

С.Г. Отгонбаатар, В.В. Новый¹

*Витебский государственный университет им. П.М. Машерова,
г. Витебск, Беларусь*

¹SPIN-код: 8336-8713

МОНИТОРИНГ И КЛАССИФИКАЦИЯ СООБЩЕНИЙ СОЦИАЛЬНОЙ СЕТИ «ТЕЛЕГРАММ»

В наше время социальные сети стали важной частью повседневной жизни людей. Они не только используются для общения друг с другом, но и для получения информации, развлечения, а в некоторых случаях даже для заработка.

Эта эволюция распространения информации впечатляет и одновременно может представлять собой опасность. Поэтому разработка программного обеспечения для мониторинга социальных сетей является актуальной задачей в современном цифровом мире.

Большинство современных социальных сетей предлагают встроенные средства мониторинга групп, каналов. Однако их возможности часто бывают либо ограничены, либо вовсе не подходящими для решения каких-то конкретных задач. Например, мессенджер «Телеграмм» не предоставляет встроенные функции для мониторинга статистики канала пользователю, у которого нет доступа к учетной записи администратора этого канала. Задачи мониторинга варьируются, от отслеживания активности пользователей, таких как реакции на сообщения, количество просмотров у определенных публикаций, количество их комментариев, до частоты публикаций, оценки эмоциональной реакции на определенную публикацию и так далее. Эти данные имеют большое значение как с точки зрения изучения мнения по определенному вопросу, реакции на рекламные материалы, так и для выявления фактов противоправного поведения или мошеннических действий.

Решение указанной проблемы включают в себя множество подзадач, таких как сбор данных, статистический анализ и представление полученной информации, определение эмоциональной окраски сообщений и комментариев к ним. Сбор данных и их ана-

лиз может быть возложен на конкретный программный продукт, который работает со встроенными возможностями данного мессенджера, расширяя их функционал.

Для выполнения этих задач, в приложении был использован язык программирования Python и библиотеки Telethon [1], Pyrogram [2] для взаимодействия с мессенджером «Телеграмм».

На данный момент приложение может круглосуточно получать и сохранять данные о каналах, такие как: ID канала, имя канала, количество подписчиков канала, общее количество сообщений, которое было отправлено. Также программа получает информацию о сообщениях, которые были отправлены в каналах такие как: ID сообщения, ID канала, дата отправления, количество просмотров всего, количество переходов, количество оставленных подписчиками комментариев и ID этих пользователей, количество и тип реакции на сообщения. Дополнительно возможна фильтрация сообщений на основе ключевых слов. В дальнейшем развитии проекта планируется добавления способов дополнительной обработки данных для улучшения анализа активности отслеживаемых каналов и классификация сообщений пользователей и реакции на них по эмоциональному тону на основе методов глубокого обучения.

Заключение. В результате выполнения работы был разработан базовый модуль получения информации из заданных каналов «Телеграмм» и их предварительной классификации для встраивания в приложения автоматизации анализа данных социальных сетей.

Литература

1. Telethon's Documentation. URL: <https://docs.telethon.dev/en/stable/#>.
2. Telegram MTProto API Framework for Python. URL: <https://docs.pyrogram.org/>.