PROJECT PRESENTATION

Vignesh Rani

Agenda

- Problem Statement
- Data Exploration
- Features
- Methodology
- Results
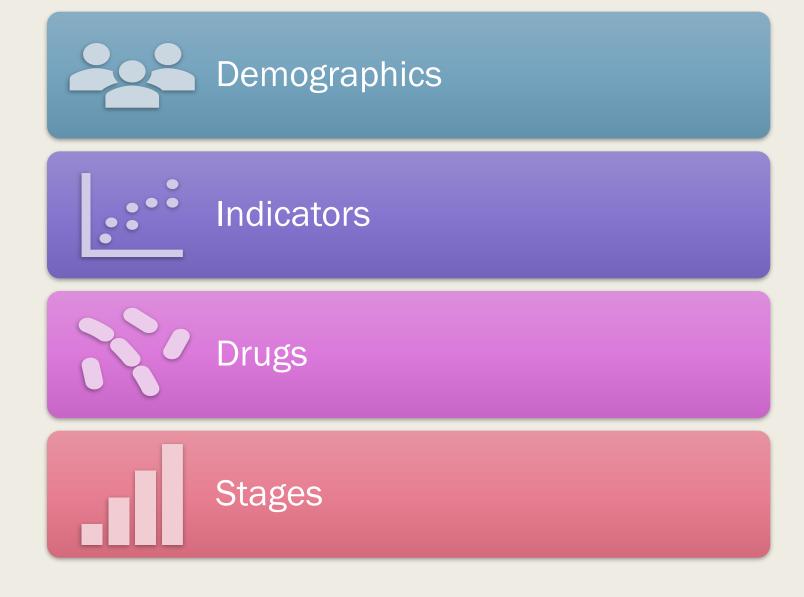- Improvements
- Conclusion
- Evaluation

Problem Statement

- Given a set of longitudinal data of different lab measurements for patients diagnosed with chronic kidney disease (CKD).

- Predict whether a patient will progress in CKD staging or not.

Data Exploration

Demographics

Indicators

Drugs

Stages

Demographics

- **300** patients

- **Categorical**: Race (5), Gender (2)

- **Discrete**: Age
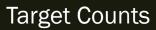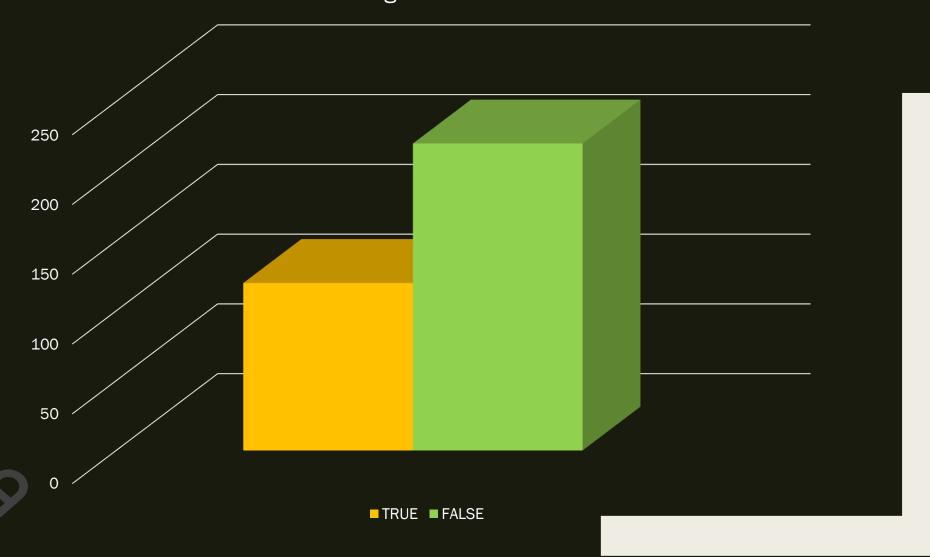
Stages

Target Counts

Demographics

Demographics

Race vs Gender

Count

Age vs Gender

Count

Indicators
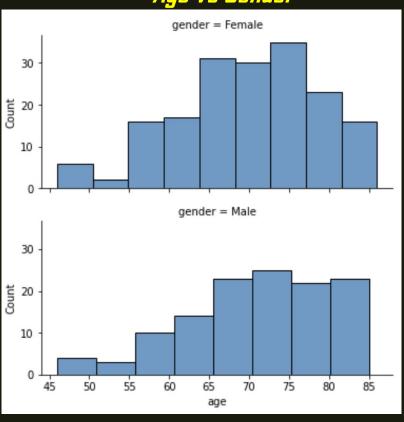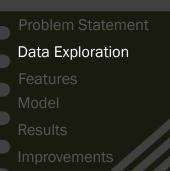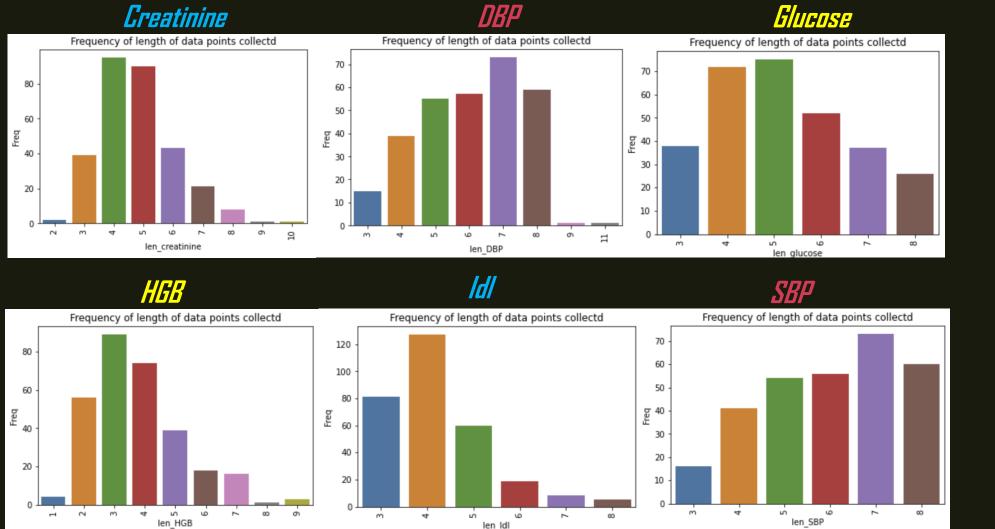
- **6** types

- **700** days

- **Continuous**: Measurement values

- Number of measurements and timestamps
  - vary for same patients per indicator
  - vary across all patients
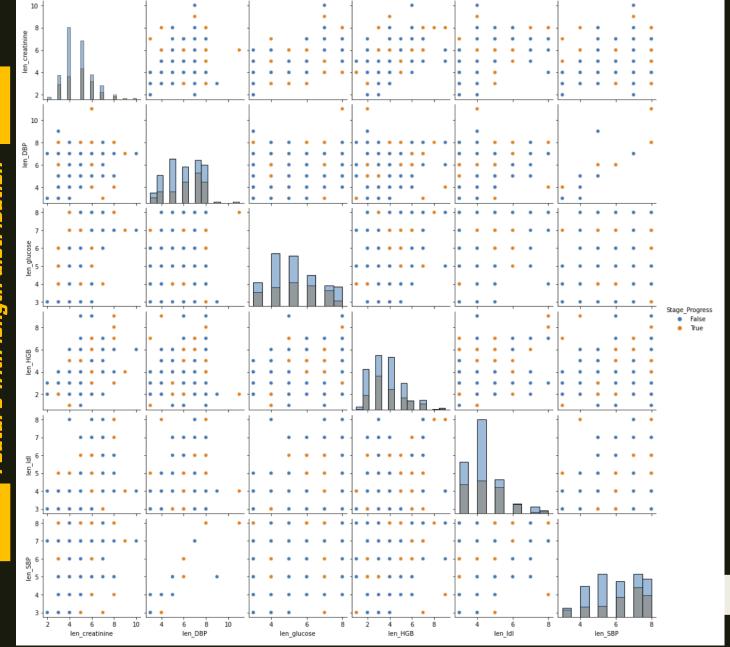
Indicators

Indicators

Feature with length distribution

Feature with length distribution

Indicators

Drugs

- **21** drugs

- **26** dosage levels

- **28** missing

- **700** days

- **Categorical**: Drug types, Dosage levels

- Varying **number** of drugs and dosage **levels**

Drugs

Number of drugs

Freq

Drugs

Drug types

Freq

Problem Statement
Data Exploration
Features
Model
Results
Improvements
Conclusion

Drugs

Drug types

Freq

Drugs

Freq
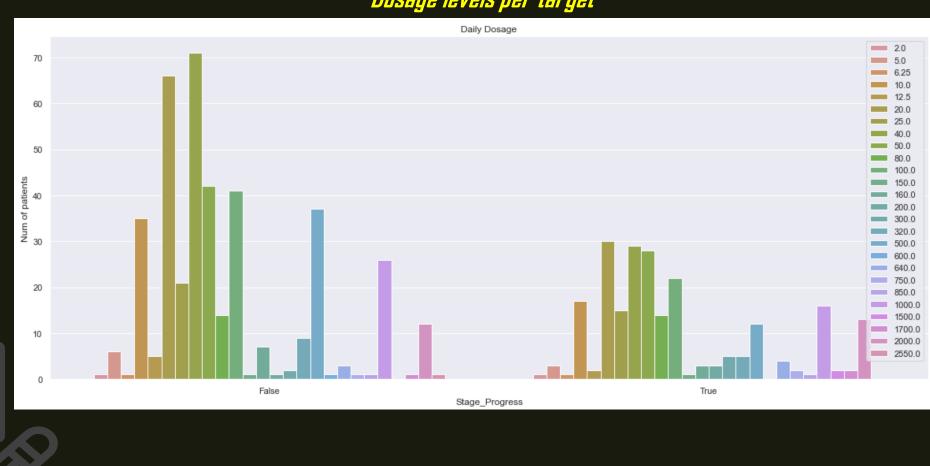
Dosage levels per target

Drugs
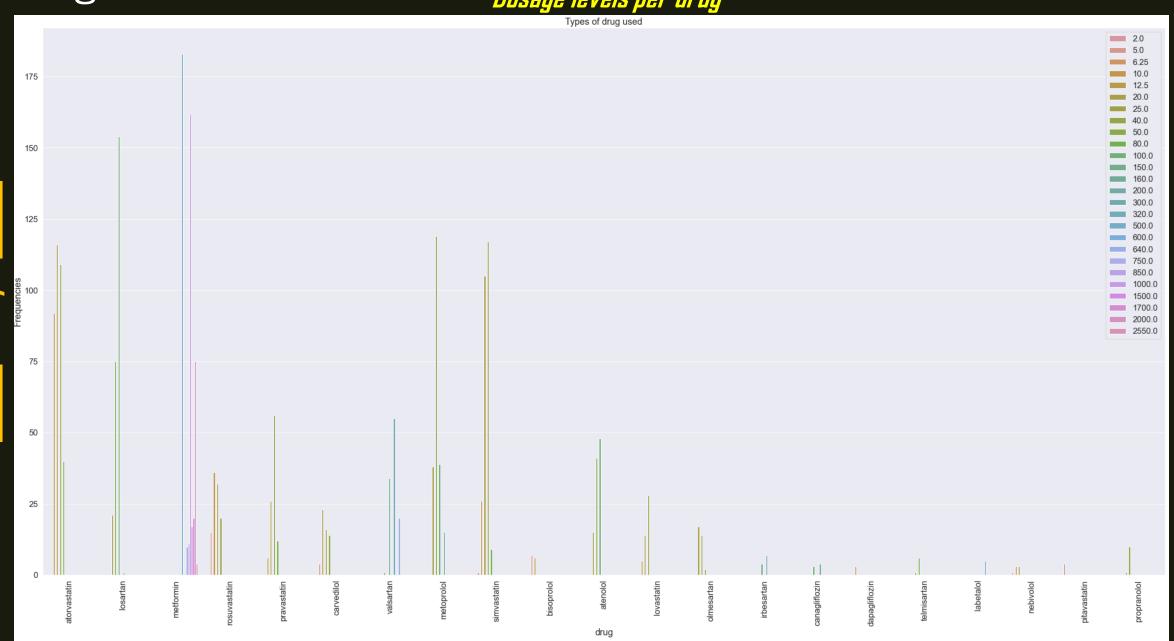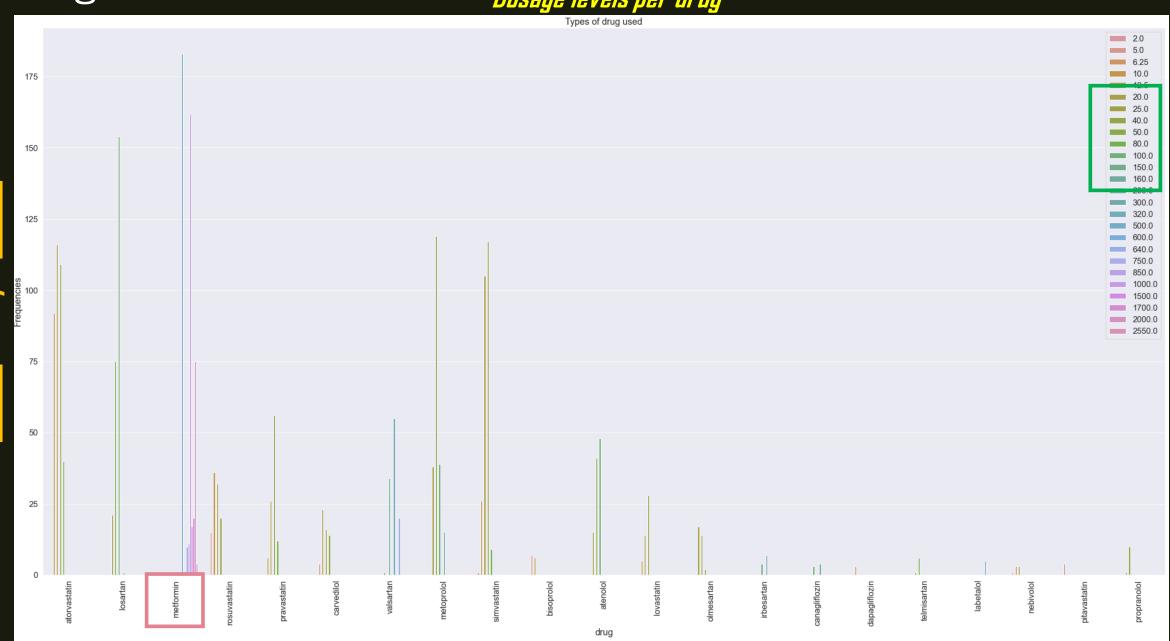
Dosage levels per drug

Drugs

Dosage levels per drug

Features

Demographics

Combine minority races

Normalize age

Indicators

Drugs

Features

Demographics

Combine minority races

Normalize age

Indicators

Interpolate the data points linearly to form 1D signal

Normalize the values of 1D signal

Drugs

Features

Demographics

Combine minority races

Normalize age

Indicators

Interpolate the data points linearly to form 1D signal

Normalize the values of 1D signal

Drugs

Reduce the number of drugs

Form 1D signals with the dosage levels

Features
Selected drugs

simvastatin

metoprolol

rosuvastatin

> 30 samples

losartan

atorvastatin

metformin

Features

Samples

Patient 49, Drug: Rosuvastatin

Patient 0 , Indicator: Glucose

Dosage

Value

Days

Features

Why 1D signal/wave?

Features

How to input many 1D signals?
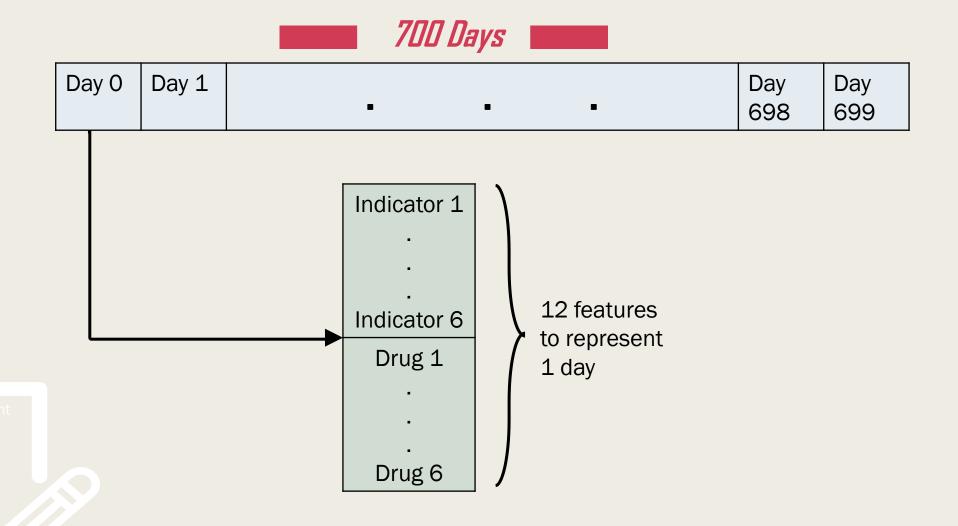
700 Days

Day 0	Day 1	. . .	Day 698	Day 699

Features

700 Days

Day 0	Day 1	. . .	Day 698	Day 699

Indicator 1
.
.
.
Indicator 6
Drug 1
.
.
.
.
Drug 6

12 features to represent 1 day

Features

Model

Wave Encoder

Wave Data
(Seq Length: 35)

[12, 1, 35] → ConvBlock → [64, 1, 18] → ConvBlock → [32 1, 9] → ConvBlock → [1, 1, 5] → GRU → [1,5]

Linear Encoder

Data on demographics

[1,3] → Linear → [1,10] → Linear → [1,5]

Concat → [1,10]

Classifier

Linear → [10,10] → Linear → [1,2] → **SoftMax**

Results
Train/Val

Results

Train/Val
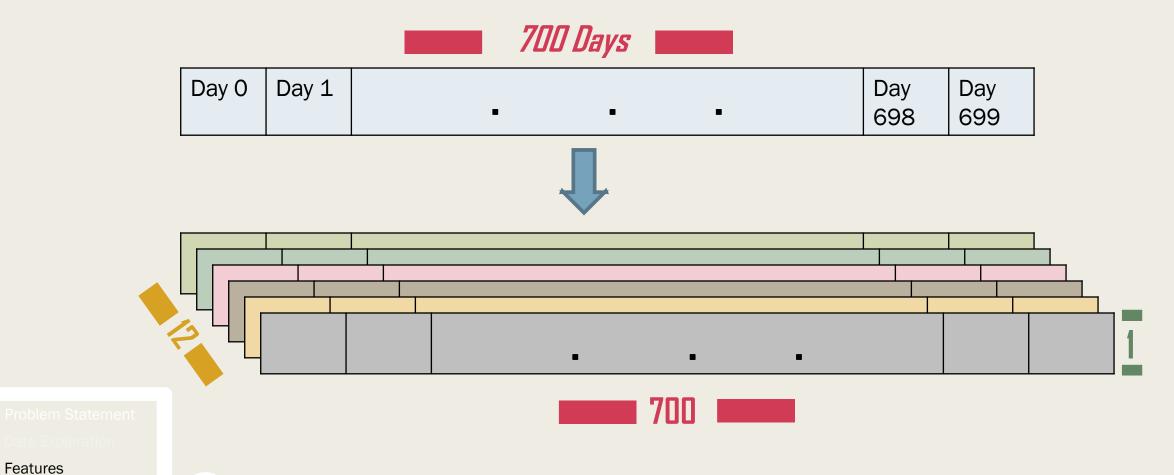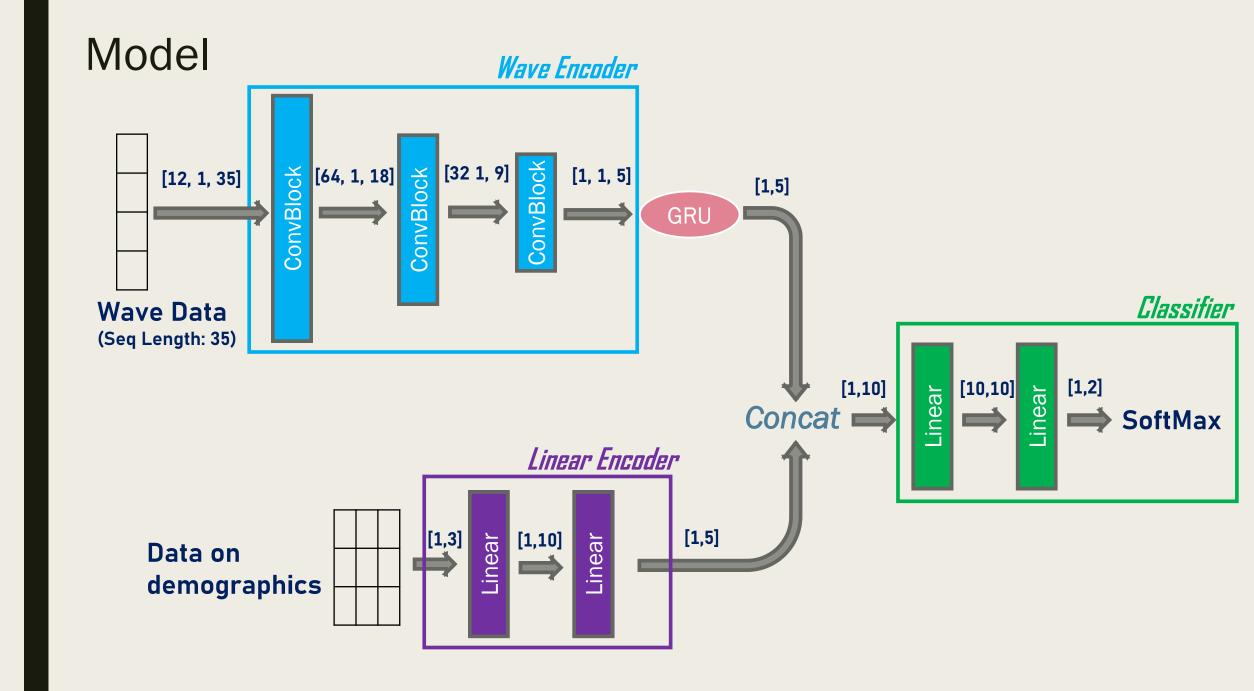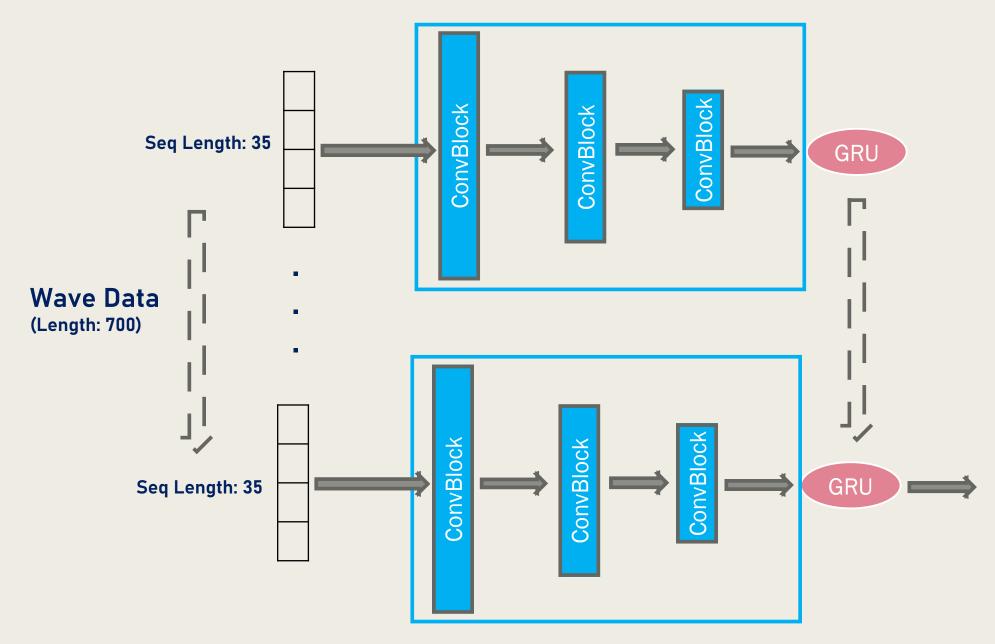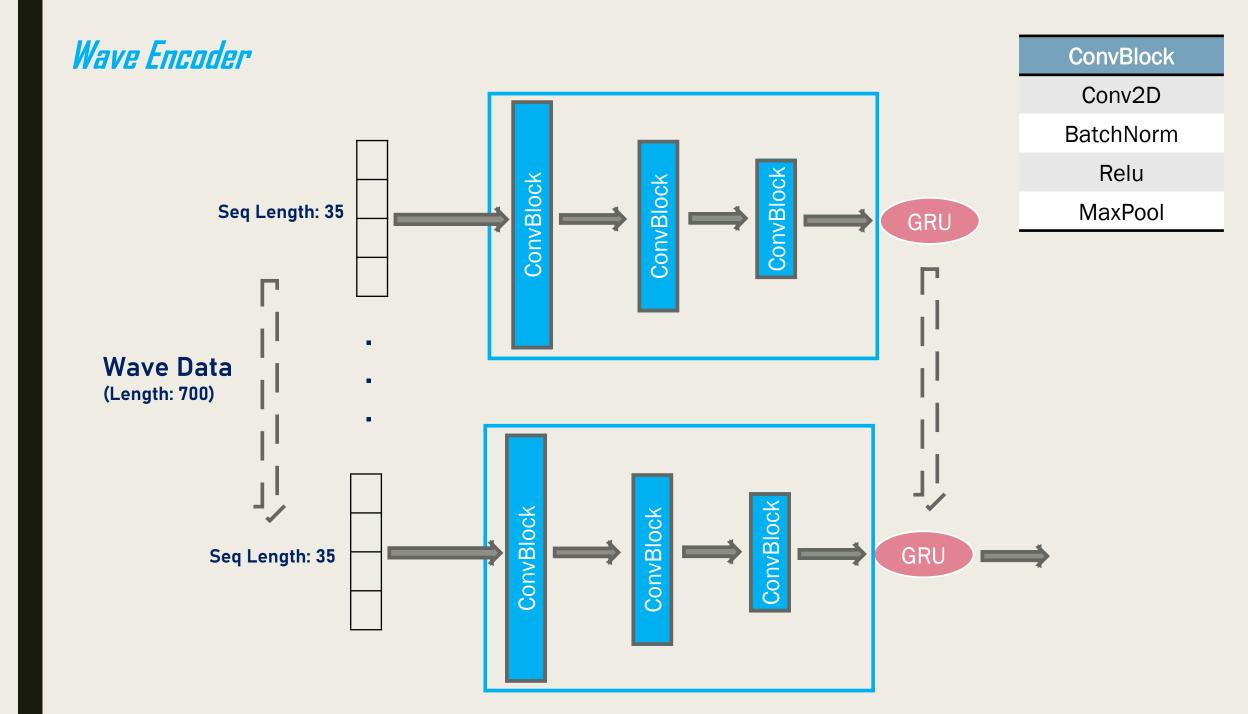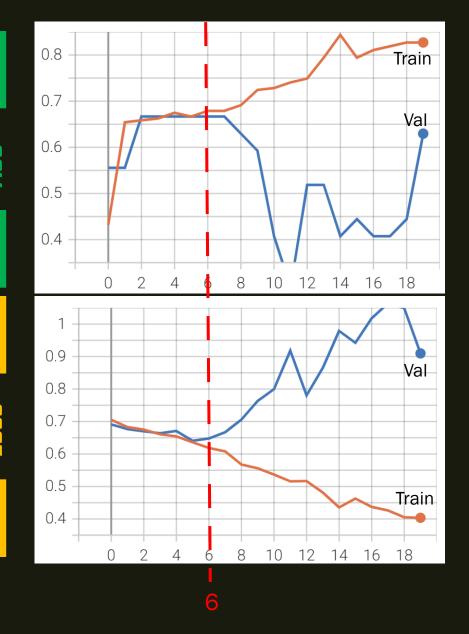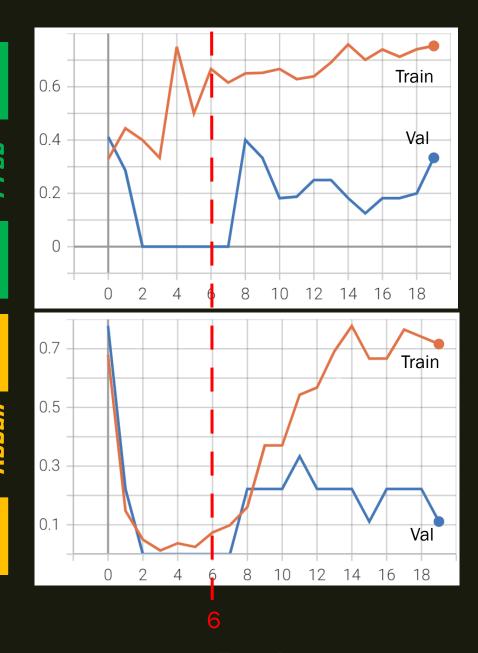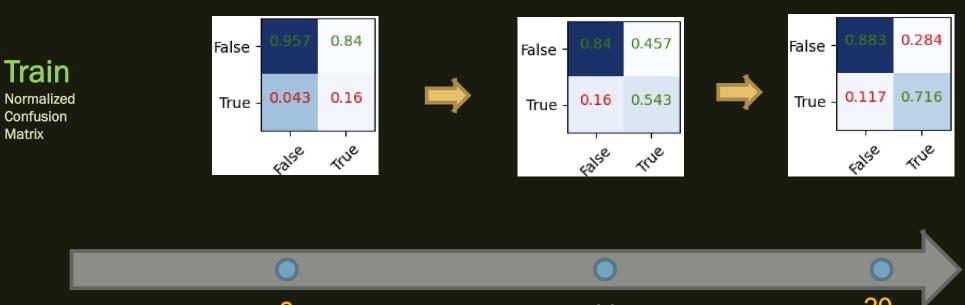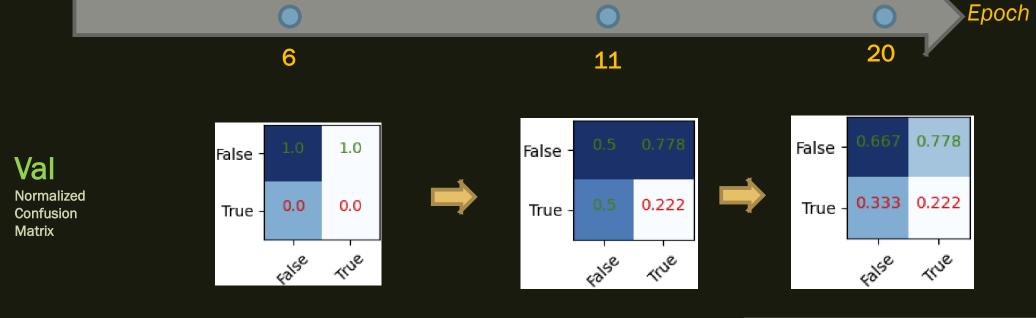
Results

Test

Normalized Confusion Matrix

	False	True
False	0.861	0.617
True	0.139	0.383

Improvements

Increasing collection of data samples

Optimizing the current model

Feature Engineering

Conclusion

Deep Learning approach

Multi modal architecture

Results not favorable

Requires further refinement

Thank You !