## Section A - a very detailed answer is required

1. Describe SVM and its training procedure. What is a kernel and what is it used for in SVM?

2. Describe forward neural networks with error propagation and the procedure of their training.

3. Describe the self-organizing maps and the procedure of their training.

4. Describe the linear classifier and its training procedure. 5. Describe the K-means algorithm, the choice of K and possible alternatives.

## Section B - a moderately detailed answer is required

1. Describe the FLDA method in detail.

2. Describe the ICA method in detail.

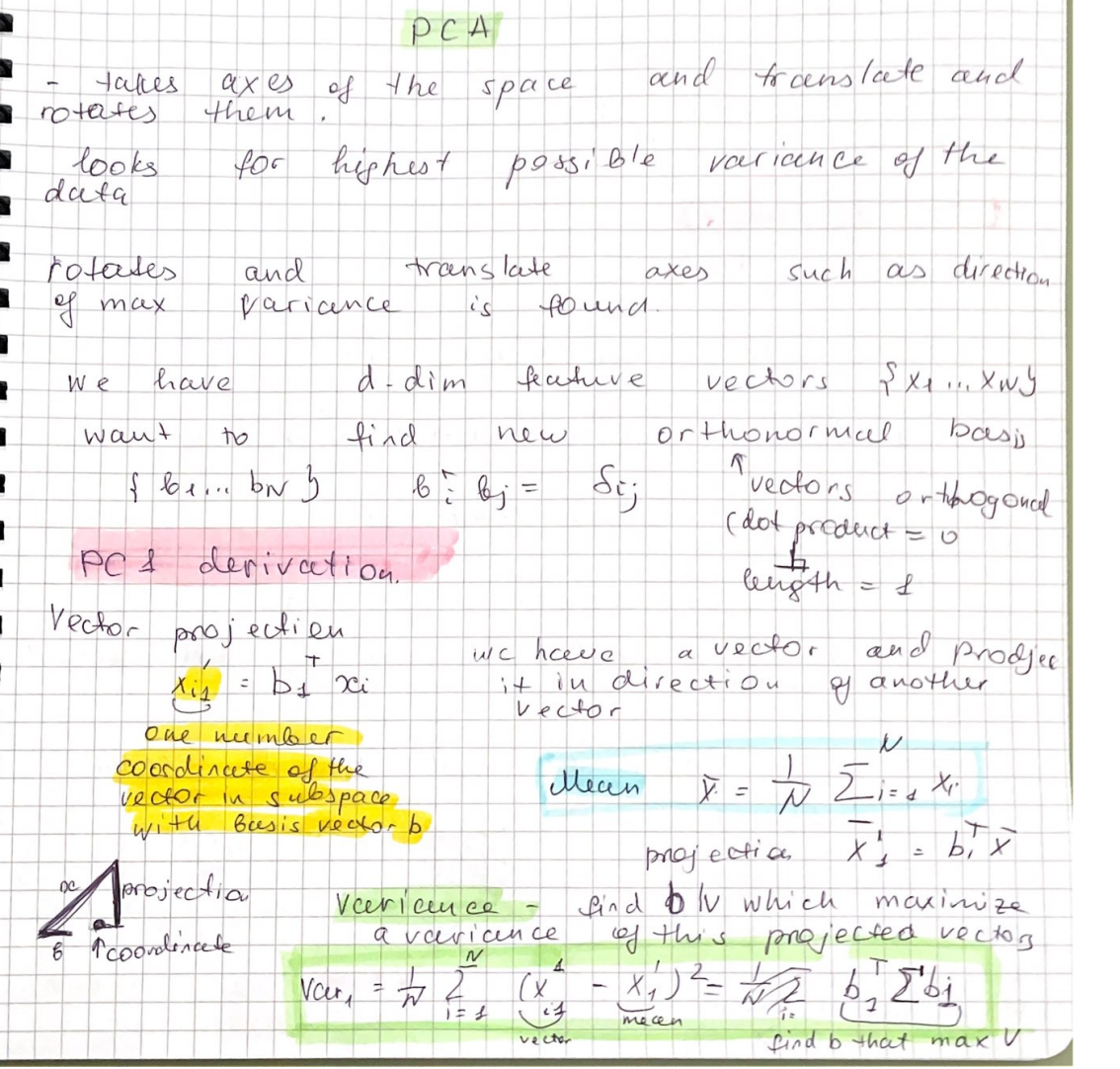
<https://towardsdatascience.com/introduction-to-ica-independent-component-analysis-b2c3c4720cd>9

3. Describe in detail the K nearest neighbors’ method.

4. What is the difference between the two basic approaches to dimension reduction (wrapper and filter), what evaluation measures are used for them?

5. What is the difference between the two basic approaches to dimension reduction (wrapper and filter), what procedures are used to construct the subsets?

6. Describe the PCA method in detail. What is the difference between PCA and ICA?



A notebook with writing on it

Description automatically generated

A paper with writing on it

Description automatically generated

Difference: in ICA – components not orthogonal

A graph of vector graphics

Description automatically generated with medium confidence

7. Describe the PCA method in detail. How do we use it to reduce the dimension?

A close-up of a piece of paper

Description automatically generated A close-up of a math note

Description automatically generated

8. Describe the PCA method in detail. How do we determine the number of significant eigenvectors?

6+7

9. Describe in detail the training of the Bayesian classifier. What do we know about the error of this classifier?

10. Describe how decision trees are created.

11. Describe the cross-validation methods and the bootstrap method. What is the main difference between them? What are they used for?

12. Describe convolutional networks in detail.

13. Describe the methods of hierarchical clustering in detail.

## Section C - the answer consists of 2-3 sentences

1. What is a confusion matrix, how is it created? What is it used for?

2. What is the ROC curve, how is it generated? What is it used for? What values can we derive from it?

3. What methods do we use to select a training and test set?

4. What is the difference between PCA and ICA?

Difference: in ICA – components not orthogonal

A graph of vector graphics

Description automatically generated with medium confidence

5. What is the difference between PCA and LDA?

6. In what ways do we normalize features?

7. What is the error function of the classifier? How is the classifier error estimated?

8. What are the filter and wrapper for? What is the difference between them?

9. Describe the classification using the Bayesian classifier.

10. Describe the classification using a linear classifier.

11. Describe the classification using decision trees.

12. Describe the SVD method. What is it used for?

We can use SVD instead of RCA

A paper with writing on it

Description automatically generated

13. Describe the measures used to evaluate features.

14. Define the terms entropy and mutual information. What are they used for when choosing features? 15. Define the terms statistical independence and consistency. What are they used for when choosing features?

16. What clustering methods do we know? Give basic division and examples of methods.

## Section D - practical questions

1. Convert data from 2D to 1D space using PCA.

https://www.youtube.com/watch?v=MLaJbA82nzk&ab\_channel=EduFlairKTUCS

2. Use the naive Bayesian classifier to classify the object.

3. Create a decision tree.

4. Calculate the number of CNN parameters and the size of the individual outputs in the layers