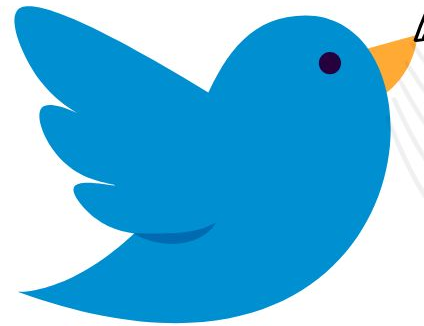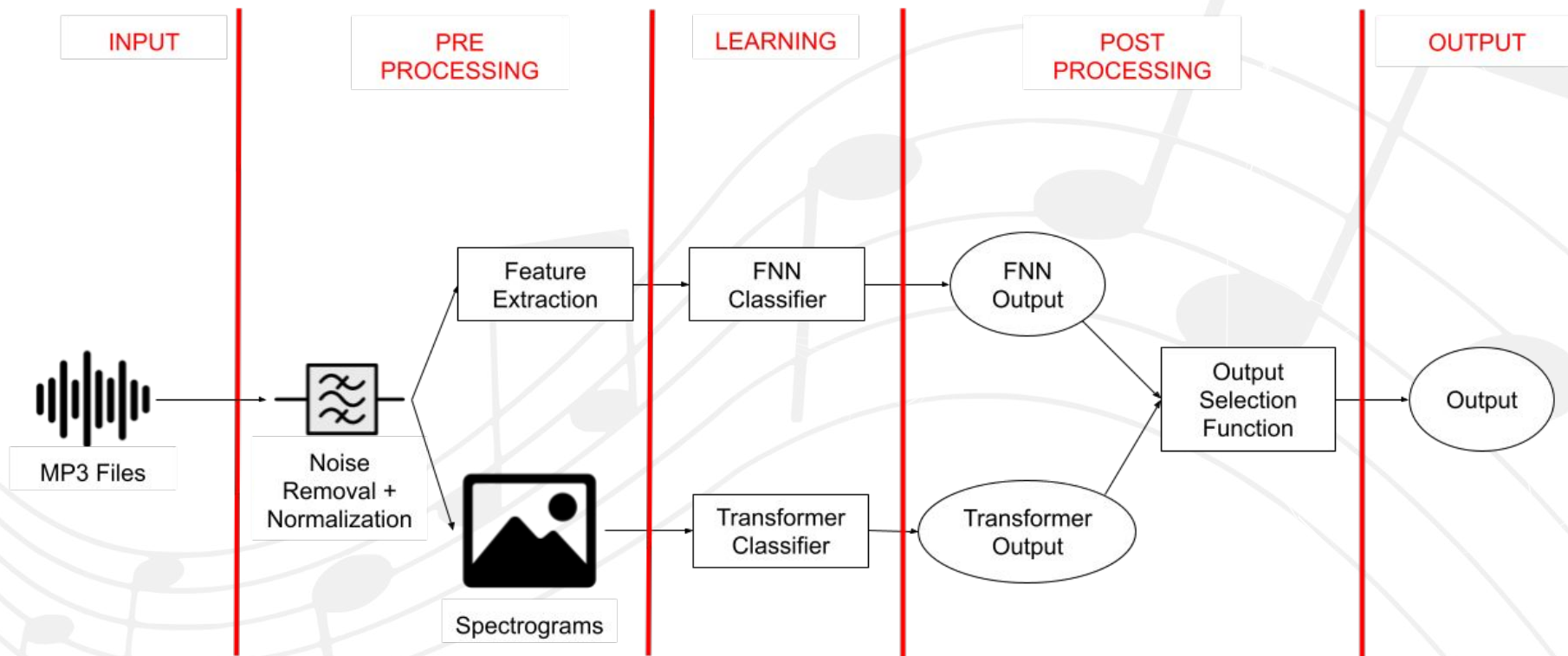# "Who Tweeted That?

Project Pitch

SYSC 5405

Group 3
Nadia Ahmed, Bardia Parmoun,
Huda Sheikh, Prianna Rahman

Review of our Meta Learning Strategy

# Review of our Meta Learning Strategy

- Created two separate classifiers:
  - Trained both on the same train/validation data
  - Audio features → FNN classifier (dense model)
    - MFCCs, Chroma, Spectral Contrast, Location
  - Spectrograms → vision transformers
- Combined predictions using Hard Voting
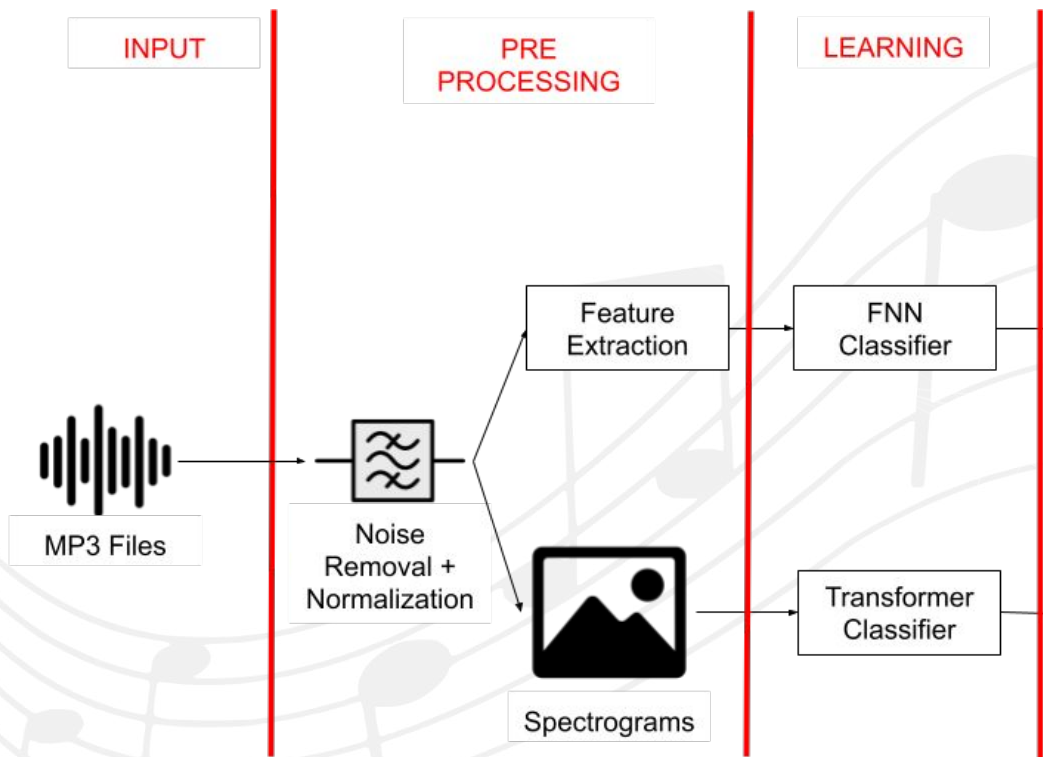
1. Data Pre-Processing
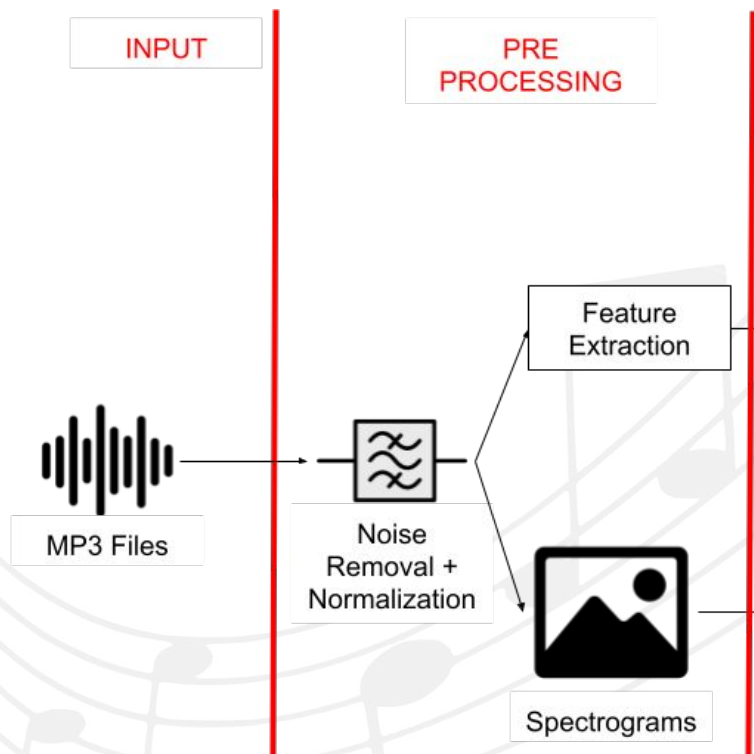
1. Data Pre-Processing

- Explored techniques:
  - Noise attenuation
  - Amplification
    - Clipping
    - Increased processing time

- Final solution:
  - Noise attenuation with Butterworth filter (3000-8000 Hz)
  - Normalization
  - Create spectrograms directly from audio files
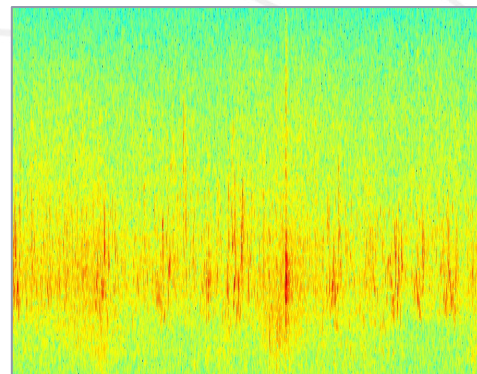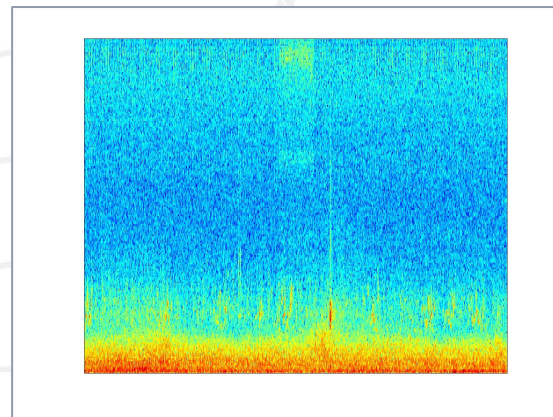
2. Training and Testing Protocol

2. Training and Testing Protocol

- Developed each classifier separately first and tried to optimize them
  - Experimented with different hyperparameters (learning rates, optimizer functions, early stopping, layers, etc.)
  - Experimented with sequential vs non-sequential data
  - Experimented with class weights → they are not needed!
  - Experimented with different feature combinations → ie. Location!
- Combined the two classifiers:
  - First, did K-Fold Cross Validation to ensure consistency
  - Then, trained the classifiers on all the data

2. Training and Testing Protocol

- Combined 2021 and 2022 data
- Performed 5-fold Cross Validation:
  - 70% for training
  - 15% for validation
  - 15% for test
- Made a final classifier
  - 85% for training
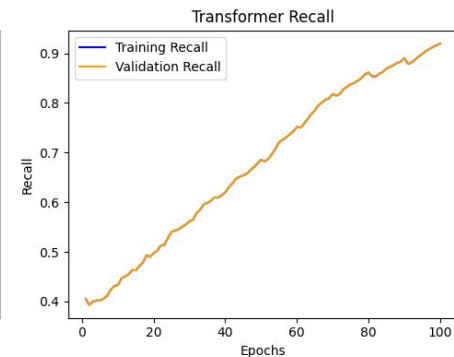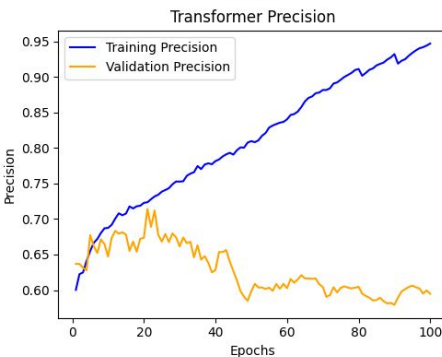 - 15% for validation
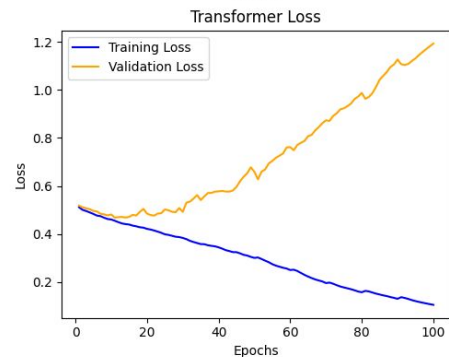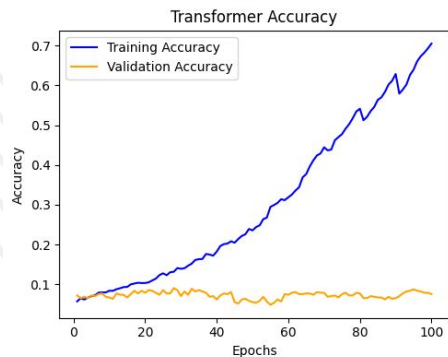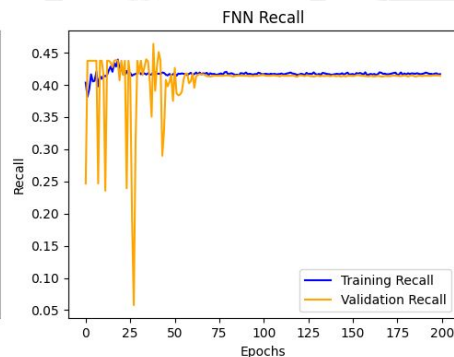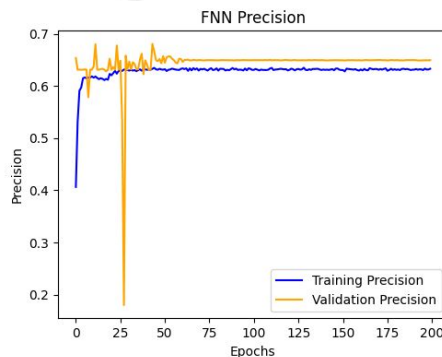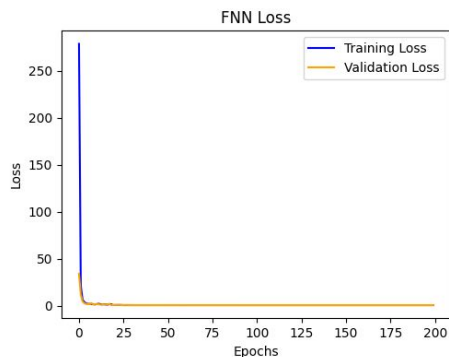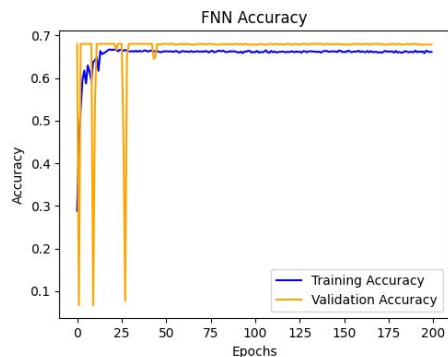
3. Post-Processing and Output
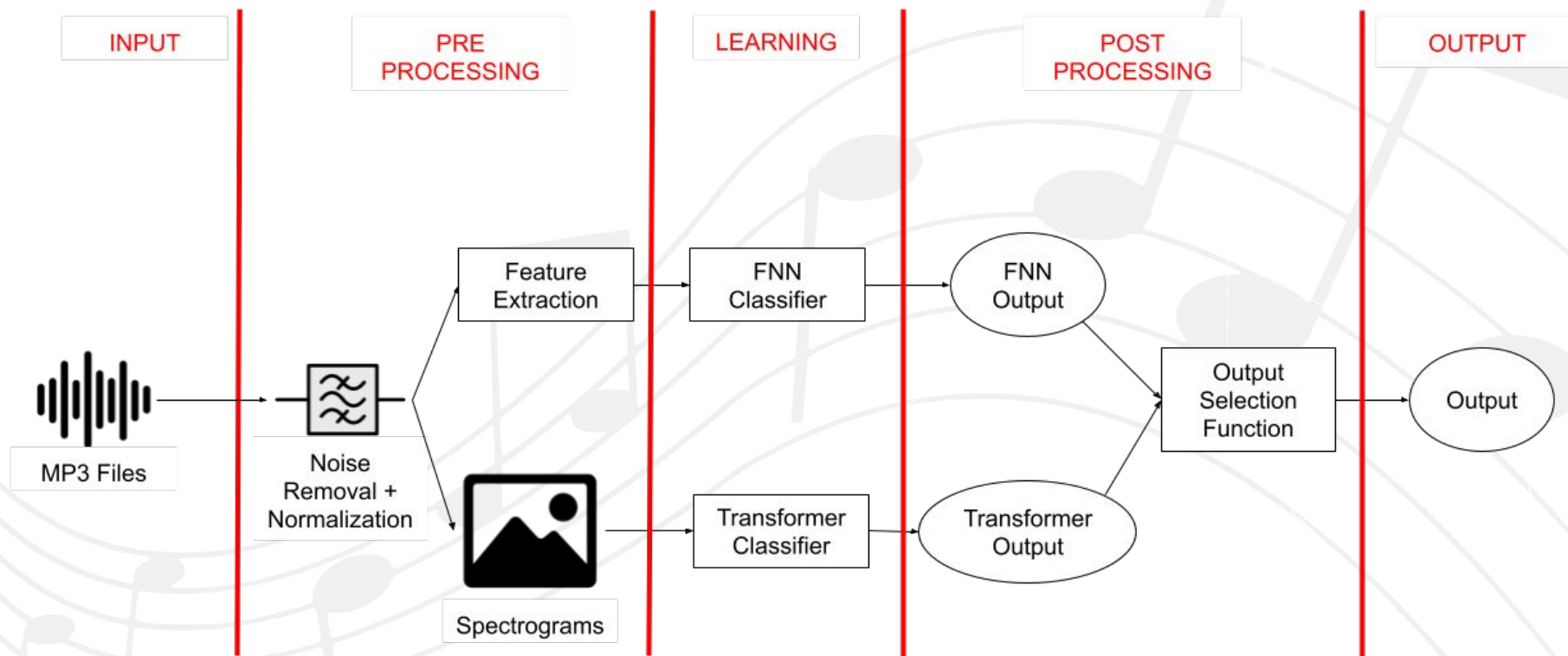
3. Output

- 100 epochs for the FNN
- 20 epochs the transformer
- 5 fold cross validation
- Only trained on 85% of the data (70% train & 15% validation)
- Pessimistic estimate!

F1 Score = 0.592 ± 0.012

3. Performance – Trained on All the Data

4. Challenges

- Class similarities
- Class imbalance
- Lack of sufficient and noise-free training data
- Relevant feature selection

Our strategy?

Improve precision as much as possible to increase F1

Recall seemed to remain unaffected → move the classifier to the **negative** side a bit and attempt to **overfit** even!

Thank you!