# Advanced Deep Learning

| | |
|---|---|
| **Lecturer** | *Paulius Rauba* |
| **Study form** | *Online* |
| **Course prerequisites** | *Statistics, linear algebra, calculus, Python* |
| **Language of instruction** | *English* |

**Introduction**

This Advanced Deep Learning course is designed as a 20-week course to get you acquainted, familiar, and comfortable with advanced methods in deep learning that are used in modern machine learning research. The primary goal is to take your existing knowledge in mathematics, statistics, and programming, and build up the skills required for advanced deep learning research. The class heavily relies on the newly released book *Deep Learning* (Bishop & Bishop, 2023) which is the new state-of-the-art account of deep learning.

This class contains two components each week: one lecture (1.5 hours) and one seminar (about 1 hours). The lecture will be covering the material from the book together with visual examples, mathematical derivations, linking intuition to theory to practice. The seminars will be mostly based on Jupyter notebooks and will involve building an improved understanding of the technical material presented before.

Each week, you can submit questions about specific parts which you do not feel comfortable with. With this, I will adjust the lecture to tailor it to your needs. You will also receive required pre-reading before each week that will help you get comfortable with the required knowledge for each course.

**Target audience**

The primary target audience for this course is those wishing to pursue higher-level studies in machine learning/computer science (MSc/PhD level) at universities, existing researchers wishing to upskill themselves, and existing machine learning practitioners wishing to transition to ML research.

**Outcomes**

By the end of this course, you will have:

- Enough expertise to kick off your own research in machine learning, including state-of-the-art projects with diffusion models, flow-based models, and others.
- Feel comfortable reading technical machine learning reports in multiple technical areas.
- Be able to explain and understand the primary innovations in machine learning with significant technical expertise.
- Be able to guide technical decisions in any machine learning-driven project by generating hypotheses about what could work and what is not likely to work.

**Course structure by each week.**

- **Week 1. Introduction. Probabilities.** Gaussian parametrization, likelihood functions, bias of likelihood function, information theory (entropy, differential entropy, maximum entropy, KL divergence, mutual information), Bayesian probabilities (posterior, Bayesian inference).
- **Week 2. Standard distributions.** Discrete and multivariate gaussian distributions. Properties of multivariate Gaussians (moments, marginal distributions, sequential estimation, mixture of Gaussians). Von Mises distributions, the exponential family and sufficient statistics, nonparametric methods for distributions.
- **Week 3. Single-layer networks for regression**. Maximum likelihood of regression models, sequential learning, basis functions, decision theory, bias-variance trade-off and its theoretical decomposition. Mathematical formalization.
- **Week 4. Single-layer networks for classification**. Perceptrons, decision theory, generative classifiers, discriminative classifiers, probit regression, canonical link functions.
- **Week 5. Deep neural networks.** Limitations of fixed basis functions, multilayer networks, universal approximation theorem, weight-space symmetries, hierarchical and distributed representations, transfer learning, contrastive learning, general network architectures, tensors, mixture density networks.

- **Week 6. Gradient descent.** Error surfaces, local quadratic approximation, optimization theory, convergence (momentum, learning rate schedule, RMSProp and Adam), batch and layer normalization.
- **Week 7. Backpropagation.** Evaluation of gradients, numerical differentiation, the Jacobian matrix, the Hessian matrix. Automatic differentiation with forward- and reverse-mode differentiation.
- **Week 8. Regularization.** Inductive biases, no free lunch theorem, weight decay, learning curves (early stopping, double descent), parameter sharing, soft weight sharing, residual connections, model averaging and dropout.
- **Week 9. Convolutional networks.** Translational equivariance, padding, multi-dimensional convolutions, pooling, saliency maps, non-max suppression, conv. Segmentation, U-net, style transfer, conv nets beyond images.
- **Week 10. Structured distributions.** Graphical models, directed graphs, factorization of graphs, conditional independence of graphical models (D-Separation, Markov blankets, graphs as filters), introduction to sequence models.
- **Week 11. Transformers.** Attention mechanism, multi-head attention, encoder, decoder transformers, sequence-to-sequence transformers.
- **Week 12. Graph neural networks.** Graph properties, adjacency matrices, permutation equivariance, neural message passing, general graph networks and geometric deep learning.
- **Week 13. Sampling.** Properties of sampling algorithms, basic sampling algorithms, Markov chain monte carlo, Langevin sampling (energy-based models, Langevin dynamics).
- **Week 14. Discrete latent variables**. Mixture of Gaussians, EM Algorithm, Evidence Lower Bound and key properties of variational inference.
- **Week 15. Continuous latent variables.** Principal component analysis, probabilistic latent variables, ELBO for PCA, nonlinear latent variable models.
- **Week 16. Generative adversarial networks**. Adversarial training, loss functions, inductive biases.
- **Week 17. Normalizing flows.** Coupling flows, autoregressive flows, continuous flows, core idea of flow matching.
- **Week 18. Autoencoders.** Deterministic autoencoders, variational autoencoders, amortized inference, reparametrization trick.
- **Week 19. Diffusion models.** Properties of diffusion models, denoising, forward encoder, reverse decoder, score matching, guided diffusion, connection to variational inference.
- **Week 20. Summary and recap.**

**Literature**

Our primary reference will be Bishop & Bishop (2023) on Deep Learning. Other material will be included each week and will be provided on a weekly basis.

**References**

Bishop, C. M., & Bishop, H. (2023). *Deep learning: Foundations and concepts*. Springer Nature.