CS109b Poster Guidelines

Why Do A Poster? Poster sessions are widely used at conferences in academia and industry to facilitate discussion. A poster is typically complementary to a talk about the topic. While a talk is typically 15-20 minutes with a short Q&A, a poster session enables longer conversations to take place with all the relevant material at hand. Poster sessions typically last a few (1-2) hours. Being able to make and present a good poster is a skill that you will likely need in many future jobs.

SEAS Design Fair. At the end of the course, you will present your work at the SEAS Design Fair (1) to your peers and course staff, (2) to students from other SEAS courses, and (3) to participants across the wider Harvard community. Doing a great poster for your work is therefore more than an academic requirement: it is a real opportunity to foster connections, broaden your network, and even get opportunities for kick-starting your post-Harvard career.

Contents. The poster presentation should cover the following questions: What is the problem your project tries to solve? What is a naïve baseline solution to this problem, and why does it not perform as well as needed? What is your approach? Why and to what extent does your approach improve on the baseline? Is there a broader take-away message from your work?

Not a Report. The goal for the poster is to serve as a visual aide to help illustrating the main points about your project. It should complement rather than repeat what you will anyway be saying, and it is meant to visually appeal to people to come and speak to you. It is *not* meant to be a self-contained description of your work, though it may include charts and figures that you did not have time to cover in the presentation.

Readability from Afar. Everything should be readable from **10 feet away** – while you interact with participants, others will be waiting for the turn to speak to you and standing farther away. You want them to be able to get a basic idea before they start talking to you so that the ensuing discussion as more interesting.

Short, Engaging Title – The title should be short, engaging, and refer to the problem that you addressed rather than to the solution. For instance, using the title “Convolutional Neural Networks for Sentiment Analysis In Movie Ads” is an arduous for a reader to parse relative to “Emotion in Movie Ads”.

Word Count. The goal of the poster session is for personal engagements, and so most readers will not have the mental energy to be reading paragraphs. The poster should engage them as much as possible with key concepts and short sentences. The word count for the poster may be as low as 200 words and ideally not above 700 words.

Clear Flow. There should be clear flow from the problem definition through to the solution and conclusion. It should be clear what the starting point is, and what precedes what.

Acknowledgments, Names, Affiliation. You should include your names, your affiliation to SEAS and/or Harvard, and an acknowledgement to any other person who helped during your project.

Be Concise and Minimalistic – The viewer will only devote a finite amount of attention to your project, so make sure they absorb as much information as possible during the time and attention that they have allocated. For instance, titles such as “introduction” force the viewer to read words that do not convey any new information about your project, and so consider replacing them with a topic, question, or statement. Generally, try to eliminate as much cognitive overhead on the part of the viewer as possible.

Engage rather than Impose. During the poster session, the goal is for you to engage with participants rather than to give a monologue about your project. Stop if they have questions, make sure they follow, make sure they don’t get bored, and feel free to express interest in them or to let the conversation diverge to other topics. Be conscious that some participants will only want a quick overview whereas others will want to go deeper. You should be able to cater well both audience types.

Template. A template that you can use as a starting point for your poster can be found on canvas as part of the Milestone #4 assignment. It's only meant as a starting point. We would like you to form your own style.

Example from Last Year.

Creating the Perfect Playlist: Content-Based Generation of Spotify Playlists

Sonu Mehta, Laura Ware, Omar Abboud

GOAL

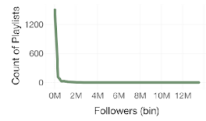
Explore methods for predicting the success of a Spotify-curated playlist, based only on data about the songs that comprise that playlist, and use these models to develop novel processes for curating successful Spotify playlists.

Features:

- Predict the success (# of followers) of a given playlist
- Identify songs that are acoustically similar to a seed song
- Generate a playlist that optimizes the grouping and sequence of tracks

DATASETS

Only a small number of playlists have over 100k followers



Spotify API

Includes data about each Spotify-curated playlist (e.g. total tracks, sequence, no. of followers) and individual tracks (e.g. audio features, popularity)

Raw Audio

30-second samples available from the Spotify API for ~25% of the 20k tracks in our dataset.

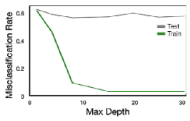
MODELS

Random Forest to Predict Number of Playlist Followers

Inputs: Acoustic features (danceability, loudness, energy, liveness, etc.), duration, popularity of songs, track order

Output: No. of playlist followers, divided into 5 bins

Tuning Random Forest Max Depth

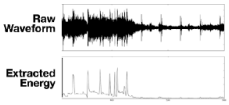


Using Raw Audio to Predict Track Popularity

Inputs: MFCCs, Chroma coefficients,Energy at each timestamp

Output: Popularity divided into 6 bins

Performance: Random forest classifier gave 75% accuracy, F-1 score of 1 for the top class



Methods for Determining Acoustic Similarity

Manhattan distance with features from Spotify API **+** *Earth Mover's Distance* based on KL Divergence and raw audio features, with vantage points to optimize the search process.

Playlist Generation Algorithm

1. Input a seed track
2. Identify candidate tracks from either: A) top tracks from related artists, B) similar tracks from related artists or C) acoustically similar songs from a vantage point database
3. Perform simulated annealing, starting with a random sample of 30 tracks
4. Return the playlist that optimizes predicted number of followers.

RESULTS

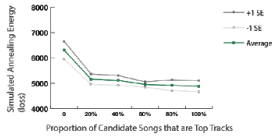
Classification Model Comparison

The mean popularity of a playlist's tracks alone is not sufficient to predict overall playlist followers; this metric is much stronger when combined with acoustic features.

Predictor Set	Classification Accuracy
Mean Popularity Only	0.35
Spotify Acoustic Features Only	0.66
Mean Popularity + Acoustic Features	0.78

Weighting of Top Tracks vs. Performance:

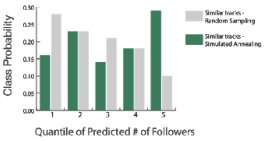
Using a higher percentage of top tracks improves performance but decreases the acoustic similarity of the resulting playlist.



CONCLUSIONS

- Combinations of songs can be reasonably optimized for popularity using simulated annealing supported by raw audio-based similarity metrics and a Random Forest predictive model of popularity.

Simulated Annealing increases optimization performance significantly



Future Work

- Take steps to mitigate the overfitting that affected our random forest classifier, which over-predicted unpopular playlists in the testing set.
- Include more features as predictors of popularity to ensure that the model is not capturing information such as release date or other non-intrinsic characteristics.

REFERENCES

Berenzweig, Adam, Beth Logan, Daniel P.W. Ellis, & Brian Whitman, A Large-Scale Evaluation of Acoustic and Subjective Music Similarity Measures. Proceedings of the ISMIR International Conference on Music Information Retrieval (Baltimore, MD), 2003, pp. 99–105.

Logan, B., "A Content-Based Music Similarity Function," (Report CRL 2001/02) Compaq Computer Corporation Cambridge Research Laboratory, Technical Report Series (Jun. 2001).