

転移学習型 NMF に基づく少数サンプルによる 宮古島産マンゴの官能評価に関する研究

内田 真帆[†] 山下 遥[‡]
上智大学[†] 上智大学[‡]

1. はじめに

マンゴは宮古島において主要な作物の1つであるが、炭疽病による品質劣化が大きな問題となっている。炭疽病は収穫日からの経日により症状を発症し、マンゴの見た目や味が変化するため、顧客の満足度に大きく影響してしまう。対策として温泉水による温湯処理が注目され、その効果に関する評価が行われている。これまで処理方法および経日変化を考慮した官能実験が行われているが、いくつかの問題点が指摘されている。一般的に官能実験における試食会は、十分なデータ収集のために大人数の被験者を集めることが望ましい。しかし現在、新型コロナウイルスの影響で十分な人を集めた食味実験の実現は困難である。そこで本研究では直行配列表を用いた実験計画を被験者の選定へと活用した張の研究[1]に基づき、被験者を削減した実験を行い、その結果を分析する。ここで張の研究[1]では、被験者削減は実現しているものの、得られる評価データ自体は被験者分しかないため、分析に十分な量とは言えない。また実験を行う前に事前アンケートを行っているが、これを活用しきれていないといった問題がある。そこで本研究では、削減した被験者が存在した場合の疑似官能データを生成することで、増加させたデータを分析する。しかし疑似官能データは、あくまで実際には存在しない仮想上のデータであるため、実際に得られたデータに対して小さい重みで分析することが望ましいと考えられる。分析結果の信頼性に関する懸念がある。そこで転移学習型 NMF を活用し、疑似データの影響を緩和したデータ分析を行う。そのとき、実在の被験者と価値観の関係と、架空の被験者を含めた被験者と評価結果の関係が類似していると仮定して、被験者の官能実験データの特徴を学習する。具体的には、架空の被験者を含んだ評価データを活用し、予め学習させた実在の被験者とその被験者の価値観（性別やマンゴが好きかなど）の関係を転移させることで、実際に得られたデータを重点的に学習しつつ、サンプルサイズの不足の問題を解決するためのアプローチを提案する。

2. 従来研究 転移学習型 NMF[2]

転移学習型 NMF[2]とは、知識の転移元である元ドメインでのデータ行列 X_s と転移された知識を活用する目標ドメインでのデータ行列 X_t で、特徴空間を保存しつつデータの構築を実現する。

以下は転移学習の手法の枠組みである。

Step1: 元ドメインにおける $G \times H$ サイズのデータ行列 X_s に NMF を適用して U_s を学習する。特徴量の数を k とする。 U_s, V_s はそれぞれ $G \times k, k \times H$ サイズの行列とし、以下のように分解されるものとする。

$$X_s \approx U_s V_s \quad (1)$$

Step2: 学習した U_s を目標ドメインに転移する。

Step2-1: 特徴行列 U_s を構成する特徴ベクトル u_1, \dots, u_k (U_s の列ベクトル)により、トピックグラフ W_s を構築する。

$$W_s = U_s^T U_s \quad (u_l^T u_l = 1, \forall l = 1, \dots, k) \quad (2)$$

Step2-2: X_t に対して、以下の式を最小化することにより TNT を適用して目標ドメインにおける行列の積 U_t, V_t を学習する。

$$J_1 = \|X_t - U_t V_t\|^2 + \mu \text{tr}(U_t (D_s - W_s) U_t^T) \quad (3)$$

(μ : 第1項と第2項どちらを重視するかのパラメータ, $D_s: W_s$ に対応する次数行列)

J_1 の最小化は、 U_t, V_t を更新し、重み行列 V_t を学習することで実現する。

3. 転移学習型 NMF に基づく少数サンプルによる宮古島産マンゴの官能評価

3.1. 実験計画に基づく被験者の選定

マンゴを官能評価するために、選定された被験者による食味実験を行う。実験は被験者の価値観として、(1)性別、(2)マンゴの嗜好度、(3)加工されたマンゴの嗜好度、(4)週5回以上フルーツを食べるか、(5)食へのこだわりがあるかの5つの属性に着目した。水準の組み合わせが網羅的に表れる8名の被験者(表1)を、属性を問う事前調査により、条件に合う被験者を選出した。

表1 直交表を用いた被験者の属性の設定

被験者	(1)	(2)	(3)	(4)	(5)
①	男	普通	普通	未満	ない
②	男	普通	好き	以上	ある
③	男	好き	普通	以上	ない
④	男	好き	好き	未満	ある
⑤	女	好き	普通	未満	ある
⑥	女	好き	好き	以上	ない
⑦	女	普通	普通	以上	ある
⑧	女	普通	好き	未満	ない

3.2. 実験概要

宮古島マンゴの処理方法と時間経過に対する味の差について以下のような実験を実施した。

無処理のマンゴ・温湯処理を施されたマンゴ(以下、温泉マンゴ)・配送直前に収穫された新鮮な無処理のマンゴ(以下、フレッシュマンゴ)

Study on sensory evaluation of small sample sized mango from Miyakojima based on NMF considering transfer learning

[†]Maho Uchida · Sophia University

[‡]Haruka Yamashita · Sophia University

の3つについて試食による食味実験を行った。経日による味の変化を考慮するために約2週間ごとに計3回行った。具体的には、6月26日、7月10日、7月23日の3日間である。この間、無処理マンゴー・温泉マンゴーは冷蔵保存した。

食味実験の評価では、10種類の評価項目（色、切る前の香り、切った後の香り、甘味、苦味、酸味、滑らかさ、みずみずしさ、総合評価、周囲の人にお勧めしたいか）の5段階評価データを収集した。

3.3. 疑似官能データの生成

本研究では、価値観の同じ被験者はマンゴーに対して類似評価をする、という仮定のもと疑似官能データを生成する。被験者 $i(i = 1, \dots, n, \dots, p)$ が q 個の属性を有すると仮定して、属性値を $Z = \{0,1\}^{p \times q}$ とする。このとき、実在する n 人の被験者の属性値が直交表の要素に対応するものとして $\{z_1, z_2, \dots, z_n\} = Z_0$ とおく。観測されていない系列(架空の $(p - n)$ 人の被験者の属性値)を $\{z_{n+1}, \dots, z_p\} = Z - Z_0$ とする。今回の実験では、3回の実験に対して、マンゴーの種類はそれぞれ3種類とした。実験回数 $a(a = 1, 2, 3)$ 、処理方法 $b(b = 1, 2, 3)$ のマンゴーの対する評価項目 $c(c = 1, \dots, 10)$ を r_{abc} とおくと、 r_{abc} 番目の官能実験の結果 $y_{ir_{abc}}$ の平均値は

$$\sum_{i=1}^n \frac{y_{ir_{abc}}}{n} = \bar{y}_{r_{abc}} \quad (4)$$

で表される。架空の被験者と同じ属性 q を持つ実在の被験者の r_{abc} 番目の官能実験の結果の平均値 $\bar{y}_{qr_{abc}}$ は以下で表される。

$$\begin{aligned} \bar{y}_{qr_{abc}x_{p=1}} &= \frac{\sum_{i=1}^n y_{iqr_{abc}} I_{ip}}{\sum_{i=1}^n I_{ip}}, \\ \bar{y}_{qr_{abc}x_{p=0}} &= \frac{\sum_{i=1}^n y_{iqr_{abc}} (1 - I_{ip})}{\sum_{i=1}^n (1 - I_{ip})} \end{aligned} \quad (5)$$

(I_{ip} : i 番目のデータの p 番目の変数が1であれば1, そうでなければ0のインジケータ変数)したがって、 $z_{i \geq n}$ の系列における予測値は以下で算出される。

$$\sum_{p=1}^p \left\{ \bar{y}_{qr_{abc}x_{p=1}} I_{ip} + \bar{y}_{qr_{abc}x_{p=0}} (1 - I_{ip}) \right\} - (p - 1) \bar{y}_q \quad (6)$$

3.4. 転移学習型 NMF に基づく分析方法

事前調査において観測した実在する被験者とその属性値からなる行列を元ドメイン $X_s = Z_0$ と位置づけて(1)式に基づきNMFを適用し、得られた被験者の特徴ベクトル $U_s \in \mathbb{R}^{n \times k}$ から(2)式で表されるトピックグラフ W_s を作成する。次に、(6)式によって得られる架空の被験者を含めた全被験者とその評価データ $X_t \in \mathbb{R}^{p \times r_{abc}}$ に対して、(3)式を最小化することで、得られた評価項目の特徴ベクトル $U_t \in \mathbb{R}^{p \times k}$ を考察する。

4. 分析結果と考察

表2にある各値は各潜在クラスへの所属度合いを表す。なお、事前分析に基づき特徴量 k を3とした。

表2 被験者の各特徴量に対する所属度

被験者/クラス	1	2	3
①	0.364	0.687	1.011
②	1.360	0.891	1.468
③	1.608	0.835	0.820
④	0.309	1.684	1.114
⑤	0.179	1.256	2.182
⑥	2.531	1.155	0.015
⑦	1.227	0.000	2.090
⑧	1.059	1.090	0.731

それぞれのクラスにおいて所属度の大きい4人に注目して、そのクラスの特徴とする。クラス1は週5回以上フルーツを食べている(4人中4人該当)。クラス2は女性でマンゴーと加工されたマンゴーどちらも好むがフルーツは週5回未満しか食べない人が所属するクラス(いずれも4人中3人該当)である。クラス3は食へのこだわりがある(4人中4人該当)。

次に、評価結果の特徴行列を解釈する。ここでは紙面の都合上、総合評価の各処理方法に関する特徴行列(表3)に限定する。

表3 総合評価に関する各特徴量に対する所属度

クラス	1	2	3
総合評価(無処理)	0.724	0.535	0.622
総合評価(温泉)	0.339	0.644	1.050
総合評価(フレッシュ)	0.351	0.635	1.148

表3より、クラス1は無処理マンゴー、クラス2は温泉マンゴー、クラス3はフレッシュマンゴーを高く評価している。よって、被験者の特徴行列のクラス解釈より、週5回以上フルーツを食べる人は無処理マンゴーを好み、女性でマンゴーと加工されたマンゴーどちらも好むがフルーツは週5回未満しか食べない人は温泉マンゴーを好み、食へのこだわりがある人はフレッシュマンゴーを好むことが分かった。

5. おわりに

本研究では直交表で被験者を選択し、宮古島産マンゴーの官能実験を行った。さらに疑似官能データを生成し、転移学習型NMFに基づき考察した。

今後は、被験者に砂糖を食べさせるなどして数値化した個人の味覚感度を踏まえた分析が検討できる。

参考文献

- [1] ZHANG SIYING: 少人数の官能実験に基づく宮古島産マンゴーの経日劣化の評価に関する研究, 2020年度卒業論文
- [2] 萩野広樹, 吉田哲也, トピックグラフに基づくNMFを用いた転移学習/情報処理学会論文誌数理モデル化と応用 Vol4 No3 73-83(2011)