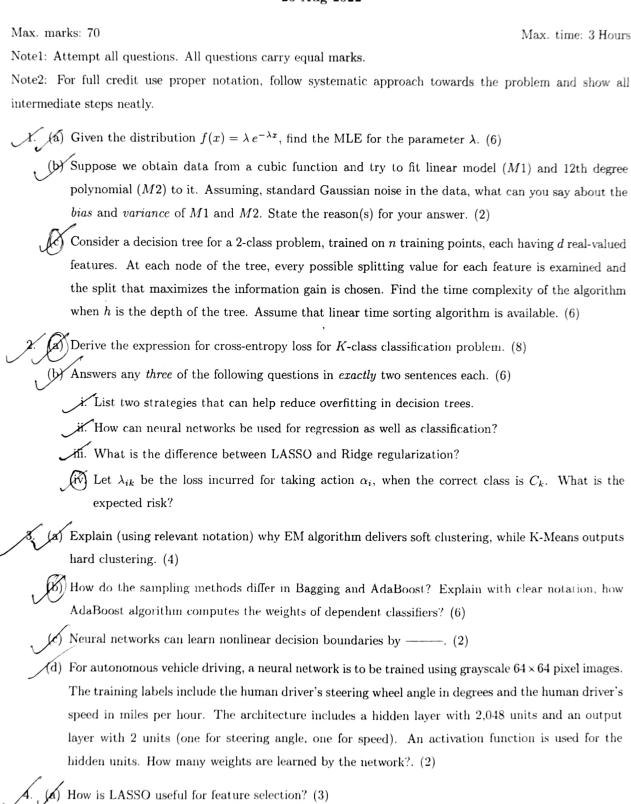
MSc Computer Science

MCSC 201: Machine Learning

UPC: 223411201

Semester II

25 Aug-2022



(b) What is color quantization? Explain how is K-Means clustering used for color quantization (5)

(c) Consider 4-dimensional training set with N instances, categorized in 3 classes (Appearance), with details of attributes given below.

Age [Young, Mid-age, elderly]

Weight [heavy, light]

Weight [heavy, light]

Dress [Western, Ethnic]

Appearance [Smart, Dashing, Unimpressive]

List the conditional probabilities that need to be computed for NB classifier. What is the size of the model? (6)

5. (a) Given the bi-variate normal data partitioned into two (practically) linearly separable classes, draw iso-probability contours when

- i. the two classes have diagonal covariance matrix with equal variances on both dimensions
- ii. the two classes have diagonal covariance matrix with unequal variances on dimensions
- iii. the two classes have arbitrary but shared covariance matrix with equal variances

The figures should be labeled neatly. (6)

Show that for two Gaussian classes with shared covariance matrix Σ , and respective means μ_1, μ_2 , the log odds is linear. (5)

How many binary classifiers are required to solve a k-class classification problem? Explain your answer. (3)