

Фан Чэньси

Ростовский государственный экономический университет

Разработка и применение интеллектуального инструмента мониторинга и анализа расходов для системной интеграции на основе LSTM-оптимизации и динамической песочницы прогнозов

Аннотация. В данном исследовании построена многомерная модель прогнозирования, объединяющая макроэкономические показатели, и визуальная система поддержки принятия решений для нужд динамического мониторинга финансового менеджмента предприятия. Модель позволяет повысить точность прогнозирования на $\pm 7\%$ (снижение ошибок на 53% по сравнению с традиционной моделью) за счет внедрения байесовской LSTM-сети и разработки интерактивного механизма причинно-следственных связей для поддержки моделирования динамических сцен. Эмпирические данные показывают, что система может снизить стоимость скользящего кредита для производственных предприятий на 23%, а риск отставания товарных запасов в розничной торговле - на 3,2 млн юаней. Инновации исследования: 1. предложение мультимодальной системы оценки финансового здоровья на основе данных 2. создание гибкого алгоритма составления бюджета, интегрирующего анализ текста политики.

Ключевые слова: финансовый менеджмент, мониторинг доходов и расходов, разработка системы, построение системы

Fan Chensi

Rostov State University of Economics

Development and application of a system-integrated intelligent income and expenditure monitoring and analysis tool based on LSTM optimisation and dynamic sandbox projections

Abstract. In this study, a multi-dimensional prediction model and visual decision support system integrating macroeconomic indicators are constructed for the dynamic monitoring needs of enterprise financial management. It achieves $\pm 7\%$ prediction accuracy improvement (53% error reduction compared with traditional model) by introducing Bayesian LSTM network, and develops interactive causal inference engine to support dynamic scene simulation. Empirical evidence shows that the system can reduce the rolling loan cost of manufacturing enterprises by 23% and the risk of inventory backlog in retail industry by 3.2 million RMB. The research innovations are: 1. proposing a multimodal data-driven financial health assessment system 2. establishing a flexible budget derivation algorithm that integrates policy text analysis.

Keywords: income and expenditure monitoring, system development, system construction, financial management

Цель исследования: создание системы мониторинга доходов и расходов бизнеса на основе данных. Интеграция 12 типов источников данных, таких как корпоративные ERP-

системы, API-интерфейсы Национального бюро статистики, отраслевые документы и т. д., использование вычислительного движка реального времени Flink для построения конвейеров данных, создание родословной данных и проведение динамического корреляционного анализа между данными о доходах и расходах, темпами роста ВВП и индексом РРІ, позволит решить проблему «информационной оболочки» в традиционной финансовой системе. Проблемы.

Значимость исследования: В диссертации предлагается инновационная система оценки достоверности данных «точка-три-четыре-уровня», которая позволяет достичь значительных результатов в области управления данными коммерческого банка за счет интеграции анализа временных рядов, инженерии знаний и технологии машинного обучения. Система строит систему оценки на основе четырех измерений: полноты, согласованности, своевременности и точности, а конкретная техническая реализация выглядит следующим образом:

В измерении целостности для интеллектуального обнаружения отсутствующих паттернов временных данных используется модель Марковской цепи Монте-Карло (MCMC). Временные зависимости отсутствующих данных изображаются с помощью построения матрицы переноса состояний, а схема заполнения отсутствующих значений генерируется с помощью байесовской техники выборки апостериорного распределения. По сравнению с традиционным методом интерполяции, модель может эффективно справляться с немонокотонными пропущенными моделями и увеличить коэффициент полноты данных до 98,7 % за счет итеративной оптимизации I этапа (интерполяция) и P этапа (обновление апостериорного распределения). Для измерения согласованности на основе онтологии построена база финансовых знаний, содержащая 120 правил, семантические отношения между финансовыми показателями формально определены на языке OWL, а автоматический вывод и проверка осуществляются с помощью механизма правил SWRL. Онтология знаний охватывает 4 основных набора понятий, таких как платежеспособность, рентабельность и т. д., и устанавливает 9 типов семантических отношений, таких как SubConceptOf, HasPart и т. д., что обеспечивает логическую согласованность финансовых данных на уровне 99,2 %.

Эффективное измерение строит конвейер данных в реальном времени через очередь сообщений Kafka и использует модель производитель-потребитель для достижения обновления данных T+0. Разработана архитектура с двумя очередями для обработки транзакционных данных OLTP (пропускная способность 150 000 элементов/сек) и пакетных данных (задержка <50 мс) соответственно, в сочетании с кэширующим слоем Redis для снижения нагрузки на доступ к базе данных, что позволило увеличить индекс свежести данных на 41 %. Модуль обнаружения аномалий в изолированном лесу, разработанный Accurasy Dimensions, рассчитывает оценки аномалий на основе длины пути путем построения 1 000 деревьев решений для рекурсивного разбиения пространства признаков транзакций. Алгоритм выполняет бесконтрольное обучение 28-мерных признаков транзакций и обнаруживает 0,67% транзакций с высоким риском, при этом уровень ложных тревог снижается на 62% по сравнению с традиционными методами.

После внедрения системы в коммерческом банке индекс достоверности данных был улучшен с 0,72 до 0,91 благодаря четырехмерной синергии, что показало следующее: время отклика на предупреждения об отсутствии кредитных данных сократилось на 83 %, цикл составления финансового отчета сократился на 57 %, а точность выявления рискованных операций в режиме реального времени достигла 92,4 %. Практика показывает, что система

оценки предоставляет количественные решения для управления финансовыми данными и имеет важное справочное значение для предотвращения системных рисков и повышения качества принимаемых решений.

Содержание исследования: Гибридная система прогнозирования, предложенная в данном исследовании (рис. 1), использует многоуровневую конструкцию, первый слой - слой разработки признаков, ключевые макроэкономические показатели отбираются алгоритмом XGBoost, который понижает уровень высокоразмерных данных, рассчитывая важность признаков (например, вклад прироста), и в итоге выбирает семь основных показателей, таких как волатильность ИПЦ, предложение M2, темпы роста добавленной стоимости промышленности, торговый баланс импорта и экспорта, темпы роста основных средств темпов роста инвестиций в основной капитал, индекса потребительского доверия и изменения процентной ставки. Процесс подтверждается бизнес-логикой, например, сильная корреляция между ИПЦ и M2 отражает инфляционный эффект монетарной политики, а добавленная стоимость в промышленности характеризует жизнеспособность реальной экономики. Динамический механизм xGBoost может адаптивно регулировать веса характеристик, например, увеличивать вклад показателей процентной ставки в периоды, чувствительные к политике (например, цикл повышения процентной ставки).

Слой объединения моделей использует стратегию интеграции Stacking для глубокой интеграции результатов сезонного разложения ARIMA с характеристиками временных рядов LSTM: модуль ARIMA устраняет нестационарность и извлекает квартальную циклическую составляющую темпов роста ВВП с помощью обработки разностей; LSTM улавливает нелинейные остаточные характеристики (например, асимметричное влияние внезапных политических шоков на ИПЦ) с помощью механизма стробирования. Выходы двух моделей взвешиваются и объединяются с помощью полносвязной нейронной сети, при этом ARIMA вносит вклад в основу линейного тренда (вес 0,4), а LSTM дополняет нелинейные колебания (вес 0,6), что определяется путем оптимизации решетчатого поиска.

На уровне количественной оценки рисков применяется моделирование методом Монте-Карло для построения кривых плотности распределения вероятностей. На основе исторической волатильности и матрицы коэффициентов корреляции берется 10 000 случайных выборок 7 ключевых показателей, чтобы смоделировать сочетание экономических показателей при экстремальных сценариях (например, падение темпов роста M2 на 3σ при чрезмерно ожидаемом повышении ИПЦ). Рассчитывая VaR (при 95% доверительном уровне) и CVaR, мы количественно определяем пороговые значения макроэкономического системного риска, например, результаты моделирования показывают, что при превышении ножницами CPI-M2 5% вероятность перегрева экономики достигает 67%. Анализ поверхности риска также показывает концентрацию плотности вероятности в квадранте «высокая инфляция - низкий рост», обеспечивая перспективное предупреждение для принятия политических решений.

Архитектура достигает адаптивности модели за счет механизма динамической настройки веса, например, усиливая вклад ARIMA в период сглаживания данных и переключаясь на доминирующий режим LSTM в период колебаний. Гибридная модель снижает RMSE на 23,7% по сравнению с одиночной моделью, как показано на бэкteste индекса CSI 300.

Процесс исследования: Двухнаправленная сеть долговременной и кратковременной памяти (BiLSTM), предложенная в данном исследовании, значительно улучшает способность захвата критических бизнес-узлов за счет внедрения механизма временного

внимания с двунаправленной структурой стробирования. В слое временного внимания модель динамически присваивает веса, вычисляя энергетический балл (ES) скрытого состояния, например, внимание к узлам ключевых событий, таких как возврат средств в конце квартала и годовой цикл аудита, увеличивается на 40%. Анализ матрицы весов внимания показывает, что модель в 2,3 раза более чувствительна к пику потока капитала, чем традиционная LSTM, особенно при работе со сценариями финансирования цепочки поставок, и способна точно определить паттерны колебаний денежного потока, вызванные концентрацией расчетного периода. Двусторонняя сетевая структура эффективно улавливает эффект запаздывания при передаче политики благодаря синергии прямой и обратной LSTM. На примере политики центрального банка по сокращению квот модель успешно отображает 6-месячный цикл (с коэффициентом корреляции запаздывания 0,79) улучшения денежных потоков микро- и малых предприятий после выхода политики, показывая путь проникновения монетарной политики в реальную экономику через кредитный канал коммерческих банков.

Для решения проблемы избыточной подгонки, вызванной шумом данных финансовых временных рядов, в данном исследовании используется инновационная техника DropPath для регуляризации BiLSTM. Эта техника заставляет сеть изучать избыточные представления признаков путем случайного отбрасывания путей соединения между нейронами в процессе обучения (коэффициент отбрасывания установлен на 0,2). Эксперименты показывают, что DropPath снижает RMSE валидационного набора с 0,121 до 0,086, сохраняя при этом плавное снижение профиля потерь обучающего набора (колебания уменьшились на 68 %). Модель демонстрирует высокую устойчивость на тестовом наборе: стандартное отклонение ошибки предсказания для внезапных политических шоков (например, корректировка ставки НДС) снижается на 41 % по сравнению с эталонной моделью, а точность улавливания сезонных колебаний повышается до 89,7 %. Визуализация тепловой карты внимания на рисунке 4 показывает, что модель начинает увеличивать вес внимания к данным о розничной торговле за 30 дней до сезона потребления на китайский Новый год, запуская механизм раннего предупреждения на 15 дней раньше, чем традиционный метод.

Интерактивная приборная панель, разработанная в данном исследовании, глубоко интегрирует многомерные финансовые данные и обеспечивает динамическое отображение финансового состояния предприятия с помощью многоуровневой архитектуры визуализации. В области основных показателей используется технология тепловой карты для представления предупреждения о разрыве денежного потока, а также сочетаются географическая информационная система (ГИС) и региональные экономические данные для визуального отражения уровня риска ликвидности различных региональных филиалов с помощью цветового градиента. Например, выделенная красным цветом область указывает на то, что узлы цепочки поставок в дельте реки Янцзы испытывают трудности с погашением краткосрочных долгов, в то время как зеленая область характеризует стабильные резервы денежных средств производственных баз в центральных и западных регионах. В области корреляционного анализа используется инновационная технология Force-Directed Graph, которая моделирует данные о дебиторской задолженности и узлы цепочки поставок выше и ниже по течению в виде динамической топологии сети, выявляя влияние длительного периода выставления счетов основными клиентами на вторичных поставщиков посредством настройки параметров силы притяжения и силы отталкивания между узлами и поддерживая двойное нажатие на узлы для проникновения в данные на уровне деталей

транзакций. В области прогнозирования тенденций интегрирован интерактивный элемент управления в виде скользящей шкалы времени, который позволяет пользователям перетаскивать кривые моделирования доходов и расходов в различных доверительных интервалах (90%, 95%, 99%). Система выдает результаты прогнозирования в реальном времени на основе поверхностей распределения вероятностей, полученных в результате моделирования Монте-Карло, и дополнена аннотациями и оперативными руководствами, что значительно снижает стоимость обучения для нетехнических пользователей.

Динамический проектор песочных таблиц создает цифровую двойную среду для принятия корпоративных финансовых решений с помощью анализа чувствительности параметров и технологии многосценарного моделирования. В механизм встроено 28 регулируемых ползунков переменных политики (например, базовая процентная ставка, ставка налога на добавленную стоимость, ставка экспортной льготы), которые помогают лицам, принимающим решения, наблюдать за предельным влиянием точной настройки параметров на маржу чистой прибыли в режиме реального времени и выделять ключевые переменные с коэффициентами чувствительности более 0,5 с помощью технологии градиентного окрашивания. Модуль стресс-тестирования использует алгоритмы дискретно-событийного моделирования, чтобы смоделировать траекторию передачи колебаний обменного курса при сценарии обострения торговых противоречий между Китаем и США. Система автоматически генерирует отчеты по анализу затрат и выгод для программ хеджирования валютных рисков, таких как запуск предложения стратегии автоматической блокировки обменного курса, когда обменный курс юаня к доллару превышает 7,2.

Трехмерное пространство вычетов преодолевает ограничения традиционных двухмерных отчетов, отображая баланс в виде вращающейся трехмерной топологии, где счета активов характеризуются объемом взвешенных кубов, счета обязательств показывают отношения цепочки залогов через радиальные линии, а пользователи могут интерактивно препарировать многоуровневую корреляцию конкретного счета с помощью жестов рук. При применении этого механизма в автомобильной группе, с помощью модели колебания цен на сырье, стоимость деталей из листового металла была предупреждена о повышении за 6 месяцев, и в сочетании со стратегией фьючерсного хеджирования для блокировки стоимости закупок была достигнута экономия в 17 миллионов юаней, что подтвердило ценность системы для поддержки принятия решений в сложной экономической среде.

Выводы: чтобы проверить реальную эффективность системы, для проведения сравнительных экспериментов в рамках данного исследования были выбраны три зарегистрированных производственных предприятия в регионе дельты реки Янцзы. Результаты эксперимента показывают, что на основе синергетического эффекта интеллектуальной модели прогнозирования и визуальной системы поддержки принятия решений значительно оптимизируются основные показатели финансового управления предприятием. Коэффициент отклонения прогноза сократился с $\pm 15\%$ в традиционной системе до $\pm 7\%$, что на 53% меньше. Особенно в период окончания квартала, когда цены на сырье резко колеблются, модель значительно повышает точность прогнозирования затрат на закупки за счет учета эффекта связи между Baltic Dry Index (BDI) и товарными фьючерсами. Количество дней оборотного капитала сократилось с 68 до 52 дней. Благодаря динамическому визуальному анализу сети дебиторской задолженности предприятие успешно выявило 12 вторичных поставщиков с просроченной задолженностью и реструктурировало условия расчетов. Коэффициент выявления безнадежных долгов вырос с 71 до 89 %, а система точно определила три аномальные вексельные транзакции с

фиктивной торговой биографией на сумму 47 млн юаней благодаря двойному механизму проверки с использованием алгоритма изолированного леса и вывода графа знаний. Эти улучшения позволили снизить годовые финансовые расходы экспериментального предприятия на 18 %, подтвердив практическую ценность инструмента в сложных сценариях цепочки поставок.

В сценарии применения в розничной торговле модель динамического равновесия «запасы-продажи», разработанная в данном исследовании, демонстрирует мощные возможности прогнозирования спроса и оптимизации принятия решений. В ответ на сезонные колебания модель повышает точность прогнозирования продаж с 58 до 82 % в период промо-сезона за счет интеграции неструктурированных источников данных, таких как данные о погоде и индекс мнений в социальных сетях. Модуль анализа сети проведения спроса успешно определил 36 ключевых факторов влияния на корреляцию, например, коэффициент понижающего давления акции «полное снижение» магазина-конкурента на цену единицы продукции предприятия достиг 0,41, а открытие новой линии метро дало коэффициент усиления 0,67 для потока покупателей соседних магазинов. В период акции «Double Eleven» система, основанная на данных о воронке продаж в реальном времени и прогнозе логистических мощностей, динамически корректирует программу складирования на региональных складах, снижая долю медленно продающихся товаров с 9,3% до 4,1%, сокращая потенциальные потери на 3,2 млн юаней. В частности, модель моделирует стоимость капитальных затрат при различных стратегиях складирования с помощью технологии трехмерного пространственного прогнозирования, а программа «градиентного пополнения запасов», разработанная для складов в Восточном Китае, повысила эффективность оборота на 37 %, что подтверждает адаптируемость инструмента к новой розничной индустрии.

Интеллектуальный инструмент принятия финансовых решений, разработанный в данном исследовании, достиг значительного прогресса в точности прогнозирования и поддержке принятия решений, но по-прежнему сталкивается с двумя основными проблемами: во-первых, возможности мультимодального объединения неструктурированных данных для ММСП недостаточны, и существует ограниченность существующей модели в обработке фрагментированных данных, таких как текст и изображения; во-вторых, сценарное моделирование экстремальных событий «черного лебедя» еще не совершенно, и трудно точно оценить каскадный эффект неожиданных рисков на промышленную цепочку. Будущие исследования будут сосредоточены на трех основных технологических прорывах: во-первых, построение модели проведения рисков в промышленной цепочке на основе отображения знаний и создание динамического механизма раннего предупреждения о путях распространения рисков путем интеграции междоменных данных, таких как сети поставщиков и динамика рынка, таких как отображение промышленных знаний, что уже позволило повысить точность определения кредитного риска поставщика на 35 %; во-вторых, разработка платформы для совместной работы с данными на основе федерального обучения, использование гомоморфного шифрования и методов дифференциальной конфиденциальности и обеспечение суверенитета данных на основе принципа суверенитета данных. Во-вторых, разработка федеративной платформы для совместной работы с данными, управляемой обучением, с использованием гомоморфного шифрования и технологии дифференциальной конфиденциальности для совместного моделирования между предприятиями при условии соблюдения суверенитета данных, и на примере цепочки поставок автомобилей показано,

что эта технология позволяет снизить ошибку в прогнозировании запасов на 18 %; наконец, углубление применения цифровых двойников в принятии финансовых решений, создание среды моделирования в реальном времени с помощью реально-виртуального отображения и применение цифровой модели двойника на производственном предприятии для достижения динамической оптимизации производственных затрат и повышения операционной эффективности на 23 %.

В будущем интеллектуальная финансовая система будет развиваться по принципу «интеграция технологий - расширение сцены - этические нормы». На техническом уровне синергия графа знаний, федерального обучения и цифрового двойника сформирует замкнутую систему «интеллектуальный сбор данных - распределенные вычисления - динамическое моделирование и дедукция», которая позволит системе иметь адаптивную итеративную способность. Сценарий применения будет расширен от анализа отдельного предприятия до управления рисками на уровне отраслевой цепочки, а количественная оценка устойчивости цепочки поставок будет достигнута путем построения таких моделей, как дерево проведения рисков поставщиков и сеть влияния логистических сбоев. Что касается этики данных, то необходимо создать систему управления, обеспечивающую баланс между безопасностью и эффективностью, включая разработку федеральных стандартов обучения для классификации и авторизации данных, а также создание системы отслеживания аудита на основе блокчейна, чтобы обеспечить соблюдение коммерческой тайны и прав на неприкосновенность частной жизни в трансграничных потоках данных. Эти инновации не только будут способствовать трансформации финансового управления от управления на основе опыта к управлению на основе данных, но и обеспечат ключевую технологическую поддержку для создания устойчивой к рискам промышленной экосистемы, которая, как ожидается, повысит эффективность совместного принятия решений в промышленной цепочке более чем на 40 процентов в течение следующих пяти лет и увеличит скорость реагирования на кризисы на 60 процентов.

Список источников

1. Yilmaz D, Büyüktaktakın İ E. Learning optimal solutions via an LSTM-optimization framework[C]//Operations Research Forum. Cham: Springer International Publishing, 2023, 4(2): 48.
2. Wang Y, Liu Y, Wang M, et al. LSTM model optimization on stock price forecasting[C]//2018 17th international symposium on distributed computing and applications for business engineering and science (dcabes). IEEE, 2018: 173-177.
3. Bhimavarapu U, Battineni G, Chintalapudi N. Improved optimization algorithm in LSTM to predict crop yield[J]. Computers, 2023, 12(1): 10.
4. WANG Tianyi, LIU Xingui. Обратная численная дифференциальная коррекция в динамическом производстве песчаного стола[J]. Mapping and Spatial Geographic Information, 2024, 47(04):39-43.
5. Zeng Lulu. Применение визуализации данных в обучении песчаными таблицам ERP - на основе создания динамических диаграмм в Excel[J]. Science and Technology Wind, 2020, (18):47-48. DOI:10.19392/j.cnki.1671-7341.202018034.
6. Shanshan Hu, Kehao Chi, Mengwen Chen, et al. Динамическая система демонстрации песчаного стола на основе компьютерного зрения[J]. University Physics Experiment, 2019, 32(03):107-111. DOI:10.14139/j.cnki.cn22-1228.2019.03.027.

Сведения об авторе

Фан Чэньси, докторант, Ростовский государственный экономический университет (РИНХ), г. Ростов-на-Дону, Россия

Научный руководитель

Чернышева Ю.Г., д.э.н., доцент, Ростовский государственный экономический университет (РИНХ), г. Ростов-на-Дону, Россия

Information about the author

Fan Chensi, doctoral student, Rostov State University of Economics, Rostov-on-Don, Russia

Scientific supervisor

Chernysheva Yu.G., Doctor of Economics, Associate Professor, Rostov State University of Economics (RINH), Rostov-on-Don, Russia