Towards Better Healthcare: Amalgamation of Time Series, Viome and Image Data

Dhruv Patel, Haikoo Khandor, Nishant Basu

Texas A&M University, College Station
Team Name: Model Mavericks

December 6, 2024

Introduction

- Novel deep learning approach for predicting lunch calorie intake
- Integration of multiple data modalities:
  - Continuous Glucose Monitoring (CGM)
  - Viome data
  - Food images
- Specialized neural network model with multimodal attention fusion

Methodology

- Data Preprocessing and Encoding
- Model Architecture:
  - CGM Encoder (LSTM)
  - Image Encoder (CNN)
  - Demo Encoder (MLP)
- Multimodal Attention Fusion
- Prediction Layer

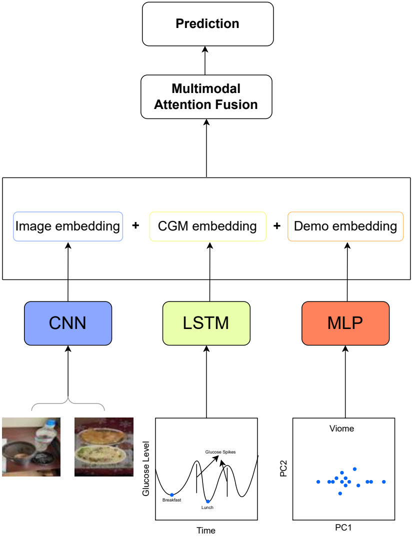
Data Preprocessing

- CGM Data:
  - Feature extraction (mean, std dev, etc.)
  - Scaling and tensor conversion
- Viome Data:
  - Encoding categorical variables
  - Scaling numerical variables
  - PCA for microbiome data
- Food Images:
  - Resizing and normalization
  - Gaussian blurring and color conversion

Model Architecture

- CGM Encoder:
  - LSTM network (input size: 1, hidden size: 64, 2 layers)
  - Fully connected layer for fixed-size encoding
- Image Encoder:
  - CNN with multiple convolutional layers
  - ReLU activation, batch normalization, max-pooling
- Demo Encoder:
  - MLP with linear layers, ReLU activations, batch normalization

Model Architecture



Multimodal Attention Fusion

- Attention Mechanism:
  - Modality-specific embedding layers
  - Multi-head attention layer
- Fusion Process:
  - Layer normalization
  - Feed-forward neural network
- Prediction Layer:
  - Linear predictor for lunch calorie intake

Training Process

- Loss Function: Root Mean Square Relative Error (RMSRE)
- Optimizer: Adam with learning rate scheduling
- Gradient clipping for stable convergence
- Early stopping to prevent overfitting

Dataset Description

- Collected over 10 days for 40+ participants
- Components:
  - Demographic and Viome Data
  - Continuous Glucose Monitoring Data
  - Meal Images
  - Nutritional Labels

Results

- Training Performance:
  - Initial Training Loss: 1.0285
  - Final Training Loss: 0.7162
  - Initial Validation Loss: 1.1354
  - Final Validation Loss: 0.7876
- Early stopping after 14 epochs
- Rapid initial improvement (17.5% for training, 32.7% for validation)

Conclusion and Future Work

- Successful integration of CGM, food images, and viome data
- Rapid convergence and consistent performance
- Future work:
  - Refine model architecture
  - Expand dataset
  - Conduct extensive real-world testing
- Potential for improving prediction accuracy and generalizability