

Скоринговые модели могут использоваться банками не только при оценке заемщиков, но и для проверки контрагентов. При этом серьезным препятствием для применения полноценных скоринговых моделей до сих пор является необходимость своевременного доступа к обновляющейся в режиме онлайн информации. Набор и качество информационных источников в этом году существенно увеличились. Как банки могут использовать эти источники для автоматизации процесса принятия управленческих решений?

Скоринг для обеспечения банковской безопасности: как использовать новые источники данных?

Индексы платежной дисциплины: Paydex и кредитные истории

Первый из новых инструментов — индекс платежной дисциплины, так называемый Paydex (от payment index). Он рассчитывается в системе СПАРК на основании методики D&B («Дан энд Бредстрит») и собираемых в России данных, которые уже охватили 0,5 млн компаний, присваивая компании скоринговый балл на основании того, как она оплачивает свои счета.

Конечно, качество обслуживания долга компанией можно узнать из ее кредитной истории, но важно не только то, как она обслуживает кредиты, но и то, как она платит за электричество, телефон, газ, прочие товары и услуги, которые потребляет.

Откуда можно получить эту информацию? В каждой стране — участнице глобального проекта индексов платежной дисциплины формируются «клубы платежной дисциплины». В них вступают компании, которые сообщают о том, как их контрагенты выполняют принятые на себя финансовые обязательства.

Как показывают последние исследования, и в России, и за рубежом этот индекс обладает отличной предсказательной силой. Но нужно сразу сказать, что индексы платежной дисциплины — не единственное, что следует анализировать. Например, в Германии и Японии failure



Илья МУНЕРМАН,
*Исследовательский
центр Интерфакс-ЛАБ,
директор*

Илья МУНЕРМАН

score (скоринг, который предсказывает вероятность дефолта (PD)) включает два важнейших фактора: Paydex и кредитную историю. Российский аналог failure score — индекс финансового риска (ИФР), рассчитываемый в системе СПАРК-Интерфакс. Вы видите, как компания платит банкам и как она платит остальным контрагентам. На основании этого простой логистической регрессией выводите вероятность дефолта — и все, скоринг готов. Но, к сожалению, в российских условиях требуется проделывать больше работы, потому что немногие компании соглашаются быть участниками индексов платежной дисциплины.

Что же касается кредитных историй, то у этого массива данных есть несколько особенностей:

- во-первых, они не всегда достоверны: часто из-за ошибок проблемных банков в кредитные истории заносится недостоверная информация, выявление и исправление которой требуют определенных навыков и ресурсов;

- во-вторых, банк не может обрабатывать кредитную историю без согласия субъекта кредитной истории, а это согласие не всегда удается получить. Из хороших новостей нужно отметить следующее: законодательство сегодня позволяет работать с кредитными историями не только банкам, но и обычным компаниям при наличии согласия субъекта; бюро кредитных историй могут публиковать агрегированные данные и проверять модели сторонних разработчиков, сообщая статистические характеристики модели, но не раскрывая данные, на основе которых модель тестировалась. Все это способствует либерализации рынка скоринговых моделей для бизнеса.

Сегодня бюро кредитных историй могут публиковать агрегированные данные и проверять модели сторонних разработчиков, сообщая статистические характеристики модели.

Судебные решения и исполнительные производства, исковая нагрузка

Еще один важный источник информации для построения современных скорингов — судебные решения и информация судебных приставов. Анализируя судебные решения, можно получить огромное количество информации для включения в скоринговые модели. Вы знаете категорию дела, то есть знаете, что именно оспаривается, знаете сумму, которая оспаривается, и знаете, на какой стадии находится судебный процесс. Сегодня лингвисты с помощью экономистов распознают судебные дела и добавляют к карточке судебного дела большое количество формализованной информации. Именно смысловой процессинг судебных решений и экстракция из них ценной информации позволяют рассчитать важнейший показатель — исковую нагрузку компании. Анализируя предмет судебного спора, его участников, их историю по подобным спорам, историю решений

Скоринг для обеспечения банковской безопасности: как использовать новые источники данных?

данного суда и судьи и т.д., вполне можно спрогнозировать вероятность выигрыша или проигрыша, срок процесса и возможные активы или обязательства, которые появятся у участников процесса в результате вынесения судебного решения.

Что же касается информации судебных приставов об исполнительных производствах, то их очень много. Однако не имеет смысла анализировать все типы исполнительных производств, вполне достаточно сконцентрироваться на основных: это штрафы ГИБДД, налоги, зарплата, кредиты и т.д. Очевидно, что правильная интерпретация этих данных очень важна. Так, задолженность по зарплате — это существенный негатив, а незначительная задолженность по штрафам ГИБДД — скорее позитив, чем негатив, так как свидетельствует об активном ведении бизнеса.

Данные о персонале

Из массива данных, раскрытых ФНС России в сентябре этого года, банкам стала доступна численность персонала. Этот показатель важен с самых разных точек зрения. Минимальная и нулевая численность вызывает подозрения в добросовестности компании, значительная численность не всегда является эффективным показателем, так как, например, холдинговая структура распределяет персонал между различными юридическими лицами, что не всегда дает корректную оценку реальной производительности труда. Часто данные о персонале искажаются за счет кэптивных компаний и дружественных индивидуальных предпринимателей.

Тем не менее, некоторая информация все же стала доступнее, чем раньше, когда банки экспертным путем определяли, сколько в компании работает людей, и позволяет рассчитывать много новых коэффициентов, в том числе связанных с производительностью труда.

Второй показатель — это, конечно, вакансии. Динамика вакансий, новые вакансии, старые вакансии, зарплата, которая предлагается в этих вакансиях, — это тоже интересно, и, конечно, недобросовестный контрагент или банкрот вряд ли будет нанимать персонал на HeadHunter. Во многих развитых странах, участвующих в системах глобальных скорингов Labor Data, это сегодня один из важных инструментов прогнозирования событий, которые будут происходить с компанией.

Данные о закупках

Крупные компании так или иначе вынуждены проводить тендерные закупки, в том числе на электронных торговых площадках. Отсюда следствие: огромное количество информации о закупке товаров

Задолженность контрагента по зарплате — это существенный негатив, а незначительная задолженность по штрафам ГИБДД — скорее позитив, чем негатив, так как свидетельствует об активном ведении бизнеса.

Илья МУНЕРМАН

и услуг на электронных торговых площадках доступно из публичных источников информации. Сегодня агрегаторы электронных торговых площадок предоставляют возможность анализировать множество данных: кто участвовал, с кем участвовал, кто выиграл, что поставил, по какой цене, остался ли потребитель доволен и т.д.

И на самом деле, если вы видите, что компания неоднократно выигрывала тендеры у некоего крупного клиента с отличным финансовым состоянием, потом перестала выигрывать эти тендеры, а одновременно появился судебный иск данного крупного потребителя к компании, то ваша скоринговая модель должна отреагировать на это в режиме реального времени. Уже сегодня есть массивы данных, которые можно для этого использовать, и технологические решения, позволяющие все это делать.

Среди задач, решаемых в сегменте анализа данных о торгах, есть также оценка и анализ стоимости закупленных объектов, в том числе с вычислением агрегированных удельных показателей этой стоимости, например стоимости одного километра дороги или стоимости школы в расчете на одного учащегося. Интересны также частные скоринги поставщиков и крупных потребителей, которые оценивают прозрачность, комфортность данного контрагента, частоту судебных дел, которые с ним связаны, и т.д.

Большой интерес представляют и рекомендательные B2B-системы. Ведь если в интернет-ритейле данный инструмент достаточно понятен, то в B2B такие системы пока редкость. Причем для бизнеса данные системы не только обеспечивают комфорт при работе закупщика или подрядчика, но и исполняют контрольную функцию: если директор по закупкам регулярно игнорирует предложения рекомендательной системы, то вполне возможно, что он находится в сговоре с поставщиками.

Данные о недвижимости

Нужно отметить, что в России очень плохая ситуация с данными о недвижимости ввиду проблем с кадастром недвижимости. Здесь наша страна отстает даже от соседей по Восточной Европе, очень сильно отстает от Эстонии, Польши, Финляндии, Португалии и многих других развитых стран, в которых кадастр и кадастровая стоимость недвижимости рассчитаны достаточно точно, что позволяет строить более эффективную налоговую политику. Помимо того, многие страны переводят кадастр на блокчейн, чем Россия пока точно не может похвастаться.

Тем не менее, даже в условиях слабости государственных сервисов информирования о недвижимости такие сервисы, как Google, Яндекс,

Даже в условиях слабости государственных сервисов информирования о недвижимости такие сервисы, как Google, Яндекс, 2ГИС, предоставляют очень много информации, которая раньше была недоступна.

Скоринг для обеспечения банковской безопасности: как использовать новые источники данных?

2ГИС, предоставляют очень много информации, которая раньше была недоступна. Нужно отметить две существенные проблемы. Первая — отсутствие данных, которые бы позволили точно понять, какая недвижимость находится на балансе конкретного собственника, хотя бы юридического лица. Вторая — недостаток конкуренции на рынке решений по автоматизированной оценке недвижимости как для целей государственной кадастровой оценки, так и для автоматизированной оценки залога в режиме реального времени.

Интернет-данные

Если у компании есть сайт, на котором она что-то продает, то вы уже знаете о ней очень многое: посещаемость сайта, средний чек и пр. Для этого есть много различных сервисов, как дешевых, так и дорогих. Для представителей данного сегмента бизнеса скоринговые модели подстраиваются под дополнительное факторное пространство этих компаний.

Новости и социальные сети

Еще недавно считалось, что этот источник информации очень ценен для финансового анализа, что репутация компании не менее значима, чем ее финансовые показатели. Несмотря на важность этого тезиса, необходимость лингвистического анализа большого количества художественных текстов создает проблемы с вовлечением этих источников в семейство скоринговых моделей. Кроме того, зачастую журналистский подход к новости не совпадает с подходом финансового аналитика: журналист считает, что новость, которую никто не перепечатал, просто неинтересная, а вот аналитик думает по-другому.

Сведения о товарных знаках и договорах на их использование полезны для анализа стоимости брендов не только в крупном, но и в малом бизнесе.

Даркнет

Для любого скоринга, который связан с безопасностью, даркнет является важнейшим источником данных, так как именно там можно узнать, кто и как пытается его взломать. Сейчас появились стартапы, которые позволяют искать информацию в даркнете, не используя специальные средства доступа к нему, что облегчает жизнь специалистам по безопасности.

Объекты интеллектуальной собственности

Сведения о товарных знаках и договорах на их использование полезны для анализа стоимости брендов не только в крупном, но и в малом бизнесе. В частности, можно привести такой пример из практики нашей компании.

Илья МУНЕРМАН

Пример 1 Анализ ИФР для франчайзи

В соответствии с мировой практикой отличным решением для человека, который по тем или иным причинам надумал открыть малый бизнес, но не хочет сильно рисковать, является покупка франшизы. И вот здесь начинающего предпринимателя подстерегают опасности. Во многих странах мира, например в США, деятельность по продаже франшиз населению приравнена к работе со сбережениями физлиц и требует сдачи значительного объема отчетности, на основе которой вычисляется коэффициент «выживаемости» сети и оценивается средняя доходность входящих в нее франчайзи. К сожалению, в России потенциальный франчайзи часто получает недостоверную информацию по франчайзинговой сети и, как результат, теряет свои вложения.

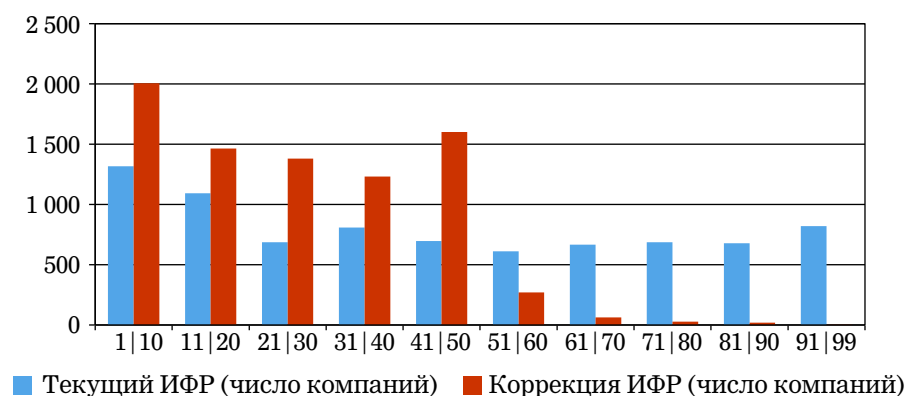
Видя такую плачевную ситуацию, мы решили создать в СПАРК подсистему корректировки индексов должной осмотрительности (ИДО) и финансового риска (ИФР) для распознанных нами франчайзинговых сетей. Так, нами был разработан алгоритм машинного обучения распознавания франчайзинговых сетей.

Сейчас нам удалось распознать около 300 сетей, объединяющих более 10 тыс. франчайзи, но работа продолжается, и мы планируем еще больше сетей и франчайзи.

Попадание компании в список франчайзи известной сети приводит к тому, что ИФР для таких компаний существенно корректируется (рис. 1). Это неудивительно, ведь на самом деле принадлежность

Рисунок 1

Изменение ИФР для франчайзи



Скоринг для обеспечения банковской безопасности: как использовать новые источники данных?

к успешной франшизе экстраполирует опыт успешных франчайзи данной сети на вновь пришедшую компанию.

Транзакции и данные онлайн-касс

Относительно новым источником информации о компании стали ее транзакции: данные можно получить, анализируя движения по счету и динамику показателей онлайн-кассы. Несмотря на огромный потенциал этих источников как с точки зрения анализа финансового состояния, так и с точки зрения безопасности и комплаенса, их использование упирается в серьезные ограничения — юридические аспекты, связанные с защитой банковской и коммерческой тайны.

Многие крупные компании создают своих собственных операторов фискальных данных (ОФД), чтобы избежать утечек к конкурентам. Тем не менее, для работы с такими данными есть несколько подходов. Первый — это использование агрегированных показателей, второй — искусственное зашумление данных и обучение модели на них, третий, как говорилось выше про кредитные истории, — использование этих данных для оценки скорингов и проверки гипотез, не сопровождаемое их раскрытием разработчикам или обменом данными с конкурентами.

Анализ налогового бремени и производительности

Этот массив данных также относится к данным ФНС России. Теперь можно получить информацию о спецрежимах и налоговом бремени компании, причем как общем, так и частном по отдельным налогам. Нужно сказать, что это революционное нововведение принесло банкам возможность рассчитывать множество относительных показателей, например отклонение совокупного или частного налогового бремени от среднеотраслевого или среднерегионального уровня. Этот показатель можно нормировать на рубль выручки или одного работника. Через зарплатные налоги и отчисления в фонды можно рассчитать средний уровень зарплаты через численность и сравнить его опять же со среднеотраслевым и региональным уровнями.

Тем не менее, поскольку сервис функционирует с сентября, нужно относиться к нему с осторожностью, так как наш анализ показывает, что данные на нем обладают огромной дисперсией, не позволяющей получать всегда достоверные оценки. Эти факторы очень важны, но только в совокупности с другими показателями.

Илья МУНЕРМАН

Традиционные финансовые коэффициенты онлайн

Сегодня повсеместное использование бухгалтерских онлайн-сервисов открыло возможность вернуть к жизни традиционные финансовые коэффициенты, от которых отказались из-за огромного лага от текущей даты до последней финансовой отчетности. Тем не менее, современная тенденция все-таки состоит не в ручном анализе коэффициентов, а в использовании их как исходных данных для обучающейся в режиме реального времени скоринговой системы. Но благодаря бухгалтерским онлайн-сервисам скоринговая система увидит коэффициенты в текущем времени, а не на дату последней отчетности. Настройка этого займет определенное время и силы, но расчеты ИФР значительно улучшатся.

В заключение хотелось бы привести два примера использования всех перечисленных выше данных.

Сегодня повсеместное использование бухгалтерских онлайн-сервисов открыло возможность вернуть к жизни традиционные финансовые коэффициенты, от которых отказались из-за огромного лага от текущей даты до последней финансовой отчетности.

Пример 2

Динамика индекса должной осмотрительности (ИДО)

Как показывает анализ динамики ИДО, за последние 10 лет количество недобросовестных компаний снизилось с 1,5 млн до 700 тыс., однако структура данных организаций существенно изменилась. Если раньше большинство из них было направлено на создание «однодневки», предназначенных для ухода от налога, то сегодня общее количество недобросовестных компаний, предназначенных для этих целей, сократилось до 200 тыс., по данным ФНС России. К сожалению, приходится констатировать, что активность мошенников, которая раньше была направлена против государства, теперь угрожает малому бизнесу и банкам.

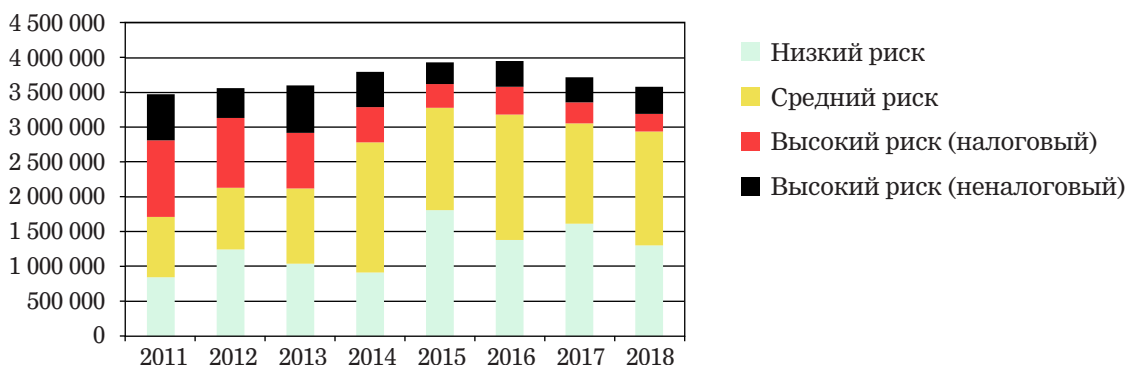
Мы теперь делим группу компаний с высоким риском на две подгруппы (рис. 2): во-первых, компании, которые потенциально могут использоваться для налоговых нарушений, во-вторых, потенциально мошеннические, неблагонадежные компании. В частности, распространено мошенничество с клонами известных компаний, их «представителями», которые в обмен на нереально большую скидку получают от малого предпринимателя аванс и скрываются с ним.

Многие наши клиенты жалуются на новые виды мошенничества, не связанные с «обналичкой». Таким образом, анализ значительного числа факторов для выявления недобросовестных компаний остается актуальной задачей, и проявлять бдительность и осторожность требуется по-прежнему.

Скоринг для обеспечения банковской безопасности: как использовать новые источники данных?

Рисунок 2

Распределение российских компаний по уровням риска



Пример 3

Особенности мониторинга дилерской сети

Очень востребованной услугой среди производителей и продавцов различных товаров сегодня является мониторинг дилерской сети. Эта новая услуга состоит в обогащении данных, которые используются в индексе финансового риска, дополнительными данными самого производителя. В частности, можно дополнительно учесть историю поставок и своевременность оплат — этот источник данных не требует согласования с дилером. Можно обязать его регулярно предоставлять отчетность и рассчитывать коэффициенты на актуальные даты. Наконец, можно получить согласие на обработку кредитной истории как от самого дилера, так и от его учредителей и директоров.

Данная услуга актуальна для компаний, которые хотят кредитовать своих дилеров, не прибегая при этом к услугам факторинговых компаний или банков. Самостоятельный расчет кредитного лимита и его мониторинг в режиме реального времени привели к тому, что компании могут экономить деньги на услугах финансовых посредников и развивать свою дилерскую сеть самостоятельно.

Мы рассказали только об основных новых источниках данных и привели всего три примера их применения, однако регулярная практика и прикладные научные исследования постоянно открывают десятки возможностей увеличить эффективность бизнеса за счет новых данных. Использование этих возможностей является для большинства банков не просто необходимостью, но зачастую и вопросом выживания. 