

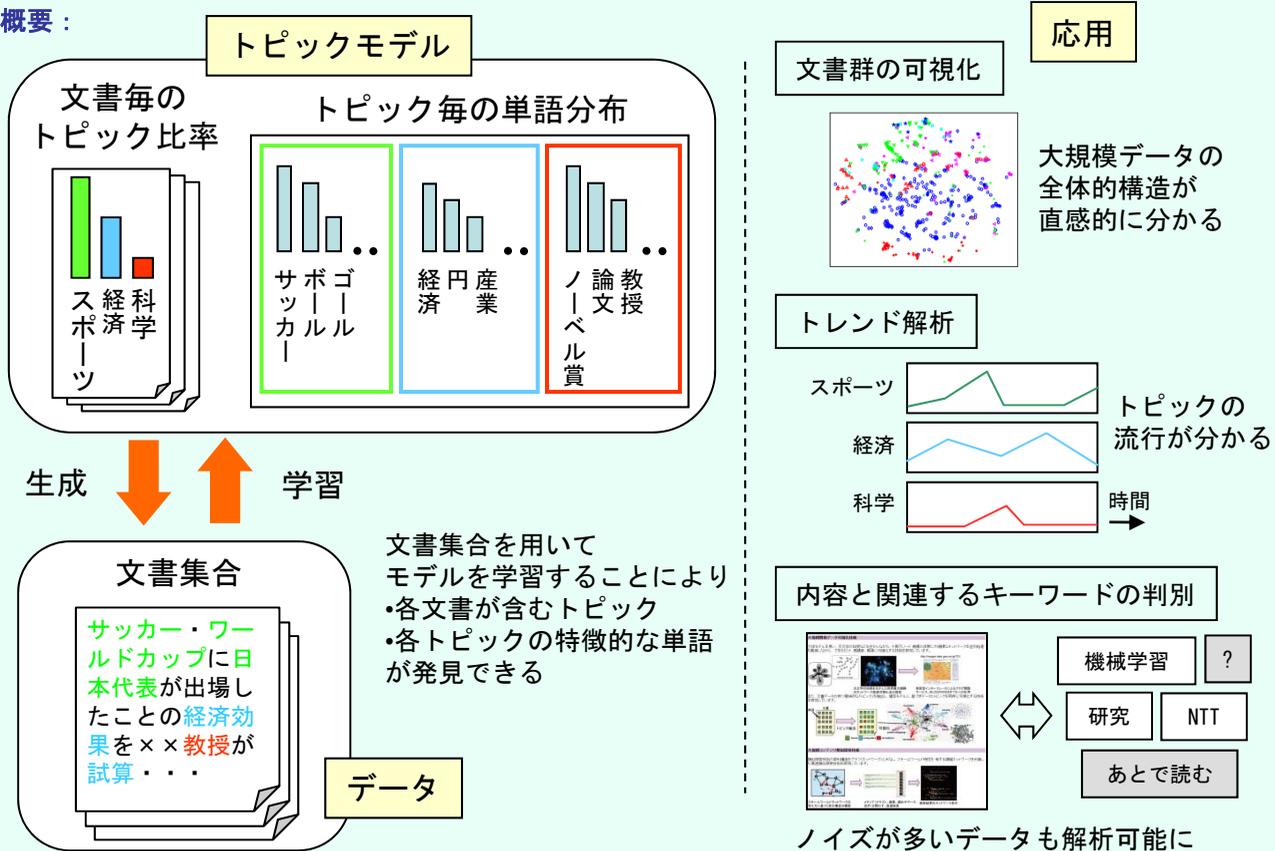
トピックモデルに基づくデータマイニング

背景・課題：インターネットをはじめとして、私たちの周りには膨大な情報があふれており、その中から不要な情報を捨て、有益な知識を取り出すことは、重要な研究課題となっています。トピックモデルを用いたデータマイニングはこの課題の解決法として近年注目されています。

アプローチ：トピックモデルとは、文書が潜在意味（トピック）に基づいて生成される過程を確率的に表現したモデルです。データを用いてモデルを学習することで自動的にトピックを抽出できます。トピックモデルは、文書、画像、購買データなど様々なデータでその有効性が確認されています。

到達点：トピックモデルに基づく3つの技術を開発しました。1) 大規模データの直感的な理解を可能にする可視化技術、2) ユーザ行動履歴データを用いたトレンド解析技術、3) 内容と関連するキーワードの自動判別技術。それぞれ従来技術と比較し優位性を示しました。

概要：



関連文献

Tomoharu Iwata, Takeshi Yamada, Naonori Ueda, "Probabilistic Latent Semantic Visualization: Topic Model for Visualizing Documents," KDD2008
Tomoharu Iwata, Shinji Watanabe, Takeshi Yamada, Naonori Ueda, "Topic Tracking Model for Analyzing Consumer Purchase Behavior," IJCAI2009
Tomoharu Iwata, Takeshi Yamada, Naonori Ueda, "Modeling Social Annotation Data with Content Relevance using a Topic Model," NIPS2009

連絡先

岩田 具治 (Tomoharu Iwata)

協創情報研究部 知能創発環境研究グループ

