|  |  |
| --- | --- |
|  |  |

ОПИСАНИЕ КОМПЕТЕНЦИИ

«Машинное обучение и большие данные»

2025 г.

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

**Описание компетенции**

Компетенция «Машинное обучение и большие данные» направлена на подготовку специалистов в области анализа данных и искусственного интеллекта – одной из ключевых технологий цифровой экономики. Машинное обучение представляет собой быстроразвивающийся раздел искусственного интеллекта, посвящённый разработке алгоритмов, способных обучаться на основе данных и принимать решения без прямого программирования.

Профессиональная деятельность включает сбор, обработку и анализ больших объёмов данных, построение, обучение и оптимизацию моделей машинного обучения, а также интерпретацию результатов для получения практически значимых выводов. В рамках компетенции формируются навыки корректной предобработки данных, проведения разведочного анализа, решения задач классификации, регрессии и кластеризации, а также разработки программных продуктов на основе аналитики больших данных.

Специалисты в области машинного обучения и больших данных крайне востребованы в различных отраслях российской экономики. Согласно анализу рынка труда, спрос на таких специалистов уже сегодня сопоставим с уровнем востребованности разработчиков программного обеспечения, особенно в малом и среднем бизнесе.

Компетенция основана на профессиональном стандарте 06.042 и интегрирована в образовательные программы СПО по специальностям 09.02.13 Интеграция решений с применением технологий искусственного интеллекта и 09.02.11 Разработка и управление программным обеспечением, что обеспечивает соответствие подготовки кадров требованиям работодателей.

Крупные российские компании, такие как ООО «Код Безопасности», «СКБ Контур», «Главстрой Санкт-Петербург», а также корпорации «Росатом», «СИБУР Холдинг» и «Ростех», активно участвуют в развитии компетенции, предоставляют реальные производственные задачи для соревнований и признают стандарты подготовки.

Компетенция «Машинное обучение и большие данные» способствует формированию кадрового резерва под запросы экономики, синхронизируя образовательные программы с реальными потребностями индустрии.

Специалист по машинному обучению и большим данным работает с комплексными и неструктурированными данными, решая задачи прогнозирования, классификации, сегментации и автоматизации аналитических процессов.

Ключевые этапы профессиональной деятельности включают специалиста в области машинного обучения и больших данных включают:

* Планирование и организацию аналитических работ с использованием технологий больших данных;
* Парсинг и предобработку данных – сбор, очистка и трансформация данных для последующего анализа;
* Разведочный анализ данных (EDA) – выявление закономерностей, аномалий и значимых признаков;
* Построение, обучение и оптимизация моделей машинного обучения;
* Разработку программного продукта и средств интеграции, включая создание сервисов на основе аналитики;
* Поддержку и сопровождение готовых решений в промышленной среде.

В профессиональной деятельности используются современные технологии и инструменты обработки данных и машинного обучения:

* Языки программирования: Python, SQL;
* Библиотеки и фреймворки: Pandas, NumPy, Scikit-learn, TensorFlow, PyTorch, XGBoost;
* Средства визуализации: Matplotlib, Seaborn, Plotly;
* Платформы для работы с большими данными: Apache Spark, Hadoop;
* Среды разработки: Jupyter Notebook, PyCharm, VS Code;
* Инструменты управления данными и интеграции: Git, Docker, API.

Работа ведётся с высоконагруженными и сложными наборами данных, требующими применения как традиционных, так и передовых методов анализа.

Машинное обучение и анализ больших данных применяются в широком спектре отраслей:

* Финансовый сектор (банки, страхование) – для оценки рисков, выявления мошенничества;
* Строительство и девелопмент – для анализа рынка и оптимизации проектов;
* Информационная безопасность – для анализа угроз и защиты ИТ-инфраструктуры;
* Промышленность и энергетика – для прогнозирования отказов оборудования;
* Сельское хозяйство, здравоохранение, логистика, образование – для автоматизации и повышения эффективности процессов.

Компетенция активно внедряется в образовательную практику и уже охватывает более чем 20 регионов России.

С 2017 года проведено более 200 соревнований – от квалификационных отборочных и региональных до межрегиональных, отраслевых и международных. Межрегиональные этапы проходят на площадках колледжей и техникумов, таких как Саранский электромеханический колледж республики Мордовия, и включают задания от реальных работодателей.

Компетенция «Машинное обучение и большие данные» способствует укреплению технологического суверенитета России, поддержке талантливой молодёжи и созданию преемственности между образованием и производством.

Компетенция продолжает развиваться под руководством менеджера компетенции Белоусова Павла Анатольевича, при поддержке экспертного сообщества и партнёров отрасли. Цель компетенции – подготовка высококвалифицированных специалистов, способных решать сложные задачи в условиях цифровой экономики.

**Нормативные правовые акты**

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

* ФГОС СПО по специальности 09.02.13 Интеграция решений с применением технологий искусственного интеллекта, утверждён Приказом Министерства просвещения Российской Федерации от 24.12.2024 №1025.
* ФГОС СПО по специальности 09.02.11 Разработка и управление программным обеспечением, утверждён Приказом Министерства просвещения Российской Федерации от 24.02.2024 №138.
* Профессиональный стандарт Специалист по большим данным, утверждён Министерством труда и социальной защиты Российской Федерации от 06.07.2020 №405н;
* Отраслевые/корпоративные стандарты: 06.042 Специалист по большим данным, 06.001 Программист
* ГОСТы Единой системы программной документации (ЕСПД);
* СанПиН 2.4.4.3172-14 «Санитарно-эпидемиологические требования к устройству, содержанию и организации режима работы образовательных организаций дополнительного образования детей»;
* СанПиН 2.2.2/2.4.134003 «Гигиенические требования к персональным электронно-вычислительным машинам и организации работы».

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

|  |  |
| --- | --- |
| **№ п/п** | **Виды деятельности** |
| 1 | Разработка кода для искусственного интеллекта |
| 2 | Разработка, администрирование и защита баз данных |
| 3 | Обучение готовых моделей искусственного интеллекта |
| 4 | Разработка и интеграция модулей программного обеспечения |
| 5 | Проектирование и разработка информационных систем |
| 6 | Разработка бизнес-приложений |

|  |  |
| --- | --- |
| **№ п/п** | **Трудовые функции** |
| 1 | Планирование и организация аналитических работ с использованием технологий больших данных |
| 2 | Подготовка данных для проведения аналитических работ по исследованию 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 сервисов на основе аналитики больших данных |