Capstone 2 Project Proposal

Predicting Chords and Identifying Common Harmony in Jazz Standards

Jeremy Ferrara

Problem Identification

The goal of this project is to analyze common and uncommon harmonic choices in jazz standards, and to build a model that can recommend or predict the next chord in an unfinished song based on its key, rhythm, and era.

This analysis can help musicians, educators, and jazz enthusiasts better understand jazz harmony and generate realistic chord progressions.

Context

Jazz musicians and composers often draw on patterns in harmonic progressions to create or interpret music. However, while many chord progressions (like ii–V–I) follow predictable harmonic patterns, jazz standards also contain surprising, non-diatonic chord changes that give the genre its distinct sound.

Understanding these patterns on a large scale can:

- Aid music educators in teaching harmony more effectively.
- Provide musicians and jazz students insight into how harmonic norms are used and broken in jazz.

This project builds on my own experience as a touring musician, guitarist, and data scientist, aiming to combine musical intuition with computational methods.

Criteria for Success

The project will be successful if it can:

1. Produce a model that can predict the correct next chord with above-random accuracy in a known jazz standard given the key, era, rhythm, and first few chords.

2. Quantify “common” and “uncommon” chord movements.

Deliverables include a GitHub repo containing the work/code completed for each step of the project, including a slide deck and project report.

Scope of Solution Space

This project will focus on identifying common and uncommon chord progressions in jazz and building a predictive model for chords in an unfinished song. Specifically, only popular jazz standards dating back to the early 1900s will be used.

Constraints

- **Musical ambiguity**: Some chords may have multiple valid interpretations (e.g., enharmonic equivalents).
- **Text data**: The dataset consists only of the chords, rhythm, composer, era, and time signature for each song. Audio, sub-genre, and instrumental arrangement are not included.

Stakeholders

- My Springboard mentor Eric Callahan
- Any educator, musician, or jazz enthusiast interested in this analysis or model

---

Data Sources

Source: Jazz Standards Dataset – Mike Oliphant GitHub Repository

Description: This dataset contains chord progressions for over one thousand well-known jazz songs in JSON format. Each song is stored as a sequence of chords with metadata such as title, key, and composer.