

**ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ И ПРИНЯТИЕ РЕШЕНИЙ
ARTIFICIAL INTELLIGENCE AND DECISION MAKING**

УДК 004.8

DOI: 10.18413/2518-1092-2022-8-2-0-5

**Крайновских В.И.¹
Комарова А.А.²
Басов О.О.²****МЕТОД ВЫЯВЛЕНИЯ КОСВЕННЫХ ПРИЗНАКОВ
КОРРУПЦИОННЫХ ДЕЯНИЙ ПО ВИДЕОЗАПИСЯМ
ВЫСТУПЛЕНИЙ ГОССЛУЖАЩИХ**

¹) Федеральное государственное автономное образовательное учреждение высшего образования «Национальный исследовательский университет «Высшая школа экономики», ул. Кантемировская дом 3, корп.1, лит. А, г. Санкт–Петербург, 194100, Россия

²) Федеральное государственное автономное образовательное учреждение высшего образования «Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики», Кронверкский пр., д. 49, г. Санкт-Петербург, 197101, Россия

e-mail: vikraynova@edu.hse.ru

Аннотация

На данный момент проблема противодействию коррупционным действиям в сфере государственной службы по–прежнему не теряет свою актуальность. Предполагается, что при совершении коррупционных деяний люди проявляют определенные вербальные и невербальные сигналы, с помощью которых можно выявить коррупционные признаки. В статье рассматривается проблема нарушения антикоррупционного законодательства в государственных и муниципальных учреждениях с точки зрения психоэмоционального состояния чиновников. Предлагается использовать методы машинного обучения для анализа видео и аудио записей выступлений с неподготовленной речью чиновников различного уровня власти, где они отвечают на вопросы журналистов и общественности, с целью определения эмоций и выявления агрессии, неуверенности, а также уклончивости в ответах. Результаты исследования могут быть полезны для государственных органов, занимающихся борьбой с коррупцией, а также для общественности, заинтересованной в прозрачности и честности деятельности государственных служащих.

Ключевые слова: противодействие коррупции; государственное управление; машинное обучение; вербальный анализ; анализ аудио-видеозаписей; анализ психоэмоционального состояния

Для цитирования: Крайновских В.И., Комарова А.А., Басов О.О. Метод выявления косвенных признаков коррупционных деяний по видеозаписям выступлений госслужащих // Научный результат. Информационные технологии. – Т.8, №2, 2023. – С. 35-45. DOI: 10.18413/2518-1092-2022-8-2-0-5

**Krainovskikh V.I.¹
Komarova A.A.²
Basov O.O.²****THE METHOD OF IDENTIFYING INDIRECT SIGNS
OF CORRUPTION ACTS BASED ON VIDEO RECORDINGS
OF SPEECHES OF CIVIL SERVANTS**

¹) Federal State Autonomous Educational Institution of Higher Education "National Research University "Higher School of Economics", 3 Kantemirovskaya str., building 1, lit. A, St. Petersburg, 194100, Russia

²) Saint Petersburg National Research University of Information Technologies, Mechanics and Optics, 49 Kronverkskiy prospekt, St. Petersburg, 197101, Russia

e-mail: vikraynova@edu.hse.ru

Abstract

At the moment, the problem of countering corruption in the field of public service still does not lose its relevance. It is assumed that when committing acts of corruption, people show certain verbal and non-verbal signals, with the help of which it is possible to identify signs of corruption.

The article deals with the problem of violation of anti-corruption legislation in state and municipal institutions from the point of view of the psycho-emotional state of officials. It is proposed to use machine learning methods to analyze video and audio recordings of speeches with unprepared speech of officials of various levels of government, where they answer questions from journalists and the public, in order to determine emotions and identify aggression, uncertainty, and evasiveness in answers. The results of the study may be useful for government agencies involved in the fight against corruption, as well as for the public interested in transparency and honesty of the activities of civil servants.

Keywords: anti-corruption; public administration; machine learning; verbal analysis; analysis of audio-video recordings; analysis of psycho-emotional state

For citation: Krainovskikh V.I., Komarova A.A., Basov O.O. The method of identifying indirect signs of corruption acts based on video recordings of speeches of civil servants // Research result. Information technologies. – Т.8, №2, 2023. – P. 35–45. DOI: 10.18413/2518-1092-2022-8-2-0-5

ВВЕДЕНИЕ

Проблема коррупции остается одной из самых актуальных в мире, и Россия не является исключением. По данным МВД России, только за первые шесть месяцев 2022 года число коррупционных преступлений в России увеличилось на 9,2% по сравнению с предыдущим годом [25]. К таким нарушениям относятся получение взяток, злоупотребление должностным положением и другие незаконные способы обогащения государственных служащих.

Коррупция является неотъемлемой частью системы теневой экономики, которая уменьшает доходы государства. В результате многие экономические транзакции происходят вне официальной экономики и не облагаются налогами. Международный Валютный Фонд оценивает потери мировой экономики от коррупции от 1,5 трлн. до 2 трлн. долларов ежегодно [24]. Кроме того, коррупционные последствия могут быть катастрофическими для экономического и социального развития страны. Они искажают нормы морали и подрывают авторитет государственной власти в глазах граждан. Однако, несмотря на все эти негативные последствия, коррупция продолжает существовать во многих странах, и борьба с ней остается одной из важнейших задач в области государственного управления.

Хоть данный феномен и является трудно наблюдаемым, однако сегодня все больше стран стараются пресечь коррупционные деяния путем внедрения современных информационных технологий. Россия не стала исключением и также рассматривает искусственный интеллект как вариант профилактики коррупционных действий [26]. На сегодняшний день применяются информационно-коммуникационные технологии, связанные с формированием транспарентности органов публичной власти. Такими примерами могут служить электронные порталы, на которых государственные служащие могут обнародовать свои доходы и имущество, онлайн-реестры государственных закупок, системы электронного документооборота и даже приложения для мобильных устройств, позволяющие гражданам сообщать о коррупционных проявлениях. Однако, применение технологий не является панацеей от коррупции и не гарантирует полное ее исключение.

Существует связь между определенными языковыми, поведенческими и невербальными проявлениями на аудио-видеозаписях и совершением коррупционных деяний. Например, наличие специфических лингвистических выражений, сигналов и жестов, которые могут указывать на протекцию, взяточничество, подкуп или другие формы коррупции. Данная гипотеза предполагает, что во время совершения коррупционных деяний люди могут проявлять определенное поведение, использовать специфическую лексику, жестикуляцию или другие невербальные средства, которые могут быть записаны на аудио-видеозаписях. Анализ этих записей может помочь выявить определенные признаки, которые указывают на возможное совершение коррупционных деяний.

Анализ видео и аудио записей с выступлениями чиновников, в которых они отвечают на вопросы журналистов и общественности, может стать эффективным инструментом в борьбе с коррупцией. Такой анализ ответов помогает определять степень уклончивости, неуверенности и агрессии государственных служащих, что может навести на следы коррупционных деяний. Связь

этих характеристик с реальными нарушениями антикоррупционного законодательства может дать более полное представление об уровне коррупции в государственных и муниципальных учреждениях и помочь в разработке эффективных стратегий борьбы с данной проблемой, и в дальнейшем может обеспечить прозрачность и честность в работе государственных и муниципальных учреждений в будущем.

Данная работа направлена на разработку метода выявления признаков нарушения антикоррупционного законодательства в видеозаписях выступлений госслужащих, основанного на определении таких косвенных признаков как неуверенность, наличие агрессии и уклончивость в ответах госслужащих при анализе уровня коррупции в государственных и муниципальных учреждениях.

ОСНОВНАЯ ЧАСТЬ

В рамках данного исследования была принята гипотеза о наличии взаимосвязи между тремя психоэмоциональными характеристиками, которые может проявлять человек в случаях нарушения антикоррупционного законодательства: агрессия, неуверенность и уклончивость в ответах. Был проведен сбор данных, состоящий из видео выступлений чиновников разного уровня власти, для определения поведенческих маркеров трех психоэмоциональных характеристик.

Методика выявления признаков нарушения антикоррупционного законодательства в видеозаписях чиновников состоит из трех модулей: детекции агрессии, неуверенности и уклончивости в ответах, и основана на анализе видеоматериалов с неподготовленной речью чиновника. На вход подается видеоматериал, из которого извлекаются видео и аудио каналы. Аудиоканал подвергается процессу транскрибации, для получения текстового значения выступления. После чего необходимые потоки данных подаются на вход трем модулям для дальнейшего анализа. Структура разрабатываемого метода представлена на Рисунке 1:

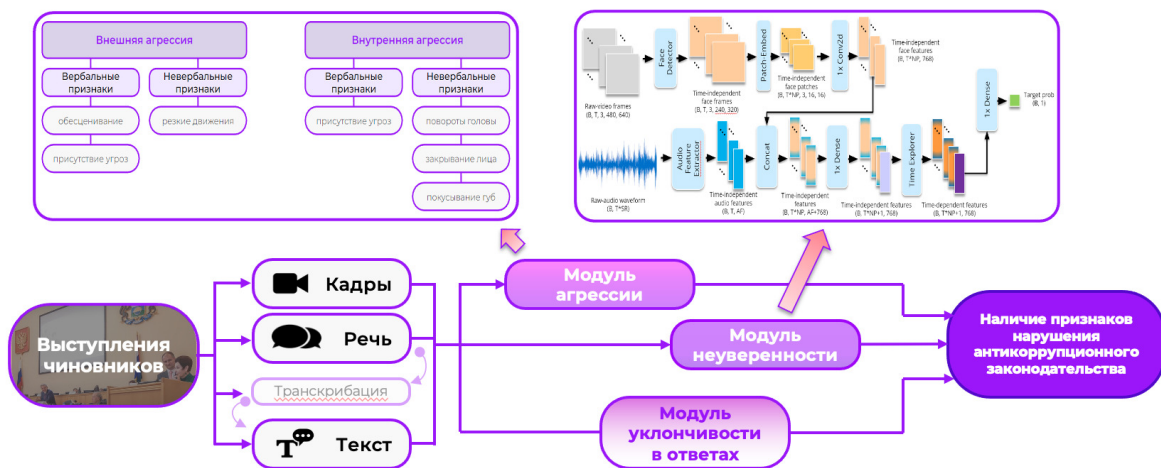


Рис. 1. Структура разрабатываемого метода выявления косвенных признаков нарушения антикоррупционного законодательства в видеозаписях выступлений госслужащих

Fig. 1. Structure of the developed method for identifying indirect signs of violation of anti-corruption legislation in video recordings of speeches of civil servants

Выборка данных состоит из различных видео материалов с неподготовленной речью чиновников различного уровня, где они отвечают на различные вопросы: записи встреч и переговоров с представителями бизнеса, СМИ и общественностью, а также записи ответов на вопросы, связанных с их должностью, доходом и коррупционными действиями. Схема с классификацией используемых ресурсов с неподготовленной речью чиновников представлена на рисунке 2:



Рис. 2. Схема с используемыми ресурсами неподготовленной речи чиновников
Fig. 2. A diagram with the resources used for the unprepared speech of officials

На Рисунке 3 представлена классификация вопросов, которые используются в дальнейшем анализе видеофрагментов из выступлений госслужащих:

Характеристики вопросов	
Категория вопросов	Описание
Финансовые вопросы	Относятся к финансовым аспектам деятельности и могут выявлять наличие взяточничества, раскредитования, отмывания денег и других финансовых махинаций
Контакты с лицами, связанными с коррупцией	Относятся к общению и взаимодействию с лицами, связанными с коррупцией, такими как чиновники, представители частных компаний и т.д.
Использование влияния и ресурсов	Относятся к использованию личного или служебного влияния и ресурсов в корыстных целях, в том числе использование служебного положения или ресурсов для получения личной выгоды
Запросы на взятки и неправомерные выгоды	Относятся к прямому запросу на получение денежных средств, подарков или других неправомерных выгод в обмен на предоставление услуг
...	...

Рис. 3. Рассматриваемые категории вопросов, задающиеся чиновникам
Fig. 3. Considered categories of questions asked to officials

Для дальнейшего анализа требуется, чтобы каждый фрагмент записей описывался набором невербальных и вербальных маркеров поведения. Ниже представлены методы, использующиеся для извлечения указанных маркеров.

ДЕТЕКЦИЯ АГРЕССИИ

Данный модуль позволяет определить степень агрессии говорящего. Сперва уточним, что подразумевается под агрессией: агрессия – это поведение, целью которого является причинение вреда объекту нападения (будь то живое существо или неодушевленный предмет), оно несовместимо с общепринятыми нормами и правилами поведения в обществе; проявлением такого поведения может быть как физический ущерб, так и вызванные негативные эмоции или психологический дискомфорт у объекта нападения (чувство страха, беспокойства, тревоги, депрессии и т. д.) [2].

Агрессию лучше всего рассматривать с точки зрения двух характеристик: внутреннего и внешнего проявления. Для определения агрессии следует учитывать вербальные и невербальные сигналы, они разделяются следующим образом:

Невербальные сигналы:

- темп речи;
- резкое повышение/понижение громкости речи;
- эмоциональный окрас речи: злость, отвращение, спокойствие, радость, страх, удивление, печаль.

Вербальные сигналы:

- наличие повелительного наклонения;
- наличие нецензурной лексики и жаргона в тексте;
- наличие уменьшительно-ласкательных\уменьшительно-унижительных конструкций;
- наличие оскорблений;
- наличие угроз;
- наличие токсичности.

Токсичность в высказываниях описывается как качество высказывания, которое проявляется в язвительных, агрессивных, оскорбительных, унижающих или провоцирующих высказываниях, которые могут вызвать негативные эмоции, оскорбить, унижить или причинить вред другому человеку или группе людей.

Набор данных содержал записи с выступлений чиновников, на которых присутствовали посторонние шумы, разговоры других людей и т.д. Поэтому перед анализом данных было необходимо провести предварительную обработку, включающую удаление шумов и нормализацию громкости [14].

Темп речи измеряется в количестве произнесенных слов в минуту. Для вычисления этого показателя используется спектральный поток сигнала, который вычисляется с помощью функций из библиотеки «Librosa» [15]. Затем производится подсчет количества пиковых значений, которые считаются началом каждого слова.

Для определения "резкого" изменения громкости применяется эвристический метод, заключающийся в измерении угла наклона между соседними отсчетами амплитудной огибающей и абсолютным расстоянием до максимального значения. Если амплитудное значение отсчета близко к максимальному значению по всей записи и изменение происходит быстрее, чем переход к другим значениям внутри записи, то это изменение громкости считается "резким".

Для определения класса эмоций была обучена сверточная нейронная сеть с полносвязным классификатором, аудио–сигнал в которой представлен в виде мел-кепстральных коэффициентов, подающиеся на вход сети [8]. Для того, чтобы в ходе обучения не возникло ситуации переобучения на одном языке [11], были выбраны следующие наборы данных: CaFE [9], Emo-DB [5], RAVDESS [13], SAVEE [17]. В полученный набор были включены короткие аудиозаписи, отнесенные к одному из семи классов эмоций, которые соотносятся с классификацией эмоций Р. Вудвортста [1]: злость, скука, тревога, радость, печаль, отвращение и нейтральное состояние (отсутствие эмоции). Для классификации эмоций используется фрагмент длительностью 3 секунды, а в случае, если он имеет меньшую длительность, его заполняют нулями до необходимого размера.

Для оценки вербальных признаков, подающихся из текста, необходимо было провести процедуру транскрипции речи. С этой целью, был выбран набор инструментов Vosk.

Был собран словарь из 140 тыс. слов, состоящий из максимально всевозможных морфологических вариантов слов нецензурной лексики, с целью ее детектирования из речи. Каждое слово входящего текста проверяется по данному словарю и подсчитывается количество нецензурной лексики в тексте.

Для детекции слов, находящихся в повелительном наклонении, была использована библиотека rufmorphu2, с помощью которой возможен полный морфологический разбор слова [10].

Для детекции уменьшительно-ласкательных и пренебрежительных выражений, окончания слов были отброшены с использованием стеммера (процесс нахождения основы слова) Портера, адаптированного для русского языка. Далее была проверка на наличие уменьшительно-

ласкательных суффиксов у слова, включая все возможные морфологические варианты, такие как "ик", "ек", "к", "ец", "иц", "оск", "ечк", "оньк", "еньк", "ышк", "инш", "ушк" и "юшк".

Обнаружение оскорблений и угроз в тексте было выполнено дообучение моделей бинарной классификации [3] с помощью нейронной сети, на базе архитектуры RuBertTiny [20].

Для детекции угроз:

- тесты, содержащие угрозу;
- тексты, не содержащие угроз.

Для детекции оскорблений:

- тексты, содержащие оскорбления;
- тексты, не содержащие оскорбления.

С помощью данных из социальной сети «Одноклассники» были взяты 248290 комментария, относящиеся к трем категориям: нейтральные сообщения, сообщения с оскорблениями и сообщения с угрозами.

Также была обучена модель RuBertTiny для детекции токсичности в речи [17]. Текст, содержащий дискриминирующие и оскорбительные сообщения по отношению к кому-либо, принято считать токсичным. Была обучена модель классификации на нескольких различных наборах данных [7].

ДЕТЕКЦИЯ НЕУВЕРЕННОСТИ

Данный модуль предназначен для того, чтобы понять степень уверенности говорящего. Уверенность формируется на основе успешного решения социальных задач и достижения поставленных целей.

Анализ данной характеристики будет проходить по двум сигналам: вербальным и невербальным. В каждом из них были выделены следующие показатели:

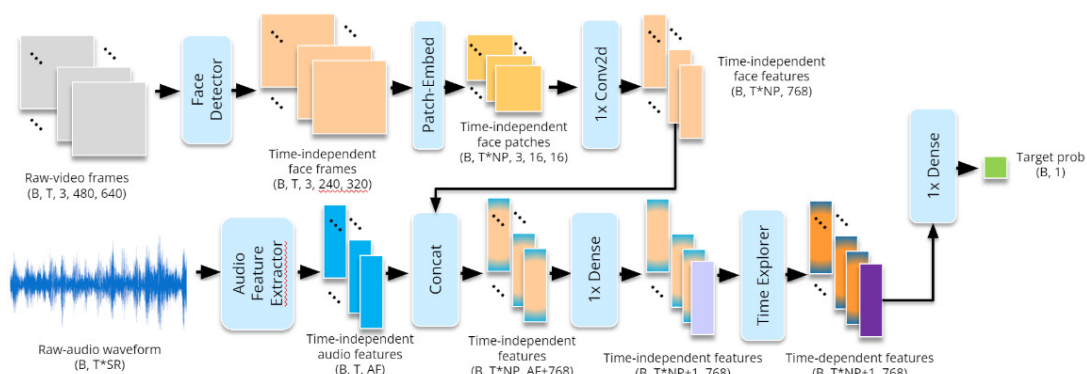
Невербальные сигналы:

- темп речи;
- резкое повышение/понижение громкости речи;
- эмоциональный окрас речи: радость, злость, нейтральное состояние.

Вербальные сигналы:

- содержание речи

Модуль работает с входными видеозаписями чиновников, в ходе чего из них извлекаются последовательности из кадров и сэмплов длительностью 10 секунд для дальнейшего анализа. В работе модуля используется нейронная сеть–трансформер (Рисунок 4) [19], с целью обработки мультимодального потока видео [22, 21, 23].



*Рис. 4. Пайплайн определения уверенности человека с помощью сети-трансформер
Fig. 4. Pipeline for determining a person's confidence using a transformer network*

Существует два способа, чтобы произвести извлечение необходимых признаков как из видеоматериалов, так и из аудиосигналов. Рассмотрим каждый более подробно, начнем с видеосигналов:

1) Для получения вектора мимики лица рассматриваемого человека используется сверточная нейронная сеть BlazeFace вместе с модулем Face Mesh [4], который извлекает вектор из 478 точек лица. Сеть принимает на вход изображение кадра видеопоследовательности в тензорном представлении, сжатом до 256x256 пикселей.

2) В альтернативном сценарии вместо точек лица используются более высокоуровневые признаки, полученные непосредственно из изображения лица человека реального размера, которое в дальнейшем векторизуется с помощью сверточной нейронной сети TinaFace.

Теперь рассмотрим существующие сценарии для извлечения признаков из аудиосигнала:

1) Извлечение интерпретируемых атрибутов из аудиопотока, таких как эмоции, дрожание голоса, тон, частота, амплитуда, количество и длительность пауз.

2) Преобразование аудиопотока в спектрограмму, по которой вычисляются MFCC коэффициенты [16, 6].

Следующим этапом является объединение аудио и видео сигналов в один поток мультимодальных признаков, основываясь на соответствующей их размерности [12]. Затем, к получившемуся потоку добавляется обучаемый вектор, и они вместе передаются в сеть-трансформер для обучения относительно временной размерности.

Модель содержит два нейрона с функцией активации сигмоид, которые предоставляют информацию о видеофрагменте:

1) Первый нейрон: происходит вычисление вероятности отнесения видеофрагмента к классу 0, что является оценкой уверенности в данной классификации;

2) Второй нейрон: происходит вычисление вероятности отнесения видеофрагмента к классу 1, что является оценкой неуверенности в данной классификации.

Согласно формулам, которые изображены ниже, происходит объединение результатов предсказаний для видеозаписей:

$$T = \frac{1}{n} * \sum_{i=0}^1 \left(\frac{T_i}{T_i + F_i} \right),$$

$$F = \frac{1}{n} * \sum_{i=0}^1 \left(\frac{F_i}{F_i + T_i} \right).$$

где T - значение уверенности первого нейрона (уверенность в высказывании);

F - значение уверенности второго нейрона (неуверенность в высказывании);

n - количество видеофрагментов в видео.

Эта формула позволяет оценить вклад каждого нейрона в предсказание модели, вместо того чтобы рассчитывать независимую оценку уверенности каждого нейрона:

- Значение в пределах 50% соответствует низкой степени уверенности эксперта.
- Значения от 50% до 70% соответствуют средней степени уверенности эксперта.
- Значения, превышающие 70%, соответствует более высокой степени уверенности.

ДЕТЕКЦИЯ УКЛОНЧИВОСТИ В ОТВЕТАХ

Уклончивость в ответах означает избегание прямых, ясных и конкретных ответов на вопросы или ведение обсуждения в сторону или избегание точного ответа на вопросы. Это может проявляться в использовании общих или неопределенных формулировок, уходе от прямых ответов, в различных альтернативах или отвлеченных рассуждениях, вмешательстве в детали или недостаточном предоставлении информации. Анализ научной литературы по данной теме позволило выделить и классифицировать наиболее распространенные способы избегания ответов.

Проанализировав ряд источников, была разработана схема на рисунке 5, в которой представлены различные способы уклончивости в ответах. В данной схеме визуализированы типичные приемы, которыми спикеры могут избегать прямых и конкретных ответов на вопросы. Классификация способов уклончивости включает такие категории, как ответ вопросом на вопрос, сомнение в задаваемом вопросе, отказ от ответа или игнорирование вопроса, и неполный ответ.



Рис. 5. Схема со способами избегания ответа на вопрос
Fig. 5. A diagram with ways to avoid answering the question

Разрабатываемый алгоритм выявления уклончивости в ответах представляет собой сложную систему, основанную на двух обученных моделях, представляет собой обобщенный подход, который может использоваться для анализа диалогов или речи спикера.

Первая модель, использующая трансформер BERT, выполняет классификацию вопросительных и утвердительных предложений [27]. Для дообучения этой модели был использован датасет, состоящий из около 81 тысячи утвердительных предложений из датасета SPAADIA и 131 тысячи вопросительных предложений из датасета SQuAD. В результате тестирования, данная модель показала высокое значение F-меры, составляющей 0.96 для вопросительных предложений, что в данном случае свидетельствует о ее эффективности.

Далее, ответ человека в качестве контекста передается во вторую модель, которая в свою очередь с определенной степенью уверенности выделяет этот ответ из речи. Эта модель была построена с использованием техники fine-tuning на основе трансформера RoBERTa. Данная архитектура RoBERTa аналогична архитектуре BERT, однако отличается подходом к обучению модели. RoBERTa использует увеличенный корпус для обучения и генерацию динамических mask-токенов, что способствует ее более точному обучению. В качестве датасета для дообучения данной модели был выбран SQuAD датасет, содержащий вопросы и ответы на них. Значение F1-меры для дообученной модели составило 0.83, что указывает на ее высокую производительность и способность выделять ответы на вопросы из контекста речи спикера.

Можно предположить, что при слишком высокой уверенности модели ее ответы будут чрезмерно категоричными и необходимо будет учитывать возможность ложных срабатываний. С другой стороны, при слишком низкой уверенности модели ответы будут неопределенными и непригодными для использования. Поэтому для достижения наилучших результатов важно подобрать оптимальную границу значения уверенности, при которой модель будет давать наиболее точные и пригодные для использования ответы. Эта граница была определена экспериментально, путем анализа результатов работы модели при разных значениях уверенности и выбора наилучшего значения на основе полученных данных.

Выходом разрабатываемого алгоритма являются списки с вопросами/ответами для каждого спикера, сопровождающиеся метками "уклончивый" или "не уклончивый", а также значениями уверенности модели. Для каждого вопроса/ответа алгоритм определяет, является ли он уклончивым или нет, на основе выводов первой модели, осуществляющей классификацию предложений. Значение уверенности модели также возвращается в виде числовой оценки, отражающей степень уверенности алгоритма в своих выводах.

Таким образом, данная обобщенная модель, состоящая из двух шагов и двух обученных моделей, позволяет эффективно анализировать диалоги или речь спикера, классифицировать предложения и выделять ответы на вопросы из контекста, открывая новые возможности для автоматического анализа и обработки естественного языка.

ЗАКЛЮЧЕНИЕ

В исследовании был предложен метод выявления косвенных признаков нарушения антикоррупционного законодательства на основе анализа видеозаписей выступлений чиновников. Данный метод предполагает использование единого использования трех модулей: детекции

агрессии, неуверенности и уклончивости при ответах, для автоматического анализа видеоматериалов и выявления потенциальных нарушений антикоррупционных норм.

Разработанный метод основывается на анализе различных признаков, таких как жесты, мимика, тон голоса, паузы и другие невербальные элементы выступлений. Эти признаки могут свидетельствовать о потенциальных нарушениях антикоррупционных законов, таких как неправдивые заявления, скрытые сделки, взяточничество и другие формы коррупционной деятельности.

Комбинируя результаты этих трех модулей, можно получить информацию о возможных нарушениях антикоррупционных норм и оценить степень риска. Такой подход позволяет существенно ускорить процесс обнаружения потенциальных нарушений и улучшить эффективность мониторинга коррупционных проявлений в государственных органах.

Однако стоит отметить, что разработанный метод является предварительным инструментом и не может служить основой для однозначного заключения о наличии или отсутствии нарушений. Для более точной оценки необходимо проводить дополнительные исследования, включая анализ сторонних источников информации и проверку результатов метода на большом объеме видеоматериалов с различными контекстами и условиями.

Список литературы

1. Овсянникова В. В. К вопросу о классификации эмоций: категориальный и многомерный подходы // Финансовая аналитика: проблемы и решения. – 2013. – №. 37. – С. 43-48.
2. Рогов Е. И. Настольная книга практического психолога: Учебное пособие // М.: ВЛАДОС. – 1998. – С. 134-142.
3. Самигулин Т. Р., Смирнов И. З., Лаушкина А. А. Определение маркеров агрессивного поведения человека на основе анализа аудио и текстового каналов // Научный результат. Информационные технологии. – 2022. – Т. 7. – №. 2. – С. 55-62.
4. Bazarevsky V. et al. Blazeface: Sub-millisecond neural face detection on mobile gpus //arXiv preprint arXiv:1907.05047. – 2019.
5. Burkhardt F. et al. A database of German emotional speech // Interspeech. – 2005. – Т. 5. – С. 1517-1520.
6. Chow A., Louie J. Detecting lies via speech patterns. – 2017.
7. Devyatkin D. A. et al. Intelligent analysis of manifestations of verbal aggressiveness in network community texts //Scientific and Technical Information Processing. – 2014. – Т. 41. – С. 377-389.
8. Goupil L. et al. Listeners' perceptions of the certainty and honesty of a speaker are associated with a common prosodic signature // Nature communication. – 2021. – Т. 12. – №. 1. – С. 861.
9. Gournay P., Lahaie O., Lefebvre R. A canadian french emotional speech dataset //Proceedings of the 9th ACM multimedia systems conference. – 2018. – С. 399-402.
10. Korobov M. Morphological analyzer and generator for Russian and Ukrainian languages //Analysis of Images, Social Networks and Texts: 4th International Conference, AIST 2015, Yekaterinburg, Russia, April 9–11, 2015, Revised Selected Papers 4. – Springer International Publishing, 2015. – С. 320-332.
11. Kossaiji J. et al. Sewa db: A rich database for audio-visual emotion and sentiment research in the wild // IEEE transactions on pattern analysis and machine intelligence. – 2019. – Т. 43. – №. 3. – С. 1022-1040.
12. Laushkina, Anastasia & Smirnov, Ivan & Medvedev, Anatoly & Laptev, Andrey & Sinko, Mikhail. (2022). Detecting incongruity in the expression of emotions in short videos based on a multimodal approach. Cybernetics and Physics. 210-216.
13. Luna-Jiménez C. et al. A Proposal for Multimodal Emotion Recognition Using Aural Transformers and Action Units on RAVDESS Dataset // Applied Sciences. 2021. – Vol. 12. – № 1. – P. 327.
14. Marcolla F., de Santiago R., Dazzi R. Novel Lie Speech Classification by using Voice Stress // Proceedings of the 12th International Conference on Agents and Artificial Intelligence. SCITEPRESS – Science and Technology Publications. – 2020. – С. 742–749.
15. McFee B. et al. librosa: Audio and music signal analysis in python //Proceedings of the 14th python in science conference. – 2015. – Т. 8. – С. 18-25.
16. Oviatt S. et al. (ed.). The Handbook of Multimodal-Multisensor Interfaces: Signal Processing, Architectures, and Detection of Emotion and Cognition-Volume 2. – Association for Computing Machinery and Morgan & Claypool, 2018.

17. S. Haq P.J.B.J. Multimodal Emotion Recognition / ed. Wang W. IGI Global, 2010. С. 398–423.
18. Saeed H. H., Shahzad K., Kamiran F. Overlapping toxic sentiment classification using deep neural architectures //2018 IEEE international conference on data mining workshops (ICDMW). – IEEE, 2018. – С. 1361-1366.
19. Sinko M. et al. Method of constructing and identifying predictive models of human behavior based on information models of non-verbal signals //Procedia Computer Science. – 2022. – Т. 212. – С. 171-180.
20. Tenney I., Das D., Pavlick E. BERT rediscovers the classical NLP pipeline // arXiv preprint arXiv:1905.05950. – 2019.
21. Tsai Y. H. H. et al. Learning factorized multimodal representations // arXiv preprint arXiv:1806.06176. – 2018.
22. Tsai Y. H. H. et al. Multimodal transformer for unaligned multimodal language sequences //Proceedings of the conference. Association for Computational Linguistics. Meeting. – NIH Public Access, 2019. – Т. 2019. – С. 6558.
23. Tzirakis P. et al. End-to-end multimodal emotion recognition using deep neural networks //IEEE Journal of selected topics in signal processing. – 2017. – Т. 11. – №. 8. – С. 1301-1309.
24. В МВФ оценили потери мировой экономике от коррупции [Электронный ресурс]. – Режим доступа: <https://www.rbc.ru/economics/18/09/2017/59bfead89a794704d063c4f0> (дата обращения: 03.04.2023).
25. Краткая характеристика состояния преступности в Российской Федерации за январь-июнь 2022 года. URL: <https://xn--b1aew.xn--p1ai/reports/item/31209853/> (дата обращения: 03.04.2023).
26. НИУ ВШЭ предложил использовать искусственный интеллект для предотвращения коррупции в РФ [Электронный ресурс]. – Режим доступа: <https://tass.ru/ekonomika/14304599> (дата обращения: 03.04.2023).
27. Bert for Sequence Classification (Question vs Statement)_ [Электронный ресурс]. – Режим доступа: https://sparknlp.org/2021/11/04/bert_sequence_classifier_question_statement_en.html (дата обращения: 15.04.2023).

References

1. Ovsyannikova V.V. On the question of the classification of emotions: categorical and multidimensional approaches // Financial analytics: problems and solutions. – 2013. – No. 37. – P. 43-48.
2. Rogov E.I. The handbook of a practical psychologist: a textbook //Moscow: VL. – 1998. – P. 134-142.
3. Samigullin T.R., Smirnov I.Z., Laushkina A.A. Determination of markers of aggressive human behavior based on the analysis of audio and text channels //Scientific result. Information technology. – 2022. – Vol. 7. – No. 2. – P. 55-62.
4. Bazarevsky V. et al. Blazeface: Sub-millisecond neural face detection on mobile gpus //arXiv preprint arXiv:1907.05047. – 2019.
5. Burkhardt F. et al. A database of German emotional speech // Interspeech. – 2005. – Т. 5. – P. 1517-1520.
6. Chow A., Louie J. Detecting lies via speech patterns. – 2017.
7. Devyatkin D. A. et al. Intelligent analysis of manifestations of verbal aggressiveness in network community texts //Scientific and Technical Information Processing. – 2014. – Т. 41. – P. 377-389.
8. Goupil L. et al. Listeners' perceptions of the certainty and honesty of a speaker are associated with a common prosodic signature // Nature communication. – 2021. – Т. 12. – №. 1. – P. 861.
9. Gournay P., Lahaie O., Lefebvre R. A canadian french emotional speech dataset // Proceedings of the 9th ACM multimedia systems conference. – 2018. – P. 399-402.
10. Korobov M. Morphological analyzer and generator for Russian and Ukrainian languages // Analysis of Images, Social Networks and Texts: 4th International Conference, AIST 2015, Yekaterinburg, Russia, April 9–11, 2015, Revised Selected Papers 4. – Springer International Publishing, 2015. – P. 320-332.
11. Kossaiji J. et al. Sewa db: A rich database for audio-visual emotion and sentiment research in the wild // IEEE transactions on pattern analysis and machine intelligence. – 2019. – Т. 43. – №. 3. – P. 1022-1040.
12. Laushkina, Anastasia & Smirnov, Ivan & Medvedev, Anatoly & Laptev, Andrey & Sinko, Mikhail. (2022). Detecting incongruity in the expression of emotions in short videos based on a multimodal approach. Cybernetics and Physics. P. 210-216.
13. Luna-Jiménez C. et al. A Proposal for Multimodal Emotion Recognition Using Aural Transformers and Action Units on RAVDESS Dataset // Applied Sciences. 2021. – Vol. 12. – № 1. – P. 327.
14. Marcolla F., de Santiago R., Dazzi R. Novel Lie Speech Classification by using Voice Stress // Proceedings of the 12th International Conference on Agents and Artificial Intelligence. SCITEPRESS – Science and Technology Publications. – 2020. – P. 742–749.

15. McFee B. et al. librosa: Audio and music signal analysis in python // Proceedings of the 14th python in science conference. – 2015. – Т. 8. – P. 18-25.
16. Oviatt S. et al. (ed.). The Handbook of Multimodal-Multisensor Interfaces: Signal Processing, Architectures, and Detection of Emotion and Cognition-Volume 2. – Association for Computing Machinery and Morgan & Claypool, 2018.
17. S. Haq P.J.B.J. Multimodal Emotion Recognition / ed. Wang W. IGI Global, 2010. P. 398–423.
18. Saeed H. H., Shahzad K., Kamiran F. Overlapping toxic sentiment classification using deep neural architectures // 2018 IEEE international conference on data mining workshops (ICDMW). – IEEE, 2018. – P. 1361-1366.
19. Sinko M. et al. Method of constructing and identifying predictive models of human behavior based on information models of non-verbal signals // Procedia Computer Science. – 2022. – Т. 212. – P. 171-180.
20. Tenney I., Das D., Pavlick E. BERT rediscovers the classical NLP pipeline // arXiv preprint arXiv:1905.05950. – 2019.
21. Tsai Y. H. H. et al. Learning factorized multimodal representations // arXiv preprint arXiv:1806.06176. – 2018.
22. Tsai Y. H. H. et al. Multimodal transformer for unaligned multimodal language sequences // Proceedings of the conference. Association for Computational Linguistics. Meeting. – NIH Public Access, 2019. – Т. 2019. – P. 6558.
23. Tzirakis P. et al. End-to-end multimodal emotion recognition using deep neural networks // IEEE Journal of selected topics in signal processing. – 2017. – Т. 11. – №. 8. – P. 1301-1309.
24. The IMF estimated the losses to the world economy from corruption [Electronic resource]. – URL: <https://www.rbc.ru/economics/18/09/2017/59bfead89a794704d063c4f0> (date of application: 03.04.2023).
25. Brief description of the state of crime in the Russian Federation for January-June 2022. URL: <https://xn--b1aew.xn--p1ai/reports/item/31209853> / (accessed 03.04.2023).
26. HSE has proposed using artificial intelligence to prevent corruption in the Russian Federation [Electronic resource]. – URL: <https://tass.ru/ekonomika/14304599> (date of application: 03.04.2023).
27. Bert for Sequence Classification (Question vs Statement) [Electronic resource]. – URL: https://sparknlp.org/2021/11/04/bert_sequence_classifier_question_statement_en.html (accessed: 04/15/2023).

Крайновских Вероника Игоревна, студент 4–го курса бакалавриата

Комарова Алёна Алексеевна, магистрант факультета цифровой трансформации

Басов Олег Олегович, доктор технических наук, доцент, профессор факультета цифровой трансформации

Krainovskikh Veronika Igorevna, 4th year Bachelor's student

Komarova Alyona Alekseevna, Master's student of the Faculty of Digital Transformation

Oleg Basov Olegovich, Doctor of Technical Sciences, Associate Professor, Professor of the Faculty of Digital Transformation