

# 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 5, 2022

## 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

## レシピを減らす

### 1. 類似度の高いレシピを減らす

→ python で、文字列の類似度を比較する difflib というライブラリがある。それを使って類似度の高いレシピを削除するプログラムを作る。

### 2. 嫌いな食材が入っているレシピを取り除く

→使用者に取り除く食材を入力してもらい、その食材が含まれるレシピを削除する。python には文字列を検索することができるのでそれを使う。

## プログラムを改良する

並列処理の方法を変える

→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

## まとめ

- ・プログラム高速化する方向で研究を進める
- ・説明した提案手法は調べただけなのでまだ実装に至っていない

## これからやること

- ・python の知識が不足していたので、python についての学習
- ・レシピサイトのブラッシュアップ
- ・説明したプログラムの作成
- ・ほかにプログラムで改善できるところを考える
- ・ほかの並列処理について調べ、処理速度を改善できないか考える