M.TECH DATA SCIENCE AND MACHINE LEARNING - SEMESTER I

UE20CS905 - MACHINE LEARNING - I

INSTRUCTIONS
• All questions are compulsory. • Section A should be handwritten in the answer script provided and signed at the end of the same. • Section B and C are coding questions which have to be answered in the system.

Section – A

1. What is Machine learning? State any two types of machine learning.
2. Explain the type of supervised Machine Learning models with example
3. State few applications of Machine Learning.
4. Explain some issues with Machine Learning.
5. How can you handle overfitting and underfitting?
6. Explain data pre-processing steps.
7. What is Linear Regression in Machine Learning?
8. Explain the assumptions of Linear Regression.
9. What are the benefits of using Linear Regression?
10. Write the expression for the cost function, which need to be minimized in Linear regression with RIDGE regularization.
11. What is the difference between Classification and Regression problem?
12. What is Multicollinearity? How to detect the presence of multicollinearity and which variables are involved in it?
13. Write about different variance measures involved in the operation of Linear Regression.
14. How the problem of overfitting can be reduced in Linear Regression? What is bias variance tradeoff?
15. If $y = 2x_1 + 12x_2 + 3x_3 + 5$ is the linear regression equation, then explain how the coefficients of x_1 and x_2 affect the value of y.
16. Discuss the use of Rsquared and adjusted Rsquared in a Linear Regression model?
17. How adjusted R-square is differing from R-square? Brief the role of adjusted R-square in feature selection process.
18. What is k-fold cross validation? Write briefly about the procedure.
19. Explain Reciprocal Transformation Technique.
20. Explain the procedure involved in Forward Feature Selection.

21. Explain Gradient Descent in brief
22. What is Lasso Regularization?
23. A linear regression model is build with three independent variable price, advertisement cost and promotion cost to predict unit sales of mobile phone. Say the p value for the t-test of the variable 'advertisement cost' is 0.02. What is your inference on this?
24. The RMSE of the regression model which predicting the CTC salary is 12324 and the RMSE of the other regression model which predicting the age of the person is 55. Comment on the performance of these two models. [output column is not scaled or transformed]
25. If we increase the value of lambda, what will happen to the estimated coefficients in RIDGE model?

Section – B & C

1. House Price Prediction model.
2. Airline companies dynamic pricing model.
3. Engineering graduate salary model
4. Salary of the student's prediction model
5. Traffic volume Prediction model
6. Stock Market Prediction model
7. Loan Default Prediction model
8. Market Sales Forecasting model
9. Advertising model
10. Health Insurance model