SRM Institute of Science and Technology
School of Computing
DEPARTMENT OF DATA SCIENCE AND BUSINESS SYSTEMS
SRM Nagar, Kattankulathur – 603203, Chengalpattu District, Tamilnadu

Academic Year: 2024-25 (ODD)

Test: CLAT-1 [Portion: UNIT 1 & First half of 2nd Unit]	**Date: 23.02.2024**
Course Code & Title: 21CSE426T & Financial Machine Learning	**Duration:** 100 Minutes
Year & Sem: III Year & VI Sem	**Max. Marks:** 40

Part A [Answer all / 10 * 1 = 10 Marks]

1. What is the primary benefit of using machine learning in algorithmic trading?
A) Slower trade execution B) The ability to manually calibrate trading strategies C) Advanced strategies that can adapt in real-time D) Limiting the role of automated trading instructions

2. Which of the following applications in finance is most concerned with using algorithms to calibrate a financial portfolio to the goals and risk tolerance of a user?
A) Fraud Detection B) Robo-Advisors C) Trade Settlement D) Asset Price Prediction

3. Which of the following is NOT mentioned as an application of (NLP) in finance?
A) Sentiment analysis B) Derivative pricing C) Chatbots D) Document processing

4. In a feedforward neural network, how are the weights updated during training?
A) Using gradient descent B) By random initialization C) Through the activation function D) By normalizing inputs

5. What is the primary function of the activation function in an artificial neural network?
To initialize weights B) To introduce non-linearity C) To calculate the loss D) To perform backpropagation

6. Which of the following kernels is most commonly used in SVM for non-linear data?
A) Polynomial B) Sigmoid C) RBF (Radial Basis Function) D) Linear

7. Which algorithm is referred to as a "lazy learner" because it requires no training phase?
A) KNN B) SVM C) Logistic Regression D) Linear Regression

8. _____ represents the cost function used in linear regression?
A) $RSS = \sum_{i=1}^{n}(y_i - \hat{y}_i)$ B) $RSS = \sum_{i=1}^{n}(y_i - \beta_0 - \sum_{j=1}^{p}\beta_j x_{ij})^2$ C) $RSS = \sum_{i=1}^{n}(y_i - x_i)$ D) $RSS = \sum_{i=1}^{n}\frac{(y_i - \hat{y}_i)^2}{n}$

9. $CostFunction = RSS + \lambda * \sum_{j=1}^{p}\beta_j^2$ is the _____ regularization
a. Lasso   b. Rigid   c. Elastic net   d. Elastic grid

10. Using kernels, the original data is projected into a _____ to classify the data better.
a. Lower dimension b. Higher dimension   c. hyperplane   d. boundaries

Part B [Answer any 2 / 3 * 15= 30 marks]

1. You are tasked with developing a SVM model to classify emails as spam or not spam based on the email's content and metadata.
- Explain the concept of the kernel trick in SVM and why it might be useful for this classification problem. Also, discuss the choice of kernel (linear, polynomial, RBF) you would use and why.
- Write a Python code snippet using sklearn to implement the SVM model with your chosen kernel. Include the steps for training and testing the model on a dataset of emails.
- Discuss the steps involved in tuning the hyperparameters of the SVM model using GridSearchCV, and explain how you would evaluate the model's performance.

2. You are analyzing customer data for a retail company to predict whether a customer will make a purchase during a sale event. The input features include customer demographics, browsing behavior, and past purchase history.
- Explain the concept of the logistic function and how logistic regression can be used for binary classification in this scenario.
- Write a Python code snippet using sklearn to implement logistic regression for this task. Include the steps for training the model and interpreting the coefficients.
- Discuss the steps involved in checking for multicollinearity among the features and how you would handle it if detected.

3. Assume you are working on a recommendation system where you need to suggest movies to users based on their past ratings. You decide to use the KNN algorithm for this task.
- Discuss the importance of the value of 'k' in the KNN algorithm and its impact on the model's performance.
- Write a Python code snippet using sklearn to implement a KNN-based recommendation system, including the selection of the optimal 'k' value.
- Describe the steps involved in cross-validation and how you would evaluate the performance of the recommendation system.