ООО ТЕРРИКСА

ML LLM инженер по обработке документов

от 170 000р

- Требуемый опыт работы: От 1 до 3 лет
- Занятость: Частичная
- График: Удаленная работа
- Образование: Не имеет значения

О вакансии

TERRIXA — интеллектуальная система для работы с техническими документами и производственными данными. Ищем ML/LLM-инженера, который будет развивать подготовку данных, классификацию запросов и оценку качества ответов, а также участвовать в тестировании сквозных сценариев анализа документов и отчётов по скважинам.

Чем предстоит заниматься

- Формировать, размечать и актуализировать наборы пользовательских запросов и эталонных ответов для обучения и тестирования.
- Поддерживать сбалансированные выборки по типам задач: общий поиск по документам, точные значения из таблиц, графовые связи, расчёты и вызов агентов.
- Развивать и оценивать классификацию намерений пользователя и маршрутизацию запросов между компонентами системы.
- Проверять качество ответов на технических документах и суточных отчётах.
- Участвовать в проверке сводного анализа группы отчётов: по одной скважине за период и по нескольким скважинам за период.
- Оценивать полноту и корректность сводных результатов: выполненные работы, статусы, выявленные проблемы и динамика.
- Тестировать сквозной цикл обработки документов: загрузка, извлечение данных, запись в базы, выполнение запросов и получение результата, в том числе в Excel.
- Разбирать ошибки и случаи отсутствия результата, группировать причины и предлагать улучшения качества.
- Настраивать и оценивать локальный запуск моделей на GPU/CPU, включая Ollama или vLLM, параметры параллельной обработки и таймауты.
- Работать совместно с разработчиками парсера, backend-командой и тестировщиками; документировать эксперименты и результаты.

Что важно для этой работы

- Опыт работы ML-, NLP- или LLM-инженером.
- Уверенный Python и практические навыки обработки данных, в том числе с pandas.
- Опыт подготовки датасетов, разметки данных, оценки классификаторов и анализа ошибок.
- Понимание LLM/RAG-систем, поиска по документам и подходов к оценке качества ответов.
- Практический опыт работы с SQL и структурированными данными.
- Умение тестировать пайплайны обработки документов и находить причины некорректных результатов.
- Навыки работы в Linux-среде и чтения логов.
- Умение ясно фиксировать ход эксперимента, результаты и ограничения.

Будет преимуществом

- Опыт с Hugging Face Transformers, scikit-learn, PyTorch или аналогичными инструментами.
- Практика запуска LLM через Ollama или vLLM и настройки инференса на GPU.
- Опыт с векторными и графовыми базами данных.
- Опыт обработки PDF, DOCX, Excel и документов со сложными таблицами.
- Знакомство с нефтегазовой тематикой, буровыми процессами и суточными рапортами.
- Опыт нагрузочного тестирования на больших наборах документов.

Задачи на первые месяцы

- Разобраться в текущих наборах данных и сценариях классификации запросов.
- Согласовать воспроизводимый набор тестов и метрики качества для ключевых типов поиска и отчётных сценариев.
- Регулярно предоставлять команде разбор ошибок с примерами и предложениями по улучшению.
- Помочь стабилизировать оценку работы моделей в документном пайплайне и отчётных сценариях.

Адрес размещения вакансии:

Москва, Арбат, д. 24

Контактная информация:

- +7 (931) 521-07-00
- Людмила
- office@terrixa.com

Другие вакансии этой компании

Инженер-электронщик / инженер по наладке оборудования
от 150 000р
Собирать, проверять и настраивать датчики, электронные платы, блоки АЦП и другое оборудование.Выполнять монтаж и ручную пайку компонентов, в том числе SMD.Диагностировать и устранять неисправности плат, блоков, кабелей и разъёмов.Проводить измерения,

проверять питание, сигналы и работоспособность оборудования.Тестировать оборудование после сборки или ремонта.Подключать платы с микроконтроллерами, загружать готовые прошивки и настраивать параметры по инструкции.Работать со схемами и технической документацией.Подключать компьютеры, блоки сбора данных и другое оборудование в локальную сеть.Настраивать сетевые параметры и проверять связь между устройствами.Искать причины отсутствия связи и ошибок передачи данных.Помогать с подключением и настройкой оборудования на буровых объектах.Помогать с установкой и настройкой программы, драйверов и подключаемого оборудования.Разбирать обращения пользователей и помогать восстановить работу системы.Определять, где возникла проблема: в оборудовании, подключении, сети или настройках программы.Собирать информацию об ошибках и передавать сложные вопросы разработчикам.
деревня Крёкшино, Терминальный проезд, д. 5с13

Инженер по удалённому мониторингу бурения
от 90 000р
Контроль качества данных, контроль передачи данныхКонтроль процесса бурения, первичный анализ ситуацийВзаимодействие с представителями Заказчика по линии технической поддержкиУдаленный контроль процесса бурения по данным, поступающим с буровых: глубина, нагрузка на долото, скорость проходки, расход, крутящий момент, обороты и другие доступные показатели, включая показатели ННБКонтроль работоспособности Amirig: доступность сервисов, своевременное поступление и регистрация данных, корректность визуализации параметровВыявление пропусков, задержек, недостоверных данных, фиксация замеченных отклонений и своевременное реагирование в соответствии с регламентами ЗаказчикаОтслеживание событий и предупреждений систем, входящих в платформу Amirig, анализ событий с учетом операций на скважине; своевременное оповещение ответственным специалистам, представителям Заказчика по установленным регламентам.Взаимодействие с инженерами буровой бригады, специалистами ГТИ, ННБ, технической поддержки и представителями Заказчика по вопросам работы платформы, передачи данныхВедение журнала наблюдений и инциндентов, подготовка кратких сводок и отчетовУчастие в проверке работы систем после настройки, обновлений, передача разработчикам обратной связи: замечаний, предложений по улучшению качества сервиса
Москва, Новый Арбат, д. 24