

ベイジアンネットによる映画コンテンツ推薦方式の検討

小野智弘* 本村陽一** 麻生英樹** 芳賀麻誉美***

*株式会社 KDDI 研究所 **産業技術総合研究所 ***女子栄養大学

1. はじめに

近年、多数のコンテンツからユーザ個人に適合したものを推薦する方式やサービスが目立っており、amazon の書籍の推薦や Tsutaya Online の映画や DVD の推薦等効果を挙げている例も報告されている。これまでに推薦を行う枠組みとして、ユーザのコンテンツに対する購買や評価履歴からユーザの嗜好傾向を探り、嗜好傾向の類似した他人が好むコンテンツを推薦する協調フィルタリング方式や、コンテンツそのものや属性群からユーザの嗜好構造を反映させる特徴の抽出を行い、ユーザの好きなコンテンツ群の特徴と類似した特徴を持つコンテンツを推薦するコンテンツベース（属性ベース）の方式が提案されている。これらは一定の効果があるものの、前者は大量の履歴データが必要なこと、後者は適切な特徴抽出が困難といった欠点があるため、両者を組み合わせた方式も検討されている。また推薦システムの課題として推薦コンテンツの提示だけでは納得感が生まれにくく、心的根拠やユーザの状況等の推薦に至った根拠を示すことが望まれているが、ユーザ～コンテンツの共起パターンの類似性で判定する前者方式や、コンテンツの属性や特徴の類似性で判定する後方式では実現は困難である。筆者らはこれらの問題を解決するために、ユーザの映画視聴に関するアンケート、映画評価履歴、コンテンツ属性等を用いてベイジアンネットによりユーザの映画コンテンツ嗜好をモデル化し推薦を行う方式の研究を行っている[1]。本発表ではこの経過を報告する。

2. 目的

筆者らの想定する推薦システムの概要を図1に示す。システムの基本的枠組みはユーザ U_i からコ

ンテンツ推薦要求を受け付けると、候補となるコンテンツ C_j について、その属性群および U_i に関する属性情報を入力とし、ベイジアンネットモデルによりユーザ U_i がコンテンツ C_j を V_{ij} と評価する確率 $P(V_{ij} | U_i, C_j)$ を求め、 V_{ij} =良いとなる確率の高い順にコンテンツを推薦する。

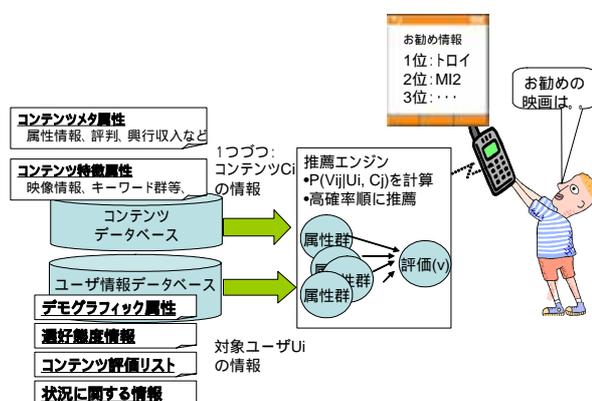


図1: 推薦システムの概要図

本研究の目的はこのような推薦システムを実現するためのベイジアンネットによる映画嗜好予測モデルを構築することである。ベイジアンネットを用いた理由は、1)確率推論アルゴリズムによりベイズ推定が実行可能、2)推論によって構築したモデルがそのまま推薦システムで利用可能、3)推薦しながらモデルが更新できるためである。一方共分散構造分析などの他のグラフィカルモデルは推薦システムにそのまま利用できない。なお、ベイジアンネットの説明は文献[2]を参照されたい。

3. 方法

3.1. データ

モデル構築に用いたデータは 2004 年 12 月に

1616 人への Web アンケート調査により取得した。概要は以下の通り。

- 1) リクルート条件： 提示した 100 コンテンツ中 20 作品以上視聴経験のある被験者
- 2) 調査項目（各ユーザ）：
 - (ア) デモグラフィック・ライフスタイル属性：年齢、性別、職業等 27 項目
 - (イ) 映画視聴に関する態度属性：映画コンテンツ（自宅での鑑賞も含む）の鑑賞頻度、映画選択時の重視項目、映画を見る主要目的（感動したい等 7 項目）
 - (ウ) ユーザのコンテンツ評価履歴：視聴経験のある 20 作品（ランダムに抽出）について、評価{良い・悪い}、その時の気分（感動した等 7 項目）
- 3) コンテンツ属性（100 コンテンツ分）：
 - (ア) メタ属性：ジャンル、収録時間等
 - (イ) 特徴属性：映画紹介文からテキストマイニングで抽出したキーワード群

3.2 モデル構築方法

ベイジアンネットワークモデルにおけるモデル構築は、当該分野の知識を活用した手動による構築と半自動による構築の 2 通りがある。後者の場合は構造に関する大まかな仮説に基づき変数間の依存関係（親子関係）の探索を繰り返しながらモデル構築を行う。筆者らは、仮説として図 2 の外枠に示すような評価構造モデルを用いた。これはコンテンツへの評価はコンテンツに対する感じ方に影響を受け、その感じ方はコンテンツ属性とユーザ属性の双方の影響を受けていることを示している。各種属性をこの仮説に基づきグループ分けし、相互情報量に基づく依存関係探索により変数を選択した。なおこの過程はベイジアンネットワーク構築ツール [3] を用いた。

4. モデル構築経過

構築中のモデルを図 2 に示す。ユーザ属性群（22 ノード）とコンテンツ属性群（6 ノード）がそれ

ぞれ共通属性（7 ノード）に結びつき、さらに総合評価へ結びついている。

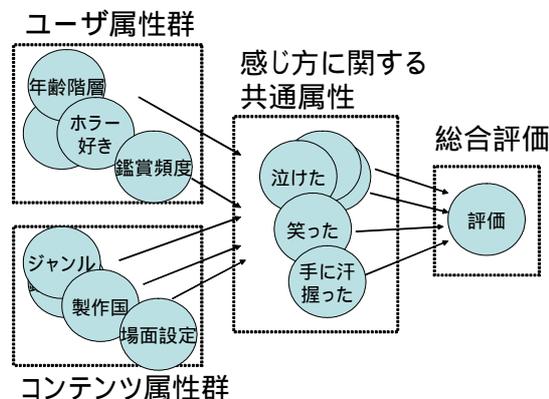


図 2 構築中の映画コンテンツ嗜好モデル

5. 考察と今後の課題

嗜好のモデル化では総合評価に強い影響を与える変数を抽出することが重要である。これまでの研究ではすでに取りあげたユーザ属性やコンテンツ属性についてこの影響の強さを評価との相関などから客観的に評価し選択することが多い。しかし、評価に対して各属性がどのように評価に影響を与えるかという問題は心理的に複雑なメカニズムを考慮しなければならず、線形の相関関係によるモデル化では不十分であると考えベイジアンネットワークによるモデル化を行っている。現在構築中のベイジアンネットワークモデルは既存のユーザ属性やコンテンツ属性を利用しているため、今後は評価グリッド法に基づく定性調査・定量調査により重要変数を抽出し、抽出した変数を用いてモデル構築を実施する予定である。

参考文献

- 1) 小野、本村、麻生、"ベイジアンネットワークによる映画コンテンツ推薦方式の検討," 電子情報通信学会技術報告(信学技報), vol.104, no.NC2004-66, pp.55-60, 2004
- 2) 本村、"ベイジアンネットワーク", 信学技報, no.NC2003-38, pp.25-30, 2003
- 3) 本村、"ベイジアンネットワークソフトウェア Bayonet", 計測と制御, vo.42, No.8, pp.693-694, 2003