План работы по написанию статьи

**«Использование косинусной меры близости векторов признаков (Embedding vector) в задачах идентификации диктора при проведении фоноскопической экспертизы»**

1. Введение – на 0.5 - 2 страницах объяснить, почему важно иметь надежные инструменты идентификации по голосу. Показать, что в судебных заседаниях ведется запись под протокол, и потом верифицированный голос используется в качестве образца для сравнения с голосом в записи телефонных переговоров или в записи на диктофоне. Сформулировать задачу, решаемую в статье – какая минимальная длительность фрагмента необходима для проведения автоматизированного анализа с целью идентификации диктора? Насколько достоверными будут является полученные результаты экспертизы? Сформулировать цель, предмет, методы исследований (машинный эксперимент и статистический анализ).
2. Описать имеющийся задел – написать реферат по данным <https://www.elibrary.ru/>, ключевые слова – фоноскопия, экспертиза, идентификация диктора по голосу….

* СРАВНЕНИЕ МЕТОДОВ ПРИНЯТИЯ РЕШЕНИЯ ПО ВЫХОДНЫМ ДАННЫМ МНОГОСЛОЙНОЙ НЕЙРОННОЙ СЕТИ В ЗАДАЧЕ ИДЕНТИФИКАЦИИ ДИКТОРА ПО ГОЛОСУ. Бучнева Т.И., Кудряшов М.Ю. В сборнике: Современные тенденции в науке и образовании. Сборник научных трудов по материалам Международной научно-практической конференции: в 5 частях. ООО "АР-Консалт". 2015. С. 160-163.
* АЛГОРИТМ ИДЕНТИФИКАЦИИ ЛИЧНОСТИ ПО ГОЛОСУ ДЛЯ САНКЦИОНИРОВАНИЯ ДОСТУПА К ИНФОРМАЦИИ Чернецова Е.А., Шишкин А.Д. Международный научно-исследовательский журнал. 2019. № 2 (80). С. 59-64.
* ВОЗМОЖНОСТИ ИДЕНТИФИКАЦИИ ЛИЧНОСТИ ПО ГОЛОСУ И ЗВУЧАЩЕЙ РЕЧИ ПО ДИАЛЕКТНЫМ ОСОБЕННОСТЯМ Березина О.В. В сборнике: АКТУАЛЬНЫЕ ВОПРОСЫ РАЗВИТИЯ ОБЩЕСТВЕННЫХ НАУК И ЮРИСПРУДЕНЦИИ. Сборник статей II Международной научно-практической конференции. Петрозаводск, 2022. С. 91-95.
* ИСПОЛЬЗОВАНИЕ МОБИЛЬНОГО ПРИЛОЖЕНИЯ BIRDNET ДЛЯ ИДЕНТИФИКАЦИИ ПО ГОЛОСУ МУХОЛОВКОВЫХ В ПРОЦЕССЕ ВНЕУРОЧНОЙ РАБОТЫ ОБУЧАЮЩИХСЯ Коробко А.А. В сборнике: МЕТОДИКА ОБУЧЕНИЯ ДИСЦИПЛИНАМ ЕСТЕСТВЕННО-НАУЧНОГО ЦИКЛА: ПРОБЛЕМЫ И ПЕРСПЕКТИВЫ. материалы XXI Всероссийской научно-практической конференции студентов, аспирантов и молодых ученых. Красноярский государственный педагогический университет им. В.П. Астафьева. 2022. С. 83-84.
* АВТОМАТИЧЕСКАЯ ИДЕНТИФИКАЦИЯ ЧЕЛОВЕКА ПО ГОЛОСУ С ИСПОЛЬЗОВАНИЕМ АЛГОРИТМА НА ОСНОВЕ МОДЕЛИ ГАУССОВЫХ СМЕСЕЙ Лукьянов Д.И., Михайлова А.С. В книге: Новые информационные технологии в научных исследованиях. материалы XХII Всероссийской научно-технической конференции студентов, молодых ученых и специалистов. Рязанский государственный радиотехнический университет. 2017. С. 100-102.
* СИСТЕМА ИДЕНТИФИКАЦИИ ДИКТОРА ПО ГОЛОСУ НА ОСНОВЕ GMM-UBM ПОДХОДА Костыренко В.С., Мазова А.В., Власов И.А., Петров В.Ю., Ревякина В.Я. Молодой ученый. 2019. № 25 (263). С. 27-31.

Задача обзора – показать, что тематика актуальна, и что существуют разные подходы к решению задачи.

1. Обоснование выбора инструмента – в работе используется технология сравнения косинусных мер векторов признаков. Алгоритм вычисления признаков реализован в пакете расширения Python Speechbrain.

(https://speechbrain.readthedocs.io/en/latest/API/speechbrain.inference.speaker.html#module-speechbrain.inference.speaker).

Для формирования вектора признаков в пакете Speechbrain используется модель ECAPA-TDNN, которая обучена более чем 9000 часов речи и обеспечивает точность верификации> 95% на тестах VoxCeleb .

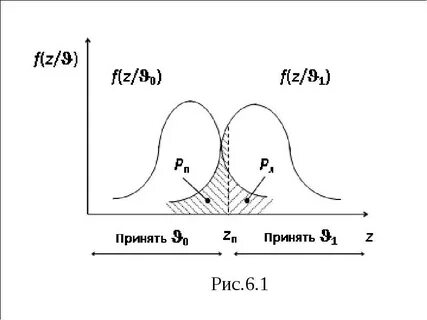
Для оценки сходства векторов признаков использовалась косинусная мера близости, вычисляемая по формуле

,

Где A и B – соответствующие вектора признаков

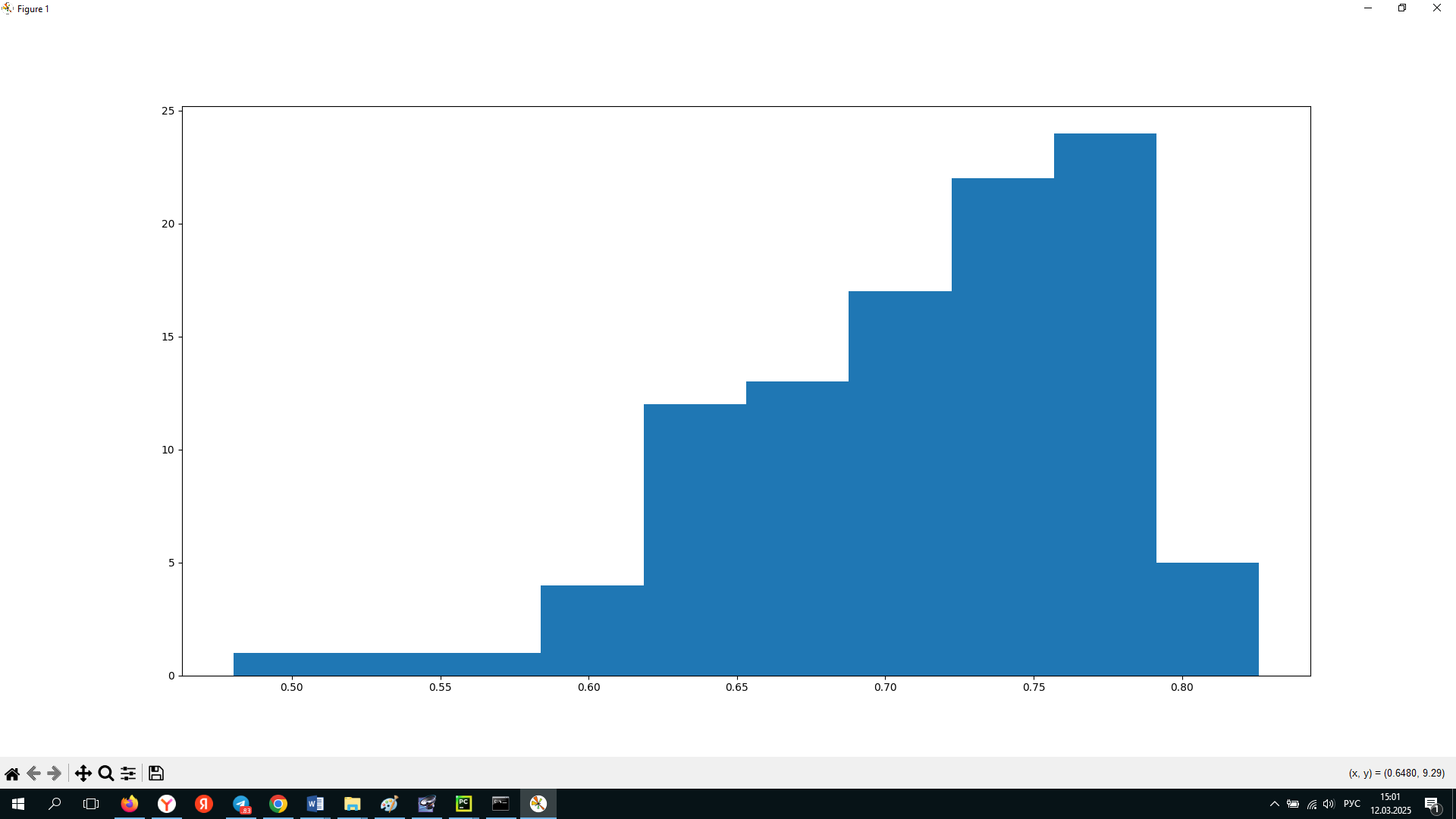
Необходимо раскрыть подробно описание модели, что за признаки и как она вычисляет, как с ней работать. Показать, что модель может быть использована для проведения фоноскопической экспертизы.

1. Более детально сформулировать проблему, решаемую в статье – как доказать состоятельность анализа? Сколько времени должна быть запись? Насколько велика вероятность принятия ошибочного решения? Привести примеры когда модель будет давать ошибки.
2. Раскрыть инструменты для доказательства корректности анализа – оценка доверительного интервала/доверительной вероятности для значений косинусной меры близости, при различной длительности анализируемых фрагментов речи. Критерий Неймана-Пирсона. Раскрыть – что такое доверительный интервал, и в чем суть критерия. Показать, что если значения косинусной меры близости для разных дикторов лежат в интервале, например 0-0.2 а для одного диктора – 0.7-1, то полученные оценки можно считать научно обоснованными.

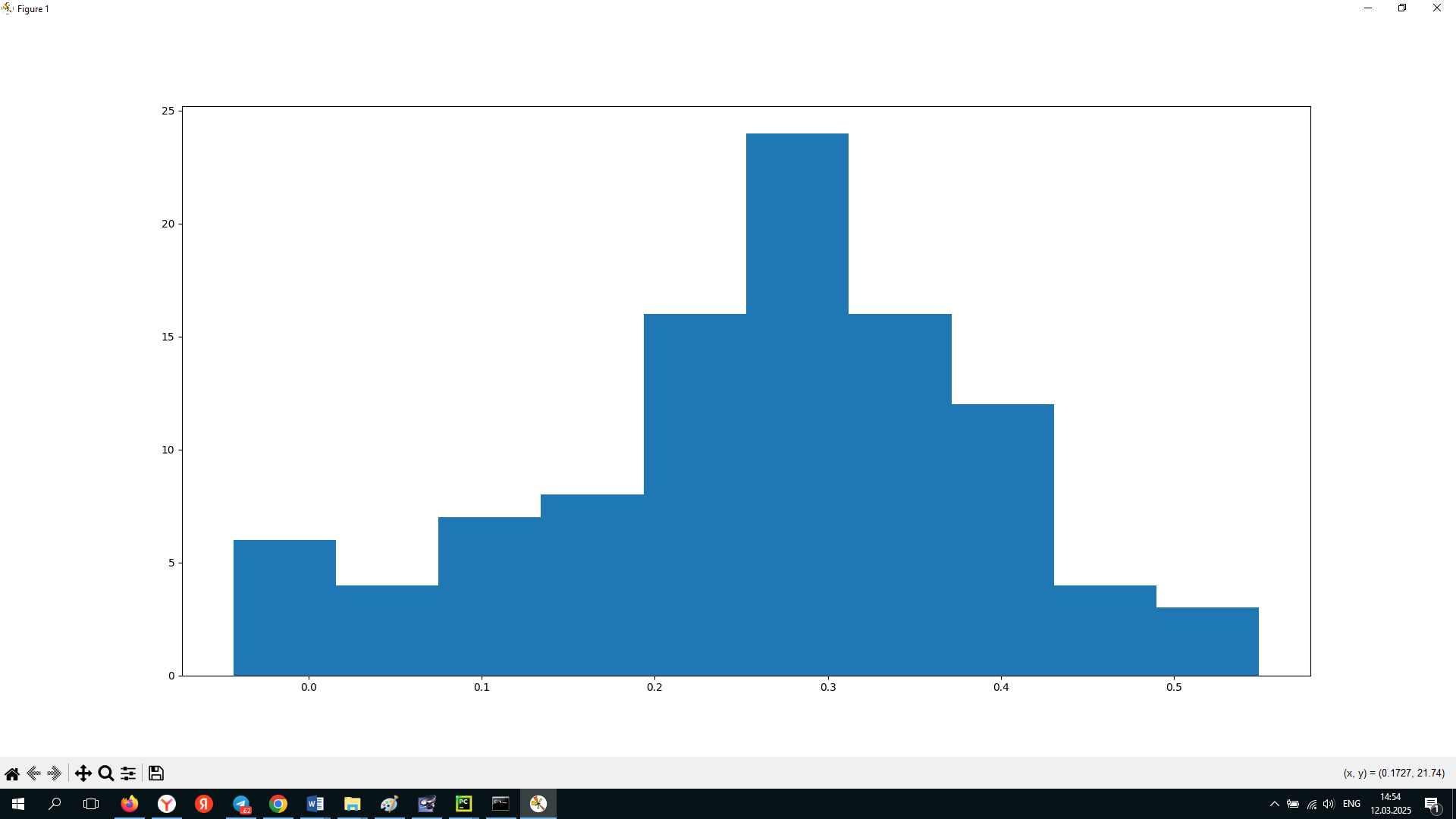


1. Описать методику тестирования технологии идентификации по голосу. Например, методика может быть следующая – взять звуковой файл, в котором запись только одного голоса. Взять для сравнения 100 пар фрагментов из этого файла, длительностью 1 с. Построить распределение значений косинусной меры близости. Зачем взять 100 пар фрагментов, длительностью 5 с. и повторить анализ. Также провести анализ распределения косинусной меры близости для файлов, содержащих голоса разных дикторов. Начертить блок схему проведения эксперимента.
2. Провести эксперимент согласно описанной методике. Необходимые файлы – исходные записи, программный код, таблицы результатов выгрузить в открытый репозиторий.

Примерные результаты эксперимента – распределения значений косинусной меры близости для одного и того же диктора, при сравниваемых фрагментах 5 с



Распределение значений косинусной меры близости для длительности фрагментов анализа 1 с



Рекомендуется оформить графики в векторном формате.

1. Сделать выводы – при использовании технологии … для оценки принадлежности голоса диктору с доверительной вероятностью…требуется исследование фрагмента речи длительностью не менее…, в противном случае возможны ошибки первого/второго рода, вероятность которых при длине исследуемого фрагмента менее 1 с составляет….
2. Оформить список литературы, не менее 15 пунктов.
3. Оформить на иностранном языке реферат, аннотацию, ключевые слова.