

ИСПОЛЬЗОВАНИЕ НЕЙРОСЕТЕВЫХ ТЕХНОЛОГИЙ ДЛЯ ДЕТЕКТИРОВАНИЯ ЛОЖНОЙ ИНФОРМАЦИИ

Тюрнев Александр Сергеевич

к.т.н., доцент, Иркутский национальный
исследовательский технический университет
tyurnevas@bk.ru

USING NEURAL NETWORKS TECHNOLOGIES TO DETECT FALSE INFORMATION

A. Tyurnev

Summary. The article discusses various methods for identifying false information. One of the new directions in this area is the use of neural network technologies to detect false information.

In the research presented in the article such methods like analysis and synthesis were applied, as well as a systematic approach were used to design the structure of a neural network and the system as a whole.

Based on the results of the research, the structure of the future system was proposed. The system will allow identifying false information with high accuracy. To develop such a system, it will also be necessary to prepare a large dataset for training the proposed models. These datasets can be selected both on real crimes and by questioning students at a university to identify illegal actions.

The results of the research may be of practical interest both for law enforcement agencies and for large organizations, including the administration of educational organizations.

Keywords: false information detection, handwriting expertise, neural networks.

Аннотация. В статье рассматриваются различные методы выявления ложной информации. Одно из новых направлений в этой сфере — использование нейросетевых технологий для выявления ложной информации.

В исследовании преобладают такие методы, как анализ и синтез, а также системный подход к построению структуры нейронной сети и системы в целом.

По результатам исследования была предложена структура будущей системы, которая позволит выявлять ложную информацию с высокой точностью. Для разработки такой системы потребуется также подготовить большой набор данных для обучения предлагаемых моделей. Этот набор данных можно подобрать как на реальных преступлениях, так и при анкетировании обучающихся в вузе для выявления правонарушений.

Результаты исследования могут представлять практический интерес как для правоохранительных органов, так и для крупных организаций, в том числе и для администрации отечественных образовательных организаций.

Ключевые слова: детекция лжи, почерковедческая экспертиза, нейронные сети.

Введение

В последнее время происходит бурный рост систем искусственного интеллекта и нейросетевых технологий в различных сферах. Не исключением является и исследования по выявлению ложной информации. Существуют традиционные методы детектирования лжи. Например, в правоохранительных органах для этой цели используют детекторы лжи (полиграфы). Они дают достаточно точные результаты, однако, и они могут ошибаться. Связано это, скорее всего, с неточностью получения биометрических параметров устройством либо неполным набором инструкций, заложенных в полиграф.

Нейросетевые технологии могут позволить повысить качество и точность детектирования, так как могут выявлять закономерности (например, связанность множества показателей), которые не может «увидеть» человек [1].

Более того, нейронные сети способны к обучению и самообучению. То есть даже если нейросеть «ошиблась» один раз, и мы указали сети на эту ошибку, то в следующий раз при схожей ситуации нейросеть с высокой долей вероятности уже выполнит верное детектирование.

Немаловажным аспектом исследования является выявление заинтересованных лиц. Такими заинтересованными лицами могут выступать работники правоохранительных органов, ведущие расследования преступлений или выполняющие функции дознания. Однако, кроме правоохранительных органов в подобном исследовании могут быть заинтересованы предприятия и организации, причем как государственные, так и частные. В последнее время многие организации прибегают к проверке своих сотрудников на полиграфе [2].

Выявление ложной информации может найти свое применение и в образовательной среде для выявления фактов злоупотреблений сотрудников обучающей организации по отношению к обучаемым, а также правонарушений внутри коллектива обучающихся.

В первую очередь необходимо изучить принципы, на которых строятся различные подходы к детектированию лжи.

Одним из самых важных достижений в области детектирования лжи стало появление полиграфа, который анализирует физиологические процессы.

До появления самого полиграфа предпринималось множество попыток измерять изменения в человеческом теле в момент, когда человек говорит неправду [3].

Например, плетизмограф — прибор для измерения кровенаполнения сосудов и изменений пульса, который позволял фиксировать изменения частоты сердцебиения при предъявлении исследуемому образов, внушающих страх.

Позднее аналогичные исследования привели к созданию гидросфигмографа — устройства, с помощью которого фиксировались изменения кровяного давления допрашиваемого.

В дальнейшем сходные исследования были направлены и на другие физиологические процессы — частота дыхания, изменения систолического давления крови, электрические токи сердца, проводимость кожи и другие.

Все эти исследования привели в итоге к созданию устройства, получившего название детектор лжи или полиграф [4].

Современные полиграфы фиксируют сразу множество параметров, снимаемых с человеческого тела (хотя основные методики остались прежними — анализ ритмов сердца, давления, дыхания, сопротивления кожи).

Однако, и они до сих пор не дают 100 % результатов и могут «ошибаться» [5]. Причем ошибки могут быть разного рода. Прибор может некорректно реагировать на истинные высказывания, относя их к ложным. И наоборот, ложные высказывания прибор может «пропустить». Также существуют различные методики «обмана» полиграфа, так как его можно «сбить с толку».

Полиграф не учитывает психоэмоциональное состояние испытуемого, из-за чего результаты могут искажаться [6]. Тем не менее, постоянно предпринимаются попытки совершенствования технологий детектирования лжи. Так, в последние годы все активнее стали появляться попытки применения нейросетевых технологий и других методов машинного анализа для детектирования лжи.

Анализ существующих разработок по применению систем искусственного интеллекта для выявления ложной информации

В работе [7] были проанализированы существующие системы детектирования лжи и алгоритмы нейросетевого детектирования лжи, выявлены их недостатки. В работе предложены и реализованы алгоритмы нейросетевого полиграфного аппарата.

В работе [8] рассмотрена система детекции ложной информации на основе обработки изображений методом видеоусиления Эйлера для измерения пульса человека.

В работах [9, 10, 11] приводится обзор исследований психофизиологических маркеров, позволяющих идентифицировать правдивую и ложную информацию. Дополнительно рассмотрена технология Eye-tracking как альтернативный или дополнительный метод тестирования. Также отмечены тенденции перспективности использования технологии Eye-tracking совместно с параметрическими системами. Представлены преимущества использования технологий регистрации движений глаз (Eye-tracking) и ритмов сердца. Кроме отечественных публикаций по этой тематике также была рассмотрена статья [12, 13], где в качестве альтернативы предлагается исследование реакции зрачков (расширение зрачков значительно больше, когда участники «выдумывают» ответы, чем когда они говорят правду). Размер зрачка увеличивается с увеличением когнитивной нагрузки. Исследование показало, что ложное воспоминание связано с увеличением размера зрачка и требует большей когнитивной обработки, чем правдивое воспоминание.

Концепция будущей системы

По результату анализа источников был сделан очень важный вывод. Большинство предлагаемых в рассмотренных источниках разработок реализуют только одну методику детектирования, в отдельных случаях — сочетание одного-двух аспектов разных методик. Поэтому качество таких работ сопоставимо с полиграфом (а зачастую — гораздо ниже).

В перспективе хотелось бы разработать комплексную систему, реализующую сразу несколько аспектов. На текущий момент выбор остановлен на анкетном методе, параметрическом подходе и анализе почерка. Дополнительно предполагается рассмотреть возможность анализа голоса или анализа поведенческих характеристик по видео (как, например, в [14, 15, 16]). В дальнейшем можно попытаться использовать eye-трекеры, предложить использование тепловизоров, так как это может позволить анализировать изменения температуры тела (или, например, только лица), наполнения кровеносных сосудов на лице (когда человек лжет — у него «щеки горят») и т.п.

Первым методом, выбранным для реализации, стал анализ почерка. В работе [17] описана концепция диагностики ложной информации при анализе почерка.

При детектировании ложной информации участвует два человека: задающий вопрос и дающий ответ на поставленный вопрос [18]. Выявление ложных высказыва-

ний строится с помощью опросника. Вопросы в нем разделены на три группы (таблица 1).

Первая группа — нейтральная — содержит вопросы общего характера (ФИО, пол, возраст и т.п.). При ответе на вопросы данной группы реципиент находится в расслабленном состоянии. Текст из нейтральной зоны может использоваться как некий эталон.

Вторая группа — контрольная. Испытуемый отвечает на вопросы, которые относятся к делу непосредственно. При этом у виновного человека чаще всего меняется почерк вследствие психомоторной реакции на неудобные

вопросы. В качестве маркеров, на которые надо обращать внимание, могут выступать изменение ширины и высоты букв, угла их наклона и прочее. У невиновного человека почерк обычно почти не меняется. При этом значения указанных маркеров отклоняются в допустимых пределах.

Третья группа — проверочная — необходима для подтверждения гипотезы о виновности испытуемого. Вопросы в этой группе также являются нейтральными. Испытуемый обычно выходит из состояния психологической нагрузки, а его почерк принимает прежний вид. В случае причастности подозреваемого к преступле-

Таблица 1.

Пример опросника для анализа почерка

Название	Тема	Вопросы (текст)
Ориентировочная зона	Незначима	Начальнику УМВД...
Нейтральная		Ф.И.О., дата рождения, место рождения, место жительства, место работы, семейное положение, образование, с кем проживает ...
Социальная оценка	Тревожность	— Я считаю себя законопослушным(ой) гражданином(ой)? В данных обстоятельствах обязуюсь говорить правду. — Считаю, что на данный момент по сделанному мною заявлению о преступлении ... (Дает оценку работе сотрудников полиции, которые расследуют данное преступление)
Знание	Значимая (для всех)	ПИШУТ САМИ-БЕЗ ДИКТОВКИ На данный момент о преступлении... (могу сказать, мне известно, могу написать следующее), данную информацию я узнал(а) от... с чьих слов, время (когда стало известно)
Алиби		— В момент совершения данного преступления, я находился(ась) ... — Занималась ... — Мое присутствие может подтвердить....
	Значимая (для преступника)	3 блока ДИКТУЕМ БЫСТРО
Контрольная 1		1. Основание подозревать: — У меня нет веских оснований подозревать свою дочь в оговоре ... — Я уверена в том, что моя дочь рассказала мне правду ...
Проверочная 1		2. Знание преступника: — Я достоверно точно не знаю, где еще, кроме указанного в заявлении события, моя дочь могла ... — Я никогда ... в присутствии своей дочери.
Проверочная 2		3. Личное участие: — Я официально заявляю, что в своем заявлении все изложила честно и правдиво, без преувеличений. — У меня не было умысла оговорить ... — Я не заставляла свою дочь обманывать работников правоохранительных органов.
Контрольная 2		Пишут про себя — У меня никогда не возникало желание и мысли обманывать своих родственников. — Я никогда не совершала никаких противоправных деяний — Я никогда не обманывал работников правоохранительных органов
Жертвенно-релевантная		— Я все написала честно и правдиво, в здравом уме, твердой памяти и хорошем самочувствии, без какого-то ни было давления со стороны специалиста, проводившего опрос. Данное мною объяснение готова подтвердить, пройдя тестирование с использованием полиграфа (детектора лжи)
Нейтральная		— Данное объяснение написано собственноручно, в помещении _____. Время. Дата (месяц прописываем словом). Подпись.

нию значения маркеров почерка обычно равны в нейтральной и проверочной зонах, но сильно отличаются в контрольной. Если в трех группах почерк практически не меняется, то испытуемый скорее всего невиновен. В случае, когда значения маркеров почерка близки по значениям в контрольной и проверочной группах вопросов, это может означать, что необходимо провести дополнительные проверки подозреваемого, так как причастность или непричастность не доказана.

В работе [17] в качестве маркера почерка был рассмотрен только средний угол наклона символов в слове.

В нашем исследовании мы попытаемся расширить набор маркеров анализа текста. На текущий момент предлагается рассмотреть 2 варианта реализации.

Первый подход заключается в проведении аналогии с самой методикой, описанной выше. Текст разделяется на три группы. Заранее известно, какая из групп является нейтральной, контрольной и проверочной. Далее анализируется каждая группа по очереди.

Для начала рассматривается текст первой группы. Текст разбивается на слова, слова — на буквы. Начертания каждой буквы необходимо проанализировать нейронной сетью и получить набор параметров, характеризующих каждую букву. Получим некоторое усредненное значение для каждой буквы. Эти значения будут являться эталонными. Использовать в качестве эталона начертания букв из известных наборов данных не представляется перспективным, так как у каждого человека даже в спокойном состоянии почерк может сильно отличаться от некоторого общепринятого.

Имея полученный таким образом эталонный шаблон каждой буквы, можно переходить к анализу контрольной и проверочной группы. Для каждой из них будут получены некоторые отклонения. В дальнейшем и будут анализироваться эти отклонения.

Второй подход заключается в применении нейронных сетей для непосредственного анализа. Текст все также разбивается на три группы. А далее нейронная сеть сама сравнивает каждую из пар «нейтральная» — «контрольная», «контрольная» — «проверочная», «нейтральная» — «проверочная». Затем анализируются результаты этих сравнений. Выдвигается гипотеза, что нейронная сеть сама сможет выявить отклонения в характеристиках текста в разных группах.

Немаловажным фактором применения данной методики является определение психотипа испытуемого [19]. Психотип определяется анкетным способом, поэтому данная методика также будет реализована посредством нейросети. Разные люди разных психотипов будут по-разному реагировать на задаваемые вопросы. Поэтому в дальнейшем планируется в качестве дополнительного параметра на входе нейронной сети по анализу текстов подавать и определенный заранее психотип испытуемого. Более того необходимо учитывать и категории испытуемых (беременные женщины, лица, находящиеся под воздействием алкоголя или наркотиков, имеющие некоторые заболевания (как психические или психологические, так и физические, физиологические), беременные женщины и другие) [20].

Заключение

Дальнейшим развитием системы станет анализ результатов проверок на полиграфе с помощью еще одной нейронной сети.

Таким образом, предлагается реализовать три нейронные сети:

1. нейронная сеть для анализа анкет испытуемых для выявления психотипа личности;
2. нейронная сеть для анализа почерка при проведении опроса;
3. нейронная сеть для анализа параметрических результатов исследования на полиграфе.

В идеальном варианте хотелось бы получить набор данных одних и тех же испытуемых по всем трем направлениям исследования, чтобы иметь возможность сравнить результаты применения разных методик. В дальнейшем все три нейронные сети планируется свести в единую систему, которая на входе будет получать необходимые данные исследований, а на выходе отвечать на главный вопрос — говорит человек правду или лжет.

В качестве одного из вариантов подготовки наборов данных рассматривается возможность проведения добровольного анонимного анкетирования обучающихся на предмет выявления признаков коррупционной составляющей в университете, нарушений морально-этических норм поведения внутри студенческого коллектива (агрессивное поведение, травля, моральное, психологическое и физическое воздействие), фактов распространения запрещенных веществ среди обучающихся и т.п.

ЛИТЕРАТУРА

1. Ясницкий Л.Н. Интеллектуальные системы. Учебное пособие. Москва: Издательство «БИНОМ. Лаборатория знаний»; 2016. 221 с. EDN: WCGNOR
2. Оглоблин С.И., Молчанов А.Ю. Инструментальная детекция лжи (проверка на полиграфе). акад. курс. Ярославль: Нуанс; 2004. 464 с. EDN: QWFBAF
3. Кузнецова А.А., Рогожина В.В. Полиграф — из истории до наших дней. Вестник современных исследований. 2018;11.8(26):55–57. EDN: GCVXRA
4. Смольникова И.В. История появления полиграфа и его применения в практике раскрытия и расследования преступлений. Криминалистика: вчера, сегодня, завтра. 2023;1(25):197–208. EDN: CDWHOQ. DOI: 10.55001/2587-9820.2023.74.72.018
5. Шиллер А.В. Искажения и ошибки моделирования эмоций в искусственном интеллекте. Ценности и смыслы. 2020;5(69):93–107. EDN: UQPGOE. DOI: 10.24411/2071-6427-2020-10047
6. Гусева Н.В. Учет основных особенностей психотипов личности обследуемых при проведении психофизиологических исследований с использованием полиграфа. Научно-методический электронный журнал «КОНЦЕПТ». 2017;31:1456–1460. EDN: YPIWKF
7. Ясницкий Л.Н., Петров А.М., Сичинава З.И. Сравнительный анализ алгоритмов нейросетевого детектирования лжи. Известия высших учебных заведений. Поволжский регион. Технические науки. 2010;1(13):64–72. EDN: MNLIFZ
8. Захаров А.С., Кирличников А.П., Ляшева С.А., Шлеймович М.П. Система детектирования лжи на основе обработки изображений. Вестник технологического университета. 2016;18:132–134. EDN: WYBSIB
9. Селезнева Е.И., Воронова А.А. Eye-tracking как альтернативная или дополнительная к полиграфу технологий детектирования лжи. Перспективы: сборник научных статей магистрантов и аспирантов. 2021;10:154–160. EDN: UICUIO
10. Селезнева Е.И., Тихомирова Е.А. Динамика диаметра зрачка и ритма сердца при сообщении заведомо ложной информации на фоне индуцированной когнитивной нагрузки. Нейроинформатика-2021. Сборник научных трудов XVIII Международной научно-технической конференции. 2021:57–67. EDN: UOOUNS
11. Бровкин Д.Н. Особенности движений глаз и ритма сердца при калибровочном тексте полиграфа. Надежды: сборник научных статей студентов. 2021;15:73–76. EDN: DBPUWF
12. Dionisio D.P., Granholm E., Hillix W.A., Perrine W.F. Differentiation of deception using pupillary responses as an index of cognitive processing. Psychophysiology. 2001;38(2):205–211. DOI: 10.1111/1469-8986.3820205
13. Cook A.E., Hacker D.J., Webb A.K., Osher D., Kristjansson S.D., Woltz D.J., Kircher J.C. Lyin' Eyes: Ocular-motor Measures of Reading Reveal Deception. Journal of Experimental Psychology: Applied. 2012;18(3):301–313. DOI: 10.1037/a0028307
14. Медведев А.А., Лаптев А.А. Алгоритм выделения невербальных маркеров поведения человека на видео. Научный результат. Информационные технологии. 2022;7(2):58–64. EDN: HWUGOB. DOI: 10.18413/2518-1092-2022-7-2-0-8
15. Корсун О.Н., Юрко В.Н. Методы контроля психофизиологического состояния оператора по эмоциональному выражению лица и анализу характеристик морганий на основе глубоких сверточных сетей. Вестник Московского государственного технического университета им. Н.Э. Баумана. Серия приборостроение. 2021;1(134):120–134. EDN: QPUUSZ. DOI: 10.18698/0236-3933-2021-1-120-134
16. Gowri S.M., Rafeeq A., Devipriya S. Detection of real-time facial emotions via deep convolution neural network. Proceedings — 5th International Conference On Intelligent Computing and Control Systems, ICICCS 2021. 2021:1033–1037. EDN: RUWQQR. DOI: 10.1109/ICICCS51141.2021.9432242
17. Михайлов А.В. Распознавание ложной информации в рукописном тексте с помощью нейронных сетей. Информационные технологии в науке, управлении, социальной сфере и медицине (IT UBIQUITOUS): материалы VII Международной научной конференции, 17-21 октября 2022 года, г. Томск, 24–25 октября 2022 года, г. Иркутск. 2023:206–212. Режим доступа: <https://koha.lib.tsu.ru/cgi-bin/koha/opac-detail.pl?biblionumber=1001066> (дата доступа 25.08.2023)
18. Рывкин С.Ю., Евтушенко О.В. Криминалистически значимые изменения почерка и подписи. Аллея науки. 2019;12(39):743–748. EDN: EDZXJO
19. Евсеева О.Е., Николаев А.Ю. Допустимость проведения психофизиологического исследования с применением полиграфа лицам с различными особенностями личности. Вопросы экспертной практики. 2019;51:189–194. EDN: QFWITP
20. Сезонова Т.В., Лунина О.Н. Психофизическая экспертиза с применением полиграфа как метод инструментальной детекции лжи при раскрытии преступлений. Научный вестник орловского юридического института МВД России имени В.В. Лукьянова. 2022;2(91):202–209. EDN: BQQZIL

© Тюрнев Александр Сергеевич (tyurnevas@bk.ru)

Журнал «Современная наука: актуальные проблемы теории и практики»