References

Task1

[**https://stats.stackexchange.com/questions/181/how-to-choose-the-number-of-hidden-layers-and-nodes-in-a-feedforward-neural-netw**](https://stats.stackexchange.com/questions/181/how-to-choose-the-number-of-hidden-layers-and-nodes-in-a-feedforward-neural-netw)

**https://towardsdatascience.com/exploring-activation-functions-for-neural-networks-73498da59b02**

Task 2

https://datascience.stackexchange.com/questions/30676/role-derivative-of-sigmoid-function-in-neural-networks

<http://www.chioka.in/how-does-the-number-of-hidden-neurons-affect-a-neural-networks-performance/>

<http://proceedings.mlr.press/v9/glorot10a/glorot10a.pdf>

<https://www.quora.com/Why-is-initializing-weights-by-small-numbers-bad-in-neural-networks>

<https://www.quora.com/What-is-an-intuitive-explanation-of-the-Xavier-Initialization-for-Deep-Neural-Networks>

https://stats.stackexchange.com/questions/164876/tradeoff-batch-size-vs-number-of-iterations-to-train-a-neural-network