

Assignment-4

Topic: Support Vector Machine

(Dr. Debendra Muduli, Assistant Professor, CGU, Odisha.)

1. In this Assignment, we will explore the following concepts been studied in class :
SVM (with different kernels like Linear, Polynomials and RBF) with Feature reduction and without Feature reduction techniques. Here, you will considered PCA as a feature dimensionality reduction technique.

It involves a 3-step work:

- a. Preprocessing
- b. Feature Dimensionality Reduction
- c. Reporting performance metrics

Pre-processing:

- i. Read the given dataset.
- ii. Standardized the dataset (You can use a standard scaling technique in this study for standardizing the data set. Use StandardScaler)
- iii. Create a train-test split of 70-30.

Feature Dimensionality Reduction

- i. Apply Principle Component Analysis (PCA) to reduce the dimension of the feature vector.

Performance Metrics

- i. Now apply SVM to validate your proposed model performance. (**Note: Consider the degree of the polynomial is 2**)

Table-1 (Without Feature Dimensionality Reduction)

Methods	Dataset Name (.....)						
		Accuracy	Sensitivity	Specificity	Precision	F-1 Score	**MCC
SVM (Linear)							
SVM (Polynomials)							
SVM (RBF)							

Table-2 (With Feature Dimensionality Reduction technique (PCA))

Methods	Dataset Name (.....)						
	Number of Features	Accuracy	Sensitivity	Specificity	Precision	F-1 Score	**MCC
PCA+ SVM (Linear)							
PCA+ SVM (Polynomials)							
PCA+ SVM (RBF)							

- ii. Draw the Confusion Matrix of the best classifier with highest classification Accuracy.
- iii. Compared the ROC curves obtained by different classifiers (Linear SVM, Polynomial SVM and RBF SVM)

Submission Instructions:

1. Please submit your codes with results filename as **MLAssignment-4_Group_No_.py** for e.g. **MLAssignment-4_Group-1.py (Python File)**
2. Submit a write-up document with the desired results and plots and observations in pdf format, naming it as **MLAssignment-4_Group-1.pdf (PDF File)**.
3. Save the python file and the report in a folder and give the folder name as **group name**, and upload in the given Google Drive link.
<https://forms.gle/5Y1UhTvRCdEcutqd8-G-3>
4. Dead line of the submission: **15/11/2023**.

Please check your group number from the given list.

Group No.	Regd.No.	Name	Dataset
Group-1	2101020500	DATLA LOKESH REDDY	CGU_Dataset-1
	2101020501	SARDA PRASAD SAHOO	
	2101020502	ARPIT POLEI	
	2101020503	GOLAP CHAND MALLIK	
	2101020504	DIVYA SINGH	
	2101020505	ANURAG KUMAR	
	2101020506	ARPIT ANAND	
	2101020507	TANIYA ANSHU	
Group-2	2101020509	SAJAN KUMAR	CGU_Dataset-2
	2101020510	BISWAJIT NAYAK	
	2101020511	RAJESH KUMAR GOUDA	
	2101020512	BISWAJEET SAMAL	
	2101020513	RASHMI GOSWAMI	
	2101020514	AAYUSH ASHOK KASHYAP	
	2101020515	BISWARANJAN PRADHAN	
	2101020516	PUJA MAHAPATRA	
Group-3	2101020686	KUMAR DEBADATTA DAS	CGU_Dataset-3
	2101020701	SRIMANTA KUMAR ACHARJYA	
	2101020702	OMM PRATIK PARIDA	
	2101020703	DIBYASHA PRIYADARSHINI	
	2101020704	DIGVIJAY PRADHAN	
	2101020705	KSHIRABDI TANAYA RATH	
	2101020706	ARPITA MOHANTY	
	2101020707	BISWARANJAN SAHOO	
Group-4	2101020708	DIBYAJYOTI BEHERA	CGU_Dataset-4
	2101020709	BISMAY BISWAJIT SWAIN	

	2101020710	ANANYA PRATYOTPURNA NAYAK	
	2101020711	APARNA SANTOSHI PATRA	
	2101020712	PRATIK KUMAR SAHU	
	2101020713	JYOTI RANJAN PUHAN	
	2101020714	SWAPNIL KUMAR PATEL	
	2101020715	JATIN KUMAR PATRA	
Group-5	2101020716	RAJNISH KUMAR	CGU_Dataset-5
	2101020717	ARUN KUMAR BARIK	
	2101020718	SATYA PRAKASH SAHU	
	2101020719	DIBYA RANJAN SARANGI	
	2101020720	MOHIT RAJ EKKA	
	2101020728	SHUBHAJIT DAS	
	2101020729	SWAYAM PRASAD SAHOO	
	2101020735	YASH	
Group-6	2101020742	RISHEETA PANDA	CGU_Dataset-6
	2101020743	DIVYA	
	2101020744	PRAKRUTI JENA	
	2101020745	BISWA SWARUP TRIPATHY	
	2101020748	AMRUTANSHU RATH	
	2101020756	Sushant Jha	
	2101020757	SAMRAT ADITYAKESHARI PARIDA	
	2101020758	SHANKAR SINGH MAHANTY	
Group-7	2101020759	R NITISH KUMAR PATRA	CGU_Dataset-7
	2101020760	DEBASIS KUMAR GHADAI	
	2101020761	SHREYA SUMAN	
	2101020772	SWAYAMDIPTA MISHRA	
	2101020773	GURUPRASAD BHUYAN	
	2101020775	ANKUR KUMARI	
	2101020776	SHRISTY KUMARI	
	2101020779	HARSH KUMAR	
Group-8	2101020783	TUSHAR INDRA	CGU_Dataset-8
	2101020784	NIKHIL PANIGRAHI	
	2101020788	SB ASHIRBAD	
	2101020794	BIBHUDATTA JENA	
	2101020795	AYUSHI KASHYAP	
	2101020796	PRABIN KUMAR BISWAL	
	2101020798	SHASHWATI JHA	
	2101020802	Sameep Kumar Singh	
Group-9	2101020803	Sanil Prasad Gorkhali	CGU_Dataset-9
	2101020818	SHIVAM MAHESH	
	2101020819	ARNAV RANA	
	2101020826	Roshan Sah	
	2101020827	Sandesh Bhatta	
	2101020828	Girijesh Kumar Rajak	
	2101020829	Aayush Adhikari	
	2101020830	Aman Kushwaha	
Group-10	2101020834	Ramila Gurung	CGU_Dataset-10
	2101020847	BIGHNESH BARIK	
	2101020848	Aaron Sapkota	
	2101020849	Binayak Ojha (G)	

	2101020850	Rahul Kumar Gupta	
	2101020851	Roman Keshari	
	2101020852	Abhinav Kumar Singh	
	2101020856	Suman Bashyal	
Group-11	2101020858	Samir Kumar Majhi	CGU_Dataset-11
	2101020811	INDRADEV SAW	
	2201030001	IRFAN PATHAN .	
	2201030002	SOUJANYA MOHAPATRA	
	2201030003	SHIVAM NARAYAN .	
	2201030004	RACHITA	
	2201030005	ABHILIPSA MOHANTY	
	2201030006	AKSHANSH SINGH.	
	2201030007	ANKIT KUMAR KHANDA	
	2201030008	ASISH KUMAR BARIK	

Best of Luck.....Keep Learning