

Feature 49501

Activation Range

Language: Vietnamese	0.043-0.258	0.258-0.474	0.474-0.689	0.689-0.904	0.904-1.119	1.119-1.335	1.335-1.55
Model: meta-llama/Llama-3.2-1B	1.55-1.765	1.765-1.981	1.981-2.196				
Layer: model.layers.5.mlp							
SAE Model: EleutherAI/sae-Llama-3.2-1B-131k							
Selected Token Probability: 0.589							
Entropy: 1.191							

Interpretation

"The highlighted tokens are characteristic morphemes, suffixes, or letter clusters commonly found in Romanian and related Eastern European place names, river names, and surnames, often marking grammatical or regional features."

Score Type	Accuracy	Precision	Recall	F1 score	TPR	TNR	FPR	FNR
detection	0.52	0.583	0.14	0.226	0.14	0.9	0.1	0.86
fuzz	0.55	0.6	0.3	0.4	0.3	0.8	0.2	0.7

Vietnamese

#examples: [('xnli', 995), ('flores', 989)]

xnli-229. <|begin_of_text|>Chiến dịch hồi giáo cho phân vùng được dẫn dắt bởi luật sư bombay được đào tạo Luân Đôn , muhammad ali jinnah . Muhammad Ali Jinnah đã đến Luân Đôn để chiến dịch cho phân vùng .

xnli-386. <|begin_of_text|>Chiến tranh không phải là người apache , chắc chắn là vậy . Chắc chắn đó không phải là người apache , chắc chắn rồi .

xnli-172. <|begin_of_text|>Căn nhà này rất nhỏ bé và đơn giản , với một phòng ngủ , một nhà bếp nhỏ , và một vài phòng xã hội . Ngôi nhà rất lớn và tự hào hơn phòng ngủ , một nhà bếp khổng lồ , và một hồ bơi olympic đầy đủ .

xnli-707. <|begin_of_text|>Đây là tất cả những gì trong sạch và khó chịu , rung động với những người khó chịu , nhưng cũng có những câu hỏi khó khăn và những câu trả lời không được đáp ứng . Điều này thật khó chịu và khó chịu .

Text Examples for Each Interval

interval 1

Range: 1.981-2.196

#examples: 4

paws-x-921. <|begin_of_text|>Cirja # 7 überholte 2 TD Scramble und erreichte außerdem 2 TD mit Florin Oltean # 85 und mit Dan Crasnic # 93.

paws-x-933. <|begin_of_text|>Der Fluss Crișul Mic ist ein Nebenfluss des Flusses Doba in Rumänien.

paws-x-921. <|begin_of_text|>Cirja 7は2 TD Scrambleを通過し、さらに2 TD 1をFlorin Oltean 85に、1 TDをDan Crasnic 93に達成しました。

paws-x-921. <|begin_of_text|>Cirja 7通过了2次TD Scramble，并且还获得了2个TD，其中一个Florin Oltean 85，另一个是Dan Crasnic 93。

interval 2

Range: 1.765-1.981

#examples: 7

paws-x-933. <|begin_of_text|>La rivière Crișul Mic est un affluent de la rivière Doba en Roumanie.

paws-x-478. <|begin_of_text|>Der Fluss Fădimac ist ein Nebenfluss des Flusses Bega in Rumänien.

paws-x-921. <|begin_of_text|>Cirja # 7 aprobó 2 TD Scramble y también logró 2 TD uno para Florin Oltean # 85 y otro para Dan Crasnic # 93.

paws-x-933. <|begin_of_text|>El río Crișul Mic es un afluente del río Doba en Rumania.

paws-x-766. <|begin_of_text|>La rivière Valea Turcului est un affluent de la rivière Azuga en Roumanie.

paws-x-780. <|begin_of_text|>La rivière Vidăcut ou rivière Hidecut est un affluent de la rivière Eliseni, en Roumanie.

paws-x-868. <|begin_of_text|>Olt River 또는 Pârâul Sec는 루마니아의 Seaca 강 지류입니다.