

ГОВОРИ ПО РУССКИ (понимая, что говоришь)

Accuracy - доля правильных ответов $\frac{TP+TN}{TP+TN+FP+FN}$.

AUC (Area Under Curve) - площадь под кривой.

AUC-ROC - площадь под кривой ошибок.

Augmentation - аугментация, пополнение, искусственное создание новых прецедентов как слегка измененных копий существующих данных (для обучения ИНС).

Average precision (AP) - средняя точность.

Average-pooling - подвыборка по среднему значению (в Python).

Baseline - **маш.об.:** опорное решение, базовый метод, общеупотребимая программная реализация решения задачи.

Breakeven point - точка баланса (то же, что R-precision).

Case - прецедент, объект выборки.

Characteristic - признак.

Confusion matrix - матрица ошибок.

Cross-entropy loss - то же, что LogLoss.

Data Mining - интеллектуальный анализ данных.

Dataset - выборка, набор данных.

Detection - выявление, обнаружение.

Dropout - метод отсева (в регуляризации ИНС).

Embedding - векторное представление.

End-to-end обучение - **обр.изобр.:** технология, при которой на вход ИНС поступают необработанные исходные изображения, а на выходе получают готовые ответы.

False Negative (FN) - **маш. об.:** число ошибочно классифицированных объектов отрицательного класса, **стат.:** ошибка II-го рода, **воен.:** ложная тревога.

False Positive (FP) - маш. об.: число ошибочно классифицированных объектов положительного класса, стат.: ошибка I-го рода, воен.: пропуск цели.

False Positive Rate (FPR) - доля неверно классифицированных объектов

$$\frac{FP+FN}{TP+TN+FP+FN}$$

Feature - 1) признак; 2) полезное свойство, особенность, возможность.

Features selection - выбор информативных признаков.

F-score, F_β , - F-оценка, среднее гармоническое точности и полноты с весом точности (precision) β ; F1 - $\beta = 1$.

Gain - выигрыш (напр. Information gain).

IoU (Intersection Over Union) - мера $\frac{|A \cap B|}{|A \cup B|}$ совпадения (близости) множеств A и B .

Lift - маш. об.: прирост концентрации.

Logistic Loss (LogLoss) - логистическая функция потерь.

Loss - потеря.

mAP at K (mean Average Precision at K) - (именованная) средняя точность (на K элементах).

Maximum Likelihood Estimation (MLE) - оценка максимального правдоподобия.

Max-pooling - подвыборка по максимальному значению (в Python).

Mean Absolute Error (MAE) - средняя абсолютная ошибка.

Mean Squared Error (MSE) - средняя квадратичная ошибка.

Non-maximum suppression - не-максимальное подавление, **обр. изобр.:** алгоритм фильтрации, устраняющий дублирующие гипотезы определения положения объекта.

Oversampling - метод избыточной выборки (увеличение размера малочисленного, минорного класса).

Patch - «заплата», многозначный термин: 1) программное средство, используемое для устранения ошибок или изменения ПО; 2) приём

исправления изображения, совмещающий заливку и выделение; 3) область изображения.

Pooling - объединение элементов.

PR-curve (Precision-Recall Curve) - PR-кривая, кривая «точность-полнота».

Precision - точность, $\frac{TP}{TP+FP}$.

Proposal - гипотеза, **обр. изобр.:** определенная область изображения, в котором предположительно находится объект.

R-CNN (Region Convolution Neural Network) - одна из первых структур ИНС, применяемых для определения нахождения объекта на изображении.

Recall - полнота, $\frac{TP}{TP+FN}$.

Reward - награда, премия (в обучении с подкреплением).

ROC (Receiver Operating Characteristic curve) - кривая ошибок.

R-precision - R-точность, точка баланса.

Supervised learning - обучение с учителем.

True Negative (TN) - **маш. об.:** число правильно классифицированных объектов отрицательного класса, **воен.:** правильное необнаружение.

True Positive (TP) - **маш. об.:** число правильно классифицированных объектов положительного класса, **воен.:** правильное обнаружение.

True Positive Rate (TPR) - доля верно классифицированных объектов $\frac{TP+TN}{TP+TN+FP+FN}$.

Undersampling - метод недостаточной выборки (уменьшение размера доминирующего класса).