

Дополнительная общеобразовательная программа по дисциплине «Технология» на основе стандарта Ворлдскиллс по компетенции «Машинное обучение и большие данные»

1. Цели реализации программы

Дополнительная общеобразовательная программа по дисциплине «Технология» разработана для школьников 9-11 классов с целью ознакомления с современным уровнем IT-оборудования и программного обеспечения, возможностями создания программных продуктов прикладного назначения, с учетом стандарта Ворлдскиллс по компетенции «Машинное обучение и большие данные». Направлена на сознательный выбор одной из профессий информационно-коммуникационного профиля. Может использоваться для обучения взрослых.

2. Требования к результатам обучения. Планируемые результаты обучения

2.1. Характеристика нового вида профессиональной деятельности, трудовых функций и (или) уровней квалификации

№ п/п	Содержание совершенствуемой или вновь формируемой компетенции
1	Сбор, обработка и подготовка репозитория хранения данных
2	Предобработка и анализ больших данных
3	Подготовка обучающих и тестовых выборок
4	Разработка и построение математического моделей машинного обучения
5.	Разработка прикладных решений
6.	Разработка отчетной аналитической документации, разработка документации пользователей

Программа разработана в соответствии с:

- спецификацией стандартов Ворлдскиллс по компетенции «Машинное обучение и большие данные»;
- профессиональным стандартом «Программист» (утвержден приказом Минтруда России от 18.11.2013 № 679н).

Медицинские ограничения регламентированы Перечнем медицинских противопоказаний Минздрава России.

2.2. Требования к результатам освоения программы

В результате освоения дополнительной профессиональной программы у слушателя должны быть сформированы компетенции, в соответствии с разделом 2.1. программы.

В результате освоения программы слушатель должен

знать:

- стандарты Ворлдскиллс и спецификацию стандартов Ворлдскиллс по компетенции;
- требования охраны труда и техники безопасности;
- тенденции и разработки в отрасли, включая новые технологии, методы, языки, условные обозначения и технические навыки;

- статистические методы обработки данных;
- регрессионный анализ;
- методы и алгоритмы машинного обучения;
- дискриминантный анализ;
- кластерный анализ;
- нейронные сети (топология);
- основы программирования на Python и R;
- основы работы с базами данных;
- основы ImageProcessing;
- Big Data Visualization, Large Data Visualization, научную и информационную визуализацию;

- современные аппаратные средства и архитектуры для анализа и визуализации данных;
- важность тщательного документирования разработанных решений;

уметь:

- соблюдать требования охраны труда и техники безопасности;
- планировать производственный график на каждый день в соответствии с доступным временем и принимать во внимание временные ограничения и крайние сроки;
- применять исследовательские технологии и навыки, чтобы иметь представление о самых последних отраслевых рекомендациях;
- использовать коммуникационные навыки при работе в команде для сотрудничества с другими специалистами для получения желаемых результатов, успешной работы над групповым решением проблем;
- использовать навыки управления проектами в расстановке приоритетов и графика выполнения задач, распределении ресурсов между задачами;
- использовать аналитические навыки для анализа и синтеза сложной или неоднородной информации, определять тривиальные и нетривиальные зависимости данных;
- использовать современные программные продукты для построения математических моделей;
- использовать программное обеспечение для проектирования и моделирования;
- работать в операционной системе Linux;
- обрабатывать информацию и данные;
- выявлять аномалии в данных;
- масштабировать модели;
- строить деревья решений;
- производить компонентный анализ;
- применять математические методы в решении практических задач;
- разрабатывать алгоритмы машинного обучения и анализа данных;
- применять методы анализа данных и машинного обучения;
- выполнять распознавание изображений;
- реализовывать нейронные сети;
- реализовывать ботов;
- разрабатывать алгоритмы и математические модели;
- осуществлять оценку разработанных решений;
- представлять сложные структуры, а также объекты особого интереса, особые точки, аттракторы, сингулярности;
- выполнять верификацию и валидацию визуализации;
- интерпретировать результаты решения;
- интерпретировать большие данные и изображения;
- проявлять профессионализм в подготовке документации;
- разрабатывать документацию пользователей;
- работать с технической документацией на английском языке.

3. Категория лиц и содержание программы

Категория слушателей: дети 14-17 лет и взрослые без требования к уровню образования, имеющие навыки работы с ПК.

Трудоемкость обучения: 144 академических часа.

Форма обучения: очная с элементами ЭО и ДОТ.

3.1. Учебный план

№	Наименование модулей	Всего, ак. час.	В том числе			Форма контроля
			лекции	практ. занятия	промежут. и итог. контроль	
1	2	3	4	5	6	7
1.	Модуль 1. Стандарты Ворлдскиллс и спецификация стандартов Ворлдскиллс по компетенции «Машинное обучение и большие данные». Разделы спецификации	2	2			
2.	Модуль 2. Требования охраны труда и техники безопасности	2	2			
3.	Модуль 3. Современные профессиональные технологии	9	9			
4.	Модуль 4. Загрузка, подготовка, разведочный анализ и предобработка данных	40	4	36		
5.	Модуль 5. Решение задач машинного обучения	40	4	36		
6.	Модуль 6. Документирование аналитических решений, управление работой и решение проблем	40	4	36		
7.	Итоговая аттестация (тестирование, контрольная работа)	11			11	Тест КР
	ИТОГО:	144	25	108	11	

3.2. Учебно-тематический план

№	Наименование модулей	Всего, ак. час.	В том числе			Форма контроля
			лекции	практ. занятия	промежут. и итог. контроль	
1	2	3	4	5	6	7
1.	Модуль 1. Стандарты Ворлдскиллс и спецификация стандартов Ворлдскиллс по компетенции «Машинное обучение и большие данные». Разделы спецификации	2	2			
1.1	Актуальное техническое описание по компетенции. Спецификация стандарта Ворлдскиллс по компетенции «Машинное обучение и большие данные»	2	2			
2.	Модуль 2. Требования охраны труда и техники безопасности	2	2			
2.1	Требования охраны труда и техники безопасности	1	1			
2.2	Специфичные требования охраны труда, техники безопасности и окружающей среды по компетенции «Машинное обучение и большие данные»	1	1			
3.	Модуль 3. Современные технологии в профессиональной сфере	9	9			

3.1	Современные технологии в профессиональной сфере. Облачные сервисы для построения и обучения систем искусственного интеллекта. Сервисы и технологии для анализа больших данных.	9	9			
4.	Модуль 4. Загрузка, подготовка, разведочный анализ и предобработка данных	40	4	36		
4.1	Типы данных. Чтение данных из файлов и таблиц	11	1	10		
4.2	Подготовка данных	11	1	10		
4.3	Статистический анализ данных	9	1	8		
4.4	Предобработка данных	9	1	8		
5.	Модуль 5. Решение задач машинного обучения	40	4	36		
5.1	Решение задачи классификации	11	1	10		
5.2	Решение задачи регрессии	11	1	10		
5.3	Кластерный анализ данных	9	1	8		
5.4	Искусственные нейронные сети для решения задач распознавания образов	9	1	8		
6.	Модуль 6. Разработка прикладного решения и создание интерфейса пользователя. Документирование аналитических решений	40	4	36		

6.1	Создание графического интерфейса пользователя	9	1	8		
6.2	Внедрение модели машинного обучения в графический интерфейс	8	1	7		
6.3	Документирование выбора математического аппарата и алгоритма обучения	8	1	7		
6.4	Документирование результатов обучения и апробации	8	1	7		
6.5	Презентация результатов работы	7		7		
7.	Итоговая аттестация	11			11	
7.1	Тестирование. Контрольная работа	11			11	
	ИТОГО:	144	25	108	11	

3.3. Учебная программа

Модуль 1. Стандарты Ворлдскиллс и спецификация стандартов Ворлдскиллс по компетенции «Машинное обучение и большие данные». Разделы спецификации.

Тема 1.1 Актуальное техническое описание по компетенции. Спецификация стандарта Ворлдскиллс по компетенции «Машинное обучение и большие данные».

Лекция.

Актуальное техническое описание по компетенции.

Спецификация стандарта Ворлдскиллс по компетенции «Машинное обучение и большие данные».

Модуль 2. Требования охраны труда и техники безопасности.

Тема 2.1. Требования охраны труда и техники безопасности.

Лекция.

Общие требования охраны труда и техники безопасности. Трудовая деятельность человека. Основные принципы обеспечения безопасности и ОТ. Основные положения трудового права. Правовые основы охраны труда. Государственное регулирование в сфере охраны труда. Основы управления охраной труда в организации.

Тема 2.2. Специфичные требования охраны труда, техники безопасности и окружающей среды по компетенции.

Лекция.

Специфичные требования охраны труда, техники безопасности и окружающей среды по компетенции «Машинное обучение и большие данные».

Модуль 3. Современные технологии в профессиональной сфере.

Тема 3.1. Современные технологии в профессиональной сфере Облачные сервисы для построения и обучения систем искусственного интеллекта. Сервисы и технологии для анализа больших данных.

Лекция.

Облачные сервисы для построения и обучения систем искусственного интеллекта. Использование сервисов Microsoft Azure, Orange Data Mining, Amazon (Amazon SageMaker, Amazon Rekognition, Amazon Polly, Amazon Comprehend, Amazon Lex, Amazon Transcribe). Использование сервисов google cloud и google collaboratory для реализации систем искусственного интеллекта. Сервисы и технологии для анализа больших данных: Apache Hadoop, MapReduce, Hbase.

Модуль 4. Загрузка, подготовка, разведочный анализ и преобразование данных.

Тема 4.1. Типы данных. Чтение данных из файлов и таблиц.

Лекция.

Типы данных. Чтение данных формата csv и xml. Основные библиотеки для загрузки данных.

Практическое занятие.

Чтение данных формата csv и xml. Основные библиотеки для загрузки данных. Чтение текстовых файлов.

Тема 4.2. Подготовка данных.

Лекция.

Подготовка данных для их анализа. Изучение признаков (переменных) и оценка их влияния друг на друга и целевую переменную. Корреляционный анализ данных.

Практическое занятие.

Выполнение корреляционного анализа переменных и отбор признаков из данных.

Тема 4.3. Статистический анализ данных.

Лекция.

Статистический анализ данных.

Практическое занятие.

Расчет основных статистик во время проведения разведочного анализа данных.

Тема 4.4. Преобразование данных.

Лекция.

Разведочный анализ данных. Методы визуализации зависимостей между переменными.

Практическое занятие.

Выполнение разведочного анализа набора данных. Визуализация зависимостей переменных. Отбор информативных признаков.

Модуль 5. Решение задач машинного обучения.

Тема 5.1. Решение задачи классификации.

Лекция.

Типы задач машинного обучения. Виды моделей машинного обучения. Примеры решения задач классификации.

Практическое занятие.

Построение и обучение моделей машинного обучения. Кросс-валидация. Анализ результатов предсказания модели машинного обучения. Примеры решений задачи классификации.

Тема 5.2. Решение задачи регрессии.

Лекция.

Решение задач прогнозирования и регрессии. Анализ временных рядов.

Практическое занятие.

Линейные модели для задач регрессии. Деревья регрессии. Градиентный бустинг для задач регрессии. Примеры решений задач прогнозирования и регрессии. Анализ временных рядов.

Тема 5.3. Кластерный анализ данных.

Лекция.

Кластерный анализ данных.

Практическое занятие.

Изучение основных алгоритмов и моделей для проведения кластеризации в данных. Оценка результатов работы алгоритмов и их точности. Кластерный анализ исходного набора данных.

Тема 5.4. Искусственные нейронные сети для решения задач распознавания образов.

Лекция.

Использование искусственных нейронных сетей для решения задач распознавания образов и типы нейронных сетей. Основные параметры искусственных нейронных сетей. Обучение нейронных сетей.

Практическое занятие.

Использование искусственных нейронных сетей для решения задачи распознавания рукописных цифр и изображений.

Модуль 6. Разработка прикладного решения и создание интерфейса пользователя. Документирование аналитических решений.

Тема 6.1. Создание графического интерфейса пользователя.

Лекция.

Основы создания графического интерфейса пользователя (GUI).

Практическое занятие.

Создание графического интерфейса пользователя с помощью библиотеки Tkinter и PyQt.

Тема 6.2. Внедрение модели машинного обучения в графический интерфейс.

Лекция.

Внедрение модели машинного обучения в графический интерфейс.

Практическое занятие.

Внедрение модели машинного обучения в графический интерфейс.

Тема 6.3. Документирование выбора математического аппарата и алгоритма обучения.

Лекция.

Документирование выбора математического аппарата и алгоритма обучения.

Практическое занятие.

Документирование выбора математического аппарата и алгоритма обучения.

Тема 6.4. Документирование результатов обучения и апробации.

Лекция.

Документирование результатов обучения и апробации.

Практическое занятие.

Документирование результатов обучения и апробации.

Тема 6.5. Презентация результатов работы.

Практическое занятие.

Презентация результатов работы.

3.4. Календарный учебный график (порядок освоения модулей)*

Период обучения (недели, учебные годы)**	Наименование модуля
1 неделя (учебный год)	Модуль 1. Название
2 неделя (учебный год)	
	Итоговая аттестация

* Составляется под конкретную учебную группу.

**Точный порядок реализации модулей (дисциплин) обучения определяется в расписании занятий.

4. Организационно-педагогические условия реализации программы

4.1. Материально-технические условия реализации программы

Наименование помещения	Вид занятий	Наименование оборудования, программного обеспечения
<i>1</i>	<i>2</i>	<i>3</i>
Аудитория	Лекции	Компьютер, мультимедийный проектор, экран, доска, флипчарт
Мастерская	Лабораторные и практические занятия, тестирование, демонстрационный экзамен	Рабочее место преподавателя: компьютерный стол, компьютерный стул, компьютер-сервер. Рабочее место: компьютерный стол, компьютерный стул, компьютер в сборе (системный блок, 3 монитора, клавиатура, мышь, наушники, ноутбук, источник бесперебойного питания). Монитор/телевизор LCD. МФУ лазерное. Проектор. Доска интерактивная. Флипчарт. Оборудование, оснащение рабочих мест, инструменты и расходные материалы – в соответствии с инфраструктурным листом по компетенции Ворлдскиллс Оборудование, оснащение рабочих мест, инструменты и расходные материалы – в соответствии с инфраструктурным листом по компетенции Ворлдскиллс

4.2. Учебно-методическое обеспечение программы

- техническое описание компетенции;
- комплект оценочной документации по компетенции;
- печатные раздаточные материалы для слушателей;
- учебные пособия, изданных по отдельным разделам программы;
- профильная литература;
- отраслевые и другие нормативные документы;
- электронные ресурсы и т.д.

- официальный сайт оператора международного некоммерческого движения WorldSkills International - Союз «Молодые профессионалы (Ворлдскиллс Россия)» (электронный ресурс) режим доступа: <https://worldskills.ru>;
- единая система актуальных требований Ворлдскиллс (электронный ресурс) режим доступа: <https://esat.worldskills.ru>.
- 1. Андреас, Мюллер Введение в машинное обучение с помощью Python. Руководство для специалистов по работе с данными: моногр. / Мюллер Андреас. - М.: Альфа-книга, 2017. - **697** с.
- 2. Барский, А.Б. Логические нейронные сети / А.Б. Барский. - М.: Бином. Лаборатория знаний / Интернет-Университет Информационных Технологий (ИНТУИТ), 2017. - **336** с.
- 3. Домингос, П. Верховный алгоритм. Как машинное обучение изменит наш мир / П. Домингос. - М.: Манн, Иванов и Фербер, 2016. - **190** с.
- 4. Домингос, Педро Верховный алгоритм: как машинное обучение изменит наш мир / Педро Домингос. - Москва: **Гостехиздат**, 2015. - **989** с.
- 5. Круглов, В.В. Искусственные нейронные сети. Теория и практика / В.В. Круглов, В.В. Борисов. - М.: Горячая линия - Телеком; Издание 2-е, стер., 2002. - 382 с.
- 6. Нейронные сети. Statistica Neural Networks. Методология и технологии современного анализа данных. - М.: Горячая линия - Телеком, 2008. - 392 с.
- 7. Плас, Джейк Вандер Python для сложных задач. Наука о данных и машинное обучение. Руководство / Плас Джейк Вандер. - М.: Питер, 2018. - **527** с.
- 8. Себастьян, Рашка Python и машинное обучение / Рашка Себастьян. - М.: ДМК Пресс, 2017. - **809** с.
- 9. Татузов, А. Л. Нейронные сети в задачах радиолокации / А.Л. Татузов. - М.: Радиотехника, 2009. - 432 с.
- 10. Шибзухов, З.М. Конструктивные методы обучения сигма-пи нейронных сетей / З.М. Шибзухов. - М.: Наука, 2016. - **297** .

4.3.Кадровые условия реализации программы

Реализацию программы осуществляют преподаватели, прошедшие повышение квалификации по направлению подготовки. Приветствуется наличие у них сертификатов экспертов Ворлдскиллс по соответствующей компетенции одной из категорий:

- сертифицированный эксперт Ворлдскиллс;
- сертифицированных эксперт-мастер Ворлдскиллс;
- эксперт с правом проведения чемпионата;
- экспертов с правом оценки демонстрационного экзамена.

5. Оценка качества освоения программы

Итоговая аттестация проводится в форме тестирования и выполнения контрольной практической работы, для которой могут использоваться задания из Комплекта оценочной документации (КОД) № № 1.1.,1.2.,1.3 по компетенции «Машинное обучение и большие данные».

6. Составители программы

Программа составлена Данковской Еленой Владимировной, зав. мастерской «Машинное обучение и большие данные», на основе примерной программы, рекомендованной Белоусовым П.А., сертифицированным экспертом Ворлдскиллс по компетенции «Машинное обучение и большие данные», менеджером компетенции.