Kaily Mejia

THE QUEST FOR SUPERIOR MUSIC RECOMMENDATIONS

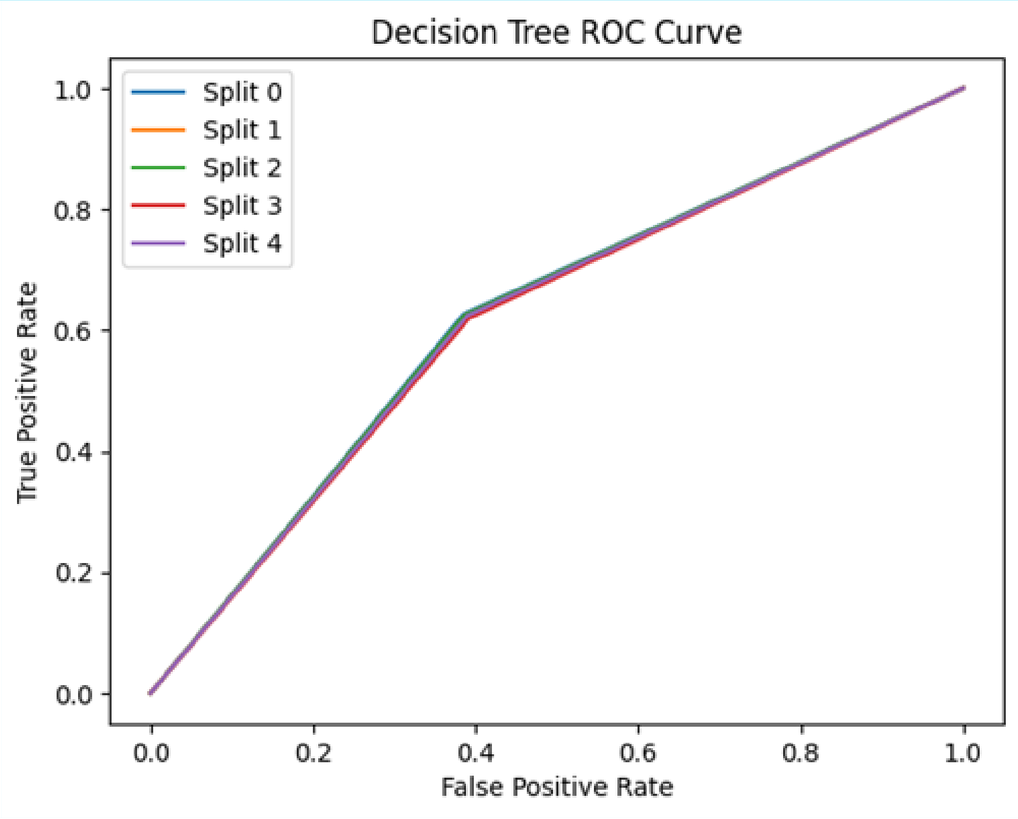
INTRODUCTION

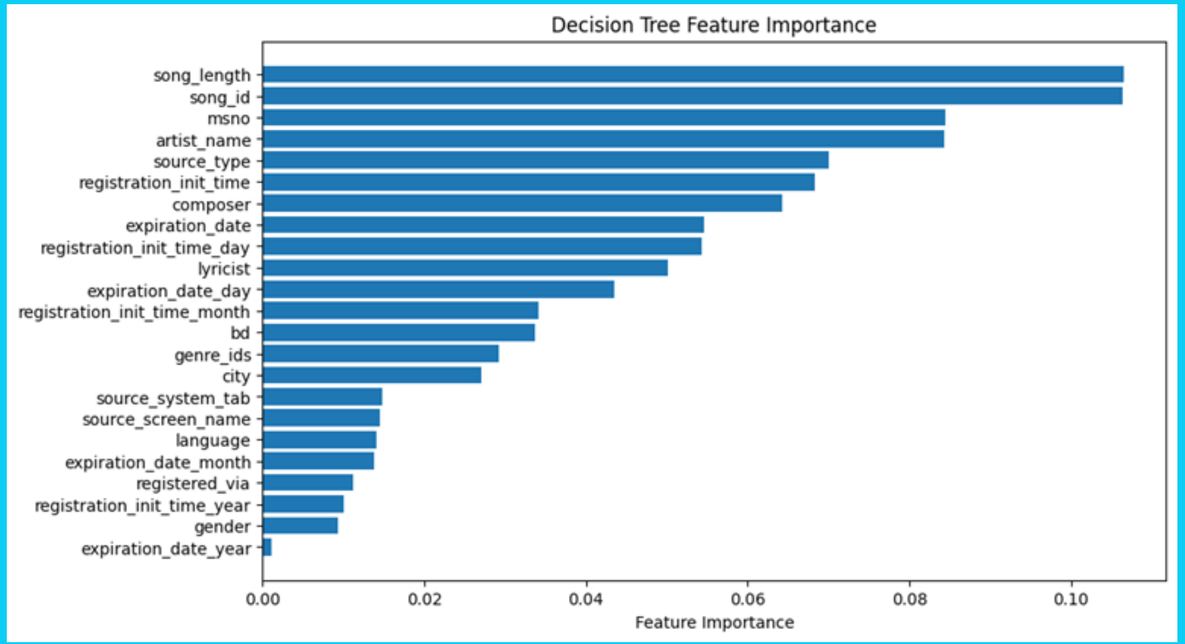
- Data provided by KKBox, Asia's leading streaming service
- Challenging you to build a better music recommendation system using KKbox's donated dataset
- Current recommendation algorithms struggle with giving personalized recommendations
- Current algorithm uses a collaborative filtering-based algorithm with matrix factorization and word embedding

01/ DATASET

- Competition provided 6 data file
  - members
  - sample submission
  - extra song info
  - songs
  - test
  - train
- Analysis done with data from the train file joined with member and song file
- There were plenty of null values in most of the features
  - categorical data nulls filled in with “unknown”
  - numerical data nulls filled in with mean from the feature column

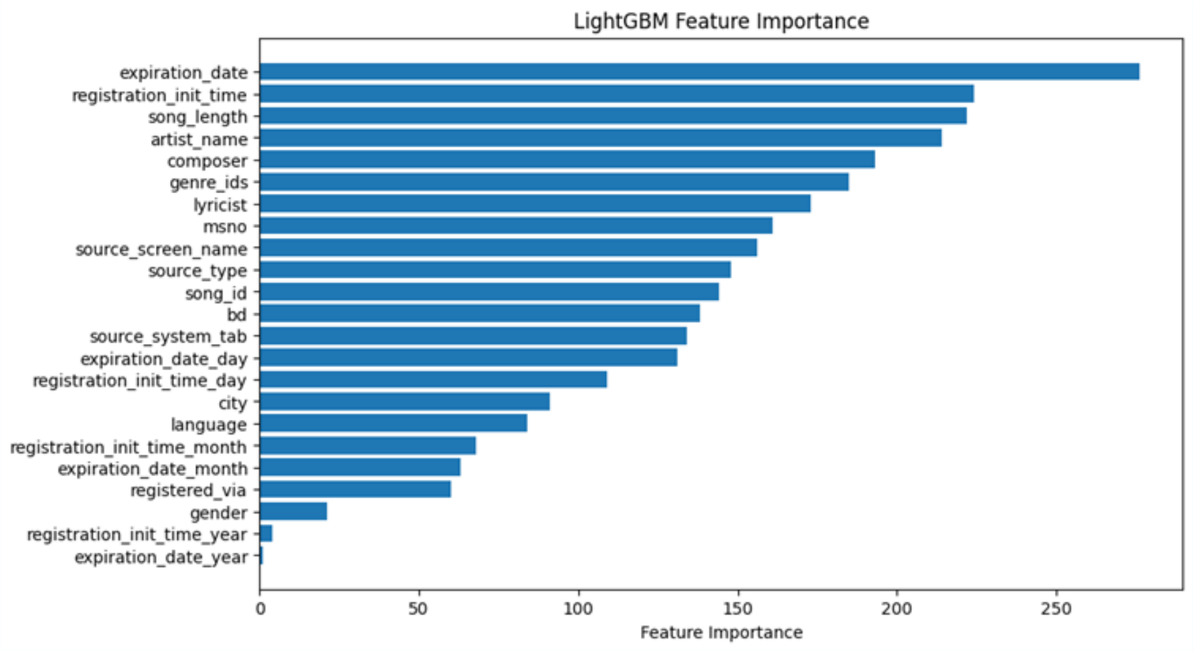
DECISION TREE MODEL

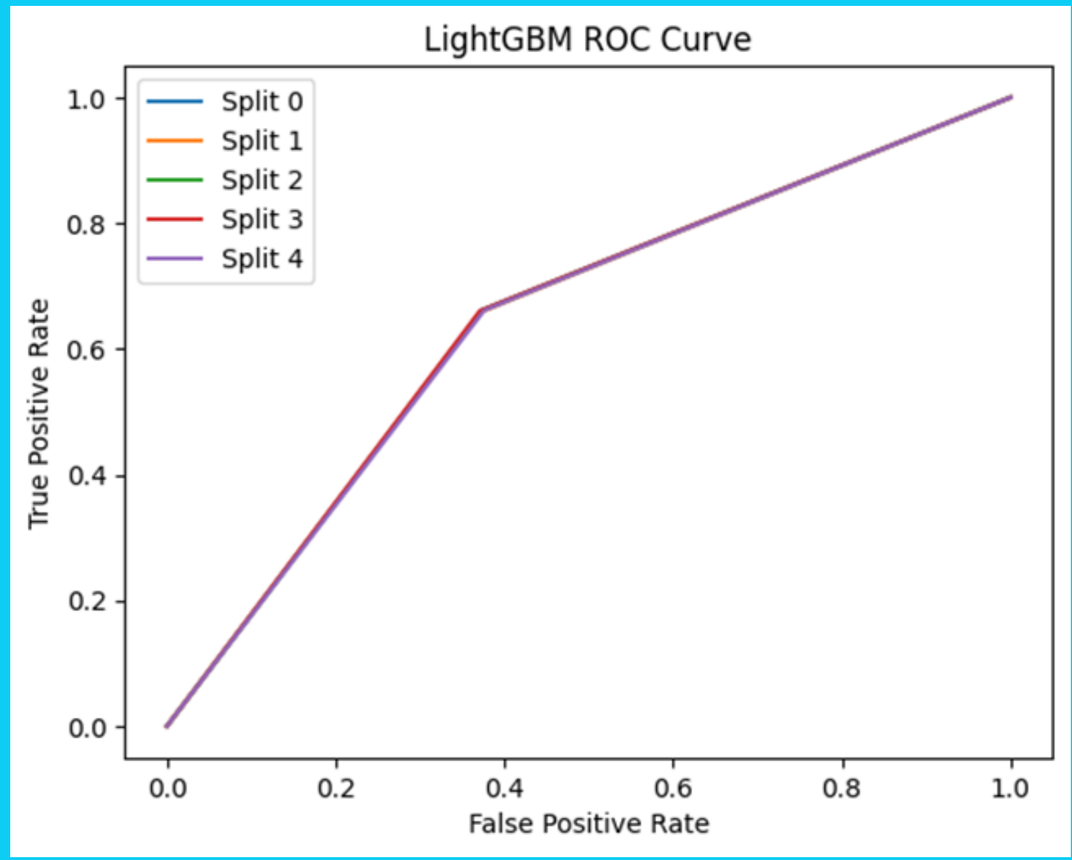




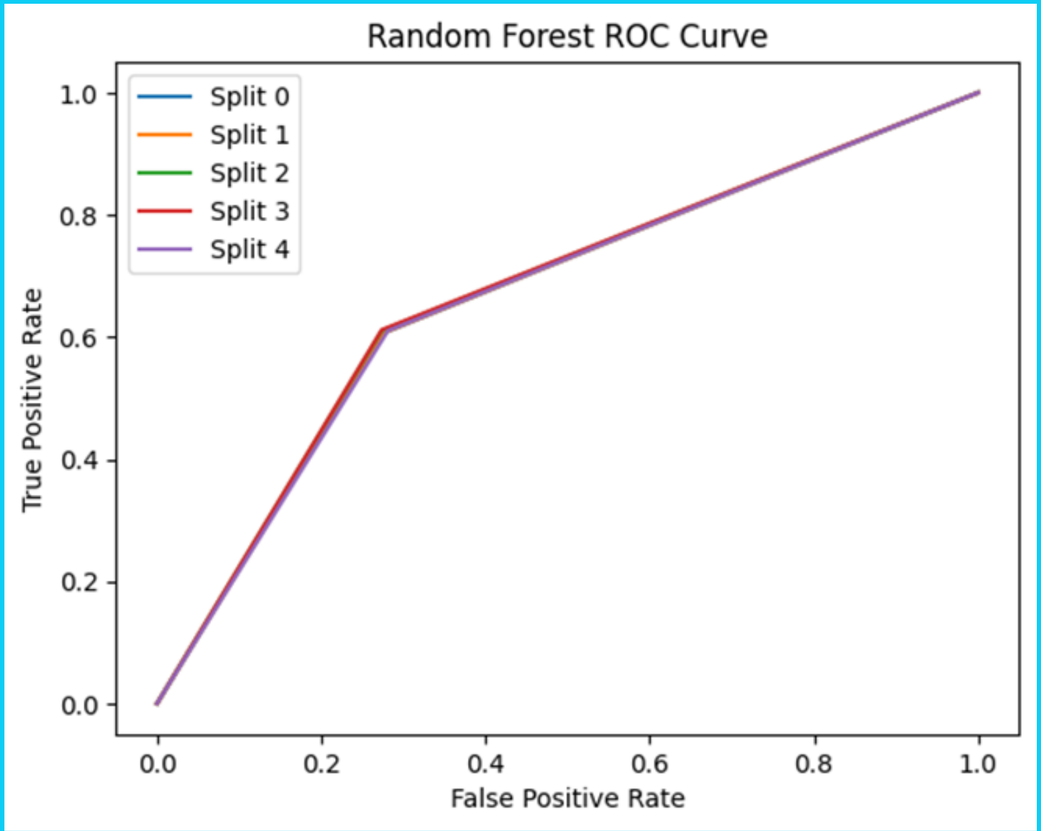
02 / MODELS

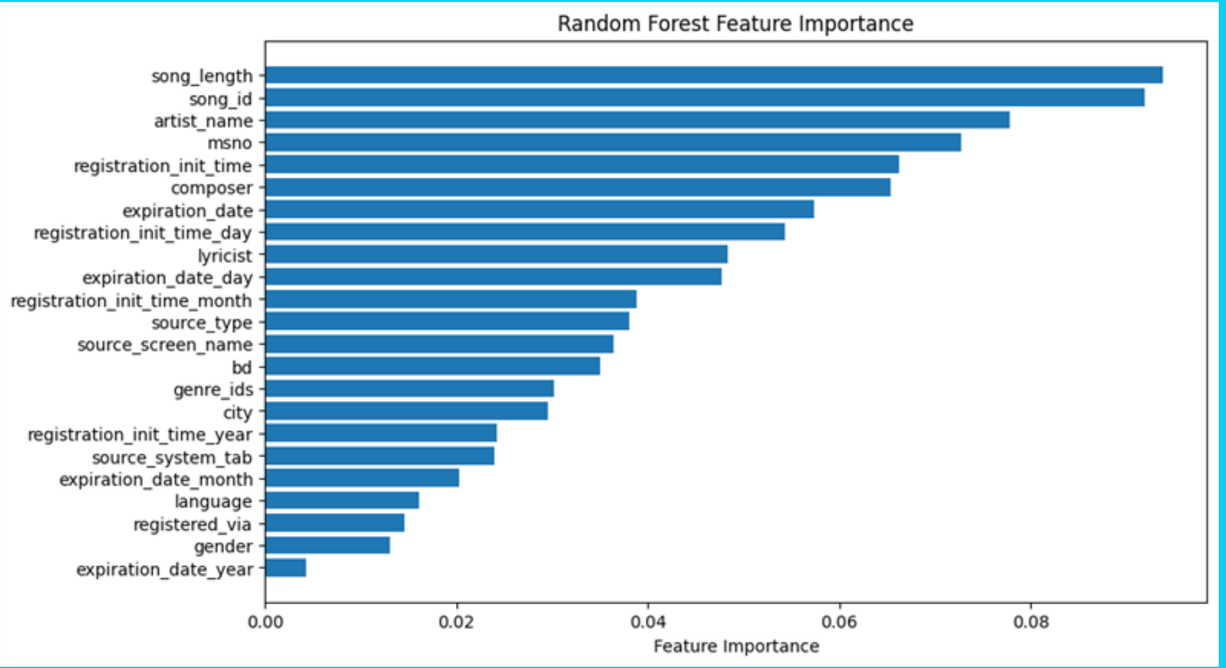
LIGHTGBM MODEL





RANDOM FOREST MODEL





Summary of Performance Metrics:

Split	Accuracy	Precision	Recall	F1 Score
0	0.6208	0.6214	0.6262	0.6238
1	0.6189	0.6210	0.6244	0.6227
2	0.6192	0.6210	0.6248	0.6229
3	0.6140	0.6153	0.6196	0.6175
4	0.6172	0.6185	0.6239	0.6212

Summary of Performance Metrics:

Split	Accuracy	Precision	Recall	F1 Score
0	0.6444	0.6414	0.6612	0.6512
1	0.6436	0.6427	0.6581	0.6503
2	0.6438	0.6418	0.6612	0.6514
3	0.6448	0.6430	0.6604	0.6516
4	0.6426	0.6403	0.6611	0.6505

Summary of Performance Metrics:

Split	Accuracy	Precision	Recall	F1 Score
0	0.6650	0.6891	0.6064	0.6451
1	0.6646	0.6897	0.6074	0.6459
2	0.6680	0.6933	0.6103	0.6492
3	0.6687	0.6934	0.6117	0.6500
4	0.6642	0.6873	0.6102	0.6465

- Highest accuracy came from Random Forest Model
- Highest precision came from Random Forest Model
- Highest recall came LightGBM Model
- Highest F-Score came from LightGBM Model
- Overall the best model depends on the goals of the algorithm
- minimizing false positives (precision) is critical - Random Forest
- capturing as many positive instances as possible (recall) is critical - LightGBM

03/ CONCLUSION

THANKS