

Рабочий лист №1

Дата "01" 02 2025 г.
 (заполняется оргкомитетом)

Шифр ПН-40
 (заполняется оргкомитетом)

Оценка работы

(таблица заполняется по итогам проверки работы членами жюри олимпиады)

№ задания	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	Итого (итоговый балл, подпись председателя жюри)
Балл																88
№ задания	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	93
Балл																

Математика
 (название олимпиады, заполняется участником)

Программное обеспечение
 (профиль олимпиады, заполняется участником)

Задачи-1:

Источники данных: Телефонные звонки

Аудиофайлы (звонки, записи)

Источники аудио-текстовых расшифровок переговоров

Компоненты приема и обработки данных:

Запись и прием аудио данных

Преобразование данных в пригодный формат для обработки (wav, mp3)

Первичная аудио (улучшение шумов) и т.д.

Преобразование аудио (текстовый) в текст.

Компоненты хранения данных:

Хранение в метаданных текстовых расшифровок

данных; хранение в базе данных аудио, записи расшифровок

Хранение скринов данных

Компоненты расширения в текст:

Алгоритм автоматического расширения текста

Поддержка аудио

Обработка ошибок распознавания

Ком-ты анализа данных: Поиск ключевых слов, дат

Анализ текстов; классификация разговоров с М. Морозов

Отчеты о результатах работы

Ком-ты Безопасности и контроля доступа:

шифрование данных (при хранении и передаче)

Аутентификация и авторизация пользователей

Ведение реального пользователей

Интерфейс пользователя (UI):

Интерфейс для просмотра диаграмм, отчетов.

Технологии: • **FSR**: OpenAI, Google speech to Text.

• **Хранение**: PostgreSQL, MongoDB.

• **Анализ**: Python, TensorFlow.

• **Интерфейс**: React, Django.

Определение вероятности мошенничества (пример):

1. Получить аудио звонок

2. Преобразовать текст ~~в~~ с помощью FSR

3. Анализовать текст по ключевым словам, использовать ML для классификации

4. Если вероятность мошенничества больше порога вызвать предупреждение.

Задачи 2: 1, 2, ..., - **машин**; 0, 1, ... - **избранные**. (примеры кода на Python)

Псевдокод алгоритма:

1. Получить текстовую расшифровку разговора

2. Преобразовать текст:

а. Превести к нижнему регистру

б. Убрать знаки пунктуации

в. Разбить текст на слова

text = re.sub(r'[^\w\s]', '', text)

3. Поиск ключевых слов

а. Поиск по словарю часто используемых слов мошенниками (фразы).

б. Если такие совпали перейти к шагу 4.

4. Анализ контекста:

а. определение точности текста

б. найти повторяющиеся поразительные слова / фразы

в. Анализ достоверности фраз (точность, ритм разговора)

5. Классификация:

а. Превести текст в ML-модель

б. Получить вероятность мошенничества

в. **Вывод**

6. Принятие решения:

а. При подтвержденной вероятности больше порога вызвать предупреждение.

7. Уверенность:

а. Сохранить результат в базу данных

б. отправить уведомление

Псевдокод для шага 3:

Всех слов, четких фраз = 0, 0

for по всем словам:

if (ключевое слово найдено):

четких слов += 1

for по всем фразам:

if (ключевая фраза найдена):

четких фраз += 1

Дополнительный рабочий лист
(без рабочего листа №1 недействителен)

Дата "01" 02 2025 г.
(заполняется участником)

Шифр ПИ-40
(заполняется участником)

Задача 3: Код на Python

```
import re
# Предопределенные шаблоны для поиска информации
patterns = { "Паспорт": r"\b\d{4}\s?\d{6}\b",
              "Снилс": r"\b\d{33}-\d{33}-\d{33}\s?\d{23}\d",
              "Номер карты": r"\b\d{4}\s?\d{4}\s?\d{4}\s?\d{4}\s?\d{4}\b",
              "CVC/CVV": r"\b\d{3}\b"
            }

def extn_pers_inf(text):
    results = {}
    for key, pattern in patterns.items():
        matches = re.findall(pattern, text)
        if matches:
            results[key] = matches

    return results

# Пример использования:
text = "Паспорт 1234 567890, Снилс 123-456-78901,
        номер карты 1234 5678 9012 3456, CVC 123."

personall_inf = extn_pers_inf(text)

# Вывод в консоль:
print(personall_inf)
```

patterns - регулярное выражение.

extn_pers_inf - функция для извлечения данных.

Предположительная структура: (О - оператор; К - клиент)

О -кажите номер Вашего паспорта

К - 4512 345678

О - какой у вас номер СНИЛС

К - 123-456-78900

#0 - графический номер вашей карты

$\pi/k - 1234 \cdot 5678 \cdot 9012 \cdot 3456$

7 0-назовите СВС

K - 123