

学生のレポートで使われる単語と授業中の集中度合いの関係の解析 Analysis of the relationship between words used in students reports and the level of concentration in class

上條 浩一[†] 塩尻 (齋藤) 亜希[†] 神沼 充伸[†]

Koichi Kamijo Aki Shiojiri(Saito) Atsunobu Kaminuma

1. はじめに

大学などの集団授業では、全ての学生の集中力を持続させ、講義内容の理解を促すことが教師にとって極めて重要である。しかし、学生が授業を通して集中力を保つことは容易ではない。我々の先行研究では、60 分の授業中に 5 分程度の短いテストを 2 回実施した場合、多くの学生の集中度はテスト中に比べてテスト後の講義で上昇することが観測された[1][2]。図 1 にその例を示す。ただし、図 1 が示すとおり、すべての学生に同様の傾向が見られるわけではなく、テスト後に集中度が低下する学生も存在した。

このような集中度の個人差を事前に推定することができれば、集団授業であったとしても学生ひとりひとりに合わせた学習支援や、授業構成の工夫、難易度設定の調整など、より効果的な授業設計につなげることが可能になる。

こうした個々の習熟度や認知特性の違いを把握する手がかりとして、授業後に提出されるようなレポートが有力な情報源であると考えられる。レポートに記述される内容や用語の選択は、学生の理解度や興味・注意の傾向を反映している可能性がある。

そこで本研究では、学生が書いたレポートをもとに、学生個々の授業中における集中度の変化を推定する AI モデルの構築とその評価を行った。また、集中度と、レポートで用いられる単語との関連性についても解析を行った。

2. 先行研究

Acheampong らは、Transformer, GPT, Cross-lingual Language Models (XLM)[3], および BERT によって、テキストから感情を推定している。Kamijo らは、テキストから人の性格を推定する研究を行い、0~1 で正規化した BIG5 (OCEAN) の値を MAE=0.1 程度で推定することに成功している[4][5]。また、特許の文章を分析することで、その特許が成立するかどうか[6]、対象となる医薬品が承認さ

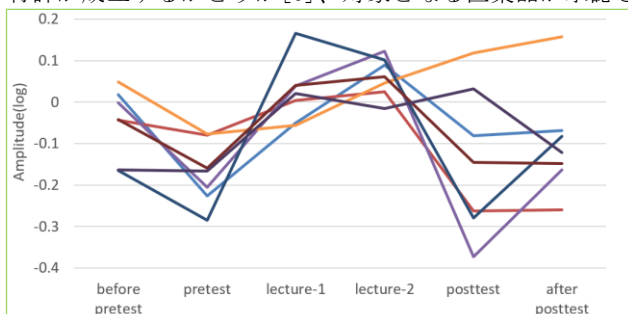


図 1 授業の各ステージ (x 軸) における集中度 (y 軸) 推移例

表 1 実験授業 (補講) の時間配分

Duration (min)	5	5	20	20	5	5
Stage	before pretest	pretest	lecture-1	lecture-2	posttest	after posttest

れるかどうか[7]、その医薬品が将来どのくらい売れるか[8]、の解析も行われている。例えば、[6]においては、米国で開発された医薬品の場合、F1 スコア=0.989 で、特許の成立不成立の推定が可能であった。これらの知見も踏まえ、本研究では集中度の推定にテキストを用いている。

脳波情報を使って集中レベルを測定する研究もいくつか存在する[9-11]。Lokare らは、学生が異なる集中レベルを必要とするさまざまな課題をこなしている間に EEG 装置で記録した学生の脳信号を使い、集中レベルを分類した[12]。Kamijo らは、授業中に学生に装着した脳波計より取得できる α 波、 β 波の変化を測定し、教師の行動と学生の集中度の関係の解析を行っている[1][2]。

このように、テキストを用いた推定や授業中の脳波情報を用いた研究は行われているが、テキスト情報から集中度の変化を推定する研究は見当たらない。

3. 研究方法

3.1 実験 (補講授業) の実施

実験は、東京国際工科専門職大学情報工学科の Python プログラミング授業・演習の補講として行われた。補講は、1 回 60 分で、3 日間行われ、同学科の 2 年生 121 名中希望者 38 名が参加した。各日において、朝 (11-12 時)、昼 (13-14 時)、夜 (18-19 時) の 3 回に分けて実施した。各日における授業の内容は同じため、参加者は、朝、昼、夜のうち都合がよい 1 回 (3 日で 3 回) を受講した。

3.2 授業の構成

本授業の内容は、1 年生で既に学習している Python の補講である。各授業では、最初に、5 分間の簡単な復習 (before pretest) が行われた。その後、5 分間の pretest と 40 分間の lecture (主に pretest の解説) が行われた。Lecture は、解析の便宜上、前半 20 分 (lecture-1) と後半 20 分 (lecture-2) に分けた。さらに、学生が lecture の内容を理解したかどうかを確認するため、5 分間の posttest が行われ、最後に wrap-up (after posttest) が行われた (表 1)。

[†] 東京国際工科専門職大学
International Professional University of Technology in Tokyo

表 2 5つの基本脳波の特徴

名称 (波)	周波数 (Hz)	状態
Gamma (γ)	>35	興奮
Beta (β)	12-35	集中, 緊張
Alpha (α)	8-12	リラックス, やや集中
Theta (θ)	4- 8	ウトウト
Delta (δ)	0.5- 4	睡眠

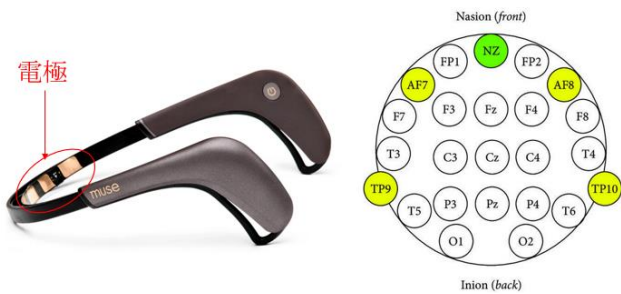


図 2 Muse2 (左) および Muse2 の EEG sensors (右, right, TP9, TP10, AF7, AF8)

3.3 集中度測定方法

表 2 は、5 つの基本的な脳波の特徴を示している [13]。Yoshita らは、 α 波と β 波の振幅の比 (β/α) が、人間の集中レベルを観察するための効果的な方法であることを示しており [14]、今回の実験では、 β/α を集中度として使用した。

今回用意できる脳波計は 3 台であり、各授業で最大 3 人の有志学生が脳波計を装着した。脳波の測定には、Muse2 (Interaxon 社製) [15] (図 2 左) を使用した。Muse2 は、10-20 方式 [16] に準拠しており、複数の電極 (TP9, TP10, AF7, AF8, 図 2 右) を通して、脳波を所得する。

脳波計で脳波を測定する各学生には、授業開始前に、脳波測定アプリの Mind Monitor application [17] がインストールされたスマートフォン (Google Pixel 4a) が渡された。Muse2 は 300Hz で脳波をサンプリングし、Bluetooth 経由で、スマートフォン内の Mind Monitor application に脳波データを送信した。さらに、Mind Monitor application は、取得した脳波データを、Wi-Fi 経由でクラウドデータベースにアップロードした。

3.4 脳波データ前処理

3.3 節で取得した脳波の測定結果は、以下のように処理された：まず、300Hz でサンプリングしたデータを 75Hz にダウンサンプリングした。次に、帯域幅 5~37Hz のバンドパスフィルタを適用した。その後、1 秒間隔で 4 秒ウィンドウ内に記録された各電極の μV (マイクロボルト) 単位の脳波データの出力値の標準偏差を計算した。標準偏差が 20 より大きい電極データはノイズとみなされ、無効なデータとした [18]。さらに、1 時間の授業で、無効なデータを除去した後に残った有効データが 30 分未満であった場合も、その電極の 1 時間分のデータは全て無効とした。そして、各学生特有のベースライン情報を除去するため、各電極の

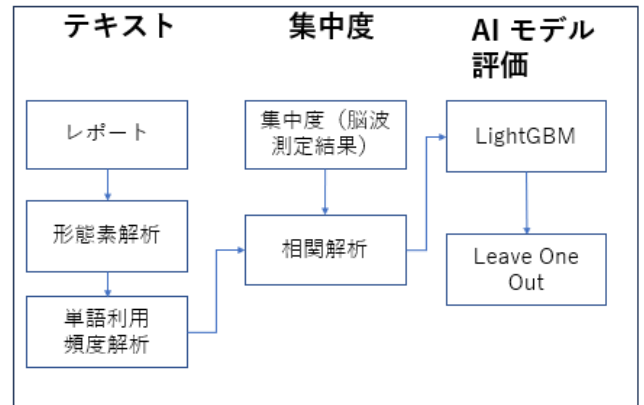


図 3 集中度推定 AI モデル

有効データの平均値を計算し、各データからその平均値を引く処理 (baseline removal) を行った。

4. 集中度推定モデル

図 3 に AI モデルの概要を示す。各補講で 3 人の学生の脳波を測定したが、脳波が正しく測定でき、かつ、レポートを本研究で使える量を提出している学生数 (以下、 N) は 11 人であった。この学生 11 人のレポートから得られる単語利用頻度を AI モデルの説明変数とし、各授業フェーズでの集中度の推定を行う。AI アルゴリズムとしては、LightGBM を用いた。LightGBM は、先行研究 [6-8] において推定精度が高く、かつ計算コストが低いという利点がある。

この実験では、学習に用いるデータサイズが 11 人分と少ないため、深層学習アルゴリズムは利用しなかった。AI モデルの評価は、Leave-One-Out (LOO) により行った。具体的には、10 人のレポートから算出されるデータを説明変数とし、残り 1 人分の集中度を推定する、という操作を 11 人分繰り返し、平均値を評価結果とした。

レポートを説明変数に変換するために、まず、各学生のレポートに対して形態素解析を行い、その利用頻度の解析を行った。具体的には、学生 i のレポートについて、形態素解析後の単語 w の利用頻度を学生 i のレポートに含まれる全単語数で割ったものを、 u_w^i とした。ここで、 $\sum_w u_w^i = 1$ が成立する。日本語のステミングやレマタイゼーションを含む形態素解析には、MeCab を使用した。 w においては、集中度推定に寄与するものであれば何でもよいと考え、ストップワードも含めた。推定に寄与しない単語は、後述の手順でふるい落とされる。また、 w は、全学生のレポートを通じて 2 回以上使用された単語に限定した。

次に、各単語 w の N ($=11$) 人のユーザ毎の利用頻度ベクトル u_w を以下の様に定義した。

$$u_w = [u_w^1, \dots, u_w^N]^T \quad (1)$$

ここで、 T は転置である。次に、 y_s^i を、学生 i 授業のステージ s (例: $s = \text{pretest}$) における集中度として、 u_w^i と y_s^i の間のピアソンの r 値、 r_{ws}^i を求めた。ただし、

$$\begin{aligned} u_w^i &= [u_w^1, \dots, u_w^{i-1}, u_w^{i+1}, \dots, u_w^N]^T \\ y_s^i &= [y_s^1, \dots, y_s^{i-1}, y_s^{i+1}, \dots, y_s^N]^T \end{aligned} \quad (2)$$

表3 実験で使われた1人分のレポートに含まれる文数、単語数の最大・最小・平均値

	文数	単語数
max	190	6727
min	33	2492
avg	143.0	5354.3

表4 実験パラメータ (Tr, Tp以外はLightGBM)

Parameter		Parameter	
min_gain_to_split	0	num_leaves	2
min_data_in_leaf	1	boosting_type	gbdt
min_child_samples	1	random_seed	42
n_estimators	1000	Tr	0.5
learning_rate	0.01	Tp	0.01
colsample_bytree	0.6		

である。さらに、各 i に対して、 i 以外の学生の集中度との相関が高い単語の集合 $\Omega_s^i(Tr, Tp)$ を以下のように求めた。

$$\Omega_s^i(Tr, Tp) = \{w \mid |r_{ws}^i| \geq Tr, p_{ws}^i \leq Tp\} \quad (3)$$

ここで、 p_{ws}^i は r_{ws}^i に対する p 値、 Tr, Tp は各々 $|r_{ws}^i|, p_{ws}^i$ に対する閾値である。負の絶対値の大きい r 値は、その単語が小さい集中度の推定に寄与することを意味する。

次に、 $\Omega_s^i(Tr, Tp)$ 内の w を $|r_{ws}^i|$ の大きい順に並べ替えた $\Omega_s^i(Tr, Tp)$ を定義する。

$$\Omega_s^i(Tr, Tp) = [w_{s1}^i, \dots, w_{sn}^i]^T$$

$$n = |\Omega_s^i(Tr, Tp)| \leq |\Omega_s^i(Tr, Tp)| \quad (4)$$

さらに、ベクトル \mathbf{x}_s^i を次のように定義する。

$$\mathbf{x}_s^i = [u_{w_{s1}^i}^i, \dots, u_{w_{sn}^i}^i]^T \quad (5)$$

そして、以下のデータセットを用いて、 \mathbf{x}_s^i より \mathbf{y}_s^i の推定を行う。

$$\mathbf{X}_s^i = [\mathbf{x}_s^1, \dots, \mathbf{x}_s^{i-1}, \mathbf{x}_s^{i+1}, \dots, \mathbf{x}_s^N]$$

$$\mathbf{y}_s^i = [y_s^1, \dots, y_s^{i-1}, y_s^{i+1}, \dots, y_s^N] \quad (6)$$

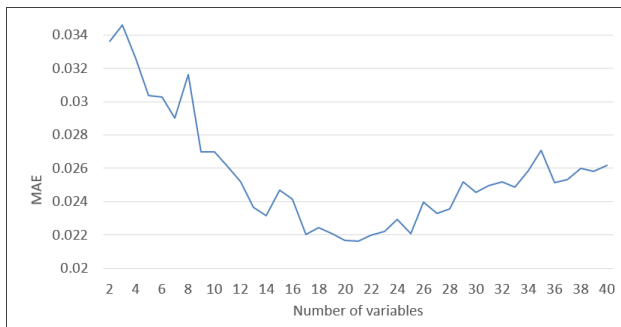


図4 説明変数の数 (n) と推定精度の関係

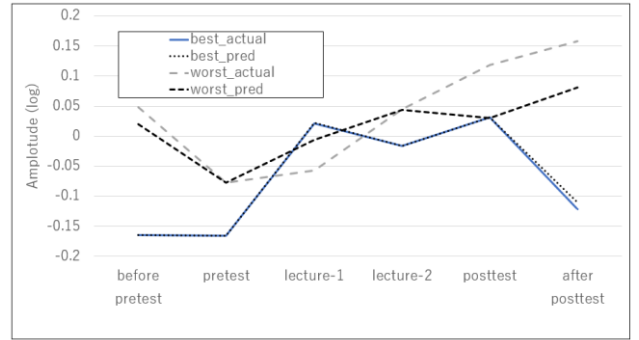


図5 誤差が最小(best)の場合と最大(worst)の場合の集中度の実測値と推定値

5. 実験結果

5.1 集中度推定

4章で議論した方法を用いて、各レポートから各学生の集中度を推定し、LOOで評価を行った。文章は1年生必修の「情報工学概論」の5回(1回につき1,500文字以上)のレポートを使用した。レポートのテーマの一例として「AIで一番興味があることを書きなさい」がある。表3に、実験の説明変数に用いた11名の学生について、各学生の5回のレポートに含まれる文数と単語数の最大値、最小値、平均値を、表4に、実験におけるLightGBMのパラメータ、および Tr, Tp の値を示す。LightGBMのパラメータは、optunaを用いて決定した。

評価には、以下で計算される平均絶対誤差(MAE)を用いた。

$$MAE_s = \frac{1}{N} \sum_{i=1}^N |\hat{y}_s^i - y_s^i| \quad (7)$$

ここで、 \hat{y}_s^i は学生 i のステージ s の集中度の推定値である。

図4は、説明変数の数 (n) を2から40まで変化させた場合の推定精度の全学生、全ステージのMAEの平均値であり、 $n=21$ のときMAEは最小(0.0216)となった。対象学生11名全員の全ステージにおける集中力の平均値と標準偏差は、それぞれ0.0216と0.1173であり、MAEの最小値は標準偏差の0.1814倍となる。図5は、11名の学生の中で、各ステージの集中度の実測値と推定値の差の合計が一番小さかった学生と一番大きかった学生の各値を示している。

5.2 集中度と使用単語の相関の解析

次に、5.1節の推定において、集中度と利用頻度との相関 (r 値) が高い単語を調査した (p 値 < 0.01)。表5に各授業中のフェーズにおける集中度との r 値が高い5つの単語およびその品詞を示す。また、図6に、集中度との相関の高い best 10, best 50 および best 90 の単語の各品詞の出現率(全単語中の利用割合)を示す。Best10, best50, best90 となるにつれ、ほぼすべてのステージで名詞の出現率が下がっており、名詞の集中度の推定に対する貢献度が高く、逆に、動詞の貢献度が低いことが判る。

表5 集中度との相関が高い単語例

before pretest			pretest		
word	品詞	r	word	品詞	r
1 工学	名詞	0.860	進める	動詞	0.833
2 す	動詞	0.718	用いる	動詞	0.831
3 判別	名詞	0.685	予測	名詞	0.794
4 前提	名詞	0.684	関係	名詞	0.793
5 データ	名詞	0.683	データ	名詞	0.788

lecture-1			lecture-2		
word	品詞	r	word	品詞	r
1 多く	名詞	0.822	一	名詞	0.569
2 以下	名詞	0.822	人数	名詞	0.559
3 大事	形状詞	0.821	方	名詞	0.552
4 新しい	形容詞	0.798	三	名詞	0.541
5 全て	名詞	0.785	世の中	名詞	0.527

posttest			after posttest		
word	品詞	r	word	品詞	r
1 予想	名詞	0.765	表示	名詞	0.904
2 物体	名詞	0.762	連携	名詞	0.884
3 表示	名詞	0.740	災害	名詞	0.838
4 収集	名詞	0.726	協力	名詞	0.835
5 J	補助記号	0.720	「	補助記号	0.835

6. おわりに

本研究では、学生の書いた文章から、各段階における学生の集中度を推定した。その結果、非常に高い精度で各ステージでの集中度を推定できることがわかった。また、集中度との名詞利用率との相関が高く、名詞を多く使う学生の集中度は、相対的に高くなる可能性があることが示唆された。これらの結果を応用することで、授業における学生の集中度や理解度を向上させることが期待される。例えば、集中度が低くなることが予想される学生を教員の近くや前の方に座らせる、クラス分けの際に、集中度の推定値で分け、集中度の低い学生のクラスでは積極的に発言させる、といった方法が考えられる。また、集中度に応じて、テストの内容を変えるなどの個別教育も考えられる。

今回の実験結果は、授業の内容、構成、学年および講師を限定して得られたものである。今後は、様々な異なる条件下で実験を行い、集中度や理解度の傾向や推定の精度を確認し、各条件に適した質の高い授業の実現を目指したい。

謝辞

本研究は、科学研究費補助金 (22K02868, 24K06270, 25K06531) の助成を受けたものである。

参考文献

- [1] Kamijo, K., Shiojiri (Saito), A., Ohzeki, K., Suzuki, M., and Sugaya, M., "Effect of Teacher's Behavior on Students' Concentration Level," *Proceedings of the International Conference on Future of Teaching and Education* (2022).
- [2] 塩尻(斎藤)亜希, 上條浩一, 鈴木圭, 大関和夫, 鈴木雅実ほか 脳波・小テスト・アンケートを用いた少人数授業における集中度と理解度の推移. 教育システム情報学会 2022 年度第 4 回研究会, 研究報告 JSiSE7 (2022).
- [3] Conneau, Alexis, and Guillaume Lample, "Cross-lingual language model pretraining," *Advances in neural information processing systems* 32 (2019).
- [4] Kamijo, Koichi, Tetsuya Nasukawa, and Hideya Kitamura, "Personality estimation from Japanese text," *Proceedings of the Workshop on Computational Modeling of People's Opinions, Personality, and Emotions in Social Media (PEOPLES)* (2016).



図6 集中度との相関が高い品詞の出現頻度：相関の大きい best 10,50,90 での比較

- [5] 那須川哲哉, 上條浩一, 日本語における筆者の性格推定の取組み. 言語処理学会第23回年次大会発表論文集: 807-810 (2017).
- [6] Kamijo, Koichi, Yaeko Mitsumori, Hiroshi Kato, Akiko Kato, "Comparison of Patent Grants within Patent Families-Why are some patents within the same patent family granted in one country but not in another?" *PICMET2022*, IEEE, pp. 1-10 (2022).
- [7] Kamijo, Koichi, Yaeko Mitsumori, Hiroshi Kato, Akiko Kato, "Drug Approval Prediction using Patents," *PICMET2023*, IEEE (2023).
- [8] Kamijo, Koichi, Y.Mitsumori, H.Kato, A.Kato, "Annual pharmaceutical sales estimates using patents," *PICMET 2024* (2024).
- [9] Giannitrapani, D., "The role of 13-hz activity in mentation," *The EEG of Mental Activities*, 149-152. doi:10.1159/000416426 (1988).
- [10] Antonenko, P., & Niederhauser, D., "The effects of loads on cognitive load and learning in a hypertext environment," *Computers in Human Behavior*, 26, 140-150. doi:10.1016/j.chb.2009.10.014 (2010).
- [11] Antonenko, P., Paas, F., Grabner, R., & Gog, T., "Using Electroencephalography to Measure Cognitive Load," *Educational Psychology Review*, 22(4), 425-438. doi:10.1007/s10648-010-9130-y (2010).
- [12] Lokare, V. T., & Netak, L. D., "Concentration level prediction system for the students based on physiological measures using the EEG device," *Intelligent Human Computer Interaction: 12th International Conference, IHCI 2020*, Daegu, South Korea, Part I 12, 24-33. doi: 10.1007/978-3-030-68449-5_3 (2021).
- [13] Priyanka, A. A., Gawali, B. W., and Mehrotra, S. C., "Technological, basics of EEG recording and operation of apparatus. *Introduction to EEG-and Speech-Based Emotion Recognition* : 19-50 (2016).
- [14] Yoshita, K., Sakamoto, Y., Miyaji, I., and Yamada, K., "Analysis comparison of brain waves at the learning status by simple electroencephalography," *KES'2012, Proceedings, Knowledge-Based Intelligent Information and Engineering Systems* : 1817-1826 (2012).
- [15] EEGLAB (2022) <https://eeqlab.org/> (accessed 2025.06.06)
- [16] 人見健文, 池田昭夫, 脳波の基礎知識. *臨床神経生理学*, 42(6) : 365-370 (2014).
- [17] Mind Monitor, https://www.researchgate.net/figure/EEG-sensors-TP9-AF7-AF8-and-TP10-of-the-Muse-headband-on-the-international-standard_fig2_331733067 (accessed 2025.06.06).
- [18] Krigolson, O. E., Williams, C. C., Norton, A., Hassall, C. D., & Colino, F. L. Choosing, "MUSE: Validation of a low-cost, portable EEG system for ERP research," *Frontiers in neuroscience*, 11, 109. doi: 10.3389/fnins.2017.00109 (2017).