

身体・天候状況に基づくオリジナルコーヒー提案システムの改善と評価

清水 大聖[†] 中島 毅[‡]芝浦工業大学[†]

1. 研究背景

日本におけるコーヒーの消費量は 1996 年から 2016 年にかけて約 15 万トン増大している。その中で、場所別の喫飲杯数が最も多かったのが家庭であると報告されており [1]、家庭での個人が飲むコーヒーのおいしさが益々追及されるようになってきている。個人にとってのコーヒーのおいしさは、元々個人が持つ嗜好に加え、コーヒー喫飲時の体調や周囲の環境に依存していると考えられる。これらの要因を考慮して、個人がいつでも「最高」と思えるようなオリジナルコーヒーを作ることができる手軽なサポートが求められている。

2. 先行研究とその課題

先行研究は個人の身体データと天候データを基に、ユーザが一番おいしいと感じられるコーヒーのブレンド比率を推薦するシステムを提案している [2]。

図 1 は、先行研究の試作システムの構成である。

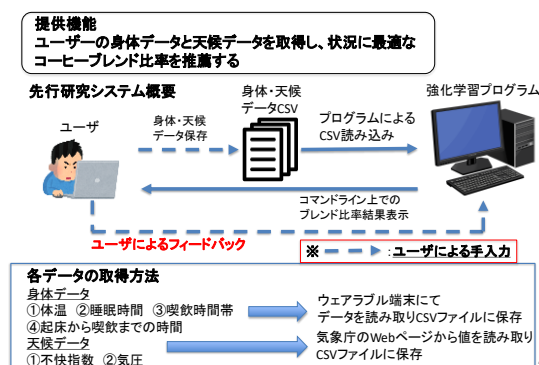


図 1 先行研究の試作システム概要図

このシステムは以下の 3 つのデータを入力する。

- ① 天候状況 (気圧, 不快指数)
- ② 身体データ (体温, 睡眠時間, 喫飲時間帯, 起床から喫飲までの時間)
- ③ ユーザによる各味覚値のフィードバック

システムはこれらのデータを基に Q 学習 [3] により学習し、ユーザに適したコーヒーのブレンド比率を推薦する。Q 学習は強化学習の一つで、実際の報酬と見込みの差分から得られる Q 値 (最適行動価値関数) が最大になるように Q-Table と呼ばれる表に格納し、更新する。これらの試行を繰り返すことで学習が完了する。

先行研究には以下の 3 つの課題がある。

- (A) データ収集の自動化や UI の不便さの改善
- (B) Q 学習の収束状況の評価
- (C) 適切な強化学習法の適用

(B) について、先行研究では Q 学習の実行と共に作成される Q 値の収束に対しての評価ができていなかった。

また (C) について本応用で Q 学習の出力の精度を高めるには Q-Table をすべて埋め、かつ状態数に応じて一定の値に収束している必要がある [3]。これを満たすには多くの試行回数が必要となるため Q 学習に代わる強化学習で収束性の評価を行う。

3. 研究目的と各課題に対するアプローチ

本研究は、先行研究の (A) ~ (C) の課題を解決する事を目的とする。以下が課題に対するアプローチである。

- ① 課題 A に対して：先行研究システムの改善をする。主な改善点として Android アプリを作成し、ブレンド比率取得とフィードバックを効率的に行えるようにする。
- ② 課題 B, C に対して：少ないデータでも良い結果を得ることができる学習法として Deep Q Network を採用し、システムの実装を行う。このシステムと先行研究の Q 学習版について 50 回時点と、100 回時点において有効性の評価を行うことで Q 値収束性について評価を行う。

4. データ収集と UI の改善及びその評価

4.1 実施したシステム改善

課題 A に対応した改善システムの概要図を以下の図 2 に示す。これを基に改善システムの実装を行った。



図 2 改善システムの概要図

図 2 の改善点を以下に示す。

- ① スマートフォンアプリを実装。システムから出力されるブレンドの要求, 結果表示, フィードバックをアプリ上で実行。

②身体情報をウェアラブル端末から自動取得.

③天候データを WebAPI を用いて自動取得.

②, ③のデータを入力としてシステムが Q-Table の状態表からユーザがおいしいと思うブレンド比率を取得する.

4.2 改善システムの機能的評価について

改善システムの機能的評価としてアンケート評価を行った. アンケート評価結果を以下の図 3 に示す.

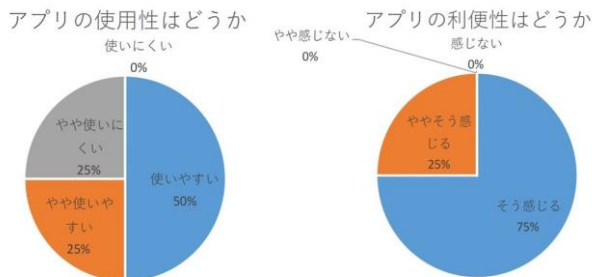


図 3 アンケート結果について：使用性

図 3 の結果からアプリの使用性は「使いやすい」, 「やや使いやすい」合わせて 75%を超える結果となった. また利便性についても「利便性を感じる」が 75%を超える結果となった. アンケート結果より, システムの改善について肯定的な意見が得られたため, 有用性があると感じられた. しかし, 本アプリが文字での結果表示であり, イラストなどで表示した方が分かりやすいのではないかなどの提案があったため, 修正する必要があると感じた.

5. 学習方法の改善と収束性確認実験

5.1 Deep Q Network の適用

Deep Q Network は Q 値をニューラルネットワークで近似することによって学習を行う強化学習法であり [4], Q 学習では困難であった高次元データに対応が可能である. Deep Q Network を用いたシステム実装において以下の 2 つの点を工夫した.

①報酬の設定: 味覚値を用いた学習を行うことで個人の嗜好にあったコーヒーを出力させるため, ユーザが一番おいしいと感じたブレンド比率の味覚値と出力したブレンド比率のそれぞれの味覚値の距離を利用して報酬を決定するように設定した.

②ミニバッチ学習: データ同士の時間的相関を少なくすることで偏りを抑えた学習ができるようにするため, 1Episode ごとに学習するのではなく, メモリに各ステップの内容を保存しておき, 一定の数に達したときにメモリからランダムに内容を取り出していき学習させるよう設定した.

5.2 学習の収束性確認のための実験

目的: Q 学習と Deep Q Network を用いてブレンド比率推薦時に Q 値が試行回数によってどのように収束していくかを確認する.

被験者: 22 歳～23 歳の 男性 3 名

実験手順:

① 初回時, 被験者は 10 種類の豆から一番おいしい

コーヒー豆を選択 (以下メイン豆種). サブ豆種 2 種類はメイン豆種から自動的に決定する.

② 実験時に身体データと天候データを取得し, それぞれのデータを保存する. さらにメイン豆種をベースとしたブレンド率 6 種類のコーヒーをそれぞれ啜飲してもらい, 一番おいしいブレンド比率, 一番まずいブレンド比率を記録 (試行回数 100 回).

③ 50 回と 100 回の試行が終わった時点でシステムがおいしいブレンド比率を出力できているかそれぞれ 10 回ずつ実施.

④ 収集したデータを基に 2 つの学習法に基づいた Q 値収束についての評価をそれぞれ行い比較.

5.3 実験結果

Q 学習と Deep Q Network を使用したシステムの実験結果 (テスト 10 回) を以下の表 1 に示す.

表 1 被験者別正答数 (率)

	被験者 A	被験者 B	被験者 C
Q 学習版	4 回 (40%)	3 回 (30%)	3 回 (30%)
改善システム	5 回 (50%)	6 回 (60%)	5 回 (50%)

表 1 ではシステムが出力したブレンド比率とユーザがその時に一番おいしいと感じたブレンド比率が合致した割合を示している. Q 学習での割合が, 30%～40%, Deep Q Network では 50%～60%となった. 以上の結果から Q 学習を使用したシステムよりも Deep Q Network を使用したシステムの方がおよそ 20%～30% 正答率が高く, 改善の効果を確認できた.

6. まとめと今後の課題

ユーザの身体状況・天候状況のデータを収集し, それらのデータからユーザが一番おいしいと感じるブレンド比率を算出するシステムと UI 部分の改善・評価を行った.

また Q 学習と Deep Q Network の 2 つの強化学習を用いて比較評価を行った結果, Deep Q Network を使用した方が正答率が高い結果となった. 実験により Deep Q Network の損失関数が減少傾向であったため, どちらの学習法も試行回数をさらに増やし, 収束性の評価を行いたいと考える.

参考文献

[1] 全日本コーヒー協会統計資料:

<http://coffee.aica.or.jp/data>

[2] 濱田: 身体・天候状況に基づくオリジナルコーヒー提案システム, 芝浦工業大学工学部情報工学科, 2018 年度卒業論文

[3] ゼロから Deep まで学ぶ強化学習:

<https://qiita.com/icoxfog417/items/242439ecd1a477ece312>

[4] DeepMind Technologies (2013) Playing Atari with Deep Reinforcement Learning

[5] ReNom 損失関数の紹介:

https://www.renom.jp/ja/notebooks/tutorial/basic_algorithm/lossfunction/notebook.html