

ILP を用いた体脂肪量の減少につながる生活習慣規則の抽出

牛窪翔† 金盛克俊† 大和田勇人†

†東京理科大学 理工学部 経営工学科

1. はじめに

生活習慣病は日本では重大な問題として広く知られており、患者は若者にまで及ぶと言われている。食事、運動、睡眠と肥満が生活習慣病に大きく影響を与える要因と言われており、生活習慣病を予防するためには適切な体脂肪量を維持する必要がある。しかし、体脂肪量の増減に寄与する要因は完全には明らかにされていない。そこで、本研究では生活習慣に関する時系列データから帰納論理プログラミング(ILP)を用いて体脂肪量の減少につながる生活習慣規則を抽出する。ILP は述語論理を用いた機械学習手法の 1 つであり、他の機械学習手法と比較して背景知識に複数の表が考慮でき、様々な特徴量間の関係を記述できること、また学習の結果から抽出された規則がどのようなルールの基で生成されたか簡単に理解することが出来るメリットを持つ。そのため本研究では ILP を用いる。

2. データセット

本研究では 2015 年 9 月 17 日から 10 月 31 日まで男性 5 人の被験者を対象として生活習慣に関する時系列データを収集した。各被験者におけるデータ数は表 1 の通りである。

表 1. 各被験者におけるデータ数

Subject id	The number of data
1	42
2	44
3	45
4	31
5	45

生活習慣に関する特徴量は毎日の睡眠に関する特徴量(睡眠時間, 睡眠の質), 運動に関する特徴量(消費カロリー), 食事に関する特徴量(摂取カロリー, 摂取ビタミン及びミネラル等を含む計 31 個)と毎日の体脂肪量が含まれている。データの形式は図 1 の通りである。

Subject id	Date	Body fat mass	Sleeping hours	Sleep quality	Bedtime	Stress	Calorie consumption	Calories	...	Salt equivalent
1	9/17	11.4	7	0	29	0	2143	1313	...	7.92
2	9/17	6.95	7	1	25	-1	2344	1194	...	10.2
⋮										
5	10/31	11.3	4	1	27	0	1907	1706	...	8.14

31 type of features about nutrition intake

図 1. 生活習慣に関する時系列データの概要

3. データ前処理

本研究では前述のとおり ILP を用いる。しかし、ILP では連続値の取り扱いが難しい。そこで、データを連続値から離散値もしくはカテゴリに変換する必要がある。本研究では以下のステップで変換を行った。

- Step 1: ΔB を前日との体脂肪量の増減と定義し、体脂肪量に関して差分を取る。
- Step 2: ΔB が 0 より大きい(体脂肪量の増加)場合、値を 0 に変換する。 ΔB が 0 以下(体脂肪量の減少)の場合、値を 1 に変換する。
- Step 3: 睡眠, 運動, 食事に関する連続値を含む特徴量は平均値と標準偏差に基づいて、以下のカテゴリに分類する。

- ave ($|\text{値}-\text{平均値}| < \text{標準偏差} \times 0.2$)
- low ($|\text{値}-\text{平均値}| < \text{標準偏差}$)
- mid ($|\text{値}-\text{平均値}| < \text{標準偏差} \times 2.0$)
- high ($|\text{値}-\text{平均値}| \geq \text{標準偏差} \times 2.0$)

以上の処理を行うことにより、表 2 のように連続値のないデータとなる。

表 2. データ処理を施した時系列データの概要

Subject id	Date	ΔB	Sleeping hours	Sleep quality	Bedtime	Stress	Calorie consumption	Calories	...	Salt equivalent
1	9/18	1	ave	0	low	0	low	low	...	ave
2	9/18	0	ave	1	low	-1	ave	low	...	low
⋮										
5	10/31	1	low	1	mid	0	mid	mid	...	low

4. 提案手法

本研究では ILP システムとして GKS[1]を用いる。GKS は背景知識データ, 正事例データ, 負事例データの 3 つのデータを入力とすることで、学習結果として正事例に多く共通するルールを出力

することが出来る。

背景知識に取り入れる述語は表2のデータを用いて、睡眠時間を例として次のように表す。カッコ内は引数を表す。

Sleeping_hours(Id_and_date, Day_before, Val)

第1引数は被験者とその日時を表し、第2引数はその日時から何日前を参照することを表し、時系列に関する情報を背景知識に追加している。第3引数はその日時から何日前かの睡眠時間の離散値もしくはカテゴリを表す。他の特徴量についても同様に定義し、背景知識に取り入れる。ただし、本研究では背景知識に取り入れる特徴量を以下の表3のように組み合わせて実験を行った。

表3. 背景知識に取り入れる述語の組み合わせ

A set of predicates	Description
Sleeping hours, Sleep quality, Bedtime, Stress, Calorie consumption, Protein, Fat, Carbohydrate	Consider sleep, exercise and three major nutrients
Vitamin D, Vitamin E, Vitamin B1, Vitamin B2, Niacin, Vitamin B6, Vitamin B12, Pantothenic acid, Biotin	Consider vitamins
Potassium, Calcium, Magnesium, Phosphorus, Iron, Zinc, Manganese, Iodine, Selenium, Chromium, Molybdenum	Consider minerals
Sleeping hours, Sleep quality, Bedtime, Calorie consumption, Protein, Fat, Carbohydrate, Calcium, Vitamin D	Consider vitamin D and calcium and on the basis of related work

正事例及び負事例については、 ΔB が1となる場合を正事例、0を不事例とし、5人の被験者あわせて正事例109、不事例85のデータとなる。

5. 結果

提案手法による予測精度を表4に示す。

表4. 各述語の組み合わせにおける結果

TP	FP	FN	TN	Accuracy	Precision	Recall	F1 score
109	27	0	58	0.861	0.801	1.0	0.890
105	29	4	56	0.830	0.784	0.963	0.864
108	31	1	54	0.835	0.777	0.991	0.871
109	28	0	57	0.856	0.796	1.0	0.886

表3中のTP, FP, FN, TNはTrue Positive, False Positive, False Negative, True Negativeを表す。本研究では睡眠、運動と三大栄養素に関する述語からなる背景知識を用いた場合、F値が一番高い結果が得られた。これらから生成されたルールの中で、最もスコアの高い生活習慣規則を示す。ルールのスコア(Score)は(1)のようにそのルールが導く正事例の数(T)から負事例の数(F)及び仮説を構成するリテラル数(C)を引いた値となる。

$$\text{Score} = T - (F + C) \tag{1}$$

このスコアが最も高かったルールを以下に示す。

- 3日前に十分な睡眠を取っており、昨日の炭水化物摂取量が少ない
(正事例18, 負事例1, リテラル数2)
- 3日前及び昨日のタンパク質摂取量が高く、2日前の脂肪摂取量が少ない。
(正事例18, 負事例0, リテラル数3)

この結果から、睡眠時間と炭水化物摂取量の組み合わせもしくはタンパク質摂取の継続と脂肪摂取量が体脂肪量の減少に関係があることがわかった。睡眠時間の不足は体脂肪量の増加をもたらすこと、また炭水化物を少なく摂取することは体脂肪量の減少に効果があることが報告されている[2][3]。また継続してタンパク質を摂取することは体重の減少をもたらすこと、また脂肪量を少なく摂取することは体重の減少に効果があることが報告されている[4]。そのため、本研究で抽出されたルールは妥当性があると考えられる。

6. まとめと今後の課題

本研究ではILPを用いて体脂肪量の減少につながる生活習慣規則の抽出を行った。その結果、睡眠と炭水化物量の減少、またタンパク質摂取の継続と脂肪摂取量の減少が体脂肪量減少に影響があることが示された。今後の課題として、被験者数及びデータ数を拡張し、体脂肪量の減少に寄与する新たな要因を抽出することが必要であると考えられる。

参考文献

[1] Hayato Ohwada, Hiroyuki Nishiyama, Fumio Mizoguchi, "Concurrent Execution of Optimal Hypothesis Search for Inverse Entailment," Lecture Notes in Artificial Intelligence, vol. 1866, no. 4, pp.165-173, Aug 2000.

[2] Hassan. S. Dashti, Frank. AJL. Scheer, Paul. F. Jacques, Stefania. Lamon-Fava, "Short Sleep Duration and Dietary Intake: Epidemiologic Evidence, Mechanisms, and Health Implications," Advances in Nutrition, vol. 6, pp. 648-659, Nov 2015.

[3] Nordmann. AJ, Nordmann. A, Briel. M, Keller. U, Yancy. WS Jr, Brehm. BJ, Bucher. HC, "Effects of low-carbohydrate vs low-fat diets on weight loss and cardiovascular risk factors: a meta-analysis of randomized controlled trials," Arch Intern Med, vol. 166, no. 3, pp. 285-293, Feb 2006.

[4] Damien. P. Belobrajdic, Jan. Frystyk, Nilani Jeyaratnaganathan, Ulrick. Espelund, Allan. Flyvbjerg, Peter. M. Clifton and Manny. Noakes, "Moderate energy restriction-induced weight loss affects circulating IGF levels independent of dietary composition," European Journal of Endocrinology, vol. 162, no. 6, pp. 1075-1082, Jun 2010.