

ИСПОЛЬЗОВАНИЕ РЕГРЕССИОННОГО АНАЛИЗА ДЛЯ ОЦЕНКИ ВЛИЯНИЯ ВЕСА МОДЕЛИ И ТОЧНОСТИ ПРОГНОЗИРОВАНИЯ НА ПРОИЗВОДИТЕЛЬНОСТЬ ВЕБ-ПРИЛОЖЕНИЯ

USING REGRESSION ANALYSIS TO ASSESS THE IMPACT OF MODEL WEIGHT AND FORECASTING ACCURACY ON WEB APPLICATION PERFORMANCE

A. Ilyin
S. Plotnikov

Summary. The article examines the relationship between the model weight and the quality of prediction using the regression analysis method. The study is aimed at identifying patterns and obtaining optimal parameters for further development of the author's method of predictive analysis of user behavior. The results proved a clear relationship between these parameters, thereby proving the hypothesis about the influence of the prediction model weight and the attenuation depth on the performance of a web application.

Keywords: predictive analysis, web application, Markov chains, regression analysis, probability theory.

Ильин Андрей Юрьевич

МИРЭА — Российский технологический университет, Москва
frayz2507@gmail.com

Плотников Сергей Борисович

к.т.н. доцент, МИРЭА — Российский технологический университет, Москва
plotnikovsb@mail.ru

Аннотация. В статье рассматривается взаимосвязь веса модели и качество предсказания методом регрессионного анализа. Исследование направлено на выявление закономерностей и получения оптимальных параметров для дальнейшей разработки авторского метода предиктивного анализа пользовательского поведения. Результаты доказали явную связь этих параметров, чем доказали гипотезу о влиянии веса модели предсказания и глубины затухания на производительность работы веб-приложения.

Ключевые слова: предиктивный анализ, веб-приложение, Марковские цепи, регрессионный анализ, теория вероятностей.

Введение

Современные веб-приложения играют ключевую роль в цифровой экономике, предоставляя широкий спектр услуг — от интернет-магазинов до социальных сетей. Быстродействие веб-страниц является критически важным параметром, от которого зависит удовлетворённость пользователей [1]. Кроме того, исследование компании Akamai показало, что задержка загрузки страницы в 100 миллисекунд снижает уровень десктопной конверсии на 2,4 %, а мобильной — на 7,1 %. Секундная задержка снижает конверсию с десктопов на 21,8 %, со смартфонов — на 20,5 %, а с планшетов — на 17,8 % [2]. Эти данные подчёркивают, что для создания конкурентоспособного продукта необходимо постоянно оптимизировать скорость взаимодействия веб-приложения с пользователем.

С цифровой трансформацией требования к качеству и быстродействию веб-приложений постоянно возрастают. Компании вынуждены модернизировать свою инфраструктуру, чтобы соответствовать ожиданиям пользователей, которые теперь требуют не только функциональности, но и высокой производительности. Цифровая трансформация требует внедрения иннова-

ционных методов разработки и использования новых алгоритмов, способных анализировать поведение пользователей и предугадывать их потребности [3]. Это обусловлено тем, что современный рынок требует быстрой адаптации и постоянного совершенствования сервисов.

В условиях такой динамики появляется необходимость использования алгоритмов предиктивного анализа, позволяющих не только анализировать текущие данные, но и прогнозировать будущие действия пользователей. Среди методов предиктивного анализа можно выделить несколько подходов, каждый из которых имеет свои преимущества. Одним из эффективных методов являются марковские цепи, которые позволяют моделировать вероятностное поведение пользователей, переходящих из одного состояния в другое. Марковские цепи широко применяются для оценки и предсказания последовательностей действий в веб-приложениях, что помогает заранее оптимизировать загрузку контента (например, переходы между страницами) [4].

Другим эффективным методом является регрессионный анализ, который применяется для оценки зависимости различных параметров системы. С его помощью можно установить, как изменение вычислительной на-

грузки или других характеристик модели влияет на время отклика приложения. Регрессионный анализ позволяет не только выявить ключевые зависимости, но и количественно оценить вклад каждого параметра в общую производительность системы [5].

Наконец, современные нейронные сети представляют собой ещё один перспективный подход в области предиктивного анализа. Благодаря своей способности обучаться на больших объёмах данных, нейронные сети могут эффективно предсказывать поведение пользователей, даже при наличии сложных нелинейных зависимостей. Этот метод особенно актуален в условиях постоянного роста данных и необходимости адаптации к быстро меняющимся условиям эксплуатации веб-приложений [6].

В условиях быстрого роста объёма данных и повышения требований к быстродействию веб-приложений, становится очевидным, что для эффективного предсказания поведения пользователей необходимы алгоритмы, которые бы минимизировали вычислительную нагрузку, сохраняя высокий уровень точности.

Применение методов марковских цепей позволит эффективно сохранять баланс между поставленными целями. Однако сокращение объёма обрабатываемых данных требует глубокого анализа влияния упрощений на точность прогнозов, как свидетельствуют исследования в области оптимизации алгоритмов [7, 8] и современные подходы в машинном обучении.

Актуальность поиска компромиссных решений между производительностью и качеством предсказаний становится очевидной в условиях цифровой трансформации и возрастания требований к веб-приложениям. Такой подход, основанный на интеграции инновационных методов анализа данных и оптимизации алгоритмов, позволяет не только улучшить пользовательский опыт, но и обеспечить стабильность работы систем в динамично изменяющейся среде [9].

Расчёт оптимальных параметров

При разработке авторской модели на основе метода, предложенного в введении, ключевое преимущество заключается в сниженной вычислительной нагрузке. Однако уменьшение «веса» модели (за счёт меньшего количества обрабатываемых данных, определяемого параметром *historyLength* (HL), и более резкого затухания через параметр *decayRate* (DR)) может негативно сказаться на точности предсказания т.к. предсказание приводит к избыточной загрузке компонентов, что увеличивает общее время отклика системы.

В разработанной модели точность предсказания *A* определяется на основе ретроспективного анализа

исторических данных. Для каждого взаимодействия система делала предсказание следующего действия, и точность вычислялась по формуле 1.

$$A = \frac{\text{Количество корректных предсказаний}}{\text{Общее количество предсказаний}} \quad (1)$$

При анализе исторических данных для каждой конфигурации получены следующие значения:

- для минимальной конфигурации: $A=0,88$ (88 % корректных предсказаний);
- для средней конфигурации: $A=0,93$ (93 % корректных предсказаний);
- для максимальной конфигурации: $A=0,95$ (95 % корректных предсказаний).

Данные значения, получены в ходе 200 экспериментальных запусков на каждую конфигурацию в системах различной сложности, после которых были усреднены и вычислены по формуле (1).

Эти значения отражают, что более «тяжёлая» модель (с наибольшим количеством обрабатываемых данных) учитывает больше информации и, соответственно, точность предсказаний повышается, однако за счёт увеличения базового времени загрузки.

Модель тестировалась в трёх конфигурациях, для которых были получены следующие измеренные значения:

Минимальная конфигурация:

- параметры: $HL = 10$, $DR = 0,7$;
- базовое время загрузки компонента $Y=0,354$ сек;
- вычисленная точность предсказания $A=0,88$ (88 % корректных предсказаний).

Средняя конфигурация:

- Параметры: $HL = 50$, $DR = 0.8$.
- Базовое время загрузки компонента $Y=0.36$.
- Вычисленная точность предсказания $A=0.93$ (93 % корректных предсказаний).

Максимальная конфигурация:

- Параметры: $HL = 100$, $DR = 0.9$.
- Базовое время загрузки компонента $Y=0.374$ сек.
- Вычисленная точность предсказания $A=0.95$ (95 % корректных предсказаний).

Прослеживается прямая зависимость — от увеличения обработки данных, повышается длительность ожидания загрузки компонента, что имеет отрицательный эффект на изначально поставленную цель разработки модели предиктивного анализа.

Для оценки влияния неверных предсказаний введем понятие «штрафа» — P , представляющего дополнитель-

ное время загрузки из-за лишнего запроса. В эксперименте значение этого параметра составило $P=0.56$ сек. Данная величина также была получена путём вычитания времени загрузки одного компонента при верном предсказании из времени загрузки компонента с неверным предсказанием, когда параллельно загружаются необходимый компонент и компонент, который не нужен. Штраф рассчитывается по формуле 2.

$$Penalty = (1 - A) \times P. \quad (2)$$

Вычислим штраф для каждой конфигурации:

- для минимальной конфигурации:

$$Penalty = (1 - 0,88) \times 0,56 = 0,0672;$$

- для средней конфигурации:

$$Penalty = (1 - 0,93) \times 0,56 = 0,0392;$$

- для максимальной конфигурации:

$$Penalty = (1 - 0,95) \times 0,56 = 0,028.$$

Можно заметить, что значение штрафа прямо пропорционально уменьшается относительно сложности

модели (учитывая данный фактор и вывод из предыдущего блока вычислений).

Эффективное время загрузки Y_{eff} учитывает как базовое время Y , так и дополнительную задержку из-за неверных предсказаний, таким образом этот параметр вычисляется по формуле 3.

$$Y_{eff} = Y + Penalty. \quad (3)$$

Подставив значения в формулу 3, получим следующие значения, отображенные в таблице 1.

Таблица 1.

Полученные данные

Конфигурация	HL	RD	Y (сек)	A	Penalty (сек)	Y_{eff} (сек)
Минимальная	10	0.7	0.354	0.88	0.0672	0.4212
Средняя	50	0.8	0.36	0.93	0.0392	0.3992
Максимальная	100	0.9	0.374	0.95	0.028	0.402

Из данных таблицы следует, что несмотря на то, что базовое время загрузки минимальной конфигурации

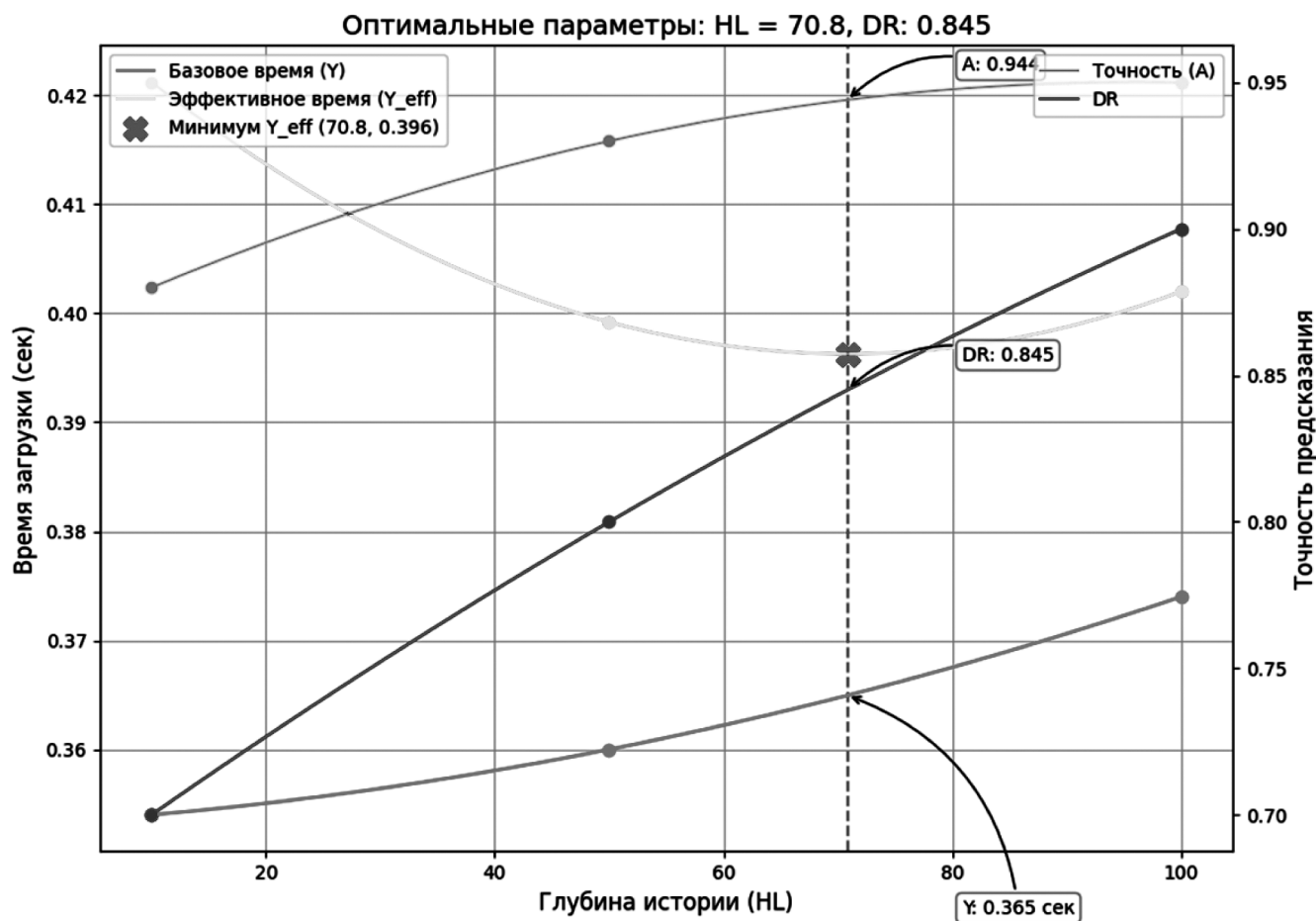


Рис. 1. Получение оптимальных параметров

наименьшее из всех, её относительно меньшая точность приводит к большему штрафу, и итоговое эффективное время составляет 0.4212 сек, что является наихудшим результатом среди всех. Средняя конфигурация, с $HL = 50$ и $DR = 0.8$, показывает оптимальное соотношение всех критериев: базовое время 0.36 сек и высокая точность (93 %) приводят к минимальному $Y_{eff} = 0.3992$ сек. Максимальная конфигурация, хоть и обладает наивысшей точностью, имеет немного большее базовое время, что даёт $Y_{eff} = 0.402$ сек.

На основе полученных данных, был построен график, который представлен на Рисунке 1.

Данный график был построен на основе вычислений, проведенных ранее. Точки на графике — результаты вычислений, через которые проведены оси, а также они были «сглажены» путем интерполяции данных с использованием полиномиальной аппроксимации. Далее была найдена минимальная точка экстремума на оси Y_{eff} и на основе этой точки были получены наилучшие значения для работы модели, а именно $historyLength (HL) = 70,8$ и $decayRate (DR) = 0,845$ при котором Y_{eff} будет принимать значение 0.396 сек.

Таким образом, были рассчитаны параметры для настройки модели предиктивного анализа, основанной на марковских цепях и механизме затухания. Используя эту модель веб-приложение, получит максимальную точность анализа действий пользователя, что, в свою очередь, приведёт к требуемой высокой скорости загрузки его компонентов.

Заключение

В работе проведено исследование параметров модели предсказательного анализа пользовательского поведения в клиент-серверном приложении, включающее построение регрессионной зависимости между ключевыми параметрами модели, такими как глубина истории (HL), коэффициент отсева (DR), время базовой загрузки компонента (Y), точность предсказания (A), а также штраф за неверные предсказания (P).

В ходе математического эксперимента определена закономерность, согласно которой увеличение глубины

истории модели влечет за собой как повышение точности предсказаний, так и увеличение времени загрузки компонентов. С использованием регрессионного анализа и последующего поиска оптимальных значений параметров определен баланс, при котором достигается минимальное значение эффективного времени загрузки (Y_{eff}), что критически важно для обеспечения высокой производительности всех систем, в которых будет интегрирована модель предиктивного анализа.

Результаты эксперимента показали, что при увеличении точности предсказания (A) происходит снижение количества неверных вызовов компонентов, что в свою очередь уменьшает избыточную нагрузку на серверную часть системы. В процессе анализа было проведено сглаживание полученных зависимостей для выявления минимальной точки скорости загрузки, что позволило определить оптимальное значение HL, при котором достигается наилучший компромисс между быстродействием и глубиной предсказательного анализа в модели предиктивного анализа, которая может быть использована в любом проекте и будет работать одинаково относительно тяжести компонентов веб-приложения.

Таким образом, в результате расчетов было получено оптимальное значение конфигурации предсказательной модели, при котором достигается максимальная эффективность работы системы:

- минимизация времени загрузки компонента при сохранении высокой точности предсказаний;
- уменьшение избыточных вызовов благодаря корректному прогнозированию пользовательских действий;
- достижение баланса между вычислительной сложностью модели и ее предсказательной мощностью.

Результаты проведенных исследований могут быть использованы для дальнейшего совершенствования системы предсказательного анализа пользовательского поведения, а также стать основой для разработки более адаптивных и интеллектуальных механизмов автоматического управления загрузкой компонентов в клиент-серверных приложениях.

ЛИТЕРАТУРА

1. Скорость сайта и пользовательский опыт: Важная взаимосвязь: Suhrob Baxtiyorov. [Электронный ресурс]. URL: <https://dora.uz/ru/blog/skorost-saita-i-polzovatel'skii-opyt-vaznaia-vzaimosviaz> (дата обращения: 10.02.2025).
2. Akamai: секундная задержка загрузки сайта снижает конверсию на 20,5%. [Электронный ресурс]. URL: <https://m.seonews.ru/events/akamai-sekundnaya-zaderzhka-zagruzki-sayta-snizhaet-konversiyu-na-20-5/> (дата обращения: 10.02.2025).
3. Разработка методов и алгоритмов идентификации моделей поведения пользователей: Тимошенко А.И. URL: <http://elar.urfu.ru/handle/10995/135788> (дата обращения: 13.02.2025).
4. Скрытая марковская цепь, условная Марковская цепь и Марковская цепь Блекуэлла: Бежаева З.И., Куликов В.Л., Олехова Е.Ф., Оселедец В.И. URL: <https://www.elibrary.ru/item.asp?id=35296619> (дата обращения: 13.02.2025).

5. The Elements of Statistical Learning: Data Mining, Inference, and Prediction: Trevor H., Robert T., Jerome F. URL: <https://hastie.su.domains/ElemStatLearn/> (дата обращения: 19.02.2025).
6. Deep Learning: Ian G., Yoshua B., Aaron C. URL: <https://www.deeplearningbook.org/> (дата обращения: 19.02.2025).
7. Алгоритм последовательных упрощений при оценке временной сложности алгоритмов с применением полумарковских моделей: Ларкин Е.В., Привалов А.Н., Клепиков А.К. URL: <https://cyberleninka.ru/article/n/algoritm-posledovatelnyh-uproscheniy-pri-otsenke-vremennoy-slozhnosti-algoritmov-s-primeneniem-polumarkovskih-modeley> (дата обращения: 19.02.2025).
8. Марковские цепи в задачах диагностики и прогнозирования: Куравский Л.С. URL: https://psyjournals.ru/nonserialpublications/mmdp_2017/mmdp_2017.pdf (дата обращения: 19.02.2025).
9. Оптимизация интеграции технологий искусственного интеллекта в веб-приложения для улучшения пользовательского опыта: Бугрова О.А., Копша О.Ю. URL: <https://www.xn----8sbempclcwd3bmt.xn--p1ai/article/19298> (дата обращения 19.02.2025).

© Ильин Андрей Юрьевич (frayz2507@gmail.com); Плотников Сергей Борисович (plotnikovsb@mail.ru)

Журнал «Современная наука: актуальные проблемы теории и практики»