

СОГЛАСОВАНО:

Главный эксперт



Н.Н. Гордеев

СОГЛАСОВАНО:

Индустриальный партнер

АО «Научно-исследовательский институт «Вектор»,

заместитель директора Центра защиты информации



Д.В. Магницкий

М.П.



# ОПИСАНИЕ КОМПЕТЕНЦИИ «МАШИННОЕ ОБУЧЕНИЕ И БОЛЬШИЕ ДАННЫЕ»

**Наименование компетенции:** «Машинное обучение и большие данные»

**Формат участия в соревновании:** индивидуальный

**Описание компетенции.**

Машинное обучение определяется как построение компьютерных алгоритмов, которые могут автоматически улучшаться с помощью опыта и использования данных.

В контексте больших данных машинное обучение используется для обработки больших объемов данных, поиску закономерностей и извлечению из них ключевой актуальной информации.

Основная цель специалиста по машинному обучению и большим данным – создание информационных технологий нового поколения, обеспечивающих экономически эффективное извлечение полезной информации из больших объемов разнообразных данных путем высокой скорости их сбора, обработки и анализа, и применение этих технологий в информационно-аналитической деятельности, в системах управления и принятия решений, а также для разработки на их основе новых продуктов и услуг.

Задачи компетенции заключаются в постановке и нахождении путей решения прикладных задач с использованием современных инструментов работы с большими данными.

Машинное обучение (machine learning) лежит в основе многих инновационных технологий различных секторов экономики. Лучшие практики использования машинного обучения:

1. В промышленности (минимизация простоев на производстве, системы управления производством, выявление угроз безопасности);
2. В финансах (обнаружение мошенничества для безопасных транзакций, алгоритмическая торговля – анализ настроений, финансовые консультации и управление портфелем, оценка кредитоспособности);

3. В здравоохранении (диагностика заболеваний; автоматические роботизированные операции; улучшение клиентского сервиса);

4. В ритейле и маркетинге (прогнозирование действий покупателей, персонализированные предложения и реклама, прогнозирование спроса и автоматизация закупок);

5. В логистике (экономия топлива и повышение производительности транспорта, предотвращение сбоев в поставках).

Актуальность компетенции «Машинное обучение и большие данные» связана в первую очередь с растущим количеством информации, необходимостью её обработки и анализа, а также бурным развитием инноваций.

Используя язык программирования Python, среду разработки Jupyter Notebook, инструменты хранения и извлечения данных SQL, PySpark, инструменты визуализации данных и библиотеки numpy, Scikit-Learn, Keras, Catboost, специалисты этой компетенции применяют алгоритмы машинного обучения для задач классификации и регрессии, оценивают качество обученных моделей, настраивают параметры и применяют нейронные сети для решения прикладных задач.

### **Нормативные правовые акты**

Поскольку Описание компетенции содержит лишь информацию, относящуюся к соответствующей компетенции, его необходимо использовать на основании следующих документов:

#### **1. ФГОС СПО**

– 09.02.07 Информационные системы и программирование (Приказ Минобрнауки России от 09.12.2016 N 1547 (ред. от 17.12.2020) Зарегистрировано в Минюсте России 26 декабря 2016 г. N 44936);

#### **2. Профессиональный стандарт**

– 06.042 Специалист по большим данным (УТВЕРЖДЕН приказом Министерства труда и социальной защиты Российской Федерации от 6 июля 2020 года N 405н);

- 6.001 «Программист», утвержден Министерством труда и социальной защиты РФ от 20 июля 2022 года № 424н;
- 6.015 «Специалист по информационным системам», Утвержден Приказом Министерства труда и социальной защиты Российской Федерации №896н от 18.11.2014;
- 6.022 «Системный аналитик», утвержден приказом Министерства труда и социальной защиты РФ 28.10.2014 № 809н;

### 3. ЕКС

- Техник-программист, Единый квалификационный справочник должностей руководителей, специалистов и других служащих (ЕКС), 2019, утвержден Постановлением Минтруда РФ от 21.08.1998 N 37 (редакция от 15.05.2013);
- Инженер-программист (программист), Единый квалификационный справочник должностей руководителей, специалистов и других служащих (ЕКС), 2019, утвержден Постановлением Минтруда РФ от 21.08.1998 N 37 (редакция от 15.05.2013)

### 4. Квалификационные характеристики (профессиограмма):

- Содержание труда: осуществляет сбор, обработку и анализ больших данных с использованием существующей в организации методологической и технологической инфраструктуры. Управляет этапами жизненного цикла методологической и технологической инфраструктуры анализа больших данных в организации, разработкой продуктов, услуг и решений на основе больших данных. Разрабатывает и внедряет новые методы и технологии исследования больших данных.
- Профессионально важные качества: хорошее зрение; оперативная память; устойчивое концентрированное внимание; стойкая работоспособность; дисциплинированность; аналитическое мышление.
- Квалификационные требования: СПО, ВУЗы.

– Медицинские противопоказания: Снижение остроты зрения; нарушение функций щитовидной железы; сердечно-сосудистые заболевания; болезни крови.

## 5. ГОСТы

– ГОСТ Р 59895-2021 «Технологии искусственного интеллекта в образовании. Общие положения и терминология» (утв. и введен в действие приказом Федерального агентства по техническому регулированию и метрологии от 26 ноября 2021 г. N 1617-ст);

– ГОСТ Р 52872-2019 «Интернет-ресурсы и другая информация, представленная в электронно-цифровой форме. Приложения для стационарных и мобильных устройств, иные пользовательские интерфейсы» (утв. и введен в действие Приказом Федерального агентства по техническому регулированию и метрологии от 29 августа 2019 г. N 589-ст);

– ГОСТ Р ИСО/МЭК 90003-2014 «Разработка программных продуктов. Руководящие указания по применению ИСО 9001:2008 при разработке программных продуктов» (утв. приказом Федерального агентства по техническому регулированию и метрологии от 23 октября 2014 г. N 1405-ст)

– ГОСТ Р ИСО/МЭК 15910-2002 «Информационная технология. Процесс создания документации пользователя программного средства» (Постановлением Госстандарта России от 25 июня 2002 г. N 249-ст).

## 6. Санитарно-эпидемиологические правила и нормативы СанПиН 2.2.2/2.4.2620-10

– «Гигиенические требования к персональным электронно-вычислительным машинам и организации работ» Изменения N 2 к СанПиН 2.2.2/2.4.1340-03 (утв. постановлением Главного государственного санитарного врача РФ от 30 апреля 2010 г. N 48)

– Санитарные правила СП 2.2.3670-20 «Санитарно-эпидемиологические требования к условиям труда» (постановление Главного государственного санитарного врача Российской Федерации от 18.05.2009 N 30 "Об утверждении СП 2.2.9.2510-09" (зарегистрировано Минюстом России 09.06.2009, регистрационный N 14036)).

Перечень профессиональных задач специалиста по компетенции определяется профессиональной областью специалиста и базируется на требованиях современного рынка труда к данному специалисту.

№ п/п	Виды деятельности/трудовые функции
1	Подготовка данных для проведения аналитических работ по исследованию больших данных
2	Планирование и организация аналитических работ с использованием технологий больших данных
3	Проведение аналитического исследования с применением технологий больших данных в соответствии с требованиями заказчика
4	Разработка продуктов на основе встроенной аналитики больших данных
5	Разработка инфраструктурных решений на основе аналитики больших данных
6	Подготовка данных для проведения аналитических работ по исследованию больших данных