# DINH HOANG VIET PHUONG – 301123263

# LAB 01 REPORT

4-11. Libraries are imported for numerical operations, plotting, data generation, decomposition, model selection, and linear models.

n\_samples = 1500: This sets the number of data points in the Swiss roll dataset to 1500.

noise = 0.05: This adds some noise (variance) to the generated dataset.

X, color = make\_swiss\_roll(n\_samples, noise=noise): Generatesdhvietphuonggergerhgreh the Swiss roll dataset with the given parameters.

20-28. This block of code plots the generated Swiss roll dataset in a 3D space.

A graph of a graph with colored dots

Description automatically generated with medium confidence

33-42. Applies kPCA with a linear kernel to the dataset and then plots the transformed data.

A diagram of a spiral

Description automatically generated

44-54. Applies kPCA with an RBF kernel to the dataset and then plots the transformed data.

A diagram of a graph

Description automatically generated

56-66. Applies kPCA with a sigmoid kernel to the dataset and then plots the transformed data.

A diagram of a diagram

Description automatically generated with medium confidence

73. Splits the dataset into a training set and a test set.

75-79. Sets up a pipeline with kPCA and logistic regression. This means that the data will first be transformed by kPCA and then be fed into the logistic regression for classification.

81-85. This sets up the parameter grid for GridSearchCV. It will try different combinations of gamma values and kernels to find the best for kPCA.

87-92. Initializes and fits GridSearchCV to find the best hyperparameters for the model. The classification task here is set as binary, where data points with colors above the median are one class, and those below are another.

Best parameters found: {'kpca\_\_gamma': 0.05, 'kpca\_\_kernel': 'rbf'}

Best cross-validation score: 0.68

93-96. Prints the best hyperparameters found by the grid search and the corresponding cross-validation score.

Test set accuracy with best parameters: 0.64

99-107. This block of code sets up a grid of values in the original feature space.

109-111. Transforms the grid using the best kPCA model and predicts the classes using logistic regression.

113-122. Plots the decision boundary of the best model. Points in the training set are marked with circles while points in the test set are marked with squares. The decision boundary is displayed as a contour plot.

A diagram of a decision boundary

Description automatically generated with medium confidence