Domande Machine Learning

**2nd call July 2023**

* PCA

Describe the Principal Component Analysis (PCA) dimensionality reduction approach, focusing on

* + The goal of the model and its formulation
  + The training objective function
  + The relationship between the data distribution and the PCA solution
  + Practical considerations when employing PCA
* LDA
  + The goal of the model and its formulation
  + The training objective function
  + The relationship between the data distribution and the LDA solution
  + Practical considerations when employing LDA
* Q1 describe and compare PCA and LDA covering the following aspects:
  + Goals of the 2 models and their formulations
  + Training objectives of the 2 models
  + characteristics of the PCA principal components and the LDA discriminant directions
  + how the 2 models can be employed for classification

Q2)- Describe the GMM in the context of density estimation and pattern classification , covering the following aspects:

\* definition of the model interpretation of the model parameters and formulation of the GMM as a latent variable model

\* estimation of the models params

\* how the model can be used to solve classification problems, including open-set tasks.

\* potential issues of GMMs, possible way to address this issues and possible variations of the models.

He asked mvg standard, tied and naive bayes

Then LR, comparison with SVM losks function and how to extend them to build non linear decision boundaries

Describe the binary logistic regression model for classification, focusing on

The model assumptions and the classification rule of the logistic regression model

The probabilistic interpretation of the model and of the classification score

The training objective and its different interpretations

Limitations of the standard model and possible extensions to address these limitations

Describe Gmm in context of density estimations

* ﻿Estimation of their model parameters
* ﻿Issues of gmm and how to address them

﻿how is gmm suitable for open set tasks or something like that

He asked mvg standard, tied and naive bayes

Then LR, comparison with

SVM loss function and how to extend them to build non linear decision boundaries

MVG and Binary Logistic Regression

comparison between svm loss function and logistic loss function

Pca and Lda:

Their training objectives

Goals

How they can be employed in classification tasks

Characteristics of principal components and linear discriminant directions

But like I showed the maximum likelihood formula and then I showed how it applies for both Gaussian and logistic regression

Also I showed how the formula for the probably changes with the

different Gaussian classifiers (tied, naive ecc)

Domande teoria

3. How could you implement non - linear models, how would you use and make work non linear model

4. normalization changes PCA directions?

5. what is PCA doing? how do you compute PCA

7. DCF how to compute

8. what are generative approaches?

9. SVM, how to use it non linearly

10. What we use to estimate the parameters for a gaussian distribution in a MVG

11. What is the loglikelihood ratio?

12. How do you obtain the threshold from the training data

13. Assumption made from the logistic regression model

14. How the logistic regression model works and what is the function we are trying to minimize?

15. How do you train the gaussian classifier model?

16. How LDA works, assumptions and limitations.

17. What rank is the within covariance matrix?

18. GMM what they are?

19. Relationship between LDA and tied gaussian covariance classifier

20. SVM in general how do they work how to make them non linear

21. How do you train a GMM, what do they achieve and how do they work

22. What we do when we apply the EM algorithm

23. What are responsibilities

24. How accuracy and DCF are related

25. How did you apply kfold to compute the minDCF

26. What are the meaning of SVM primal and dual solutions?

27. What the hyperparameters in SVM represent?

28. How do you calculate actual DCF and what we use it for?

29. When we calibrate scores using prior-weighted Linear regression, which data should we give as input to the function to estimate w and b? (I'm not speaking about the "x" that are the scores that we have to calibrate)

30. How do we model discrete values?

31. How do we make decisions?

32. Relationship between SVM and Logistic Regression + differences

32. What do Lagrangian multipliers represent

33. What does the kernel function represent

34. What is dimensionality reduction

35. What is the reconstruction error

36. In LDA, why do we first whiten Sw and then Sb

37. How would you do a naive Bayes GMM

38. What are models that describe discrete data

39. How would you compute optimal decisions

40. Qualitative definition of Bayes risk (not the formula)

41. Generative vs discriminative approach

42. Hinge loss vs Logistic loss

43. Speak about multiclass Logistic Regression

44. What is RBF kernel

45. What is KNN

46. In the report how do you explain the differences between the results with validation and evaluation sets