MACHINE LEARNING FUNDAMENSTALS - LS2026
HOMEWORK: MLE AND PLUGIN BAYES CLASSIFIER

Assignment (4 points) *Consider a multi-class classification problem with the following setup. The input features lie in $\mathcal{X} = \mathbb{R}^n$, the class label is from $\mathcal{Y} = \{0, 1, \ldots, Y-1\}$, and the loss function $\ell\colon \mathcal{Y} \times \mathcal{Y} \to \mathbb{R}_+$. You are provided with the training sample $T_m = ((\boldsymbol{x}_i, y_i) \in \mathcal{X} \times \mathcal{Y} \mid i = 1, \ldots, m)$. The task involves:*

a) Modeling the training data. *Fit the training data using mixture of multivariate normal distributions*

$$p(\boldsymbol{x}, y; \boldsymbol{\theta}) = \frac{p(y)}{(2\pi)^{n/2}|\boldsymbol{V}|^{1/2}} \exp\Big[-\frac{1}{2}(\boldsymbol{x} - \boldsymbol{\mu}_y)^T \boldsymbol{V}^{-1}(\boldsymbol{x} - \boldsymbol{\mu}_y)\Big] .$$

whose parameters $\boldsymbol{\theta}$ involve i) the prior probabilities $p(y)$, $y \in \mathcal{Y}$, ii) the mean vectors $\boldsymbol{\mu}_y \in \mathbb{R}^n$, $y \in \mathcal{Y}$, and iii) a common covariance matrix $\boldsymbol{V} \in \mathbb{R}^{n \times n}$. The parameters of the mixture model should be estimated using the Maximum-Likelihood estimation method.

b) Constructing a classifier. *Use the estimated model $p(\boldsymbol{x}, y; \boldsymbol{\theta}_m)$ to create a plug-in Bayes classifier:*

$$h(\boldsymbol{x}) \in \operatorname*{Argmin}_{\hat{y} \in \mathcal{Y}} \sum_{y \in \mathcal{Y}} p(\boldsymbol{x}, y; \boldsymbol{\theta}_m)\, \ell(y, \hat{y})$$

This classifier should not only provide predictions for the class labels but also quantify the uncertainty associated with these predictions by computing the conditional risk:

$$r(\boldsymbol{x}) = \sum_{y \in \mathcal{Y}} p(y \mid \boldsymbol{x}; \boldsymbol{\theta}_m)\, \ell(y, h(\boldsymbol{x})) .$$