# План курса лекций «компьютерное зрение»

## Лекция 0. Python

1. Вводная беседа
2. Основы python

Источник: https://docs.python.org/3/tutorial/

## Лекция 1. Анализ табличных данных

1. skitit-learn
2. xgbboost
3. Сравнение линейной регрессии и xgbboost на конкретном примере обработки данных

Источники:

<https://github.com/dmlc/xgboost/tree/master/demo#machine-learning-challenge-winning-solutions>

<https://machinelearningmastery.com/gentle-introduction-xgboost-applied-machine-learning/>

<https://habr.com/en/company/ods/blog/327250/>

<https://dl.acm.org/doi/pdf/10.1145/2939672.2939785?download=true>

## Лекция 2. Свёрточные нейронные сети и классификация изображений

1. Вводная часть про обучение нейронных сетей, какие проблемы приходится решать
2. MNIST и LeNet
3. Задача ImageNet

Источники:

<https://arxiv.org/pdf/1609.04747.pdf>

<https://www.eecis.udel.edu/~shatkay/Course/papers/NetworksAndCNNClasifiersIntroVapnik95.pdf>

<https://arxiv.org/pdf/1502.03167.pdf>

<http://www.vlfeat.org/matconvnet/matconvnet-manual.pdf>

<http://www.image-net.org>

Николенко и др., Глубокое обучение

Goodfellow

## Лекция 3. Нейросетевые детекторы положения объектов на изображении

1. Region proposals via selective search R-CNN
2. Fast R-CNN
3. Faster R-CNN
4. YOLO, SSD

<https://towardsdatascience.com/r-cnn-fast-r-cnn-faster-r-cnn-yolo-object-detection-algorithms-36d53571365e>

<http://openaccess.thecvf.com/content_iccv_2015/papers/Girshick_Fast_R-CNN_ICCV_2015_paper.pdf>

<http://papers.nips.cc/paper/5638-faster-r-cnn-towards-real-time-object-detection-with-region-proposal-networks>

<http://openaccess.thecvf.com/content_cvpr_2017/html/Redmon_YOLO9000_Better_Faster_CVPR_2017_paper.html>

<https://arxiv.org/pdf/1512.02325.pdf>

## Лекция 4. Нейросетевые методы поиска особых точек OpenPose

1. [Shotton, Jamie, Ross Girshick, Andrew Fitzgibbon, Toby Sharp, Mat Cook, Mark Finocchio, Richard Moore et al. "Efficient human pose estimation from single depth images." IEEE transactions on pattern analysis and machine intelligence 35, no. 12 (2012): 2821-2840.](https://www.microsoft.com/en-us/research/wp-content/uploads/2016/02/main-39.pdf)
2. [Tompson, Jonathan, Ross Goroshin, Arjun Jain, Yann LeCun, and Christoph Bregler. "Efficient object localization using convolutional networks." In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 648-656. 2015.](https://www.cv-foundation.org/openaccess/content_cvpr_2015/papers/Tompson_Efficient_Object_Localization_2015_CVPR_paper.pdf)
3. [Ramakrishna, Varun, Daniel Munoz, Martial Hebert, James Andrew Bagnell, and Yaser Sheikh. "Pose machines: Articulated pose estimation via inference machines." In European Conference on Computer Vision, pp. 33-47. Springer, Cham, 2014.](https://kilthub.cmu.edu/articles/Pose_Machines_Articulated_Pose_Estimation_via_Inference_Machines/6558671/files/12040949.pdf)
4. [Cao, Zhe, Gines Hidalgo, Tomas Simon, Shih-En Wei, and Yaser Sheikh. "OpenPose: realtime multi-person 2D pose estimation using Part Affinity Fields." arXiv preprint arXiv:1812.08008 (2018).](https://arxiv.org/pdf/1812.08008.pdf)
5. [Sun, Ke, Bin Xiao, Dong Liu, and Jingdong Wang. "Deep high-resolution representation learning for human pose estimation." In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 5693-5703. 2019.](https://github.com/microsoft/human-pose-estimation.pytorch)

## Лекция 5. GANs

1. [Gui, Jie, Zhenan Sun, Yonggang Wen, Dacheng Tao, and Jieping Ye. "A Review on Generative Adversarial Networks: Algorithms, Theory, and Applications." arXiv preprint arXiv:2001.06937 (2020).](https://arxiv.org/pdf/2001.06937.pdf)
2. [Kingma, Diederik P., and Max Welling. "Auto-encoding variational bayes." arXiv preprint arXiv:1312.6114 (2013).](https://arxiv.org/pdf/1312.6114.pdf)
3. [Pu, Yunchen, Zhe Gan, Ricardo Henao, Xin Yuan, Chunyuan Li, Andrew Stevens, and Lawrence Carin. "Variational autoencoder for deep learning of images, labels and captions." In Advances in neural information processing systems, pp. 2352-2360. 2016.](http://papers.nips.cc/paper/6528-variational-autoencoder-for-deep-learning-of-images-labels-and-captions.pdf)
4. [Makhzani, Alireza, Jonathon Shlens, Navdeep Jaitly, Ian Goodfellow, and Brendan Frey. "Adversarial autoencoders." arXiv preprint arXiv:1511.05644 (2015).](https://arxiv.org/pdf/1511.05644.pdf)
5. [Goodfellow, Ian, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. "Generative adversarial nets." In Advances in neural information processing systems, pp. 2672-2680. 2014.](http://papers.nips.cc/paper/5423-generative-adversarial-nets.pdf)
6. [Chen, Xi, Yan Duan, Rein Houthooft, John Schulman, Ilya Sutskever, and Pieter Abbeel. "Infogan: Interpretable representation learning by information maximizing generative adversarial nets." In Advances in neural information processing systems, pp. 2172-2180. 2016.](http://papers.nips.cc/paper/6399-infogan-interpretable-representation-learning-by-information-maximizing-generative-adversarial-nets.pdf)
7. [Reed, Scott E., Zeynep Akata, Santosh Mohan, Samuel Tenka, Bernt Schiele, and Honglak Lee. "Learning what and where to draw." In Advances in neural information processing systems, pp. 217-225. 2016.](https://arxiv.org/pdf/1610.02454.pdf)
8. [Isola, Phillip, Jun-Yan Zhu, Tinghui Zhou, and Alexei A. Efros. "Image-to-image translation with conditional adversarial networks." In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 1125-1134. 2017.](https://arxiv.org/pdf/1611.07004.pdf)
9. [Karras, Tero, Samuli Laine, and Timo Aila. "A style-based generator architecture for generative adversarial networks." In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 4401-4410. 2019.](http://openaccess.thecvf.com/content_CVPR_2019/papers/Karras_A_Style-Based_Generator_Architecture_for_Generative_Adversarial_Networks_CVPR_2019_paper.pdf)

## Лекция 6. Способы подготовки данных для обучения нейронных сетей

1. Доверительный интервал для оценки достоверности классификации
2. Оценки объёмов тестирующих выборок
3. Источники данных
4. Платформы mturk, toloka
5. Симуляционные данные
6. Трюки при обучении (pseudo-labeling, аугментация)

Источники:

<https://ru.coursera.org/lecture/stats-for-data-analysis/dovieritiel-nyie-intiervaly-s-pomoshch-iu-kvantiliei-yboDc>

<https://sebastianraschka.com/blog/2018/model-evaluation-selection-part4.html>

<https://www.mturk.com>

<https://toloka.yandex.ru/tasks>

<https://github.com/immersive-limit/Unity-ComputerVisionSim>

## Лекция 7. Методы ускорения нейросетевых вычислений

1. [Пример кода с использованием SIMD-инструкций](https://gist.github.com/vermorel/7ad35212df44f3a79bca8ab5fe8e7622#file-fast_convolve_1d-cpp-L123)
2. [Библиотека Openvino](https://software.intel.com/en-us/openvino-toolkit/choose-download)
3. [Howard, Andrew G., Menglong Zhu, Bo Chen, Dmitry Kalenichenko, Weijun Wang, Tobias Weyand, Marco Andreetto, and Hartwig Adam. "Mobilenets: Efficient convolutional neural networks for mobile vision applications." arXiv preprint arXiv:1704.04861 (2017).](https://arxiv.org/pdf/1704.04861.pdf)
4. [Sandler, Mark, Andrew Howard, Menglong Zhu, Andrey Zhmoginov, and Liang-Chieh Chen. "Mobilenetv2: Inverted residuals and linear bottlenecks." In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 4510-4520. 2018.](https://arxiv.org/pdf/1801.04381.pdf)
5. [Courbariaux, Matthieu, Itay Hubara, Daniel Soudry, Ran El-Yaniv, and Yoshua Bengio. "Binarized neural networks: Training deep neural networks with weights and activations constrained to+ 1 or-1." arXiv preprint arXiv:1602.02830(2016).](https://arxiv.org/pdf/1602.02830)
6. [Rastegari, Mohammad, Vicente Ordonez, Joseph Redmon, and Ali Farhadi. "Xnor-net: Imagenet classification using binary convolutional neural networks." In European conference on computer vision, pp. 525-542. Springer, Cham, 2016.](https://arxiv.org/pdf/1603.05279.pdf)
7. [BMXNet: An Open-Source Binary Neural Network Implementation Based on MXNet](https://github.com/hpi-xnor/BMXNet)

## Лекция 8. Классические методы компьютерного зрения: вычитание фона

## Лекция 9. Классические методы компьютерного зрения: вычисление точек особенностей. Усиление метода нейронными сетями

## Лекция 10. Обобщённые дескрипторы изображений, triplet loss.

## Лекция 11. Реккурентные нейронные сети в компьютерном зрении. GRU, LSTM, visual question answering

## Лекция 12. Обучение с подкреплением