Deep Learning Model for Classifying Snoring Sounds

By Aditya Pendyala

Importance of Snoring Detection

- Affects millions worldwide but often dismissed as harmless.
- Persistent snoring may signal serious health conditions like:
  - Obstructive Sleep Apnea (OSA).
  - Increased risk of cardiovascular diseases, diabetes, and accidents.
- Early and accurate detection is crucial for timely intervention.

Challenges in Current Solutions

- Polysomnography:
  - Gold standard but expensive and labor-intensive.
  - Requires overnight stays in sleep labs.
  - Inconvenient and inaccessible for many.

Background & Motivation

Everyday Applications

- Consumer health devices, e.g., smartwatches and sleep trackers.
- Monitor sleep quality at home.
- Empower individuals to take proactive health measures:
  - Seek medical advice.
  - Adjust lifestyle habits to improve sleep.

Objective & Metrics

Objective

- Develop a snoring detection system using VGG16 transfer learning on audio features like mel-spectrograms.
- Analyze model performance in distinguishing snoring vs. non-snoring audio to enable applications in clinical and consumer health settings.

Metrics

- Precision: Measures correct snoring predictions out of all predicted snoring cases. Avoids false alarms in devices.
- Recall: Ensures snoring events are detected, avoiding missed diagnoses.
- F1 Score: Balances precision and recall, suitable for imbalanced datasets.
- AUC-PR: Evaluates model effectiveness in identifying snoring across thresholds, ideal for class imbalance scenarios.
- Goal: Deliver a reliable, high-performing system for medical and personal use cases.

Dataset Overview

Snoring Data

Cleaning & Transformations

- Normalize audio signals for consistent scaling.
- Convert raw waveforms into mel spectrograms to capture meaningful frequency and temporal patterns.

Feasibility

- Well-suited for supervised learning due to clear labels.
- Contains rich, real-world features essential for snoring detection.
- Direct applications in medical diagnostics and consumer sleep monitoring.

Structure

- Labeled audio files in two folders: snoring and non-snoring.
- Balanced dataset: 500 .wav files for each class.

Data Characteristics

- Snoring Samples:
  - 363 samples: Snoring sounds from children, men, and women (no background noise).
  - 137 samples: Snoring with background non-snoring sounds.
- Non-Snoring Samples:
  - 10 categories of background sounds (e.g., baby crying, rain, television noise).

Data Challenges

- Raw audio waveforms are high-dimensional and less structured for modeling.
- Requires feature extraction (e.g., mel spectrograms) to make the data usable.

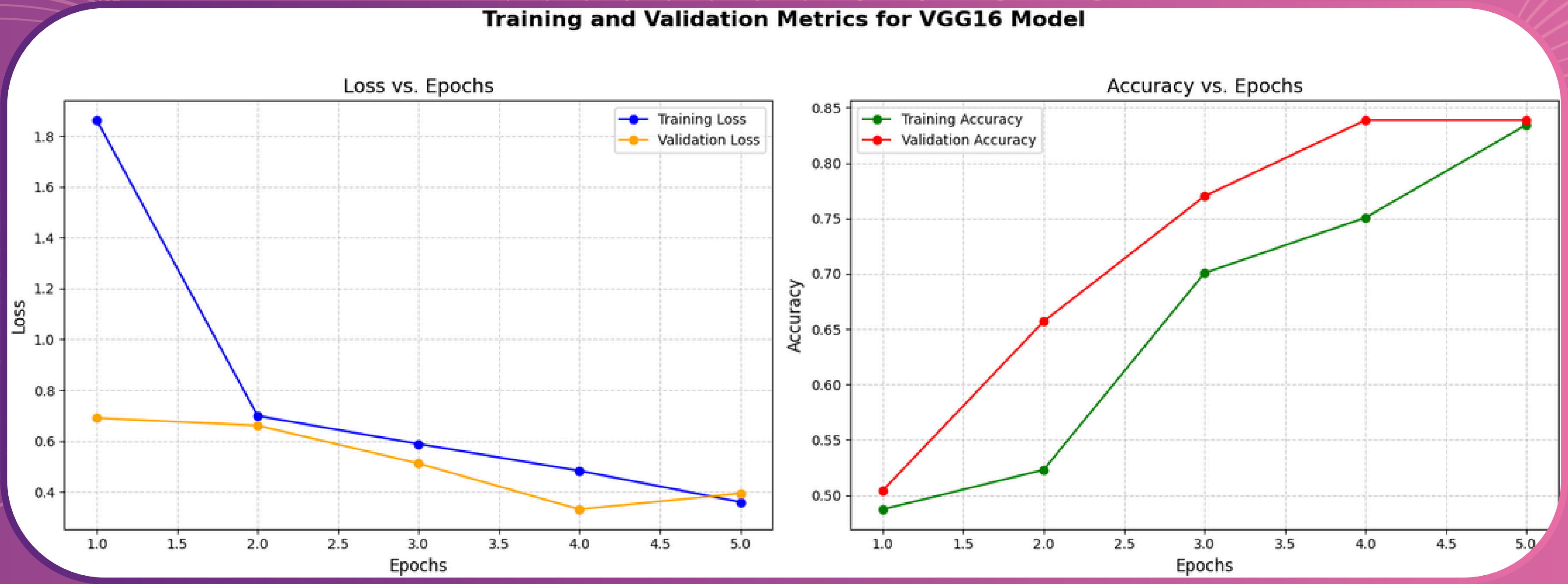
Model Comparison

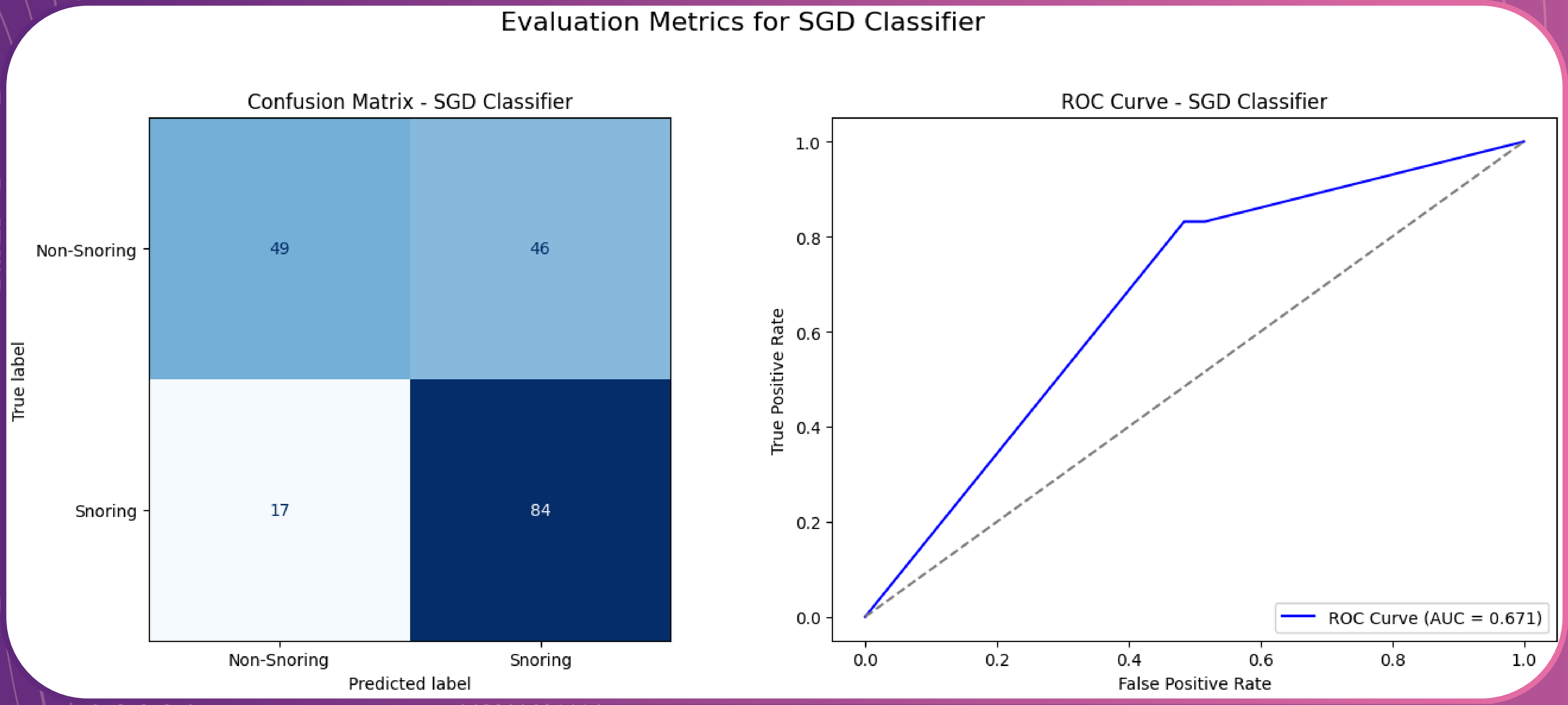
VGG16 with Transfer Learning

- Architecture:
  - Deep CNN pre-trained on ImageNet.
  - Transfer learning: Freeze convolutional layers and add custom fully connected layers.
  - Input: Mel-spectrograms treated as image-like data.
- Why VGG16?
  - Captures spatial hierarchies in Mel-spectrograms.
  - Leverages pre-trained knowledge for better performance with limited data.
- Advantages:
  - Faster training due to transfer learning.
  - Extracts complex patterns from input data.
- Performance:
  - AUC: 0.98 (excellent model performance).
  - Confusion Matrix:
    - Non-Snoring: 278 true negatives, 16 false positives.
    - Snoring: 259 true positives, 30 false negatives.

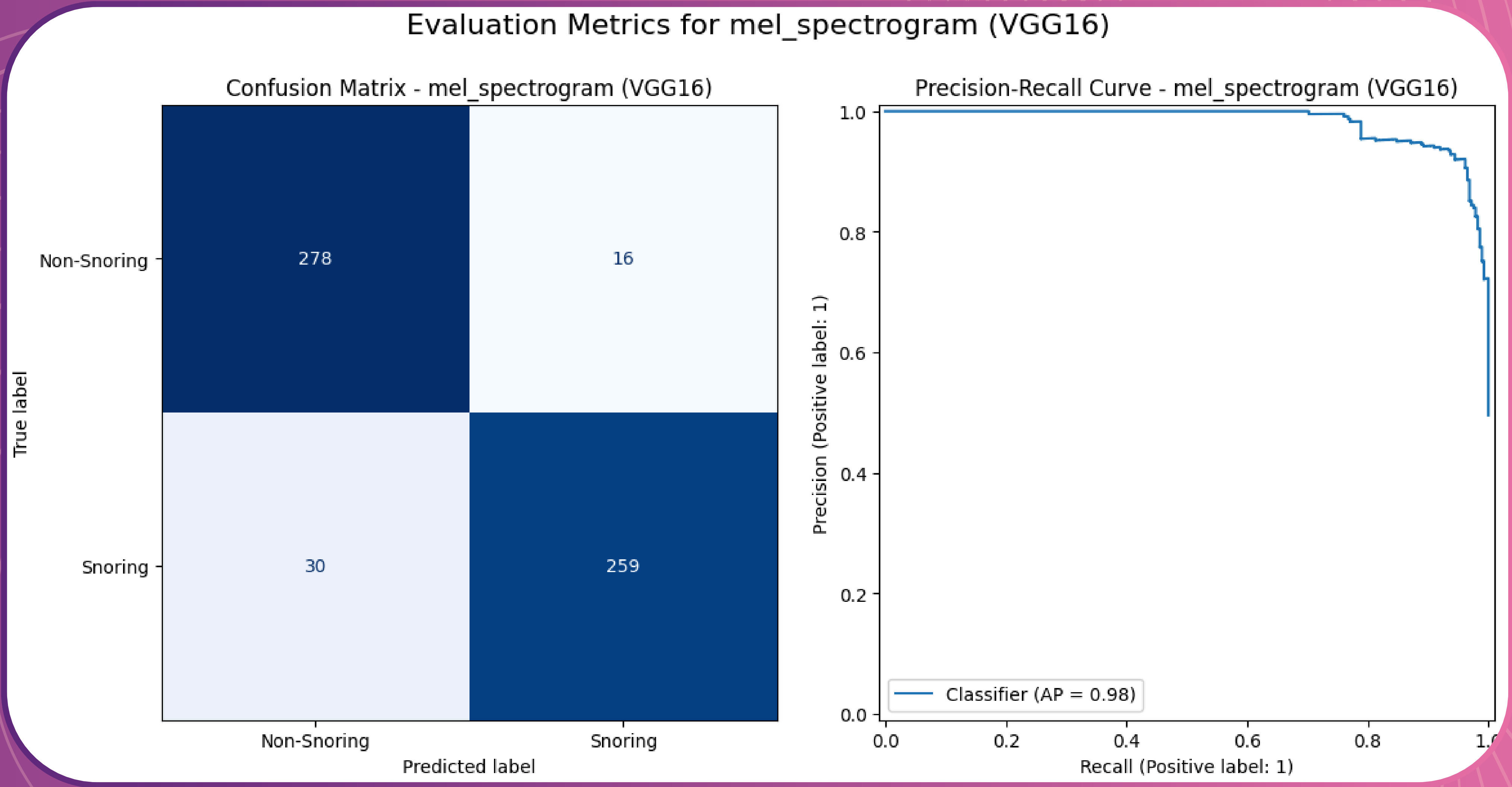
SGD Classifier

- Architecture:
  - Linear model trained with stochastic gradient descent.
  - Loss Function: 'log_loss' for binary classification.
- Why SGD?
  - A simple, computationally efficient baseline.
  - Works well for large datasets with simple decision boundaries.
- Advantages:
  - Fast training and less prone to overfitting.
  - Effective baseline for comparing with deep learning models.
- Performance:
  - Accuracy: 0.68.
  - Precision & Recall:
    - Non-Snoring: Precision = 0.74, Recall = 0.52.
    - Snoring: Precision = 0.65, Recall = 0.83.
  - Weaker in detecting non-snoring instances (lower recall).





Results
VGG16



Results
SGD

Conclusion

- VGG16 significantly outperforms the SGD Classifier in snoring detection tasks, especially due to its ability to handle the complexity of Mel-spectrograms and learn hierarchical patterns from the data.
- SGD Classifier, while a useful baseline, lacks the performance needed for this problem, especially with imbalanced class distribution in snoring vs. non-snoring samples.
- VGG16 with Transfer Learning is the preferred model for the task, offering robust, high-performance classification suited for practical applications like sleep monitoring systems.