



# Matematinis (didelių) kalbos modelių matematinis modelis ir iš jo išplaukiantys taikymai

doc. dr. Linas Petkevičius

 @linas-p  @dxlino

Informatikos institutas

VU MIF

linas.petkevicius@mif.vu.lt

2023 spalio 24

- Reikalingos sąvokos
- Uždavinio formuluotės
- Taikymai

## Duomenys: Tekstas, teksto vienetai

- Tarkime turime tekstyną  $\mathcal{D} = \{T_i\}_{i=1}^{d_N}$
- Tekstyno elementai  $T_i, i = 1, \dots, d_N$  įvairūs tekstai dažnai iš interneto - vikipedijos, enciklopedijų, moksliniai straipsniai ir pan.
- Atliekant dažniausiai pasitaikančių teksto vienetų paiešką sudaromas teksto vienetų žodynas  $\mathcal{Z} = \mathcal{Z}_{\mathcal{D}}$  t.y. baigtinis numeruojamas indeksų sarašas turintis  $d_Z$  elementų. Taigi pirma teksto transformacija  $\tilde{T}_i : T_i \rightarrow \{0, 1, \dots, d_Z - 1\}^{k_i}$ :

$$\mathcal{Z}_{\mathcal{D}} = \left\{ \begin{array}{ll} 0 : \text{"ir"}, & 1 : \text{"jis"} \\ 2 : \text{"bûti"}, & 3 : \text{"tas"} \\ \dots & \\ i : \text{"darbas"}, & i + 1 : \text{"pmogus"} \\ \dots & \\ d_Z - 6 : \text{"SEP"}, & d_Z - 5 : \text{"CLS"} \\ d_Z - 4 : \text{"PAD"}, & d_Z - 3 : \text{"UNK"} \\ d_Z - 2 : \text{"MASK"}, & d_Z - 1 : \text{"EOS"} \end{array} \right\}$$

## Duomenys: Fiktyvūs kintamieji, įterpinių matrica

Kaip įprasta statistikoje kategoriniai kintamieji, koduojami fiktyviais kintamaisiais, šiuo atveju

$$\mathbf{X}_{I,1} = \begin{pmatrix} 0 & 0 & \dots & 1 & 0 & \dots & 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & 0 & \dots & 1 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 & 0 & \dots & 0 & 0 & \dots & 1 \end{pmatrix}_{k_1 \times d_Z}$$

Praėjusio pavyzdžio atveju  $k_1 = 22, d_Z = 32000$ .

Dėl didelio žodyno dydžio  $d_Z \gg 1$ , dimensiškumas iš kart mažinamas naudojant įterpinių (angl. embeddings) matricą  $\mathbf{E}_{d_Z \times d_F}$ , sukonstruojant realius daugiamačius vektorius  $\mathbf{X}_{k_1 \times d_F} = \mathbf{X}_{k_1 \times d_Z} \mathbf{E}_{d_Z \times d_F}$ , kur  $d_F \ll d_Z$ , bei  $\mathbf{X}_{k_1 \times d_F} \in \mathbb{R}^{k_1 \times d_F}$ .

## Duomenys: Suvedimas į fiksuotą dimensiją

Teksto vienetų skaičius tekste  $k_i$  priklauso nuo teksto ilgio. Tipiškai fiksuojamas skaičius  $K$ , pvz:  $K = 512$ , nusakantis kad turėsime fiksuotą duomenų įvesties dimensiją.

- Jei tekstas trumpas, papildome - fiktyviais vienetais
- Jei tekstas ilgas, skaidome dalimis

Tarkime turime tekstą  $T_1$ , po indeksavimo, fiktyvių kintamųjų transformacijos, bei įterpinių transformacijos, bei dimensijos  $K$  fiksavimo turime:

$$\mathbf{X}_1 = \begin{pmatrix} 0 & 0 & .. & 1 & 0 & .. & 0 & 0 & .. & 0 \\ 0 & 0 & .. & 0 & 0 & .. & 1 & 0 & .. & 0 \\ .. & .. & .. & .. & .. & .. & .. & .. & .. & .. \\ 0 & 0 & .. & 0 & 0 & .. & 0 & 0 & .. & 1 \end{pmatrix}_{K \times d_Z}$$

## Tekstas: prognozavimo uždavinys

Tarkime turime tekstą  $T_1 = \text{"Atėjus šaltajam sezonui dažnas susimąstome, kaip galima natūraliomis priemonėmis sustiprinti imunitetą"}$ , arba po pavertimo į teksto vienetus  $\tilde{T}_0 = \{ 1003, 3327, 934, 592, 179, 2709, 14754, 678, 2651, 827, 389, 231, 207, 330, 180, 407, 1050, 4175, 18614, 2987, 10020, 24149 \}$  turinčiu  $k_1 = 22$  teksto vienetų<sup>1</sup>. Tai savo ruožtu leidžia, formuluoti pradinį uždavinį:

---

<sup>1</sup>Modelis <https://huggingface.co/joelito/legal-lithuanian-roberta-base>

## Tekstas: prognozavimo uždavinys

Tarkime turime tekstą  $T_1 = \text{"Atėjus šaltajam sezonui dažnas susimąstome, kaip galima natūraliomis priemonėmis sustiprinti imunitetą"}$ , arba po pavertimo į teksto vienetus  $\tilde{T}_0 = \{ 1003, 3327, 934, 592, 179, 2709, 14754, 678, 2651, 827, 389, 231, 207, 330, 180, 407, 1050, 4175, 18614, 2987, 10020, 24149 \}$  turinčiu  $k_1 = 22$  teksto vienetų<sup>1</sup>. Tai savo ruožtu leidžia, formuluoti pradinį uždavinį:

$$\{1003, 3327, 934, 592\} \Rightarrow 179$$

$$\{3327, 934, 592, 179\} \Rightarrow 2709$$

$$\{934, 592, 179, 2709\} \Rightarrow 14754$$

..

$$\{4175, 18614, 2987, 10020\} \Rightarrow 24149$$

---

<sup>1</sup>Modelis <https://huggingface.co/joelito/legal-lithuanian-roberta-base>

## Tekstas: prognozavimo uždavinys

Tarkime turime tekstą  $T_1 = \text{"Atėjus šaltajam sezonui dažnas susimąstome, kaip galima natūraliomis priemonėmis sustiprinti imunitetą"}$ , arba po pavertimo į teksto vienetus  $\tilde{T}_0 = \{ 1003, 3327, 934, 592, 179, 2709, 14754, 678, 2651, 827, 389, 231, 207, 330, 180, 407, 1050, 4175, 18614, 2987, 10020, 24149 \}$  turinčiu  $k_1 = 22$  teksto vienetų<sup>1</sup>. Tai savo ruožtu leidžia, formuluoti pradinį uždavinį:

$$\{1003, 3327, 934, 592\} \Rightarrow 179$$

$$\{3327, 934, 592, 179\} \Rightarrow 2709$$

$$\{934, 592, 179, 2709\} \Rightarrow 14754$$

..

$$\{4175, 18614, 2987, 10020\} \Rightarrow 24149$$

Suprantama galvodami apie statistinius modelius, formuluojame sąlyginį statistinį modelį  $f(\mathbf{X}|\theta), \theta \in \Theta$

<sup>1</sup>Modelis <https://huggingface.co/joelito/legal-lithuanian-roberta-base>



## Teksto generavimo modelis

Jei modeliuotume tekstą generatyviniu modeliu:

$$\begin{aligned} z_{k+1} &= \text{sekantis teksto vienetas} = f(\mathbf{x}|\boldsymbol{\theta}) = f(z_1, z_2, \dots, z_k|\boldsymbol{\theta}) = f(\text{tekstas}|\boldsymbol{\theta}) \\ \mathbf{x}_1 &= \text{Girtas vairuotojas padarė avariją ir įpuolė į stotelę. Teismas skyrė baudą} \\ \Rightarrow f(\mathbf{x}_1|\boldsymbol{\theta}) &= \text{už} \\ \Rightarrow f(\mathbf{x}_2|\boldsymbol{\theta}) &= \text{avariją} \\ \Rightarrow f(\mathbf{x}_3|\boldsymbol{\theta}) &= \text{tačiau} \\ \Rightarrow f(\mathbf{x}_4|\boldsymbol{\theta}) &= \text{teisėjai} \\ \Rightarrow f(\mathbf{x}_5|\boldsymbol{\theta}) &= \text{nustatė} \end{aligned}$$

Neuroninių tinklų parametrinis modelis: Praktikoje turime sudėtingus parametrinius modelius  $f_\theta : \mathbf{X} \rightarrow Y$ .

Naudojami dirbtiniai neuroniniai tinklai, jie turi daug tarpinių transformacijų:

$$\begin{aligned} z_{k+1} &= f(\mathbf{x}_1|\theta) = f_L(f_{L-1}(\dots f_2(f_1(\mathbf{x}|\theta_1)|\theta_2)|\theta_{L-1})|\theta_L) \\ &= [\mathbf{T}_1 = f_1(\mathbf{X}|\theta_1), \mathbf{T}_2 = f_2(\mathbf{T}_1|\theta_2), \dots, \mathbf{T}_L = z_{k+1} = f_L(\mathbf{T}_{L-1}|\theta_L)] \end{aligned}$$

kur  $L$  - transformacijų skaičius.

Toliau panagrinėkime Aukšto lygio požymių vektorius  $\mathbf{T}_{L-1} \in \mathbb{R}^{d_{T_{L-1}}}$ , pvz.  
 $d_{T_{L-1}} = 786$ .

## Daugelio klasių klasifikavimo uždavinys

Laikydami, jog  $\mathbf{y}$  įgyja reikšmes  $\{0, 1, \dots, d_Z - 1\}$ , t.y. vieną iš žodyno pvz.  $d_Z = 32000$   
Tardami, jog  $(\mathbf{X}, \mathbf{y})$  tam tikra atsitiktine pora. Tada  $\mathbf{y}$  skirstinys  $\mathbf{X}$  atžvilgiu, vertinant kokia yra salygine tikimybe:

$$p(\mathbf{X}) = \mathbb{P}_{\mathbf{X}}(\mathbf{y} = j), \quad j = 0, 1, \dots, d_Z - 1$$

Arba kitaip tariant daugelio klasių klasifikavimo uždavinys.

## Daugelio klasių klasifikavimo uždavinys

Dideliose modeliuose, šiuo atveju multinominė regresija netinka, kad  $\mathbf{X}$  nėra labai informatyvus, tai dažniausiai geresnei aproksimacijai, ryšio operatorius  $f_{\theta}(\mathbf{X})$  imamas didelis neuroninis tinklas, t.y. didelis parametrinis modelis:

$$p(\mathbf{X}) = \text{softmax}(f_{\theta}(\mathbf{X}))$$

kur  $\theta \in \Theta \subset \mathbb{R}^{d_{\theta}}$ ,  $d_{\theta} \approx 10^9 - 10^{11}$ . Generative Pre-Training (GPT), 2018 [4].

[\[PDF\] Improving language understanding by generative pre-training](#)

[A Radford](#), [K Narasimhan](#), [T Salimans](#), [I Sutskever](#) - 2018 - [cs.ubc.ca](#)

Natural language understanding comprises a wide range of diverse tasks such as textual entailment, question answering, semantic similarity assessment, and document classification. Although large unlabeled text corpora are abundant, labeled data for learning these specific tasks is scarce, making it challenging for discriminatively trained models to perform adequately. We demonstrate that large gains on these tasks can be realized by generative pre-training of a language model on a diverse corpus of unlabeled text, followed ...

☆ Išsaugoti   🔖 Cituoti   Cituoja 5462   Susiję straipsniai   Visos 9 versijos   🔗

# Monte Karlo simuliacijos

Turint, tikimybinį pasiskirstymą, galime parinkti pasirinkimus atliekant Monte Karlo simuliacijas.

$T_{k+1}$

draugo 0.3
galima 0.1
negalima 0.1
nusipirkti 0.1
parduoti 0.1
bet 0.1
ji 0.1
SEP 0.1

# Monte Karlo simuliacijos

$T_{k+1}$

$T_{k+2}$

draugo 0.3	draugo 0.1
galima 0.1	galima 0.1
negalima 0.1	negalima 0.3
nusipirkti 0.1	nusipirkti 0.1
parduoti 0.1	parduoti 0.1
bet 0.1	bet 0.1
jj 0.1	jj 0.1
SEP 0.1	SEP 0.1

# Monte Karlo simuliacijos

Turint, tikimybinį pasiskirstymą, galime parinkti pasirinkimus atliekant Monte Karlo simuliacijas.

$T_{k+1}$	$T_{k+2}$					$T_{k+7}$	$T_{k+8}$
draugo 0.3	draugo 0.1	draugo 0.1	draugo 0.1	draugo 0.1	draugo 0.1	draugo 0.1	draugo 0.1
galima 0.1	galima 0.1	galima 0.1	galima 0.1	galima 0.1	galima 0.3	galima 0.1	galima 0.1
negalima 0.1	negalima 0.3	negalima 0.1	negalima 0.1	negalima 0.1	negalima 0.1	negalima 0.1	negalima 0.1
nusipirkti 0.1	nusipirkti 0.1	nusipirkti 0.3	nusipirkti 0.1	nusipirkti 0.1	nusipirkti 0.1	nusipirkti 0.1	nusipirkti 0.1
parduoti 0.1	parduoti 0.1	parduoti 0.1	parduoti 0.1	parduoti 0.1	parduoti 0.1	parduoti 0.3	parduoti 0.1
bet 0.1	bet 0.1	bet 0.1	bet 0.3	bet 0.1	bet 0.1	bet 0.1	bet 0.1
jį 0.1	jį 0.1	jį 0.1	jį 0.1	jį 0.3	jį 0.1	jį 0.1	jį 0.1
SEP 0.1	SEP 0.1	SEP 0.1	SEP 0.1	SEP 0.1	SEP 0.1	SEP 0.1	SEP 0.3

# Monte Karlo simuliacijos, problemos

- Pasikartojančios trumpalaikės sekos
- Įvairovės stoka
- Priklausomybių tarp pasirinkimų stoka
- ..

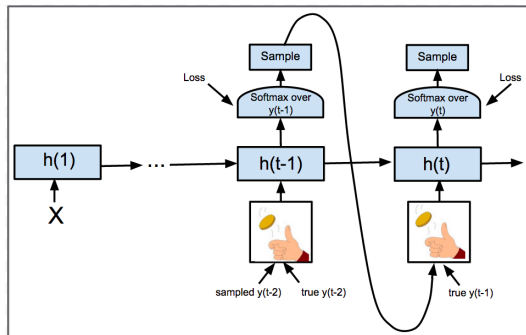


# Tobulinimai

Sekos dedamoji  $(\mathbf{X}, \mathbf{Y})$ , vertindami  $P(\mathbf{Y}|\mathbf{X})$ , jau nagrinėjame visą seką ilgio  $K$  [1]:

$$\begin{aligned}\log P(Y|X) &= \log P(y_1^K|X) \\ &= \sum_{t=1}^K \log P(y_t|y_1^{t-1}, X)\end{aligned}$$

$$\mathcal{L} = \mathcal{L}_{cls} + \lambda \mathcal{L}_{seq}$$



Metrikų parinkimai, pvz: BLEU

$$\text{BLEU} = \underbrace{\min \left( 1, \exp \left( 1 - \frac{\text{reference-length}}{\text{output-length}} \right) \right)}_{\text{brevity penalty}} \underbrace{\left( \prod_{i=1}^4 \text{precision}_i \right)^{1/4}}_{\text{n-gram overlap}}$$

$$\text{precision}_i = \frac{\sum_{\text{snt} \in \text{Cand-Corpus}} \sum_{i \in \text{snt}} \min(m_{\text{cand}}^i, m_{\text{ref}}^i)}{w_t^i = \sum_{\text{snt}' \in \text{Cand-Corpus}} \sum_{i' \in \text{snt}'} m_{\text{cand}}^{i'}}$$

- $m_{\text{cand}}^i$  - prognozuojamo teksto i-osios gramos skaičius sutampantis su lyginamu tekstu
- $m_{\text{ref}}^i$  - lyginamo teksto i-osios gramos skaičius
- $w_t^i$  - prognozuojamo teksto i-osios gramos skaičius

# Tobulinimai, žmogaus ekspertinis vertinimas

Jei turime sekų duomenis  $T^1, T^2$ ,

$$(y_1^1, y_2^1, \dots, y_K^1) \succ (y_1^2, y_2^2, \dots, y_K^2)$$

ir apibrėžiame ekspertinį reitingavimą:

$$r(y_1^1) + r(y_2^1) + \dots + r(y_K^1) \succ r(y_1^2) + r(y_2^2) + \dots + r(y_K^2)$$

2017 m. pasiūlytas [2] žmogaus ekspertinių žinių įtraukimas:

$$\hat{P}(T^1 \succ T^2) = \frac{\exp \sum_t r(y_t^1)}{\exp \sum_t r(y_t^1) + \exp \sum_t r(y_t^2)}. \quad (1)$$

Naudojant nuostolių funkciją:

$$\text{loss} = - \sum_{T^1, T^2, \mu \in \mathcal{D}} \mu_1 \log \hat{P}(T^1 \succ T^2) + \mu_2 \log \hat{P}(T^2 \succ T^1).$$

čia  $\mu_1, \mu_2$  - tekstų ekspertinis vertinimas.

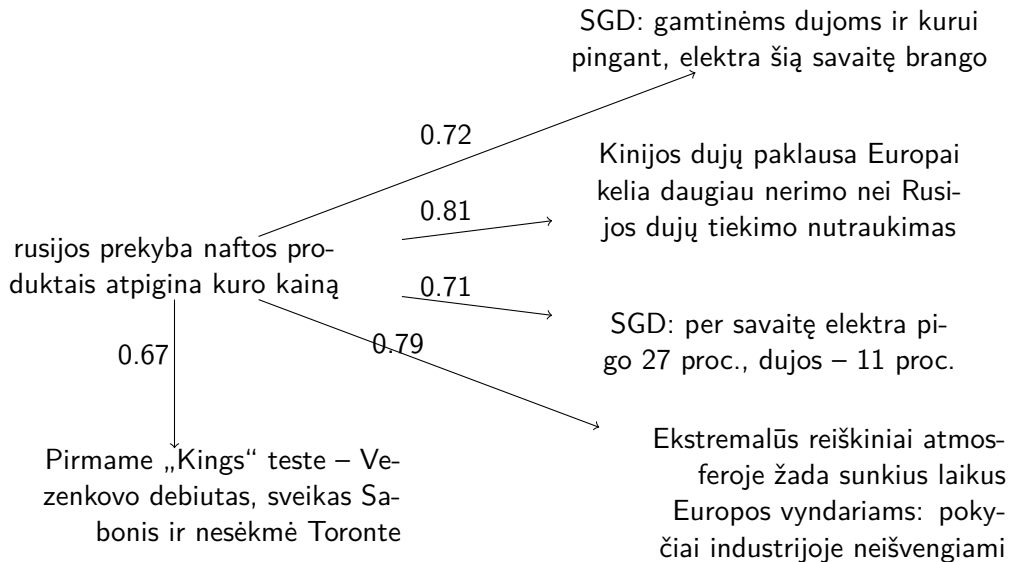
- Tolimesnio teksto prognozavimo modelis duotam pradiniam teksto fragmentui, kai sprendžiamas daugelio klasių klasifikavimo uždavinys

$$\mathcal{L} = \mathcal{L}_{cls} + \lambda \mathcal{L}_{seq} + \gamma \mathcal{L}_{eks}$$

- Parametrinis modelis 20mlrd. modelio parametrų
- Parametrai vertinti iš 300 mlrd. žodžių tekstyno

- Teksto panašumas
- Teksto interpoliacija
- "Anomalijų" paieška

# Teksto panašumas, kaip spręsti šį uždavinį?



## Aukšto lygio požymiai

Tegu  $\mathbf{x}$  - bet koks tekstas, o  $z_1, z_2, \dots, z_k$  - teksto vienetai duotame tekste.

$$\Rightarrow f(\mathbf{x}_1|\boldsymbol{\theta}) = f_L(\mathbf{T}_{L-1}^1|\theta_L) = \tilde{z}_{k+1}$$

$$\Rightarrow f(\mathbf{x}_2|\boldsymbol{\theta}) = f_L(\mathbf{T}_{L-1}^2|\theta_L) = \tilde{z}_{k+2}$$

$$\Rightarrow f(\mathbf{x}_3|\boldsymbol{\theta}) = f_L(\mathbf{T}_{L-1}^3|\theta_L) = \tilde{z}_{k+3}$$

.. ..

$$\Rightarrow f(\mathbf{x}_{k+m}|\boldsymbol{\theta}) = f_L(\mathbf{T}_{L-1}^m|\theta_L) = \tilde{z}_{k+m}$$

Sugeneruoja atsakymą  $\hat{\mathbf{y}}$  arba  $\tilde{z}_{k+1}, \tilde{z}_{k+2}, \dots, \tilde{z}_{k+m}$ , bet jei imame vidurkį gauname sakinio vidutinį aukšto lygio požymių vidurkį:

$$\tilde{\mathbf{T}}_{L-1}^{\hat{\mathbf{y}}} = \frac{1}{m} \sum_{i=1}^m \mathbf{T}_{L-1}^i$$

arba šiuo atveju - teksto reprezentaciją vektoriniu pavidalu.

# Vektorių palyginimas

Turėdami fiksuotos dimensijos vektorius, pasirinktam tekstui  $\mathbf{y}$  ir aibei kitų tekstų  $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n$ , galime juos palyginti pasirinkę atstumo ar panašumo matą. Teksto uždaviniuose vietoje įprastų atstumo matų  $L_1, L_2$ , taikomas kosinusinis panašumo skaičiavimas tarp vektorių.

$$s = \cos(\theta) = \frac{\mathbf{x} \cdot \mathbf{y}}{\|\mathbf{x}\| \|\mathbf{y}\|} = \frac{\sum_{i=1}^n x_i y_i}{\sqrt{\sum_{i=1}^n x_i^2} \sqrt{\sum_{i=1}^n y_i^2}}$$

Gavę panašumo matus  $s_1, s_2, \dots, s_n$ , galime rasti  $j = \operatorname{argmax}(s_1, \dots, s_n)$ , arba j-tąjį  $\mathbf{x}_j$  tekstą labiausiai panašų į duotą tekstą  $\mathbf{y}$ .



## Kur čia ta grėsmė? Piktybiniai naratyvai

Naratyvas:  $y$  = "rusijos prekyba naftos produktais atpigina kuro kainą".

Tekstų aibė:

- $s = 0.72 = s(y, x_1)$ ,  $x_1$  = "SGD: Gamtinėms dujoms ir kurui pingant, elektra šią savaitę brango"
- $s = 0.81 = s(y, x_2)$ ,  $x_2$  = "Kinijos dujų paklausa Europai kelia daugiau nerimo nei Rusijos dujų tiekimo nutraukimas"
- $s = 0.71 = s(y, x_3)$ ,  $x_3$  = "SDT: per savaitę elektra pigo 27 proc., dujos – 11 proc."
- $s = 0.79 = s(y, x_4)$ ,  $x_4$  = "Ekstremalūs reiškiniai atmosferoje žada sunkius laikus Europos vyndariams: pokyčiai industrijoje neišvengiami"
- $s = 0.67 = s(y, x_5)$ ,  $x_5$  = "Pirmame „Kings“ teste – Vezenkovo debiutas, sveikas Sabonis ir nesėkmė Toronte"

Viskas automatizuotai, panaudojus vektorių palyginimą. Jei kam įdomu - Python užrašų knygutė pabandyti <sup>3</sup>.

<sup>3</sup><https://colab.research.google.com/drive/1Fv1t65xklwkz51P-i9uRPgWohz9iegej?usp=sharing>

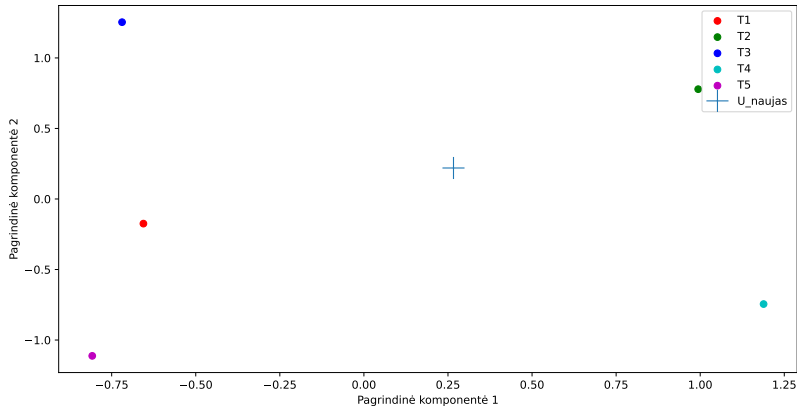
Tarkime turime aibę uždavinių  $\boldsymbol{x}_1, \boldsymbol{x}_2, \dots, \boldsymbol{x}_n$ .

Ir turime kitą tarpinį uždavinį:  $\boldsymbol{x}_{n+1}$ .

Panagrinėkime, kaip atrodo  $\boldsymbol{T}_1, \boldsymbol{T}_2, \dots, \boldsymbol{T}_n$  projekcijos ir  $\boldsymbol{T}_{n+1}$ .

# Daugiamatė vizualizacija PCA

Tegu naujas "nematytas" tekstas  $y$  = "Rusijos prekyba naftos produktais atpigina kuro kainą",



Naujo sugeneruoto uždavinio reprezentacija pagrindinių komponentių erdvėje

## Kur čia ta nauda?

Bendriniai tekstai, ir informacija, kaip daugelis mokomosios medžiagos pakankamai visur viešai prieiname. Nenuostabu, jog tai leidžia mokymo medžiagai būti interpoliuojamai.

Realiai turime tiesinę, egzistuojančių tekstų interpoliaciją (tik aukšto lygio požymių erdvėje):

$$\mathbf{T}_{naujas} = \alpha_1 \mathbf{T}_1 + \alpha_2 \mathbf{T}_2 + \dots + \alpha_n \mathbf{T}_n$$

## Išskirtys - Du požiūriai:

Tegu  $X_1, X_2, \dots, X_n$  n.a.d. su a.d  $X_i$  k.p.f  $F_i(x) \in \mathcal{F}_0$ . Čia

$$\mathcal{F}_0 = \{F(x, \theta), \theta \in \Theta \subset \mathbf{R}^m\},$$

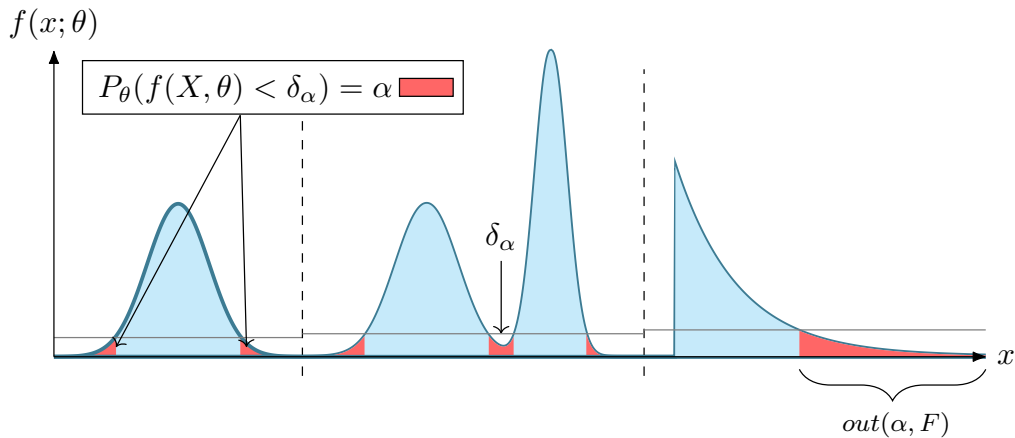
parametrinė, absoliučiai tolydi k.p.f šeima, su tankio f-ja  $f(X, \theta)$ .

- $\alpha$ -išskirčių regionas

$$out(\alpha, F) = \{x \in (F(x, \theta)) : f(x, \theta) < \delta_\alpha\};$$

čia  $\delta_\alpha$  yra  $\alpha$  f-ja tokia, kad  $P_\theta(f(X, \theta) < \delta_\alpha) = \alpha$ .

## $\alpha$ -išskirčių regionas



- teršalai (angl. contaminants):

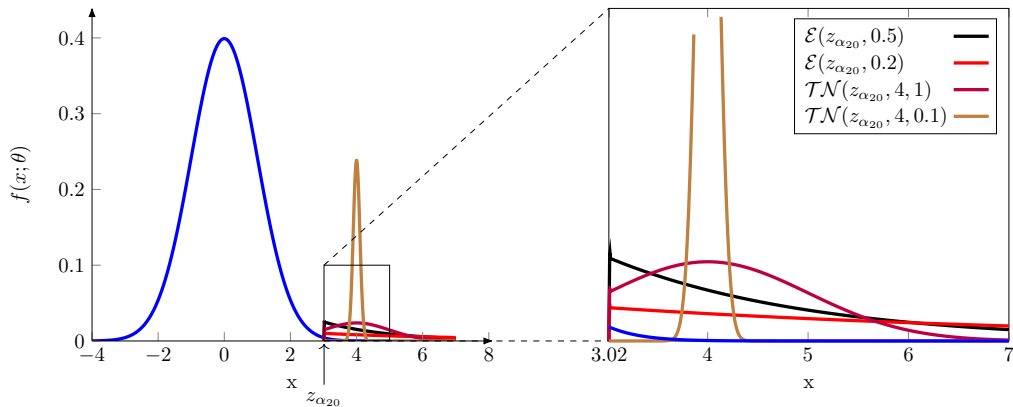
$$X_j \in F_j(x) \neq F_{isk}(x)$$

## Išskirtys-teršalai

Iš to seka, jei turime išskirtis sukoncentruotas  $\alpha$ -regione:

$$X_j \in \{F_{isk}(x) : (F_{isk}) \subset out(\alpha, F)\}$$

galime lyginti abu požiūrius.





Nagrinėkime Delfi naujienas. Kaip patikrinti, kad tekstas

- "„Hamas“ ir Izraelio karas. Izraelis patvirtino: „Hamas“ paleido dar du įkaitus"

yra panašūs į Delfi lietuviškas naujienas, o tekstai

- "some nsfw text"
- "AAAAAAAAAAAAAAAAAAAAAAAAAAAA"
- "17-year-old takes just 23 seconds to make Barcelona history"

nėra panašūs į Delfi lietuviškas naujienas.

Panagrinėkime užrašų knygutę <sup>4</sup>

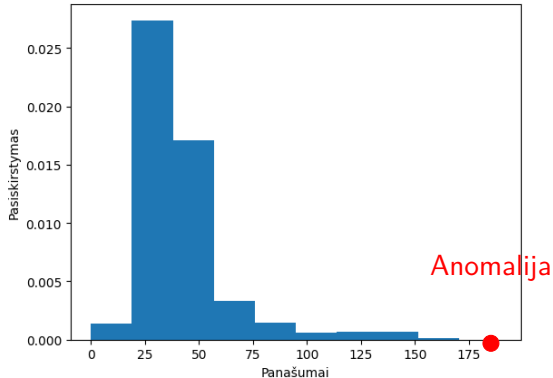
---

<sup>4</sup><https://colab.research.google.com/drive/1LencotwETILrFZDipMeQp5iMIIH3vxRDscrollTo=PXheKou0zoe>

# Išskirtys-teršalai

- Raskime nagrinėjamos aibės tekstų tarpusavio panašumus tarp  $T_i$  ir  $T_j, i, j = 1, \dots, k$ .
- Įsivertinkime empirinį pasiskirstymą  $\hat{F}_S(s)$
- Ieškome "anomalijų"

$$\begin{pmatrix} s_{1,2} & s_{1,3} & \dots & s_{1,k} \\ & s_{2,3} & \dots & s_{2,k} \\ & & \dots & \dots \\ & & & s_{k-1,k} \end{pmatrix}$$

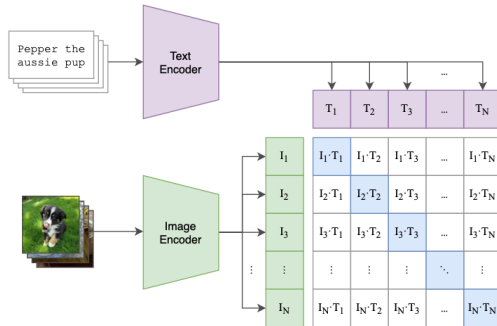


# Teksto/vaizdų modelių apjungimas

2021 m. Learning transferable visual models from natural language supervision [3].

Tegu  $f_{vaizdo} : \mathbb{R}^{d_w \times d_h \times d_c} \rightarrow \mathbb{R}^{d_e}$ ,  $f_{teksto} : \mathbb{R}^{K \times d_z} \rightarrow \mathbb{R}^{d_e}$





$$\begin{aligned}f_{vaizdo}(\mathbf{I}|\boldsymbol{\theta}_{vaizdo}) &= \mathbf{T}_{vaizdo} \\f_{teksto}(\mathbf{I}|\boldsymbol{\theta}_{teksto}) &= \mathbf{T}_{teksto} \\ \hat{\mathbf{Y}} &= \mathbf{T}_{vaizdo}^T \mathbf{T}_{teksto}\end{aligned}$$



Klausimai & Diskusija

doc. dr. Linas Petkevičius  
linas.petkevicius@mif.vu.lt  
☎ @linas-p 📷 @dxlino



-  Samy Bengio, Oriol Vinyals, Navdeep Jaitly, and Noam Shazeer.  
Scheduled sampling for sequence prediction with recurrent neural networks.  
*Advances in neural information processing systems*, 28, 2015.
-  Paul F Christiano, Jan Leike, Tom Brown, Miljan Martic, Shane Legg, and Dario Amodei.  
Deep reinforcement learning from human preferences.  
*Advances in neural information processing systems*, 30, 2017.
-  Alec Radford, Jong Wook Kim, Chris Hallacy, Aditya Ramesh, Gabriel Goh, Sandhini Agarwal, Girish Sastry, Amanda Askell, Pamela Mishkin, Jack Clark, et al.  
Learning transferable visual models from natural language supervision.  
In *International conference on machine learning*, pages 8748–8763. PMLR, 2021.
-  Alec Radford, Karthik Narasimhan, Tim Salimans, Ilya Sutskever, et al.  
Improving language understanding by generative pre-training.  
2018.