Project Proposal   Group 1

MSAI 349-0   Quincia H, Jaieu S

Prof. Doug Downey   Jack R, Ikhlas A

11/25/2018

Find out the most helpful diagnosis for predicating Alzheimer's Disease

Problem Statement

Alzheimer's disease (AD) is a type of dementia that causes problems with memory, thinking and behavior. AD is an irreversible process that typically begins after age 60. Once a patient has been diagnosed, their mental function usually declines until death. While some drug and non-drug treatments may help with both the cognitive and behavioral symptoms, currently Alzheimer's disease has no cure. In the United States, more than 5 million people aged 65 and over suffer from the disease. AD is currently ranked as the sixth leading cause of death in the United States and is the third leading cause of death for older people, just behind heart disease and cancer (NIA, 2017). The estimated national cost of patient care for Alzheimer's and other types of dementia was $236 billion in 2016. Therefore, there is both a human and economic incentive to find an effective therapy for AD. While scientists are still uncertain about the precise cause of AD, research has identified strong indicators for the disease. Due to the irreversible nature of AD, it is believed that the effectiveness of any potential therapy greatly depends on early detection and treatment.

While there have been significant advances in diagnostic testing methods for Alzheimer's that use brain scans and some technique like spinal taps, can detect certain biomarkers of the disease even in its pre-clinical stage, currently, there is no single test that can diagnose Alzheimer's disease with 100% accuracy. Doctors must use a variety of assessments and laboratory measurements to make a "differential diagnosis" (also called "Alzheimer's Diagnostic Tests"). They focus on ruling out all other possible causes for the symptoms.

In order to make the early diagnose as accurate as possible, biomedicalists have developed a complex testing protocol, which includes tens of examining categories. Different criteria are set to classify the testing results and define the phase of AD. Nowadays, the study of AD becomes a worldwide project and some shared database is built for cross-county research. One of the biggest AD database - **Alzheimer's Disease Neuroimaging Initiative database (ADNI)** has collected thousands of subjects with determined AD and their symptom change over the time.

It will be greatly beneficial to make use of the data for the purpose of developing efficient detection and treatment protocol of AD.

Data recourse

ADNI have included 400 subjects diagnosed with mild cognitive impairment (MCI, which is considered as AD), 200 subjects with early AD, and 200 control subjects. All the subjects have been under trace for 78 months at most to observe the AD phase change over time. Available features of ADNI database are patients' demographics, Medical history (disease and medication), Lab records, Cognitive test score and Imaging data. ADNI (http://adni.loni.usc.edu) (ADNI 2017) is a free and open-source database.

Project Proposal

MSAI 349-0

Prof. Doug Downey

11/25/2018

Group 1

Quincia H, Jaieu S

Jack R, Ikhlas A

Everyone can access the data, once the application is approved by the administrator. We submitted the application of ADNI access and it was approved. The data is ready to use.

Project Goal

The testing result in AD diagnosis is represented as scores. ADNI includes 8 types of diagnosis (shown in Fig. 1). Each type of the diagnosis contains multiple tests and each test is evaluated by scoring the result based on some pre-set criteria. For instance, the clinical diagnosis involves 6 subtypes, and the Neuropathology involves more than 10.

After studying the diagnosis model, we realized that most of the tests took long time and were challenging to the patients. From the testing records, we found that some of the examinees got anxiety and frustration when they were unable to perform the task and then the test had to be abandoned. For example, in the Clock drawing test the examinee is required to draw a clock to verbal instructions, which is every hard even if the examinee only has mild AD. We got a rough estimation from a Month 6 dataset (the data collected when the examinees came back at Month 6) that 81 out of 3874 examinees (2.1%) quit the program because they are unhappy to continue. The ratio became much higher after 4 years and the ADNI program was not able to extend over 5 years.
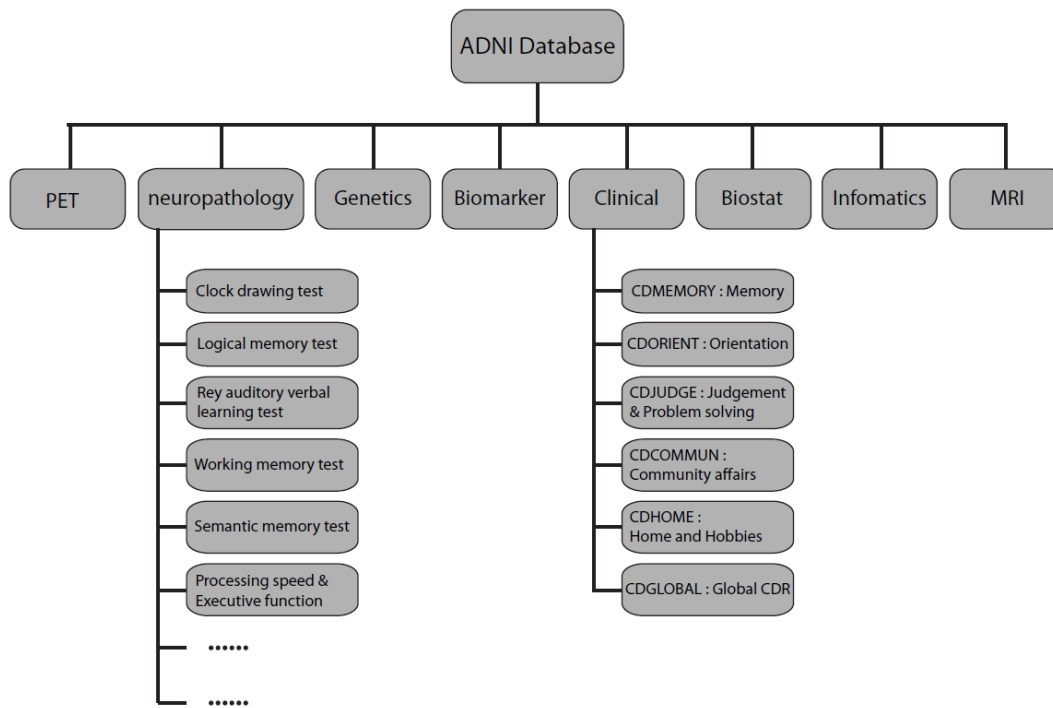

Figure 1 Schematic of diagnosis category and it subcategory in ADNI database

Project Proposal   Group 1

MSAI 349-0   Quincia H, Jaieu S

Prof. Doug Downey   Jack R, Ikhlas A

11/25/2018

The goal of our project is to use ADNI database to

1. Analyze the correlation between the subcategories of each diagnosis and the phase of AD. Here we plan to use clustering and decision tree mode, take the score of each test as the features and the phase of AD (normal, mild, moderate, severe…) as classifier, reveal the most helpful examinations for the examiner to determine the AD phase of a patient and simplify the currently used diagnosis model.
2. Analyze the scores change of the tests over time using. Suggest an efficient AD diagnosis model for long-time using purpose.

Solutions and Plans

1. Stage 1: Clean data, select the features and classifier.

2. Stage 2: Pick up several features from one diagnosis category in same period, visualize data, and analyze the correlation.

3. Stage 3: Integrate the full categories in a selected period and find out the most help diagnosis. Achieve **Goal 1**.

4. Stage 4: Get the most help diagnosis in different period. Compare the same diagnosis in different period, and get a long-time AD diagnosis model. Achieve **Goal 2**.