Abteilung Maschinelles Lernen
Institut für Softwaretechnik und theoretische Informatik
Fakultät IV, Technische Universität Berlin
Prof. Dr. Klaus-Robert Müller
Email: klaus-robert.mueller@tu-berlin.de

Exercise Sheet 12

Exercise 1: Deep SVDD (20 P)

Consider a dataset $\boldsymbol{x}_1, \dots, \boldsymbol{x}_N \in \mathbb{R}^d$, and a simple linear feature map $\phi(\boldsymbol{x}) = \boldsymbol{w}^\top \boldsymbol{x} + b$ with trainable parameters $\boldsymbol{w}$ and b. For this simple scenario, we can formulate the deep SVDD problem as:

$$\min_{\boldsymbol{w},b} \quad \frac{1}{N} \sum_{i=1}^{N} \|\boldsymbol{w}^\top \boldsymbol{x}_i + b - 1\|^2$$

where we have hardcoded the center parameter of deep SVDD to 1. We then classify new points $\boldsymbol{x}$ to be anomalous if $\|\boldsymbol{w}^\top \boldsymbol{x} + b - 1\|^2 > \tau$.

(a) *Give* a choice of parameters $(\boldsymbol{w}, b)$ that minimizes the objective above for any dataset $(\boldsymbol{x}_1, \dots, \boldsymbol{x}_N)$.

(b) We now consider a regularizer for our feature map ϕ which simply consists of forcing the bias term to $b = 0$. *Show* that under this regularizer, the solution of deep SVDD is given by:

$$\boldsymbol{w} = \Sigma^{-1} \bar{\boldsymbol{x}}$$

where $\bar{\boldsymbol{x}}$ and Σ are the empirical mean and uncentered covariance.

Exercise 2: Restricted Boltzmann Machine (30 P)

The restricted Boltzmann machine is a system of binary variables comprising inputs $\boldsymbol{x} \in \{0, 1\}^d$ and hidden units $\boldsymbol{h} \in \{0, 1\}^K$. It associates to each configuration of these binary variables the energy:

$$E(\boldsymbol{x}, \boldsymbol{h}) = -\boldsymbol{x}^\top W \boldsymbol{h} - \boldsymbol{b}^\top \boldsymbol{h}$$

and the probability associated to each configuration is then given as:

$$p(\boldsymbol{x}, \boldsymbol{h}) = \frac{1}{Z} \exp(-E(\boldsymbol{x}, \boldsymbol{h}))$$

where Z is a normalization constant that makes probabilities sum to one. Let $\mathrm{sigm}(t) = \exp(t)/(1 + \exp(t))$ be the sigmoid function.

(a) *Show* that $p(h_k = 1 \,|\, \boldsymbol{x}) = \mathrm{sigm}\big(\boldsymbol{x}^\top W_{:,k} + b_k\big)$.

(b) Show that $p(x_j = 1 \,|\, \boldsymbol{h}) = \mathrm{sigm}\big(W_{j,:}^\top \boldsymbol{h}\big)$.

(c) *Show* that

$$p(\boldsymbol{x}) = \frac{1}{Z} \exp(-F(\boldsymbol{x}))$$

where

$$F(\boldsymbol{x}) = -\sum_{k=1}^{K} \log\big(1 + \exp\big(\boldsymbol{x}^\top W_{:,k} + b_k\big)\big)$$

is the free energy and where Z is again a normalization constant.

Exercise 3: Programming (50 P)

Download the programming files on ISIS and follow the instructions.

Exercise Sheet 12

1 Deep SVDD

a) Minimum at $\vec{w} = \vec{0}$, $b = 1$

b)
$$\mathcal{J} = \min_W \frac{1}{N} \sum_{i=1}^{N} \| w^T x_i - 1 \|^2 = \min_W \frac{1}{N} \sum_{i=1}^{N} (w^T x_i - 1)^T (w^T x_i - 1)$$

$$= \min_W \frac{1}{N} \sum_{i=1}^{N} w^T x_i x_i^T w - 2 w^T x_i + 1$$

$$= \min_W w^T \frac{\sum_{i=1}^{N} x_i x_i^T}{N} w - \frac{2}{N} \sum_{i=1}^{N} x_i^T w + \frac{1}{N} \sum_{i=1}^{N} 1$$

$$\frac{\partial \mathcal{J}}{\partial W} = (\Sigma + \Sigma^T) w - 2 \frac{1}{N} \sum_{i=1}^{N} x_i = 0$$

$$\Leftrightarrow 2 \Sigma w - 2 \bar{x} = 0 \quad \Leftrightarrow \quad w = \Sigma^{-1} \bar{x}$$

2 Restricted Boltzmann Machine

a)
$$p(h_k = 1 | x) = \frac{p(x, h_k = 1)}{p(x)} = \frac{\sum_{h_{-k}} p(x, h_k = 1, h_{-k})}{p(x)}$$

$$= \frac{\sum_{h_{-k}} p(x, h_k = 1, h_{-k})}{\sum_{q \in \{0,1\}} \sum_{h_{-k}} p(x, h_k = q, h_{-k})}$$

$$= \frac{\sum_{h_{-k}} \frac{1}{Z} \exp(x^T W_{:,k} + b_k - E(x, h_{-k}))}{\sum_{q \in \{0,1\}} \sum_{h_{-k}} \frac{1}{Z} \exp(x^T W_{:,k} q + b_k q - E(x, h_{-k}))}$$

$$= \frac{\exp(x^T W_{:,k} + b_k) \sum_{h_{-k}} \exp(-E(x, h_{-k}))}{\sum_{q \in \{0,1\}} \exp(x^T W_{:,k} q + b_k q) \sum_{h_{-k}} \exp(-E(x, h_{-k}))}$$

$$= \frac{\exp(x^T W_{:,k} + b_k)}{1 + \exp(x^T W_{:,k} + b_k)} = \mathrm{sigm}(x^T W_{:,k} + b_k)$$

b) $$p(x_j=1|h) = \frac{p(x_j=1,h)}{p(h)} = \frac{\sum_{x_{-j}} p(x_j=1, x_{-j}, h)}{\sum_{q\in\{0,1\}} \sum_{x_{-j}} p(x_j=q, x_{-j}, h)}$$

$$= \frac{\sum_{x_{-j}} \frac{1}{Z} \exp(W_{j,:}^T h - E(x_{-j}, h))}{\sum_{q\in\{0,1\}} \sum_{x_{-j}} \frac{1}{Z} \exp(W_{j,:}^T h q - E(x_{-j}, h))}$$

$$= \frac{\exp(W_{j,:}^T h)}{1+\exp(W_{j,:}^T h)} = \mathrm{sigm}(W_{j,:}^T h)$$

c) $$p(x) = \sum_h p(x,h) = \sum_h \frac{1}{Z} \exp(-E(x,h))$$

$$= \sum_h \frac{1}{Z} \exp(x^T W h - b^T h) = \sum_h \frac{1}{Z} \exp(\sum_k x^T W_{:,k} h_k - b_k^T h_k$$

$$= \frac{1}{Z} \sum_h \prod_k \exp(x^T W_{:,k} h_k - b_k^T h_k)$$

$$= \frac{1}{Z} \prod_k (1 + \exp(x^T W_{:,k} - b_k))$$

$$= \frac{1}{Z} \exp(\log(\prod_k (1+\exp(x^T W_{:,k} - b_k))))$$