

PV021 Neural Networks

Exam manual

This manual specifies the knowledge demanded by the PV021 exam. Please, keep in mind that the knowledge described below is *mandatory even for the E grade*. Missing a single part automatically means F. You may repeat the exam as often as you wish (at the official exam dates); only the best grade goes into the information system.

- Slides 17 - 55: You need to know all definitions formally using the mathematical notation. You need to be able to explain and demonstrate the geometric interpretation of a neuron (a picture). Keep in mind that I may demand complete proofs (e.g., slide 44) or at least the basic idea (slides 47 and 48). You do not need to memorize the exact numbers from slide 53 but know the claims about computability.
- Slides 81 - 103 except slides 89 and 100. Everything in great detail. You need to be able to provide all mathematical details as well as an understanding of the fundamental notions such as maximum likelihood and all proofs.
- slides 105 - 155. All details of all observations and methods except:
 - Slide 110: Just know roughly what the theorem says.
 - Slide 118: You do not have to memorize all the schedules, just understand that there is a scheduling and be able to demonstrate one schedule.
 - Slide 134: You do not have to memorize all the ReLU variants, just know one.
 - Slides 136, 137: No need to memorize the exact assignment of initialization methods to the activation functions. Need to know that there is such an assignment.
 - Slides 144 and 145: The elephant is not needed.

Let me stress that apart from the above exceptions, you need to have detailed knowledge, including the mathematical formulae (e.g., for the momentum, AdaGrad, etc.), and intuitive understanding. In particular, you must know and understand how the normal LeCun initialization method (slides 127 and 128) is derived!

- Slides 179 - 186: Everything in great detail, including all the handwritten stuff. You may be asked to derive the backpropagation algorithm for CNN even though it did not explicitly appear in the lecture (but is similar to the derivation for MLP). It also helps to know the intuition for CNN from the applications (slides 163 - 177).

- Slides 200 – 227. Everything in detail. Especially all methods described in the slides 209 – 227 (gradient saliency maps, GradCAM, occlusion, LIME) with all mathematical details and intuitive understanding. You may skip the details of the slide 223 but *have to* know how the LIME works from the slide 222.
- Slides 229 - 240: All details, including all the handwritten stuff. You may be asked to derive the backpropagation algorithm for RNN even though it did not explicitly appear in the lecture (but it is similar to the derivation for MLP). LSTM is not mandatory this year.
- Slides 267 – 272. All mathematical details and intuition. In particular, you should know and understand how the self-attention layer works. You may be asked to derive the backpropagation algorithm for the self-attention layer (possibly with $\alpha_{ij} = e_{ij}$ to simplify the derivation).