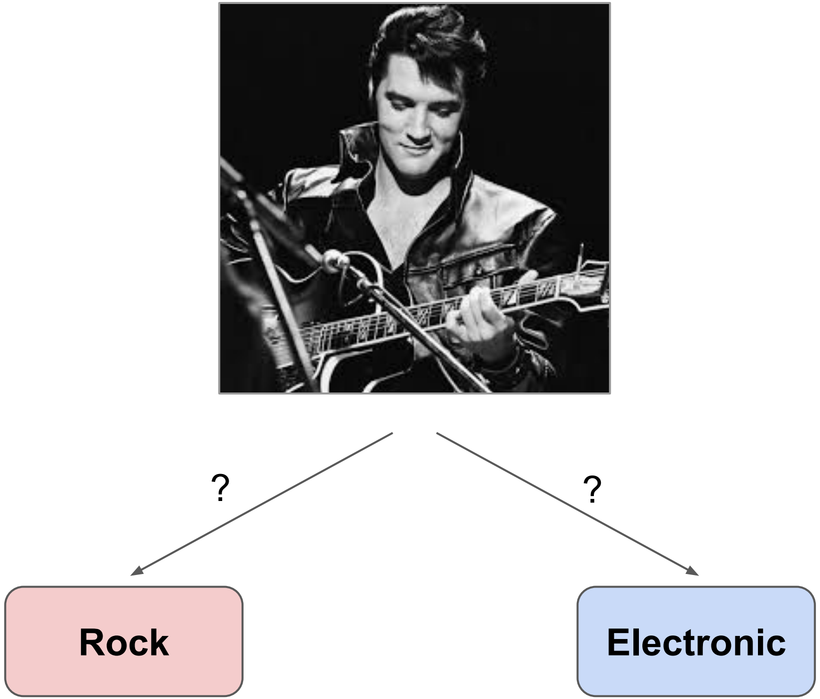
Genre Classification: A Transductive, Inductive and Deep Approach

Group 32:
Isinsu Katircioglu, Sena Kiciroglu,
George Adaimi, Okan Altingovde

Introduction

- Classify genres of the FMA dataset
- Practical applications:
  - Building genre specific playlists, song recommendations etc.
- Methods:
  - Classical ML methods: SVM, MLP, etc.
  - Transductive learning methods
  - Graph Convolutional Networks



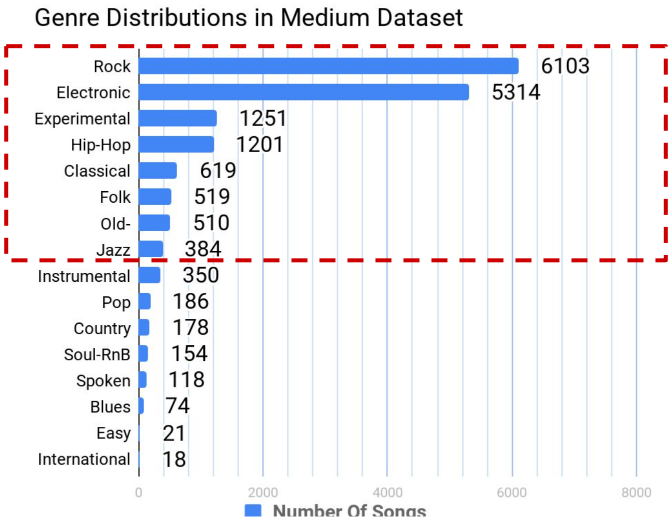
FMA Dataset

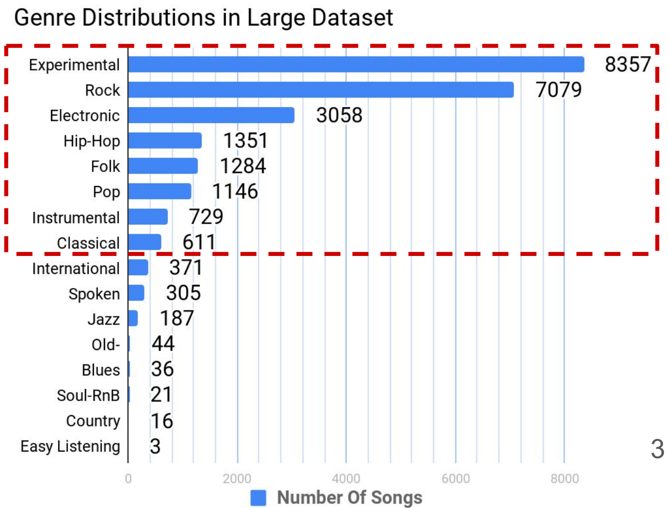
- 106k songs from 16k artists.
- Small, medium and large subsets.
- Take 8 most popular genres of each subset.

Forming the graph:

- MFCC features
- "Correlation" distance metric
- Threshold edges below 0.9
    - Remove disconnected nodes!

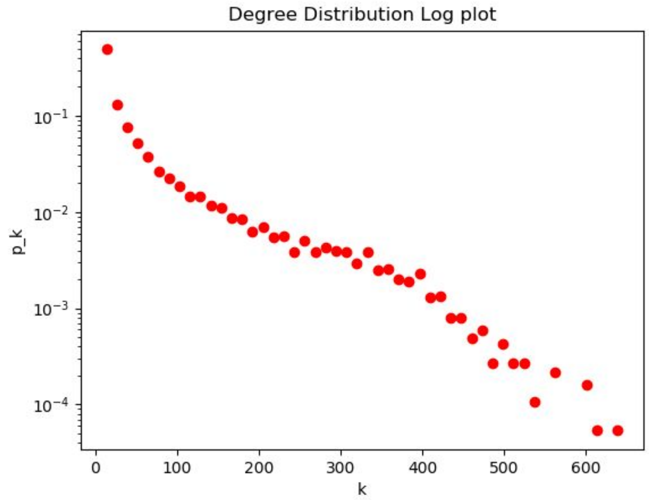
M. Defferrard, *et al.* FMA: A Dataset for Music Analysis. ISMIR 2017.





Graph Statistics

- Sparse graph
- Average degree generally higher than real life graphs
- Many connected components, but most of nodes in largest component (96%).



Number of Nodes	18648
Number of Edges	447280
Sparsity	0.2%
Average Degree	48.0
Number of Connected Components	214
Size of Largest Connected Component	18084
Average Clustering Coefficient	0.34

Pre-Processing

- **PCA** to reduce feature dimensionality
- Train a **multi layer perceptron (MLP)** to produce better features, tailored to genre classification.
    - Produces better adjacency!

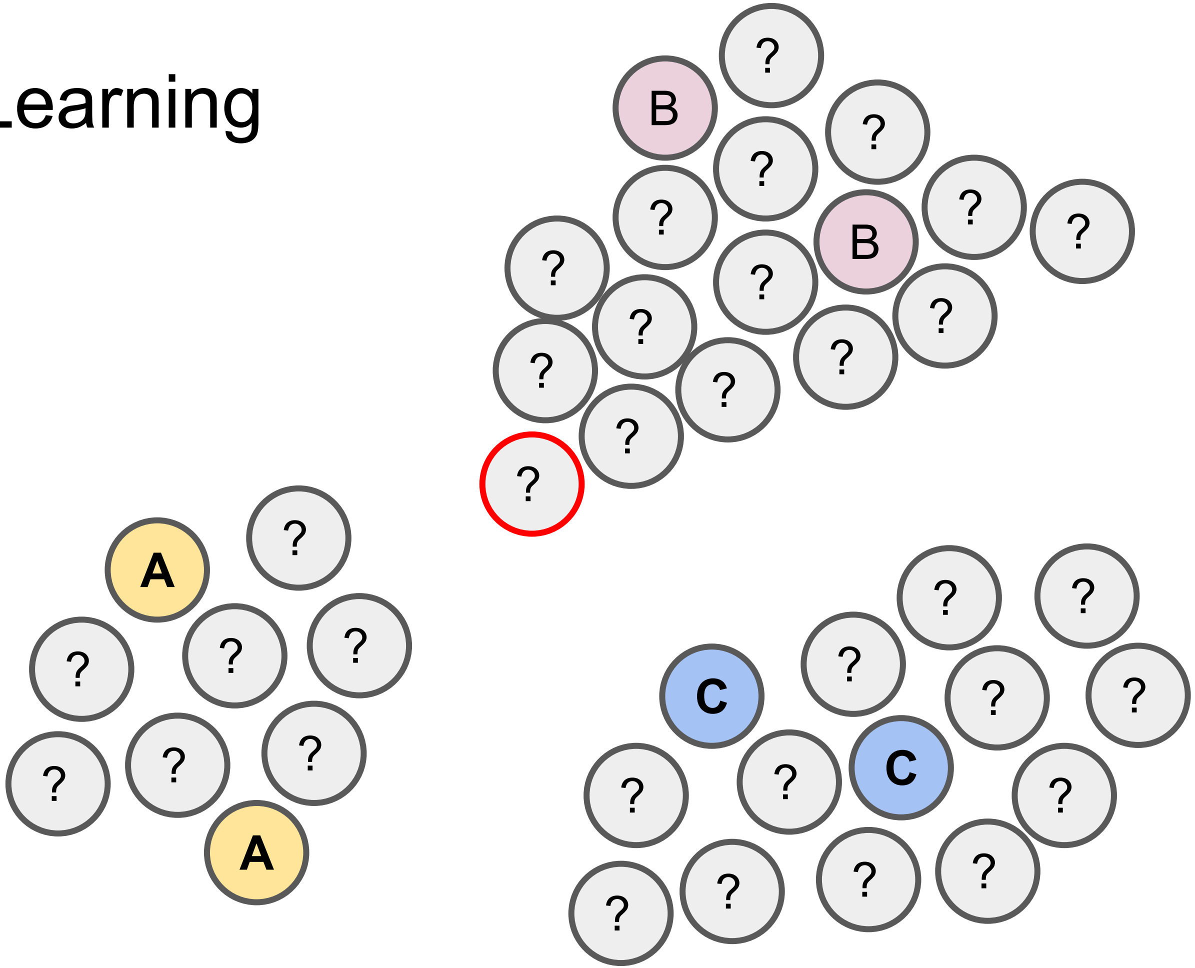
	SVM	RF	KNN
% Error with MLP & PCA	**20.38**	**21.44**	21.51
% Error w/o MLP, with PCA	21.86	26.48	**21.13**
% Error w/o PCA w/o MLP	-	26.87	27.95

- MLP-features and PCA generally reduce error

Classical Inductive ML Methods

- Most commonly used ML methods:
  - Support Vector Machine (SVM)
  - Random forest (RF)
  - K-Nearest Neighbours (KNN)
  - Multi-Layer Perceptrons (MLP)
- Separate graph using "split" in dataset into "training" and "test".
- Use cross-validation for training.

Transductive Learning

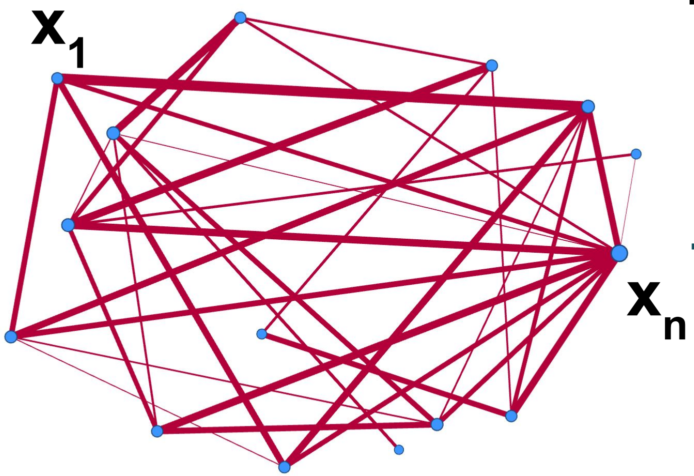


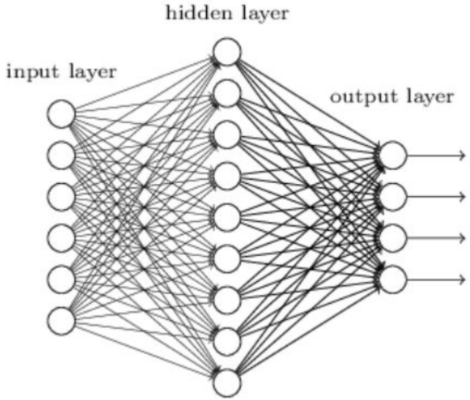
X. Zhu *et al.* ICML 2003
Y. Yamaguchi, *et al.* AAAI, 2015.
X.M. Wu, *et al*. NIPS, 2012.

Graph Neural Network

Train

Class probabilities

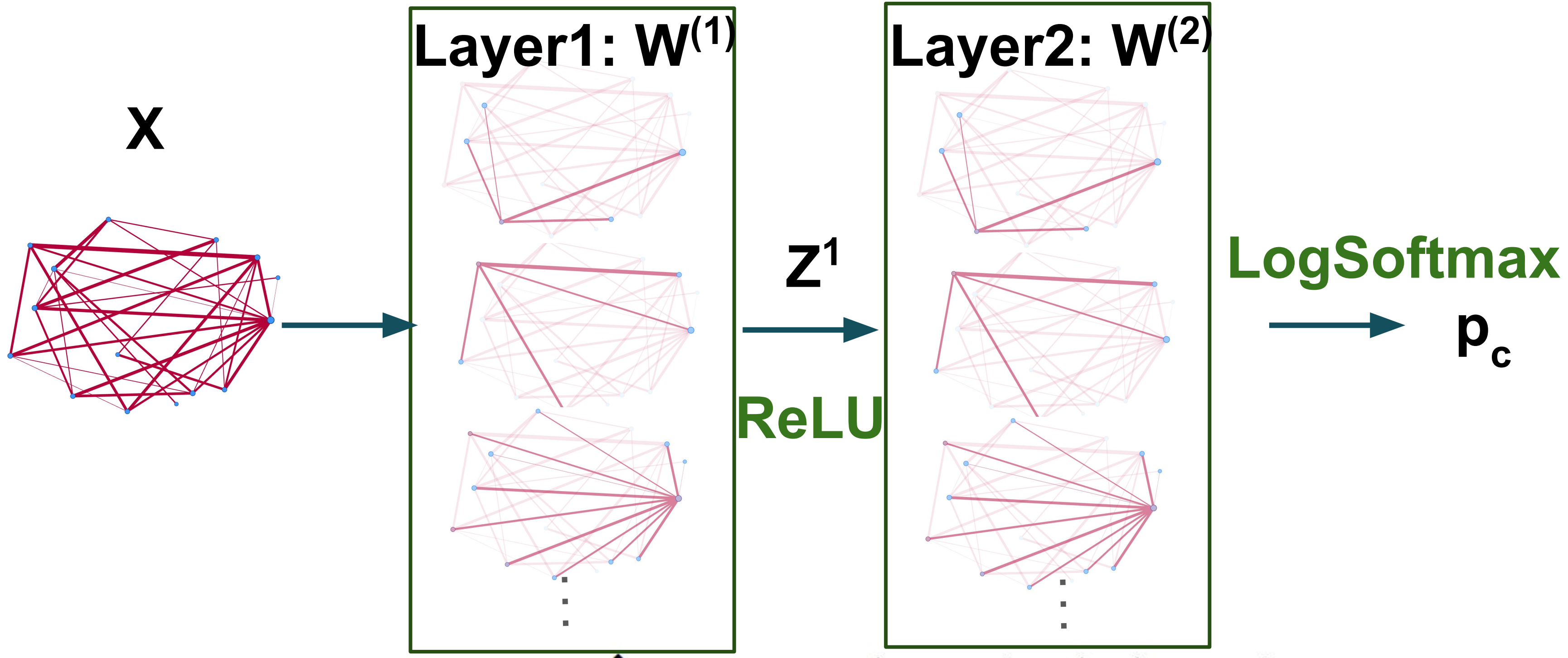




0.34	Electronic
0.28	Rock
0.07	Hip-Hop
0.31	Experimental

$$Y = f_{gcn}(X, A)$$

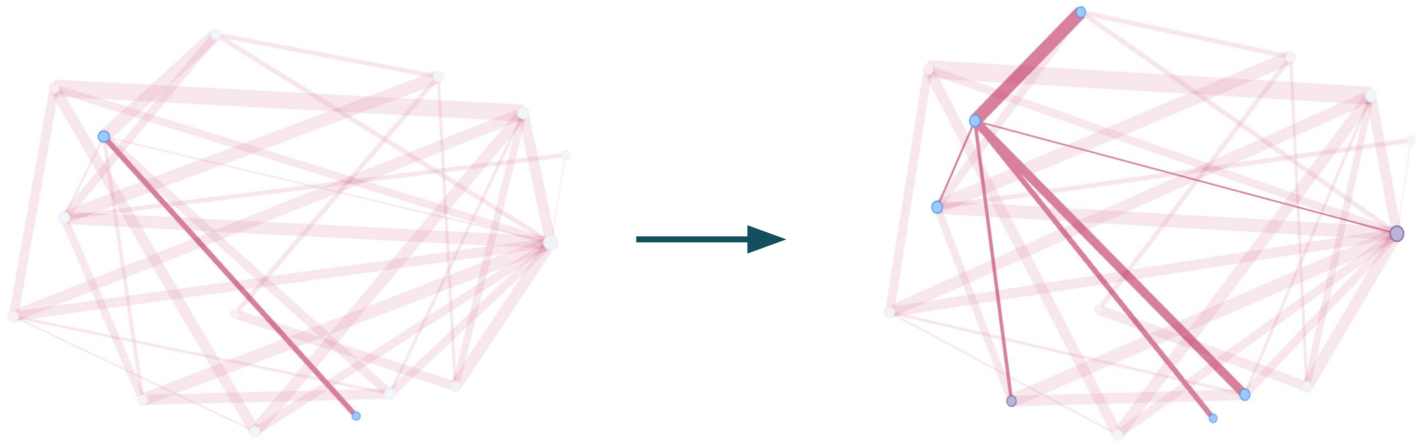
Graph Neural Network



$$Z^{l+1} = f_l(Z^l, \hat{A}) = \sigma(\hat{D}^{-1/2}\hat{A}\hat{D}^{-1/2}Z^lW^l)$$

One Step Further: Graph K-Hop

K = 2



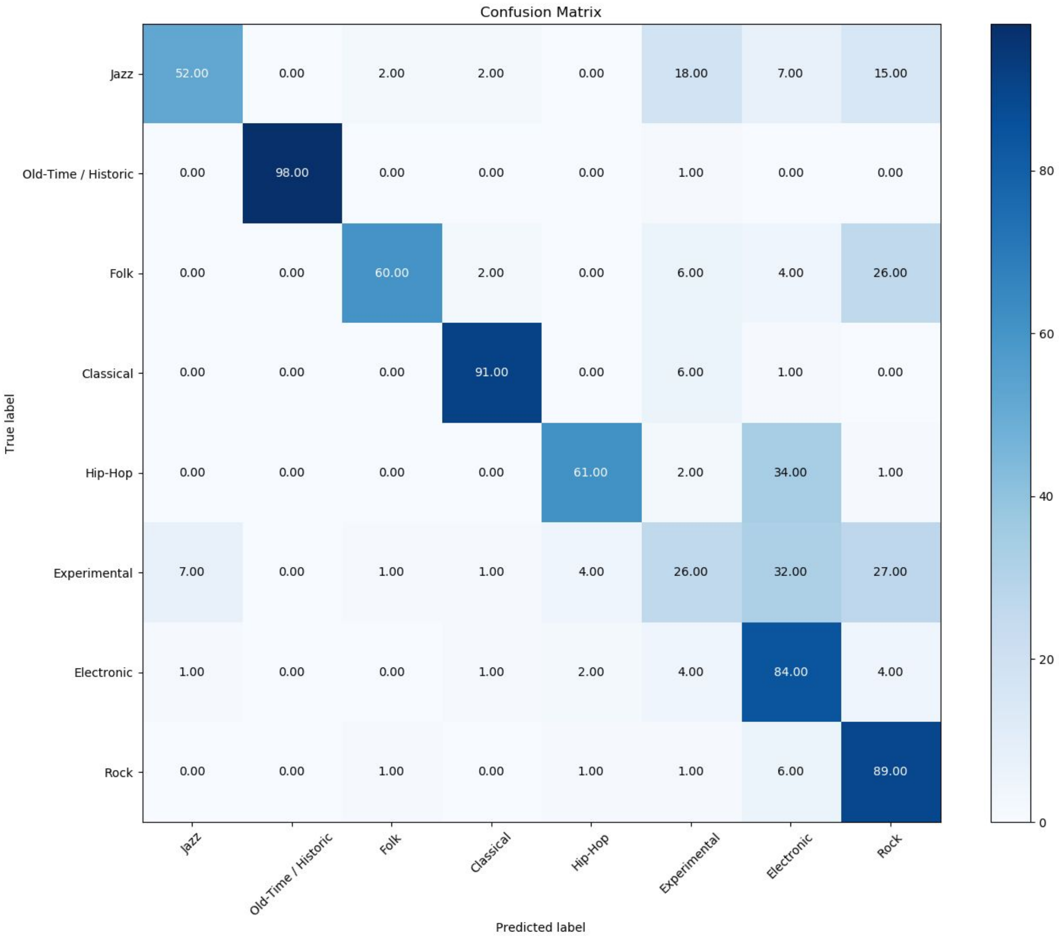
- Reach neighbors at distance 2.
- Fuse the features coming from network with one-hop adjacency and two-hop adjacency matrices.

Results

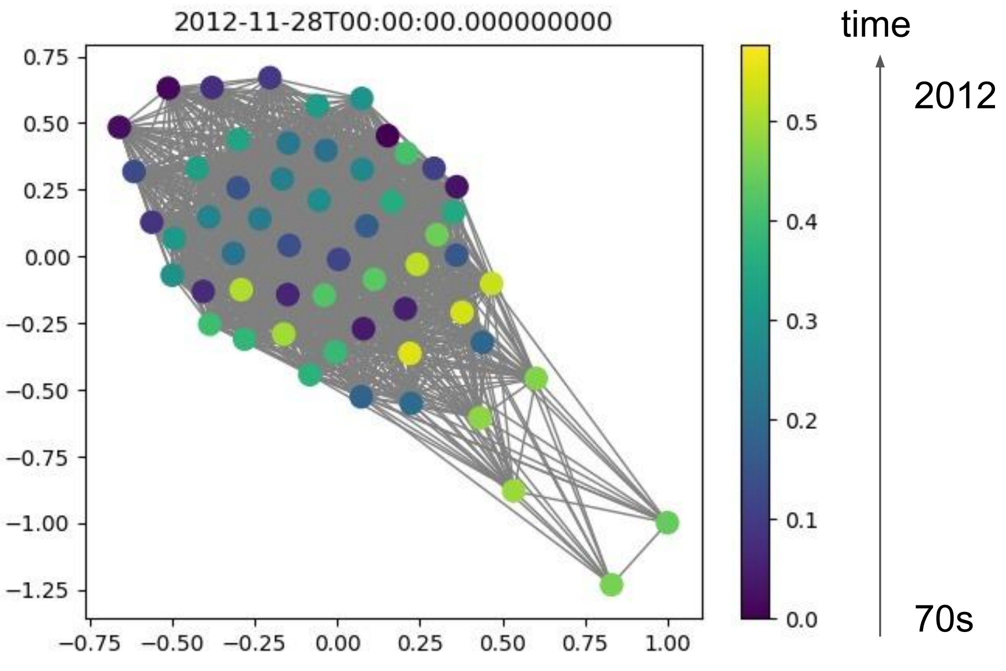
Method	Medium Dataset Error (%)	Training Time (s)	Large Dataset Error (%)	Training Time (s)
SVM	20.38	85	37.13	160
RF	21.44	46	**36.46**	36
KNN	21.51	39	37.07	48
MLP	20.59	68	38.02	90
GCN	20.53	256	37.78	241
GCN-KHop	**20.24**	731	37.56	1655
GFHF	22.15	93	38.81	140
OMNI-Prop	22.43	303	40.15	339
PORW	26.26	0.7	38.36	0.635

Results

- "Good predictions for "Old-Time/Historic", "Classical", "Rock"
- "Experimental" is confused with "Electronic" and "Rock"
- "Folk is confused with "Rock"



Future Work: Drifting in time



Rock

- Changing trends in music industry
- Technological advancements in track recording devices and instruments
- Special effects
- Predict release date of song within a genre

Discussion & Conclusion

- Challenges:
  - Forming good graph - crafting good features
  - Genre imbalance within medium and large datasets

- Summary:
  - Used inductive, transductive and deep methods for genre classification.
  - Gained experience in building **graph representations** of data.
  - Designed, trained and tested **GCNs**, which could also be interesting for our own research!