

Документ подписан простой электронной подписью

Информация о владельце:

ФИО: Хоружий Л.И. Ивановна

Должность: Директор института экономики и управления АПК

Дата подписания: 2023.08.30 09:33:14

Уникальный программный ключ:

1e90b132d9b040c67585160b015dddf2cb1e6a9



**МИНИСТЕРСТВО СЕЛЬСКОГО ХОЗЯЙСТВА
РОССИЙСКОЙ ФЕДЕРАЦИИ**

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ
**«РОССИЙСКИЙ ГОСУДАРСТВЕННЫЙ АГРАРНЫЙ УНИВЕРСИТЕТ –
МСХА имени К.А. ТИМИРЯЗЕВА»**

(ФГБОУ ВО РГАУ - МСХА имени К.А. Тимирязева)

**Институт экономики и управления АПК
Кафедра статистики и кибернетики**



УТВЕРЖДАЮ:

**Директор Института экономики и
управления АПК**

Хоружий Л.И.
Хоружий Л.И.
“30” августа 2023 г.

**РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ
Б1.В.05 Анализ больших данных (Big Data Analytics)**

для подготовки магистров

ФГОС ВО

Направление: 09.04.02 Информационные системы и технологии
Направленность: Науки о данных

Курс 2

Семестр 3

Форма обучения: очная

Год начала подготовки: 2023

Москва, 2023

Разработчики: Демичев В.В., канд. экон. наук, доцент,
Невзоров А.С., ассистент

(ФИО, ученая степень, ученое звание)


(подпись)

«28» августа 2023 г.

Рецензент: Быстренина И.Е., канд. пед. наук, доцент

(ФИО, ученая степень, ученое звание)


(подпись)

«28» августа 2023 г.

Программа составлена в соответствии с требованиями ФГОС ВО по направлению подготовки 09.04.02 Информационные системы и технологии, профессионального стандарта и учебного плана 2023 года начала подготовки

Программа обсуждена на заседании кафедры статистики и кибернетики протокол № 11 от «28» августа 2023 г.

И.о. зав. кафедрой Уколова А.В., канд. экон. наук, доцент

(ФИО, ученая степень, ученое звание)

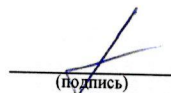

(подпись)

«28» августа 2023 г.

Согласовано:

Председатель учебно-методической
комиссии института экономики и управления АПК
Гупалова Т.Н., канд. экон. наук, доцент

(ФИО, ученая степень, ученое звание)


(подпись)

«28» августа 2023 г.

Зав. выпускающей кафедрой статистики и кибернетики
Уколова А.В., канд. экон. наук, доцент

(ФИО, ученая степень, ученое звание)


(подпись)

«28» августа 2023 г.

Заведующий отделом комплектования ЦНБ


(подпись)

СОДЕРЖАНИЕ

| | |
|---|-----------|
| АННОТАЦИЯ | 2 |
| 1. ЦЕЛЬ ОСВОЕНИЯ ДИСЦИПЛИНЫ..... | 2 |
| 2. МЕСТО ДИСЦИПЛИНЫ В УЧЕБНОМ ПРОЦЕССЕ | 2 |
| 3. ПЕРЕЧЕНЬ ПЛАНИРУЕМЫХ РЕЗУЛЬТАТОВ ОБУЧЕНИЯ ПО ДИСЦИПЛИНЕ, СООТНЕСЕННЫХ С ПЛАНИРУЕМЫМИ РЕЗУЛЬТАТАМИ ОСВОЕНИЯ ОБРАЗОВАТЕЛЬНОЙ ПРОГРАММЫ..... | 3 |
| 4. СТРУКТУРА И СОДЕРЖАНИЕ ДИСЦИПЛИНЫ | 7 |
| 4.1 РАСПРЕДЕЛЕНИЕ ТРУДОЁМКОСТИ ДИСЦИПЛИНЫ ПО ВИДАМ РАБОТ | 7 |
| ПО СЕМЕСТРАМ | 7 |
| 4.2 СОДЕРЖАНИЕ ДИСЦИПЛИНЫ..... | 7 |
| 4.3 ПРАКТИЧЕСКИЕ ЗАНЯТИЯ..... | 9 |
| 5. ОБРАЗОВАТЕЛЬНЫЕ ТЕХНОЛОГИИ | 11 |
| 6. ТЕКУЩИЙ КОНТРОЛЬ УСПЕВАЕМОСТИ И ПРОМЕЖУТОЧНАЯ АТТЕСТАЦИЯ ПО ИТОГАМ ОСВОЕНИЯ ДИСЦИПЛИНЫ | 11 |
| 6.1. ТИПОВЫЕ КОНТРОЛЬНЫЕ ЗАДАНИЯ ИЛИ ИНЫЕ МАТЕРИАЛЫ, НЕОБХОДИМЫЕ ДЛЯ ОЦЕНКИ ЗНАНИЙ, УМЕНИЙ И НАВЫКОВ И (ИЛИ) ОПЫТА ДЕЯТЕЛЬНОСТИ | 11 |
| 6.2. ОПИСАНИЕ ПОКАЗАТЕЛЕЙ И КРИТЕРИЕВ КОНТРОЛЯ УСПЕВАЕМОСТИ, ОПИСАНИЕ ШКАЛ ОЦЕНИВАНИЯ | 11 |
| 7. УЧЕБНО-МЕТОДИЧЕСКОЕ И ИНФОРМАЦИОННОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ..... | 12 |
| 7.1 ОСНОВНАЯ ЛИТЕРАТУРА | 14 |
| 7.2 ДОПОЛНИТЕЛЬНАЯ ЛИТЕРАТУРА..... | 14 |
| 7.3 МЕТОДИЧЕСКИЕ УКАЗАНИЯ, РЕКОМЕНДАЦИИ И ДРУГИЕ МАТЕРИАЛЫ К ЗАНЯТИЯМ..... | 14 |
| 8. ПЕРЕЧЕНЬ РЕСУРСОВ ИНФОРМАЦИОННО-ТЕЛЕКОММУНИКАЦИОННОЙ СЕТИ «ИНТЕРНЕТ», НЕОБХОДИМЫХ ДЛЯ ОСВОЕНИЯ ДИСЦИПЛИНЫ (МОДУЛЯ)..... | 14 |
| 9. ПЕРЕЧЕНЬ ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ И ИНФОРМАЦИОННЫХ СПРАВОЧНЫХ СИСТЕМ..... | 15 |
| 10. ОПИСАНИЕ МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЙ БАЗЫ, НЕОБХОДИМОЙ ДЛЯ ОСУЩЕСТВЛЕНИЯ ОБРАЗОВАТЕЛЬНОГО ПРОЦЕССА ПО ДИСЦИПЛИНЕ..... | 16 |
| 11. МЕТОДИЧЕСКИЕ РЕКОМЕНДАЦИИ СТУДЕНТАМ ПО ОСВОЕНИЮ ДИСЦИПЛИНЫ..... | 16 |
| 12. МЕТОДИЧЕСКИЕ РЕКОМЕНДАЦИИ ПРЕПОДАВАТЕЛЯМ ПО ОРГАНИЗАЦИИ ОБУЧЕНИЯ ПО ДИСЦИПЛИНЕ..... | 17 |

АННОТАЦИЯ

рабочей программы учебной дисциплины Б1.В.05 Анализ больших данных (Big Data Analytics) для подготовки магистров по направлению 09.04.02 Информационные системы и технологии, направленности Науки о данных

Цель освоения дисциплины: по окончании изучения дисциплины студент должен знать: теоретические и прикладные основы анализа больших данных, технологии хранения и обработки больших данных в организации: базы данных, хранилища данных, распределенная и параллельная обработка данных. По окончании изучения дисциплины студент должен уметь: использовать инструментальные средства для извлечения, преобразования, хранения и обработки данных из разнородных источников, в том числе в режиме реального времени; проводить очистку, интеграцию, преобразование и анализ больших объемов данных. Также по окончании изучения дисциплины студент должен владеть: навыками получения и обработки больших данных с поддержкой работы в режиме реального времени.

Место дисциплины в учебном плане: дисциплина включена в часть учебного плана по направлению подготовки 09.04.02 Информационные системы и технологии, формируемую участниками образовательных отношений.

Требования к результатам освоения дисциплины: в результате освоения дисциплины формируются следующие компетенции (индикаторы): ПКос-1 (ПКос-1.1; ПКос-1.2; ПКос-1.3); ПКос-2 (ПКос-2.1; ПКос-2.2; ПКос-2.3); ПКос-3 (ПКос-3.1; ПКос-3.2; ПКос-3.3).

Краткое содержание дисциплины:

Понятие большие данные. Источники больших данных в сельском хозяйстве. Возможности и трудности использования больших данных в сельском хозяйстве. Формат больших данных. Большие данные и Python. Методы обработки больших данных. Основные этапы анализа больших данных. Постановка цели исследования. Сбор данных. Подготовка и исследование данных. Моделирование данных. Визуализация и отображение результатов анализа больших данных. Экосистема больших данных. Анализ больших данных с применением статистических методов. Регрессия. Кластеризация. Классификация. Важнейшие библиотеки Python для применения статистических методов. Процесс построения моделей.

Общая трудоемкость дисциплины: 180 / 5 (часы/зач. ед.)

Промежуточный контроль: экзамен, защита КР/КП.

1. Цель освоения дисциплины

Целью дисциплины «Анализ больших данных (Big Data Analytics)» является освоение студентами теоретических и практических знаний и приобретение умений и навыков в области статистических методов, применяемых в анализе больших данных, а также формирование знаний в сфере анализа данных. Также целью данной дисциплины является приобретение умений и навыков применения на практике статистических алгоритмов анализа больших данных; по результатам изучения дисциплины студент должен уметь применять различные статистические методы и решать возникающие перед ним практические задачи.

2. Место дисциплины в учебном процессе

Дисциплина «Анализ больших данных (Big Data Analytics)» включена в часть учебного плана, формируемую участниками образовательных отношений. Дисциплина «Анализ больших данных (Big Data Analytics)» реализуется в соответствии с требованиями ФГОС ВО, ОПОП ВО и Учебного плана по направлению 09.04.02 Информационные системы и технологии.

Дисциплина «Анализ больших данных (Big Data Analytics)» изучается на

третьем курсе образовательного цикла.

Предшествующими курсами, включенными в учебный план, на которых непосредственно базируются дисциплина «Анализ больших данных (Big Data Analytics)», являются «Модели информационных процессов и систем», «Наука о данных (Data Science)», «Инструменты Data Science в R, Python, SQL».

Дисциплина «Анализ больших данных (Big Data Analytics)» может быть использована при написании выпускной квалификационной работы.

Особенностью дисциплины является реализация статистических методов анализа больших данных средствами языка программирования Python 3.

Рабочая программа дисциплины «Анализ больших данных (Big Data Analytics)» для инвалидов и лиц с ограниченными возможностями здоровья разрабатывается индивидуально с учетом особенностей психофизического развития, индивидуальных возможностей и состояния здоровья таких обучающихся.

3. Перечень планируемых результатов обучения по дисциплине, соотнесенных с планируемыми результатами освоения образовательной программы

Изучение дисциплины направлено на формирование у обучающихся профессиональных компетенций, определяемых самостоятельно (ПКос), представленных в таблице 1.

Таблица 1

Требования к результатам освоения учебной дисциплины «Анализ больших данных (Big Data Analytics)»

| № п/п | Код компетенции | Содержание компетенции (или её части) | Индикаторы компетенций | В результате изучения учебной дисциплины обучающиеся должны: | | |
|-------|-----------------|--|------------------------|--|--|--|
| | | | | знать | уметь | владеть |
| 1. | ПКос-1 | Способен составлять отчеты о проделанной работе, обзоры, готовить публикации; разрабатывать научно-методические и учебно-методические материалы, обеспечивающие реализацию учебных курсов дисциплин (модулей) программ бакалавриата и (или) ДПП в области знаний профессиональной деятельности; организовывать научно-исследовательскую, преподавать учебные курсы, дисциплины (модули) или отдельные виды учебных занятий по программам бакалавриата и/или ДПП с учетом передового отечественного и зарубежного опыта | ПКос-1.1 | основы работы с большими данными: понимание основных характеристик больших данных, их источников и проблем, связанных с их обработкой, знать методы сбора и предварительной обработки данных: умение выбирать и применять подходящие методы сбора, очистки, преобразования и интеграции данных, статистические алгоритмы для работы с текстовыми данными | - | - |
| | | | ПКос-1.2 | - | выбирать и применять подходящие методы сбора, очистки, преобразования и интеграции данных, инструменты и технологии анализа больших данных: знание и умение использовать различные инструменты и технологии для анализа больших объемов данных, включая языки программирования, базы данных, библиотеки и программные средства для работы с большими данными | - |
| | | | ПКос-1.3 | - | - | принципами и техниками машинного обучения: основы машинного обучения, его применения в анализе больших данных и умение выбирать наиболее подходящие модели ма- |

| | | | | | | |
|----|--------|---|----------|---|---|--|
| | | | | | | шинного обучения для решения конкретных задач. |
| 2. | ПКос-2 | Способен применять современные средства анализа данных и бизнес-аналитики в профессиональной деятельности | ПКос-2.1 | инструментальные средства бизнес-аналитики и пакеты прикладных статистических программ - statistica, eviews | - | - |
| | | | ПКос-2.2 | - | осуществлять сбор, обработку, анализ массовые данных, строить прогнозы с использованием современных инструментов бизнес-аналитики и пакетов прикладных статистических программ - statistica, eviews | - |
| | | | ПКос-2.3 | - | - | сбором, обработкой, анализом массовых данных, в т.ч. больших данных, эконометрического моделирования и прогнозирования с использованием современных средств бизнес-аналитики и пакетов прикладных статистических программ - statistica, eviews |
| 3. | ПКос-3 | Способен совершенствовать и разрабатывать новые методы, модели, алгоритмы, технологии и инструментальные средства работы с данными, в т.ч. большими данными | ПКос-3.1 | методы науки о данных, в т.ч. методы машинного обучения, обработки и визуализации больших данных; состояние и перспективы развития науки о данных, используемого при обработке данных программного инструментария - statistica, eviews и в среде программирования Anaconda на языке python; потребности в совершенствовании и разработке новых методов, технологий и инструменталь- | - | - |

| | | | | | | |
|--|--|--|----------|---|---|---|
| | | | | ных средств для работы с данными, в т.ч. большими; область применения науки о данных в сельском хозяйстве | | |
| | | | ПКос-3.2 | - | определять перспективную тематику научно исследовательских работ в области совершенствования и разработки новых методов, моделей, алгоритмов, технологий и инструментальных средств работы с данными; планировать и проводить аналитические и научные исследования по тематике информационных технологий, применяемых в науке о данных на основе работы с пакетами прикладных программ - statistica, eviews и в среде программирования Anaconda на языке python | - |
| | | | ПКос-3.3 | - | - | разработки новых методов, моделей, алгоритмов, технологий и инструментальных средств работы с данными на основе анализа потребностей и передового зарубежного и отечественного опыта; планирования состава и содержания, согласование перечня научно-исследовательских работ в профессиональной деятельности применяя в работе с статистическими программами - statistica, eviews и в среде программирования Anaconda на языке python |

4. Структура и содержание дисциплины

4.1 Распределение трудоёмкости дисциплины по видам работ по семестрам

Общая трудоёмкость дисциплины составляет 5 зач.ед. (180 часов), их распределение по видам работ и семестрам представлено в таблице 2.

Таблица 2

| Вид учебной работы | Трудоёмкость | |
|--|-----------------------|----------------------------|
| | час. всего/* | в т.ч. по семестрам № 3 /* |
| Общая трудоёмкость дисциплины по учебному плану | 180 | 180 |
| 1. Контактная работа: | 53,4 | 53,4 |
| Аудиторная работа | 53,4 | 53,4 |
| <i>в том числе:</i> | | |
| <i>лекции (Л)</i> | 16 | 16 |
| <i>практические занятия (ПЗ)</i> | 32 | 32 |
| <i>курсовая работа (проект) (КР/КП) (консультация, защита)</i> | 3 | 3 |
| <i>консультации перед экзаменом</i> | 2 | 2 |
| <i>контактная работа на промежуточном контроле (КРА)</i> | 0,4 | 0,4 |
| 2. Самостоятельная работа (СРС) | 102 | 102 |
| <i>самостоятельное изучение разделов, самоподготовка (проработка и повторение лекционного материала и материала учебников и учебных пособий, подготовка к практическим занятиям)</i> | 77,4 | 77,4 |
| <i>Подготовка к экзамену (контроль)</i> | 24,6 | 24,6 |
| Вид промежуточного контроля: | Экзамен, защита КР/КП | |

4.2 Содержание дисциплины

Таблица 3

Тематический план учебной дисциплины

| Наименование разделов и тем дисциплин (укрупнённо) | Всего | Аудиторная работа | | | Внеаудиторная работа СР |
|---|-------|-------------------|---------------|-----|-------------------------|
| | | Л | ПЗ всего/* | ПКР | |
| Тема 1 «Классификация» | 30 | 4 | 8 | - | 18 |
| Тема 2 «Статистические алгоритмы для работы с графовыми данными» | 20 | 2 | 4 | - | 14 |
| Тема 3 «Статистические алгоритмы для работы с текстовыми данными» | 20 | 2 | 4 | - | 14 |
| Тема 4 «Байесовские и априорные вероятности» | 20 | 2 | 4 | - | 14 |
| Тема 5 «Оценка параметров байесовских моделей» | 20 | 2 | 4 | - | 14 |
| Тема 6 «Проверка гипотез в байесовской статистике» | 20 | 2 | 4 | - | 14 |

| Наименование разделов и тем дисциплин (укрупнённо) | Всего | Аудиторная работа | | | Внеаудиторная работа СР |
|--|------------|-------------------|---------------|------------|-------------------------|
| | | Л | ПЗ всего/* | ПКР | |
| Тема 7 «А/В тестирование в байесовской статистике» | 44,6 | 2 | 4 | - | 38,6 |
| курсовая работа (проект) (КР/КП) (консультация, защита) | - | - | - | 3 | - |
| контактная работа на промежуточном контроле (КРА) | 0,4 | - | - | 0,4 | - |
| Консультации перед экзаменом | 2 | - | - | 2 | - |
| Итого по дисциплине | 180 | 16 | 32 | 5,4 | 126,6 |

Тема 1. Классификация.

Понятие классификации. Виды методов классификации. Примеры применения классификации в анализе больших данных. Построение моделей классификации и оценка их качества.

Тема 2. Статистические алгоритмы для работы с графовыми данными.

Понятие связанные и графовые данные. Методы работы с графовыми данными. Примеры использования графовых данных.

Тема 3. Статистические алгоритмы для работы с текстовыми данными.

Понятие и примеры применения глубокого анализа текста. Методы глубокого анализа текста. Дерево принятия решений. Применение библиотеки NLTK для реализации статистических методов глубокого анализа текста.

Тема 4. Байесовские и априорные вероятности.

Понятие вероятности в байесовской статистике. Теорема Байеса. Априорная и апостериорная вероятности и правдоподобие в теории Байеса. Распределение вероятностей в теории Байеса.

Тема 5. Оценка параметров байесовских моделей.

Усреднение и оценка параметров. Измерение разброса данных. Нормальное распределение и инструменты оценки параметров.

Тема 6. Проверка гипотез в байесовской статистике.

Коэффициент Байеса и апостериорный шанс. Создание байесовских А/В тестов. Понятие «сумеречной зоны».

Тема 7. А/В тестирование в байесовской статистике.

Понятие А/В тестирования. Сравнение частотного и байесовского подхода к А/В тестированию. Суть алгоритма байесовского тестирования. Байесовское тестирование на Python.

4.3 Практические занятия

Таблица 4

Содержание практических занятий и контрольные мероприятия

| Название раздела, темы | № и название лекций/ практических занятий | Формируемые компетенции (индикаторы) | Вид контрольного мероприятия | Кол-во часов |
|---|--|--|------------------------------|--------------|
| Тема 1 «Классификация» | Лекция 1 «Классификация» | ПКос-1.1; ПКос-1.2; ПКос-1.3; | - | 4 |
| | Практическая работа 1 «Применение методов классификации» | ПКос-1.1; ПКос-1.2; ПКос-1.3; ПКос-3.1; ПКос-3.2; ПКос-3.3 | Устный опрос | 8 |
| Тема 2 «Статистические алгоритмы для работы с графовыми данными» | Лекция 2 «Статистические алгоритмы для работы с графовыми данными» | ПКос-1.1; ПКос-1.2; ПКос-1.3; ПКос-2.1; ПКос-2.2 | - | 2 |
| | Практическая работа 2 «Анализ графовых данных» | ПКос-1.1; ПКос-2.2; ПКос-2.3; ПКос-3.1; ПКос-3.2; ПКос-3.3 | Устный опрос | 4 |
| Тема 3 «Статистические алгоритмы для работы с текстовыми данными» | Лекция 3 «Статистические алгоритмы для работы с текстовыми данными» | ПКос-1.1; ПКос-1.2; ПКос-1.3 | - | 2 |
| | Практическая работа 3 «Статистические алгоритмы для работы с текстовыми данными» | ПКос-1.1; ПКос-1.2; ПКос-1.3; ПКос-2.1; ПКос-2.2; ПКос-2.3; ПКос-3.1; ПКос-3.2; ПКос-3.3 | Устный опрос | 4 |
| Тема 4 «Байесовские и априорные вероятности» | Лекция 4 «Байесовские и априорные вероятности» | ПКос-1.1; ПКос-3.1; ПКос-3.2; ПКос-3.3 | - | 2 |
| | Практическая работа 4 «Расчет и определение байесовской вероятности» | ПКос-1.1; ПКос-1.2; ПКос-1.3; ПКос-3.3 | Устный опрос | 4 |
| Тема 5 «Оценка параметров байесовских моделей» | Лекция 5 «Оценка параметров байесовских моделей» | ПКос-1.1; ПКос-2.2; ПКос-2.3; ПКос-3.1; | | 2 |

| Название раздела, темы | № и название лекций/ практических занятий | Формируемые компетенции (индикаторы) | Вид контрольного мероприятия | Кол-во часов |
|--|---|--|------------------------------|--------------|
| | | ПКос-3.2; ПКос-3.3 | | |
| | Практическая работа 5 «Применение инструментов оценки параметров» | ПКос-2.3; ПКос-3.1; ПКос-3.2; ПКос-3.3 | Устный опрос | 4 |
| Тема 6 «Проверка гипотез в байесовской статистике» | Лекция 6 «Проверка гипотез в байесовской статистике» | ПКос-1.1; ПКос-1.2; ПКос-1.3; ПКос-2.1 | | 2 |
| | Практическая работа 6 «Формулировка и проверка гипотез» | ПКос-1.1; ПКос-1.2; ПКос-1.3; ПКос-2.2; ПКос-2.3; ПКос-3.2; ПКос-3.3 | Устный опрос | 4 |
| Тема 7 «А/В тестирование в байесовской статистике» | Лекция 7 «А/В тестирование в байесовской статистике» | ПКос-1.1; ПКос-1.2; ПКос-2.3; ПКос-3.1; ПКос-3.2; ПКос-3.3 | | 2 |
| | Практическая работа 7 «Создание байесовских А/В-тестов» | ПКос-1.1; ПКос-1.2; ПКос-1.3; ПКос-2.1; ПКос-2.2 | Устный опрос | 4 |

Таблица 5

Перечень вопросов для самостоятельного изучения дисциплины

| № п/п | Название раздела, темы | Перечень рассматриваемых вопросов для самостоятельного изучения |
|-------|---|--|
| 1. | Тема 1 «Классификация» | Методы классификации. Примеры применения классификации. (ПКос-1.1; ПКос-1.2; ПКос-1.3; ПКос-2.1; ПКос-2.2; ПКос-2.3; ПКос-3.1; ПКос-3.2; ПКос-3.3). |
| 2. | Тема 2 «Статистические алгоритмы для работы с графовыми данными» | Теория графов и пространственные статистические методы. (ПКос-1.1; ПКос-1.2; ПКос-1.3; ПКос-2.1; ПКос-2.2; ПКос-2.3; ПКос-3.1; ПКос-3.2; ПКос-3.3). |
| 3. | Тема 3 «Статистические алгоритмы для работы с текстовыми данными» | Многомерные статистические методы в контексте статистического анализа текста. (ПКос-1.1; ПКос-1.2; ПКос-1.3; ПКос-2.1; ПКос-2.2; ПКос-2.3; ПКос-3.1; ПКос-3.2; ПКос-3.3). |
| 4. | Тема 4 «Байесовские и априорные вероятности» | Распределение вероятностей в теории Байеса. (ПКос-1.1; ПКос-1.2; ПКос-1.3; ПКос-2.1; ПКос-2.2; ПКос-2.3; ПКос-3.1; ПКос-3.2; ПКос-3.3). |

| № п/п | Название раздела, темы | Перечень рассматриваемых вопросов для самостоятельного изучения |
|-------|--|--|
| 5. | Тема 5 «Оценка параметров байесовских моделей» | Нормальное распределение в теории Байеса. (ПКос-1.1; ПКос-1.2; ПКос-1.3; ПКос-2.1; ПКос-2.2; ПКос-2.3; ПКос-3.1; ПКос-3.2; ПКос-3.3) |
| 6. | Тема 6 «Проверка гипотез в байесовской статистике» | Примеры применения байесовских А/В-тестов. (ПКос-1.1; ПКос-1.2; ПКос-1.3; ПКос-2.1; ПКос-2.2; ПКос-2.3; ПКос-3.1; ПКос-3.2; ПКос-3.3) |
| 7. | Тема 7 «А/В тестирование в байесовской статистике» | Байесовское тестирование на Python. (ПКос-1.1; ПКос-1.2; ПКос-1.3; ПКос-2.1; ПКос-2.2; ПКос-2.3; ПКос-3.1; ПКос-3.2; ПКос-3.3) |

5. Образовательные технологии

Таблица 6

Применение активных и интерактивных образовательных технологий

| № п/п | Тема и форма занятия | | Наименование используемых активных и интерактивных образовательных технологий (форм обучения) |
|-------|---|----|---|
| 1. | Тема 1 «Классификация» | ПЗ | Разбор конкретных ситуаций |
| 2. | Тема 2 «Статистические алгоритмы для работы с графовыми данными» | ПЗ | Мозговой штурм |
| 3. | Тема 3 «Статистические алгоритмы для работы с текстовыми данными» | ПЗ | Разбор конкретных ситуаций |
| 4. | Тема 4 «Байесовские и априорные вероятности» | ПЗ | Разбор конкретных ситуаций |

6. Текущий контроль успеваемости и промежуточная аттестация по итогам освоения дисциплины

6.1. Типовые контрольные задания или иные материалы, необходимые для оценки знаний, умений и навыков и (или) опыта деятельности

Вопросы для подготовки к устным опросам

Тема 1 «Классификация»

1. Понятие классификации.
2. Виды методов классификации.
3. Примеры применения классификации в анализе больших данных.
4. Построение моделей классификации и оценка их качества.

Тема 2 «Статистические алгоритмы для работы с графовыми данными»

1. Понятие связанные и графовые данные.
2. Методы работы с графовыми данными.
3. Примеры использования графовых данных.

Тема 3 «Статистические алгоритмы для работы с текстовыми данными»

1. Понятие и примеры применения глубокого анализа текста.
2. Методы глубокого анализа текста.

3. Дерево принятия решений.
4. Применение библиотеки NLTK для реализации статистических методов глубокого анализа текста.
5. Многомерные статистические методы в контексте статистического анализа текста.

Тема 4 «Байесовские и априорные вероятности»

1. Понятие вероятности в байесовской статистике.
2. Теорема Байеса.
3. Априорная и апостериорная вероятности и правдоподобие в теории Байеса.
4. Распределение вероятностей в теории Байеса.

Тема 5 «Оценка параметров байесовских моделей»

1. Усреднение и оценка параметров.
2. Измерение разброса данных.
3. Нормальное распределение и инструменты оценки параметров.

Тема 6 «Проверка гипотез в байесовской статистике»

1. Коэффициент Байеса и апостериорный шанс.
2. Создание байесовских А/В-тестов.
3. Понятие «сумеречной зоны».

Тема 7 «А/В тестирование в байесовской статистике»

1. Понятие А/В тестирования.
2. Сравнение частотного и байесовского подхода к А/В тестированию.
3. Суть алгоритма байесовского тестирования.
4. Байесовское тестирование на Python.

6.2. Описание показателей и критериев контроля успеваемости, описание шкал оценивания

Оценка знаний также ведется на основе рейтинговой оценки студента, которая формируется как сумма баллов за участие в устном опросе на практических занятиях. Студент допускается к сдаче экзамена при достижении рейтинга 60%.

Максимальная оценка за участие в опросе также 10 баллов.

9 баллов - ставится при наличии незначительных неточностей в ответе.

8 баллов - при наличии негрубых ошибок в ответе, которые не привели к ложным выводам и неверному пониманию сути вопроса.

7 баллов - сделаны неверные выводы по применяемым методам, при этом общее понимание применяемых методов не искажено.

6-5 баллов - нарушена логика в понимании применяемых методов.

Количество баллов складывается следующим образом: 7 устных вопросов * 10 (максимальное количество) баллов = 70 баллов (максимально возможное количество набранных баллов). В процентах (количество набранных баллов / максимально возможная сумма баллов) * 100.

Участие в интерактивных занятиях может быть зачтено активным студентам как участие в опросе по теме, на котором применялись интерактивные технологии.

На экзамене студент может получить максимальное количество баллов равное 100. Далее итоговая оценка определяется следующим образом. Если текущий рейтинг студента составляет 70 баллов, а на экзамене студент получил 100 баллов («отлично»), то итоговая оценка 100 баллов + 70 баллов («отлично»).

Промежуточный контроль – экзамен.

Таблица 7

| Шкала оценивания (средний балл) | Экзамен |
|---------------------------------|---------------------|
| > 145 | Отлично |
| 128-145 | Хорошо |
| 103-127 | Удовлетворительно |
| 0-102 | Неудовлетворительно |

Положительными оценками, при получении которых дисциплина засчитывается в качестве пройденной, являются оценки «удовлетворительно», «хорошо» и «отлично».

Если получена оценка «неудовлетворительно» по дисциплине, то необходимо, после консультации с преподавателем, в течение 10 календарных дней следующего семестра подготовить ответы на ряд вопросов, предусмотренных программой обучения, и представить результаты этих ответов преподавателю.

Критерии оценивания результатов обучения

Таблица 8

| Оценка | Критерии оценивания |
|---|---|
| Высокий уровень «5» (отлично) | оценку «отлично» заслуживает студент, освоивший знания, умения, компетенции и теоретический материал без пробелов; выполнивший все задания, предусмотренные учебным планом на высоком качественном уровне; практические навыки профессионального применения освоенных знаний сформированы. Компетенции, закреплённые за дисциплиной, сформированы на уровне – высокий. |
| Средний уровень «4» (хорошо) | оценку «хорошо» заслуживает студент, практически полностью освоивший знания, умения, компетенции и теоретический материал, учебные задания не оценены максимальным числом баллов, в основном сформировал практические навыки. Компетенции, закреплённые за дисциплиной, сформированы на уровне – хороший (средний). |
| Пороговый уровень «3» (удовлетворительно) | оценку «удовлетворительно» заслуживает студент, частично с пробелами освоивший знания, умения, компетенции и теоретический материал, многие учебные задания либо не выполнил, либо они оценены числом баллов близким к минимальному, некоторые практические навыки не сформированы. Компетенции, закреплённые за дисциплиной, сформированы на уровне – достаточный. |

| | |
|---|---|
| Минимальный уровень «2» (неудовлетворительно) | оценку «неудовлетворительно» заслуживает студент, не освоивший знания, умения, компетенции и теоретический материал, учебные задания не выполнил, практические навыки не сформированы. Компетенции, закреплённые за дисциплиной, не сформированы. |
|---|---|

7. Учебно-методическое и информационное обеспечение дисциплины

7.1 Основная литература

1. Анализ данных : учебник для вузов / В. С. Мхитарян [и др.] ; под редакцией В. С. Мхитаряна. — Москва : Издательство Юрайт, 2023. — 490 с. — (Высшее образование). — ISBN 978-5-534-00616-2. — Текст : электронный // Образовательная платформа Юрайт [сайт]. — URL: <https://urait.ru/bcode/511020> (дата обращения: 16.08.2023).

2. Миркин, Б. Г. Введение в анализ данных : учебник и практикум / Б. Г. Миркин. — Москва : Издательство Юрайт, 2023. — 174 с. — (Высшее образование). — ISBN 978-5-9916-5009-0. — Текст : электронный // Образовательная платформа Юрайт [сайт]. — URL: <https://urait.ru/bcode/511121> (дата обращения: 16.08.2023).

3. Статистика : учебник для вузов / под редакцией И. И. Елисеевой. — 3-е изд., перераб. и доп. — Москва : Издательство Юрайт, 2023. — 361 с. — (Высшее образование). — ISBN 978-5-534-04082-1. — Текст : электронный // Образовательная платформа Юрайт [сайт]. — URL: <https://urait.ru/bcode/510524> (дата обращения: 16.08.2023).

7.2 Дополнительная литература

1. Эконометрика : учебник для вузов / И. И. Елисеева [и др.] ; под редакцией И. И. Елисеевой. — Москва : Издательство Юрайт, 2023. — 449 с. — (Высшее образование). — ISBN 978-5-534-00313-0. — Текст : электронный // Образовательная платформа Юрайт [сайт]. — URL: <https://urait.ru/bcode/510472> (дата обращения: 16.08.2023).

2. Дудин, М. Н. Статистика : учебник и практикум для вузов / М. Н. Дудин, Н. В. Лясников, М. Л. Лезина. — Москва : Издательство Юрайт, 2023. — 374 с. — (Высшее образование). — ISBN 978-5-9916-8908-3. — Текст : электронный // Образовательная платформа Юрайт [сайт]. — URL: <https://urait.ru/bcode/512310> (дата обращения: 16.08.2023).

3. Воронов, М. В. Науки о данных : учебник и практикум для вузов / М. В. Воронов, В. И. Пименов, И. А. Небаев. — Москва : Издательство Юрайт, 2023. — 256 с. — (Высшее образование). — ISBN 978-5-534-14916-6. — Текст : электронный // Образовательная платформа Юрайт [сайт]. — URL: <https://urait.ru/bcode/519916> (дата обращения: 16.08.2023).

7.3 Методические указания, рекомендации и другие материалы к занятиям

1. Харитонова, Анна Евгеньевна. Анализ и прогнозирование с использованием пакетов прикладных программ: практикум / А. Е. Харитонова; Российский государственный аграрный университет - МСХА имени К. А. Тимирязева (Москва), Факультет экономики и финансов, Кафедра статистики и

эконометрика. — Электрон. текстовые дан. — Москва: Росинформагротех, 2017. — 107 с.

2. Дашиева, Баярма Шагдаровна. Эконометрика: методические указания / Б. Ш. Дашиева, В. В. Демичев; Российский государственный аграрный университет - МСХА имени К. А. Тимирязева (Москва), Факультет экономики и финансов, Кафедра статистики и эконометрика. — Электрон. текстовые дан. — Москва: РГАУ-МСХА им. К. А. Тимирязева, 2018. — 32 с.

8. Перечень ресурсов информационно-телекоммуникационной сети «Интернет», необходимых для освоения дисциплины

1. Официальный сайт Python. URL: <https://www.python.org/> (открытый доступ)
2. Официальный сайт дистрибутива языков программирования Python и R Anaconda. URL: <https://www.anaconda.com/> (открытый доступ)
3. Официальный сайт Росстата. URL: <https://rosstat.gov.ru/> (открытый доступ).

9. Перечень программного обеспечения и информационных справочных систем

Таблица 9

Перечень программного обеспечения

| № п/п | Наименование раздела учебной дисциплины | Наименование программы | Тип программы | Автор | Год разработки |
|-------|---|-------------------------|---|-------------------------|----------------|
| 1 | Тема 1 «Классификация» | Excel/ Word/Anaconda | Расчетная/система управления пакетами и дистрибутив | Microsoft/Anaconda Inc. | 2007/2012 |
| 2 | Тема 2 «Статистические алгоритмы для работы с графовыми данными» | Excel/ Word/Anaconda | Расчетная/система управления пакетами и дистрибутив | Microsoft/Anaconda Inc. | 2007/2012 |
| 3 | Тема 3 «Статистические алгоритмы для работы с текстовыми данными» | Excel/ Word/Anaconda | Расчетная/система управления пакетами и дистрибутив | Microsoft/Anaconda Inc. | 2007/2012 |
| 4 | Тема 4 «Байесовские и априорные вероятности» | Excel/ Word/Anaconda | Расчетная/система управления пакетами и дистрибутив | Microsoft/Anaconda Inc. | 2007/2012 |

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

Таблица 10

Сведения об обеспеченности специализированными аудиториями, кабинетами, лабораториями

| Наименование специальных* помещений и помещений для самостоятельной работы (№ учебного корпуса, № аудитории) | Оснащенность специальных помещений и помещений для самостоятельной работы** |
|---|---|
| 1 | 2 |
| учебная аудитория для проведения занятий лекционного типа, учебная аудитория для проведения практических занятий, групповых и индивидуальных консультаций, текущего контроля и промежуточной аттестации (Учебный корпус № 2, аудитория № 102) | Мультимедиа: Проектор ACERX118 черный [mr.jpz 11.001], компьютер конфигурации: CelD-1800/512/80/DVD-R |
| учебная аудитория для проведения занятий лекционного типа, учебная аудитория для проведения занятий, групповых и индивидуальных консультаций, текущего контроля и промежуточной аттестации (Учебный корпус № 2, аудитория № 101) | Мультимедиа: монитор инв.№ 34799/3, экран настенный с электроприводом инв.№ 35641/7, системный блок инв.№ 558788/135, доска меловая, стулья-87, столы-50 |
| учебная аудитория для проведения занятий лекционного типа, учебная аудитория для проведения занятий, групповых и индивидуальных консультаций, текущего контроля и промежуточной аттестации (Учебный корпус № 2, аудитория № 202) | Мультимедиа: Системный блок 1 шт. (Инв. 556563), Монитор 1 шт. (Инв. 34799/4), парты 36 шт., скамья 36 шт., доска меловая 1 шт., экран для проектора настенно потолочный. |
| учебная аудитория для проведения практических занятий, групповых и индивидуальных консультаций, текущего контроля и промежуточной аттестации (Учебный корпус № 2, аудитория № 302) | 10 компьютеров конфигурации: INTELCorei3-2100/4096 Mb/500Gb/DVD-RW, MSWord, MSExcel, пакеты прикладных программ: STATA, R, EViews, Statistica, доступ к сети Internet, справочной системы КонсультантПлюс |
| Центральная научная библиотека имени Н.И. Железнова | Читальные залы библиотеки |
| Студенческое общежитие | Комната для самоподготовки |

11. Методические рекомендации студентам по освоению дисциплины

Приступая к изучению дисциплины «Анализ больших данных (Big Data Analytics)», студенты должны ознакомиться с учебной программой, учебной, научной и методической литературой, имеющейся в библиотеке РГАУ-МСХА им. К.А. Тимирязева, получить в библиотеке рекомендованные учебники и учебно-методические пособия, завести новую тетрадь для работы с первоисточниками.

В ходе занятий вести конспектирование учебного материала. Обращать внимание на категории, формулировки, раскрывающие содержание тех или иных явлений и процессов, научные выводы и практические рекомендации. За-

давать преподавателю уточняющие вопросы с целью уяснения теоретических положений, разрешения спорных ситуаций.

В ходе подготовки к практическим занятиям изучить основную литературу, ознакомиться с дополнительной литературой в соответствии с поставленной задачей. При этом учесть рекомендации преподавателя и требования учебной программы. Необходимо дорабатывать свой конспект, делая в нем соответствующие записи из литературы, рекомендованной преподавателем и предусмотренной учебной программой.

При подготовке к зачету/экзамену (в конце семестра) повторять пройденный материал в строгом соответствии с учебной программой. Использовать конспекты и литературу, рекомендованную преподавателем. Обратит особое внимание на темы учебных занятий, пропущенных студентом по разным причинам. При необходимости обратиться за консультацией и методической помощью к преподавателю.

Виды и формы отработки пропущенных занятий

Студент, пропустивший занятия обязан самостоятельно подготовиться к теме устного опроса, которые состоялись на практическом занятии. В рамках часов консультаций студент может ответить на вопросы пропущенного устного опроса, которые были пропущены.

12. Методические рекомендации преподавателям по организации обучения по дисциплине

Курс «Анализ больших данных (Big Data Analytics)» должен давать не абстрактно-формальные, а прикладные знания. Данная цель может быть реализована только при условии соблюдения в учебных планах преемственности учебных дисциплин. Базовые знания для изучения «Анализ больших данных (Big Data Analytics)» дают такие дисциплины, как «Инструменты Data Science в R, Python, SQL», «Наука о данных (Data Science)», «Байесовские методы в машинном обучении». Освоение основных тем данной дисциплины позволит студентам использовать знания при написании выпускной квалификационной работы, понять всю ширину науки и получить необходимые знания для последующего профессионального развития в этой области.

Студент может подготовить доклад по теме, представляющей его научный интерес, представить результаты в виде презентации. В случае надлежащего качества, его работа может быть заслушана на научном кружке кафедры или на студенческой научной конференции. По решению кафедры, студенты, занявшие призовые места на научных студенческих конференциях, могут освобождаться от сдачи зачета по этой дисциплине.

Преподаватель должен указывать, в какой последовательности следует изучать материал дисциплины, обращать внимание на особенности изучения отдельных тем и разделов, помогать отбирать наиболее важные и необходимые сведения из учебных пособий, а также давать объяснения вопросам программы курса, которые обычно вызывают затруднения. При этом преподавателю необходимо учитывать следующие моменты:

1. Не следует перегружать студентов творческими заданиями.

2. Чередовать творческую работу на занятиях с заданиями во внеаудиторное время.
3. Давать студентам четкий инструктаж по выполнению самостоятельных заданий: цель задания; условия выполнения; объем; сроки; требования к оформлению.
4. Осуществлять текущий учет и контроль за самостоятельной работой.
5. Давать оценку и обобщать уровень усвоения навыков самостоятельной, творческой работы.

Программу разработали:

Демичев В.В., кандидат экономических наук, доцент,

Невзоров А.С., ассистент



(подпись)

РЕЦЕНЗИЯ

на рабочую программу дисциплины Б1.В.05 Анализ больших данных (Big Data Analytics) ОПОП ВО по направлению 09.04.02 Информационные системы и технологии, направленность Науки о данных (квалификация выпускника – магистр)

Быстрениной Ирины Евгеньевны, доцентом кафедры прикладной информатики, кандидатом педагогических наук (далее по тексту рецензент), проведено рецензирование рабочей программы дисциплины «Анализ больших данных (Big Data Analytics)» ОПОП ВО по направлению 09.04.02 Информационные системы и технологии, направленность «Науки о данных» (магистр) разработанной в ФГБОУ ВО «Российский государственный аграрный университет – МСХА имени К.А. Тимирязева», на кафедре статистики и кибернетики (работчики – Демичев Вадим Владимирович, доцент, кандидат экономических наук, Невзоров А.С. ассистент кафедры статистики и кибернетики).

Рассмотрев представленные на рецензирование материалы, рецензент пришел к следующим выводам:

1. Предъявленная рабочая программа дисциплины «Анализ больших данных (Big Data Analytics)» (далее по тексту Программа) соответствует требованиям ФГОС ВО по направлению 09.04.02 Информационные системы и технологии. Программа содержит все основные разделы, соответствует требованиям к нормативно-методическим документам.

2. Представленная в Программе **актуальность** учебной дисциплины в рамках реализации ОПОП ВО не подлежит сомнению – дисциплина относится ко части дисциплин, формируемых участниками образовательных отношений – Б1.В.

3. Представленные в Программе **цели** дисциплины соответствуют требованиям ФГОС ВО направления 09.04.02 Информационные системы и технологии.

4. В соответствии с Программой за дисциплиной «Анализ больших данных (Big Data Analytics)» закреплены **3 профессиональных компетенций, определяемых самостоятельно**. Дисциплина «Анализ больших данных (Big Data Analytics)» и представленная Программа способна реализовать их в объявленных требованиях.

5. **Результаты обучения**, представленные в Программе в категориях знать, уметь, владеть соответствуют специфике и содержанию дисциплины и демонстрируют возможность получения заявленных результатов.

6. Общая трудоёмкость дисциплины «Анализ больших данных (Big Data Analytics)» составляет 5,0 зачётных единиц (180 часов).

7. Информация о взаимосвязи изучаемых дисциплин и вопросам исключения дублирования в содержании дисциплин соответствует действительности. Дисциплина «Анализ больших данных (Big Data Analytics)» взаимосвязана с другими дисциплинами ОПОП ВО и Учебного плана по направлению 09.04.02 Информационные системы и технологии и возможность дублирования в содержании отсутствует.

8. Представленная Программа предполагает использование современных образовательных технологий, используемые при реализации различных видов учебной работы. Формы образовательных технологий соответствуют специфике дисциплины.

9. Программа дисциплины «Анализ больших данных (Big Data Analytics)» предполагает проведение занятий в интерактивной форме.

10. Виды, содержание и трудоёмкость самостоятельной работы студентов, представленные в Программе, соответствуют требованиям к подготовке выпускников, содержащимся во ФГОС ВО направления 09.04.02 Информационные системы и технологии.

11. Представленные и описанные в Программе формы *текущей* оценки знаний (устный опрос, коллоквиум), соответствуют специфике дисциплины и требованиям к выпускникам.

Форма промежуточного контроля знаний студентов, предусмотренная Программой, осуществляется в форме зачета и экзамена в первом и втором семестрах соответственно, что соответствует статусу дисциплины, формируемой участниками образовательных отношений – Б1.В. ФГОС ВО направления 09.04.02. Информационные системы и технологии.

12. Формы оценки знаний, представленные в Программе, соответствуют специфике дисциплины и требованиям к выпускникам.

13. Учебно-методическое обеспечение дисциплины представлено: основной литературой – 3 источника (базовые учебники), дополнительной литературой – 3 наименования, Интернет-ресурсы – 3 источника и соответствует требованиям ФГОС ВО направления 09.04.02 *Информационные системы и технологии*.

14. Материально-техническое обеспечение дисциплины соответствует специфике дисциплины «Анализ больших данных (Big Data Analytics)» и обеспечивает использование современных образовательных, в том числе интерактивных методов обучения.


15. Методические рекомендации студентам и методические рекомендации преподавателям по организации обучения по дисциплине дают представление о специфике обучения по дисциплине «Анализ больших данных (Big Data Analytics)».

ОБЩИЕ ВЫВОДЫ

На основании проведенного рецензирования можно сделать заключение, что характер, структура и содержание рабочей программы дисциплины «Анализ больших данных (Big Data Analytics)» ОПОП ВО по направлению 09.04.02 Информационные системы и технологии, направленность «Науки о данных» (квалификация выпускника – магистр), разработанная Демичевым Вадимом Владимировичем, доцентом, кандидатом экономических наук, Невзоровым Александром Сергеевичем, ассистентом кафедры статистики и кибернетики, соответствует требованиям ФГОС ВО, современным требованиям экономики, рынка труда и позволит при её реализации успешно обеспечить формирование заявленных компетенций.

Рецензент:

Быстренина Ирина Евгеньевна, доцент кафедры прикладной информатики, кандидат педагогических наук


_____ «28» августа 2023 г.
(подпись)