

**University of Barishal**  
Dept. of Computer Science and Engineering  
**Course Code: CSE-4213; Course Title: Machine Learning and Data mining**  
**Admission Session: 2019-20; Exam: Mid Term-II; Marks: 20; Time: 1Hour**

1. A dataset contains “Age” ranging from 18–80 and “Income” ranging from 20,000–200,000. Which feature scaling technique would you use and why? 5
2. Compare min-max scaling and standardization with respect to handling outliers in a dataset. 5
3. Suggest two methods to handle imbalanced data and explain how they improve model performance. 5
4. You are building a KNN classifier to detect if a student will pass or fail based on study hours and attendance. Explain how the choice of k affects the model performance. 5

**University of Barishal**  
Dept. of Computer Science and Engineering  
**Course Code: CSE-4213; Course Title: Machine Learning and Data mining**  
**Admission Session: 2019-20; Exam: Mid Term-II; Marks: 20; Time: 1Hour**

1. A dataset contains “Age” ranging from 18–80 and “Income” ranging from 20,000–200,000. Which feature scaling technique would you use and why? 5
2. Compare min-max scaling and standardization with respect to handling outliers in a dataset. 5
3. Suggest two methods to handle imbalanced data and explain how they improve model performance. 5
4. You are building a KNN classifier to detect if a student will pass or fail based on study hours and attendance. Explain how the choice of k affects the model performance. 5

**University of Barishal**  
Dept. of Computer Science and Engineering  
**Course Code: CSE-4213; Course Title: Machine Learning and Data mining**  
**Admission Session: 2019-20; Exam: Mid Term-II; Marks: 20; Time: 1Hour**

1. A dataset contains “Age” ranging from 18–80 and “Income” ranging from 20,000–200,000. Which feature scaling technique would you use and why? 5
2. Compare min-max scaling and standardization with respect to handling outliers in a dataset. 5
3. Suggest two methods to handle imbalanced data and explain how they improve model performance. 5
4. You are building a KNN classifier to detect if a student will pass or fail based on study hours and attendance. Explain how the choice of k affects the model performance. 5