## **Feature 84378**

Language: Hindi

Model: meta-llama/Llama-3.2-1B

Layer: model.layers.10.mlp

SAE Model: EleutherAl/sae-Llama-

3.2-1B-131k

Selected Token Probability: 0.296

Entropy: 0.427

### **Activation Range**

3.377-3.79 3.79-4.203

# Interpretation

"The text contains references to the year 1947, the country Pakistan, and related political or administrative terms, often in the context of independence or governance, and these patterns appear across multiple languages and scripts."

Score Type	Accuracy	Precision	Recall	F1 score	TPR	TNR	FPR	FNR
detection	0.52	1.0	0.04	0.077	0.04	1.0	0.0	0.96
fuzz	0.51	1.0	0.02	0.039	0.02	1.0	0.0	0.98

### Hindi

#examples: [('xnli', 1000), ('flores', 997)]

xnli-886. <|begin\_of\_text|>म<mark>ुझे लगत</mark>ा **है कि यह क**ाम <mark>कर रह</mark>ा है , ' ' उसने एक महीने से पहले एक महीने के <mark>बारे में कहा थ</mark>ा , और उसके पित ा का एक अस ् पष ् ट प ् रभ<mark>ाव थ</mark>ा कि उसकी बेटेी के चेहरे पर कम या कम तीन zits कम थे . पिता ने सोचा कि उसकी बेटेी आमतौर पर बह ुत zits थी .

xnli-981. <|begin\_of\_text|><mark>हर जगह म**ौ**न , और ध**ीर**े विडोज �� . एक <mark>तूफ</mark>ान <mark>आ रह**ा थ**ा तो खिड ��</mark> की ध**ीर**े <mark>थ</mark>ी ।</mark>

xnli-352. <|begin\_of\_text|>वह एक चेतावनी <mark>लिख रहा था , कोई नुस ् खा नहीं . पुलिस अधिकारी एक चेतावनी टिकट <mark>लिख रहा था , ड</mark>ॉक ् ट र से कोई नुस ् खा नहीं .</mark>

xnli-844. <|begin\_of\_text|>जैसवा कि मैं <mark>लिखतवा ह</mark>ूँ , घर <mark>फटूल</mark>ों और <mark>फलवों से जीवित है</mark> , और <mark>यह फिर से बर ्फ शुर् कर रहवा है</mark> । <mark>यह</mark> गर ् म और गर ् म है .

## **Text Examples for Each Interval**

### interval 1

Range: 3.79-4.203 #examples: 1

paws-x-915. <|begin\_of\_text|>Aye<mark>za Khan (née</mark> le 15 janvier 1991 sous le nom d'Aiza Khan), également connue sous le nom de Kinza Khan, est une actrice et mannequin télévisée pakistanaise.

### interval 2

Range: 3.377-3.79 #examples: 0

#### interval 3

Range: 2.964-3.377 #examples: 8

paws-x-915. <|begin\_of\_text|>Aye<mark>za Khan</mark> (\* 15. Januar 1991 als Aiza Khan), auch als Kinza Khan bekannt, ist eine pakistanische Fernsehschauspielerin und - modell.

flores-246. <|begin\_of\_text|>Since Pakistani independence from British rule in 1947, the Pakistani President has appointed "Political Agents" to govern FATA, who exercise near-complete autonomous control over the areas.

flores-246. <|begin\_of\_text|>1947年にイギリスの支配からパ<mark>キスタンが独立して以来、パキスタン大統領はFATA</mark>を統治するために「政治エージェント」を任命しており、FATAはほぼ完全な自治権を行使しています。

paws-x-915. <|begin\_of\_text|>Ayeza Khan (nacida el 15 de enero de 1991 como Aiza Khan), también conocida como Kinza Khan, es una actriz y modelo de televisión pakistaní.