Worksheet 8: Probabilistic Neural Networks for Weather Forecasting

Goal. This worksheet mirrors the lecture on probabilistic NNs and proper scoring rules. We show how a small NN can output a *mean* μ and *scale* σ to define a Normal predictive distribution, and how to compute suitable losses (Gaussian NLL and CRPS). We end with a simple Bernoulli example (rain/no rain) and the Brier/log loss.

Example 1: Tiny MLP producing μ and σ

Setup. Inputs are two past-day temperature anomalies $x = [0.5, -0.2]$. Hidden layer (2 units) with ReLU:

$$\mathbf{h} = \mathrm{ReLU}\Big(\underbrace{\begin{bmatrix} 0.8 & -0.4 \\ 0.3 & 0.6 \end{bmatrix}}_{\mathbf{W}_h} \mathbf{x} + \underbrace{\begin{bmatrix} 0.1 \\ -0.05 \end{bmatrix}}_{\mathbf{b}_h}\Big) = \mathrm{ReLU}\left(\begin{bmatrix} 0.58 \\ -0.02 \end{bmatrix}\right) = \begin{bmatrix} 0.58 \\ 0 \end{bmatrix}.$$

Parameter heads:

$$\mu = \underbrace{\begin{bmatrix} 0.5 & -0.2 \end{bmatrix}}_{\mathbf{w}_\mu} \mathbf{h} + b_\mu = 0.29, \qquad \rho = \underbrace{\begin{bmatrix} 0.1 & 0.3 \end{bmatrix}}_{\mathbf{w}_\sigma} \mathbf{h} + b_\sigma = -0.142,$$

$\sigma = \mathrm{softplus}(\rho) = \log(1 + e^\rho) = 0.624667.$

Losses with observation $y = 0.30$. **Gaussian NLL** (per-example):

$$\ell_{\mathrm{NLL}}(\mu, \sigma; y) = \tfrac{1}{2}\log(2\pi) + \log\sigma + \tfrac{1}{2}\left(\tfrac{y-\mu}{\sigma}\right)^2 \ = \ 0.448529.$$

Gaussian CRPS (closed form):

$$\mathrm{CRPS}(\mu, \sigma; y) = \sigma\Big[z(2\Phi(z) - 1) + 2\phi(z) - \tfrac{1}{\sqrt{\pi}}\Big], \quad z = \tfrac{y-\mu}{\sigma} \ = \ 0.146045.$$

Interpretation. The network predicts an entire Normal distribution $\mathcal{N}(\mu, \sigma^2)$. Minimising NLL or CRPS encourages calibrated, sharp forecasts.

Example 2: Tiny 1D CNN feature $\to (\mu, \sigma)$

Setup. Past 3-day sequence $x = [0.2, -0.1, 0.4]$. One conv filter $k = [0.5, 0.5]$ (stride 1) and ReLU:

$$f_1 = \mathrm{ReLU}(0.5 \cdot 0.2 + 0.5 \cdot (-0.1)) = 0.05, \qquad f_2 = \mathrm{ReLU}(0.5 \cdot (-0.1) + 0.5 \cdot 0.4) = 0.15.$$

Global average feature: $\bar{f} = (f_1 + f_2)/2 = 0.1$.
Heads: $\mu = 0.8\,\bar{f} + 0.05 = 0.13$, $\rho = 0.6\,\bar{f} - 0.1 = -0.04$, $\sigma = \mathrm{softplus}(\rho) = 0.673348$.

Losses with observation $y = 0.25$. $\text{NLL} = 0.539326, \quad \text{CRPS} = 0.165867.$

Example 3: Binary event (rain) with Bernoulli output

Setup. Hidden score $z = 0.7 \cdot \text{humidity} + (-0.3) \cdot \text{wind} + 0.2$ with humidity $= 0.4$, wind $= 0.1$. Probability $p = \sigma(z) = \frac{1}{1+e^{-z}} = 0.610639$. Observation $y = 1$ (it rained).

Losses. **Brier** $= (p - y)^2 = 0.151602$; **Log loss** $= -[y \log p + (1 - y) \log(1 - p)] = 0.493249$.