Informatics in Medicine Unlocked 35 (2022) 101120

Contents lists available at ScienceDirect

Informatics in Medicine Unlocked

journal homepage: www.elsevier.com/locate/imu

Detection of mild cognitive impairment based on mouse movement data of trail making test

Gergely Hanczár [a,b], Erika Griechisch [a,b], Nóra Ovád [a], Olivér Máté Törteli [a], Gábor Tóth [b], Dávid Hanák [a,*], Balázs Vértes [c], András Horváth [d,e], Anita Kamondi [d,f]

[a] Cursor Insight Ltd., 20-22 Wenlock Road, N1 7GU, London, United Kingdom
[b] Patient Record, 20-22 Wenlock Road, N1 7GU, London, United Kingdom
[c] Precognize, Németvölgyi út 2, 1126, Budapest, Hungary
[d] National Institute of Mental Health, Neurology and Neurosurgery, Amerikai út 57, 1145, Budapest, Hungary
[e] Semmelweis University, Department of Anatomy, Histology and Embryology, Tűzoltó u. 58, 1094, Budapest, Hungary
[f] Semmelweis University, Department of Neurology, Balassa u. 6, 1083, Budapest, Hungary

ARTICLE INFO

Keywords:
Mouse movement
Mild cognitive impairment
Early detection
Machine learning

ABSTRACT

Mild cognitive impairment (MCI) has 10%–20% prevalence in the population above the age of 65, and a significant portion of these people will go on to develop dementia later in their lives. However, if MCI is detected early, preventative measures can be taken to delay the onset of severe symptoms. Current diagnostic methods for MCI are not suitable for regular wide-scale screening. Advances in machine learning algorithms in combination with digital movement data offer rich possibilities for automated MCI detection. This paper introduces a machine learning model that effectively predicts MCI based on only a few seconds of computer mouse movement. To our knowledge, studies directly comparable to ours have not been done before. On a dataset of 70 participants, we demonstrated 80% accuracy in distinguishing healthy controls from patients with MCI. This gives an opportunity to develop a cost-efficient and easy-to-use screening method that could aid the work of healthcare professionals.

1. Introduction

Dementia affects over 50 million people worldwide and causes 2.4 million deaths a year. By 2050, the number of sufferers is predicted to increase to over 150 million [1]. Dementia is currently the 7th leading cause of death. In the coming decades, however, it will become the top mortality cause among elderly people [2], unless suitable early screening and prevention methods are introduced.

10%–20% of the world population above 65 years of age has mild cognitive impairment (MCI), which is often a precursor to dementia [3]. Early detection of MCI allows preventative actions to be taken before the condition develops to a stage where it significantly affects the patient's quality of life [4]. There is evidence that lifestyle changes, non-pharmacological approaches, and medical treatment can effectively delay the onset of the most debilitating symptoms [5–7]. Slowing the progression of dementia even by as little as one year could eliminate more than 9 million cases by 2050 [8].

The current diagnostic method for MCI is a medical workup including cognitive assessment tests, neurological exams, laboratory tests, and in some cases, brain imaging. This is not suitable for wide-scale screening of the population at risk because of the time-consuming and/or expensive nature of these tools, and the fact that they require highly skilled medical personnel. Full or partial automation of neurocognitive assessment protocols could eventually replace medical staff to some extent [9–11], providing cost-efficient pre-screening methods that decrease the pressure on healthcare systems.

Machine learning (ML) based movement analysis could provide significant advances in developing an automated MCI screening tool. Studies using in-home monitoring sensor technologies have shown that movement patterns of patients can be indicative of cognitive impairment [12,13]. ML has been used for gait assessment to effectively detect MCI [14]. Eye movement analysis with machine learning algorithms also seems promising based on existing research [15]. Our focus, however, has been on examining hand movements through recording and analysing mouse movement data. This has the potential to be the basis of a non-invasive MCI detection tool that supports wide-scale screening

* Correspondence to: Mester utca 1 3/18, Budapest, H-1095, Hungary.
E-mail addresses: gergely@cursorinsight.com (G. Hanczár), erika@cursorinsight.com (E. Griechisch), nora@cursorinsight.com (N. Ovád), oliver@cursorinsight.com (O.M. Törteli), gabor@patient-record.com (G. Tóth), david@cursorinsight.com (D. Hanák), bvertes@gmail.com (B. Vértes), andras.horvath.semmelweis@gmail.com (A. Horváth), kamondianita@gmail.com (A. Kamondi).

https://doi.org/10.1016/j.imu.2022.101120
Received 15 August 2022; Received in revised form 26 October 2022; Accepted 26 October 2022
Available online 29 October 2022

Table 1
Descriptive statistics about the patients. The categorical variables are given as quantities, the continuous variables are given as mean ± standard deviation. The mean age is significantly higher in the MCI group.

	MCI	Healthy
Count	22	48
Gender (male/female/unspecified)	10/12/0	18/27/3
Age (years)	71.18 ± 5.82	66.88 ± 7.25
Education (years)	14.40 ± 2.37	15.43 ± 2.09

Table 2
Descriptive statistics about the mouse movement data. Three tasks (DTMT = Digitised Trail Making Test, PM = Pair Matching game) were performed with right and left hands.

Task	Hand		Total
	Left	Right	
DTMT-1	67	70	137
DTMT-2	67	70	137
PM	47	50	97
Total	181	190	371

in a cost-efficient way, since no special hardware is required, in contrast with other movement analysis based approaches. A computer with a mouse is sufficient.

In this paper, we propose an ML prediction model that analyses a few seconds of mouse movement data to differentiate subjects with some level of cognitive decline from healthy controls. Computer use has been examined in connection with cognitive impairment before, however, these studies investigated the association with MCI using statistical approaches [16]. Notable work relevant to our objective is an exploratory study by Seelye et al. that analysed mouse movement patterns of participants collected over the course of a week, and the results showed that these patterns were different for cognitively impaired and healthy subjects in a statistically significant way [17]. We are not aware of prior work that uses machine learning algorithms on mouse movement data to detect MCI, and thus would be directly comparable to ours.

2. Method

The analysis relies on data gathered at the National Institute of Mental Health, Neurology and Neurosurgery, Neurocognitive Research Center in Budapest, Hungary (NIMHNN). In the framework of the Hungarian National Brain Research Program II, NIMHNN's research group conducted a clinical observation called Precognize Pilot Study (PPS) between 2017 and 2021. The data used in our work were provided by NIMHNN. The data have been anonymised, thus no personally identifiable information was exchanged. The original research was authorised by the Hungarian Medical Research Council with reference number 024505/2015/OTIG. All participants signed an informed consent form.

2.1. Subjects

Seventy individuals took part in the PPS (22 patients with MCI and 48 healthy controls), who were recruited from the AlzEpi Cohort Observational Library (ACOL database) of NIMHNN. The library is part of the Euro-Fingers international database.[1]

All participants took part in detailed diagnostic procedures. The test battery included neurological examination, blood tests, neuropsychological tests, and structural MRI. Neuropsychological tests consisted of Mini Mental State Examination (MMSE), Addenbrooke Cognitive Examination (ACE), Rey Auditory Verbal Learning Test (RAVLT), and Trail Making Tests A and B (TMT-A and TMT-B). These tests had been selected because they are well established methods for sensitively identifying cognitive impairment via the analysis of different cognitive domains [18–21]. Neuropsychological tests were administered by trained neuroscientists, neurologists, or neuropsychologists. Metadata, such as *group* (0 for healthy, 1 for MCI), *age* (in years), *gender* (1 for male, 2 for female), and *education* (in years) were also collected. Every participant in the PPS was right-handed. See Table 1 for more details.

Control subjects met the following criteria: negative neurological status; cognitive performance in normal range based on the cut-off scores of the applied neuropsychological test battery; absence of significant lesions or cortical atrophy on brain MRI. The diagnosis of MCI was based on the revised Petersen criteria [22]. MCI population fulfilled the proposed criteria: negative neurological status; objective impairment in neuropsychological performance; presence of cortical thinning of the entorhinal cortex, and the reduction of total grey matter volume confirmed by MRI.

2.2. Data collection

Mouse movement information was acquired from participants with the use of a browser-based testing programme developed specifically for PPS. The tests were conducted on a single computer using the same mouse and mouse pad during the full span of the PPS. In the test, participants had to perform exercises first with their right and then with their left hands. Subjects performed two digitised tests inspired by TMT-A and a Pair Matching (PM) game. In the digitised trail making tests (DTMT), participants have to connect circled numbers from 1 to 9 in growing order by moving the mouse cursor over them. In the first test (DTMT-1), all the circles are visible during the full duration of the task. The second test (DTMT-2) repeats the full sequence, but this time only the circle that comes next in the sequence is visible, all the others are hidden. In the PM game, which is a kind of memory test, individuals are shown a grid of 16 cards and must click on them in the right order to identify the eight matching pairs. During all three exercises, the programme collects the position of the mouse on the screen (at 60 Hz), and the timestamp and position of the mouse actions performed by the participants during the game, such as scrolling the wheel or clicking on a target. For the DTMT tasks, the number of the circle that comes next in the sequence is also collected. The number of completed exercises and the used hand are shown in Table 2 for each task.

2.3. Using the Rey Auditory Verbal Learning Test

RAVLT [23] is one of the most widely used word learning tests in clinical practice to assess memory and learning abilities [24,25]. In RAVLT, subjects need to memorise a list of 15 words and recall them with 5 repetitions. The performance is expressed in the `RAVLT-sum5` score, which is the total number of correctly recalled words, hence lower scores mean lower cognitive performance. The validated Hungarian version of RAVLT [26] was administered to objectively assess memory complaints according to the Petersen criteria [27].

Previous studies revealed its prominent sensitivity in the detection of amnestic MCI [19] due to the early involvement of verbal learning-oriented memory functions. The participants of our current study were carefully examined for the presence of MCI by a multidisciplinary medical team using the detailed procedures mentioned above, and `RAVLT-sum5` showed the strongest correlation with the diagnosis. This is shown in Fig. 1, where normalised `RAVLT-sum5` and other test scores (MMSE, ACE, and RAVLT with 7 repetitions) are plotted for both groups. Therefore, we used the `RAVLT-sum5` score to ensure that participants are sampled representatively when selecting training samples for our models. This score, however, was not used directly during the model training and evaluation phases.

The `RAVLT-sum5%` value, used in the rest of this document, is the `RAVLTsum5` score expressed as a percentage of the theoretical maximum of 75. In the PPS, the range of the `RAVLT-sum5%` scores

[1] https://www.eufingers.com/, retrieved on October 12, 2022.

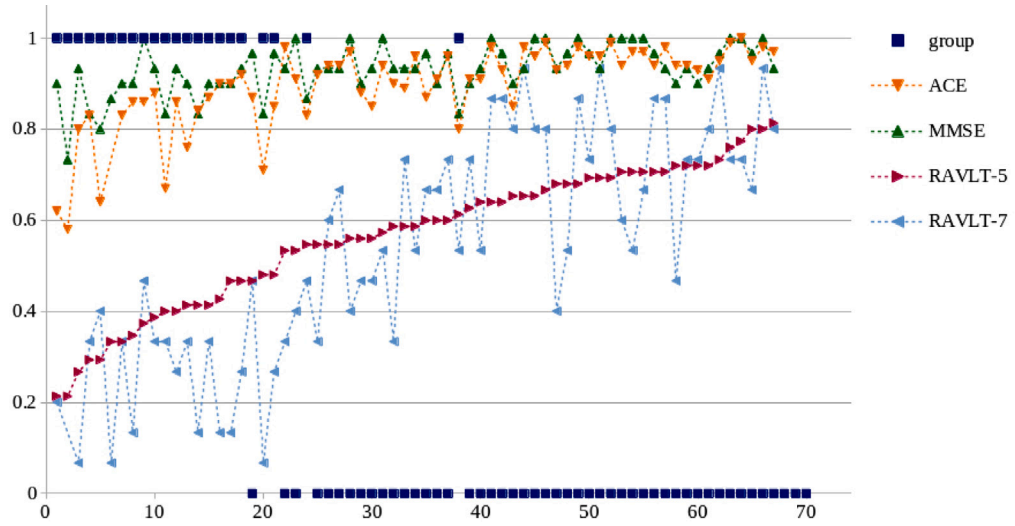

Fig. 1. Various test scores of the participants. (Group 0: healthy subjects, group 1: subjects with MCI.) `RAVLT-sum5` score shows correlation with the diagnosis.

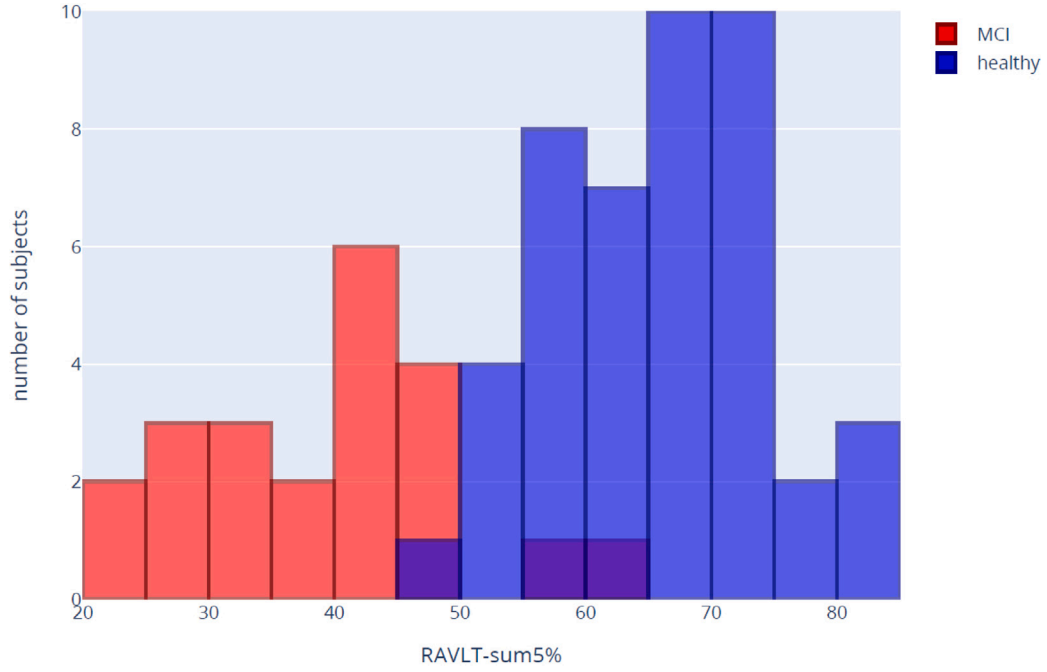

Fig. 2. Histogram of the `RAVLT-sum5%` score. Red is MCI, blue is healthy. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

was 46.67–81.33 for healthy controls and 21.33–61.33 for patients with MCI [28]. The distribution of this score is shown as two overlaid histograms in Fig. 2.

2.4. Data analysis

The main steps of the data analysis were: data cleaning (including reconstruction of the pointer device movement, which is the closest we can get to reconstructing the true movement of the subject's hand from this data), subset bundling, feature extraction using two different algorithms, data augmentation, prediction model building, model evaluation, and visualisation.

Due to the nature of the web-based data collection environment, significant data preparation had to be conducted. The collector is capable of measuring the coordinates of the mouse cursor. To analyse the movement patterns of the subjects, however, we needed to reconstruct the pointer device movements that generated the recorded cursor movements. There are several properties of the hardware and software involved that make this task non-trivial. For example, because of the operating system-level mouse acceleration feature, the relationship between mouse and cursor speed is non-linear. Cursor Insight's (CI)[2] proprietary mouse data cleaning toolchain was applied to reconstruct pointer device movement data as accurately as possible.

Owing to the limited size of the dataset, and also aiming for interpretable prediction models, decision trees became our primary choice of ML algorithm. There are numerous other classification techniques, including neural networks, that were not covered in the current study because of the aforementioned reasons.

Reasonably small decision trees are generally a good choice when there are few samples to train on, as in our case. To improve the

[2] https://cursorinsight.com/.

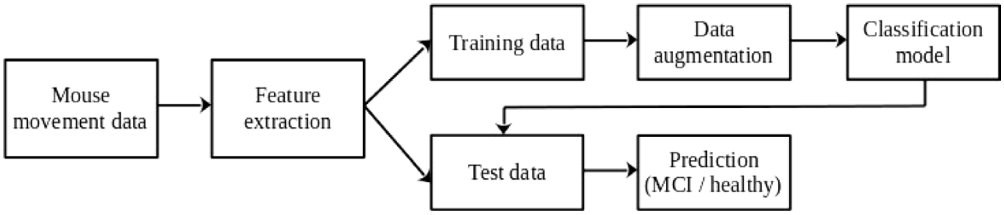

Fig. 3. Steps of the classification algorithm.

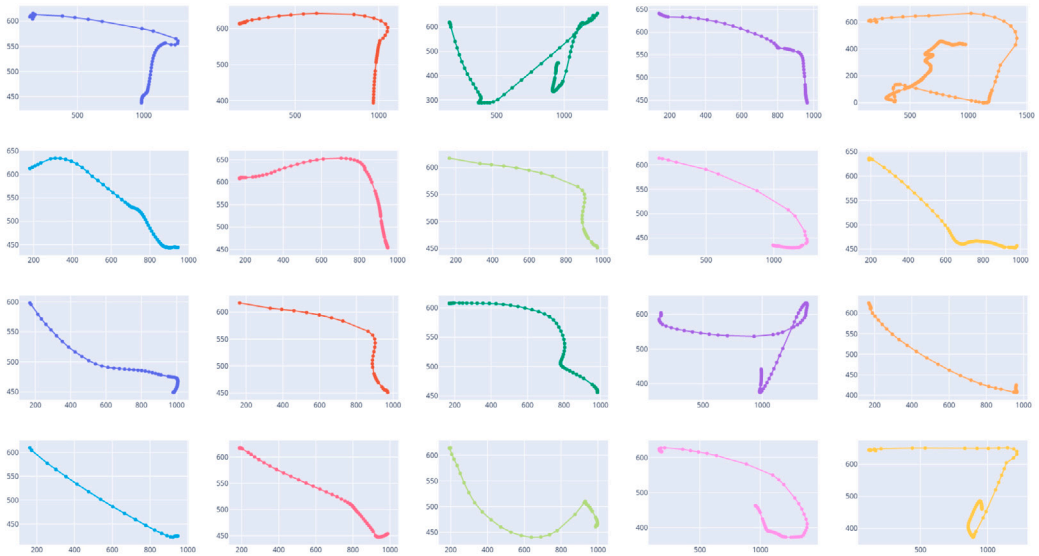

Fig. 4. x–y plots of a selection of mouse motion samples. The participants had to perform an ideally diagonal motion.

situation, a technique called *data augmentation* [29] was applied to artificially boost the number of training samples. The usual measures, such as limiting tree depth and impurity decrease, using separate training and test sets, and cross-validation (via repeated randomised data augmentation) were also taken to avoid overfitting.

In the following sections, we describe how the dataset was split into training and test sets, what feature sets were applied, and our approach to generate samples using data augmentation. The flowchart in Fig. 3 illustrates the steps of the algorithm.

2.4.1. Training and test set separation

The PPS dataset was split into two subsets: test data, consisting of the samples of 14 people (7 with diagnosed MCI and 7 healthy; 20% of participants); and training data, consisting of the samples of everyone else (80% of participants). Furthermore, we focused our attention on:

1. only one of the three tasks that the subjects had to complete (DTMT-1);
2. the exercise performed with the right hand;
3. a single stroke of mouse movement within the exercise, where every subject had to move the mouse from one fixed point on the screen to another (connecting circles number 4 and 5).

The test subjects were chosen such that the distribution of their medical condition roughly matched that of the full population of the 70 participants, according to their `RAVLT-sum5%` scores. A small selection of the inspected samples (performing an ideally diagonal motion) can be seen in Fig. 4. Note that the average duration of the considered strokes was no more than two seconds, which means 100–200 data points at most.

After separating the test set, samples from 56 people (15 MCI, 41 healthy) remained to train our model.

2.4.2. Feature sets

Two different feature sets were used to measure the performance of the applied ML algorithms: a small *baseline set* and three features (chosen by their performance in the classification) from CI's proprietary feature space. The latter, consisting of up to several thousand scalar features, has been crafted and refined in various research and commercial projects, which all relied on movement analysis, mouse movement in particular.

Before computing the features, a Savitzky–Golay filter [30] was applied to the raw time series of x and y coordinates to remove the undesired effects of discrete sampling and integer rounding, and restore the original, continuous motion patterns as closely as possible. The filter had a window radius of 10 and a polynomial order of 3. These parameters were found to be the best for smoothing out noise that did not originate from human motion, while leaving acceleration and jerk detectable in the coordinates.

Baseline feature set. The baseline set comprises 66 (6×11) features: 11 statistical values (minimum, maximum, mean, standard deviation, skewness, kurtosis, and 10th, 25th, 50th, 75th, 90th percentile values) of six different properties: $x(t)$, $y(t)$, $\Delta x(t)$, $\Delta y(t)$, $\Delta\Delta x(t)$, and $\Delta\Delta y(t)$. Here $x(t), y(t)$ denote the coordinates of the mouse cursor at a given time t, the derivatives $\Delta x(t), \Delta y(t)$ express the instantaneous velocity, and the second derivatives $\Delta\Delta x(t), \Delta\Delta y(t)$ express the instantaneous acceleration.

Reduced CI feature set. In the case of the PPS, the larger CI feature set included 441 features for each sample. With preliminary experiments, however, we identified three features which were sufficient for building a basic model. In the following, this feature set has been used, and we will refer to it as *reduced CI feature set*. The reduction decreases the complexity of the problem and helps to avoid overfitting.

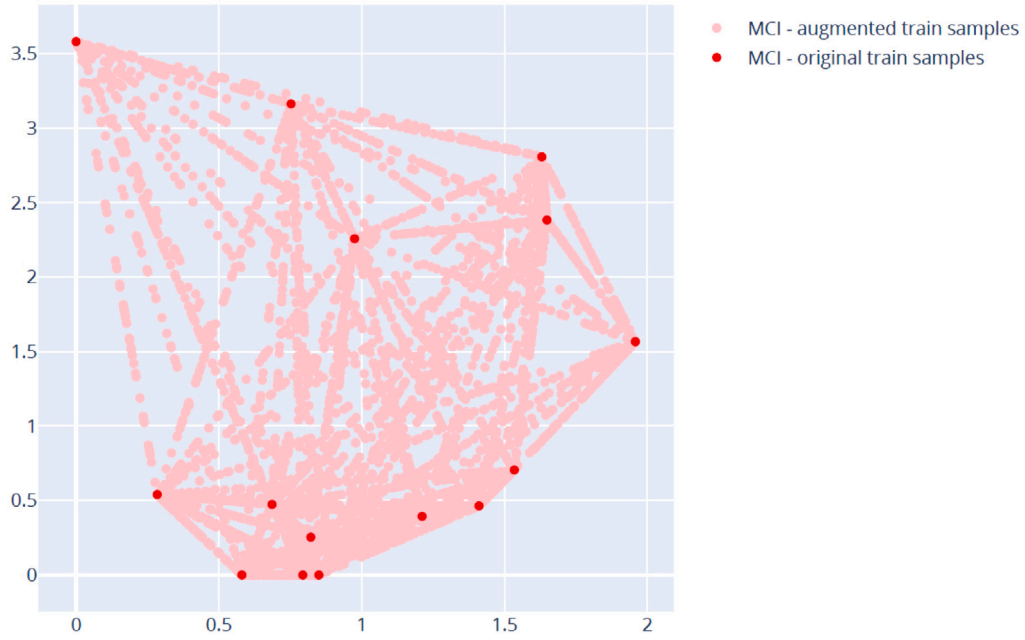

Fig. 5. Augmentation of training data for two features. Each colour coded dot represents an instance: red — original, pink — augmented/generated. The generation method works similarly in higher dimensions. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

2.4.3. Training data generation with data augmentation

To generate more training data, data augmentation by linear combination [29] was used after extracting features from the original training samples. Although there are undoubtedly more sophisticated augmentation techniques, this approach was used because of its simplicity, scalability, and domain independence, and because it generates a robust sample space with large amounts of new features.

For every generated feature representation, two of the original samples were randomly selected, and a convex combination of them was computed. The weights used to combine the input features were chosen randomly from a uniform distribution over $[0, 1]$. This method guarantees that all generated feature representations are within the convex hull of the representations of the original training samples. Note, however, that the generated representations are all located along the multidimensional segments connecting the vertices of the original samples and are not evenly distributed inside this convex hull, see Fig. 5 for a 2-D example.

2.4.4. Classifier

For the classification of the samples, a decision tree was used with a maximum depth of 2. This ensures that the results are interpretable, which is preferred over black box methods for medical applications [31]. Other parameters of the decision tree were: *n_subfeatures = -1, min_samples_leaf = 4, min_samples_split = 4, min_purity_increase = 0.1*. (n_subfeatures = −1 means that the number of features to consider at random per split is the square root of the number of features.)

In theory, higher accuracy could be achieved given sufficient amounts of training data using random forests instead of single trees. With the PPS dataset, however, the application of forests would result in overfitting and less interpretable models.

A visual representation of one such classification tree is shown in Fig. 6, which displays a projection of 14 test samples onto the 2-D plane. The horizontal axis shows the values of the primary feature chosen as the root of the tree, while the vertical axes to the left and right of the vertical divide show the values of features chosen for the left and the right branches of the tree, respectively. The actual scales of the axes are irrelevant. Blue dots represent healthy subjects, whereas red dots represent subjects with diagnosed MCI. Likewise, blue areas are predicted as healthy, whereas red areas are predicted as patients with MCI by the model.

Table 3
Averaged confusion matrix using the baseline feature set.

		Predicted	
		Healthy	MCI
Diagnosed	Healthy	3.02 (22%)	3.98 (28%)
	MCI	1.01 (7%)	5.99 (43%)

Table 4
Averaged confusion matrix using the reduced CI feature set.

		Predicted	
		Healthy	MCI
Diagnosed	Healthy	5.32 (38%)	1.68 (12%)
	MCI	1.15 (8%)	5.85 (42%)

3. Results

Two specific parameters of the experiments were grid searched. 1000 to 5000 (in steps of 1000) training and an equal number of test samples per class (healthy vs. MCI) were generated with randomised data augmentation, which was repeated 50 to 500 times (in steps of 50), using a different (but deterministic) random seed every time. For each experiment, both feature sets were evaluated the given number of times using random forests, and the aggregate mean of the confusion matrices was calculated. The results presented below were computed using an augmentation of 5000 samples and 200 iterations of randomised augmentation. Fewer than 200 repetitions resulted in larger deviations in the cells of the averaged confusion matrices, and more than that did not cause a significant improvement. Using 5000 training and test samples was still feasible, and it also produced the least variance in the results. As a result of the uniform size of the augmented sets, the ratio of MCI to healthy subjects both in the training and test data came out as 50%–50%.

Using the baseline feature set, healthy participants seem to be classified randomly. In contrast, only 1 out of 7 MCI participants were misclassified on average, see the averaged confusion matrix in Table 3.

Using the reduced CI feature set, an 80% averaged accuracy could be achieved, see the averaged confusion matrix in Table 4.

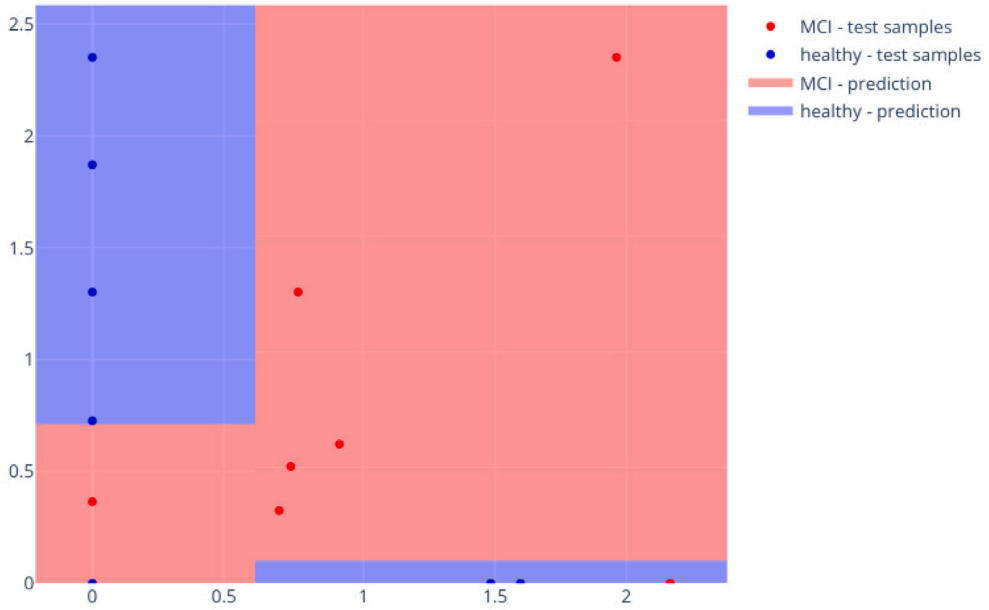

Fig. 6. A projection of the 14 test samples onto the 2-D plane (dots coloured according to diagnosis), and the segmentation of the plane into prediction classes (areas coloured according to prediction); blue — healthy, red — MCI. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 5
Comparison of metrics calculated from the averaged confusion matrices.

	Baseline	Reduced CI
Sensitivity	85.6%	83.6%
Specificity	43.1%	76%
Precision	60.1%	77.7%
Accuracy	64.4%	79.8%
F1-score	70.6%	80.5%

From these numbers, we calculated some commonly used metrics to compare the performance of the two feature sets, see Table 5. When using the reduced CI feature set, 83.6% of patients who truly have MCI will get correctly diagnosed, and 76% of patients who do not have MCI will be correctly classified as healthy. Thus, the false positive rate is 24%, the false negative rate is 16.4%. While the sensitivity is slightly better in the case of the baseline feature set, specificity and other metrics are significantly worse.

Due to the small number of samples, runtime of our algorithm was not significant, hardware and software requirements were also minimal. It is important to mention that the small sample size is not ideal and a weakness of our result. It is likely that we would be able to build a more robust and accurate model using a larger dataset, which in turn would be more resource intensive.

4. Conclusion and future work

People with MCI and healthy people were distinguished using only a few seconds of computer mouse data, dynamic features, and a single decision tree.

In previous research projects, the proprietary CI feature set had successfully been used to train models to recognise or classify large numbers of individuals based on their motion patterns, given that there were enough samples to learn from. In this case, however, considering the small amount of samples at our disposal, special care was necessary to avoid model overfitting. To achieve that, we employed data augmentation and used a restricted feature space, applied an augmented training set and representative test set, and took the customary, industry standard measures to ensure that overfitting stayed within acceptable limits in the generated ML model.

The results indicate that MCI can be predicted effectively from mouse movement data using machine learning techniques. The approximately 80% accuracy, even with a fairly low confidence, is a major achievement, considering that only a few seconds of mouse navigation were examined. The 83.6% sensitivity and 76% specificity values are also remarkable. Our result holds the promise of developing a cost-efficient and automated MCI screening tool, which bears special significance, since current diagnostic methods are not suitable for wide-scale screening. Generating interpretable models is an added benefit, given that such methods are preferred in medical applications, as doctors have higher confidence in them than in black box solutions.

The next step could be a larger scale study with significantly more collected samples, as this would produce ML models with potentially higher accuracy rates. Furthermore, these models could also be tuned to achieve (or stay below) specific target false positive and/or false negative rates, and thus reach a target sensitivity/specificity ratio that fits the health economy optimum.

Having a larger dataset would also open up feasible opportunities to research certain areas in greater depth in the future. These areas, that push the boundaries of our current research, include, but are not limited to:

- various other machine learning and classification algorithms, such as support vector machines and neural networks;
- more sophisticated data augmentation techniques, which capture the dimensionality and arrangement of the samples better;
- the effect of data augmentation on accuracy and generalisation capability;
- a more thorough grid search of the hyperparameters;
- proper N-fold cross-validation.

That being said, acquiring large amounts of labelled data is not a trivial task, as such data are not readily available in medical circles. Nevertheless, the option is worth investigating.

5. About the participants

Cursor Insight has developed a proprietary toolchain which is able to analyse time series via computation of a large-scale feature set and application of ML technologies. This toolchain has proved to be effective on human motion data analysis for both academic and real-life

applications in various areas in the past decade. In 2015, CI won 1th prize in the largest global competition on automatic on-line (dynamic) handwritten signature verification on skilled forgeries [32]. In 2016, CI won the Mouse Dynamics Challenge competition of Balabit [33], a leading cyber-security company in Hungary. Since 2016, CI has been operating signature verification systems in sectors with high security demands. **Patient Record**, a subsidiary of CI, is focusing on healthcare applications.

Precognize was founded by Balázs Vértes with healthcare at its focus, to research the possibility of early detection of Alzheimer's Disease via hand movements.

Anita Kamondi, MD, PhD, DSc is the director of neurology services and head of the Department of Neurology at the National Institute of Mental Health, Neurology and Neurosurgery, Budapest, Hungary, and full professor of neurology at the Department of Neurology, Semmelweis University, Budapest. **András Horváth, MD, PhD** is the head of the Neurocognitive Research Center at the National Institute of Mental Health, Neurology and Neurosurgery.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Gergely Hanczár, PhD reports a relationship with Cursor Insight that includes: board membership and employment. Erika Griechisch, PhD reports a relationship with Cursor Insight that includes: employment. Nóra Ovád reports a relationship with Cursor Insight that includes: employment. Olivér Máté Törteli reports a relationship with Cursor Insight that includes: employment. Gábor Tóth reports a relationship with Patient Record that includes: board membership. Dávid Hanák reports a relationship with Cursor Insight that includes: employment. Cursor Insight as a company has financial interest in movement analysis, mouse movement in particular. One of the primary focus areas is the potential medical applications of movement analysis.

Acknowledgement

The authors of the manuscript "*Detection of mild cognitive impairment based on mouse movement data of trail making test*" submitted for publication to *Informatics in Medicine Unlocked* would like to acknowledge the support and contribution of:

- the National Institute of Mental Health, Neurology and Neurosurgery, Neurocognitive Research Center in Budapest, Hungary for conducting the clinical observation called Precognize Pilot Study, and providing us the anonymised data;
- Ágnes Chripkó, PhD, for thoroughly revising the manuscript based on the received reviews;
- Bence Borbély, PhD, for an internal review of the manuscript before submission;
- Tamás Zelczer, CEO of Cursor Insight, for proofreading the manuscript.

References

[1] Guerchet M, Prince M, Prina M. Numbers of people with dementia worldwide. 2020, URL: https://www.alzint.org/resource/numbers-of-people-with-dementia-worldwide/, retrieved on October 12, 2022.

[2] Prince M, Ali G-C, Guerchet M, Prina AM, Albanese E, Wu Y-T. Recent global trends in the prevalence and incidence of dementia, and survival with dementia. Alzheimers Res Ther 2016;8:1–13.

[3] Bai W, Chen P, Cai H, Zhang Q, Su Z, Cheung T, et al. Worldwide prevalence of mild cognitive impairment among community dwellers aged 50 years and older: a meta-analysis and systematic review of epidemiology studies. Age Ageing 2022;51. http://dx.doi.org/10.1093/ageing/afac173.

[4] Chehrehnegar N, Nejati V, Shati M, Rashedi V, Lotfi M, Adelirad F, et al. Early detection of cognitive disturbances in mild cognitive impairment: a systematic review of observational studies. Psychogeriatrics 2020;20:212–28. http://dx.doi.org/10.1111/psyg.12484.

[5] Sperling R, Mormino E, Johnson K. The evolution of preclinical Alzheimer's disease: implications for prevention trials. Neuron 2014;84:608–22.

[6] Rasmussen J, Langerman H. Alzheimer's disease–why we need early diagnosis. Degener Neurol Neuromuscul Dis 2019;9:123.

[7] Livingston G, Huntley J, Sommerlad A, Ames D, Ballard C, Banerjee S, et al. Dementia prevention, intervention, and care: 2020 report of the Lancet Commission. Lancet 2020;396:413–46.

[8] Brookmeyer R, Johnson E, Ziegler-Graham K, Arrighi HM. Forecasting the global burden of Alzheimer's disease. Alzheimers Dement 2007;3:186–91.

[9] Oliveira MOd, Brucki SMD. Computerized Neurocognitive Test (CNT) in mild cognitive impairment and Alzheimer's disease. Dement Neuropsychol 2014;8:112–6.

[10] Égerházi A, Berecz R, Bartók E, Degrell I. Automated Neuropsychological Test Battery (CANTAB) in mild cognitive impairment and in Alzheimer's disease. Prog Neuropsychopharmacol Biol Psychiatry 2007;31:746–51.

[11] Snyder PJ, Jackson CE, Petersen RC, Khachaturian AS, Kaye J, Albert MS, et al. Assessment of cognition in mild cognitive impairment: a comparative study. Alzheimers Dement 2011;7:338–55.

[12] Lussier M, Lavoie M, Giroux S, Consel C, Guay M, Macoir J, et al. Early detection of mild cognitive impairment with in-home monitoring sensor technologies using functional measures: A systematic review. IEEE J Biomed Health Inform 2019;23:838–47. http://dx.doi.org/10.1109/JBHI.2018.2834317.

[13] Khan T, Jacobs PG. Prediction of mild cognitive impairment using movement complexity. IEEE J Biomed Health Inform 2021;25:227–36.

[14] Ghoraani B, Boettcher LN, Hssayeni MD, Rosenfeld A, Tolea MI, Galvin JE. Detection of mild cognitive impairment and Alzheimer's disease using dual-task gait assessments and machine learning. Biomed Signal Process Control 2021;64:102249.

[15] Lagun D, Manzanares C, Zola SM, Buffalo EA, Agichtein E. Detecting cognitive impairment by eye movement analysis using automatic classification algorithms. J Neurosci Methods 2011;201:196–203.

[16] Stringer G, Couth S, Brown LJE, Montaldi D, Gledson A, Mellor J, et al. Can you detect early dementia from an email? A proof of principle study of daily computer use to detect cognitive and functional decline. Int J Geriatr Psychiatry 2018;33:867–74.

[17] Seelye A, Hagler S, Mattek N, Howieson DB, Wild K, Dodge HH, et al. Computer mouse movement patterns: A potential marker of mild cognitive impairment. Alzheimers Dement (Amst) 2015;1:472–80. http://dx.doi.org/10.1016/j.dadm.2015.09.006.

[18] Arevalo-Rodriguez I, Smailagic N, Figuls MRi, Ciapponi A, Sanchez-Perez E, Giannakou A, et al. Mini-mental state examination (MMSE) for the detection of Alzheimer's disease and other dementias in people with mild cognitive impairment (MCI). In: Cochrane database syst rev. 2015.

[19] Alladi S, Arnold R, Mitchell J, Nestor PJ, Hodges JR. Mild cognitive impairment: applicability of research criteria in a memory clinic and characterization of cognitive profile. Psychol Med 2006;36:507–15.

[20] Crawford S, Whitnall L, Robertson J, Evans JJ. A systematic review of the accuracy and clinical utility of the Addenbrooke's cognitive examination and the Addenbrooke's cognitive examination—Revised in the diagnosis of dementia. Int J Geriatr Psychiatry 2012;27:659–69.

[21] Stanczak DE, Lynch MD, McNeil CK, Brown B. The expanded trail making test: rationale, development, and psychometric properties. Arch Clin Neuropsychol 1998;13:473–87.

[22] Petersen RC, Roberts RO, Knopman DS, Boeve BF, Geda YE, Ivnik RJ, et al. Mild cognitive impairment: ten years later. Arch Neurol 2009;66:1447–55.

[23] Rey A. Auditory verbal learning test, psychological appraisal of children with cerebral deficits. 1964.

[24] Ferreira Correia A, Campagna Osorio I. The Rey Auditory Verbal Learning Test: Normative Data Developed for the Venezuelan Population. Arch Clin Neuropsychol 2013;29:206–15. http://dx.doi.org/10.1093/arclin/act070.

[25] Hall J, O'Carroll RE, Frith CD. 7 - neuropsychology. In: Johnstone EC, Owens DC, Lawrie SM, McIntosh AM, Sharpe M, editors. Companion to psychiatric studies. 8th ed.. Churchill Livingstone; 2010, p. 121–40. http://dx.doi.org/10.1016/B978-0-7020-3137-3.00007-3.

[26] Kónya A, Verseghi A, Rey T. Rey: emlékezeti vizsgálatok. Budapest: Pszicho-Teszt; 1995.

[27] Petersen RC. Mild cognitive impairment as a diagnostic entity. J Intern Med 2004;256:183–94.

[28] Csukly G, Sirály E, Fodor Z, Horváth A, Salacz P, Hidasi Z, et al. The differentiation of amnestic type MCI from the non-amnestic types by structural MRI. Front Aging Neurosci 2016;8:52.

[29] Chawla NV, Bowyer KW, Hall LO, Kegelmeyer WP. SMOTE: synthetic minority over-sampling technique. J Artif Intell Res 2002;16:321–57.

[30] Press WH, Teukolsky SA. Savitzky-Golay smoothing filters. Comput Phys 1990;4:669–72.

[31] McKelvey T, Ahmad M, Teredesai A, Eckert C. Interpretable machine learning in healthcare. IEEE Intell Inform Bull 2018;19.

[32] Malik MI, Ahmed S, Marcelli A, Pal U, Blumenstein M, Alewijns L, et al. ICDAR2015 competition on signature verification and writer identification for on-and off-line skilled forgeries (SigWIcomp2015). In: 2015 13th international conference on document analysis and recognition. ICDAR, IEEE; 2015, p. 1186–90.

[33] Füöp Á, Kovács L, Kurics T, Windhager-Pokol E. Balabit Mouse Dynamics Challenge data set. 2016, URL: https://github.com/balabit/Mouse-Dynamics-Challenge, retrieved on October 12, 2022.