

ОЦЕНКА ВЕРОЯТНОСТИ ЗАКРЫТИЯ КОММЕРЧЕСКИХ БАНКОВ С ПРИМЕНЕНИЕМ БАЙЕСОВСКОГО ПОДХОДА

Багаев И.В., Васенина М.С., Кудрявцев П.А.

Вятский государственный университет,
Институт компьютерных и физико-математических наук,
кафедра математического моделирования.

Россия, 610000, г.Киров, ул.Московская, д.36

E-mail: ivan.bagaev1993@gmail.com, vasenina_masha@mail.ru, pas6693@yandex.ru

По данным интернет-портала Банки.ру¹ с 1991 года по 2016 год более двух с половиной тысяч коммерческих банков прекратило свою деятельность вследствие ликвидации или отзыва лицензии. На конец октября 2016 года свою деятельность продолжают 624 банка, давших согласие на раскрытие информации. Отчетность всех организаций на все отчетные даты находится в открытом доступе на сайте Центрального Банка и может быть загружена любым желающим для анализа.

Используя два вышеупомянутых ресурса, была сформирована выборка, содержащая данные о всех действующих и закрытых коммерческих банках, которые включают в себя дату регистрации, объем уставного капитала, место расположения головного офиса, а также данные финансовой отчетности - оборотная ведомость по счетам бухгалтерского учета, отчет о финансовых результатах, расчет собственных средств (капитала) и информация об обязательных нормативах. Для закрывшихся банков - дата и причина закрытия. Большинство данных или ежемесячные, или ежеквартальные. Очевидно, что весь объем информации, содержащейся в отчетных формах, - избыточен, поэтому основной проблемой, как и в любой прикладной задаче анализа данных, является формирование наиболее релевантного набора объясняющих переменных.

Для оценки вероятности закрытия использовались стандартные подходы машинного обучения. Были сформированы два набора данных - обучающий и тестовый. Первый включает в себя информацию о банках, закрытых с 2013 по 2015 год, второй, необходимый для проверки качества модели, - аналогичные данные за 2016 год. Для закрывшихся организаций в выборку включаются последние имеющиеся отчетные данные. Особенностью является то, что действующие банки могут быть включены и в обучающий и в тестовый набор, используя при этом данные на разные отчетные даты.

В качестве базовой модели использовалась байесовская логистическая регрессия. Байесовский вывод в настоящее время представляет собой одно из наиболее перспективных направлений в статистике и анализе данных. Одним из его преимуществ для данной задачи является то, что на выходе модели имеются не конкретные значения искомых параметров, а некоторое их распределение, что позволяет более эффективно оценивать неопределенность относительно конечного результата.

¹ <http://www.banki.ru/banks/memory/>