



自然言語のデータマイニング

[キーワード: テキストマイニング、計量言語学、機械学習]

教授 石田基広

<研究の概要>

日本語あるいは欧米語によるテキストの構成を、数理統計モデルによって近似する方法を探索している。最近の主な研究テーマは以下の通り。

1. 文の構成要素(文字、単語、文)の確率分布
2. 文要素の時系列変化
3. 書き手の判別
4. テキストジャンルの推定
5. 確率プログラミング
6. 応用統計学

<主要研究業績>

- ・奥村晴彦・石田基広(2018)『Rで楽しむベイズ統計入門』技術評論社
- ・石田基広(2017)『Rによるテキストマイニング入門』森北出版
- ・石田基広・市川大祐・瓜生真也・湯谷啓明(2017)『Rによるスクレイピング入門』C&R研究所
- ・石田基広(2016)『R言語逆引きハンドブック』C&R出版
- ・松浦健太郎・石田基広(2016)『StanとRでベイズ統計モデリング』共立出版
- ・石田基広(2015)『新米探偵、データ分析に挑む』ソフトバンククリエイティブ
- ・石田基広(2014)『とある弁当屋の統計技師—データ分析のはじめかた』共立出版
- ・石田基広・小林雄一郎(2013)『Rで学ぶ日本語テキストマイニング』ひつじ書房

専門分野 : 自然言語のデータマイニング(テキストマイニング)

E-mail: ishida.Motohiro@tokushima-u.ac.jp

Tel: :

Fax: :

HP: :

詳細情報 : <http://pub2.db.tokushima-u.ac.jp/ERD/person/60564/profile-ja.html>