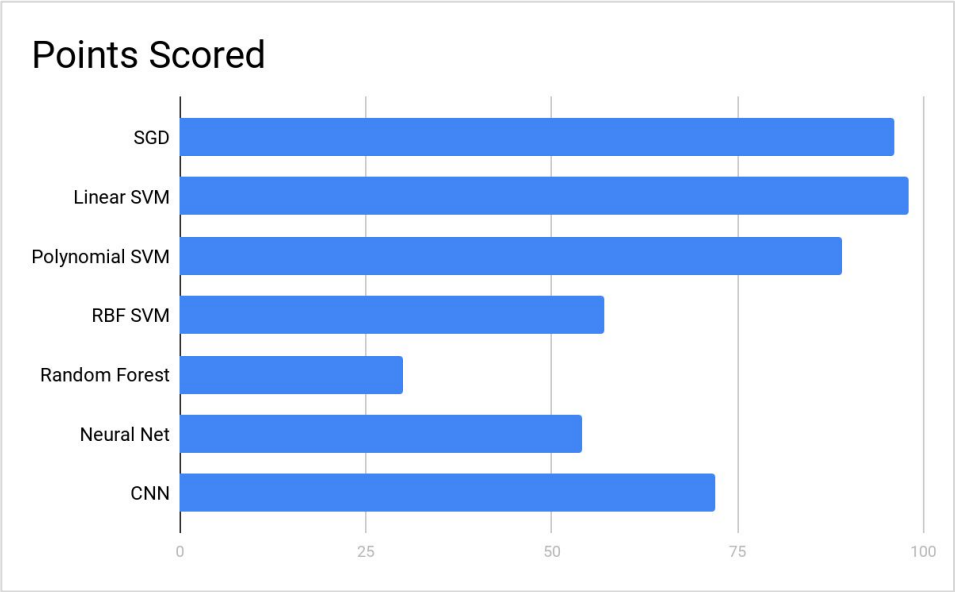
Musical Instrument Classification

Spencer Gass
Tim Coulter
Eli Johnson
Sam Tanner

Project Summary

- Evaluate ability of classification methods to determine musical instruments
- Evaluated Nsynth data set, with over 300,000 recordings, of 11 different instruments, played for 4 seconds at different notes
- Classifiers
  - Stochastic Gradient Descent
  - Support Vector Machine
    - Linear
    - Polynomial
    - Radial Basis Function
  - Random Forest
  - Neural Networks



Feature Selection

- NSynth data divided into:
  - 289,305 training samples
  - 12,678 validation samples
  - 4,096 test samples.
- Contains 11 different instruments, not evenly represented
- Valuable Features:
  - Mel-frequency cepstral coefficients (MFCC)
  - Spectral centroid
  - Spectral bandwidth
  - Spectral slope
  - Zero crossing rate
- Validation runs on small Neural Network to select transform parameters

Family	Acoustic	Electronic	Synthetic	Total
Bass	200	8,387	60,368	68,955
Brass	13,760	70	0	13,830
Flute	6,572	35	2,816	9,423
Guitar	13,343	16,805	5,275	35,423
Keyboard	8,508	42,645	3,838	54,991
Mallet	27,722	5,581	1,763	35,066
Organ	176	36,401	0	36,577
Reed	14,262	76	528	14,866
String	20,510	84	0	20,594
Synth	0	0	5,501	5,501
Vocal	3,925	140	6,688	10,753
Total	108,978	110,224	86,777	305,979

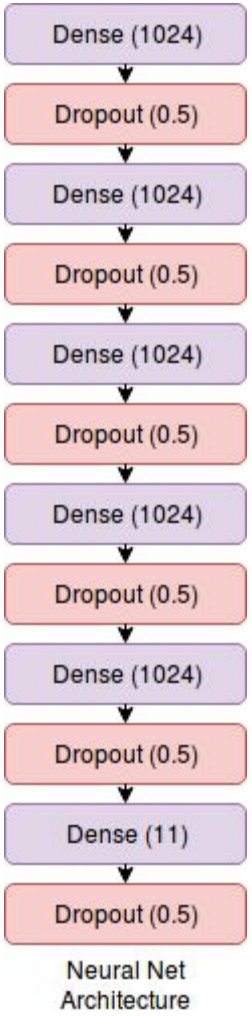
Neural Networks - Spencer Gass

Neural Net

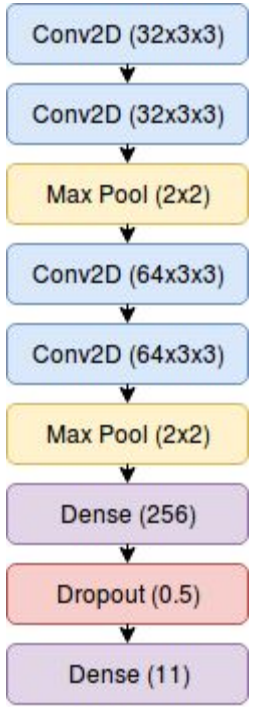
- Determined hyper parameters with a holdout set
- 38 million parameters
- 54.81% accuracy on test set

Convolutional Neural Net

- Adapted from VGGnet. Hyper parameters were refined with a holdout set.
- 600 thousand parameters
- 72.96% accuracy on test set



Neural Net Architecture



CNN Architecture

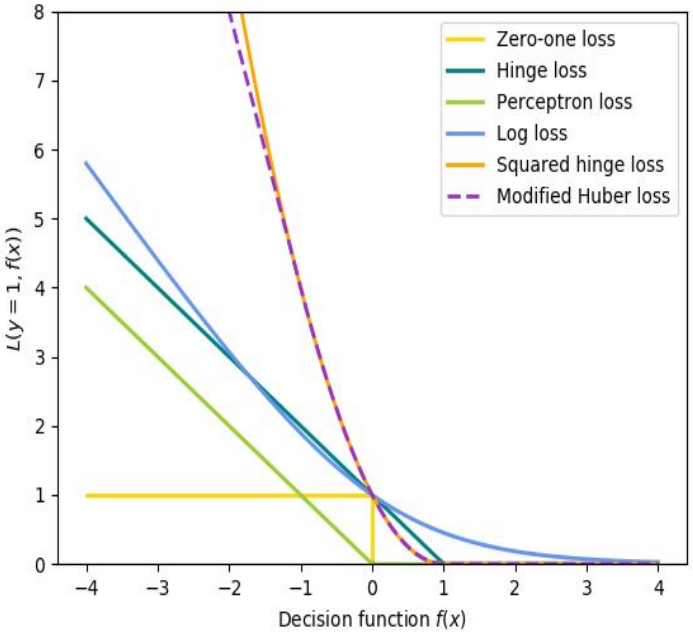
Linear Kernel SVM and SGD- Tim Coulter

Linear SVM

- Linear SVM trained on reduced training set (12,000 samples)
- 98% accuracy on the test set with Linear SVM

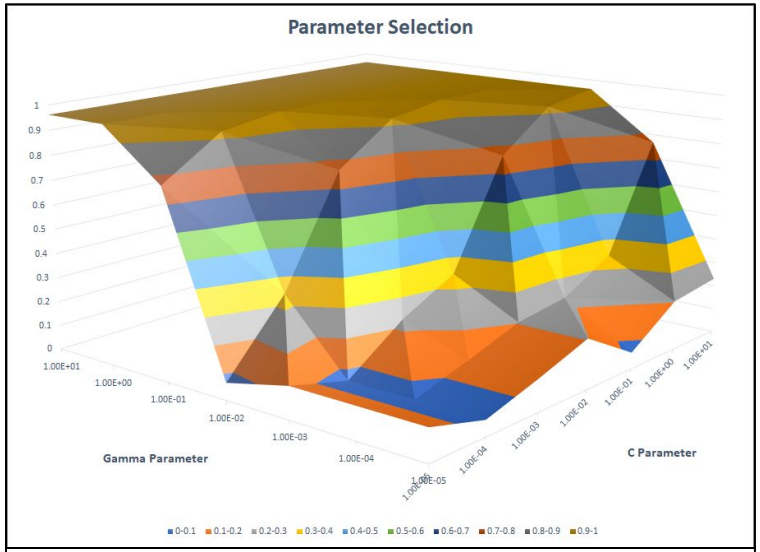
Stochastic Gradient Descent

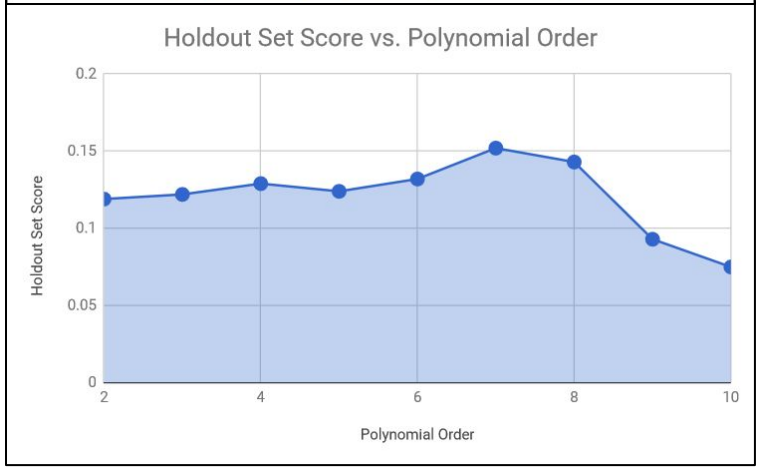
- Partial fitting to utilize whole training set
- Emulated linear SVM with hinge loss model
- 96% accuracy with reduced training set and 45% accuracy with full training set



Polynomial Kernel SVM
Eli Johnson

- No stochastic training tools available
  - Used subset of training data (12,000 points)
- First determined C and Gamma using holdout set
  - C = 1
  - Gamma = 1E-3
- Trained classifiers with 2nd-10th degree kernels
  - Evaluated score on second holdout set
  - 7th-order polynomial selected
- 89.1% score on test set





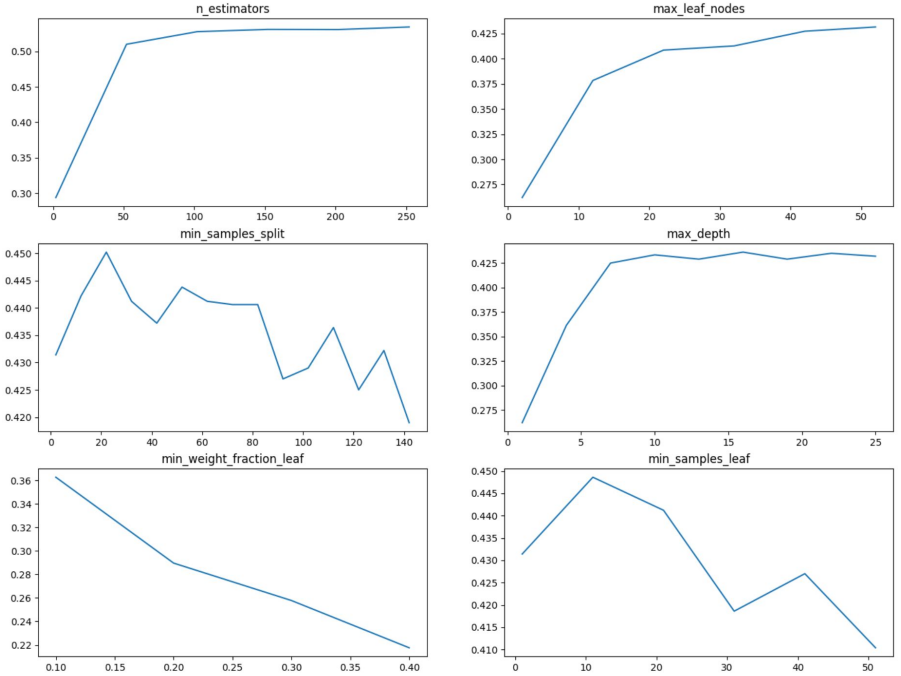
RBF Kernel SVM & Random Forest - Sam Tanner

RBF Kernel

- C and Gamma values determined using GridSearch on a holdout set
- Trained on reduced feature set
- 289,205 Samples
- %56.69 accuracy on test set

Random Forest

- Parameters graphed to limit range for GridSearch
- Trained on full feature set
- 289,205 Samples
- %29.74



Discussion

With more time we could extend to more challenging classification problems e.g. more instrument categories, note/velocity detection.

Conclusion: Linear classifiers can out performed more complex models even on a large, high dimensional data set.