2830

Ordering Concepts Based on Common Attribute Intensity

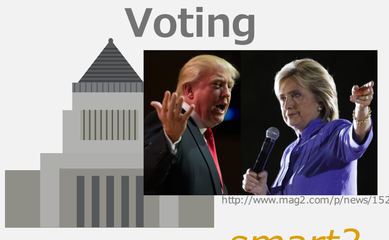
Tatsuya Iwanari♠ Naoki Yoshinaga♣ Nobuhiro Kaji♥ Toshiharu Nishina♦ Masashi Toyoda♣ Masaru Kitsuregawa♣,★

♠Univ. of Tokyo ♣IIS, Japan ♥Yahoo Japan Corporation ♦Rakuten Inc. ★NII, Japan

Introduction

Ordering concepts from various perspectives is essential to make decisions in a daily life

Voting



http://www.mag2.com/p/news/152554

reliable? smart?

Task scheduling

Vector Open Stock, https://www.vectoropenstock.com/

important? urgent?

Choosing a restaurant (others' reviews)

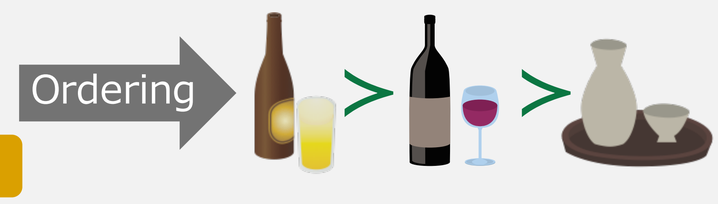
tasty? cheap?

Question: Can we derive our views on concept ordering from what we have written?

Task Setting

a set of concepts + an adjective

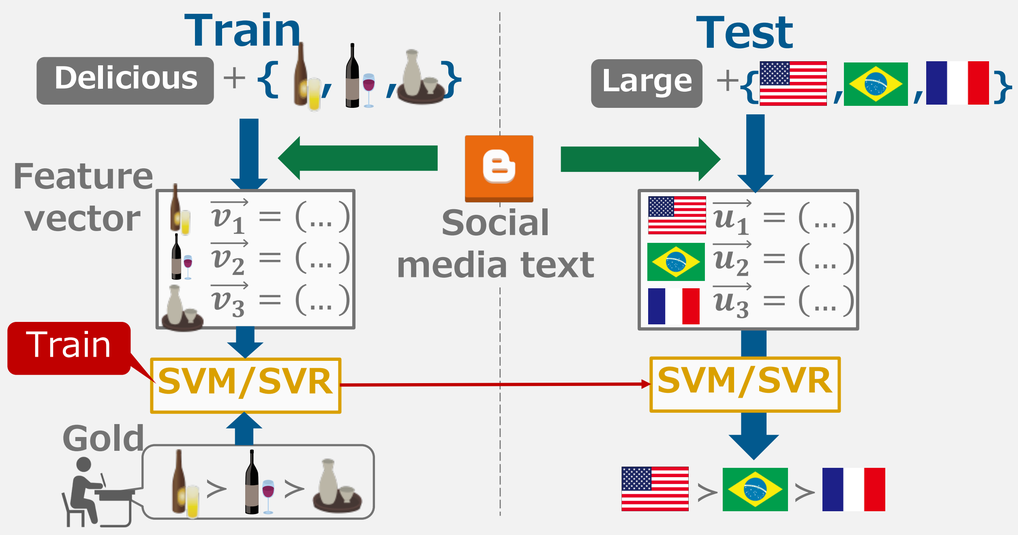
concepts **ordered by attribute intensity**



What is the gold standard? the ordering maximizing averaged Spearman's ρ against human orderings

Proposal

Extract **pieces of ordering evidence** from social media and integrate them in **supervised learning**



Evidence	Example
Co-occurrence	Look how **large** that **elephant** is!
Dependency	**Ants** are so **small** that elephants can't kill them
Simile	He is as **brave** as a **lion**
Comparative	**Elephants** are **larger** than **dogs**

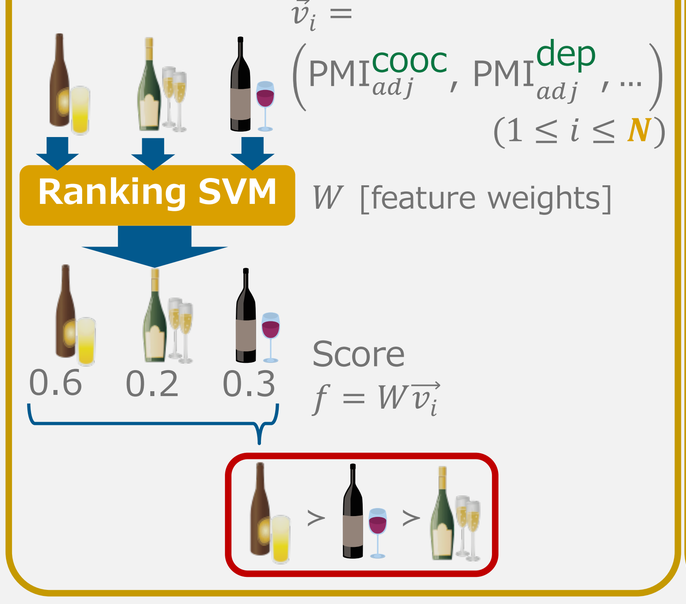
counts to feature values

e.g., $\overrightarrow{\text{elephant}}_{\text{large}} = \left(\text{PMI}^{\text{cooc}}_{\text{large}}, \text{PMI}^{\text{dep}}_{\text{large}}, \cdots\right)$

$\log \frac{p(\text{noun–adjective co–occurence})}{p(\text{noun})\,p(\text{adjective})}$

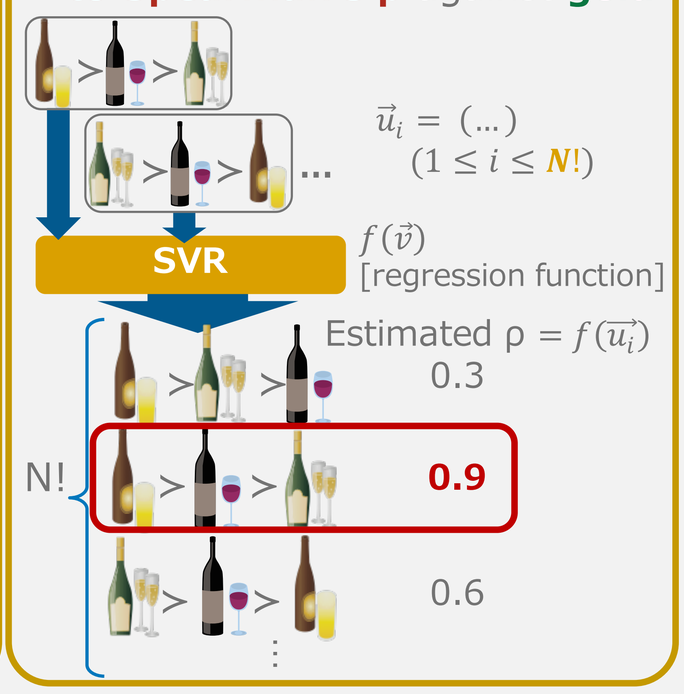
Ranking SVM [Joachims '02]

- Generate **a feature vector for each concept**
- Learn SVM to minimize **# of incorrect orderings of two items** against **gold**



SVR [Drucker et al. '97] (Support Vector Regression)

- Generate **a feature vector for each ordering**
- Learn SVR to **map an ordering to Spearman's ρ** against **gold**



Evaluation

[Data]

- Blog articles (Japanese)
  - 2005-2013
  - > 1 million users
  - ~ 2 billion sentences

[Queries]

- 35 queries (nouns & adj)
  - concept to instance nouns
  - objective to subjective adj

[Methods]

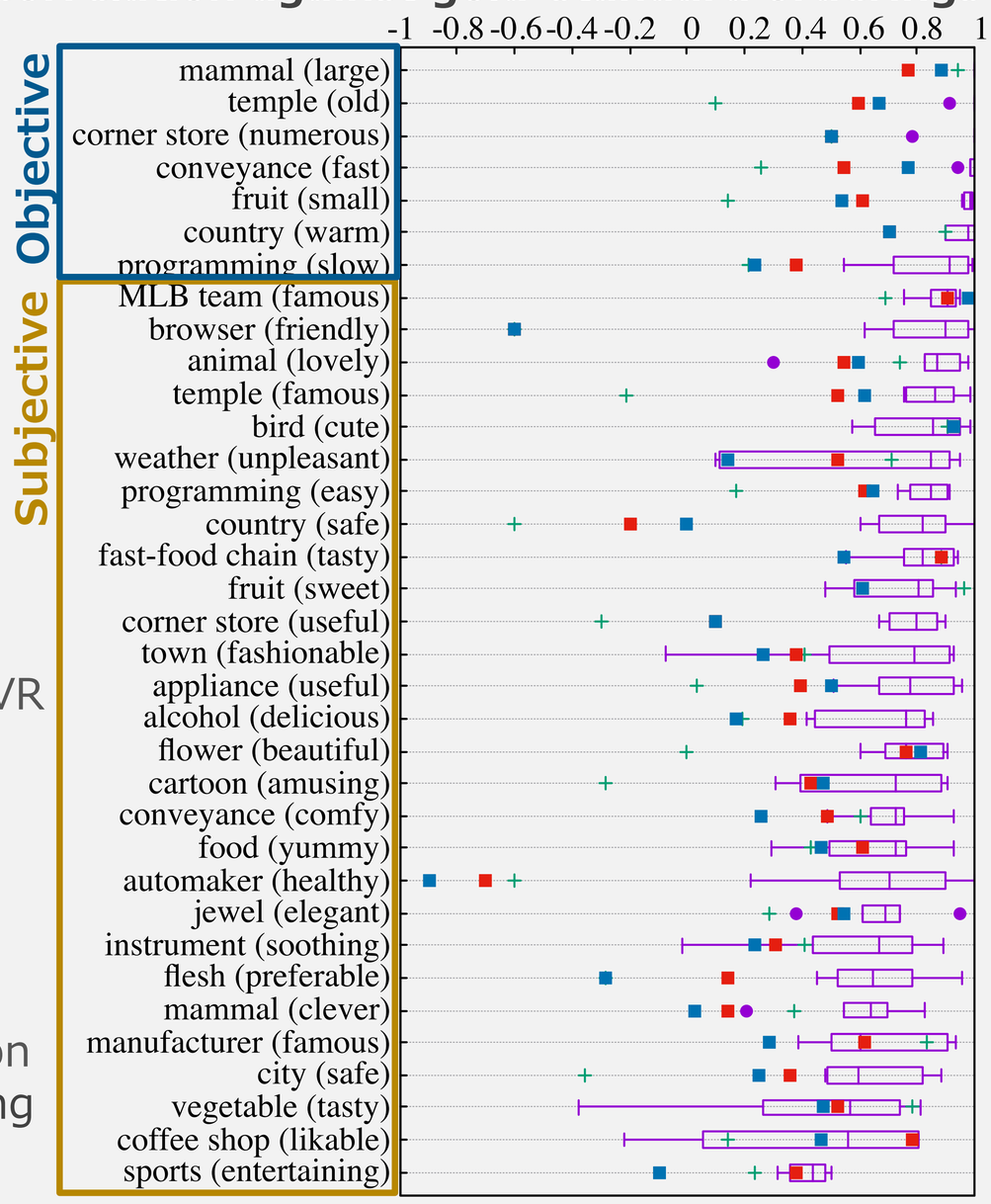
- **Proposed**: Ranking SVM / SVR
  - 4 types of evidence
  - Linear kernel
- **Baseline**: PMI of noun-adj co-occurrence [Turney '02]

[Train & Test]

- Leave one out cross-validation
- **Gold**: the ordering maximizing averaged Spearman's ρ against 7 human orderings

Human orderings showed high correlation

Correlation against gold-standard orderings



Method	Baseline	Ranking SVM	SVR	Human ave. ρ
ave. ρ	0.274	**0.441**	0.366	0.750

Ablation test for Ranking SVM

Features	ave. ρ
All features	**0.441**
⊖ co-occurrence	0.391
⊖ dependency	0.407
⊖ simile	**0.292**
⊖ comparative	0.424

Simile works well

Aggregating evidence is effective

Examples

Thank your for taking time!!

Want to see more?

Please visit: http://www.tkl.iis.u-tokyo.ac.jp/~nari/ijcai-16/