

участники: Россети, Техснабэкспорт, Таврида Электрик, QIWI. Цели к 2035 году: 40 млрд долл. США в год объем выручки Российской Федерации на глобальном рынке.

7. ТЕХНЕТ • Цифровое проектирование и моделирование; • Новые материалы; • Аддитивные технологии; • Робототехника; • Big data и Plot. Компании-участники: ОАК, НПО «Сатурн», Волгабас, CompMechLab [2]. Цели к 2035 году: 10-е место Российской Федерации в рейтинге Global Manufacturing Competitiveness Index.

8. КРУЖКОВОЕ ДВИЖЕНИЕ • Цифровое управление талантами, наставничество; • Технологические соревнования; • Проектные школы, ярмарки, фестивали. Компании-участники: университеты, кванториумы, «Сириус», ЦМИТы. Цели к 2025 году: 500 тыс. участников движения.

Подводя итог вышесказанному, отметим, что Национальная технологическая инициатива (НТИ) это развивающаяся система, подстраивающаяся под самые современные рыночные тенденции.

#### Библиографические ссылки

1. Информбюро 20.35 – официальный медиа-ресурс Национальной Технологической Инициативы [Электронный ресурс]. – Режим доступа : <https://ntinews.ru/>. – Дата доступа : 02.02.2020.

2. Официальный сайт Национальной технологической инициативы [Электронный ресурс]. – Режим доступа : <https://nti2035.ru/>. – Дата доступа : 28.01.2020.

УДК 330.43

## МЕТОДЫ ОЦЕНКИ КАЧЕСТВА СКОРИНГОВОЙ МОДЕЛИ

А. И. Ткачѳв<sup>1)</sup>, И. А. Карачун<sup>2)</sup>

<sup>1)</sup> *Ведущий специалист Управления финансовой стабильности Национального банка, аспирант экономического факультета Белорусского государственного университета, г. Минск*

<sup>2)</sup> *Кандидат экономических наук, доцент, заведующий кафедрой цифровой экономики Белорусского государственного университета, г. Минск*

Важным заключительным этапом построения скоринговой системы является этап ее валидации, который заключается в проверке достоверности полученной модели на обучающей и контрольной выборке. Валидация скоринговой модели должна производиться на основе системы критериев качества. В работе предлагается комплексная методика оценки кредитных скоринговых карт, основанная на таких критериях как обычное и кумулятивное распределение кредитов, статистка Колмогорова – Смирнова, ROC-кривая и Lift-кривая.

**Ключевые слова:** кредитный риск; сравнение моделей; ROC-кривая; оценка кредитоспособности; математические методы.

## METHODS FOR EVALUATING THE QUALITY OF THE SCORING MODEL

A. I. Tkachev<sup>1)</sup>, I. A. Karachun<sup>2)</sup>

<sup>1)</sup> *Leading Specialist of the Financial Stability Department of the National Bank, Post-Graduate Student of the Faculty of Economics, Belarusian state University, Minsk*

<sup>2)</sup> *PhD in Economics, Associate Professor, Head of the Department of Digital Economy, Belarusian State University, Minsk*

An Important final stage of building a scoring system is its validation stage, which consists in verifying the validity of the obtained model on a training and control sample. Validation of the scoring model should be based on a system of quality criteria. The paper proposes a comprehensive method for evaluating credit scoring

cards based on such criteria as the usual and cumulative distribution of loans, the Kolmogorov – Smirnov statistic, the ROC curve and the Lift curve.

**Key words:** credit risk; model comparison; ROC curve; credit rating; mathematical methods.

В июле 2004 г. Базельским комитетом был выпущен документ: Revised Framework on International Convergence of Capital Measurement and Capital Standards (Пересмотренная схема между-народного объединения подходов и стандартов расчета капитала (далее – Базель II)). Для целей исследования дефолтом принято считать наличие у клиента просроченной задолженности свыше 90 дней по основному долгу или процентам, что соответствует Базель II.

В соглашении Базель II дефолт конкретного контрагента считается произошедшим в случае, если имело место одно (или несколько одновременно) из следующих событий:

- банк считает, что заемщик не в состоянии полностью погасить свои кредитные обязательства без принятия банком решения о реализации обеспечения;
- заемщик имеет просрочку более чем 90 дней по погашению любых существенных обязательств перед банком [1]

Базельский комитет определяет такой параметр риска как среднегодовая вероятность дефолта (PD – Probability of default). PD – это оценка вероятности того, что произойдет дефолт, то есть что кредит не будет выплачен.

Банки применяют скоринговые модели в процессе розничного кредитования с целью выявления потенциально более рискованных заемщиков и сокращения потерь, вызванных отказом должников платить по своим обязательствам. Чем более точно скоринговая модель разделяет «хороших» и «плохих» заемщиков, тем меньше потери и больше прибыль банка.

Заключительным этапом построения модели логистической регрессии является проверка ее достоверности и апробация на реальных данных. О степени валидации (от англ. validity – доказанность, обоснованность, пригодность) модели логистической регрессии говорит способность ее правильно классифицировать заемщиков, способность модели отличать «хороших» заемщиков от «плохих». Модель должна давать корректные прогнозы не только на обучающей совокупности, но и на практике при ее применении.

Одна из стратегий валидации модели – формирование случайным образом двух выборок: обучающей – по ней строится модель, и тестовой – используется для проверки модели. Проверку достоверности модели, как правило, производят на обучающей и контрольной выборке в пропорциях примерно 70–80 % и 30–20 % соответственно от исходных данных для построения модели. Хорошая модель должна давать приемлемые результаты точности и на обучающей, и на контрольной выборке. Схожие показатели, полученные на обеих выборках – признак того, что на практике модель будет выдавать верные прогнозы.

Более сложная стратегия валидации модели может предполагать формирование трех и более выборок. Например, первая выборка используется для оценки параметров модели. Вторая выборка – для проверки модели. Если получаются значительные отклонения результатов по обучающей и тестовой выборке, то из них удаляются выбросы или переменные, влияющие на эти отклонения, а затем строится новая модель по объединенной первой и второй выборке. Результаты новой модели проверяются на заранее зарезервированной третьей выборке.

При построении скоринговых карт с применением логрессионных моделей в первую очередь проверяется статистическая значимость модели в целом и значимость коэффициентов в отдельности, а также расчет доверительных интервалов. Тем не менее стоит отметить, что эти процедуры подходят не для всех скоринговых карт. В ситуации, когда включаются новые экспертные критерии, на первый план выходит вопрос оценки модели не только статистическими методами. Нужны методы, пригодные для любых скоринговых карт. Выделим четыре таких метода:

1. Построение кривых распределения скоринговых баллов для хороших и плохих кредитов. Чем меньше пересечений между данными распределениями, тем более эффективна скоринговая карта.

## 2. Тест Колмогорова – Смирнова (К-С).

Он представляет из себя распределение кумулятивных процентов для хороших и плохих кредитов. Вертикальное расстояние между точками кумулятивных кривых есть уровень статистики Колмагорова – Смирнова (К-С). График показывает зависимость скорингового балла от доли хороших кредитов накопленным итогом и доли плохих кредитов накопленным итогом. Эта статистика измеряет способность скоринговой карты разделять классы, и на практике ее диапазон варьирования – от 25 до 75. Чем выше значение статистики К-С, тем лучше качество скоринговой карты. В месте максимального значения К-С лежит оптимальный скоринговый балл, минимизирующий суммарную ошибку классификации I и II рода, рисунок 1.

## 3. Lift-кривая.

Формируется на основе лифт-фактора, который в машинном обучении при решении задачи оптимизации массовой рассылки был впервые определен как показатель, отражающий увеличение числа откликов относительно числа почтовых отправлений. Lift-кривая строится следующим образом: по горизонтальной оси откладывается размер выборки в процентах от общего числа наблюдений в списке примеров, упорядоченном по убыванию вероятности наступления положительного исхода. По вертикальной оси фиксируется кумулятивное число положительных исходов в каждой подвыборке [2].

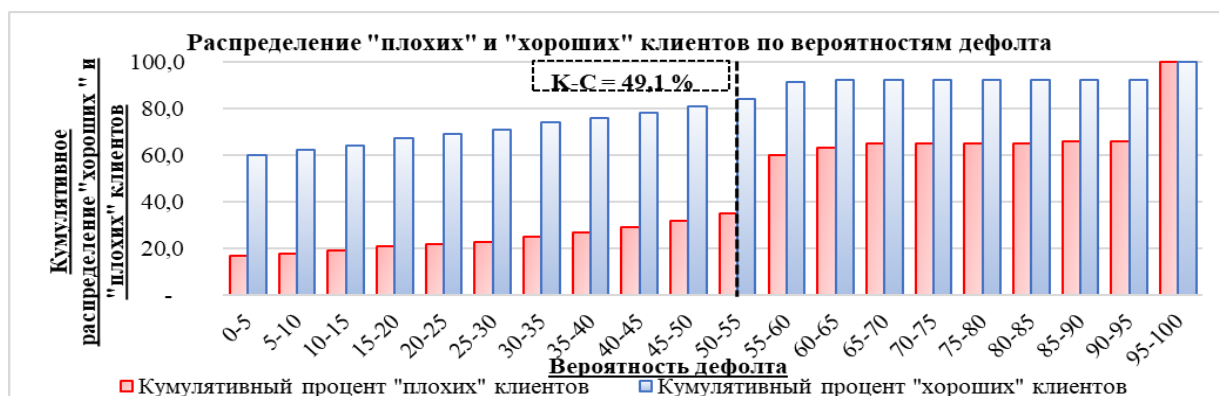


Рисунок 1 – Кумулятивное распределение клиентов, статистический тест К-С

Примечание – Источник [3].

4. В результатах модели в случае бинарной классификации («положительный» или «отрицательный») наблюдение может принадлежать любому из следующих типов в зависимости от вероятности прогнозирования модели:

- TP (истинно положительный) – если наблюдение правильно классифицировано как «положительное»;
- FP (ложноположительный) – если наблюдение неправильно классифицировано как «положительное»;
- TN (истинно отрицательный) – если наблюдение правильно классифицировано как «отрицательное»;
- FN (ложноотрицательный) – если наблюдение ошибочно классифицируется как «отрицательное».

Точность модели характеризуется двумя показателями: чувствительностью (Sensitivity) и специфичностью (Specificity). Чувствительность измеряется долей истинно положительных результатов – долей положительных результатов у «плохих» заемщиков, рассчитывается по формуле  $TP / (TP + FN)$ . Специфичность измеряется долей истинно отрицательных результатов у «хороших» заемщиков, рассчитывается по формуле  $TN / (TN + FP)$ . Пары значений чувствительности и 1-специфичности на всем диапазоне разделяющих значений соответствуют координатам ROC-кривой, как на рисунке 2 [3].

Таким образом, совокупно применяя данные методы, можно выбрать модель, подходящую наилучшим образом под заданные стратегией банка риск-аппетиты.

Провести сравнение и выбрать подходящую методику оценки можно вне зависимости от применяемых при построении модели методов и баз данных. Например, мож-

но проверить пригодность использования готовых моделей, построенных на базе данных материнского банка или дочерних банков в иных странах.

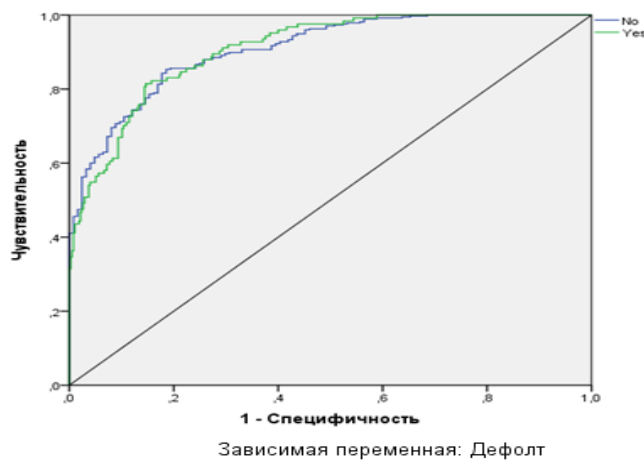


Рисунок 2 – ROC-кривая (кривая ошибок)

Примечание – Источник [4].

#### Библиографические ссылки

1. Basel Committee on Banking Supervision. Basel II: Revised international capital [Electronic resource] // Bank for International Settlements. – June, 2004. – Mode of access : <http://www.bis.org/publ/bcbs107.pdf>. – Date of access : 24.09.2018.
2. Уланов С. В. Оценка качества и сравнение скоринговых карт / С. В. Уланов // Экономические науки. – 2009. – № 9(58). – С. 330–335.
3. Ткачѳв А. И. Системы кредитного скоринга. Матричный подход / А. И. Ткачѳв, А. В. Шипунов // Банкаўскі веснік. – 2019. – № 10(674). – С. 37–46.
4. Ткачѳв, А. И. Применение экономико-математических моделей для принятия решений при кредитовании / А. И. Ткачѳв, А. В. Шипунов // Банкаўскі веснік. – 2018. – № 10(663). – С. 43–50.

УДК 656:004 (476)

## ЭКОНОМЕТРИЧЕСКАЯ МОДЕЛЬ ОЦЕНКИ АВТОМОБИЛЯ КЛАССА J

**Н. Ю. Трифонов<sup>1)</sup>, В. В. Коржуков<sup>2)</sup>**

<sup>1)</sup> Кандидат физико-математических наук, доцент, почётный оценщик Республики Казахстан, доцент кафедры экономики торговли и услуг Белорусского государственного экономического университета, г. Минск

<sup>2)</sup> Магистрант электромеханического факультета Белорусско-Российского университета, г. Могилѳв

Описываются результаты эконометрического моделирования рыночной стоимости автомобилей класса J на основе сбора с помощью программы-парсера информации с популярного интернет-сайта. Полученная модель в виде статистически значимого регрессионного уравнения предназначена для использования в практике оценочной деятельности.

**Ключевые слова:** дорожные транспортные средства; интернет-данные; оценка стоимости; парсер; регрессионная модель; рыночное обесценивание; статистическая выборка.