DRAFT 2020-12-07
The latest version is available at erik.bjareholt.com/thesis/thesis.pdf

—

M.Sc. Thesis

Classifying brain activity using low-cost biosensors and automated time tracking

Erik Bjäreholt
(erik@bjareho.lt, dat13ebj@student.lu.se)

2020-12-07

Abstract

Add: Abstract

Contents

Notes & TODOs

1 Introduction

Improve: Needs more

2 Background

People spend more time than ever using computing devices. Work, entertainment, and services, have been steadily moving online over the last few decades and this trend is expected to continue. While several studies have been tracking how people spend time on their devices a wider study of how people's app usage is changing over time and how it varies with demographics, is not publicly available.

Furthermore, how different device activities affect the user behaviorally and neurologically is of interest for many areas of research, including:

- psychological well-being, such as depression and social anxiety [1][2], stress [3], self-esteem, life satisfaction, loneliness, and depression [4].
- the impact of screen time on children and adolescents [5].
- attention span among media multitasking adults [3].
- enhancing personal productivity [6].

Understanding device use and the underlying cognitive processes are essential when designing for motivation, engagement and wellbeing in digital experiences [7].

This becomes especially relevant for knowledge workers, such as software developers, who spend the majority of their working time on computing devices.

2.1 Automated time trackers

Automated time-trackers have been developed for computing devices, with various applications such as tracking productivity, managing excessive use of social networking sites (SNSs), and studying user behavior.

2.1.1 Commercial use

Companies like RescueTime [8], Hubstaff [9], and others offer automated time tracking as a service. These services let the user track their screen time by installing a program on their device which tracks the active application and sends the data to their servers for storage and analysis. The user can then view their data in a dashboard on the service's website. Some of these services, like

RescueTime and Hubstaff, are marketed towards teams and professionals, who want to keep track of individual and team productivity.

However, these services have some issues for use by researchers and individuals alike. Notably, their collection of detailed and non-anonymized behavioral data into a centralized system bring significant privacy concerns, especially in cases where the data is shared with a team or an employer.

Other limitations of these services, such as low temporal resolution and limited event detail, cause additional issues for certain tasks that are timing-sensitive (such as ERPs), or preprocessing steps that can take advantage of high level of detail (like classifying activity).

2.1.2 Research use

Previous research has been published which used automated time trackers, such as TimeAware [6] and ScreenLife [10]. However, these previous contributions are — like the commercial services — not open source nor permissively licensed, and therefore not available for external research use nor further development.

2.1.3 ActivityWatch

The free and open source automated time tracker ActivityWatch [11] addresses aforementioned issues with other software around source availability/licensing, privacy, temporal resolution, event detail, and cross-platform support.

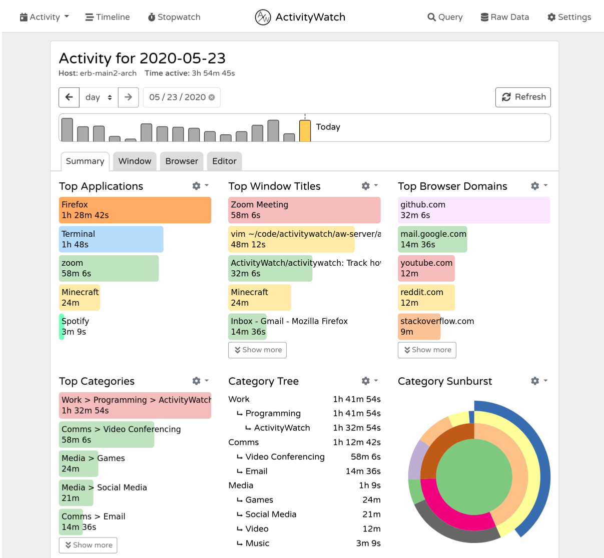

Figure 1: ActivityWatch activity dashboard. Showing top applications, window titles, browser domains, and categories.

2.2 EEG and low-cost biosensors/functional brain imaging

Functional brain imaging methods such as fMRI, fNIRS, and EEG, have been used to study the relationship between cognitive or physical activity, and brain activity [12][13][14]. The more accurate methods such as fMRI are costly and inflexible/impractical for many uses.

However, the recent availability of low-cost biosensors such as EEG, HEG, and fNIRS, enables studying brain activity during real-life tasks. As an example it has been shown that it is possible to classify what task a participant is undertaking using fMRI [12], which has been replicated using EEG and low-cost biosensors [14].

But they are not without their limitations — among them a notably low signal-to-noise ratio [15] — yet visual evoked potentials (VEPs) have been shown to be sufficient for high-speed BCI applications [16].

To combat the low signal-to-noise ratio, machine learning methods have been employed with varying degrees of success. Examples from previous research include Convolutional Neural Networks (CNNs), which have been successful in classifying time series in general [17], and EEG data in particular [18]. As well as Hierarchical Convolutional Neural Networks (HCNNs), which have been used for EEG-based emotion recognition [19].

Add: Applications to software engineers

2.3 Aim of the thesis

Add: Insert stuff from goal document

2.4 Related work

It has previously been shown that both fMRI[12] and single-channel EEG[14] provides enough information to classify wether the subject was reading prose or code. However, accuracy with single-channel EEG was poor, and notably outperformed by a HRV monitor. One of the goals of this thesis was to investigate if better equipment could improve this accuracy, and wether the code vs prode compresension task generalizes across more varied activities, such as organic device use.

Add: Insert mention of preprint that Fucci mentioned?

3 Methods

3.1 Collection of device activity data

All device activity is collected using the automated time tracker ActivityWatch[20].

Add: Summary of ActivityWatch

3.2 Collection of EEG data

EEG data was collected during both organic use and under controlled conditions.

actually perform controlled experiments

3.2.1 Devices

- Muse S
- OpenBCI Cyton
- Neurosity Notion 1 (soon)
- Neurosity Notion 2 (maybe?)

3.3 Riemannian geometry

The state of the art in many EEG classification tasks involves the use of Riemannian geometry.

Add: Explanation of riemannian geometry, from this tutorial we're working on

[21]

4 Results

Our classifier performance is. . .

5 Conclusions

Our results show. . .

6 Discussion

6.1 Ethical considerations

Add: Discussion around ethics of data collection, how it's dealt with in ActivityWatch, and how it should perhaps be dealt with here.

Acknowledgements

- My advisor
- The NeuroTechX crowd, specifically Morgan Hough and John Griffiths
- The people at the LTH Department for Automatic Control
- My brother Johan Bjäreholt, for working with me on ActivityWatch all these years.

The Oxford English Dictionary defines 'thesis' as "a long essay or dissertation involving *personal research*, written by a candidate for a university degree". I can't think of more "personal research" than research in quantified self with personal data.

References

[1] Maarten H. W. Selfhout et al. "Different types of Internet use, depression, and social anxiety: The role of perceived friendship quality". In: *Journal of Adolescence* 32.4 (Aug. 1, 2009), pp. 819–833. ISSN: 0140-1971. DOI: 10.1016/j.adolescence.2008.10.011. URL: http://www.sciencedirect.com/science/article/pii/S0140197108001218 (visited on 05/13/2020).

[2] Dhavan Shah et al. "Nonrecursive Models of Internet Use and Community Engagement: Questioning Whether Time Spent Online Erodes Social Capital". In: *Journalism & Mass Communication Quarterly* 79.4 (Dec. 1, 2002). Publisher: SAGE Publications Inc, pp. 964–987. ISSN: 1077-6990. DOI: 10.1177/107769900207900412. URL: https://doi.org/10.1177/107769900207900412 (visited on 05/13/2020).

[3] Gloria Mark, Yiran Wang, and Melissa Niiya. "Stress and multitasking in everyday college life: an empirical study of online activity". In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '14. Toronto, Ontario, Canada: Association for Computing Machinery, Apr. 26, 2014, pp. 41–50. ISBN: 978-1-4503-2473-1. DOI: 10.1145/2556288.2557361. URL: https://doi.org/10.1145/2556288.2557361 (visited on 05/22/2020).

[4] Chiungjung Huang. “Time Spent on Social Network Sites and Psychological Well-Being: A Meta-Analysis”. In: *Cyberpsychology, Behavior, and Social Networking* 20.6 (June 1, 2017). Publisher: Mary Ann Liebert, Inc., publishers, pp. 346–354. ISSN: 2152-2715. DOI: 10.1089/cyber.2016.0758. URL: https://www.liebertpub.com/doi/10.1089/cyber.2016.0758 (visited on 05/13/2020).

[5] Kaveri Subrahmanyam et al. “The impact of computer use on children’s and adolescents’ development”. In: *Journal of Applied Developmental Psychology*. Children in the Digital Age 22.1 (Jan. 1, 2001), pp. 7–30. ISSN: 0193-3973. DOI: 10.1016/S0193-3973(00)00063-0. URL: http://www.sciencedirect.com/science/article/pii/S0193397300000630 (visited on 05/13/2020).

[6] Young-Ho Kim et al. “TimeAware: Leveraging Framing Effects to Enhance Personal Productivity”. In: *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. CHI ’16. San Jose, California, USA: Association for Computing Machinery, May 7, 2016, pp. 272–283. ISBN: 978-1-4503-3362-7. DOI: 10.1145/2858036.2858428. URL: https://doi.org/10.1145/2858036.2858428 (visited on 05/13/2020).

[7] Dorian Peters, Rafael A. Calvo, and Richard M. Ryan. “Designing for Motivation, Engagement and Wellbeing in Digital Experience”. In: *Frontiers in Psychology* 9 (2018). Publisher: Frontiers. ISSN: 1664-1078. DOI: 10.3389/fpsyg.2018.00797. URL: https://www.frontiersin.org/articles/10.3389/fpsyg.2018.00797/full?report=reader (visited on 05/23/2020).

[8] *RescueTime: Automatic Time-Tracking Software*. Library Catalog: www.rescuetime.com. URL: https://www.rescuetime.com/ (visited on 05/23/2020).

[9] *Hubstaff: Time Tracking and Productivity Monitoring Tool*. Library Catalog: hubstaff.com. URL: https://hubstaff.com/ (visited on 05/23/2020).

[10] John Rooksby et al. “Personal Tracking of Screen Time on Digital Devices”. In: *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. CHI ’16. San Jose, California, USA: Association for Computing Machinery, May 7, 2016, pp. 284–296. ISBN: 978-1-4503-3362-7. DOI: 10.1145/2858036.2858055. URL: https://doi.org/10.1145/2858036.2858055 (visited on 05/13/2020).

[11] *ActivityWatch: Open source automated time tracker*. In collab. with Erik Bjäreholt and Johan Bjäreholt. original-date: 2016-04-27T15:26:09Z. May 23, 2020. URL: https://github.com/ActivityWatch/activitywatch (visited on 05/23/2020).

[12] Benjamin Floyd, Tyler Santander, and Westley Weimer. “Decoding the Representation of Code in the Brain: An fMRI Study of Code Review and Expertise”. In: *2017 IEEE/ACM 39th International Conference on Software Engineering (ICSE)*. 2017 IEEE/ACM 39th International Conference on Software Engineering (ICSE). ISSN: 1558-1225. May 2017, pp. 175–186. DOI: 10.1109/ICSE.2017.24.

[13] Keum-Shik Hong, Noman Naseer, and Yun-Hee Kim. "Classification of prefrontal and motor cortex signals for three-class fNIRS–BCI". In: *Neuroscience Letters* 587 (Feb. 5, 2015), pp. 87–92. ISSN: 0304-3940. DOI: 10.1016/j.neulet.2014.12.029. URL: http://www.sciencedirect.com/science/article/pii/S0304394014009744 (visited on 05/22/2020).

[14] Davide Fucci et al. "A Replication Study on Code Comprehension and Expertise using Lightweight Biometric Sensors". In: *2019 IEEE/ACM 27th International Conference on Program Comprehension (ICPC)*. 2019 IEEE/ACM 27th International Conference on Program Comprehension (ICPC). ISSN: 2643-7171. May 2019, pp. 311–322. DOI: 10.1109/ICPC.2019.00050.

[15] D. J. McFarland and J. R. Wolpaw. "EEG-based brain–computer interfaces". In: *Current Opinion in Biomedical Engineering*. Synthetic Biology and Biomedical Engineering / Neural Engineering 4 (Dec. 1, 2017), pp. 194–200. ISSN: 2468-4511. DOI: 10.1016/j.cobme.2017.11.004. URL: http://www.sciencedirect.com/science/article/pii/S246845111730082X (visited on 05/13/2020).

[16] Martin Spüler. "A high-speed brain-computer interface (BCI) using dry EEG electrodes". In: *PLoS ONE* 12.2 (Feb. 22, 2017). ISSN: 1932-6203. DOI: 10.1371/journal.pone.0172400. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5321409/ (visited on 05/13/2020).

[17] Bendong Zhao et al. "Convolutional neural networks for time series classification". In: *Journal of Systems Engineering and Electronics* 28.1 (Feb. 2017). Conference Name: Journal of Systems Engineering and Electronics, pp. 162–169. ISSN: 1004-4132. DOI: 10.21629/JSEE.2017.01.18.

[18] Robin Tibor Schirrmeister et al. "Deep learning with convolutional neural networks for EEG decoding and visualization". In: *Human Brain Mapping* 38.11 (2017). _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/hbm.23730, pp. 5391–5420. ISSN: 1097-0193. DOI: 10.1002/hbm.23730. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/hbm.23730 (visited on 05/22/2020).

[19] Jinpeng Li, Zhaoxiang Zhang, and Huiguang He. "Hierarchical Convolutional Neural Networks for EEG-Based Emotion Recognition". In: *Cognitive Computation* 10.2 (Apr. 1, 2018), pp. 368–380. ISSN: 1866-9964. DOI: 10.1007/s12559-017-9533-x. URL: https://doi.org/10.1007/s12559-017-9533-x (visited on 05/22/2020).

[20] Erik Bjäreholt and Johan Bjäreholt. *ActivityWatch - An open-source automated time tracker*. Language: eng. Nov. 24, 2020. DOI: 10.5281/zenodo.3794887. URL: https://zenodo.org/record/3794887 (visited on 12/03/2020).

[21] Alexandre Barachant et al. "Classification of covariance matrices using a Riemannian-based kernel for BCI applications". In: *Neurocomputing* 112 (July 1, 2013), pp. 172–178. ISSN: 0925-2312. DOI: 10.1016/j.neucom.2012.12.039. URL: https://doi.org/10.1016/j.neucom.2012.12.039 (visited on 12/01/2020).

Further Reading

John Rooksby et al. “Personal tracking as lived informatics”. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI ’14. Toronto, Ontario, Canada: Association for Computing Machinery, Apr. 26, 2014, pp. 1163–1172. ISBN: 978-1-4503-2473-1. DOI: 10.1145/2556288.2557039. URL: https://doi.org/10.1145/2556288.2557039 (visited on 05/13/2020).

Tsung-Sheng Chang and Wei-Hung Hsiao. “Time Spent on Social Networking Sites: Understanding User Behavior and Social Capital”. In: *Systems Research and Behavioral Science* 31.1 (2014). _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/sres.2169, pp. 102–114. ISSN: 1099-1743. DOI: 10.1002/sres.2169. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/sres.2169 (visited on 05/13/2020).

Yu Zhang et al. “Multi-kernel extreme learning machine for EEG classification in brain-computer interfaces”. In: *Expert Systems with Applications* 96 (Apr. 15, 2018), pp. 302–310. ISSN: 0957-4174. DOI: 10.1016/j.eswa.2017.12.015. URL: http://www.sciencedirect.com/science/article/pii/S0957417417308291 (visited on 05/13/2020).

Ji Peng Koh. “Brain computer interface via EEG signals”. In: (2018). Accepted: 2018-04-23T02:21:24Z. URL: https://dr.ntu.edu.sg//handle/10356/73955 (visited on 05/13/2020).

Gianluca Di Flumeri et al. “The Dry Revolution: Evaluation of Three Different EEG Dry Electrode Types in Terms of Signal Spectral Features, Mental States Classification and Usability”. In: *Sensors (Basel, Switzerland)* 19.6 (Mar. 19, 2019). ISSN: 1424-8220. DOI: 10.3390/s19061365. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6470960/ (visited on 05/13/2020).

Hermann Hinrichs et al. “Comparison between a wireless dry electrode EEG system with a conventional wired wet electrode EEG system for clinical applications”. In: *Scientific Reports* 10.1 (Mar. 23, 2020), pp. 1–14. ISSN: 2045-2322. DOI: 10.1038/s41598-020-62154-0. URL: https://www.nature.com/articles/s41598-020-62154-0 (visited on 05/13/2020).

Lamyaa Sadouk. “CNN Approaches for Time Series Classification”. In: *Time Series Analysis - Data, Methods, and Applications* (Nov. 5, 2018). Publisher: IntechOpen. DOI: 10.5772/intechopen.81170. URL: https://www.intechopen.com/books/time-series-analysis-data-methods-and-applications/cnn-approaches-for-time-series-classification (visited on 05/22/2020).

Christian Mühl et al. “A survey of affective brain computer interfaces: principles, state-of-the-art, and challenges”. In: *Brain-Computer Interfaces* 1.2 (Apr. 3, 2014). Publisher: Taylor & Francis _eprint: https://doi.org/10.1080/2326263X.2014.912881, pp. 66–84. ISSN: 2326-263X. DOI: 10.1080/2326263X.2014.912881. URL: https://doi.org/10.1080/2326263X.2014.912881 (visited on 05/22/2020).

Manuela Züger et al. “Reducing Interruptions at Work: A Large-Scale Field Study of FlowLight”. In: *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. CHI ’17. Denver, Colorado, USA: Association for Computing Machinery, May 2, 2017, pp. 61–72. ISBN: 978-1-4503-4655-9. DOI: 10.1145/3025453.3025662. URL: https://doi.org/10.1145/3025453.3025662 (visited on 06/13/2020).

Ji Young Jung et al. "Caffeine Maintains Arousal Level and Prevents Change of Electroencephalogram Spectral Powers with Time at Rest". In: *Journal of Korean Sleep Research Society* 11.1 (June 30, 2014). Publisher: Korean Sleep Research Society, pp. 5–10. ISSN: 1738-608X, 2288-4912. DOI: 10.13078/jksrs.14002. URL: http://www.e-jsm.org/journal/view.php?number=171 (visited on 08/03/2020).

PeterPutman. *The Effects of a Single Administration of a Moderate Dose of Caffeine on Cognitive Control and Spontaneous EEG Theta/Beta Ratio*. Clinical trial registration NCT02940808. submitted: October 17, 2016. clinicaltrials.gov, Mar. 7, 2017. URL: https://clinicaltrials.gov/ct2/show/NCT02940808 (visited on 07/31/2020).

Search: EEG | Kaggle. URL: https://www.kaggle.com/search?q=eeg (visited on 09/03/2020).

brainflow-dev/brainflow. original-date: 2018-07-15T12:42:26Z. Sept. 13, 2020. URL: https://github.com/brainflow-dev/brainflow (visited on 09/13/2020).

mne-tools/mne-python. original-date: 2011-01-28T03:31:13Z. Sept. 12, 2020. URL: https://github.com/mne-tools/mne-python (visited on 09/13/2020).

Craik A, He Y, and Contreras-Vidal Jl. *Deep learning for electroencephalogram (EEG) classification tasks: a review*. Journal of neural engineering. ISSN: 1741-2552 Issue: 3 Publisher: J Neural Eng Volume: 16. June 2019. DOI: 10.1088/1741-2552/ab0ab5. URL: https://pubmed.ncbi.nlm.nih.gov/30808014/ (visited on 09/21/2020).

Roy Y et al. *Deep learning-based electroencephalography analysis: a systematic review*. Journal of neural engineering. ISSN: 1741-2552 Issue: 5 Publisher: J Neural Eng Volume: 16. Aug. 14, 2019. DOI: 10.1088/1741-2552/ab260c. URL: https://pubmed.ncbi.nlm.nih.gov/31151119/ (visited on 09/21/2020).

NeuroTechX/eeg-notebooks. original-date: 2020-08-11T00:25:01Z. Sept. 20, 2020. URL: https://github.com/NeuroTechX/eeg-notebooks (visited on 09/21/2020).

Andrew Melnik et al. "Systems, Subjects, Sessions: To What Extent Do These Factors Influence EEG Data?" In: *Frontiers in Human Neuroscience* 11 (2017). Publisher: Frontiers. ISSN: 1662-5161. DOI: 10.3389/fnhum.2017.00150. URL: https://www.frontiersin.org/articles/10.3389/fnhum.2017.00150/full (visited on 09/21/2020).

Timeflux. URL: https://timeflux.io/ (visited on 09/21/2020).

Daniela Girardi et al. "Recognizing Developers' Emotions while Programming". In: *arXiv:2001.09177 [cs]* (Jan. 24, 2020). DOI: 10.1145/3377811.3380374. arXiv: 2001.09177. URL: http://arxiv.org/abs/2001.09177 (visited on 09/23/2020).

Arnd Meiser et al. "The Sensitivity of Ear-EEG: Evaluating the Source-Sensor Relationship Using Forward Modeling". In: *Brain Topography* (Aug. 24, 2020). ISSN: 1573-6792. DOI: 10.1007/s10548-020-00793-2. URL: https://doi.org/10.1007/s10548-020-00793-2 (visited on 10/05/2020).

Alexander Craik, Yongtian He, and Jose L. Contreras-Vidal. “Deep learning for electroencephalogram (EEG) classification tasks: a review”. In: *Journal of Neural Engineering* 16.3 (Apr. 2019). Publisher: IOP Publishing, p. 031001. ISSN: 1741-2552. DOI: 10.1088/1741-2552/ab0ab5. URL: https://doi.org/10.1088%2F1741-2552%2Fab0ab5 (visited on 10/29/2020).

Yannick Roy et al. “Deep learning-based electroencephalography analysis: a systematic review”. In: *Journal of Neural Engineering* 16.5 (Aug. 2019). Publisher: IOP Publishing, p. 051001. ISSN: 1741-2552. DOI: 10.1088/1741-2552/ab260c. URL: https://doi.org/10.1088%2F1741-2552%2Fab260c (visited on 10/29/2020).

Laurent Vézard et al. “EEG classification for the detection of mental states”. In: *Applied Soft Computing* 32 (July 1, 2015), pp. 113–131. ISSN: 1568-4946. DOI: 10.1016/j.asoc.2015.03.028. URL: http://www.sciencedirect.com/science/article/pii/S1568494615001866 (visited on 11/10/2020).

Axel Uran et al. “Applying Transfer Learning To Deep Learned Models For EEG Analysis”. In: *arXiv:1907.01332 [cs, eess, stat]* (July 2, 2019). arXiv: 1907.01332. URL: http://arxiv.org/abs/1907.01332 (visited on 11/23/2020).

Zitong Wan et al. “A review on transfer learning in EEG signal analysis”. In: *Neurocomputing* 421 (Jan. 15, 2021), pp. 1–14. ISSN: 0925-2312. DOI: 10.1016/j.neucom.2020.09.017. URL: http://www.sciencedirect.com/science/article/pii/S0925231220314223 (visited on 11/23/2020).

G. Gopan K, N. Sinha, and D. B. Jayagopi. “Alcoholic EEG Analysis Using Riemann Geometry Based Framework”. In: *2019 27th European Signal Processing Conference (EUSIPCO)*. 2019 27th European Signal Processing Conference (EUSIPCO). ISSN: 2076-1465. Sept. 2019, pp. 1–5. DOI: 10.23919/EUSIPCO.2019.8902506.

Hubert Banville et al. “Uncovering the structure of clinical EEG signals with self-supervised learning”. In: *Journal of Neural Engineering* (2020). ISSN: 1741-2552. DOI: 10.1088/1741-2552/abca18. URL: http://iopscience.iop.org/article/10.1088/1741-2552/abca18 (visited on 12/02/2020).