BIRZEIT UNIVERSITY

Department of Electrical and Computer Engineering
Second Semester, 2023/2024
Intelligent Systems Lab, ENCS5141

Case Study #2: Comparative Analysis of Classification Techniques: Support Vector Machines (SVM) and Multilayer Perceptron (MLP)

In this case study, you will study and compare the performance of two popular classification techniques: support vector machines (SVM) and multilayer perceptron (MLP). Through this study, you will grasp the strengths and limitations of each algorithm and gain insight into their applicability in different scenarios. The study will be conducted on a dataset comprising over 24,000 records (samples) that depict banknotes across various accessibility scenarios. Each entry in the dataset represents a unique set of 256 features and includes corresponding labels for currency types (such as AUD, USD, CAD, etc.), denominations (e.g., 10, 100, 50, etc.), and orientations (front=1, back=2).

Table: Example rows from the dataset

v_0	v_1	v_2	....	v_254	v_255	Currency	Denomination and Orientation
5.144637	0.0	0.7171247	....	4.7246137	0.0	AUD	100_1
2.6717074	0.0	0.31792498	....	2.6489065	0.6563814	AUD	100_2
2.4650123	0.0	0.19788292	....	0.8239467	1.5399158	AUD	100_2
0.2279757	1.1194861	2.6831365	....	0.65020627	2.4121835	AUD	10_1
0.45672822	0.0	2.1918573	....	1.0260967	1.9409924	CAD	10_1

The dataset can be accessed through the "BankNotesDataset.csv" file in the GitHub repository: https://github.com/mkjubran/ENCS5141Datasets.

Train and evaluate both SVM and MLP models to predict currency, denomination, and orientation. In other words, the classifiers should predict the currency, denomination, and orientation of the banknote based on the provided features. To accomplish this, we propose two options:

Option 1: Create three separate classifiers, each for predicting one label (currency, denomination, or orientation).

Option 2: Merge currency, denomination, and orientation into a new label and create a single classifier to predict this new combined label.

You must implement and compare both options using SVM and MLP. Also, utilize hyperparameter tuning methods like grid search or other optimization techniques. Additionally, apply appropriate dimensionality reduction techniques to reduce data size while retaining essential information. Finally, fine-tune and compare the performance of both models before and after dimensionality reduction.

Submissions:

- You need to submit the code in .ipynb format. You can obtain this file in Google Colab by navigating to the File menu and selecting Download > Download .ipynb.
- Additionally, write a report detailing the case study. Ensure adherence to the report preparation guidelines outlined in the "ENCS5141 Case Study Report Guidelines.pdf" document. If you opt to write the report using LaTeX, utilize the provided report template "ENCS5141 Sample Report.tex"."

Important Notes:

- Make sure to add descriptive comments and headings using markup language, such as Markdown, in your Google Colab notebook or Jupiter notebook.
- **Deadline**: Tuesday, 30 April 2024 at 11:59 pm. Please submit your case study solution and report through Ritaj as a reply to this message.
- **Late Submission Policy**: One mark (out of the 40 marks assigned to case studies in the course outline) will be deducted for every day of late submission of the case study report. No submissions will be accepted beyond the third day past the due date.