

**«Финансовый университет при Правительстве Российской Федерации»
(Финуниверситет)**

Калужский филиал Финуниверситета

Кафедра «Бизнес-информатика и высшая математика»

«УТВЕРЖДАЮ»

**Директор Калужского филиала
Финуниверситета**



В.А. Матчинов

И.В. Винокуров

**ПРОЕКТИРОВАНИЕ АНАЛИТИЧЕСКИХ
ИНФОРМАЦИОННЫХ СИСТЕМ**

Рабочая программа дисциплины

для студентов, обучающихся по направлению подготовки

38.03.05 «Бизнес-информатика»

Образовательная программа

«Цифровая трансформация управления бизнесом»

Очная форма обучения

*Рекомендовано Ученым советом Калужского филиала Финуниверситета
(протокол № 03 от 29 июня 2023 г.)*

*Одобрено кафедрой «Бизнес-информатика и высшая математика»
Калужского филиала Финуниверситета
(протокол № 12 от 29 июня 2023 г.)*


КАЛУГА 2023

Рабочая программа предназначена для преподавания дисциплины «Проектирование аналитических информационных систем» студентам, обучающимся по направлению подготовки 38.03.05 «Бизнес-информатика», образовательная программа «Цифровая трансформация управления бизнесом», по очной форме обучения.


В рабочей программе излагаются планируемые результаты освоения дисциплины, содержание дисциплины, тематика и содержание семинаров и практических занятий, технологии их проведения. Приводится перечень учебно-методического обеспечения для самостоятельной работы обучающихся, фонд оценочных средств для проведения промежуточной аттестации обучающихся, перечень основной и дополнительной литературы, а также ресурсов информационно-телекоммуникационной сети Интернет.

СОГЛАСОВАНО:

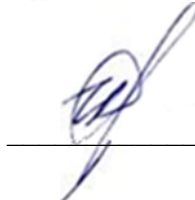
Заместитель директора
по учебно-методической работе
«29» июня 2023 г.

 /Орловцева О.М./

Начальник учебно-методического отдела
«29» июня 2023 г.

 /Толстикова В.С./

Заведующий кафедрой
«Бизнес-информатика и высшая математика»
«29» июня 2023 г.

 /Дробышева И.В./

Содержание

1. Наименование дисциплины.....	4
2. Перечень планируемых результатов освоения образовательной программы (перечень компетенций) с указанием индикаторов их достижения, соотнесённых с планируемыми результатами обучения по дисциплине	4
3. Место дисциплины в структуре образовательной программы.....	4
4. Объем дисциплины в зачетных единицах и в академических часах с выделением объема аудиторной (лекции, семинары) и самостоятельной работы обучающихся	5
5. Содержание дисциплины, структурированное по темам (разделам) дисциплины с указанием их объемов (в академических часах) и видов учебных занятий	5
5.1. Содержание дисциплины.....	5
5.2. Учебно-тематический план.....	6
5.3. Содержание семинаров, практических занятий	7
6. Перечень учебно-методического обеспечения для самостоятельной работы обучающихся по дисциплине.....	7
6.1. Перечень вопросов, отводимых на самостоятельное освоение дисциплины, формы внеаудиторной самостоятельной работы.....	7
6.2. Перечень вопросов, заданий, тем для подготовки к текущему контролю успеваемости.....	8
7. Фонд оценочных средств для проведения промежуточной аттестации обучающихся по дисциплине.....	8
7.1 Перечень компетенций с указанием индикаторов их достижения в процессе освоения дисциплины	8
7.2 Типовые контрольные задания или иные материалы, необходимые для оценки индикаторов достижения компетенций, умений и знаний	8
8. Перечень основной и дополнительной литературы, необходимой для освоения дисциплины.....	10
9. Перечень ресурсов сети Интернет, необходимых для освоения дисциплины....	10
10. Методические указания для обучающихся по освоению дисциплины	11
11. Перечень информационных технологий, используемых при осуществлении образовательного процесса по дисциплине, включая перечень необходимого программного обеспечения и информационных справочных систем	11
11.1 Комплект лицензионного программного обеспечения:	11
11.2 Современные профессиональные базы данных и информационные справочные системы:	11
11.3 Сертифицированные программные и аппаратные средства защиты информации: не предусмотрены	11
12. Описание материально-технической базы, необходимой для осуществления образовательного процесса по дисциплине.....	11

1. Наименование дисциплины

«Проектирование аналитических информационных систем».

2. Перечень планируемых результатов освоения образовательной программы (перечень компетенций) с указанием индикаторов их достижения, соотнесённых с планируемыми результатами обучения по дисциплине

В результате изучения дисциплины у студентов должны быть сформированы следующие компетенции:

Таблица 1

Код компетенции	Наименование компетенции	Индикаторы достижения компетенции	Результаты обучения (владения, умения и знания), соотнесённые с компетенциями/индикаторами достижения компетенции
ПКН-3	Способность применять аналитические системы и консультировать по вопросам разработки и развития аналитических систем работы с данными	1. Применяет аналитические системы работы с данными	Уметь применять аналитические модели и системы для решения практических задач
		2. Проводит анализ рынка	Уметь проводить анализ рынка с использованием аналитических моделей
		3. Консультирует по вопросам применения аналитических систем работы с данными	Знать современные способы представления данных и умение их использовать при разработке аналитических систем
ПКП-3	Способность предлагать различные варианты инфраструктурных решений для поддержки ИТ/ИС	1. Анализирует текущий уровень инфраструктурных решений предприятия/организации	Знать особенности применения и сферы использования ИТ-решений разного вида и назначения Уметь выполнять анализ текущих ИТ-решений организации на соответствие её потребностям
		2. Формирует и обосновывает варианты технологического слоя архитектуры предприятия/организации	Знать виды современных ИТ-решений их назначение и особенности для формирования технологического слоя архитектуры организации Уметь формировать и обосновывать структуру и компоненты технологического слоя архитектуры организации

3. Место дисциплины в структуре образовательной программы

Дисциплина «Проектирование аналитических информационных систем» относится к элективной дисциплине модуля «Информационно-аналитические технологии», отражающего специфику ВУЗа по направлению подготовки 38.03.05 «Бизнес-информатика», ОП «Цифровая трансформация управления бизнесом».

4. Объем дисциплины в зачетных единицах и в академических часах с выделением объема аудиторной (лекции, семинары) и самостоятельной работы обучающихся

Таблица 2

Вид учебной работы по дисциплине	Всего (в з/е и часах)	Семестр 7 (в часах)
Общая трудоемкость дисциплины	3 з/е, 108 ч.	108 ч.
Контактная работа – аудиторные занятия	50	50
Лекции	16	16
Семинары, практические занятия	34	34
Самостоятельная работа	58	58
Вид текущего контроля	Контрольная работа	Контрольная работа
Вид промежуточной аттестации	зачёт	зачёт

5. Содержание дисциплины, структурированное по темам (разделам) дисциплины с указанием их объемов (в академических часах) и видов учебных занятий

5.1. Содержание дисциплины

Тема 1. Представление данных в аналитических моделях

Способы и особенности представления данных в виде списков (list), множеств (set), серий данных (Series) и датафреймов (DataFrame) библиотеки Pandas; векторов (array) и матриц (matrix) библиотеки Numpy. Формирование наборов данных для обучения и тестирования моделей. Понятие аугментации данных и способы её реализации. Визуализация данных с использованием библиотеки matplotlib. Среда для создания и исследования моделей – Google Colaboratory и Visual Studio Code.

Тема 2. Проектирование регрессионных моделей аналитических систем

Прогнозирование значений наблюдаемой величины с использованием линейной и полиномиальной регрессий. Регрессор LinearRegression библиотеки Scikit-Learn.

Тема 3. Проектирование моделей кластеризации аналитических систем

Прогнозирование принадлежности одному или нескольким классам с использованием логистической регрессии. Регрессор LogisticRegression библиотеки Scikit-Learn. Кластеризация методом k-средних. Регрессор KMeans библиотеки Scikit-Learn. Кластеризация с использованием методов дерева принятия решений и случайного леса. Регрессоры DecisionTreeClassifier и RandomForestClassifier библиотеки Scikit-Learn. Классификация методом опорных векторов. Регрессор SVM библиотеки Scikit-Learn.

5.2. Учебно-тематический план

Таблица 3

№	Наименование тем(разделов) дисциплины	Трудоемкость в часах					Формы текущего контроля успеваемости
		Всего	Аудиторная работа			Самостоятельная работа	
			Общая	Лекции	Семинары, практические занятия		
Тема 1. Представление данных в аналитических моделях							
1	Способы и особенности представления данных аналитических моделей	12	6	2	4	6	Выполнение и защита практических работ
2	Разделение данных на обучающих и тестовый наборы. Аугментация и визуализация данных	14	6	2	4	8	Выполнение и защита практических работ
Тема 2. Проектирование регрессионных моделей аналитических систем							
3	Проектирование и реализация линейной регрессии	12	6	2	4	6	Выполнение и защита практических работ
4	Проектирование и реализация полиномиальной регрессии	12	6	2	4	6	Выполнение и защита практических работ
Тема 3. Проектирование моделей кластеризации аналитических систем							
5	Проектирование и реализация модели логистической регрессии	14	6	2	4	8	Выполнение и защита практических работ
6	Проектирование и реализация модели кластеризации методом k-средних	14	6	2	4	8	Выполнение и защита практических работ
7	Проектирование и реализация моделей дерева принятия решений и случайного леса	16	8	2	6	8	Выполнение и защита практических работ
8	Проектирование и реализация модели метода опорных векторов	14	6	2	4	8	Выполнение и защита практических работ
В целом по дисциплине		108	50	16	34	58	Контрольная работа

5.3. Содержание семинаров, практических занятий

Таблица 4

Наименование тем (разделов) дисциплины	Перечень вопросов для обсуждения на семинарских, практических занятиях, рекомендуемые источники из разделов 8, 9	Формы проведения занятий
Тема 1. Представление данных в аналитических моделях	<ul style="list-style-type: none"> Формирование и визуализация данных для аналитической модели заданного типа. <p>Основная литература: 1,2 Дополнительная литература: 7</p>	Компьютерный практикум
Тема 2. Проектирование регрессионных моделей аналитических систем	<ul style="list-style-type: none"> Проектирование и реализация линейной и полиномиальной регрессионных моделей для обработки экспериментальных данных. <p>Основная литература: 1,2 Дополнительная литература: 7</p>	Компьютерный практикум
Тема 3. Проектирование моделей кластеризации аналитических систем	<ul style="list-style-type: none"> Проектирование и реализация моделей: логистической регрессии, кластеризации методом k-средних, дерева принятия решений и случайного леса, метода опорных векторов. <p>Основная литература: 1,2 Дополнительная литература: 7</p>	Компьютерный практикум

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

6.1. Перечень вопросов, отводимых на самостоятельное освоение дисциплины, формы внеаудиторной самостоятельной работы

Таблица 5

Наименование тем (разделов) дисциплины	Перечень вопросов, отводимых на самостоятельное освоение	Формы внеаудиторной самостоятельной работы
Тема 1. Представление данных в аналитических моделях	<ul style="list-style-type: none"> Изучение основных функциональных возможностей библиотек Scikit-Learn, TensorFlow и Keras. <p>Основная литература: 1,2 Дополнительная литература: 7</p>	Изучение методических материалов по теме в электронном виде и рекомендуемых разделов основной и дополнительной литературы, интернет – источников. Подготовка к практическим работам
Тема 2. Проектирование регрессионных моделей аналитических систем	<ul style="list-style-type: none"> Матрица ошибок confusion matrix и способы её визуализации. <p>Основная литература: 1,2 Дополнительная литература: 7</p>	Изучение методических материалов по теме в электронном виде и рекомендуемых разделов основной и дополнительной литературы, интернет – источников. Подготовка к практическим работам

		ским работам.
Тема 3. Проектирование моделей кластеризации аналитических систем	<ul style="list-style-type: none"> Метрики точности регрессионных моделей – Accuracy, Precision, F1-score и AUC-ROC. <p>Основная литература: 1,2 Дополнительная литература: 7</p>	Изучение методических материалов по теме в электронном виде и рекомендуемых разделов основной и дополнительной литературы, интернет – источников. Подготовка к практическим работам

6.2. Перечень вопросов, заданий, тем для подготовки к текущему контролю успеваемости

Примерные темы для контрольной работы:

Проектирование аналитической системы для <заданной предметной области>.

Критерии балльной оценки по контрольной работе содержатся в соответствующих методических рекомендациях кафедры.

7. Фонд оценочных средств для проведения промежуточной аттестации обучающихся по дисциплине

7.1 Перечень компетенций с указанием индикаторов их достижения в процессе освоения дисциплины

Перечень компетенций представлен в разделе 2, который характеризует перечень планируемых результатов обучения по дисциплине, соотнесенных с планируемыми результатами освоения образовательной программы.

7.2 Типовые контрольные задания или иные материалы, необходимые для оценки индикаторов достижения компетенций, умений и знаний

Таблица 6

Компетенция	Типовые задания
ПКН-3 Способность применять аналитические системы и консультировать по вопросам разработки и развития аналитических систем работы с данными	1. Применяет аналитические системы работы с данными Задание 1. Оцените величину дохода организации с использованием линейной регрессии Задание 2. Оцените величину прибыли организации с использованием полиномиальной регрессии
	2. Проводит анализ рынка Задание 1. Оцените динамику рынка с использованием модели случайного леса Задание 2. Оцените динамику рынка с использованием метода опорных векторов
	3. Консультирует по вопросам применения аналитических систем работы с данными Задание 1. Выберите регрессионную модель для эффективного решения оценки рисков Задание 2. Выберите модель кластеризации для

	распознавания текущей ситуации
ПКП-3 Способность предлагать различные варианты инфраструктурных решений для поддержки ИТ/ИС	1. Анализирует текущий уровень инфраструктурных решений предприятия/организации Задание 1. Предложите метод аналитики для повышения эффективности работы обучаемой модели ИС Задание 2. Осуществите анализ инфраструктурных решений организации с целью повышения эффективности их работы
	2. Формирует и обосновывает варианты технологического слоя архитектуры предприятия/организации Задание 1. Спроектируйте и реализуйте ИС, реализующую выбранную модель аналитики Задание 2. Исследуйте работу обучаемой модели в ИС и предложите способы повышения её эффективности

Критерии балльной оценки различных форм текущего контроля успеваемости содержатся в соответствующих методических рекомендациях кафедры.

Примерные вопросы к зачёту:

1. Приведите основные библиотеки Python, используемые для обработки экспериментальных данных. Опишите их назначения и основные функциональные возможности.
2. Опишите основные этапы работы с моделями данных в Google Colaboratory и Visual Studio Code.
3. Раскройте понятие линейной регрессии. Приведите типы задач, в которых используется линейная регрессия. Опишите модель линейной регрессии, реализованную в библиотеке Scikit-Learn.
4. Раскройте понятие логистической регрессии. Приведите типы задач, в которых используется логистическая регрессия. Опишите модель логистической регрессии, реализованную в библиотеке Scikit-Learn.
5. Раскройте понятие дерева принятия решений. Приведите типы задач, в которых используется дерево принятия решений. Опишите модель дерева принятия решений, реализованную в библиотеке Scikit-Learn.
6. Раскройте понятие кластеризации. Приведите типы задач, в которых используется кластеризация. Опишите реализацию модели кластеризации в библиотеке Scikit-Learn.
7. Раскройте понятие бинарной классификации. Приведите типы задач, в которых используется бинарная классификация. Опишите реализации моделей бинарной классификации в библиотеке Scikit-Learn.
8. Раскройте понятие многоклассовой классификации. Приведите типы задач, в которых используется многоклассовая классификация. Опишите модели многоклассовой классификации, реализованные в библиотеке Scikit-Learn.
9. Раскройте основные достоинства и недостатки известных вам моделей кластеризации.

10. Раскройте понятие метрики точности модели. Приведите известные вам метрики точности и опишите реализацию их вычисления на языке Python.

8. Перечень основной и дополнительной литературы, необходимой для освоения дисциплины

Основная литература:

1. Маккинли, У. Python и анализ данных / Уэс Маккинли. – Саратов: Профобразование, 2019. – 482 с.
2. Протоdjяконов, А.В. Алгоритмы Data Science и их практическая реализация на Python: учебное пособие / А. В. Протоdjяконов, П. А. Пылов. – Садовников В.Е. – М., Вологда: Инфра-Инженерия, 2022. – 392 с. – URL: <https://www.iprbookshop.ru/124000.html>
3. Гудфеллоу Я., Бенджио И., Курвилль А. Глубокое обучение / пер. с англ. А. А. Слинкина. – 2-е изд., испр. – М.: ДМК Пресс, 2018. – 652 с. – URL: <https://e.lanbook.com/book/107901.html>
4. Хливненко, Л. В. Практика нейросетевого моделирования: учебное пособие для вузов / Л. В. Хливненко, Ф. А. Пятакович. – СПб: Лань, 2021. – 200 с. – URL: <https://e.lanbook.com/book/173811.html>
5. Мюллер, А, Гвидо, С. Введение в машинное обучение с помощью Python. Руководство для специалистов по обработке данных / Андреас Мюллер, Сара Гвидо. – СПб.: ООО “Альфа-книга”, 2017. – 480 с.

Дополнительная литература:

6. Клетте, Р. Компьютерное зрение. Теория и алгоритмы: учебник / Рейнхальд Клетте; перевод с англ. А. А. Слинкина. – М: ДМК Пресс, 2019. – 506 с. – URL: <https://e.lanbook.com/book/131691.html>
7. Грас, Д. Data Science. Наука о данных с нуля / Джоэл Грас. – СПб: БХВ-Петербург, 2020. – 411 с.

9. Перечень ресурсов сети Интернет, необходимых для освоения дисциплины

1. Электронная библиотека Финансового университета <http://elib.fa.ru/>
2. Электронно-библиотечная система BOOK.RU <http://www.book.ru>
3. Электронно-библиотечная система «Университетская библиотека ОНЛАЙН» <http://biblioclub.ru/>
4. Электронно-библиотечная система издательства «ЮРАЙТ» <https://www.urait.ru/>
5. Электронно-библиотечная система издательства «Лань» <https://e.lanbook.com/>
6. Библиотека Scikit-Learn <https://scikit-learn.org>
7. Создание приложений с использованием машинного обучения Hugging-Face <https://huggingface.co/>
8. Социальная сеть специалистов по обработке данных и машинному обуче-

нию Kaggle <https://www.kaggle.com>

9. Библиотека для машинного обучения AutoGluon <https://auto.gluon.ai>

10. Методические указания для обучающихся по освоению дисциплины

Рекомендации по освоению дисциплины приведены в «Методических рекомендациях для студентов бакалавриата по освоению дисциплин образовательных программ высшего образования», утвержденных приказом № 1040 ректора Финуниверситета от 11 мая 2021 г.

11. Перечень информационных технологий, используемых при осуществлении образовательного процесса по дисциплине, включая перечень необходимого программного обеспечения и информационных справочных систем

11.1 Комплект лицензионного программного обеспечения:

1. Операционная система – Windows 8 или выше, Linux.
2. Среды разработки – Google Colaboratory и Visual Studio Code.

11.2 Современные профессиональные базы данных и информационные справочные системы:

№	Название рекомендуемых технических и компьютерных средств обучения	Наименование разделов и тем
1	Правовая база данных «КонсультантПлюс»	Все темы
2	Справочно-правовая система «Гарант»	Все темы
3	www.skrin.ru – Система комплексного раскрытия информации «СКРИН»	Все темы
4	http://www.iteam.ru/publications/strategy – Технологии корпоративного управления	Все темы
5	Информационная система СПАРК	Все темы
6	Информационная система Bloomberg	Все темы
7	Информационная система Thomson Reuters	Все темы
8	https://spravochnick.ru/informacionnye_tehnologii/ – Информационные технологии	Все темы

11.3 Сертифицированные программные и аппаратные средства защиты информации: не предусмотрены

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

Помещения для проведения лекций, семинарских занятий, групповых и индивидуальных консультаций, текущего контроля и промежуточной аттестации, а также помещения для самостоятельной работы.