

Создана нейросеть для поиска фейковых новостей



Группа программистов разработала алгоритмы, помогающие отличать фейковые новости от настоящих. Этому они научили нейронные сети, которые теперь не только не уступают в данной области человеку, но и значительно его превосходят.

Павел Посохов недавно окончил Пятигорский государственный университет по направлению «Интеллектуальные системы в гуманитарной сфере». Идея создать ресурс, который бы проверял новости на достоверность, ему пришла, когда он участвовал в конкурсе RuCode. В финале программистам предложили придумать технологию, способную находить ложные новости.

– Это был кейс от Сбербанка. Моя модель в результате заняла 15-е место, то есть от победителей я оказался далеко. Однако позже рассказал о своей придумке одногруппникам. Их заинтересовала идея, что машинная модель может определять, правдивая новость или ложная, – рассказал Посохов.

Поддержала проект преподаватель университета **Ольга Тимченко**, которая не только вела в вузе несколько дисциплин, но и стала наставником ребят. Она помогала с научными исследованиями, отправляла команду молодых программистов на соревнования и конкурсы.

В итоге важным достижением стало третье место на конкурсе молодых предпринимателей, который прошел минувшей зимой во время ежегодной национальной выставки ВУЗПРОМЭКСПО в Москве. Пятигорским студентам пришлось соревноваться с разработчиками серьезных проектов, в том числе современного

мусороперерабатывающего завода и электромобиля.

– По итогам конкурса мы получили 15 тысяч рублей на всю команду, что далеко не главное. Важнее наша репутационная победа. Мы подали документы на оформление программы для ЭВМ, так как это более простой и доступный способ закрепить права на разработку по сравнению с получением патентов, – пояснил Посохов.

В основе технологии лежат две нейронные сети – на базе Google Colab и GPT-2. Мировые гиганты в области IT предоставляют базу для таких проектов, чтобы энтузиасты на существующих мощностях могли создавать и обучать свои алгоритмы. В качестве фейковых новостей Посохов использовал сообщения на сайте «Панорама» – известного ресурса, где ежедневно публикуются подобные тексты, написанные при этом профессионалами-журналистами. Кстати, именно оттуда материалы брали организаторы первого конкурса, в котором участвовал Павел. Поэтому он также использовал данный любопытный ресурс.

Из новостей сформировали базу, с помощью которой программисты и обучали две нейронные сети. Она состояла из 15 тысяч текстов. Позже Павел добавил разделение по рубрикам. Процесс обучения проходил достаточно долго – одна из сетей усвоила материал за сутки, а вторая – за шесть дней.

Сейчас на запрос пользователя обе программы отвечают всего за три секунды. Их точность значительно выше, чем у человека. Разработчики даже проводили сравнительный тест с участием одноклассников. В 100 процентах случаев машина оказалась сообразительнее. Однако испытуемым, как и нейросети, давали только статьи без каких-либо пояснений и дополнительной информации, так как важно было определить подлинность новости по синтаксису и семантике.

– В перспективе возможно использовать наши алгоритмы, чтобы проверять на достоверность сообщения в социальных сетях. Но тут проблема в том, чтобы найти базу данных для обучения нейросети. Благодаря «Панораме» в интернете есть большой

массив ложных новостей. А с сообщениями и постами пользователей в соцсети сложнее – необходимо разрешение людей и доступ к их переписке, – продолжает Павел.

Можно, кстати, обрабатывать и другой медиаконтент – видео, аудио и фотографии. Конечно, для их анализа потребуются уже иные модели, но создать их вполне реально. К слову, проектом уже заинтересовались, но кто именно, Посохов пока не говорит.

Фейкньюс активно используют многие крупные игроки разных рынков. Ведь слухи способны оказать огромное влияние на инвесторов и покупателей ценных бумаг. Классикой жанра стала схема Pump&Dump («накачка и сброс»), которая позволяет быстро увеличить стоимость акций, а затем продать их с хорошей прибылью.

Так, в апреле 2001 года представители американской энергетической корпорации, находящейся на грани банкротства, рассылали потенциальным покупателям сообщения о том, что акции компании будут расти, аналогичные заявления они размещали в соцсетях. В итоге манипуляторы смогли продать акции по завышенной цене на сумму более миллиарда долларов, а затем корпорация стала банкротом. Одним из наиболее известных примеров фейковой новости в бизнесе стал фальшивый пресс-релиз об отставке финансового директора французской строительной компании Vinci. Публикация привела к обрушению акций предприятия на 20 процентов.

– Проверка информации на достоверность требуется сейчас везде. Основные клиенты в данном случае – коммерческие структуры, банки, фондовые рынки и биржи. Информация способна обрушить акции, а позже выяснится, что новость была фейком. Такие алгоритмы, как наши, помогут компаниям обезопасить себя в подобных случаях, – заключил Посохов.

Проект выпускников из Пятигорска – одно из перспективных направлений. Сейчас на рынке нет крупных коммерческих предложений в этой сфере. Возможно, конечно, что кто-то

предлагает отдельным заказчикам закрытый продукт, но публично об этом не сообщается. Поэтому сейчас все в руках разработчиков. От того, насколько качественно и быстро они смогут реализовать свой проект, зависит успех всего дела.

По тексту Романа Кияшко

