


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



« 26 » июле 2023г.

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ

ОП.14 ОСНОВЫ МАШИННОГО ОБУЧЕНИЯ

09.02.07 Информационные системы и программирование

Москва 2023г.

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.

1.2. Цель и планируемые результаты освоения дисциплины:

В рамках программы учебной дисциплины студентами осваиваются умения и знания

Код общих и профессиональных компетенций	Умения	Знания
ОК 01	– работать с различными источниками данных: CSV, XML и XLS;	– языка Python для анализа данных и машинного обучения;
ОК 02	– подготавливать данные для анализа;	– библиотеки NumPy;
ОК 03	– визуализировать результаты анализа;	– библиотеки Pandas;
ОК 04	– выбирать оптимальный алгоритм для анализа;	– библиотеки Matplotlib;
ОК 09	– использовать язык R для решения задач машинного обучения;	– среды программирования Jupyter;
ОК 10	– применять на практике алгоритмы машинного обучения для решения аналитических задач;	– основные концепции анализа данных и машинного обучения;
ПК 6.5	– создавать аналитические панели;	– основ языка программирования R;
ПК 10.1	– работать с нейронными сетями.	– алгоритмов и задач машинного обучения;
ПК 10.2		– нейронных сетей.
ПК 11.1		

2. Структура и содержание учебной дисциплины

2.1. Объем дисциплины и виды учебной работы

Вид учебной работы	Объем в часах
Объем образовательной программы учебной дисциплины	112
Объем работы студентов во взаимодействии с преподавателем	94
в том числе:	
теоретическое обучение	52
практические занятия	42
лабораторные занятия	-
контрольные работы	-
курсовая работа (проект) <i>(если предусмотрено)</i>	-
самостоятельная работа	6
консультации	-
Промежуточная аттестация в форме экзамена	12

2.2. Тематический план и содержание учебной дисциплины

Наименование разделов и тем	Содержание учебного материала и формы организации деятельности студентов	Объем в часах	Коды компетенций, формированию которых способствует элемент программы
1	2	3	4
Раздел 1. Начало анализа данных		30	
Тема 1.1. Основы анализа данных	Содержание учебного материала	12	ОК 01-03, ОК 09 ПК 10.1, ПК 11.1,
	1. Основные концепции анализа данных. 2. Основы работы с Jupyter Notebook. 3. Библиотека NumPy. Полезные инструменты. 4. Библиотека Pandas. Возможности для Data Science.	4	
	В том числе практических занятий:	8	
	1. Практическое занятие «Использование библиотеки NumPy».	4	
	2. Практическое занятие «Использование библиотеки Pandas».	4	
Тема 1.2. Предобработка данных	Содержание учебного материала	6	ОК 01-03, ОК 09, ОК 10, ПК 10.1, ПК 11.1
	1. Очистка данных от выбросов, пропусков и дубликатов 2. Преобразование разных форматов данных	4	
	В том числе практических занятий:	2	
	3. Практическое занятие «Анализ клиентов банка».	2	
Тема 1.3. Исследовательский и статистический анализ данных	Содержание учебного материала	12	ОК 01-03, ОК 09 ПК 10.1, ПК 11.1,
	1. Исследование основных свойств данных, поиск закономерностей, распределений и аномалий 2. Библиотеки SciPy и Matplotlib 3. Анализ взаимосвязей в данных методами статистики	6	
	В том числе практических занятий:	6	
	4. Практическое занятие «Анализ популярности заправок».	4	
	5. Практическое занятие «Оптимизация воронок продаж для ускорения работы отдела маркетинга».	2	
Раздел 2.	Основы машинного обучения	46	

Тема 2.1. Введение в машинное обучение	Содержание учебного материала	12	ОК 01-03, ОК 09 ПК 10.1, ПК 11.1,
	1. Основные концепции машинного обучения	4	
	2. Задачи классификации и регрессии		
	В том числе практических занятий:	8	
	6. Практическое занятие «Создание первого проекта с машинным обучением»	4	
Тема 2.2. Вспомогательные инструменты Data Science	7. Практическое занятие «Прогноз вероятности ухода клиента из банка»	4	
	Содержание учебного материала	4	ОК 01-03, ОК 09, ПК 6.5, ПК 10.1, ПК 10.2, ПК 11.1,
1. Работа с bash, virtualenv, docker.	4		
2. Управление git-репозиторием			
Тема 2.3. Математика машинного обучения	Содержание учебного материала	12	ОК 01-02, ОК 09
	1. Алгоритмы и структуры данных: сложность алгоритма, алгоритмы на графах, динамическое программирование	8	
	2. Линейная алгебра: векторы, матрицы, расстояния		
	3. Численные методы: приближенные алгоритмы, алгоритмы оптимизации, градиентный спуск		
	4. Алгоритмы машинного обучения: решающие деревья, бустинг и бэггинг, линейные и модели		
В том числе практических занятий:	4		
Тема 2.4. Системы хранения данных	8. Практическое занятие «Метод преобразования данных для защиты личной информации клиентов»	2	
	9. Практическое занятие «Разработка модели для определения стоимости автомобиля с пробегом»	2	
	Содержание учебного материала	6	ОК 01-04, ОК 09, ОК 10, ПК 6.5, ПК 10.1, ПК 10.2, ПК 11.1,
	1. Анализ данных на SQL		
2. Методы библиотеки PySpark	4		
В том числе практических занятий:	2		
10. Практическое занятие «Анализ спроса на авиабилеты в города, где проводятся фестивали»	2		
	Содержание учебного материала	4	ОК 01-03, ОК 09

Тема 2.5. Обучение без учителя	1. Задачи кластеризации 2. Поиск аномалий	4	ПК 10.1
Тема 2.6. Машинное обучение для текстов	Содержание учебного материала	8	ОК 01-03, ОК 09 ПК 10.1, ПК 10.2, ПК 11.1,
	1. Алгоритм TF-IDF 2. Языковые представления word2vec и BERT	2	
	В том числе практических занятий:	6	
	11. Практическая работа «Оценка токсичности комментариев»	6	
Раздел 3. Основы deep learning		18	
Тема 3.1. Компьютерное зрение	Содержание учебного материала	22	ОК 01-04, ОК 09, ОК 10, ПК 10.1, ПК 11.1,
	1. Нейронные сети 2. Метод градиентного спуска 3. Регуляризация нейронных сетей 4. Свёрточные нейронные сети 5. Библиотека Keras 6. Библиотека Tensorflow	12	
	В том числе практических занятий:	6	
	12. Практическое занятие «Построение модели анализа фотографии»	2	
	13. Практическое занятие «Оптимизация нейронной сети»	4	
Консультация		-	
Промежуточная аттестация в форме экзамена		12	
Всего:		112	

3. Условия реализации программы 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 обучения	Критерии оценки	Методы 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</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>освоено частично, но пробелы не 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>	
---	--	--