

ことばのアーキテクチャ

高村大也



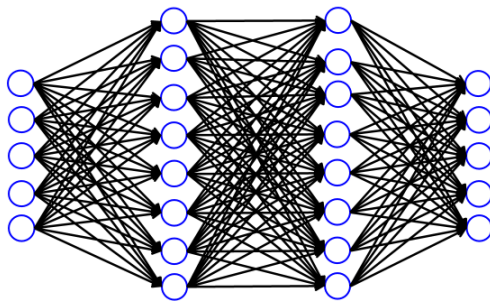
高村大也 (たかむら ひろや)

東京工業大学教授・産業技術総合研究所 人工知能研究センター 知識情報研究チーム長

1997年東京大学工学部計数工学科卒業。2000年同大学院工学系研究科計数工学専攻修了(1999年はオーストリアウィーン工科大学にて研究)。2003年奈良先端科学技術大学院大学情報科学研究科博士課程修了。博士(工学)。2003年から2010年まで東京工業大学助手のち助教(2006年にはイリノイ大学にて客員研究員)。2010年より准教授, 2017年より東京工業大学教授および産業技術総合研究所人工知能研究センター知識情報研究チーム長を兼務。

計算言語学, 自然言語処理を専門とし, 特に機械学習の応用に興味を持つ。

「言語処理のための機械学習入門」著。



$$H(\beta, x, W) = \frac{\beta}{2} \sum_{ij} w_{ij} x_i x_j + \alpha \sum_{i \in L} (x_i - a_i)^2,$$
$$\min_{w, \xi \geq 0} \frac{1}{2} \|w\|^2 + \frac{C}{N} \sum_i \xi_i, \quad \text{s.t. } \xi_i \geq \xi_j, \quad \forall i, j,$$
$$\text{s.t. } \forall i, \forall S \in \mathcal{U}, S_i: \quad \sum_i x_i \leq p,$$
$$w \cdot \phi(S_i, D_i) \geq w \cdot c_i, \quad \sum_i z_{ij} = 1; \quad \forall j,$$
$$z_{ij} = x_i; \quad \forall i,$$
$$x_i \in \{0, 1\}; \quad \forall i,$$
$$z_{ij} \in \{0, 1\}; \quad \forall i, j,$$
$$s_i \leftarrow \arg \max_{s_i \in \mathcal{U}} (f(S_i)),$$
$$\text{If } c_i + \sum_{s_i \in S} \xi_i \leq K:$$
$$S \leftarrow S \cup \{s_i\},$$
$$U \leftarrow U - \{s_i\},$$
$$\text{end}$$
$$s_i \leftarrow \arg \max_{s_i} f(s_i),$$
$$\text{If } f(S) \geq f(\{s_i\}) \text{ then } p_i \leftarrow p_i + \frac{\exp(\text{score}(p_i, s_i))}{\sum_{k=1}^n \exp(\text{score}(p_k, s_i))},$$
$$\text{otherwise: output } \{s_i\}.$$

(左下)ニューラルネットワークの模式図、(右下)言語処理で使われる数式の例

機械学習に代表される人工知能技術は、画像認識や音声認識を含む多くの分野において目覚ましい発展が見られます。人間の言語をコンピュータで扱う自然言語処理においても盛んに研究が行われており、機械翻訳などが実世界で応用されつつありますが、「言語を理解する」あるいは「言語で伝える」ことは、コンピュータにはとっては超難問です。本講演では、機械学習の基本的な考え方から始め、文章に含まれる感情を捉える技術や、文章を生成する技術など、自然言語処理の最先端の研究を紹介します。同時に、コンピュータに言語を理解させることの難しさについても考えてみたいと思います。

東京大学総合研究博物館小石川分館

〒112-0001 東京都文京区白山3-7-1
Tel. 03-5777-8600(ハローダイヤル)

開館時間: 10:00-16:30 (入館 16:00 まで)
入館料: 無料
休館日: 月・火・水曜日
(いずれも祝日の場合は開館)、年末年始、その他博物館が定める日

アクセス: 地下鉄丸ノ内線茗荷谷駅より徒歩 8 分
http://www.um.u-tokyo.ac.jp/architectonica/index_jp.html

建築博物館とは?
「アーキテクチャ」をテーマにさまざまな分野の研究者が講演を行い、関連した標本を「アーキテクトニカ・コレクション」として展示していくシリーズイベントです。

建築ミュージアム / アーキテクトニカ
KOISHIKAWA Annex.
UMUT
東京大学総合研究博物館小石川分館