

Functional Description

adVAE aims to:

- 1. Aid Alzheimer's research by creating high-quality multimodal synthetic data (gene expression, EEG, MRI)
- 2. Easily integrate with existing biomedical workflows

User Interface

Command line interface (CLI)

Functionality

- Preprocess input data used to train the VAE
- 2. Latent space representation of various data modalities
- 3. Generation of synthetic data using latent space variables

Task Prioritization

- 1. Identify example data (gene expression, EEG and MRI)
- Create VAE model architecture for one modality (gene expression)
- 3. Train the model using the example gene expression dataset
- 4. Perform latent space optimization
- Validate the results by calculating KL loss and reconstruction error of generated dataset
- 6. Extend the model to include other modalities
- 7. Train, validate, and test the model using the bulk data

Goals and Milestones

- 1. Finish all design documentation
- 2. Set up GitHub repository
- 3. Create modules
- 4. Start implementing code
- 5. Create model for one modality
- 6. Extend model for other modalities
- 7. Peer review
- 8. Refine model and update design documentation

Solutions

- 1. Tensorize all input data
- 2. Create separate modules/classes for handling different data modalities
- 3. Validate by calculating KL loss and reconstruction loss

