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 - 54: 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. Keep in mind that I may demand complete proofs (e.g., slide 44) or at least the basic idea (slide 47 and the content of the green board). You do not need to memorize the exact numbers from slide 52 but know the claims about computability.

- Slides 77 - 110 except slide 103. 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.

  In particular, remember that demonstrating the example of log-likelihood from slides 107 and 108 is not enough! You need to know and understand the full version from slide 109.

- slides 112 - 157. All details of all observations and methods except:

  - Slide 117: Just know roughly what the theorem says.
  - Slide 125: You do not have to memorize all the schedules, just understand that there is a scheduling and be able to demonstrate one schedule.
  - Slide 132: You do not need to memorize 1.7159 but have to know why the number is this.
  - Slide 143: You do not have to memorize all the ReLU variants, just know one.
  - Slides 144, 145: No need to memorize the exact assignment of initialization methods to the activation functions. Need to know that there is such an assignment.

  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 (slide 137 and the green board) is derived!

- Slides 198 - 210: 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 160 - 196).

- Slides 222 – 247. Everything in detail. Especially all methods described in the slides 229 – 247 (gradient saliency maps, GradCAM, occlusion, LIME) with all mathematical details and intuitive understanding.

- Slides 248 - 263: 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). You have to know how LSTM is defined (formally and intuitively).

- Slides 274 – 278. All mathematical details and intuition. In particular, you should know and understand how the self-attention layer works.