



9. Ivaschenko A., Anikin D., Chuvakov A., Sitnikov P., Surnin O. Computer vision system for deviant behavior identification // Proceedings of the 5th Annual Science Fiction Prototyping Conference 2021, Ghent, Belgium, April 15-16, 2021. – pp. 30 – 33

В.А. Байрамов, И.А. Лёзин

## АВТОМАТИЗИРОВАННАЯ ИНФОРМАЦИОННАЯ СИСТЕМА РАСПОЗНАВАНИЯ ПОДДЕЛЬНЫХ ОТЗЫВОВ В СЕТИ ИНТЕРНЕТ

(Самарский университет)

В наши дни онлайн торговля приобретает всё более широкое распространение и большую популярность за счёт экономии времени клиентов. Отзывы о продуктах в интернете играют всё более важную роль в жизни клиентов и представляют собой новый особый тип информации. Последние исследования показали, что 52% потребителей ищут информацию о продукте в интернете, в то время как 24% из них делают это непосредственно перед совершением покупки [1]. Не секрет, что имеет распространение факт намеренного инвестирования продавцов в положительные отзывы на свой товар, чтобы повысить его популярность. Самая распространённая причина заказа поддельного отзыва – малый поток клиентов [2]. В итоге у покупателя может сложиться неоправданно завышенное впечатление о продукте, что подтолкнёт его к совершению покупки.

Одной из задач исследования является разработка модели, которая предоставит возможность классифицировать отзывы на основе информации о пользователе, который размещает эти самые отзывы. Эту информацию можно разбить на 4 категории:

1. Публичные данные о пользователе:
  - 1.1. Имя;
  - 1.2. Описание профиля;
  - 1.3. Наличие фотографий;
  - 1.4. Дата регистрации;
2. Данные социального взаимодействия:
  - 2.1. Количество друзей/подписчиков;
  - 2.2. Количество подписок;
3. Данные отзывов:
  - 3.1. Средняя скорость, с которой пользователь оставляет отзывы;
  - 3.2. Количество отзывов с высокой оценкой относительно общего количества;
4. Лингвистические характеристики отзыва, рассчитанные с помощью LIWC.

Также цель исследования состоит в разработке алгоритмов работы автоматизированной информационной системы. Разрабатываемая система должна



иметь функционал по автоматизированному сбору данных для обучения. Для этого разработан алгоритм, представленный на рисунке 1.

Алгоритм подразумевает анализ всех отзывов на определённый продукт, извлечение перечисленных выше данных и сохранение их в файл для дальнейшей работы.

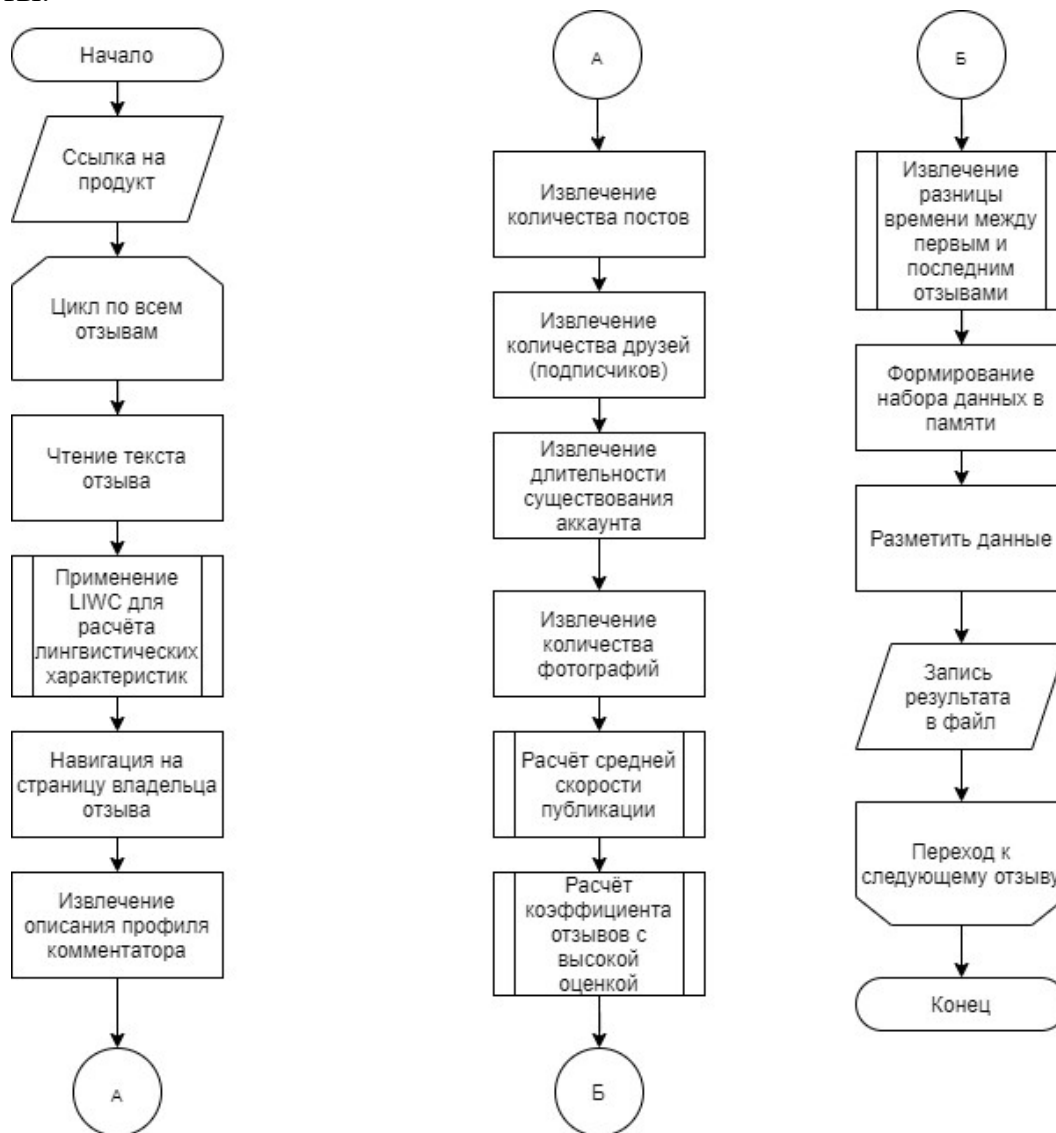


Рисунок 1 – Алгоритм работы системы при сборе данных для обучения

После сбора данных систему необходимо обучить, для этого разработан алгоритм, представленный на рисунке 2.

Для реализации системы было решено использовать алгоритм классификации под названием PU learning. Это алгоритм бинарной классификации, который позволяет проводить обучение на смеси размеченных и неразмеченных данных [3]. Он хорошо подходит для задачи классификации отзывов, т.к. получить набор полностью размеченных данных для обучения не представляется возможным.

После обучения система готова к использованию конечным пользователем. Алгоритм работы системы при классификации представлен на рисунке 3.

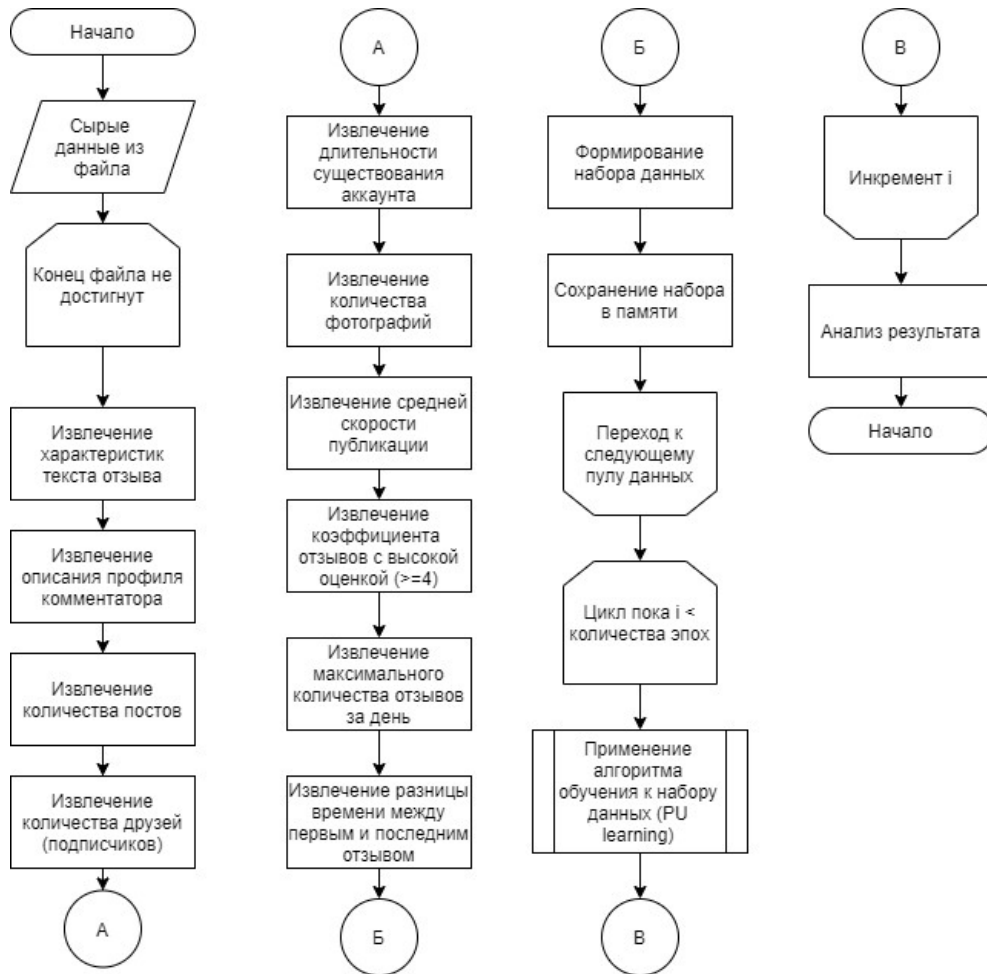


Рисунок 2 – Алгоритм работы системы при обучении

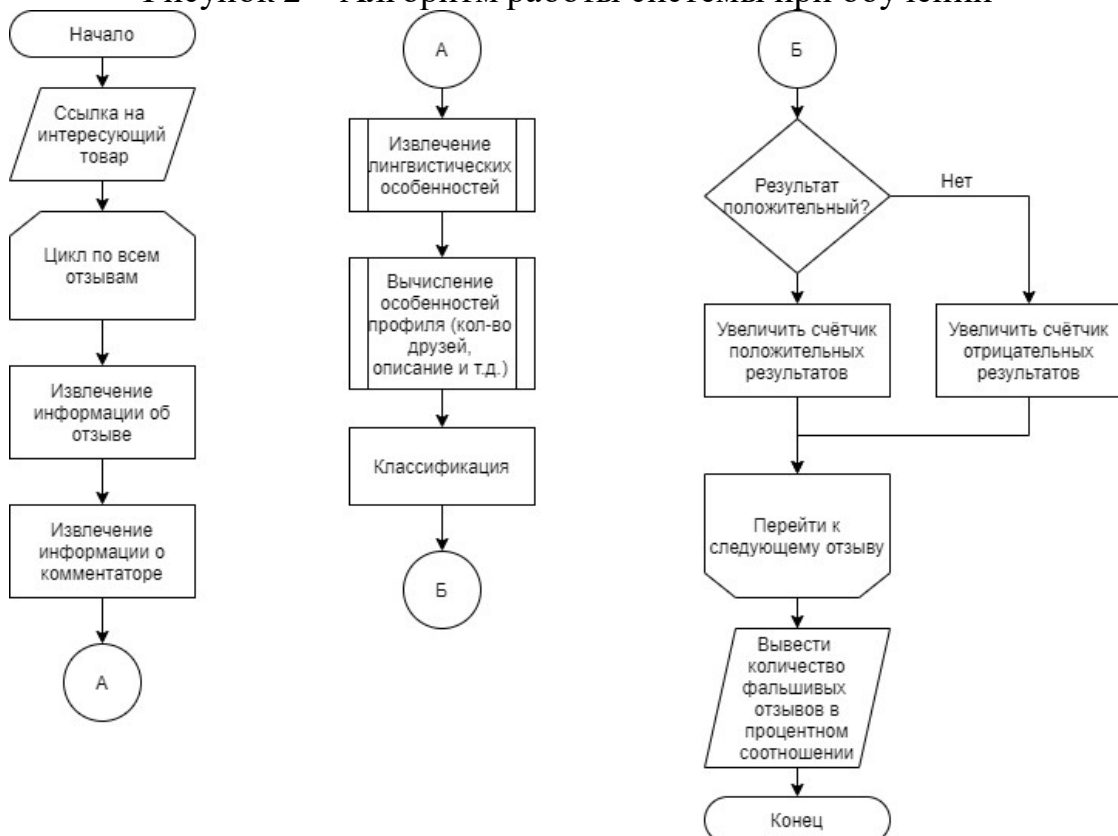


Рисунок 3 – Алгоритм работы системы при классификации



Итогом исследования является программный комплекс, осуществляющий классификацию отзывов на определённый продукт. И на основе результатов пользователь сможет сделать вывод о качестве рецензий на товар.

### Литература

1. Sung Ho Na, Impact of online consumer reviews on product sales: Quantitative analysis of the source effect [Текст] / Sung Ho Na, Soon yong Bae, and Lee Kyeong Son // Applied Mathematics & Information Sciences – 2015. – Вып. 9. – с. 373-387;
2. Как устроена фабрика заказных отзывов [Электронный ресурс] – URL: <https://texterra.ru/blog/kak-ustroena-fabrika-zakaznykh-otzyvov-opyt-30-kopirayterov.html> (Дата обращения 04.04.2021);
3. PU learning – Dealing with a negative class hidden in unlabeled data [Электронный ресурс]. – URL: <https://towardsdatascience.com/pu-learning-e2059f4f9b52> (Дата обращения 04.04.2021).

А.Р. Белов

## ПРЕДИКТИВНАЯ ЗАГРУЗКА МОДУЛЕЙ И КОМПОНЕНТОВ ВЕБ-ПРИЛОЖЕНИЙ

(Самарский университет)

Машинное обучение открывает возможности для будущего прогресса во многих областях информационной сферы. И, одно из новейших направлений в веб-сфере – адаптация интерфейса приложения к поведению пользователя. Это делается для улучшения пользовательского опыта. Но, этот аспект увеличивает количество составных компонентов приложения, тем самым и размер кодовой базы, а чем больше её объём, тем медленнее идёт загрузка по сети. Таким образом, мы неявно ухудшаем другие показатели веб-приложения, что недопустимо.

Для решения проблемы медленной загрузки приложений по сети начинает использоваться предиктивная загрузка компонентов, которые потенциально могут понадобиться пользователю. Эта концепция стала возможной при помощи комбинации предсказательной аналитики и «цепей Маркова». Основная идея данного направления – прогнозировать действия и перемещения пользователя в рамках конкретного веб-приложения и загружать только те части, которые могут понадобиться пользователю при взаимодействии.

Как правило, структура разрабатываемого приложения является субъективной со стороны архитектора и порой не учитывает по какому принципу приложение разделяется на модули и компоненты. Из-за этого, те ресурсы, которые могут понадобиться пользователю в интерфейсе в первую очередь не всегда учитываются при проектировании системы. Масштабирование принятия решений по вопросам востребованности компонентов для каждого конкретного слу-