# 1 ПРАКТИЧЕСКИЕ РАБОТЫ

Методические материалы по практическим работам

А.Е. Сулавко

Машинное обучение в приложениях биометрии

## Общие положения и методика оценивания результатов

В процессе обучения по дисциплине «Машинное обучение в приложениях биометрии» студент обязан выполнить 6 практических работ. Под выполнением практических работ подразумевается подготовка к работе, проведение экспериментальных исследований, подготовка отчета и его защита на практическом занятии. Выполнение практических работ студентами осуществляется в бригадах до 2 человек. Оформление отчета студентами осуществляется индивидуально в соответствии с принятыми в СПбГЭТУ правилами оформления студенческих работ. Отчет оформляется после выполнения экспериментальных исследований и представляется преподавателю на проверку. После проверки отчет либо возвращается (при наличии замечаний) на доработку, либо подписывается к защите. Практические работы защищаются студентами индивидуально. Каждый студент получает вопрос по теоретической и практической части, после чего ему предоставляется время для подготовки ответа. На защите практической работы студент должен показать: понимание и умение объяснять особенности применяемых методов, возможные области их применения и т.д., прогнозировать реакции исследуемого объекта на различные воздействия, навыки и умения, приобретенные при выполнении практических работы. При обсуждении ответа преподаватель может задать несколько уточняющих вопросов. В случае если студент демонстрирует достаточное знание вопроса, работа считается зачтенной. Текущий контроль включает в себя выполнение, сдачу в срок отчетов и их защиту по всем практическим работам, по результатам которой студент получает допуск к экзамену. Критерием оценки работы на практическом занятии является оценка, выставляемая по 5-ти балльной шкале в соответствии со следующими критериями: оценка в 5 баллов выставляется при отличном выполнении задания, то есть при наличии полных (с детальными пояснениями и выкладками), оригинальных и правильных решений задач, дополненных при необходимости документами, полученными в результате реализации (проверки) решения, верных ответов и высококачественного оформления работы. Оценка в 4 балла выставляется при правильном выполнении задания, то есть при наличии полных (с пояснениями и выкладками), оригинальных и правильных решений задач, дополненных при необходимости документами, полученными в результате реализации (проверки) решения, верных ответов. Оценка в 3 балла выставляется при наличии отдельных неточностей в ответах (включая грамматические ошибки) или неточностях в решении задач непринципиального характера (описки и случайные ошибки арифметического характера). Оценка в 2 и ниже баллов выставляется в случаях, когда в ответах и в решениях задач имеются неточности и ошибки, свидетельствующие о недостаточном понимании вопросов и требующие дополнительного обращения к тематическим материалам.

## Содержание отчетов

* Титульный лист
* Цель работы
* Формулировка задания
* Краткое описание выполнения основных этапов работы
* Выводы по работе

## Практические работы

### Выбор инструментов машинного обучения для проведения практических занятий. Изучение стека технологий машинного обучения

**Цель работы:** изучение подходов машинного обучения, применяемого для задач биометрического распознавания, а также определение пула нейросетевых архитектур для проведения практических занятий курса.

**Задание:**

Определить типы нейросетевых архитектур, используемых в задачах распознавания биометрических образов и подготовить сравнительную таблицу, содержащую достоинства и недостатки каждого из подходов для задач распознавания одного биометрического параметра (параметр назначается преподавателем).

**Контрольные вопросы:**

1. Какие методы машинного обучения Вам известны? Какие из них применяются в приложениях биометрии?
2. Какая из задач, решаемая методами машинного обучения, чаще всего применяется в приложениях биометрии?
3. Что представляют собой гиперпараметры в машинном обучении? Приведите пример.
4. Какие функции активации Вы знаете?
5. Какие архитектуры нейронных сетей чаще всего применяются в приложениях биометрии? Почему? Приведите примеры.

### Обучение байесовских классификаторов («наивная» схема и ее модификации с учетом присвоения весов признакам) для идентификации и верификации пользователей по клавиатурному почерку с использованием открытых наборов данных. Проведение вычислительного эксперимента по оценке вероятностей ошибок

**Цель работы:** изучение работы байесовских классификаторов для идентификации/верификации пользователей по клавиатурному почерку и формирование навыков проведения вычислительного эксперимента по оценке вероятностей ошибок первого и второго рода.

**Задание:**

Сформировать эталоны для «наивного» байесовского классификатора и произвести процедуру идентификации, используя открытый набор данных (выдается преподавателем). Для выполнения задания может быть использован программный пакет AIC desktop.

Сформировать эталоны для модификации байесовского классификатора и произвести процедуру верификации, используя открытый набор данных (выдается преподавателем). Для выполнения задания может быть использован программный пакет AIC desktop.

Для обоих типов классификаторов построить характеристические кривые вероятностей ошибок (для «осторожного» Байеса изменяя значение весовых коэффициентов признаков таким образом, чтобы минимизировать ошибки 1-го и 2-го типа).

**Контрольные вопросы:**

1. Что в широком смысле представляет собой байесовский классификатор, и каким образом он функционирует?
2. В чем заключается отличие модификации «наивного» байесовского классификатора от классической схемы?
3. Каким образом оценивается информативность признаков биометрических образцов?
4. Что позволяет изменять параметр, отвечающий за изменение количества эталонов в процессе моделирования?
5. Какие способы и подходы к вычислению весовых коэффициентов признаков Вы можете назвать для модификации байесовского классификатора?

### Анализ и визуализации биометрических образов. Построение спектрограмм голосовых паролей с использованием программных пакетов для анализа звука. Определение частоты основного тона диктора. Анализ ЭЭГ с помощью специализированных пакетов

**Цель работы:** изучение подходов к визуализации биометрических образов на примере построения спектрограмм голосовых паролей и анализа ЭЭГ, а также получение навыков работы с признаками указанных биометрических параметров.

**Задание:**

Сформировать набор спектрограмм голосовых паролей, взятых из открытой базы образцов (например, AIC-spkr-130). Провести эксперимент по определению частоты основного тона диктора. Провести анализ образцов ЭЭГ (например, AIC-eeg-19). Извлечь индивидуальные признаки, характеризующие испытуемых и позволяющие их идентифицировать. Для выполнения задания может быть использован программный пакет AIC desktop.

**Контрольные вопросы:**

1. Какие особенности голосового сигнала человека отражают спектрограммы?
2. Дайте определение частоты основного тона.
3. Какие архитектуры нейронных сетей, на Ваш взгляд, могут использоваться для выделения признаков из спектрограмм голосового сигнала человека? Почему?
4. Какие уникальные признаки можно выделять из изображений ЭЭГ человека?
5. Какие подходы к построению спектрограмм биометрических образов Вы знаете? В чем особенности Фурье-анализа и вейвлет-анализа?

### Построение и обучение многослойных нейронных сетей для идентификации подписантов по изображению и динамике воспроизведения подписи и использованием открытых наборов данных. Проведение вычислительного эксперимента по оценке вероятностей ошибок

**Цель работы:** получение навыков обучения глубоких нейронных сетей с целью идентификации подписантов.

**Задание:**

Сформировать эталоны для глубокой нейронной сети и произвести процедуру идентификации, используя открытый набор данных изображений подписей (например, AIC-sign-130). Построить характеристические кривые вероятностей ошибок. Для выполнения задания может быть использован программный пакет AIC desktop.

**Контрольные вопросы:**

1. Какие условия необходимо соблюдать при формировании базы рукописных образов подписантов (согласно ISO/IEC 19795-3)?
2. Каким образом происходит процесс извлечения признаков из рукописного образа подписанта, и какие параметры при этом вычисляются?
3. Какие подходы к извлечению признаков Вы использовали?
4. Каким образом оценивается информативность биометрического образа (подписи)?
5. Какую архитектуру искусственной нейронной сети Вы выбрали и почему?

### Обучение нейросетевых преобразователей биометрия-код для верификации личности диктора по голосовому паролю. Проведение вычислительного эксперимента по оценке вероятностей ошибок

**Цель работы:** изучение процесса обучения НПБК для верификации пользователя по голосовому паролю с использованием открытого набора данных.

**Задание:**

Произвести обучение классическогонейросетевого преобразователя биометрия-код (на базе линейных нейронов) и построить характеристические кривые вероятностей ошибок. Использовать открытые наборы данных голосовых паролей, например, AIC-sign-130.

**Контрольные вопросы:**

1. Перечислите особенности формирования базы образцов голосовых паролей. На какие аспекты стоит обратить особое внимание?
2. Каким образом производится оценка информативности голосовых признаков субъектов?
3. НПБК с какими параметрами был использован в Вашей работе? Почему?
4. Каким образом осуществлялось извлечение признаков?
5. Какие меры близости, используемые для работы с голосовыми признаками, Вы знаете? Какие меры близости используются при тестировании НПБК по ГОСТ Р 52633.3?

### Построение и обучение ансамблей нейронных сетей для распознавания субъектов по особенностям строения ушного канала. Проведение вычислительного эксперимента по оценке вероятностей ошибок

**Цель работы:** получение навыков построения и обучения ансамбля нейронных сетей для распознавания субъектов по особенностям строения ушного канала.

**Задание:**

Сформировать эталоны для гибридной нейронной сети и произвести процедуру идентификации, используя открытый набор данных строения ушного канала (например, AIC-ears-75). Построить характеристические кривые вероятностей ошибок. Для выполнения задания может быть использован программный пакет AIC desktop.

**Контрольные вопросы:**

1. С помощью какого метода получают данные о внутреннем строении наружного уха?
2. Каким образом Вы извлекали ключевые особенности эхо-сигнала, по которым можно с высокой точностью идентифицировать личность человека?
3. В чем заключается преимущество использования ансамбля нейронных сетей перед классическими байесовскими классификаторами?
4. В чем заключается преимущества использования информации о внутреннем строении наружного уха в качестве биометрического образа?
5. Какие методы объединения базовых классификаторов в ансамбль Вы знаете? Какой из них был использован в данной практической работе?