|  |
| --- |
| Пензенский государственный университет  Факультет вычислительной техники  Кафедра «Вычислительная техника» |
| Отчет  по лабораторной работе №9  по дисциплине «»  на тему «»  Вариант № |
|  |
|  |
| Проверил:  Карамышева Н.С. |
| Пенза |

**Задание.**

1. Получить описание заданной предметной области и набор соответствующих ей данных (варианты заданий приведены в конце пособия).

2. Провести исследование в соответствии с планом работы в рамках первой фазы исследования.

3. Подготовить отчёт по результатам первой фазы работ.

Отчёт должен содержать информацию в соответствии со следующей структурой:

1. Бизнес цель.

2. Ресурсы и риски.

3. Цель анализа данных.

4. План проекта.

**Ход работы.**

1 Задача определения подлинности банкнот (class: 5-й столбец набора данных) по набору из четырёх признаков, вычисленных по изображениям купюр.

Ссылка на репозиторий:

<https://archive.ics.uci.edu/ml/datasets/banknote+authentication>

**Бизнес цель проекта**

Основными заказчиками данного проекта могут являться банковские организации, государственные учреждения по контролю денежного оборота, а также следственные комитеты, государственные службы и органы правопорядка.

Помимо вышеперечисленных организаций основными пользователями могут выступать сети торговли в качестве потенциальных заказчиков ПО или модели определения, которая будет встроена в небольшое устройство (детектор или сканер).

Целью проекта является повышение эффективности распознавания поддельных банкнот, уменьшение затрат и автоматизация.

Существует множество решений и методик по распознаванию фальшивых банкнот, все они достаточно эффективны, однако алгоритмы засекречены.

Общие же методы распознавания представляют собой сравнения характерных признаков, которыми могут выступать как частотные и статистические параметры, а также общие характеристики изображения (яркость, контрастность).

**Оценка ситуации (ресурсы и риски)**

**Ресурсы**

В качестве данных для анализа и проектирования представлен реальный датасет banknote authentication Data set предоставляющий частотные признаки, полученные с помощью вейвлет-преобразования, как фальшивых, так и реальных купюр.

Другие датасеты найдены не были.

**Риски**

В проекте существуют общие риски по срыву сроков выполнения, недостаток финансирования. Помимо этого, возможным препятствием по созданию эффективного решения может быть недостаток обучающих данных или их некорректность. Подобным проблемы могут быть выявлены на второй фазе, когда представленный датасет будет исследован подробнее.

**Цель анализа данных**

Задача выделить корреляцию среди представленных в датасете признаков. Построить модель предсказания по имеющимся параметрам для автоматизации процесса определения фальшивых банкнот.

Существуют множество метрик оценки качества классификации. В этом проекте будут использоваться следующие: precision и recall, а также F1-Score.

Проект можно считать успешным, если результат метрик accuracy, precision и recall будет больше 80%.

**План проекта**

Планом проекта будет являться разработка модели предсказания (возможно для последующего её встраивания в детекторы для автоматического анализа банкнот).

Качество модели отслеживается с помощью установленных выше метрик.

Критерием устаревания модели может быть изменение общепринятых стандартов печати валюты, что приведёт к изменению признаков, отличающих оригинальные купюры от поддельных. Для восстановления актуальности потребуется создание или поиск новых данных и построение новой модели.

По окончанию работы должен быть сформирован финальный отчет, устраивающий заказчика работы. Принятие отчета заказчиком означает успешное завершение проекта.