

睡眠日誌アプリを利用した将来の元気度の予測 Cheerful status prediction using “Sleep Diary” application

秋富 穰^{*1}
Jou Akitomi

梶山 征央^{*1}
Ikuko Kajiyama

岡島 義^{*2}
Isa Okajima

山口 美峰子^{*1}
Mineko Yamaguchi

^{*1} NEC ソリューションイノベータ株式会社
NEC Solution Innovators, Ltd.

^{*2} 東京家政大学
Tokyo Kasei University

Recently, the problem of insomnia is a social issue. To solve this problem, we develop “Sleep Diary” application depend on the theory of cognitive behavioral therapy for insomnia: CBT-I. It is suggested that users continuously use this application may improve their sleep habituation. However, the continuation of behavior change is difficult for many people. So, we consider that it is required extrinsic motivation method to continuously use the application. In this study, we focused on the cheerful status prediction function as extrinsic motivation method. As a result, our method using random forest classified cheerful status with practical accuracy. Particularly effective features were chosen from average and standard deviation of 7 days sleep data. We will confirm effectiveness of this method.

1. はじめに

睡眠の問題を抱えると、精神面や身体面の両面に悪影響を与えることが知られており、疲労感や集中力の低下による産業事故のリスクを増大させ、生産性の低下など、社会的な問題を引き起こすと言われている[厚生労働省 2014].

睡眠改善の手法の一つとしては、不眠のための認知行動療法 (Cognitive Behavioral Therapy for Insomnia: CBT-I) があり、これまでにその有効性が示されている[Okajima 2011]. 我々はこの CBT-I 技法の一部である睡眠日誌の記録と睡眠衛生についての習慣や知識を提供するためのアプリである「睡眠日誌」アプリを試作開発した。ユーザは、睡眠日誌アプリの利用により、日々の睡眠を記録し、良い睡眠習慣について学び、日々の睡眠の状態を振り返る。ユーザがこのアプリの利用を継続することができれば、ユーザの睡眠習慣が改善されることが期待される。実際に、本アプリの長期利用継続者においては、睡眠の改善効果があるだけでなく、睡眠に関係するその他の生活習慣である夜間のカフェインの摂取や飲酒などについても、摂取頻度が低下するという改善効果があることが示唆されている[石井 2018].

このようなユーザの生活習慣の改善のための行動を如何に継続させるかということは、ユーザの行動変容を考える上で最も大きな課題の一つである。この課題に対する最初のアプローチとして、我々はまず本アプリのユーザが利用継続するか否かを予測する方法の検討を行った。この検討の結果、ユーザが入力した最初の数日間の睡眠データを用いることで、数週間後にそのユーザがアプリの利用を継続しているか否かが予測可能であることが示唆された [秋富 2018]。そこで本研究では、次のステップであるユーザへのアプリ利用の動機付けの方法として、睡眠データによって将来のパフォーマンス (元気度) を予測することに注目した。この機能が実現すれば、アプリの利用の継続により睡眠のデータが蓄積されることで、そのユーザにとって将来の元気度が分かるというメリットが発生するため、アプリの利用の継続を促すための手段と成りえると期待される。

本稿では、これまでに蓄積されている睡眠日誌アプリのデータから、ユーザの翌日の元気度を予測する方法について検討を行った。

2. 元気度

睡眠日誌アプリでは、記録するデータが大きく二つに分けられる。一つは、朝起きたタイミングに記録を推奨している前日の就寝時から当日の起床時までの睡眠に関するデータ (入床時刻、離床時刻など) や、それに付随する生活習慣に関するデータ (夜にカフェインを摂取した、夜に飲酒をしたなど) である。もう一つは、日中の活動が終わった後に一日を振り返るタイミングでの記録を推奨している、その日の元気度のデータである。睡眠日誌アプリのユーザは、この二つのデータを参照することにより、日中の元気度が高くなるための良い睡眠習慣や、自分の元気度が低くなる悪い睡眠習慣について気づきを得ることが出来る。

睡眠日誌アプリでは、ユーザの主観に基づくデータの記録を重視しており、元気度については、アプリから予め用意されている5段階の中からユーザが任意の段階を手入力することで記録される。ここで元気度と呼んでいる値は、基本的には主に眠りの問題において日中の活動に支障があった程度が記録されることを想定しているが、入力内容において特に厳格なルールは定めていないため、ユーザによっては単に疲労の度合いを記録していたり、仕事の生産性のようなものを記録していたりすることがあると考えられる。本稿においては、問題を単純化するため、ユーザの意図は考慮せず、単に元気度として取り扱った。

3. 将来の元気度の予測

本研究では、機械学習の手法の一つであるランダムフォレストを用いて、睡眠データから将来の元気度の予測を試みた。具体的な問題設定としては、あるユーザが記録した一週間分の睡眠データを用いて特徴量を設計し、最後に睡眠データを記録した日の翌日の元気度を予測することとした。

連絡先: 秋富 穰, NEC ソリューションイノベータ株式会社,
東京都江東区新木場 1-18-7, 03-5534-2619,
jou-akitomi@vt.jp.nec.com

なお、睡眠日誌アプリは利用を開始する際にユーザからデータの研究利用についての同意を得ており、データは全て個人を特定できない形で処理した上で使用した。

3.1 使用したデータ

予測に用いたデータは、全て睡眠日誌アプリの入力データから抽出した。抽出したデータ項目は、大きく分類すると、個人属性に関するデータ（年代、性別、主訴の種類など）と、睡眠そのものに関するデータ（睡眠時間、入床時刻、入眠時刻、覚醒時刻、離床時刻など）と、睡眠に関係する生活習慣のデータ（夜にカフェインを摂取したか否か、夜に飲酒をしたか否かなど）と、日中の元気度のデータの四種類である。これらのデータは、全てユーザごと、日ごとに紐付いているが、データの入力には任意であるため、一定の割合で欠損データが存在する。本研究では、予測に用いるユーザの質をある程度一定に保つため、60日以上利用を継続しているユーザの中からデータの欠損が少ない737名（男性370名、女性367名）に限定し、その7日目から60日目までのデータについて抽出を行った。

3.2 実験内容

抽出したデータについて連続7日分を1セットとして扱い、各データの特定期間の平均、標準偏差などを計算することにより特徴量を設計し、最終日の翌日の元気度に基づいてBad, Normal, Goodの3種類のラベルを付加した。ラベルは、元気度1~2点をBad, 3点をNormal, 4~5点をGoodとした。その結果、ユーザの年代別の元気度のラベルの割合は以下の表1の通りであった。

表1. 年代別の元気度の割合 (%)

年代	Bad	Normal	Good
10	23.6	37.6	38.8
20	26.6	42.1	31.3
30	25.4	43.1	31.5
40	21.6	41.6	36.9
50	22.2	41.1	36.7
60	12.9	36.3	50.8
70	5.5	18.8	75.6

表1の通り、抽出された元データはどの年代においてもBadの割合が少ない不均衡データであったため、SMOTE(Synthetic Minority Over-sampling Technique)アルゴリズムによってデータ数の偏りの調整を行った後にランダムフォレストの入力データとした。入力データのうち、80%を学習データとして用い、残り20%を予測データとして用いた。

4. 結果

作成した入力データを用いて、各データセットの元気度がBad, Normal, Goodのどれであるかの予測を行った。予測結果を以下の表2に示す。

表2. 元気度の予測結果

ラベル	正解率	適合率	再現率	F 値
Bad	0.887	0.880	0.771	0.822
Normal	0.864	0.778	0.859	0.817
Good	0.917	0.859	0.873	0.866

また、予測に用いられた特徴量の重要度を、ジニ係数を用いて算出した。この算出された重要度を基に、データのカテゴリ別に、どのカテゴリの特徴量の重要度が高かったかを調査した。カテゴリは、個人属性データ（個人属性）、7日間データ（7日間）、3日間データ（3日間）、2日間データ（2日間）の4カテゴリに分類した。その結果を以下の表3に示す。

表3. 特徴量の重要度順位（カテゴリ別） (%)

重要度順位	個人属性	7日間	3日間	2日間
~20	44.4	37.1	8.6	0.0
~40	0.0	20.0	20.0	17.1
~60	0.0	14.3	25.7	17.1
~80	0.0	5.7	22.9	28.6
~100	55.6	17.1	5.7	20.0
~120	0.0	5.7	17.1	17.1

5. 考察

表2によるとBadの再現率とNormalの適合率が他よりも若干低い傾向にあったが、正解率は全体的に高い傾向にあった。今回の問題設定にて予測した元気度は特に直接的に病気に係るなど深刻な健康データであるとは言えないため、今回の試みにおける予測精度は実用的には十分である可能性が高い。

また、表3によると、今回の予測結果において重要な特徴量は、個人属性データや7日間データのカテゴリに多く含まれているということが示唆された。時系列の予測問題において多くは予測するタイミングの直近のデータの重要度が高くなる傾向があるため、今回の結果はユーザの睡眠の状態とそれに関連する元気度という課題設定において特に特徴的な結果であると考えられる。ユーザの元気度により強く影響を与えるのが直近の睡眠の状態ではなく、それより前の7日間の睡眠の状態であると示唆されたことは幾つかの睡眠研究の知見と合致しており、非常に興味深い。

6. おわりに

本稿では、ユーザの健康的な生活習慣の継続支援のための介入方法の一つとして、蓄積された睡眠データを用いた将来の元気度の予測という課題に取り組んだ。将来の元気度の予測という課題に関しては、予測方法や特徴量の設計方法などに未だ改善の余地はあるが、現時点においても十分に実用的と考えられる予測精度を達成することが出来た。今後は実際に予測システムを睡眠日誌アプリに実装し、継続率が向上するか否かの評価をすることが重要であると考えられる。

また、今回取り組んだ将来の元気度の予測という課題には、ユーザの継続支援の他にも様々な応用の可能性があると考えられる。今後は予測精度の向上に取り組むと共に、元気度の予測を応用した新たな課題への取り組みを進めていきたい。

謝辞

本研究について大きな貢献をしていただいた NEC ソリューションイノベータの村上紘士さんをはじめ、本研究に協力していただいた皆様に感謝をいたします。

参考文献

- [厚生労働省 2014] 厚生労働省健康局：健康づくりのための睡眠指針 2014, 2014.
- [Okajima 2011] Isa Okajima, Yoko Kodama and Yuichi Inoue: A meta-analysis on the treatment effectiveness of cognitive behavioral therapy for primary insomnia, *Sleep and Biological Rhythms*, Japanese Society of Sleep Research, 2010.
- [石井 2018] 石井美穂, 秋富穰, 梶山征央, 岡島義, 山口美峰子: 職域での「睡眠日誌」アプリケーションの利用可能性の検討, 産業衛生学雑誌 (産業医学) (60) 382, 2018.
- [秋富 2018] 秋富穰, 梶山征央, 石井美穂, 岡島義, 山口美峰子: 睡眠日誌アプリの開発と睡眠習慣改善行動の継続の予測, 日本人工知能学会全国大会 (32 回) 講演論文集 3Pin1-52, 2018.