

Տրանսֆերային ուսուցման որոշակի մեթոդի ընդհանրացման սխալանքի գնահատման մասին

Գևորգ Մինասյան

Երևանի Պետական Համալսարան

31 Մայիսի 2019

ԱԿՆԱՐԿ

1 Տրանսֆերային ուսուցում

- Ինչ է տրասֆերային ուսուցումը
- Խորր տրասֆերային ուսուցում

2 Սահմանումներ

- Վերահսկվող առաջադրանքներ
- Վերահսկվող ներկայացումների գնահատման չափը

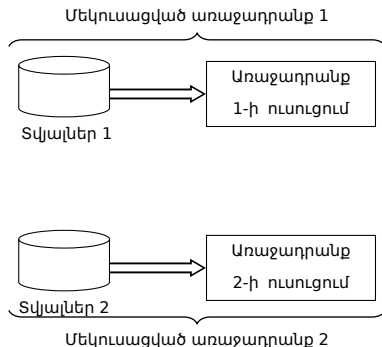
Տրանսերային ուսուցումն և սովորական
մեքենայական ուսուցման միջև
տարբերությունը

Տրանսֆերային ուսուցումն և սովորական մեքենայական ուսուցման միջև տարբերությունը

Սովորական մեքենայական ուսուցում

Տրանսֆերային ուսուցումն և սովորական մեքենայական ուսուցման միջև տարբերությունը

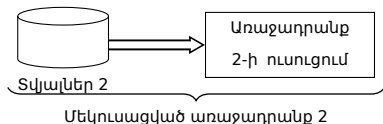
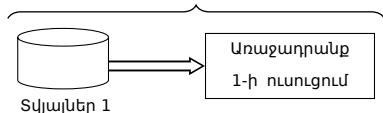
Սովորական մեքենայական ուսուցում



Տրանսֆերային ուսուցումն և սովորական մեքենայական ուսուցման միջև տարբերությունը

Սովորական մեքենայական ուսուցում

Մեկուսացված առաջադրանք 1

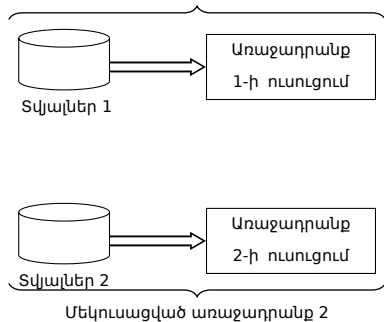


Տրանսֆերային ուսուցում

Տրանսերային ուսուցումն և սովորական
մեքենայական ուսուցման միջև
տարբերությունը

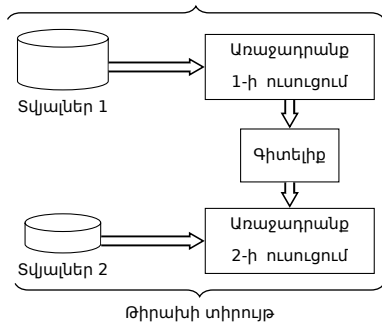
Սովորական մեքենյական ուսուցում

Մեկուսացված առաջադրանք 1

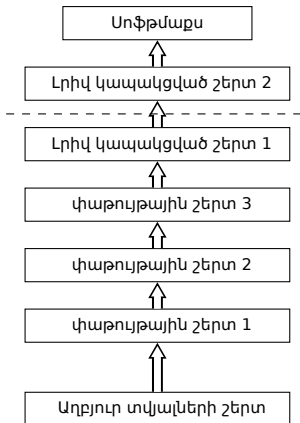


Տրանսֆերային ուսուցում

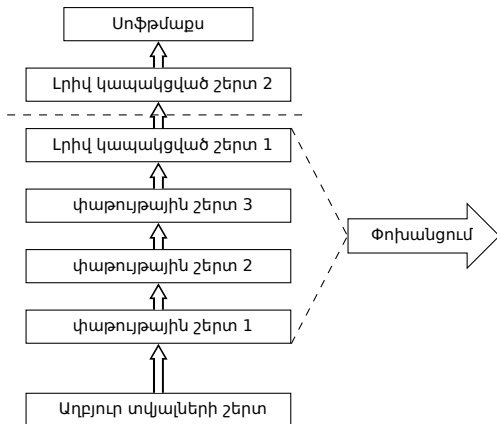
Աղբյուրի տիրույթ



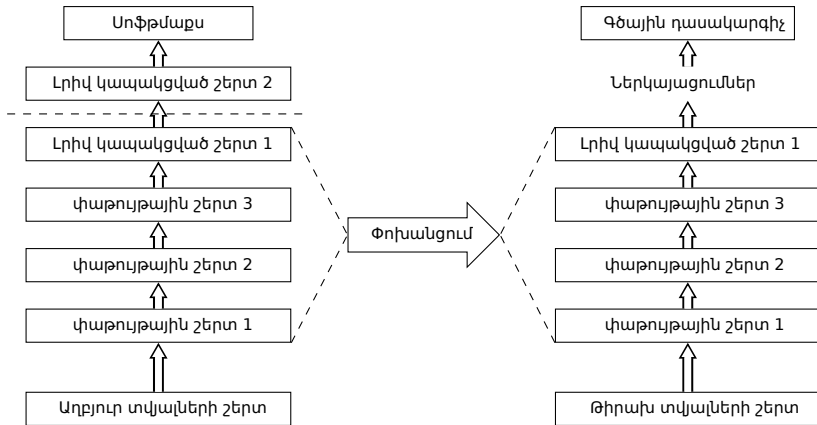
Տրանսֆերային ուսուցումը վարժեցված նեյրոնային ցանցի միջոցով



Տրանսֆերային ուսուցումը վարժեցված նեյրոնային ցանցի միջոցով



Տրանսֆերային ուսուցումը վարժեցված նեյրոնային ցանցի միջոցով



Նշանակումներ

- \mathcal{X} բոլոր հնարավոր տվյալների օրինակների բազմություն

Նշանակումներ

- \mathcal{X} բոլոր հնարավոր տվյալների օրինակների բազմություն
- \mathcal{C} բոլոր հնարավոր դասերի բազմություն

Նշանակումներ

- \mathcal{X} բոլոր հնարավոր տվյալների օրինակների բազմություն
- \mathcal{C} բոլոր հնարավոր դասերի բազմություն
- \mathcal{F} տվյալների ներկայացումների ֆունկցիաների դաս

$$f: \mathcal{X} \rightarrow \mathbb{R}^d, \forall f \in \mathcal{F}$$

Նշանակումներ

- \mathcal{X} բոլոր հնարավոր տվյալների օրինակների բազմություն
- \mathcal{C} բոլոր հնարավոր դասերի բազմություն
- \mathcal{F} տվյալների ներկայացումների ֆունկցիաների դաս

$$f: \mathcal{X} \rightarrow \mathbb{R}^d, \forall f \in \mathcal{F}$$

- $\exists R > 0 \forall f \in \mathcal{F} \|f(x)\| \leq R \forall x \in \mathcal{X}$

Նշանակումներ

Նշանակումներ

- \mathcal{T} վերահսկվող առաջադրանքը բաղկացած է $k + 1$ հատ

$$\{c_1, c_2, \dots, c_{k+1}\} \subseteq \mathcal{C}$$

միմյանցից տարբեր դասերից

Նշանակումներ

- \mathcal{T} վերահսկվող առաջադրանքը բաղկացած է $k + 1$ հատ

$$\{c_1, c_2, \dots, c_{k+1}\} \subseteq \mathcal{C}$$

միմյանցից տարբեր դասերից

- $\mathcal{P}(\mathcal{T})$ վերահսկվող առաջադրանքների դիտարկվելու հավանականային բաշխումը

Նշանակումներ

- \mathcal{T} վերահսկվող առաջադրանքը բաղկացած է $k + 1$ հատ

$$\{c_1, c_2, \dots, c_{k+1}\} \subseteq \mathcal{C}$$

միմյանցից տարբեր դասերից

- $\mathcal{P}(\mathcal{T})$ վերասիկվող առաջադրանքների դիտարկվելու հավանականային բաշխումը
- $\mathcal{P}(\mathcal{T} \mid |\mathcal{T}| = k + 1)$ $k + 1$ դասերից բաղկացած վերասիկվող առաջադրանքների դիտարկվելու հավանականային բաշխումը

Նշանակումներ

Նշանակումներ

- $\mathcal{D}_c(x)$ c դասին համապատասխան հավանականային բաշխումն է, ցույց է տալիս, թե x օրինակը ինչքանով է համապատասխան c դասին

Նշանակումներ

- $\mathcal{D}_c(x)$ c դասին համապատասխան հավանականային բաշխումն է, ցույց է տալիս, թե x օրինակը ինչքանով է համապատասխան c դասին
- $\mathcal{D}_{\mathcal{T}}(x, c) = \mathcal{D}_{\mathcal{T}}(c)\mathcal{D}_c(x)$ \mathcal{T} վերահսկվող առաջադրանգի պիտակավորված տվյալների հավանականային բաշխումն է

Նշանակումներ

- $\mathcal{D}_c(x)$ c դասին համապատասխան հավանականային բաշխումն է, ցույց է տալիս, թե x օրինակը ինչքանով է համապատասխան c դասին
- $\mathcal{D}_{\mathcal{T}}(x, c) = \mathcal{D}_{\mathcal{T}}(c)\mathcal{D}_c(x)$ \mathcal{T} վերահսկվող առաջադրանգի պիտակավորված տվյալների հավանականային բաշխումն է
- $S = \{(x_1, y_1), \dots, (x_M, y_M) \mid x_i \in \mathcal{X}, y_i \in \mathcal{T}\}$ \mathcal{T} առաջադրանքի պիտակավորված օրինակների ուսուցման բազմությունն է ընտրված միմյանցից անկախ և $\mathcal{D}_{\mathcal{T}}(x, c)$ հավանականային բաշխումից

Նշանակումներ

Նշանակումներ

- \mathcal{T} առաջադրանքի բազմադաս դասակարգիչը ֆուկցիա է՝ $g: \mathcal{X} \rightarrow \mathcal{R}^{k+1}$, որի արժեքի կորդինատները ինդեքսավորված են այդ առաջադրանքի դասերով:

Նշանակումներ

- \mathcal{T} առաջադրանքի բազմադաս դասակարգիչը ֆուկցիա է՝ $g: \mathcal{X} \rightarrow \mathcal{R}^{k+1}$, որի արժեքի կորդինատները ինդեքսավորված են այդ առաջադրանքի դասերով:
- $l(\{g(x)_y - g(x)_{y'}\}_{y \neq y'})$ -ը $(x, y) \in \mathcal{X} \times \mathcal{T}$ կետում g դասակարգիչով պայմանավորված կորուստն է

Նշանակումներ

- \mathcal{T} առաջադրանքի բազմադաս դասակարգիչը ֆուկցիա է՝ $g: \mathcal{X} \rightarrow \mathcal{R}^{k+1}$, որի արժեքի կորդինատները ինդեքսավորված են այդ առաջադրանքի դասերով:
- $l(\{g(x)_y - g(x)_{y'}\}_{y \neq y'})$ -ը $(x, y) \in \mathcal{X} \times \mathcal{T}$ կետում g դասակարգիչով պայմանավորված կորուստն է
- $l(v) = \max\{0, 1 + \max_i\{-v_i\}\}$

Նշանակումներ

- \mathcal{T} առաջադրանքի բազմադաս դասակարգիչը ֆունկցիա է՝ $g: \mathcal{X} \rightarrow \mathcal{R}^{k+1}$, որի արժեքի կորդինատները ինդեքսավորված են այդ առաջադրանքի դասերով:
- $l(\{g(x)_y - g(x)_{y'}\}_{y \neq y'})$ -ը $(x, y) \in \mathcal{X} \times \mathcal{T}$ կետում g դասակարգիչով պայմանավորված կորուստն է
- $l(v) = \max\{0, 1 + \max_i\{-v_i\}\}$
- $l(v) = \log_2(1 + \sum_i e^{-v_i})$

Նշանակումներ

Նշանակումներ

- \mathcal{T} առաջադրանքի համար g դասակարգիչի կորուստը հետևյալն է՝

$$L(\mathcal{T}, g) = \mathbb{E}_{(x, c) \sim \mathcal{D}_{\mathcal{T}}} [l(\{g(x)_c - g(x)_{c'}\}_{c \neq c'})]$$

Նշանակումներ

- \mathcal{T} առաջադրանքի համար g դասակարգիչի կորուստը հետևյալն է՝

$$L(\mathcal{T}, g) = \mathbb{E}_{(x, c) \sim \mathcal{D}_{\mathcal{T}}} [l(\{g(x)_c - g(x)_{c'}\}_{c \neq c'})]$$

- $\mathcal{V} = \{W: \|W\|_{\infty} \leq Q \text{ և } Q > 0\}$

Նշանակումներ

- \mathcal{T} առաջադրանքի համար g դասակարգիչի կորուստը հետևյալն է՝

$$L(\mathcal{T}, g) = \mathbb{E}_{(x, c) \sim \mathcal{D}_{\mathcal{T}}} [l(\{g(x)_c - g(x)_{c'}\}_{c \neq c'})]$$

- $\mathcal{V} = \{W: \|W\|_{\infty} \leq Q \text{ և } Q > 0\}$
- $g(x) = Wf(x)$, որտեղ $W \in \mathcal{V}$

Նշանակումներ

- \mathcal{T} առաջադրանքի համար g դասակարգիչի կորուստը հետևյալն է՝

$$L(\mathcal{T}, g) = \mathbb{E}_{(x, c) \sim \mathcal{D}_{\mathcal{T}}} [l(\{g(x)_c - g(x)_{c'}\}_{c \neq c'})]$$

- $\mathcal{V} = \{W: \|W\|_{\infty} \leq Q \text{ և } Q > 0\}$
- $g(x) = Wf(x)$, որտեղ $W \in \mathcal{V}$
- $L(\mathcal{T}, f) = \inf_{W \in \mathcal{V}} L(\mathcal{T}, f, W)$

Նշանակումներ

Նշանակումներ

Վերահսկիչ միջին կորուստ

$k + 1$ դասերից բաղկացած առաջադրանքների վերահսկիչ միջին կորուստը f ներկայացման համար սահմանվում է որպես՝

$$L(f) = \mathbb{E}_{\mathcal{T} \sim \mathcal{P}} [L(\mathcal{T}, f) \mid |\mathcal{T}| = k + 1]$$

Նշանակումներ

Նշանակումներ

Էմպիրիկ վերահսկիչ միջին կորուստ

Դիցուք ունենք միմյանցից անկախ $\mathcal{P}(\mathcal{T} \mid |\mathcal{T}| = k + 1)$ բաշխումից ընտրված N հատ առաջադրանքներ՝ $\mathcal{T}_1, \dots, \mathcal{T}_N$: Էմպիրիկ վերահսկիչ միջին կորուստը f ներկայացման համար հետևյալն է՝

$$\hat{L}(f) = \frac{1}{N} \sum_{i=1}^N L(\mathcal{T}_i, f)$$

Verbatim

Example (Theorem Slide Code)

```
\begin{frame}  
\frametitle{Theorem}  
\begin{theorem}[Mass--energy equivalence]  
$E = mc^2$  
\end{theorem}  
\end{frame}
```

Շնորհակալություն