EmotioNotes Dataset: Decoding Emotions in Classical Music through Concert Program Notes

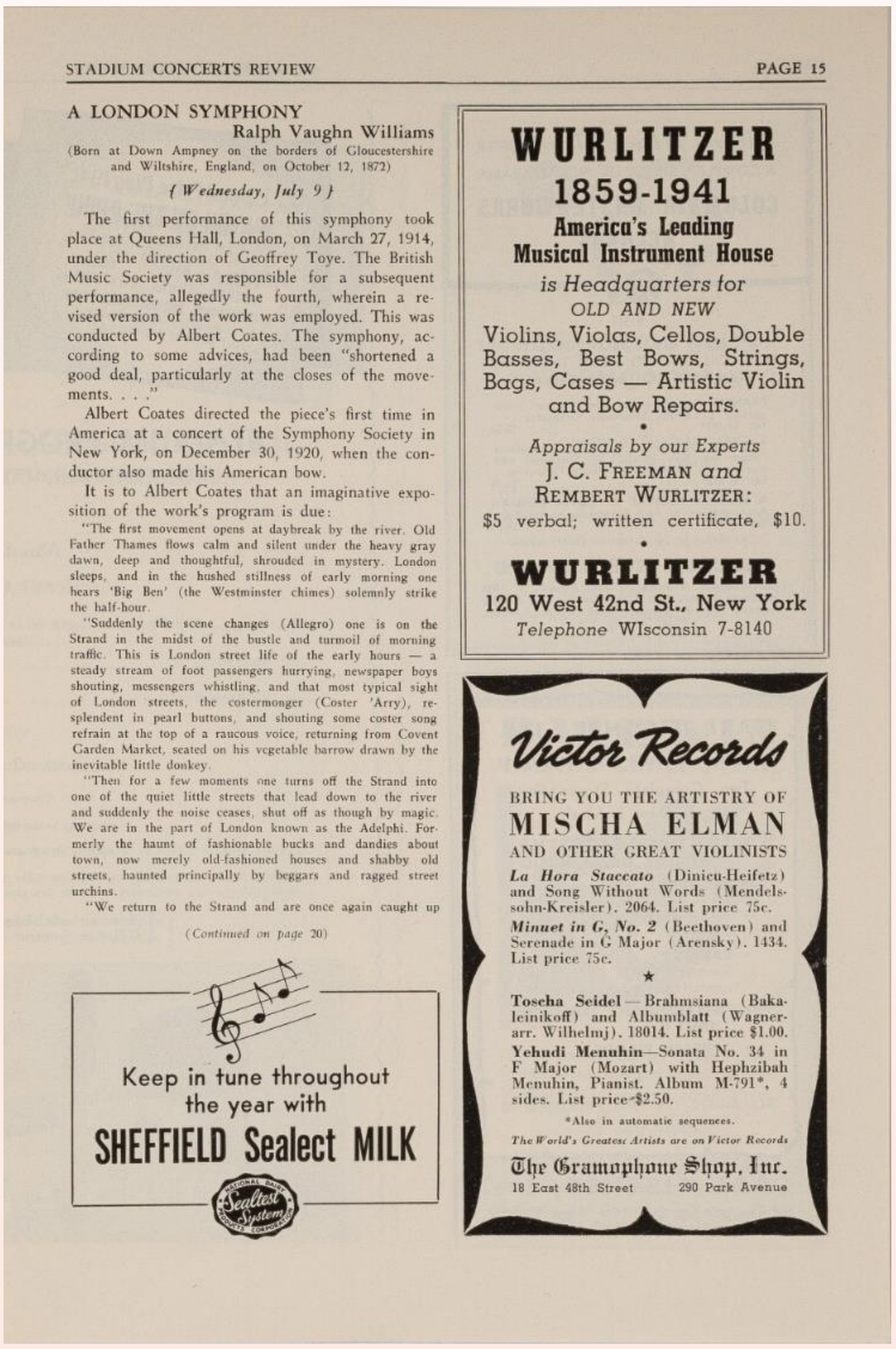
Pratik Khanal, Patrick J. Donnelly
Oregon State University, Corvallis, Oregon, USA

Abstract

The lack of large, public datasets of music and emotion is a major challenge in emotion recognition from music. We present EmotioNotes, a novel dataset designed to enable large-scale emotion recognition in classical music. We collect concert program notes from the New York Philharmonic Society's archive as unorganized text generated using Optical Character Recognition. We use a large language model, `Llama 3.1` with task-specific prompting to extract program notes. We assign each program note with emotion scores using a dataset of emotion values of English words. By extracting and annotating over 31,000 program notes, we estimate emotional content through valence, arousal, and dominance (VAD) scores. We further train a transformer-based model, `Longformer` to predict these scores directly from text. Our results demonstrate strong correlations between predicted and calculated scores, offering a scalable, language-based approach to music emotion recognition and supporting future work in music recommendation, therapy, and musicology.

Introduction

- Music has ability to profoundly shapes our emotional experiences.
- Datasets with emotional labeling for classical music is notably scarce.
- This lack of data slows progress in Music Emotion Recognition (MER).
- Descriptive program notes for classical pieces offer a unique opportunity.
- By using these notes as emotional proxies, we can scale automatic annotation without the heavy burden of human annotation studies.
- **EmotioNotes**: Our novel solution leverages program notes to facilitate robust, large-scale emotional dataset for use in emotion recognition tasks.



STADIUM CONCERTS REVIEW PAGE 15

A LONDON SYMPHONY
Ralph Vaughn Williams
(Born at Down Ampney on the borders of Gloucestershire and Wiltshire, England, on October 12, 1872)
{ Wednesday, July 9 }

The first performance of this symphony took place at Queens Hall, London, on March 27, 1914, under the direction of Geoffrey Toye. The British Music Society was responsible for a subsequent performance, allegedly the fourth, wherein a revised version of the work was employed. This was conducted by Albert Coates. The symphony, according to some advices, had been "shortened a good deal, particularly at the closes of the movements. . . ."

Albert Coates directed the piece's first time in America at a concert of the Symphony Society in New York, on December 30, 1920, when the conductor also made his American bow.

It is to Albert Coates that an imaginative exposition of the work's program is due:

"The first movement opens at daybreak by the river. Old Father Thames flows calm and silent under the heavy gray dawn, deep and thoughtful, shrouded in mystery. London sleeps, and in the hushed stillness of early morning one hears 'Big Ben' (the Westminster chimes) solemnly strike the half-hour.

"Suddenly the scene changes (Allegro) one is on the Strand in the midst of the bustle and turmoil of morning traffic. This is London street life of the early hours — a steady stream of foot passengers hurrying, newspaper boys shouting, messengers whistling, and that most typical sight of London streets, the costermonger (Coster 'Arry), resplendent in pearl buttons, and shouting some coster song refrain at the top of a raucous voice, returning from Covent Garden Market, seated on his vegetable barrow drawn by the inevitable little donkey.

"Then for a few moments one turns off the Strand into one of the quiet little streets that lead down to the river and suddenly the noise ceases, shut off as though by magic. We are in the part of London known as the Adelphi. Formerly the haunt of fashionable bucks and dandies about town, now merely old-fashioned houses and shabby old streets, haunted principally by beggars and ragged street urchins.

"We return to the Strand and are once again caught up

(Continued on page 20)

WURLITZER
1859-1941
America's Leading Musical Instrument House
is Headquarters for
OLD AND NEW
Violins, Violas, Cellos, Double Basses, Best Bows, Strings, Bags, Cases — Artistic Violin and Bow Repairs.
Appraisals by our Experts
J. C. FREEMAN and REMBERT WURLITZER:
$5 verbal; written certificate, $10.
WURLITZER
120 West 42nd St., New York
Telephone WIsconsin 7-8140

Victor Records
BRING YOU THE ARTISTRY OF
MISCHA ELMAN
AND OTHER GREAT VIOLINISTS
La Hora Staccato (Dinicu-Heifetz) and Song Without Words (Mendelssohn-Kreisler). 2064. List price 75c.
Minuet in G, No. 2 (Beethoven) and Serenade in G Major (Arensky). 1434. List price 75c.
★
Toscha Seidel — Brahmsiana (Bakaleinikoff) and Albumblatt (Wagner-arr. Wilhelmj). 18014. List price $1.00.
Yehudi Menuhin—Sonata No. 34 in F Major (Mozart) with Hephzibah Menuhin, Pianist. Album M-791*, 4 sides. List price $2.50.
*Also in automatic sequences.
The World's Greatest Artists are on Victor Records
The Gramophone Shop, Inc.
18 East 48th Street 290 Park Avenue

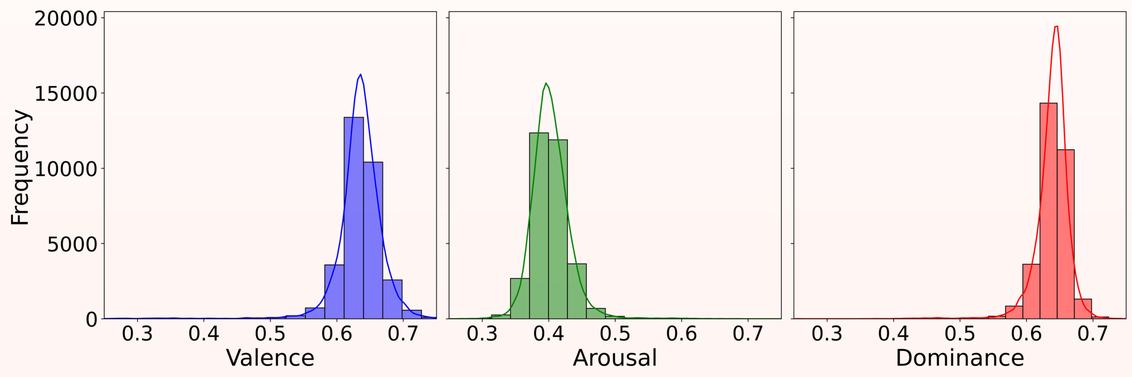
Keep in tune throughout the year with
SHEFFIELD Sealect MILK
Sealtest System

Methods

- **Data Collection & OCR**: Over 14,000 scanned New York Philharmonic program booklets processed with OCR to extract raw text.
- **Program Note Extraction**: LLaMA 3.1 (70B parameters, 128K context window) to filter ads and irrelevant text, capturing clean program notes.
- **Emotion Annotation**: Applied Warriner et al. VAD word norms for valence (pleasantness), arousal (intensity), and dominance (control).
- **Score Computation**: Lemmatized words, matched them to the VAD lexicon, then averaged them to obtain normalized emotion scores.
- **Model Training**: Trained `Longformer` for regression using tokenized notes to predict 3-dim emotion scores. 70/15/15 split with MSE loss.

Dataset

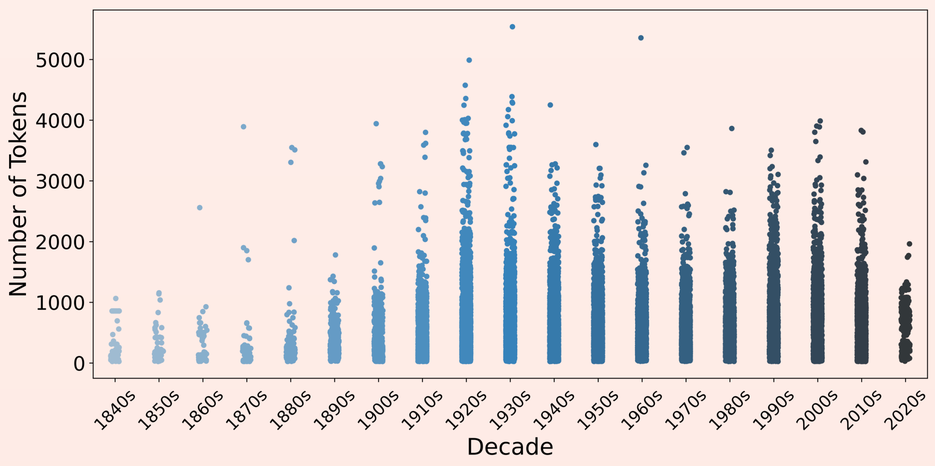
- Our process retained 31,966 program notes after the LLM filtering, from 9,569 concerts, and 1,531 different composers.
- Program notes range from concise summaries to multi-pages. essays.
- We used BLEU score (threshold: 0.006) to remove incorrect or "hallucinated" entries, removing 2,614 incorrectly produced notes.
- Data source spans 180 years, featuring hundreds of composers.
- Valence 0.634±0.039, Arousal 0.403±0.030, Dominance 0.639±0.027.



Results

- We create a novel dataset of 31,966 program notes, each annotated with estimations of valence, arousal, and dominance (VAD).
- We found the highest correlation for valence (pleasantness), moderate for arousal (intensity), and the lowest for dominance (control).

	r	ρ	R^2
Valence	0.877	0.856	0.751
Arousal	0.823	0.841	0.630
Dominance	0.766	0.754	0.553



Future Work

- Use retrieval-augmented generation with `Llama 3.1` followed by a reranking strategy and then compare results to those reported here.
- Collect copyright-free recordings of classical music in our dataset.
- Train a model on audio data using current emotion scores as proxies.
- Enable personalized music recommendation based on listeners' mood.