

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ  
**ФГАОУ ВО «СИБИРСКИЙ ФЕДЕРАЛЬНЫЙ УНИВЕРСИТЕТ»**

СОГЛАСОВАНО:

АНО ДПО «ШАД»

\_\_\_\_\_ С.Н. Федотов

« \_\_\_\_ » \_\_\_\_\_ 2021 г.



УТВЕРЖДАЮ:

Директор

НОЦ «Институт непрерывного образования»

Е.В. Мошкина

« \_\_\_\_ » \_\_\_\_\_ 2021 г.

СОГЛАСОВАНО:

Директор корпоративных программ

Департамента по работе с

персоналом СПАО «Ингосстрах»

\_\_\_\_\_ А.Н. Скуратов

« \_\_\_\_ » \_\_\_\_\_ 2021 г.

ДОПОЛНИТЕЛЬНАЯ ПРОФЕССИОНАЛЬНАЯ ПРОГРАММА  
ПРОФЕССИОНАЛЬНОЙ ПЕРЕПОДГОТОВКИ

**«Машинное обучение и анализ данных»**

Красноярск 2021

# **I. ОБЩАЯ ХАРАКТЕРИСТИКА ПРОГРАММЫ**

## **1.1. Аннотация программы**

Программа профессиональной переподготовки разработана СФУ совместно с компаниями АНО ДПО «ШАД» и СПАО «Ингосстрах» с учетом профессионального стандарта 06.042 «Специалист по большим данным» и реализуется в сетевом партнерстве этих организаций. Такая кооперация позволяет объединить лучшее от университетского образования, онлайн-обучения и реальной бизнес-практики в уникальный образовательный продукт – сетевую программу переподготовки кадров.

Обучение на программе предполагает самостоятельное освоение онлайн-курсов специализации «Машинное обучение и анализ данных» на coursera.org; очные занятия семинарского типа с преподавателями СФУ; интенсивы с приглашенными специалистами-практиками от компаний АНО ДПО «ШАД» и СПАО «Ингосстрах»; групповую работу над решением реальных кейсов от компаний АНО ДПО «ШАД» и СПАО «Ингосстрах» под руководством опытных менторов. Сопровождение и поддержка студентов оказываются в онлайн-формате. Содержание программы представлено модулями, раскрывающими процесс формирования компетентности слушателей.

Слушатели, освоившие программу профессиональной переподготовки, получают диплом о профессиональной переподготовке с правом ведения профессиональной деятельности в сфере создания и применения технологий больших данных.

## **1.2. Цель программы**

Цель реализации программы – подготовка конкурентоспособных специалистов для профессиональной деятельности в области машинного обучения и анализа данных в соответствии с требованиями профессионального стандарта 06.042 «Специалист по большим данным».

## **1.3. Компетенции (трудовые функции) в соответствии с профессиональным стандартом (формирование новых или совершенствование имеющихся)**

Программа ориентирована на формирование компетенций слушателей в области машинного обучения и анализа данных по обобщенным трудовым функциям профессионального стандарта 06.042 «Специалист по большим данным» (утвержден приказом Министерства труда и социальной защиты Российской Федерации от 6 июля 2020 года N 405н):

**А. Анализ больших данных с использованием существующей в организации методологической и технологической инфраструктуры:**

А/02.6 Планирование и организация аналитических работ с использованием технологий больших данных.

А/03.6 Подготовка данных для проведения аналитических работ по исследованию больших данных.

А/04.6 Проведение аналитического исследования с применением технологий больших данных в соответствии с требованиями заказчика.

**В. Управление этапами жизненного цикла методологической и технологической инфраструктуры анализа больших данных в организации:**

В/05.7 Управление получением, хранением, передачей, обработкой больших данных.

## **1.4. Планируемые результаты обучения**

Слушатель, освоивший программу переподготовки, будет обладать профессиональными компетенциями, включающими в себя способность:

РО1. Программировать на языке программирования Python на уровне, достаточном для решения задач анализа данных.

PO2. Применять для решения задач анализа данных средства таких библиотек, как NumPy, SciPy, Matplotlib, Pandas и Scikit-learn.

PO3. Понимать математические основы алгоритмов машинного обучения и методов анализа данных с выявлением следующих из них ограничений и определением на их основе свойств построенных моделей.

PO4. Формулировать и решать задачи обучения на размеченных данных (регрессия, классификация) с помощью различных методов машинного обучения.

PO5. Формулировать и решать задачи поиска структуры в данных.

PO6. Оценивать качество и адекватность моделей машинного обучения применительно к решаемым задачам.

PO7. Улучшать качество работы моделей машинного обучения с помощью методов регуляризации, кросс-валидации, понижения размерности.

PO8. Проверять статистические гипотезы относительно данных, оценивая зависимости методами корреляционного и регрессионного анализа.

PO9. Решать прикладные задачи анализа данных с учетом специфики предметной области.

PO10. Реализовывать полный цикл работы с данными, включая их сбор, предподготовку, описательный и визуальный анализ, подбор и реализацию подходящих методов анализа, оценку и интерпретацию результатов, презентацию итогов проекта.

### **1.5. Категория слушателей**

Обучающиеся 3-4 курсов бакалавриата и старших курсов специалитета; обучающиеся магистратуры, аспирантуры; выпускники образовательных организаций высшего образования, чья область: математика / компьютерные науки / информатика и вычислительная техника / информационные системы и технологии / анализ данных и машинное обучение и т.п.

### **1.6. Требования к уровню подготовки поступающего на обучение**

Базовые знания математики и основ программирования.

### **1.7. Условия приема на программу:**

- соответствие категории слушателей;
- прохождение вступительных испытаний на программу;
- возможность посещения очных занятий на базе СФУ в г. Красноярск.

В проект на конкурсной основе отбирается 30-40 кандидатов, показавших на вступительных испытаниях лучший результат, для дальнейшего обучения на программе.

### **1.8. Вступительные испытания**

Анкетирование кандидатов и констест на платформе Яндекс.Констест.

**1.9. Продолжительность обучения:** 312 академических часов (контактная и самостоятельная работа):

- онлайн – 238 часов самостоятельной работы, 22 часа контактной работы;
- очно – 52 часа контактной работы.

**1.10. Форма обучения:** очно-заочная с использованием электронного обучения и дистанционных образовательных технологий.

**1.11. Требования к материально-техническому обеспечению, необходимому для реализации дополнительной профессиональной программы повышения квалификации (требования к аудитории, компьютерному классу, программному обеспечению)**

Компьютерный мультимедийный класс на 20-30 ПК со стабильным интернет-соединением, в том числе через wi-fi; маркерная доска; проектор с возможностью подключения ноутбука; компьютеры с установленным ПО для студентов: Python 3.

**1.12. Документы об образовании:** сертификат об освоении специализации, диплом о профессиональной переподготовке установленного образца.

## II. ОСНОВНОЕ СОДЕРЖАНИЕ ПРОГРАММЫ

### 2.1. Учебно-тематический план

| №<br>п/п                       | Наименование и содержание<br>разделов и тем программы    | Количество часов          |          | Средства ЭО и ДОТ                                                                                                                                                                                                             | Резуль-<br>таты<br>обучения         |
|--------------------------------|----------------------------------------------------------|---------------------------|----------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------|
|                                |                                                          | Контакт-<br>ная<br>работа | СРС      |                                                                                                                                                                                                                               |                                     |
| Часть 1. Индивидуальная работа |                                                          |                           |          |                                                                                                                                                                                                                               |                                     |
| 1.                             | Математика и Python для анализа<br>данных                |                           | 3 недели | Онлайн-курс<br>специализации<br>на coursera.org:<br><a href="https://www.coursera.org/specializations/machine-learning-data-analysis#about">https://www.coursera.org/specializations/machine-learning-data-analysis#about</a> | PO1,<br>PO2,<br>PO3                 |
|                                | Введение                                                 | 6                         | 28       |                                                                                                                                                                                                                               |                                     |
| 1.1.                           | Библиотеки Python и линейная<br>алгебра                  |                           |          |                                                                                                                                                                                                                               |                                     |
| 1.2.                           | Оптимизация и матричные<br>разложения                    |                           |          |                                                                                                                                                                                                                               |                                     |
| 1.3.                           | Случайность                                              |                           |          |                                                                                                                                                                                                                               |                                     |
| 2.                             | Обучение на размеченных данных                           |                           | 4 недели | Онлайн-курс<br>специализации<br>на coursera.org:<br><a href="https://www.coursera.org/specializations/machine-learning-data-analysis#about">https://www.coursera.org/specializations/machine-learning-data-analysis#about</a> | PO2,<br>PO3,<br>PO4,<br>PO6,<br>PO7 |
| 2.1.                           | Машинное обучение и линейные<br>модели                   | 8                         | 36       |                                                                                                                                                                                                                               |                                     |
| 2.2.                           | Борьба с переобучением и<br>оценивание качества          |                           |          |                                                                                                                                                                                                                               |                                     |
| 2.3.                           | Линейные модели: классификация и<br>практические аспекты |                           |          |                                                                                                                                                                                                                               |                                     |
| 2.4.                           | Решающие деревья и композиции<br>алгоритмов              |                           |          |                                                                                                                                                                                                                               |                                     |
| 2.5.                           | Нейронные сети и обзор методов                           |                           |          |                                                                                                                                                                                                                               |                                     |
| 3.                             | Поиск структуры в данных                                 |                           | 4 недели | Онлайн-курс<br>специализации<br>на coursera.org:<br><a href="https://www.coursera.org/specializations/machine-learning-data-analysis#about">https://www.coursera.org/specializations/machine-learning-data-analysis#about</a> | PO2,<br>PO3,<br>PO5,<br>PO6,<br>PO7 |
| 3.1.                           | Кластеризация                                            | 10                        | 36       |                                                                                                                                                                                                                               |                                     |
| 3.2.                           | Понижение размерности и<br>матричные разложения          |                           |          |                                                                                                                                                                                                                               |                                     |
| 3.3.                           | Визуализация и поиск аномалий                            |                           |          |                                                                                                                                                                                                                               |                                     |
| 3.4.                           | Тематическое моделирование                               |                           |          |                                                                                                                                                                                                                               |                                     |
|                                | Промежуточная аттестация                                 | 2                         | 8        |                                                                                                                                                                                                                               |                                     |
| 4.                             | Построение выводов по данным                             |                           | 4 недели | Онлайн-курс<br>специализации<br>на coursera.org:<br><a href="https://www.coursera.org/specializations/machine-learning-data-analysis#about">https://www.coursera.org/specializations/machine-learning-data-analysis#about</a> | PO3,<br>PO8<br>PO9                  |
| 4.1.                           | Интервалы и гипотезы                                     | 8                         | 42       |                                                                                                                                                                                                                               |                                     |
| 4.2.                           | АБ-тестирование                                          |                           |          |                                                                                                                                                                                                                               |                                     |
| 4.3.                           | Закономерности и зависимости                             |                           |          |                                                                                                                                                                                                                               |                                     |
| 4.4.                           | Неделя задач                                             |                           |          |                                                                                                                                                                                                                               |                                     |

| №<br>п/п  | Наименование и содержание<br>разделов и тем программы                   | Количество часов          |                 | Средства ЭО и ДОТ                                                                                                                                                                                                    | Резуль-<br>таты<br>обучения |
|-----------|-------------------------------------------------------------------------|---------------------------|-----------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------|
|           |                                                                         | Контакт-<br>ная<br>работа | СРС             |                                                                                                                                                                                                                      |                             |
| 4.5.      | Групповая работа над проектом                                           | 8                         |                 | ZOOM, система вебинаров и видеоконференций СФУ                                                                                                                                                                       |                             |
| 5.        | <b>Прикладные задачи анализа данных</b>                                 |                           | <b>4 недели</b> | Онлайн-курс специализации на coursera.org: <a href="https://www.coursera.org/specializations/machine-learning-data-analysis#about">https://www.coursera.org/specializations/machine-learning-data-analysis#about</a> | PO9                         |
| 5.1.      | Бизнес-задачи                                                           | 10                        | 42              | ZOOM, система вебинаров и видеоконференций СФУ                                                                                                                                                                       |                             |
| 5.2.      | Анализ медиа                                                            |                           |                 |                                                                                                                                                                                                                      |                             |
| 5.3.      | Анализ текстов                                                          |                           |                 |                                                                                                                                                                                                                      |                             |
| 5.4.      | Рекомендации и ранжирование                                             |                           |                 |                                                                                                                                                                                                                      |                             |
| 5.5.      | Групповая работа над проектом                                           | 8                         |                 |                                                                                                                                                                                                                      |                             |
| <b>6.</b> | <b>Анализ данных</b>                                                    |                           | <b>3 недели</b> | Онлайн-курс специализации на coursera.org: <a href="https://www.coursera.org/specializations/machine-learning-data-analysis#about">https://www.coursera.org/specializations/machine-learning-data-analysis#about</a> | PO10                        |
| 6.1.      | Вступление.                                                             | 6                         | 24              | ZOOM, система вебинаров и видеоконференций СФУ                                                                                                                                                                       |                             |
| 6.2.      | Идентификация интернет-пользователей.                                   |                           |                 |                                                                                                                                                                                                                      |                             |
| 6.3.      | Прогнозирование оттока клиентов.                                        |                           |                 |                                                                                                                                                                                                                      |                             |
| 6.4.      | Прогнозирование временных рядов на карте.<br>Анализ тональности отзывов |                           |                 |                                                                                                                                                                                                                      |                             |
| 6.5.      | Групповая работа над проектом                                           | 6                         |                 |                                                                                                                                                                                                                      |                             |
| 7.        | <b>Аттестация</b>                                                       | 2                         | 8               | ZOOM, система вебинаров и видеоконференций СФУ                                                                                                                                                                       | PO1–<br>PO10                |
|           | <b>ИТОГО</b>                                                            | <b>74</b>                 | <b>238</b>      |                                                                                                                                                                                                                      |                             |
|           | <b>ВСЕГО</b>                                                            | <b>312</b>                |                 |                                                                                                                                                                                                                      |                             |

## 2.2. План учебной деятельности

| Результаты обучения | Учебные действия/формы текущего контроля                                                                                                                                                                                                | Используемые ресурсы/инструменты/технологии                                                                                                                                                                                                                              |
|---------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| PO1                 | <i>На семинарских занятиях:</i><br>составление конспектов, решение мини-задач на программирование.<br><i>В онлайн-курсе:</i><br>просмотр видео-лекций, выполнение тренировочных тестов/ выполнение тестов и заданий на программирование | Онлайн-курс специализации на coursera.org: <a href="https://www.coursera.org/specializations/machine-learning-data-analysis#about">https://www.coursera.org/specializations/machine-learning-data-analysis#about</a> ;<br>ZOOM; система вебинаров и видеоконференций СФУ |
| PO2                 | <i>На семинарских занятиях:</i><br>решение мини-задач на программирование.<br><i>В онлайн-курсе:</i><br>просмотр видео-лекций, выполнение тренировочных тестов / выполнение тестов и заданий на программирование                        | Онлайн-курс специализации на coursera.org: <a href="https://www.coursera.org/specializations/machine-learning-data-analysis#about">https://www.coursera.org/specializations/machine-learning-data-analysis#about</a> ;<br>ZOOM; система вебинаров и видеоконференций СФУ |
| PO3                 | <i>На семинарских занятиях:</i><br>составление конспектов, дискуссии, решение мини-задач на программирование.                                                                                                                           | Онлайн-курс специализации на coursera.org:                                                                                                                                                                                                                               |

| Результаты обучения | Учебные действия/формы текущего контроля                                                                                                                                                                                                                                                                                             | Используемые ресурсы/ инструменты/технологии                                                                                                                                                                                                                                |
|---------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|                     | <i>В онлайн-курсе:</i><br>просмотр видео-лекций, выполнение тренировочных тестов                                                                                                                                                                                                                                                     | <a href="https://www.coursera.org/specializations/machine-learning-data-analysis#about">https://www.coursera.org/specializations/machine-learning-data-analysis#about</a> ;<br>ZOOM; система вебинаров и видеоконференций СФУ                                               |
| PO4                 | <i>На семинарских занятиях:</i><br>разбор кейсов, составление конспектов.<br><i>В онлайн-курсе:</i><br>просмотр видео-лекций, выполнение тренировочных тестов, изучения дополнительного текстового материала.<br><i>На семинарских занятиях:</i> опросы, тесты; <i>в онлайн-курсе:</i> тесты, задания на программирование            | Онлайн-курс специализации на coursera.org:<br><a href="https://www.coursera.org/specializations/machine-learning-data-analysis#about">https://www.coursera.org/specializations/machine-learning-data-analysis#about</a> ;<br>ZOOM; система вебинаров и видеоконференций СФУ |
| PO5                 | <i>На семинарских занятиях:</i><br>разбор кейсов, составление конспектов.<br><i>В онлайн-курсе:</i><br>просмотр видео-лекций, выполнение тренировочных тестов, изучения дополнительного текстового материала                                                                                                                         | Онлайн-курс специализации на coursera.org:<br><a href="https://www.coursera.org/specializations/machine-learning-data-analysis#about">https://www.coursera.org/specializations/machine-learning-data-analysis#about</a> ;<br>ZOOM; система вебинаров и видеоконференций СФУ |
| PO6,<br>PO7         | <i>На семинарских занятиях:</i><br>разбор кейсов, составление конспектов.<br><i>В онлайн-курсе:</i><br>просмотр видео-лекций, выполнение тренировочных тестов, изучения дополнительного текстового материала                                                                                                                         | Онлайн-курс специализации на coursera.org:<br><a href="https://www.coursera.org/specializations/machine-learning-data-analysis#about">https://www.coursera.org/specializations/machine-learning-data-analysis#about</a> ;<br>ZOOM; система вебинаров и видеоконференций СФУ |
| PO8                 | <i>На семинарских занятиях:</i><br>разбор кейсов, составление конспектов.<br><i>В онлайн-курсе:</i><br>просмотр видео-лекций, выполнение тренировочных тестов, изучения дополнительного текстового материала<br><i>На онлайн-занятиях по работе над проектом:</i><br>разбор кейсов, выполнение групповых или индивидуальных проектов | Онлайн-курс специализации на coursera.org:<br><a href="https://www.coursera.org/specializations/machine-learning-data-analysis#about">https://www.coursera.org/specializations/machine-learning-data-analysis#about</a> ;<br>ZOOM; система вебинаров и видеоконференций СФУ |
| PO9                 | <i>На семинарских занятиях:</i><br>разбор кейсов.<br><i>В онлайн-курсе:</i><br>просмотр видео-лекций, выполнение тренировочных тестов, изучения дополнительного текстового материала.<br><i>На онлайн-занятиях по работе над проектом:</i> разбор кейсов, выполнение групповых или индивидуальных проектов                           | Онлайн-курс специализации на coursera.org:<br><a href="https://www.coursera.org/specializations/machine-learning-data-analysis#about">https://www.coursera.org/specializations/machine-learning-data-analysis#about</a> ;<br>ZOOM; система вебинаров и видеоконференций СФУ |
| PO10                | <i>На семинарских занятиях:</i><br>разбор кейсов.<br><i>В онлайн-курсе:</i><br>просмотр видео-лекций, выполнение тренировочных тестов, изучения дополнительного текстового материала.<br><i>На онлайн-занятиях по работе над проектом:</i> разбор кейсов, выполнение групповых или индивидуальных проектов                           | Онлайн-курс специализации на coursera.org:<br><a href="https://www.coursera.org/specializations/machine-learning-data-analysis#about">https://www.coursera.org/specializations/machine-learning-data-analysis#about</a> ;<br>ZOOM; система вебинаров и видеоконференций СФУ |

### 2.3. Виды и содержание самостоятельной работы

Выполнение самостоятельной работы слушателями предполагается в онлайн-режиме в рамках онлайн-курсов специализации на coursera.org: <https://www.coursera.org/specializations/machine-learning-data-analysis#about>.

Самостоятельно слушателями изучаются представленный теоретический материал, кейсы, задачи, примеры по машинному обучению и анализу данных, выполняются задания по программированию, оцениваются задания по программированию, выполненные сокурсниками (взаимооценка).

## III. УЧЕБНО-МЕТОДИЧЕСКИЕ МАТЕРИАЛЫ

### 3.1. Учебно-методическое обеспечение, в т.ч. электронные ресурсы в корпоративной сети СФУ и сети Интернет

Онлайн-курсы специализации на coursera.org: <https://www.coursera.org/specializations/machine-learning-data-analysis#about>.

### 3.2. Программное обеспечение (информационные обучающие системы, системы вебинаров, сетевые ресурсы хостинга видео, изображений, файлов, презентаций и др.)

1. Сервис видеоконференций ZOOM.
2. Система вебинаров и видеоконференций СФУ (<https://webinar.sfu-kras.ru/>).
3. Python 3.

## IV. ОЦЕНКА КАЧЕСТВА ОСВОЕНИЯ ПРОГРАММЫ

### 4.1. Формы аттестации, оценочные материалы, методические материалы

Программа предусматривает:

- промежуточную аттестацию по итогам прохождения первых трех онлайн-курсов специализации и работы на семинарских занятиях (учитывается посещаемость и активность работы на занятиях);
- итоговую аттестацию по итогам прохождения всех шести курсов специализации, работы на семинарских занятиях (учитывается посещаемость и активность работы на занятиях), выполнения проекта – решения кейса реальной практики от АНО ДПО «ШАД» и СПАО «Ингосстрах» по анализу данных (далее – кейс) под руководством менторов.

Методические материалы:

- видеолекции, дополнительный теоретический материал в текстовой форме, представленные в онлайн-курсе;
- презентации теоретического материала, Jupyter notebooks, разработанные преподавателями семинарских занятий.

Оценочные материалы:

- тесты;
- опросы;
- задания на программирование;
- кейсы реальных задач IT-практики.

*Слушатели, не прошедшие промежуточную аттестацию по каким-либо причинам, но успешно завершившие онлайн-курс(ы) специализации, могут получить удостоверение(я) СФУ о повышении квалификации объемом 36 часов, соответствующее тому онлайн-курсу специализации, который был успешно пройден.*

#### 4.2. Требования и содержание итоговой аттестации

Основанием для аттестации является портфолио слушателя программы, составленное из:

- сертификат о прохождении специализации на Coursera;
- отметки о зачете по результатам работы на семинарских занятиях;
- письменного отчета о решении кейсов по анализу данных с отметкой менторов об успешной защите.

Отчет о решении кейса по анализу данных должен содержать:

- титульный лист;
- постановку задачи;
- сведения об использованных методах решения;
- описания этапов выполнения проекта/решении кейса и промежуточных результатах;
- представление итоговых результатов, их интерпретацию;
- список использованных источников;
- приложение с программным кодом.

Программу составили:

Канд. физ.-мат. наук, доцент кафедры  
прикладной математики и компьютерной  
безопасности



Т.А. Кустицкая

Канд. пед наук, доцент кафедры прикладной  
математики и компьютерной безопасности



Р.В. Есин

Руководитель программы:

Начальник отдела информатизации УД



Г.В. Александрова