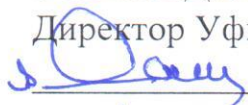


Федеральное государственное образовательное бюджетное
учреждение высшего образования
**«ФИНАНСОВЫЙ УНИВЕРСИТЕТ
ПРИ ПРАВИТЕЛЬСТВЕ РОССИЙСКОЙ ФЕДЕРАЦИИ»**
(Финансовый университет)

Уфимский филиал Финуниверситета
Кафедра «Математика и информатика»

УТВЕРЖДАЮ

Директор Уфимского филиала

 Р.М. Сафуанов

« 1 » сентября 2021 г.

МАШИННОЕ ОБУЧЕНИЕ

Рабочая программа дисциплины

для студентов, обучающихся по направлению подготовки
09.03.03 Прикладная информатика,
образовательная программа «Прикладная информатика»,
(ИТ-сервисы и технологии обработки данных в экономике и финансах)

Рекомендовано Ученым советом филиала
(протокол № 39 от «31» августа 2021г.)

Одобрено кафедрой «Математика и информатика»
(протокол № 16 от «30» июня 2021г.)

Уфа 2021

Составитель рабочей программы дисциплины «Машинное обучение» для студентов, обучающихся по направлению подготовки 09.03.03 «Прикладная информатика», образовательная программа: «Прикладная информатика» (ИТ-сервисы и технологии обработки данных в экономике и финансах), доцент кафедры «Математика и информатика» Уфимского филиала Финуниверситета, канд. техн. наук, доцент Фархиева С.А.

ОГЛАВЛЕНИЕ

	Стр.
1. Наименование дисциплины	4
2. Перечень планируемых результатов освоения образовательной программы (перечень компетенций) с указанием индикаторов их достижения и планируемых результатов обучения по дисциплине	4
3. Место дисциплины в структуре образовательной программы	5
4. Объем дисциплины (модуля) 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	6
6. Перечень основной и дополнительной учебной литературы, необходимой для освоения дисциплины	10
7. Перечень ресурсов информационно-телекоммуникационной сети «Интернет», необходимых для освоения дисциплины	10
8. Методические указания для обучающихся по освоению дисциплины	11
9. Перечень информационных технологий, используемых при осуществлении образовательного процесса по дисциплине, включая перечень необходимого программного обеспечения и информационных справочных систем	11
10. Описание материально-технической базы, необходимой для осуществления образовательного процесса по дисциплине	12

1. Наименование дисциплины

Машинное обучение

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

Код компетенции	Наименование компетенции	Индикаторы достижения компетенции	Результаты обучения (умения и знания), соотнесенные с компетенциями / индикаторами достижения компетенции
ОПК-6	Способен анализировать и разрабатывать организационно-технические и экономические процессы с применением методов системного анализа и математического моделирования	1. Системно выбирает математические методы для решения прикладных задач анализа и моделирования в экономической сфере.	Знать: основные понятия и методы машинного обучения, методологию проведения научных исследований в сфере финансовых технологий. Уметь: выбирать методы машинного обучения в рамках проведения научных исследований в профессиональной сфере.
		2. Разрабатывает алгоритмы решения прикладных задач анализа и моделирования в экономической сфере с использованием математических методов.	Знать: основные принципы разработки алгоритмов решения прикладных задач анализа и моделирования обучения в экономической сфере с использованием технологий машинного обучения. Уметь: анализировать, применять результаты анализа, разрабатывать этапы внедрения методов машинного обучения в профессиональной сфере.
		3. Применяет подходы системного анализа при разработке математических и аналитических моделей в экономической сфере с использованием математических методов.	Знать: возможности применения подходов системного анализа при разработке методов машинного обучения, используемые для оценки тенденций экономического развития. Уметь: применять системные подходы методов машинного обучения для решения прикладных задач в области экономики и финансов.
ОПК-8	Способен принимать участие в управлении проектами создания информационных систем на стадиях жизненного цикла	1. Владеет современными методологиями управления проектами в области разработки программного обеспечения.	Знать: основные современные методологии управления проектами с использованием методов и технологий машинного обучения. Уметь: применять методологию интеллектуального анализа данных и выбирать адекватные методы машинного обучения для решения прикладных задач разработки программного обеспечения.

		2. Владеет навыками планирования и организации работ на различных этапах жизненного цикла разработки программного обеспечения.	<p>Знать: основные этапы жизненного цикла разработки программного обеспечения для реализации методов классификации, регрессии, кластеризации и поиска аномалий.</p> <p>Уметь: организовать работы на различных этапах жизненного цикла разработки программного обеспечения методов машинного обучения для решения прикладных задач, интерпретировать полученные результаты, оценивать качество построенных моделей, давать рекомендации по их усовершенствованию и использованию на практике.</p>
--	--	--	---

3. Место дисциплины в структуре образовательных программ

Дисциплина «Машинное обучение» относится к общепрофессиональному циклу обязательной части образовательной программы «Прикладная информатика» (ИТ-сервисы и технологии обработки данных в экономике и финансах) по направлению подготовки 09.03.03 — Прикладная информатика.

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

Таблица 1

Очная форма обучения

Вид учебной работы по дисциплине	Всего (в з/е и часах)	Семестр 4 (в часах)
Общая трудоемкость дисциплины	4/144	144
Контактная работа - Аудиторные занятия	68	68
<i>Лекции</i>	16	16
<i>Семинары, практические занятия</i>	52	52
Самостоятельная работа	76	76
Вид текущего контроля	Контрольная работа	Контрольная работа
Вид промежуточной аттестации	Экзамен	Экзамен

Заочная форма обучения

Вид учебной работы по дисциплине	Всего (в з/е и часах)	Семестр 5 (в часах)
Общая трудоемкость дисциплины	4/144	144
Контактная работа - Аудиторные занятия	14	14
<i>Лекции</i>	2	2
<i>Семинары, практические занятия</i>	12	12
Самостоятельная работа	130	130
Вид текущего контроля	Контрольная работа	Контрольная работа
Вид промежуточной аттестации	Экзамен	Экзамен

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

5.1. Содержание дисциплины

1. От данных к ценности: основы наук о данных

Основные определения: интеллектуальный анализ данных, большие данные, машинное обучение.

Методы и задачи интеллектуального анализа данных, машинного обучения и обработки больших данных.

Области применения методов и технологий интеллектуального анализа данных, машинного обучения и обработки больших данных.

Примеры задач машинного обучения: поиск информации в интернете, распознавание изображений, лиц, эмоций, пола, возраста, ..., распознавание речи, языка, эмоциональной окраски текстов, прогнозирование продаж, прогнозирование оттока клиентов, кредитный скоринг, рекомендательные системы и др.

Основные характеристики больших данных и их влияние на сбор, хранение, обработку и анализ данных (4V).

Критерии аналитических задач, решение которых предпочтительно с использованием технологий Big Data.

Цикл обработки данных: поиск данных, сбор данных, очистка данных, трансформация данных, интеллектуальный анализ данных, интерпретация и практическое применение результатов.

Статистические основы обработки данных: краткий обзор выборочного метода, методов описательной статистики, построения интервальных оценок и проверки гипотез.

Обзор современных технологий машинного обучения: R, Python, Spark, Microsoft Azure ML.

2. Практическое использование моделей классификации

Принципы разработки и оценки систем машинного обучения. Основные классы моделей машинного обучения: обучение с учителем (классификация и регрессия) и без учителя (кластеризация и поиск аномалий).

Задача классификация с обучением. Задача кредитного скоринга.

Модель логистической регрессии и ее компьютерная реализация.

Методы оценки качества моделей классификации: доля правильных ответов, точность, полнота, F1, AUC. Цены ошибок первого и второго рода.

Модели деревьев и лесов решений. Усиление (бустинг) деревьев решений. Нейронные сети. Компьютерная реализация методов классификации.

Недообучение и переобучение в моделях классификации.

Основы технологии улучшения моделей машинного обучения. Подбор оптимальных параметров моделей. Важность подготовки данных. Генерация синтетических признаков. Работа с пропущенными данными. Работа с несбалансированными выборками.

3. Практическое использование моделей регрессии

Модель множественной линейной регрессии. Методы оценки качества моделей регрессии. Коэффициент детерминации, средняя абсолютная ошибка предсказания, средняя относительная ошибка предсказания.

Методы классификации и их компьютерная реализация.

Недообучение и переобучение в моделях регрессии.

Основы технологии улучшения моделей машинного обучения в задачах регрессии.

Задача прогнозирования продаж.

4. Практическое использование моделей кластерного анализа и поиска аномалий

Кластерный анализ и его компьютерная реализация. Задача сегментирования потребителей. Иерархические агломеративные алгоритмы. Метод К-средних.

Методы оценки качества моделей кластерного анализа. Расстояния между кластерами, расстояния между объектами внутри кластеров. Характеристики центров кластеров.

Задача снижения размерности факторного пространства. Метод главных компонент и его компьютерная реализация. Геометрическое и экономическое содержание метода главных компонент. Использование методов снижения размерности для улучшения качества моделей машинного обучения. Использование методов снижения размерности для выявления латентных факторов.

Поиск аномалий. Задача поиска мошеннических транзакций.

5. Анализ текстов, изображений, эмоций и рекомендательные системы

Принципы анализа текстовой информации.

Принципы анализа графической информации.

Принципы анализа эмоциональной окраски текстов.

Принципы создания рекомендательных систем.

Интеллектуальные сервисы и чат-боты.

Перспективы развития систем обработки больших данных и машинного обучения.

Финансовые технологии, основанные на обработке данных и машинном обучении: интеллектуальные кредитные сервисы, интеллектуальные страховые сервисы, интеллектуальные сервисы интернета вещей.

Современные практические исследования Финансового университета в области обработки данных и машинного обучения: индикаторы смены состояний финансовых рынков, анализ сложных сетей в управлении городами, энергетике, на транспорте, предиктивный ремонт оборудования, системы скоринга.

5.2. Учебно-тематический план

Очная/заочная форма обучения

Таблица 2

№ п/п	Наименование тем (разделов) дисциплины	Трудоёмкость в часах					Формы теку- щего контроля успеваемости
		Всего	Контактная работа - Аудиторная работа			Само- стоя- тельная работа	
			Общая	Лек- ции	Семина- ры, прак- тические занятия		
1	От данных к цен- ности: Основы наук о данных	22	10/1	2/1	8/0	12/21	Лабораторные работы (постро- ение компью- терной системы визуализаци и данных, анализ конкретной си- туации и обсуж- дение результа- тов)
2	Практическое использование моделей классификации	28	12/3	2/1	10/2	16/25	
3	Практическое ис- пользование моде- лей регрессии	32	16/4	4/0	12/4	16/28	
4	Практическое ис- пользование моде- лей кластерного анализа и поиска аномалий	32	16/4	4/0	12/4	16/28	

5	Анализ текстов, изображений, эмоций и рекомендательные системы	30	14/2	4/0	10/2	16/28	
	В целом по дисциплине	144	68/14	16/2	52/12	76/130	Согласно учебному плану: контрольная работа

5.3. Содержание семинаров, практических занятий

Таблица 3

Наименование тем (разделов) дисциплины	Перечень вопросов для обсуждения на семинарских, практических занятиях, рекомендуемые источники из разделов 6,7 (указывается раздел и порядковый номер источника)	Формы проведения занятий
1. От данных к ценности: основы наук о данных	Построение компьютерной системы визуализации данных, анализ конкретной ситуации и обсуждение результатов. Построение компьютерной системы визуализации данных, анализ конкретной ситуации и обсуждение результатов Рекомендуемые источники: из раздела 6: 6.1,6.2, 6.3, 6.4. из раздела 7: 7.1- 7.10.	Лабораторная работа, интерактивная форма.
2. Практическое использование моделей классификации	Построение системы прогнозирования продаж. Построение системы кредитного скоринга. Оптимизация системы кредитного скоринга. Рекомендуемые источники: из раздела 6: 6.1,6.2, 6.3, 6.4. из раздела 7: 7.1- 7.10.	Лабораторная работа, интерактивная форма.
3. Практическое использование моделей регрессии	Построение системы прогнозирования продаж. Построение системы прогнозирования продаж» (построение компьютерной системы машинного обучения, анализ конкретной ситуации и обсуждение результатов). Рекомендуемые источники: из раздела 6: 6.1,6.2, 6.3, 6.4. из раздела 7: 7.1- 7.10.	Лабораторная работа, интерактивная форма.

4. Практическое использование моделей кластерного анализа и поиска аномалий	Кластерный анализ и его компьютерная реализация. Иерархические агломеративные алгоритмы. Метод К-средних. Методы оценки качества моделей кластерного анализа. Расстояния между кластерами, расстояния между объектами. Поиск аномалий. Задача поиска мошеннических транзакций. Сегментирование потребителей. Рекомендуемые источники: из раздела 6: 6.1, 6.2, 6.3, 6.4. из раздела 7: 7.1- 7.10.	Лабораторная работа, интерактивная форма.
5. Анализ текстов, изображений, эмоций и рекомендательные системы	Финансовые технологии, основанные на обработке данных и машинном обучении: интеллектуальные кредитные сервисы, интеллектуальные страховые сервисы, интеллектуальные сервисы интернета вещей. Построение рекомендательного сервиса. Анализ эмоциональной окраски записей в социальных сетях. Рекомендуемые источники: из раздела 6: 6.1, 6.2, 6.3, 6.4. из раздела 7: 7.1- 7.10.	Лабораторная работа, интерактивная форма.

6. Перечень основной и дополнительной учебной литературы, необходимой для освоения дисциплины

Основная литература:

1. Миркин, Б. Г. Введение в анализ данных : учебник и практикум / Б. Г. Миркин. — Москва : Издательство Юрайт, 2020. — 174 с.— URL: <https://ezpro.fa.ru:3217/bcode/450262>
2. Воронина, В. В. Теория и практика машинного обучения : учебное пособие / В. В. Воронина. — Ульяновск : УлГТУ, 2017. — 290 с.— URL: <https://ezpro.fa.ru:3178/book/165053>

Дополнительная литература:

3. Бессмертный, И. А. Системы искусственного интеллекта : учебное пособие для вузов / И. А. Бессмертный. — 2-е изд., испр. и доп. — Москва : Издательство Юрайт, 2021. — 157 с.— URL: <https://ezpro.fa.ru:3217/bcode/470638>
4. Станкевич, Л. А. Интеллектуальные системы и технологии : учебник и практикум для вузов / Л. А. Станкевич. — Москва : Издательство Юрайт, 2021. — 397 с. — URL: <https://ezpro.fa.ru:3217/bcode/469517>

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

1. Электронная библиотека Финансового университета (ЭБ) <http://elib.fa.ru/>

2. Электронно-библиотечная система BOOK.RU <http://www.book.ru>
3. Электронно-библиотечная система «Университетская библиотека ОН-ЛАЙН» <http://biblioclub.ru/>
4. Электронно-библиотечная система Znanium <http://www.znanium.com>
5. Электронно-библиотечная система издательства «ЮРАЙТ» <https://urait.ru/>
6. Электронно-библиотечная система издательства Проспект <http://ebs.prospekt.org/books>
7. Электронно-библиотечная система издательства «Лань» <https://e.lanbook.com/>
8. Электронная библиотека Издательского дома «Гребенников» <https://grebennikon.ru/>
9. Деловая онлайн-библиотека Alpina Digital <http://lib.alpinadigital.ru/>
10. Научная электронная библиотека eLibrary.ru <http://elibrary.ru>

8. Методические указания для обучающихся по освоению дисциплины

Наименование методических материалов для обучающихся	Год утверждения	Местонахождение материала (ссылка на ИОП, информационный стенд кафедры/филиала, др.)
Методические указания к лекциям	2021	http://www.fa.ru/fil/ufa/about/ums/Pages/info.aspx
Методические указания к практическим занятиям	2021	http://www.fa.ru/fil/ufa/about/ums/Pages/info.aspx
Методические указания к самостоятельной работе	2021	http://www.fa.ru/fil/ufa/about/ums/Pages/info.aspx
Методические указания к контрольной работе	2021	http://www.fa.ru/fil/ufa/about/ums/Pages/info.aspx

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

9.1. Комплект лицензионного программного обеспечения:

Продукты компании Microsoft, включая ОС Windows и Office.

9.2. Современные профессиональные базы данных и информационные справочные системы

Электронное периодическое издание Справочная Правовая Система Консультант Бюджетные организации: версия Проф.

9.3. Сертифицированные программные и аппаратные средства защиты информации

Сертифицированные программные и аппаратные средства защиты информации – не используются.

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

Учебная аудитория для проведения всех видов занятий, предусмотренных программой бакалавриата, оснащенная оборудованием и техническими средствами обучения.