

Отчеты об операциях фондовых отделений контор Государственного банка как источник по истории регулирования фондового рынка Российской империи в начале XX в.: кластеризация и OCR-распознавание табличных данных

Научный руководитель – Бородкин Леонид Иосифович

Галушко Илья Николаевич

Аспирант

Московский государственный университет имени М.В.Ломоносова, Исторический факультет, Москва, Россия

E-mail: galushko.iolk@yandex.ru

История системы регулирования фондового рынка Российской империи интересна как призма, через которую мы можем детально проследить, как государство реагировало на вызовы времени, создавая правовые и институциональные рамки для функционирования рынка ценных бумаг. Совершенствование биржевого законодательства, регулирование деятельности акционерных обществ, контроль за эмиссией ценных бумаг, управление предложением облигаций государственного займа – все эти сложные и масштабные задачи были направлены на защиту интересов инвесторов и поддержание стабильности финансовой системы; на упрочение образа Российской империи как надежного заёмщика и выгодного экономического партнера.

Кроме того, изучение системы регулирования фондового рынка помогает понять, как формировались отношения между государством, предпринимательскими слоями и обществом в условиях бурно протекающего процесса модернизации [1]. Переход к диверсифицированной и капиталоемкой экономике, ограниченная демократизация доступа к фондовому рынку через распространение онкольного кредита [2], усиливающаяся власть прессы над умами инвестирующей и спекулирующей публики – к началу XX века жизнь фондового рынка Российской империи становится крайне многоаспектным процессом, вовлекающим всё больше акторов и денежных ресурсов; и затрагивающим всё большие области государственных интересов. В представленном докладе рассматривается специализированная задача Государственного банка по управлению ликвидностью облигаций государственного займа. Как показывают материалы отчетов фондового отделения Московской конторы Государственного банка, в зависимости от потребностей бюджета Государственный банк мог стимулировать или наоборот понижать спрос населения на бумаги государственных займов через систему скидок и надбавок к рыночной цене при продаже и покупке облигаций в своих отделениях [3]. Особый интерес вызывают статистические приложения к годовым отчетам фондового отделения Государственного банка за период с 1895 по 1913 года. Сводные таблицы содержат информацию о клиентской активности по покупке и продаже облигаций государственных займов во всех конторах Государственного банка с разбивкой по уездам. Поскольку одной из ключевых задач Государственного банка было поддержание ликвидности и курса правительственных займов, представленный материал служит ценным источником информации как об особенностях политики регулирования рынка государственных облигаций, так и о специфике обращения этих бумаг среди населения. Это очень важный пласт отечественной финансовой истории, так как успешное размещение займов было необходимо для поддержания бюджета и стабильности эмиссионного права. Кроме того, в портфелях населения и региональных кредитных обществ государственные займы служили наиболее распространенным активом для хранения сбережений.

Объем имеющихся данных об операциях фондовых отделений контор Государственного банка позволяет нам применить в исследовании метод кластеризации. Мы предполагаем, что составление и анализ устойчивого набора кластеров позволит глубже понять особенности функционирования региональных экономик Российской империи. В докладе рассмотрены ключевые аспекты построения модели: подготовка датасета и отбор признаков (feature_engineering), выбор алгоритма кластеризации (DBSCAN или K-means), предобработка данных для нивелирования эффекта различий в объемах операций на уровне губерний (стандартизация, или удаление среднего и масштабирование дисперсии; этот шаг необходим, так как прямое сравнение столичных районов и, например, сибирских контор приведет к искажению обучения модели). Отдельной частью доклада является сравнение различных методов оптического распознавания таблиц дореволюционной печати. Мы проанализировали эффективность доступных сегодня методов распознавания сканов дореволюционных таблиц. Были рассмотрены: FineReader, Transkribus, Google Tesseract с дообучением на датасете «Russian old orthography OCR» [4], промтинг LLM (Gigachat, Deepseek, YandexGPT).

Помимо решения сугубо исторических задач, представленное исследование способно включиться в ряд больших дискуссий об эволюции и функционировании регулирующих органов в сфере фондовой торговли. Рынок ценных бумаг Российской империи – это крайне интересный полигон с немалым количеством данных, которые сегодня постепенно вводятся в научный оборот. Другой важной целью доклада видится дальнейшая апробация методов машинного обучения и искусственного интеллекта в исторических исследованиях. И если кластерный анализ уже давно закрепился в качестве продуктивного исследовательского инструмента, то оптическое распознавание исторических источников только входит в практику работы историков. Как показал наш опыт, эта технология может давать ценные результаты и существенно ускорять обработку при использовании хорошо обученных моделей.

Источники и литература

- 1) Бородкин Л.И. Концепции модернизации и модерности в контексте российских трансформаций XIX–XX вв. // Уральский Исторический Вестник, №4, 2017. С. 6-15.
- 2) Саломатина С.А. Комерческие банки в России: динамика и структура операций, 1864-1917. М.: РОССПЭН, 2004. С. 136-139.
- 3) ЦГАМ. Ф. 450. ОП. 2. Д. 101
- 4) URL: <https://huggingface.co/datasets/nevmenandr/russian-old-orthography-ocr>