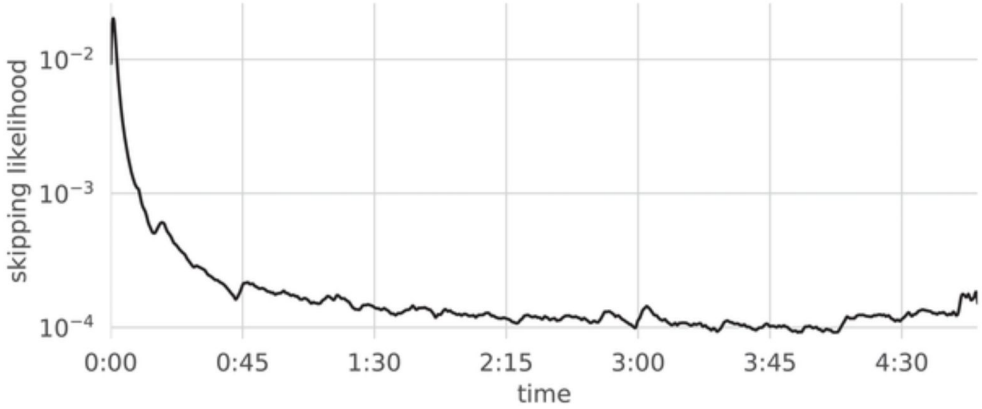
A Neural Network Approach To Predicting Song Skip Behavior

Sophie Johnson and Aditi Patil

Implications of Skip Song Behavior

Instances of users skipping songs raise questions about how this “implicit feedback signal” reflects user satisfaction and future engagement with the streaming service (Meggetto et al., 2021).



Montecchio et al., 2020.

Spotify Sequential Skip 2019 Challenge

- Joint effort by Spotify, CrowdAI, and 12th ACM International Conference on Web Search and Data Mining in Melbourne Australia
- Challenge focused on session-based sequential skip prediction
- Dataset of distinct million unique user listening sessions

How can one predict user song skip behavior with the most accuracy?

We hypothesize that metadata regarding user behavior are integral in understanding song skip behavior, and believe Long-Short Term Memory (LSTM) and Gradient Boosting Tree models will achieve high accuracy in doing so.

Related Work

- Questions of which features to include in model and what kind of model?
- Wide variety of approaches: metric learning, sequence learning models, boosting trees, multi-RNN
- 2nd place submission in competition used Multi-RNN approach; included all data features in model (Hansen et. al., 2018).
- Amplified accuracy when incorporating user-log features into model that before primarily relied on acoustic features, suggesting user-logs “contain useful information for sequential skip predictions “ (Chang et al., 2019).

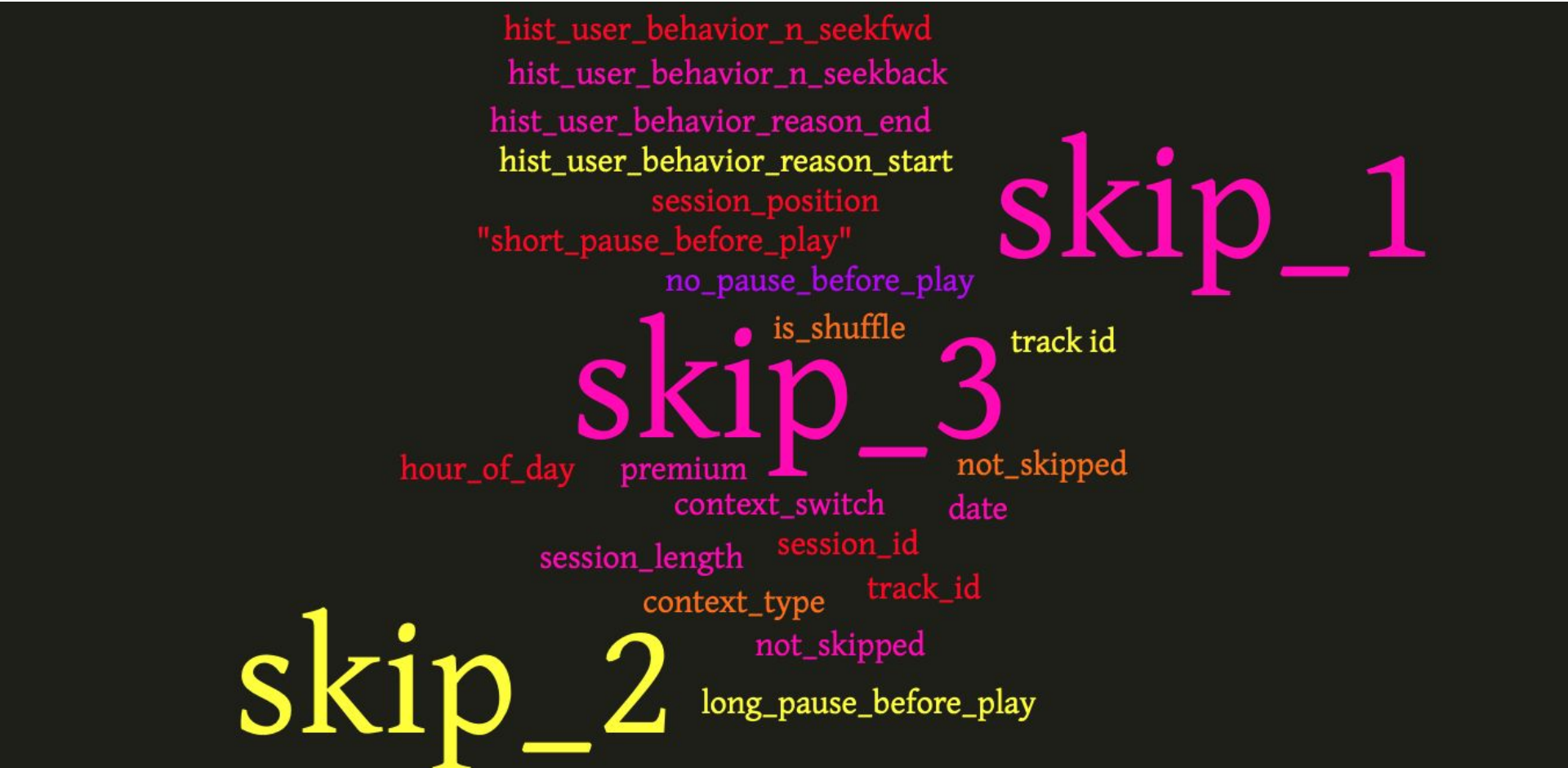
The Dataset

130 million unique Spotify listening sessions, defined by 21 features

50704 distinct music tracks

21 listening session variables

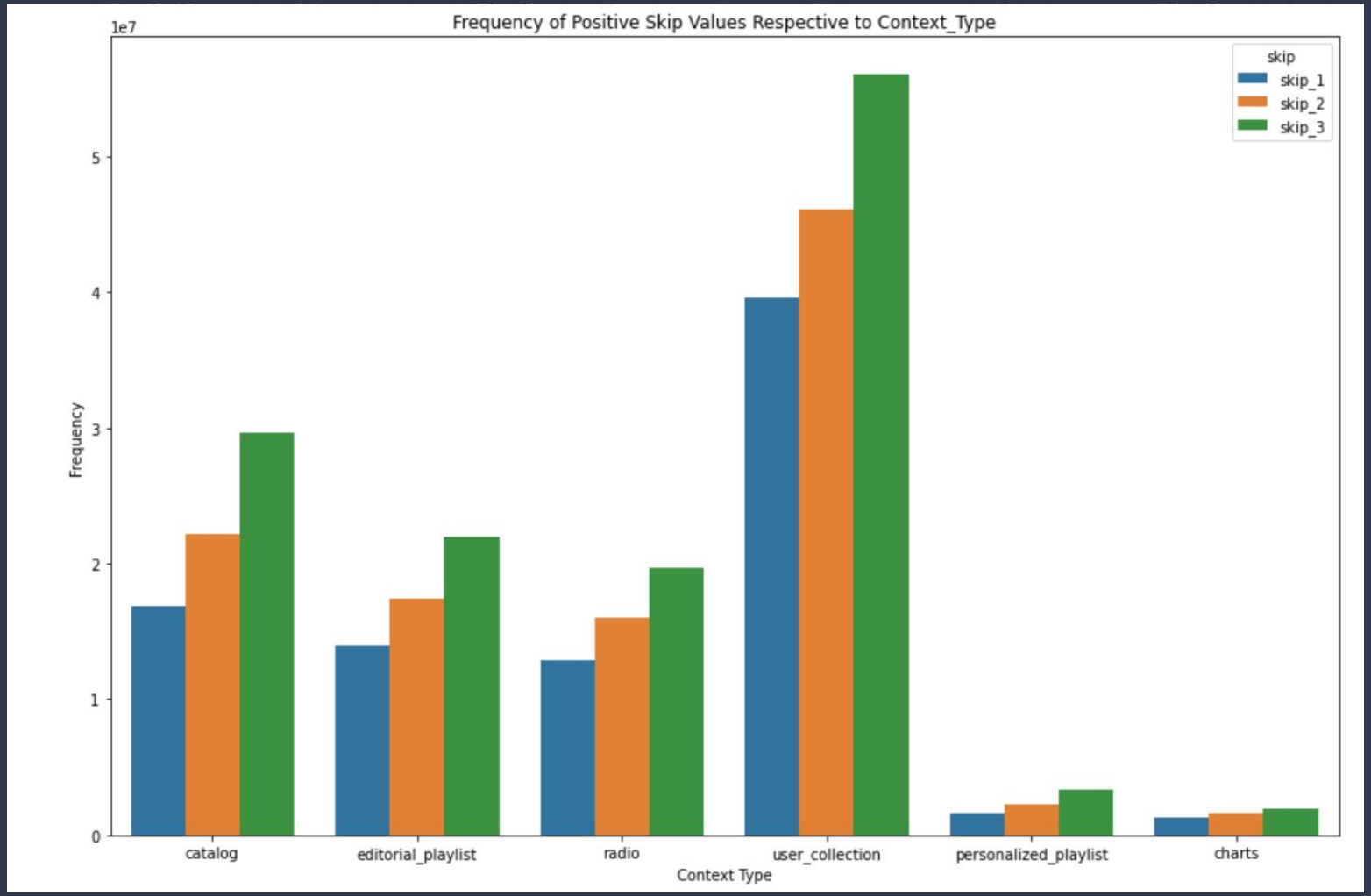
Listening Session Dataset Variables



Correlations: Mean Chi-Square Statistical Values

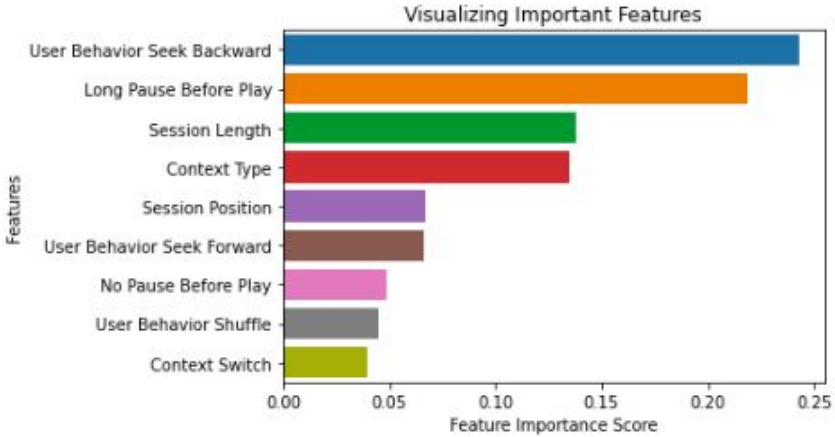
	context_type	session_length	hist_user_behavior_n_seekfwd	hist_user_behavior_n_seekback	hist_user_behavior_is_shuffle
skip_1	39012.925	32016.846	11051.312	39950.939	11177.027
skip_2	23871.195	25712.025	2336.529	15218.364	9186.128
skip_3	13009.489	17462.549	27514.597	5344.180	4703.993

	context_switch	no_pause_before_play	premium	session_position	long_pause_before_play	hour_of_day
skip_1	24523.615	95281.331	429.888	25132.850	68313.739	3219.610
skip_2	14158.195	28699.031	168.557	18547.076	13060.201	3349.045
skip_3	4659.594	18129.427	382.767	2261.416	26772.830	4107.117


Frequency of Positive Skip Values Respective to Context_Type
1e7
skip
skip_1
skip_2
skip_3
Frequency
0
1
2
3
4
5
catalog
editorial_playlist
radio
user_collection
personalized_playlist
charts
Context Type

Baseline Model: Random Forest Classifier

- Construct a decision tree for every sample and receive a resulting prediction, conduct a vote for each predicted result, and for the final prediction, select the prediction result with the most votes.
- Random Forest Classifier performed at an accuracy level of 64%.
- Visualized RFC model's feature importance to determine helpful insights for two other models



Gradient Boosting Trees

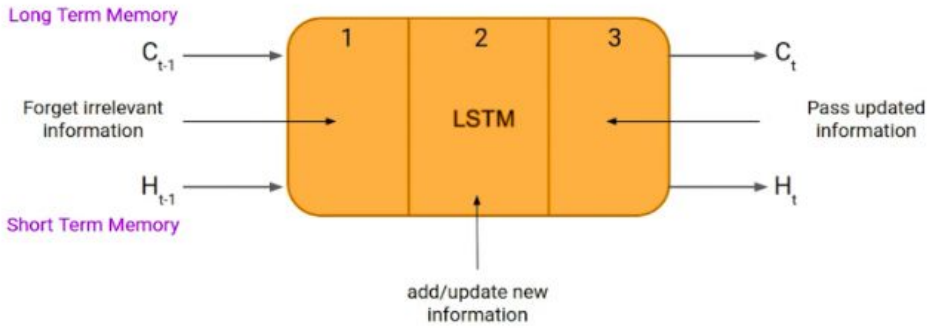
- Gradient Boosting Trees is similar to Random Forest, except that each successive predictor tries to improve on its predecessor by reducing the errors
- Learning rate = 0.1, maximum depth = 5, and logarithmic loss function
- Accuracy level of 66%, slightly more than RFC

	precision	recall	f1-score
0 (skipped)	0.54	0.11	0.18
1 (not skipped)	0.66	0.95	0.78

Table 1: Precision/Recall/F1 score of class prediction for the Gradient Boosting Tree classifier.

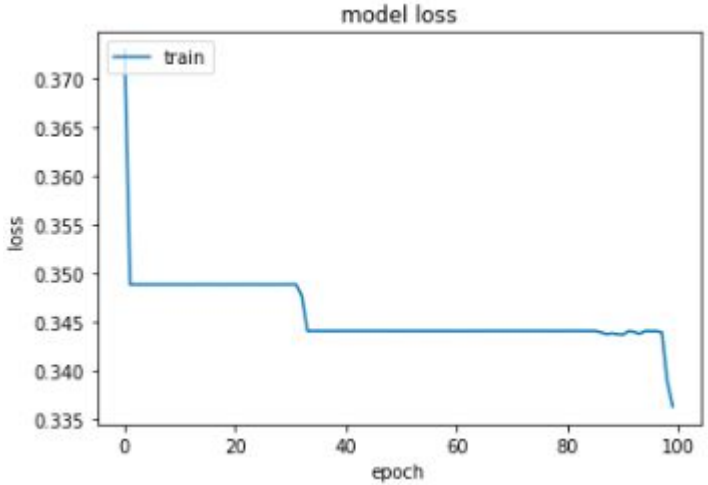
Long Short Term Memory

- A type of recurrent neural network that is beneficial for learning order importance in sequential prediction.
- Important for this project because listening sessions are given in order, and skip behavior can be dependent on previous listening patterns.
- Architecture: 1 Embedding input layer, 4 dense hidden layers with ReLU activation function, 1 output dense layer.



Final Results / Research Takeaways

- LSTM performs with accuracy level of 64%, and has a significantly longer training time.
- Decision Tree models are robust to smaller datasets, while LSTM is not necessarily.
- User metadata regarding their song listening behaviors are important features when predicting song skips.
- Incorporating acoustic data and training on full dataset would boost performance further.



References

References

Brian Brost, Rishabh Mehrotra, and Tristan Jehan. 2019. Spotify sequential skip prediction challenge. In *Proceedings of the 28th ACM International Conference on Information and Knowledge Management (CIKM '19)*. Association for Computing Machinery.

Sungkyun Chang, Seungjin Lee, and Kyogu Lee. 2019. Sequential skip prediction with few-shot in streamed music contents. arXiv:1901.08203.

Andres Ferraro, Dmitry Bogdanov, and Xavier Serra. 2019. Skip prediction using boosting trees based on acoustic features of tracks in session. arXiv:1903.11833.

Christian Hansen, Casper Hansen, Stephen Alstrup, Jakob Grue Simonsen, and Christina Lioma. 2019. Modeling sequential music track skips using a multi-rnn approach. arXiv:1903.08408.

Francesco Meggetto, Crawford Revie, John Levine, and Yashar Moshfeghi. 2021. On skipping behaviour types in music streaming sessions. In *Proceedings of the 30th ACM International Conference on Information Knowledge Management (CIKM '21)*.

Spotify. Listening is everything. *Spotify*.

Ping Xuan, Chang Sun, Tiangang Zhang, Yilin Ye, Tonghui Shen, and Yihua Dong. 2019. Gradient boosting decision tree-based method for predicting interactions between target genes and drugs. *Frontiers in Genetics*, 10.