### Оценка дискриминирующей способности модели PD Lifetime

Дискриминирующая способность модели PD Lifetime отражает ее способность разделять объекты, которые подвержены дефолту, и те, которые останутся недефолтными на каждом временном горизонте. Это ключевой аспект оценки, поскольку модель должна эффективно отличать заемщиков с высоким риском от заемщиков с низким риском дефолта.

#### Основные тесты для оценки дискриминирующей способности

1. **Кривая ROC и AUC (Area Under Curve)**  
   ROC-кривая строится для каждого временного горизонта () путем сравнения доли истинно положительных срабатываний (True Positive Rate, TPR) и доли ложноположительных срабатываний (False Positive Rate, FPR) при различных порогах вероятности дефолта.

* **AUC (площадь под кривой):**
  + AUC измеряет дискриминацию модели; она варьируется от 0.5 (случайное угадывание) до 1.0 (идеальная дискриминация).
  + Чем выше AUC, тем лучше модель отличает дефолтные и недефолтные объекты.
* **Интерпретация:**
  + AUC > 0.8 — высокая дискриминирующая способность.
  + AUC в диапазоне 0.7–0.8 — удовлетворительная.
  + AUC < 0.7 — низкая дискриминирующая способность, требующая доработки модели.
* ROC-кривые и AUC рассчитываются отдельно для каждого горизонта (, и так далее), чтобы оценить дискриминацию на разных этапах жизненного цикла актива.

1. **Кривая PR (Precision-Recall) и AUC PR**  
   Precision-Recall (PR) кривая применяется, если данные несбалансированы (например, доля дефолтов значительно ниже доли недефолтов).
   * **Precision (точность):** доля истинных дефолтов среди всех объектов, классифицированных как дефолтные.
   * **Recall (полнота):** доля истинных дефолтов, которые были правильно идентифицированы моделью.

* AUC для PR-кривой показывает дискриминирующую способность в условиях несбалансированных данных, где обычный ROC-анализ может быть менее информативным.

1. **Коэффициент Каппы (Cohen’s Kappa)**  
   Этот показатель измеряет, насколько результаты классификации (дефолт/недефолт) превосходят случайное угадывание. Формула:

* где:
  + — доля совпадений предсказаний с наблюдениями,
  + — ожидаемая доля совпадений при случайном распределении.
* Значение близкое к 1 указывает на высокую дискриминирующую способность модели.

1. **KS-статистика (Kolmogorov-Smirnov)**  
   KS-статистика используется для измерения максимальной разницы между кумулятивными распределениями вероятностей дефолта для дефолтных и недефолтных объектов.
   * Для каждого временного горизонта () вычисляются два кумулятивных распределения:
     + — для объектов, которые дефолтнули.
     + — для объектов, которые не дефолтнули.
   * KS-статистика определяется как максимальная вертикальная разница между этими распределениями.

* **Интерпретация:**
  + Чем выше значение KS, тем лучше модель различает дефолтные и недефолтные объекты.
  + Значения KS > 0.3 считаются хорошими, KS между 0.2 и 0.3 — удовлетворительными.

1. **Gini-коэффициент**  
   Gini-коэффициент — это альтернативный показатель, связанный с AUC ROC. Он рассчитывается как:

* Значения Gini выше 0.4 обычно свидетельствуют о хорошей дискриминирующей способности.

1. **Сегментный анализ дискриминации**  
   Для более глубокого анализа дискриминирующей способности модель проверяется на разных сегментах заемщиков (например, по отраслям, регионам, возрастным группам или уровням кредитного рейтинга).
   * ROC-кривые и AUC рассчитываются отдельно для каждого сегмента.
   * Сравнение результатов между сегментами позволяет выявить группы, где модель хуже различает дефолтные и недефолтные объекты.

#### Рекомендации по улучшению дискриминирующей способности

* Если дискриминация модели на определенных горизонтах или в определенных сегментах недостаточна, возможны следующие меры:
  + Улучшение качества входного параметра , так как он оказывает значительное влияние на прогнозы .
  + Пересмотр архитектуры модели для учета дополнительных факторов риска.
  + Использование различных алгоритмов для учета нелинейных зависимостей в данных.

Эти тесты обеспечивают объективную оценку способности модели PD Lifetime точно отличать заемщиков с высоким риском дефолта от низкорисковых заемщиков на протяжении всего срока жизни кредита.