

(19)



**Евразийское
патентное
ведомство**

(11) **047243**

(13) **B1**

(12) **ОПИСАНИЕ ИЗОБРЕТЕНИЯ К ЕВРАЗИЙСКОМУ ПАТЕНТУ**

(45) Дата публикации и выдачи патента
2024.06.24

(51) Int. Cl. **G06N 20/00** (2019.01)
A61B 5/05 (2021.01)

(21) Номер заявки
202392414

(22) Дата подачи заявки
2023.09.25

(54) **СПОСОБ И СИСТЕМА АВТОМАТИЧЕСКОЙ ПОЛИГРАФИЧЕСКОЙ ПРОВЕРКИ**

(31) **2023100807**

(56) US-A1-20210248511
RU-C2-2708807
US-A1-20220232024
RU-C2-2767157
RU-C1-2421123

(32) **2023.01.16**

(33) **RU**

(43) **2024.06.21**

(71)(73) Заявитель и патентовладелец:
**ПУБЛИЧНОЕ АКЦИОНЕРНОЕ
ОБЩЕСТВО "СБЕРБАНК
РОССИИ" (ПАО СБЕРБАНК) (RU)**

(72) Изобретатель:
**Асонов Дмитрий Валерьевич, Крылов
Максим Андреевич, Рябкина
Анастасия Евгеньевна, Литвинов
Евгений Вячеславович, Митрофанов
Максим Алексеевич, Михайлов
Максим Алексеевич (RU)**

(74) Представитель:
Герасин Б.В. (RU)

(57) Изобретение относится к автоматизированному способу и системе автоматической полиграфической проверки с помощью алгоритмов машинного обучения. Техническим результатом от реализации заявленного способа является повышение точности полиграфической проверки. Указанный технический результат достигается благодаря осуществлению компьютерно-реализуемого способа автоматической полиграфической проверки, выполняемого с помощью вычислительной системы, содержащей модель машинного обучения, при этом способ выполняет этапы, на которых получают записи полиграфических проверок, содержащие, по меньшей мере, сигналы датчиков с временными шкалами, на которых промаркированы начало и конец вопроса; получают дополнительные данные, содержащие, по меньшей мере, возраст проверяемого, пол, должностную информацию; осуществляют обработку полученных сигналов и дополнительных данных с помощью модели машинного обучения (МО) и определяют, что ответ является ложью, если выходное значение выше или равно пороговому значению, или ответ является правдой, если выходное значение ниже порогового значения.

B1

047243

047243

B1

Область техники

Заявленное техническое решение относится к автоматизированному способу и системе автоматической полиграфической проверки с помощью алгоритмов машинного обучения.

Уровень техники

Классические полиграфические скрининги регулярно используются значимыми предприятиями такими как банки, правоохранительные структуры и федеральные органы власти. Основное беспокойство научных сообществ заключается в том, что эти скрининги склонны к содержанию ошибок. Однако эти ошибки могут быть следствием не только метода, но и человека (полиграфолога).

Безопасность клиентских денег и данных (например, транзакций) заложена в основе банковской культуры и репутации. В качестве одного из инструментов защиты клиентов, только с согласия кандидатов и сотрудников, и в соответствии с законодательством, банк использует полиграфические скрининги (ПС). Они применяются при найме кандидатов на рискованных направлениях, чтобы предотвратить наем ненадежного человека. Чтобы обнаружить нарушение, сотрудники на особо рискованных позициях регулярно проходят проверку. ПС включает следующие темы: наркотические вещества, зависимость от азартных игр, инсайдерская торговля, разглашение конфиденциальной информации, взяточничество, коррупция, незаконные присвоение средств и мошенничество. Финансовая отрасль - не единственная, использующая ПС; другие примеры - такие важные отрасли, как авиация, промышленность и правоохранительные структуры во всем мире [1,2].

Классический полиграф - это устройство, записывающее сердечно-сосудистую активность (такую как сердечный пульс), грудное и брюшное дыхание, гальваническая реакция кожи (электрическая активность кожи или ЭАК) и дрожь. Полиграфолог задает вопросы испытуемому, на которые получает ответы "да" или "нет". Обзоры классического полиграфа и методологии построения вопросов представлены в [3, 4, 5].

Нетрадиционные исследования обнаружения лжи используют анализ видео и аудио [6] (включая мимику лица [7, 8], реакцию зрачка [9] и задержку между вопросом и ответом [10]), электромиографию (ЭМГ) [11], электроэнцефалографию (ЭЭГ) [12], магнитно-резонансную томографию [13, 14] или письменные последовательности (динамика нажатия клавиш) [15] в дополнение к классическим полиграфическим данным.

Некоторые из этих исследований даже получили возможность на освоение новой области, так детектор лжи iBorderCtrl тестируется в европейских аэропортах [16, 17] или VeriPol применяется испанской полицией на делах о страховых требованиях [18, 19]. Классический полиграф остается инструментом выбора в традиционных задачах как скрининг при найме и уголовные или внутренние расследования.

Полиграф имеет длинную историю критики от ученых из области психологии и права, а также и со стороны общества и правительства [1, 23]. Основная обеспокоенность заключается в том, что эта методика надежно не определяет ложь и правду. И все же "парадоксально, хотя Конгресс выражает глубокое беспокойство по поводу эффективности данной технологии, ЕРРА разрешает использование детекторов лжи в случае, если точность результата имеет первостепенное значение: национальная оборона, безопасность и законные текущие расследования" [22].

Критика данной методики предоставляет много аргументов, почему полиграфический скрининг может потерпеть неудачу при обнаружении лжи или отметить правду как ложь. Например, "Полиграфические тестирования оценивают не обман, а ситуации, которые построены так, чтобы вызвать и оценить страх" [24].

Правдивый младший менеджер может бояться, что его назовут коррупционером больше, чем хладнокровный старший менеджер боится быть пойманным на лжи полиграфологом. Еще один пример конструктивной критики - призыв к стандартизации процедуры полиграфического скрининга и обучения полиграфолога [25]. Ошибки полиграфолога могут происходить, например, когда полиграфолог неопытен, уставший, отвлечен или предвзят [26].

Существует простое решение проверки качества: всегда проводить проверку еще одним полиграфологом, который подтвердит или опровергнет заключение предыдущего полиграфолога [27]. Чтобы провести проверку полиграфического отчета, другому полиграфологу требуется пересмотреть запись скрининга, включающую полиграмму (графическое представление данных с датчиков, связанных с вопросами полиграфолога и ответами испытуемого), иногда аудио- и видеозапись и сравнить его заключение с оригинальным. Данная проверка занимает минимум половину времени от скрининга. Стандартный скрининг длится минимум два часа. Таким образом, повторная проверка стоит и времени, и денег. По этой причине, отделения внутренней безопасности проводят повторные проверки редко или не проводят их вообще. Другая причина, по которой проверки вторым полиграфологом могут быть не эффективны: второй полиграфолог может допустить ту же самую ошибку, которую допустил оригинальный полиграфолог.

Общим недостатком существующих решений в данной области является присутствие человеческого фактора при полиграфической проверке, что негативно сказывается на точности и скорости проверки, а также отсутствие автоматизированного процесса повторной проверки.

Раскрытие изобретения

В заявленном техническом решении предлагается новый подход, к автоматической полиграфической проверке с использованием моделей машинного обучения (МО).

Эффективность данного решения подтверждается существенным приростом точности и скорости проведения автоматической полиграфической проверки.

Таким образом, решается техническая проблема точной и скоростной автоматической полиграфической проверки.

Техническим результатом, достигающимся при решении данной проблемы, является повышение точности полиграфической проверки.

Дополнительным техническим результатом, достигающимся при решении данной проблемы, является повышение скорости полиграфической проверки.

Также дополнительным техническим результатом, достигающимся при решении данной проблемы, является автоматизация процесса полиграфической проверки.

Указанные технические результаты достигаются благодаря осуществлению компьютерно-реализуемого способа автоматической полиграфической проверки, выполняемого с помощью вычислительной системы, содержащей модель машинного обучения, при этом способ выполняет этапы, на которых

получают записи полиграфических проверок, содержащие, по меньшей мере, сигналы датчиков с временными шкалами, на которых промаркированы начало и конец вопроса;

получают дополнительные данные, содержащие, по меньшей мере, возраст проверяемого, пол, должностную информацию и информацию о геомагнитных бурях;

осуществляют обработку полученных сигналов и дополнительных данных с помощью модели машинного обучения (МО), причем в ходе указанной обработки осуществляется

определение временных интервалов для извлечения переменных на основе временных меток начала и конца вопроса и временной метки ответа, и на основе типа и темы вопроса;

извлечение переменных из каждого сигнала на определенных временных интервалах;

обработка полученных переменных из сигналов, при которой выполняется нормализация и конкатенация обработанных переменных и построение на их основе вектора;

разделение дополнительных данных на категориальные и численные переменные;

обработка полученных переменных из дополнительных данных, при которой выполняется векторизация категориальных переменных и нормализация численных переменных;

конкатенация обработанных переменных, извлеченных из каждого сигнала, и обработанных дополнительных переменных и построение на их основе вектора;

подача упомянутого вектора в модель МО для получения выходного значения модели МО;

сравнение выходного значения модели с заданным пороговым значением и

определяют, что ответ является ложью если выходное значение выше или равно пороговому значению или ответ является правдой если выходное значение ниже порогового значения.

В одном из частных вариантов реализации способа модель МО определяет ответ как правдивый или ложный не по одному полученному ответу, а по совокупности вопросов и ответов, связанных с темой полиграфической проверки, при этом получаемые переменные сигналы объединяются в единый вектор, либо усредняются перед передачей в модели МО.

В другом частном варианте реализации способа модель МО имеет тип градиентный бустинг, случайный лес, или нейронная сеть.

В другом частном варианте реализации способа модель МО обучена на одной из тем для проверок или их комбинации, где темами для проверок являются наркотические вещества, получение дополнительного вознаграждения, разглашение конфиденциальной информации, долговые обязательства, сторонний доход, уголовные правонарушения, административные правонарушения, нарушения внутренних нормативных документов (ВНД).

В другом частном варианте реализации способа записи полиграфических проверок содержат сигналы с датчиков, включающие по меньшей мере одно из: частота сердечного сокращения (ЧСС), кожно-гальваническая реакцию (КГР), артериальное давление, верхнее и нижнее дыхание, пьезоплетизмограмму, фотоплетизмограмму, термических, движения зрачка или их комбинации.

В другом частном варианте реализации способа дополнительные данные содержат по меньшей мере одно из: идентификационный номер полиграфологов, идентификационный номер полиграфов, информацию о погодных условиях, результаты электроэнцефалограммы, магнитно-резонансной томографии, функциональной ближней инфракрасной спектроскопии, информацию о геомагнитных бурях или их комбинации.

Кроме того, заявленный технический результат достигается за счет системы автоматической полиграфической проверки, содержащей

по меньшей мере один процессор;

по меньшей мере одну память, соединенную с процессором, которая содержит машиночитаемые инструкции, которые при их выполнении по меньшей мере одним процессором обеспечивают выполне-

ние способа автоматической полиграфической проверки.

Краткое описание чертежей

Признаки и преимущества настоящего изобретения станут очевидными из приводимого ниже подробного описания изобретения и прилагаемых чертежей.

Фиг. 1 иллюстрирует пример реализации способа автоматической полиграфической проверки с помощью одной модели МО.

Фиг. 2 иллюстрирует пример реализации способа автоматической полиграфической проверки с помощью двух моделей МО.

Фиг. 3 иллюстрирует пример реализации способа автоматической полиграфической проверки с помощью двух ансамблей моделей МО.

Фиг. 4 иллюстрирует пример реализации способа автоматической полиграфической проверки с помощью трех ансамблей моделей МО.

Фиг. 5 иллюстрирует график, где модель и полиграфолог сошлись и не сошлись во мнениях в зависимости от оценки модели.

Фиг. 6 иллюстрирует распределение оценок для двух релевантных тем.

Фиг. 7 иллюстрирует пример реализации базовой модели.

Фиг. 8 иллюстрирует пример реализации модели, обученной на одной теме.

Фиг. 9 иллюстрирует пример реализации универсальной модели.

Фиг. 10 иллюстрирует пример свертки сигналов.

Фиг. 11 иллюстрирует общий вид системы автоматической полиграфической проверки.

Осуществление изобретения

Данное техническое решение может быть реализовано на компьютере, в виде автоматизированной информационной системы (АИС) или машиночитаемого носителя, содержащего инструкции для выполнения вышеупомянутого способа.

Техническое решение может быть реализовано в виде распределенной компьютерной системы.

В данном решении под системой подразумевается компьютерная система, ЭВМ (электронно-вычислительная машина), ЧПУ (числовое программное управление), ПЛК (программируемый логический контроллер), компьютеризированные системы управления и любые другие устройства, способные выполнять заданную, четко определённую последовательность вычислительных операций (действий, инструкций).

Под устройством обработки команд подразумевается электронный блок либо интегральная схема (микроспроцессор), исполняющая машинные инструкции (программы)/

Устройство обработки команд считывает и выполняет машинные инструкции (программы) с одного или более устройств хранения данных, например таких устройств, как оперативно запоминающие устройства (ОЗУ) и/или постоянные запоминающие устройства (ПЗУ). В качестве ПЗУ могут выступать, но, не ограничиваясь, жесткие диски (HDD), флеш-память, твердотельные накопители (SSD), оптические носители данных (CD, DVD, BD, MD и т.п.) и др.

Программа - последовательность инструкций, предназначенных для исполнения устройством управления вычислительной машины или устройством обработки команд.

Для данного технического решения была построена базовая модель второго мнения, обученная на исторических данных из реальных ответов 2094 полиграфических скринингов, включающих поле "Выявлена ложь", заполненное полиграфологами, которые проводили скрининг.

Данное техническое решение демонстрирует метрики качества базовой модели в табл. 1.а, в колонке "все темы". Основные метрики это ROC AUC (receiver operating characteristic, рабочая характеристика приёмника) и TPR (целевой рейтинг target rating point) при FPR=0.05.

Табл. 2 и 3 отражают важность признаков, основанных на 10 физиологических сигналах и признаков пол и возраст испытуемого; детали построения признаков второго уровня будут описаны ниже. Фиг. 5 отображает график, где модель и полиграфолог сошлись и не сошлись во мнениях в зависимости от оценки модели.

Также было измерено качество модели, применив ее к каждой из семи тем скрининга (табл. 1а).

Оценка эффективности модели влечет за собой использование примерно 100 человеко-часов, потому что это потребует задействовать настоящего полиграфолога, который будет проверять множество полиграфических скринингов, отобранных на основании оценки модели.

Таблица 1. Качество предсказаний трех моделей на 7 темах, измеренное при помощи стратифицированной кросс-валидации: а) Базовая модель (признаки, построенные на основе 10 физиологических сигналов, возраст и пол испытуемого); б) Базовая модель с альтернативными данными; с) Базовая модель с признаком "Должность".

	Все темы	Нарк. вещества	Получение ден. вознаграгр	Разгл. конф. Инф.	Долг. обязат-ва	Сторонний доход	Уголовн. и адм. правонаруш.	Нарушение ВНД
ROC AUC (Std.)	a. 0.75 (0.02)	0.79 (0.02)	0.71 (0.04)	0.73 (0.05)	0.64 (0.03)	0.81 (0.03)	0.83 (0.06)	0.69 (0.03)
	b. 0.80 (0.01)	0.85 (0.02)	0.81 (0.03)	0.77 (0.05)	0.74 (0.07)	0.80 (0.09)	0.90 (0.03)	0.7 (0.03)
	c. 0.79 (0.01)	0.84 (0.02)	0.82 (0.01)	0.76 (0.05)	0.69 (0.06)	0.80 (0.02)	0.90 (0.03)	0.69 (0.04)
Recall (Std.)	a. 0.28 (0.01)	0.34 (0.01)	0.27 (0.10)	0.30 (0.08)	0.12 (0.06)	0.26 (0.09)	0.16 (0.08)	0.2 (0.07)
	b. 0.30 (0.02)	0.42 (0.03)	0.25 (0.09)	0.15 (0.07)	0.20 (0.05)	0.27 (0.09)	0.54 (0.19)	0.13 (0.05)
	c. 0.29 (0.02)	0.40 (0.03)	0.22 (0.05)	0.24 (0.07)	0.15 (0.04)	0.37 (0.06)	0.47 (0.14)	0.09 (0.05)
Precision (Std.)	a. 0.14 (0.01)	0.32 (0.06)	0.14 (0.04)	0.10 (0.03)	0.04 (0.02)	0.08 (0.02)	0.05 (0.03)	0.10 (0.04)
	b. 0.19 (0.02)	0.45 (0.08)	0.17 (0.06)	0.06 (0.03)	0.12 (0.03)	0.10 (0.04)	0.12 (0.04)	0.09 (0.04)
	c. 0.18 (0.02)	0.41 (0.06)	0.16 (0.02)	0.11 (0.04)	0.08 (0.03)	0.12 (0.02)	0.11 (0.03)	0.06 (0.03)
F1_score (Std.)	a. 0.19 (0.01)	0.32 (0.04)	0.18 (0.06)	0.14 (0.04)	0.06 (0.03)	0.12 (0.03)	0.08 (0.04)	0.14 (0.05)
	b. 0.23 (0.02)	0.43 (0.05)	0.2 (0.07)	0.09 (0.04)	0.14 (0.04)	0.14 (0.05)	0.18 (0.06)	0.1 (0.04)
	c. 0.22 (0.02)	0.40 (0.04)	0.18 (0.03)	0.15 (0.05)	0.10 (0.03)	0.18 (0.02)	0.16 (0.04)	0.07 (0.04)
Accuracy (Std.)	a. 0.92 (0.00)	0.90 (0.01)	0.93 (0.00)	0.94 (0.00)	0.93 (0.01)	0.94 (0.01)	0.93 (0.01)	0.92 (0.01)
	b. 0.93 (0.01)	0.92 (0.01)	0.94 (0.00)	0.94 (0.01)	0.94 (0.01)	0.94 (0.01)	0.93 (0.01)	0.91 (0.01)
	c. 0.93 (0.01)	0.92 (0.01)	0.94 (0.01)	0.94 (0.01)	0.94 (0.01)	0.94 (0.01)	0.93 (0.01)	0.92 (0.01)
TNR (Std.)	a. 0.95 (0.00)	0.94 (0.01)	0.95 (0.00)	0.95 (0.00)	0.95 (0.01)	0.95 (0.01)	0.94 (0.01)	0.94 (0.01)
	b. 0.95 (0.01)	0.96 (0.01)	0.96 (0.00)	0.95 (0.01)	0.96 (0.01)	0.95 (0.01)	0.94 (0.01)	0.94 (0.01)
	c. 0.95 (0.01)	0.96 (0.01)	0.96 (0.00)	0.96 (0.01)	0.96 (0.01)	0.96 (0.01)	0.94 (0.01)	0.95 (0.01)
FPR (Std.)	a. 0.05 (0.00)	0.06 (0.01)	0.05 (0.00)	0.05 (0.00)	0.05 (0.01)	0.05 (0.01)	0.06 (0.01)	0.06 (0.01)
	b. 0.05 (0.01)	0.04 (0.01)	0.04 (0.00)	0.05 (0.01)	0.04 (0.01)	0.05 (0.01)	0.06 (0.01)	0.06 (0.01)
	c. 0.05 (0.01)	0.04 (0.01)	0.04 (0.00)	0.04 (0.01)	0.04 (0.01)	0.04 (0.01)	0.06 (0.01)	0.05 (0.01)
Кол-во выявленных	261	137	48	36	32	24	24	47

Таблица 2. Важность признаков модели первого уровня

	Признак	Важность
1	АБСОЛЮТНЫЙ_ОБЪЕМ_КРОВИ_min_min_mean	0.89
2	ФОТОЭЛЕКТРИЧЕСКИЙ ПЛЕТИЗМОГРАФ_mean_diff_mean	0.83
3	АБСОЛЮТНЫЙ_ОБЪЕМ_КРОВИ_min_mean_mean	0.75
4	АБСОЛЮТНЫЙ_ОБЪЕМ_КРОВИ_max_max_mean	0.72
5	АБСОЛЮТНЫЙ_ОБЪЕМ_КРОВИ_max_max_max	0.70
6	ДОПОЛНИТЕЛЬНЫЙ_ДАТЧИК_std_max_max	0.68
7	АБСОЛЮТНЫЙ_ОБЪЕМ_КРОВИ_min_mean_max	0.66
8	АБСОЛЮТНЫЙ_ОБЪЕМ_КРОВИ_mean_mean_mean	0.65

9	АБСОЛЮТНЫЙ_ОБЪЕМ_КРОВИ_mean_max _mean	0.61
10	ЭЛЕКТРОДЕРМАЛЬНАЯ АКТИВНОСТЬ_std_mean_max	0.57
11	АБСОЛЮТНЫЙ_ОБЪЕМ_КРОВИ_mean_max _max	0.56
12	ДОПОЛНИТЕЛЬНЫЙ ДАТЧИК_std_mean_max	0.54
13	ЭЛЕКТРОДЕРМАЛЬНАЯ АКТИВНОСТЬ_mean_diff_max	0.52
14	ДОПОЛНИТЕЛЬНЫЙ ДАТЧИК_std_mean_mean	0.50
15	БРЮШНОЕ ДЫХАНИЕ_mean_diff_max	0.49
16	ДОПОЛНИТЕЛЬНЫЙ ДАТЧИК_std_max_mean	0.48
17	ДРОЖЬ_mean_max_max	0.47
18	ДРОЖЬ_min_mean_mean	0.47
19	ТОНИЗИРУЮЩАЯ ЭЛЕКТРОДЕРМАЛЬНАЯ АКТИВНОСТЬ_amplitude_mean_max	0.47
20	ТОНИЗИРУЮЩАЯ ЭЛЕКТРОДЕРМАЛЬНАЯ АКТИВНОСТЬ_std_mean_max	0.46

Таблица 3. Важность признаков модели второго уровня

1	Признак	Важность
2	Возраст проверяемого	28.56
3	Минимальная вероятность тестов	19.53
4	Максимальная вероятность тестов	17.53
5	Средняя вероятность тестов	16.50
6	Разница вероятностей тестов	8.96
7	Пол проверяемого	8.93

Также было предположено, что информация о геомагнитных бурях на Земле [28] и погодных условиях в городе в день скрининга может улучшить качество предсказания модели. Основой для этого предположения было то, что во время геомагнитных бурь и при разных погодных условиях люди могут вести себя по-другому, что может выражаться в небольшом изменении физиологических сигналов или данные датчика могут быть немного смещены, или оба этих случая.

Также был получен ID полиграфолога из расчета того, что эти данные могут помочь модели, так как разные полиграфологи могут вызывать немного разные физиологические реакции у испытуемых, или полиграфическое оборудование, назначенное каждому полиграфологу может иметь разную погрешность измерений. Также были собраны данные о должностях испытуемых, потому что люди с разным образованием и подготовкой могут говорить правду и ложь по-разному.

Таблица 1.b отражает качество базовой модели, обученной с альтернативными данными, а табл. 4 показывает важность этих признаков. Улучшения от каждого источника относительно модели, построенной только на данных физиологических сигналов, показаны в табл. 5.

Как показывает эксперимент, все альтернативные данные улучшают качество модели, однако были оставлены только возраст, пол, должность и геомагнитные бури для эксплуатации (табл. 1.e).

Погода демонстрирует высокий прирост качества и высокую важность, и было опасение, что это может быть по техническим причинам: данные имеют сильный дисбаланс по доле выявленных по городам. Чтобы исключить смещение, вызванное дисбалансом городов, были взяты данные только по одному городу, но погода все равно имела высокое значение важности.

Таким образом, можно сделать вывод, что погода - это значимые альтернативные данные. Однако на всей выборке погода может нести в себе информацию о городе и оценка модели может быть смещен-

ной из-за дисбаланса, выявленного по городам. В то время как признак ID полиграфолога демонстрирует умеренный прирост качества, природа данного признака требует более глубокого изучения до того, как использовать этот признак при эксплуатации модели.

Например, если на самом деле улучшению качества способствует не ID полиграфолога, а ID полиграфа, то при смене устройства одним полиграфологом будет получена неверная оценка модели.

Таблица 4. Важность признаков для модели второго уровня с альтернативными данными

	Признак	Важность
1	Возраст проверяемого	8.52
2	Минимальная вероятность тестов	6.98
3	Максимальная вероятность тестов	6.64
4	Идентификатор полиграфолога	6.59
5	Средняя вероятность тестов	6.26
6	Давление	5.83
7	Текущая должность	5.55
8	Ветер	5.25
9	Точка росы	4.83
10	Назначаемая должность	4.75
11	Влажность	4.62
12	Разница вероятностей тестов	4.24
13	Текущий департамент	4.18
14	Облачность	4.17
15	Температура	4.00
16	Время	3.63
17	Назначаемый департамент	3.08
18	Скорость ветра	3.04
19	Пол проверяемого	2.41
20	Категория проверяемого	2.31

Таблица 5. Прирост качества от каждого источника дополнительных данных

	Без альт. данных	Возраст испытуемого	Пол испытуемого	Должность	Погода	Геомагн. бури	ID полиграфолога
Все темы	0.75 (0.02)	+4%	+2%	+8%	+4%	0%	+3%
Наркотические вещества	0.79 (0.02)	+6%	+3%	+11%	+9%	0%	+5%
Получение денежного возн.	0.71 (0.04)	-2%	+1%	+7%	-3%	+1%	+5%
Разглашение информации	0.73 (0.05)	+5%	+4%	+7%	+4%	+3%	+1%
Долговые обязательства	0.64 (0.03)	0%	-2%	+6%	+3%	-2%	+1%
Сторонний доход	0.81 (0.03)	+6%	+6%	+8%	+3%	-1%	+6%
Уголовные и адм. правонарушения	0.83 (0.06)	+3%	+5%	+11%	+6%	+4%	+3%
Нарушение ВНД	0.69 (0.03)	+5%	-1%	+2%	0%	0%	+3%

Если обучить отдельную модель для каждой темы, то результаты будут лучше, чем результаты базовой модели с геомагнитными бурями. Табл. 6 демонстрирует, что было получено +2% ROC AUC (6% относительного прироста), если модель обучалась только на теме наркотические вещества. Таким обра-

зом, разделение на темы поможет модели лучше обучаться и делать заключение.

Таблица 6. Качество предсказания модели, построенной на теме «Наркотические вещества».

Измерено при помощи, стратифицированной кросс-валидации.

	Все темы	Нарк. вещества	Получение ден. вознаграгр	Разгл. конф. Инф.	Долг. обязат-ва	Сторонний доход	Уголовн. и адм. правонаруш.	Нарушение ВНД
ROC AUC	0.76 (0.02) $\Delta = 0\%$	0.86 (0.02) $\Delta = +3\%$	0.74 (0.01) $\Delta = +1\%$	0.76 (0.05) $\Delta = +1\%$	0.67 (0.06) $\Delta = -2\%$	0.66 (0.07) $\Delta = -9\%$	0.78 (0.03) $\Delta = +1\%$	0.69 (0.02) $\Delta = +3\%$
Recall	0.26 (0.03) $\Delta = +1\%$	0.38 (0.04) $\Delta = -3\%$	0.10 (0.05) $\Delta = -6\%$	0.30 (0.11) $\Delta = +3\%$	0.06 (0.03) $\Delta = -3\%$	0.17 (0.08) $\Delta = +6\%$	0.21 (0.09) $\Delta = -2\%$	0.19 (0.06) $\Delta = -3\%$
Precision	0.26 (0.02) $\Delta = -2\%$	0.62 (0.04) $\Delta = +6\%$	0.19 (0.08) $\Delta = -3\%$	0.18 (0.06) $\Delta = 0\%$	0.06 (0.04) $\Delta = -10\%$	0.10 (0.04) $\Delta = -3\%$	0.11 (0.04) $\Delta = -4\%$	0.20 (0.09) $\Delta = +3\%$
F1_score	0.26 (0.02) $\Delta = -1\%$	0.47 (0.04) $\Delta = 0\%$	0.13 (0.06) $\Delta = -4\%$	0.23 (0.08) $\Delta = +2\%$	0.06 (0.03) $\Delta = -4\%$	0.12 (0.05) $\Delta = -1\%$	0.14 (0.05) $\Delta = -3\%$	0.18 (0.06) $\Delta = 0\%$
Accuracy	0.91 (0.00) $\Delta = 0\%$	0.88 (0.01) $\Delta = +1\%$	0.91 (0.01) $\Delta = +1\%$	0.92 (0.01) $\Delta = 0\%$	0.91 (0.00) $\Delta = -2\%$	0.92 (0.00) $\Delta = 0\%$	0.92 (0.01) $\Delta = -1\%$	0.88 (0.01) $\Delta = +1\%$
TNR	0.95 (0.00) $\Delta = 0\%$	0.96 (0.00) $\Delta = +1\%$	0.97 (0.00) $\Delta = +1\%$	0.95 (0.01) $\Delta = 0\%$	0.96 (0.01) $\Delta = -1\%$	0.95 (0.01) $\Delta = +1\%$	0.94 (0.00) $\Delta = -1\%$	0.93 (0.01) $\Delta = +1\%$
FPR	0.05 (0.00) $\Delta = 0\%$	0.04 (0.00) $\Delta = -1\%$	0.03 (0.00) $\Delta = -1\%$	0.05 (0.01) $\Delta = 0\%$	0.04 (0.01) $\Delta = +1\%$	0.05 (0.01) $\Delta = -1\%$	0.06 (0.00) $\Delta = +1\%$	0.07 (0.01) $\Delta = -1\%$
Кол-во выявленных	258	136	46	36	32	23	23	47

Модель, обученная на наркотических вещества, о которой велась речь в предыдущем параграфе, была применена ко всем остальным шести темам (табл. 6). В сравнении с универсальной моделью качество модели на наркотических веществах колеблется от темы к теме. Это доказывает, что модель, обученная на одной теме, может быть использована на остальных темах пусть и с незначительными падениями качества для некоторых тем.

В табл. 1 показано, что модель на некоторых темах (таких как наркотические вещества и уголовные правонарушения) работает лучше, чем на других темах (таких как сторонний доход и нарушение ВНД). Это наблюдение соответствует давней проблеме скринингов: люди не могут с уверенностью ответить на вопрос, когда полностью не уверены в ответе.

Например, в банке присутствует сотни ВНД, в каждом из которых множество страниц, и это, не упоминая о различных версиях, поэтому люди могут быть не уверены, нарушали ли они хотя бы один нормативный документ. Похожая ситуация наблюдается со сторонним доходом: некоторые люди начинают задаваться вопросом по типу "если я получал деньги от родственников, является ли это сторонним доходом?" и т.п.

Также данное наблюдение может быть использовано для улучшения качества базовой модели. Если будут найдены темы, на которых люди не уверены, то модель должна быть не уверена, обучившись на этих темах.

Неопределенные темы были удалены из всей выборки, однако в табл. 7 видно, что это действие значимо не улучшило качество модели, и качество скоринга "размытых" тем упало либо осталось неизменным. Это связано с тем, что доля выявленных в скринингах по "размытым" темам в обучающей выборке незначительна.

Таблица 7. Базовая модель без тем "Нарушение ВНД" и "Сторонний доход" в обучающей выборке

	Все темы	Нарк. вещества	Получение ден. вознаграгр	Разгл. конф. Инф.	Долг. обязат-ва	Сторонний доход	Уголовн. и адм. правонаруш.	Нарушение ВНД
ROC AUC	0.81 (0.01) $\Delta = +2\%$	0.87 (0.01) $\Delta = +3\%$	0.82 (0.03) $\Delta = 0\%$	0.76 (0.04) $\Delta = 0\%$	0.74 (0.06) $\Delta = +5\%$	0.81 (0.03) $\Delta = +1\%$	0.93 (0.02) $\Delta = +3\%$	0.67 (0.03) $\Delta = +2\%$
Recall	0.34 (0.02) $\Delta = +5\%$	0.43 (0.03) $\Delta = +3\%$	0.29 (0.06) $\Delta = +7\%$	0.27 (0.09) $\Delta = +3\%$	0.29 (0.04) $\Delta = +14\%$	0.25 (0.11) $\Delta = -12\%$	0.6 (0.05) $\Delta = +13\%$	0.21 (0.04) $\Delta = +12\%$
Precision	0.18 (0.02) $\Delta = +1\%$	0.42 (0.07) $\Delta = +1\%$	0.2 (0.04) $\Delta = +4\%$	0.09 (0.03) $\Delta = -2\%$	0.14 (0.03) $\Delta = +6\%$	0.06 (0.03) $\Delta = -6\%$	0.13 (0.06) $\Delta = +2\%$	0.09 (0.02) $\Delta = +3\%$
F1_score	0.23 (0.02) $\Delta = +5\%$	0.42 (0.05) $\Delta = +2\%$	0.23 (0.04) $\Delta = +5\%$	0.13 (0.04) $\Delta = -2\%$	0.18 (0.03) $\Delta = +8\%$	0.09 (0.04) $\Delta = -9\%$	0.20 (0.06) $\Delta = +4\%$	0.12 (0.03) $\Delta = +5\%$
Accuracy	0.93 (0.01) $\Delta = 0\%$	0.92 (0.01) $\Delta = 0\%$	0.94 (0.00) $\Delta = 0\%$	0.93 (0.00) $\Delta = -1\%$	0.94 (0.01) $\Delta = 0\%$	0.94 (0.00) $\Delta = 0\%$	0.93 (0.01) $\Delta = 0\%$	0.90 (0.01) $\Delta = -2\%$
TNR	0.95 (0.01) $\Delta = 0\%$	0.95 (0.01) $\Delta = -1\%$	0.96 (0.00) $\Delta = 0\%$	0.95 (0.01) $\Delta = -1\%$	0.95 (0.01) $\Delta = -1\%$	0.95 (0.01) $\Delta = -1\%$	0.94 (0.01) $\Delta = 0\%$	0.92 (0.01) $\Delta = -3\%$
FPR	0.05 (0.01) $\Delta = 0\%$	0.05 (0.01) $\Delta = +1\%$	0.04 (0.00) $\Delta = 0\%$	0.05 (0.01) $\Delta = +1\%$	0.05 (0.01) $\Delta = +1\%$	0.05 (0.01) $\Delta = +1\%$	0.06 (0.01) $\Delta = 0\%$	0.08 (0.01) $\Delta = +3\%$
Кол-во выявленных	261	137	48	36	32	24	24	47

Построенная универсальная модель, которую описывали выше, не использует название темы в качестве признака для обучения и оценки. Причина этого заключается в том, что была необходима модель, которая может оценивать любую тему в скрининге, а не только 7 тем, которые присутствуют в обучающей выборке.

Ниже в табл. 11 показано, как использование названия темы в роли дополнительного признака влияет на качество модели. Показано, что знание темы помогает модели оценивать "размытую" тему нарушение ВНД в то время, как качество по другим темам осталось неизменным.

Перед добавлением признака "название темы" была сбалансирована выборка по количеству выявленных риск-факторов в скринингах по темам. Эта балансировка заключается в уменьшении числа выявленных по теме Наркотические вещества в 4 раза. В табл. 11 показано, что это сокращение уменьшает качество темы Наркотические вещества, которая раньше всегда была необъяснимым лидером.

Качество по теме Наркотические вещества сравнилось с ближайшими по качеству темами, такими как получение денежного вознаграждения и уголовные, и административные правонарушения. Это наблюдение объясняет предыдущее лидерство темы НВ - это было по причине того, что по данной теме имелось наибольшее число объектов минорного класса (выявлено сокрытие информации) по сравнению с другими темами.

Было измерено улучшение качества от добавления новых, дополнительных скринингов с выявленными риск-факторами, и проведены эксперименты с различными вариациями ансамблей. Результаты представлены в табл. 8.

Суммарно ансамблирование и дополнительные данные увеличили AUC на 5% по всем темам и на 11% на выбранных темах.

Таблица 8. Влияние ансамблирования и дополнительных данных на AUC

Универсальная модель	Базовая модель	Ансамбль	+ 189 Выявленных	Ансамбль и + 189 Выявленных
Все темы	0.79 (0.01)	0.83 (0.01) $\Delta = +4\%$	0.84 (0.01) $\Delta = +5\%$	0.84 (0.01) $\Delta = +5\%$
Наркотические вещества	0.84 (0.02)	0.88 (0.01) $\Delta = +4\%$	0.84 (0.01) $\Delta = 0\%$	0.88 (0.01) $\Delta = +4\%$
Получение денежного вознаграждения	0.80 (0.03)	0.82 (0.04) $\Delta = +2\%$	0.88 (0.02) $\Delta = +8\%$	0.86 (0.02) $\Delta = +6\%$
Разглашение конфиденциальной информации	0.72 (0.05)	0.79 (0.05) $\Delta = +7\%$	0.85 (0.01) $\Delta = +13\%$	0.83 (0.02) $\Delta = +11\%$
Долговые обязательства	0.75 (0.03)	0.80 (0.05) $\Delta = +5\%$	0.84 (0.06) $\Delta = +9\%$	0.82 (0.06) $\Delta = +7\%$
Сторонний доход	0.79 (0.07)	0.83 (0.05) $\Delta = +4\%$	0.85 (0.04) $\Delta = +6\%$	0.83 (0.04) $\Delta = +4\%$
Уголовные и административные правонарушения	0.90 (0.01)	0.90 (0.01) $\Delta = 0\%$	0.91 (0.04) $\Delta = +1\%$	0.91 (0.04) $\Delta = +1\%$
Нарушение ВНД	0.69 (0.02)	0.71 (0.03) $\Delta = +2\%$	0.75 (0.02) $\Delta = +6\%$	0.73 (0.03) $\Delta = +4\%$

Теперь можно сравнить работу двух продвинутых моделей: Универсальной модели (ансамбль с альтернативными данными) и модели, обученной на теме НВ (модель на одной теме с альтернативными данными). Тестирование поиска ошибок полиграфологов проводится среди 2094 архивных скринингов. Поиск ошибок полиграфологов проводится в двух темах, которые наиболее распространены в скринингах. Были отобраны скрининги, в которых полиграфолог сделал заключение, что тема не выявлена, но модель уверенно утверждает, что выявлена ложь.

Основываясь на 15 наивысших оценках модели, обученной на теме Наркотические вещества, было отобрано 15 выводов полиграфологов с заключением "Не выявлено" в качестве кандидатов на ошибку полиграфолога по теме Наркотические вещества. Таким же путем, основываясь на оценках универсальной модели, было отобрано 15/5/5 заключений "Не выявлено" для тем "Получение денежного вознаграждения"/"Разглашение конфиденциальной информации"/"Уголовные и административные правонарушения". Таким образом, получилось 40 заключений по 36 скринингам.

Данные 36 скринингов были проверены через слепую повторную проверку у двух полиграфологов. Полиграфологи не знали результаты скринингов и не передавали результаты друг другу. Причина проведения двух проверок заключается в том, что в случае расхождения выводов оригинального заключения и одной из повторных проверок, будет слово одного полиграфолога против другого.

В эксперименте участвовали чрезвычайно опытные полиграфологи, и существует предположение, что доля ошибок полиграфологов среди всех скринингов колеблется от 0,0 до 1,0%.

Обзор двух повторных проверок представлен в табл. 9. Распределения оценок для двух релевантных тем представлены на фиг. 6.

Путем проверки 39 выводов в 35 скрининга (из 2094 скринингов) было обнаружено 30 проблемных заключений, в которых ошибка полиграфолога подтверждена двумя проверками (13 заключений) или одна из проверок не согласилась с оригинальным заключением (17 заключений). Оставшиеся 9 заключений - ошибки модели, т.е. оригинальное заключение было верное, что и подтвердили обе повторные проверки.

Было ожидаемо, что будут случаи расхождения результатов двух повторных проверок, т.к. бывают трудные случаи, в которых решение не очевидно. Для подобных трудных случаев собирается консилиум, на которых полиграфологи обсуждают их конфликтующие заключения и приходят к единому мнению.

Было замечено, что в некоторых проблемных скринингах (заключение "Выявлено" в одной или в обеих перепроверках), полиграфологи, которые проводили перепроверку, поместили, что экзаменуемый оказывал противодействие. Пропуск противодействия является ошибкой полиграфолога по определению.

Таблица 9. Результаты двух повторных проверок для наиболее вероятных кандидатов на ошибку полиграфолога (ЕЕ). МЕ - ошибка модели (подозрение ошибки полиграфолога опровергнуто), CN - требуется консилиум (проверка 1 и проверка 2 не сошлись во мнениях).

Скрининг	Нарк. в-ва		Получение ден. возн-я		Разгл. конф. инф-ии		Уголовн. и адм. правонаруш.		Итого
	score	result	score	result	score	result	score	result	
1	0.959	МЕ	0.769	CN			0.840	МЕ	
2	0.933	CN							
3	0.932	CN							
4	0.931	ЕЕ							
5	0.903	CN			0.767	CN			
6	0.900	ЕЕ							
7	0.896	МЕ							
8	0.893	ЕЕ							
9	0.891	CN							
10	0.883	CN							
11	0.867	ЕЕ							
12	0.860	ЕЕ							
13	0.858	ЕЕ							
14	0.857	МЕ							
15	0.853	CN					0.838	МЕ	
16			0.888	CN					
17			0.826	CN					
18			0.804	ЕЕ					
19			0.798	CN					
20			0.787	CN					
21			0.765	МЕ					
22			0.739	ЕЕ					
23			0.734	МЕ					
24			0.732	ЕЕ					
25			0.729	CN					
26			0.729	МЕ					
27			0.728	ЕЕ					
28			0.727	ЕЕ					
29					0.817	CN			
30					0.803	ЕЕ			
31					0.769	CN			
32					0.767	CN			
33							0.813	ЕЕ	
34							0.808	CN	
35							0.799	МЕ	
Ошибки полиграфолога	6		5		1		1		13
Требуется консилиум	6		6		4		1		17
Ошибки модели	3		3		0		3		9

Описание данных

Датасет состоит из 2094 исторических полиграфических проверок типа скрининг, включающие в себя метки "Выявлено сокрытие информации" и "Не выявлено сокрытие информации", предоставленные полиграфологами, которые проводили скрининги. Эти скрининги проводились на кандидатах и персона-

ле банка на рискованных направлениях с их согласия и в соответствии с законодательством, перед наймом, повышением или каждый год, в зависимости от позиции. Полиграфический скрининг включает в себя подмножество из 14 тем, включая наркотические вещества и получение денежного вознаграждения.

Запись скрининга включает в себя физиологические сигналы испытуемого, аудио и вопросы в текстовом формате. Каждый вопрос имеет три временных метки для каждого предъявления: начало вопроса, заданного полиграфологом, конец вопроса и момент ответа. Каждый вопрос классифицируется на четыре типа.

Типы вопросов:

Типы вопросов / на Русском	Описание
Проверочный / Пв	Вопрос напрямую связан с темой исследования
Сравнения / Вс	Вопросы, с которыми производится сравнение проверочных вопросов
Нейтральный / Нв	Вопросы, которые предполагают отсутствие или незначительное эмоциональное влияние на испытуемого
Жертвующий / 0, Жв	Вопросы, которые делают возможным уменьшение вероятности ошибки, которая может быть результатом оценки первого вопроса в последовательности

Примеры вопросов:

Тема скрининга (риск фактора)	Описание (пример вопроса)
Наркотические вещества	<i>Употребляли ли Вы незаконные наркотические вещества за последние 12 месяце?</i>
Получение денежного вознаграждения	<i>Получали ли Вы когда-либо незаконное вознаграждение от клиента?</i>
Разгл. конф. инф-ии	<i>Использовали ли вы информацию, которая содержит банковскую тайну в личных целях?</i>
Долговые обязательства	<i>Испытываете ли Вы трудности с долговыми обязательствами в данный момент?</i>
Сторонний доход	<i>Получаете ли Вы доход вне компании, о котором не сообщили?</i>
Углов. и адм. правонарушения	<i>Признавал ли Вас когда-либо суд виновным в уголовном преступлении?</i>
Нарушение ВНД	<i>Вы когда-либо нарушали Внутренний (Корпоративный) Нормативный Документ?</i>

В дополнении к физиологическим сигналам табл. 10 фиксируется пол и возраст испытуемого. Скрининг проводился на полиграфе Polyconius, моделью 7.

Таблица 10 (Описание полиграфических сигналов)

	Все темы	Нарк. вещества	Получение ден. вознаграгр	Разгл. конф. Инф.	Долг. обязат-ва	Сторонний доход	Уголовн. и адм. правонаруш.	Нарушение ВНД
ROC AUC (Std.)	a. 0.67 (0.02)	0.80 (0.04)	0.66 (0.04)	0.70 (0.04)	0.66 (0.06)	0.67 (0.03)	0.65 (0.03)	0.53 (0.04)
	b. 0.67 (0.02)	0.67 (0.10)	0.68 (0.04)	0.67 (0.04)	0.65 (0.04)	0.67 (0.06)	0.74 (0.05)	0.65 (0.06)
	c. 0.68 (0.02)	0.69 (0.05)	0.66 (0.06)	0.71 (0.03)	0.64 (0.03)	0.67 (0.04)	0.70 (0.04)	0.66 (0.03)
Recall (Std.)	a. 0.12 (0.01)	0.21 (0.08)	0.08 (0.04)	0.21 (0.1)	0.09 (0.06)	0.07 (0.04)	0.07 (0.07)	0.10 (0.04)
	b. 0.16 (0.03)	0.19 (0.10)	0.12 (0.03)	0.09 (0.03)	0.18 (0.12)	0.16 (0.04)	0.15 (0.09)	0.23 (0.07)
	c. 0.17 (0.01)	0.10 (0.06)	0.09 (0.06)	0.19 (0.05)	0.12 (0.05)	0.05 (0.02)	0.15 (0.09)	0.27 (0.04)
Precision (Std.)	a. 0.04 (0.00)	0.09 (0.04)	0.04 (0.02)	0.05 (0.02)	0.02 (0.02)	0.03 (0.02)	0.01 (0.01)	0.02 (0.01)
	b. 0.06 (0.01)	0.05 (0.03)	0.06 (0.02)	0.04 (0.02)	0.05 (0.02)	0.07 (0.02)	0.05 (0.04)	0.08 (0.02)
		0.15 (0.09)	0.06 (0.04)			0.02 (0.02)	0.05 (0.04)	0.08 (0.01)

Сигнал	Описание
ABDOMINAL_RESP	Брюшное дыхание
ABS_BLOOD_VOLUME	Абсолютный объем крови
BLOOD_VOLUME	Относительный объем крови
EDA	Электродермальная активность
HEART_RATE	Частота сердцебиения
OPTIONAL	Дополнительный датчик на выбор полиграфолога
PLE	Фотоэлектрический плевтизограф
THORACIC_RESP	Грудное дыхание
TONIC_EDA	Тонизирующая электродермальная активность
TREMOR	Дрожь

Таблица 11. (Три модели, построенные на урезанной, но сбалансированной выборке:
а) Модель на одной теме; б) Базовая модель без признака "тема"; с) Базовая модель с признаком "тема"

	с. 0.06 (0.00)			0.07 (0.03)	0.06 (0.03)			
F1_score (Std.)	a. 0.06 (0.00)	0.10 (0.04)	0.06 (0.02)	0.08 (0.03)	0.04 (0.02)	0.04 (0.03)	0.02 (0.02)	0.04 (0.02)
	b. 0.08 (0.01)	0.08 (0.04)	0.07 (0.02)	0.05 (0.02)	0.07 (0.03)	0.10 (0.02)	0.07 (0.05)	0.12 (0.03)
	c. 0.09 (0.01)	0.12 (0.07)	0.07 (0.05)	0.10 (0.01)	0.08 (0.03)	0.02 (0.02)	0.07 (0.06)	0.13 (0.02)
Accuracy (Std.)	a. 0.93 (0.01)	0.94 (0.01)	0.93 (0.01)	0.94 (0.01)	0.94 (0.01)	0.94 (0.01)	0.93 (0.01)	0.91 (0.01)
	b. 0.94 (0.00)	0.94 (0.00)	0.93 (0.01)	0.94 (0.01)	0.93 (0.00)	0.95 (0.00)	0.94 (0.00)	0.93 (0.01)
	c. 0.94 (0.00)	0.96 (0.01)	0.92 (0.01)	0.94 (0.01)	0.92 (0.00)	0.95 (0.01)	0.93 (0.00)	0.92 (0.01)
TNR (Std.)	a. 0.95 (0.01)	0.95 (0.01)	0.94 (0.01)	0.95 (0.01)	0.96 (0.01)	0.96 (0.00)	0.95 (0.01)	0.92 (0.01)
	b. 0.95 (0.00)	0.96 (0.00)	0.95 (0.01)	0.95 (0.01)	0.95 (0.01)	0.96 (0.00)	0.95 (0.01)	0.94 (0.01)
	c. 0.95 (0.00)	0.97 (0.01)	0.94 (0.01)	0.95 (0.01)	0.93 (0.01)	0.96 (0.01)	0.94 (0.01)	0.93 (0.01)
FPR (Std.)	a. 0.05 (0.01)	0.05 (0.01)	0.06 (0.01)	0.05 (0.01)	0.04 (0.01)	0.04 (0.00)	0.05 (0.01)	0.08 (0.01)
	b. 0.05 (0.00)	0.04 (0.00)	0.05 (0.01)	0.05 (0.01)	0.05 (0.01)	0.04 (0.00)	0.05 (0.01)	0.06 (0.01)
	c. 0.05 (0.00)	0.03 (0.01)	0.05 (0.01)	0.05 (0.01)	0.07 (0.01)	0.04 (0.01)	0.06 (0.01)	0.07 (0.01)
Number of DI	158	30	30	30	30	24	24	30

Обработка данных

Основная задача - оценить вывод полиграфолога (Выявлено или Не Выявлено) для определенной темы в скрининге.

Для построения модели были представлены данные в следующем формате: каждая строка в наборе данных представляет из себя запись скрининга по конкретной теме в конкретном тесте для одного экзаменуемого. Целевая переменная размечена как "Выявлено" или "Не Выявлено". Может наблюдаться смещение из-за такой разметки, так как испытуемый может солгать не на всех текстах по теме в течении скрининга.

Физиологические сигналы были извлечены из временных окон согласно временным меткам вопросов. Таким образом, изначально строка - это временной ряд физиологического сигнала для конкретного предъявления по конкретной теме.

Для каждого предъявления проверочного и контрольного вопросов сгенерировали базовые статистики: минимальное, максимальное, среднее, амплитуду и стандартное отклонение. Далее использовались минимальное, максимальное, среднее и стандартное отклонение в качестве агрегатных функций на каждом шаге. Данные каждого предъявления были сгруппированы по вопросам. Был создан дополнительный признак, характеризующий разницу между первым и последующими предъявлениями. Похожим образом вопросы были сгруппированы по темам внутри каждого теста. В конце концов каждая

строка включает в себя 600 признаков, извлеченных из записи полиграфического скрининга по конкретной теме в конкретном тесте (фиг. 10).

Модели

В данном техническом решении был использован ансамбль типа стэкинг, который состоял из двух уровней моделей, основанных на алгоритме градиентного бустинга. Двухуровневая структура требовалась для того, чтобы избежать "проклятия размерности". Модель первого уровня была обучена на 600 признаках, построенных на физиологических сигналах для каждой темы, делая заключение Выявлено/Не выявлено для каждого теста внутри скрининга. На вход второй модели подавались агрегированные по каждой теме выходные значения первой модели.

Модель второго уровня имела следующие признаки:

pred_proba_max - максимальная вероятность выявления лжи среди тестов;

pred_proba_mean - средняя вероятность выявления лжи;

pred_proba_min - минимальная вероятность обнаружения лжи среди тестов;

pred_proba_diff - разница между максимальным и средним значением вероятностей.

Эти вероятности объединяются с альтернативными данными (биографические данные, данные о магнитных бурях и т.п.) Полученный набор данных подается на вход второй модели, которая выдает вероятность лжи на скрининге по определенной теме.

Базовая модель

Данная модель (фиг. 7) не получает информацию о темах в момент обучения и предсказания. Информация о темах используется для дальнейшей агрегации внутри скрининга. Например, в первом скрининге происходит агрегация по теме "наркотические вещества" по всем тестам, то же самое происходит, например, с темой "инсайдерская торговля" во втором скрининге.

Модель, обученная на одной теме

Логика построения модели и создания признаков такая же, как и в базовой модели. Различие заключается в том, что для обучения берутся признаки только для одной темы. Отобраны данные только по одной теме перед тем, как применяется модель первого уровня. На фиг. 8 показан порядок обучения модели для темы наркотические вещества, в то время как остальные темы убраны из тренировочного набора данных.

Универсальная модель

Было принято решение использовать лучшие стороны моделей, описанных выше. Таким образом, был построен ансамбль из существующих архитектур (фиг. 9). После серии экспериментов лучший результат показала архитектура, в которой усредняется степень уверенности следующих моделей:

базовая модель - алгоритм: градиентным бустинг; с альтернативными данными; модель на одной теме - алгоритм: градиентный бустинг; с альтернативными данными; базовая модель - алгоритм: случайный лес.

Данный ансамбль был применен ко всем темам, кроме темы Наркотические вещества. Помимо привычных преимуществ ансамблей было рационально использовать разные архитектуры (градиентный бустинг и случайный лес) для того, чтобы исключить чистую ошибку модели, и выделены будут те скрининги, в которых присутствует ошибка полиграфолога.

Как показано на фиг. 1 способ автоматической полиграфической проверки (100), выполняется с помощью вычислительной системы, содержащей модель машинного обучения, и состоит из нескольких взаимосвязанных этапов.

На этапе 101 получают записи полиграфных проверок, содержащие, по меньшей мере, сигналы датчиков с временными шкалами, на которых промаркированы начало и конец вопроса.

Далее на этапе 102 получают дополнительные данные, содержащие по меньшей мере возраст проверяемого, пол, должностную информацию.

Далее на этапе 103 осуществляют обработку полученных сигналов и дополнительных данных с помощью модели машинного обучения (МО), причем в ходе указанной обработки осуществляется:

определение временных интервалов для извлечения переменных на основе временных меток начала и конца вопроса и временной метки ответа, и на основе типа и темы вопроса. На данном этапе в зависимости от логики построения модели выделяются временные метки только тех вопросов (и ответов), физиологические данные которых будут участвовать в обучении и тестировании модели. Например, в универсальной модели используются вопросы по всем темам, поэтому будут выбраны все метки;

извлечение переменных из каждого сигнала на определенных временных интервалах. На данном этапе из базы данных извлекаются физиологические сигналы на выбранных ранее интервалах;

обработка полученных переменных из сигналов, при которой выполняется нормализация и конкатенация обработанных переменных и построение на их основе вектора. На данном этапе каждое значение численной переменной становится равным значению от 0 до 1;

разделение дополнительных данных на категориальные и численные переменные.

обработка полученных переменных из дополнительных данных, при которой выполняется векторизация категориальных переменных и нормализация численных переменных. На данном этапе каждое значение численной переменной становится равным значению от 0 до 1;

конкатенация обработанных переменных, извлеченных из каждого сигнала, и обработанных дополнительных переменных, и построение на их основе вектора. На данном этапе из всех выбранных релевантных для конкретного вывода полиграфолога данных получают один вектор, одинаково применимый как для тренировки модели (при наличии вывода полиграфолога), так и для принятия моделью решения (выявлено/не выявлено);

подача упомянутого вектора в модель МО для получения выходного значения модели МО. На данном этапе получают скор модели в интервале от 0 до 1;

сравнение выходного значения модели с заданным пороговым значением.

И на этапе 104 определяют, что ответ является ложью если выходное значение выше или равно пороговому значению или ответ является правдой если выходное значение ниже порогового значения.

В одном из частных вариантов реализации способа модель МО определяет ответ как правдивый или ложный не по одному полученному ответу, а по совокупности вопросов и ответов, связанных с темой полиграфической проверки, при этом получаемые переменные сигналы объединяются в единый вектор, либо усредняются перед передачей в модели МО.

В другом частном варианте реализации способа модель МО имеет тип градиентный бустинг, случайный лес или нейронная сеть.

В другом частном варианте реализации способа модель МО обучена на одной из тем для проверок или их комбинации, где темами для проверок являются наркотические вещества, получение дополнительного вознаграждения, разглашение конфиденциальной информации, долговые обязательства, сторонний доход, уголовные правонарушения, административные правонарушения, нарушения внутренних нормативных документов (ВНД).

В другом частном варианте реализации способа записи полиграфических проверок содержат сигналы с датчиков, включающие по меньшей мере одно из: частота сердечного сокращения (ЧСС), кожно-гальванические реакции (КГР), артериальное давление, верхнее и нижнее дыхание, пьезоплетизмограмму, фотоплетизмограмму, термических, движения зрачка или их комбинации.

В другом частном варианте реализации способа дополнительные данные содержат по меньшей мере одно из: идентификационный номер полиграфологов, идентификационный номер полиграфов, информацию о погодных условиях, результаты электроэнцефалограммы, магнитно-резонансной томографии, функциональной ближней инфракрасной спектроскопии, информацию о геомагнитных бурях или их комбинации.

Реализация данного технического решения на базе одной модели МО, позволяет автоматизировать процесс полиграфических проверок и с высокой точностью выявлять сокрытие информации.

Как показано на фиг. 2 в другом частном варианте реализации, способ автоматической полиграфической проверки (200), выполняется с помощью вычислительной системы, содержащей по меньшей мере две модели машинного обучения, и состоит из нескольких взаимосвязанных этапов.

На этапе 201 получают записи полиграфных проверок, содержащие, по меньшей мере, сигналы датчиков с временными шкалами, на которых промаркированы начало и конец вопроса.

Далее на этапе 202 получают дополнительные данные, содержащие, по меньшей мере, возраст проверяемого, пол, должностную информацию.

Далее на этапе 203 осуществляют обработку полученных сигналов с помощью первой модели машинного обучения (МО), причем в ходе указанной обработки осуществляется

определение временных интервалов для извлечения переменных на основе временных меток начала и конца вопроса и временной метки ответа, и на основе типа и темы вопроса. На данном этапе в зависимости от логики построения модели выделяются временные метки только тех вопросов (и ответов), физиологические данные которых будут участвовать в обучении и тестировании модели. Например, в универсальной модели используются вопросы по всем темам, поэтому будут выбраны все метки;

извлечение переменных из каждого сигнала на определенных временных интервалах. На данном этапе из базы данных извлекаются физиологические сигналы на выбранных ранее интервалах;

обработка полученных переменных из сигналов, при которой выполняется нормализация и конкатенация обработанных переменных и построение на их основе вектора. На данном этапе каждое значение численной переменной становится равным значению от 0 до 1;

подача упомянутого вектора в 1-ю модель МО для получения выходного значения 1-й модели МО. На данном этапе получают скор модели в интервале от 0 до 1;

передача выходного значения 1-й модели МО на вход 2-й модели МО.

Далее на этапе 204 с помощью второй модели машинного обучения (МО) осуществляют обработку выходного значения 1-й модели МО и дополнительных данных, причем в ходе указанной обработки осуществляется

разделение дополнительных данных на категориальные и численные переменные;

обработка полученных переменных из дополнительных данных, при которой выполняется векторизация категориальных переменных и нормализация численных переменных. На данном этапе каждое значение численной переменной становится равным значению от 0 до 1;

конкатенация обработанных дополнительных переменных, а также выходного значения 1-й модели

МО и построение на их основе вектора. На данном этапе из всех выбранных релевантных для конкретного вывода полиграфолога данных получают один вектор, одинаково применимый как для тренировки модели (при наличии вывода полиграфолога), так и для принятия моделью решения (выявлено/не выявлено);

подача вектора во 2-ю модель МО для получения выходного значения 2-й модели МО. На данном этапе получают скор модели в интервале от 0 до 1;

сравнение выходного значения 2-й модели с заданным пороговым значением.

И на этапе 205 определяют, что ответ является ложью, если выходное значение выше или равно пороговому значению или ответ является правдой, если выходное значение ниже порогового значения.

В одном из частных вариантов реализации способа модели МО определяют ответ как правдивый или ложный не по одному полученному ответу, а по совокупности вопросов и ответов, связанных с темой полиграфической проверки, при этом получаемые переменные сигналы объединяются в единый вектор, либо усредняются перед передачей в модели МО.

В другом частном варианте реализации способа 1-я модель МО имеет тип градиентный бустинг, случайный лес или нейронная сеть.

В другом частном варианте реализации способа 2-я модель МО имеет тип градиентный бустинг, случайный лес или нейронная сеть.

В другом частном варианте реализации способа модели МО обучены на одной из тем для проверок или их комбинации, где темами для проверок являются наркотические вещества, получение дополнительного вознаграждения, разглашение конфиденциальной информации, долговые обязательства, сторонний доход, уголовные правонарушения, административные правонарушения, нарушения внутренних нормативных документов (ВНД).

В другом частном варианте реализации способа записи полиграфических проверок содержат сигналы с датчиков, включающие по меньшей мере одно из: частота сердечного сокращения (ЧСС), кожно-гальванические реакции (КГР), артериальное давление, верхнее и нижнее дыхание, пьезоплетизмограмму, фотоплетизмограмму, термических, движения зрачка или их комбинации.

В другом частном варианте реализации способа дополнительные данные содержат по меньшей мере одно из: идентификационный номер полиграфологов, идентификационный номер полиграфов, информацию о погодных условиях, результаты электроэнцефалограммы, магнитно-резонансной томографии, функциональной ближней инфракрасной спектроскопии, информацию о геомагнитных бурях или их комбинации.

Реализация данного технического решения на базе двух моделей МО, позволяет автоматизировать процесс полиграфических проверок и с высокой точностью выявлять сокрытие информации. Данная реализация позволяет повысить точность выявления сокрытия информации за счет предварительного обесчета физиологических данных в модели первого уровня и использования уже агрегатов из первой модели (с добавлением доп. нефизиологических данных) в модели 2-го уровня.

Как показано на фиг. 3 в другом частном варианте реализации, способ автоматической полиграфической проверки (300) выполняется с помощью вычислительной системы, содержащей по меньшей мере два ансамбля моделей машинного обучения, и состоит из нескольких взаимосвязанных этапов.

На этапе 301 получают записи полиграфических проверок, содержащие, по меньшей мере, сигналы датчиков с временными шкалами, на которых промаркированы начало и конец вопроса.

Далее на этапе 302 получают дополнительные данные, содержащие, по меньшей мере, возраст проверяемого, пол, должностную информацию.

Далее на этапе 303 осуществляют обработку полученных сигналов с помощью первого ансамбля моделей МО, обученного на одной теме, причем в ходе указанной обработки осуществляется

обработка сигналов первой моделью МО, в ходе которой выполняется

определение временных интервалов для извлечения переменных на основе временных меток начала и конца вопроса и временной метки ответа, и на основе типа и темы вопроса. На данном этапе в зависимости от логики построения модели выделяются временные метки только тех вопросов (и ответов), физиологические данные которых будут участвовать в обучении и тестировании модели. Например, в универсальной модели используются вопросы по всем темам, поэтому будут выбраны все метки;

извлечение переменных из каждого сигнала на определенных временных интервалах. На данном этапе из базы данных извлекаются физиологические сигналы на выбранных ранее интервалах;

обработка полученных переменных из сигналов, при которой выполняется нормализация и конкатенация обработанных переменных и построение на их основе вектора. На данном этапе каждое значение численной переменной становится равным значению от 0 до 1;

подача упомянутого вектора в 1-ю модель МО для получения выходного значения 1-й модели МО. На данном этапе получают скор модели в интервале от 0 до 1;

передачу выходного значения 1-й модели МО на вход 2-й модели МО.

с помощью второй модели МО осуществляют обработку выходного значения 1-й модели МО и дополнительных данных, причем в ходе указанной обработки осуществляется

разделение дополнительных данных на категориальные и численные переменные;

обработка полученных переменных из дополнительных данных, при которой выполняется векторизация категориальных переменных и нормализация численных переменных. На данном этапе каждое значение численной переменной становится равным значению от 0 до 1;

конкатенация обработанных дополнительных переменных, а также выходного значения 1-й модели МО и построение на их основе вектора. На данном этапе из всех выбранных релевантных для конкретного вывода полиграфолога данных получают один вектор, одинаково применимый как для тренировки модели (при наличии вывода полиграфолога), так и для принятия моделью решения (выявлено/не выявлено);

подача упомянутого вектора во 2-ю модель МО для получения выходного значения 2-й модели МО. На данном этапе получают скор модели в интервале от 0 до 1;

подача выходного значения 2-ой модели МО в третью модель МО для формирования выходного значения первого ансамбля.

Далее на этапе 304 осуществляют обработку полученных сигналов с помощью второго ансамбля моделей МО, обученного на комбинации тем, причем в ходе указанной обработки осуществляется

обработка сигналов первой моделью МО в ходе которой выполняется

определение временных интервалов для извлечения переменных на основе временных меток начала и конца вопроса и временной метки ответа и на основе типа и темы вопроса. На данном этапе в зависимости от логики построения модели выделяются временные метки только тех вопросов (и ответов), физиологические данные которых будут участвовать в обучении и тестировании модели. Например, в универсальной модели используются вопросы по всем темам, поэтому будут выбраны все метки;

извлечение переменных из каждого сигнала на определенных временных интервалах. На данном этапе из базы данных извлекаются физиологические сигналы на выбранных ранее интервалах;

обработка полученных переменных из сигналов, при которой выполняется нормализация и конкатенация обработанных переменных и построение на их основе вектора. На данном этапе каждое значение численной переменной становится равным значению от 0 до 1;

подача упомянутого вектора в 1-ю модель МО для получения выходного значения 1-й модели МО. На данном этапе получают скор модели в интервале от 0 до 1;

передача выходного значения 1-й модели МО на вход 2-й модели МО.

с помощью второй модели МО осуществляют обработку выходного значения 1-й модели МО и дополнительных данных, причем в ходе указанной обработки осуществляется

разделение дополнительных данных на категориальные и численные переменные;

обработка полученных переменных из дополнительных данных, при которой выполняется векторизация категориальных переменных и нормализация численных переменных. На данном этапе каждое значение численной переменной становится равным значению от 0 до 1;

конкатенация обработанных дополнительных переменных, а также выходного значения 1-й модели МО и построение на их основе вектора. На данном этапе из всех выбранных релевантных для конкретного вывода полиграфолога данных получают один вектор, одинаково применимый как для тренировки модели (при наличии вывода полиграфолога), так и для принятия моделью решения (выявлено/не выявлено);

подача упомянутого вектора во 2-ю модель МО для получения выходного значения 2-й модели МО. На данном этапе получают скор модели в интервале от 0 до 1;

подача выходного значения 2-ой модели МО в третью модель МО для формирования выходного значения второго ансамбля.

Далее на этапе 305 с помощью третьей модели МО осуществляют обработку выходных значений первого и второго ансамблей МО, причем в ходе указанной обработки осуществляется

конкатенация обработанных выходных значений первого и второго ансамблей и построение на их основе вектора.

подача упомянутого вектора в 3-ю модель МО для получения выходного значения 3-й модели МО. На данном этапе получают скор модели в интервале от 0 до 1;

сравнение выходного значения 3-й модели с заданным пороговым значением.

И на этапе 306 определяют, что ответ является ложью, если выходное значение выше или равно пороговому значению или ответ является правдой, если выходное значение ниже порогового значения.

В одном из частных вариантов реализации способа модели МО обучены на одной из тем для проверок или их комбинации, где темами для проверок являются: наркотические вещества, получение дополнительного вознаграждения, разглашение конфиденциальной информации, долговые обязательства, сторонний доход, уголовные правонарушения, административные правонарушения, нарушения внутренних нормативных документов (ВНД).

В другом частном варианте реализации способа модели МО первого и второго ансамблей имеют тип, выбираемый из группы: градиентный бустинг, случайный лес или нейронная сеть.

В другом частном варианте реализации способа третья модель МО имеет тип, выбираемый из группы: логистическая регрессия, случайный лес, градиентный бустинг или усреднение.

В другом частном варианте реализации способа записи полиграфических проверок содержат сигнала

лы с датчиков, включающие по меньшей мере одно из: частота сердечного сокращения (ЧСС), кожно-гальваническая реакции (КГР), артериальное давление, верхнее и нижнее дыхание, пьезоплетизмограмму, фотоплетизмограмму, термические, движения зрачка или их комбинации.

В другом частном варианте реализации способа дополнительные данные содержат по меньшей мере одно из: идентификационный номер полиграфологов, идентификационный номер полиграфов, информацию о погодных условиях, результаты электроэнцефалограммы, магнитно-резонансной томографии, функциональной ближней инфракрасной спектроскопии, информацию о геомагнитных бурях или их комбинации.

Реализация данного технического решения на базе двух ансамблей моделей МО, позволяет автоматизировать процесс полиграфических проверок и с высокой точностью выявлять сокрытие информации. Данная реализация позволяет повысить точность выявления сокрытия информации за счет различий архитектур между двумя ансамблями и за счет добавления верхнеуровневой модели.

Как показано на фиг. 3 в другом частном варианте реализации, способ автоматической полиграфической проверки (400) выполняется с помощью вычислительной системы, содержащей по меньшей мере три ансамбля моделей машинного обучения, и состоит из нескольких взаимосвязанных этапов.

На этапе 401 получают записи полиграфических проверок, содержащие, по меньшей мере, сигналы датчиков с временными шкалами, на которых промаркированы начало и конец вопроса.

Далее на этапе 402 получают дополнительные данные, содержащие, по меньшей мере, возраст проверяемого, пол, должностную информацию.

Далее на этапе 403 осуществляют обработку полученных сигналов с помощью первого ансамбля моделей МО, обученного на одной теме, причем в ходе указанной обработки осуществляется

обработка сигналов первой моделью МО, в ходе которой выполняется

определение временных интервалов для извлечения переменных на основе временных меток начала и конца вопроса и временной метки ответа и на основе типа и темы вопроса. На данном этапе в зависимости от логики построения модели выделяются временные метки только тех вопросов (и ответов), физиологические данные которых будут участвовать в обучении и тестировании модели. Например, в универсальной модели используются вопросы по всем темам, поэтому будут выбраны все метки;

извлечение переменных из каждого сигнала на определенных временных интервалах. На данном этапе из базы данных извлекаются физиологические сигналы на выбранных ранее интервалах;

обработка полученных переменных из сигналов, при которой выполняется нормализация и конкатенация обработанных переменных и построение на их основе вектора. На данном этапе каждое значение численной переменной становится равным значению от 0 до 1;

подача упомянутого вектора в 1-ю модель МО для получения выходного значения 1-й модели МО. На данном этапе получают скор модели в интервале от 0 до 1;

передача выходного значения 1-й модели МО на вход 2-й модели МО;

с помощью второй модели МО осуществляют обработку выходного значения 1-й модели МО и дополнительных данных, причем в ходе указанной обработки осуществляется

разделение дополнительных данных на категориальные и численные переменные;

обработка полученных переменных из дополнительных данных, при которой выполняется векторизация категориальных переменных и нормализация численных переменных. На данном этапе каждое значение численной переменной становится равным значению от 0 до 1;

конкатенация обработанных дополнительных переменных, а также выходного значения 1-й модели МО и построение на их основе вектора. На данном этапе из всех выбранных релевантных для конкретного вывода полиграфолога данных получают один вектор, одинаково применимый как для тренировки модели (при наличии вывода полиграфолога), так и для принятия моделью решения (выявлено/не выявлено);

подача упомянутого вектора во 2-ю модель МО для получения выходного значения 2-й модели МО. На данном этапе получают скор модели в интервале от 0 до 1;

подача выходного значения 2-ой модели МО в третью модель МО для формирования выходного значения первого ансамбля.

Далее на этапе 404 осуществляют обработку полученных сигналов с помощью второго ансамбля моделей МО, обученного на комбинации тем, причем в ходе указанной обработки осуществляется

обработка сигналов первой моделью МО в ходе которой выполняется

определение временных интервалов для извлечения переменных на основе временных меток начала и конца вопроса и временной метки ответа и на основе типа и темы вопроса. На данном этапе в зависимости от логики построения модели выделяются временные метки только тех вопросов (и ответов), физиологические данные которых будут участвовать в обучении и тестировании модели. Например, в универсальной модели используются вопросы по всем темам, поэтому будут выбраны все метки;

извлечение переменных из каждого сигнала на определенных временных интервалах. На данном этапе из базы данных извлекаются физиологические сигналы на выбранных ранее интервалах;

обработка полученных переменных из сигналов, при которой выполняется нормализация и конкатенация обработанных переменных и построение на их основе вектора. На данном этапе каждое значение

численной переменной становится равным значению от 0 до 1;

подача упомянутого вектора в 1-ю модель МО для получения выходного значения 1-й модели МО. На данном этапе получают скор модели в интервале от 0 до 1;

передача выходного значения 1-й модели МО на вход 2-й модели МО;

с помощью второй модели МО осуществляют обработку выходного значения 1-й модели МО и дополнительных данных, причем в ходе указанной обработки осуществляется

разделение дополнительных данных на категориальные и численные переменные;

обработка полученных переменных из дополнительных данных, при которой выполняется векторизация категориальных переменных и нормализация численных переменных. На данном этапе каждое значение численной переменной становится равным значению от 0 до 1;

конкатенация обработанных дополнительных переменных, а также выходного значения 1-й модели МО и построение на их основе вектора. На данном этапе из всех выбранных релевантных для конкретного вывода полиграфолога данных получают один вектор, одинаково применимый как для тренировки модели (при наличии вывода полиграфолога), так и для принятия моделью решения (выявлено/не выявлено);

подача упомянутого вектора во 2-ю модель МО для получения выходного значения 2-й модели МО.

На данном этапе получают скор модели в интервале от 0 до 1;

подача выходного значения 2-ой модели МО в третью модель МО для формирования выходного значения второго ансамбля.

Далее на этапе 405 осуществляют обработку полученных сигналов с помощью третьего ансамбля моделей машинного обучения, обученного на комбинации тем, причем в ходе указанной обработки осуществляется

обработка сигналов первой моделью МО, в ходе которой выполняется

определение временных интервалов для извлечения переменных на основе временных меток начала и конца вопроса и временной метки ответа, и на основе типа и темы вопроса. На данном этапе в зависимости от логики построения модели выделяются временные метки только тех вопросов (и ответов), физиологические данные которых будут участвовать в обучении и тестировании модели. Например, в универсальной модели используются вопросы по всем темам, поэтому будут выбраны все метки;

извлечение переменных из каждого сигнала на определенных временных интервалах. На данном этапе из базы данных извлекаются физиологические сигналы на выбранных ранее интервалах;

обработка полученных переменных из сигналов, при которой выполняется нормализация и конкатенация обработанных переменных и построение на их основе вектора. На данном этапе каждое значение численной переменной становится равным значению от 0 до 1;

подача упомянутого вектора в 1-ю модель МО для получения выходного значения 1-й модели МО.

На данном этапе получают скор модели в интервале от 0 до 1;

передача выходного значения 1-й модели МО на вход 2-й модели МО;

с помощью второй модели МО осуществляют обработку выходного значения 1-й модели МО и дополнительных данных, причем в ходе указанной обработки осуществляется

разделение дополнительных данных на категориальные и численные переменные;

обработка полученных переменных из дополнительных данных, при которой выполняется векторизация категориальных переменных и нормализация численных переменных. На данном этапе каждое значение численной переменной становится равным значению от 0 до 1;

конкатенация обработанных дополнительных переменных, а также выходного значения 1-й модели МО и построение на их основе вектора. На данном этапе из всех выбранных релевантных для конкретного вывода полиграфолога данных получают один вектор, одинаково применимый как для тренировки модели (при наличии вывода полиграфолога), так и для принятия моделью решения (выявлено/не выявлено);

подача упомянутого вектора во 2-ю модель МО для получения выходного значения 2-й модели МО.

На данном этапе получают скор модели в интервале от 0 до 1;

подача выходного значения 2-ой модели МО в третью модель МО для формирования выходного значения третьего ансамбля.

Далее на этапе 406 с помощью третьей модели МО осуществляют обработку выходных значений первого, второго и третьего ансамблей МО, причем в ходе указанной обработки осуществляется

конкатенация обработанных выходных значений первого, второго и третьего ансамблей и построение на их основе вектора. На данном этапе каждое значение численной переменной становится равным значению от 0 до 1;

подача упомянутого вектора в 3-ю модель МО для получения выходного значения 3-й модели МО.

На данном этапе получают скор модели в интервале от 0 до 1;

сравнение выходного значения 3-й модели с заданным пороговым значением.

И на этапе 407 осуществляют определение того, что ответ является ложью если выходное значение выше или равно пороговому значению или ответ является правдой, если выходное значение ниже порогового значения.

В одном из частных вариантов реализации способа модели МО обучены на одной из тем для проверок или их комбинации, где темами для проверок являются наркотические вещества, получение дополнительного вознаграждения, разглашение конфиденциальной информации, долговые обязательства, сторонний доход, уголовные правонарушения, административные правонарушения, нарушения внутренних нормативных документов (ВНД).

В другом частном варианте реализации способа модели МО первого, второго и третьего ансамблей имеют тип, выбираемый из группы: градиентный бустинг, случайный лес или нейронная сеть.

В другом частном варианте реализации способа третья модель МО имеет тип, выбираемый из группы: логистическая регрессия, случайный лес, градиентный бустинг или усреднение.

В другом частном варианте реализации способа записи полиграфических проверок содержат сигналы с датчиков, включающие по меньшей мере одно из: частота сердечного сокращения (ЧСС), кожно-гальваническая реакции (КГР), артериальное давление, верхнее и нижнее дыхание, пьезоплетизмограмму, фотоплетизмограмму, термических, движения зрачка или их комбинации.

В другом частном варианте реализации способа дополнительные данные содержат по меньшей мере одно из: идентификационный номер полиграфологов, идентификационный номер полиграфов, информацию о погодных условиях, результаты электроэнцефалограммы, магнитно-резонансной томографии, функциональной ближней инфракрасной спектроскопии, информацию о геомагнитных бурях или их комбинации.

Реализация данного технического решения на базе трех ансамблей моделей МО позволяет автоматизировать процесс полиграфических проверок и с высокой точностью выявлять сокрытие информации. Данная реализация позволяет повысить точность выявления сокрытия информации за счет различий архитектур между тремя ансамблями и за счет добавления верхнеуровневой модели.

На фиг. 11 представлен пример общего вида вычислительной системы (500), которая обеспечивает реализацию заявленного способа или является частью компьютерной системы, например, сервером, персональным компьютером, частью вычислительного кластера, обрабатывающим необходимые данные для осуществления заявленного технического решения.

В общем случае система (500) содержит объединенные общей шиной информационного обмена один или несколько процессоров (501), средства памяти, такие как ОЗУ (302) и ПЗУ (503), интерфейсы ввода/вывода (504), устройства ввода/вывода (1105) и устройство для сетевого взаимодействия (506).

Процессор (501) (или несколько процессоров, многоядерный процессор и т.п.) может выбираться из ассортимента устройств, широко применяемых в настоящее время, например, таких производителей, как Intel™, AMD™, Apple™, Samsung Exynos™, MediaTek™, Qualcomm Snapdragon™ и т.п. Под процессором или одним из используемых процессоров в системе (500) также необходимо учитывать графический процессор, например GPU NVIDIA или Graphcore, тип которых также является пригодным для полного или частичного выполнения способа, а также может применяться для обучения и применения моделей машинного обучения в различных информационных системах.

ОЗУ (502) представляет собой оперативную память и предназначено для хранения исполняемых процессором (501) машиночитаемых инструкций для выполнения необходимых операций по логической обработке данных. ОЗУ (502), как правило, содержит исполняемые инструкции операционной системы и соответствующих программных компонент (приложения, программные модули и т.п.). При этом, в качестве ОЗУ (502) может выступать доступный объем памяти графической карты или графического процессора.

ПЗУ (503) представляет собой одно или более устройств постоянного хранения данных, например жесткий диск (HDD), твердотельный накопитель данных (SSD), флэш-память (EEPROM, NAND и т.п.), оптические носители информации (CD-R/RW, DVD-R/RW, BlueRay Disc, MD) и др.

Для организации работы компонентов системы (500) и организации работы внешних подключаемых устройств применяются различные виды интерфейсов В/В (504). Выбор соответствующих интерфейсов зависит от конкретного исполнения вычислительного устройства, которые могут представлять собой, не ограничиваясь: PCI, AGP, PS/2, IrDa, FireWire, LPT, COM, SATA, IDE, Lightning, USB (2.0, 3.0, 3.1, micro, mini, type C), TRS/Audio jack (2.5, 3.5, 6.35), HDMI, DVI, VGA, Display Port, RJ45, RS232 и т.п.

Для обеспечения взаимодействия пользователя с вычислительной системой (300) применяются различные средства (505) В/В информации, например клавиатура, дисплей (монитор), сенсорный дисплей, тач-пад, джойстик, манипулятор, мышь, световое перо, стилус, сенсорная панель, трекбол, динамики, микрофон, средства дополненной реальности, оптические сенсоры, планшет, световые индикаторы, проектор, камера, средства биометрической идентификации (сканер сетчатки глаза, сканер отпечатков пальцев, модуль распознавания голоса) и т.п.

Средство сетевого взаимодействия (506) обеспечивает передачу данных посредством внутренней или внешней вычислительной сети, например, Интранет, Интернет, ЛВС и т.п. В качестве одного или более средств (306) может использоваться, но не ограничиваясь Ethernet карта, GSM модем, GPRS модем, LTE модем, 5G модем, модуль спутниковой связи, NFC модуль, Bluetooth и/или BLE модуль, Wi-Fi модуль и др.

Представленные материалы заявки раскрывают предпочтительные примеры реализации техниче-

ского решения и не должны трактоваться как ограничивающие иные, частные примеры его воплощения, не выходящие за пределы испрашиваемой правовой охраны, которые являются очевидными для специалистов соответствующей области техники.

Источники информации

- [1] M. Harris, "The Lie Generator: Inside the Black Mirror World of Polygraph Job Screenings," in *Wired.com*, 2018.
- [2] B. Banerjee and G. Chatterjee, "The world of lie detection: a study into state of lie detection usage by state and society in Asia, Africa and Europe," 2021.
- [3] S. E. Fienberg, J. J. Blascovich, J. T. Cacioppo, R. J. Davidson, P. Ekman, D. L. Faigman and e. al., "The polygraph and lie detection," in *National Research Council, The National Academies Press*, Washington, DC, 2003.
- [4] A. Slavkovic, "Evaluating polygraph data," in *Carnegie Mellon University*, 2002.
- [5] J. Synnott, D. Dietzel and M. Ioannou, "The Polygraph: History, Methodology and Current Status," in *Reviewing Crime Psychology*, 2020.
- [6] G. Krishnamurthy, N. Majumder, S. Poria and E. Cambria, "A Deep Learning Approach for Multimodal Deception Detection," in *19th International Conference on Computational Linguistics and Intelligent Text Processing (CICLing)*, Hanoi, Vietnam, 2018.
- [7] D. Avola, L. Cinque, G. L. Foresti and D. Pannone, "Automatic Deception Detection in RGB videos using Facial Action Units," in *ICDSC 2019: Proceedings of the 13th International Conference on Distributed Smart Cameras*, Trento, 2019.
- [8] N. Samadiani, G. Huang, B. Cai, W. Luo, C.-H. Chi, Y. Xiang and J. He, "A Review on Automatic Facial Expression Recognition Systems Assisted by Multimodal Sensor Data," in *Sensors, Vol. 19, Issue 8, Sp. Issue Sensor Applications on Face Analysis*, 2019.
- [9] A. K. Webb, C. R. Honts, J. C. Kircher, P. Bernhardt and A. E. Cook, "Effectiveness of pupil diameter in a probable-lie comparison question test for deception," in *Legal and Criminological Psychology, Vol. 14, Issue 2*, 2010.
- [10] J. J. Walczyk, K. T. Mahoney, D. Doverspike and D. A. Griffith-Ross, "Cognitive Lie Detection: Response Time and Consistency of Answers as Cues to Deception," in *Journal of Business and Psychology, Vol. 24*, 2009.
- [11] A. Shuster, L. Inzelberg, O. Ossmy, L. Izakson, Y. Hancin and D. Levy, "Lie to my face: An electromyography approach to the study of deceptive behavior," in *Brain and Behavior, vol.11 issue 12*, 2021.
- [12] V. Abootalebi, M. H. Moradi and M. A. Khalilzadeh, "A new approach for EEG feature extraction in P300-based lie detection," in *Computer Methods and Programs in Biomedicine, vol.94, issue 1*, 2009.
- [13] A. Kozel, K. Johnson, Q. Mu, E. Grenesko, S. Laken and M. George, "Detecting Deception Using Functional Magnetic Resonance Imaging," in *Biological Psychiatry, Vol. 58, Issue 8*, 2005.
- [14] M. J. Farah, J. B. Hutchinson, E. A. Phelps and A. D. Wagner, "Functional MRI-based lie detection: scientific and societal challenges," in *Nature Reviews Neuroscience, Vol. 14*, 2014.
- [15] M. Monaro, C. Galante, R. Spolaor, Q. Q. Li, L. Gamberini, M. Conti and G. Sartori, "Covert lie detection using keyboard dynamics," in *Nature, Scientific Reports, Vol.8*, 2018.
- [16] L. Sousedikova, M. Hromada and M. Adamek, "Analysis of Artificial Intelligence Lie Detector Developed for Airport Security," in *Tomas Bata University in Zlin*, 2021.

- [17] J. Sánchez-Monedero and L. Dencik, "The politics of deceptive borders: 'biomarkers of deceit and the case of iBorderCtrl,'" in *Information, Communication & Society*, Vol. 25 Issue 3, 2022.
- [18] L. Quijano-Sánchez, F. Liberatore, J. Camacho-Collados and M. Camacho-Collados, "Applying automatic text-based detection of deceptive language to police reports: Extracting behavioral patterns from a multi-step classification model to understand how we lie to the police," in *Knowledge-Based Systems*, Vol. 149, 2018.
- [19] "Police use a computer to expose false testimony. A lie-detection system being used by Spanish police highlights concerns about algorithms.," in *Nature, Editorial*, 2018.
- [20] G. Ben-Shakhar and W. Iacono, "Fallacies in the estimation of the validity of the Comparison Question Polygraph Test: A reply to Ginton (2020)," in *Investigative Psychology and Offender Profiling*, Vol.18, Issue 3, 2021.
- [21] D. Grubin and L. Madsen, "Lie detection and the polygraph: A historical review," in *The Journal of Forensic Psychiatry & Psychology*, 2005.
- [22] C. Hinkle, "The Modern Lie Detector: AI-Powered Affect Screening and the Employee Polygraph Protection Act (EPPA)," in *The Georgetown Law Journal* Vol.109, Georgetown, 2021.
- [23] J. Bittle, "Lie detectors have always been suspect. AI has made the problem worse.," in *MIT Technology review*, 2020.
- [24] L. Saxe, "Science and the CQT polygraph - A theoretical critique," in *Integrative Physiological and Behavioral Science*, 1991.
- [25] A. M. Perkey, "Recommendations for uniform polygraph examinations for preemployment screening of law enforcement applicants," in *University of Wisconsin-Platteville*, 2021.
- [26] W. Egerton, "Use of the Polygraph to Screen Police Candidates," in *Law Enforcement Management Institute of Texas (LEMIT)*, North Richland Hills, Texas, 2020.
- [27] D. Baur, "Federal Psychophysiological Detection of Deception Examiner Handbook," in *Counterintelligence Field Activity Technical Manual*, 2006.
- [28] J. Matzka, O. Bronkalla, K. Tornow, K. Elger and C. Stolle, "Geomagnetic Kp index V. 1.0.," in *GFZ Data Services*. <https://doi.org/10.5880/Kp.0001>, Potsdam, Germany, 2021.
- [29] C. R. Honts and S. Amato, "Automation of a screening polygraph test increases accuracy," in *Psychology, Crime & Law*, Vol. 13, Issue 2, 2007.
- [30] A. Mambreyan, E. Punskeya and H. Gunes, "Dataset Bias in Deception Detection," in *26TH International Conference on Pattern Recognition*, Montréal Québec, 2022.
- [31] M. Abouelenien, V. Pérez-Rosas, R. Mihalcea and M. Burzo, "Detecting Deceptive Behavior via Integration of Discriminative Features from Multiple Modalities," in *IEEE Transactions on Information Forensics and Security*, Vol. 12, Issue 5, 2016.
- [32] Interfax, "Bill on possible ban on transfer abroad of Russians' personal data being submitted to State Duma," in <https://interfax.com/newsroom/top-stories/77833/>, 2022.
- [33] M. Handler and N. Hernandez, "Introduction to the NCCA ASCII Standard," in *Polygraph & Forensic Credibility Assessment: A Journal of Science and Field Practice*, 2019.

ФОРМУЛА ИЗОБРЕТЕНИЯ

1. Компьютерно-реализуемый способ автоматической полиграфической проверки, выполняемый с помощью вычислительной системы, содержащей модель машинного обучения, при этом способ выполняет этапы, на которых

получают записи полиграфических проверок, содержащие, по меньшей мере, сигналы датчиков с временными шкалами, на которых промаркированы начало и конец вопроса;

получают дополнительные данные, содержащие, по меньшей мере, возраст проверяемого, пол, должностную информацию;

осуществляют обработку полученных сигналов и дополнительных данных с помощью модели машинного обучения (МО), причем в ходе указанной обработки осуществляется

определение временных интервалов для извлечения переменных на основе временных меток начала и конца вопроса и временной метки ответа, и на основе типа и темы вопроса;

извлечение переменных из каждого сигнала на определенных временных интервалах;

обработка полученных переменных из сигналов, при которой выполняется нормализация и конкатенация обработанных переменных и построение на их основе вектора;

разделение дополнительных данных на категориальные и численные переменные;

обработка полученных переменных из дополнительных данных, при которой выполняется векторизация категориальных переменных и нормализация численных переменных;

конкатенация обработанных переменных, извлеченных из каждого сигнала, и обработанных дополнительных переменных и построение на их основе вектора;

подача упомянутого вектора в модель МО для получения выходного значения модели МО; сравнение выходного значения модели с заданным пороговым значением и определяют, что ответ является ложью, если выходное значение выше или равно пороговому значению или ответ является правдой если выходное значение ниже порогового значения.

2. Способ по п.1, характеризующийся тем, что модель МО определяет ответ как правдивый или ложный не по одному полученному ответу, а по совокупности вопросов и ответов, связанных с темой полиграфической проверки, при этом получаемые переменные сигналы объединяются в единый вектор либо усредняются перед передачей в модели МО.

3. Способ по п.1, характеризующийся тем, что модель МО имеет тип градиентный бустинг, случайный лес или нейронная сеть.

4. Способ по п.1, характеризующийся тем, что модель МО обучена на одной из тем для проверок или их комбинации, где темами для проверок являются наркотические вещества, получение дополнительного вознаграждения, разглашение конфиденциальной информации, долговые обязательства, сторонний доход, уголовные правонарушения, административные правонарушения, нарушения внутренних нормативных документов (ВНД).

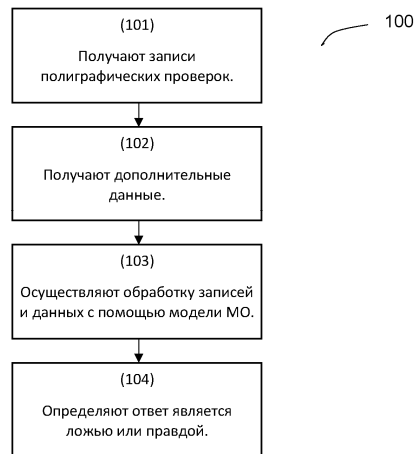
5. Способ по п.1, характеризующийся тем, что записи полиграфических проверок, содержат сигналы с датчиков, включающие по меньшей мере одно из: частота сердечного сокращения (ЧСС), кожно-гальваническая реакция (КГР), артериальное давление, верхнее и нижнее дыхание, пьезоплетизмограмма, фотоплетизмограмма, термических, движения зрачка или их комбинации.

6. Способ по п.1, характеризующийся тем, что дополнительные данные содержат по меньшей мере одно из: идентификационный номер полиграфологов, идентификационный номер полиграфов, информацию о погодных условиях, результаты электроэнцефалограммы, магнитно-резонансной томографии, функциональной ближней инфракрасной спектроскопии, информацию о геомагнитных бурях или их комбинации.

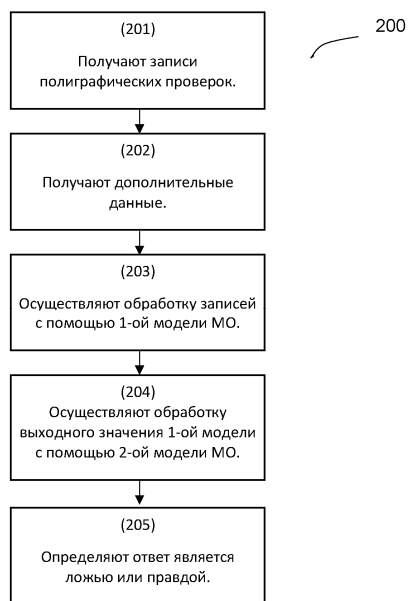
7. Система автоматической полиграфической проверки, содержащая по меньшей мере один процессор;

по меньшей мере одну память, соединенную с процессором, которая содержит машиночитаемые инструкции, которые при их выполнении по меньшей мере одним процессором обеспечивают выполнение способа по любому из пп.1-6.

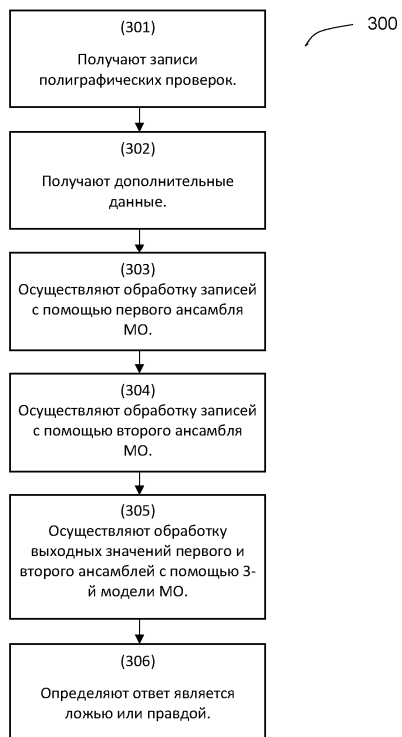
ЧЕРТЕЖИ К ОПИСАНИЮ.



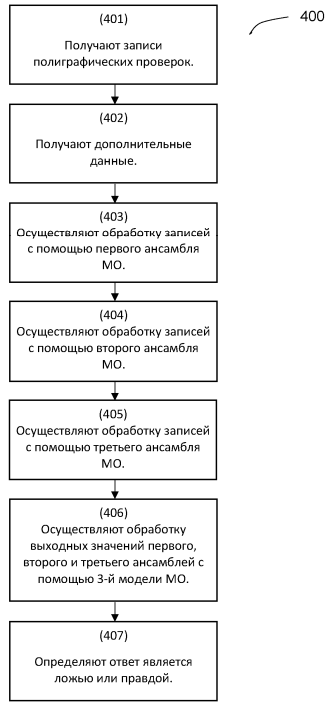
Фиг. 1



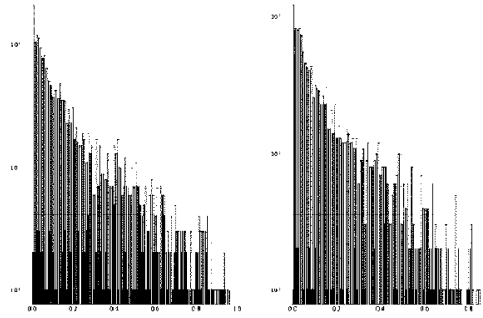
Фиг. 2



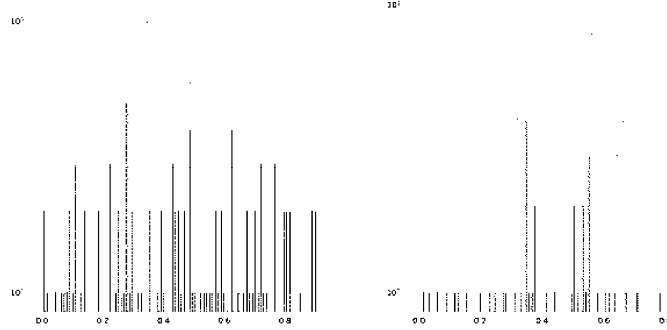
Фиг. 3



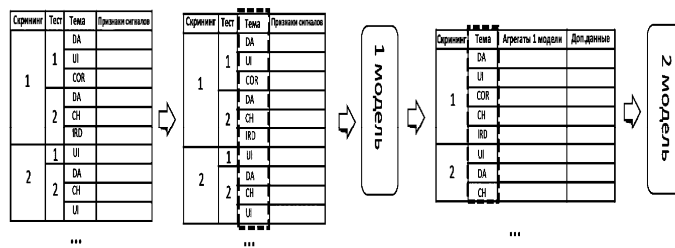
Фиг. 4



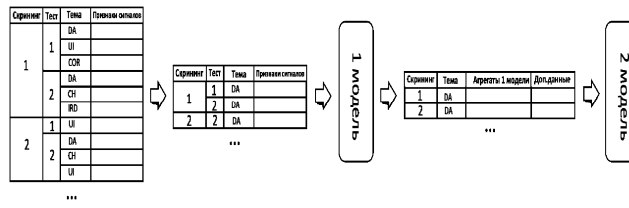
Фиг. 5



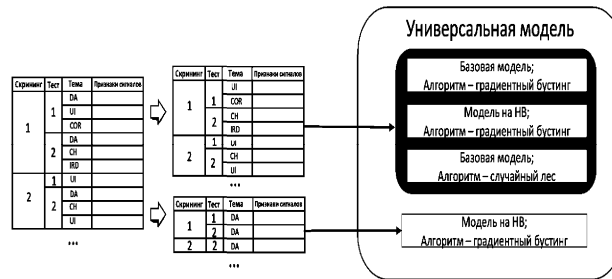
Фиг. 6



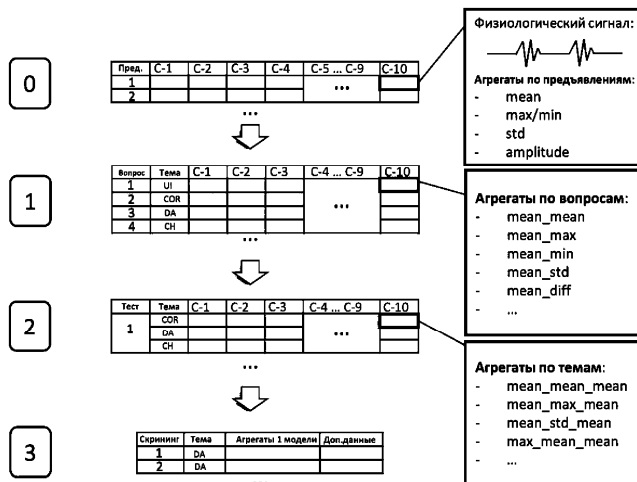
Фиг. 7



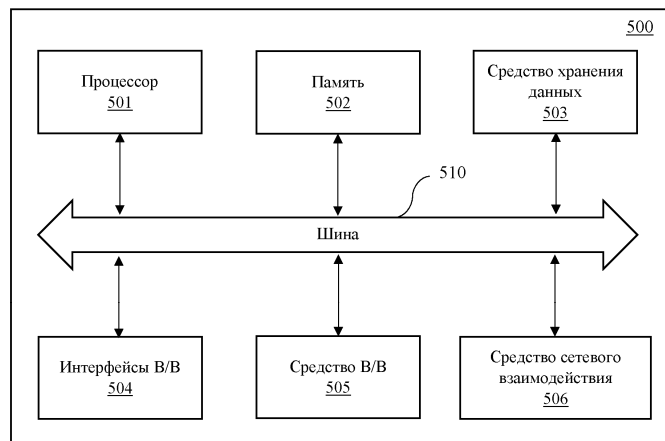
Фиг. 8



Фиг. 9



Фиг. 10



Фиг. 11