

Feature 32154

Language: Japanese
Model: meta-llama/Llama-3.2-1B
Layer: model.layers.8.mlp
SAE Model: EleutherAI/sae-Llama-3.2-1B-131k
Selected Token Probability: 0.396
Entropy: 1.069

Activation Range

0.092-0.202	0.202-0.312	0.312-0.422	0.422-0.532	0.532-0.642	0.642-0.752	0.752-0.862
0.862-0.972	0.972-1.082	1.082-1.192				

Interpretation

"The highlighted tokens are primarily Bulgarian morphemes or word segments, especially those forming relative pronouns (such as \"която\", \"който\", \"което\", \"кои\"), as well as common suffixes and inflections (like \"ъ\", \"оя\", \"ой\", \"r\", \"точ\", \"еч\", \"еств\", \"ал\", \"ото\"). These segments often appear at the boundaries of words or as part of grammatical constructions, indicating a focus on morphological and syntactic markers in Bulgarian text."

Score Type	Accuracy	Precision	Recall	F1 score	TPR	TNR	FPR	FNR
detection	0.73	0.96	0.48	0.64	0.48	0.98	0.02	0.52
fuzz	0.71	0.957	0.44	0.603	0.44	0.98	0.02	0.56

Japanese

#examples: [('paws-x', 991), ('flores', 997)]

paws-x-138. <|begin_of_text|>最大濃度に達した場合、毒性毒性が発生するまでに1〜2日の◆◆延があるかもしれません。

paws-x-957. <|begin_of_text|>ESAによってeNLP（European NLP Solver）とも呼ばれるWORHPは、連続的な数学的最適化問題を数値的に解くための非線形ソフトウェアライブラリです。

paws-x-883. <|begin_of_text|>Pidouxは、Pablo Casalsの `` Jackie ``でチェロ奏者PabloLarraínとして登場しました。

paws-x-519. <|begin_of_text|>もう1つのマーストン社の製品ラインは1931年に始まり、後に英国のシー◆◆ルとして知られるマーストンシー◆◆ルとして最初に販売されたマリン船外機エンジンで始まりました。

Text Examples for Each Interval

interval 1

Range: 1.082-1.192
#examples: 5

flores-412. <|begin_of_text|>Кавалерията е армия, която се сражава на кон. Асирийската кавалерия се сражавала на голите гърбове на конете си, тъй като седлото още не било измислено.

flores-276. <|begin_of_text|>В ролята му на председател на подкомисията за надзор и разследване, която е под егидата на Комисията по енергетика и търговия на Парламента Стърнс разследва дали данъците се използват за финансиране на аборти чрез Планирано Родителство.

flores-719. <|begin_of_text|>Карибска държава, която заема източната половина на остров Испаньола, който дели с Хаити, е Доминиканската република (на испански: Република Доминикана).

flores-952. <|begin_of_text|>Въпреки че според вас е логично да изберете авиокомпанията, с която летите най-много, трябва да знаете, че предлаганите привилегии често са различни и друга авиокомпания в същия съюз може да бъде много по-щедра по отношение на точките за редовни клиенти.

xnli-274. <|begin_of_text|>знам , че не познавам никого с ума си , който казва , че го правя , защото искам да го направя . знам , че има хора , които мислят , че го правя , защото го желая .

interval 2

Range: 0.972-1.082
#examples: 24

xnli-601. <|begin_of_text|>когато американският дипломат , който наблюдава международните мирни наблюдатели в косово , обвини сръбската полиция в клането , югославската (т.е. американският дипломат е изключително доволен от работата и постиженията си .