

# Seminar/project topics

proposed topics are grouped in five areas:

1. understanding deep neural networks
2. overview on other ML techniques
3. natural language processing
4. advanced computer vision
5. machine learning in physics
6. datasets and pretrained models

# Seminars: understanding deep neural networks

## 1. Techniques that allow understand more on how and why deep architectures work

- <https://distill.pub/2020/circuits/early-vision/>
- <https://ai.googleblog.com/2015/06/inceptionism-going-deeper-into-neural.html>
- <https://distill.pub/2017/feature-visualization/>
- [D. Erhan, Y. Bengio, A. Courville, P. Vincent. , Visualizing higher-layer features of a deep network \(2009\).](#)
- <http://arxiv.org/pdf/1506.02753.pdf>
- <https://arxiv.org/pdf/1312.6034v2.pdf>
- <https://github.com/tensorflow/lucid>

## 2. Fooling deep neural networks: deep learning models are highly sensitive to carefully prepared adversarial attacks

How to generate adversarial examples? Can we use them to improve network stability?

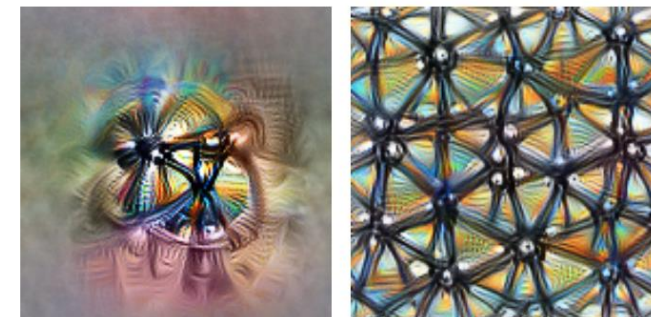
- Szegedy, Christian, et al. “Intriguing properties of neural networks” *arXiv:1312.6199* (2013).
- Goodfellow, Ian J., Jonathon Shlens, and Christian Szegedy. “Explaining and harnessing adversarial examples” *arXiv:1412.6572* (2014).
- Papernot, Nicolas, et al. “The limitations of deep learning in adversarial settings.” *2016 IEEE European symposium on security and privacy (EuroS&P)*. IEEE (2016).
- Su, Jiawei, Danilo V. Vargas, and Kouichi Sakurai. “One pixel attack for fooling deep neural networks.” *IEEE Transactions on Evolutionary Computation* 23.5 828 (2019).

## 3. Adversarial Attacks on LLMs: <https://lilianweng.github.io/posts/2023-10-25-adv-attack-llm/>

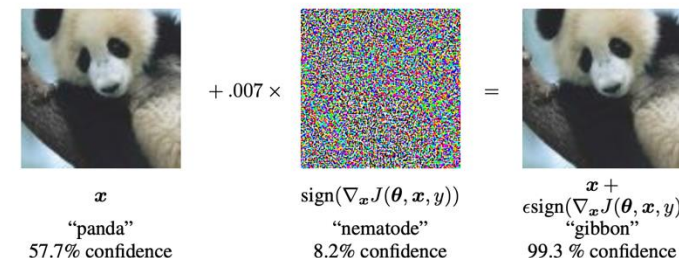
## 4. Defense strategies against adversarial attacks

Overview on various techniques for defending against adversarial examples for attacking deep neural networks.

- Szegedy, Christian, et al. “Intriguing properties of neural networks” *arXiv:1312.6199* (2013).
- Xie, Cihang, et al. “Mitigating adversarial effects through randomization.” *arXiv:1711.01991* (2017).
- Das, Nilaksh, et al. “Keeping the bad guys out: Protecting and vaccinating deep learning with jpeg compression.” *arXiv:1705.02900* (2017).
- Xie, Cihang, et al. “Feature denoising for improving adversarial robustness.” *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (2019).
- Goodfellow, Ian J., Jonathon Shlens, and Christian Szegedy. “Explaining and harnessing adversarial examples” *arXiv:1412.6572* (2014).



**Feature visualization** answers questions about what a network—or parts of a network—are looking for by generating examples.



# Seminars: overview on other ML techniques

## 1. Reinforcement Learning

Basics and applications

- <https://arxiv.org/pdf/cs/9605103.pdf>
- <https://mpatacchiola.github.io/blog/2016/12/09/dissecting-reinforcement-learning.html>
- <https://deepsense.ai/what-is-reinforcement-learning-the-complete-guide/>

## 2. Hidden Markov model

Definition and example applications

- <https://jonathan-hui.medium.com/machine-learning-hidden-markov-model-hmm-31660d217a61>
- <http://www.cs.sjsu.edu/~stamp/RUA/HMM.pdf>

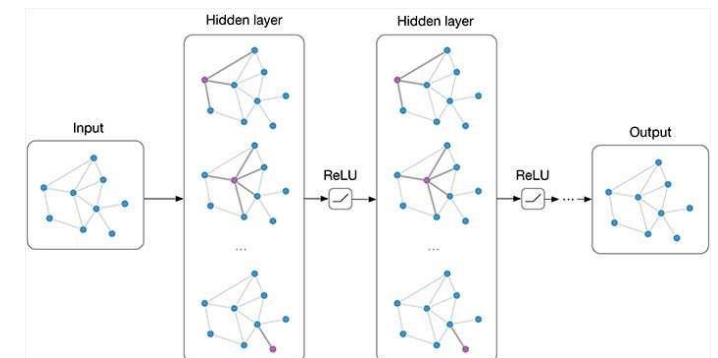
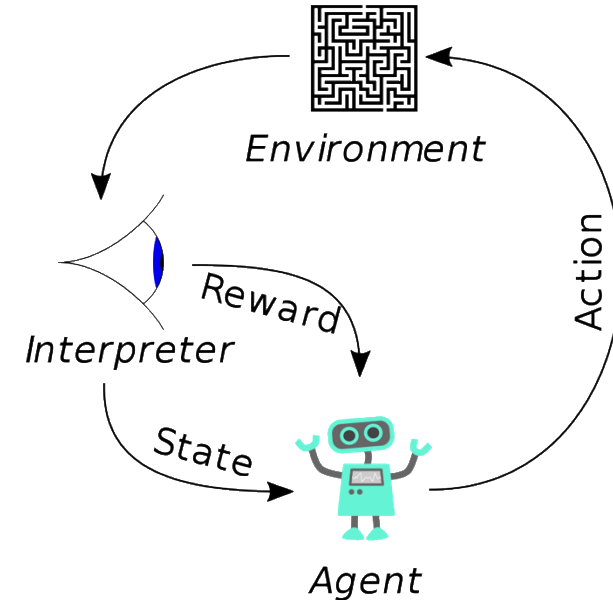
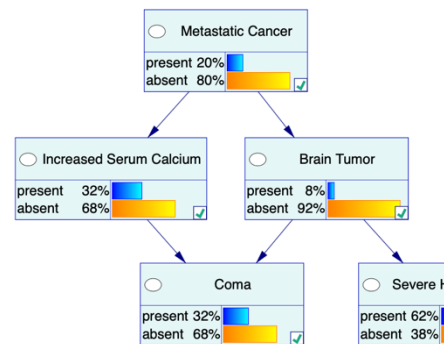
## 3. Recommender systems

Idea, applications and the Netflix Price

- <https://arxiv.org/pdf/1203.4487.pdf>
- <https://towardsdatascience.com/introduction-to-recommender-systems-6c66cf15ada>

## 4. Bayesian networks

- <https://repo.bayesfusion.com/bayesbox.html>
- <http://www.eng.tau.ac.il/~bengal/BN.pdf>
- <http://www.niedermayer.ca/node/35>



# Seminars: natural language processing, NLP

## 1. Optical character recognition (OCR)

- problem definition
- classical or neural network approach: an overview
- presenting one particular engine, e.g. state-of-the-art *Tesseract* (what deep learning model is used inside?)
- <https://tesseract-ocr.github.io/tessdoc/>
- <https://huggingface.co/spaces?category=ocr>

## 2. Automatic speech recognition (ASR)

- problem definition
- possible applications
- databases (e.g. *Librispeech*)
- an overview on single specific algorithm, e.g. *ContextNet* or *Deep Speech*
- <https://arxiv.org/pdf/2005.03191.pdf>
- <https://arxiv.org/pdf/1512.02595.pdf>

## 3. Machine Translation

- problem definition, example methods
- <https://huggingface.co/spaces?category=language-translation>

## 4. Question Answering

- problem definition, example methods
- <https://huggingface.co/spaces?category=question-answering>

# Seminars: computer vision

## 1. Point Feature Matching

- problem definition
- classical and deep learning approaches
- *SuperGlue* algorithm: <https://arxiv.org/abs/1911.11763>

## 2. Pose estimation

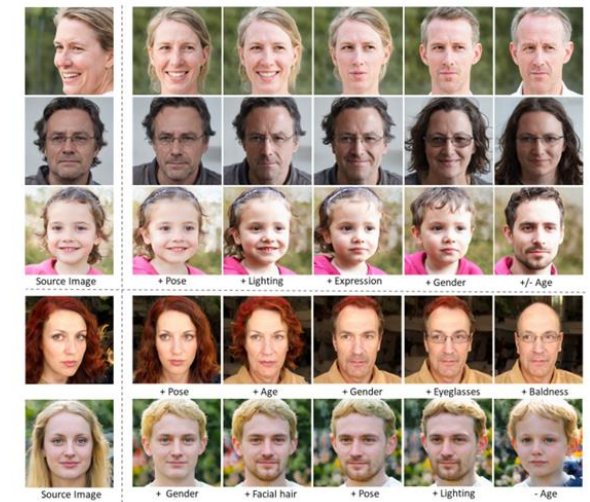
- problem definition and deep-learning solution, e.g.:
- <https://arxiv.org/pdf/1803.08225.pdf>

## 3. Super-Resolution

- idea, applications, algorithms
- <https://arxiv.org/abs/1809.00219>

## 4. Editing GANs latent space

- <https://medium.com/codex/how-to-edit-images-with-gans-controlling-the-latent-space-of-gans-afde630e53d1>
- <https://www.unite.ai/editing-a-gans-latent-space-with-blobs>



# Seminars: machine learning in physics

## 1. Physics-informed neural networks

- <https://benmoseley.blog/my-research/so-what-is-a-physics-informed-neural-network>
- <https://maziarraissi.github.io/PINNs>
- seminal paper: <https://www.sciencedirect.com/science/article/abs/pii/S0021999118307125>

## 2. Group-equivariant neural networks

- original paper: <https://proceedings.mlr.press/v48/cohenc16.pdf>
- <https://uvagedl.github.io>
- any  $O(n)$ ,  $E(n)$ -equivariant network: <https://github.com/DavidRuhe/clifford-group-equivariant-neural-networks>

## 3. Neural network quantum states

- <https://www.science.org/doi/10.1126/science.aag2302>

## 4. Classical shadows

- <https://www.science.org/doi/10.1126/science.abk3333>
- [https://pennylane.ai/qml/demos/tutorial\\_classical\\_shadows](https://pennylane.ai/qml/demos/tutorial_classical_shadows)

## 5. Neural tangent kernels

- <https://arxiv.org/abs/1806.07572>
- <https://lilianweng.github.io/posts/2022-09-08-ntk/>

# Seminars: machine learning in physics (cont.)

## 6. Graph Convolutional Neural Networks for quantum chemistry and prediction of material properties

- seminal paper <https://journals.aps.org/prl/abstract/10.1103/PhysRevLett.120.145301>
- SchNet: <https://github.com/atomistic-machine-learning/SchNet>

## 7. GNNs for materials discovery

- <https://www.nature.com/articles/s41586-023-06735-9>

## 8. Big data for materials discovery

- <https://next-gen.materialsproject.org/>
- detection of Topological Materials with ML: <https://www.topologicalquantumchemistry.com/mltqc/>
- hydrogen storage: <https://advanced.onlinelibrary.wiley.com/doi/full/10.1002/adma.202413430>
- topological invariants learning: <https://www.nature.com/articles/s41524-019-0224-x>

## 9. GAN networks for metamaterials design:

- <https://www.science.org/doi/10.1126/sciadv.aaz4169>
- <https://www.nature.com/articles/s41524-023-01036-1>

## 10. AlphaTensor: discovering novel algorithms

- <https://www.nature.com/articles/s41586-022-05172-4>
- <https://deepmind.google/discover/blog/discovering-novel-algorithms-with-alphatensor/>

## 11. Quantum Convolutional Neural Networks: <https://arxiv.org/abs/1810.03787>

## 12. Need more advanced topics? <https://lilianweng.github.io/>

- inspiring blog with in-depth analysis of various aspects of deep learning

# Projects: interesting datasets

1. <https://huggingface.co/>

- popular web service with interesting datasets and ready-to-use models (SOTA in many cases)

2. <https://www.kaggle.com/>

- machine-learning competitions supported by plenty of datasets,
- e.g. butterfly classification <https://www.kaggle.com/datasets/phucthaiv02/butterfly-image-classification>

3. <https://paperswithcode.com/>

- ML scientific papers supported by models + datasets

4. Classical datasets, [https://en.wikipedia.org/wiki/List\\_of\\_datasets\\_for\\_machine-learning\\_research](https://en.wikipedia.org/wiki/List_of_datasets_for_machine-learning_research)

- MNIST, Fashion-MNIST
- Iris
- CIPHAR
- ImageNet
- other somewhat older: <https://github.com/jbrownlee/Datasets>

5. **Need more advanced topics?** <https://lilianweng.github.io/>

- inspiring blog with in-depth analysis of various aspects of deep learning