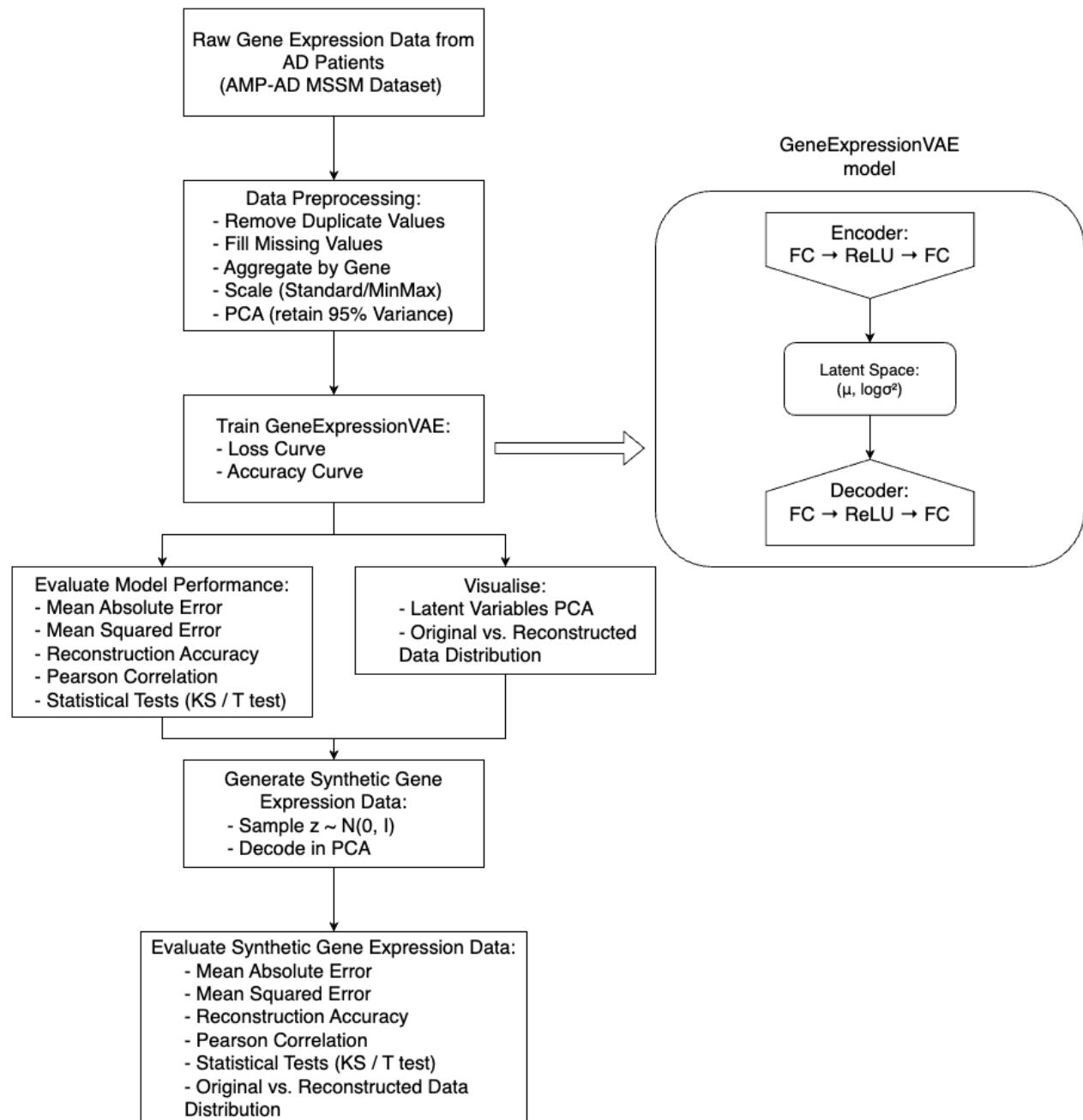
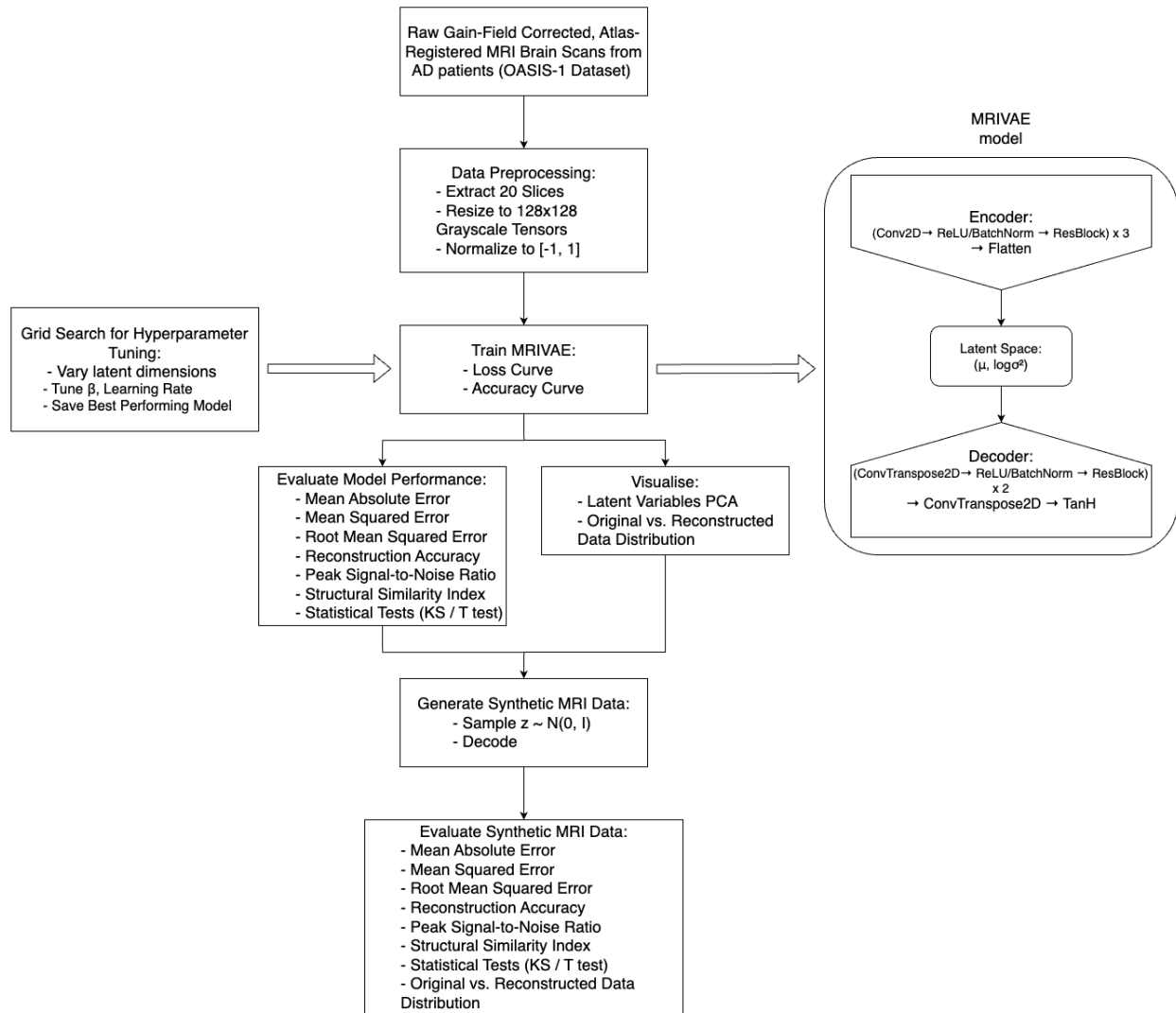


adVAE: Alzheimer's Data Variational Autoencoder

Design Documents



Model Architecture – GeneExpressionVAE Pipeline



Model Architecture – MRIVAE Pipeline

Functional Description
<p>adVAE aims to:</p> <ol style="list-style-type: none"> 1. Aid Alzheimer's research by creating high-quality multimodal synthetic data (gene expression, EEG, MRI) 2. Easily integrate with existing biomedical workflows

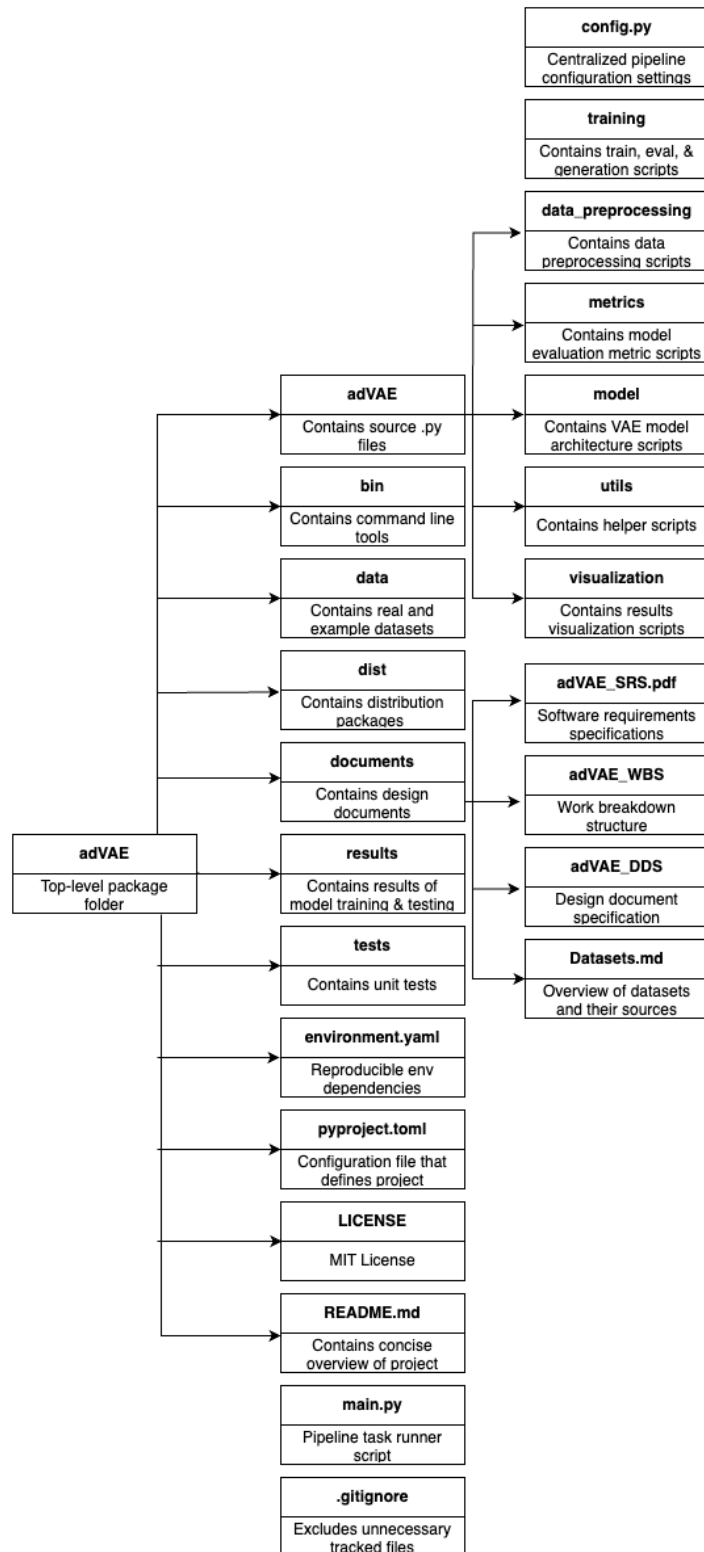
User Interface	Functionality
Command line interface (CLI)	<ol style="list-style-type: none"> 1. Preprocess input data used to train the VAE 2. Latent space representation of various data modalities 3. Generation of synthetic data using latent space variables

Task Prioritization
<ol style="list-style-type: none"> 1. Identify example data (gene expression, EEG and MRI) 2. Create VAE model architecture for one modality (gene expression) 3. Train the model using the example gene expression dataset 4. Perform latent space optimization 5. 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
<ol style="list-style-type: none"> 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
<ol style="list-style-type: none"> 1. Tensorize all input data 2. Create separate modules/classes for handling different data modalities 3. Validate by calculating KL loss and reconstruction loss

DDS



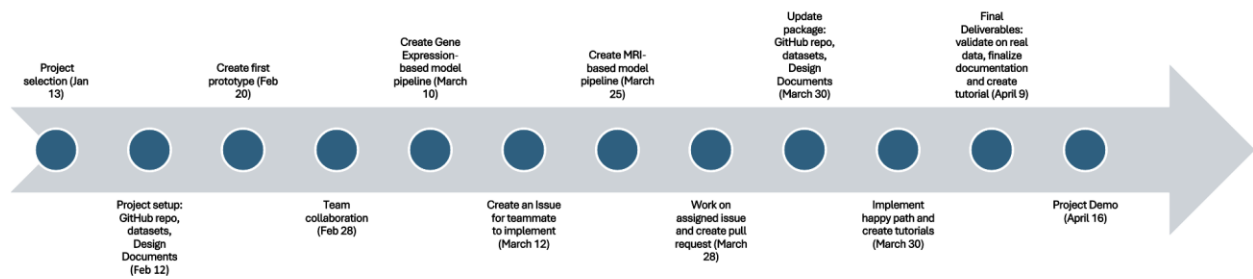
Module Structure

```

[adVAE]
LICENSE
README.md
[adVAE]
  __init__.py
  config.py
  [data_preprocessing]
    __init__.py
    [gene_expression]
      __init__.py
      dataset.py
      modified_preprocess.py
      preprocess.py
    misc_utils.py
    [mri]
      __init__.py
      dataset.py
      preprocess.py
  [metrics]
    __init__.py
    performance.py
    vae_loss.py
  [models]
    __init__.py
    gene_expression_vae.py
    mri_vae.py
  [training]
    __init__.py
    evaluate_gene_expression.py
    evaluate_mri.py
    evaluate_synthetic_gene_expression.py
    evaluate_synthetic_mri.py
    generate_synthetic_gene_expression.py
    generate_synthetic_mri.py
    train_gene_expression.py
    train_mri.py
  [utils]
    __init__.py
    data_loader.py
    generator.py
    utils.py
  [visualization]
    __init__.py
    plot_utils.py
    visualize_gene_expression.py
    visualize_mri.py
[bin]
[dist]
  adVAE-0.1.0-py3-none-any.whl
  advae-0.1.0.tar.gz
[documents]
  Datasets.md
  adVAE_DDS.pdf
  adVAE_SRS.pdf
  adVAE_WBS.pdf
environment.yaml
main.py
pyproject.toml
[tests]
  __init__.py
  test_data_loader.py
  test_datasets.py
  test_evaluate.py
  test_generator.py
  test_metrics.py
  test_models.py
  test_preprocess.py
  test_training.py
  test_visualization.py
[tutorials]
  example_dataset_tutorial.ipynb
  real_dataset_tutorial.ipynb

```

Module Structure 2



Project Timeline