

ПРИМЕНЕНИЕ НЕЙРОННЫХ СЕТЕЙ В ИССЛЕДОВАНИЯХ НА ПОЛИГРАФЕ¹

© 2022 г. Л. А. Деревягин^а, В. В. Макаров^{а,*}, А. Ю. Молчанов^а,
В. И. Цурков^б, А. Н. Яковлев^а

^а МФТИ (НИУ), Москва, Россия

^б ФИЦ ИУ РАН, Москва, Россия

*e-mail: viktor.makarov@phystech.edu

Поступила в редакцию 16.02.2022 г.

После доработки 24.02.2022 г.

Принята к публикации 28.03.2022 г.

Предлагается использование машинного обучения для автоматизации работы полиграфолога с архитектурами нейронных сетей из библиотеки scikit-learn с применением VotingClassifier и трансформера. При этом повышается эффективность анализа полиграмм, осуществляется согласование по признакам, уменьшается количество ошибочных выводов по ответам испытуемого. Применяются следующие показатели (каналы): электрокожное сопротивление (кожногальваническая реакция), кровонаполняемость сосудов (плетизмограмма), дыхательные ритмы.

DOI: 10.31857/S0002338822040059

Введение. Несмотря на постоянное развитие систем безопасности, человеческий фактор все еще остается одним из самых незащищенных элементов. Использование полиграфа позволяет уменьшить такие риски. Психофизиологические исследования решают следующие прикладные задачи: выявление негативных факторов в прошлом опыте кандидатов на должность, проверка на лояльность и соблюдение внутренних регламентов организации, осуществление оперативно-розыскной деятельности, проведение корпоративных и антикоррупционных расследований и т.д.

Первым начал использовать методики выявления стресса выдающийся итальянский криминалист Ч. Ломброзо, который в 1895 г. изложил в своей широко известной книге “Преступный человек” практический опыт применения гидроплетизмографа (устройство для измерения пульса и кровенаполнения сосудов) в ходе проверки фигуранта по уголовному делу об ограблении. При проведении психофизиологического исследования автор зафиксировал видимые изменения динамики артериального давления в ответ на предъявление стимулов, связанных с ограблением.

В 1888 г. Ч. Фере, обследуя больную с жалобами на электрические покалывания в кистях рук и ступнях, обнаружил, что при пропускании слабого тока через предплечье происходили отклонения стрелки включенного в цепь гальванометра в моменты сенсорных или эмоциональных воздействий. Независимо от Ч. Фере в 1890 г. И. Тарханов показал, что электрические сдвиги наблюдаются и без приложения внешнего тока. Таким образом, он открыл кожный потенциал, величина которого тоже изменяется в ответ на сенсорные и эмоциональные раздражители. Подобные эффекты были названы кожно-гальванической реакцией (КГР).

Среди пионеров техники полиграфных проверок необходимо назвать сотрудника полиции штата Калифорния Д. Ларсона, который в 1921 г. сконструировал первый прообраз современного профессионального полиграфа (данный аппарат одновременно регистрировал изменения динамики относительного артериального давления, пульса и дыхания).

Одним из первых исследователей канала КГР был знаменитый К. Юнг, который рассматривал данный сигнал как объективное физиологическое “окно” в сферу бессознательного, подле-

¹ Работа выполнена при частичной финансовой поддержке РФФИ (грант № 21-51-53019).

жащего изучению через психоанализ. К. Юнг первым выявил прямую зависимость между величиной КГР и силой эмоционального переживания.

Итак, полиграф является медико-биологическим прибором для одновременного измерения динамики физиологических показателей:

- динамики дыхательных циклов;
- электрокожного сопротивления (КГР);
- артериального давления;
- периферического кровонаправления (ФПГ);
- двигательной активности (тремор).

Как известно, полиграф работает так: имеется круг вопросов из некоторой предметной области. К испытуемому индивиду подключаются датчики для вышеуказанных измерений. Принимается ответ: либо “да”, либо “нет”. Все вопросы повторяются не менее 3 раз. Полиграфолог анализирует полученные функции с помощью системы трехбалльной оценки и окончательно констатирует о справедливости ответа [1].

Суть такой системы заключается в том, что реакциям на каждый вопрос теста выставляются баллы в следующем порядке:

- (сильная) максимальная по выраженности и близкие к ней реакции — 2 балла,
- (средняя) вторая по выраженности реакция и близкие к ней — 1 балл,
- (слабая) все остальные реакции — 0 баллов.

Критерии визуальной оценки значимости вопросов по каждому показателю отличаются. Они подробно описаны в профильной литературе. Реакция на каждый вопрос обсчитывается отдельно по респираторному каналу, кардиоканалу и электро-дермальному каналу. В итоге получается три числа — по одному на показатель. Для принятия решения суммируются баллы по трем каналам по всем повторениям.

Например, если в результате трех повторений вопроса получилось 11 баллов, то можно говорить о значимости реакции и возможном сокрытии информации, если 2 балла, то имеет место слабая реакция (это указывает на то, что индивид не испытывает стресс и ему нечего скрывать).

1. Краткий обзор существующих подходов и постановка задачи. К данному моменту применяют только прямые методы обсчета полиграмм [2–4]. В результате вычисляются и сравниваются следующие характеристики:

- длина линии сигнала верхнего (грудного) и нижнего (диафрагмального) дыхания,
- количество дыхательных циклов,
- максимальная амплитуда КГР,
- площадь под графиком сигнала КГР,
- смещение средней линии плетизмограммы,
- другие характеристики.

К сожалению, индивидуальные физиологические особенности каждого человека привносят неравномерные искажения по одному или нескольким сигналам. Прямые методы обсчетов не позволяют сделать автоматическую подстройку к таким изменениям, поэтому точность таких подходов является недостаточной для самостоятельного принятия решения по результатам проведенного психофизиологического исследования. Далее ставится задача исправить этот недостаток при помощи нейронных сетей и машинного обучения.

2. Подготовка данных. Каждая реакция представляет 12 с записанной полиграммы с частотой сигнала 20 Гц, т.е. 240 точек по каждому каналу:

- КГР (электрическая активность кожи),
- ФПГ,
- грудное и диафрагмальное дыхание.

По каждой реакции группа профессиональных полиграфологов проставляет оценку от 0 до 2 баллов в соответствии с описанными выше правилами. Таким образом, было обработано 90 психофизиологических исследований с разными испытуемыми и получено 8000 классифицированных отрезков по каждому каналу. Временные значения распределены равномерно: в каждой записи находится 240 численных значений.

3. Сравнительное тестирование архитектур. Рассматриваются встроенные архитектуры библиотеки scikit-learn [5]. Сначала выбираем подготовленные “сырые” необработанные данные

поочередно как в равном количестве, так и в несбалансированном отношении 7:3. Применяем следующие алгоритмы классификации данных, реализованные в указанной выше библиотеке. Они создают для каждого из трех показателей соответствующую модель для расчета результатов полиграфного тестирования.

SVM (support vector machine) – набор схожих алгоритмов обучения с учителем, использующихся для задач классификации и регрессивного анализа.

GPC (gaussian process) – случайный процесс (набор случайных величин, индексированных по времени или пространству), такой, что каждый конечный набор этих случайных величин имеет многомерное нормальное распределение, т.е. каждая их конечная линейная комбинация обычно распределена.

GNB (gaussian naive bayes) – простой вероятностный классификатор, основанный на применении теоремы Байеса со строгими предположениями о независимости и нормальном распределении признаков в наборе данных.

AdaBoost (с параметрами: количество оценщиков – 2500, коэффициент обучения – 0.001) – адаптивный классификатор в том смысле, что последующие слабые ученики настраиваются в пользу тех экземпляров, которые были неправильно классифицированы предыдущими классификаторами.

MLP (multi-layer perceptron) – класс упреждения искусственной нейронной сети.

Gradient Boosting [6] – метод машинного обучения для регрессии, классификации и других задач, который создает модель прогнозирования в виде ансамбля слабых моделей прогнозирования, обычно деревьев решений.

DecisionTree – средство поддержки принятия решений, использующееся в машинном обучении, анализе данных и статистике.

RandomForest – алгоритм машинного обучения, заключающийся в применении комитета решающих деревьев.

ExtraTrees [7] – ансамблевый алгоритм машинного обучения, который объединяет прогнозы из многих деревьев решений.

В итоге найдено 27 моделей, по одной на каждый показатель и на каждый алгоритм классификаций. Модели получают на вход прямые данные с полиграфа, а на выходе выдают класс, соответствующий силе реакции: 0, 1 или 2 балла.

Выбираем три наиболее точных классификатора по каждому каналу. Предположим, что при оценке определенной реакции на вопрос с вероятностью 100% первые два алгоритма получили класс “0 баллов”, а третий – класс “2 балла”. Для адекватного принятия решения, особенно в случаях несогласованности между классификаторами, применяем надстройку в виде VotingClassifier [8], которая присваивает весовые коэффициенты каждому из них. Итоговый результат будет рассчитан по формуле

$$\max(w_1 p_{11} + w_2 p_{21} + w_3 p_{31}, w_1 p_{12} + w_2 p_{22} + w_3 p_{32}, w_1 p_{13} + w_2 p_{23} + w_3 p_{33}),$$

где w_i – весовой коэффициент i -го классификатора, p_{ij} – вероятность принадлежности реакции к классу j , полученная при помощи классификатора i .

Пусть в указанном выше примере коэффициенты равны 0.2, 0.35 и 0.45 соответственно. Тогда по этим результатам мы принимаем следующее решение: класс “0 баллов” – вероятность 55%, класс “2 балла” – вероятность 45%. Поэтому в таком случае выбирается класс “0 баллов”. Аналогично рассматриваются другие ситуации для выбора оптимального классификатора.

Точность классификации силы реакции по трем классам на архитектуре VotingClassifier, включающей в себя три наилучших алгоритма, составила 52%, а из пяти – 59%.

Для повышения точности остановимся подробнее на реализации классификатора для двух классов при помощи модели трансформера, измененной в соответствии с задачей. По аналогии с рекуррентными нейронными сетями (РНС) трансформеры предназначены для обработки последовательностей, таких, как текст на естественном языке, и решения таких задач, как машинный перевод и автоматическое реферирование. В отличие от РНС, трансформеры не требуют обработки последовательностей по порядку. Например, если входные данные – это текст, то трансформеру не требуется обрабатывать конец текста после обработки его начала.

В классическом варианте трансформер представляет собой две части – кодировщик и декодировщик. В нашем случае была использована структура, состоящая в части кодировщика из чередующихся слоев внимания (внимание на основе скалярного произведения) и многослойного



Рис. 1. Проекция ручной разметки для канала ФПГ на обучающей выборке (слева направо: классы “2 балла”, “1 балл”, “0 баллов”)

перцептрона, а также декодировщика (простого классификатора). Обе части трансформера реализованы стандартными средствами библиотеки Keras.

Данные на вход подаются в размерности (26; 240), т.е. одновременно подается 26 образцов данных размером 240 значений. Затем данные кодируются при помощи одномерного (1D) PatchEncoder с шириной (8; 2), т.е. линейно трансформируются проецированием на вектор указанной выше размерности.

Кодированные данные поступают на группу чередующихся слоев внимания и перцептронов, а затем поступают на слой SeqPool для субдискретизации. Этап подвыборки последовательности нормируется (LayerNormalization) средствами библиотеки Keras и данные с него затем поступают на полносвязный слой в 600 нейронов с активацией SeLu. Результат работы кодировщика на этом этапе обрабатывается слоем FeaturePooling с количеством карт черт, равным двум.

Главные характеристики:

размер набора данных (batch size): 60,

количество блоков (чередование внимания и перцептрона): 6,

размерность слоя внимания: 312,

выброс слоя внимания: 0.1 (10%),

функция трансформации ядра: softmax,

количество наборов матриц весов (запросов, ключей, значений): 6.

Точность обучения на два класса (сильная и слабая реакции на заданный вопрос) составляет для показателей:

плетизмограммы – 86.8% +–3%,

кожно-гальванического сопротивления – 95.3% +–3%,

дыхания – 72.7% +–3%.

4. Пути повышения точности. Также были предприняты попытки обучить классификатор трансформера сразу на три класса, однако они показали заметно худшую точность и большой процент ошибок в сравнении классов “1 балл”–“0 баллов” и “1 балл”–“2 балла”. Используемый метод заметно уменьшает долю неясных средних реакций в пользу выраженных слабых и сильных.

Для визуализации и оценки результатов работы FeaturePooling данные были обработаны нейронной сетью без последнего слоя. Получившийся набор векторов был нормирован и кластеризован в t-SNE [9]. Градиент функции потерь будет в таком случае равен

$$\frac{\partial Cost}{\partial y_i} = 4 \sum_j (p_{ij} - q_{ij})(y_i - y_j)(1 - \|y_i - y_j\|^2)^{-1},$$

где p_{ij} – сходство точек, q_{ij} – сходство точек отображения.

Для проверки результатов работы сетей кластеризуем обучающую и валидационную выборку для каждой из характеристик в t-SNE и сравним с результатами работы трансформера. На рис. 1, 2 представлены соответствующие результаты по каналу ФПГ. Можно заметить, что классифицированные машиной точки “1 балл” не разбросаны хаотично, а образуют сплошные непрерывные линии после обработки трансформером. Кроме того, точки категории “1 балл” намного реже



Рис. 2. Пример результата работы архитектуры сети трансформер для канала ФПГ на обучающей выборке (слева направо: классы “2 балла”, “1 балл”, “0 баллов”)

попадают в кластеры “0 баллов” и “2 балла” по сравнению с ручной разметкой. Это означает наличие возможности повышения точности обучения на три класса.

Заключение. В дальнейших исследованиях архитектура будет улучшена для повышения точности классификации на три класса значимости реакции (“0 баллов”, “1 балл”, “2 балла”), а также проработан вопрос выявления артефактных (некорректно записанных) сигналов. В настоящий момент проводится интеграция данных нейронных сетей в разработанное авторами программное обеспечение полиграфа “Финист”. Получены результаты по эффективности для оценки факторов риска кандидатов при трудоустройстве. Интересным направлением развития является интеграция полиграфа с дополнительными биометрическими модальностями, показывающими хорошую чувствительность в близких задачах, а именно с трехмерной картой лица [10] и зрачковой реакцией [11].

СПИСОК ЛИТЕРАТУРЫ

1. *Оглоблин С.И., Молчанов А.Ю.* Инструментальная “детекция лжи”: академический курс. Ярославль: Ньюанс, 2004. 464 с.
2. *Gordon N.* Essentials of Polygraph and Polygraph Testing. Boca Raton: CRC Press, 2016. 320 p.
3. *Леонтьев К.А., Панин С.Д., Холодный Ю.И.* Оценка результатов тестирования на полиграфе методами регрессионного анализа // Наука и образование: электронный журнал МГТУ им. Н.Э. Баумана. 2014. № 10. С. 230–243.
4. *Минакова Н.Н., Божич Е.В.* Применение методов многомерного анализа данных при обработке полиграмм для изучения биофизических характеристик // Изв. АлтГУ. 2018. № 1 (99). С. 34–38.
5. *Porcu V.* Python for Data Mining Quick Syntax Reference. N.Y.: Apress Media LLC, 2018. 260 p.
6. *Greenwell B., Boehmke B.* Hands-On Machine Learning with R. Boca Raton: CRC Press, 2019. 488 p.
7. *Abdelkader B., Jaziri R., Bernard G.* Deep Cascade of Extra Trees // Pacific-Asia Conference on Knowledge Discovery and Data Mining (PAKDD). Macau, China, 2019. P. 117–129.
8. *Hung C., Lee S.* Adaptive Distance-Based Voting Classification // Intern. Conf. on Machine Learning and Cybernetics (ICMLC). Tianjin, China, 2013. P. 1671–1677.
9. *Попова И.А., Попова А.А., Соболева Е.Д.* Визуализация многомерных наборов данных при помощи алгоритмов снижения пространства признаков PCA и T-SNE // Научно-образовательный журнал для студентов и преподавателей “StudNet”. 2020. № 11.
10. *Knyaz V.A., Matveev I.A., Muryinin A.B.* Applying Computer Stereovision Algorithms to Study of Correlation Between Face Asymmetry and Human Vision Pathology // Pattern Recognition and Image Analysis. 2009. V. 19. № 4. P. 679–686.
11. *Varchenko N.N., Gankin K.A., Matveev I.A.* Using Binocular Pupillometry Method for Evaluating Functional State of Person // Sports Technology. 2015. V. 8. P. 67–75.