Neural networks: Assignment 2

March 6, 2024

Please submit your assignments carefully written down in LaTeX or jupyter notebook, as applicable.

Problem 1: The name of softmax

Let $z \in \mathbb{R}^d$. Calculate $\lim_{C\to\infty} \operatorname{softmax}(Cz)$ and $\lim_{C\to-\infty} \operatorname{softmax}(Cz)$.

Problem 2: ReLU of a gaussian

Let $G \sim \mathcal{N}(0,1)$ be a gaussian random variable. Recall that for $0 \leq \alpha \leq 1$ the leaky ReLU activation is given as

$$\mathrm{ReLU}_\alpha(x) = \begin{cases} x & \text{if } x \geq 0, \\ \alpha x & \text{if } x < 0. \end{cases} .$$

Let $Y = \mathrm{ReLU}(G)$. Calculate $\mathbb{E}[Y]$ and $\mathrm{Var}[Y]$.

Problem 3: Power of linear neural networks

Let $d \geq 3$. Define the function $\chi : \{-1,1\}^d \to \{-1,1\}$ as

$$\chi(x_1, \ldots, x_d) = \prod_{i=1}^{d} x_i .$$

Let $n \geq 1, W \in \mathbb{R}^{n\times d}, b \in \mathbb{R}^n, v \in \mathbb{R}^n$. Define a linear neural network $\Phi(x) = v^T \cdot (Wx + b)$.

We say that Φ computes χ if $\Phi(x) = \chi(x)$ for every $x \in \{-1,1\}^d$. For every $d \geq 3$, either find n, W, b, v such that Φ computes χ or prove that the task is impossible.

Problem 4: Early stopping and variable learning rate

Take the code from `network_second.ipynb` file. Modify it so that implements:

- **Early stopping**, where the user does not have to specify the number of epochs in advance. For example, the network can stop if for 5 epochs in a row the validation data accuracy has not improved.
- **Variable learning rate**. For example, wait until the network does not improve according to the early stopping rule. Then divide the learning rate by 10. Repeat that three times.

You can implement the procedures as described or make your own version. Report if the network performance improved on the MNIST dataset.

Problem 5: Fashion MNIST

Fashion MNIST is a dataset for image recognition of items of clothing. It can be accessed directly in keras by `tf.keras.datasets.fashion_mnist.load_data()`.

Access the dataset in keras. Explore the data and train a convolutional model to solve it. What testing accuracy can you achieve?