Rezultati zero-shot in few-shot učenja

* Stopenjsko few-shot učenje: 1,2,3,4,6,8 primerov primerov na (pod)kategorijo.
* Granularnost nalog: 4 kategorije in 12 podkategorij.
* Pristop učenja: tradicionalni učni pozivi in učenje s pomočjo sheme ontologije (v json formatu).
* Velikosti podatkovnih množic: 48, 96, 192, 372 primerov, ki so enakomerno porazdeljeni na posamezne (pod)kategorije.

1. **Model GaMS:**

EKSPERIMENTALNI PREGLED

━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━

• Testiranih pogojev: 112

• Uspešnih pogojev: 77

• Uspešnost eksperimentov: 68.8%

• Dostopne kategorije: 12

NAJBOLJŠA KONFIGURACIJA

━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━

• Ime: few\_shot\_traditional\_prompts\_fine\_12sub\_48\_epc2

• Natančnost: 1.000

• Razpoznavanje: 0.067

• Pristop: traditional\_prompts

• Granularnost: fine\_12sub

• Primeri na kategorijo: 2

• Velikost podatkov: 48

STATISTIČNA ANALIZA Z VIZUALIZACIJAMI

━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━

• Povprečna natančnost: 0.236 ± 0.215

• Mediana natančnosti: 0.233

• Razpon natančnosti: 0.000 - 1.000

• Povprečno razpoznavanje: 0.377 ± 0.294

• Vizualizacije: 9 integriranih grafov z interpretacijskimi vodili

Osnovne statistike:

Povprečna natančnost: 0.236 ± 0.215

Najboljša natančnost: 1.000

Razpon natančnosti: 0.000 - 1.000

Povprečni razpon razčlenjevanja: 0.377 ± 0.294

Najbolj uspešna kombinacija:

Ime: few\_shot\_traditional\_prompts\_fine\_12sub\_48\_epc2

Natančnost: 1.000

Config: {'learning\_approach': 'traditional\_prompts', 'granularity': 'fine\_12sub', 'shot\_type': 'few\_shot', 'examples\_per\_category': 2, 'dataset\_size': 48}

A group of graphs and charts

Description automatically generated

**GaMS-1B Learning Curves: Dataset Size Effect GaMS-1B Few-Shot Learning Progressions**

A graph of a graph of a graph

Description automatically generated with medium confidence

PODROBNEJŠA ANALIZA KATEGORIJ IN PODKATEGORIJ

======================================================================

PREGLED USPEŠNOSTI GLAVNIH KATEGORIJ:

LEKSIKALNO PARAFRAZIRANJE:

• Experiments: 41

• Mean accuracy: 0.667 ± 0.201

• Performance range: 0.125 - 1.000

- zero\_shot: 0.717 (n=4)

- few\_shot: 0.661 (n=37)

• Best config: few\_shot, traditional\_prompts

(Examples: 8, Size: 192)

PRAGMATIČNO PARAFRAZIRANJE:

• Experiments: 40

• Mean accuracy: 0.738 ± 0.112

• Performance range: 0.412 - 1.000

- zero\_shot: 0.778 (n=4)

- few\_shot: 0.733 (n=36)

• Best config: few\_shot, traditional\_prompts

(Examples: 6, Size: 48)

SINTAKTIČNO PARAFRAZIRANJE:

• Experiments: 39

• Mean accuracy: 0.760 ± 0.160

• Performance range: 0.333 - 1.000

- zero\_shot: 0.794 (n=4)

- few\_shot: 0.756 (n=35)

• Best config: few\_shot, traditional\_prompts

(Examples: 3, Size: 96)

SEMANTIČNO PARAFRAZIRANJE:

• Experiments: 45

• Mean accuracy: 0.467 ± 0.221

• Performance range: 0.000 - 1.000

- zero\_shot: 0.289 (n=4)

- few\_shot: 0.485 (n=41)

• Best config: few\_shot, traditional\_prompts

(Examples: 8, Size: 192)

PREGLED USPEŠNOSTI PODKATEGORIJ:

1. aktivno\_pasivna\_sprememba

Performance: 0.796 ± 0.145 (Easy, Stable)

Range: 0.400 - 0.938

Experiments: 16

Main category: sintaktično\_parafraziranje

Best: 0.938 (few\_shot, 8 ex)

2. nominalizacija

Performance: 0.791 ± 0.234 (Easy, Variable)

Range: 0.000 - 1.000

Experiments: 14

Main category: sintaktično\_parafraziranje

Best: 1.000 (few\_shot, 6 ex)

3. jedrnatost\_razširjanje\_konteksta

Performance: 0.789 ± 0.114 (Easy, Stable)

Range: 0.500 - 0.933

Experiments: 21

Main category: leksikalno\_parafraziranje

Best: 0.933 (few\_shot, 6 ex)

4. hipernim\_hiponim

Performance: 0.773 ± 0.158 (Easy, Stable)

Range: 0.500 - 0.938

Experiments: 14

Main category: semantično\_parafraziranje

Best: 0.938 (few\_shot, 8 ex)

5. pretvorba\_stavčnih\_vrst

Performance: 0.748 ± 0.192 (Easy, Stable)

Range: 0.000 - 0.929

Experiments: 24

Main category: leksikalno\_parafraziranje

Best: 0.929 (few\_shot, 2 ex)

6. sprememba\_fokusa

Performance: 0.744 ± 0.223 (Easy, Variable)

Range: 0.000 - 1.000

Experiments: 21

Main category: pragmatično\_parafraziranje

Best: 1.000 (few\_shot, 3 ex)

7. zamenjava\_sopomenk

Performance: 0.743 ± 0.124 (Easy, Stable)

Range: 0.500 - 0.900

Experiments: 20

Main category: semantično\_parafraziranje

Best: 0.900 (few\_shot, 6 ex)

8. sprememba\_modalnosti

Performance: 0.741 ± 0.254 (Easy, Variable)

Range: 0.000 - 0.938

Experiments: 16

Main category: pragmatično\_parafraziranje

Best: 0.938 (few\_shot, 8 ex)

9. antonim\_z\_negacijo

Performance: 0.738 ± 0.150 (Easy, Stable)

Range: 0.400 - 1.000

Experiments: 26

Main category: semantično\_parafraziranje

Best: 1.000 (zero\_shot, 0 ex)

10. prilagajanje\_tona

Performance: 0.734 ± 0.138 (Easy, Stable)

Range: 0.429 - 0.900

Experiments: 19

Main category: pragmatično\_parafraziranje

Best: 0.900 (few\_shot, 6 ex)

11. časovno\_aspektualne\_spremembe

Performance: 0.732 ± 0.198 (Easy, Stable)

Range: 0.333 - 1.000

Experiments: 30

Main category: leksikalno\_parafraziranje

Best: 1.000 (few\_shot, 8 ex)

12. sprememba\_vrstnega\_reda

Performance: 0.636 ± 0.233 (Easy, Variable)

Range: 0.286 - 1.000

Experiments: 11

Main category: sintaktično\_parafraziranje

Best: 1.000 (few\_shot, 2 ex)

1. **Model LlaMa-3.1-8B:**

EKSPERIMENTALNI PREGLED

━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━

• Testiranih pogojev: 112

• Uspešnih pogojev: 92

• Uspešnost eksperimentov: 82.1%

• Dostopne kategorije: 12

NAJBOLJŠA KONFIGURACIJA

━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━

• Ime: few\_shot\_traditional\_prompts\_coarse\_4main\_192\_epc1

• Natančnost: 1.000

• Razpoznavanje: 0.050

• Pristop: traditional\_prompts

• Granularnost: coarse\_4main

• Primeri na kategorijo: 1

• Velikost podatkov: 192

STATISTIČNA ANALIZA Z VIZUALIZACIJAMI

━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━

• Povprečna natančnost: 0.238 ± 0.188

• Mediana natančnosti: 0.231

• Razpon natančnosti: 0.000 - 1.000

• Povprečno razpoznavanje: 0.569 ± 0.318

• Vizualizacije: 9 integriranih grafov z interpretacijskimi vodili

Osnovne statistike:

Povprečna natančnost: 0.238 ± 0.188

Najboljša natančnost: 1.000

Razpon natančnosti: 0.000 - 1.000

Povprečni razpon razčlenjevanja: 0.569 ± 0.318

Najbolj uspešna kombinacija:

Name: few\_shot\_traditional\_prompts\_coarse\_4main\_192\_epc1

Accuracy: 1.000

Config: {'learning\_approach': 'traditional\_prompts', 'granularity': 'coarse\_4main', 'shot\_type': 'few\_shot', 'examples\_per\_category': 1, 'dataset\_size': 192}

A group of graphs and charts

Description automatically generated

**LlaMA 3.1 Learning Curves: Dataset Size Effect LlaMA 3.1 Few-Shot Learning Progressions**

A graph of a person's body

Description automatically generated with medium confidence PODROBNEJŠA ANALIZA KATEGORIJ IN PODKATEGORIJ

======================================================================

PREGLED USPEŠNOSTI GLAVNIH KATEGORIJ:

SINTAKTIČNO PARAFRAZIRANJE:

• Experiments: 51

• Mean accuracy: 0.573 ± 0.237

• Performance range: 0.067 - 1.000

- zero\_shot: 0.738 (n=8)

- few\_shot: 0.543 (n=43)

• Best config: few\_shot, traditional\_prompts

(Examples: 1, Size: 192)

PRAGMATIČNO PARAFRAZIRANJE:

• Experiments: 48

• Mean accuracy: 0.669 ± 0.180

• Performance range: 0.000 - 0.875

- zero\_shot: 0.813 (n=7)

- few\_shot: 0.644 (n=41)

• Best config: zero\_shot, schema\_guided

(Examples: 0, Size: 192)

LEKSIKALNO PARAFRAZIRANJE:

• Experiments: 51

• Mean accuracy: 0.702 ± 0.134

• Performance range: 0.231 - 0.923

- zero\_shot: 0.665 (n=8)

- few\_shot: 0.709 (n=43)

• Best config: few\_shot, schema\_guided

(Examples: 6, Size: 96)

SEMANTIČNO PARAFRAZIRANJE:

• Experiments: 49

• Mean accuracy: 0.613 ± 0.191

• Performance range: 0.200 - 0.929

- zero\_shot: 0.388 (n=7)

- few\_shot: 0.650 (n=42)

• Best config: few\_shot, schema\_guided

(Examples: 2, Size: 96)

PREGLED USPEŠNOSTI PODKATEGORIJ:

1. hipernim\_hiponim

Performance: 0.908 ± 0.040 (Easy, Stable)

Range: 0.800 - 0.950

Experiments: 26

Main category: semantično\_parafraziranje

Best: 0.950 (zero\_shot, 0 ex)

2. nominalizacija

Performance: 0.897 ± 0.129 (Easy, Stable)

Range: 0.222 - 0.950

Experiments: 29

Main category: sintaktično\_parafraziranje

Best: 0.950 (zero\_shot, 0 ex)

3. sprememba\_vrstnega\_reda

Performance: 0.867 ± 0.160 (Easy, Stable)

Range: 0.333 - 1.000

Experiments: 15

Main category: sintaktično\_parafraziranje

Best: 1.000 (few\_shot, 6 ex)

4. aktivno\_pasivna\_sprememba

Performance: 0.863 ± 0.180 (Easy, Stable)

Range: 0.154 - 0.950

Experiments: 29

Main category: sintaktično\_parafraziranje

Best: 0.950 (zero\_shot, 0 ex)

5. prilagajanje\_tona

Performance: 0.841 ± 0.079 (Easy, Stable)

Range: 0.500 - 0.929

Experiments: 29

Main category: pragmatično\_parafraziranje

Best: 0.929 (few\_shot, 6 ex)

6. pretvorba\_stavčnih\_vrst

Performance: 0.840 ± 0.177 (Easy, Stable)

Range: 0.000 - 0.950

Experiments: 38

Main category: leksikalno\_parafraziranje

Best: 0.950 (zero\_shot, 0 ex)

7. antonim\_z\_negacijo

Performance: 0.833 ± 0.188 (Easy, Stable)

Range: 0.000 - 1.000

Experiments: 46

Main category: semantično\_parafraziranje

Best: 1.000 (few\_shot, 1 ex)

8. sprememba\_modalnosti

Performance: 0.814 ± 0.242 (Easy, Variable)

Range: 0.056 - 0.950

Experiments: 31

Main category: pragmatično\_parafraziranje

Best: 0.950 (zero\_shot, 0 ex)

9. časovno\_aspektualne\_spremembe

Performance: 0.810 ± 0.121 (Easy, Stable)

Range: 0.462 - 0.941

Experiments: 42

Main category: leksikalno\_parafraziranje

Best: 0.941 (few\_shot, 8 ex)

10. jedrnatost\_razširjanje\_konteksta

Performance: 0.764 ± 0.191 (Easy, Stable)

Range: 0.133 - 0.923

Experiments: 39

Main category: leksikalno\_parafraziranje

Best: 0.923 (few\_shot, 2 ex)

11. sprememba\_fokusa

Performance: 0.755 ± 0.227 (Easy, Variable)

Range: 0.000 - 0.929

Experiments: 38

Main category: pragmatično\_parafraziranje

Best: 0.929 (few\_shot, 2 ex)

12. zamenjava\_sopomenk

Performance: 0.751 ± 0.266 (Easy, Variable)

Range: 0.067 - 1.000

Experiments: 29

Main category: semantično\_parafraziranje

Best: 1.000 (few\_shot, 2 ex)