

Федеральное государственное образовательное бюджетное
учреждение высшего образования
«Финансовый университет при Правительстве Российской
Федерации»
(Финансовый университет)
Колледж информатики и программирования

УТВЕРЖДАЮ

Заместитель директора по
учебной работе

 Н.Ю. Долгова
« 30 » июня 2022г.

РАБОЧАЯ ПРОГРАММА УЧЕБНОЙ ДИСЦИПЛИНЫ

ОП.14 ОСНОВЫ МАШИННОГО ОБУЧЕНИЯ

09.02.07 Информационные системы и программирование

Москва 2022г.

Рабочая программа учебной дисциплины разработана на основе Федерального государственного образовательного стандарта среднего профессионального образования (далее – ФГОС) по специальности среднего профессионального образования (далее – СПО) специальности 09.02.07 Информационные системы и программирование.

Разработчики:

Морозова М.В., преподаватель Колледжа информатики и программирования

Рецензент:

Эдгулова Елизавета Каральбиевна, кандидат физико – математических наук, преподаватель колледжа информационных технологий и экономики КБГУ

(ФИО, ученая степень, звание, должность)

Рабочая программа учебной дисциплины рассмотрена и рекомендована к утверждению на заседании предметной (цикловой) комиссии Информационных систем и программирования
(наименование ПЦК)

Протокол от « 12 » мая 2022г. № 10

Председатель ПЦК  Н.Г. Титов
(подпись)

СОДЕРЖАНИЕ

1. ОБЩАЯ ХАРАКТЕРИСТИКА РАБОЧЕЙ ПРОГРАММЫ УЧЕБНОЙ ДИСЦИПЛИНЫ
2. СТРУКТУРА И СОДЕРЖАНИЕ УЧЕБНОЙ ДИСЦИПЛИНЫ
3. УСЛОВИЯ РЕАЛИЗАЦИИ УЧЕБНОЙ ДИСЦИПЛИНЫ
4. КОНТРОЛЬ И ОЦЕНКА РЕЗУЛЬТАТОВ ОСВОЕНИЯ УЧЕБНОЙ ДИСЦИПЛИНЫ

1. ОБЩАЯ ХАРАКТЕРИСТИКА РАБОЧЕЙ ПРОГРАММЫ УЧЕБНОЙ ДИСЦИПЛИНЫ

1.1. Место дисциплины в структуре основной образовательной программы

Учебная дисциплина «ОП.14 Основы машинного обучения» является вариативной частью общепрофессионального цикла основной профессиональной образовательной программы в соответствии с ФГОС по специальности 09.02.07 Информационные системы и программирование.

Учебная дисциплина «ОП.14 Основы машинного обучения» обеспечивает формирование профессиональных и общих компетенций по всем видам деятельности ФГОС специальности 09.02.07 Информационные системы и программирование. Особое значение дисциплина имеет при формировании и развитии общих, профессиональных компетенций и личностных результатов:

1.1.1. Перечень общих компетенций

Код	Наименование общих компетенций
ОК 01	Выбирать способы решения задач профессиональной деятельности, применительно к различным контекстам.
ОК 02	Осуществлять поиск, анализ и интерпретацию информации, необходимой для выполнения задач профессиональной деятельности.
ОК 03	Планировать и реализовывать собственное профессиональное и личностное развитие.
ОК 04	Работать в коллективе и команде, эффективно взаимодействовать с коллегами, руководством, клиентами.
ОК 09	Использовать информационные технологии в профессиональной деятельности.
ОК 10	Пользоваться профессиональной документацией на государственном и иностранном языках.

1.1.2. Перечень профессиональных компетенций

Код	Наименование профессиональных компетенций
ПК 6.5	Осуществлять техническое сопровождение, обновление и восстановление данных информационной системы в соответствии с техническим заданием.
ПК 10.1	Обрабатывать статический и динамический информационный контент.
ПК 10.2	Разрабатывать технические документы для управления информационными ресурсами.
ПК 11.1	Осуществлять сбор, обработку и анализ информации для проектирования баз данных.

1.1.3. Перечень личностных результатов

Код	Наименование личностных результатов
ЛР 4	Проявляющий и демонстрирующий уважение к людям труда, осознающий ценность собственного труда. Стремящийся к формированию в сетевой среде личностного и профессионального конструктивного «цифрового следа».
ЛР 9	Соблюдающий и пропагандирующий правила здорового и безопасного образа жизни, спорта; предупреждающий либо преодолевающий зависимости от алкоголя, табака, психоактивных веществ, азартных игр и т.д. Сохраняющий психологическую устойчивость в ситуативно сложных или стремительно меняющихся ситуациях.
ЛР 13	Демонстрирующий умение эффективно взаимодействовать в команде, вести диалог, в том числе с использованием средств коммуникации
ЛР 14	Демонстрирующий навыки анализа и интерпретации информации из различных источников с учетом нормативно-правовых норм
ЛР 15	Демонстрирующий готовность и способность к образованию, в том числе самообразованию, на протяжении всей жизни; сознательное отношение к непрерывному образованию как условию успешной профессиональной и общественной деятельности.
ЛР 16	Соответствующий ожиданиям работодателей: креативно мыслящий, эффективно сотрудничающий с другими людьми, осознанно выполняющий профессиональные требования, распределяющий время и другие ресурсы для выполнения поставленной задачи в установленный срок, ответственный, дисциплинированный, целеустремленный, стрессоустойчивый.
ЛР 18	Демонстрирующий способность использовать в цифровой среде различные цифровые средства, позволяющие во взаимодействии с другими людьми

	достигать поставленных целей; предупреждающий собственное и чужое деструктивное поведение в сетевом пространстве.
--	---

1.2. Цель и планируемые результаты освоения дисциплины:

В рамках программы учебной дисциплины обучающимися осваиваются умения и знания

Код ОК, ПК	Умения	Знания
ОК 01 ОК 02 ОК 03 ОК 04 ОК 09 ОК 10 ПК 6.5 ПК 10.1 ПК 10.2 ПК 11.1	<ul style="list-style-type: none"> – работать с различными источниками данных: CSV, XML и XLS; – подготавливать данные для анализа; – визуализировать результаты анализа; – выбирать оптимальный алгоритм для анализа; – использовать язык R для решения задач машинного обучения; – применять на практике алгоритмы машинного обучения для решения аналитических задач; – создавать аналитические панели; – работать с нейронными сетями. 	<ul style="list-style-type: none"> – языка Python для анализа данных и машинного обучения; – библиотеки NumPy; – библиотеки Pandas; – библиотеки Matplotlib; – среды программирования Jupyter; – основные концепции анализа данных и машинного обучения; – основ языка программирования R; – алгоритмов и задач машинного обучения; – нейронных сетей.

2. СТРУКТУРА И СОДЕРЖАНИЕ УЧЕБНОЙ ДИСЦИПЛИНЫ

2.1. Объем учебной дисциплины и виды учебной работы

Вид учебной работы	Объем в часах
Объем образовательной программы учебной дисциплины	112/64
Объем работы обучающихся во взаимодействии с преподавателем	94/58
в том числе:	
теоретическое обучение	52/16
практические занятия	42/42
лабораторные работы	-
контрольные работы	-
курсовая работа (проект) <i>(если предусмотрено)</i>	-
самостоятельная работа	6/6
консультации	4
Промежуточная аттестация в форме экзамена	8

2.2. Тематический план и содержание учебной дисциплины

Наименование разделов и тем	Содержание учебного материала и формы организации деятельности обучающихся	Объем в часах	Коды компетенций, формированию которых способствует элемент программы
1	2	3	4
Раздел 1. Начало анализа данных		30/21	
Тема 1.1. Основы анализа данных	Содержание учебного материала	12/10	ОК 01-03, ОК 09 ПК 10.1, ПК 11.1, ЛР 4, ЛР 9, ЛР 13, ЛР 14, ЛР 15, ЛР 16, ЛР 18
	1. Основные концепции анализа данных.	1/0	
	2. Основы работы с Jupyter Notebook.	1/0	
	3. Библиотека NumPy. Полезные инструменты.	2/2	
	4. Библиотека Pandas. Возможности для Data Science.	2/2	
	В том числе практических занятий:	6/6	
	1. Практическое занятие «Использование библиотеки NumPy».	2/2	
Тема 1.2. Предобработка данных	2. Практическое занятие «Использование библиотеки Pandas».	4/4	ОК 01-03, ОК 09, ОК 10, ПК 10.1, ПК 11.1, ЛР 4, ЛР 14, ЛР 15, ЛР 16, ЛР 18
	Содержание учебного материала	6/4	
	1. Очистка данных от выбросов, пропусков и дубликатов	2/1	
	2. Преобразование разных форматов данных	2/1	
	В том числе практических занятий:	2/2	
Тема 1.3. Исследовательский и статистический анализ данных	3. Практическое занятие «Анализ клиентов банка».	2/2	ОК 01-03, ОК 09 ПК 10.1, ПК 11.1, ЛР 4, ЛР 14, ЛР 15, ЛР 16, ЛР 18
	Содержание учебного материала	12/7	
	1. Исследование основных свойств данных, поиск закономерностей, распределений и аномалий	2/0	
	2. Библиотеки SciPy и Matplotlib	2/1	
	3. Анализ взаимосвязей в данных методами статистики	2/0	
	В том числе практических занятий:	6/6	
	4. Практическое занятие «Анализ популярности заправок».	4/4	
Раздел 2.	Основы машинного обучения	48/29	ОК 01-03, ОК 09 ПК 10.1, ПК 11.1, ЛР 4, ЛР 9, ЛР 13,
	Тема 2.1. Введение в машинное обучение	8/4	
	1. Основные концепции машинного обучения	2/0	

	2. Задачи классификации и регрессии	2/0	ЛР 14, ЛР 15, ЛР 16, ЛР 18
	В том числе практических занятий:	4/4	
	6. Практическое занятие «Создание первого проекта с машинным обучением»	2/2	
	7. Практическое занятие «Прогноз вероятности ухода клиента из банка»	2/2	
Тема 2.2. Вспомогательные инструменты Data Science	Содержание учебного материала	8/7	ОК 01-03, ОК 09, ПК 6.5, ПК 10.1, ПК 10.2, ПК 11.1, ЛР 4, ЛР 9, ЛР 13, ЛР 14, ЛР 15, ЛР 16, ЛР 18
	1. Работа с bash, virtualenv, docker.	1/0	
	2. Управление git-репозиторием	1/1	
	Самостоятельная работа обучающихся: <i>Работа с git-репозиторием</i>	6/6	
Тема 2.3. Математика машинного обучения	Содержание учебного материала	14/6	ОК 01-02, ОК 09, ПК 10.1, ПК 11.1, ЛР 4, ЛР 13, ЛР 14, ЛР 15, ЛР 16, ЛР 18
	1. Алгоритмы и структуры данных: сложность алгоритма, алгоритмы на графах, динамическое программирование	2/0	
	2. Линейная алгебра: векторы, матрицы, расстояния	2/0	
	3. Численные методы: приближенные алгоритмы, алгоритмы оптимизации, градиентный спуск	2/0	
	4. Алгоритмы машинного обучения: решающие деревья, бустинг и бэггинг, линейные и модели	2/0	
	В том числе практических занятий:	6/6	
	8. Практическое занятие «Метод преобразования данных для защиты личной информации клиентов»	4/4	
	9. Практическое занятие «Разработка модели для определения стоимости автомобиля с пробегом»	2/2	
Тема 2.4. Системы хранения данных	Содержание учебного материала	6/6	ОК 01-04, ОК 09, ОК 10, ПК 6.5, ПК 10.1, ПК 10.2, ПК 11.1, ЛР 4, ЛР 9, ЛР 13, ЛР 14, ЛР 15, ЛР 16, ЛР 18
	1. Анализ данных на SQL	2/2	
	2. Методы библиотеки PySpark	2/2	
	В том числе практических занятий:	2/2	
	10. Практическое занятие «Анализ спроса на авиабилеты в города, где проводятся фестивали»	2/2	
Тема 2.5. Обучение без учителя	Содержание учебного материала	2/0	ОК 01-03, ОК 09, ПК 10.1, ПК 11.1, ЛР 4, ЛР 9, ЛР 13,
	1. Задачи кластеризации	1/0	
	2. Поиск аномалий	1/0	

			ЛР 14, ЛР 15, ЛР 16, ЛР 18
Тема 2.6. Машинное обучение для текстов	Содержание учебного материала	10/6	ОК 01-03, ОК 09 ПК 10.1, ПК 10.2, ПК 11.1, ЛР 4, ЛР 9, ЛР 13, ЛР 14, ЛР 15, ЛР 16, ЛР 18
	1. Алгоритм TF-IDF	2/0	
	2. Языковые представления word2vec и BERT	2/0	
	В том числе практических занятий:	6/6	
	11. Практическая работа «Оценка токсичности комментариев»	6/6	
Раздел 3. Основы deep learning		22/14	
Тема 3.1. Компьютерное зрение	Содержание учебного материала	22/14	ОК 01-04, ОК 09, ОК 10, ПК 10.1, ПК 11.1, ЛР 4, ЛР 9, ЛР 13, ЛР 14, ЛР 15, ЛР 16, ЛР 18
	1. Нейронные сети	2/0	
	2. Метод градиентного спуска	2/0	
	3. Регуляризация нейронных сетей	2/0	
	4. Свёрточные нейронные сети	2/0	
	5. Библиотека Keras	2/2	
	6. Библиотека Tensorflow	2/2	
	В том числе практических занятий:	10/10	
	12. Практическое занятие «Построение модели анализа фотографии»	6/6	
	13. Практическое занятие «Оптимизация нейронной сети»	4/4	
Консультация		4	
Промежуточная аттестация в форме экзамена		8	
Всего:		112/64	

3. УСЛОВИЯ РЕАЛИЗАЦИИ ПРОГРАММЫ УЧЕБНОЙ ДИСЦИПЛИНЫ

3.1. Материально-техническое обеспечение

Для реализации программы учебной дисциплины должны быть предусмотрены следующие специальные помещения (в соответствии с ФГОС и ПООП): лаборатория программирования и баз данных

- оборудование учебного кабинета: 25 рабочих мест кабинета, рабочее место преподавателя (АРМ), парты для обучающихся.
- технические средства обучения: мультимедиа-проектор, ПК с подключением к сети Интернет.

- оборудование лаборатории и рабочих мест лаборатории: ПК по количеству учащихся, подключение к локальной сети, подключение к сети Интернет.

Лицензионное программное обеспечение общего и профессионального назначения.

3.2. Информационное обеспечение реализации программы

Для реализации программы библиотечный фонд образовательной организации должен иметь печатные и/или электронные образовательные и информационные ресурсы, рекомендуемых для использования в образовательном процессе

3.2.1. Печатные издания

Законодательные и нормативные акты:

1. ГОСТ 7.1. – 2003. Библиографическая запись. Библиографическое описание. Общие требования и правила составления. – М.: ИПК Издательство стандартов, 2004. – 169 с.
2. ГОСТ 7.32 – 2001. Система стандартов по информацию, библиотечному и издательскому делу. Отчет о научно-исследовательской работе. Структура и правила оформления. – М.: ИПК Издательство стандартов, 2001. – 21 с.
3. ГОСТ 7.82-2001 Библиографическая запись. Библиографическое описание электронных ресурсов. Общие требования и правила составления. – М.: ИПК Издательство стандартов, 2001. – 21 с.
4. Единая система программной документации. – М.: Стандартинформ, 2005. – 128 с.

3.2.2. Основная литература:

1. Кондрашов, Ю.Н., Анализ данных и машинное обучение на платформе MS SQL Server : учебное пособие / Ю.Н. Кондрашов. — Москва : Русайнс, 2021. — 303 с. — ISBN 978-5-4365-7924-5. —

URL:<https://book.ru/book/941049> (дата обращения: 08.06.2022). — Текст : электронный.

2. Воронина, В. В. Теория и практика машинного обучения : учебное пособие / В. В. Воронина. — Ульяновск : УлГТУ, 2017. — 290 с. — ISBN 978-5-9795-1712-4. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/165053> (дата обращения: 08.06.2022). — Режим доступа: для авториз. пользователей.

3.2.3. Электронные издания (электронные ресурсы):

1. Распределенные представления слов и фраз Миколов Томас. — [Электронный ресурс]. — Режим доступа: <https://papers.nips.cc/paper/5021-distributed-representations-of-words-and-phrases-and-their-compositionality>
2. Draw.io — [Электронный ресурс]. — Режим доступа: <https://startpack.ru/application/draw-io663>
3. Инструкция по работе с TensorFlow Object Detection API. — [Электронный ресурс]. — Режим доступа: <https://habr.com/company/nixsolutions/blog>
4. Многопоточность на примерах – модуль threading. — [Электронный ресурс]. — Режим доступа: <https://python-scripts.com/threading>
5. Метаклассы и метапрограммирование в Python. — [Электронный ресурс]. — Режим доступа: <https://gitjournal.tech/metaklassy-i-metaprogrammirovanie-v-python/>
6. Keras: the Python deep learning API. — [Электронный ресурс]. — Режим доступа: <https://keras.io/>
7. Библиотеки для глубокого обучения: Keras. — [Электронный ресурс]. — Режим доступа: <https://habr.com/ru/company/ods/blog/325432/>
8. Методы оптимизации нейронных сетей. — [Электронный ресурс]. — Режим доступа: <https://habr.com/ru/post/318970/>

4. КОНТРОЛЬ И ОЦЕНКА РЕЗУЛЬТАТОВ ОСВОЕНИЯ УЧЕБНОЙ ДИСЦИПЛИНЫ

Контроль и оценка результатов освоения учебной дисциплины осуществляется преподавателем в процессе проведения практических занятий и лабораторных работ, тестирования, а также выполнения обучающимися индивидуальных заданий, проектов, исследований.

Результаты (освоенные общие компетенции)	Основные показатели оценки результата	Формы и методы контроля и оценки
<p><i>Перечень умений, осваиваемых в рамках дисциплины:</i></p> <ul style="list-style-type: none"> – работать с различными источниками данных: CSV, XML и XLS; – подготавливать данные для анализа; – визуализировать результаты анализа; – выбирать оптимальный алгоритм для анализа; – использовать язык R для решения задач машинного обучения; – применять на практике алгоритмы машинного обучения для решения аналитических задач; – создавать аналитические панели; – работать с нейронными сетями. 	<p>«Отлично» - теоретическое содержание курса освоено полностью, без пробелов, умения сформированы, все предусмотренные программой учебные задания выполнены, качество их выполнения оценено высоко.</p> <p>«Хорошо» - теоретическое содержание курса освоено полностью, без пробелов, некоторые умения сформированы недостаточно, все предусмотренные программой учебные задания выполнены, некоторые виды заданий выполнены с ошибками.</p> <p>«Удовлетворительно» - теоретическое содержание курса освоено частично, но пробелы не</p>	<ul style="list-style-type: none"> – компьютерное тестирование по отдельным темам дисциплины; – текущий контроль в форме защиты практических работ; – экзамен по дисциплине.

<p><i>Перечень знаний, осваиваемых в рамках дисциплины:</i></p> <ul style="list-style-type: none"> – язык Python для анализа данных и машинного обучения; – библиотека NumPy; – библиотека Pandas; – библиотека Matplotlib; – среда программирования Jupyter; – основные концепции анализа данных и машинного обучения; – основ языка программирования R; – алгоритмы и задачи машинного обучения; – нейронные сети. 	<p>носят существенного характера, необходимые умения работы с освоенным материалом в основном сформированы, большинство предусмотренных программой обучения учебных заданий выполнено, некоторые из выполненных заданий содержат ошибки.</p> <p>«Неудовлетворительно» - теоретическое содержание курса не освоено, необходимые умения не сформированы, выполненные учебные задания содержат грубые ошибки.</p>	
---	--	--