

新規性について

水上 和秀

富山県立大学 電子・情報工学科

July 27, 2022

プログラムの現状と課題

近年、生活習慣病を患う人々が増加している。生活習慣病とは「食習慣、運動習慣、休養、喫煙、飲酒、ストレスなどの生活習慣を原因として発症する疾患の総称」のことであり、深刻な疾患に深く関与している。その要因には、偏った栄養による食生活や運動不足、飲酒、喫煙などが挙げられる。これを防ぐためには、過度な喫煙や飲酒を控えることや、しっかりと栄養を食事から摂取することが重要視されている。また、学校給食や病院食の現場で献立作成業務を担当している栄養士は、摂るべき栄養素の計算や、食事にかかる金額などを設定する作業を、繰り返し見直す必要があるため、献立作成の負荷は高いことがわかる。



図1 生活習慣病を起因とする疾患

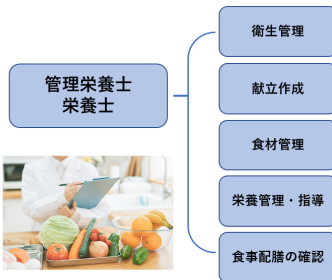


図2 栄養士の主な業務内容

2.2 Web 上のレシピデータを活用

3/7

おさらい
おさらい

提示するレシピはレシピサイト「ボブとアンジー」から、主菜と副菜カテゴリに含まれる料理と、朝食でキーワード検索した料理レシピデータ（必要材料、摂取栄養量、カロリーなど）をスクレイピングし、参照する。食品価格動向を調査しているサイト「小売物価統計調査による価格調査」から様々な食品とその価格データをスクレイピングする。次に、料理レシピデータの食材と食材価格データの食材を照らし合わせて食材コストを計算する。



図7 レシピサイト・ボブとアンジーにおける料理レシピ情報

スクレイピングする主なデータ

- ・料理レシピ名
- ・調理時間
- ・摂取カロリー
- ・摂取栄養量
- ・必要食材名
- ・必要食材量
- ・作り方
- ・画像URL
- ・食材価格
- ・販売単位
- ・食材名

小売物価統計調査による価格推移

HOME / キャベツの価格推移・価格推移 / 全国

全国のキャベツ 1 k g
価格推移 / 過去84ヵ月



WEBスクレイピング

データ抽出・出力

全国のスーパーで定まっているキャベツ 1 k g 単位の平均は124円。
2015年1月～2021年12月(過去84ヵ月)の期間で全国内のキャベツが最も高かった最安値月(2018年2月)で394円。逆に最もキャベツが安かった最安値月は2023年2月で180円となっています。
全国エリアでキャベツの最安値月(2018年2月)と最安値月(2023年2月)との価格差は276.4321円となっています。
キャベツ 1 k g の2015年1月～2021年12月の価格推移とグラフは下記をご覧ください。
出典元：総務省統計局 小売物価統計調査(2021年12月)
関連新聞「キャベツのよきと新報(3月17日)」

図8 食品価格推移調査サイトの例



図9 Webデータ活用の流れ

プログラムの現状と課題

安藤さんの卒論では 6438 個のレシピデータから献立を作成していたので、プログラムを起動してから出力まで pc1 台の場合 5 時間 34 分、pc4 台で並列処理を行っても 2 時間 34 分かかっていた。これでは時間がかかりすぎているのでプログラムの高速化についてより効率的な手法を取り入れる必要がある。

新規性

プログラムを改良し、高速化を図る
レシピの追加、嫌いな食材を含むレシピを表示しないようにする

レシピの追加について

- eatSMART に軽食が乗っていたので csv に変換するプログラムを作った
- 食材コストを計算することができていないため、レシピを追加しても出力されない。
→ difflib について調べてコストを計算できるようにする

嫌いな食材を含むレシピを表示しないようにすることについて
まだ実装に至ってません

並列処理の方法の変更について

- ・安藤さんの卒論では Dask というライブラリを使って並列処理を行っていた。しかし大量のデータを扱う際、dask よりも Spark を使ったほうが処理速度が速い
→ dask から spark に置き換えるのに時間がかかっており、まだ実装できていない

まとめ

- ・プログラム高速化する方向で研究を進める
- ・軽食のレシピのプログラムを作った

これからやること

- ・とりあえず安藤さんのプログラムの理解することを進める。
- ・軽食を表示できるようにする
- ・数理最適化の論文について調べ、他に数理最適化を活かせる分野があれば今のプログラムを応用できないか考える
- ・大学院試験の勉強を引き続きする