7 Activation functions and neural networks

Select all statements which are true.

Select one or more alternatives

- [x] The ReLU activation usually works better than sigmoid activation function for hidden units because the sigmoid activations are sparse.
- [ ] Increasing the training set size generally does not hurt an algorithm's performance, and it may help significantly.
- [x] In logistic regression, the weights w should be initialized randomly rather than to all zeros, because if you initialize to all zeros, logistic regression will fail to learn a useful decision boundary as it will fail to 'break symmetry'.
- [x] The tanh activation usually works better than sigmoid activation function for hidden units because the mean of its output is closer to zero, and so it centers the data better for the next layer.
- [ ] Increasing the size of a neural network generally does not hurt an algorithm's performance, and it may help significantly.

Nullstill

Maks poeng: 9