

## **МЕТОДЫ ЗАЩИЩЕННОСТИ СЕТЕЙ ДЕТСКИХ ОБРАЗОВАТЕЛЬНЫХ УЧРЕЖДЕНИЙ ОТ ИНФОРМАЦИИ, КОТОРАЯ МОЖЕТ ПРИЧИНИТЬ ВРЕД ЗДОРОВЬЮ И РАЗВИТИЮ ДЕТЕЙ**

*И.В. Алексеев, А.В. Лукьянов, А.Н. Носков, С.Е. Косенков*

*Ярославский государственный университет им. П. Г. Демидова*

Ярославль

Машинное обучение возникло на стыке прикладной статистики, оптимизации, дискретного анализа и за последние 30 лет оформилось в самостоятельную математическую дисциплину. Методы машинного обучения составляют основу ещё более молодой дисциплины - интеллектуального анализа данных (data mining). Машинное обучение это не только математическая, но и практическая, инженерная дисциплина.

Чистая теория, как правило, не приводит сразу к методам и алгоритмам, применимым на практике. Чтобы заставить их хорошо работать, приходится изобретать дополнительные эвристики, компенсирующие несоответствие сделанного в теории предположений условиям реальных задач. Практически ни одно исследование в машинном обучении не обходится без эксперимента на модельных или реальных данных, подтверждающего практическую работоспособность метода.

В мире и в России идет активная работа по разработке, оптимизации и верификации методов информационного поиска, к одному из которых относится автоматическая классификация или кластеризация документов. Несомненно, что данное направление является важнейшим направлением научно-прикладных исследований, поскольку приложения на основе таких систем способны существенно улучшить работу огромного количества отраслей науки и экономики, а объемы информации испытывают экспоненциальный рост, в частности в сети Интернет.

Предлагаемое нами исследование направлено на разработку или адаптацию существующих алгоритмических моделей к применению в области защиты подрастающего поколения от информации, которая может нанести вред их здоровью и развитию.

В области исследований, адаптации и независимой оценки методов автоматической классификации, в частности, очень заметен постоянно действующий Российский семинар по Оценке Методов Информационного Поиска (РОМИП), а также Университетская информационная система РОССИЯ (УИС РОССИЯ), которая создана и поддерживается, как база электронных ресурсов для исследований и образования в области экономики, социологии, политологии, международных отношений и других гуманитарных наук. За рубежом активность работ по информационному поиску поддерживается на высоком уровне, в частности, при помощи The Text REtrieval Conference (TREC), регулярно проводимой институтом стандартизации (NIST) и американским министерством обороны. Цель – содействие исследованиям по системам информационного поиска, в частности путем предоставления инфраструктуры для тестирования алгоритмов, путем усиления общения между экономикой, наукой и бизнесом, обеспечивая обмен идеями и обеспечение узких предметных областей для обеспечения предметной области задачам информационного поиска.

Данный проект направлен на проведение исследований и применение методов машинного обучения для категоризации интернет контента с целью обеспечения защиты детей от негативного воздействия на их психическое и физическое здоровье информации, которая может нанести такой вред. Система будет работать как под управлением экспертов, для которых задача максимально автоматизируется, путем предложения готовых вариантов и автоматической фиксации результатов, а также в полностью автоматическом режиме, не исключая, однако, коррекцию результата вручную экспертом.

Мы использовали и при необходимости модифицировали под данную задачу определенную совокупность методов и подходов: метод автоматической классификации текстовых документов, используя разные подходы и технологии, методику применения ранее созданных черных и белых списков, а также списков категорий, алгоритмы анализа графической и видео-информации.

Результирующие (мета) данные хранятся в СУБД в текстовой и машинно-читаемой форме и предназначены для распространения и применения в фильтрующих контент системах - шлюзах, прокси серверах, устанавливаемых непосредственно в сетях получателей контента.

Сравнение существующих алгоритмов производилось на основании собственного опыта работы с большими объемами данных. Приведем достоинства и недостатки некоторых из вышеперечисленных алгоритмов.

Алгоритмы построения деревьев решений. К достоинствам метода относится простота интерпретации, высокая точность, возможность построения масштабируемых алгоритмов (которые способны работать на сверхбольших базах данных), быстрый процесс обучения, возможна обработка пропущенных значений, возможность работы с категориальными типами данных. К недостаткам: проблема повторений некоторых

частей дерева (replication problem) при построении деревьев решений, создание в некоторых случаях трудно интерпретируемых правил и переобучение классификатора на отдельных обучающих коллекциях

Метод опорных векторов. К достоинствам метода относится то, что методом решается задача квадратичного программирования, имеющая единственное решение (вместо многоэкстремальной задачи) — методы оптимизации в этом случае существенно более эффективны; автоматически определяется число нейронов скрытого слоя — оно равно числу опорных векторов (преимущество перед нейронными сетями); принцип оптимальной разделяющей гиперплоскости приводит к максимизации ширины разделяющей полосы между классами (следовательно, к более уверенной классификации), в то время как градиентные нейро-сетевые методы выбирают положение разделяющей гиперплоскости произвольным образом. Недостатки метода таковы: неустойчивость по отношению к шуму в исходных данных; для классификации используется не все множество объектов, а лишь их небольшая часть, которая находится на границах

Метод k ближайших соседей. Достоинства: простота использования полученных результатов, простота интерпретации, отсутствие предварительной обработки данных. Недостатки: существует сложность выбора метрики близости, вычислительная трудоемкость (если не применить соответствующие методы оптимизации), неэффективный расход памяти.

Метод байесовских сетей. Достоинства: простота реализации, низкие вычислительные затраты при обучении и классификации; в модели определяются зависимости между всеми переменными — это позволяет легко обрабатывать ситуации, в которых значения некоторых переменных неизвестны; байесовский метод позволяет естественным образом совмещать закономерности, выведенные из данных, и, например, экспертные знания, полученные в явном виде; использование байесовских сетей позволяет избежать проблемы переучивания (overfitting), то есть избыточного усложнения модели, что является слабой стороной многих методов (например, деревьев решений и нейронных сетей). К недостаткам относится низкое качество классификации; невозможна непосредственная обработка непрерывных переменных — требуется их преобразование к интервальной шкале, чтобы атрибуты были дискретными (однако такие преобразования иногда могут приводить к потере значимых закономерностей).

Нейронные сети (многослойный персептрон). Достоинства: достаточно высокая эффективность, легко реализуется на вычислительных устройствах с параллельной архитектурой, высокая степень общности. К недостаткам методики отсутствие гарантии сходимости, возможно попадание в многочисленные локальные минимумы, приходится заранее фиксировать число нейронов скрытого слоя, склонность к переобучению.

Далее была проведена разработка системы анализа эффективности алгоритмов для использования ее в случае классификации неструктурированных HTML и текстовых документов. Эта система реализована в виде набора сценариев на языках bash и Perl и с ее помощью проведены дальнейшие исследования.

Были созданы тестовые наборы описаний, базирующиеся на реальной информации, которая содержится в общедоступных сетях. Для этого были привлечены люди - эксперты, которые путем применения ряда правил к экспертной оценке этой информации провели первичную классификацию и в дальнейшем контролировали корректность ее категоризации при применении автоматизированных методов. Для этой цели были разработаны автоматизированные рабочие места экспертов с удобными интерфейсами для выполнения этой работы.

## ЛИТЕРАТУРА

1. Алексеев И.В., Благовещенский И.А., Захарова М.Н., Лукьянов А.В. Разработка мобильного варианта среда управления ИТ Парком и взаимодействия его участников // Интернет и современное общество: Сборник тезисов докладов. Материалы XIV Всероссийской объединенной конференции «Интернет и современное общество». Санкт-Петербург, 12–14 октября 2011 г. СПб., 2011. Т. 1. С. 5—6.
2. Алексеев И.В., Захарова М.Н., Капранов В.А., Лукьянов А.В. Разработка программного решения для мониторинга и контроля трафика в распределенной информационной среде контекстно-зависимой образовательной мультимедиа информации с использованием интернет протокола IPv6 // Интернет и современное общество: Сборник тезисов докладов. Материалы XIV Всероссийской объединенной конференции «Интернет и современное общество». Санкт-Петербург, 12–14 октября 2011 г. СПб., 2011. Т. 1. С. 6—7.
3. Алексеев И.В. Межсетевой экран «Интернет Контроль Сервер» поддерживает высокие стандарты безопасности // Information security. Информационная безопасность. 2012. №2. С. 43.