**CS109B Final Project**

**Milestone 4**

**Group 26**

Dominick Delucia , Nina Iftikhar, Nisreen Shiban and Hany Bassily

Deep Learning Classifier for Movie Genres

**Purpose:**

Main objective of the Milestone is to build a deep learning classifier for movie genres. This part details the Convolutional net architecture adapted for this purpose

**Data preprocessing:**

Al poster images (9988) were transformed into numpy tensor with RGB decomposition. The images were cropped and rescaled with same aspect ratio to a reduced size of 85 x 128

Reduced tensors were transferred to an AWS EBS storage available to be mounted at any instance initiation. For this purpose, a private instance was implemented and shared among the group

**Selected Architectures**

All architectures were derived from the main architecture introduced in the lab and applied with the MINST dataset. Additional layer and rearrangement were done after few initial trials with runs that did not exceed 20 epochs which led to the first version

1. **Version 1:**

Data preprocessing:

- Channel rearrangement

- 14 labels multi label classification

- 20% validation

- Centered features

Main Architecture

- Multi layer CNN :

- Conv2D - Relu depth 32 and 7x7 kernel - He uniform initialization

- MaxPool 2x2

- Conv2D - Relu depth 32 and 5x5 kernel - He uniform initialization

- MaxPool 2x2

- Conv2D - Relu depth 16 and 3x3 kernel - He uniform initialization

- MaxPool 2x2

- Conv2D - Relu depth 16 and 3x3 kernel - He uniform initialization

- FC 128 , Relu , He uniform initialization

- FC 64 , Relu , He uniform initialization

- SGD optimzer with 1e-6 learning rate and 0.99 momentum

- Binary cross entropy loss function

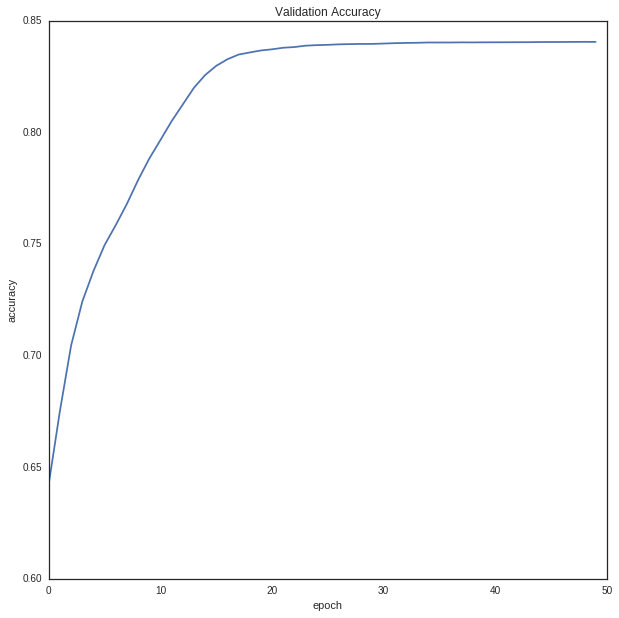
- Batch size 64

- Training convergence after 50 epochs

**Main Features:**

* Combination of batch size and learning rate proves to give reasonable training behavior
* Initialization introduced for consistency in training
* Loss function was selected to be Binary cross entropy to fit th multilabel classification problem

**Performance:**﻿﻿﻿﻿﻿



1. **Version 2:**

Data preprocessing:

- Channel rearrangement

- 14 labels multi label classification

- data augmentation by shift and zoom for 8000 training samples and 2000 test samples

- Centered features

Main Architecture

- Multi layer CNN :

- Conv2D - Relu depth 32 and 7x7 kernel - He uniform initialization

- MaxPool 2x2

- Conv2D - Relu depth 32 and 5x5 kernel - He uniform initialization

- MaxPool 2x2

- Conv2D - Relu depth 64 and 3x3 kernel - He uniform initialization

- MaxPool 2x2

- Conv2D - Relu depth 64 and 3x3 kernel - He uniform initialization

- FC 128 , Relu , He uniform initialization

- FC 64 , Relu , He uniform initialization

- SGD optimzer with 1e-6 learning rate and 0.99 momentum

- Binary cross entropy loss function

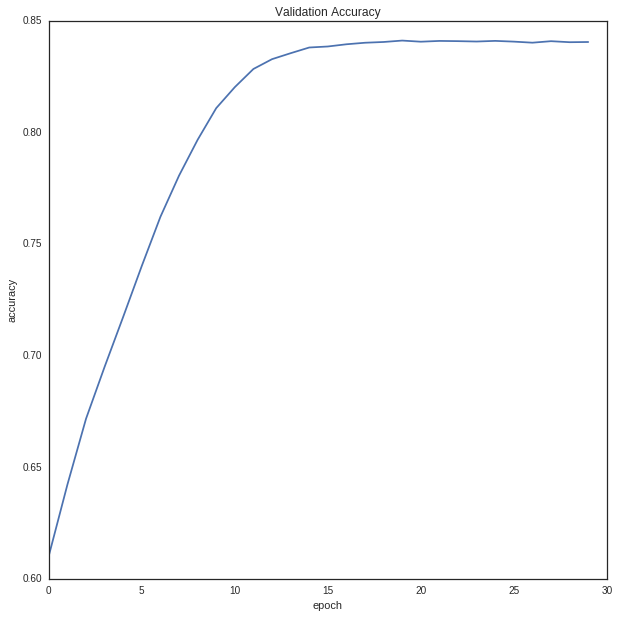
- Batch size 64

- Training convergence after 30 epochs

**Main Features:**

* Data augmentation was introduced to give more flexibility in increasing the complexity of the model without high risk of overfitting. Results were more depth in the conv layers

**Performance:**

﻿****

1. **Version 3:**

Data preprocessing:

- Channel rearrangement

- 14 labels multi label classification

- Data augmentation by shift and zoom for 8000 training sample and 2000 test sample

- Centered features

Main Architecture

- Multi layer CNN :

- Conv2D - Relu depth 32 and 7x7 kernel - He uniform initialization

- MaxPool 2x2

- Conv2D - Relu depth 32 and 5x5 kernel - He uniform initialization

- MaxPool 2x2

- Conv2D - Relu depth 64 and 3x3 kernel - He uniform initialization

- MaxPool 2x2

- Conv2D - Relu depth 64 and 3x3 kernel - He uniform initialization

- FC 128 , Relu , He uniform initialization

- FC 64 , Relu , He uniform initialization

- Adam optimzer with 1e-6 learning rate

- Binary cross entropy loss function

- Batch size 64

- Training convergence after 30 epochs

**Main Features:**

* This version is to test the Adam optimization which is clearly did not contribute a lot
* Decision point: Current architecture with the SGD optimization is basis for any other variation

**Performance:**

* Figure not available. Decay is not consistent

1. **Version 4:**

Data preprocessing:

- Channel rearrangement

- 14 labels multi label classification

- data augmentation by shift and zoom for 8000 training samples and 2000 test samples

- Centered features

Main Architecture

- Multi layer CNN with Batch Normalization Layers :

- Conv2D - Relu depth 32 and 7x7 kernel - He uniform initialization

- Batch Norm

- MaxPool 2x2

- Conv2D - Relu depth 32 and 5x5 kernel - He uniform initialization

- Batch Norm

- MaxPool 2x2

- Conv2D - Relu depth 64 and 3x3 kernel - He uniform initialization

- Batch Norm

- MaxPool 2x2

- Conv2D - Relu depth 64 and 3x3 kernel - He uniform initialization

- Batch Norm

- FC 128 , Relu , He uniform initialization

- Batch Norm

- FC 64 , Relu , He uniform initialization

- Batch Norm

- SGD optimzer with 1e-4 learning rate and 0.99 momentum

- Binary cross entropy loss function

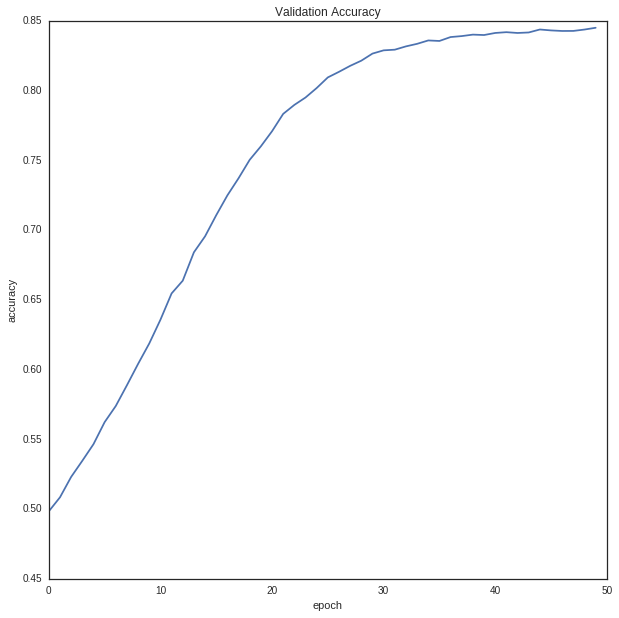
- Batch size 128

- Training convergence after 50 epochs

**Main Features:**

* Batch normalization was introduced after each ReLu activated layer
* Batch increased to improve statistical inference

**Performance:**

****

1. **Version 5:**

Data preprocessing:

- Channel rearrangement

- 14 labels multi label classification

- data augmentation by shift and zoom for 8000 training samples and 2000 test samples

- Centered features

Main Architecture

- Multi layer CNN with Batch Normalization Layers and Dropout :

- Conv2D - Relu depth 32 and 7x7 kernel - He uniform initialization

- 30 % Dropout

- Batch Norm

- MaxPool 2x2

- Conv2D - Relu depth 32 and 5x5 kernel - He uniform initialization

- 30 % Dropout

- Batch Norm

- MaxPool 2x2

- Conv2D - Relu depth 64 and 3x3 kernel - He uniform initialization

- 30 % Dropout

- Batch Norm

- MaxPool 2x2

- Conv2D - Relu depth 64 and 3x3 kernel - He uniform initialization

- 30 % Dropout

- Batch Norm

- FC 128 , Relu , He uniform initialization

- 50 % Dropout

- Batch Norm

- FC 64 , Relu , He uniform initialization

- 50 % Dropout

- Batch Norm

- SGD optimzer with 1e-4 learning rate and 0.99 momentum

- Binary cross entropy loss function

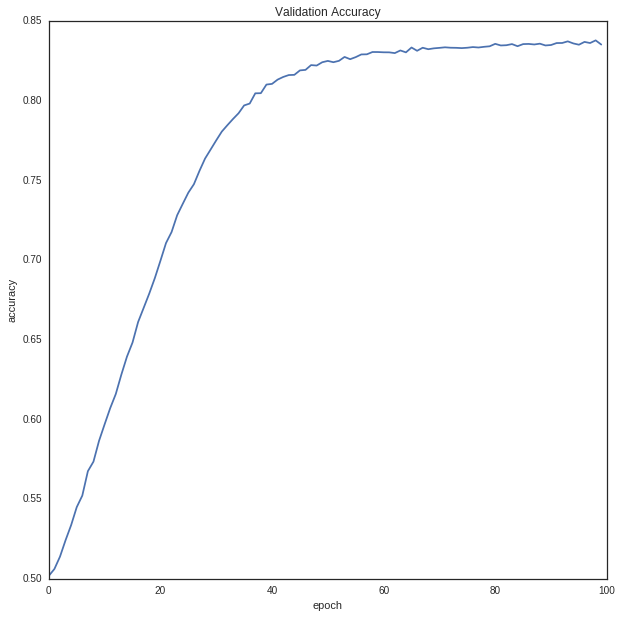
- Batch size 128

- Training convergence after 100 epochs

**Main Features:**

* Aggressive dropout ratios were chosen to test the effect – no change in accuracy
* Training was extended to 100 epochs to account for the drop out

**Performance:**

****

1. **Version 6:**

Data preprocessing:

- Channel rearrangement

- 14 labels multi label classification

- data augmentation by shift and zoom for 8000 training samples and 2000 test samples

- Centered features

Main Architecture

- Multi layer CNN with Batch Normalization Layers and Dropout and 0.1 L2 Kernel Regularization

- Conv2D - Relu depth 32 and 7x7 kernel - He uniform initialization - 0.1 L2 Kernel Regularization

- 30 % Dropout

- Batch Norm

- MaxPool 2x2

- Conv2D - Relu depth 32 and 5x5 kernel - He uniform initialization - 0.1 L2 Kernel Regularization

- 30 % Dropout

- Batch Norm

- MaxPool 2x2

- Conv2D - Relu depth 64 and 3x3 kernel - He uniform initialization - 0.1 L2 Kernel Regularization

- 30 % Dropout

- Batch Norm

- MaxPool 2x2

- Conv2D - Relu depth 64 and 3x3 kernel - He uniform initialization - 0.1 L2 Kernel Regularization

- 30 % Dropout

- Batch Norm

- FC 128 , Relu , He uniform initialization - 0.1 L2 Kernel Regularization

- 50 % Dropout

- Batch Norm

- FC 64 , Relu , He uniform initialization - 0.1 L2 Kernel Regularization

- 50 % Dropout

- Batch Norm

- SGD optimzer with 1e-4 learning rate and 0.99 momentum

- Binary cross entropy loss function

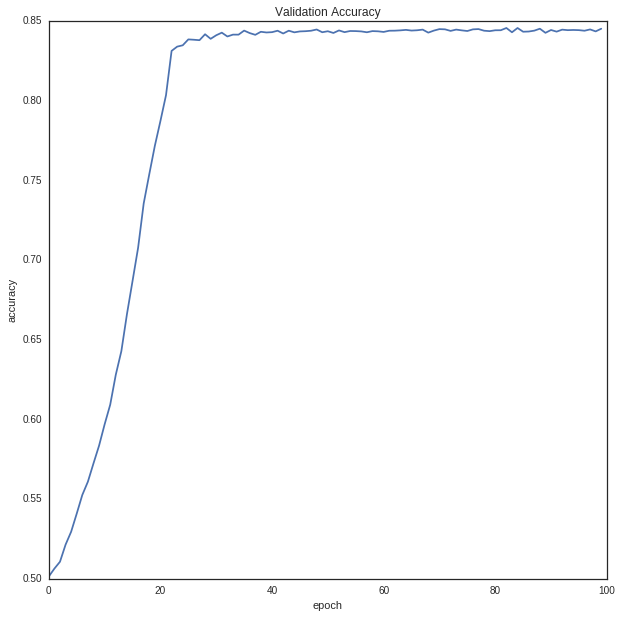
- Batch size 128

- Training convergence after 100 epochs

**Main Features:**

* Two values of the L2 regularizer were selected 0.01 and 0.1 and no difference in the overall accuracy was noted
* Similar training approach was applied with faster training convergence as an effect of restricting the weights freedom to adapt during the backpropagation adaptation
* Slower training rate could be an improvement for this version
* This is the main objective of a cross validation effort that will be subsequently implemented

**Performance:**

****

**Final Conclusion:**

* No improvement was observed when trying to adapt all the previous versions and the accuracy remained at about 84% on the validation/test set
* This could be an artifact of the data
* Some improvement to be considered is to merge other features with the image features, perform cross validation and try other architecture less deeper but different initializations