

УДК 004.457

РАЗРАБОТКА ПРОЕКТА АВТОМАТИЗАЦИИ РАЗВЕРТЫВАНИЯ ИНФРАСТРУКТУРЫ ДЛЯ РЕШЕНИЯ ЗАДАЧ МАШИННОГО ОБУЧЕНИЯ

Тихонов Л.А. (Национальный Исследовательский Университет ИТМО)

Научный руководитель – преподаватель практики Брюханов К.В.

(Национальный Исследовательский Университет ИТМО)

Введение. В настоящее время объем данных растет, и это доставляет неудобства при работе с ними. MLOps — набор практик и инструментов, которые помогают стандартизировать и повысить эффективность процессов работы с машинным обучением. Это направление закрывает задачи трекинга, версионирования и мониторинга, возникающих при создании и развертывании систем на основе машинного обучения.

Основная часть. Инфраструктура организована на основе:

- Hadoop - версии 3.3 в качестве инструмента для хранения большого количества тренировочных данных. Вместо одного большого компьютера для обработки и хранения данных Hadoop предлагает использовать для параллельного анализа огромных наборов данных кластеры на базе стандартного аппаратного обеспечения.
- Spark – версии 3 для организации работы конвейера непрерывной поставки компонентов машинного обучения. Также инструменты Spark ML позволяют ускорять разработку и настройку моделей.
- Bash скрипты – для связи компонентов между собой и автоматического запуска.

Выводы. В данной работе была развернута инфраструктура для эффективной работы моделей машинного обучения. Осуществлена автоматизация рутинных действий ML разработчика.

Список использованных источников:

1. Брюханов К.В., Петров В.В., Авксентьева Е.Ю. Анализ существующих проблем в организации непрерывной доставки и развертывании программного продукта // Современная наука: актуальные проблемы теории и практики. Серия: Естественные и технические науки - 2020. - № 5. - С. 79-82

2. Руководство по MLOps. [Электронный ресурс]. –

URL: <https://medium.com/sciforce/ml-ops-comprehensive-beginners-guide-c235c77f407f>.

3. Документация по Hadoop. [Электронный ресурс]. –

URL: <https://hadoop.apache.org/>

4. Документация по Spark. [Электронный ресурс]. –

URL: <https://spark.apache.org/>

Тихонов Л.А. (автор)

Подпись

Брюханов К.В. (научный руководитель)

Подпись