

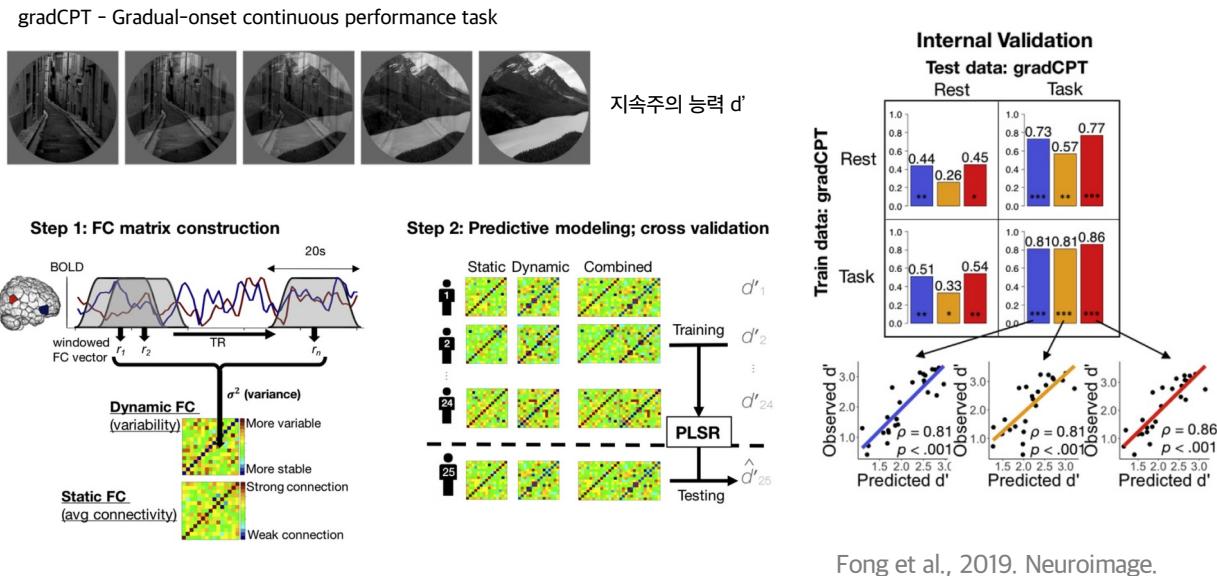
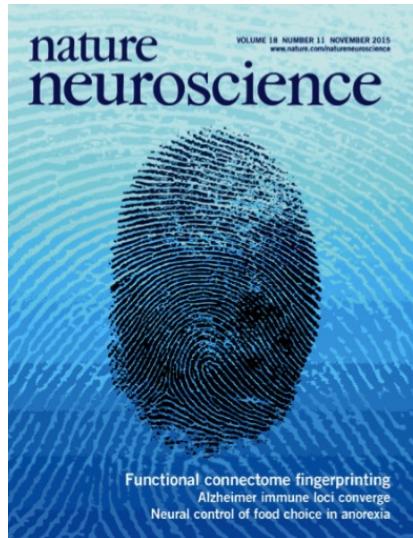
40

요약

fMRI 전뇌 기능적 연결성 분석을 활용한 인지 능력 예측 모델링의 예비 연구

- 기간 : 2020.10 - 2021.01, 약 4개월
- 요약
 - 휴식 또는 간단한 과제 수행 중 획득된 fMRI 기능적 연결성 데이터로 인지 기능을 예측하는 예측 모델링의 예비적 분석 수행
 - Human Connectome Project 공개 데이터와 KBRI 데이터를 활용
 - 연결성으로 인지 측정치를 예측하는 예측 모델링 절차 구현 및 예비 검증
- 주요 역할
 - 데이터 처리 및 통계 분석, 예측 모델링
- 도구
 - FSL (뇌영상 데이터 처리), Bash Shell Script (처리 절차 자동화)
 - Matlab (데이터 전처리, 연결성 분석), Python (통계 분석 및 예측 모델링, 시각화)
- 성과
 - 예측 모델링 절차 구체화
 - 예비 분석을 통해 실험 설계 및 구현에 기여
 - 다차원 데이터 분석 및 모델링 노하우 습득

뇌의 연결성 패턴으로 인지 능력의 개인차 예측하기



FC Fingerprints : 개인마다 지문과 같이 독특한 연결성 패턴이 있으며, 개인차를 예측하는데 사용 가능

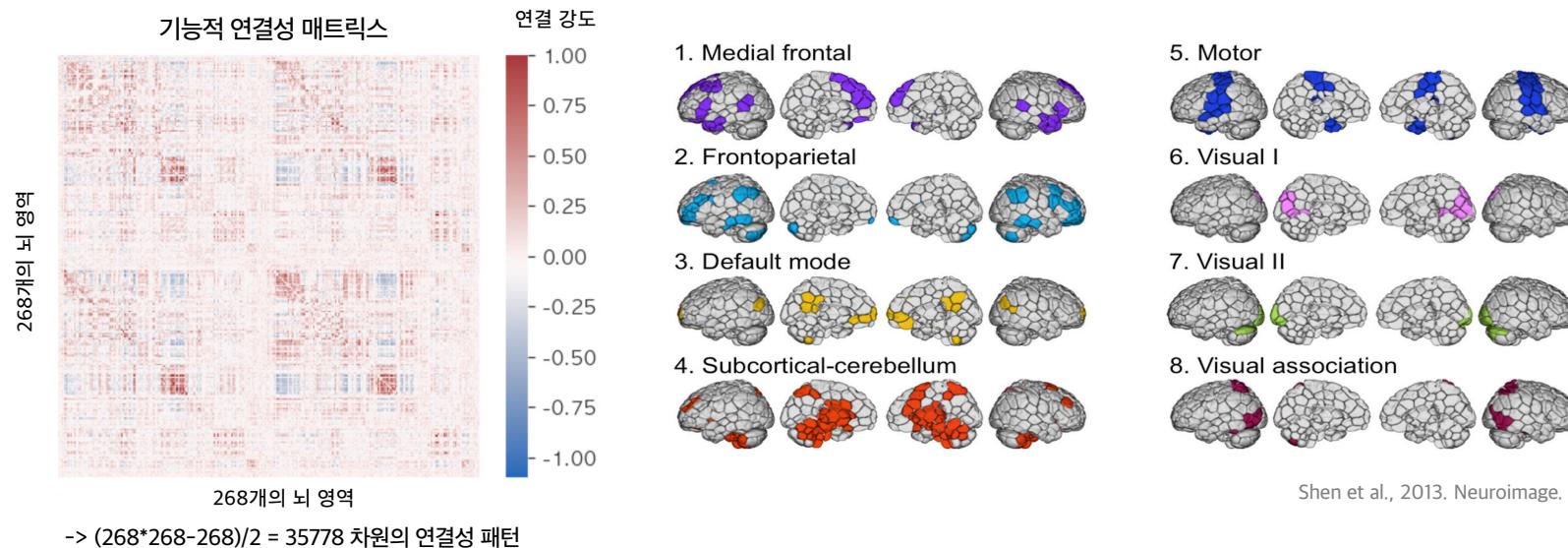
개인의 연결성 패턴으로 지속 주의 능력을 예측한 선행 연구를 참조, HCP 공개 데이터로 예측 모델링 절차 구현

42

방법

데이터 분석 과정

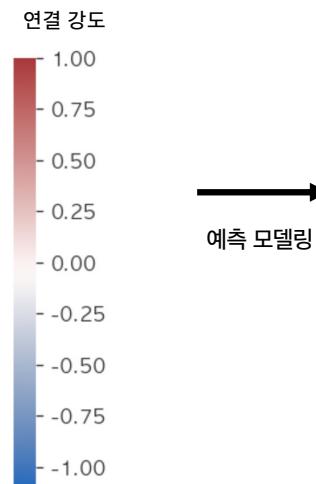
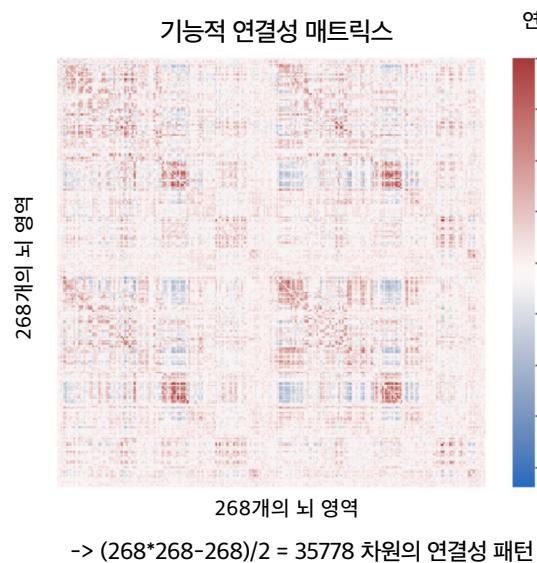
PROJECT 4



Shen 268 parcels을 활용하여 추출된 n = 337의 HCP 공개 데이터 활용

42 방법

데이터 분석 과정



- 개인차의 행동 지표 - 7개의 과제 수행치
 - Fluid intelligence (Raven's Matrix)
 - Working memory (list sorting task)
 - Picture vocabulary (verbal IQ)
 - Reading (verbal IQ)
 - Processing speed
 - Episodic memory (picture sequence)
 - Measure of delay discounting

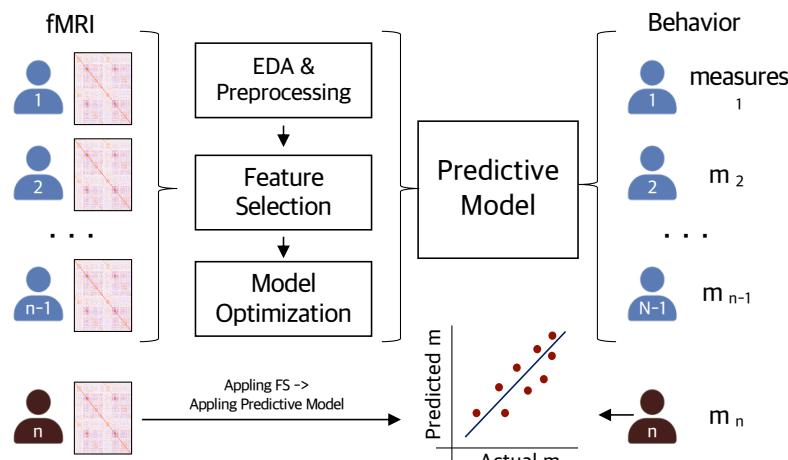
Rest, Working memory, Language, Gambling, Social의 5개 state에서 획득된 연결성 패턴 중,
Working memory (n-back task 수행) state의 패턴 선택. 패턴으로 7개 과제 수행치를 예측하는 모델링 절차 구현

42

방법

데이터 분석 과정

- Connectome-based Predictive Modeling 절차 개요



1. Exploratory Data Analysis & Preprocessing

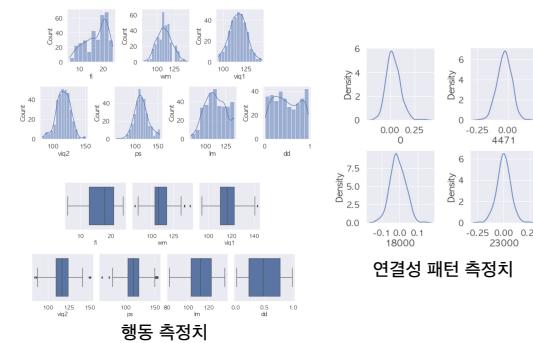
- 1) 데이터 개수, 차원, 길이, 유형 (int, float, ...) 등 확인
- 2) 통계치 확인 : 중심경향치, 산포도 등
- 3) 결측치 여부 확인 및 처리

- 제거 또는 대체 (중심 경향치, Linear Impute, KNN impute)



4) 분포 확인 & 이상치 여부 확인 및 처리

- 제거 (+-SD, Quantile), 대체 (비율 상의 상,하한 최고치)



연결성 패턴 측정치

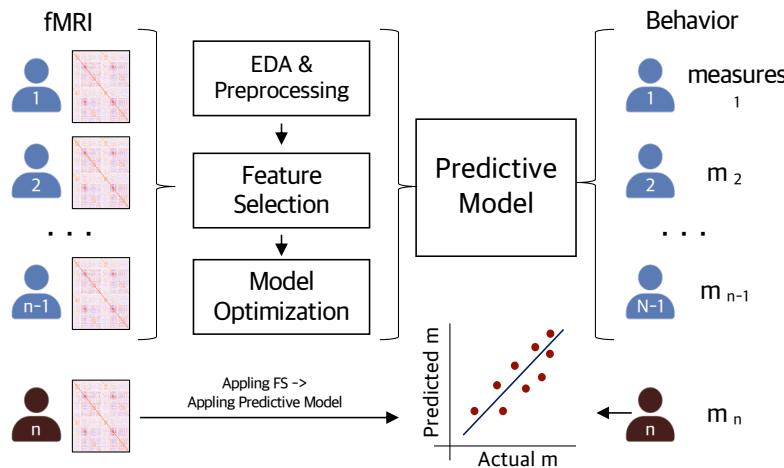
행동 측정치

PROJECT 4

42 방법

데이터 분석 과정

- Connectome-based Predictive Modeling 절차 개요



2. Feature Engineering

1) Feature Scaling

- Standardization, Normalization 등, 주로 Standardization

```
# feature scaling
scaler = StandardScaler()
targ_col = X_train_r.columns.values
X_train_r = scaler.fit_transform(X_train_r)
X_train_r = pd.DataFrame(X_train_r)
X_train_r.columns = targ_col
X_test_r = scaler.transform(X_test_r)
X_test_r = pd.DataFrame(X_test_r)
X_test_r.columns = targ_col
```

2) Feature Selection

- Filter - Correlation, Variance, Regression & Anova
- Wrapper - Stepwise, Model-based
- Embedded - Lasso, ElasticNet
- fMRI 연구에서는 주로 Regression & ANOVA 사용

```
elif fs_name == 'f-Reg':
    Threshold = 0.05
    selector_freq = SelectKBest(f_regression, k='all')
    selector_freq.fit(X_train_r, y_train)
    is_support_freq = selector_freq.pvalues_ < Threshold

    X_train_r_fs = np.array(X_train_r[X_train_r.columns[is_support_freq].values])
    X_test_r_fs = np.array(X_test_r[X_test_r.columns[is_support_freq].values])
    y_train_fs = np.array(y_train)
    y_test_fs = np.array(y_test)
```

3) Feature Extraction

- PCA, PLSR 등
- 337개의 Case와 35778의 서로 상관이 있는 feature dimension을 고려.
X Feature의 차원 축소에 Y과의 관계를 고려하는 Partial Least Squares Regression 사용

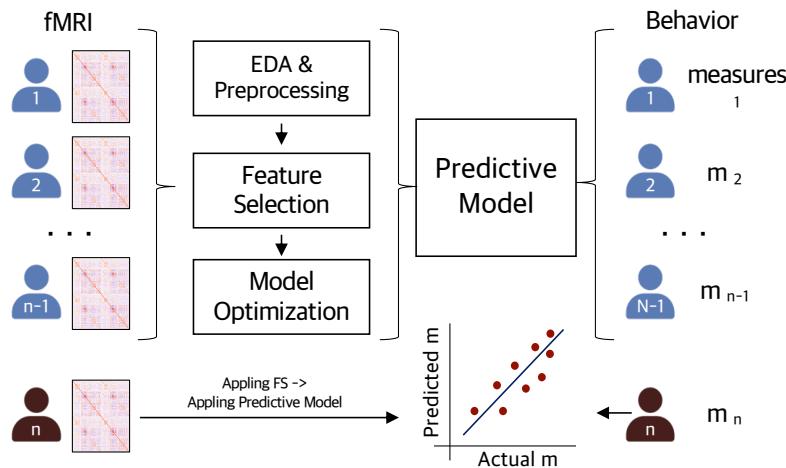
42

방법

데이터 분석 과정

PROJECT 4

- Connectome-based Predictive Modeling 절차 개요



3. Model Optimization

1) Train-Validation Split

- Hold-out, K-fold, Leave-one-out, Leave-one-group-out

2) Hyperparameter tuning

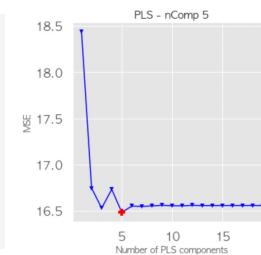
- 선택한 모델의 hyperparameter를 grid search 방식으로 바꾸어가며 train-validation 수행 확인
- Metric으로 MeanSquaredError, R2 등 사용
- PLSR 모델에서는 Component의 수

```
def optimise_pls_cv(X, y, n_comp):
    # Define PLS object
    pls = PLSRegression(n_components=n_comp)

    # Cross-validation
    y_cv = cross_val_predict(pls, X, y, cv=12)

    # Calculate scores
    r2 = r2_score(y, y_cv)
    mse = mean_squared_error(y, y_cv)
    rpd = y.std() / np.sqrt(mse)

    return (y_cv, r2, mse, rpd)
```



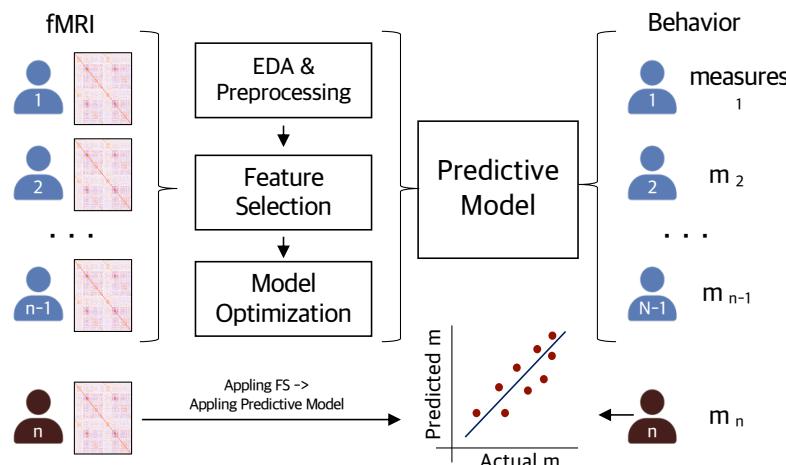
42

방법

데이터 분석 과정

PROJECT 4

- Connectome-based Predictive Modeling 절차 개요



4. Predictive Modeling

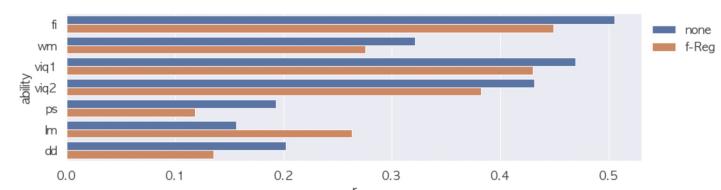
- 1) 1~n-1 의 참가자의 FC 패턴과 행동 측정치 간의 PLSR model을 fitting

- 2) Fitting된 PLSR model을 n 번째 참가자의 FC 패턴에 적용

```
# Model Fitting & Test
model_r = PLSRegression(n_components=nc)
model_r.fit(X_train_r_fs, y_train_fs)
y_pred_r = model_r.predict(X_test_r_fs)
```

- 3) 전체 참가자에 걸쳐 모든 절차를 반복

- 4) 예측된 행동 측정치와 실제 행동 측정치 간의 관계 확인

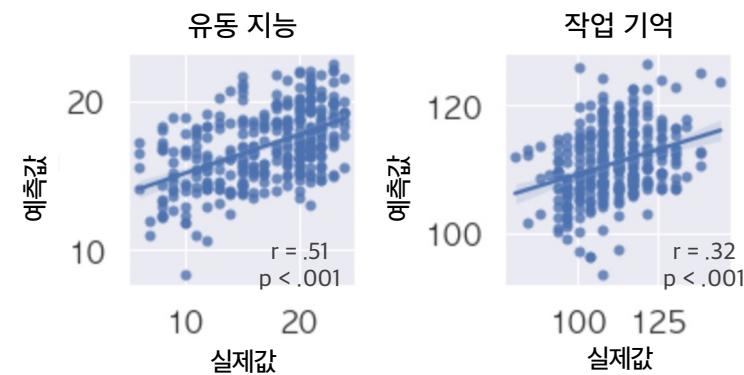
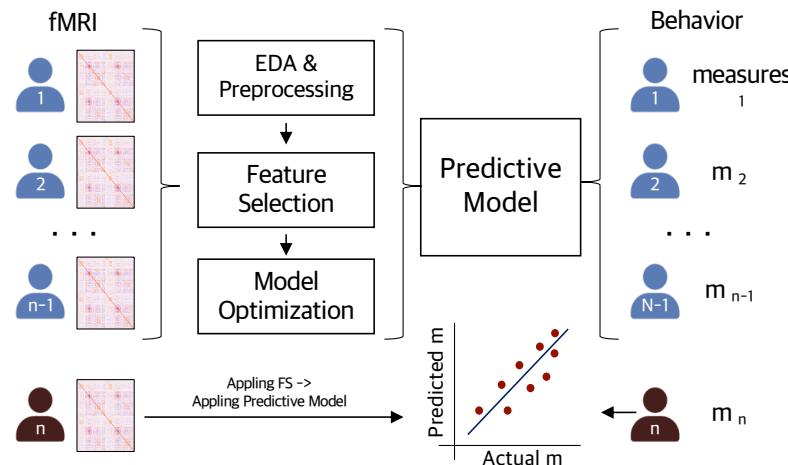


43

결과

FC 패턴과 행동 측정치 간 모델링 결과

- Connectome-based Predictive Modeling 절차 개요



- 성과
- 후속 연구를 위한 예측 모델링 절차 구체화
- 실험 설계에 기여
- 다차원 데이터 분석 및 모델링 노하우 습득