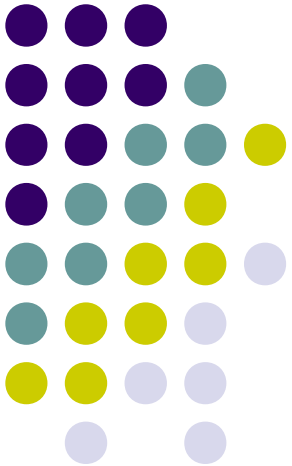
Pattern Recognition

Lab3:
LDA
Support Vector Machines

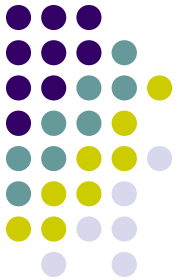
Francesco Tortorella

University of Cassino and Southern Latium

Cassino, Italy

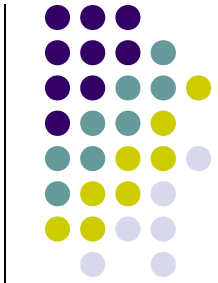


Lab 3.1



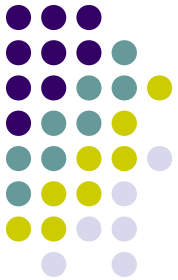
- Read the file ‘pima-indians-diabetes.data’
- Part the set into two subsets: assume one (PimaTr) as training set and the other one (PimaTest) as test set
- Starting from the training set, build a quadratic classifier in the original feature space.
- Compute the LDA projections using only the training data.
  - Project both training and test data using the LDA weights
  - Classify the test data in the LDA subspace. Repeat the above steps several times. What is the average AUC?
- Consider different sizes of the training set (25%, 50%, 75%) and analyze how the AUC changes.

Lab 3.2



- Read the file ‘pima-indians-diabetes.data’
- Part the set into three subsets:
  - Training set (PimaTr) [30%]
  - Validation set (PimaVal) [30%]
  - Test set (PimaTest) [40%]
- Use the training set for building a Support Vector Machine and the validation set for choosing the optimal value of its parameters. Evaluate the AUC on the test set.
- Execute the steps above with different kernels of the SVM:
  - Linear
  - Polynomial
  - RBF
- Discuss the results obtained

Lab 3.3



- Read the file ‘pima-indians-diabetes.data’
- Part the set into three subsets:
  - Training set (PimaTr) [30%]
  - Validation set (PimaVal) [30%]
  - Test set (PimaTest) [40%]
- Compute the LDA projections using only the training data.
  - Project both training, validation and test data using the LDA weights
  - Apply the SVM models considered in Lab 3.2 in the LDA subspace and evaluate the AUC obtained on the test set.
- Discuss your results