Министерство науки и высшего образования Российской Федерации
федеральное государственное бюджетное образовательное учреждение
высшего образования
«Казанский национальный исследовательский технологический университет»

УТВЕРЖДАЮ
Проректор по УР
_____ Бурмистров А.В.
«01» 04 2019 г.

РАБОЧАЯ ПРОГРАММА

По дисциплине Б1.В.ДВ.05.01 «Интеллектуальный анализ данных»

Направление подготовки 01.03.05 «Статистика»

Профиль подготовки «Бизнес-статистика и прогнозирование»

Квалификация выпускника *БАКАЛАВР*

Форма обучения очная

Институт, факультет Институт управления инновациями, Факультет промышленной политики и бизнес-администрирования

Кафедра бизнес-статистики и математических методов в экономике

Курс, семестр 4 курс, 8 семестр

	Часы	Зачетные единицы
Лекции	18	0,5
Практические (семинарские) занятия	-	-
Лабораторные занятия	18	0,5
Самостоятельная работа	99	2,75
Экзамен, семестр	Экзамен (45)	1,25
Всего	180	5

Казань, 2019 г.

Рабочая программа составлена с учетом требований Федерального государственного образовательного стандарта подготовки высшего образования (приказ № 140 от 16 февраля 2017 г.) по направлению 01.03.05 «Статистика» для профиля «Бизнес-статистика и прогнозирование», на основании учебных планов набора обучающихся 2019 г.

Типовая программа по дисциплине отсутствует.

Разработчик программы доцент кафедры бизнес-статистики и математических методов в экономике _____ Гадельшина Г.А.

Рабочая программа рассмотрена и одобрена на заседании кафедры бизнес-статистики и математических методов в экономике протокол от 11.06. 2019 г. № 10

Зав. кафедрой, проф._____Аксянова А.В.

У Т В Е Р Ж Д Е Н О

Протокол заседания методической комиссии факультета промышленной политики и бизнес-администрирования от 13.06 2019 г. № 10

Председатель комиссии, профессор _____ Тузиков А.Р.

Начальник УМЦ, доцент _____ Китаева Л.А.

1. Цели освоения дисциплины

Целями освоения дисциплины «Интеллектуальный анализ данных» являются

а) формирование знаний о задачах, решаемых с помощью интеллектуального анализа данных, разделах Data Mining и используемых методах, а также сферах его применения;

б) обучение методике применения интеллектуального анализа данных;

в) обучение способам применения вычислительной техники, прикладных программ, математических и статистических методов при решении задач интеллектуального анализа данных;

г) раскрытие сущности процессов, лежащих в основе использования определенных моделей Data Mining.

2. Место дисциплины в структуре ООП ВО

Дисциплина «Интеллектуальный анализ данных» относится к вариативной части ООП (дисциплина по выбору) и формирует у обучающихся по направлению подготовки 01.03.05 «Статистика» набор знаний, умений, навыков и компетенций, необходимых для выполнения аналитической деятельности. Для успешного освоения дисциплины «Интеллектуальный анализ данных» бакалавр по направлению подготовки 01.03.05 «Статистика» должен владеть материалом дисциплин:

а) Б1.Б.21 «Теория вероятностей и математическая статистика»,
б) Б1.В.14 «Экономическая информатика»,
в) Б1.Б.27 «Эконометрика»,
г) Б1.В.08 «Анализ временных рядов и прогнозирование»,
д) Б1.Б.29 «Системы статистической обработки данных»,
е) Б1.В.ДВ.02.01 «Базы данных и технологии больших данных».

Знания, полученные при изучении дисциплины «Интеллектуальный анализ данных» могут быть использованы при прохождении преддипломной практики и выполнении выпускной квалификационной работы по направлению подготовки 01.03.05 «Статистика».

3. Компетенции обучающегося, формируемые в результате освоения дисциплины

1. способностью самостоятельно осваивать новые методы прикладной и математической статистики для их использования в аналитической работе (ПК-3);
2. способностью осознанно применять методы математической и дескриптивной статистики для анализа количественных данных, содержательно интерпретировать полученные результаты, (ПК-4).

В результате освоения дисциплины обучающийся должен:

1) Знать:
   а) основные методы Data Mining;
   б) типы задач, решаемых с помощью Data Mining;
   в) этапы интеллектуального анализа данных.
2) Уметь:
   а) применять методы и модели Data Mining в прикладных исследованиях;
   б) оценивать качество моделей;
   в) интерпретировать результаты моделирования;
   г) прогнозировать дальнейшее поведение системы.
3) Владеть:
   а) инструментальными методами интеллектуального анализа данных;
   б) одним из специализированных программных пакетов для применения методов Data Mining.

4. Структура и содержание дисциплины

Общая трудоемкость дисциплины составляет 5 зачетных единиц, 180 часов.

№ п/п	Раздел дисциплины	Семестр	Виды учебной работы (в часах)					Оценочные средства для проведения промежуточной аттестации
			Лекции	Практ. зан.	Лабораторные занятия	СРС	Всего	
1.	Основные задачи и методы интеллектуального анализа данных	8	1			2	3	Выполнение/сдача лабораторных работ, выполнение расчетного задания, посещение занятий
2.	Сбор и подготовка исходных данных.	8	1		2	8	11	
3.	Разведочные и графические методы для интерактивного исследования данных	8	2		2	8	12	
4.	Анализ ассоциаций	8	2		2	12	16	
5.	Деревья принятия решений	8	2		2	12	16	
6.	Анализ независимых компонент	8	2		2	12	16	
7.	Обобщенные методы кластерного анализа	8	2		2	15	19	
8.	Методы машинного обучения	8	2		4	16	22	
9.	Обобщенные аддитивные модели	8	2		2	12	16	
10.	Нейронные сети	8	2			2	4	
	Форма аттестации						45	Экзамен
	ИТОГО		18		18	99	180	

5. Содержание лекционных занятий

№ п/п	Раздел / Тема лекционного занятия	Часы	Краткое содержание	Формируемые компетенции
1.	Основные задачи и методы интеллектуального анализа данных	1	Понятие интеллектуального анализа данных. Задачи Data Mining. Этапы исследования. Сферы применения интеллектуального анализа данных.	ПК-3 ПК-4
2.	Сбор и подготовка исходных данных.	1	Использование различных источников данных. Технология получение данных из баз данных с помощью SQL запросов. Предобработка данных. Очистка данных.	ПК-3 ПК-4
3.	Разведочные и графические методы для интерактивного исследования данных	2	Исследование типов данных. OLAP-технологии. Исследование взаимосвязей между переменными. Инструменты Drill-Down ППП Statistica. Построение графиков, частотных таблиц, вывод описательной статистики.	ПК-3 ПК-4
4.	Анализ ассоциаций	2	Сферы применения анализа ассоциаций. Ассоциативные правила. Поддержка. Достоверность. Полезность. Алгоритм Apriori. Структура входных данных.	ПК-3 ПК-4

№ п/п	Раздел / Тема лекционного занятия	Часы	Краткое содержание	Формируемые компетенции
			Переменные с многомерными откликами и многомерной дихотомией.	
5.	Деревья принятия решений	2	Деревья принятия решений. Общерегрессионные и классифицирующие деревья. CHAID деревья. Стохастический градиентный метод бустинга. Алгоритм случайных лесов. Многомерные адаптивные сплайны.	ПК-3 ПК-4
6.	Анализ независимых компонент	2	Задачи разделения сигналов. Области применения.	ПК-3 ПК-4
7.	Обобщенные методы кластерного анализа	2	Методы кластерного анализа. Метод k-средних. EM метод. Выбор числа кластеров. Оценка качества кластеризации	ПК-3 ПК-4
8.	Методы машинного обучения	2	Метод опорных векторов. Метод k ближайших соседей. Байесовские процедуры классификации.	ПК-3 ПК-4
9.	Обобщенные аддитивные модели	2	Методы множественной регрессии. Виды функций связи. Виды функций распределения предикторов.	ПК-3 ПК-4
10.	Нейронные сети	2	Обзор нейронных сетей. Этапы нейросетевого моделирования. Обучение сети. Проведение классификации.	ПК-3 ПК-4

6. Содержание практических занятий

Учебным планом направления 01.03.05 не предусмотрено проведение практических занятий по дисциплине «Интеллектуальный анализ данных».

7. Содержание лабораторных занятий

Целью проведения лабораторных работ является освоение и систематизирование лекционного материала, формирование и укрепление навыков использования компьютера при решении задач анализа данных (пакете STATISTICA).

№ п/п	Раздел дисциплины	Часы	Наименование лабораторной работы	Компетенции
1.	Сбор и подготовка исходных данных.	2	Получение исходных данных	ПК-3 ПК-4
2.	Разведочные и графические методы для интерактивного исследования данных	2	Интерактивное бурение	
3.	Анализ ассоциаций	2	Анализ ассоциаций	
4.	Деревья принятия решений	2	Построение деревьев принятия решений	
5.	Анализ независимых компонент	2	Анализ независимых компонент	
6.	Обобщенные методы кластерного анализа	2	Кластерный анализ	
7.	Методы машинного обучения	2	Метод опорных векторов	
8.		2	Метод k ближайших соседей.	
9.	Обобщенные аддитивные модели	2	Множественная регрессия с нелинейной функцией связи	

Лабораторные занятия проводятся в компьютерных классах с использованием специального оборудования: персональных компьютеров типа IBM PC, работающих в среде Windows с установленными компонентами: MS Office (Word, Excel, Statistica) с подключением в локальную сеть.

8. Самостоятельная работа бакалавра

№ п/п	Темы, выносимые на самостоятельную работу	Часы	Форма СР	Формируемые компетенции
1.	Основные задачи и методы интеллектуального анализа данных	2	Освоение теоретического материала Подготовка отчетов по лабораторным работам.	ПК-3 ПК-4
2.	Сбор и подготовка исходных данных.	8		
3.	Разведочные и графические методы для интерактивного исследования данных	8		
4.	Анализ ассоциаций	12		
5.	Деревья принятия решений	12		
6.	Анализ независимых компонент	12		
7.	Обобщенные методы кластерного анализа	15		
8.	Методы машинного обучения	16		
9.	Обобщенные аддитивные модели	12		
10.	Нейронные сети	2		

9. Использование рейтинговой системы оценки знаний.

При оценке результатов деятельности студентов в рамках дисциплины «Интеллектуальный анализ данных» используется рейтинговая система в соответствие с Положением о балльно-рейтинговой системе оценки знаний студентов и обеспечения качества учебного процесса (утверждено решением УМК Ученого совета ФГБОУ ВО «КНИТУ» от 04.09.2017 г.). Рейтинговая оценка формируется на основании текущего и промежуточного контроля. Максимальное и минимальное количество баллов по различным видам учебной работы описано в положении о балльно-рейтинговой системе.

Виды деятельности	Минимальное кол-во баллов	Максимальное кол-во баллов
Посещение занятий	6	10
Выполнение/сдача лабораторных работ	18	30
Выполнение расчетного задания	12	20
Экзамен	24	40
Итого	60	100

10. Оценочные средства для текущего контроля успеваемости, промежуточной аттестации по итогам освоения дисциплины

Для оценки сформированности компетенций используются:
- лабораторные работы,
- презентация и доклад по результатам расчетного задания,
- экзаменационные билеты.

Полный перечень оценочных средств представлен в фонде оценочных средств по дисциплине «Интеллектуальный анализ данных».

11. Информационно-методическое обеспечение дисциплины

11.1 Основная литература

При изучении дисциплины «Интеллектуальный анализ данных» в качестве основных источников информации рекомендуется использовать следующую литературу.

№	Основные источники информации	Кол-во экз.
1.	Соловьева, С. В. Информационные технологии в профессиональной деятельности: инструментарий бизнес-аналитики : практикум / С.В. Соловьева, Ю.П. Александровская, Ю.В. Хайрутдинова ; Казанский нац. исслед. технол. ун-т .— Казань : Изд-во КНИТУ, 2017 .— 102 с.	156 экз. в УНИЦ КНИТУ 8 экз. на кафедре БСМЭ
2.	Мардас А.Н. Эконометрика : Учебник и практикум / Мардас А.Н. — 2-е изд., испр. и доп .— М. : Издательство Юрайт, 2018 .— 180 .	ЭБС «Юрайт» Доступ из любой точки интернета после регистрации с IP-адресов КНИТУ http://www.biblio-online.ru/book/C3F5B1E3-0900-4ADD-8864-D98F195BB173
3.	Костюнин В.И. Эконометрика: Учебник и практикум / Костюнин В.И. — М. : Издательство Юрайт, 2018 .— 285 .	ЭБС «Юрайт» Доступ из любой точки интернета после регистрации с IP-адресов КНИТУ http://www.biblio-online.ru/book/14E91936-EC2D-4AC8-A80F-787D7FFCD41C
4.	Александровская, Ю.П. Многомерный статистический анализ в экономике [Учебники] : учеб. пособие / Ю.П. Александровская ; Казанский нац. исслед. технол. ун-т .— Казань : Изд-во КНИТУ, 2017 .— 94 с.	66 экз. в УНИЦ КНИТУ 9 экз. на кафедре БСМЭ

11.2 Дополнительная литература

В качестве дополнительных источников информации рекомендуется использовать следующую литературу

№	Дополнительные источники информации	Кол-во экз.
1.	Ниворожкина, Л.И. Статистические методы анализа данных: учебник / Л.И. Ниворожкина [и др.] ; Ростовский гос. экон. ун-т ; под общ. ред. Л.И. Ниворожкиной .— М. : РИОР : ИНФРА-М, 2017 .— 331 с.	1 экз. в УНИЦ КНИТУ ЭБС «Znanium.com» Доступ из любой точки интернета после регистрации с IP-адресов КНИТУ http://znanium.com/go.php?id=556760
2.	Боровиков В. STATISTIKA. Искусство анализа данных на компьютере / В. Боровиков .— 2-е изд., [испр. и доп.] .— М. [и др.] : Питер, 2003 .— 688 с. : ил.+ ПРИЛОЖЕНИЕ на CD .— (Для профессионалов) .	3 экз. в УНИЦ КНИТУ
3.	Вуколов Э.А. Основы статистического анализа. Практикум по статистическим методам и исследованию операций с использованием пакетов STATISTICA и EXCEL: учебное пособие. – 2-е изд., испр. и доп. – М.: ФОРУМ, 2013. – 464 с.	1 экз. в УНИЦ КНИТУ

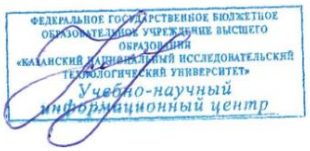
№	Дополнительные источники информации	Кол-во экз.
4.	Карякин, А. А. Методы статистической обработки экспериментальных данных: учеб. пособие / А.А. Карякин, О.Е. Карякина ; Северный (Арктический) федерал. ун-т им. М.В. Ломоносова .— Архангельск, 2015 .— 101 с.	1 экз. в УНИЦ КНИТУ

11.3 Электронные источники информации

При изучении дисциплины «Интеллектуальный анализ данных» использование электронных источников информации:

1. Электронный каталог УНИЦ КНИТУ. – Доступ свободный: http://ruslan.kstu.ru/
2. ЭБС «Юрайт». – Доступ из любой точки интернета после регистрации с IP-адресов КНИТУ: https://biblio-online.ru
3. ЭБС «BOOK.RU» - Доступ из любой точки интернета после регистрации с IP-адресов КНИТУ: http://www.book.ru/
4. Научная электронная библиотека (РУНЭБ) - Доступ из любой точки интернета после регистрации с IP-адресов КНИТУ: http://elibrary.ru

Согласовано:
Зав. сектором ОКУФ



1. Журнал «ФИНАНСЫ И КРЕДИТ». Сайт журнала «Финансы и Кредит». – http://www.fin-izdat.ru/journal/fc/
2. Журнал «АКТУАЛЬНЫЕ ПРОБЛЕМЫ ЭКОНОМИКИ И УПРАВЛЕНИЯ». Сайт журнала – http://emtp.guap.ru
3. Журнал «ИННОВАЦИОННАЯ ЭКОНОМИКА: ИНФОРМАЦИЯ, АНАЛИТИКА, ПРОГНОЗЫ». Сайт журнала – http://morvesti.ru/archive/ie/2012/06.php
4. Журнал «ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ И МАТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ В ЭКОНОМИКЕ, ТЕХНИКЕ, ЭКОЛОГИИ, ОБРАЗОВАНИИ, ПЕДАГОГИКЕ И ТОРГОВЛЕ». Сайт журнала – http://www.sibsau.ru
5. Журнал «ЭКОНОМИКА И ПРОГНОЗИРОВАНИЕ». Сайт журнала – http://www.ief.org.ua/IEF_rus/EP.htm
6. Федеральная служба государственной статистики www.gks.ru
7. Территориальный орган Федеральная служба государственной статистики по Республике Татарстан www.tatstat.ru
8. Портал Знаний StatSoft. Сайт http://statistica.ru

11.4. Современные профессиональные базы данных и информационные справочные системы.

База данных ScienceDirect - www.sciencedirect.com
Университетская информационная система Россия - uisrussia.msu.ru

12. Материально-техническое обеспечение дисциплины (модуля).

В качестве материально-технического обеспечения дисциплины используются мультимедийные средства и средства мониторинга.

1. Лекционные занятия:
    a. комплект электронных презентаций,
    b. аудитория, оснащенная презентационной техникой (проектор, ноутбук),
2. Лабораторные занятия:
    a. компьютерный класс,
    b. ППП MS Office(Word, Excel),
    c. ППП EViews / Statistica (любой один из перечисленных ППП).
3. Прочее
    a. рабочее место преподавателя, оснащенное компьютером с доступом в Интернет,
    b. рабочие места студентов, оснащенные компьютерами с доступом в Интернет.

13. Образовательные технологии

Учебным планом направления 01.03.05 по дисциплине «Интеллектуальный анализ данных» предусмотрено проведение занятий в интерактивной форме в объеме 18 часов (групповые дискуссии, проблемные лекции, разбор кейсовых задач).