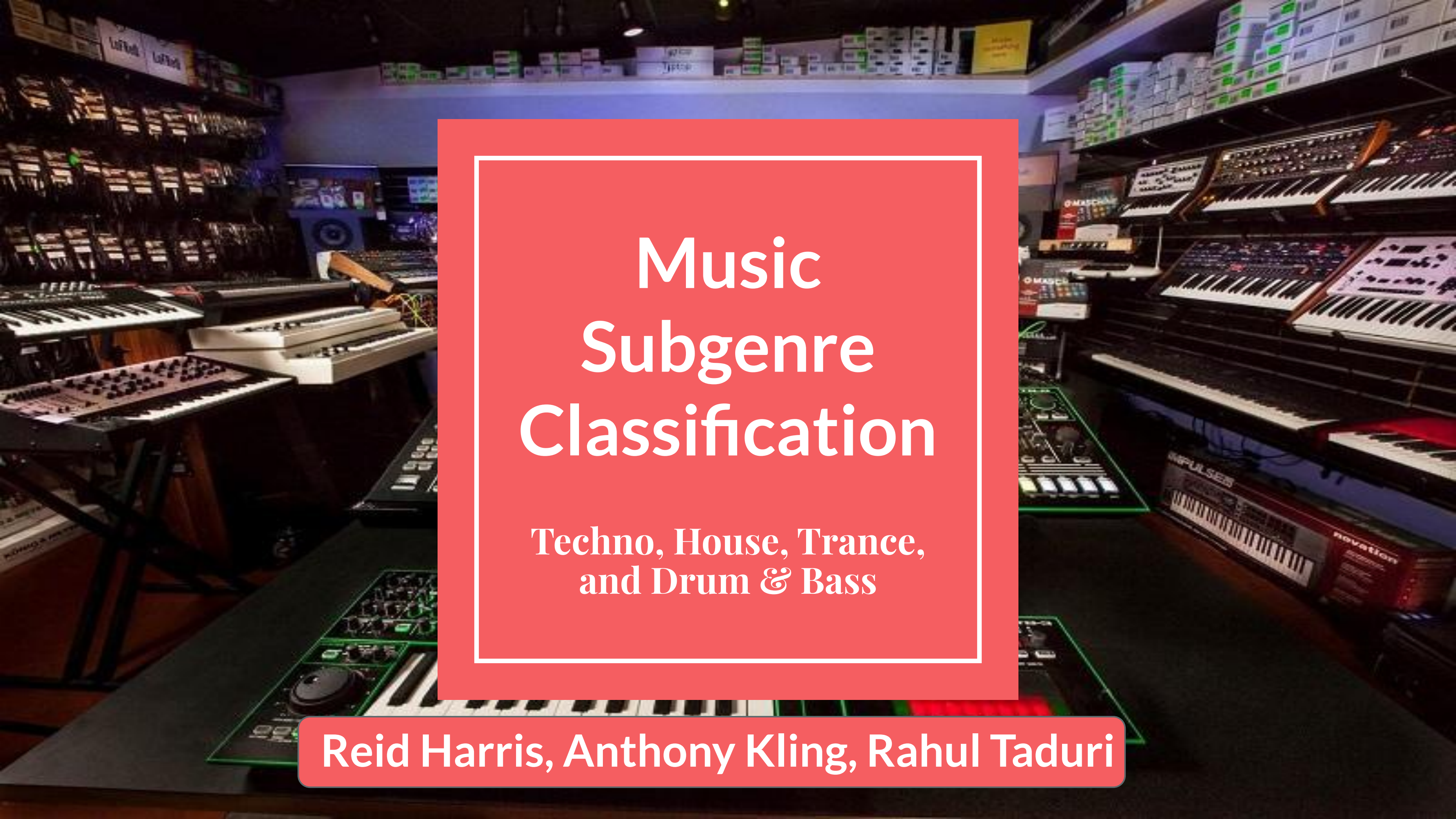


Music Subgenre Classification

Techno, House, Trance, and Drum & Bass

Reid Harris, Anthony Kling, Rahul Taduri

Introduction of Problem

- We tackle the supervised learning problem of multi-label genre classification for songs, focusing on four prominent subgenres of electronic music: techno, house, trance, and drum & bass.
- These subgenres are distinct and well-defined, but sometimes challenging to discern.

Techno

House

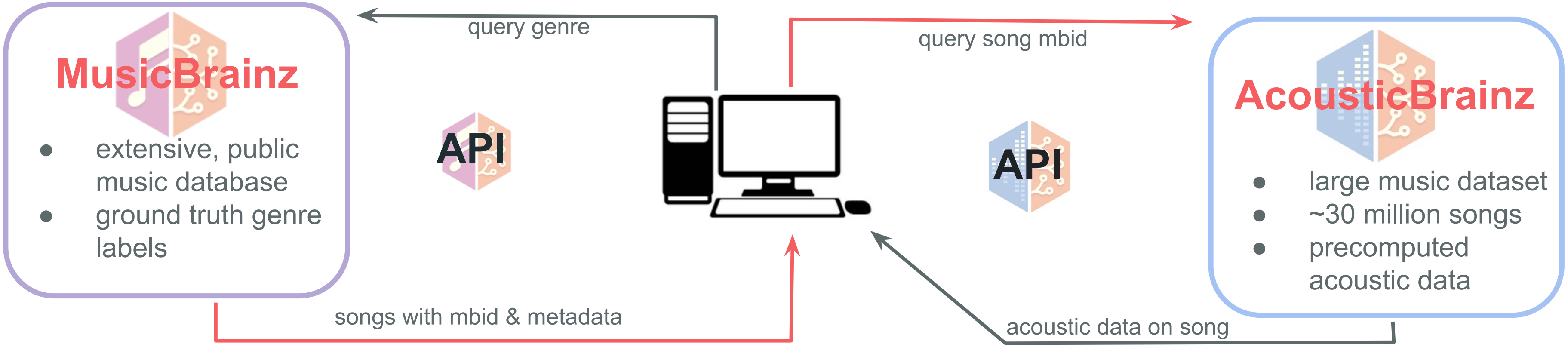
Trance

Drum & Bass

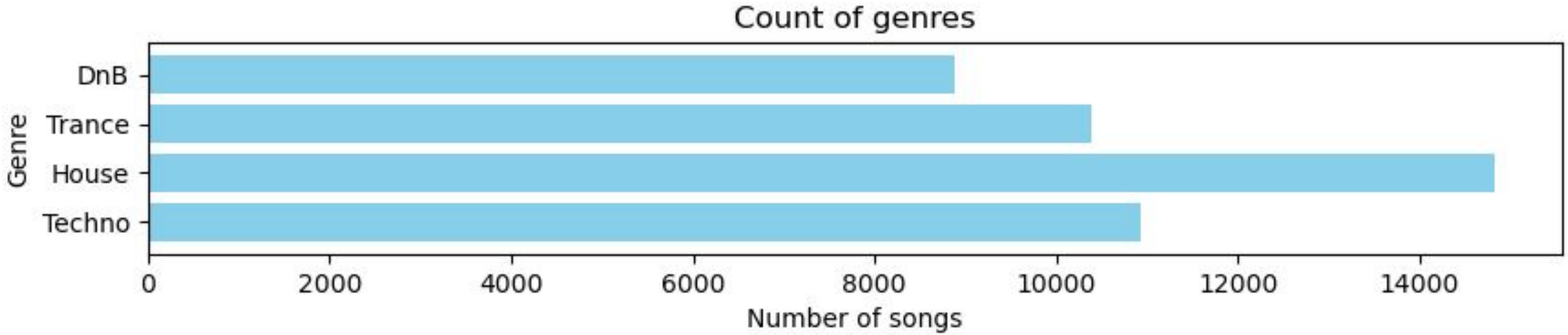
Techno + House

Dataset Collection

We used a subset of the AcousticBrainz dataset.



~80000 mbids ➜ 40000 songs in AcousticBrainz ➜ 37000 songs after cleaning



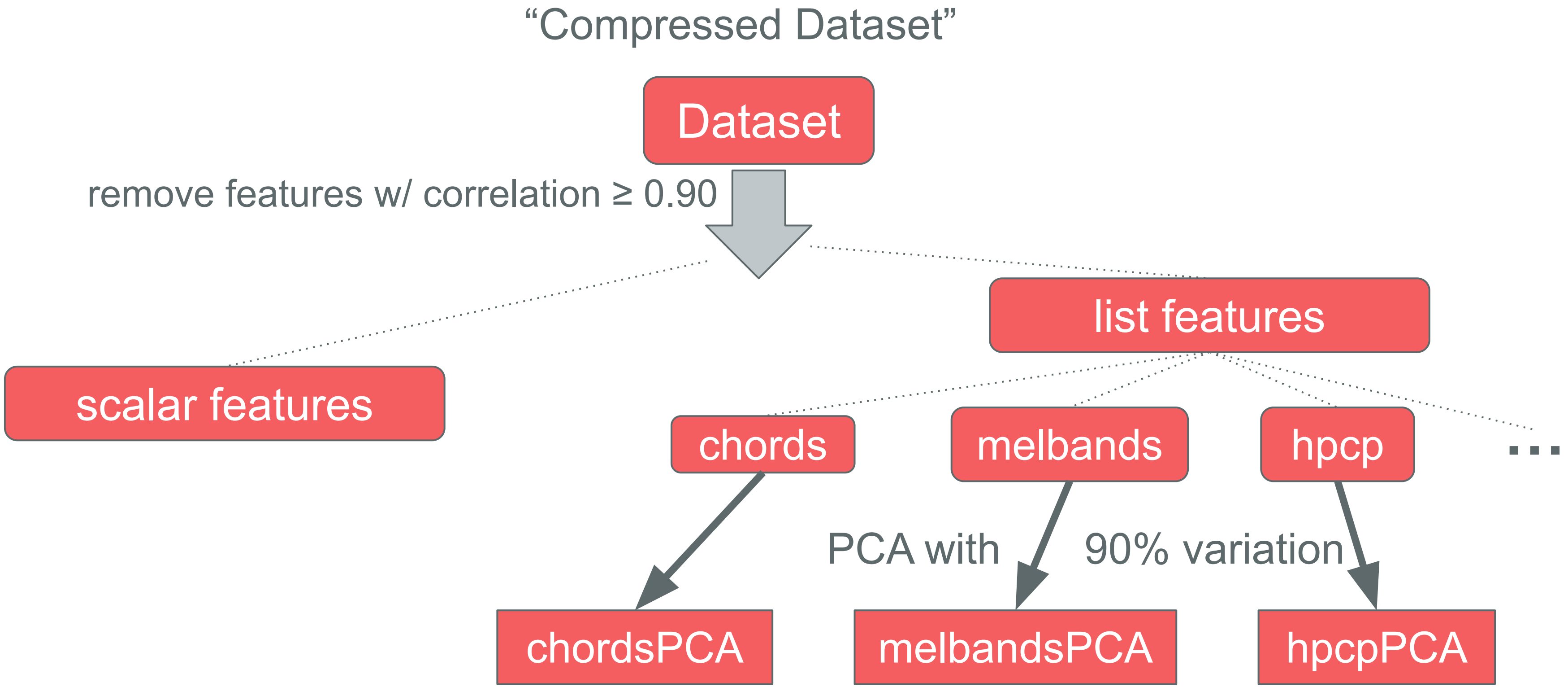
Feature Importance

Acoustic information extracted from AcousticBrainz had four categories: Low-level, Tonal, Rhythm and Metadata.

- Lowlevel: Bands (Bark, Mel, ERB, Spectral), loudness, silence, noise, etc...
- Rhythm: Beats position, BPM, beats loudness, dancebility, etc...
- Tonal: HPCP, etc...
- Metadata: Mostly textual and categorical information. 'Genre' is the multi class target variable.
- Most numerical data had lists resulting over 2500 columns when expanded.
- Lot of the features were dependent on each other showing high correlation.
- PCA on entire numerical data, limiting to 200 PC's explained 87% variance.
- Compressed dataset was also created which narrowed input features to 476.

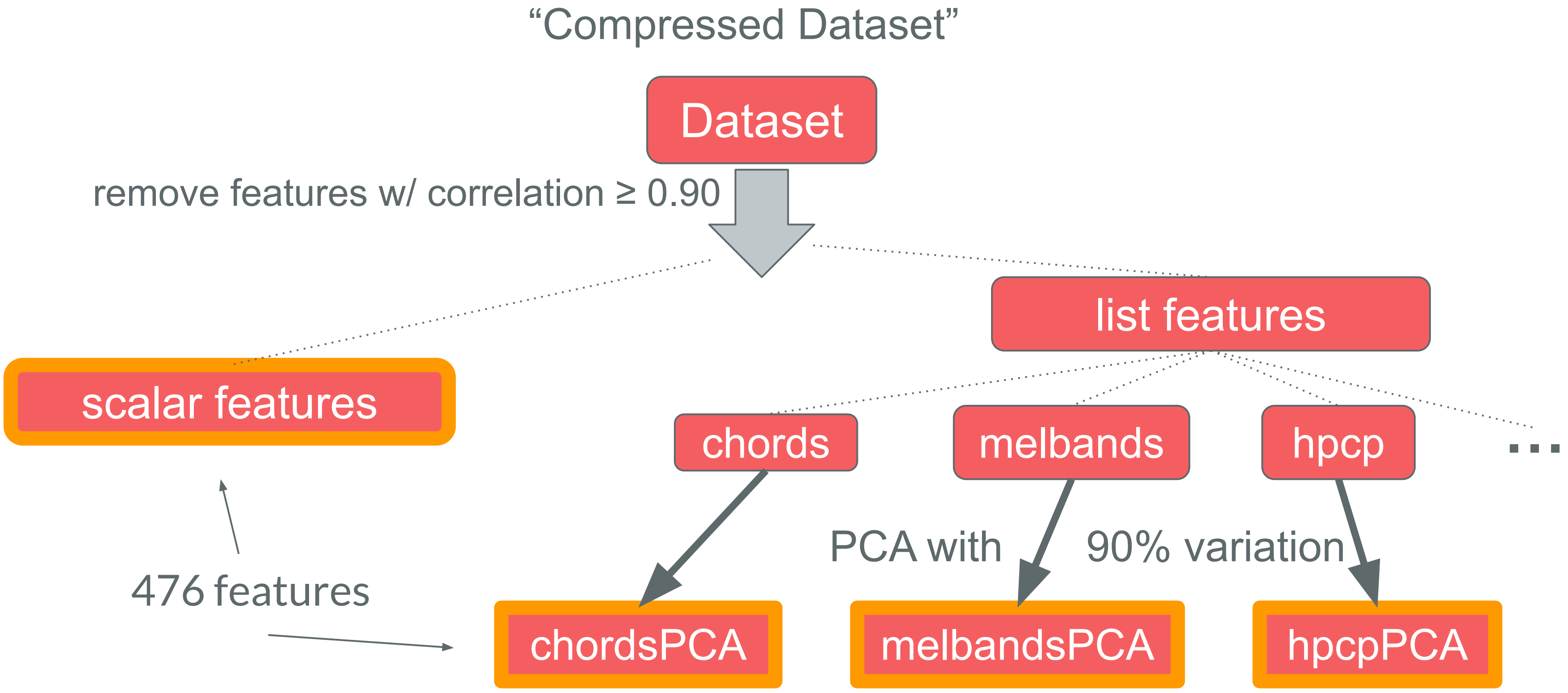
EDA & Feature Reduction

- Removed duplicates
- Remove all songs with **all** 4 genre labels (inaccurate).



EDA & Feature Reduction

- Removed duplicates
- Remove all songs with **all** 4 genre labels (inaccurate).



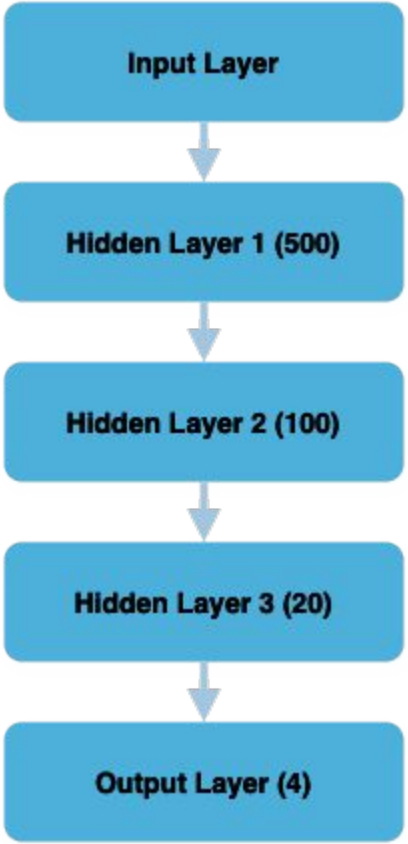
Classical ML Models

Model	Validation Accuracy (House/Trance/Techno/DnB)	F1 Score			
		House	Trance	Techno	Drum and Bass
Random Forest	0.64/0.75/0.76/0.77	0.27	0.33	0.31	0.12
KNN	0.68/0.73/**0.80**/0.81	0.57	0.18	**0.68**	0.44
Gaussian NB	0.49/0.72/0.48/0.81	0.57	0.41	0.5	0.58
Bernoulli NB	0.68/0.74/**0.80**/0.84	0.56	0.48	0.62	0.63
Bernoulli NB Oversampling	0.66/0.67/0.77/0.81	0.6	0.55	0.64	0.65
Bernoulli NB Undersampling	0.66/0.67/0.77/0.81	0.6	0.55	0.64	0.65
XGBoost	**0.79**/ **0.88**/ **0.80**/ **0.91**	**0.72**	**0.77**	0.60	**0.79**

- Results from compressed dataset.
- Skewness in target variable did not affect model accuracy.
- XGBoost performed better than most classical ML models.

Neural Networks

- For our neural network, we used 3 hidden layers of sizes 500, 100, and 20.
- We used 4 different datasets :
    - Raw data, including all numerical data from AcousticBrainz (containing 2614 features)
    - 3 other datasets in which similar lists of numerical data were compressed into 603, 476, and 379 features
- For the raw data, we randomly split the full dataset into a training set (30000) and a test set (6885)
- A 5-fold cross-validation was performed by randomly splitting the training set into 5 sets of 6000 songs each.



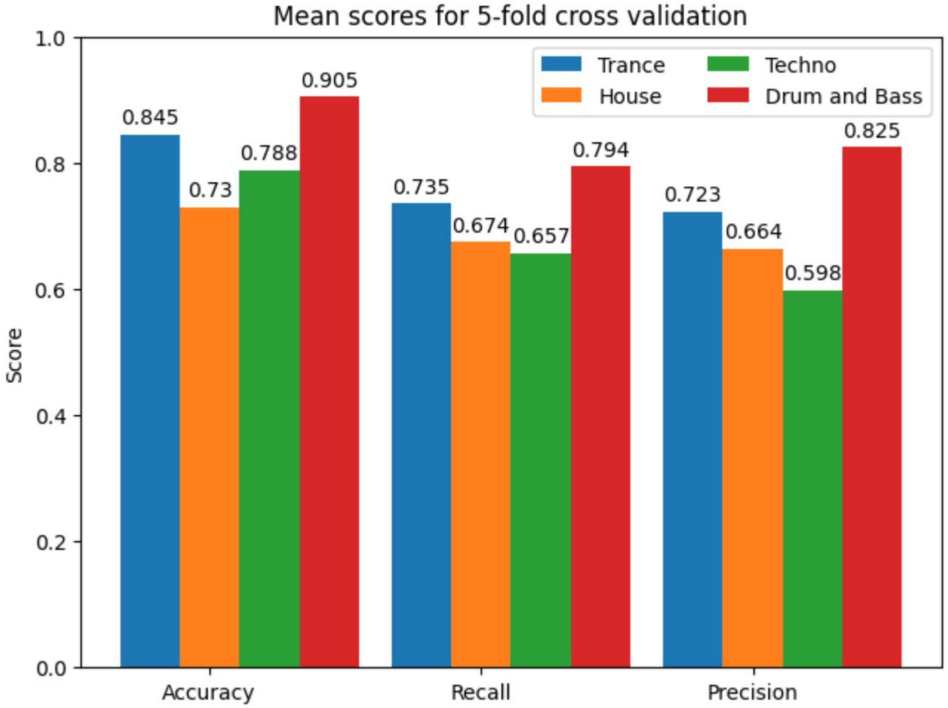
Neural Network Training

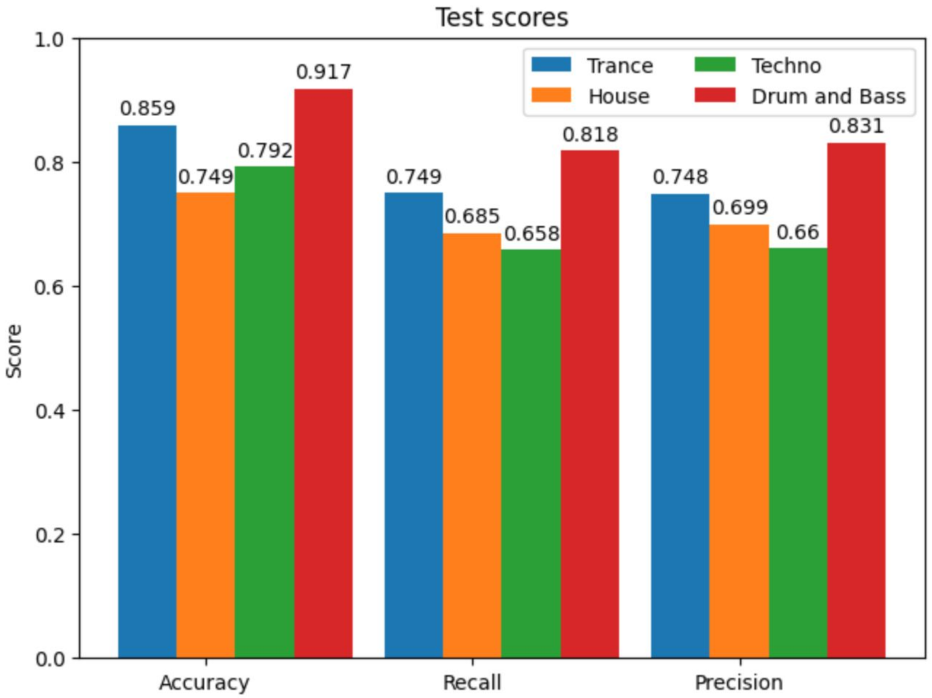
- Training each neural network used SGD with a learning rate of 0.005 and momentum of 0.9.
- For each epoch, the training data was sent in batches of size 128 in order to speed of training.
- The network was trained for a total of 50 epochs.
- The loss function used was the binary cross entropy loss (BCELoss):

$$\mathcal{L}(\hat{y}, y) = -\frac{1}{N}\sum_{i=1}^{N}(y_i \log(\hat{y}_i) + (1 - y_i)\log(1 - \hat{y}_i))$$

Neural Network Results

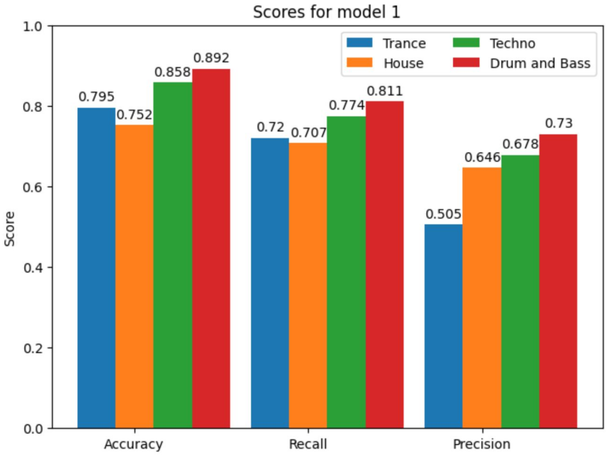
- The mean accuracy, precision, and recall scores for the cross-validation and the scores for the test set are shown below for each of the 4 classes.

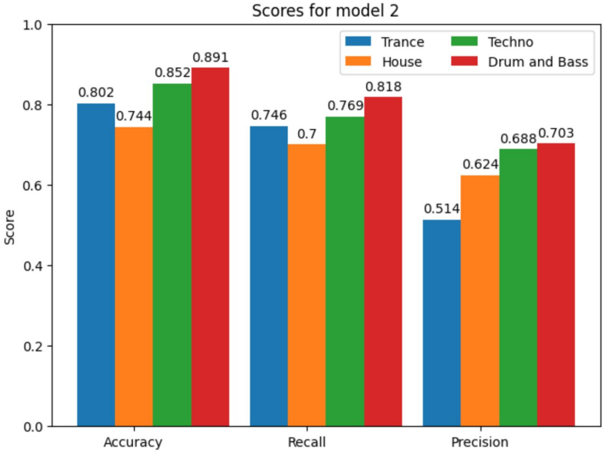


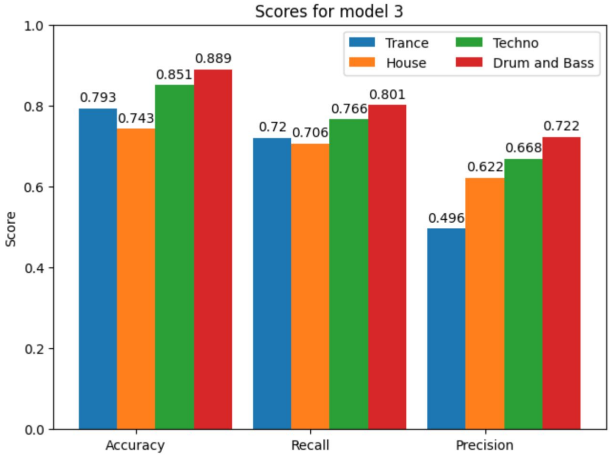


Neural Network Results

- The neural network was trained on each of the three compressed datasets and the results are given below.
- All measures were approximately the same or worse, except for the Techno music accuracies and the recall scores, for which the compressed models performed better than the network trained on the raw data.







Conclusion & Future Directions

Accuracy/ F1	Techno	House	Trance	DnB
XGBoost	**0.80/ 0.69**	**0.79/ 0.72**	**0.88/ 0.77**	0.91/ 0.79
Neural Network	0.79/ 0.62	0.75/ 0.69	0.86/ 0.73	**0.92/ 0.82**

- We can make our genre labels even more accurate by combining genre labels from other music databases.
- Try other groups of subgenres.
- Obtain more data. Due to time and computational restraints, we stopped querying via the API after ~40,000 songs.

Thank you