|  |  |  |
| --- | --- | --- |
| ITU Logo | INTERNATIONAL TELECOMMUNICATION UNION**TELECOMMUNICATIONSTANDARDIZATION SECTOR**STUDY PERIOD 2017-2020 | FG-AI4H-B-013-R1 |
| **ITU-T Focus Group on AI for Health** |
| **Original: English** |
| **WG(s):** | Plenary | New York, 15-16 November 2018 |
| **DOCUMENT** |
| **Source:** | Laboratory for Research in Neuroimaging, Department of Clinical Neurosciences, Faculty of Biology and Medicine, UNIL Centre Hospitalier Universitaire Vaudois (CHUV) |
| **Title:** | Proposal: Using machine learning and AI for validation of Alzheimer’s disease biomarkers for use in the clinical practice |
| **Purpose:** | Discussion |
| **Contact:** | Kherif Ferah, vice-director LREN, CHUVSwitzerland | Tel: +41 79 556 11 06Email: Ferath.kherif@chuv.ch  |
| **Contact:** | Marc LecoultreMLLab.aiSwitzerland | Tel: +41 79 321 09 29Fax: +41 22 364 30 69Email: ml@mllab.ai  |

|  |  |
| --- | --- |
| **Abstract:** | This document contains a use-case proposal for using machine learning and AI for validation of Alzheimer’s disease biomarkers for use in the clinical practice. R1 of this document includes answers to the proposal submission questionnaire. |

**Organization Name:** Laboratory for Research in Neuroimaging, Department of Clinical Neurosciences, Faculty of Biology and Medicine, UNIL Centre Hospitalier Universitaire Vaudois (CHUV) (Switzerland)

**Contact Name:** Kherif Ferah, Marc Lecoultre

**Contact Email Address:** Ferath.kherif@chuv.ch, ml@mllab.ai

**Contact Phone Number:** +41 79 556 11 06

**Project Title:** Using machine learning and AI for validation of Alzheimer’s disease biomarkers for use in the clinical practice

**Overview**

*Please give a general overview of the project, and describe what health problem it is attempting to contribute to solve.*

The present use cases address Alzheimer’s disease (AD) and related syndromes (defined by DSM V as ‘Neurocognitive disorders’), as these brain pathologies represent an important global health challenge, acknowledged as such by the WHO in May 2017. AD affects about 47 million people worldwide and the affected population will probably double at least over the next 30 years. Currently we can estimate the direct costs for western countries as about one billion Euros per year/ million inhabitants. Most likely, milder and prodromal syndromes (“Mild Cognitive Impairment”) are 2-3 times more prevalent than the full-blown dementia cases.

While the scientific advances in Alzheimer’s disease are expanding continuously, an ever-growing gap is developing between primary evidence i.e. clinical data and the biology- or imaging-based research findings.

Several biomarkers using machine learning tools exist for Alzheimer’s disease however, few studies on the clinical validity and generalizability of these biomarkers using “real-world patient data” exist and therefore there is little information on their clinical utility. From a clinical perspective, the most important application is to use routinely collected data at the hospitals for: Improving the classification of different dementia subtypes using differential patterns of cortical atrophy associated with cognitive decline using predictive machine learning algorithms, including benchmarking and cross-validation of learned models.

**Impact**

*Please explain the significance of the problem and describe the potential impact of the project. Please also provide a brief overview of existing work in the area of the project, and describe the current state of the art how the problem is currently addressed.*

With the introduction of electronic health records (EHR) and picture archiving and communication systems (PACS), clinical researchers can acquire information from groups of patients in their hospital after receiving informed consent from patients. However, due to data protection regulations concerning patient-information privacy and security, individual hospitals set up their own EHR and PACS systems to collect patient data. Patient medical data is therefore scattered across large numbers of hospitals, clinics and private practices, as with previous paper-based medical records. The integration of dispersed EHR and PACS systems currently remains a big test, not only from the patient data-protection point-of-view, but also from incompatible IT solutions (different codes and structures). Consequently, clinical researchers only access data stored in systems in their own hospitals (1).

We need to efficiently acquire all the data into central storage and organise and analyse it to extract its research potential. Success will come from collaborative efforts to generate machine learning tools necessary to analyse these high quantities of diverse data as well as strategies to share and investigate rich data sources across institutions and validate the results in the hope of improving clinical investigation.

The strategy is to use computational and machine learning approaches (from data pre-processing, brain feature extraction to data mining) and create a meeting place for neuroscience and IT for collaborative brain disease research as well as benefitting clinicians on a daily basis.

**Data Availability**

Data will include both real world patient’s data and data collected from research cohorts. The data will include clinical scores, diagnostic, cognitive measures and biological measures (PET, MRI, fMRI, lab results). A representative sample will be created and will be use for the creation of the models. The models will be then validated (see benchmarking methods below) on the real world patients data.

We have already collected patients on more than 6 000 patients on dementia (one of the largest patients’ cohort) different stages of the disease (subjective complains, mild impairments or demented).

**Benchmarking**

*Please describe what you expect participants in the benchmarking process to submit. Please also describe how the submissions should be evaluated, and why.*

The tools developed need to be useful for

* Clinicians, for objective diagnoses and treatment of brain disease
* Scientists, for the application and testing of new models and methods
* Pharmaceutical or biotech researchers, for disease target discovery

The benchmarking process will be based using the state of the art methods for the methods used by the ML community, but also methodology recommended for clinical trial. Thus, the algorithm will not be only evaluated for their performance, i.e. accuracy but also in term of:

1. Analytical validity,
2. Clinical validity, Assessment of clinical validity involves measurement of biomarker’s or other health-relevant feature’s clinical performance, including: (1) clinical sensitivity (ability to identify those who have or will get the disease), (2) clinical specificity (ability to identify those who do not have or will not get the disease), (3) positive predictive value (PPV)
3. clinical utility: is the most important (does the results impact the clinical path of the patients, does it affect the relatives, …)

**Organizer Details**

*Please describe why your organization is interested in this project, and if you have run similar projects / benchmarks / challenges before.*

My group has extensive expertise in collecting clinical data, curating and pre-processing these data to the highest standard. We have also developed own methods for machine learning. But we believe in fruitful collaboration between "business experts", and machine learning expert will bring high value from the data and personalised tools for disease prognoses. We have established several small-scale collaborations with these goals in mind. We have demonstrated for example by benchmarking different state of the art ML tools that in fact, some methods do not scale to the size of our data and between those methods that support large-scale data similar results were obtained. We also demonstrated that deep learning method are highly performant with our data but need to be tuned to the specificity of medical data (noisy data, missing data, unclear label, heterogenous data, etc.).

Our organisation is highly interested in the next step, which is AI for the diagnostic of brain diseases. AI can and will solve many health-related issues, however, currently; AI for healthcare suffer many problems that medical doctors can help explain.

Annex A
Answers to the proposal submission questionnaire

1. **Relevance** - How relevant is the health problem to be addressed?

The present use cases address Alzheimer’s disease (AD) and related syndromes (defined by DSM V as ‘Neurocognitive disorders’), as these brain pathologies represent an important global health challenge, acknowledged as such by the WHO in May 2017. AD affects about 47 million people worldwide and the affected population will probably double at least over the next 30 years. Currently we can estimate the direct costs for western countries as about one billion Euros per year/ million inhabitants. Most likely, milder and prodromal syndromes (“Mild Cognitive Impairment”) are 2-3 times more prevalent than the full-blown dementia cases.

While the scientific advances in Alzheimer’s disease are expanding continuously, an ever-growing gap is developing between primary evidence i.e. clinical data and the biology- or imaging-based research findings.

The number of brain health related data (imaging, labs, clinical) collected daily in clinical and research establishments is enormous.

There is a pressing demographic and economic need to answer questions about the compensated period prior to dementia, the incidence and natural history of pathological change in the compensated state, how to make more precise diagnoses based on brain change rather than behavioural expression, how to monitor rate of pathological brain changes on sufficient numbers of people such that the results are generalisable.

1. **Impact** - What level of impact will a benchmark in the context of the proposed project have?

The repercussions of the results will be important because there is preliminary evidence to suggest that the dementias can be differentiated by early distribution of brain atrophy (Seeley WW et al., 2009) it should therefore be possible to identify purer cohorts of the different dementia-associated diseases than is now possible to identify, test and develop new specific disease-modifying drugs.

1. **Existing work** - Does the project start from scratch, or are there preliminary experiences?

“Industrial scale” image analysis is a promising answer. It is documented that support vector machines (SVM) detect AD accurately in single T1 weighted images, as validated by pathology (see Kloppel et al., 2008a). Thus, proof of principle has been demonstrated. These classification techniques are computer-dependant and automated, so in addition to removing error variance due to misdiagnosis, experimenter bias is minimised. However, what is needed additionally is an understanding of disease mechanisms, ultimately to enable accurate pre-clinical classification.

1. **Feasibility** - Is the project feasible, based on the current state of the art?

Pattern recognition, classification and machine learning methods have been introduced recently for individual diagnosis and prognosis (Kloppel et al., 2008a). Support vector machines have proved very useful for the diagnosis of Alzheimer’s disease where performance exceeds that of full, combined clinical methods including neuroradiological interpretation (Kloppel et al., 2008b). These techniques are especially powerful with large sample sizes and a variety of inter-related data.

1. **Data Availability** - Is there sufficient data available? How much of it can be openly available? How much of it as part of the non-disclosed data set?

The primary data are already available and growing in volume. Data will include both real world patient’s data and data collected from research cohorts. The data will include clinical scores, diagnostic, cognitive measures and biological measures (PET, MRI, fMRI, lab results).

We have already collected patients on more than 6 000 patients on dementia (one of the largest patients’ cohort) different stages of the disease (subjective complains, mild impairments or demented). A large representative sample will be created and will be use for the creation of the models. The models will be then validated (see benchmarking methods below) on the real world patients data.

1. **Data Quality** - Is the available data of high quality?

Data quality of these real world data is comparable to the data quality from the high standard research cohorts.

1. **Annotation / Label Quality** - Are the annotations / labels of the data of high quality?

the data have been annotated using best clinical practice and described using Common data Element and international classifications.

1. **Data Provenance** - Has the data been obtained in a professional and ethically correct way?

The data have been obtained under each country ethical and regulation policies, including patient informed consents.

1. **Benchmarking** - Do the applicants have a clear proposal about what exactly should be evaluated / measured?

The tools developed need to be useful for

* Clinicians, for objective diagnoses and treatment of brain disease
* Scientists, for the application and testing of new models and methods
* Pharmaceutical or biotech researchers, for disease target discovery

The benchmarking process will be based using the state of the art methods for the methods used by the ML community, but also methodology recommended for clinical trial. Thus, the algorithm will not be only evaluated for their performance, i.e. accuracy but also in term of:

**Analytical validity**,

**Clinical validity**, Assessment of clinical validity involves measurement of biomarker’s or other health-relevant feature’s clinical performance, including: (1) clinical sensitivity (ability to identify those who have or will get the disease), (2) clinical specificity (ability to identify those who do not have or will not get the disease), (3) positive predictive value (PPV)

**Clinical utility**: is the most important (does the results impact the clinical path of the patients, does it affect the relatives, …)

1. **Organizers** - Can the Focus Group work with the applicants, and do they have the time / resources to work with the Focus Group on the problem?

**YES**- My group has extensive expertise in collecting clinical data, curating and pre-processing these data to the highest standard. We have also developed own methods for machine learning. But we believe in fruitful collaboration between "business experts", and machine learning expert will bring high value from the data and personalised tools for disease prognoses.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_