

Площадь под ROC-кривой (Area Under The ROC Curve)

Синонимы: AUC, Область под ROC-кривой

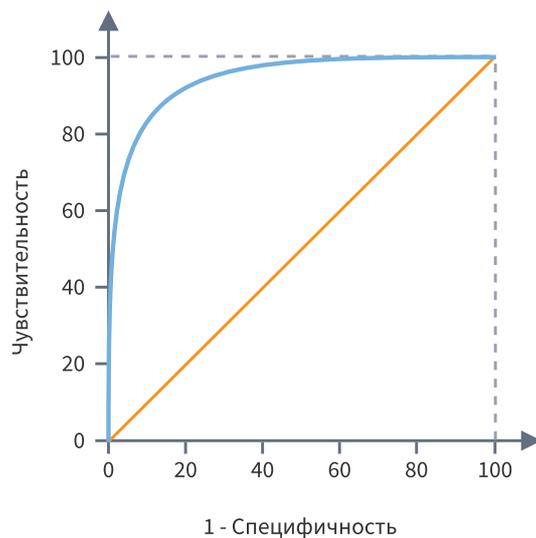
Разделы: [Метрики](#)

Площадь под ROC-кривой — часть координатной плоскости под графиком ROC-кривой. В математической статистике и машинном обучении является мерой качества модели бинарной классификации (логистической регрессии, машины опорных векторов, линейного дискриминантного анализа и т.д.).

ROC-кривая — графический инструмент оценки точности моделей бинарной классификации. Она позволяет найти оптимальный баланс между чувствительностью и специфичностью модели, который соответствует точке ROC-кривой, наиболее близкой к координате (0,1), в которой чувствительность и специфичность равны 1, когда ложно-положительные и ложно-отрицательные классификации отсутствуют.

В случае идеальной модели график ROC-кривой проходит через точку (0,1), и площадь под ним максимальна и равна 1. По мере снижения точности модели (т.е. росте числа ложно-положительных и ложно-отрицательных классификаций) кривизна ROC-кривой уменьшается, при этом уменьшается и AUC.

$AUC = 0,7$ можно интерпретировать как вероятность 70%, что модели удастся разделить классы. Как только ROC-кривая выродится в прямую линию, когда $AUC = 0,5$, соответствующая модель станет бесполезной в том смысле, что она будет работать не лучше классификатора, который присваивает классы случайным образом (путем простого угадывания). Дальнейшему уменьшению параметра AUC будут соответствовать классификаторы, работающие хуже случайного, когда вероятность неправильной классификации выше, чем вероятность правильной.



Когда $AUC = 0$, такой классификатор всегда распознает положительный пример как отрицательный, т.е. вероятность ошибки составляет 100%. На практике в зависимости от значения AUC эффективность модели классифицируется следующим образом:

- $0,8 \leq AUC \leq 1,0$ — модель работает превосходно;
- $0,6 \leq AUC < 0,8$ — модель работает хорошо;
- $0,5 < AUC < 0,6$ — модель работает удовлетворительно;
- $AUC \leq 0,5$ — модель не работает.

В Logiном существует специализированный визуализатор качество бинарной классификации, в котором строятся диаграммы и таблицы с результатами моделирования. С его помощью производится оценка полученных метрик, в том числе и AUC ROC — площадь под ROC-кривой.