SPECTA: Speech-based Predictive Enrolment in Clinical Trials for Alzheimer's

Hali Lindsay,[1] Nicklas Linz,[2] Johannes Tröger[1*]

[1] German Research Center for Artificial Intelligence, DFKI GmbH, Campus D3 2, D-66123 Saarbrücken, Germany;
[2] ki elements UG (haftungsbeschränkt), Am Holzbrunnen 1a, D-66121 Saarbrücken, Germany
*Corresponding author: johannes.troeger@dfki.de

ABSTRACT
Alzheimer's Disease (AD) is a global issue with a severe socio-economic impact. In the search of a cure, pharmaceutical companies face the challenge of effectively on-boarding early-onset AD patients to prove the effectiveness of a new generation of secondary prevention drugs targeting the preclinical AD stage. SPECTA proposes a speech-driven neurocognitive pre-screening solution that uses automatic speech analysis and machine learning models to pre-select potential trial candidates based on the ordinary telephone. Results show that the fully automated solution performs on par with face-to-face assessment identifying early onset AD patients with an AUC of 0.92.

Keywords: Dementia; Remote Screening; Speech Analysis; Cognitive pre-screening.

1. INTRODUCTION

Alzheimer's Disease (AD) is a global issue with an economic impact that increases as life expectancy increases. While currently there is no cure, pharmaceutical companies are faced with many challenges in developing drugs that could aid in combatting this neurodegenerative disease. Within the last decade, 99.6% of all AD drug trials failed mainly because pharmaceutical industry runs trials with participants that are too late in AD's progression. Leading experts announced a new era of clinical trials: secondary prevention of AD, targeting patients at a stage of AD with no clinical symptoms but at-risk (AR-AD). In this new era, the major challenge is the clinical trial enrolment: companies now must onboard large cohorts with very tight specifications, but can't rely on medical records and clinics' databases for this pre-clinical phase of AD development. Population-wide screening incurs heavy costs, 99% of which is spent on screening the population that does not fit the criteria of the study to find a select few individuals.

The SPECTA project proposes a speech-driven neurocognitive pre-screening and monitoring solution that uses automatic speech analysis and machine learning models to pre-select and monitor potential AR-AD candidates at a preclinical stage for the enrolment into clinical trials for novel AD drugs. This cost-effective, wide-spread screening application hinges on speech-based neuropsychological AD diagnostics to enable a pre-screening and monitoring population-wide based on the ordinary telephone. For pharmaceutical industry SPECTA could be a real game-changer for the next decade of clinical trials, potentially resulting in more efficient drug development life cycles.

From the SPECTA project, we report on a solution and a feasibility study demonstrating very encouraging performance in an all speech-based telephone screening scenario for early dementia detection for enrolment in clinical trials, outperforming previous state of the art machine learning solutions and performing on par with face to face assessment models.

2. STATE OF THE ART

Current state-of-the-art for telephone-based dementia screening used simulated phone-quality audio samples of classic neuropsychological tests to differentiate between healthy controls and mild cognitive impairment (MCI). Results reported were an 0.85 area under the receiver operator curve (AUC) from the machine learning classification experiment [1]. While these results are promising, the classification accuracy could be improved and challenges of using automatic speech recognition (ASR) with an elderly population over the phone still pose a modern issue.

3. BREAKTHROUGH CHARACTER OF THE PROJECT

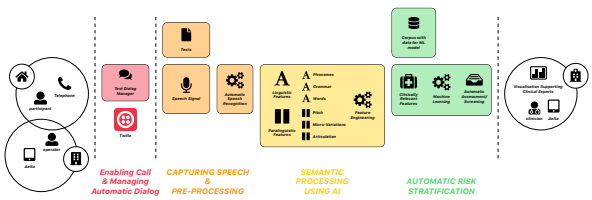
Previous solutions mainly lack scalability and cost-effectiveness. Here, the proposed SPECTA solution represent a breakthrough in two aspects: 1) there is no complex technical setup needed as SEPCTA relies on ordinary telephone only and therefore can reach out to a maximum screening population and 2) speech input is automatically assessed through an advanced combination of validated computational linguistic markers of cognition and machine learning models allowing a extremely low-cost screening workflow.

4. PROJECT RESULTS

The project results are twofold: a solution for automatic scalable early AD screening over the telephone and a feasibility study demonstrating the advanced performance of such a solution as compared to a classical face-to-face setup.

Solution Architecture

The automatic phone-based screening solution proposed by this project, includes four different components: interfaces (phone towards the patient and iPad App Delta towards the clinician), an automatic dialogue management component, a computational linguistics pre-processing engine (automatic speech recognition & feature extraction) and a ML-powered decision support component (see also the architectural overview below).

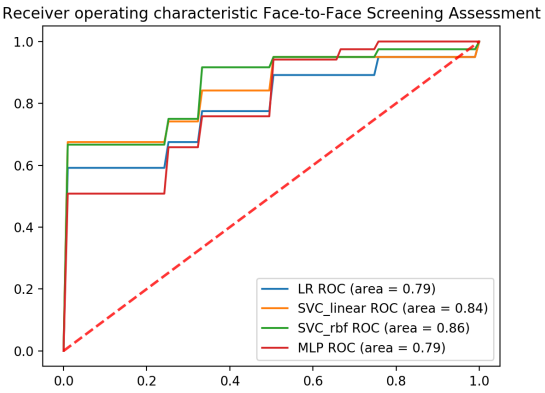


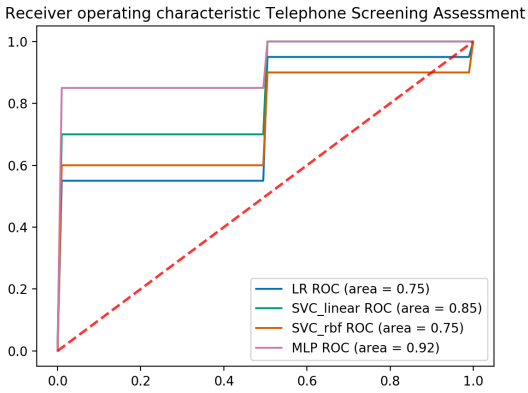
The instructions of the cognitive assessments are automated and given over a voice interface. Twilio, an external software, is used to design this interface and manage interactions of different scenarios. In this testing scenario, audio is recorded from the phone call.

Once the audio has been recorded the linguistic pre-processing engine is used to examine the input. The first step to analysing this data is automatic speech recognition. Once a transcript is obtained from the audio recordings, the scoring of the test is then automated and housed in the backend of the application. The screening results are then stored and displayed for the clinician through the front end of the application.

Comparing face-to-face and fully automated setup

In order to prove the automatic solution's feasibility, SPECTA compared two modes of screening: a cognitive screening in person with a clinician and a follow up screening six months later with a semi-automated telephone call relying on above-shown architecture. In





this case, a clinician initiates the system to administer cognitive tasks over the phone and the system reads out prompts and records the patients answers.

For both the face-to-face and the automated telephone assessment, data was collected from 70 individuals (MCI=34). Of these 70 persons, 31 completed the follow-up telephone screening (MCI=14). During both the face to face diagnostic and telephone screening scenario both the Semantic verbal fluency task (SVF) to measure executive function and the Word List Test (WLT) to measure memory function were administered—once together with a clinician face-to-face and once automatically over the phone. Different versions of the test are administered between time points to mitigate a learning effect.

After collecting the data, for the automated remote screening mode the tests were automatically scored by the application and for the face-to-face mode test data was manually corrected by a clinician.

To compare the face to face and telephone scenarios, both modes are individually explored by building binary classification models from the collected data. For machine learning experiments, a range of classification models were used linear regression, support vector machines and multi-layer perceptron. Models were built and tested separately for the different administrations and five-fold cross validation is used and the average score is reported over the folds. Machine learning results are visualised with the area under the receiver operator curve in the figures above, where 1.0 is a perfect classification.

The result from the classification experiments show the results on-par or outperforming current state-of-the-art methods. For the face to face screening, the best models is the support vector machine with radial-bass Kernel with an area under curve (AUC) of 0.86. The telephone-based screening out performs the face-to-face model with the best modelling being the light-weight neural model and achieves and AUC of 0.92. The core results of this project are a system that can be used to administer, score and collect testing either in a clinic or over the phone. Expanding the system to work with Dutch language resources including automatic speech recognition and clinical materials. The feasibility of building models that are scalable and cost efficient for the on-boarding of clinical trials in either a binary classification scenario and a forecasting of the clinical dementia rating scale. As well as building classification models that outperform current state-of-the-art for screening for dementia over the phone.

5. FUTURE PROJECT VISION

The SPECTA project showed the feasibility of applying speech-based screening for early stages of Alzheimer's in the context of recruitment for pharmaceutical studies. The constructed prototype solution together with first scientific results show great potential in scaling the project to a fully-fledged product for commercialisation or to apply similar technology in different disease areas such as depression or stress.

5.1. Technology Scaling

To scale the SPECTA prototype on a technical level, two major areas need to be separated: the technical scalability of the developed application and the regulatory requirements for medical devices.

The constructed prototype was developed following classical development processes and standards ensuring its technical scalability, with a minimal level of technical depth. The main component of the system, the predictive model, could be continually improved with the addition of more data.

The second barrier to real world applications is the regulatory requirements that might be posed to the system when used outside of research. The final product would maybe not be characterised as a medical device. Still, pharmaceutical companies have rigorous requirements concerning the certification and methodology used by their suppliers. This aspect will be the largest challenge in scaling the SPECTA system beyond a prototype.

5.2. Project Synergies and Outreach

As the largest technical risk in scaling the prototype to a product are the requirements posed by pharmaceutical companies, they are the most logical partner to involve in a Phase 2 of the ATTRACT program.

In a Phase 2 of the ATTRACT program together with a pharmaceutical partner, SPECTA could generate

outreach on major pharma events geared towards AD like the Alzheimer's Europe conference or the Clinical Trials on Alzheimer's international conference.

5.3. Technology application and demonstration cases

In a Phase 2 of the ATTRACT program, our goal would be twofold: to continue the development of the prototype and the research results obtained during Phase 1; specifically to collect larger data sets and more evidence that validates and improves the performance of predictive models for AD. Secondly, to use the collected results and expand the tools applicability to a second disease area, most likely depression.

This could be achieved by partnering with a pharmaceutical company as well as clinical organisations. Likely, the focus would be on partnering with study centres that are already involved in the collection of longitudinal data in the observed disease areas. This has the distinct advantage that speech recordings can be easily added to existing cohorts with minimal effort and cost.

5.4. Technology commercialisation

Results from the SPECTA project have already gaged interest from large pharmaceutical companies active in the disease area of Alzheimer's. Based on first conversations, a pilot project is being constructed to proof the feasibility of the SPECTA approach in a real-world environment and to get buy in from various stakeholders involved in the buying process.

Pilot projects are a common part of the pre-sales process with novel trial technologies. Adaptation and sales cycles can be particularly long in this area (>18 months). To kickstart commercialisation, a capital investment is very likely needed.

5.5. Envisioned risks

The project will face two main risk when adopted in ATTRACT Phase 2: gaining access to larger amounts of data to validate the novel developed biomarkers from Phase 1 and to meet the regulatory challenges posed when moving from a prototype to a product used outside of research. Both can be mitigated by the choice of partners for ATTRACT Phase 2. Having a strong clinical site/CRO involved that already has standing access to the desired patient population will guarantee access to the needed clinical data for validation. Cooperating with a pharmaceutical company in the field can ensure the necessary technical and regulatory requirements are met to the industry standard.

5.6. Liaison with Student Teams and Socio-Economic Study

For an ATTRACT Phase 2 project, SPECTA would open up the possibility through DFKI and Saarland University Chair of Language Science and Technology to jointly design a Master Seminar on pathological speech analysis for important psychiatric disorders such as AD or depression.

This Master's (MSc) seminar would be facilitated by the DFKI researcher and SPECTA final report author Hali Lindsay, who has experience in designing such a seminar. Recently, positive experience has been collected with a remote hands-on data science seminar where attendees are introduced to the data handling and preprocessing, feature extraction using computational linguistics and building machine learning models from the latent features analysis. This seminar included an interactive code package that enabled a broad scientific audience from medical to computer science students to participate remotely with minimal technical effort and prerequisites. The seminar is aimed at facilitating conversation and collaboration between fields to build computational models in a way that are viable for clinical efforts, furthering the goals of the SPECTA project.

6. ACKNOWLEDGEMENT

This project has received funding from the ATTRACT project funded by the EC under Grant Agreement 777222.

7. REFERENCES

[1] Johannes Tröger, Nicklas Linz, Alexandra König, Philippe Robert, and Jan Alexandersson. 2018. Telephone-based Dementia Screening I: Automated Semantic Verbal Fluency Assessment. In Proceedings of the 12th EAI International Conference on Pervasive Computing Technologies for Healthcare (PervasiveHealth '18). Association for Computing Machinery, New York, NY, USA, 59–66. DOI:https://doi.org/10.1145/3240925.3240943