

ПРИМЕНЕНИЕ МЕТОДОВ МАШИННОГО ОБУЧЕНИЯ ДЛЯ АНАЛИЗА БАНКОВСКОГО СЕКТОРА

Борисова Л.Р., Жукова Г.С.

Финансовый университет при Правительстве РФ, Москва, Россия

borisovalr@mail.ru, galsevzhukova@mail.ru

Аннотация. Исследуется взаимосвязь рентабельности капитала и рентабельности активов банков Российской Федерации, изучаются вопросы влияния наиболее информативных и основных показателей банков на показатель прибыли на февраль 2022 года. Используются методы машинного обучения для совместного исследования нормативов ликвидности и основных финансовых показателей на рентабельность капитала банков.

Ключевые слова: банк, прибыльность, рентабельность, финансовые показатели, машинное обучение, регрессия.

Введение

Финансовая система для жизни страны не случайно сравнивается с кровеносной системой для жизни живого организма. Основой финансовой системы является банковская. Банки, предоставляя кредиты, техническую помощь, гранты и инвестиции в акционерный капитал, содействуют социально-экономическому развитию регионов страны. Не случайно в самом развитом регионе – Москве – сосредоточено большинство банков. Все крупные банки – Сбербанк, ВТБ, Россбанк, Альфа-банк – зарегистрированы в Москве. Банки предоставляют консультационные услуги, сосредотачивают финансовые ресурсы с помощью операций по совместному финансированию с привлечением официальных, коммерческих и экспортных кредитных источников. Для бизнеса важно контролировать свои операции на корпоративном уровне. Поэтому практически каждая крупная промышленная или торговая компания имеет свой региональный банк. Теоретические вопросы финансового анализа в банках остаются до настоящего времени актуальными как для управления коммерческим банком, так и для клиентов, выбирающих надежный банк для размещения личных сбережений.

1. Регрессионный анализ

1.1. Прибыльность банков

Прибыльность банка влияет на финансовую стабильность [1]. Самыми прибыльными банками по данным Центрального Банка России являются Сбербанк, ВТБ, Газпромбанк, Россельхозбанк, Московский Кредитный Банк, Банк Открытие, Совкомбанк, Райффайзенбанк и Росбанк. Не вызывает сомнения предположение, что доверие населения и бизнеса – это основа прибыльности банков в современном мире, основной характеристикой которого является цифровизация. Большую роль в устойчивом развитии стран мира играют банки развития. В частности, Азиатский банк развития уже более 50 лет отбирает, финансирует и реализует проекты в странах Азиатско-Тихоокеанского региона. Деятельность этого банка направлена не просто на обеспечение банковской стабильности, а на инклюзивное и устойчивое экономическое развитие, которое помогает достичь финансовой устойчивости, что в конечном итоге должно привести к уменьшению бедности в данном регионе [2].

Ранее [3] по данным на 1 августа 2021 года была исследована зависимость прибыли банков от 5 основных показателей (в тыс. рублей): активов банка (x_1), кредитного портфеля (x_2), просроченной задолженности в кредитном портфеле (x_3), вкладов физических лиц (x_4) и вложения в ценные бумаги (x_5), сделан вывод, что прибыль прямо пропорциональна вкладам клиентов в банк, активам банка и обратно пропорциональна вложениям в ценные бумаги и выданным кредитам, пока от этих сделок не получена прибыль. Интересно посмотреть, как изменится уравнение множественной регрессии, если проанализировать данные на февраль 2022 года.

Если за переменную y обозначить прибыль (в тыс. рублей), то уравнение имеет вид:

$$y = 59431,6 + 0,00125x_1 - 0,00267x_2 - 0,02775x_3 + 0,01597x_4 - 0,01851x_5$$

Все коэффициенты в данном уравнении кроме свободного члена – статистически значимы (p -value=0). Множественный коэффициент корреляции равен 0,98, статистически значим (по результатам 347 наблюдений). Из данного уравнения следует, что прибыль возрастает при

увеличении активов банка и вкладов физических лиц, а кредитование и ценные бумаги прибыль не приносят.

1.2. Рентабельность банков

Проверим, есть ли корреляция между рентабельностью капитала и рентабельностью активов нетто. В анализе принимали участие 332 банка, для которых были данные по этим показателям. Как и следовало ожидать, показатели тесно связаны между собой. Коэффициент корреляции равен 0,84, он статистически значим.

2. Использование методов машинного обучения

2.1. Классификация банков по рентабельности

Методы машинного обучения могут быть применены к анализу при наличии пропусков в данных. Пропусков не должно быть только в признаке, который выбирается в качестве группирующего. Рассмотрим в таком качестве рентабельность капитала. Исключим из дальнейшего анализа рентабельность активов нетто, так как выше была показана высокая корреляция этих показателей рентабельности. В качестве зависимых признаков рассмотрим все основные показатели: активы (нетто), чистую прибыль, капитал (по форме 123), кредитный портфель, просроченную задолженность в кредитном портфеле, вклады физических лиц, вложения в ценные бумаги, кредиты физическим лицам (в тыс. руб.), нормативы ликвидности Н1 и Н3 (в %). Норматив Н1 регулирует риск несостоятельности банка, определяет требования по минимальной величине собственных средств банка, необходимых для покрытия кредитного, операционного и рыночного рисков. Он определяется как отношение размера собственных средств банка и суммы активов, взвешенных по уровню риска. Н3 – норматив текущей ликвидности, то есть возможность банка быстро рассчитаться по своим обязательствам в срок от 30 дней до определенной даты отчета [<https://www.klerk.ru/buh/articles/560341/>].

Как уже было отмечено выше, в России насчитывается всего с десяток самых крупных банков. Но база всех банков России, представленная на сайте банки.ру, содержит более 350 банков. Разумеется, большинство из них – средние и мелкие. Как было отмечено в работе [4] любая кредитная организация в условиях внешней неопределенности каждый день решает задачи, связанные с риском ликвидности.

При решении поставленной задачи были использованы как классические методы, вошедшие в компьютерный комплекс Data Master Azforus (бустинг адаптивный и градиентный, деревья решений, методы опорных векторов и др.), так и разработанный указанной выше компанией, любезно предоставившей свой программный комплекс метод статистически взвешенных синдромов (СВС) [5].

При распознавании использовали данные для 361 банка (для которых были данные по рентабельности капитала). Все используемые в анализе банки были разделены на 2 группы: с отрицательной (первая группа) и положительной рентабельностью (2 группа).

В таблице 1 представлены основные параметры, характеризующие качество распознавая. На рис 1. представлена диаграмма качества.

Таблица 1. Качество классификации для целевой функции – рентабельности капитала (в %)

Accuracy (Правильность)	0,9143
Precision (Точность)	0,8125
Sensitivity (Чувствительность)	0,8125
Specificity (Специфичность)	0,9444

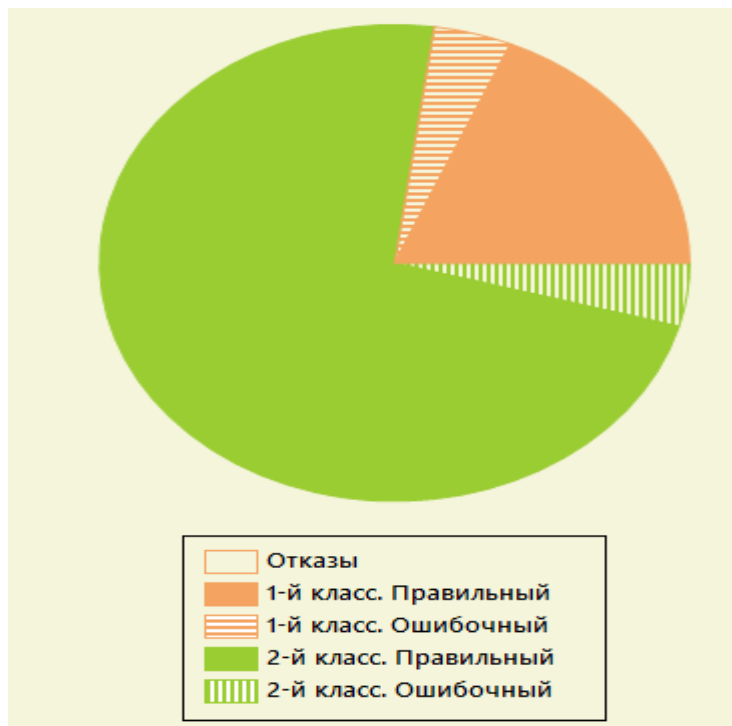


Рис. 1. Диаграмма качества классификации данных по рентабельности капитала, рассчитанная по методу SVC

В результате проведенного распознавания получили площадь под ROC-кривой, равную 0,9 (рис. 2).

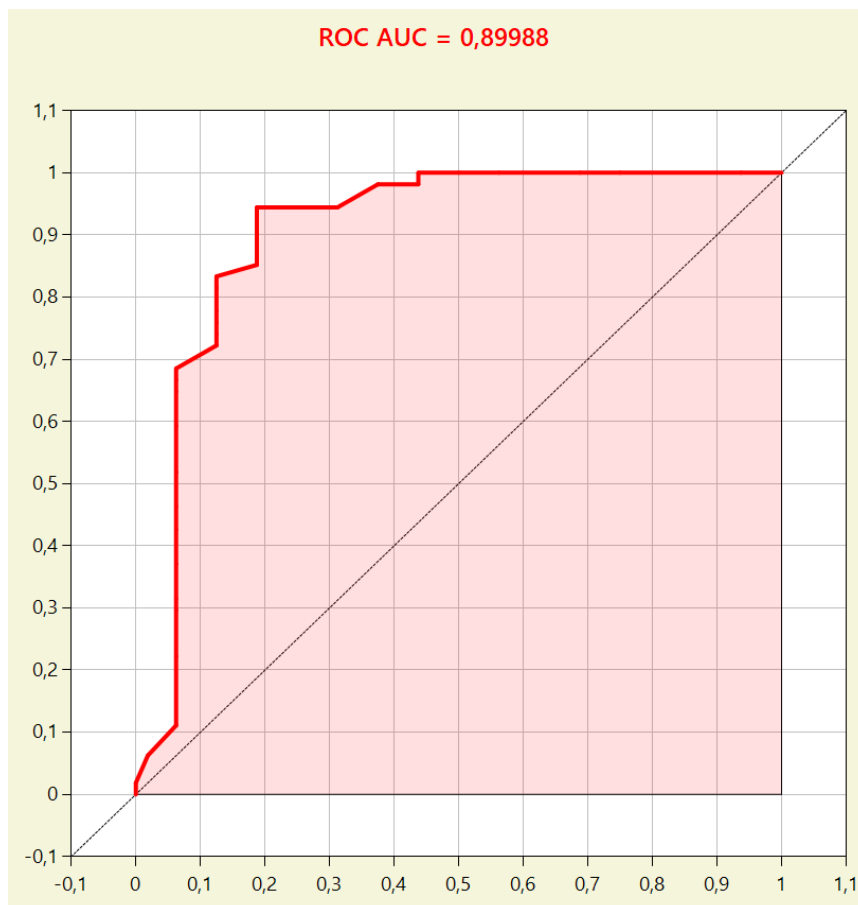


Рис. 2. ROC-кривая при использовании SVC-метода

Отметим, что качество классификации проверяли на чистом контроле, когда в выборку было включено 20% от всех участвующих в исследовании банков. Было всего две ошибки в распознавании: два банка были распознаны в соответствии с указанными выше показателями как банки с положительной рентабельностью, хотя на февраль 2022 года рентабельность у них была отрицательная. Это Московский кредитный банк, на сегодня (июль 2023 года) входящий в десятку самых высокорейтинговых банков и небольшой московский банк Максима, популярный у вкладчиков.

3. Заключение

Полученные в работе результаты свидетельствуют роли временного фактора в исследовании прибыльности и рентабельности банков при использовании статистических методов исследования. Изменение политической и экономической жизни страны существенно сказывается на вклад основных показателей деятельности банков в прибыльность и рентабельность. Изменение – существенное, так как один и тот же показатель в разное время влияет на прибыль банка с разным знаком. Этот вывод сделан на основе сравнительного анализа использования регрессионного анализа. Из построенных в 2021 году и сейчас двух уравнений множественной регрессии зависимости прибыльности банков от основных показателей их деятельности можно заключить, что выбранные для составления уравнения регрессии показатели существенно влияют на прибыль банков, так как высокий коэффициент корреляции между показателями и прибылью остался неизменным (0,98). Данный вывод подтверждает устойчивость банковской системы в России. Ранее также применялся регрессионный анализ для финансового планирования [6]. Следует из результатов анализа, проведенного ранее [3] и сейчас, несмотря на устойчивость банковской системы, наблюдается высокая волатильность основных параметров этой системы в динамическом развитии. С другой стороны, полученные результаты подчеркивают ведущую роль стабильности финансового сектора в целом в устойчивости банковской системы. Доверие общественности к конкретному банку невозможно без предварительного улучшения финансовой стабильности всего банковского сектора.

Машинное обучение – перспективное направление в анализе данных. Его преимущество по сравнению с регрессионными моделями состоит в том, что можно использовать для обработки данные при наличии пропусков. Важно только для машинного обучения, чтобы целевая функция, заданная таблично (группирующий признак) не имела пропусков. Использование методов машинного обучения позволяет выявить важные показатели, характеризующие деятельность того или иного банка, которые в момент обработки данных, оказались нехарактерными для деятельности банка (выбросами, возможно), что приводит к неверным выводам и наводит на мысль, что одного статистического анализа недостаточно, необходимы динамические, возможно, нелинейные, модели использовать. Отметим, что в работе [7] подчеркивается, что существует большая потребность в дополнительных исследованиях, анализе применения другими исследователями и аналитика методов машинного обучения, поскольку не было проведено тщательной оценки того, как машинное обучение использовалось в банковской сфере. Машинное обучение хорошо зарекомендовало себя при анализе финансовой системы предприятий, связи функционирования финансовой системы с реальной экономикой в конкретных регионах, в частности, в Арктическом регионе нашей страны [8].

В заключении, вслед за автором статьи [9] можно с уверенностью сказать, что технологии машинного обучения все больше покоряют банковский сектор. Не за горами то время, когда для любого клиента банка (физического или юридического лица) в открытом доступе будут доступны программы анализа привлекательности конкретного банка для финансовых взаимоотношений с ним исходя из прогноза деятельности банка, который возможно будет самым выполнить на основе технологий машинного обучения.

Литература

1. Broby D. Financial technology and the future of banking // Financial Innovation. – 2021.– P. 7–47.
2. Hasanul B., Rabiul A. Is digital financial inclusion good for bank stability and sustainable economic development? Evidence for emerging Asia // Asian Development Bank Institute – 2021.– № 1242.
3. Борисова Л.Р., Жукова Г.С. Исследование зависимости прибыли банков России от основных показателей методами машинного обучения // Самоуправление. 2021. № 6. – С. 180-183.
4. Жилан О.Д. Особенности оценки ликвидности банка на основе обязательных нормативов// Baikal Research Journal. – 2019. – Vol.9 – N. 3 – P. 1–9.
5. Senko OV, Kuznetsova A.V. A recognition method based on collective decision-making using systems of regularities of various types // Pattern Recogn. Image Anal. –2010. –Vol. 20. –N2. –P.152–162.

6. *Kuzmenko O., Kyrkach S.* The use of regression analysis in the financial planning of banks, mathematical formalization of the stages of financial planning in banks // *Banks and bank system.* –2014. –Vol.9. –N2. – P.120–126.
7. *Hiran Kumar M., Megavarshini G., Sreenivasan A., Suresh M.* Machine Learning in the Banking Sector // *Proceedings of the International Conference on Industrial Engineering and Operations Management.* Manila, Philippines, March 7-9, 2023.
8. *Borisova L.R., Kuznetsova A.v., Senko O.V., Sergeva N.V.* Comparison of Arctic zone RF companies with different Polar Index ratings by economic criteria with the help of machine learning tools // *Computer Research and Modelling.* –2020. –Vol.12. –N1. – P.201–205.
9. *Donepudi P.* Machine Learning and Artificial Intelligence in Banking // *Engineering International.* – 2017. – Vol.5. – N.2. –P.83-86.