A study on the characteristics of student behavior and emotion

Liu Chen 2020521032, Zhang Yufei 2021533169

May 15, 2024

Abstract

Real-time recorded information has better correlation and prediction accuracy in inferring individual emotions than delayed recorded information. Based on real-time user information data and sentiment indicators reported weekly from College Experience Study Dataset, this project attempts to discover the time-series-based characteristic relationship between student behavior and emotion. Further, we focus on students' gait and audio characteristics, hoping to predict students' emotional state through these characteristics, and provide guidance for students' stress reduction.

1 Introduction

The relationship between behavioral traits and emotions is an interesting area of research that explores how people's behavioral patterns are affected by their emotional states. For college students, the impact of emotional state on academic performance is also significant, so understanding the relationship between behavioral traits and emotions can provide valuable insights into students' emotional health and may help develop interventions to support their mental health.

Previous research has shown that emotional states can have a significant impact on an individual's physical movement and behavior. For example, research has shown that individuals experiencing positive emotions tend to exhibit more open and extended body postures, while individuals experiencing negative emotions may exhibit narrower and more closed body posturesss [1]. These findings suggest that emotions affect not only facial expressions, but also body movements.

For this reason, our work is based on Dartmouth College's StudentLife dataset. We examined whether there was a correlation between activity characteristics and mood by analyzing data collected from students' daily activities, based primarily on walking movement characteristics (including daily steps, stride length, location, etc.) and audio characteristics (including Mean audio amplitude, Standard deviation of audio amplitude, Duration of detected conversations, Number of conversation and Ratio of detection of voice to number of inferences.). This analysis can help identify potential patterns and associations that may exist between behavioral variables.

College Experience Study Dataset tracks behavioral data of Dartmouth students which were collected over 200 undergraduate students throughout their four-year college experience, before and after the end of the COVID-19 pandemic using passive sensor data, questionnaires, and interviews. The population information collected by this dataset is shown in the **table 1, 2**. And passive sensing data the dataset has is shown in **table 3**, we focus on walking and audio.

In addition, we mainly use PHQ4-score the dataset collected as a standard to measure students' emotional indicators, It is included in the Ecological Momentary Assessments(EMA) of the dataset. PHQ4 is an ultra-short scale used to screen symptoms of depression and anxiety, and its titles are shown in **table 4**. For subscores of depression and anxiety, if the score reaches 3 points or higher, it is usually

Sex	Count	Percentage
Female	129	69.35%
Male	57	30.65%

Table 1: Sex of students in dataset

Race	Count and Percentage
White	110 59.14%
Asian	43 23.12%
Black or African American	6 3.23%
American Indian/Alaska Native	6 3.23%
More than one race	15 8.05%
Not reported	6 3.23%

Table 2: Race of students in dataset

regarded as a positive standard for screening.[18]

2 Related Work

At present, there have been some efforts to collect life data of subjects by means of questionnaires, such as action track, action distance, web browsing, dining place and other information, so as to study the feasibility of using daily behavior information to predict emotional changes and assess mental state. But there is a certain loss of accuracy with this kind of backward data collection. Due to the problem design, forgetting details and other factors, it is impossible to restore the real information. To solve this problem, the StudentLife dataset was built. The data set uses the built-in APP of the mobile phone to collect the physical information of individuals in real time, which improves the accuracy of the data on the basis of previous research. [2] In 2013, the first version of the dataset was released as a test version, with a wide variety of data and a short time span. In 2016, a second dataset was created using a variety of statistical tools to analyze the relationship between measures such as sleep duration, activity duration, concentration, and individual mood. [3] In 2024, the authors published a third dataset and related work, which followed 220

Category	Features
Physical Activity	Walking / sedentary (still) / biking / running duration
	Duration in vehicle, number of steps
Mobility	Semantic Locations
	Distance travelled, time spent at home, workout places, study places, social places, dorms, others' dorm, greek houses, max distance from campus, number of locations visited
Phone usage	Number of phone locks & unlocks, duration of phone unlock
Audio Plays	Number of audio plays, duration of audio plays
Sleep	Sleep duration, sleep start time, sleep end time

Table 3: Passive sensing data of dataset

Problem	Not at all	Several days	More than half the days	Nearly every day
Feeling nervous, anxious or on edge	0	1	2	3
Not being able to stop or control worrying	0	1	2	3
Feeling down, depressed, or hopeless	0	1	2	3
Little interest or pleasure in doing things	0	1	2	3

Table 4: Question and score of PHQ4

university students for four years and combined magnetic resonance imaging to analyze the correlation of various life behaviors and emotions. [4,5,12] The focus of the three versions of the StudentLife dataset is different. In the dataset published in 2013, emphasis was placed on the broadness of the data types, with individual mood indicators assessed only before and after the project; In the dataset published in 2016, the authors increased the number of emotional assessments; In the dataset published in 2024, the frequency of mood assessments was increased to once a week.

The StudentLife dataset is the first to use mobile phones to record mobile perception data to show students emotions. [2] Subsequent work has attempted to use perceptual data to predict human behavior. [6,7,8,9] including predictions about people's mental health. [10,11] In these works, mobile perception data was used to analyze the correlation with users' mental health, and a large number of machine learning models and deep learning models were used. For example, in the paper published at the same time of the third edition of the dataset, the author used deep learning algorithms to predict students' mental health with behavioral data [12]. In summary, there is a lot of work and algorithms being developed to analyze and predict the relationship between behavioral data and mental health.

As for movement and gait characteristics, existing studies mainly focus on the gait characteristics of elderly people with diseases [13,14]. Few studies have focused on the physical response of emotions in students or adults, or have focused on contributing to the recovery of specific diseases. There are few articles on emotion discrimination in healthy adults, using motion capture [15] and Near-infrared spectroscopy [16] to identify different emotions through subtle features. Our difference is that we base our discussion on daily data collection over a long period of time, and the subject does not need to make other specified actions.

Existing studies of speech detection of mental illness are mostly based on acoustic signal processing detection of small samples of patients[19], but the analysis of ambient sounds and social frequency can also study the characteristics of social support of subjects, thereby predicting emotional development.[20]

3 Problem statement

As mentioned earlier, the analysis of behavioral data may have some guidance for the prevention of mental illness, and on this basis, we wanted to focus on the relationship between movement characteristics , audio detected and mood based on the College Experience Study Dataset published in 2024. Our goal was to analyze the correlation between movement characteristics such as number of steps, stride length, location of movement and emotion , audio detected and emotion in existing datasets.

About fitting the relationship, we propose two kinds of methods. In first methods, we propose that testees can be divided into several clusters. Feature extraction from clusters can be more effective for the fitting of relationship. In second methods, we use the predictive property of Time sequence data to fit relationship between movement characteristics and emotion of each individual to give a predictive model for each testee.

4 Method

4.1 Fitting relationship based on clustering

Kshape is a clustering algorithm specicficly used to cluster Time Sequence data. Cross-correlation distance is calculated from two time sequence to quantify the distance between them. Cross-correlation distance can reflect important features of time sequence data, like the Translation invariance and scale invariance. Like K-means, K-shape indicate some centers and minimize the distance between each time sequence data and centers. So emotion data is clustered to identify individual with different resistance to the emotion changes; movement data is clustered to identify different pattern of movement.

Finally, each uid will have a emotion cluster code and a movement cluster code. Each time sequence contains all records of one item of each uid, for example, four years' phq4 score with time labels of uid1.

4.2 Evaluating the relationship between movement and emotion

DTW(Dynamic Time Warping) is used to calculate the closet distanc e between two time sequence. Since clustered data is still messy for correlation calculation, here we define a correlation based on clustering and DTW distance:

$$DTWD(i,j) = \frac{DTW(i,j)}{len(i) * len(j)}$$

$$correlation = \frac{\sum_i \sum_j DTWD^2(i,j)}{\sum_i \sum_j (DTWD_s(i,j) + DTWD_s(i,j))/2}$$

In which, $DTW(u,v)$ means DTW distance between time sequence u and time sequence v, and u,v are each separately vectors in emotion cluster i and movement cluster j. DTWD means the normalized distance based on length of each time sequence.

Since DTW is the direct distance between two vectors, DTW is greatly influcenced by length of vectors, making the comparison between clusters difficult. $DTWD_s$,which is also $DTWD_{shuffle}$, means randomly re-distributed clusters for each time sequence, and do recalculation to make background noise.

4.3 Audio data

In view of the fact that the data set collects parameters of various activities on an hourly basis, with a high dimension, but the EMA scale is collected several days apart, we only select Audio parameters measured on a daily basis for analysis.We first select the features of interest, and draw the correlation matrix after cleaning.(Figure 1) Therefore, we found that audio records had a strong correlation and could be compressed, but call records had a weak correlation with each other. In view of the large number of missing values in the call records, we abandoned call-related features. For EMA data, we also plotted correlations for all features. (Figure 2) In view of the strong correlation between PHQ4 score and single-item scores, we choose to use PHQ4 score as a direct measure of students' emotional state. Then we compress the audio feature into a one-dimensional vector,

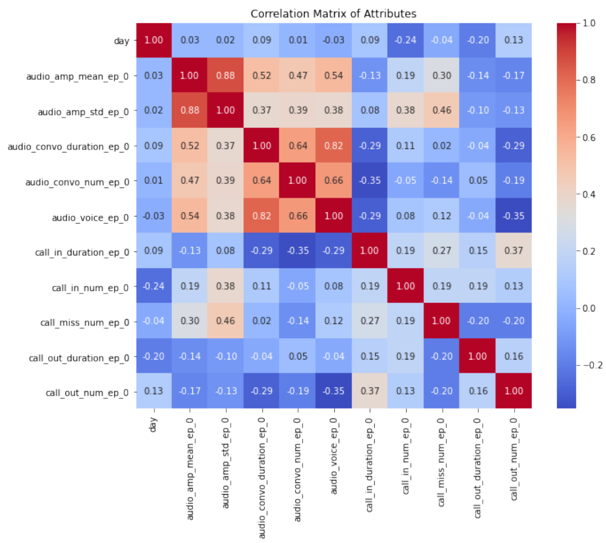

Figure 1: The correlation matrix of the audio class

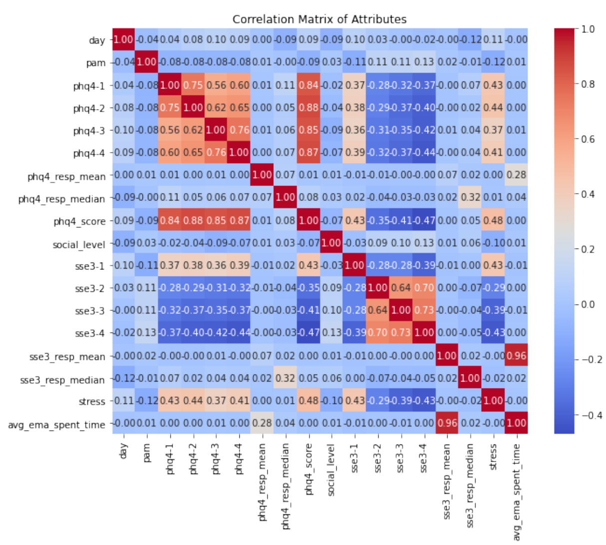

Figure 2: Correlations for all features of EMA

and use different methods and models to try to fit and predict the PHQ4 score after interpolation using the audio vector. In this process, we found that everyone's emotional pattern was different, and it was difficult to cluster them under general conditions using clustering methods, so we used the method of training the model separately for different students when training the model.Then we compress the audio feature into a one-dimensional vector, and use different methods and models to try to fit and predict the PHQ4 score after interpolation using the audio vector. In this process, we found that everyone's emotional pattern was different, and it was difficult to cluster them under general conditions using clustering methods, so we used the method of training the model separately for different students when training the model. The specific experimental procedure will be mentioned in **Part 5**.

4.4 Step data

As for gait data, we observe the Sensing and Step data, and extract students who only use walking as a mode of travel. By associating the number of steps taken daily with the walking time, we can obtain the stride length data of these students through indirect calculation. As mentioned earlier, the pattern of posture data can reflect emotional states to some extent, so we assume that significant stride changes correspond to unstable emotional states.

5 Experiments

5.1 Clustering of time sequence shows bad results

Each time sequence is composed with 4 years' phq4 score sampled in each week of each uid(testee) as inputs, Kshape algorithm is used for clustering to get a 20-cluster classification. Due the complexity of emotion time sequence data, the clustering result is poor, gives less feature information. Clustering of movement information gives better results: Calculating the DTWD distance matrix, $DTWD_{shuffle}$ distance matrix and corrected DTWD distance matrix. Here

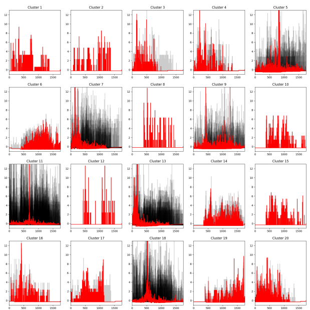

Figure 3: Clustering of emotion time sequence data

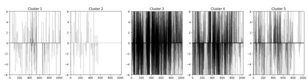

Figure 4: Clustering of movement time sequence data

shows the DTWD distance between phq4 and strike: Here, $DTWD$ and $DTWD_{shuffle}$ are calculated separately as described in methods. Then $\frac{DTWD}{DTWD_{shuffle}}$ is done to remove noise caused by the sampling methods. As shown in results, whatever before or after denoising, DTWD can not tell a difference between different clusters, indicating that attempting to extract information through clustering is not suitable for this database.

5.2 Audio data

We compress the audio features into one-dimensional vectors using PCA, complete the PHQ4 score values using linear interpolation, and then regularize them respectively. Linear Regression, Random Forest and Gradient Boosting Regression Tree were respectively used to build models for each uid to predict PHQ4 score, and then the mean square error between the real regularized PHQ score and the predicted value was calculated respectively. The performance of the model is evaluated as the average and sum of all Uids (N=98).The final MSE obtained is shown in **Figure 3**. We found that GBRT fitted the PHQ4 score well and decided to test the prediction by splitting the data into a training set and a test set.Dividing the data set into a training set (80%) and a post-test set (20%), the GBRT model obtained an average MSE value of 1.414.(n_estimators=110, random_state=42)

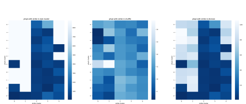

Figure 5: DTWD distance bwtween phq4 and strike

	Linear Regression	Random Forest	Gradient Boosting Regression Tree
Average MSE	0.769	0.162	0.005
Sum MSE	75.411	15.890	0.527

Figure 6: MSE of different model

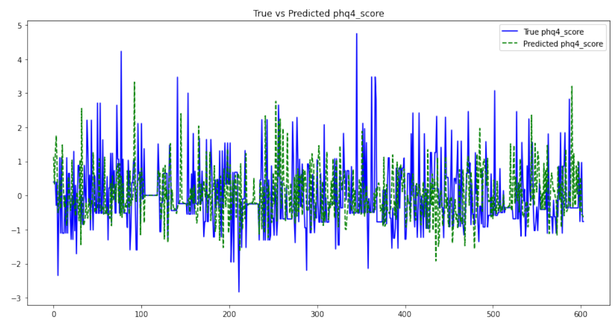

Figure 7: Predicted PHQ4 score and true PHQ4 score

5.3 Step data

6 Conclusion and Discussion

We found that Gradient Boosting Regression Tree can better predict the PHQ4 score state through daily Audio collection, which means that daily passive collection can replace scale collection to a certain extent, which provides a new way to judge students' emotional state. However, due to the connection between daily social status and emotional status, we believe that the information in Audio data still has a repetitive function in predicting scores, and the collected data can still be further compressed. In addition, the emotional change pattern of each person is indeed significantly different, but the cost of individual modeling is large, and how to better cluster is still a challenge.

7 References

[1] Frontiers | Emotional characteristic analysis of human gait while real-time movie viewing https://www.frontiersin.org/articles/10.3389/frai.2022.989860/full

[2] Rui Wang, Fanglin Chen, Zhenyu Chen, Tianxing Li, Gabriella Harari, Stefanie Tignor, Xia Zhou, Dror Ben-Zeev, and Andrew T. Campbell. 2014. StudentLife: assessing mental health, academic performance and behavioral trends of college students using smartphones. In Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp '14). Association for Computing Machinery, New York, NY, USA, 3–14. https://doi.org/10.1145/2632048.2632054

[3] Rui Wang, Weichen Wang, Alex daSilva, Jeremy F. Huckins, William M. Kelley, Todd F. Heatherton, and Andrew T. Campbell. 2018. Tracking Depression Dynamics in College Students Using Mobile Phone and Wearable Sensing. Proc. ACM Interact. Mob. Wearable Ubiquitous Technol. 2, 1, Article 43 (March 2018), 26 pages. https://doi.org/10.1145/3191775

[4] Wang W, Nepal S, Huckins JF, Hernandez L, Vojdanovski V, Mack D, Plomp J, Pillai A, Obuchi M, Dasilva A, Murphy E, Hedlund E, Rogers C, Meyer M, Campbell A. First-Gen Lens: Assessing Mental Health of First-Generation Students across Their First Year at College Using Mobile Sensing. Proc ACM Interact Mob Wearable Ubiquitous Technol. 2022 Jul;6(2):95. doi: 10.1145/3543194. Epub 2022 Jul 7. PMID: 36561350; PMCID: PMC9770714.

[5] Huckins JF, daSilva AW, Wang W, Hedlund E, Rogers C, Nepal SK, Wu J, Obuchi M, Murphy EI, Meyer ML, Wagner DD, Holtzheimer PE, Campbell AT. Mental Health and Behavior of College Students During the Early Phases of the COVID-19 Pandemic: Longitudinal Smartphone and Ecological Momentary Assessment Study. J Med Internet Res. 2020 Jun 17;22(6):e20185. doi: 10.2196/20185. PMID: 32519963; PMCID: PMC7301687.

[6] Saeed Abdullah, Elizabeth L. Murnane, Mark Matthews, Matthew Kay, Julie A. Kientz, Geri Gay, and Tanzeem Choudhury. 2016. Cognitive rhythms: unobtrusive and continuous sensing of alertness using a mobile phone. In Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp '16). Association for Computing Machinery, New York, NY, USA, 178–189. https://doi.org/10.1145/2971648.2971712

[7] N. D. Lane, E. Miluzzo, H. Lu, D. Peebles, T. Choudhury and A. T. Campbell, "A survey of mobile phone sensing," in IEEE Communications Magazine, vol. 48, no. 9, pp. 140-150, Sept. 2010, doi: 10.1109/MCOM.2010.5560598.

[8] Jonathan Liono, Zahraa S. Abdallah, A. K. Qin, and Flora D. Salim. 2018. Inferring Transportation Mode and Human Activity from Mobile Sensing in Daily Life. In Proceedings of the 15th EAI International Conference on Mobile and Ubiquitous Systems: Computing, Networking and Services (MobiQuitous '18). Association for Computing Machinery, New York, NY, USA, 342–351. https://doi.org/10.1145/3286978.3287006

[9] Tauhidur Rahman, Mary Czerwinski, Ran Gilad-Bachrach, and Paul Johns. 2016. Predicting "About-to-Eat" Moments for Just-in-Time Eating Intervention. In Proceedings of the 6th International Conference on Digital Health Conference (DH '16). Association for Computing Machinery, New York, NY, USA, 141–150. https://doi.org/10.1145/2896338.2896359

[10] Mehrab Bin Morshed, Koustuv Saha, Richard Li, Sidney K. D'Mello, Munmun De Choudhury, Gregory D. Abowd, and Thomas Plötz. 2019. Prediction of Mood Instability with Passive Sensing. Proc. ACM Interact. Mob. Wearable Ubiquitous Technol. 3, 3, Article 75 (September 2019), 21 pages. https://doi.org/10.1145/3351233

[11] Arvind Pillai, Subigya Kumar Nepal, Weichen Wang, Matthew Nemesure, Michael Heinz, George Price, Damien Lekkas, Amanda C. Collins, Tess Griffin, Benjamin Buck, Sarah Masud Preum, Trevor Cohen, Nicholas C. Jacobson, Dror Ben-Zeev, and Andrew Campbell. 2024. Investigating Generalizability of Speech-based Suicidal Ideation Detection Using

Mobile Phones. Proc. ACM Interact. Mob. Wearable Ubiquitous Technol. 7, 4, Article 174 (December 2023), 38 pages. https://doi.org/10.1145/3631452

[12] Subigya Nepal, Wenjun Liu, Arvind Pillai, Weichen Wang, Vlado Vojdanovski, Jeremy F. Huckins, Courtney Rogers, Meghan L. Meyer, and Andrew T. Campbell. 2024. Capturing the College Experience: A Four-Year Mobile Sensing Study of Mental Health, Resilience, and Behavior of College Students during the Pandemic. Proc. ACM Interact. Mob. Wearable Ubiquitous Technol. 8, 1, Article 38 (March 2024), 37 pages. https://doi.org/10.1145/3643501

[13] Kyrdalen IL, Thingstad P, Sandvik L, Ormstad H. Associations between gait speed and well-known fall risk factors among community-dwelling older adults. Physiother Res Int. 2019 Jan;24(1):e1743. doi: 10.1002/pri.1743. Epub 2018 Sep 10. PMID: 30198603.

[14] Dragašević-Mišković NT, Bobić V, Kostić M, Stanković I, Radovanović S, Dimitrijević K, Svetel M, Petrović I, Đurić-Jovičić M. Impact of depression on gait variability in Parkinson's disease. Clin Neurol Neurosurg. 2021 Jan;200:106324. doi: 10.1016/j.clineuro.2020.106324. Epub 2020 Oct 21. PMID: 33129594.

[15] Roether, Claire L et al. "Critical features for the perception of emotion from gait."Journal of vision vol. 9,6 15.1-32. 26 Jun. 2009, doi:10.1167/9.6.15

[16] Schneider S, Christensen A, Häußinger FB, Fallgatter AJ, Giese MA, Ehlis AC. Show me how you walk and I tell you how you feel - a functional near-infrared spectroscopy study on emotion perception based on human gait. Neuroimage. 2014;85 Pt 1:380-390. doi:10.1016/j.neuroimage.2013.07.078

[17] Gross MM, Crane EA, Fredrickson BL. Effort-Shape and kinematic assessment of bodily expression of emotion during gait. Hum Mov Sci. 2012;31(1):202-221. doi:10.1016/j.humov.2011.05.001

[18] https://www.mcgill.ca/schoolofmedicine-learneraffairs/pgme/wellness-support/self-screening-tools/anxiety-depression-phq-4

[19] 刘振泰, 向春妮, 刘陈陵, 钟宝亮, 黄海, 彭志昆, 吕筑, 丁钟. 基于语音的抑郁检测研究综述 [J]. 信号处理, 2023, 39(4): 616-631. DOI: 10.16798/j.issn.1003-0530.2023.04.003

[20]daSilva AW, Huckins JF, Wang W, Wang R, Campbell AT, Meyer ML. Daily perceived stress predicts less next day social interaction: Evidence from a naturalistic mobile sensing study. Emotion. 2021;21(8):1760-1770. doi:10.1037/emo0000994