Using Mouse Trajectories for Predictions in Web Surveys

Tobias Wistuba[1], Lisa Bondo Andersen[2], Ailin Liu[2], Felix Henninger[2], Arie Kapteyn[3], Bart Orriens[3], Stefan Schneider[3], Frauke Kreuter[2], Sonja Greven[1]

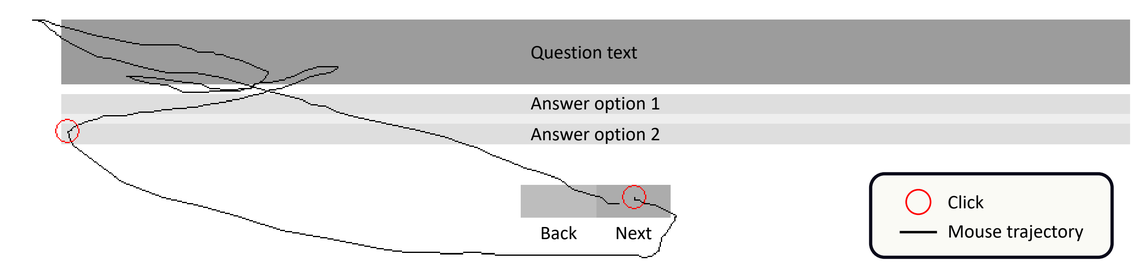
1 Humboldt Universität zu Berlin, 2 Ludwig-Maximillian-Universität München, 3 University of Southern California

Basic Idea

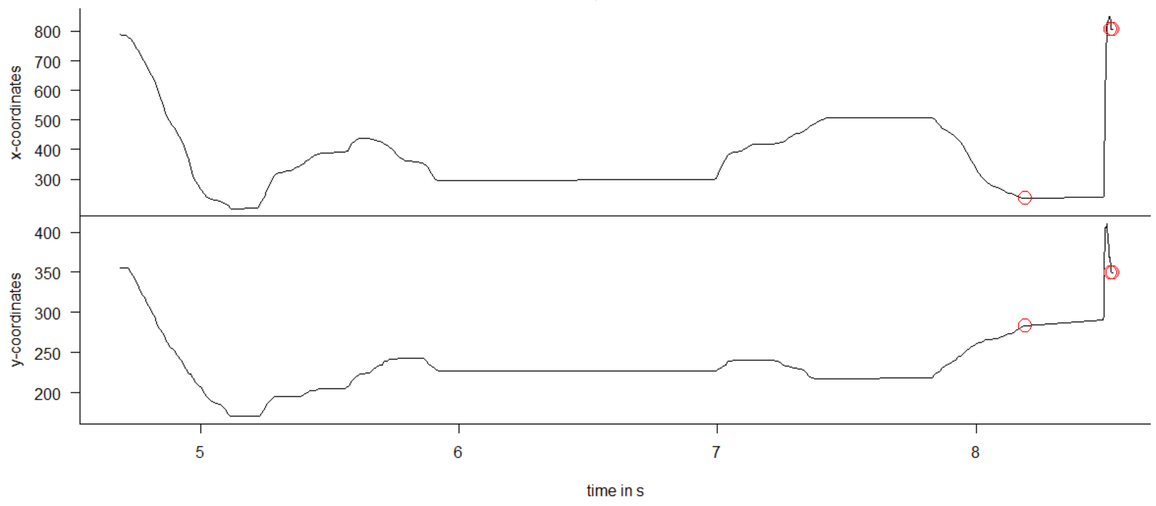
- Mouse trajectories can be collected „for free“ in web surveys as triples of shape (x-pos., y-pos, timestamp)
- Using Deep Learning to detect respondents with difficulties
  → Offering further information/help to improve data quality

The Data

Observed:

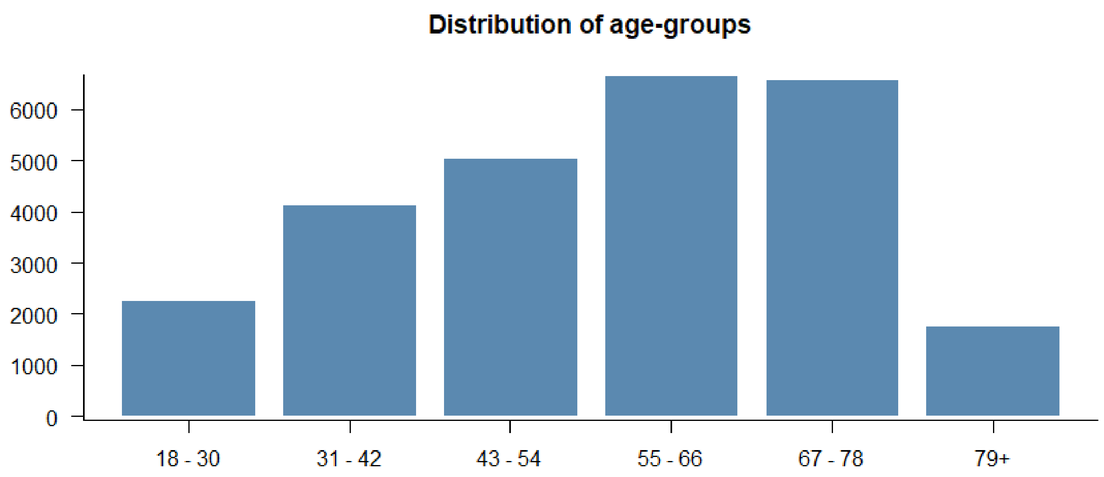


Recorded data



Current Progress

- Predicting self-assessed difficulty ratings of questions still work-in-progress
- Currently, predicting age-classes solely based on mouse trajectories via different neural network architectures for ordinal responses
- Extraction of mouse trajectory characteristics like velocity and acceleration
  → Using those as inputs besides the triples of (x-pos, y-pos, timestamp) as additional vectors or encoded in RGB channels of images (as CNN input)



Results

- Visualization of results via confusion matrices and selection of performance measures either averaged with equal weights (**Mac.**) or weighted by the number of observations per class (**Weighted**)
- Results shown for default fully connected neural network with 600k parameters (as baseline) and 2D-CNN with 100k parameters (because of its success in text recognition and generally image based rasks)
  → Comparison of time series/functional input vs image input

Results fully connected neural network

	Pred 18–30	Pred 31–42	Pred 43–54	Pred 55–66	Pred 67–78	Pred 79+	True
True 18–30	104	164	98	41	14	6	427
True 31–42	120	350	229	93	38	4	834
True 43–54	49	327	346	193	117	10	1042
True 55–66	27	224	427	385	300	24	1387
True 67–78	8	92	254	378	509	49	1290
True 79+	3	11	30	82	174	24	324
Pred	311	1168	1384	1172	1152	117	0.324

Metric	Value
Accuracy	0.324
Accuracy +/− 1 *	0.769
Mac. Precision	0.31
Weighted Precision	0.329
Mac. Recall	0.29
Weighted Recall	0.324
Mac. F1−Score	0.291
Weighted F1−Score	0.32
Mac. RPS−Score	0.733
Weighted RPS−Score	0.68
Mac. MAE	1.064
Weighted MAE	0.971
Mac. RMSE	2.008
Weighted RMSE	1.713

* +/- 1 age group deviation still counts as correct classivication

Results 2D-CNN

	Pred 18–30	Pred 31–42	Pred 43–54	Pred 55–66	Pred 67–78	Pred 79+	True
True 18–30	76	211	132	40	5	0	464
True 31–42	83	291	223	82	24	1	704
True 43–54	86	318	396	197	66	1	1064
True 55–66	52	254	448	451	200	43	1448
True 67–78	20	103	310	406	288	97	1224
True 79+	1	27	75	146	107	44	400
Pred	318	1204	1584	1322	690	186	0.291

Metric	Value
Accuracy	0.291
Accuracy +/− 1 *	0.723
Mac. Precision	0.288
Weighted Precision	0.31
Mac. Recall	0.268
Weighted Recall	0.291
Mac. F1−Score	0.263
Weighted F1−Score	0.287
Mac. RPS−Score	0.845
Weighted RPS−Score	0.767
Mac. MAE	1.153
Weighted MAE	1.061
Mac. RMSE	2.238
Weighted RMSE	1.941

* +/- 1 age group deviation still counts as correct classivication

Conclusion

- Age-group classification shows ability to learn from mouse trajectories
- Fully connected neural network shows good performance which can be matched with 2D-CNN with fewer parameters
  → Promising for development of own architectures
- Still challenges that need to be overcome for difficulty assessment

- Fernández-Fontelo, A., Kieslich, P. J., Henninger, F., Kreuter, F., & Greven, S. (2023). Predicting question difficulty in web surveys: A machine learning approach based on mouse movement features. Social Science Computer Review, 41(1), 141-162.
- Fernández-Fontelo, A., Henninger, F., Kieslich, P. J., Kreuter, F., & Greven, S. (2022). Classification ensembles for multivariate functional data with application to mouse movements in web surveys. arXiv preprint arXiv:2205.13380.
- Cao, W., Mirjalili, V., & Raschka, S. (2020). Rank consistent ordinal regression for neural networks with application to age estimation. Pattern Recognition Letters, 140, 325-331.a