

# Теоретические вопросы

---

## Теоретическое задание 1.

Я – начинающий аналитик, которому нужно постигнуть машинное обучение. Как мне может помочь в этом облако?

Облачные вычисления предоставляют доступ к мощным вычислительным ресурсам без необходимости их покупки. Это позволяет сэкономить средства и время, которые можно потратить на изучение машинного обучения. Такие ресурсы предоставляются в формате сервисов или платформ, что упрощает их использование. Рассмотрим несколько облачных инструментов, которые могут быть полезны для освоения машинного обучения.

### Google Colab и Kaggle

**Google Colab** – это облачный сервис для работы с Jupyter Notebook, где уже установлены ключевые библиотеки для машинного обучения, такие как TensorFlow, Keras, PyTorch и другие. Он предоставляет возможность использования GPU и TPU для ускорения вычислений. Бесплатная версия имеет ограничения по времени и доступному оборудованию, но подписка на Colab Pro расширяет эти возможности, позволяя проводить небольшие исследования.

Альтернативой может служить **Kaggle Notebook**, где доступ к GPU и TPU предоставляется бесплатно на несколько часов в сутки. Kaggle также предлагает удобное хранилище данных, избавляя от необходимости повторной загрузки после перезапуска ноутбука.

### Google Cloud Datalab

**Google Cloud Datalab** – это инструмент для работы с большими объемами данных и их визуализации, также основанный на Jupyter Notebook. В отличие от Colab, он предоставляет возможности работы с данными в BigQuery и поддерживает SQL. Услуга является платной, но доступен бесплатный пробный период. Подробнее о Datalab можно узнать [здесь](#).

### Amazon SageMaker

**Amazon SageMaker** – комплексный сервис для машинного обучения, включающий инструменты для обработки данных, визуализации и обучения моделей. Он поддерживает предобученные модели для таких задач, как классификация изображений и распознавание речи. Сервис предоставляет подробные руководства, которые помогут освоить его использование. Подробнее ознакомиться с возможностями SageMaker можно [здесь](#).

### Yandex Cloud

**Yandex Cloud** – облачная платформа с широким набором сервисов для работы с данными и машинным обучением. Среди них Vision для обработки изображений, SpeechKit для распознавания

речи, Yandex GPT API для генерации текста и другие. При регистрации пользователям предоставляется грант в размере 4000 рублей, который можно использовать в течение полугода. Более подробная информация доступна [здесь](#).

---

Эти облачные решения позволяют начинающим аналитикам эффективно изучать машинное обучение, предоставляя доступ к мощным инструментам без значительных затрат. Выбор сервиса зависит от конкретных задач и предпочтений, однако каждый из них значительно упрощает процесс обучения и экспериментов.

## Теоретическое задание 2.

Допустим, у вас развернута большая инфраструктура: сотни виртуальных машин, десятки разных сервисов. Как будете мониторить состояние инфраструктуры? Интересуют логи, отказы, нагрузка и тд. Приведите пример для любого из провайдеров

Для мониторинга инфраструктуры в Yandex Cloud используется сервис **Yandex Monitoring**. Он предоставляет широкий спектр возможностей, таких как сбор и отображение метрик, анализ логов, построение графиков и создание дашбордов.

### Алерты

Yandex Monitoring позволяет настроить автоматические уведомления о событиях, таких как сбои или превышение порога нагрузки. Оповещения могут отправляться через email, SMS или пуш-уведомления в мобильное приложение Yandex Cloud. Также можно предоставить доступ нескольким пользователям для получения уведомлений. Подробности о настройке алертов доступны [здесь](#).

### Логи и сбор метрик

Для сбора метрик и логов используется **Unified Agent**. Он совместим с различными операционными системами Linux и может быть установлен на физические компьютеры, виртуальные машины или запущен через Docker. Информация о возможностях и настройке Unified Agent находится [здесь](#).

### Нагрузка и сбои

Нагрузка отслеживается с помощью **Application Load Balancer**, который может распределять запросы между серверами, защищать сервисы от несанкционированного доступа и обеспечивать безопасность трафика через TLS. Подробнее об этом инструменте можно узнать [здесь](#).

### Дашборды и аналитика

Сервис предоставляет возможность создавать настраиваемые дашборды, которые отображают ключевые метрики и логи. С помощью языка запросов Yandex Monitoring можно выполнять

сложные выборки и анализ данных, например, строить графики нагрузки CPU за последний час. Подробнее о языке запросов [здесь](#).

Yandex Monitoring — это универсальный инструмент, предоставляющий удобный функционал для мониторинга и анализа состояния инфраструктуры.

## **Вывод**

Эти вопросы показали мне, как облачные сервисы помогают изучать машинное обучение и эффективно управлять инфраструктурой. Я узнала о доступных инструментах, таких как Google Colab, Kaggle, Yandex Monitoring, которые упрощают работу с данными, моделями и мониторингом систем.