Drop-out

Intuitive idea

- Leave out neurons randomly (except for output layer)
  - For every training step different neurons are dropped with probability p (typically 10% - 50%)
    - Input value of dropped neuron is multiplied by 0
  - Dropped neurons change every training step
    - → drop-out trained neurons do not co-adapt to their neighbours
    - → drop-out trained neurons cannot rely on individual input neurons
  - Neurons and whole neural network are less sensitive to small input variations (more resilient)
  - During training, the weights of every input link are divided by (1-p) to compensate different number of active neurons during training in comparison to the resulting network

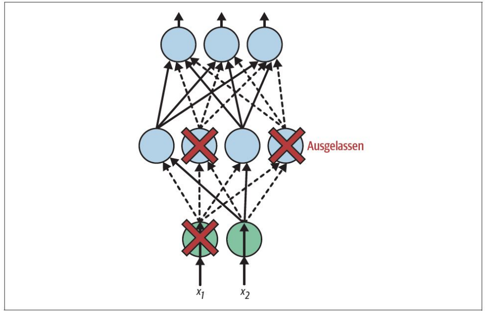
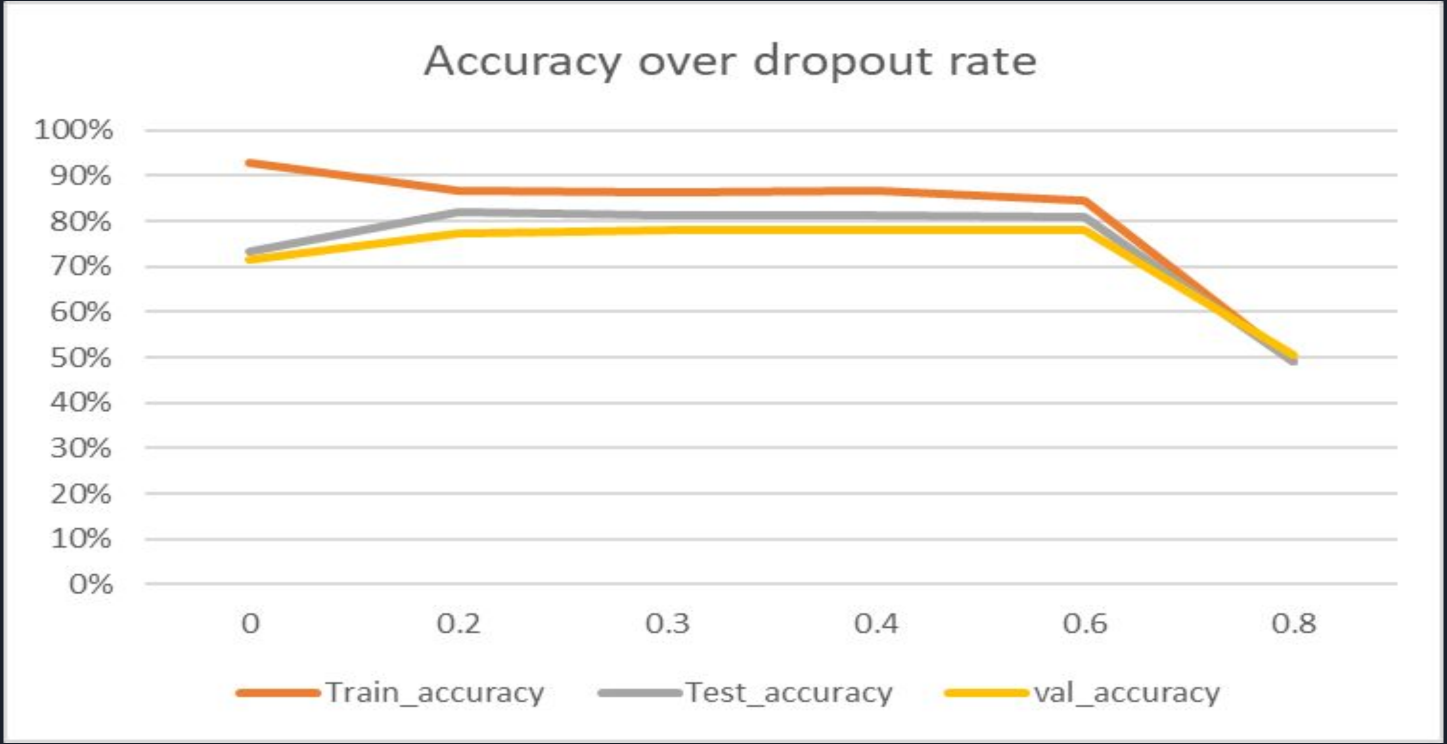


Abbildung 11-10: Bei der Drop-out-Regularisierung wird bei jeder Trainingsiteration eine zufällige Untermenge aller Neuronen in einer oder mehreren Schichten – außer der Ausgabeschicht – »gedroppt«; diese Neuronen geben in dieser Iteration 0 zurück (dargestellt durch die gestrichelten Pfeile).

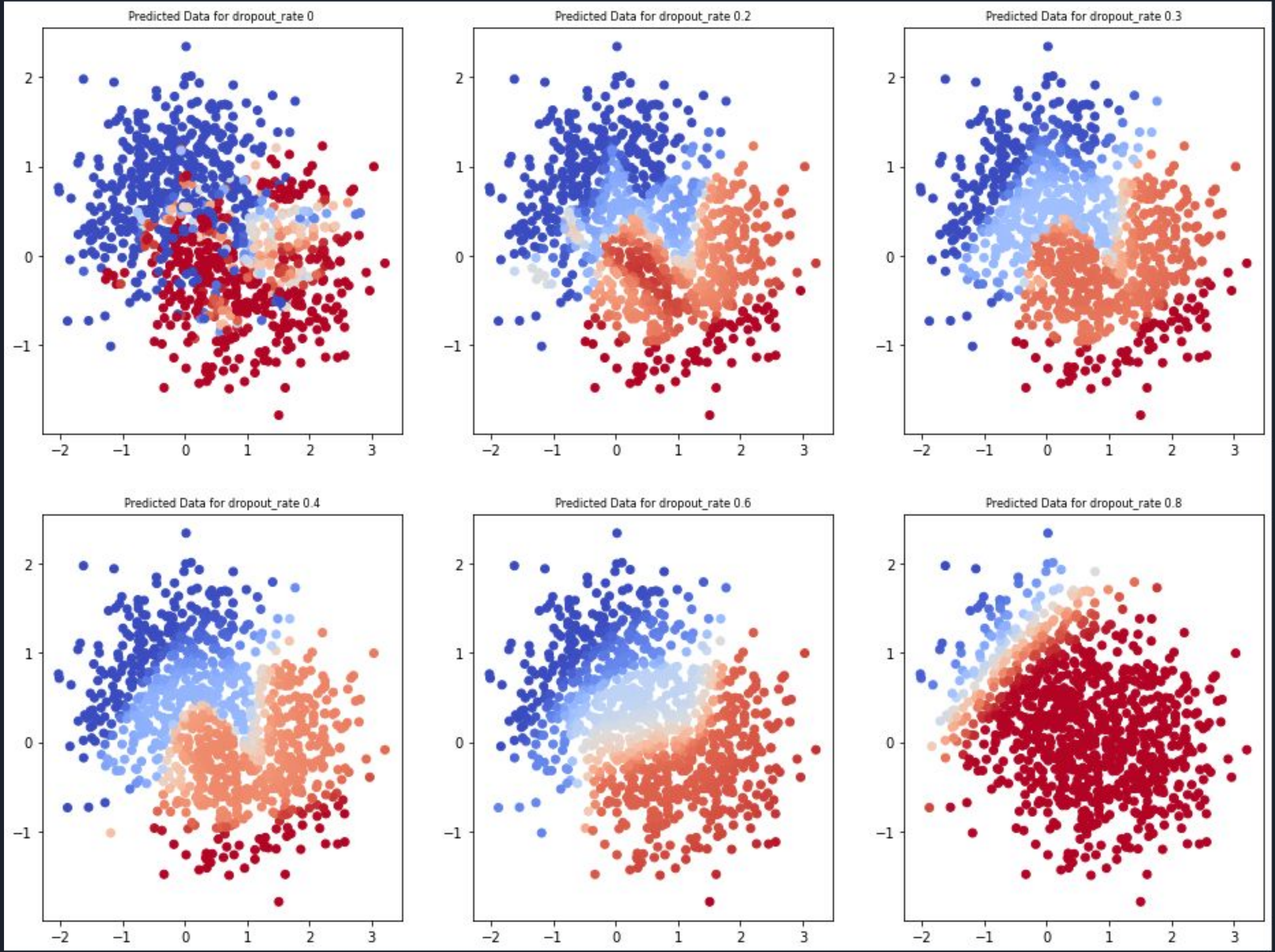
- Alternative view
  - In every training step an individual neural network is generated
    2^N possible networks (N=number of potential dropped neurons)
  - Resulting network can be viewed as ensemble of all the individual neural networks

Ergebnisse (1/2) - Von Overfitting zu Underfitting



Ergebnisse (2/2) - Differenzen in Train accuracy (durch batches?)

	Evaluation from learning curves		Evaluation by evaluate		Evaluation after optimisation
dropout rate	Train_accuracy	Test_accuracy	Train_accuracy	Test_accuracy	val_accuracy
0	93%	73,20%	93,00%	73,20%	71,60%
0,2	85,60%	82,00%	86,80%	82,00%	77,20%
0,3	86,00%	81,20%	86,20%	81,20%	78,00%
0,4	84,80%	81,20%	86,80%	81,20%	78,00%
0,6	82,20%	80,80%	84,60%	80,80%	78,00%
0,8	67,60%	51,20%	49,20%	49,20%	50,40%



Predicted Data for dropout_rate 0

Predicted Data for dropout_rate 0.2

Predicted Data for dropout_rate 0.3

Predicted Data for dropout_rate 0.4

Predicted Data for dropout_rate 0.6

Predicted Data for dropout_rate 0.8