**МИНОБРНАУКИ РОССИИ**

**федеральное государственное бюджетное**

**образовательное учреждение высшего образования**

**«Череповецкий государственный университет»**

|  |
| --- |
| УТВЕРЖДАЮ  Врио ректора  \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ А.Н. Стрижов  М.П.  «\_\_\_\_\_» \_\_\_\_\_\_\_\_\_\_\_\_\_\_ 2020 г. |
| дата утверждения |

**Паспорт Образовательной программы**

«Инструменты анализа данных и машинного обучения»

|  |  |
| --- | --- |
| **Версия программы** | **1** |
| **Дата Версии** | **12.10.2020** |

1. **Сведения о Провайдере**

|  |  |  |
| --- | --- | --- |
| 1.1 | Провайдер | федеральное государственное бюджетное образовательное учреждение высшего образования "Череповецкий государственный университет" |
| 1.2 | Логотип образовательной организации | C:\Users\Привет\Downloads\ЧГУ_лого_вер.jpg |
| 1.3 | Провайдер ИНН | 3528051834 |
| 1.4 | Ответственный за программу ФИО | Ершов Евгений Валентинович |
| 1.5 | Ответственный должность | Директор института информационных технологий |
| 1.6 | Ответственный Телефон | +79211353167 |
| 1.7 | Ответственный Е-mail | evershov@chsu.ru |

1. **Основные Данные**

|  |  |  |
| --- | --- | --- |
| **№** | **Название** | **Описание** |
| 2.1 | Название программы | Инструменты анализа данных и машинного обучения |
| 2.2 | Ссылка на страницу программы | <https://www.chsu.ru/do/pk/dp/-/asset_publisher/k1Dp/content/instrumenty-analiza-dannyh-i-masinnogo-obucenia?> |
| 2.3 | Формат обучения | Онлайн |
|  | Подтверждение от ОО наличия возможности реализации образовательной программы с применением электронного обучения и (или) дистанционных образовательных технологий с возможностью передачи данных в форме элементов цифрового следа | Подтверждаю |
| 2.4 | Уровень сложности | Базовый |
| 2.5 | Количество академических часов | 72 |
|  | Практикоориентированный характер образовательной программы: не менее 50 % трудоёмкости учебной деятельности отведено практическим занятиям и (или) выполнению практических заданий в режиме самостоятельной работы (кол-во академических часов) | 54 |
| 2.6 | Стоимость обучения одного обучающегося по образовательной программе, а также предоставление ссылок на 3 (три) аналогичные образовательные программы иных организаций, осуществляющих обучение, для оценки объективности стоимости или обоснование уникальности представленной образовательной программы в случае отсутствия аналогичных образовательных программ на рынке образовательных услуг | 30 000  <https://cs.hse.ru/dpo/ml>  <https://mipt.ru/cdpo/programs/software/supervised_learning.php?ELEMENT_ID=1855780>  https://www.specialist.ru/course/ds2 |
| 2.7 | Минимальное количество человек на курсе | 100 |
| 2.8 | Максимальное количество человек на курсе | 120 |
| 2.9 | Данные о количестве слушателей, ранее успешно прошедших обучение по образовательной программе | 100 |
| 2.10 | Формы аттестации | Зачет |
|  | Указание на область реализации компетенций цифровой экономики, к которой в большей степени относится образовательная программа, в соответствии с Перечнем областей | Искусственный интеллект |

1. **Аннотация программы**

Цифровая трансформация — это процесс интеграции цифровых технологий во все аспекты бизнес-деятельности. Она означает отказ от устаревших технологий, ускорение производственных и бизнес-процессов, повышение их качества, поддерживаемое информационными технологиями. Она предполагает смещение акцента из центров обработки данных на предприятия, от аналитиков данных к специалистам, сотрудникам предприятий, выполняющим производственные функции.

Предлагаемая программа предназначена для подготовки таких сотрудников, не являющихся аналитиками или статистиками, но способных создавать математические и информационные модели с использованием цифровых инструментов, что может приносить значительный результат, так их глубокое знание предметной области имеет преимущество перед навыками анализа данных. Наличие таких специалистов дает предприятию или организации дополнительные возможности в цифровой трансформации, под которой, частности, понимается оцифровка ключевого оборудования в цепочке создания стоимости. В том числе сбор, хранение и обработка всех значимых данных, связанных как с работой агрегатов, так и с параметрами выпускаемой продукции, создание оптимизационных и предиктивных моделей, цифровых двойников, автоматизация операций управления оборудованием.

Программа совмещает два последовательных модуля. Первый модуль «Анализ данных и поиск решений» ориентирован на развитие компетенций роли «CitizenDataScientist» <https://www.sas.com/ru_ru/insights/articles/analytics/how-to-find-and-equip-citizen-data-scientists.html>. Модуль может рассматриваться как самостоятельный. Второй модуль, «Машинное обучение и техническое зрение», предполагает предварительное изучение модуля «Анализ данных и поиск решений» и предназначен для развития компетенций роли «BusinessTranslator» <https://www.mckinsey.com/business-functions/mckinsey-analytics/our-insights/analytics-translator> .

Для освоения программы слушателям необходимо иметь базовые знания и умения в области информатики.

**МИНОБРНАУКИ РОССИИ**

**федеральное государственное бюджетное**

**образовательное учреждение высшего образования**

**«Череповецкий государственный университет»**

|  |
| --- |
| УТВЕРЖДАЮ  Врио ректора  \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ А.Н. Стрижов  М.П.  «\_\_\_\_\_» \_\_\_\_\_\_\_\_\_\_\_\_\_\_ 2020 г. |
| дата утверждения |

**ДОПОЛНИТЕЛЬНАЯ ПРОФЕССИОНАЛЬНАЯ ПРОГРАММА**

**- программа повышения квалификации**

|  |
| --- |
| Инструменты анализа данных и машинного обучения |
| (наименование программы) |

(72 час)

г. Череповец, 2020 год

**ОБЩАЯ ХАРАКТЕРИСТИКА ПРОГРАММЫ**

**1.Цель программы**

Целью освоения программы является развитие у сотрудников промышленных предприятий информационных компетенций, способствующих цифровой трансформации предприятий и производств, а также организаций, заинтересованных во внедрении цифровых технологий.

**2.Планируемые результаты обучения:**

2.1.Знание (осведомленность в областях)

2.1.1. Принципы алгоритмизации на уровне концепции - машинное обучение, математическая оптимизация, компьютерное зрение, графический анализ данных

2.1.2. Основные типы графиков и способы их интерпретации

2.2. Умение (способность к деятельности)

2.2.1. Транслировать требования бизнеса в формат, понятный для IT-подразделения (способен сформулировать алгоритм, который должен реализовать IT-специалист)

2.2.2. Определять точки применения цифровых инструментов и оценивать потенциал от их внедрения

2.2.3. Проводить графический анализ данных с помощью программного обеспечения (KNIME,R,EXCEL) с целью генерации идей

2.2.4. Решать задачи оптимизации с помощью пакета "Поиск решений" в Excel

2.2.5. Формулировать целевую функцию для оптимизации и перечень ограничений

2.3.Навыки (использование конкретных инструментов)

2.3.1 Навыки использования специализированных программных средств необходимых для расчетов

2.3.2. Навыки построения графиков

2.3.3. Навыки реализации алгоритмов на выбранном уровне

**3. Категория слушателей**

Образование: Лица с высшим или (и) средним профессиональным образованием

Для освоения программы слушателям необходимо иметь базовые знания и умения в области информатики

**4.Учебный план программы «**Инструменты анализа данных и машинного обучения»

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **№ п/п** | **Модуль** | **Всего, час** | **Виды учебных занятий** | | |
| **лекции** | **практические занятия** | **самостоятельная работа** |
| 1 | Анализ данных и поиск решений | 26 | 6 | 16 | 4 |
| 2 | Машинное обучение и техническое зрение | 46 | 12 | 20 | 14 |
| Итоговая аттестация | |  | зачет | | |

**5.Календарный план-график реализации образовательной** программы

|  |  |  |  |
| --- | --- | --- | --- |
| **№ п/п** | **Наименование учебных модулей** | **Трудоёмкость (час)** | **Сроки обучения** |
| **1** | Анализ данных и поиск решений | 26 | 01.11.2020-06.11.2020 |
| **2** | Машинное обучение и техническое зрение  И | 46 | 07.11.2020-15.11.2020 |
| **Всего:** | | 72 | 01.11.2020-15.11.2020 |

**6.Учебно-тематический план программы «**Инструменты анализа данных и машинного обучения**»**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **№ п/п** | **Модуль / Тема** | **Всего, час** | **Виды учебных занятий** | | | **Формы контроля** |
| **лекции** | **практические занятия** | **самостоятельная работа** |
| 1 | Анализ данных и поиск решений | 26 | 6 | 16 | 4 | Входной тест, промежуточный тест, |
| 1.1 | Обзор существующих цифровых технологий обработки данных, их применения. Кто такой CitizenDataScientist. | 6 | 2 | 2 | 2 | Практическая работа  Входной тест |
| 1.2 | Подготовка данных и первичный графический анализ данных. | 16 | 4 | 10 | 2 | Лабораторная работа  Промежуточный тест |
| 1.3 | Итоговое тестирование по модулю 1 | 4 |  | 4 |  | Тест |
| 2 | Машинное обучение и техническое зрение | 46 | 12 | 20 | 14 | Входной тест, промежуточный тест, итоговый тест |
| 2.1 | Машинное обучение. Примеры. Основные этапы решения задачи методом машинного обучения. Методы решения задач | 14 | 4 | 6 | 4 | Лабораторная работа  Входной тест |
| 2.2 | Методы оптимизации. Примеры задач. Применение алгоритмов линейного программирования и эвристических алгоритмов | 14 | 4 | 6 | 4 | Лабораторная работа  Промежуточный тест |
| 2.3 | Компьютерное зрение в Knime | 14 | 4 | 4 | 6 | Лабораторная работа  Промежуточный тест |
| 2.4 | Итоговое тестирование по модулю 2 | 4 |  | 4 |  |  |
| 3 | Итоговая аттестация |  |  |  |  | зачет |
| 4 | Итого | 72 | 18 | 36 | 18 |  |

**7.Учебная (рабочая) программа повышения квалификации «**Инструменты анализа данных и машинного обучения**»**

**Модуль 1.** Анализ данных и поиск решений (26 час.)

**Тема 1.1** Обзор существующих цифровых технологий обработки данных, их применения. Кто такой CitizenDataScientist.**(**6 час.**)**

Обзор существующих цифровых технологий обработки данных, их применения: аналитика больших данных, прогнозирование, кластеризация, машинное обучение. Примеры. Кто такой CitizenDataScientist. Почему он необходим при цифровой трансформации производств.

**Тема 1.2** Подготовка данных и первичный графический анализ данных (16 час.)

Шкалы измерений и классификация данных. Качественные и количественные переменные. Виды статистических графиков. Трудности при сборе данных, важность корректных данных. Описание данных. Меры среднего и их применимость, меры вариации. Зависимость и независимость переменных. Инструменты графического анализа: график плотности распределения, диаграмма рассеяния, матрица корреляций, box-plot, временные ряды и тренды, график «тепловая карта». Программы, позволяющие использовать инструменты (KNIME,R,EXCEL). Оценка достоверности полученных результатов.

**Тема 1.3.** Итоговое тестирование по модулю 1 (4 час.) Итоговый тест

**Модуль 2.** Машинное обучение и техническое зрение**(**46 час.**)**

**Тема 2.1.** Машинное обучение. Примеры. Основные этапы решения задачи методом машинного обучения. Методы решения задач. (14 час.)

Что такое машинное обучение. Примеры задач машинного обучения. Этапы решения задачи методом машинного обучения. Анализ и обработка данных и машинное обучение. Задачи регрессии в машинном обучении (на примере задачи отдачи от алюминия). Линейная регрессия, гребневая регрессия. Задачи классификации в машинном обучении. Расширенный анализ данных.

**Тема 2.2.** Методы оптимизации. Примеры задач. Применение алгоритмов линейного программирования и эвристических алгоритмов. (14 час.)

Что такое оптимизация. Оптимизация в EXCEl. Примеры использования. Формулировка целевой функции, ограничений. Выбор метода. Постановка и методы решения задачи линейного программирования. Симплекс-метод (на примере задачи оптимизации расхода ферросплавов). Эвристические алгоритмы (на примере задачи оптимизации металлозавалки).

**Тема 2.3.** Компьютерное зрение в KNIME. (14 час.)

Примеры задач компьютерного зрения. Основные этапы решения задачи технического зрения. Искусственные нейронные сети (на примере задачи фиксации разливки без защитной трубы). Работа в программе KNIME. Интерфейс и принцип работы программы. Методы и алгоритмы, реализованные в программе KNIME.

**Тема 2.4.** Итоговое тестирование по модулю 2 (4 час.) Итоговый тест

**Описание практико-ориентированных заданий и кейсов**

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Номер темы/модуля** | **Наименование практического занятия** | **Описание** |
| 1.1 | Тема 1.1. | Построение схемы информационных потоков | Постройте схему информационных потоков для своего подразделения (для решения одной из задач). Какие данные вы получаете, собираете, храните, передаете. Где это происходит? Какие из описанных в примерах решений могли бы быть полезны для вас? Что вы могли бы сделать прямо сейчас? |
| 1.2 | Тема 1.2 | Описание данных | Откройте файл с данными и описание ситуации. Опишите переменные, определите распределение, используя Excel, вычислите описательные характеристики, постройте диаграммы и графики, которые уместны для вашего набора данных. Можно ли говорить о зависимости отдельных переменных? Свой ответ подтвердите графически, и, если возможно, расчётами. |
| 1.3 | Тема 2.1. | Постановка задачи машинного обучения | Используя имеющиеся у вас данные и описание ситуации, определите, какую задачу можно решить методами машинного обучения. Проанализируйте, достаточно ли у вас данных. Какие из алгоритмов машинного обучения могут быть вам полезны. Попробуйте применить их, используя KNIME. Оцените достоверность полученного результата. |
| 1.4. | Тема 2.2. | Постановка и решение задачи оптимизации | Используя материал своей задачи, сформулируйте целевую функцию. В «пакете анализа» Excel выберите подходящий метод. Оцените полученный результат с точки зрения точности и достоверности. |
| 1.5. | Тема 2.3. | Прототип системы компьютерного зрения | Разработка прототипа системы компьютерного зрения с помощью программного обеспечения KNIME |

**8.Оценочные материалы по образовательной программе**

**8.1. Вопросы тестирования по модулям**

|  |  |  |  |
| --- | --- | --- | --- |
| **№ модуля** | **Вопросы входного тестирования** | **Вопросы промежуточного тестирования** | **Вопросы итогового тестирования** |
| 1.1 | 1. Какие данные необходимы для выполнения вашей работы:  🞏 Результаты измерений приборов контроля  🞏 Данные финансового контроля  🞏 Значения параметров технологического процесса  🞏 Результаты мониторинга выполнения работ  🞏 Другое  2. Выберите цифровые инструменты, подходящие для вашей работы  🞏 Визуализация данных  🞏 Анализ статистических характеристик  🞏 Регрессионный анализ  🞏 Мне не знакомы эти инструменты  3. Какие диаграммы можно построить, если у вас есть данные по результатам сдачи экзамена?  🞏 Круговая  🞏 Гистограмма  🞏 Столбиковая диаграмма  🞏 Другое  4. По каким значениям строят гистограмму?  🞏 По любым  🞏 Непрерывным, числовым  5. Для обработки данных можно использовать  🞏 Только специальные программы  🞏 Excel  🞏 Paint  🞏 Orange  🞏 Knime  🞏 Не могу сказать | 1. Цифровая трансформация это  🞏 Применение цифровых технологий с целью повышения эффективности деятельности  🞏 Преобразование модели бизнеса, продуктов, услуг предприятия с использованием цифровых технологи  🞏 Внедрение цифровых технологий во все процессы организации  2. Предназначением …(заполнить пропуск)……… является ясное, прозрачное и недвусмысленное представление информации для улучшения понимания и процесса коммуникации. Основной задачей данного направления является преобразование информации на различных уровнях сложности в более доступные для широкой аудитории формы. Выберите пропущенное направление  🞏 информационного дизайна  🞏 визуализации информации  🞏 визуализации данных  🞏 научной визуализации  🞏 визуальной аналитики  🞏 нет ответа  3. Дана частотная гистограмма распределения доходов населения в некотором населённом пункте, полученная по результатам маркетингового исследования. Выберите статистическую характеристику для оценки положения доходов (возможен только один ответ):  image  🞏 Арифметическое среднее  🞏 Мода  🞏 Медиана  🞏 Геометрическое среднее  4. Найдите моду и медиану для данных, представленных частотным распределением   |  |  |  | | --- | --- | --- | | х | f | Mo | | 1 | 7 | Me | | 2 | 9 |  | | 3 | 18 |  | | 4 | 12 |  | | 5 | 6 |  |   5. Статистическое испытание гипотез проводится  🞏 По генеральной совокупности  🞏 По выборочным данным  🞏 Не имеет значения  6. Какая мера центральной тенденции чувствительна к выбросам?  🞏 Среднее  🞏 Мода  🞏 Медиана  7. Что из перечисленного НЕ является мерой разброса?  🞏 Дисперсия  🞏 Медиана  🞏 Стандартное отклонение  🞏 Размах  8. Даны две выборки значений, измеренных в номинальной шкале. Для установления связи между выборками можно использовать:  🞏 Коэффициент Хи-квадрат  🞏 Коэффициент корреляции Пирсона  🞏 Тест Манна-Уитни  🞏 Коэффициент корреляции Спирмена  🞏 нет ответа  9. Какое значение коэффициент корреляции НЕ может принимать?  🞏 -3.56  🞏 -0.78  🞏 1.02  🞏 0.45  🞏 нет ответа  10. Если исходные данные содержат квадратичную зависимость для пары переменных, можно ли использовать уравнение линейной регрессии для аппроксимации этих данных?  🞏 Да  🞏 Нет | 1. Укажите техники визуализации данных, относящиеся к геометрическому типу. Выберите один или несколько ответов:  🞏 Каркасные диаграммы  🞏 Кодирование геометрическими формами  🞏 Параллельные координаты  🞏 Матрицы диаграмм рассеяния  🞏 Лица Чернова  🞏 нет ответа  2. Диаграмма Ящик-с-усами (box-and-whisker) для выборки выглядит как одна единственная точка:  🞏 Это невозможно  🞏 Это возможно, если выборка состоит из одного значения  🞏 Это возможно, если выборка состоит из нескольких различных значений  🞏 нет ответа  3. Цифровая трансформация это  🞏 Применение цифровых технологий с целью повышения эффективности деятельности  🞏 Преобразование модели бизнеса, продуктов, услуг предприятия с использованием цифровых технологи  🞏 Внедрение цифровых технологий во все процессы организации  4. Какие из типов диаграмм подойдут для описания категориальных переменных  🞏 Гистограммы  🞏 Столбиковые диаграммы  🞏 Кумулятивные частотные графики  🞏 Секторные (круговые) диаграммы  5. Какое из средних наиболее восприимчиво к экстремальным значениям в наборе данных  🞏 Мода  🞏 Медиана  🞏 Среднее арифметическое  6. В каком случае для вычисления среднего арифметического необходима средняя точка интервала  🞏 Когда исходные данные представлены отдельными значениями  🞏 Когда исходные данные представлены частотным распределением отдельных значений  🞏 Исходные данные представлены частотным распределение сгруппированных значений  7. Частотная диаграмма для некоторой выборки показывает два пика, в таком случае всегда верно, что …  🞏 Среднее совпадает с медианой  🞏 Среднее совпадает с модой  🞏 Среднее не совпадает с модой  🞏 Медиана совпадает с модой  8. На картинке приведен вид несимметричного распределения. Подпишите моду , медиану, межквартильный размах  9. Если мы знаем доверительный интервал параметра совокупности, то мы можем сказать  🞏 точное значение параметра находится в этом интервале  🞏 точное значение параметра с заранее известной вероятностью находится в этом интервале  🞏 можно найти точное значение параметра, если взять достаточно большую выборку  10. Какое испытание позволяет провести критерий хи-квадрат  🞏 О средних значениях  🞏 О доле  🞏 На независимость |
| 1.2 | 1. Искусственный интеллект –это  🞏 Программа, которая может заменить человека в решении многих классов задач  🞏 Программа, пригодная для решения конкретных задач  🞏 Не знаю  2. Что нужно для того, чтобы обучить программу  🞏 Данные, соответствующие задаче, для решения которой она написана  🞏 Любые данные, относящиеся к задаче  🞏 Чем больше различных данных загрузим, тем лучше  🞏 Не знаю  3. Как можно получить признаки  🞏 На основе данных  🞏 Предположить и сопоставить свою гипотезу с данными  🞏 Не знаю  4. Какое решение является оптимальным  🞏 Которое позволяет найти минимум или максимум значений  🞏 Которое является самым подходящим  🞏 Любое, которое устраивает | 1. Какие задачи можно решить с помощью искусственного интеллекта на сегодняшний день?  🞏 Повышение качества изображений  🞏 Игра в футбол  🞏 Сумма выдаваемого банковского кредита  🞏 Лента новостей в соцсетях.  2. Какие признаки относятся к качественным? Необходимо выбрать все правильные ответы.  🞏 Количественные  🞏 Порядковые  🞏 Интервальные  🞏 Номинальные  3. Какой признак является порядковыми?  🞏 Уровень образования  🞏 Вес  🞏 Национальность  🞏 Площадь квартиры  4. Предположим, что Вы проводите исследование. Один из исследуемых признаков “Уровень образования”. Какие меры разброса Вы можете вычислить для него?  🞏 Размах  🞏 Межквартильный размах  🞏 Стандартное отклонение  🞏 Дисперсия  5. Какие виды гипотез существуют? Необходимо выбрать все верные ответы.  🞏 Нулевая  🞏 Положительная  🞏 Альтернативная  🞏 Нетрадиционная  🞏 нет ответа  6. Какие из приведенных ниже задач можно решить с помощью классификации?  🞏 Определение цены квартиры в зависимости от района  🞏 Определение жанра книги  🞏 Определение пола человека по весу и росту  🞏 Определение возраста моллюска по внешним данным  7. Какое количество классов возможно в задаче бинарной классификации?  🞏 1  🞏 2  🞏 10  🞏 100  🞏 (свое значение)  8. Какие из приведенных ниже задач можно решить с помощью кластеризации?  🞏 Распределение текстов на темы  🞏 Решение о выдаче кредита банком  🞏 Сегментация целевой аудитории  🞏 Определение возраста моллюска по внешним данным  9. Методы оптимизации позволяют:  🞏 Найти минимальное или максимальное значение целевой функции при выполнении условий ограничений  🞏 Найти решение, которое больше всего устраивает потребителя с точки зрения его критериев  🞏 Найти модель, наиболее удобную для проведения расчетов  10. Если функция представлена таблично, какие методы оптимизации использовать НЕЛЬЗЯ  🞏 Аналитические  🞏 Численные  🞏 Эмпирические | 1. Чем похожи концепт машинного обучения и обучение в реальной жизни?  🞏 Искусственный интеллект обучается на готовых данных, так же происходит и в реальной жизни  🞏 Искусственный интеллект, как и человек, находит взаимосвязи в доступной ему информации, позволяющие предугадывать ответ верно  🞏 В обоих случаях увеличение доступных данных неизменно ведет к улучшению качества обучения  🞏 Решается одинаково широкий круг задач  2. Выберите верные утверждения  🞏 Одна из задач машинного обучения – научиться делать прогнозы для объектов  🞏 Объекты описываются с помощью признаков  🞏 Признаки описываются с помощью объектов  🞏 Одна из задач машинного обучения –научиться делать прогнозы для признаков  3. Что из этого корректные названия типов признаков  🞏 Номинальные (категориальные) признаки  🞏 Бинарные признаки  🞏 Нетривиальные признаки  🞏 Устойчивые признаки  🞏 Числовые (количественные) признаки  4. Какая из задач является детерминированной (то есть, количество ответов определено при постановке задачи)  🞏 Кластеризация  🞏 Классификация.  🞏 Сегментация  5. Выберите все верные утверждения:  🞏 С помощью корреляции можно сделать вывод о влиянии одного признака на другой  🞏 Корреляция вычисляется между тремя и более переменными  🞏 Высокий коэффициент корреляции всегда позволяет говорить о наличии взаимосвязи между исследуемыми признаками  🞏 Коэффициент корреляции вычисляется только для двух переменных  🞏 Корреляция может отражать только положительную взаимосвязь  🞏 Корреляция может отражать как положительную, так и отрицательную взаимосвязь  6. Какие из этих задач являются задачами классификации  🞏 Поиск групп похожих пользователей интернет-магазина  🞏 Прогноз температуры на следующий день  🞏 Прогноз оценки студента по пятибалльной шкале на экзамене по машинному обучению в следующую сессию  🞏 Разделение книг, хранящихся в электронной библиотеке, на научные и художественные  7. Какая из этих фраз точно описывает переобучение  🞏 Переобучение – это ситуация, в которой алгоритм часто отказывается от построения прогноза на новых данных  🞏 Переобучение – это ситуация, в которой алгоритм выдает недетерминированные ответы на новых данных (то есть при разных запусках на одном и том же объекте можно получить разные предсказания)  🞏 Переобучение – это ситуация, в которой алгоритм показывает хорошее качество на обучающей выборке, но при этом плохо работает на новых данных  8. Методы оптимизации позволяют:  🞏 Найти минимальное или максимальное значение целевой функции при выполнении условий ограничений  🞏 Найти решение, которое больше всего устраивает потребителя с точки зрения его критериев  🞏 Найти модель, наиболее удобную для проведения расчетов  9. Каким методом можно аналитически решить задачу линейного программирования  🞏 Жордано-Гаусса  🞏 Симплексным  🞏 Потенциалов  10. Задача сегментации изображения может это задача  🞏 классификации  🞏 кластеризации  🞏 ни то, ни другое |
|  |  |  |  |

**8.2. Входной, промежуточный и итоговый тесты считаются пройденными в случае не менее 60 % правильных ответов.**

**8.3. примеры контрольных заданий по модулям или всей образовательной программе**

Модуль 1:

1.Какие данные вы используете для выполнения своей работы?

2.Какие числовые инструменты вы знаете?

3.Назовите цифровые инструменты, которые дадут лучшие результаты

4.Какие характеристики описывают ваш набор данных лучше?

5.Что помогло бы представить ваши данные визуально?

6.В каких случаях пригодится построение гистограмм?

Поясните, почему

7.Какое ПО для обработки данных вам известно?

8. Какие программы вы используете?

9. Что такое выборка?

10. Какие меры среднего вы знаете?

11. Какие меры разброса вы знаете?

12. Что такое квартиль?

13. Третий квартиль -это (0.25 -квантиль, 0.5 - квантиль, 0.75 – квантиль)

14. В каких пределах находится коэффициент корреляции?

15. Какие виды гипотез существуют?

16. Что такое информационный дизайн?

17.Что такое визуальная аналитика?

18. Когда используется научная визуализация?

19. Что должна учитывать репрезентативная выборка?

20. В каком случае будет необходимо вычислить среднее арифметическое, моду и медиану, а когда достаточно среднего арифметического?

21. Что такое размах?

22. Какое значение коэффициент корреляции не может принимать (-3,56; -0,74; 1,2; 0,45)?

23. Коэффициент корреляции, равный 0.97, свидетельствует о...(...прямой сильной связи,...обратной слабой связи, ...прямой слабой связи,...обратной сильной связи)

24. Что такое нулевая гипотеза?

25. Какие техники визуализации вам известны?

26. Перечислите техники визуализации геометрического типа.

27. В каком случае необходимы матрицы диаграмм рассеяния?

Модуль 2:

1.Как вы понимаете, что такое искусственный интеллект?

2. Что необходимо программе, чтобы обучиться?

3. Какие задачи решает машинное обучение?

4.Какое решение вы считаете оптимальным?

5. Какие алгоритмы, инструменты вы применяете, чтобы найти оптимальное решение?

6. Встречались ли вам задачи компьютерного зрения?

7. Что включает в себя техническое зрение?

8.Какие задачи можно решить с помощью искусственного интеллекта на сегодняшний день? (Повышение качества изображений. Игра в футбол. Сумма выдаваемого банковского кредита. Лента новостей в соцсетях.).

9. Какая зависимость устанавливается в задаче регрессии?

10.Что такое классификация в машинном обучении?

11. Какое количество классов возможно в задаче бинарной классификации?

12.По каким принципам должен разбивать объекты на группы хороший алгоритм кластеризации?

13. Что такое обучение без учителя?

14. Для какого класса задач применима логистическая регрессия?

15. Что такое случайный лес?

16. Какое решение считаем оптимальным?

17. Запишите свою задачу в форме, пригодной для решения в EXCEL

18.Чем похожи концепт машинного обучения и обучение в реальной жизни?

19. Какие из перечисленных задач могут быть решены как регрессия? (Оценка стоимости автомобиля, Полезность отзыва на товар, Время задержки рейса самолета)

20. В чем отличие задач классификации и регрессии?

21. Пусть регрессионная формула стоимости автомобиля записывается как P = L^2+10\*L+5, где L — длина в метрах. На сколько отличается стоимость машин с длинами 3 и 5 метров?

22. Какие базовые алгоритмы можно использовать в бэггинге?

23. Какие базовые алгоритмы можно использовать в случайном лесе?

24. Запишите в общем виде целевую функцию и ограничения.

25. Что такое эвристический алгоритм?

26. Что включает в себя система технического зрения?

**8.4. тесты и обучающие задачи (кейсы), иные практикоориентированные формы заданий**

Модуль 1:

Входной тест:

1. Какие данные необходимы для выполнения вашей работы:

🞏 Результаты измерений приборов контроля

🞏 Данные финансового контроля

🞏 Значения параметров технологического процесса

🞏 Результаты мониторинга выполнения работ

🞏 Другое

2. Выберите цифровые инструменты, подходящие для вашей работы

🞏 Визуализация данных

🞏 Анализ статистических характеристик

🞏 Регрессионный анализ

🞏 Мне не знакомы эти инструменты

3. Какие диаграммы можно построить, если у вас есть данные по результатам сдачи экзамена?

🞏 Круговая

🞏 Гистограмма

🞏 Столбиковая диаграмма

🞏 Другое

4. По каким значениям строят гистограмму?

🞏 По любым

🞏 Непрерывным, числовым

5. Для обработки данных можно использовать

🞏 Только специальные программы

🞏 Excel

🞏 Paint

🞏 Orange

🞏 Knime

🞏 Не могу сказать

Промежуточный тест:

1. Цифровая трансформация это

🞏 Применение цифровых технологий с целью повышения эффективности деятельности

🞏 Преобразование модели бизнеса, продуктов, услуг предприятия с использованием цифровых технологи

🞏 Внедрение цифровых технологий во все процессы организации

2. Предназначением …(заполнить пропуск)……… является ясное, прозрачное и недвусмысленное представление информации для улучшения понимания и процесса коммуникации. Основной задачей данного направления является преобразование информации на различных уровнях сложности в более доступные для широкой аудитории формы.Выберите пропущенное направление

🞏 информационного дизайна

🞏 визуализации информации

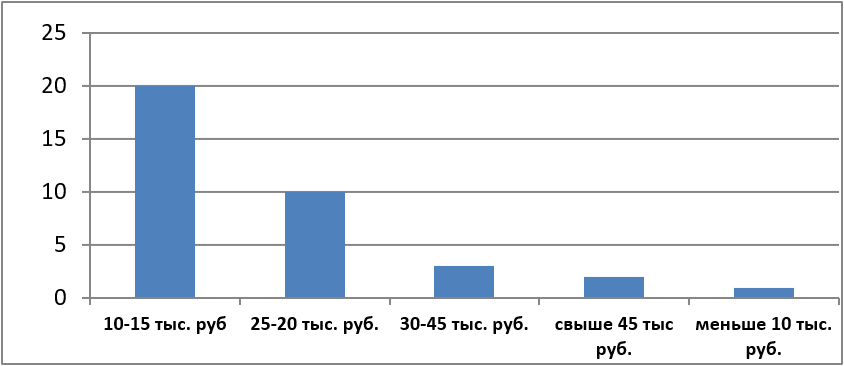
🞏 визуализации данных

🞏 научной визуализации

🞏 визуальной аналитики

🞏 нет ответа

3. Дана частотная гистограмма распределения доходов населения в некотором населённом пункте, полученная по результатам маркетингового исследования. Выберите статистическую характеристику для оценки положения доходов (возможен только один ответ):



🞏 Арифметическое среднее

🞏 Мода

🞏 Медиана

🞏 Геометрическое среднее

4. Найдите моду и медиану для данных, представленных частотным распределением

|  |  |  |
| --- | --- | --- |
| х | f | Mo |
| 1 | 7 | Me |
| 2 | 9 |  |
| 3 | 18 |  |
| 4 | 12 |  |
| 5 | 6 |  |

5. Статистическое испытание гипотез проводится

🞏 По генеральной совокупности

🞏 По выборочным данным

🞏 Не имеет значения

6. Какая мера центральной тенденции чувствительна к выбросам?

🞏 Среднее

🞏 Мода

🞏 Медиана

7. Что из перечисленного НЕ является мерой разброса?

🞏 Дисперсия

🞏 Медиана

🞏 Стандартное отклонение

🞏 Размах

8. Даны две выборки значений, измеренных в номинальной шкале. Для установления связи между выборками можно использовать:

🞏 Коэффициент Хи-квадрат

🞏 Коэффициент корреляции Пирсона

🞏 Тест Манна-Уитни

🞏 Коэффициент корреляции Спирмена

🞏 нет ответа

9. Какое значение коэффициент корреляции НЕ может принимать?

🞏 -3.56

🞏 -0.78

🞏 1.02

🞏 0.45

🞏 нет ответа

10. Если исходные данные содержат квадратичную зависимость для пары переменных, можно ли использовать уравнение линейной регрессии для аппроксимации этих данных?

🞏 Да

🞏 Нет

**Итоговый тест (1 часть)**

1. Укажите техники визуализации данных, относящиеся к геометрическому типу. Выберите один или несколько ответов:

🞏 Каркасные диаграммы

🞏 Кодирование геометрическими формами

🞏 Параллельные координаты

🞏 Матрицы диаграмм рассеяния

🞏 Лица Чернова

🞏 нет ответа

2. Диаграмма Ящик-с-усами (box-and-whisker) для выборки выглядит как одна единственная точка:

🞏 Это невозможно

🞏 Это возможно, если выборка состоит из одного значения

🞏 Это возможно, если выборка состоит из нескольких различных значений

🞏 нет ответа

3. Цифровая трансформация это

🞏 Применение цифровых технологий с целью повышения эффективности деятельности

🞏 Преобразование модели бизнеса, продуктов, услуг предприятия с использованием цифровых технологи

🞏 Внедрение цифровых технологий во все процессы организации

4. Какие из типов диаграмм подойдут для описания категориальных переменных

🞏 Гистограммы

🞏 Столбиковые диаграммы

🞏 Кумулятивные частотные графики

🞏 Секторные (круговые) диаграммы

5. Какое из средних наиболее восприимчиво к экстремальным значениям в наборе данных

🞏 Мода

🞏 Медиана

🞏 Среднее арифметическое

6. В каком случае для вычисления среднего арифметического необходима средняя точка интервала

🞏 Когда исходные данные представлены отдельными значениями

🞏 Когда исходные данные представлены частотным распределением отдельных значений

🞏 Исходные данные представлены частотным распределение сгруппированных значений

7. Частотная диаграмма для некоторой выборки показывает два пика, в таком случае всегда верно, что …

🞏 Среднее совпадает с медианой

🞏 Среднее совпадает с модой

🞏 Среднее не совпадает с модой

🞏 Медиана совпадает с модой

8. На картинке приведен вид несимметричного распределения. Подпишите моду , медиану, межквартильный размах

9. Если мы знаем доверительный интервал параметра совокупности, то мы можем сказать

🞏 точное значение параметра находится в этом интервале

🞏 точное значение параметра с заранее известной вероятностью находится в этом интервале

🞏 можно найти точное значение параметра, если взять достаточно большую выборку

10. Какое испытание позволяет провести критерий хи-квадрат

🞏 О средних значениях

🞏 О доле

🞏 На независимость

Модуль 2:

Входной тест

1. Искусственный интеллект – это

🞏 Программа, которая может заменить человека в решении многих классов задач

🞏 Программа, пригодная для решения конкретных задач

🞏 Не знаю

2. Что нужно для того, чтобы обучить программу

🞏 Данные, соответствующие задаче, для решения которой она написана

🞏 Любые данные, относящиеся к задаче

🞏 Чем больше различных данных загрузим, тем лучше

🞏 Не знаю

3. Как можно получить признаки

🞏 На основе данных

🞏 Предположить и сопоставить свою гипотезу с данными

🞏 Не знаю

4. Какое решение является оптимальным

🞏 Которое позволяет найти минимум или максимум значений

🞏 Которое является самым подходящим

🞏 Любое, которое устраивает

Промежуточный тест:

1. Какие задачи можно решить с помощью искусственного интеллекта на сегодняшний день?

🞏 Повышение качества изображений

🞏 Игра в футбол

🞏 Сумма выдаваемого банковского кредита

🞏 Лента новостей в соцсетях.

2. Какие признаки относятся к качественным? Необходимо выбрать все правильные ответы.

🞏 Количественные

🞏 Порядковые

🞏 Интервальные

🞏 Номинальные

3. Какой признак является порядковыми?

🞏 Уровень образования

🞏 Вес

🞏 Национальность

🞏 Площадь квартиры

4. Предположим, что Вы проводите исследование. Один из исследуемых признаков “Уровень образования”. Какие меры разброса Вы можете вычислить для него?

🞏 Размах

🞏 Межквартильный размах

🞏 Стандартное отклонение

🞏 Дисперсия

5. Какие виды гипотез существуют? Необходимо выбрать все верные ответы.

🞏 Нулевая

🞏 Положительная

🞏 Альтернативная

🞏 Нетрадиционная

🞏 нет ответа

6. Какие из приведенных ниже задач можно решить с помощью классификации?

🞏 Определение цены квартиры в зависимости от района

🞏 Определение жанра книги

🞏 Определение пола человека по весу и росту

🞏 Определение возраста моллюска по внешним данным

7. Какое количество классов возможно в задаче бинарной классификации?

🞏 1

🞏 2

🞏 10

🞏 100

🞏 (свое значение)

8. Какие из приведенных ниже задач можно решить с помощью кластеризации?

🞏 Распределение текстов на темы

🞏 Решение о выдаче кредита банком

🞏 Сегментация целевой аудитории

🞏 Определение возраста моллюска по внешним данным

9. Методы оптимизации позволяют:

🞏 Найти минимальное или максимальное значение целевой функции при выполнении условий ограничений

🞏 Найти решение, которое больше всего устраивает потребителя с точки зрения его критериев

🞏 Найти модель, наиболее удобную для проведения расчетов

10. Если функция представлена таблично, какие методы оптимизации использовать НЕЛЬЗЯ

🞏 Аналитические

🞏 Численные

🞏 Эмпирические

**Итоговый тест (2 часть):**

1. Чем похожи концепт машинного обучения и обучение в реальной жизни?

🞏 Искусственный интеллект обучается на готовых данных, так же происходит и в реальной жизни

🞏 Искусственный интеллект, как и человек, находит взаимосвязи в доступной ему информации, позволяющие предугадывать ответ верно

🞏 В обоих случаях увеличение доступных данных неизменно ведет к улучшению качества обучения

🞏 Решается одинаково широкий круг задач

2. Выберите верные утверждения

🞏 Одна из задач машинного обучения – научиться делать прогнозы для объектов

🞏 Объекты описываются с помощью признаков

🞏 Признаки описываются с помощью объектов

🞏 Одна из задач машинного обучения – научиться делать прогнозы для признаков

3. Что из этого корректные названия типов признаков

🞏 Номинальные (категориальные) признаки

🞏 Бинарные признаки

🞏 Нетривиальные признаки

🞏 Устойчивые признаки

🞏 Числовые (количественные) признаки

4. Какая из задач является детерминированной (то есть, количество ответов определено при постановке задачи)

🞏 Кластеризация

🞏 Классификация.

🞏 Сегментация

5. Выберите все верные утверждения:

🞏 С помощью корреляции можно сделать вывод о влиянии одного признака на другой

🞏 Корреляция вычисляется между тремя и более переменными

🞏 Высокий коэффициент корреляции всегда позволяет говорить о наличии взаимосвязи между исследуемыми признаками

🞏 Коэффициент корреляции вычисляется только для двух переменных

🞏 Корреляция может отражать только положительную взаимосвязь

🞏 Корреляция может отражать как положительную, так и отрицательную взаимосвязь

6. Какие из этих задач являются задачами классификации

🞏 Поиск групп похожих пользователей интернет-магазина

🞏 Прогноз температуры на следующий день

🞏 Прогноз оценки студента по пятибалльной шкале на экзамене по машинному обучению в следующую сессию

🞏 Разделение книг, хранящихся в электронной библиотеке, на научные и художественные

7. Какая из этих фраз точно описывает переобучение

🞏 Переобучение – это ситуация, в которой алгоритм часто отказывается от построения прогноза на новых данных

🞏 Переобучение – это ситуация, в которой алгоритм выдает недетерминированные ответы на новых данных (то есть при разных запусках на одном и том же объекте можно получить разные предсказания)

🞏 Переобучение – это ситуация, в которой алгоритм показывает хорошее качество на обучающей выборке, но при этом плохо работает на новых данных

8. Методы оптимизации позволяют:

🞏 Найти минимальное или максимальное значение целевой функции при выполнении условий ограничений

🞏 Найти решение, которое больше всего устраивает потребителя с точки зрения его критериев

🞏 Найти модель, наиболее удобную для проведения расчетов

9. Каким методом можно аналитически решить задачу линейного программирования

🞏 Жордано-Гаусса

🞏 Симплексным

🞏 Потенциалов

10. Задача сегментации изображения может это задача

🞏 классификации

🞏 кластеризации

🞏 ни то, ни другое

Кейс 1. Построение схемы информационных потоков. Постройте схему информационных потоков для своего подразделения (для решения одной из задач). Какие данные вы получаете, собираете, храните, передаете. Где это происходит? Какие из описанных в примерах решений могли бы быть полезны для вас? Что вы могли бы сделать прямо сейчас?

Кейс 2. Описание данных. Откройте файл с данными и описание ситуации. Опишите переменные, определите распределение, используя Excel, вычислите описательные характеристики, постройте диаграммы и графики, которые уместны для вашего набора данных. Можно ли говорить о зависимости отдельных переменных? Свой ответ подтвердите графически, и, если возможно, расчётами.

Кейс 3. Постановка задачи машинного обучения. Используя имеющиеся у вас данные и описание ситуации, определите, какую задачу можно решить методами машинного обучения. Проанализируйте, достаточно ли у вас данных. Какие из алгоритмов машинного обучения могут быть вам полезны. Попробуйте применить их, используя KNIME. Оцените достоверность полученного результата.

Кейс 4. Постановка и решение задачи оптимизации. Используя материал своей задачи, сформулируйте целевую функцию. В «пакете анализа» Excel выберите подходящий метод. Оцените полученный результат с точки зрения точности и достоверности.

Кейс 5. Прототип системы компьютерного зрения. Разработка прототипа системы компьютерного зрения с помощью программного обеспечения KNIME.

**8.5. Результаты обучения оцениваются в форме зачет/незачет по результатам итогового тестирования по модулям 1 и 2.**

При неудовлетворительном результате тестирования (менее 60% правильных ответов) слушатель может пройти итоговую аттестацию повторно в соответствии с принятым в ЧГУ Положением об организации и осуществлении образовательной деятельности по дополнительным профессиональным программам.

**9.Организационно-педагогические условия реализации программы**

**9.1. Кадровое обеспечение программы**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **№**  **п/п** | **Фамилия, имя, отчество (при наличии)** | **Место основной работы и должность, ученая степень и ученое звание (при наличии)** | **Ссылки на веб-страницы с портфолио (при наличии)** | **Фото в формате jpeg** | **Отметка о полученном согласии на обработку персональных данных** |
| 1 | Юдина Ольга Вадимовна | доцент кафедры математического и программного обеспечения ЭВМ, кандидат технических наук |  | **D:\Users\iino7\Desktop\ЮдинаО.В.jpg** | Да |
| 2 | Ершов Евгений Валентинович | директор Института информационных технологий, заведующий кафедрой математического и программного обеспечения ЭВМ, доктор технических наук, профессор |  | **D:\Users\iino7\Desktop\Ершов Е.В..jpg** | Да |
| 3 | Виноградова Людмила Николаевна | доцент кафедры математического и программного обеспечения ЭВМ, кандидат технических наук |  | **D:\Users\iino7\Desktop\Виноградова Л.Н..jpg** | Да |

**9.2.Учебно-методическое обеспечение и информационное сопровождение**

|  |  |
| --- | --- |
| **Учебно-методические материалы** | |
| Методы, формы и технологии | Методические разработки,  материалы курса, учебная литература |
| Реализация образовательной программы с применением электронного обучения и (или) дистанционных образовательных технологий с возможностью передачи данных в форме элементов цифрового следа | Методические разработки, материалы курса размещены на образовательном портале ЧГУ (edu.chsu.ru)  Основная литература:  Вальке, А.А. Электронные средства сбора и обработки информации: учебное пособие / А.А. Вальке, В.А. Захаренко ; Минобрнауки России, Омский государственный технический университет. - Омск: Издательство ОмГТУ, 2017. - 112 с. : табл., схем., ил. - Библиогр. в кн. - ISBN 978-5-8149-2519-0; То же [Электронный ресурс]. - URL: <http://biblioclub.ru/index.php?page=book&id=493448> (22.04.2019).  Блюмин, А.М. Мировые информационные ресурсы : учебное пособие / А.М. Блюмин, Н.А. Феоктистов. - 3-е изд., перераб. и доп. - Москва : Издательско-торговая корпорация «Дашков и К°», 2016. - 384 с. : ил. - (Учебные издания для бакалавров). - Библиогр. в кн. - ISBN 978-5-394-02411-5 ; То же [Электронный ресурс]. - URL: <http://biblioclub.ru/index.php?page=book&id=453024>  Кравченко, Ю.А. Тенденции развития компьютерных технологий: учебное пособие / Ю.А. Кравченко, Э.В. Кулиев, Д.В. Заруба; Министерство образования и науки РФ, Южный федеральный университет, Инженерно-технологическая академия. - Таганрог: Издательство Южного федерального университета, 2017. - 107 с. : схем., ил. - Библиогр. в кн. - ISBN 978-5-9275-2360-3; То же [Электронный ресурс]. - URL: <http://biblioclub.ru/index.php?page=book&id=493214> (19.06.2019).…  Павлов, С.И. Системы искусственного интеллекта: учебное пособие / С.И. Павлов. - Томск: Томский государственный университет систем управления и радиоэлектроники, 2011. - Ч. 1. - 175 с. - ISBN 978-5-4332-0013-5; То же [Электронный ресурс]. - URL: <http://biblioclub.ru/index.php?page=book&id=208933> (19.06.2019).  Дополнительная литература:  Интеллектуальный анализ данных в управлении производственными системами (подходы и методы)=Intelligentdataanalysisinthemanagementofproductionsystems (approachesandmethods) :монография / Л.А. Мыльников, Б. Краузе, М. Кютц и др. - Москва : БИБЛИО-ГЛОБУС, 2017. - 334 с. : табл., граф., схем., ил. - Библиогр.: с. 311-328. - ISBN 978-5-9500501-7-6; То же [Электронный ресурс]. - URL: <http://biblioclub.ru/index.php?page=book&id=499006> (19.06.2019). |
|  |  |

|  |  |
| --- | --- |
| **Информационное сопровождение** | |
| Электронные  образовательные ресурсы | Электронные  информационные ресурсы |
| образовательный портал ЧГУ (edu.chsu.ru) | Электронный каталог библиотеки университета http://marc.chsu.ru/Default.asp  Научная электронная библиотека «eLibrary» http://elibrary.ru/defaultx.asp  Информационная система «Единое окно» http://window.edu.ru/  Электронная библиотека «Лань» (бесплатный пакет) http://e.lanbook.com/  Новая электронная библиотека в ЧГУ http://biblioclub.ru/  <https://www.hpe.com/ru/ru/what-is/digital-transformation.html> сайт кампании «Хьюлетт Паккард»  <https://www.sas.com/ru_ru/insights/articles/analytics/how-to-find-and-equip-citizen-data-scientists.html>.  <https://www.mckinsey.com/business-functions/mckinsey-analytics/our-insights/analytics-translator> . |
|  |  |

**9.3.Материально-технические условия реализации программы**

|  |  |
| --- | --- |
| Вид занятий | Наименование оборудования,  программного обеспечения |
| Лекция | MicrosoftTeams  Комплекс «Активный экран» ACTIVboard 78, Проектор 3DAcerH5380, MicrosoftWindowsPro 7 RUS |
| Практическая работа | Microsoft Teams  Персональныйкомпьютер (Intel Core i3-9100F, NVidia GeForce CTX 1660, DDR2 8Gb, HDD 1Tb)  1. Microsoft Windows Pro 7 RUS; 2. KNIME, 3. R |
| Лабораторная работа | Microsoft Teams  Персональныйкомпьютер (Intel Core i3-9100F, NVidia GeForce CTX 1660, DDR2 8Gb, HDD 1Tb)  1. Microsoft Windows Pro 7 RUS; 2. KNIME, 3. R |

**Составители программы:**

Юдина Ольга Вадимовна, доцент кафедры математического и программного обеспечения ЭВМ, кандидат технических наук

Ершов Евгений Валентинович, директор Института информационных технологий, заведующий кафедрой математического и программного обеспечения ЭВМ, доктор технических наук, профессор

Виноградова Людмила Николаевна, доцент кафедры математического и программного обеспечения ЭВМ, кандидат технических наук

**Приложение. Паспорт компетенций**

наименование дополнительной профессиональной образовательной программы повышения квалификации: **Инструменты анализа данных и машинного обучения**

наименование организации, реализующей дополнительную профессиональную образовательную программу повышения квалификации: **федеральное государственное бюджетное образовательное учреждение высшего образования "Череповецкий государственный университет"**

Компетенция 1

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 1. | Наименование компетенции | | Анализ больших данных с использованием существующей в организации методологической и технологической инфраструктуры | |
| 2. | Указание типа компетенции | общекультурная/  универсальная |  | |
| общепрофессиональная |  | |
| профессиональная | + | |
| профессионально-специализированная |  | |
| 3. | Определение, содержание и основные сущностные характеристики компетенции | | Цифровые технологии обработки данных, их применения: аналитика больших данных, прогнозирование, кластеризация, машинное обучение. Цифровая трансформация производства.  Шкалы измерений и классификация данных. Качественные и количественные переменные. Виды статистических графиков. Трудности при сборе данных, важность корректных данных. Описание данных. Меры среднего и их применимость, меры вариации. Зависимость и независимость переменных.  Машинное обучение. Примеры задач машинного обучения. Этапы решения задачи методом машинного обучения. | |
| 4. | Дескриптор знаний, умений и навыков по уровням | | Уровни сформированности компетенции  обучающегося | Индикаторы |
| Выявление, формирование и согласование требований к результатам аналитических работ с применением технологий больших данных | | Начальный уровень  (Компетенция недостаточно развита. Частично проявляет навыки, входящие в состав компетенции. Пытается, стремится проявлять нужные навыки, понимает их необходимость, но у него не всегда получается.) | И1: Владеть навыками выявления, формирования и согласования требований заказчика к результатам анализа, определения возможностей применения анализа данных больших данных в предметной области и конкретных задачах заказчика  И2: Уметь проводить анализ больших данных в соответствии с утвержденными требованиями к результатам аналитического исследования  И3: Знать современные методы и инструментальные средства анализа больших данных |
| Планирование и организация аналитических работ с использованием технологий больших данных | | Базовый уровень  (Уверенно владеет навыками, способен, проявлять соответствующие навыки в ситуациях с элементами неопределён-ности, сложности.) | И1: Владеть навыками разработки, обсуждения и утверждения содержания аналитических работ с использованием технологий больших данных  И2: Уметь проводить аналитические работы с использованием технологий больших данных  И3: Знать методы интерпретации и визуализации анализа больших данных |
| Подготовка данных для проведения аналитических работ по исследованию больших данных | | Продвинутый  (Владеет сложными навыками, способен активно влиять на происходящее, проявлять соответствующие навыки в ситуациях повышенной сложности.) | И1: Владеть навыками получения и фильтрации больших объемов данных из гетерогенных источников  И2: Уметь использовать инструментальные средства для извлечения, преобразования, хранения и обработки данных из разнородных источников, в том числе в режиме реального времени  И3: Знать методы извлечения информации и знаний из гетерогенных, мультиструктурированных, неструктурированных источников, в том числе при потоковой обработке |
| Проведение аналитического исследования с применением технологий больших данных в соответствии с требованиями заказчика | | Профессиональный  (Владеет сложными навыками, создает новые решения для сложных проблем со многими взаимодействую-щими факторами, предлагает новые идеи и процессы, способен активно влиять на происходящее, проявлять соответствующие навыки  в ситуациях повышенной сложности.) | И1: Владеть навыками разработки, поверки, оценки используемых моделей больших данных  И2: Уметь решать задачи классификации, кластеризации, регрессии, прогнозирования, снижения размерности и ранжирования данных  И3: Знать основы машинного обучения (классификация, кластеризация, обнаружение выбросов, фильтрация), алгоритмы машинного обучения (обучение с учителем, обучение без учителя, полууправляемое обучение, обучение с подкреплением) |
| 5. | Характеристика взаимосвязи данной компетенции с другими компетенциями/ необходимость владения другими компетенциями для формирования данной компетенции | | Владение технологиями разработки программного обеспечения и технологиями хранения больших данных | |
| 6. | Средства и технологии оценки | | Анализ выполнения практических и лабораторных (кейсовых) заданий, тестирование | |

Компетенция 2

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 1. | Наименование компетенции | | Управление разработкой продуктов, услуг и решений на основе больших данных | |
| 2. | Указание типа компетенции | общекультурная/  универсальная |  | |
| общепрофессиональная |  | |
| профессиональная | + | |
| профессионально-специализированная |  | |
| 3. | Определение, содержание и основные сущностные характеристики компетенции | | Инструменты графического анализа: график плотности распределения, диаграмма рассеяния, матрица корреляций, box-plot, временные ряды и тренды, график «тепловая карта». Оценка достоверности полученных результатов.  Анализ и обработка данных и машинное обучение. Задачи регрессии в машинном обучении. Линейная регрессия, гребневая регрессия. Задачи классификации в машинном обучении. Расширенный анализ данных.  Оптимизация в EXCEl. Примеры использования. Формулировка целевой функции, ограничений. Выбор метода. Постановка и методы решения задачи линейного программирования. Симплекс-метод. Эвристические алгоритмы.  Примеры задач компьютерного зрения. Основные этапы решения задачи технического зрения. Искусственные нейронные сети. | |
| 4. | Дескриптор знаний, умений и навыков по уровням | | Уровни сформированности компетенции  обучающегося | Индикаторы |
| Разработка продуктов на основе встроенной аналитики больших данных | | Начальный уровень  (Компетенция недостаточно развита. Частично проявляет навыки, входящие в состав компетенции. Пытается, стремится проявлять нужные навыки, понимает их необходимость, но у него не всегда получается.) | И1: Владеть навыками разработки бизнес-плана создания нового продукта на основе встроенной аналитики больших данных  И2: Уметь осуществлять математическое и информационное моделирование  И3: Знать технические средства и среды сбора, хранения и обработки больших данных, машинное обучение |
| Разработка сервисов на основе аналитики больших данных | | Базовый уровень  (Уверенно владеет навыками, способен, проявлять соответствующие навыки в ситуациях с элементами неопределён-ности, сложности.) | И1: Владеть навыками определения источников больших данных и условий их использования или приобретения  И2: Уметь разрабатывать программно-аппаратные компоненты и системы на основе технологий больших данных  И3: Знать существующий опыт разработки и использования продуктов и услуг на основе технологий больших данных |
| Разработка инфраструктурных решений на основе аналитики больших данных | | Продвинутый  (Владеет сложными навыками, способен активно влиять на происходящее, проявлять соответствующие навыки в ситуациях повышенной сложности.) | И1: Владеть навыками организации разработки, согласования и утверждения рабочего проекта инфраструктурного решения на основе технологий больших данных  И2: Уметь проводить сравнительный анализ методов и программного обеспечения функций и поддержки бизнес-процессов, процессов принятия решений и аналитических задач на основе технологий больших данных  И3: Знать методологию системного анализа и ее практическое применение для анализа деятельности организации |
| Разработка сложных инфраструктурных решений на основе аналитики больших данных | | Профессиональный  (Владеет сложными навыками, создает новые решения для сложных проблем со многими взаимодействую-щими факторами, предлагает новые идеи и процессы, способен активно влиять на происходящее, проявлять соответствующие навыки  в ситуациях повышенной сложности.) | И1: Владеть навыками разработки плана выполнения рабочего проекта сложного инфраструктурного решения на основе технологий больших данных  И2: Уметь проводить презентации по итогам проектных работ  И3: Знать показатели эффективности технологий больших данных |
| 5. | Характеристика взаимосвязи данной компетенции с другими компетенциями/ необходимость владения другими компетенциями для формирования данной компетенции | | Системный анализ, управление поддержкой принятия решений, управление программными проектами | |
| 6. | Средства и технологии оценки | | Анализ выполнения практических и лабораторных (кейсовых) заданий, тестирование | |

**VI. Иная информация о качестве и востребованности образовательной программы** (результаты профессионально-общественной аккредитации образовательной программы, включение в системы рейтингования, призовые места по результатам проведения конкурсов образовательных программ и др.) (при наличии)

Нет

**V. Рекомендаций к программе от работодателей**:

- Рекомендательное письмо МАУ «Центр муниципальных информационных ресурсов и технологий»

-Рекомендательное письмо Общество с ограниченной ответственностью «Малленом Системс»

**VI.Указание на возможные сценарии профессиональной траектории граждан** по итогам освоения образовательной программы

Сценарий 1: Гражданин, состоящий на учете в Центре занятости или безработный после освоения, программы получит предложения по трудоустройству

Сценарий 2: Гражданин, работающий по найму в организации, на предприятии после освоения программы разовьет профессиональные качества (развитие компетенций в текущей сфере деятельности)

**VII.Дополнительная информация**

Нет

**VIII.Приложенные Скан-копии**

Утвержденная рабочая программа (подпись, печать, в формате pdf)