

**Министерство науки и высшего образования Российской Федерации  
Федеральное государственное бюджетное образовательное учреждение  
высшего образования  
«Бурятский государственный университет имени Доржи Банзарова»**

**ПРОГРАММА**

**вступительного экзамена по направлению подготовки**

**01.04.02 Прикладная математика и информатика**

**«Интеллектуальные системы в здравоохранении»**

г. Улан-Удэ, 2024

## **ОСНОВНЫЕ ТРЕБОВАНИЯ**

Программа вступительного испытания составлена в соответствии с требованиями ФГОС ВО по направлению подготовки.

Цель вступительного испытания: определить уровень подготовки, поступающих в БГУ по направлению 01.04.02 Прикладная математика и информатика «Интеллектуальные системы в здравоохранении».

В содержательную основу программы вступительных испытаний включены следующие дисциплины: Теория вероятностей и статистика, Математический анализ, Линейная алгебра, Структуры и алгоритмы обработки данных, Объектно-ориентированное программирование, Базы данных, Машинное обучение.

Вступительное испытание проходит в форме компьютерного тестирования.

Время, отведенное на тестирование - 45 мин.

Максимальная сумма баллов – 100.

Минимальный балл для сдачи экзамена по программам магистратуры - 65 баллов.

## **СОДЕРЖАНИЕ ПРОГРАММЫ**

### **ТЕОРИЯ ВЕРОЯТНОСТЕЙ И СТАТИСТИКА**

1. Эксперимент со случайным исходом. Понятие о вероятности. Вероятность события как предел относительной частоты его появления в серии независимых испытаний. Понятие закона больших чисел.
2. Пространство элементарных исходов как математическая модель эксперимента со случайным исходом. Алгебра событий. Несовместные события.
3. Вероятность в пространстве элементарных исходов. Классическая вероятность. Простейшие комбинаторные схемы. Вероятность суммы событий.
4. Условные вероятности. Вероятность произведения событий. Независимые события.
5. Формула полной вероятности. Формула Байеса
6. Выборочный метод. Типы данных. Первичные описательные статистики. Виды распределения данных. Нормальное распределение. Проблема статистического вывода. Гипотеза. Статистическая значимость. Критерий. Интерпретация. Выбор метода. Классификация методов.
7. Корреляция. Параметрические (корреляция Пирсона) и непараметрические (Спирмена, Кенделла, номинативная корреляция) методы исследования взаимосвязей. Диаграмма рассеяния. Методы исследования нелинейной связи. Частная корреляция. Анализ корреляционных матриц.
8. Критерий Стьюдента для зависимых и независимых выборок. Проверка условий применения. Критерий Вилкоксона. Критерий Манна-Уитни. Критерий Краскала-Уоллеса. Критерий Фридмана.

9. Однофакторный дисперсионный анализ. Множественные сравнения. Многофакторный дисперсионный анализ. Многомерный ANOVA.
10. Простая регрессия. Множественная регрессия. Нелинейная регрессия. Методы оценки параметров модели. Метод наименьших квадратов.

#### ЛИТЕРАТУРА

1. ТЕОРИЯ ВЕРОЯТНОСТЕЙ И МАТЕМАТИЧЕСКАЯ СТАТИСТИКА: Учебник/Гмурман В.Е.. —М.: Издательство Юрайт, 2016. —479 с. Режим доступа: <http://www.biblio-online.ru/book/6AB2913B-ECB0-45DF-AFE5-DBD442DECA4C>
2. ТЕОРИЯ ВЕРОЯТНОСТЕЙ И МАТЕМАТИЧЕСКАЯ СТАТИСТИКА: Учебник/Сидняев Н.И.. —М.: Издательство Юрайт, 2016. —219 с. Режим доступа: <http://www.biblio-online.ru/book/D943B16A-85DD-4E7C-BD46-16AB6E525178>
3. Теория вероятностей и математическая статистика: учеб.-метод. комплекс для студентов специальности 010501.65 Прикладная математика и информатика и направления бакалавриата 010501.62 Прикладная математика и информатика/В. В. Кибирев; М-во образования и науки Рос. Федерации, Бурят. гос. ун-т. —Улан-Удэ: Изд-во Бурятского госуниверситета, 2012. —130, [1] с.
4. ТЕОРИЯ ВЕРОЯТНОСТЕЙ И МАТЕМАТИЧЕСКАЯ СТАТИСТИКА: Учебник и практикум/Ивашев-Мусатов О.С.. —М.: Издательство Юрайт, 2016. —224 с. Режим доступа: <http://www.biblio-online.ru/book/BB281C15-FCBA-4362-B6C4-7A861A37F8A8>
5. Цыбиков А. С. Прикладной анализ данных: учебно-методическое пособие для обучающихся по направлениям подготовки 02.03.03 Математическое обеспечение и администрирование информационных систем, 02.03.01 Математика и компьютерные науки, 01.03.02 Прикладная математика и информатика, 09.03.03 Прикладная информатика/М-во образования и науки Рос. Федерации, Бурят. гос. ун-т; [сост. А. С. Цыбиков]. —Улан-Удэ: Изд-во Бурятского госуниверситета, 2017 Ч. 1. —2017. —145 с. (Электронный ресурс ИРБИС")

#### МАТЕМАТИЧЕСКИЙ АНАЛИЗ

1. Вещественные числа. Десятичная запись вещественного числа. Свойства вещественных чисел. Аксиома Архимеда. Свойство непрерывности.
2. Верхняя и нижняя грани числового множества, их характеристические свойства. Теорема о существовании верхней (нижней) грани у ограниченного сверху (снизу) числового множества.
3. Ограниченные отображения, верхняя и нижняя грани отображения.
4. Бесконечно малые, бесконечно большие последовательности, их свойства. Арифметические свойства сходящихся последовательностей.
5. Предельный переход в неравенствах. Теорема о существовании предела у ограниченной монотонной последовательности. Число “e”.
6. Теорема Больцано–Вейерштрасса о существовании частичного предела у ограниченной числовой последовательности. Верхний и нижний пределы последовательности.
7. Критерий Коши сходимости последовательности.
8. Предел функции в точке по Гейне и по Коши; эквивалентность этих определений. Односторонние пределы в точке. Арифметические операции над функциями, имеющими предел.
9. Критерий Коши существования предела функции в точке.

10. Замечательные пределы.
11. Бесконечно малые и бесконечно большие функции, теоремы о них. Сравнение бесконечно малых функций. Эквивалентные бесконечно малые функции.
12. Теорема о пределе сложной функции.
13. Непрерывность функции в точке. Односторонняя непрерывность. Арифметические операции над непрерывными функциями.
14. Свойство устойчивости знака непрерывной в точке функции.
15. Свойство локальной ограниченности непрерывной в точке функции.
16. Непрерывность элементарных функций.
17. Точки разрыва функции и их классификация. Теорема о точках разрыва монотонной на отрезке функции.
18. Первая теорема Коши (о прохождении непрерывной функции через нуль при смене знаков).
19. Вторая теорема Коши (о промежуточных значениях непрерывной на отрезке функции).
20. Первая теорема Вейерштрасса (об ограниченности непрерывной на отрезке функции).
21. Вторая теорема Вейерштрасса (о достижении верхней и нижней границей непрерывной на отрезке функцией).
22. Равномерная непрерывность. Теорема Кантора о равномерной непрерывности функции, непрерывной на отрезке.
23. Свойства открытых и замкнутых множеств. Компакт.
24. Теорема о равномерной непрерывности функции, непрерывной на компакте.
25. Производная, ее геометрический смысл. Односторонние производные.
26. Непрерывность функции, дифференцируемой в точке.
27. Производная суммы, произведения и частного двух функций.
28. Производная сложной и обратной функций. Дифференцирование функции, заданной параметрически.
29. Производные элементарных функций.
30. Производные высших порядков. Формула Лейбница.
31. Дифференциал функции, геометрический смысл дифференциала. Правила вычисления дифференциала. Инвариантность формы первого дифференциала.
32. Дифференциалы высших порядков. 33. Лемма Дарбу о возрастании или убывании функции в точке.
34. Теорема Ферма о локальном экстремуме функции.
35. Теорема Ролля о нуле производной.
36. Теорема Лагранжа (формула конечных приращений).
37. Теорема Коши (обобщенная формула конечных приращений).
38. Первое и второе правило Лопиталю.
39. Формулы Тейлора и Маклорена с остаточным членом в форме Пеано.
40. Формула Тейлора с остаточным членом в форме Лагранжа.
41. Разложение элементарных функций по формуле Маклорена.
42. Необходимое и достаточное условия локального экстремума.
43. Выпуклость графика функции. Достаточное условие выпуклости графика функции.
44. Необходимое и достаточное условия точки перегиба.
45. Асимптоты графика функции. Необходимое и достаточное условия существования наклонной асимптоты.

46. Первообразная и неопределенный интеграл. Основные свойства неопределенного интеграла. Таблица интегралов.
47. Замена переменной в неопределенном интеграле. Формула интегрирования по частям.
48. Интегрирование рациональных дробей.
49. Интегрирование тригонометрических выражений, универсальная тригонометрическая подстановка.
50. Интегрирование простейших иррациональных функций.

#### ЛИТЕРАТУРА

1. Краткий курс математического анализа: учебника для студентов высших учебных заведений/Л. Д. Кудрявцев. —Москва: ФИЗМАТЛИТ, 2015. —443 с. Режим доступа: [http://e.lanbook.com/books/element.php?pl1\\_id=71994](http://e.lanbook.com/books/element.php?pl1_id=71994)
2. МАТЕМАТИЧЕСКИЙ АНАЛИЗ: Учебник и практикум/Баврин И.И.. —М.: Издательство Юрайт, 2016. —327 с. Режим доступа: <http://www.biblio-online.ru/book/F574610E-5839-4AB7-8238-DFAB3C30FC21>
3. МАТЕМАТИЧЕСКИЙ АНАЛИЗ: Учебное пособие для бакалавров/Кытманов А.М.. —М.: Издательство Юрайт, 2016. —607 с. Режим доступа: <http://www.biblio-online.ru/book/6A8265F0-92EA-4A50-9D69-A2E3AA55DC00>
4. МАТЕМАТИЧЕСКИЙ АНАЛИЗ В 2 Ч. ЧАСТЬ 1 В 2 КН. КНИГА 1: Учебник/Ильин В.А., Садовничий В.А., Сендов Б.Х.. —М.: Издательство Юрайт, 2016. —331 с. Режим доступа: <http://www.biblio-online.ru/book/08178B51-9F3B-46B4-8C87-E6856F05D504>
5. МАТЕМАТИЧЕСКИЙ АНАЛИЗ В 2 Ч. ЧАСТЬ 1 В 2 КН. КНИГА 2: Учебник/Ильин В.А., Садовничий В.А., Сендов Б.Х.. —М.: Издательство Юрайт, 2016. —328 с. Режим доступа: <http://www.biblio-online.ru/book/CD0F2087-8E37-4572-9347-EB692E66B0C4>
6. Сборник задач по математическому анализу. —Москва: Физматлит, 2009. —496 с. Режим доступа: [http://e.lanbook.com/books/element.php?pl1\\_cid=25&pl1\\_id=2227](http://e.lanbook.com/books/element.php?pl1_cid=25&pl1_id=2227)
7. Лекции по математическому анализу: учебник для студентов вузов, обучающихся по направлениям 010400 «Прикладная математика и информатика», 010900 «Прикладные математика и физика»/О. В. Бесов. —Москва: ФИЗМАТЛИТ, 2015. —480 с. Режим доступа: [http://e.lanbook.com/books/element.php?pl1\\_id=72002](http://e.lanbook.com/books/element.php?pl1_id=72002)

#### ЛИНЕЙНАЯ АЛГЕБРА

1. Системы линейных уравнений. Метод Гаусса решения систем линейных уравнений. Определитель матрицы. Свойства определителя. Метод Крамера решения систем линейных уравнений.
2. Линейные пространства. Линейная зависимость и линейная независимость систем векторов. Базис и ранг системы векторов. Матрица перехода от одного базиса к другому. Координаты вектора в базисе. Изменение координат вектора при изменении базиса.
3. Кольцо многочленов. Делимость многочленов. Наибольший общий делитель двух многочленов. Алгоритм Евклида нахождения наибольшего общего делителя.
4. Линейные преобразования линейных пространств. Матрица линейного преобразования в базисе. Изменение матрицы линейного преобразования при изменении базиса.
5. Собственные числа и собственные векторы линейного преобразования. Характеристический многочлен линейного преобразования. Нахождение собственных чисел и собственных векторов линейного преобразования.

6. Евклидовы пространства. Симметрические преобразования. Нахождение ортонормированного базиса, состоящего из собственных векторов симметрического преобразования.

#### ЛИТЕРАТУРА

1. Линейная алгебра/И. А. Мальцев. —Москва: Лань, 2010. —384 с. Режим доступа: [http://e.lanbook.com/books/element.php?p11\\_cid=25&p11\\_id=610](http://e.lanbook.com/books/element.php?p11_cid=25&p11_id=610)
2. Курс высшей алгебры: учеб./А.Г. Курош. —Москва: Лань, 2013. —431 с. Режим доступа: [http://e.lanbook.com/books/element.php?p11\\_cid=25&p11\\_id=30198](http://e.lanbook.com/books/element.php?p11_cid=25&p11_id=30198)
3. Линейная алгебра/Б. А. Горлач. —Москва: Лань, 2012. —480 с. Режим доступа: [http://e.lanbook.com/books/element.php?p11\\_cid=25&p11\\_id=4042](http://e.lanbook.com/books/element.php?p11_cid=25&p11_id=4042)

## СТРУКТУРЫ И АЛГОРИТМЫ ОБРАБОТКИ ДАННЫХ

Алгоритмы теории чисел

Линейные контейнеры

Нелинейные контейнеры

Сортировки

Геометрия

Поиск

Рекурсия

Перебор

Линейные алгоритмы

Динамическое программирование

Структуры данных

Основы теории графов

Кратчайшие пути

Остовные деревья

Наименьший общий предок

Максимальный поток

Строковые алгоритмы и структуры

Паросочетания

#### ЛИТЕРАТУРА

1. Алгоритмы и программы. Язык С++: учебное пособие для студентов, обучающихся по направлению «Прикладная информатика»/Е. А. Конова, Г. А. Поллак. —Москва: Лань, 2016. —384 с. Режим доступа: [http://e.lanbook.com/books/element.php?p11\\_id=72986](http://e.lanbook.com/books/element.php?p11_id=72986)
2. Универсальное кодирование. Теория и алгоритмы/Штарьков Ю.М.. —Москва: Физматлит, 2013 Режим доступа: [http://e.lanbook.com/books/element.php?p11\\_id=59667](http://e.lanbook.com/books/element.php?p11_id=59667)

## БАЗЫ ДАННЫХ

1. Антипаттерны SQL баз данных. Понятие антипаттернов баз данных. Простые деревья. Хранение древовидных структу. Хранение многозначных атрибутов. Антипаттерн "Один размер для всех случаев". Способы распознавания, допустимые способы использования, решения. Антипаттерн "Объект-Ключ-Значение". Способы распознавания, допустимые способы использования, решения. Антипаттерн "Клонирование таблиц (столбцов)".

Способы распознавания, допустимые способы использования, решения. Антипаттерн "Полиморфные ассоциации". Способы распознавания, допустимые способы использования, решения. Многостолбчатые атрибуты. Антипаттерн "Пропуск ограничений". Способы распознавания, допустимые способы использования, решения. Хранение файлов. Антипаттерн "Ограничение столбца конкретными значениями". Способы распознавания, допустимые способы использования, решения. Создание динамических SQL-запросов. Антипаттерн "Считываемые пароли". Способы распознавания, допустимые способы использования, решения.

2. SQL базы данных. Введение в SQL базы данных. Принципы работы, язык запросов, внутреннее устройство.

3. Целостность и упорядочивание данных. Поддержка целостности данных. Внешние и родительские ключи. Ограничения внешнего ключа.

4. Программируемые объекты баз данных. Представления. Типы представлений

5. Согласованность данных. Транзакции. Требования ACID. "Грязные" транзакции. Блокировки данных. Типы блокировок. Области базы данных (Разделы). Масштабирование.

6. NoSQL базы данных. NoSQL базы данных. Особенности SQL-баз данных. Преимущества, недостатки. Модели распределения. Требования к транзакционным системам (ACID). Проблема согласованности. Теорема CAP. NoSQL базы данных. Основные характеристики NoSQL-баз данных. Модели данных. Map-Reduce (отображение-свертка). Базы данных "ключ-значение". Документные базы данных. Семейство столбцов. Графовые базы данных. Миграции схем. Многовариантная персистентность. Индексирование NoSQL-баз данных. Понятие репликации данных. Принципы сегментирования. Сегментированные кластеры.

## ЛИТЕРАТУРА

1. БАЗЫ ДАННЫХ: Учебник и практикум/Нестеров С.А.. —М.: Издательство Юрайт, 2016. —230 с. Режим доступа: <http://www.biblio-online.ru/book/B5E199E0-F0B1-4B55-AF98-9B7BC4841BCC>

2. Базы данных: учебно-методическое пособие для обучающихся по направлениям подготовки 02.03.03 Математическое обеспечение и администрирование информационных систем, 09.03.03 Прикладная информатика, 01.03.02 Прикладная математика и информатика, 02.03.01 Математика и компьютерные науки/М-во образования и науки Рос. Федерации, Бурят. гос. ун-т; [сост.: Г. И. Занданова, О. А. Лобсанова]

3. MongoDB в действии/Бэнкер К.. —Москва: ДМК Пресс, 2012 Режим доступа: [http://e.lanbook.com/books/element.php?pl1\\_cid=25&pl1\\_id=4156](http://e.lanbook.com/books/element.php?pl1_cid=25&pl1_id=4156)

4. БАЗЫ ДАННЫХ: ПРОЕКТИРОВАНИЕ. ПРАКТИКУМ: Учебное пособие/Стружкин Н.П., Годин В.В.. —М.: Издательство Юрайт, 2016. —291 с. Режим доступа: <http://www.biblio-online.ru/book/61C51717-0D09-4A82-8DC8-42BAD8347E3D>

## ОБЪЕКТНО-ОРИЕНТИРОВАННОЕ ПРОГРАММИРОВАНИЕ

1. Классы в языке C#. Определение класса, методы, аргументы. Классы и объекты. Статические методы и атрибуты класса. Конструкторы. Перегрузка методов. Сборка мусора

2. Наследование в языке C#. Синтаксис наследования. Модификаторы доступа. Наследование и конструкторы. Конструктор гарантирует инициализацию. Восходящее преобразование. Полиморфизм.
3. Абстрактные классы. Интерфейсы.
- 4.. Лямбда выражения. Параметризация и обобщения
5. Коллекции в языке C#.
6. Исключительные ситуации. Вызов и обработка исключений. Создание собственных исключений
7. Понятие пространства имён
8. Основы ООП в диаграммах UML - классы, наследование, полиморфизм, абстрактные классы, интерфейсы.
9. Паттерн Стратегия
10. Паттерн Наблюдатель
11. Паттерн Декоратор
12. Паттерн Простая Фабрика. Паттерн Фабричный метод
13. Паттерн Абстрактная фабрика
14. Паттерн Команда
15. Паттерн Адаптер. Паттерн Фасад.

#### ЛИТЕРАТУРА

1. ОБЪЕКТНО-ОРИЕНТИРОВАННОЕ ПРОГРАММИРОВАНИЕ: Учебное пособие/Тузовский А.Ф.. —М.: Издательство Юрайт, 2016. —206 с. Режим доступа: <http://www.biblio-online.ru/book/5B61CA55-D3ED-4574-977E-B869CAFF31D0>
2. ПРОГРАММИРОВАНИЕ НА VISUAL C# 2013: Учебное пособие/Казанский А.А.. — М.: Издательство Юрайт, 2016. —191 с. Режим доступа: <http://www.biblio-online.ru/book/C173A2E1-DD88-40D3-8974-B0A2DE2355CF>
3. ПРОГРАММИРОВАНИЕ НА VISUAL C# 2013: Учебное пособие/Казанский А.А.. — М.: Издательство Юрайт, 2016. —191 с. Режим доступа: <http://www.biblio-online.ru/book/517E2E45-BE20-42EB-955F-05A5E655C43C>

#### МАШИННОЕ ОБУЧЕНИЕ

1. Постановка задач обучения по прецедентам. Объекты и признаки. Типы шкал: бинарные, номинальные, порядковые, количественные. Типы задач: классификация, регрессия, прогнозирование, ранжирование. Основные понятия: модель алгоритмов, метод обучения, функция потерь и функционал качества, принцип минимизации эмпирического риска, обобщающая способность, скользящий контроль. Линейные модели регрессии и классификации. Метод наименьших квадратов. Полиномиальная регрессия. Примеры прикладных задач.
2. Гипотезы компактности и непрерывности. Обобщённый метрический классификатор. Метод ближайших соседей kNN и его обобщения. Подбор числа k по критерию скользящего контроля. Метод окна Парзена с постоянной и переменной шириной окна. Метод потенциальных функций и его связь с линейной моделью классификации. Задача отбора эталонов. Полный скользящий контроль (CCV), формула быстрого вычисления для метода 1NN. Профиль компактности. Отбор эталонных объектов на основе минимизации функционала полного скользящего контроля. Непараметрическая регрессия. Локально взвешенный метод наименьших квадратов. Ядерное сглаживание. Оценка Надарая-

Ватсона с постоянной и переменной шириной окна. Выбор функции ядра и ширины окна сглаживания. Задача отсева выбросов. Робастная непараметрическая регрессия. Алгоритм LOWESS.

3. Постановка задачи кластеризации. Примеры прикладных задач. Типы кластерных структур. Алгоритм k-средних и EM-алгоритм для разделения гауссовской смеси. Алгоритм DBSCAN. Агломеративная кластеризация, Алгоритм Ланса-Вильямса и его частные случаи. Алгоритм построения дендрограммы. Определение числа кластеров.

4. Оптимальная разделяющая гиперплоскость. Понятие зазора между классами (margin). Случаи линейной разделимости и отсутствия линейной разделимости. Связь с минимизацией регуляризованного эмпирического риска. Кусочно-линейная функция потерь. Задача квадратичного программирования и двойственная задача. Понятие опорных векторов. Рекомендации по выбору константы C. Функция ядра (kernel functions), спрямляющее пространство, теорема Мерсера. Способы конструктивного построения ядер. Примеры ядер. SVM-регрессия. Регуляризации для отбора признаков: LASSO SVM, Elastic Net SVM, SFM, RFM. Метод релевантных векторов RVM

5. Задача регрессии, многомерная линейная регрессия. Метод наименьших квадратов, его вероятностный смысл и геометрический смысл. Сингулярное разложение. Проблемы мультиколлинеарности и переобучения. Регуляризация. Гребневая регрессия через сингулярное разложение. Методы отбора признаков: Лассо Тибширани, Elastic Net, сравнение с гребневой регрессией. Метод главных компонент и декоррелирующее преобразование Карунена-Лоэва, его связь с сингулярным разложением. Спектральный подход к решению задачи наименьших квадратов.

6. Метод Ньютона-Рафсона, метод Ньютона-Гаусса. Обобщённая аддитивная модель (GAM): метод настройки с возвращениями (backfitting) Части-Тибширани. Логистическая регрессия. Метод наименьших квадратов с итеративным пересчётом весов (IRLS). Пример прикладной задачи: кредитный скоринг. Бинаризация признаков. Скоринговые карты и оценивание вероятности дефолта. Риск кредитного портфеля банка. Обобщённая линейная модель (GLM). Экспоненциальное семейство распределений. Неквадратичные функции потерь. Метод наименьших модулей. Квантильная регрессия. Пример прикладной задачи: прогнозирование потребительского спроса. Робастная регрессия, функции потерь с горизонтальными асимптотами

7. Линейный классификатор, модель МакКаллока-Питтса, непрерывные аппроксимации пороговой функции потерь. Метод стохастического градиента SG. Метод стохастического среднего градиента SAG. Эвристики: инициализация весов, порядок предъявления объектов, выбор величины градиентного шага, «выбивание» из локальных минимумов. Проблема мультиколлинеарности и переобучения, регуляризация или редукция весов (weight decay). Вероятностная постановка задачи классификации. Принцип максимума правдоподобия. Вероятностная интерпретация регуляризации, совместное правдоподобие данных и модели. Принцип максимума апостериорной вероятности. Гауссовский и лапласовский регуляризаторы. Логистическая регрессия. Принцип максимума правдоподобия и логарифмическая функция потерь. Метод стохастического градиента для логарифмической функции потерь. Многоклассовая логистическая регрессия. Регуляризованная логистическая регрессия. Калибровка Платта.

8. Байесовская теория классификации. Оптимальный байесовский классификатор. Генеративные и дискриминативные модели классификации. Наивный байесовский классификатор. Линейный наивный байесовский классификатор в случае

экспоненциального семейства распределений. Мультиномиальный наивный байесовский классификатор для классификации текстов. Метод парзеновского окна. Выбор функции ядра. Выбор ширины окна, переменная ширина окна. Нормальный дискриминантный анализ. Квадратичный дискриминант. Вид разделяющей поверхности. Подстановочный алгоритм, его недостатки и способы их устранения. Линейный дискриминант Фишера. Связь линейного дискриминанта Фишера с методом наименьших квадратов. Смесь многомерных нормальных распределений. Сеть радиальных базисных функций (RBF) и применение EM-алгоритма для её настройки. Сравнение RBF-сети и SVM с гауссовским ядром.

9. Восстановление плотности распределения. Параметрическое оценивание плотности. Многомерное нормальное распределение, геометрическая интерпретация. Выборочные оценки параметров многомерного нормального распределения. Проблемы мультиколлинеарности и переобучения. Регуляризация ковариационной матрицы. Матричное дифференцирование. Вывод оценок параметров многомерного нормального распределения. Непараметрическое оценивание плотности. Ядерная оценка плотности Парзена-Розенблатта. Одномерный и многомерный случаи. Смесь распределений. EM-алгоритм как метод простых итераций. Обобщённый EM-алгоритм. Стохастический EM-алгоритм. Детали реализации EM-алгоритма. Критерий останова. Выбор начального приближения. Выбор числа компонентов смеси. Пошаговая стратегия. Иерархический EM-алгоритм.

10. Биологический нейрон, модель МакКаллока-Питтса как линейный классификатор. Функции активации. Проблема полноты. Задача исключающего или. Полнота двухслойных сетей в пространстве булевых функций. Алгоритм обратного распространения ошибок. Быстрые методы стохастического градиента: Поляка, Нестерова, AdaGrad, RMSProp, AdaDelta, Adam, Nadam, диагональный метод Левенберга-Марквардта. Проблема взрыва градиента и эвристика gradient clipping. Метод случайных отключений нейронов (Dropout). Интерпретации Dropout. Обратный Dropout и L2-регуляризация. Функции активации ReLU и PReLU. Проблема «паралича» сети. Эвристики для формирования начального приближения. Метод послойной настройки сети. Подбор структуры сети: методы постепенного усложнения сети, оптимальное прореживание нейронных сетей (optimal brain damage).

11. Критерии качества классификации: чувствительность и специфичность, ROC-кривая и AUC, точность и полнота, AUC-PR. Внутренние и внешние критерии. Эмпирические и аналитические критерии. Скользящий контроль, разновидности эмпирических оценок скользящего контроля. Критерий непротиворечивости. Разновидности аналитических оценок. Регуляризация. Критерий Акаике (AIC). Байесовский информационный критерий (BIC). Оценка Вапника-Червоненкиса. Сложность задачи отбора признаков. Полный перебор. Метод добавления и удаления, шаговая регрессия. Поиск в глубину, метод ветвей и границ. Усечённый поиск в ширину, многорядный итерационный алгоритм МГУА. Генетический алгоритм, его сходство с МГУА. Случайный поиск и Случайный поиск с адаптацией (СПА) Лабораторная работа. 4(0) ч. Нейросетевая реализация линейной модели (регрессия, классификация) с помощью TensorFlow. Оценка качества алгоритма

12. Понятие логической закономерности. Параметрические семейства закономерностей: конъюнкции пороговых правил, синдромные правила, шары, гиперплоскости. Переборные алгоритмы синтеза конъюнкций: стохастический локальный поиск, стабилизация,

редукция. Двухкритериальный отбор информативных закономерностей, парето-оптимальный фронт в  $(p,n)$ -пространстве. Статистический критерий информативности, точный тест Фишера. Сравнение областей эвристических и статистических закономерностей. Разнообразие критериев информативности в  $(p,n)$ -пространстве. Решающее дерево. Жадная нисходящая стратегия «разделяй и властвуй». Алгоритм ID3. Недостатки жадной стратегии и способы их устранения. Проблема переобучения. Вывод критериев ветвления. Мера нечистоты (impurity) распределения. Энтропийный критерий, критерий Джини. Редукция решающих деревьев: предредукция и постредукция. Алгоритм C4.5. Деревья регрессии. Алгоритм CART. Небрежные решающие деревья (oblivious decision tree). Решающий лес. Случайный лес (Random Forest). Решающий пень. Бинаризация признаков. Алгоритм разбиения области значений признака на информативные зоны. Решающий список. Жадный алгоритм синтеза списка. Преобразование решающего дерева в решающий список.

13. Основные понятия: базовый алгоритм, корректирующая операция. Простое голосование (комитет большинства). Стохастические методы: бэггинг и метод случайных подпространств. Случайный лес (Random Forest). Взвешенное голосование. Преобразование простого голосования во взвешенное. Алгоритм AdaBoost. Экспоненциальная аппроксимация пороговой функции потерь. Процесс последовательного обучения базовых алгоритмов. Теорема о сходимости бустинга. Идентификация нетипичных объектов (выбросов). Теоретические обоснования. Обобщающая способность бустинга. Базовые алгоритмы в бустинге. Решающие пни. Сравнение бэггинга и бустинга. Алгоритм ComBoost. Обобщение на большое число классов. Лабораторная работа. 4(0) ч. Реализация стохастических методов: бэггинг и метод случайных подпространств. Реализация случайного леса (Random Forest). Взвешенное голосование.

14. Обоснования глубоких нейронных сетей: выразительные возможности, скорость сходимости при избыточной параметризации. Свёрточные нейронные сети (CNN) для изображений. Свёрточный нейрон. Pooling нейрон. Выборка размеченных изображений ImageNet. ResNet: остаточная нейронная сеть (residual NN). Сквозные связи между слоями (skip connection). Свёрточные сети для сигналов, текстов, графов, игр.

15. Рекуррентные нейронные сети (RNN). Обучение рекуррентных сетей: Backpropagation Through Time (BPTT). Сети долгой кратковременной памяти (Long short-term memory, LSTM). Рекуррентные сети Gated Recurrent Unit (GRU) и Simple Recurrent Unit (SRU).

16. Нейронная сеть Кохонена. Конкурентное обучение, стратегии WTA и WTM. Самоорганизующаяся карта Кохонена. Применение для визуального анализа данных. Искусство интерпретации карт Кохонена. Автокодировщик. Линейный AE, SAE, DAE, CAE, RAE, VAE, AE для классификации, многослойный AE. Пред-обучение нейронных сетей (pre-training). Перенос обучения (transfer learning). Многозадачное обучение (multi-task learning). Самостоятельное обучение (self-supervised learning). Дистилляция моделей или суррогатное моделирование. Обучение с использованием привилегированной информации (learning using privileged information, LUPI). Генеративные состязательные сети (generative adversarial net, GAN).

## ЛИТЕРАТУРА

1. Машинное обучение (курс лекций, К.В.Воронцов). Режим доступа: <https://bit.ly/ML-Vorontsov>

2. Машинное обучение. Вводная лекция. К.В. Воронцов, Школа анализа данных, Яндекс.  
Режим доступа:  
<https://www.youtube.com/watch?v=SZkrxWhI5qM&list=PLJOzdkh8T5krxc4HsHbB8g8f0hu7973fK>

3. Машинное обучение. Наука и искусство построения алгоритмов, которые извлекают знания из данных: 8 научно-популярная литература/П. Флах ; пер. с англ. А. А. Слинкин.  
—Москва: ДМК Пресс, 2015. —400 с. Режим доступа:  
[http://e.lanbook.com/books/element.php?pl1\\_id=69955](http://e.lanbook.com/books/element.php?pl1_id=69955)