文書・文間及びカテゴリ間の関係を 考慮したレーティング予測

知能数理研究室 12056 外山 洋太

研究背景

多カテゴリにおける商品レビューのレーティング予測

- ▶ カテゴリ:レーティングの付く各項目のこと
- ▶ 文同士の位置関係及びカテゴリ間の関係が重要

ホテルの雰囲気はとてもよく食事もおいし 総合 ☆☆☆☆☆ かったです。部屋についても、窓からの見 サービス 晴らしがよく海がとても綺麗でした。チェッ 立地 クイン当日、入口のフロアの汚れが気に 部屋 なりましたが、翌日にはきちんと清掃され 設備・アメニティ ていました。機会があれば、また利用した 風呂 食事

研究背景

文同士の位置関係

(例1) 食事が美味しかった。 しかし、それよりも良かったことがある。 部屋からの眺めが素晴らしかった。 強調

(例2) 部屋からの眺めが素晴らしかった。 交換 しかし、それよりも良かったことがある。 食事が美味しかった。

3

研究背景

カテゴリ間の関係

- ▶ 食事○ ⇒ サービス○
- ▶ 設備・アメニティ ⇒ サービス ○



関連研究

隠れ状態を用いたホテルレビューのレーティング予測

- ▶ 文毎のレーティングからレビュー全体のレーティングを予測
- ▶ カテゴリ間の繋がりを手調整で変化させて考慮

パラグラフベクトル

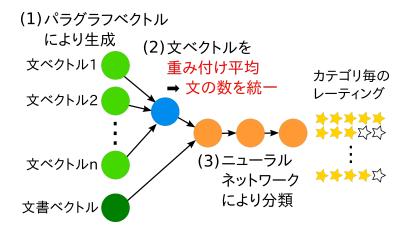
- ▶ 文や文書を、その意味を表す実数ベクトルに変換する手法
- ▶ 評判分類において優れる

ニューラルネットワーク

- ▶ 神経回路を模した機械学習手法
- ▶ 入力の要素間の複雑な関係を考慮

提案手法

レーティング予測の流れ



実験及び結果

実験設定

- ▶ 7カテゴリにおける 0~5点のレーティング予測の正答率を測定
- ▶ データセット:楽天トラベルにおけるレビュー約330,000件

結果

▶ 提案手法が従来手法より高い正答率を 示した

手法	正答率
従来手法	0.4832
提案手法	0.5030