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_ 2021г.

РАБОЧАЯ ПРОГРАММА УЧЕБНОЙ ДИСЦИПЛИНЫ

ОП.14 Основы машинного обучения

09.02.07 Информационные системы и программирование

Москва 2021г.

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, квалификация-Программист

Разработчики:
Морозова М.В., преподаватель Колледжа информатики и программирования

Рецензент:
Семенова О.А., директор «ЧПОУ Московкий городской открытый колледж», кандидат педагогических наук.

Рабочая программа учебной дисциплины рассмотрена и рекомендована к утверждению на заседании предметной (цикловой) комиссии
Программирования и баз данных
(наименование ПЦК)

Протокол от «17» июня 2021 г. № 11

Председатель ПЦК _____ А.И. Пестов
(подпись)

РЕЦЕНЗИЯ
НА РАБОЧУЮ ПРОГРАММУ УЧЕБНОЙ ДИСЦИПЛИНЫ
ОП.14 Основы машинного обучения
специальности 09.02.07 Информационные системы и программирование, составленную преподавателем Морозовой М.В.

Рабочая программа, разработанная на основе Федерального государственного образовательного стандарта по специальности 09.02.07 Информационные системы и программирование, рассчитана на 110 часов – учебной нагрузки, из них 52 часа – теоретически занятия, 42 часа практических занятий, 6 часов самостоятельной работы студентов, 2 часа-консультации, промежуточная аттестация в форме экзамена-8 часов.

Рабочая программа включает в себя:
1. Титульный лист;
2. Паспорт рабочей программы по дисциплине;
3. Структуру и содержание;
4. Условия реализации рабочей программы по дисциплине;
5. Контроль и оценку результатов освоения, -

что соответствует типовым требованиям к рабочей программе и требованиям ФГОС СПО при формировании как общих, так и профессиональных компетенций.

Рабочая программа предусматривает выполнение практических работ и приобретение профессиональных компетенций по осуществлению сбора, обработки и анализа информации для проектирования баз данных. Программа ориентирована на формирование умений: 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.

В программе представлены такие формы контроля, как устный опрос, анализ практических работ, решений задач, тестирование для установления уровня обученности по данной дисциплине. Комплекс форм и методов контроля и оценки освоенных компетенций конкретизирован с учетом специфики обучения по программе и образует систему достоверной и объективной оценки результатов его освоения.

Содержание учебного материала соответствует требованиям ФГОС к знаниям и умениям, целям и современным научным представлениям по данной дисциплине и отвечает принципам единства теоретического и практического обучения.

В программе соблюдена логическая последовательность раскрытия материала.

Программа составлена квалифицированно, демонстрирует профессионализм и высокий уровень методической подготовки.

Рецензент Семенова О.А., директор «ЧПОУ Московский городской открытый колледж», кандидат педагогических наук.



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 «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, квалификация Программист.

Учебная дисциплина «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 и профессиональных 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.

ПК 11.1. Осуществлять сбор, обработку и анализ информации для проектирования баз данных.

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 11.1	-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.	-языка Python для анализа данных и машинного обучения; -библиотеки NumPy; -библиотеки Pandas; -библиотеки Matplotlib; -среды программирования Jupyter; -основные концепции анализа данных и машинного обучения; -основы языка программирования R;

Код ОК, ПК	Умения	Знания
		-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	110
Объем работы обучающихся во взаимодействии с преподавателем	94
в том числе:	
теоретическое обучение	52
практические занятия	42
лабораторные работы	-
контрольные работы	-
курсовая работа (проект) *(если предусмотрено)*	-
самостоятельная работа	6
консультации	2
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**	
Тема 1.1. Основы анализа данных	**Содержание учебного материала:**	**12**	ОК 01-03, ОК 09 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.	**1** **1** **2** **2**	
	В том числе практических занятий:	**6**	
	Практическое занятие «Использование библиотеки NumPy». **Практическое занятие** «Использование библиотеки Pandas».	**2** **4**	
Тема 1.2. Предобработка данных	**Содержание учебного материала:**	**6**	ОК 01-03, ОК 09, ОК 10, ПК 11.1
	1. Очистка данных от выбросов, пропусков и дубликатов 2. Преобразование разных форматов данных	**2** **2**	
	В том числе практических занятий:	**2**	
	Практическое занятие «Анализ клиентов банка».	**2**	
Тема 1.3. Исследовательский и статистический анализ данных	**Содержание учебного материала:**	**12**	ОК 01-03, ОК 09, 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	**2** **2** **2**	
	В том числе практических занятий:	**6**	
	Практическое занятие «Анализ популярности заправок». **Практическое занятие** «Оптимизация воронок продаж для ускорения работы отдела маркетинга».	**2** **4**	
Раздел 2. Основы машинного обучения		**48**	
Тема 2.1. Введение в машинное обучение	**Содержание учебного материала:**	**8**	ОК 01-03, ОК 09 ПК 11.1
	1. Основные концепции машинного обучения 2. Задачи классификации и регрессии	**2** **2**	

	В том числе практических занятий:	**4**	
	Практическое занятие «Создание первого проекта с машинным обучением» **Практическое занятие** «Прогноз вероятности ухода клиента из банка»	**2** **2**	
Тема 2.2. Вспомогательные инструменты Data Science	**Содержание учебного материала:**	**10**	ОК 01-03, ОК 09, ПК 11.1
	1.Работа с bash, virtualenv, docker. 2. Управление git-репозиторием	**2** **2**	
	Самостоятельная работа обучающихся: Работа с git-репозиторием	**6**	
Тема 2.3. Математика машинного обучения	**Содержание учебного материала**	**14**	ОК 01-02, ОК 09 ПК 11.1
	1. Алгоритмы и структуры данных: сложность алгоритма, алгоритмы на графах, динамическое программирование 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	**2** **2** **2**	
	В том числе практических занятий:	**6**	
	Практическое занятие «Метод преобразования данных для защиты личной информации клиентов» **Практическое занятие** «Разработка модели для определения стоимости автомобиля с пробегом»	**4** **2**	
Тема 2.4. Системы хранения данных	**Содержание учебного материала:**	**6**	ОК 01-04, ОК 09, ОК 10, ПК 11.1
	1.Анализ данных на SQL 2.Методы библиотеки PySpark	**2** **2**	
	В том числе практических занятий:	**2**	
	Практическое занятие «Анализ спроса на авиабилеты в города, где проводятся фестивали»	**2**	
Тема 2.5. Обучение без учителя	**Содержание учебного материала:**	**2**	ОК 01-03, ОК 09 ПК 11.1
	Задачи кластеризации. Поиск аномалий	**2**	
Тема 2.6. Машинное обучение для текстов	**Содержание учебного материала:**	**8**	ОК 01-03, ОК 09 ПК 11.1
	Алгоритм TF-IDF Языковые представления word2vec и BERT	**2**	

	В том числе практических занятий:	**6**	
	Практическое занятие «Оценка токсичности комментариев»	**6**	
Раздел 3. Основы deep learning		**22**	
Тема 3.1. Компьютерное зрение	**Содержание учебного материала:**	**22**	ОК 01-04, ОК 09, ОК 10, ПК 11.1
	Нейронные сети Метод градиентного спуска Регуляризация нейронных сетей Свёрточные нейронные сети Библиотека Keras Библиотека Tensorflow	**2** **2** **2** **2** **2** **2**	
	В том числе практических занятий:	**10**	
	Практическое занятие «Построение модели анализа фотографии» **Практическое занятие** «Оптимизация нейронной сети»	**6** **4**	
Консультация		**2**	
Промежуточная аттестация в форме экзамена		**8**	
Всего:		**110**	

3. УСЛОВИЯ РЕАЛИЗАЦИИ ПРОГРАММЫ УЧЕБНОЙ ДИСЦИПЛИНЫ

3.1. Материально-техническое обеспечение

Для реализации программы учебной дисциплины должна быть предусмотрена лаборатория программирования и баз данных, оснащенная следующим образом:
1.Специализированная мебель:
Стол студенческий двухместный – 8 шт.
Стол одноместный – 12 шт.
Стулья студенческие – 16 шт.
Стулья компьютерные – 12 шт.
Стол (учительский) – 1 шт.
Стул (учительский) – 1 шт.
Доска маркерная – 1 шт.
Шкаф – 1 шт.
2. Технические средства обучения:
Мультимедиа-проектор – 1 шт.
Экран – 1 шт.
Компьютер студенческий – 12 шт.
Компьютер преподавателя – 1 шт.
3. Перечень лицензионного программного обеспечения:
1) Антивирусная защита: ESET NOD32
2) Windows, Microsoft Office
3) Project Expert, Microsoft SQL Server, Microsoft Visual Studio, 1С:Предприятие (учебная версия), Консультант Плюс, EclipseIDEforJavaEEDevelopers, .NETFrameworkJDK 8, MicrosoftSQLServerExpressEdition, Microsoft Visio Professional, MySQLInstallerforWindows, NetBeans, SQLServerManagementStudio, MicrosoftSQLServerJavaConnector, AndroidStudio, IntelliJIDEA
4. Компьютеры подключены к локальной вычислительной сети, информационно-образовательной среде Финуниверситета и сети Интернет.
5. Учебно-наглядные и методические пособия, учебно-методическая документация.

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

Нормативно-правовые документы:

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. Бринк Х., Ричардс Д., Феверолф М. Машинное обучение. – Спб.: Питер, 2017. – 336 с.
2. Хэдли У., Гарретт Г. Язык R в задачах науки о данных: импорт, подготовка, обработка, визуализация и моделирование данных. – М.: Вильямс, 2018. – 592 с.

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. Анализ тональности текстов с помощью сверточных нейронных сетей. – [Электронный ресурс]. – Режим доступа: https://habr.com/ru/company/mailru/blog/417767/
9. Методы оптимизации нейронных сетей. – [Электронный ресурс]. – Режим доступа: https://habr.com/ru/post/318970/

3.2.4. Дополнительные источники:

1. Чакон С., Штрауб Б. Git для профессионального программиста. – СПб.: Питер, 2016. – 496 с.
2. Geron A. Hands-on machine learning with Scikit-Learn and TensorFlow. – Sebastopol, CA: O'Reilly, 2017. – 751 с.
3. Плас Д.В. Python для сложных задач: наука о данных и машинное обучение. – СПб.: Питер, 2018. – 576 с.
4. Бенгфорт Б., Билбро Р., Охеда Т. Прикладной анализ текстовых данных на Python. Машинное обучение и создание приложений обработки естественного языка. – СПб.: Питер, 2019. – 368 с.

4. КОНТРОЛЬ И ОЦЕНКА РЕЗУЛЬТАТОВ ОСВОЕНИЯ УЧЕБНОЙ ДИСЦИПЛИНЫ

Результаты обучения	Критерии оценки	Методы оценки
Перечень знаний, осваиваемых в рамках дисциплины: – 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. Перечень умений, осваиваемых в рамках дисциплины: – 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.	«Отлично» - теоретическое содержание курса освоено полностью, без пробелов, умения сформированы, все предусмотренные программой учебные задания выполнены, качество их выполнения оценено высоко. «Хорошо» - теоретическое содержание курса освоено полностью, без пробелов, некоторые умения сформированы недостаточно, все предусмотренные программой учебные задания выполнены, некоторые виды заданий выполнены с ошибками. «Удовлетворительно» - теоретическое содержание курса 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. «Неудовлетворительно» - теоретическое содержание курса не освоено, необходимые умения не	компьютерное тестирование по отдельным темам дисциплины; текущий контроль в форме защиты практических работ; анализ результатов наблюдений за деятельностью обучающегося в процессе освоения образовательной программы; экзамен по дисциплине.

	сформированы, выполненные учебные задания содержат грубые ошибки.	