midihum: A Machine Learning-Based MIDI Humanizing Tool

Erich Grunewald
Independent Researcher
Berlin, Germany
erichgrunewald@gmail.com

***Abstract*—We present midihum, a machine learning system for MIDI piano performance humanization through velocity prediction. The model employs XGBoost gradient boosted trees trained on 2.6K piano competition performances from the International Piano-e-Competition dataset. Feature engineering produces approximately 400 features derived from note characteristics, including logarithmic duration derivatives, pitch and octave information, time-based pitch oscillators, and temporal spacing patterns between notes. The framework modifies only velocity values while preserving original note timing and pitch information.**

I. Introduction

midihum addresses the challenge of adding humanization to mechanically generated MIDI piano performances. The system targets velocity modification, transforming uniform dynamics into varied patterns that reflect natural musical performance characteristics. The tool is applicable to both user-composed music and historical piano roll recordings, which capture precise timing and pitch information but often lack velocity data due to the mechanical nature of the perforation process.

II. Methodology

The system uses XGBoost gradient boosted trees trained on piano competition performance data. Feature engineering generates approximately 400 features from raw MIDI data.

Feature categories include: (1) Logarithmic transformations of note durations using $\log(|duration| + 1)$; (2) Pitch features including raw MIDI pitch values, pitch class (modulo 12), octave, and intervals between consecutive notes; (3) Moving averages with window sizes of 15, 30, and 75 time units, and their differences from current values; (4) Time intervals between notes, also logarithmically transformed.

The system adapts technical indicators from financial analysis, including Ichimoku Cloud components (Tenkan-sen, Kijun-sen, Senkou spans) and exponential weighted moving averages similar to MACD calculations. These indicators were included to capture non-linear patterns and temporal dependencies in musical sequences. Additional features include chord size, song-level statistics (mean, standard deviation, min, max), and previous note values. The model outputs velocity values in the standard MIDI range (1-127) while keeping original timing and pitch unchanged.

III. Datasets

Training utilized the International Piano-e-Competition dataset containing 2.6K piano performances in MIDI format. The dataset spans European art music tradition from Baroque to Romantic periods, providing 2.6K performances across various piano works. Data preprocessing included MIDI parsing and cleaning steps.

IV. RenCon 2025 Submission

For this competition submission, the trained midihum model was applied directly to the provided MusicXML scores after conversion to MIDI format. No post-processing, quantization, manual correction, or human intervention was applied to the outputs. The system operates deterministically from the input MIDI data without controllability parameters or human-involved design choices. Three pieces were submitted: Rachmaninoff Op. 21 No. 7, Beethoven WoO 80 variations, and Handel HWV 483 (free choice). All outputs preserve original note timing with only velocity modifications applied.

V. Source Code

Source code is available at https://github.com/erwald/midihum.