

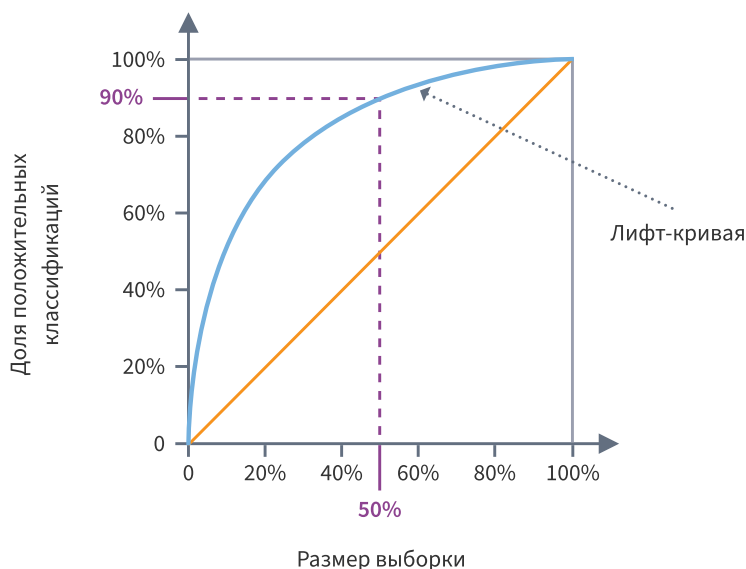
Лифт-кривая (Lift-Curve)

Разделы: [Метрики](#), [Визуализация](#)

Лифт-кривая представляет собой визуальное средство для оценки предсказательной способности и сравнения моделей бинарной классификации, а также для оптимизации издержек классификации.

Предположим, что модель бинарной классификации формирует на выходе оценки принадлежности наблюдения к классу, а также определяет порог для разделения классов. Лифт-кривая строится в системе координат, по горизонтальной оси которой откладывается число наблюдений, ранжированное по значению предсказанной вероятности положительного исхода классификации, а по вертикали — число положительных классификаций.

Таким образом, каждая точка лифт-кривой показывает ожидаемое число положительных классификаций для соответствующего размера выборки. Сама же лифт-кривая отражает прирост (**лифт**) положительных классификаций при увеличении выборки.



Диагональная линия на графике (как и для ROC-кривой) соответствует «бесполезному» классификатору, который выполняет предсказания случайным образом. Очевидно, что модель с такой лифт-кривой не обеспечивает лифт, и смысла наращивать объем выборки для получения лучшей модели не имеет смысла. Если лифт-кривая проходит ниже «линии бесполезности», то такая модель оказывается хуже случайной и использовать ее бессмысленно.

Любая модель, для которой лифт-кривая проходит выше диагональной линии, обеспечивается лифт. И он тем больше, чем выше кривизна лифт-кривой. Высокий лифт означает, что даже небольшое увеличение выборки приводит к значительному росту доли положительных классификаций.

На практике данное свойство лифт-кривой используется для оптимизации издержек, связанных с проведением маркетинговых компаний. Пусть, например, ставится задача попытаться вернуть клиентов, прекративших пользоваться услугами компании за последний месяц (коих насчитывается, скажем, 100). Для этого с каждым клиентом совершается контакт (рассылка, звонок, SMS и т.д.) стоимостью 20 рублей. Тогда стоимость всего маркетингового мероприятия будет 2000 рублей. Пожизненная стоимость клиента 30 рублей, вернулось 40 клиентов, принесла компании доход 1200 рублей, т.е. маркетинговая кампания оказалась провальной.

Однако, если построить модель бинарной классификации, предсказывающую возврат клиента (положительный исход) и построить лифт-кривую, то можно увидеть, что контакт с 50% клиентов — это контакт с 90% вернувшихся. Уменьшив число контактов вдвое, стоимость кампании снизится до 1000 рублей, а прибыль (с 36 клиентов) составит 1080 рублей, что позволит вывести кампанию в плюс.

Еще одним приложением лифт-кривых является сравнение качества моделей бинарной классификации. Для этого обычно используется разновидность лифт-кривых, называемая гайн-кривой (gain-curve — от англ. gain — прирост, усиление) или кумулятивной лифт-кривой.

Отличием гайн-кривой является то, что по вертикальной оси откладывается не число положительно классифицированных примеров, а чувствительность. При увеличении размера выборки увеличивается количество ложно-положительных наблюдений. Соответственно, увеличивается и риск ложно-положительной классификации. При уменьшении выборки, наоборот, число ошибок классификации снижается. Задача заключается в том, чтобы выбрать такой размер выборки, для которого точка на гайн-кривой максимально приближена к точке (0,1).

В Loginom существует специализированный визуализатор качество бинарной классификации, который отображает различные диаграммы и их показатели (в том числе лифт-кривую) и таблицы с результатами проведенной классификации на основе логистической регрессии.

С помощью лифт-кривой происходит оценка качества скоринговой модели в решении Loginom Scorecard Modeler.