|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| BUBT THEORY COURSE OUTLINE **Bangladesh University of Business and Technology (BUBT)**  **Faculty of Engineering& Applied Sciences (FEAS)**  **Department of Computer Science and Engineering (CSE)** | | | | | | | |
|  |  | | | | | | |
| 1 | Program | | | B.Sc. Engg. in CSE | | | |
| **2** | **Course Code** | | | CSE 475 | | | |
| **3** | **Course Title** | | | Data Mining | | | |
| **4** | **Course Type** | | | Core Course | | | |
| **5** | **Academic Session** | | | Fall 2021-22 | | | |
| **6** | **Credit Hour** | | | 3.0 | | | |
| **7** | **Intake** | | | 38 | | | |
| **8** | **Section** | | | 5 | | | |
| **9** | **Pre-requisites** | | | CSE 207-Database Systems | | | |
| **10** | **Campus** | | | Permanent Campus | | | |
| **11** | **Course Teacher** | | | |  |  |  | | --- | --- | --- | | **Name:** Badhan Chandra Das | **Designation**: Lecturer | | | **Specialization**: Social Network Analysis, Machine Learning, Data Mining, Natural Language Processing. | | | | **Room No.** | **Email:** badhan\_das@bubt.edu.bd | **Cell No**. 01670140193 | | | | |
| **12** | **Class Schedule** | | | |  |  |  | | --- | --- | --- | | **Class Day** | **Class Hours** | **Class Room** | | Tuesday | 2:30 PM – 03.00 PM | 320 (B-2) | | Thursday | 10:30 AM – 11:30 PM | 910 (B-2) | | | | |
| **13** | **Counselling Schedule** | | | |  |  |  | | --- | --- | --- | | **Class Day** | **Class Hours** | **Class Room** | | Sunday | 10:00 AM – 11.00 AM | 911(B-2) | | Monday | 10:00 AM – 11.00 AM | 911(B-2) | | | | |
| **14** | **Course Objectives** | | | This course introduces fundamental data mining techniques and data warehousing principles. The course is designed for students to evaluate different models used for OLAP and data pre-processing. This course enables the students to select appropriate data mining algorithms for respective applications. The students will learn to analyze and construct knowledge from data. | | | |
| **15** | **Course Synopsis** | | | Introduction: Basic Concepts and Motivation, Technologies, Applications; Data Preprocessing: Data Objects and Attribute Types, Measuring Data Similarity and Dissimilarity, Data Cleaning, Data Integration, Data Reduction, Data Transformation; Data Warehousing: Data Cube, OLAP, Data Warehouse Design; Data Cube Technology; Mining Frequent Patterns, Associations and Correlations; Classification: Decision Tree Induction, Rule-Based Classification, Model Evaluation and Selection; Cluster Analysis: Basic Concepts, Clustering Methods: Partitioning and Hierarchical Methods; Outlier Analysis. | | | |
| **16** | **Text Book** | | | 1. Data Mining Concepts and Techniques -Jiawei Han, Micheline Kamber, Jian Pei | | | |
| **17** | **Reference Book** | | | 1. Data Mining Techniques - Michael Berry | | | |
| **18** | **Course Outcomes (COs)** | | | Upon completing this course students will be able to:  **CO1**: **Understand** different data mining concepts, techniques and its applications.  **CO2: Explain** different data mining algorithms and models.  **CO3**: **Apply** different data preprocessing and data mining techniques on real world data.  **CO4:**  **Analyze** different data mining solutions involving different data mining techniques. | | | |
| **Mapping of COs to POs** | | | |  |  |  |  |  |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | | CO | PO1 | PO2 | PO3 | PO4 | PO5 | PO6 | PO7 | PO8 | PO9 | PO10 | PO11 | PO12 | | CO1 | √ |  |  |  |  |  |  |  |  |  |  |  | | CO2 | √ |  |  |  |  |  |  |  |  |  |  |  | | CO3 |  |  | √ |  |  |  |  |  |  |  |  |  | | CO4 |  |  |  | √ |  |  |  |  |  |  |  |  | | | | |
| |  |  |  |  |  | | --- | --- | --- | --- | --- | | **CO No.** | **PO No.** | **Bloom’s**  **Domain / Level** | **Delivery Methods / Activities** | **Assessment Tools** | | CO1 | PO1 | Cognitive / Understanding | Class Lecture | Midterm | | CO2 | PO1 | Cognitive / Understanding | Class Lecture | Midterm and Final | | CO3 | PO3 | Cognitive / Applying | Class Lecture | Midterm and Final | | CO4 | PO2 | Cognitive/ Analyzing | Class Lecture | Final | | | | | | | |
| **19** | **Teaching Strategy** | Maximum topics will be covered from the textbook. For the rest of the topics, reference books will be followed. Some class notes will be uploaded on the web. White board will be used for most of the time. Multimedia projector and a PC will be used for the convenience of the students to understand codes practically. Students must participate in classroom discussions for case studies, problems solving and project developments. | | | | | |
| **20** | **Assessment and Marks Distribution:** | |  |  |  | | --- | --- | --- | | Attendance and Class Performance | **:** | 10% | | Assignment/Presentation | **:** | 10% | | Class Test | **:** | 10% | | Midterm Examination | **:** | 30% | | Final Examination | **:** | 40% | | | | | | |
| **21** | **Lecture Plan (Weekly Schedule)** | | | | | | |
| |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | | **Week** | **Lecture #** | **Selected Topics** | **Chapter #** | **COs** | **Assessment** | | 1 | 1 | Introduction to data mining, objectives, Evolution of information technology | [Jiawei Han] Ch1 | CO1 | **Mid Term Exam**  **30** | | 2 | Knowledge discovery process, kinds of data and pattern mining, Data mining functionality | Ch1 | CO1 | | 2 | 3 | Data mining technology, applications and major issues. | Ch1 | CO1 | | 4 | Data objects and attributes types. Basic statistical descriptions of data | Ch2 | CO2 | | 3 | 5 | Measuring Data Similarity and Dissimilarity | Ch2 | CO3 | | 6 | Measuring Data Similarity and Dissimilarity (Cont.), Data Cleaning. **[CT-1]** | Ch2 | CO3 | | 4 | 7 | Data reduction, Data transformation | Ch3 | CO1 | | 8 | Data discretization, Data warehouse: basic concepts | Ch3 | CO2 | | 5 | 9 | Data Warehouse Modeling: OLAP | Ch4 | CO2 | | 10 | Data warehouse design and usage, Data warehouse implementation | Ch4 | CO2 | | 6 | 11 | Data cube technology : Preliminary Concepts**,**  Computation Methods | Ch5 | CO1 | | 12 | Mining frequent patterns: Basic concepts, Frequent item sets, closed item sets, association rules | Ch6 | CO1 | | 7 | 13 | Frequent item set mining methods : Apriori Algorithm, FP Growth Algorithm | Ch6 | CO3 | | 14 | **Review class for Midterm Examination** |  |  | | **8** | **Midterm Examination** | | | | | | 9 | 15 | Pattern evaluation methods | Ch6 | CO2 | Final Exam  40 | | 16 | Classification: basic concepts, general approach to classification. Decision tree induction, ID3 Algorithm | Ch8 | CO2 | | 10 | 17 | Ensemble Classification: Bagging, Boosting, [**CT-2]** | Ch8 | CO3 | | 18 | Ensemble Classification: Random Forrest, Rule based classification | Ch8 | CO3 | | 11 | 19 | Model Evaluation and Selection | Ch8 | CO4 | | 20 | Techniques to improve classification accuracy | Ch8 | CO4 | | 12 | 21 | Cluster analysis: Partitioning methods | Ch10 | CO3 | | 22 | Hierarchical methods | Ch10 | CO3 | | 13 | 23 | Analysis of different clustering methods | Ch10 | CO4 | | 24 | Outlier analysis and outlier detection methods [**CT-3]** | Ch12 | CO4 | | 14 | 25 | Outlier analysis and outlier detection methods (Cont.) | Ch12 | CO4 | | 26 | **Review class for Semester Final Term** |  |  | | **15** | **Final Exam** | | | | | | | | | | | |
| **22** | **Overall CO Assessment Criteria** | | Assessment methods of COs are given below:   |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | | **Assessment Area** | **CO** | | | | **Assessment Area Mark** | |  | **CO1** | **CO2** | **CO3** | **CO4** |  | | Class Participation |  |  |  |  |  | | Assignment/Presentation |  |  |  |  |  | | Class Test |  |  |  |  |  | | Midterm Exam | 10 | 10 | 10 |  | 30 | | Final Exam |  | 10 | 20 | 10 | 40 | | **Total Mark** | **10** | **20** | **30** | **10** | **70** | | | | | |
| **23** | **Rubrics** | | |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | | **COs**  **(Bloom’s Level)** | **Excellent**  **(80%-100%)** | **Good**  **(70%-79%)** | **Satisfactory**  **(60%-69%)** | **Poor**  **(40%-59%)** | **Unsatisfactory**  **(0-39%)** | **Marks**  **(70)** | | CO1  (Understanding) | Answer is complete and sufficient detail provided to support issues related to the question. And also deals fully with the entire question. | Answer is brief with sufficient detail provided to support issues were introduced. And most of the basic details are included but some are missing. | Answer is brief with insufficient detail provided to support issues were introduced. | Answer is incomplete and excessive discussion of unrelated issues.  And serious gaps in the basic details. | None of the relevant details were included or didn’t answer. |  | | CO2  (Understanding) | Answer is complete and sufficient detail provided to support issues related to the question. And also deals fully with the entire question. | Answer is brief with sufficient detail provided to support issues were introduced. And most of the basic details are included but some are missing. | Answer is brief with insufficient detail provided to support issues were introduced. | Answer is incomplete and excessive discussion of unrelated issues.  And serious gaps in the basic details. | None of the relevant details were included or didn’t answer. |  | | CO3  (Applying) | The question is answered appropriately by applying the suggested method in the question. | The question is answered briefly by applying the suggested method in the question. | The question is answered correctly by applying the suggested method in the question but some steps are missing. | The question is answered incompletely by applying the suggested method in the question but some steps are correct. | No attempt to implement the suggested method. |  | | CO4  (Analyzing) | A clear, complete, and properly ordered chain of analyzing steps (i.e. proper explanation of the procedure) is followed to answer the question. | The chain of analyzing steps is complete and correctly ordered but lack of expected explanation. | One or more intermediate analyzing steps are missing or unclear, but the correctness of the analysis is not compromised. | One or more intermediate analyzing steps are missing or unclear to answer the question. | The stated chain of analysis does not lead to the stated question. |  | | | | | |
| **24** | **Grading Policy** | | The following chart will be followed for grading. This has been customized from the guideline provided by the School of Engineering and Computer Science.   |  |  |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | | A+ | A | A- | B+ | B | B- | C+ | C | D | F | | ≥ 80 | 75-<80 | 70-<75 | 65-<70 | 60-<65 | 55-<60 | 50-<55 | 45-<50 | 40-<45 | <40 | | | | | |
| **25** | **Additional Course Policies** | | Assignments | | There will be at least two assignments. Average marks of the assignments will be counted. No late homework will be accepted.  ***Any kind of copy/manipulation in assignment will carry zero mark.***  Two or more copied assignments will carry zero mark in all assignments. **Zero tolerance will be shown in this regard.** Solutions to assignment problems will be provided through web and on hand. | | |
| Class Test | | There will be at least three class tests (CT).Best two of three or best three of four CTs will be counted. Both of regular and surprise CTs can be conducted. | | |
| Exams | | CT, Mid-term and final exam will be closed book, closed notes. Mobile phone is strictly prohibited in exam hall. Students are insisted to carry their own watch and synchronize time during exam hours. | | |
| Test Policy | | If a student is absent from class test anyway and made no report to the class teacher personally beforehand, his/her score for that test will be zero. No make-up for the class test will be allowed as 2 of 3 or 3 of 4 CTs are being considered. No make-up for Mid-exam will be entertained without physical presence and recommendation of the guardian along with written permission of the department. Make-up of Mid-exam may be much harder than the regular one. | | |
| **26** | **Additional Information** | | 1. Academic Calendar Summer 2020: http://www.bubt.edu.bd/academics/academic-calendar. 2. Academic Policies: http://www.bubt.edu.bd/academics/academic-rules-a-regulations. 3. Grading & Evaluation: <http://www.bubt.edu.bd/academics/academic-rules-a-regulations>. 4. Proctorial Rules: http://www.bubt.edu.bd/administrator/proctors-office. | | | | |
| **27** | **Bloom’s Taxonomy for Teaching-Learning** | | | | | | |
|  | Bloom's Taxonomy is a set of three hierarchical models used to classify educational learning objectives into levels of complexity and specificity. The three lists cover the learning objectives in Cognitive, Affective and Psychomotor domains. The Cognitive domain list has been the primary focus of most education and is frequently used to structure curriculum learning objectives, assessments and activities. The three domains and respective levels are illustrated below.   |  |  |  | | --- | --- | --- | | **Cognitive [C]** (Knowledge-based) | **Affective [A]** (Emotion-based) | **Psychomotor [P]** (Action-based) | | 1. Remembering | 1. Receiving | 1. Imitating | | 1. Understanding | 1. Responding | 1. Manipulating | | 1. Applying | 1. Valuing | 1. Précising | | 1. Analyzing | 1. Organizing | 1. Articulating | | 1. Evaluating | 1. Characterizing | 1. Naturalizing | | 1. Creating |  |  | | | | | | | |
| **28** | **Descriptions of Cognitive Domain (Anderson and Krathwohl’s Taxonomy 2001):**  The **cognitive domain** involves the development of our mental skills and the acquisition of knowledge. | | | | | | |
|  | |  |  |  |  | | --- | --- | --- | --- | | **Level** | **Category** | **Meaning** | **Keywords** | | C1 | Remembering | Recognizing or recalling knowledge from memory. Remembering is when memory is used to produce or  retrieve definitions, facts, or lists, or to recite previously learned information. | Define, describe, draw, find, identify, label, list, match, name, quote, recall, recite, tell, write | | C2 | Understanding | Constructing meaning from different types of functions be they written or graphic messages or activities like interpreting, exemplifying, classifying, summarizing, inferring, comparing, or explaining. | Classify, compare, exemplify, conclude, demonstrate, discuss, explain, identify, illustrate, interpret, paraphrase, predict, report | | C3 | Applying | Carrying out or using a procedure through executing, or implementing. Applying relates to or refers to situations where learned material is used through products like models, presentations, interviews or simulations. | Apply, change, choose, compute, dramatize, implement, interview, prepare, produce, role play, select, show, transfer, use | | C4 | Analyzing | Breaking materials or concepts into parts, determining how the parts relate to one another or how they interrelate, or how the parts relate to an overall structure or purpose. Mental actions included in this function are differentiating, organizing, and attributing, as well as being able to distinguish between the components or parts. When one is analyzing, he/she can illustrate this mental function by creating spreadsheets, surveys, charts, or diagrams, or graphic representations. | Analyze, characterize, classify, compare, contrast, debate, deconstruct, deduce, differentiate, discriminate, distinguish, examine, organize, outline, relate, research, separate, structure | | C5 | Evaluating | Making judgments based on criteria and standards through checking and critiquing. Critiques, recommendations, and reports are some of the products that can be created to demonstrate the processes of evaluation. | Appraise, argue, assess, choose, conclude, critique, decide, evaluate, judge, justify, predict, prioritize,  prove, rank, rate, select, Monitor | | C6 | Creating | Putting elements together to form a coherent or functional whole ;reorganizing elements into a new pattern or structure through generating, planning, or producing. Creating requires users to put parts together in a new way, or synthesize parts into something new and different creating a new form or product. This process is the most difficult mental function. | Construct, design, develop, generate, hypothesize, invent, plan, produce, compose, create, make, perform, plan, produce | | | | | | | |
| **29** | **Graduate Attributes (Program Outcomes) for B.Sc. in Engineering Program based on Washington Accord** | | | | | | |
|  | Program Outcomes (POs) are narrower statements that describe what students are expected to know and be able to do by the  Time of graduation. These relate to the knowledge skills and attitudes that students acquire while progressing through the program. The students of the B.Sc. in EEE program are expected to achieve the following graduate attributes or program outcomes at the time of graduation.  **PO1–Engineering knowledge (Cognitive):** Apply the knowledge of mathematics, science, engineering fundamentals and an engineering specialization to the solution of complex engineering problems.  **PO2–Problem analysis (Cognitive):** Identify, formulate, research the literature and analyze complex engineering problems and reach substantiated conclusions using first principles of mathematics, the natural sciences and the engineering sciences.  **PO3–Design/development of solutions (Cognitive, Affective):** Design solutions for complex engineering problems and design system components or processes that meet the specified needs with appropriate consideration for public health and safety as well as cultural, societal and environmental concerns.  **PO4–Investigation (Cognitive, Psychomotor):** Conduct investigations of complex problems, considering design of experiments, analysis and interpretation of data and synthesis of information to provide valid conclusions.  **PO5–Modern tool usage (Psychomotor, Cognitive):** Create, select and apply appropriate techniques, resources and modern engineering and IT tools including prediction and modeling to complex engineering activities with an understanding of the limitations.  **PO6–The engineer and society (Affective):** Apply reasoning informed by contextual knowledge to assess societal, health, safety, legal and cultural issues and the consequent responsibilities relevant to professional engineering practice.  **PO7–Environment and sustainability (Affective, Cognitive):** Understand the impact of professional engineering solutions in societal and environmental contexts and demonstrate the knowledge of, and need for sustainable development.  **PO8–Ethics (Affective):** Apply ethical principles and commit to professional ethics, responsibilities and the norms of the engineering practice.  **PO9–Individual work and teamwork (Psychomotor, Affective):** Function effectively as an individual and as a member or leader of diverse teams as well as in multidisciplinary settings.  **PO10–Communication (Psychomotor, Affective)**: Communicate effectively about complex engineering activities with the engineering community and with society at large. Be able to comprehend and write effective reports, design documentation, make effective presentations and give and receive clear instructions.  **PO11–Project management and finance (Cognitive, Psychomotor):** Demonstrate knowledge and understanding of the engineering and management principles and apply these to one’s own work as a member or a leader of a team to manage projects in multidisciplinary environments.  **PO12–Life-long learning (Affective, Psychomotor):** Recognize the need for and have the preparation and ability to engage in independent, life-long learning in the broadest context of technological change. | | | | | | |
| **30** | **Social & Moral Capital** | | | | | | |
|  | Our promises are based on the three cardinal principles:   1. What we do believe (b) What we do practice, and (c) What we will promote   However, students are advised to undertake the following commitments for moral development. | | | | | | |
| 1. To be punctual and attentive in class 2. To maintain inclusive learning environment 3. To ensure mutual respect 4. To be cooperative in group learning. 5. To be innovative and Creative 6. To follow dress code and wearing ID card 7. To be always proactive | | | | | 1. Try to follow and review day to day class 2. To avoid conspiracy 3. To prioritize honesty & faith 4. To be motivated for asking question and encourage feedback 5. To develop attitude for speaking in English 6. Do not ignore to carry out any assignments or commitments 7. To be clean and decent in all levels. | 1. To be sincere for class preparation 2. Do not forget to switch-off the cell phone in class 3. Do not forget to carry course pack and learning stuffs in class 4. To maintain loyalty and trust to the university 5. Must avoid unfair means and plagiarism in exam, reports and assignments 6. Must maintain eco-friendly environment in the campus. |

## Prepared by: Checked by: Approved by: