

# ТЕСТОВОЕ ЗАДАНИЕ

для стажера на позицию

**Разработчик Machine learning**

## Задача:

Необходимо реализовать RAG-систему поиска информации по технической документации. Провести исследование и настройку системы с точки зрения повышения качества выдаваемых результатов.

Комплект документации доступен по ссылке: <https://infotecs.ru/downloads/documents/vipnet-coordinator-hw-5/>

## Комментарии:

Задание исследовательское, а следовательно – творческое.

В решении мы ожидаем увидеть кругозор кандидата в области Machine Learning и LLM, а также умение самостоятельно проводить исследования и демонстрировать результаты.

1. Необходимо самостоятельно сформировать все необходимые выборки из приведенных документов.
2. Самостоятельно составить бенчмарк и показать на нем результаты своих исследований и этапов настройки.
3. По результатам составить отчет, включающий описание разработанной системы, описание исследований и выборок.

## Ограничение:

Допускается использовать только маленькие модели, не более 8млрд параметров.

## ТРЕБОВАНИЯ К ПРИСЫЛАЕМЫМ РЕШЕНИЯМ

Присылаемое решение должно содержать:

- Все файлы проекта с исследованиями и программной реализацией системы (.ipynb, .py).
- Отчет, включающий описание разработанной системы, инструкцию по установке и использованию, описание проведенных исследований по настройке и улучшению качества системы (.docx, .pdf);
- Сформированные выборки в любом текстовом формате (.json, .txt, .csv,...)
- Дополнительно можно дать ссылку на репозиторий на github.
- Решение необходимо загрузить в архиве.

Максимальное время  
на выполнение  
задания

**1 неделя**

