Neuroninis tinklas su matriciniais įėjimais

Povilas Daniušis
vadovas doc. P. Vaitkus

2007 m. gegužės 30 d.

Uždavinys ir modelis

- $T = ((X_1, y_1), (X_2, y_2), ..., (X_M, y_M))$, kur X_i - $m \cdot n$ eilės matricos (įėjimai), o y_i - skaliarai (išėjimai).

Uždavinys ir modelis

- $T = ((X_1, y_1), (X_2, y_2), ..., (X_M, y_M))$, kur X_i - $m \cdot n$ eilės matricos (įėjimai), o y_i - skaliarai (išėjimai).
- Pvz. X - Rentgeno nuotrauka, y - diagnozė

Uždavinys ir modelis

- $T = ((X_1, y_1), (X_2, y_2), ..., (X_M, y_M))$, kur X_i - $m \cdot n$ eilės matricos (įėjimai), o y_i - skaliarai (išėjimai).
- Pvz. X - Rentgeno nuotrauka, y - diagnozė
- Reikalingas modelis darbui su matricomis!

Uždavinys ir modelis

- $T = ((X_1, y_1), (X_2, y_2), ..., (X_M, y_M))$, kur X_i - $m \cdot n$ eilės matricos (įėjimai), o y_i - skaliarai (išėjimai).
- Pvz. X - Rentgeno nuotrauka, y - diagnozė
- Reikalingas modelis darbui su matricomis!
- Siūlome tokį modelį:

$$\widehat{y}(X) = \sum_{i=1}^{N} \alpha_i \sigma(\mathbf{u}_i X \mathbf{v}_i' + b_i) \qquad (1)$$

Uždavinys ir modelis

- $T = ((X_1, y_1), (X_2, y_2), ..., (X_M, y_M))$, kur X_i - $m \cdot n$ eilės matricos (įėjimai), o y_i - skaliarai (išėjimai).
- Pvz. X - Rentgeno nuotrauka, y - diagnozė
- Reikalingas modelis darbui su matricomis!
- Siūlome tokį modelį:

$$\widehat{y}(X) = \sum_{i=1}^{N} \alpha_i \sigma(\mathbf{u}_i X \mathbf{v}_i' + b_i) \tag{1}$$

Uždavinys ir modelis

- $T = ((X_1, y_1), (X_2, y_2), ..., (X_M, y_M))$, kur X_i - $m \cdot n$ eilės matricos (įėjimai), o y_i - skaliarai (išėjimai).
- Pvz. X - Rentgeno nuotrauka, y - diagnozė
- Reikalingas modelis darbui su matricomis!
- Siūlome tokį modelį:

$$\widehat{y}(X) = \sum_{i=1}^{N} \alpha_i \sigma(\mathbf{u}_i X \mathbf{v}_i^{'} + b_i) \tag{1}$$

- pvz. $\sigma(x) = 1/1 + exp(-x)$.

Uždavinys ir modelis

- $T = ((X_1, y_1), (X_2, y_2), ..., (X_M, y_M))$, kur X_i - $m \cdot n$ eilės matricos (įėjimai), o y_i - skaliarai (išėjimai).
- Pvz. X - Rentgeno nuotrauka, y - diagnozė
- Reikalingas modelis darbui su matricomis!
- Siūlome tokį modelį:

$$\widehat{y}(X) = \sum_{i=1}^{N} \alpha_i \sigma(\mathbf{u}_i X \mathbf{v}_i' + b_i) \tag{1}$$

- pvz. $\sigma(x) = 1/1 + exp(-x)$.
- $(... \ \sigma(< \mathbf{w}_i, \mathbf{x} > +b_i) \ )$

Istorija

- $\widehat{y}(x) = \sum_{i=1}^{N} \alpha_i \sigma(< \mathbf{w}_i, \mathbf{x} > + b_i)$, Rumelhart et el. 1986.

Istorija

- $\widehat{y}(x) = \sum_{i=1}^{N} \alpha_i \sigma(< \mathbf{w}_i, \mathbf{x} > + b_i)$, Rumelhart et el. 1986.
- $\widehat{y}(X) = \mathbf{u} X \mathbf{v}'$, Cai 2006.

Kaip rasti parametrus?

$$\widehat{y}(X) = \sum_{i=1}^{N} \alpha_i \sigma(\mathbf{u}_i X \mathbf{v}_i' + b_i) \tag{2}$$

$$Err(..) = \frac{1}{2M} \sum_{(X,y)\in T} (\widehat{y}(X) - y)^2 \rightarrow min \tag{3}$$

- Gradientinis metodas?

Kaip rasti parametrus?

$$\widehat{y}(X) = \sum_{i=1}^{N} \alpha_i \sigma(\mathbf{u}_i X \mathbf{v}_i^{'} + b_i) \tag{2}$$

$$Err(..) = \frac{1}{2M} \sum_{(X,y)\in T} (\widehat{y}(X) - y)^2 \rightarrow min \tag{3}$$

- Gradientinis metodas?
- Extreme learning machines?

Kaip rasti parametrus?

$$\widehat{y}(X) = \sum_{i=1}^{N} \alpha_i \sigma(\mathbf{u}_i X \mathbf{v}_i^{'} + b_i) \tag{2}$$

$$Err(..) = \frac{1}{2M} \sum_{(X,y) \in T} (\widehat{y}(X) - y)^2 \rightarrow min \tag{3}$$

- Gradientinis metodas?
- Extreme learning machines?
- Stochastiniai algoritmai?

Privalumai

- Netiesinis.

Privalumai

- Netiesinis.
- Parametrų skaičius ($< \mathbf{w}_i, \mathbf{x} >$, $\mathbf{u}X\mathbf{v}'$)

Privalumai

- Netiesinis.
- Parametrų skaičius ($< \mathbf{w}_i, \mathbf{x} >$, $\mathbf{u}X\mathbf{v}'$)
- Specialiai darbui su matricomis.

Aproksimacinės galimybės

- Su vektoriniais įėjimais uždarinys sutampa su C(K), o su matriciniais - ne (Kreines - Vostrecov, 1961) - trūkumas.

Aproksimacinės galimybės

- Su vektoriniais įėjimais uždarinys sutampa su C(K), o su matriciniais - ne (Kreines - Vostrecov, 1961) - trūkumas.
- Kada sutampa? Kai yra du paslėpti sluoksniai:

$$\sum_{i=1}^{K} d_i \sigma \left(\sum_{j=1}^{K_i} c_{i,j} \sigma(\mathbf{u}_{i,j} X \mathbf{v}'_{i,j} + b_{i,j}) + \gamma_i \right) \tag{4}$$

(Pinkus 1999, Kolmogorov 1950).

Aproksimacinės galimybės

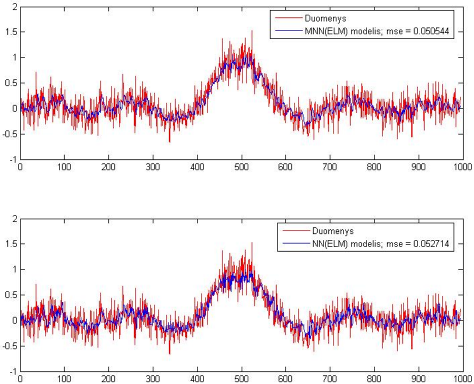
- Su vektoriniais įėjimais uždarinys sutampa su C(K), o su matriciniais - ne (Kreines - Vostrecov, 1961) - trūkumas.
- Kada sutampa? Kai yra du paslėpti sluoksniai:

$$\sum_{i=1}^{K} d_i \sigma\left(\sum_{j=1}^{K_i} c_{i,j} \sigma(\mathbf{u}_{i,j} X \mathbf{v}'_{i,j} + b_{i,j}) + \gamma_i\right) \tag{4}$$

(Pinkus 1999, Kolmogorov 1950).

- $\exists \sigma \in C^{\infty}(\mathbb{R})$, kad jų reiks $(4n+3) \times (2n+1)$ (Pinkus).

Kaip atrodo?


Duomenys
MNN(ELM) modelis; mse = 0.050544
Duomenys
NN(ELM) modelis; mse = 0.052714

Papildoma informacija

- Skaidrės darytos su LaTeX paketu **Beamer**.
- MNN paketas suprogramuotas su MATLAB. Jį atsisiųsti galima iš `http://pou.hacker.lt/MNN_package.rar`

Pabaiga

Ačiū už dėmesį