Рис. 1: ReCAPTHCA первой версии

Вы наверняка помните то время, когда интернет сайты постоянно заставляли вас разбирать, что написано на данных картинках. Сайты требовали это, чтобы избежать злоупотребления функционалом сайта со стороны ботов, предупреждая таким образом регистрацию множества "пользователей" отправку кучи сообщений или злоупотребления иным функционалом сайта или попытки перебора паролей.

Однако несколько лет назад вдруг подобные надписи пропали. Почему? Правоохранительные органы пересажали всех злоумышленников? Или они сами решили уйти в монастырь замаливать грехи. Вряд ли. Произошло то, что всегда происходит в таких случаях: всегда найдётся кто-то, кто захочет злоупотребить вашей системой во вред вам или во вред пользователям, ради кражи их данных, создания повышенной нагрузки на сервера и соответственно замедление или вообще прекращения их работы и других целей. Подобные люди никуда не делись и не денутся.

С появлением сайтов появились и скрипты, которые имитировали действия реальных пользователей, не являясь ими, создателям сайтов это не понравилось, поэтому они попытались отличить ботов от реальных пользователей. Первоначальной идеей было создать задание, с которым бы легко справлялись люди, но плохо справлялись роботы, такой задачей стала задача распознавания символов. Действительно для вас не составляет никакого труда разобрать, что значат все эти закорючки, линии и кружочки из которых состоят слова в этой книге (по крайней мере, я очень на это надеюсь). Но до недавних пор эта задача была для компьютеров совсем не тривиальной.

Так и появилось то, что называется капчей. Капча (от CAPTCHA — англ. Completely Automated Public Turing test to tell Computers and Humans Apart — полностью автоматизированный публичный тест Тьюринга для различения компьютеров и людей) — компьютерный тест, используемый для того, чтобы определить, кем является поль-

зователь системы: человеком или компьютером. Термин появился в 2000 году. Основная идея теста: предложить пользователю такую задачу, которая с лёгкостью решается человеком, но крайне сложна и трудоёмка для компьютера. По состоянию на 2013 год, каждый день пользователями по всей планете вводится примерно 320 миллионов «капчей». [1]

Но разумеется злоумышленники (и не обязательно злоумышленники, но и многие исследователи из научного интереса) не сдались и решили попытаться обойти защиту. Мы рассмотрим хронологию обхода ReCAPTCHA - пожалуй самую известную реализацию, изображение которой и было представлено выше, на рис. 1.

- 1 августа 2010 Chad Houck добился точности "взлома" капчи в 31.8%.

- 26 мая 2012 Adam, C-P and Jeffball - 99.1%. Следует уточнить, что ими была взломана аудио версия, при этом за несколько часов до презентации Google выпустил обновление точность упала до 60.95%.

- 27 июня 2012 - мексиканские студенты Claudia Cruz, Fernando Uceda, and Leobardo Reyes добились точности 82%.

- Август 2017 - исследователи из университета Мериленда: Kevin Bock, Daven Patel, George Hughey, Dave Levin - 85.15%.

Исследователи утверждали, что Гугл постоянно менял свою капчу, при этом часто возвращаясь к предыдущей версии[**wiki:captcha_is_hard**]. Однако, всё же сложность капчи постепенно росла в конце концов требование, с которым она создавалась, — быть сложной в решении для программ и лёгкой для людей, стало выполняться всё хуже. Так в августе 2012 — более 90% пользователей находят капчу сложной для ввода[**wiki:captcha_is_hard**]. В мае 2016 была прекращена поддержка данного типа капчи, то есть она перестала обновляться, в 31 марта 2018 выключена окончательно.[**google:choose_recaptcha**]

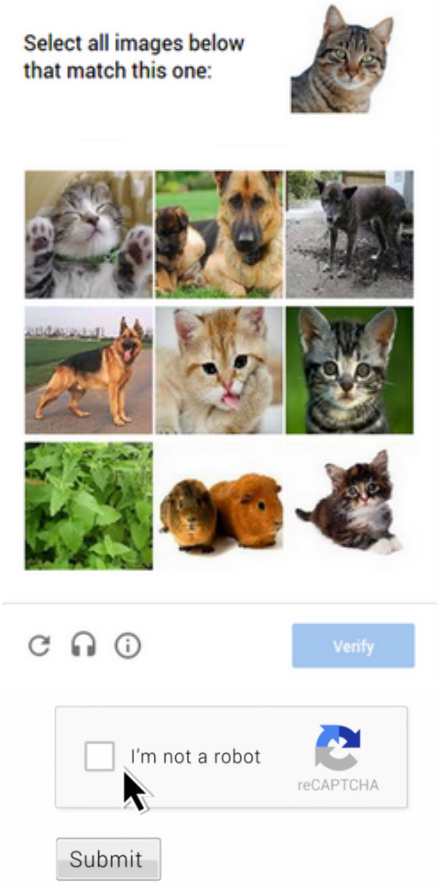

Рис. 2: ReCAPTHCA второй версии

После этого гугл поменял тип своей капчи: теперь необходимо выбрать картинки с котиками, ибо все любят котиков. Таким образом гугл убивает сразу двух зайцев: отличает людей от ботов и создаёт обучающую выборку для своих алгоритмов распознавания текста и картинок. Перед этим была промежуточная версия, в которой выводилась фотография номера дома, соответственно пользователю нужно было разобрать этот номер, во славу google.maps.

Сейчас гугл развивает систему, которая анализирует действия пользователей: движения мыши, клики, нажатия клавиш. Люди не могут похвастать монотонностью действий, если боту нет смысла вести курсор прерывисто или не по прямой, нажимать клавиши с разной скоростью, а не середине ввода вообще отвлечься на звонок или сходить за чаем (в принципе, можно вообще не водить мышью или эмулировать физические нажатия клавиш, а использовать JS для совершения действий, но это уже совсем другая история). Создатели ботов тут же стали изменять свои алгоритмы, чтобы быть более человекоподобными (и автору самому приходилось этим заниматься), разумеется гугл тоже не стоит на месте, с каждым разом обходить защиту становится всё сложнее. Следует сказать, что эта капча, пожалуй, является самой дружелюбной к обычным пользователям, а мы ведь не хотим отпугивать пользователей необходимостью постоянно вводить свою электронную почту, sms коды подтверждения, и выбирать на каких картинках изображено что либо (кроме котиков, разумеется, на котиков всегда прият-

но смотреть). Тут же человек сразу начинает пользоваться сайтом, скрипты в фоне анализируют его действия и их характер и в случае подозрения выводят уже обычную капчу или блокируют активность.

Как мы видим, алгоритмы машинного обучения заставили гугл постоянно менять свою систему, чтобы добиться необходимой эффективности защиты, во многом добиваясь этого использованием тех же алгоритмов: для того, чтобы распознавать, на какой именно части панорамы улицы находится номер улицы, и, следовательно, какую именно, часть следует отослать пользователям на проверку, и вывести метрики для отличия ботов от людей.

Рассмотрим ещё несколько сфер, в которых сейчас первенство за системами машинного обучения.

Шахматы

Пожалуй, первый широко известный случай победы машинного интеллекта над человеком является пример шахмат.

- В феврале 1996 года Гарри Каспаров победил шахматный суперкомпьютер Deep Blue со счетом 4-2. Этот матч выдающийся тем, что первую партию выиграл Deep Blue, автоматически став первым компьютером, победившим чемпиона мира по шахматам в турнирных условиях.

- В мае 1997 года Deep Blue II выигрывает матч у Гарри Каспарова со счётом $3\frac{1}{2} : 2\frac{1}{2}$.

- В 2000 году коммерческие шахматные программы Junior и Fritz смогли свести в ничью матчи против предыдущих мировых чемпионов Гарри Каспарова и Владимира Крамника.

- В ноябре-декабре 2006 года чемпион мира Владимир Крамник играл с программой DeepFritz. Матч закончился выигрышем машины со счётом 2-4.

Сейчас же даже у лучших гроссмейстеров нет шансов против среднего компьютера[**chess_human_has_no_chance**]. При этом

программы продолжают играть против друг друга и совершенствоваться в уровне игры и производительности алгоритмов.

Го

Шахматы пали перед алгоритмами и вычислительной мощью компьютеров, но многие годы считалось, что уж в других сферах компьютеры ещё долго не смогут составить конкуренцию человеку. Выражались мысли, что поле для игры слишком большое, что компьютер не сможет точно оценить текущую позицию, все возможные варианты её развития, выбрать оптимальный, у него просто не хватит на это "мозгов"[**go_is_hard**][**go_is_hard2**].

Конечно, доля правды в этих оценках была, притом немалая, однако, насчёт прогнозов, когда компьютеры смогут посоревноваться на равных с человеком, и причины, которые ему пока мешают были в корне неверными.

В Го существует огромное количество возможных вариантов развития игры и позиций, для того, чтобы перебрать их все компьютеру потребуется время больше всей жизни вселенной, но ведь человек тоже не перебирает все возможные варианты, почему это должна делать программа? Человек тоже весьма ограничен в своих способностях к оценке ситуации и вариантов её развития, не перебирает все возможности, выбирая лучшую, не делает этого и компьютер, для того, чтобы обойти человека не нужно играть идеально, нужно играть лишь лучше человека. На простой механический перебор всех вариантов хода в шахматах тоже уйдёт время всей жизни вселенной, компьютер не перебирает все ходы в шахматах, не делает этого и в Го. Пожалуй, самой главной проблемой, препятствующей созданию эффективных алгоритмов, — являлось отсутствие хороших математических моделей, описывающих игру, с развитием и созданием этих моделей, появились и успехи у игровых программ:

- В октябре 2015 года программа AlphaGo, разработанная компанией DeepMind выиграла у трехкратного чемпиона Европы Фань Хуэя (2 профессиональный дан) матч из пяти партий со счётом 4—1. Это первый в истории случай, когда компьютер выиграл в

го у профессионала в равной игре.

- В марте 2016 года AlphaGo победила профессионала 9 дана Ли Седола в четырёх партиях из пяти.
- В мае 2017 года на саммите «Future of Go Summit» AlphaGo выиграла три партии из трёх в мини-матче с одним из сильнейших игроков в мире, лидером мирового рейтинга Эло Кэ Цзе.

Сейчас можно сказать, что все пошаговые настольные игры "сдались"программам, человек может победить в игре, только если правила были специально модифицированы правила, чтобы игроку было легче, или игра недостаточно заинтересовала исследователей, чтобы они разработали эффективные алгоритмы. Следующий шаг — компьютерные игры в реальном времени.

По числу своих возможностей и вариантов развития компьютерные игры, пожалуй, даже превышают настольные игры, однако, как мы выяснили, на примере шахмат и го, это лишь временное неудобство, а не принципиальная проблема, следующей ключевой особенностью является игра в реальном времени, в случае пошаговой игры, у компьютера (и у человека) есть время над обдумыванием хода вплоть до нескольких часов, в случае игры в реальном времени зачастую быстрое, но неоптимальное решение, лучше хорошего, но запоздавшего, а два решения лучше одного. Кроме этих проблем есть и другие, например, неполнота информации, в случае шахмат оба игрока видят всю доску и знают всю информацию об игре, однако, в компьютерные играх существует так называемый туман войны, который скрывает часть игровой карты, которая слишком далеко от существ или зданий игрока, таким образом, всегда приходится действовать в условиях неопределённости: вот в этом лесу, совсем рядом с твоей базой или персонажем уже могли затаиться враги, готовые атаковать, а могли и не затаиться, что есть, как раз таки наоборот, там небольшой разведывательный отряд, уничтожение которого, отбросит противника назад.

И в этом направлении достигнуты реальные успехи, например, на International 2017 (август 2017) программная система побила одного

Рис. 3: Пример программы распознавания номеров

из лучших игроков в мире Dendi в игре Dota 2 один-на-один[**dota:ai_vs_dendi**]. И сейчас разрабатывается бот для игры пять-на-пять, в стандартном режиме для Dota 2[**dota:ai_vs_5**].

Другая команда в это время работает над ботом для компьютерной игры StarCraft[**starcraft:bot_vs_human**][**starcraft:bot_vs_human2**]. Они обе используют одинаковую модель: бот просматривает повторы игр, анализирует кадры и стоит свою стратегию, во время игры используется модель обучения с подкреплением(reinforcement learning). Суть которой заключается в том, что бот принимает решения, а потом получает ответ от системы(игры), правильно он поступил или нет, и таким образом обучается.

Распознавание номеров

Кто из вас когда либо получал так называемые письма счастья из ГАИ, в которых сообщалось, что вы превысили скоростной режим и должны заплатить штраф? Это стало возможным благодаря каме-

рам автоматической фиксации, сейчас мы переживаем настоящий бум, их появляется всё больше и больше. Что сложно назвать удивительным, учитывая, что они работают[**speed_cameras_and_safety_2005**][**spe**
Кроме того, в условиях России они весьма быстро окупаются[**speed_cameras_pr**

Автор предлагает сейчас отвлечься от прочтения книги и задуматься, как бы вы реализовывали функционал камер, фиксирующих превышение скоростного режима, если бы к вам пришла госавтоинспеция с такой просьбой. Задача, вроде бы, не самая сложная, но с наскока её не решишь, на рис. 3 представлен пример работы такой программы в реальном времени, вам предлагается подуматься, что за магия стоит за этими картинками...

Если вы достаточно продвинулись в изучении машинного обучения, вы можете решить просто создать свёрточную нейронную сеть (особый тип нейронных сетей, которые анализируют картинку и находят участки, которые удовлетворяют некоторым критериям), которая сразу выдаст вам последовательность букв и цифр, которые вы потом объедините в номер. Однако такое решение "в лоб" обладает рядом недостатков: вам придётся создать большую обучающую выборку, так как сети будет сложно сразу понять, что является, например, буквой "А" а что обычной кляксой, пятном, бликом или ржавчиной, кроме того, большую проблему создадут следующие авто:

Действительно, на автомобиле можете быть размещено огромное количество разнообразных надписей (как правило рекламных) и наша система не должна срабатывать на них. Как же этого добиться? Обычно в таких случаях задачу решают не сразу в лоб, от начала до конца, а разбивая на подзадачи:

1. Найти на фотографии потока все участки, которые соответствуют номерным пластинам.

2. Улучшить изображение, чтобы облегчить выполнение следующего этапа.

3. Разрезать изображение номера на отдельные буквы и цифры.

4. Распознать их.

Мы более детально рассмотрим эту задачу в следующих главах.

Распознавание лиц

Похожей, но, пожалуй, более сложной задачей является задача распознавания лиц.

Рис. 4: Пример работы системы распознавания лиц

Системы распознавания лиц обладают огромным практическим потенциалом. Алгоритмы, которые обнаруживают лица, встроены во все фотоаппараты и телефоны, устройства, которые не обладают большими вычислительными мощностями, а главное, питающимися от батарей, на основании чего, можно сделать вывод, что эти алгоритмы не несут больших дополнительных расходов. Следующим шагом можно назвать алгоритмы, распознающие улыбки, их уже встраивают в камеры, что избавляет от нужды нажимать на кнопки для фотографирования, достаточно просто улыбнуться[1]. Alibaba недавно запустила новый сервис "Smile to Pay"в ресторанах KFC, зарегистрировавшись в их сервисе Alipay, вы можете подтвердить платёж на кассе простой улыбкой: специальные камеры сфотографируют ваше лицо, которое потом будет распознано, связано с вашим аккаунтом, с которого спишут необходимую сумму[**smile-to-pay**].

[1]В данной книге мы не будем уделять большое внимание данным алгоритмом, если эта тема вам интересна, то своё знакомство вы можете начать и изучения метода Виолы — Джонса, а узнать больше про алгоритмы распознавания улыбки в следующих работах 1, 2, 3, 4 и 5

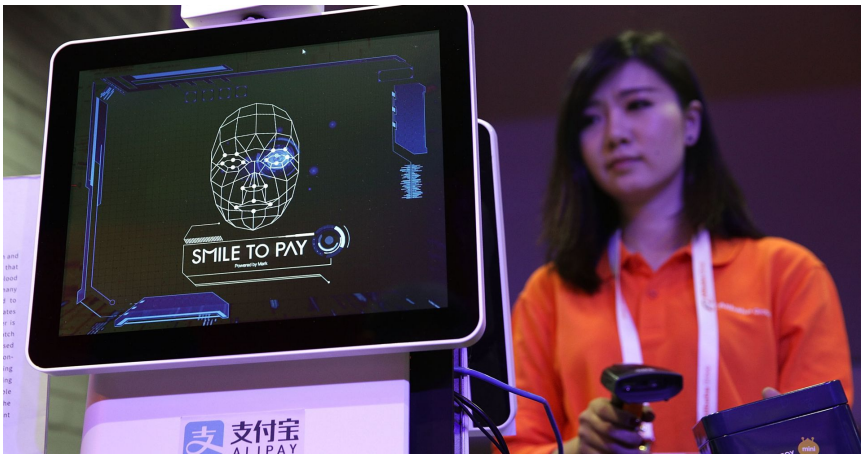

Рис. 5: Система Smile to Pay

Так же в Китае всё большую силу набирают системы слежения с автоматическим распознаванием лиц и возможностью поиска конкретного человека, например в в Гуяйне (провинция Гуйчжоу) в базу внесены все жители города, не составляет большого труда узнать, где сейчас находится человек и об истории его перемещений, о возможностях этой системы говорит тот факт, что в недавнем эксперименте журналиста BBC нашли через 7 минут, после внесения его лица в базу[**china-face-recognition**].