Leaky ReLU Activation Function

$$f(x) = \begin{cases} 0.01x & \text{for x} < 0 \\ x & \text{for x} \geq 0 \end{cases}$$

Figure 29-10. *Leaky ReLU activation function*

Maxout

The Maxout activation function generalizes the ReLU and leaky ReLU functions and hence takes advantage of the efficiency of ReLU while avoiding its pitfalls of some neurons dying out. In any case, a trade-off needs to be made, because Maxout increases the parameter size of each neuron during training.

As a rule of thumb, different types of activation functions are not mixed in the same network. Also, ReLU is typically used for the hidden layers, and the softmax activation is used for classification problems at the output layer since this layer returns a probability of membership of a particular class.

This chapter provided an overview on how to train a predictive model using neural networks. This chapter ends Part 5 on introducing deep learning. The chapters in Part 6 will cover deep learning algorithms and their implementation with TensorFlow and Keras.

PART VI

Deep Learning in Practice