2026

1st Quizz:

MLE/Gaussians + Bayes (10 min)

1. Recall the definition of a Bernoulli random variable of parameter $p \in [0, 1]$
2. Give the values of reals a, b, c and d so that the density of a Gaussian of mean $\mu \in \mathbb{R}$ and variance $\sigma^2 > 0$, equals $\frac{1}{\sqrt{2\pi\sigma^a}} \exp\left(c\frac{(x-\mu)^b}{\sigma^d}\right)$
3. What's the density of a mixture of $K \in \mathbb{N}$ Gaussians, with parameters $(\pi_k, \mu_k, \Sigma_k)_{k=1}^K$?
4. Assume you have realizations $x_1, ..., x_n$ of iid variables $X_1, ..., X_n$ following a Bernoulli distribution of parameter $\theta \in [0, 1]$. Give the formula for the maximum likelihood estimator of θ as a function of $x_1, ..., x_n$. Give the expression of the likelihood. Prove the formula for the MLE.

2nd Quizz:

Bayes/MAP/AE/VAE (10 min)

1. Recall Bayes formula and the law of total probability.
2. Your are given observations $x_1, ..., x_n \in \mathbb{R}^d$. Describe the structure (components, dimensions, training loss) of an autoencoder. How should the dimension of the code be with respect to the dimension of the data? Why?
3. What's a *denoising* autoencoder? What's its interest?
4. Recall the differences between an autoencoder and a variational autoencoder. How is a variational autoencoder trained in practice?

3rd Quizz: : GAN/CFM

1. Describe the principle of a GAN. What's its main drawback? What's "mode collapse"?

About Flow Matching:

2. What is the definition space of a velocity field, and its arrival space (from which space to which space does the function v go)?
3. Write down the loss of Flow Matching.
4. What is the conditional velocity field? What is the marginal target velocity field?
5. Explain how to sample new points from p_{data} after training

22/10/2026

GANs (10min). You want to train a generative adversarial network (GAN), to map a simple distribution p_0, to a more complex distribution p_{data}, i.e., you want to learn a generator G and a discriminator D.

- From which space to which space does the generator G go? what about the discriminator D?
- Using the fact that D does binary classification between real and fake samples, re-derive the GAN min-max optimization problem:

$$\min_G \max_D \mathcal{L}(G, D) = \min_G \max_D \mathbb{E}_{x \sim p_{\text{data}}} \log D(x) + \mathbb{E}_{z \sim p_0} \log(1 - D(G(z)))$$

- For a fixed generator G, rederive the "optimal" discriminator D_G (one can introduce the pushforward measure $p_G := G_{\#} p_0$, i.e., if $z \sim p_0$, then $G(z) \sim p_G$)

Diffusion (10min). In **diffusion** (DDPM, Denoising Diffusion Probabilistic Models)

- explain (and provide any relevant details) the forward (noising) process
- explain (and provide any relevant details) the backward process
- give the overall goal of learning the diffusion model
- give the actual training loss that is optimized