

Context-Aware SVM に基づく状況依存型情報推薦方式の提案

奥 健太[†] 中島 伸介[†] 宮崎 純[†] 植村 俊亮[†]

[†] 奈良先端科学技術大学院大学 情報科学研究科 〒630-0101 奈良県生駒市高山町 8916-5

E-mail: †{kenta-o,shin,miyazaki,uemura}@is.naist.jp

あらまし ユーザの置かれている状況 (コンテキスト) に依存した情報推薦として, Context-Aware Support Vector Machine (C-SVM) に基づく状況依存型情報推薦方式を提案する。ユーザの要求はそのときのコンテキストに依存するが, これまでに協調フィルタリング等の情報推薦手法が数多く提案されているものの, ユーザの詳細なコンテキストを考慮したものはほとんどない。本研究では, ユーザに対する推薦情報の判定方法として, 2 クラス分類問題に適用される Support Vector Machine (SVM) を用いる。さらに, SVM の特徴空間にユーザの判断に影響するコンテキスト軸を追加した C-SVM を開発する。この C-SVM を用いて, コンテキスト依存型情報推薦システムを実現する。また, C-SVM をモバイル環境における飲食店情報推薦システムに適用した実験を行い, 情報推薦にコンテキストを考慮することの有効性を検証した。

キーワード 情報検索, 情報推薦, 状況依存 (Context-Aware), SVM, 協調フィルタリング

Context-Aware SVM for Context-Dependent Information Recommendation

Kenta OKU[†], Shinsuke NAKAJIMA[†], Jun MIYAZAKI[†], and Shunsuke UEMURA[†]

[†] Graduate School of Information Science, Nara Institute of Science and Technology

E-mail: †{kenta-o,shin,miyazaki,uemura}@is.naist.jp

Abstract The purpose of this study is to propose Context-Aware Support Vector Machine (C-SVM) for Context-Dependent information recommendation. User's demands depend on his/her contexts. However, currently, there are few recommendation methods considering user's contexts. Thus, in this study, a new type of SVM –Context-Aware SVM– is used to decide which information to recommend to the user. Usually, a SVM is applied to the problem of classification. We added axes of contexts to the feature space of the SVM. In this study, we apply C-SVM to a restaurant information recommendation system in a mobile environment. Our experiments show the effectiveness of C-SVM.

Key words Information Retrieval, Recommendation, Context-Aware, SVM, Collaborative Filtering

1. はじめに

膨大な情報源の中からユーザの嗜好に合った情報を見つけ出し, それをユーザに能動的に提供する情報推薦技術が注目されている。人の嗜好は, その人の置かれている状況に応じて変化するものである。例えば, ユーザの嗜好も含めた, ユーザの気分や体調, 行動目的などのような内的状況や, 季節や時間, ユーザのいる場所, 同伴者などのような外的状況など, あらゆる状況が, 人の情報選択の判断に対して影響を及ぼすと考える。しかしながら, 従来の情報推薦技術には, 情報推薦においてユーザの状況を考慮したものは少ない。

我々は, 人の情報選択の判断に対して影響を及ぼす外的及び内的状況を「コンテキスト」と定義し, そのようなコンテ

クトに依存した情報推薦システムの実現を目的とする。本研究では, コンテキスト依存型情報推薦システムを実現するための技術として, Context-Aware Support Vector Machine (以下, C-SVM) 及び Context-Aware Support Vector Machine に基づく協調フィルタリング (以下, C-SVM-CF) を提案する。

C-SVM は, Support Vector Machine (SVM) の特徴空間にユーザの情報選択の判断に影響するコンテキスト軸を追加したものである。SVM は 2 クラス分類問題に適用される手法であり, この SVM に基づいてユーザに対する推薦情報の判定を行う。このような SVM の特徴空間に対して, コンテキスト軸を追加することにより, 学習データ集合に対するユーザの好みをユーザのコンテキストも含めて分類することができ, ユーザのコンテキストに依存した推薦情報を判定することができる

える。

C-SVM-CF は、C-SVM をベースにしたコンテキスト依存型の協調フィルタリングである。協調フィルタリングは「嗜好が類似するユーザは同じものに対して興味を持つ」という仮説に基づく情報推薦手法であり、ある情報に対して類似ユーザが与えた評価を参考にし情報推薦を行う。C-SVM-CF におけるユーザ間の類似度は、類似度を求めたいユーザそれぞれの C-SVM の識別平面の類似度から求められる。このように、ユーザ類似度の計算に C-SVM をベースにすることにより、コンテキストも含めたユーザの好みに基づいて、類似ユーザを判定することができる。

本稿では、C-SVM 及び C-SVM-CF それぞれの説明及び特徴を示す。また、C-SVM 及び C-SVM-CF をモバイル環境における飲食店情報推薦システムに適用した実験を行い、その実験結果から情報推薦においてコンテキストを考慮することによる有効性を示す。

2. 関連研究および関連手法

2.1 情報推薦技術の関連研究

従来の情報推薦技術として、a) ユーザプロフィールを用いた情報推薦、b) 協調フィルタリング (Collaborative Filtering) を用いた情報推薦、などがある。

a) ではユーザの年齢や性別などのような個人属性や、過去の行動履歴などから抽出した嗜好や習慣などをユーザプロフィールとして作成し、そのユーザプロフィールに適合する情報を探し出す。これはユーザの嗜好を反映した情報推薦を目指すものである。

WebWatcher [1] や WebMate [2] は、ユーザの Web ページへのアクセス履歴などからユーザプロフィールを作成し、そのユーザプロフィールに基づいて情報推薦を行うものである。しかし、たとえ推薦された情報がユーザの嗜好に合っていたとしても、ユーザがその情報を望むコンテキストでなければ受け入れられない。

一方、b) は複数のユーザが協調して利用することを前提とした情報推薦手法であり、「嗜好が類似している他のユーザ」が高く評価している情報を推薦するという方法である。GroupLens [3] や MovieLens [4]、amazon.com [5] は、ネットニュースや映画、書籍などを対象とした協調フィルタリングシステムである。対象ユーザだけのプロフィールを用いる a) に比べ、b) は一度人間の判断に基づいて評価された情報を対象とするため、推薦情報に対する信頼度は比較的高いといえる。しかしながら、やはりユーザのコンテキストに合った情報でなければ、推薦情報はそのユーザには満足されにくい。

また、協調フィルタリングにおけるユーザ間の類似度は、ニュースの記事や映画、商品などに対する評価の相関係数によって求める。しかし、ここでは評価時のユーザのコンテキストを考慮していないため、求められたユーザ間の類似度は、やはり正確なものとはいえない。

本研究は、情報推薦においてユーザのコンテキストを考慮することにより、ユーザの満足度をより向上させることができる

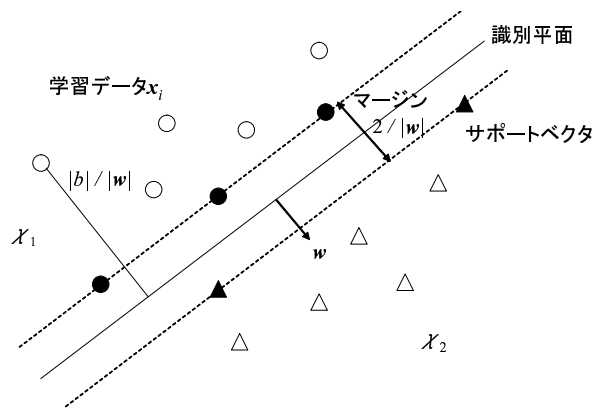


図1 SVMの特徴空間

という考えに基づき、コンテキスト依存型の情報推薦システムの実現を目指すものである。

2.2 クラス分類手法

本研究ではユーザの経験により蓄積された学習データ集合をユーザにとって満足なデータと不満足なデータに分類するという観点から、クラス分類手法を用いる。クラス分類手法により未知データに対してユーザに推薦すべきか否かの判定を行う。

クラス分類手法には、ニューラルネットワークや k 最近傍法、決定木、ベイジアンフィルタ、Support Vector Machine (SVM) などがある。

このうち本研究では以下のような利点を持つ SVM を用いる。

- 少ない学習データから識別関数を構成できる
- 膨大な次元数を持つ学習データにも対応可能
- 解が一意に定まる
- 最適化すべき学習パラメータが少ない

SVM は2クラスの分類手法の一つであり、1995年に Vapnik ら [6] によって提案された。SVM は主に画像認識や自然言語処理などの分野で利用されている。

図1のように学習データが分布する空間を SVM の特徴空間という。図中の“○”、“●”、“△”、“▲”は学習データを表現し、それぞれの形状は正データ、負データを表す。SVM では、学習データ集合に対する複数の識別境界候補の中で、図1に示すようなマージンを最大にする識別平面を求める。このマージン最大化が SVM の特徴であり、これにより解が一意に求められる。図1の“●”及び“▲”はサポートベクタとよばれ、実際にはこのサポートベクタだけから識別平面が構成される。

このような SVM により、ユーザの学習データ集合を満足、若しくは不満足に2クラスに分類する識別モデルを構築する。その識別モデルに基づいて、未知のデータがユーザにとって満足されるものか否かの推測を行う。ここで、満足されると推測されたデータがユーザに対する推薦情報として判定される。なお、本稿では SVM における識別モデルを SVM モデルと呼ぶ。

3. 提案手法

3.1 Context-Aware SVM (C-SVM)

Context-Aware SVM (C-SVM) は、図4のように、SVM の特徴空間にユーザの情報選択の判断に影響するコンテキスト

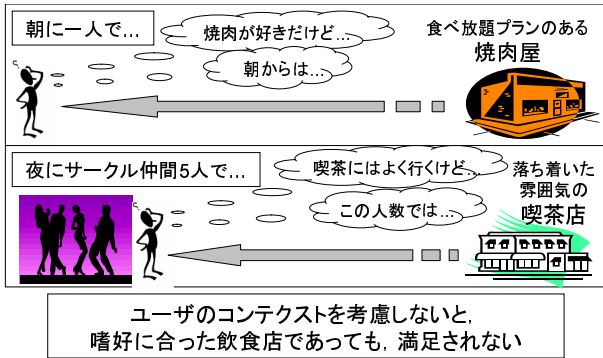


図 2 従来の情報推薦技術のイメージ

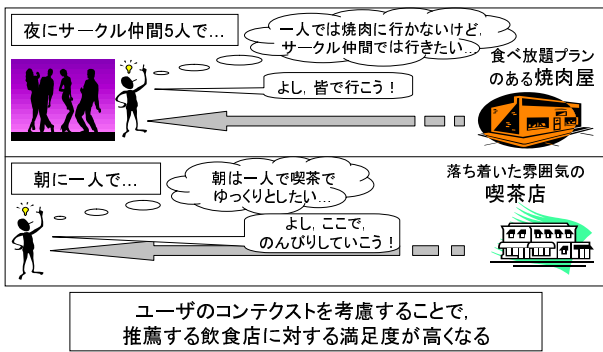


図 3 C-SVM による情報推薦のイメージ

軸 c を追加したものである。このように、コンテキスト軸を追加することにより、学習データ集合に対するユーザの好みに対し、ユーザのコンテキストも含めて、分類する識別平面を構成することができ、ユーザのコンテキストに依存した推薦情報を判定することができる。

C-SVM では、コンテキストをパラメータ化して表現し、あらゆるコンテキストにおけるユーザの好みを一つの特徴空間上で記述することができる。これにより、ユーザの情報選択に影響を及ぼすあらゆるコンテキストを総合的に考慮したうえで、推薦情報を判定することが可能となる。

ここでは、飲食店情報の推薦を例に、従来の情報推薦と比較しながら、C-SVM による情報推薦の特徴について述べる。

図 2 は従来の情報推薦技術のイメージである。従来の情報推薦技術はユーザの好みに合った情報を推薦しようというものである。図 2 では、焼肉が好きでユーザに焼肉屋を、喫茶が好きでユーザに喫茶店を推薦している。しかし、ユーザが朝に一人にいるときに食べ放題プランのある焼肉屋を、サークル仲間 5 人いるときに落ち着いた雰囲気のある喫茶店を推薦されても嬉しくない。たとえユーザに対して推薦した飲食店が、そのユーザの嗜好に合っていたとしても、そのときのユーザのコンテキストに適したものでなければ、満足されない。

これに対し、C-SVM ではコンテキストを考慮した情報推薦を行う。図 3 は C-SVM による情報推薦のイメージである。先の例に対し、ユーザが夜にサークル仲間 5 人いるときに、食べ放題プランのある焼肉屋が推薦されると嬉しい。また、ユーザが朝に一人いるときに、落ち着いた雰囲気のある喫茶店が推薦されると嬉しい。このように、飲食店の推薦においてユーザの

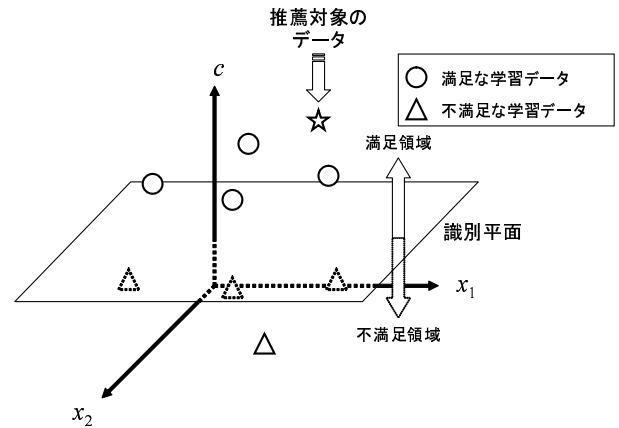


図 4 C-SVM の概念図

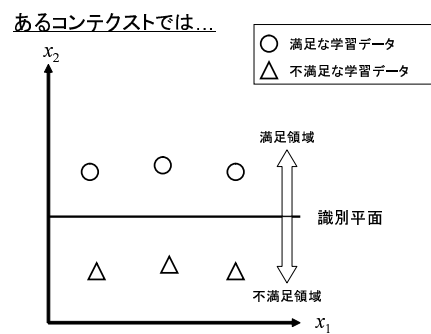


図 5 学習データ集合に対する満足、不満足の分類 (コンテキスト 1)

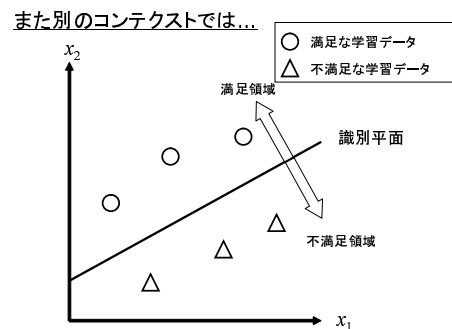


図 6 学習データ集合に対する満足、不満足の分類 (コンテキスト 2)

コンテキストを考慮することで、推薦する飲食店に対する満足度が高くなる。

以下、図 4、図 5 及び図 6 を用いて、C-SVM について説明する。

図 5 及び図 6 は、あるユーザの、学習データ集合に対する満足、不満足を分類したものである。図中の“○”、“△”は学習データを表現し、それぞれユーザにとって満足、不満足なデータを表す。また、学習データは便宜上 2 次元のパラメータ (x_1, x_2) で表現される。このパラメータは学習データの特徴を表現するものであり、本来ならば多次元で表現されるものであるが、ここでは説明のため 2 次元で表現されるものとする。

図 5 は、あるコンテキストにおける分類、図 6 は、また別のコンテキストにおける分類を示したものである。このように、学習データ集合に対する満足、不満足の分類がコンテキストご

と異なる場合、SVM の一つの特徴空間では、あらゆるコンテキストに対して有効な識別平面をひくことはできない。この問題に対して、特徴空間をコンテキストごとに用意する方法も考えられるが、あるコンテキストにおける学習データは、別のコンテキストでは学習データとして扱うことができなくなってしまう。

我々が提案する C-SVM では、SVM の特徴空間にコンテキスト軸を追加することにより、この問題を解決する。

図 4 の x_1, x_2 は、前述したように学習データのパラメータ軸である。ここで、C-SVM はコンテキスト軸を表す c 軸を追加する。これにより、コンテキストごとに異なる、学習データ集合に対するユーザの満足、不満足を分類することができる。なお、本来はコンテキスト軸も多次元で表現されるものであるが、ここでは説明のため便宜上 1 次元とした。本稿では C-SVM における SVM モデルを C-SVM モデルと呼ぶ。

C-SVM による情報推薦方法を、図 4 の C-SVM モデルで考える。ユーザのあるコンテキストにおいて、識別平面から満足領域にあるデータ (図中の “☆”) がユーザに対する推薦対象の情報として判定される。

3.2 Context-Aware SVM に基づく協調フィルタリング (C-SVM-CF)

Context-Aware SVM に基づく協調フィルタリング (C-SVM-CF) は、C-SVM をベースにしたコンテキスト依存型の協調フィルタリングである。C-SVM-CF は、コンテキストも含めた好みが類似しているユーザのお気に入り、かつ、ユーザのコンテキストに合致した情報を推薦する。

C-SVM-CF におけるユーザ類似度は、類似度を求めたいユーザそれぞれの C-SVM の特徴空間における識別平面の類似度から求められる。こうすることにより、コンテキストも含めたユーザの好みに基づいて、類似ユーザを判定することができる。これは、単に同じ好みを持つというだけでなく、“同じコンテキストにおいて、かつ同じ好みを持つユーザ”を類似ユーザとして扱うことにより、より正確に類似ユーザを判定することができるという考えに基づく。

C-SVM-CF の特徴は、情報そのものに対して各ユーザが与えた評価を直接比較するのではなく、各ユーザの C-SVM の識別平面の類似度から類似ユーザを判定するため、必ずしも各ユーザが同一の情報に対して評価を行っている必要はないということである。

ここでも、飲食店情報の推薦を例に、従来の協調フィルタリングと比較しながら、C-SVM-CF による情報推薦の特徴について述べる。

図 7 及び図 8 は、従来の協調フィルタリング及び C-SVM-CF それぞれにおいて扱うユーザ類似度の考え方である。図 7 のように、従来の協調フィルタリングでは、ユーザの互いの好みの店だけの類似からユーザ類似度を求めている。しかし、コンテキストに基づいた好みを考慮しないと、好みの店の類似だけでは正確には類似ユーザを判定できない。

これに対し、C-SVM-CF では図 8 のように、コンテキストも含めて好みの店が類似しているユーザを類似ユーザとして扱

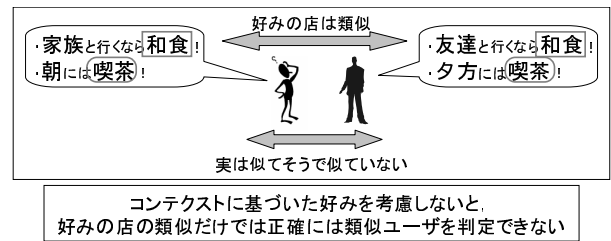


図 7 従来の協調フィルタリングにおけるユーザ類似度

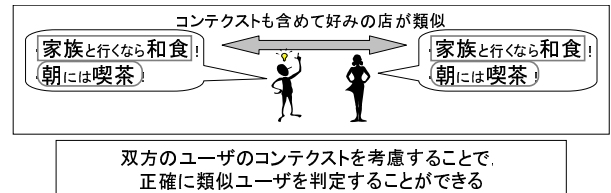


図 8 C-SVM-CF におけるユーザ類似度

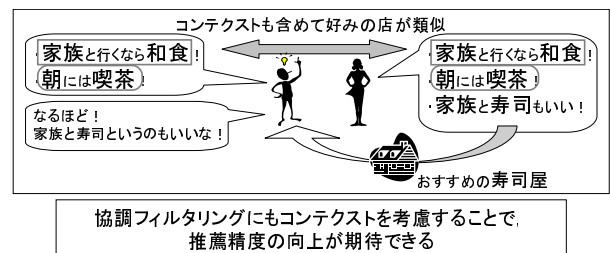


図 9 C-SVM-CF による情報推薦のイメージ

う。このように、双方のユーザのコンテキストを考慮することで、正確に類似ユーザを判定できると考える。

さらに、図 9 は C-SVM-CF による情報推薦のイメージである。コンテキストも含めた好みの店が類似しているユーザに基づいて、ユーザが家族と一緒にいるときに寿司屋を推薦している。このように、協調フィルタリングにもコンテキストを考慮することで、推薦精度の向上が期待できる。

C-SVM-CF におけるユーザ u とユーザ v との間のユーザ類似度 $sim(u, v)$ は式 (1) により求められる。

$$sim(u, v) = \frac{1}{2} \left(\frac{M_{v \rightarrow u}}{N_u} + \frac{M_{u \rightarrow v}}{N_v} \right) \times 100(\%) \quad (1)$$

ここで、 N_u はユーザ u の学習データ数、 $M_{v \rightarrow u}$ はユーザ v の C-SVM モデルによりユーザ u の学習データ集合を分類したときのクラスが一致したデータ数である。このユーザ類似度が、ある一定の閾値を超えたユーザ同士を類似ユーザとして扱う。

いま、ユーザ A とユーザ B の 2 人のユーザを考える。それぞれ、学習データ集合に対して、図 10 のような C-SVM モデルをもつ。ここでは説明の都合上、簡略化のため、学習データの特徴を表現する 1 次元のパラメータ x 及び 1 次元のコンテキストパラメータ c で構成される 2 次元の特徴空間を考える。図中の “○”, “△”, “●”, “▲” は学習データを表現し、それぞれ、ユーザ A にとって満足、不満足なデータ、ユーザ B にとって満足、不満足なデータを表す。

図 11 のように、各ユーザの学習データを互いの C-SVM モデルに写像したとき、 $N_A = 6, N_B = 5, M_{B \rightarrow A} = 5, M_{A \rightarrow B} = 4$

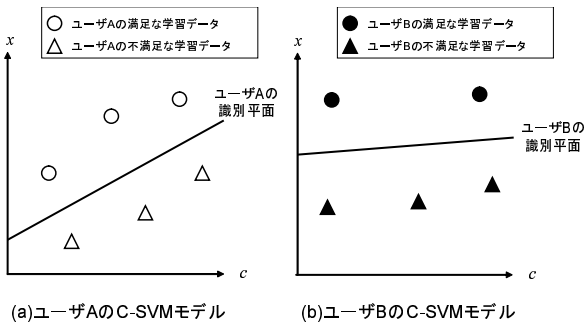


図 10 ユーザ A 及び B の C-SVM モデル

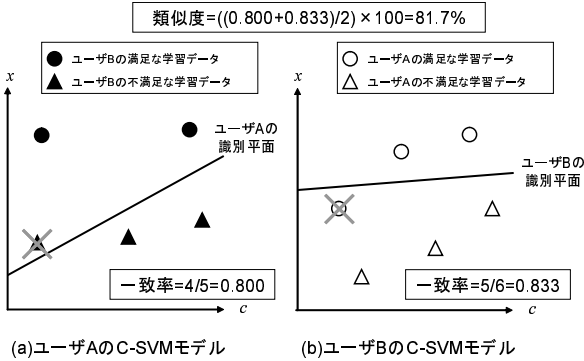


図 11 ユーザ類似度の計算

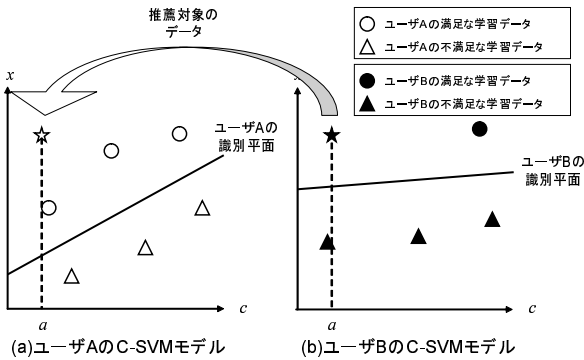


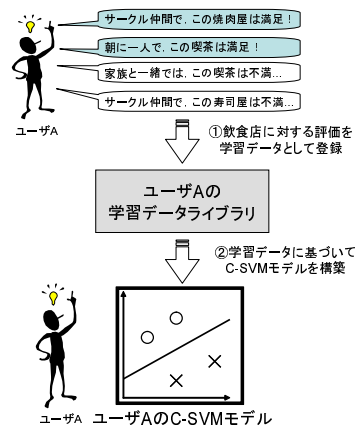
図 12 C-SVM-CF による情報推薦

であるから、ユーザ A とユーザ B との間のユーザ類似度 $sim(A, B)$ は、 $((5/6 + 4/5)/2) \times 100 = 81.7(\%)$ となる。ここで、類似ユーザ判定の閾値を 80% とすると、ユーザ A とユーザ B は類似ユーザとして扱われる。

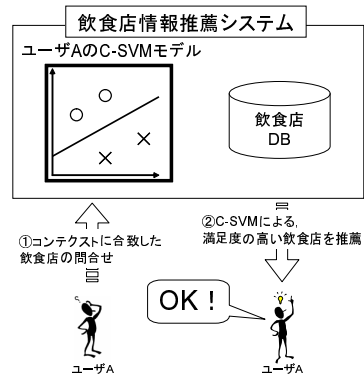
C-SVM-CF による情報推薦方法を図 12 で考える。いま、ユーザ A がコンテキスト a に置かれているとする。そのとき、ユーザ B の C-SVM モデルでは、同じくコンテキスト a において、ある学習データ (図中の “★”) に対して満足と評価していた。ここで、そのユーザ B の持つ学習データ (“★”) をユーザ A の C-SVM モデルに写像してみると、ユーザ A の C-SVM モデルにおいても、識別平面から満足領域に属する (図中の “☆” 参照)。したがって、このデータ (“★”) が、コンテキスト a において、ユーザ A に対する推薦情報として判定される。

3.3 飲食店情報推薦システムへの応用

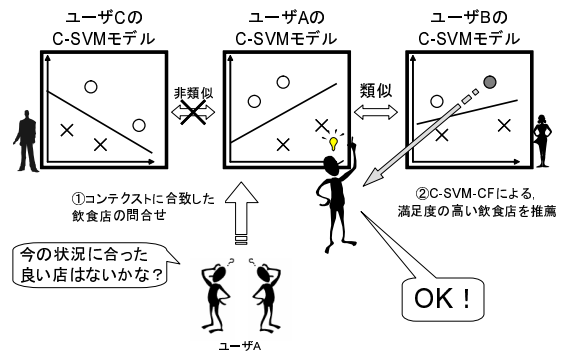
本研究では C-SVM 及び C-SVM-CF をモバイル環境における飲食店情報推薦システムへ適用する。飲食店情報の選択の際



(i) C-SVM モデルの構築



(ii) C-SVM による飲食店情報の推薦



(iii) C-SVM-CF による飲食店情報の推薦

図 13 飲食店情報推薦システムのイメージ

にはユーザのコンテキストによる影響を受けやすい。また、飲食店情報は日常的に利用機会が多く、ユーザごとに嗜好の違いが表れやすい。

図 13 は C-SVM 及び C-SVM-CF を飲食店情報推薦システムに適用したイメージである。

(i) は C-SVM モデルの構築のイメージである。ユーザ A が飲食店に対する評価を行った際、その評価対象の飲食店データと、そのときのユーザ A のコンテキストも含めて、学習データとして学習データライブラリに登録する。その学習データに基づいて、ユーザ A の C-SVM モデルを構築する。

(ii) は C-SVM による飲食店情報推薦のイメージである。ユーザ A はそのときのコンテキストに合致した飲食店情報をシス

テムに問い合わせる。(i)で構築したユーザAのC-SVMモデルにより、ユーザAに対して飲食店情報を推薦する。

(iii)はC-SVM-CFによる飲食店情報推薦のイメージである。(i)で構築したユーザAのC-SVMモデルに基づいて、ユーザAと類似するユーザを見つける。ユーザAはそのコンテキストに合致した飲食店情報を問い合わせる。C-SVM-CFにより、ユーザAと類似するユーザBのC-SVMモデルに基づいて、ユーザAに対して飲食店情報を推薦する。

4. 検証実験と考察

C-SVM及びC-SVM-CFの有効性を検証するために、C-SVM及びC-SVM-CFを飲食店情報推薦システムに適用した実験を行った。本実験の目的は以下のとおりである。

- 従来のSVMモデルとC-SVMモデルのテストデータに対する識別性能の違いを評価する
- 従来のSVMによる情報推薦に対する満足度と比較することによって、C-SVM及びC-SVM-CFによる情報推薦の有効性を検証する
- 学習データ数に対する、C-SVMモデルの識別性能及びC-SVMによる情報推薦に対する被験者の満足度の変化をみる。

4.1 実験環境

本実験では『Yahoo!グルメ』[7]を参考に、実験用の飲食店データベースを作成した。飲食店データの登録対象エリアとして、「心齋橋、なんば」エリアを選択した。「心齋橋、なんば」エリアは、大阪に位置する歓楽街であり、居酒屋、喫茶店など多様な種類の飲食店を有する。なお、『Yahoo!グルメ』において、「心齋橋、なんば」エリアに登録されている飲食店件数は938件(2005年12月1日現在)である。

飲食店の特徴を表現するパラメータ(以下、飲食店パラメータ)として、『Yahoo!グルメ』を参考に、表1に示した17個のパラメータを用いる。それぞれ、設備やサービスなどの有無を{0,1}の2値で設定することにより、飲食店の個性を表現する。

本実験の被験者は8名とした。いずれの被験者も、飲食店データの登録対象エリアである「心齋橋、なんば」エリアについては土地勘はあるものの、当然ながらすべての飲食店について把握しているわけではない。

本実験では表2に示す14個のコンテキストパラメータを用いる。各パラメータの設定値は表2のとおりである。ただし、コンテキストパラメータの項目及び設定値は、今回の実験用に暫定的に設定しただけであるため、コンテキストパラメータの選定については別途検討課題とする。

なお、SVMの学習データとして与えるパラメータ値の範囲は{0~1}をとりうる。そのため、表2に示したコンテキストパラメータの設定値を{0~1}の範囲に正規化する。

本実験ではSVMライブラリとしてLIBSVM[8]を用いる。LIBSVMにより学習データ集合からSVMモデルを構築する。LIBSVMには、SVMモデルの識別性能(accuracy)を求めするためのツール(*v*-fold cross-validation)及びSVMの最適な学習パラメータを選定するためのツール(grid-search)が付属している。

表1 飲食店パラメータ

駐車場あり	ランチサービスあり	パーティ, 2次会におすすめ
禁煙席あり	食べ放題プランあり	デートにおすすめ
個室あり	テイクアウト可	家族で食事におすすめ
カラオケあり	クーポンサービスあり	夜景がきれい
ライブあり	接待におすすめ	海がみえる
深夜営業 OK	宴会におすすめ	

表2 コンテキストパラメータ

カテゴリ	パラメータ	設定値
時間	月	1~12月
	時刻	0~23時
	曜日	月~日
スケジュール	エリアタイプ	歓楽街, 駅周辺, 観光地, 該当なし
	予算	0~10000円
	休日	一日休み, 休憩時間, 休日前, 該当なし
同伴者	人数(男性)	0~10人
	人数(女性)	0~10人
	最低年齢	0~100歳
	最高年齢	0~100歳
	関係	家族, 恋人, 友人, 上司, 部下, 該当なし
	身分	学生, 社会人, 該当なし
外的要因	天気	晴れ, 曇り, 雨
	気温	-5~40

4.2 実験手順

本実験は以下の手順にしたがって行う。

(1) 学習データの登録からSVMモデル及びC-SVMモデルの構築

(2) 構築したモデルに基づく情報推薦

以下、各段階について詳細に述べる。

4.2.1 学習データの登録からSVMモデル及びC-SVMモデルの構築

被験者に対して推薦対象となる情報をSVMにより判定するためには、あらかじめ被験者のSVMモデルを構築しておく必要がある。SVMモデルは被験者によって蓄積された学習データ集合から構築される。ここでは、被験者による学習データの登録手順について述べる。

被験者による学習データの登録の流れを以下に示す。

(i) 被験者自身が遭遇するコンテキスト(スケジュール及び同伴者)を、あらかじめ5パターン設定する。

(ii) 設定した5パターンのコンテキストを想定し、飲食店データベースからランダムに提示される20件の飲食店に対して、満足度の評価を行う。すなわち5パターンのコンテキストそれぞれに対して、20件の飲食店となり、合計100回の実験を行う。

被験者は提示された飲食店情報に対する満足度を10段階で評価する。10段階評価のうち、ある閾値以上を満足、閾値未満を不満足として、学習データを登録する。この閾値のとり方については検討する余地があるが、今回は閾値を6とした。

SVM の学習データは、満足、不満足を表現するクラスラベル $\{+1, -1\}$ と、表 1 に示す飲食店パラメータ値 $\{0, 1\}$ から構成される。C-SVM の学習データは、これに加え、表 2 に示すコンテキストパラメータの設定値を正規化した値 $\{0 \sim 1\}$ から構成される。

被験者によって登録された学習データ集合から、LIBSVM により SVM モデル及び C-SVM モデルを構築する。構築したモデルの識別性能は、 v -fold cross-validation ($v = 5$) により得る。

また、C-SVM-CF による情報推薦を行うために、各被験者の C-SVM モデルから、各被験者間のユーザ類似度を計算しておく。

4.2.2 構築したモデルに基づく情報推薦

構築した SVM モデル及び C-SVM モデルに基づいて、被験者に対して飲食店情報の推薦を行う。被験者は推薦された飲食店情報に対して満足度の評価を行う。

被験者に対する情報推薦の流れを以下に示す。

(i) 10 パターンのコンテキストそれぞれにおいて推薦される 20 件の飲食店に対して、満足度の評価を行う。

(ii) 10 パターンのコンテキストそれぞれに対して、20 件の飲食店となり、合計 200 回の評価を行う。

各コンテキストパターンにおける情報推薦は、以下の 4 種の推薦手法により行い、各種推薦手法による推薦情報に対する被験者の満足度から比較検証を行う。

- (a) Random 提示
- (b) SVM による情報推薦
- (c) C-SVM による情報推薦
- (d) C-SVM-CF による情報推薦

つづいて、学習データ数に対する C-SVM モデルの識別性能及び推薦情報に対する被験者の満足度の変化をみる。被験者 1 名を対象に、学習データ数 $\{100, 150, 200, \dots, 500\}$ それぞれにおける、C-SVM モデルの識別性能及び推薦情報に対する被験者の満足度を得る。

4.3 結果及び考察

4.3.1 識別モデル構築にコンテキストを考慮することによる識別性能の評価

構築した SVM モデル及び C-SVM モデルの識別性能を比較し、識別モデルの構築にコンテキストを考慮することによる識別性能を評価した。ここでは、識別モデルの構築に用いた学習データをそのままテストデータとして識別性能を得た。したがって、既知データに対する識別性能であるため、識別性能が 100% となることが望ましい。

図 14 は被験者別の SVM モデル及び C-SVM モデルの既知データに対する識別性能を比較したものである。C-SVM モデルではいずれの被験者も 100% に近い識別性能が得られたのに対し、SVM モデルでは平均 70% にも満たなかった。

この実験結果は、識別モデルの構築にユーザのコンテキストを考慮した C-SVM モデルでは、コンテキストごとに変化するユーザの好みも正確に分類することができることを示唆している。一方、ユーザのコンテキストを考慮していない SVM モデ

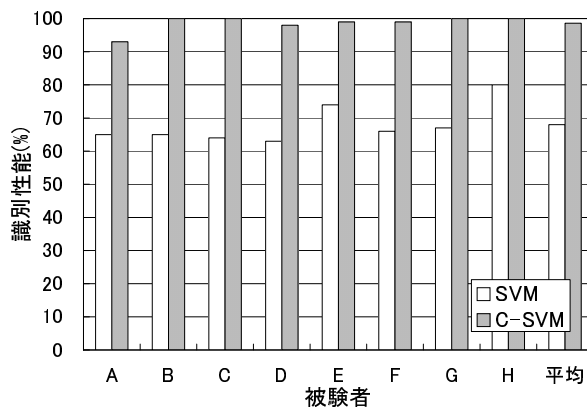


図 14 SVM モデル及び C-SVM モデルの識別性能

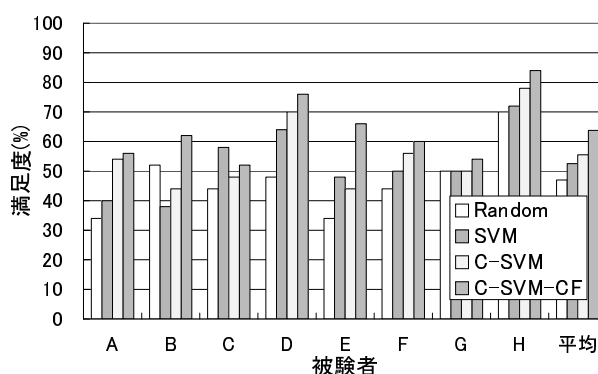


図 15 推薦情報に対する被験者の満足度

ルでは、ユーザの好みコンテキストごとに変化するということに対応できていないため、識別性能が平均 70% という結果となったと考える。

4.3.2 C-SVM 及び C-SVM-CF による情報推薦の有効性の検証

本実験では、a) Random, b) SVM, c) C-SVM, d) C-SVM-CF の 4 種の推薦手法により被験者に対して情報推薦を行った。これらの 4 種の推薦手法による推薦情報に対する被験者の満足度を比較することによって、C-SVM 及び C-SVM-CF による情報推薦の有効性を検証した。

なお、今回の被験者におけるユーザ類似度は、42.5% ~ 62.5% となり、極端に類似度が高い、若しくは低いということはない。このうち、ユーザ類似度 50% 以上を類似ユーザとして、C-SVM-CF による情報推薦を行った。

図 15 は 4 種の推薦手法による推薦情報に対する被験者の満足度を比較したものである。平均的にみると、C-SVM-CF に対する満足度が最も高く、次いで、C-SVM, SVM, Random 提示の順に高かった。

今回の実験は、実験対象エリアを「心齋橋、なんば」エリアとし、被験者は 8 名とも「心齋橋、なんば」エリアについて土地勘がある、というような条件下において C-SVM-CF による情報推薦を行った。この実験結果は、類似ユーザが推薦対象エリアにおけるいくつかの飲食店データに対して評価済みである、

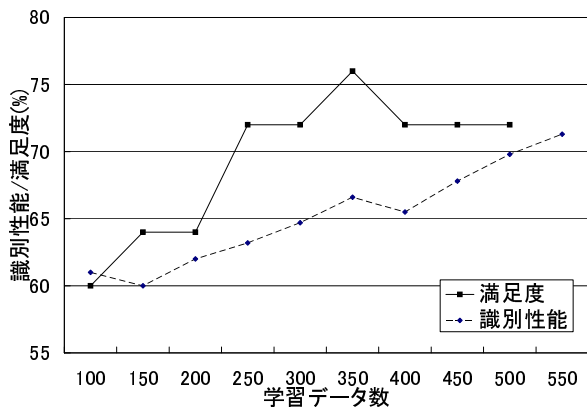


図 16 学習データ数に対する C-SVM モデルの識別性能及び被験者の満足度の変化

という条件下では、C-SVM-CF が有効に働くことを示した。

今回の実験結果では、C-SVM-CF に対する満足度は平均して 64% であったが、被験者数を増やし、類似ユーザが増えることによって、より満足度を向上させることができると期待できる。

4.3.3 学習データ数に対する C-SVM モデルの識別性能及び満足度の変化の分析

学習データ数を増加させることによる、C-SVM モデルの識別性能及び満足度の変化をみた。

図 16 は 1 名の被験者に対して、学習データ数を {100, 150, 200, 250, ..., 500} のように増加させたときの、C-SVM モデルの識別性能及び被験者の満足度の変化を示したものである。

この実験結果から、C-SVM モデルの識別性能は、学習データ数の増加とともに向上していることが分かる。したがって、学習データ数が蓄積されることによって、より識別性能の高い C-SVM モデルが構築されることが期待できる。

一方、C-SVM による推薦情報に対する被験者の満足度も、学習データ数の増加とともに向上している。この結果から、今回の実験条件のように、推薦対象エリアが同一という条件下では、その推薦対象エリアにおける学習データが蓄積されることにより、C-SVM による情報推薦の有効性が高まるということが期待できる。

ただし、被験者の満足度は 70% を超えたあたりで横ばいの状態になった。これは今回は飲食店パラメータ及びコンテキストパラメータを暫定的に設定したため、それぞれのパラメータが飲食店及びコンテキストを適切に表現できていないことが原因であると考えられる。今後は、C-SVM による情報推薦に対するユーザの満足度を向上させるために、飲食店パラメータ及びコンテキストパラメータに対する十分な検討が必要である。

5. おわりに

本研究では、コンテキスト依存型の情報推薦システムを実現するための技術として、Context-Aware SVM (C-SVM) 及び Context-Aware SVM に基づく協調フィルタリング (C-SVM-

CF) を提案した。

本研究では、C-SVM 及び C-SVM-CF を飲食店情報推薦システムに適用した実験を行い、情報推薦においてコンテキストを考慮することによる有効性を検証した。

実験結果より、以下の事項を示した。

- 識別モデルの構築にユーザのコンテキストも含めた C-SVM モデルでは、コンテキストごとに変化するユーザの好みも正確に分類することができる

- 限られた推薦対象エリアにおいて、ユーザ自身の学習データを蓄積することにより、推薦情報に対するユーザの満足度をより向上させることが期待できる

- 類似ユーザが推薦対象エリアにおけるいくつかの飲食店データに対して評価済みである、という条件下では、C-SVM-CF が有効に働く

なお、今回は C-SVM 及び C-SVM-CF を飲食店情報推薦システムに適用したが、他にも音楽情報推薦や旅行情報推薦などのように、ユーザのコンテキストが情報選択の判断基準に影響を及ぼすような分野に適用可能であると考えられる。

今後の検討課題として、以下の事項が挙げられる。

- コンテキストパラメータの直交性の検証
- より詳細なコンテキストパラメータ (過去の行動履歴、ユーザの目的、現在地からの距離、周辺店舗の特性、同伴者の嗜好など) の検討

- 被験者数の増加や実験対象エリアの多様化などによる、多様な条件下における C-SVM 及び C-SVM-CF の挙動の検証

- 多クラス分類への対応、複数飲食店データベースへの対応、飲食店パラメータの検討

謝 辞

本研究の一部は、科学技術振興機構戦略的創造研究推進事業 (CREST) 「高度メディア社会の生活情報技術」、科学研究費補助金基盤研究 (A) (2) (課題番号: 15200010)、若手研究 (B) (課題番号: 17700132) の支援による。ここに記して謝意を表す。

文 献

- [1] T. Joachims, D. Freitag and T. Mitchell: "WebWatcher: A Tour Guide for the World Wide Web", *Proc. of IJCAI'97* (1997).
- [2] L. Chen and K. Sycara: "WebMate: A Personal Agent for Browsing and Searching", *Proc. of the 2nd International Conference on Autonomous Agent (Agent'98)*, pp. 132-139 (1998).
- [3] P. Resnick, N. Iacovou, M. Suchak, P. Bergstrom and J. Riedl: "GroupLens: An Open Architecture for Collaborative Filtering of Netnews", *Proc. of ACM 1994 Conference on Computer Supported Cooperative Work (CSCW'94)*, pp. 175-186 (1994).
- [4] GroupLens Research, "movielens", <http://movielens.umn.edu/login>
- [5] amazon, "amazon.com", <http://www.amazon.com/>
- [6] C. Cortes and V. Vapnik: "Support-Vector Networks", *Machine Learning* Vol. 20, No. 3, pp. 273-297 (1995).
- [7] Yahoo Japan, "Yahoo!グルメ", <http://gourmet.yahoo.co.jp/gourmet/>
- [8] C-C. Chang and C-J. Lin: "LIBSVM: a Library for Support Vector Machines" (2005).