**Mild Cognitive Impairment Prediction and Diagnosis Procedure Recommendation using Machine Learning**

Sajjad Fouladvand, PhD1, Morteza Noshad, PhD1, Jonathan H Chen, MD, PhD1*,*2

1Stanford Center for Biomedical Informatics, Stanford University, Stanford, CA, USA;

2Division of Hospital Medicine, Stanford University, Stanford, CA, USA;

# **Abstract**

# **Introduction**

[WHAT IS MCI AND WHY IT IS IMPORTANT TO PREDICT IT]

[WHY IT IS IMPORTANT TO PREDICT MCI WITH AI INSTEAD OF TRADITIONAL METHODS AND SCREENING]

[LITRATURE REVIEW ON MCI PREDICTION USING AI]

[WHAT DID WE DO IN THIS WORK]

# **Materials and Methods**

## Data Set

## [GENERAL INFORMATION ABOUT SHC\_CORE DATA]

[DETAILED EXPLANATION ON COHORT SELECTION AND INCLUSION EXCLUSION CRITERIA]

## Data Pre-processing

[DESCRIBTION OF DATA PREPROCESSINGS: CONVERTING TO STATIONARY, FEATURE GROUPING AND …]

**Table 1**. Patient Characteristics among Cases and Controls.

|  |  |  |
| --- | --- | --- |
| **Variable** | **Case** | **Control** |
| **Demographic** |  |  |
| Age | 70 (62, 82) | 72 (66, 83) |
| Female | 2,524 (54.43%) | 2,524 (54.43%) |
| Race |  |  |
| Asian | 544 (11.73%) | 719 (15.50%) |
| Black | 262 (5.65%) | 167 (3.60%) |
| Native American | 17 (0.37%) | 23 (0.28%) |
| Pacific Islander | 35 (0.75%) | 36 (0.78%) |
| White | 3,091 (66.66%) | 2,672 (57.62%) |
| Unknown | 140 (3.02%) | 435 (9.35%) |
| Other | 548 (11.82%) | 595 (12.83%) |
| **Diagnoses** |  |  |
| Nervous system signs and symptoms (CCS=SYM010) | 1,476 (31.83%) | 390 (8.41%) |
| Medical examination/evaluation  (CCS=FAC014) | 2,242 (48.35%) | 994 (21.44%) |
| Exposure, encounters, screening or contact with infectious disease  (CCS=FAC016) | 1,666 (35.93%) | 625 (13.48%) |
| Musculoskeletal pain, not low back pain  (CCS=MUS010) | 1,955 (42.16%) | 837 (18.05%) |
| Sleep wake disorders  (CCS=NVS016) | 1,350 (29.11%) | 494 (10.65%) |
| **Medications** |  |  |
| ANTIHYPERLIPIDEMIC-HMGCOA REDUCTASE INHIB(STATINS)  **(M4D)** | 1,943 (41.90%) | 1,065 (22.97%) |
| ANTICONVULSANTS  **(H4B)** | 1,212 (26.14%) | 408 (8.80%) |
| OPIOID ANALGESIC AND NON-SALICYLATE ANALGESICS  **(H3U)** | 1,836 (39.59%) | 841 (18.14%) |
| ANTIEMETIC/ANTIVERTIGO AGENTS  **(H6J)** | 1,896 (40.89%) | 863 (18.61%) |
| **Procedure** |  |  |
| METABOLIC PANEL, COMPREHENSIVE (1369) | 2,852 (61.50%) | 1,219 (26.29%) |
| TSH (2220) | 2204 (47.53%) | 757 (16.32%) |
| UnMapped External Results (464825) | 47 (1.01%) | 615 (13.26%) |
| CBC WITH DIFFERENTIAL (475) | 2,869 (61.87%) | 1357 (29.26%)) |
| OTHER ORDER SCANNED REPORT (147416) | 2,126 (45.85%) | 927 (19.99%) |

# **Experimental Results**

**Table 2.** Performance of MCI prediction using machine learning on unseen test sets.

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Model | Accuracy | | Precision | | Recall | | F1-score | | AUC | |
| bln | imb | bln | imb | bln | imb | bln | imb | bln | imb |
| Logistic Regression | 0.72 | 0.63 | 0.77 | 0.03 | 0.64 | 0.64 | 0.70 | 0.05 | 0.79 | 0.70 |
| XGBoost | 0.77 | 0.53 | 0.79 | 0.02 | 0.74 | 0.74 | 0.77 | 0.05 | 0.85 | 0.71 |
| Random Forest | 0.77 | 0.48 | 0.78 | 0.02 | 0.77 | 0.77 | 0.78 | 0.04 | 0.86 | 0.70 |

**Table 2.** Performance of MCI prediction using machine learning on unseen test sets.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Model | Accuracy | Precision | Recall | F1-score | AUC |
| Logistic Regression |  |  |  |  |  |
| XGBoost |  |  |  |  |  |
| Random Forest |  |  |  |  |  |

**DISCUSSION**

# **Conclusion**

# [SUMMARY OF THIS WORK. LIMITATIONS. FUTURE WORK]

# **References**