

Курсови проекти

Защо проекти?

- 60% от крайната оценка идва от проекта
- Практически опит с реални LLM задачи
- Портфолио за бъдеща работа
- Възможност за творчество и експериментиране

Формат на проекта

- Индивидуален или в екип (до 2-3 човека)
- Код + кратък доклад (3-5 страници)
- Демонстрация/презентация (10-15 минути)
- Срок: края на семестъра

Идея 1: Fine-tuning на GPT-2 за български език

Описание:

- Fine-tune GPT-2 (или подобен модел) върху български текстови данни
- Възможни dataset-и: български речник, Wikipedia BG, новинарски статии

Технически акценти:

- Работа с Hugging Face Transformers
- Подготовка на данни, tokenization
- Training loop, checkpointing
- Оценка на perplexity

Очакван резултат:

- Модел, който генерира смислен български текст

Идея 2: RAG система за счетоводни документи

Описание:

- Retrieval-Augmented Generation за отговаряне на въпроси
- Фокус: българско счетоводство, данъчно законодателство, НАП указания

Технически акценти:

- Парсване и chunking на PDF документи
- Vector database (ChromaDB, FAISS, Pinecone)
- Embedding модели (multilingual)
- Интеграция с LLM (OpenAI API / Ollama)

Очакван резултат:

- Чатбот, който отговаря на счетоводни въпроси с цитиране на източници

Идея 3: Sentiment анализ на български отзиви

Описание:

- Класификация на потребителски отзиви (положителни/отрицателни/неутрални)
- Източници: Pazaruvaj.com, booking отзиви, социални мрежи

Технически акценти:

- Web scraping или готов dataset
- Fine-tuning на BERT/mBERT за класификация
- Сравнение с baseline (Bag of Words, TF-IDF)

Очакван резултат:

- Модел с измерена точност (accuracy, F1)
- Демо интерфейс за live класификация

Идея 4: Chatbot за образователни цели

Описание:

- Чатбот асистент за конкретен предмет/курс
- Използва материали от курса като контекст

Технически акценти:

- RAG или fine-tuning подход
- Prompt engineering за образователен тон
- Обработка на conversation history

Очакван резултат:

- Бот, който отговаря на въпроси по определена тема
- Оценка от реални потребители

Идея 5: Автоматично обобщаване на новини

Описание:

- Summarization на български новинарски статии
- Възможност за multi-document summarization

Технически акценти:

- Extractive vs Abstractive summarization
- Работа с mT5 или подобен модел
- ROUGE метрики за оценка

Очакван резултат:

- Система за генериране на кратки резюмета
- Сравнителен анализ на различни подходи

Идея 6: Named Entity Recognition за български

Описание:

- Разпознаване на именувани същности (имена, организации, места)
- Приложение: анализ на новини, юридически документи

Технически акценти:

- Анотиране на данни (или използване на съществуващи)
- Fine-tuning на sequence labeling модел
- BIO tagging схема

Очакван резултат:

- NER модел с precision/recall метрики
- Визуализация на разпознатите entities

Идея 7: Превод между езици

Описание:

- Machine translation: BG ↔ EN или друга езикова двойка
- Може да е специализиран домейн (технически, медицински)

Технически акценти:

- Encoder-decoder архитектура
- Работа с паралелни корпуси
- BLEU score оценка

Очакван резултат:

- Преводач за специфичен домейн
- Сравнение с Google Translate

Идея 8: AI Agent за конкретна задача

Описание:

- Изграждане на агент с достъп до инструменти
- Примери: код асистент, research асистент, travel planner

Технически акценти:

- Tool calling (function calling)
- ReAct или подобен pattern
- Orchestration logic

Очакван резултат:

- Работещ агент за практическа задача
- Демонстрация на multi-step reasoning

Критерии за оценка

Критерий	Тежест	Описание
Техническа сложност	35%	Адекватно използване на ML/LLM техники
Презентация	25%	Ясно обяснение, демо, отговори на въпроси
Резултати и анализ	25%	Метрики, сравнения, интерпретация
Код и документация	15%	Четим код, README, reproducibility

Техническа сложност - какво значи?

По-високо се оценява:

- Fine-tuning на модел (не само API calls)
- Собствена обработка на данни
- Сравнителен анализ на подходи
- Работа с по-малко документирани инструменти

По-ниско:

- Само wrapper около OpenAI API
- Copy-paste от tutorials без разбиране
- Липса на базови метрики

Презентация - съвети

- Започнете с **проблема**, който решавате
- Покажете **работещо демо**
- Обяснете **ключовите технически решения**
- Бъдете готови за **въпроси** за архитектурата
- 10-15 минути презентация + 5 минути въпроси

Резултати и анализ

- **Baseline сравнение:** покажете подобрене спрямо прост подход
- **Метрики:** accuracy, F1, BLEU, ROUGE, perplexity – каквото е релевантно
- **Грешки и ограничения:** анализирайте какво не работи
- **Честност:** не криете проблеми, обяснете ги

Код и документация

- **README** с инструкции за стартиране
- **requirements.txt** или **pyproject.toml**
- Код структуриран логично (не един огромен notebook)
- Коментари за неочевидни части

Процес и срокове

Етап	Срок
Избор на тема и одобрение	Седмица 3-4
Checkpoint (progress report)	Седмица 8
Финална презентация	Седмица 12-13
Предаване на код и доклад	1 седмица след презентация

Checkpoint (междинен отчет)

- Кратко описание на прогреса (1-2 страници)
- Какво е направено, какво предстои
- Евентуални проблеми и блокери
- **Не се оценява**, но е задължителен

Избор на тема

- Можете да предложите **собствена идея** (препоръчително!)
- Консултирайте се преди да започнете
- По-добре амбициозен проект с частичен успех, отколкото тривиален с пълен

Често задавани въпроси

Q: Трябва ли да използвам GPU?

- Не е задължително. Fine-tuning на малки модели може и на CPU
- Google Colab предоставя безплатен GPU достъп

Q: Мога ли да използвам платени API-та?

- Да, но покажете че разбирате какво се случва "под капака"

Q: Какво ако проектът не работи добре?

- Анализът на неуспехите е също ценен
- Обяснете защо не работи и какво бихте пробвали другояче

Ресурси

- Hugging Face Hub: модели и datasets
- Kaggle: datasets и notebooks
- Papers with Code: state-of-the-art методи
- LangChain / LlamaIndex: за RAG системи
- Ollama: локално стартиране на модели

Въпроси?

Контакт:

- По време на лекции
- Email
- Office hours

Успех с проектите!