Riyaz Readiness Predictor: A Machine Learning Tool for Hindustani Classical Vocal Performance Assessment

Name: Mileesha Vannan
Registration Number: 2024SEPVUGP0019
Course: Hackathon 3
University: Vijaybhoomi University
Date: 9th February 2026

Abstract

Hindustani Classical vocal training places strong emphasis on riyaz (daily practice), swara shuddhata (pitch accuracy), laya (rhythmic stability), and improvisational development. In music schools, decisions about whether a student is ready for jury exams, baithaks, or recitals are often based on subjective assessment by instructors. While this traditional approach is valuable, it can benefit from additional data-driven support.

This project proposes the development of a machine learning–based decision support system called the *Riyaz Readiness Predictor*. The system predicts whether a student is "Ready" or "Needs More Riyaz" for a performance evaluation based on structured practice and academic indicators such as riyaz hours, taal discipline, swara accuracy, repertoire size, and past jury performance. The solution includes a trained predictive model, a dashboard for analysis and prediction, and a clear plan for deployment and maintenance within a university music school setting.

1. Problem Statement

In Hindustani Classical vocal education, consistent practice and technical discipline are essential for achieving performance readiness. However, instructors must often rely solely on qualitative judgment to decide whether a student is prepared for important evaluations such as jury exams, baithaks, or recitals. This process can vary across teachers and lacks a standardized analytical support system.

The problem addressed in this project is:

How can a university music school use structured student data to predict whether a Hindustani Classical vocal student is ready for performance evaluation?

The objective is to design a machine learning system that assists faculty by predicting performance readiness based on measurable training indicators. The system will not replace the guru–shishya tradition, but rather support instructors by highlighting students who may need additional riyaz or focused guidance before appearing for evaluations.

2. Data Description

The project uses **synthetically generated data** that reflects realistic student training behavior in a Hindustani Classical vocal program. Large datasets are not uploaded to the repository. Instead, a data generator script is included that produces training data based on defined rules and distributions.

Input Features

Feature Name	Description
`daily_riyaz_hours`	Average hours of daily practice
`taal_discipline_score`	Rhythm consistency (0–10)
`swara_accuracy`	Pitch accuracy based on teacher rubric (0–10)
`bandish_repertoire`	Number of bandishes known
`alaap_consistency`	Stability and flow in alaap (0–10)
`layakari_comfort`	Comfort with tempo variations (0–10)
`past_jury_marks`	Previous jury exam score (0–100)

Target Variable

- `ready`

    ○ 1 → Ready for Performance

    ○ 0 → Needs More Riyaz

The dataset is generated using controlled randomness so that higher riyaz hours, better swara accuracy, stronger taal discipline, and higher past jury marks increase the probability of being labeled "Ready".

3. Models and Evaluation

The project trains and compares **three traditional machine learning models**:

1. **Logistic Regression** – Baseline, interpretable model

2. **Random Forest Classifier** – Handles non-linear relationships

3. **Gradient Boosting / XGBoost** – High-performance ensemble model

Training & Testing

- Stratified train/test split

- Cross-validation (k=5)

- Hyperparameter tuning using GridSearchCV

Evaluation Metrics

- Accuracy

- Precision

- Recall (important for identifying students who are not ready)

- F1-Score

- Confusion Matrix

The best-performing model is selected for deployment in the dashboard.

4. Dashboard and Prediction Interface

A **Streamlit-based dashboard** will be developed to support instructors and administrators.

Dashboard Features

- Dataset overview and summary

- Feature distribution plots

- Model comparison table

- Feature importance visualization

Prediction Panel

Faculty can enter:

- Riyaz hours

- Taal score

- Swara accuracy

- Bandish count

- Alaap score

- Layakari comfort

- Past jury marks

The system returns:

- Readiness Probability

- Predicted Label: **Ready / Needs More Riyaz**

- Improvement Suggestions (e.g., increase alaap consistency, focus on taal practice)

5. Expected Output

- A trained machine learning model for performance readiness prediction

- A reproducible data generator script

- An interactive dashboard for analysis and predictions

- Clear evaluation metrics and reports

- A documented GitHub repository with full project history

6. Project Timeline

Week	Task
Week 1	Finalize problem & proposal
Week 2	Data generation + EDA
Week 3	Model training & evaluation
Week 4	Dashboard development
Week 5	Testing & documentation
Week 6	Final submission & deployment prep

7. Maintenance & Updation Plan

- Retraining every **3 months** using updated student data

- Performance monitoring via recall and accuracy metrics

- Data drift checks each semester

- Dashboard updates annually based on faculty feedback

- Owner: Music Department Data Assistant / Project Lead

8. Git Repository

A public GitHub repository will be created containing:

- Data generator script

- Model training code

- Evaluation notebooks

- Dashboard application

- Documentation

Repository URL:

[https://github.com/yourusername/riyaz-readiness-predictor]