

СИСТЕМА ПОДДЕРЖКИ ПРИНЯТИЯ РЕШЕНИЙ НА ОСНОВЕ ОБУЧЕНИЯ С ПОДКРЕПЛЕНИЕМ ДЛЯ АВТОМАТИЗАЦИИ ПРОЦЕССА ОДОБРЕНИЯ КРЕДИТОВ

A DECISION SUPPORT SYSTEM BASED ON REINFORCEMENT LEARNING TO AUTOMATE THE LOAN APPROVAL PROCESS

Z. Askerov

Summary. This article is devoted to the development and implementation of a decision support system based on reinforcement learning to automate the loan approval process. The paper discusses key aspects of reinforcement learning, including agent behavior modeling, state and action space definition and analysis, and strategy updating mechanisms based on received rewards. Examples of the system's use in the banking sector, features of integration with existing banking systems, and challenges related to data protection and privacy are presented.

Keywords: reinforcement learning, decision support system, loan approval automation, artificial intelligence, data protection.

Аскеров Заур Ханахмедович

аспирант, Национальный Исследовательский
Ядерный Институт МИФИ, Москва
zaur_askerov_1998@mail.ru

Аннотация. Статья посвящена разработке и внедрению системы поддержки принятия решений на основе обучения с подкреплением для автоматизации процесса одобрения кредитов. В работе рассматриваются ключевые аспекты обучения с подкреплением, включая моделирование поведения агентов, определение и анализ пространства состояний и действий, а также механизмы обновления стратегий на основе полученных наград. Приводятся примеры использования системы в банковской сфере, особенности интеграции с существующими банковскими системами, а также обсуждаются вызовы, связанные с защитой данных и конфиденциальностью.

Ключевые слова: обучение с подкреплением, система поддержки принятия решений, автоматизация кредитования, искусственный интеллект, защита данных.

В современном мире, где скорость и точность принятия финансовых решений становятся максимально важны, автоматизация финансовых услуг с использованием методов искусственного интеллекта становится не просто тенденцией, а необходимостью. Одним из наиболее перспективных направлений в этой области является применение обучения с подкреплением (Reinforcement Learning, RL) — метода машинного обучения, который позволяет системам автоматически оптимизировать процессы принятия решений на основе взаимодействия с окружающей средой [8].

Целью данной работы является разработка и внедрение системы поддержки принятия решений на основе обучения с подкреплением для автоматизации процесса одобрения кредитов. Эта система предназначена для повышения эффективности кредитных операций путем автоматизации процессов анализа и одобрения заявок на кредиты, что позволяет существенно сократить время обработки заявок и уменьшить человеческий фактор при оценке кредитоспособности клиентов [4].

В работе рассматриваются основные компоненты системы обучения с подкреплением, включая агентов, среду, награды и стратегии, а также методы их взаимодействия и обучения. Особое внимание уделяется вопросам интеграции системы с существующими банковскими

системами, а также проблемам защиты данных и конфиденциальности, возникающим в процессе автоматизации кредитных операций [5].

Обучение с подкреплением является направлением машинного обучения, где агенты автоматически определяют идеальные действия в конкретной среде для максимизации накопленной награды [3]. Этот процесс включает в себя обучение оптимальной стратегии действий через испытания, ошибки и корректировку на основе полученного опыта [2].

Основные компоненты обучения с подкреплением:

1. Агенты:

Агенты в RL — это субъекты или программы, которые выполняют действия в среде. Агенты принимают решения на основе текущего состояния среды с целью максимизации суммарной будущей награды. Они обучаются выбирать действия, которые увеличивают вероятность достижения желаемых результатов, часто через методы проб и ошибок.

2. Среда:

Среда в RL описывает мир или контекст, в котором действует агент. Среда отвечает на действия агента, пре-

доставляя новые состояния и награды. Состояния могут включать различные наблюдаемые параметры, необходимые агенту для оценки текущих условий и выбора следующего действия.

3. Награды:

Награда — это сигнал, который агент получает от среды после выполнения действия. Награда может быть положительной или отрицательной и служит указателем на «хорошее» или «плохое» действие в контексте задачи. Обучение с подкреплением направлено на максимизацию суммарной награды, которую агент получает со временем.

4. Стратегии:

Стратегия (или политика) определяет, как агент выбирает действие на основе наблюдаемого состояния среды. Стратегия может быть детерминированной, где для каждого состояния существует конкретное действие, или стохастической, где результаты действий не всегда предсказуемы. Цель обучения — найти стратегию, которая максимизирует суммарную награду [1].

В различных сферах, включая игры (например, Go и шахматы), автономные транспортные средства (управление беспилотными автомобилями), робототехнику (автоматизация роботизированных задач) и здравоохранение (оптимизация лечебных процедур), обучение с подкреплением показало свою эффективность.

В финансовых услугах RL может использоваться для автоматизации и оптимизации решений о кредитовании, управлении инвестиционными портфелями и ценообразовании в реальном времени. Применимость RL к финансовым услугам обусловлена его способностью адаптироваться к изменяющимся условиям рынка и принимать решения на основе сложных, непредсказуемых переменных рыночной среды [6].

Основа системы — это модель, которая обрабатывает кредитные данные, прогнозируя результаты заявок на основе исторических данных о клиентах. Это включает информацию от кредитных бюро, данные о предыдущих кредитах и информацию, предоставленную клиентами при подаче заявки.

Система использует алгоритмы Q-обучения, обучаясь на симуляции кредитного бизнес-процесса. Это позволяет системе оптимизировать порог одобрения в реальном времени, постоянно корректируя его в зависимости от изменений в данных и рыночной среде [2].

Система интегрирована с банковскими системами для обеспечения бесперебойного взаимодействия. Ин-

терфейсы позволяют системе обмениваться данными с банком, автоматически принимать решения о кредитах на основе текущего порога одобрения, установленного алгоритмами RL.

На рисунке 1 показана диаграмма работы системы принятия решений.



Рис. 1. Диаграмма работы

Клиент начинает процесс, отправляя запрос на подачу заявки на кредит, а система отвечает подтверждением приема заявки. Это визуализация является хорошим началом для описания процесса обмена информацией в автоматизированной системе кредитования.

Модель данных включает:

1. Сбор данных: Интеграция с банковскими системами позволяет системе автоматически собирать данные о клиентах, включая кредитные истории, демографические данные и информацию о предыдущих займах.
2. Обработка данных: Данные обрабатываются для создания профилей заемщиков, которые включают оценки кредитоспособности, основанные на историческом поведении и возвращаемости кредитов.

Алгоритмы обучения включают [10]:

1. Обучение с подкреплением (Reinforcement Learning, RL): Система использует Q-обучение для динамической адаптации порога кредитного скоринга, оптимизируя его для максимизации прибыли.
2. Адаптация к изменениям: Алгоритмы постоянно анализируют входящие данные о новых займах и их исходах, корректируя стратегии в реальном времени для учета новой информации и изменений в экономической среде.

Интерфейсы включают [7]:

1. API для интеграции: система обеспечивает API-интерфейсы для интеграции с банковскими и кредитными системами, что позволяет обмениваться данными в реальном времени.

2. Интерфейс пользователя: для операторов кредитных отделов предусмотрены интерфейсы, позволяющие визуализировать текущие параметры модели, истории одобрений и отказов, а также управлять настройками порогов скоринга.

Для обучения системы, использующей обучение с подкреплением для автоматизации процесса одобрения кредитов, были использованы датасеты, сгенерированные в рамках симуляции. Эти данные основываются на историческом периоде с 24 мая 2015 года по 9 июля 2017 года, включающем 112 наблюдений, каждое из которых соответствует одной неделе.

Описание данных:

В каждой симулированной неделе генерировались следующие данные:

- Количество новых и повторных заявок на кредиты от клиентов.
- Сумма займа и срок, на который он выдан.
- Потенциальная прибыль от каждого займа.
- Параметры, указывающие, просрочит ли клиент платеж или вернет долг после процедуры взыскания.
- Предсказания кредитного скоринга, сгенерированные моделью компании на основе этих данных.

Процесс обучения:

Обучение проводилось в течение 100 тренировочных эпизодов. Каждый эпизод включал следующие шаги:

1. Генерация состояния (например, недельная норма одобрения) и передача его агенту.
2. Агент предсказывал Q-значения для действий (пороги одобрения), используя модель функции ценности.
3. На основе предсказанных Q-значений и политики обучения (Boltzmann-Q) выбиралось действие (порог одобрения на следующую неделю).
4. Среда генерировала новое состояние на основе выбранного действия и выдавала награды (прибыль).
5. Обновление модели функции ценности на основе полученных наград и истории действий.

Использование сгенерированных данных вместо реальных данных в обучении моделей машинного обучения, особенно в сфере кредитования, часто обусловлено вопросами конфиденциальности и защиты данных:

1. Реальные данные клиентов часто содержат конфиденциальную информацию, такую как личные идентификаторы, финансовые данные, историю транзакций и кредитную историю. Использование этих данных без соответствующих мер безопасно-



Рис. 2. Алгоритм обучения

сти и согласия субъектов данных может нарушить законодательные требования, такие как GDPR или другие местные нормы защиты данных.

2. Законы о защите данных во многих странах требуют, чтобы данные обрабатывались законно, справедливо и прозрачно. Использование реальных данных без ясного согласия может подвергнуть организации юридическим рискам, включая штрафы и ущерб репутации.
3. Генерация данных позволяет исследователям создавать сценарии, которые точно отражают различные аспекты кредитного процесса, при этом обеспечивая полный контроль над распределением и характеристиками данных. Это помогает избежать случайных искажений, которые могут возникнуть при использовании реальных данных.
4. Использование реальных данных может поднимать этические вопросы, особенно если данные использовались без адекватного информирования субъектов о том, как и для каких целей они будут использоваться.
5. Сгенерированные данные могут быть спроектированы таким образом, чтобы они отражали разнообразные условия и сценарии, что улучшает способность модели к обобщению на новые, неизведанные данные.

Для реализации системы автоматизации процесса одобрения кредитов с использованием обучения с подкреплением были применены различные технологии и инструменты, основанные на текущих тенденциях в области искусственного интеллекта и машинного обучения.

Основные технологии и инструменты [9]:

1. Язык программирования Python остается основным инструментом для разработки систем обучения с подкреплением, благодаря своей гибкости и обширной поддержке библиотек машинного обучения.
2. TensorFlow и PyTorch. Эти фреймворки глубокого обучения используются для создания и тренировки нейронных сетей, которые могут оценивать возможные действия и их последствия в контексте задачи одобрения кредитов.
3. Ray RLlib. Эта библиотека предлагает удобные инструменты для распределенного обучения с подкреплением, поддерживая различные алгоритмы RL и предоставляя единый API для обучения и взаимодействия с моделями. RLlib используется в различных областях, включая финансовую торговлю и оптимизацию сетей, что делает его идеальным выбором для решения сложных оптимизационных задач в разных отраслях.
4. Имитационное моделирование (Simulation Modeling). Для обучения агентов RL часто исполь-

зуются симуляторы или «цифровые двойники», которые воспроизводят условия реального мира и предоставляют платформу для тренировки агентов без риска для реальных операций.

5. Облачные вычисления. Платформы, такие как Amazon SageMaker, предоставляют инфраструктуру для тренировки и развертывания моделей машинного обучения в масштабе, что критически важно для обработки больших объемов данных и сложных вычислений, связанных с RL.

Построение модели окружающей среды для агента ИИ в системе поддержки принятия решений на основе обучения с подкреплением включает разработку симуляционной среды, которая воссоздает условия реального мира, с которыми сталкивается агент. Сам процесс представлен ниже:

1. Определение пространства состояний и действий.

Среда, разработанная на базе класса `gym.Env`, определяет пространства состояний и действий, а также устанавливает обмен состоянием, действием и наградой между агентом и средой.

Здесь `action_space` и `observation_space` определяют в зависимости от типа действий (дискретные или непрерывные), что позволяет моделировать различные сценарии взаимодействия агента с средой.

2. Взаимодействие с агентом.

Среда получает действия от агента, генерирует новые состояния на основе этих действий и вычисляет награды. Это процесс включает динамическое вычисление характеристик портфеля займов и получаемых наград.

Метод `_step` обрабатывает действие, отправленное агентом, и возвращает новое состояние, награду и информацию о том, завершен ли эпизод.

3. Настройка параметров и инициализация.

Перед началом обучения агента параметры среды настраиваются для отражения специфических условий, таких как искажения в данных или особенности кредитного портфеля.

Конструктор класса `SimulationEnv` позволяет задать искажения, которые могут имитировать различные экономические условия или изменения в поведении заемщиков.

В ходе данной работы была разработана и протестирована система поддержки принятия решений на основе обучения с подкреплением, предназначенная для

автоматизации процесса одобрения кредитов. Эта система продемонстрировала свою эффективность в оптимизации кредитных операций, позволяя существенно сократить время обработки заявок и улучшить точность принятия решений.

Основные преимущества данной системы заключаются в способности адаптироваться к меняющимся условиям рынка и принимать решения на основе анализа большого количества данных. Это делает её перспективным инструментом для банковских организаций, стремящихся к автоматизации и повышению эффективности своих операций.

В процессе работы были выявлены и решены несколько ключевых проблем, связанных с защитой дан-

ных и конфиденциальностью, что подтвердило возможность безопасного использования системы в реальных условиях. Кроме того, система интегрирована с существующими банковскими системами, что упрощает её внедрение и использование в практике кредитования.

Для дальнейших исследований предлагается рассмотреть возможность расширения области применения системы поддержки принятия решений на основе обучения с подкреплением, включая другие сферы финансовых услуг, такие как управление инвестиционными портфелями или оценка рисков. Также интересным направлением может стать разработка более сложных моделей, способных учитывать дополнительные факторы, влияющие на кредитные решения, и взаимодействие с другими системами искусственного интеллекта.

ЛИТЕРАТУРА

1. Абдуллаев, Э.А. Искусственный интеллект: текущие достижения и перспективы / Э.А. Абдуллаев. — Текст: непосредственный // Молодой ученый. — 2023. — № 33 (480). — С. 9–10. — URL: <https://moluch.ru/archive/480/105512/> (дата обращения: 23.04.2024).
2. Ванг З., Ду Я., Чжан Ю., Фанг М., Хуанг Б. МАССА: Офлайн многоагентное обучение с подкреплением с каузальным распределением кредита // Сайт ar5iv.org. Дата обновления: 27.02.2024. URL: <https://ar5iv.labs.arxiv.org/html/2312.03644>.
3. Ванг Ч., Сяо З. Глубокое обучение для оценки кредитоспособности с использованием трансформера встроенных особенностей // Applied Sciences. 2022. Т. 12. № 21. Стр. 10995. DOI: <https://doi.org/10.3390/app122110995>.
4. Данг Т.К., Тран Т.С., Туан Л.М., Тиеп М.В. Машинное обучение на основе методов пересэмплирования и глубокого обучения с подкреплением для систем обнаружения мошенничества с кредитными картами // Applied Sciences. 2021. Т. 11. № 21. Стр. 10004. DOI: <https://doi.org/10.3390/app112110004>.
5. Искусственные нейронные сети: учебник / В.В. Цехановский, Е.Ю. Бутырский, Н.А. Жукова [и др.] ; под ред. В.В. Цехановского. — Москва: КноРус, 2024. — 350 с. — ISBN 978-5-406-13273-9. — URL: <https://book.ru/book/954274> (дата обращения: 23.04.2024). — Текст: электронный.
6. Искусственные нейронные сети. Практикум: учебное пособие / Е.Ю. Бутырский, Н.А. Жукова, В.Б. Мельников [и др.] ; под ред. В.В. Цехановского. — Москва: КноРус, 2024. — 381 с. — ISBN 978-5-406-12340-9. — URL: <https://book.ru/book/951505> (дата обращения: 23.04.2024). — Текст: электронный.
7. Можжухина, А.В. Повышение эффективности размещения элементов БИС на основе алгоритмов машинного обучения / А.В. Можжухина. — Текст: непосредственный // Исследования молодых ученых: материалы XXXIII Междунар. науч. конф. (г. Казань, февраль 2022 г.). — Казань: Молодой ученый, 2022. — С. 9–15. — URL: <https://moluch.ru/conf/stud/archive/427/17004/> (дата обращения: 23.04.2024).
8. Пахса А. Финансовые технологии систем поддержки принятия решений // Журнал электрических систем и информационных технологий. 2024. Т. 11. Статья № 5. Доступ: <https://journal.example.com>. Дата обращения: 15.01.2024.
9. Пиньятелли Э., Феррет Дж., Гейст М., Меснар Т., Ван Хасселт Х., Тони Л. Обзор временного распределения кредита в глубоком обучении с подкреплением // Papers With Code. Дата публикации: 2 дек. 2023. URL: <https://paperswithcode.com/paper/a-survey-of-temporal-credit-assignment-in>.
10. Сивамайил К., Раджасекар Э., Альджафари Б., Николовски С., Ваиравасундарам С., Ваиравасундарам И. Систематическое исследование приложений на основе обучения с подкреплением // Energies. 2023. Т. 16. № 3. Стр. 1512. DOI: <https://doi.org/10.3390/en16031512>.

© Аскеров Заур Ханахмедович (zaur_askerov_1998@mail.ru)
Журнал «Современная наука: актуальные проблемы теории и практики»