Optimization of Day Out Task Prediction

Pedram, M.[2], Brown, K.D.[1], Cook, D.[2], & Schmitter-Edgecombe, M.[1]

[1]Department of Psychology and [2]School of Electrical Engineering and Computer Science

Introduction

- Using a data driven approach, we attempted to determine what tests and cognitive domains (memory, executive function, and processing speed) contribute more to various performance measures on a natural assessment of functional ability in a mixed sample of older adults.

Methods

Sample

- 452 Older adults (M_{age} = 67.3, Range: 47 - 97)
  - 274 healthy older adults
  - 75 with Mild Cognitive Impairment (MCI)
  - 13 with Dementia
  - 19 with Parkinson's Disease (PD)
  - 14 with PDMCI
  - 57 with other medical conditions

Neuropsychological Domains and Tests

- Executive Functioning
  - Trails B
  - Verbal fluency (FAS)
  - Category fluency
  - Category switching fluency
  - Design fluency (DF)
    - Open dots & switching conditions
  - Prospective memory measure (PM)
  - Temporal order memory measure (TO)
- Memory
  - MAS list acquisition
  - MAS short delay (SD) list recall
  - MAS long delay (LD) list recall
  - MAS delayed prose recall
- Speed of Processing
  - Trails A
  - SDMT written format
  - SDMT oral format
  - Design fluency
    - Solid dots (DF)
- Functional Ability (Day Out Task)
  - Accuracy
  - Total time
  - Sequencing
  - Efficiencies
  - Inefficiencies
  - Inaccuracies

Methods

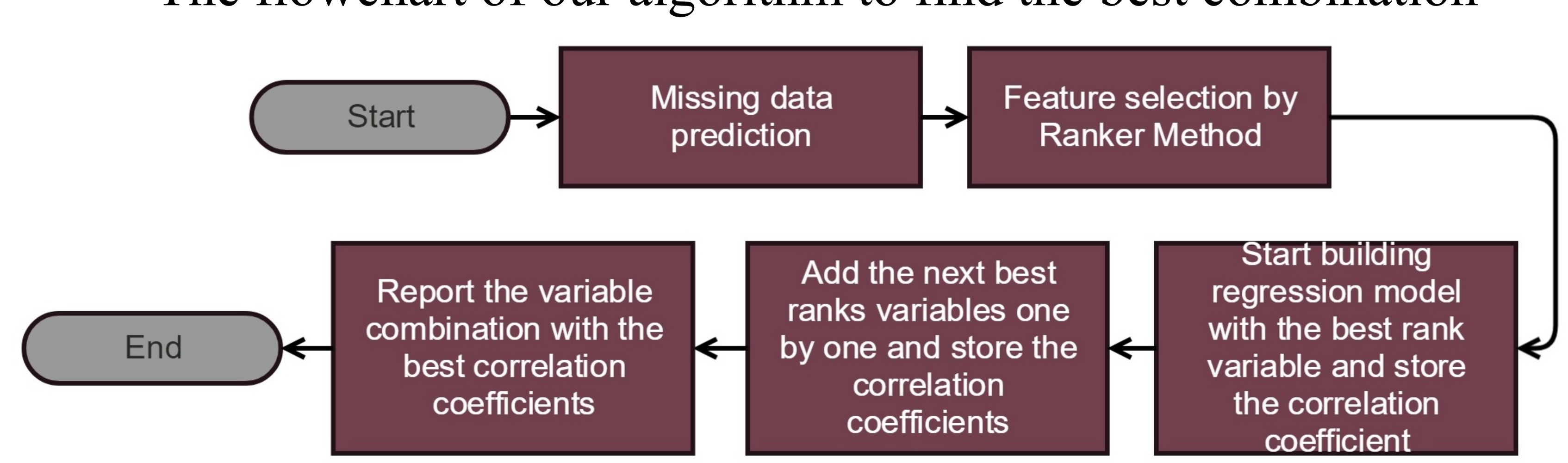
- We used the following 3 different methods for predicting missing data: predicting each input variable with other input variables, data imputation, and joint prediction.
- Experiments have been done with 10 folds cross validation.
- Machine learning methods and correlation coefficients (*r*) of the regression model for predicting each input variable can be found in the table below.

Test	Regression Model	Correlation Coefficient	Test	Regression Model	Correlation Coefficient
Trails B	Random Forest	0.7464	MAS LD list recall	Random Forest	0.8964
FAS	Linear Regression	0.5668	MAS SD list recall	Linear Regression	0.8764
Category fluency	Linear Regression	0.7390	MAS list acquisition	SVM	0.8269
Category fluency switching	SVM	0.6492	MAS delayed prose recall	MP5	0.5536
DF open dots	SVM	0.7737	TO measure	Random Forest	0.5923
DF switching	Linear Regression	0.6111	Trails A	Random Forest	0.6546
PM measure	SVM	0.5112	SDMT written	Linear Regression	0.8938
SDMT oral	Linear Regression	0.8958	DF solid dots	Mp5	0.7305

- The table below contains strengths (*r*) of our prediction for each DOT outcome variable.

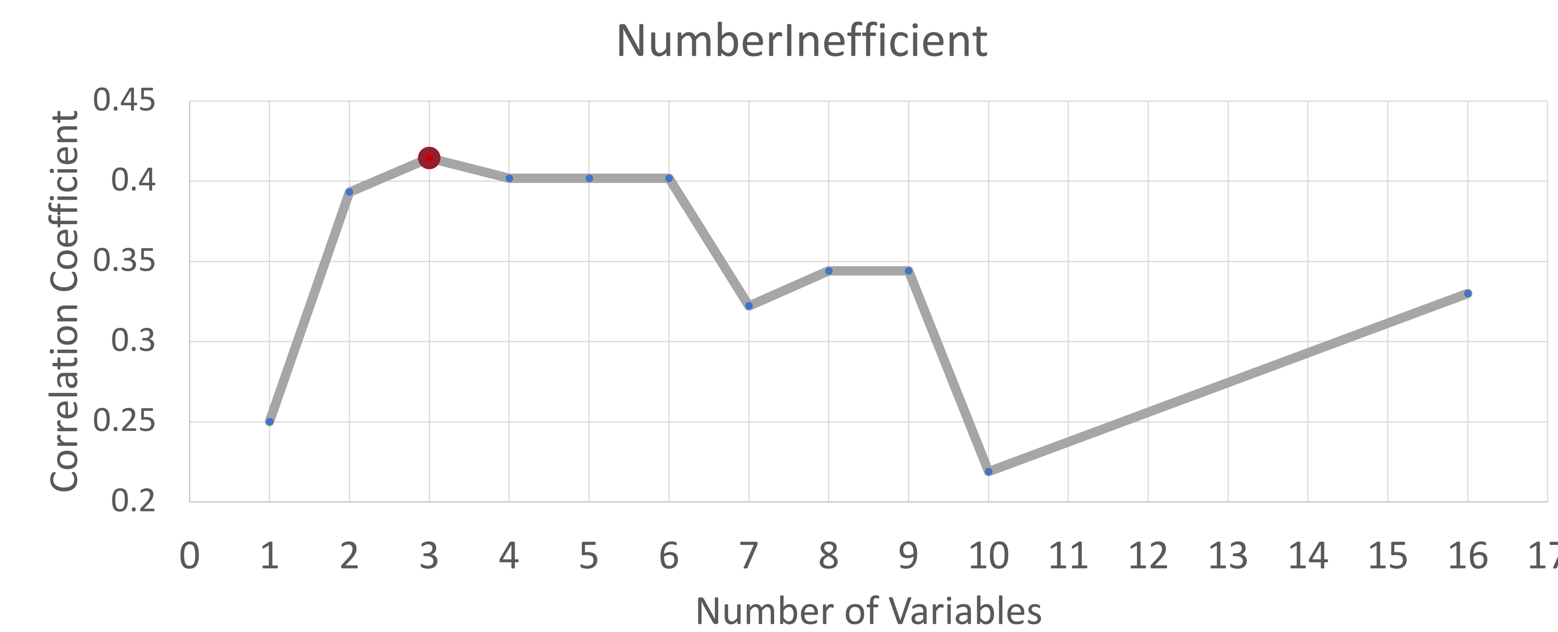
DOT OUTCOME VARIABLES	RAW DATA	PREDICTED DATA	DATA IMPUTATION	JOINT PREDICTION
EFFICIENCIES	0.3928	0.4082	0.3918	0.416
INEFFICIENCIES	0.3346	0.236	0.3401	0.3357
INACCURACIES	0.3758	0.4197	0.3903	0.4059
SEQUENCING	0.333	0.362	0.3878	0.3914
ACCURACY	0.4741	0.4979	0.4759	0.4858
TOTAL TIME	0.4512	0.4294	0.4594	0.4676
AVERAGE	**0.3935**	**0.3922**	**0.4075**	**0.417**

- The flowchart of our algorithm to find the best combination

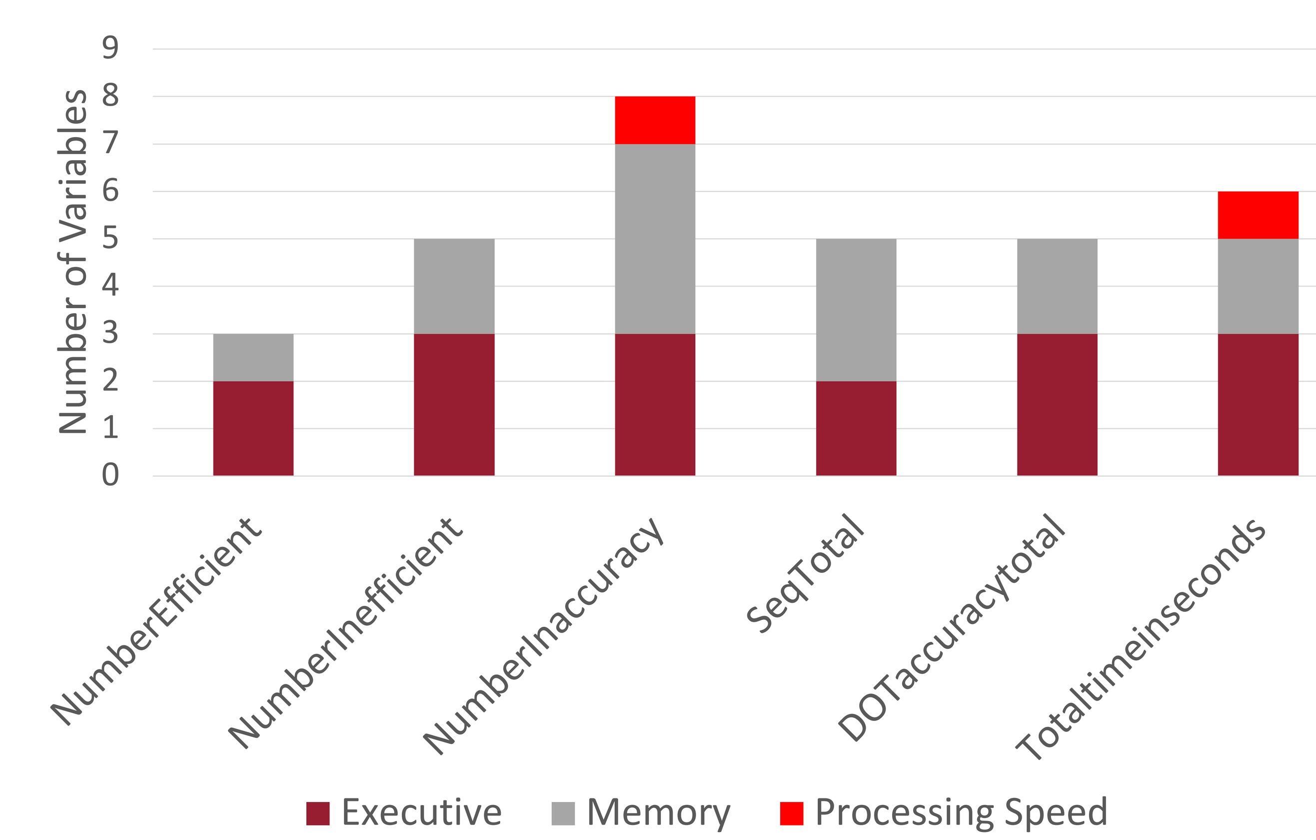


Results

- The graph below demonstrates how our algorithm finds the best correlation coefficient for an output variable.



- In the chart below, the optimum number of tests that we need to predict each outcome variable is displayed.
- The domain distribution for required variables to predict each outcome is also shown.



Conclusion and Future Works

- We determined the optimum number of tests for predicting specific performance measures on the DOT.
- Tests of executive functioning & memory are more important in predicting functional ability on the DOT than tests of processing speed.
- Using these results, we plan to build a model based off of performances on these tests to predict DOT performance.