Далее приведен список тем, обсуждаемых на занятиях школы, а также набор наиболее полезных ссылок, содержащих материалы по темам.

**Общие материалы:**

Лекции К.В. Воронцова по машинному обучению

<http://www.machinelearning.ru/wiki/index.php?title=%D0%9C%D0%B0%D1%88%D0%B8%D0%BD%D0%BD%D0%BE%D0%B5_%D0%BE%D0%B1%D1%83%D1%87%D0%B5%D0%BD%D0%B8%D0%B5_%28%D0%BA%D1%83%D1%80%D1%81_%D0%BB%D0%B5%D0%BA%D1%86%D0%B8%D0%B9%2C_%D0%9A.%D0%92.%D0%92%D0%BE%D1%80%D0%BE%D0%BD%D1%86%D0%BE%D0%B2%29>

pdf:<http://www.machinelearning.ru/wiki/images/6/6d/Voron-ML-1.pdf>

видолекции курса: <https://yandexdataschool.ru/edu-process/courses/machine-learning>

вводный курс на курсере: https://www.coursera.org/learn/vvedenie-mashinnoe-obuchenie

Книга 'The Elements of Statistical Learning'

http://statweb.stanford.edu/~tibs/ElemStatLearn/

Книга ‘Pattern Recognition and Machine Learning’

<https://www.microsoft.com/en-us/research/people/cmbishop/#prml-book>

Проект Metacademy (<https://www.metacademy.org/>) - представление знаний в области машинного обучения в виде графа зависимостей концептов. Каждый из концептов содержит краткое описание и ссылки на литературу по теме.

Например, svm: <https://www.metacademy.org/graphs/concepts/support_vector_machine#focus=4wuz9t3u&mode=explore>

Иллюстрация общего конвейера решения задачи обучения с учителем

<https://raw.githubusercontent.com/rasbt/pattern_classification/master/Images/supervised_learning_flowchart.png>

Выбор модели

<http://scikit-learn.org/stable/tutorial/machine_learning_map/>

<https://hackbrightacademy.com/content/uploads/2014/05/ml_algorithms.png>

<http://dlib.net/ml_guide.svg>

<https://www.youtube.com/watch?v=BOxC1_xHP9A>

Регрессия

<https://ru.wikipedia.org/wiki/%D0%9B%D0%B8%D0%BD%D0%B5%D0%B9%D0%BD%D0%B0%D1%8F_%D1%80%D0%B5%D0%B3%D1%80%D0%B5%D1%81%D1%81%D0%B8%D1%8F>

<https://ru.wikipedia.org/wiki/%D0%9B%D0%BE%D0%B3%D0%B8%D1%81%D1%82%D0%B8%D1%87%D0%B5%D1%81%D0%BA%D0%B0%D1%8F_%D1%80%D0%B5%D0%B3%D1%80%D0%B5%D1%81%D1%81%D0%B8%D1%8F>

<http://baguzin.ru/wp/?p=6078>

<http://ufldl.stanford.edu/tutorial/supervised/SoftmaxRegression/>

Метод анализа главных компонент PCA и сингулярное разложение SVD

<http://docs.opencv.org/master/d1/dee/tutorial_introduction_to_pca.html#gsc.tab=0>

<https://en.wikipedia.org/wiki/Singular_value_decomposition>

<https://arxiv.org/pdf/1404.1100.pdf> (связь PCA и SVD)

<https://en.wikipedia.org/wiki/Eigenface> (известное применение PCA для задачи распознавания лиц)

Методы кросс-валидации

<http://scikit-learn.org/stable/modules/cross_validation.html#cross-validation>

Density estimation

<https://en.wikipedia.org/wiki/Density_estimation>

<https://en.wikipedia.org/wiki/Kernel_density_estimation>

<http://scikit-learn.org/stable/modules/density.html>

t-SNE

<https://en.wikipedia.org/wiki/T-distributed_stochastic_neighbor_embedding>

<http://scikit-learn.org/stable/modules/manifold.html#t-distributed-stochastic-neighbor-embedding-t-sne>

<http://lvdmaaten.github.io/publications/papers/JMLR_2008.pdf>

Методы предобработки данных (нормализация, стандартизация):

<http://scikit-learn.org/stable/modules/preprocessing.html>

Отбор признаков

<http://blog.kaggle.com/2016/04/08/homesite-quote-conversion-winners-write-up-1st-place-kazanova-faron-clobber/>

Решающие деревья

<https://en.wikipedia.org/wiki/Decision_tree_learning>

<http://scikit-learn.org/stable/modules/tree.html#tree-algorithms-id3-c4-5-c5-0-and-cart>

Градиентный спуск и нейронные сети

<https://ru.wikipedia.org/wiki/%D0%93%D1%80%D0%B0%D0%B4%D0%B8%D0%B5%D0%BD%D1%82%D0%BD%D1%8B%D0%B9_%D1%81%D0%BF%D1%83%D1%81%D0%BA>

<http://ufldl.stanford.edu/tutorial/supervised/MultiLayerNeuralNetworks/>

<https://www.willamette.edu/~gorr/classes/cs449/momrate.html>

<http://jmlr.org/proceedings/papers/v9/glorot10a/glorot10a.pdf>

<https://ru.wikipedia.org/wiki/%D0%A2%D0%B5%D0%BE%D1%80%D0%B5%D0%BC%D0%B0_%D0%A6%D1%8B%D0%B1%D0%B5%D0%BD%D0%BA%D0%BE>

Кластеризация

<https://en.wikipedia.org/wiki/Expectation%E2%80%93maximization_algorithm>

<https://ru.wikipedia.org/wiki/K-means>

<http://scikit-learn.org/stable/modules/clustering.html>

Бустинг

<https://en.wikipedia.org/wiki/Boosting_%28machine_learning%29>

<https://en.wikipedia.org/wiki/Gradient_boosting>

<http://xgboost.readthedocs.io/en/latest/model.html> (бустинг деревьев при помощи xgboost)

Функции потерь

<https://en.wikipedia.org/wiki/Cross_entropy#Cross-entropy_error_function_and_logistic_regression>

Стэкинг

<https://en.wikipedia.org/wiki/Ensemble_learning#Stacking>

<https://www.kaggle.com/c/otto-group-product-classification-challenge/forums/t/14335/1st-place-winner-solution-gilberto-titericz-stanislav-semenov> (пример схемы стекинга)

<http://blog.kaggle.com/2015/12/03/dato-winners-interview-1st-place-mad-professors/> (пример схемы стекинга)

Метрики качества классификации

<https://en.wikipedia.org/wiki/Precision_and_recall>

<https://en.wikipedia.org/wiki/Accuracy_and_precision>

<https://en.wikipedia.org/wiki/F1_score>

<https://en.wikipedia.org/wiki/Evaluation_of_binary_classifiers>

Сверточные нейронные сети

<http://cs231n.github.io/convolutional-networks/>

Анализ текстовых данных

<https://en.wikipedia.org/wiki/Tf%E2%80%93idf>

<https://en.wikipedia.org/wiki/Latent_semantic_analysis>

<https://en.wikipedia.org/wiki/Latent_semantic_indexing>

<https://en.wikipedia.org/wiki/Non-negative_matrix_factorization>

<https://en.wikipedia.org/wiki/Vector_space_model>

<https://en.wikipedia.org/wiki/Bag-of-words_model>

<https://en.wikipedia.org/wiki/Jaccard_index>

<https://en.wikipedia.org/wiki/Cosine_similarity>

<https://en.wikipedia.org/wiki/Word_embedding>

<https://en.wikipedia.org/wiki/Word2vec>

<http://rare-technologies.com/word2vec-tutorial/#app>

Некоторые другие понятия

<https://ru.wikipedia.org/wiki/Basic_Linear_Algebra_Subprograms>

<https://ru.wikipedia.org/wiki/%D0%96%D0%B0%D0%B4%D0%BD%D1%8B%D0%B9_%D0%B0%D0%BB%D0%B3%D0%BE%D1%80%D0%B8%D1%82%D0%BC>

<https://en.wikipedia.org/wiki/Random_forest>

<https://en.wikipedia.org/wiki/Bootstrap_aggregating>

<https://ru.wikipedia.org/wiki/HSV_%28%D1%86%D0%B2%D0%B5%D1%82%D0%BE%D0%B2%D0%B0%D1%8F_%D0%BC%D0%BE%D0%B4%D0%B5%D0%BB%D1%8C%29>

<https://ru.wikipedia.org/wiki/LAB>

<https://en.wikipedia.org/wiki/Early_stopping>

**Практические вопросы:**

Про получение вероятности из SVM

<https://en.wikipedia.org/wiki/Platt_scaling>

<http://www.csie.ntu.edu.tw/~cjlin/papers/plattprob.pdf>

Выбор параметров RBF SVM и их интуитивная интерпретация

<http://scikit-learn.org/stable/auto_examples/svm/plot_rbf_parameters.html#example-svm-plot-rbf-parameters-py>

Гайд по svm от создателей библиотеки libsvm

<https://www.csie.ntu.edu.tw/~cjlin/papers/guide/guide.pdf>

Выбор ядра svm

<https://books.google.ru/books?id=2vb-LZEn8uUC&pg=PA120&lpg=PA120&dq=linear+kernel+is+a+special+case+of+RBF&source=bl&ots=zV50172RkP&sig=D6a_GPzpW0fsjw04kWNjkCAbUL4&hl=ru&sa=X&ved=0CCAQ6AEwAWoVChMIz63Syob6yAIVgxEsCh0RggfS#v=onepage&q=linear%20kernel%20is%20a%20special%20case%20of%20RBF&f=false>

О подборе параметров алгоритмов

http://scikit-learn.org/stable/modules/grid\_search.html

https://en.wikipedia.org/wiki/Hyperparameter\_optimization

<https://scholar.google.ru/scholar?hl=ru&q=hyperparameter+optimization&btnG=>

<https://en.wikipedia.org/wiki/No_free_lunch_in_search_and_optimization>

Выбор активационной функции слоев нейронной сети

<http://cs231n.github.io/neural-networks-1/>

<https://www.reddit.com/r/MachineLearning/comments/32iyt9/question_comparison_between_softmax_and_sigmoid/>

<https://www.reddit.com/r/MachineLearning/comments/37ardy/softmax_vs_sigmoid_for_output_of_a_neural_network/>

О выборе количества слоев и количества нейронов

<https://www.quora.com/Artificial-Neural-Networks-ANNs/Artificial-Neural-Networks-How-can-I-estimate-the-number-of-neurons-and-layers>

<https://arxiv.org/abs/1206.5533>

Видеоуроки по Scikit-Learn

<http://www.datasciencecentral.com/profiles/blogs/scikit-learn-tutorial-series>