation, quartiles and percentiles, data distribution, and normal distribution). | Lecture, Question-answer, Lab Practice | Quiz/Exam, Project |
| Week 13 | CO2,CO3 | Linear and Nonlinear Dimensionality Reduction | Lecture, Question-answer, Lab Practice | Quiz/Exam, Project |
| Week 14 | CO1,CO2 | Supervised Classification. Discussion on final term project (Apply the knowledge of data science given through the lectures so far and complete the final term project.  Use R language for the final term project implementation.). | Lecture, Question-answer, Lab Practice | Quiz/Exam, Project |
| **Final term (Week 15)** | | | | |
| Week 16 | CO2,CO3 | Project Evaluation |  |  |
|  | | | | |

*\* The faculty reserves the right to change, amend, add, or delete any of the contents.*



## **X – Mapping of PO/PLO and K, P, A of this course:**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| PO Indicator ID | PO Indicators Definition (As per the requirement of WKs) | Domain | K | P | A |
| PO-d-2 | Analysis and Interpretation of collected data to provide valid conclusion acknowledging the limitations. | Cognitive Level 4 (Analyzing) | K8 |  |  |
| PO-d-3 | Investigate solution of complex engineering problem by synthesis of information to provide valid conclusions. | Cognitive Level 5 (Evaluating) | K8 | P1 P4 P5 |  |

## **XI – K, P, A Definitions**

|  |  |  |
| --- | --- | --- |
| **Indicator** | **Title** | **Description** |
| **K8** | Research Literature | Engagement with selected knowledge in the research literature of the discipline |
| **P1** | Depth of knowledge required | Cannot be resolved without in-depth engineering knowledge at the level of one or more of K3, K4, K5, K6 or K8 which allows a fundamentals-based, first principles analytical approach |
| **P4** | Familiarity of issues | Involve infrequently encountered issues |
| **P5** | Extent of applicable codes | Are outside problems encompassed by standards and codes of practice for professional engineering |



## **XII – Mapping of CO Assessment Method and Rubric**

The mapping between Course Outcome(s) (COs) and The Selected Assessment method(s) and the mapping between Assessment method(s) and Evaluation Rubric(s) is shown below:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **COs** | **Description** | **Mapped POs** | **Assessment Method** | **Assessment Rubric** |
| CO1 | Analyze the data to draw a valid conclusion following the data science principles and considering the limitations. | PO-d-2 | Project | Rubric for Project |
| CO2 | Break down the processing of collected data according to the data science principles. | PO-d-2 | Quiz/Exam | Rubric for Quiz/Exam |
| CO3 | Evaluate the solution of data science problem by considering all the requirements. | PO-d-3 | Project | Rubric for Project |
| CO4 | Justify the solution of data science problem to provide a valid conclusion. | PO-d-3 | Project | Rubric for Project |

 **XIII – Evaluation and Assessment Criteria**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **CO1:** *Analyze* the data to draw a valid conclusion following the data science principles and considering the limitations. | | | | | |
| **Assessment Criteria** | **Not Attended/ Incorrect (0)** | **Inadequate  (1-2)** | **Average (3)** | **Good  (4)** | **Excellent (5)** |
| **Evaluation Criteria** | **Evaluation Definition** | | | | |
| Problem Analysis | Clearly demonstrate the project idea and why this problem is important to consider. Design and justify the step-by-step solution of the problem to achieve a valid conclusion from the data. | | | | |
| Interpretation | Provided an analysis of the data and clearly state its outcome. Justify the outcome of the project acknowledging the limitations. | | | | |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **CO2:** *Break down* the processing of collected data according to the data science principles. | | | | | |
| **Assessment Criteria** | **Not Attended/ Incorrect (0)** | **Inadequate  (1-2)** | **Average (3)** | **Good  (4)** | **Excellent (5)** |
| **Evaluation Criteria** | **Evaluation Definition** | | | | |
| Content knowledge | Demonstrates the knowledge of the data science practice and principles. | | | | |
| Argumentation | Articulates a position or argument for the choosing the correct practice and principles of data science. | | | | |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **CO3:** *Evaluate* the solution of data science problem by considering all the requirements. | | | | | |
| **Assessment Criteria** | **Not Attended/ Incorrect (0)** | **Inadequate  (1-2)** | **Average (3)** | **Good  (4)** | **Excellent (5)** |
| **Evaluation Criteria** | **Evaluation Definition** | | | | |
| Requirements Analysis | Clearly articulates all the requirements to address the project. Presents sufficient explanation how do you accommodate all these requirements in solution design. | | | | |
| Solution Design | Design a step-by-step solution of the project. Clearly demonstrate each step of the solution design considering all the requirements of the project. | | | | |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **CO4:** *Justify* the solution of data science problem to provide a valid conclusion. | | | | | |
| **Assessment Criteria** | **Not Attended/ Incorrect (0)** | **Inadequate  (1-2)** | **Average (3)** | **Good  (4)** | **Excellent (5)** |
| **Evaluation Criteria** | **Evaluation Definition** | | | | |
| Justification | Justify how your proposed solution is sufficient to achieve a valid conclusion from the data. | | | | |
| Presentation | Clearly present your conclusion you have drawn from the data by following your proposed solution. Explain the actual outcome with the expected outcome. Draw your concluding remarks about the limitations and possible improvements. | | | | |

****

## **XIV- Course Requirements**

* Students are expected to attend at least 80% class.
* Students are expected to participate actively in the class.
* For both terms, there will be at least 2 quizzes based on the theoretical knowledge and conceptual understanding of the topic covered discussed in the classes.
* Submit report based on the given course related problems.
* Submission of assignment and/or projects should be in due time.

****

## **XV – Evaluation & Grading System\***

The following grading system will be strictly followed in this class

**Mid Term Exam:**Term Project: 50%  
Quizzes: 40%  
Attendance & Performance: 10%

**Final Term Exam:**Term Project: 60%  
Lab Performance: 30%  
Attendance & Performance: 10%

**Semester grade:** 40% midterm + 60% final term

|  |  |  |
| --- | --- | --- |
| **Letter** | **Grade Point** | **Numerical %** |
| A+ | 4.00 | 90-100 |
| A | 3.75 | 85 - < 90 |
| B+ | 3.50 | 80 - < 85 |
| B | 3.25 | 75 - < 80 |
| C+ | 3.00 | 70 - < 75 |
| C | 2.75 | 65 - < 70 |
| D+ | 2.50 | 60 - < 65 |
| D | 2.25 | 50 - < 60 |
| F | 0.00 | < 50 |
| I |  | Incomplete |
| W |  | Withdrawal |
| UW |  | Unofficially Withdrawal |

*\* The evaluation system will be strictly followed as par the AIUB grading policy.*

****

## **XVI – Textbook/ References**

* An Introduction to Data Science by Jeffrey S. Saltz and Jeffrey M. Stanton.
* R in Action by Rob Kabacoff.
* Introduction To Data Science: A python Approach to Concepts, Techniques and Applications by Laura Igual and Santi Segui.
* Lecture materials will be provided online at the course website weekly.

****

## **XVII - List of Faculties Teaching the Course**

|  |  |
| --- | --- |
| **FACULTY NAME** | **SIGNATURE** |
| MR. TOHEDUL ISLAM |  |



## **XVIII – Verification:**

|  |  |  |
| --- | --- | --- |
| **Prepared by :**  ---------------------------------  **TOHEDUL ISLAM**  *Course Convener*  Date:......................................... | **Moderated by :**  ---------------------------------  **Dr. Akinul Islam Jony**  *Point Of Contact*  *OBE Implementation Committee for CS*  Date:......................................... | |
| **Checked by:**  ....................................................  **Dr. Akinul Islam Jony**  *Head* *(Undergraduate Program)*  *Department of Computer Science*  Date:.......................................... | **Certified by:**  .........................................................  **Dr Md Abdullah Al Jubair**  *Director*,  *Faculty of Science & Information Technology*  Date:............................................... | **Approved by:**  .........................................................  **Dr. Dip Nandi**  *Associate Dean*,  *Faculty of Science & Information Technology*  Date:............................................... |