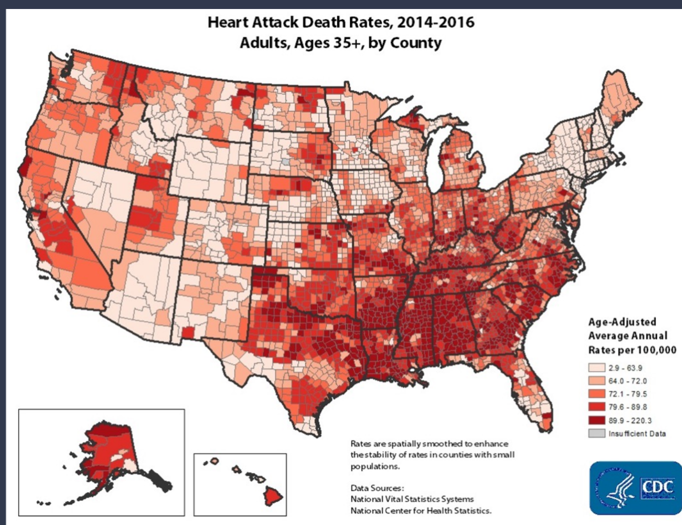
A Probabilistic Graphical Model for Major Adverse Cardiac Event Prediction

Sakin Kirti and Joshua Freeze
CSDS 491: Probabilistic Graphical Models

Contents

Introduction

Heart Disease



- Leading cause of death in the US
- Costs the US $229 billion annually
- Strongly tied to other medical disorders like diabetes and obesity
- Coronary artery disease causes 50% of heart disease-related deaths

Introduction

MACE Events

- **Major Adverse Cardiovascular Events**: a set of events used by clinicians and researchers to track and investigate deadly outcomes relating to cardiovascular health
  - Acute myocardial infarction
  - Stroke
  - Cardiovascular death

Motivation

- Current methods for MACE event prediction is formulaic
- Time-to-event analysis is done in the same way
- More advanced techniques are needed
  - Can be used in a more general sense

- Heart calcification can lead to atherosclerosis and eventually to strokes, needing bypass surgery
  - Creating early warning systems

Approach & Rationale
Data Availability

- Biomedical Imaging lab work provided the data for use
  - 23 features (general heart health)
  - 61 features (heart calcification)
- Permission from lab before using any data

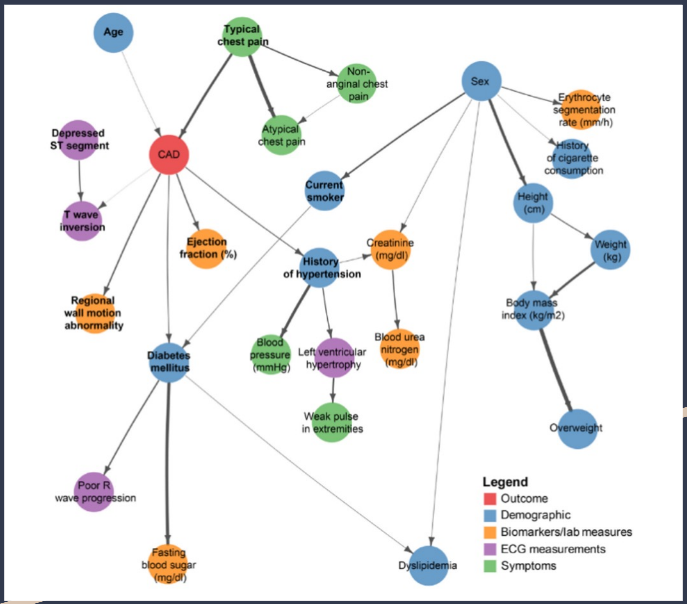
- Dataset contained little missing data <5% of 3900 samples
  - Imputation

- Dataset was imbalanced
  - 90% no MACE, 10% MACE
  - SMOTE was used to increase size of underrepresented class

Approach & Rationale
Model Metrics

- Each model was measured based on several outputs
  - Classification power
    - Risk prediction
  - Model metrics
    - Accuracy, ROC curves, AUC
  - Time-to-event analysis
    - Survival analysis

Approach & Rationale
Models

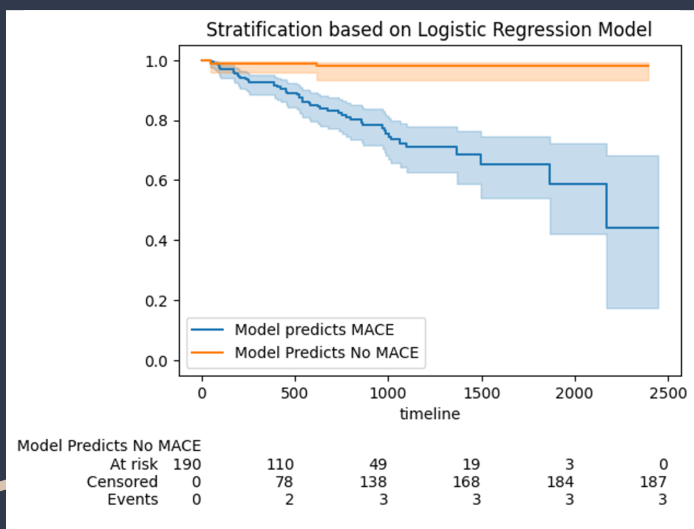


Built 3 models, wanted to understand the pros/cons of each

- Logistic regression → easy, quick
- Neural network → powerful classification
- Bayesian model → understanding of node connections

Approach & Rationale Results

Logistic Regression



Approach & Rationale

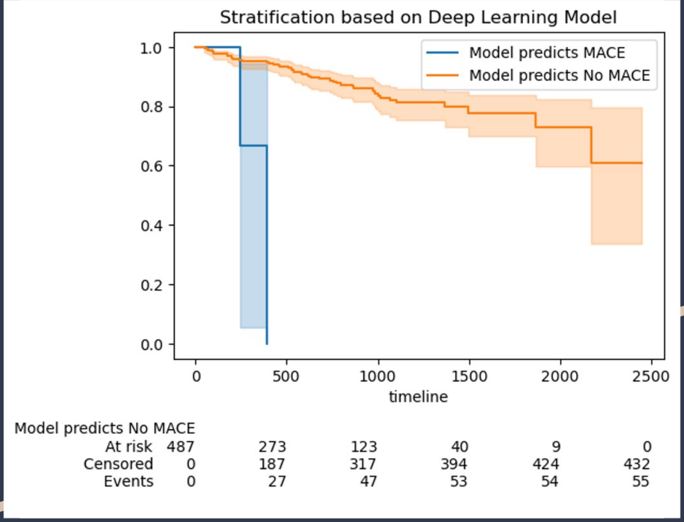
- Logistic Regression a good first model
  - Examine any correlation in the dataset
  - Simple, quick, easy (with scikit-learn)

Results

- Accuracy: 66.1%
- AUC: 0.685
  - Vastly overpredicted MACE (57 MACE, 200+ predicted)
- Survival analysis:
  - Working separation between classes

Approach & Rationale Results

Neural Network



Approach & Rationale

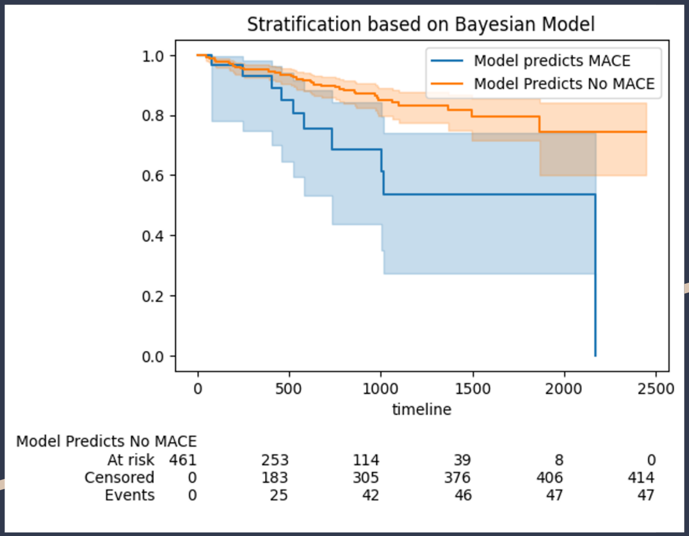
- Model architecture which output logits
- Step up from logistic regression
- Known for powerful classification
- For comparison with Bayesian model

Results

- Accuracy: 87.4%
- AUC: 0.907
  - Note: did not do well with classifying positive cases (57 MACE, only 11 detected)
  - Still significantly better than the logistic regression model
- Survival analysis
  - Classification and low numbers cause sharp drop in time-to-event analysis

Approach & Rationale Results

Bayesian Model



Approach and Rationale

- Outputs probability of MACE occurring
  - Risk score/calculation
- Node relationships based on
  - Gupta et al for general heart features
  - Conversations with Josh's PI for heart calcification data

Results

- Accuracy (rounded): 90.5%
- AUC: 0.720
- Survival Analysis
  - Accurate separation of classes

Conclusion
Project Models

3 Models built

- Logistic Regression
  - good insight into the data
- Deep Neural Network
  - classification power and comparison for Bayesian Model
- Bayesian Model
  - classification and confidence for predictions

Project-specific Outcomes

- Deep neural model did the best and created risk scores in the process (probability proxy)
- Neural network did well with classification

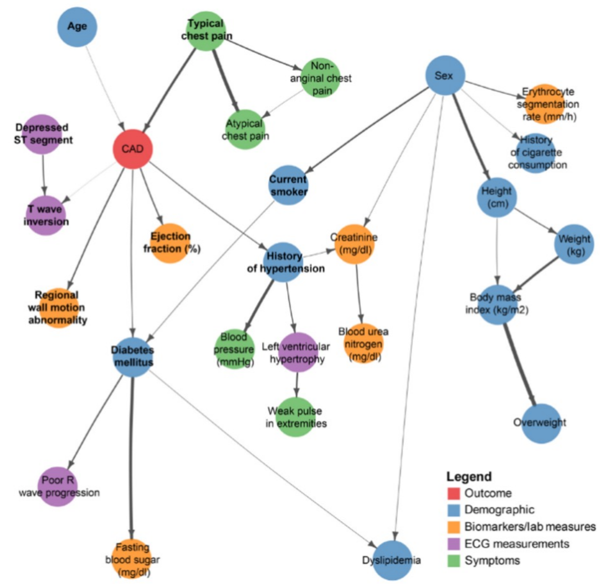
Next Steps

- Rethink node classifications and maybe reduce the number of nodes
- Spend some time on hyperparameter tuning

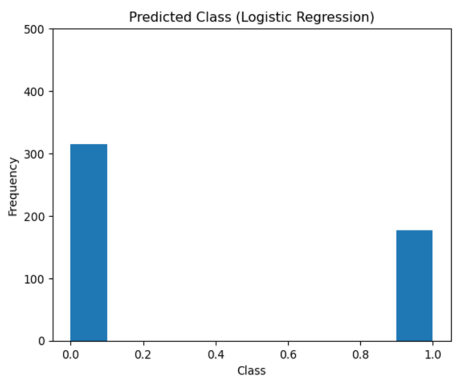
Conclusion
Project Models

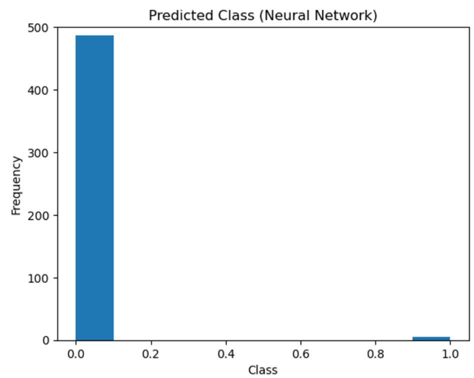
Broad Outcomes

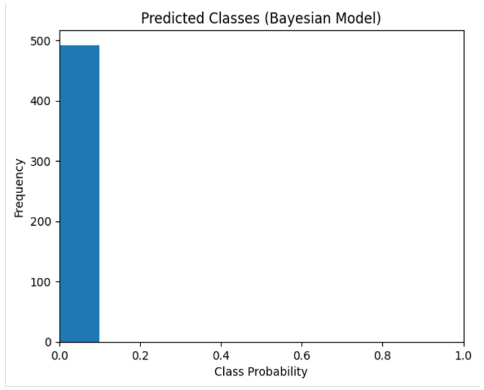
- Literature contains examples of probabilistic models for heart data
- We expand on this by adding calcification data

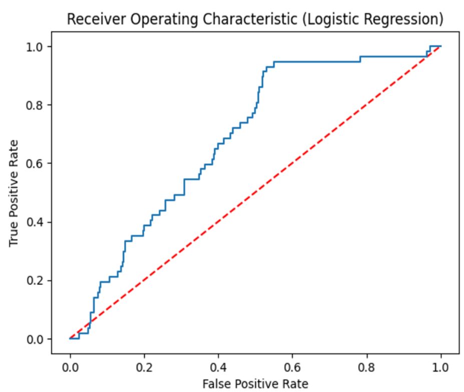


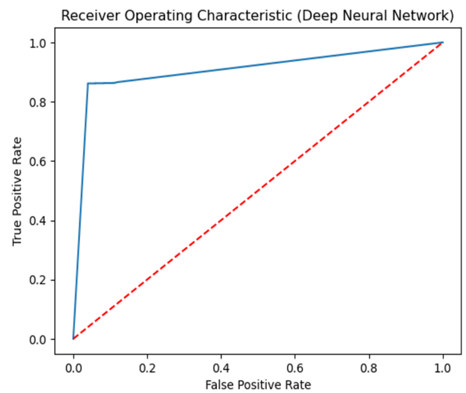
Thank you!

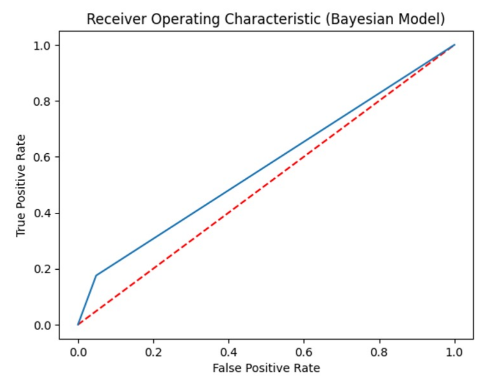




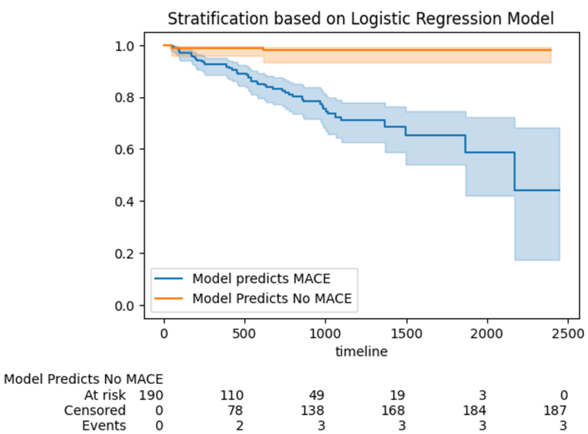


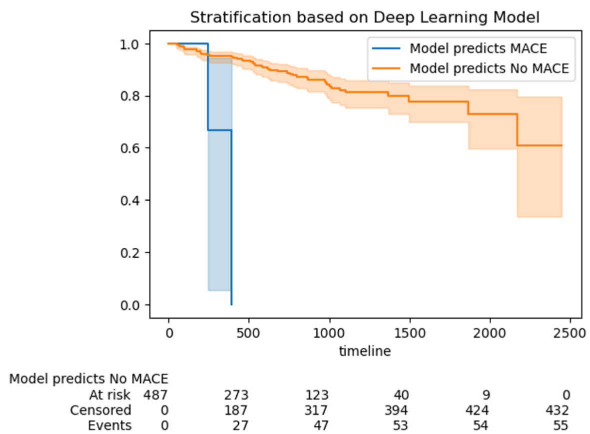


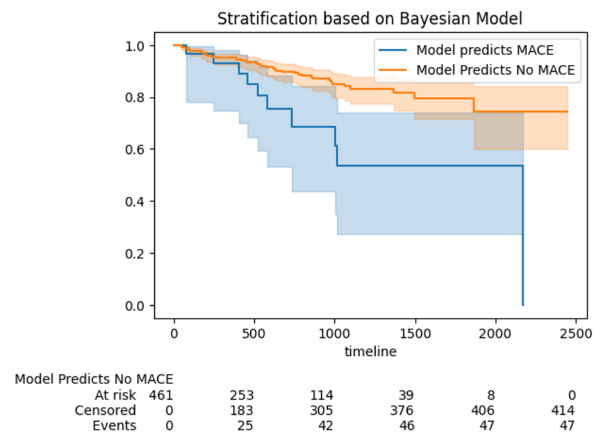




Results Summary







Results Summary