2. Рощин Д. VPN на основе протокола PPTP: как повысить безопасность? / Д. Рощин. — Текст : электронный // Системный администратор. — 2007. — № 4. — URL: https://www.elibrary.ru/download/elibrary_20394452 _21186501.pdf (дата обращения: 04.05.2022).
3. Киселев С. С. Протокол SSL/TLS как основа технологии VPN / С. С. Киселев, С. В. Пилькевич, Н. В. Салко. — Текст : электронный // Труды военно-космической академии имени А. Ф. Можайского. — 2011. — № 630. — URL: https://www.elibrary.ru/download/elibrary_21050527_ 24730965.pdf (дата обращения: 05.05.2022).
4. Продажи коммутаторов Ethernet корпоративного уровня за год. — Текст : электронный // https://www.ixbt.com : новостной ресурс. — 2021. — URL: https://www.ixbt.com/news/2021/09/10/ethernet-10-8.html (дата обращения: 14.05.2022).
5. Cisco FlexVPN DMVPN, Part 1 — Overview and Design / M. Kashin. — URL: https://packetpushers.net/cisco-flexvpn-dmvpn-high-level-design (date of the application: 10.05.2022). — Text : electronic.
6. IKEv2 VPN Protocol Explained / Miklos Zoltan. — URL: https://www.privacyaffairs.com/ikev2-vpn-protocol (date of the application: 11.05.2022). — Text : electronic.

Ю. Ю. Михайлов, Д. С. Топкасов, Т. И. Паюсова
Тюменский государственный университет, г. Тюмень

УДК 004.056

РАЗРАБОТКА ПРОГРАММНОГО КОМПЛЕКСА ДЛЯ ФИЛЬТРАЦИИ НЕЖЕЛАТЕЛЬНОГО КОНТЕНТА В ИНТЕРНЕТЕ

Аннотация. *В статье обоснована необходимость фильтрации контента, а также перечислены требования, которым должен удовлетворять программный комплекс для фильтрации нежелательного контента в Интернете.*

Ключевые слова: *обработка естественных языков, фильтрация контента, информационная безопасность, законодательство РФ, нежелательный контент.*

Введение. Информационная безопасность — это защита не только информации, но и от нее. Точнее, от ее воздействия.

Эта сфера регулируется различными федеральными законами и нормативными правовыми актами Российской Федерации.

Но в первую очередь, основания для этого обозначены в главном документе страны — Конституции России.

А именно, в 71 статье, где говорится о безопасности как личности, так и общества и государства в целом при использовании информационных технологий, информации и данных. В том же документе указана недопустимость пропаганды или призывов к дискриминации по каким-либо признакам [1].

Важными в этой сфере так же являются федеральные законы, связанные с информацией и информационными технологиями, о защите детей от вредной для развития и здоровья информации, о наркотиках, экстремизме, азартных играх и многом другом.

В реальности, несмотря на такую большую регулирующую базу (перечисленную выше), в сети Интернет можно с легкостью (даже случайно) наткнуться на нежелательные или запрещенные материалы. Просто осуществляя поиск в браузере или листая ленту новостей в социальных сетях. При этом самое опасное, когда на такую информацию натыкаются дети, потому что они легко могут поддаться ее влиянию [2].

Проблема исследования. Для спасения от такой потенциально опасной информации можно использовать существующие системы родительского контроля или фильтрации неприемлемого контента. Но такие решения слишком кардинально подходят к этому вопросу из-за логики осуществления блокировки. На основе базы данных с сайтами, к ним закрывается доступ пользователю.

В такой ситуации не совсем понятно, какими принципами руководствовались их создатели. Например, в случае с социальными сетями из-за одного комментария либо он останется доступным для ознакомления пользователями, либо будет заблокирован весь ресурс. Такое решение не совсем практично и удобно, что с большей вероятностью повлечет отказ пользователя от его использования.

Поэтому авторы данной работы рассмотрели вариант разработки программного комплекса, который будет скрывать нежелательный контент лишь содержащими его блоками текста.

Это позволит пользователям находиться на всех ресурсах, но не будут давать возможности ознакомиться с потенциально опасной и вредной информацией.

Материалы и методы. В зависимости от области использования и конечного пользователя, требования, предъявляемые к ПО для фильтрации контента, могут различаться.

Рассмотрим сценарий, когда родителю необходимо защитить своего ребенка от такой информации в Интернете, которая бы могла повлечь вред психике ребенка. Здесь важно учитывать специфику поставленной задачи. Так, в данном случае, система по фильтрации контента должна:

- быть проста в установке;
- не требовать от родителя особых умений в обращении с компьютером;
- предусматривать защиту от возможных способов обхода установленных ограничений.

Первые два пункта подразумевают, что процесс развертывания и настройки приложения должен быть автоматизирован, происходить с минимальным вмешательством пользователя и быть интуитивно понятным.

Для реализации третьего пункта необходимо предусмотреть возможные методы обхода защиты и принять соответствующие меры, чтобы не допустить реализации этих методов.

Например, для фильтрации контента на веб-страницах может быть разработано расширение для браузера, которое отправляет найденные на странице текста на сервер для анализа. Для большинства современных пользователей установка расширения является простой задачей, но все еще есть вероятность, что для некоторых родителей этот процесс будет представлять сложность.

Также стоит брать в учет меры по защите от обхода средства родительского контроля. Так, ребенок может попытаться банально отключить или удалить расширение. Чтобы не допустить этого, можно прибегнуть к использованию встроенных в операционную систему средств защиты.

Если рассматривать браузер Google Chrome, то возможно скачать с официального сайта шаблон групповых политик, который можно импортировать на устройство и в политике указать расширение, фильтрующее контент, как устанавливаемое принудительно [3].

В таком случае отключить расширение будет невозможно. Доступ к групповым политикам есть только у учетной записи администратора устройства, поэтому, ребенок должен использовать компьютер из-под учетной записи обычного пользователя.

Еще один метод обхода — использование другого браузера. В таком случае необходима настройка белого списка приложений. В ОС семейства Windows такой функционал можно реализовать, используя встроенное средство — AppLocker [4].

После того, как защита от обхода СРК была предусмотрена, необходимо автоматизировать реализацию всех принятых мер, чтобы после установки ПО пользователю не пришлось все настраивать вручную.

Рассмотрим другой случай — использование программного комплекса в других средствах защиты информации (например, Data Leak Prevention, или сокращенно DLP). Тогда должны выполняться следующие требования:

- ПО должно иметь универсальный интерфейс, позволяющий свободно обмениваться данными с разрабатываемым СЗИ;
- ПО должно быть просто в развертывании и не должно создавать конфликтов при параллельном запуске с другими приложениями;
- программный комплекс должен быть гибок в использовании, чтобы была возможность перенастроить его для выполнения конкретной задачи.

Кроме того, независимо от сценария использования, программный комплекс должен обеспечивать быстрый анализ текст с наибольшей возможной точностью. В случае использования машинного обучения для анализа текста, на указанные параметры влияет выбор модели машинного обучения и правильная предобработка текста.

Опытным путем авторами данной работы была выбрана модель логистической регрессии для решения задачи категоризации текста. Выбор производился из статистических моделей путем проведения оценки производительности модели на тестовом наборе данных. С использованием самостоятельно реализованных методов

обработки наилучший результат показала упомянутая выше модель, продемонстрировав точность в 94,14%, как можно увидеть на рис. 1.

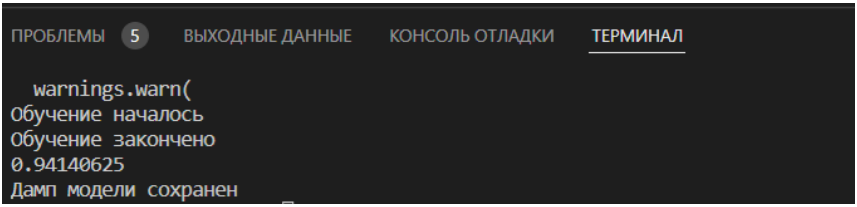

Рис. 1. Результат оценки производительности модели

Используя модель логистической регрессии, возможно спрогнозировать вероятность того, что тот или иной текст относится к некоторой категории.

На рис. 2 представлен график логистической функции $f(x) = \frac{1}{1+e^{-x}}$. В данном случае кривая разделяет плоскость на две части, каждая из которых включает в себя объекты, принадлежащие к одному из двух классов.

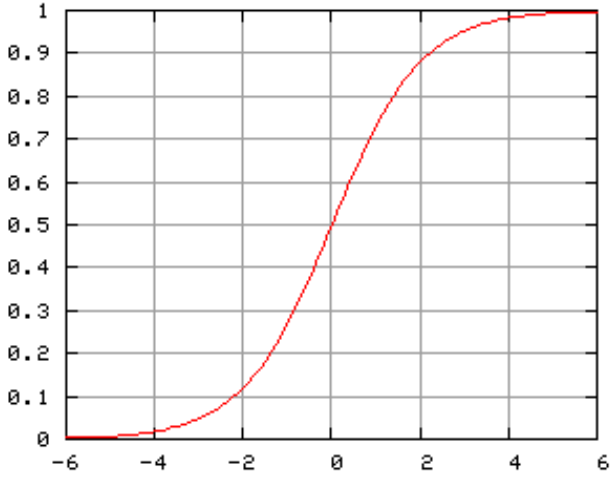

Рис. 2. График логистической функции

Обучение модели происходит подбором таких коэффициентов, при которых кривая разделяет пространство так, чтобы каждый объект из обучающей выборке принадлежал своему классу.

Результаты. Для достижения поставленной задачи была разработана серверная часть, выполняющая категоризацию текста с использованием логистической регрессии, а также клиентская часть, которая собирает текст с веб-страницы и отправляет его на сервер для анализа. Пример работы клиентской части, реализованной в виде браузерного расширения, можно увидеть на рис. 3.

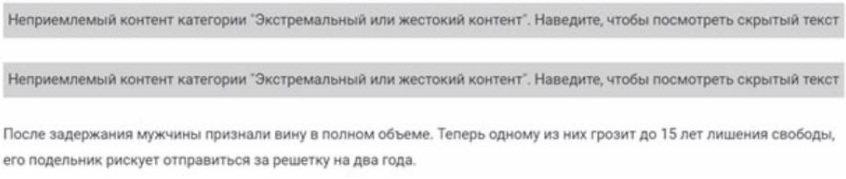

Рис. 3. Пример блокировки нежелательного контента на веб-странице

При отправке текста клиентской частью на сервер, последний обрабатывает текст и возвращает вероятности принадлежности текста к каждой из категорий, которые можно увидеть на рис. 4.

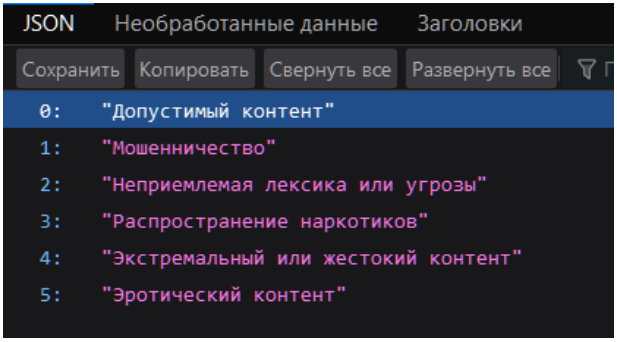

Рис. 4. Классы, используемые для категоризации контента

Благодаря использованию оптимизационного алгоритма стохастического градиентного спуска, возможно дообучение модели, с использованием новых данных. Как можно увидеть на рис. 5,

в клиентской части программного комплекса также был реализован функционал, позволяющий пользователю сообщить о некорректной работе категоризатора и указать правильную категорию.

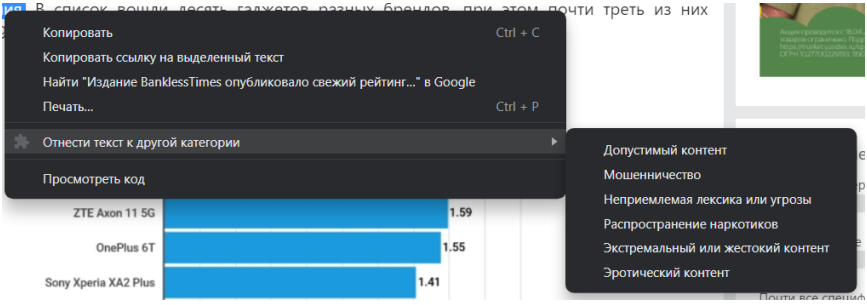

Рис. 5. Процесс корректировки категории текста

Заключение. Таким образом, в данной статье была обоснована необходимость фильтрации контента с точки зрения Конституции Российской Федерации, законодательства, а также других нормативных правовых актов нашей страны.

Рассмотрены требования, которые должны быть учтены при разработке программного комплекса для фильтрации нежелательного контента в Интернете, и описаны основные аспекты реализации.

СПИСОК ЛИТЕРАТУРЫ

1. Конституция Российской Федерации : принята 12 декабря 1993 г. Официальный текст. — Москва : Омега-Л, 2021. — 39 с. — Текст : непосредственный.
2. Радченко Л. Е. Влияние информации на нравственное, психическое здоровье и воспитание подрастающего поколения / Л. Е. Радченко. — Текст : непосредственный // Материалы Афанасьевских чтений. — 2016. — № 1 (14). — С. 215-223.
3. Jillepalli A. A. Enterprise-level hardening of web browsers for microsoft windows / A. A. Jillepalli [et al.]. — Text : direct // Internetional Journal of Computing and Digital Systems. — 2018. — № 5. — P. 261-274.
4. Митрошина Е. В. AppLocker как средство обеспечения информационной безопасности / Е. В. Митрошина. — Текст : непосредственный // Контентус. — 2016. — № 8 (49). — С. 163-166.