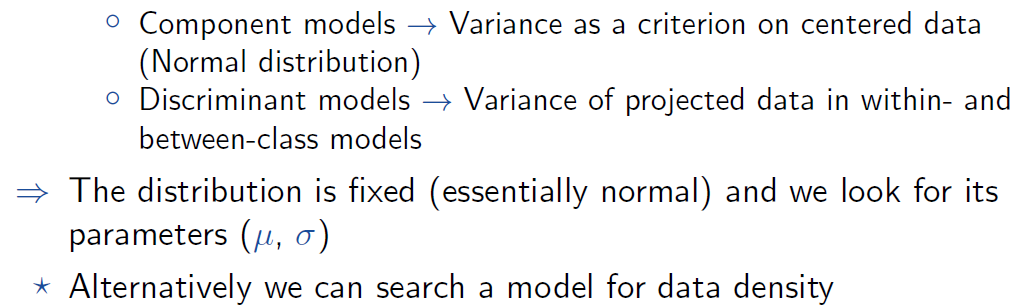
Data Science ̶ Part I

# Explain how to generate in practice samples from a given distribution. You may also discuss the particular cases of the Uniform and Gaussian distributions.

+ TP2

A picture containing diagram

Description automatically generated

# Explain in detail the method of Principal Component Analysis (PCA) including assumptions, calculation, and interpretation of the technique.

+ TP 3

# When performing Principal Component Analysis (PCA), how should one choose the number of eigenvectors to keep? Explain why and discuss the implication on the reconstruction.

Calendar

Description automatically generated

# Explain the method of Factorial Component Analysis (FCA) including assumptions, calculation, and interpretation of the technique. In particular, discuss the specificity of the data it operates on.

+ TP 3

A piece of paper with writing on it

Description automatically generated with medium confidenceA piece of paper with writing on it

Description automatically generated with medium confidence

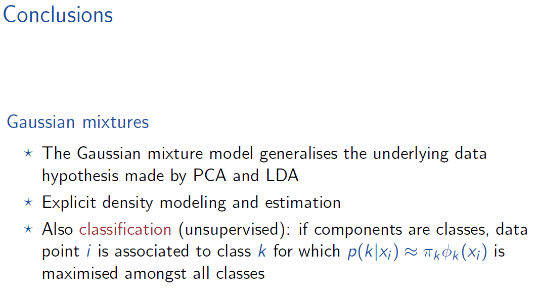
# Explain the method of Linear Discriminant Analysis (LDA) including assumptions, calculation, and interpretation of the technique. How is this technique used for classification?

+ TP 4

Table

Description automatically generated with low confidence

# Explain the Gaussian Mixture Model including assumptions, calculation, and interpretation of the model. What are its parameters? Explain how to optimize them given some data.

A picture containing chart

Description automatically generated



# What is the EM algorithm and how is it related to the Gaussian Mixture Model? In particular, explain the concept of responsibility.

+ EM-part 2 in lecture videos (review)

Table

Description automatically generated with low confidence

# Explain the problem of clustering and describe the k-means clustering algorithm. In particular, discuss the function it minimizes and its relation to a Gaussian model.

Timeline

Description automatically generated with medium confidence+ TP 4



# Explain what is the model for temporal data analysis and justify how to characterize and remove seasonality and noise. Discuss the linear prediction models.

Chart

Description automatically generated with low confidence+ TP 4



# What is an Auto-regressive prediction model? Explain how to apply it over given temporal data.

A picture containing timeline

Description automatically generated+ TP 4