# Polyphonic Music Transcription using Deep Learning Methods

*Aniruddha Zalani*
*Ayush Mittal*
*Course Project-CS365*

# What is polyphony

- Two or more independent notes playing at the same time
- Monophonic music - only one node is played at a time.

# Problem Statement

- Extract the notes played in a polyphonic piano song.
- Resynthesize the song from the transcribed notes.
- Many notes are played at once, therefore techniques of multi-class classifiers are not applicable.

# Motivation

- Many naturally occurring phenomena such as music, speech, or human motion are inherently sequential.

- Help in
  - Plagiarism detection
  - Artist identification
  - Genre classification
  - Composition assistance
  - Music tutoring system

# Related Work

- Some interesting work has been done using non-negative matrix factorization techniques [1] and [2].
- Poliner and Ellis' piano transcription system [3] consists of 87 independent support vector machine (SVM) classifiers
- However, most of the recent work involve feature learning using deep learning methods before the classification step.

# Related Work …

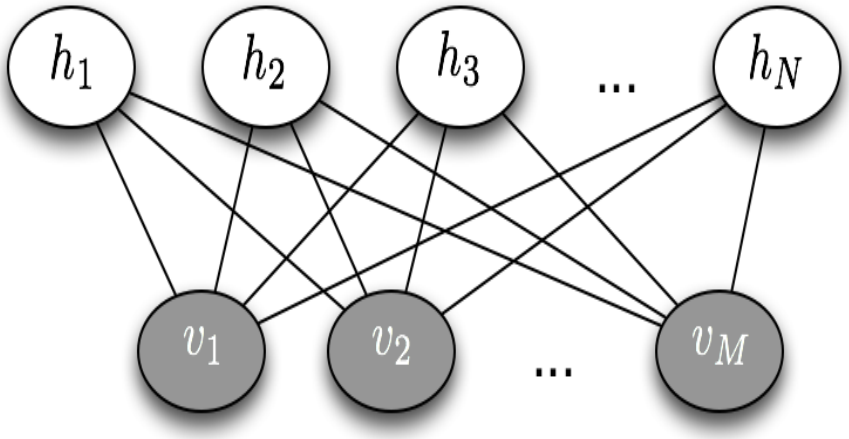
- Juhan et al., [4] trains deep belief network by “greedy layer wise stacking of RBMs”.
- They used DBN-based feature representations as input to the linear SVM for single note and multi note training.
- They used HMM-based post processing to temporally smooth the SVM output.
- We mostly follow the work by Nicholas et al., [5]

# Our Approach

- We focus on two major approaches for learning feature representations:
  - RNN-RBM based model -
    - Hessian-free optimization
  - Convolutional Deep Belief Network based model.
- In classification step we input features learned from previous step into the SVM classification method of Poliner and Ellis.
- Finally, we use HMM for temporal smoothing of the SVM output.

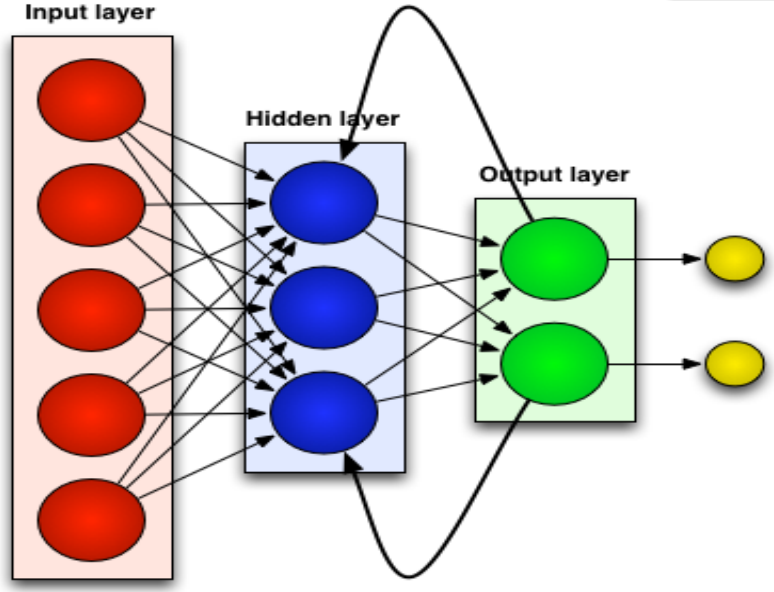
# RBM

- A generative stochastic neural network that can learn a probability distribution over its set of inputs.
- Restriction that their neurons must form a bipartite graph
- Input units features of their inputs,
- Hidden units that are trained.
- Contrastive Divergence uses two tricks to speed up the sampling process:

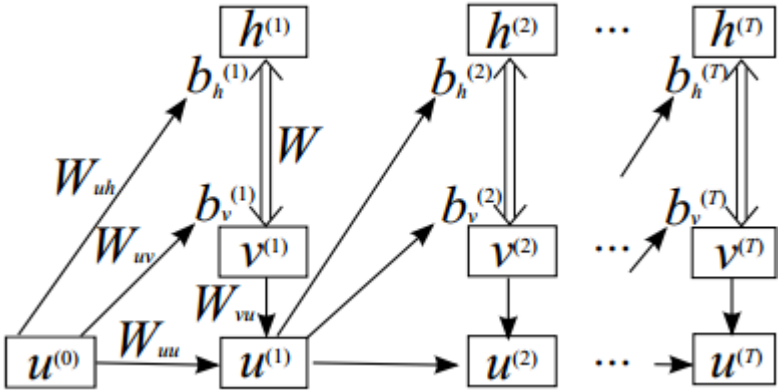


# RNN

- Connections between units form a directed cycle
- RNNs can use their internal memory to process arbitrary sequences of inputs.
- Each unit has a time-varying real-valued activation



# RNN-RBM




- Multimodal Conditional distribution of v(t) given A(t) where

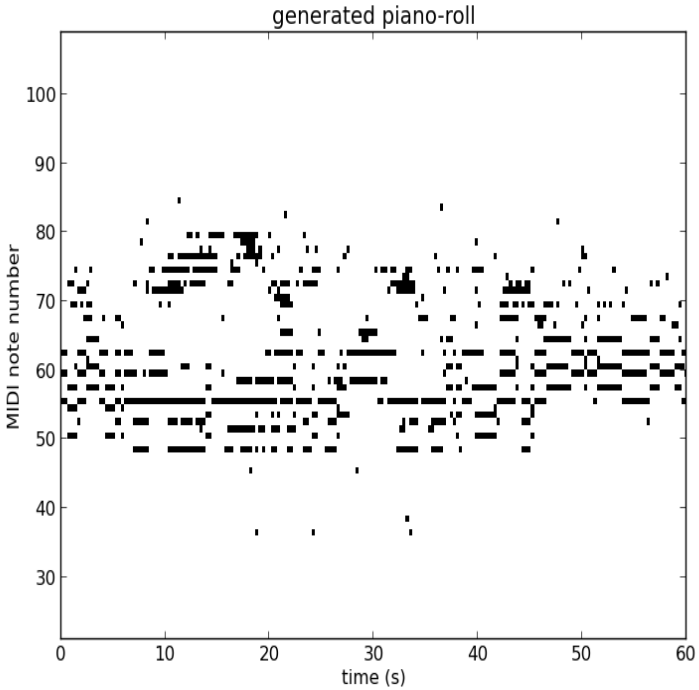
$$\mathcal{A}^{(t)} \equiv \{v_\tau | \tau < t\}$$

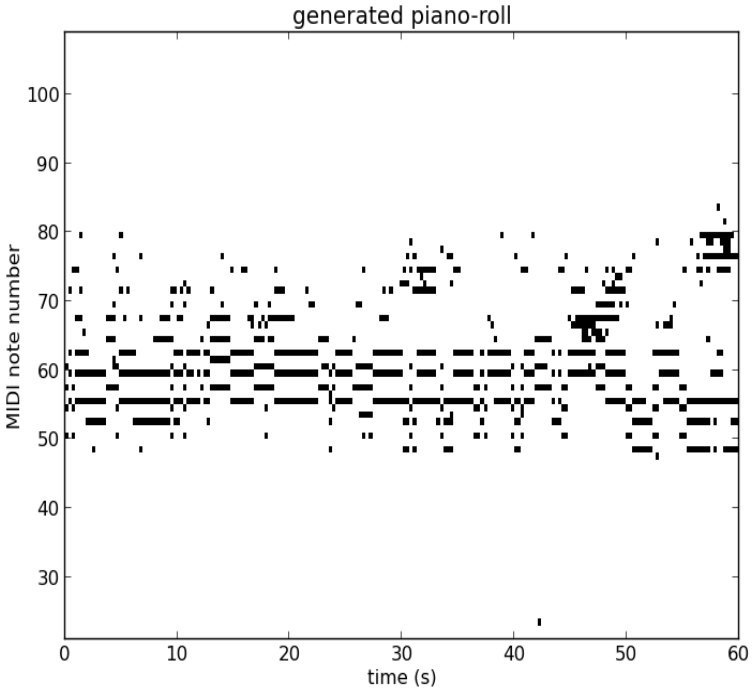
- $$P(\{v^{(t)}\}) = \sum_{t=1}^{T} P(v^{(t)}|\mathcal{A}^{(t)})$$

# Dataset

- Piano midi.de : Classical Piano midi archieve. [6]
- Nottingham: is a collection of 1200 folk tunes with chords instantiated fro, the ABC format. [7]
- MAPS: is a large piano dataset that includes various patterns of playing and pieces of music [8]
- ~70 hours of polyphonic music.

# What we have done?



# What work is left?

- Classification of notes using SVM with features learned from RNN-RBM as input to SVM.
- Post processing involving temporal smoothing using HMM and transcription.
- Trying out Convolutional Deep Belief Networks for feature discovery.

# References

[1] Arnaud , Arshia et al. \Real-Time Detection of Overlapping Sound Events with Non-Negative Matrix Factorization"

[2] Paris and Judith Non-Negative Matrix Factorization for Polyphonic Music Transcription, IEEE 2003.

[3] G. Poliner and D. Ellis: “A discriminative model for polyphonic piano transcription,” EURASIP Journal on Advances in Signal Processing,vol.2007, 2007

[4] J. Nam, J. Ngiam and H. Lee,Classification- Based Polyphonic Piano Transcription Approach Using Learned Feature Representations," ISMIR , pp. 175-180, 2011.

# Reference …

[5] N. Boulanger-Lewandowski, Y. Bengio and P.Vincent, Modeling temporal dependencies in high-dimensional sequences: Application to polyphonic music generation and transcription," ICML, 2012.

[6] http://www.piano-midi.de/

[7] http://www-etud.iro.umontreal.ca/~boulanni/icml2012

[8] ftps://ftps.tsi.telecom-paristech.fr/share/maps/

# CDBN

- Lee et al.[6] proposed the use of CDBNs in Music Information Retrieval.
-