Feature 97688

Language: Turkish

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

Layer: model.layers.10.mlp

SAE Model: EleutherAl/sae-Llama-

3.2-1B-131k

Selected Token Probability: 0.639

Entropy: 2.096

Activation Range

0.065-0.586 0.586-1.107 1.107-1.627 1.627-2.148 2.148-2.669 2.669-3.19 3.19-3.711

3.711-4.231 4.231-4.752 4.752-5.273

Interpretation

"The highlighted tokens are verb endings or suffixes in Romance languages, often marking tense, person, or number, and are typically found at the end of verbs in various conjugated forms."

| Score Type | Accuracy | Precision | Recall | F1 score | TPR | TNR | FPR | FNR |
|------------|----------|-----------|--------|----------|------|------|------|------|
| detection | 0.55 | 0.619 | 0.26 | 0.366 | 0.26 | 0.84 | 0.16 | 0.74 |
| fuzz | 0.69 | 0.788 | 0.52 | 0.627 | 0.52 | 0.86 | 0.14 | 0.48 |

Turkish

#examples: [('xnli', 978), ('flores', 997)]

xnli-517. <|begin_of_text|> Yatak odasındaki masada kilitli , bayan ınglethorp 'un bir vasiyet bul<mark>muşlar</mark> , ev<mark>lenmeden</mark> önce , servetini Alfred Inglethorp 'a bırak<mark>mış</mark> . Bayan İngleth<mark>orp 'un</mark> will' i bul<mark>dular</mark> , bu da tüm serv<mark>et</mark>ini Alfred İnglethorp 'a bır<mark>aktı</mark> .

xnli-1. <|begin_of_text|>Sezon boyunca bilirsin ve san<mark>ır</mark>ım senin sevi<mark>yes</mark>inde onları bir sonraki seviyeye kaybed<mark>ersin</mark> . Eğer ana takımı geri almaya karar verir sen yiğitlere , üçlü bir adamı hatırl<mark>amak</mark> için aramaya karar <mark>verir</mark>. yerini değiştir ve tek bir adam onun yerine geç<mark>ecek</mark> Eğer insanlar hani , aşağıdaki seviyeye .

xnli-41. <|begin_of_text|>Güven Fonu 2016 ' de nakit açıkları yayınl<mark>am</mark>aya baş<mark>ladığ</mark>ında , hükü<mark>met</mark> , Sosyal Güvenlik ' in nakit açı<mark>ğ</mark>ını , sosyal güvenlik fazl aları , halktan ödünç alarak , diğer vergileri yüks<mark>elter</mark>ek , Sosyal Güvenlik ' in nakit açı<mark>ğ</mark>ını finanse etmek için para ol<mark>malı</mark>. veya diğer hükümet har<mark>cam</mark>alarını azalt<mark>mak</mark> . Halk Genel olarak hükümetin sosyal güvenliği finanse etmek için diğer alanlarda harcamalarını azaltmayı tercih eder .

xnli-691. <|begin_of_text|>Sadece onları kurtar<mark>abiliriz</mark> diye um<mark>uyorum</mark> jon . Jon On<mark>ları</mark> öfkeli mafyadan kurt<mark>armak istedi</mark> .

Text Examples for Each Interval

interval 1

Range: 4.752-5.273 #examples: 5

flores-411. <|begin_of_text|>Le loro balestre mortali scagli<mark>av</mark>ano frecce capaci di penetrare nell'armatura dei soldati nemici. Attorno al 1000 a.C. gli Assiri introdus<mark>ser</mark>o la prima cavalleria.

flores-118. <|begin_of_text|>L'editore di videogiochi Konami ha detto oggi a un giornale giapponese che non pubb<mark>licher</mark>à il gioco Sei Giorni a Fallujah.

flores-615. <|begin_of_text|>Nuove evidenze, tuttavia, fanno pensare che i Moriori fosser o un gruppo di Maori della terraferma che dalla Nuova Zelanda migrò sulle isole Chatham, sviluppando una propria peculiare cultura pacifica.

flores-441. <|begin_of_text|> Nel XVIII secolo la Cambogia fu invasa più volte dai thailandesi, che nel 1772 distrussero la capitale Phnom Phen.

flores-800. <|begin_of_text|>Se le persone sottovalutano la potenziale pericolosità degli alci, pot<mark>reb</mark>bero avvic<mark>inars</mark>i troppo a loro e corr<mark>ere</mark> dei rischi.

interval 2

Range: 4.231-4.752 #examples: 11

flores-26. <|begin_of_text|>il secondo gol della serata è stato il suo 60° gol della stagione, e questo fa di lui il primo giocatore a realizzare 60 gol (o più) in una stagione dal 1995-96, stagione in cui Jaromir Jagr e Mario Lemieux raggi<mark>unser</mark>o ent<mark>ram</mark>bi questo trag<mark>uard</mark>o.

flores-107. <|begin_of_text|>Quando la capsula raggi<mark>unger</mark>à la Terra ed entr<mark>er</mark>à nell'atmosfera, verso le 05:00 (orario della costa est), si pre<mark>ved</mark>e che gli abit anti della California settentrionale, dell'Oregon, del Nevada e dello Utah ass<mark>isteranno</mark> ad un notevole spettacolo di luci.