Deep Bayesian Neural Network: Mathematical Framework

Technical Documentation

January 14, 2025

1 Core Probability Functions

1.1 Multivariate Normal PDF

The foundation of the likelihood computation is the multivariate normal probability density function:

$$p(x|\mu, \Sigma) = \frac{1}{(2\pi)^{d/2}|\Sigma|^{1/2}} \exp\left(-\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu)\right)$$

1.2 Posterior Probability

For class c, the posterior probability is computed as:

$$P(c|x) = \frac{P(x|c)P(c)}{\sum_k P(x|k)P(k)}$$

2 Feature Processing

2.1 Covariance Matrix

For each feature group:

$$\Sigma = \frac{1}{n-1}\sum_{i=1}^{n}(x_i-\mu)(x_i-\mu)^T + \lambda I$$

where $\lambda = 10^{-6}$ is the stability term.

2.2 Feature Standardization

$$x_{normalized} = \frac{x - \mu}{\sigma + \epsilon}$$

where $\epsilon = 10^{-8}$ for numerical stability.

3 Adaptive Learning Process

3.1 Weight Update Mechanism

For each failed case:

$$w_{new} = w_{old}(1 + \alpha(1 - \frac{P_{true}}{P_{max\_other}}))$$

where:

- α is the learning rate
- P_{true} is the posterior probability of true class
- P_{max_other} is the maximum posterior among other classes

3.2 Sample Selection

For each class c, select:

$$\begin{cases} \arg\max_{x \in \text{Failed}_c} P(c|x) & \text{highest probability failure} \\ \arg\min_{x \in \text{Failed}_c} P(c|x) & \text{lowest probability failure} \end{cases}$$

4 Error Rate Computation

The classification error rate is computed as:

$$E = \frac{1}{N} \sum_{i=1}^{N} I(y_i \neq \hat{y}_i)$$

where I is the indicator function.

5 Processing Flow

The algorithm follows this sequence:

1. Feature pair generation: $C(n, 2)$ combinations
2. Likelihood computation for each feature pair
3. Posterior probability calculation
4. Weight updates for misclassified samples
5. Error rate computation and convergence check