

МИНОБРНАУКИ РОССИИ

**Федеральное государственное бюджетное образовательное
учреждение высшего образования "Пермский
государственный национальный исследовательский
университет"**

Кафедра прикладной математики и информатики

Авторы-составители: **Гусев Андрей Леонидович**

Рабочая программа дисциплины
ИНТЕЛЛЕКТУАЛЬНЫЙ АНАЛИЗ ДАННЫХ
Код УМК 92196

Утверждено
Протокол №9
от «21» мая 2019 г.

Пермь, 2019

1. Наименование дисциплины

Интеллектуальный анализ данных

2. Место дисциплины в структуре образовательной программы

Дисциплина входит в базовую часть Блока « М.1 » образовательной программы по направлениям подготовки (специальностям):

Направление: **01.04.02** Прикладная математика и информатика
направленность Математическое и программное обеспечение вычислительных систем

3. Планируемые результаты обучения по дисциплине

В результате освоения дисциплины **Интеллектуальный анализ данных** у обучающегося должны быть сформированы следующие компетенции:

01.04.02 Прикладная математика и информатика (направленность : Математическое и программное обеспечение вычислительных систем)

ОПК.2 Способен совершенствовать и реализовывать новые математические методы решения прикладных задач

Индикаторы

ОПК.2.1 Проводит анализ и обоснование применимости конкретного математического метода для решения прикладной задачи

4. Объем и содержание дисциплины

Направления подготовки	01.04.02 Прикладная математика и информатика (направленность: Математическое и программное обеспечение вычислительных систем)
форма обучения	очная
№№ триместров, выделенных для изучения дисциплины	1
Объем дисциплины (з.е.)	4
Объем дисциплины (ак.час.)	144
Контактная работа с преподавателем (ак.час.), в том числе:	48
Проведение лекционных занятий	24
Проведение практических занятий, семинаров	24
Самостоятельная работа (ак.час.)	96
Формы текущего контроля	Входное тестирование (1) Защищаемое контрольное мероприятие (2) Итоговое контрольное мероприятие (1)
Формы промежуточной аттестации	Экзамен (1 триместр)

5. Аннотированное описание содержания разделов и тем дисциплины

Введение

Классификация основных задач, связанных с обработкой больших объемов статистической информации. Роль и место интеллектуального анализа данных. Связь интеллектуального анализа информации с математической статистикой. Теоретические основы, методы и алгоритмы, составляющие интеллектуальный анализ данных.

Постановка задач многомерного статистического анализа.

Тема 1. Задачи управления.

- 1.1. Управляющие факторы. Управляемые факторы.
- 1.2. Нелинейность I рода. Нелинейность II рода.
- 1.3. Допустимый показатель. Доля управляемости. Целевой показатель.
- 1.4. Маркерный показатель. Индикативный показатель.

Тема 2. Задачи прогноза.

- 2.1. Сжатие информационного пространства.
- 2.2. Расширение информационного пространства.
- 2.3. Оценка воздействия неизвестного управляющего вектора. Прогноз

Тема 3. Задача «чёрного ящика».

- 3.1. Задача чёрного ящика, как обобщающая задача.
- 3.2. Корреляционный анализ.
- 3.3. Регрессионный анализ.

Снижение размерности задачи и классификация

Тема 1. Обезразмеривание и нормировка.

- 1.1. Смысл обезразмеривания.
- 1.2. Нормировка исходных данных. Способы нормирования и их цель.
- 1.3. Определение признаков (факторов) связанных между собой.
- 1.4. Метод главных компонент.

Анализ данных

Тема 1. Факторный анализ.

- 1.1. Случаи, когда применим факторный анализ.
- 1.2. Методы и способы.

Тема 2. Теория классификации.

- 2.1. Способы классификации.
- 2.2. Кластерный анализ.

Тема 3. Дискриминантный анализ.

- 3.1. Случаи, когда применим дискриминантный анализ.
- 3.2. Методы и способы.

6. Методические указания для обучающихся по освоению дисциплины

Освоение дисциплины требует систематического изучения всех тем в той последовательности, в какой они указаны в рабочей программе.

Основными видами учебной работы являются аудиторские занятия. Их цель - расширить базовые знания обучающихся по осваиваемой дисциплине и систему теоретических ориентиров для последующего более глубокого освоения программного материала в ходе самостоятельной работы. Обучающемуся важно помнить, что контактная работа с преподавателем эффективно помогает ему овладеть программным материалом благодаря расстановке необходимых акцентов и удержанию внимания интонационными модуляциями голоса, а также подключением аудио-визуального механизма восприятия информации.

Самостоятельная работа преследует следующие цели:

- закрепление и совершенствование теоретических знаний, полученных на лекционных занятиях;
- формирование навыков подготовки текстовой составляющей информации учебного и научного назначения для размещения в различных информационных системах;
- совершенствование навыков поиска научных публикаций и образовательных ресурсов, размещенных в сети Интернет;
- самоконтроль освоения программного материала.

Обучающемуся необходимо помнить, что результаты самостоятельной работы контролируются преподавателем во время проведения мероприятий текущего контроля и учитываются при промежуточной аттестации.

Обучающимся с ОВЗ и инвалидов предоставляется возможность выбора форм проведения мероприятий текущего контроля, альтернативных формам, предусмотренным рабочей программой дисциплины. Предусматривается возможность увеличения в пределах 1 академического часа времени, отводимого на выполнение контрольных мероприятий.

Процедура оценивания результатов обучения инвалидов и лиц с ограниченными возможностями здоровья по дисциплине предусматривает предоставление информации в формах, адаптированных к ограничениям их здоровья и восприятия информации.

При проведении текущего контроля применяются оценочные средства, обеспечивающие передачу информации, от обучающегося к преподавателю, с учетом психофизиологических особенностей здоровья обучающихся.

7. Перечень учебно-методического обеспечения для самостоятельной работы обучающихся по дисциплине

При самостоятельной работе обучающимся следует использовать:

- конспекты лекций;
- литературу из перечня основной и дополнительной учебной литературы, необходимой для освоения дисциплины (модуля);
- текст лекций на электронных носителях;
- ресурсы информационно-телекоммуникационной сети "Интернет", необходимые для освоения дисциплины;
- лицензионное и свободно распространяемое программное обеспечение из перечня информационных технологий, используемых при осуществлении образовательного процесса по дисциплине;
- методические указания для обучающихся по освоению дисциплины.

8. Перечень основной и дополнительной учебной литературы

Основная:

1. Федин, Ф. О. Анализ данных. Часть 2. Инструменты Data Mining : учебное пособие / Ф. О. Федин, Ф. Ф. Федин. — Москва : Московский городской педагогический университет, 2012. — 308 с. — ISBN 2227-8397. — Текст : электронный // Электронно-библиотечная система IPR BOOKS : [сайт]. <http://www.iprbookshop.ru/26445>

2. Федин, Ф. О. Анализ данных. Часть 1. Подготовка данных к анализу : учебное пособие / Ф. О. Федин, Ф. Ф. Федин. — Москва : Московский городской педагогический университет, 2012. — 204 с. — ISBN 2227-8397. — Текст : электронный // Электронно-библиотечная система IPR BOOKS : [сайт]. <http://www.iprbookshop.ru/26444>

Дополнительная:

1. Пальмов, С. В. Интеллектуальный анализ данных : учебное пособие / С. В. Пальмов. — Самара : Поволжский государственный университет телекоммуникаций и информатики, 2017. — 127 с. — ISBN 2227-8397. — Текст : электронный // Электронно-библиотечная система IPR BOOKS : [сайт]. <http://www.iprbookshop.ru/75376.html>

9. Перечень ресурсов сети Интернет, необходимых для освоения дисциплины

<http://www.psu.ru/elektronnye-resursy-dlya-psu> Электронные ресурсы для ПГНИУ

<http://window.edu.ru/> Единое окно доступа к образовательным ресурсам

10. Перечень информационных технологий, используемых при осуществлении образовательного процесса по дисциплине

Образовательный процесс по дисциплине **Интеллектуальный анализ данных** предполагает использование следующего программного обеспечения и информационных справочных систем:

- доступ в режиме on-line в Электронную библиотечную систему (ЭБС);
- доступ в электронную информационно-образовательную среду университета.

Необходимое лицензионное и (или) свободно распространяемое программное обеспечение:

- приложение позволяющее просматривать и воспроизводить медиаконтент PDF-файлов «Adobe Acrobat Reader DC»;
- офисный пакет приложений «LibreOffice».

При освоении материала и выполнении заданий по дисциплине рекомендуется использование материалов, размещенных в Личных кабинетах обучающихся ЕТИС ПГНИУ (student.psu.ru).

При организации дистанционной работы и проведении занятий в режиме онлайн могут использоваться:

- система видеоконференцсвязи на основе платформы BigBlueButton (<https://bigbluebutton.org/>).
- система LMS Moodle (<http://e-learn.psu.ru/>), которая поддерживает возможность использования текстовых материалов и презентаций, аудио- и видеоконтент, а так же тесты, проверяемые задания, задания для совместной работы.
- система тестирования Indigo (<https://indigotech.ru/>).

11. Описание материально-технической базы, необходимой для осуществления образовательного процесса по дисциплине

Для лекционных занятий требуется аудитория, оснащенная презентационной техникой (проектор, экран, компьютер/ноутбук) с соответствующим программным обеспечением, меловой (и) или маркерной доской.

Для проведения практических занятий - аудитория, оснащенная презентационной техникой (проектор, экран, компьютер/ноутбук) с соответствующим программным обеспечением, меловой (и) или маркерной доской.

Для групповых (индивидуальных) консультаций - аудитория, оснащенная презентационной техникой (проектор, экран, компьютер/ноутбук) с соответствующим программным обеспечением, меловой (и) или маркерной доской.

Для проведения текущего контроля - аудитория, оснащенная меловой (и) или маркерной доской.

Самостоятельная работа студентов: аудитория, оснащенная компьютерной техникой с возможностью подключения к сети «Интернет», с обеспеченным доступом в электронную информационно-образовательную среду университета, помещения Научной библиотеки ПГНИУ.

Помещения научной библиотеки ПГНИУ для обеспечения самостоятельной работы обучающихся:

1. Научно-библиографический отдел, корп.1, ауд. 142. Оборудован 3 персональными компьютера с доступом к локальной и глобальной компьютерным сетям.
2. Читальный зал гуманитарной литературы, корп. 2, ауд. 418. Оборудован 7 персональными

компьютерами с доступом к локальной и глобальной компьютерным сетям.

3. Читальный зал естественной литературы, корп.6, ауд. 107а. Оборудован 5 персональными компьютерами с доступом к локальной и глобальной компьютерным сетям.

4. Отдел иностранной литературы, корп.2 ауд. 207. Оборудован 1 персональным компьютером с доступом к локальной и глобальной компьютерным сетям.

5. Библиотека юридического факультета, корп.9, ауд. 4. Оборудована 11 персональными компьютерами с доступом к локальной и глобальной компьютерным сетям.

6. Читальный зал географического факультета, корп.8, ауд. 419. Оборудован 6 персональными компьютерами с доступом к локальной и глобальной компьютерным сетям.

Все компьютеры, установленные в помещениях научной библиотеки, оснащены следующим программным обеспечением:

Операционная система ALT Linux;

Офисный пакет Libreoffice.

Справочно-правовая система «КонсультантПлюс»

**Фонды оценочных средств для аттестации по дисциплине
Интеллектуальный анализ данных**

**Планируемые результаты обучения по дисциплине для формирования компетенции.
Индикаторы и критерии их оценивания**

ОПК.2

Способен совершенствовать и реализовывать новые математические методы решения прикладных задач

Индикатор	Планируемые результаты обучения	Критерии оценивания результатов обучения
<p>ОПК.2.1 Проводит анализ и обоснование применимости конкретного математического метода для решения прикладной задачи</p>	<p>Умение сочетать при решении конкретных практических задач какой-либо предметной области в нужном порядке такие инструменты ИАД как обезразмеривание, нормировка, корреляционный анализ, регрессионный анализ, факторный анализ. Владение современными основами статистической классификации: группировка, кластеризация и дискриминация.</p>	<p align="center">Неудовлетворител</p> <p>Не выполнены условия на "удовлетворительно"</p> <p align="center">Удовлетворительн</p> <p>Только с помощью преподавателя может</p> <ul style="list-style-type: none"> - классифицировать конкретную практическую задачу, связанную с интеллектуальной обработкой больших объемов статистической информации (например, задача управления, задача прогноза, задача «чёрного ящика» и так далее); - устанавливать взаимосвязь интеллектуального анализа данных (информации) с прикладной математической статистикой и её современными направлениями, такими как «бутстреп», интервальная статистика, робастность, непараметрика, статистика объектов нечисловой природы. <p align="center">Хорошо</p> <p>Самостоятельно умеет</p> <ul style="list-style-type: none"> - классифицировать конкретную практическую задачу, связанную с интеллектуальной обработкой больших объемов статистической информации (например, задача управления, задача прогноза, задача «чёрного ящика» и так далее), но может допустить несущественные неточности; - устанавливать взаимосвязь интеллектуального анализа данных (информации) с прикладной математической статистикой и её современными направлениями, такими как «бутстреп», интервальная статистика, робастность,

Индикатор	Планируемые результаты обучения	Критерии оценивания результатов обучения
		<p style="text-align: center;">Хорошо</p> <p>непараметрика, статистика объектов нечисловой природы.</p> <p style="text-align: center;">Отлично</p> <p>Самостоятельно умеет</p> <ul style="list-style-type: none"> - классифицировать конкретную практическую задачу, связанную с интеллектуальной обработкой больших объемов статистической информации (например, задача управления, задача прогноза, задача «чёрного ящика» и так далее); - устанавливать взаимосвязь интеллектуального анализа данных (информации) с прикладной математической статистикой и её современными направлениями, такими как «бутстреп», интервальная статистика, робастность, непараметрика, статистика объектов нечисловой природы.

Оценочные средства текущего контроля и промежуточной аттестации

Схема доставки : Базовая

Вид мероприятия промежуточной аттестации : Экзамен

Способ проведения мероприятия промежуточной аттестации : Оценка по дисциплине в рамках промежуточной аттестации определяется на основе баллов, набранных обучающимся на контрольных мероприятиях, проводимых в течение учебного периода.

Максимальное количество баллов : 100

Конвертация баллов в отметки

«отлично» - от 81 до 100

«хорошо» - от 61 до 80

«удовлетворительно» - от 50 до 60

«неудовлетворительно» / «незачтено» менее 50 балла

Компетенция (индикатор)	Мероприятие текущего контроля	Контролируемые элементы результатов обучения
Входной контроль	Введение Входное тестирование	Знание основных понятий и методов теории вероятности и математической статистики. Умении применять эти знания на практике.
ОПК.2.1 Проводит анализ и обоснование применимости конкретного математического метода для решения прикладной задачи	Снижение размерности задачи и классификация Защищаемое контрольное мероприятие	Знание способов нормировки. Умение определять признаки, связанные между собой Владение методом главных компонент.
ОПК.2.1 Проводит анализ и обоснование применимости конкретного математического метода для решения прикладной задачи	Анализ данных Защищаемое контрольное мероприятие	Знание и умение применять на практике методы факторного анализа. Знание способов классификации. Умение применять на практике методы кластерного анализа. Умение применять на практике методы дискрименантного анализа
ОПК.2.1 Проводит анализ и обоснование применимости конкретного математического метода для решения прикладной задачи	Итоговый контроль Итоговое контрольное мероприятие	Знание основных понятий, методов, изученных в дисциплине Умение определить по конкретной проблеме, какие методы и способы интеллектуального анализа данных следует применить

Спецификация мероприятий текущего контроля

Введение

Продолжительность проведения мероприятия промежуточной аттестации: **2 часа**
 Условия проведения мероприятия: **в часы самостоятельной работы**
 Максимальный балл, выставляемый за мероприятие промежуточной аттестации: **0**
 Проходной балл: **0**

Показатели оценивания	Баллы
Письменная работа, включающая 5 задач или вопросов. Каждая задача оценивается в 2 балла. За каждую ошибку снимается от 0.5 до 1 балла	10

Снижение размерности задачи и классификация

Продолжительность проведения мероприятия промежуточной аттестации: **2 часа**
 Условия проведения мероприятия: **в часы самостоятельной работы**
 Максимальный балл, выставляемый за мероприятие промежуточной аттестации: **20**
 Проходной балл: **10**

Показатели оценивания	Баллы
Владение методом главных компонент. За каждую ошибку снижается от 2 до 5 баллов	10
Умение определять признаки, связанные между собой. За каждую ошибку снижается от 1 до 2 баллов	7
Нормирование исходных данных. За каждую ошибку снижается 1 балл	3

Анализ данных

Продолжительность проведения мероприятия промежуточной аттестации: **2 часа**
 Условия проведения мероприятия: **в часы самостоятельной работы**
 Максимальный балл, выставляемый за мероприятие промежуточной аттестации: **40**
 Проходной балл: **20**

Показатели оценивания	Баллы
Анализ данных с помощью методов факторного анализа.	15
Анализ данных с помощью методов кластерного анализа. Умение применять на практике методы дискриминантного анализа	15
Анализ данных с помощью методов дискриминантного анализа	10
За каждый неточный ответ при защите работы снимается	2

Итоговый контроль

Продолжительность проведения мероприятия промежуточной аттестации: **4 часа**
 Условия проведения мероприятия: **в часы самостоятельной работы**
 Максимальный балл, выставляемый за мероприятие промежуточной аттестации: **40**
 Проходной балл: **20**

Показатели оценивания	Баллы
Полнота содержания сообщения	30
Четкость изложения сообщения	10