

# 移住支援のための人工知能によるマッチングシステム

## AI system for relocation matching

大輪 拓也<sup>1\*</sup> 中尾 悠里<sup>1</sup> 吉良 知文<sup>2</sup> 神山 直之<sup>2</sup> 吉田 宏章<sup>1</sup> 大堀 耕太郎<sup>1</sup>  
Takuya Ohwa<sup>1</sup> Yuri Nakao<sup>1</sup> Akifumi Kira<sup>2</sup> Naoyuki Kamiyama<sup>2</sup> Hiroaki Yoshida<sup>1</sup> Kotaro Ohori<sup>1</sup>

<sup>1</sup> 株式会社富士通研究所 <sup>2</sup> 九州大学マス・フォア・インダストリ研究所  
<sup>1</sup> FUJITSULABORATORIES LTD. <sup>2</sup> Institute of Mathematics for Industry, Kyushu University

**Abstract:** 近年都市部から地方への移住希望者が増加しており、移住先に定住してもらうことは地方にとっても日本社会にとっても重要であるが、移住前の認識と移住後の実際の生活との齟齬から再び離れていってしまうなどの問題がある。我々は、このような課題を解決するための人と地域とのマッチングシステムを提案する。本システムの実現においては、コールドスタート問題への対応や、対話的に推薦を行える状況の活用などの特徴があり、これらは人工知能によるアプローチによって実現される。

## 1 はじめに

移住希望者に移住候補となる適切な地域を提示するために、少ない学習データから始めても利用者が増えていくことで適切な提示が行えるようになるシステムを提案する。近年は都市部から地方への移住希望者が増加しているが、それでも地域によっては移住を検討される機会が少ない場所もあるため、そのような地域に対する学習データを集めることは難しい。また、移住希望者が移住先の地域に関する情報を十分に把握していなければ自身の希望を明確に伝えられないので、システムの入力はそのような問題に対処する必要がある。本提案はこれらの問題を解決したシステムである。本論文の主な貢献は以下である。

- 学習データが少ない場合でも、システムの利用者が増えることで推薦精度が向上していく方法を提案し、実験によってその効果を証明した
- 自身が望む移住先を明確に伝えられなくても、システムと対話しながら検索することで好みの地域が発見できる

## 2 問題設定

### 2.1 背景

近年、都市部から地方への移住に関する注目度が高まっており、希望者が多い地域では市役所が情報を提供したり [2] 相談に対応したりしている。そのような

状況の中、移住希望者の多様なニーズと地域特性とのミスマッチによる機会損失が懸念されている。例えば、移住者が再び別の地へ移ってしまうことがあり、移住前の認識と移住後の実際の生活との齟齬が原因の一つとして考えられる。このような問題に対応するために、移住希望者に対し適切な移住候補を提供する仕組みの構築が必要である。

### 2.2 課題

システムを構築する上で主な課題は二つある。

一つ目は、学習に用いるデータが少ないことである。機会損失を減らすためには移住対象となる地域全てを提示候補にできるモデルが望ましく、できるだけ多くの地域に対する移住検討のデータを学習することがモデルの精度向上に貢献する。しかし、地域によっては移住を検討される機会が少ない場所もあるため、現実的には全ての地域に対する十分な学習データを蓄積することは困難である。また、システムを利用することで得られるデータを学習に活用することは考えられるが、導入時においてはそのようなデータも少ないので、所謂コールドスタートの問題にも対処しなければならない。

二つ目は、移住希望者が真に望む移住先を正しく伝えられないことである。例えば、移住候補の一つに“病院が充実している”という特徴をもつ地域があっても、移住希望者がそのような観点について考慮していなければ、システムに自ら伝えることはない。しかし、実際に住んでみた後に“病院についても考慮すべきであった”と考える場合もあり、満足度の低下から再び移住してしまう懸念がある。従って、システムの入力には

\*連絡先：株式会社富士通研究所  
〒211-8588 神奈川県川崎市中原区小田中 4-1-1  
E-mail: takuyaohwa@jp.fujitsu.com

移住先の特徴などの情報ではなく、移住希望者が明確に伝えられる情報であることが望ましく、そのような入力情報から適切な移住先を判断するモデルの構築が必要である。また、上記の例のような満足度低下の問題に対処するために、システムの利用を通じて移住先の特徴を理解してもらえらる仕組みが含まれていると尚良い。

### 3 システム概要

対話的な方法で地域を推薦するマッチングシステムを構築する。処理の概要は Algorithm 1 に記述した。入

---

#### Algorithm 1 マッチングシステム

---

- 1: 属性情報を入力
  - 2: お勧め地域のリストを提示
  - 3: **while** 好みの地域が決定しない **do**
  - 4:   重視する地域指標を選択
  - 5:   お勧め地域を再計算し、リストを提示
  - 6: **end while**
- 

力には、例えば性別や年齢など移住希望者が明確に回答できる情報を用いる。それらを入力すると、お勧めの地域のリストが提示される。そのリストの中に満足する地域が見当たらない場合、システムに何らかの情報を伝え、その情報を頼りにより相応しいと思われる地域を再計算しリストを提示する。これを繰り返して好みの地域を検索していく。より詳しくは次章で述べる。

### 4 提案手法

移住希望者に候補となる地域を提示するために、離散選択モデルを用いる方法が考えられる。今、移住対象となる地域の候補全体を  $R$  とし、各地域  $r \in R$  を  $n$  種類の特徴ベクトル  $\mathbf{x}(r) = (x_1(r), \dots, x_n(r))$  で表す。例えば

- $x_1(r)$ : 地域  $r$  の病院件数,
- $x_2(r)$ : 地域  $r$  の交通利便性,
- $x_3(r)$ : 地域  $r$  の学校の数

などである。このとき、移住希望者  $h$  が地域  $r$  を好むかどうかを、 $h$  の  $r$  に対する効用関数

$$U_h(r) = \beta_1^{(h)} x_1(r) + \dots + \beta_n^{(h)} x_n(r) + \alpha^{(h)}(r)$$

によって表現する。ここで、 $\beta_i^{(h)}$  は地域の特徴  $x_i$  の影響度を表す係数、 $\alpha^{(h)}$  はノイズ項である。以降では  $\beta^{(h)} = (\beta_1^{(h)}, \dots, \beta_n^{(h)})$  や  $\alpha^{(h)}$  を効用値と呼ぶことと

する。効用関数が決定すれば、それによる離散選択モデルを用いて候補の地域を提示する方法が構築できる。

この方法を実現する際は効用関数を正しく推定することが重要であるが、そのためには一人のユーザに関する複数回の試行データが必要である。例えば、各地域  $r \in R$  に対し  $i_r$  回の評価値 (ex. 採点)

$$e_{h,1}(r), e_{h,2}(r), \dots, e_{h,i_r}(r) \quad (1)$$

が得られれば、これと効用関数の二乗誤差

$$\sum_{r \in R} \sum_{i=1}^{i_r} (U_h(r) - e_{h,i}(r))^2$$

を最小化するような  $\beta^{(h)}$  や  $\alpha^{(h)}$  を決定することができる。しかし、移住に関しそのようなデータを取得することは難しい。何故なら、各地域  $r \in R$  に対する評価を複数回する負荷の問題があり、地域数が多ければ負荷はさらに大きくなる。そこで、ユーザに関する何らかの情報からユーザをクラスタリングし、そのクラスタに対して効用関数を求める方法を考える。

#### 4.1 ユーザークラスタリング

移住希望者  $h$  に対し、 $\mathbf{y}(h) = (y_1(h), \dots, y_m(h))$  は  $m$  種類の属性を表すベクトルとする。例えば

- $y_1(h)$ :  $h$  の性別,
- $y_2(h)$ :  $h$  の年齢区分,
- $y_3(h)$ :  $h$  の結婚の有無

などである。このような属性ベクトルから移住希望者  $h$  をクラスタリングし、そのクラスタ  $C$  ごとに効用関数

$$U_C(r) = \beta_1^{(C)} x_1(r) + \dots + \beta_n^{(C)} x_n(r) + \alpha^{(C)}(r)$$

を構築する。例えば、(1) のような評価値が得られたとき、

$$\sum_{h \in C} \sum_{r \in R} \sum_{i=1}^{i_r} (U_C(r) - e_{h,i}(r))^2$$

を最小化するような  $\beta^{(C)}$  や  $\alpha^{(C)}$  を決める。このようなモデルを構築することで、Algorithm 1 の行 1 から行 2 を実現する。

ここで注意点を一つ述べる。効用関数を正しく推定する目的のためには、効用値  $\beta^{(C)}, \alpha^{(C)}$  ができるだけ同じ値となるようなクラスタリングであることが望ましいが、上記のクラスタリングはあくまで属性値  $\mathbf{y}$  をもとにしたクラスタリングなので、この目的を達成するかどうかはわからない。しかしながら、データが少ない状態においては困難なため、システム初期の時点

では“同じ属性  $y$  の人は似た効用値  $\beta^{(C)}, \alpha^{(C)}$  を持つ”という仮定のもとでクラスタリングを行う。その後、システムの利用者が増えデータが蓄積されてきたら、クラスタリングを再構築する。この点については後の 4.3 章で述べる。

## 4.2 フィードバックによる再計算

Algorithm 1 の行 3 以降の処理方法を述べる。最初に提示されたリストの中に満足する地域がない場合に、フィードバックを受け取ることでリストを更新する。まず、お勧めの地域  $r_1, \dots, r_k$  のリストを提示する際、各地域の特徴  $\mathbf{x}(r_1), \dots, \mathbf{x}(r_k)$  も記載する。その値を元に、利用者は“どの特徴をもっと重視したいか”をシステムに入力する。例えば

- “もっと病院数 ( $x_1$ ) が多いほうが良い”
- “もっと交通利便性 ( $x_2$ ) が高いほうが良い”

などである。上記で入力された各特徴  $x_i$  に対し、 $\beta_i^{(C)}$  の値を変更することで効用関数を再計算し、提示する地域を更新する。

## 4.3 クラスタリングの再構築

システムの初期においては、4.1 章で述べたように利用者の属性をもとにしたクラスタリングによる離散選択モデルが使われる。その後、システムが利用されると 4.2 章の方法によって“最終的に満足する結果が得られた際に使われた効用値”が蓄積される。このデータをクラスタリングの再構築に利用する。

今、過去にこのシステムを利用したあるクラスタ  $C$  に含まれる利用者のうち、満足する地域が発見できた利用者全体を  $\bar{C} \subset C$  とし、満足する地域が提示された際に使われた効用値全体を

$$B^{(C)} = \{\beta^{(C;h)}; h \in \bar{C}\} \quad (2)$$

で表す。現在のクラスタ全体を  $C_{current}$  とし、再構築したクラスタ全体を  $C_{new}$  としたとき、 $(B^{(C)})_{C \in C_{current}}$  の分散よりも  $(B^{(C)})_{C \in C_{new}}$  の分散のほうが小さくなるようにする。クラスタの目的は効用値が似た利用者をまとめることなので、上記の方法でそれが期待できる。

## 5 関連研究

永井ら [3] は、コンパクトシティを実現する上で考慮すべき施設などの配置の問題をシミュレーションによって考察した。本件の移住支援への対策も、コンパ

クトシティを実現する際の問題に対する取り組みの一つである。コンパクトシティ化を推進する過程において、住居の移動が困難な居住者が周辺部に取り残される状況があり、稲葉 [1] は居住者が少ない地域では住民同士の交流が活発でなくなることから犯罪の発生率が高くなることを指摘している。移住支援はこのような地域の発生を解消することに貢献できる。

対話型の推薦を行う別の例としては、渡辺らが提案する観光地の推薦に対する対話型のシステムがある [4]。観光目的に関する自然言語を入力とするシステムで、推薦結果が利用者の望む内容でない場合に、システムとの対話によって推薦するアイテムを探していく。本論文においては入力の属性や出力の地域などが何らかの特徴で表されているのに対し、渡辺らの論文では入力や出力が特徴値で表されておらず、この点が大きく異なる。

## 6 実験

提案手法の効果を検証する。クラスタの再構築は、利用者ごとに実施する。フィードバックは以下のように行う。

### 6.1 効用値の更新方法

$\beta^{(C)}$  の更新方法について説明する。多くの地域  $r \in R$  に対する学習データを蓄積するためには、様々な値の  $\beta^{(C)}$  を採用すれば提示される地域もそれに伴って多様に変化するため、この目的を達成できる。例えば、 $\beta^{(C)}$  をランダムな方法で更新することが考えられる。しかし、ランダムな方法では利用者が望む結果にたどり着きにくい場合が想定される。そこで、(2) で定義された  $B^{(C)}$  を活用する方法を適用する。

まず、4.2 章で述べた“重視したい特徴”に対し、正規分布による乱数を用いて  $\beta^{(C)}$  を更新する。次に、更新した  $\beta^{(C)}$  と  $\beta^{(C;h)} \in B^{(C)}$  との距離  $d(\beta^{(C)}, \beta^{(C;h)})$  を最小にする元  $\beta$  を一つ決める。その  $\beta$  との距離がある閾値  $T$  以下となる場合、 $\beta^{(C)}$  を  $\beta$  に更新する。以上の処理を Algorithm 2 に記述した。

### 6.2 実験結果

実験結果を表 1 に示す。実験設定の詳細や表の見方を以下に記述する。

- 二つの母集団  $S_1, S_2, S_3$  で実施

---

**Algorithm 2** フィードバック

---

**Input:**  $i_1, \dots, i_M, \beta^{(C)}$ 

```
1: for  $j = 1$  to  $M$  do
2:    $X \sim N(\mu_{i_j}, \sigma_{i_j})$ 
3:    $\beta_{i_j}^{(C)} \leftarrow \beta_{i_j}^{(C)} + X$ 
4: end for
5:  $\beta = \arg \min_{\beta^{(C;h)} \in B^{(C)}} d(\beta^{(C;h)}, \beta^{(C)})$ 
6: if  $d(\beta, \beta^{(C)}) < T$  then
7:    $\beta^{(C)} \leftarrow \beta$ 
8: end if
```

**Output:**  $\beta^{(C)}$ 

---

- 属性は2種類のベクトル  $\mathbf{y} = (y_1, y_2)$  で  
 $y_1$ : 車を持つ/持たない,  
 $y_2$ : 自然が多い場所を希望する/しない  
とした
- 上記の属性に対し、実験は以下二つのクラスタに対する結果  
 $C_1 = \{ \text{車を持つ} \times \text{自然を希望} \}$   
 $C_2 = \{ \text{車を持たない} \times \text{自然を希望} \}$
- 各母集団ごと、一人ずつ順番に実施
- 1回の推薦リストで提示するお勧め地域のは5件
- フィードバックの上限を2回とし、最終的に5件中何件満足する地域が提示されたかを記録

表 1: 実験結果

順番	$S_1$		$S_2$		$S_3$
	$C_1$	$C_2$	$C_1$	$C_2$	$C_1$
1	2	2	4	4	2
2	2	3	5	4	2
3	2	3	-	4	2
4	3	3	-	4	2
5	3	3	-	5	2
6	3	3	-	-	2
7	3	3	-	-	2
8	4	3	-	-	2
9	4	-	-	-	2
10	4	-	-	-	2

母集団  $S_1$  と  $S_2$  では、クラスタ  $C_1, C_2$  の両方でシステムが改善されていく様子が確認できた。  $S_3$  については、満足する地域が提示されているが、システムの変化は確認されなかった。

## 7 まとめと今後の研究

本論文では、移住希望者と移住候補となる地域とのマッチングを実現するシステムを提案した。学習データが少ない点と希望する地域を明確に伝えられない点の二つの課題に対し、利用者のクラスタリングを用いた離散選択モデルと、対話的に推薦する方法によってシステムを構築した。実験により、利用者が増えシステムが学習されていくことで精度が向上していくことを確認した。

### 今後の研究

- クラスタリングの再構築に関し、処理速度の問題や拡張の課題がある。利用者の属性の種類が多くなれば、再構築の候補となる属性の分け方は指数的に増えていくので、処理速度が問題となる。
- 地域の特徴や利用者の属性のベクトルに関し、種類や長さを検討する必要がある。特に、稼働中のシステムに対し、ベクトル長を再検討しなければならないタイミングやそれを測る指標が何であるかを検討したい。ベクトルが長いほど効用関数の精度改善は期待できるが、処理速度が低下したりシステム利用におけるユーザの負荷が高まったりするので、長さにはトレードオフの問題があると考えられる。

### 参考文献

- [1] 稲葉 陽二: ソーシャル・キャピタル入門: 孤立から絆へ, 中央公論新社 (2011)
- [2] 「きっと満足 糸島生活」, <https://itoshimalife.city.itoshima.lg.jp/>, 2016年10月28日アクセス。
- [3] 永井 秀幸, 倉橋 節也: にぎわい施設を考慮した都市動態シミュレーション, 人工知能学会全国大会論文集, Vol. 29, pp. 1-4 (2015)
- [4] 渡辺雄介, 磯崎紘, 杉本徹: 対話型観光地推薦システムにおける頑健性の向上手法, 第73回全国大会講演論文集, No. 1, pp. 339-340 (2011)