
Ольга РИНК

Набор фактов, позволяющих идентифицировать партнера и корректно описать его текущую деятельность, — не единственный результат применения промышленных риск-решений. Другим результатом является так называемая постправда — новые смыслы, извлеченные из известных фактов. Как анализ неструктурированных текстовых массивов повышает точность индивидуальных классификаторов видов деятельности контрагентов? Что поможет оценить состояние контрагента в случае недоступности или ограничений на передачу финансовой отчетности?

Проверка зарубежных контрагентов: от правды к постправде



Ольга РИНК,
Interfax — Dun & Bradstreet, генеральный директор

Кредитные решения, продажи и закупки все чаще обосновываются расчетными показателями, риск-индикаторами, в основе моделей которых лежит инструментарий искусственного интеллекта (Artificial Intelligence, AI) и математической статистики. Экспертная составляющая скоринговых оценок постепенно замещается промышленными риск-решениями, построенными на алгоритмах обучения нейронных сетей, анализе неструктурированных текстовых массивов (Natural Language Processing, NLP), и подобными.

Данные о бизнесах в эпоху постправды

Аппарат аналитики данных (data science) и искусственного интеллекта, предназначенных для обнаружения смыслов в цифровых массивах и построения поведенческих моделей и сценариев, известен уже не один десяток лет. Однако в последнее десятилетие подходы AI заставили о себе говорить не только узкую группу ученых — специалистов по информационным технологиям, но и политиков, бизнесменов, журналистов, то есть практически всех.

Компания Deloitte признала искусственный интеллект одним из главных технических трендов 2017 г.¹ «Когнитивные вычисления,

¹ <https://www2.deloitte.com/ru/ru/pages/technology-media-and-telecommunications/articles/2017/tmt-predictions.html>.

Проверка зарубежных контрагентов: от правды к постправде

которые помогают делать выводы и прогнозы» означают «начало новой когнитивной эпохи», считают специалисты.

Доступность «больших данных» (big data), инструментов их обработки, а также появление наряду с привычной электронной почтой таких каналов прямой коммуникации, как социальные сети, мессенджеры, заметно перевернули деловую среду, социологию и многие другие сферы.

«Постправда» (post-truth) была признана в Оксфордском словаре в 2016 г. «словом года». Этот термин получил много трактовок, но по сути речь идет об обнаружении скрытых смыслов: не то чтобы их, «новых» смыслов, не существовало ранее, но появились техники их вскрытия и практического применения.

Громкими примерами стали кампании 2016 г. в поддержку Brexit (Великобритания) и республиканцев на выборах (США). За обеими стояли исследования значительных массивов социологических сведений и поведенческих характеристик, преимущественно взятых из шлюзов социальных сетей. Их результатом явилась более тонкая и подробная сегментация предпочтений аудитории и, как следствие, более точно направленный «месседжинг», различные стратегии и формулировки прямых сообщений для целевых групп.

Artificial Intelligence меняет и сферу информации о бизнесах. Доступность промышленных решений (включая математическую статистику) для разработки моделей кредитных оценок на достаточно полных базах записей о деятельности компаний позволяет все с более высокой точностью прогнозировать риски, генерировать лиды и сегментировать контрагентов для уточнения бизнес-моделей.

Конечной целью интеграторов информации стали не только расширение количества записей и более полные массивы сигналов и кредитных триггеров, которые бы позволили получить свежий отчет о состоянии бизнеса. Мы можем условно отнести эти характеристики к категории «правда» как набор фактов, позволяющих идентифицировать партнера и корректно описать его текущую деятельность.

Возможность выхода на зарубежные рынки в сочетании с применением аппарата big data изменили масштаб, важность и сложность задач по идентификации зарубежных контрагентов, управлению кредитными рисками, дебиторской задолженностью, ликвидностью, репутацией. Появились новые темы: финмониторинг, FATCA (Foreign Accounts Tax Compliance Act — Закон «О выполнении налоговых требований по зарубежным счетам»)¹, CRS (Common Reporting Stan-

Возможность выхода на зарубежные рынки в сочетании с применением аппарата big data изменили масштаб, важность и сложность задач по идентификации зарубежных контрагентов, управлению кредитными рисками, дебиторской задолженностью, ликвидностью, репутацией.

¹ <https://www.irs.gov/businesses/corporations/foreign-account-tax-compliance-act-fatca>.

Ольга РИНК

dard — стандарт международного обмена налоговой информацией)¹, проверка бенефициаров, комплаенс и санкционные списки, включая отчеты «Знай своего клиента» (Know Your Customer, KYC), «Знай своего поставщика» (Know Your Vendor, KYV), и не только.

На основе имеющихся сведений, а также анализа тенденций и закономерностей можно оценить, например, вероятность дефолта (банкротства и прекращения деятельности) на горизонте в 12–18 месяцев (в контексте статьи — «постправда», категория, включающая результаты применения AI и статистики, новые смыслы, извлеченные из известных фактов).

Конечная цель этих проектов — повышение доходности компании и снижение ее рисков, потерь; именно поэтому технологичность принятия решений в отношении условий сотрудничества на основании исходных данных и показателей рисков так важна.

Пример первый: настраиваемые классификаторы видов деятельности

Возвращаясь к сюжету о когнитивной эпохе: более 170 лет назад бизнес Льюиса Тэпана, основателя Dun & Bradstreet, начинался с кредитных отчетов, в которых оценки делались скорее исходя из репутации проверяемой структуры — мнения о порядочности и честности владельцев и руководителей компаний закладывались в основу выводов юристов-корреспондентов, которые готовили отчеты. В одном из значений credit переводится с английского языка как «репутация», поэтому, возможно, и этимология этих терминов, сути проверок могла быть близкой.

Сегодня медиапокрытие по-прежнему учитывается, в том числе и для расчета индексов деловой репутации, например в системе SCAN группы «Интерфакс», где активно используется NLP.

Однако обработка обширных текстовых массивов встречается и в других типах задач. Одним из примеров может стать создание индивидуальных справочников видов деятельности, которые используются, в частности, интеграторами маркетинговых сведений о госзакупках или потенциальных потребителях.

Каждая юрисдикция оперирует своим классификатором. Например, в Российской Федерации это ОКВЭД 2.0, в США — SIC (ISO-стандарт) или NICE-коды (американо-канадская система). У них различные принципы, структуры и степени детализации.

Когда мы хотим выбрать потенциальных покупателей или поставщиков по списку факторов в целевой нише — допустим, тепличные

¹ <http://www.oecd.org/tax/automatic-exchange/common-reporting-standard/>

Проверка зарубежных контрагентов: от правды к постправде

хозяйства Южной Кореи, Саудовской Аравии и Ирана, мы вполне можем ориентироваться на SIC как ISO-стандарт. Но иногда требуется более точный анализ — скажем, не просто производители программного обеспечения, а специализация на виртуальной реальности. В таком случае потребуется дополнительная структура кодов. Например, компания Dun & Bradstreet в некоторых странах использует NLP для создания параллельных, отличных от официальных, классификаторов, которые позволяют с точностью выше 99% выбрать нужных партнеров.

Пример второй: глобальный рэнкинг бизнесов

Исторически скоринговые показатели формировались исходя из классических подходов финансового анализа, на основании преимущественно коэффициентов отчетности в сочетании с некоторыми качественными характеристиками. К таким оценкам можно отнести индексы финансового риска и устойчивости бизнеса, рекомендованный кредитный лимит.

Наряду с традиционными подходами, предлагающими оценки текущего состояния, развивается и прогнозная аналитика (Predictive Analytics).

Почти десять лет назад G20 проанализировала¹ причины глобального финансового кризиса и среди мер управления рисками выделила рекомендации по совершенствованию техник сбора информации о бизнесах. В документах особое внимание уделено корректной идентификации компаний, а также корпоративных связей — с целью определения и мониторинга риска на контрагента и группы компаний.

Наднациональными органами была поставлена задача оценки рисков корпоративных структур на разумное время вперед, что послужило стимулом для разработки масштабных (в объеме всех компаний страны) статистических моделей прогноза вероятности дефолта бизнесов на горизонте год-полтора (Failure Score).

В 2015 г. аналитическое подразделение Dun & Bradstreet рассчитало вероятность дефолта и прекращения деятельности бизнесов на всей глобальной базе данных, охватывающей две сотни стран и более 275 миллионов записей. Полученный индекс получил название «Глобальный бизнес-рэнкинг» (Global Business Ranking, GBR). Структура факторов включает, наряду с прочим, блок макроэкономики, что позволяет сравнивать риски контрагентов в разных юрис-

¹ [http://ec.europa.eu/eurostat/statistics-explained/index.php/G20_Data_Gaps_Initiative_\(DGI\)_%E2%80%93%93_background](http://ec.europa.eu/eurostat/statistics-explained/index.php/G20_Data_Gaps_Initiative_(DGI)_%E2%80%93%93_background).

Ольга РИНК

дикциях, а также блоки собственных сведений о бизнесах (платежная дисциплина и др.).

Показатель GBR позволяет оценить состояние контрагента, например, в случае недоступности или ограничений на передачу финансовой отчетности, отсутствия моделей вероятности дефолта по конкретной стране.

Global Business Ranking используется крупными международными корпорациями: структурами рынка информационных технологий, ритейлом, финансово-промышленными группами и многими другими компаниями с географически диверсифицированной клиентской базой.

Пример третий: лидогенерация и сегментация клиентов

Риск-оценки применяются сегодня не только для целей управления кредитными рисками, управления портфелем дебиторской задолженности. Риск — важный индикатор маркетинговых листов для лидогенерации, отбора приоритетных целевых групп. Низкий риск — косвенный сигнал о возможном высоком приоритете потенциального или существующего сотрудничества. Разные интервалы рисков могут соответствовать своим бизнес-стратегиям, отличаться по условиям. Ну и, наконец, зачем тратить ресурсы на компании в «красной зоне»?

Кастомизированные (индивидуальные) прогнозные модели, позволяющие сегментировать внутреннюю клиентскую базу, сегодня часто создаются в сотрудничестве коммерческими структурами и поставщиками промышленных скоринговых решений и данных.

Тезис «Данные — основа отношений» (в английском варианте он звучит чуть изящнее — Growing relationships through data) в полной мере применим к кредитной проверке бизнесов. И сегодня задачи управления кредитными рисками, комплаенса и выявления бенефициаров, расчета вероятности мошенничества, определения самых выгодных групп клиентов или самых рискованных поставщиков могут быть решены с использованием платформ нового поколения. 