

Machine Learning Problems

(AUTUMN-2024 LAB Assignment Examination)

GR: Group

GR1. Consider a random vector, a collection of random values of n-dimension [min 'n' = 100]. Fit Poisson's distribution and Gaussian distribution for these collected values.

GR2. Use Iris and Heart Disease database to perform the following:

- Classification tasks include 50-50% training and testing using the SVM classifier.
- Apply PCA to the dataset.
- Perform classification with 50-50% training and test on 1st PCA component and 2nd PCA component.

GR3. Use Heart Disease database:

- Perform classification with 50-50% training and test before data normalisation.
- Apply the min-max and z-score normalization technique to the heart disease dataset, and then apply classification with 50-50% training and test. [Use SVM classifier]

GR4. Use the Heart Disease database - apply singular value decomposition (SVD) to analyse the decomposed features derived from the SVD technique. Apply data visualisation on the outcomes of SVD.

GR5. Use the Chronic Kidney Disease database to perform factor analysis to distinguish between the features and arrange those features based on their importance.

GR6. Use the 'Optical Recognition of Handwritten Disease' database:

- Apply several classifiers such as Naive Bayes, KNN, Random forest, and Decision Tree.
- List the performance of these classifiers in tabular form, such as:
Precision | recall | F1_score | accuracy
- Compare the F1_score to justify the best classifier for this database problem.

GR7. Suppose using the "Optical Recognition of Handwritten Disease" database SVM classifier gives better performance, then list the tabular form such as:

Nature	Learning_rate	F1_score
linear	0.1	-
	0.3	-
	0.5	-
poly	0.1	-
	0.3	-
	0.5	-
radial	0.1	-
	0.3	-
	0.5	-

Justify what will be the best nature of the SVM classifier for that database.

GR8. Use the 'Real E-state Valuation' database:

- Use linear regression, and logistic regression to justify the outcome using the database.

GR9. Use Diabetes 180 US Hospital's database, where both classification and clustering techniques can be applied on this database. Use reinforcement learning on this database to justify and analyze the features of this database.

GR10. Suppose there are 3 disease detection systems — diabetes, heart disease, thyroid disease. The diabetes system generates score '-d', heart disease generates score '-h' and thyroid system generates score '-t'.

- i. Use fuzzy inference system with (small, medium, large) or (low, moderate, high) to perform a system to derive a final score by combining the score of this disease. For this purpose, build fuzzy fusion system.

GR11.

- i. Use 'Thyroid Disease' database to perform classification with different classifiers and list in tabular form.
- ii. Visualize the clustering of all 5 features with their differences in a single plot.

GR12. Use different loss function to classify the nature of features of 'Magic Gamma Telescope' database.

GR13. Use Artificial Neural Network (ANN) classifier for the classification purpose using 'Magic Gamma Telescope' database:

- i. A (1-1-1) NN architecture.
- ii. A (1-2-1) NN architecture.

GR14. Use 'Magic Gamma Telescope' database and apply ANN with 1 input, 2 hidden, 1 output layers. Report the performance using various Activation function such as:

- i. sigmoid
- ii. tanh
- iii. ReLU
- iv. Leaky ReLU