

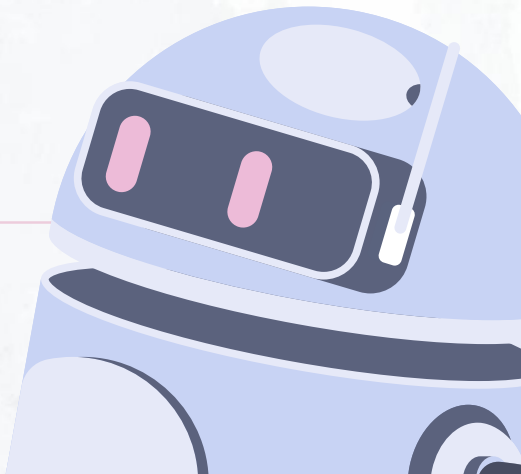
Развој интелигентног асистента за програмирање



(AI)

Ментор
др Бранко Арсић

Студент
Марија Јоловић

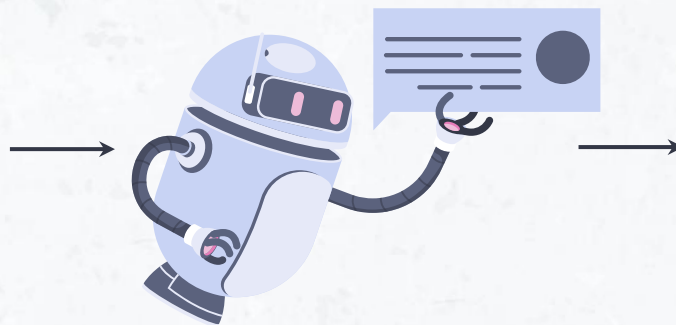


Идеја и мотивација



- Асистент за програмере
- Предвиђање наредног дела кода
- Инспирација: постојећи алати попут *GitHub Copilot*-а
- Проблем код домен-специфичних кодова

```
1 def factorial(n):  
2     ~
```



```
1 def factorial(n):  
2     if n == 0:  
3         return 1  
4     else:  
5         return n * factorial(n - 1)  
6  
7 print(factorial(5))
```

Циљеви истраживања

- Истражити и демонстрирати развој асистента
- Проучавање архитектура великих језичких модела
- Одабир модела
- Фино подешавање и евалуација
- Имплементација екстензије

01 Модели (LLM)

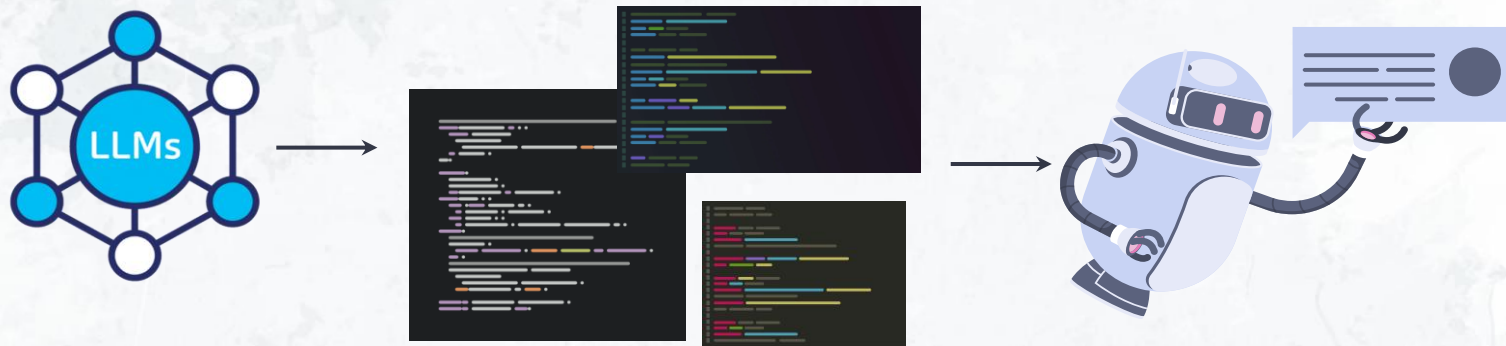
02 Подаци

03 Фино подешавање

04 Резултати

05 Екстензија

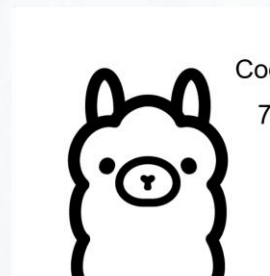
06 Демо





01 Модели

- 6 различитих модела
- Различит број параметара
- Базирани на енкодеру - предикције
- Базирани на декодеру - chat модели



CodeLlama
7B



StarCoderBase
1B
3B
7B



6.7B
DeepSeekCoder



MistralAI
7B

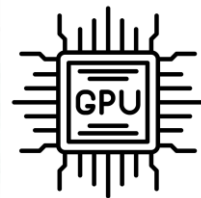
02 Скуп података



- Скуп кодова са 80+ програмских језика
- Јавно доступни кодови
- Разноврсност језика – флексибилност и боља генерализација
- Прочишћени скуп са 6000 узорака

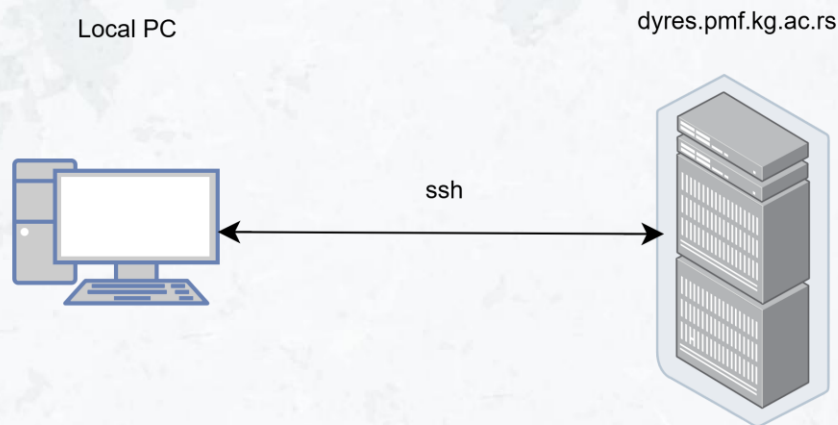


Hugging Face



03 Тренирање

- Хардверска ограничења: Две графичке картице од 16GB
- Технике оптимизације меморије – 7B модели

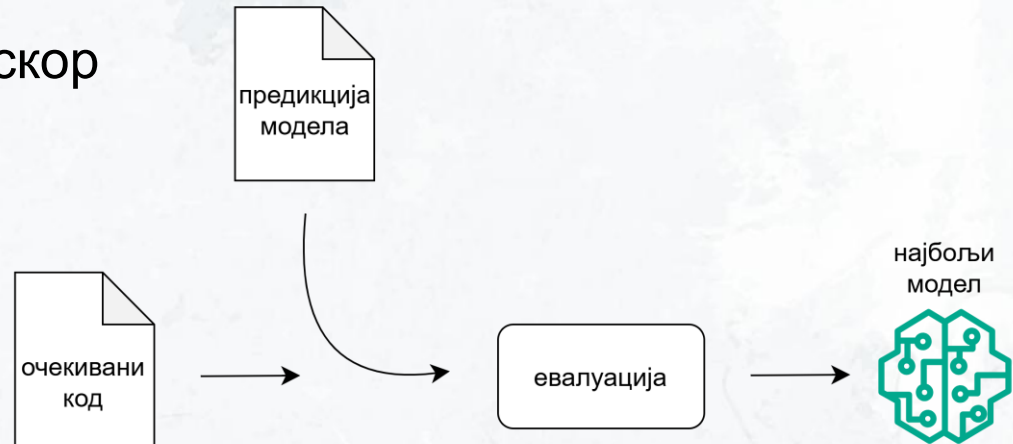


04 Резултати - метрике

- BLEU, ROUGE
- Изабрана метрика CodeBLEU за евалуацију
- Сличност текста
- Синтаксна исправност
- Логичка структура кода

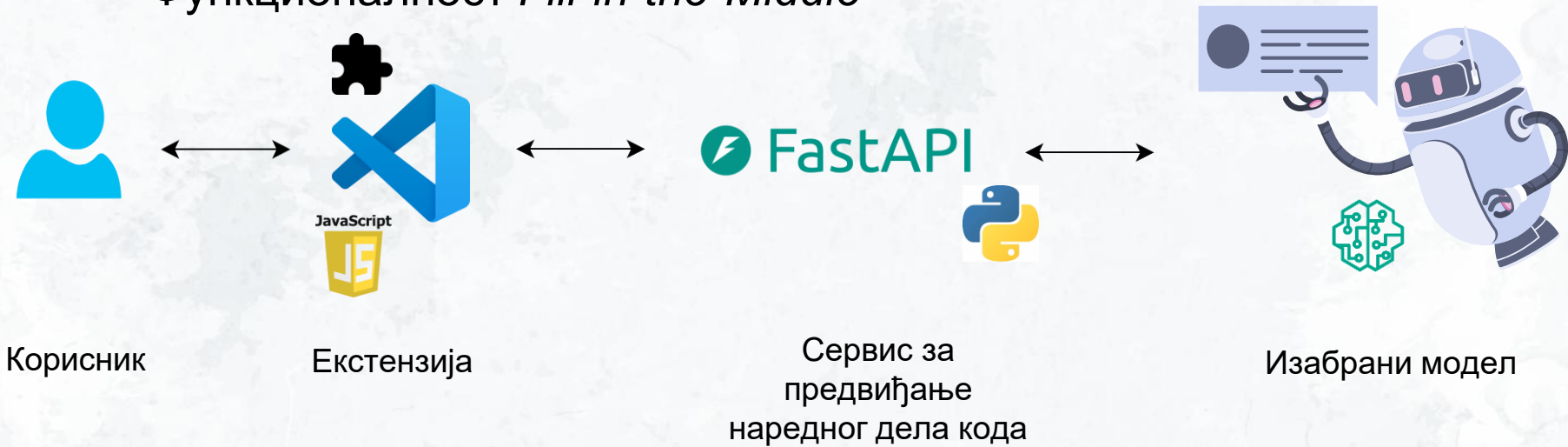
04 Резултати – поређење модела

- Мануелно и кроз метрике, на различитим упитима
- Тестни скуп – 500 узорака
- Сви модели су остварили више резултате од својих базних верзија
- StarCoder 7B – највиши скор



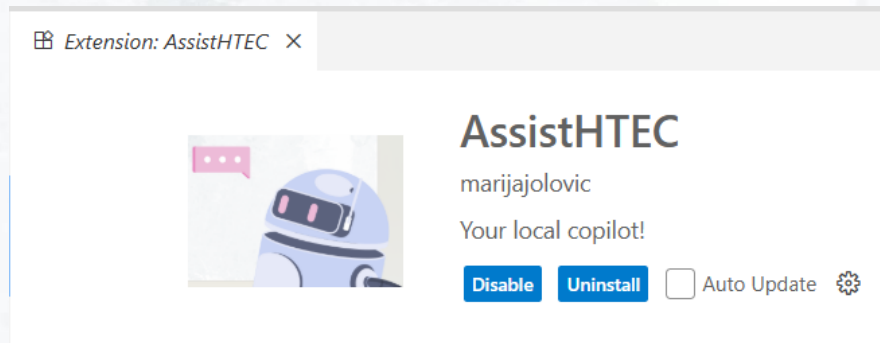
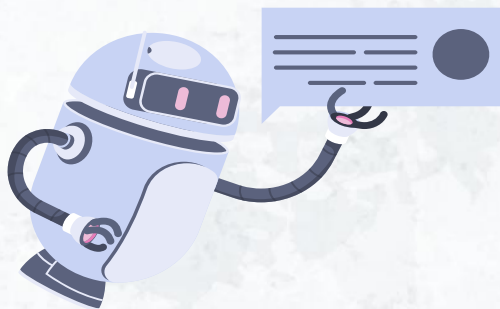
05 Употреба модела - екстензија

- Развој функционалне екстензије у радном окружењу *VS Code*
- Функционалност *Fill-in-the-Middle*



05 Употреба модела - екстензија

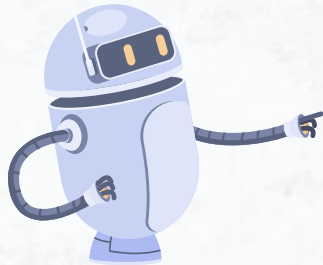
- Паковање екстензије у фајл `.vsix`



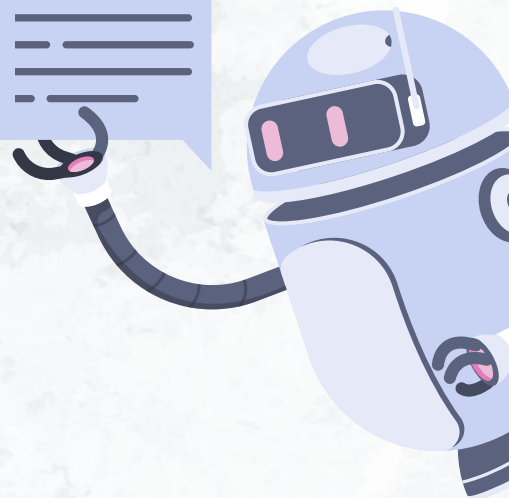
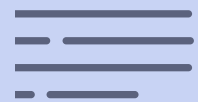
06 Демо

Закључак

- Успешно развијен сопствени AI асистент
- Додатно тренирање модела на специфичним кодовима
- Интеграција са постојећим екстензијама (*Twinnu*)
- Изазов добре евалуације



Хвала на пажњи!



→
Питања?

marija.jolovic@pmf.kg.ac.rs