Lightweight Probabilistic Deep Networks

Team Members:

- Eashan Gupta 160050045
- Ankit 160050046
- Kumar Saurav 160050057

Paper Used:

Lightweight Probabilistic Deep Networks

Results:

1. **MNIST** Database

Mode:	Accuracy:
Simple Softmax using conventional CNN	97.13
Using probabilistic output layer and only Dirichlet categorical classifier	97.57
Using ADF and Dirichlet distribution	97.58

2. **CIFAR10** database

Mode:	Accuracy:
Simple Softmax using conventional CNN	53.71
Using probabilistic output layer and only Dirichlet categorical classifier	51.43
Using ADF and Dirichlet distribution	33.89

Basic Overview:

$$\mu_z^{(i)} = E_{q(z^{(i-1)})}[f^{(i)}(z^{(i-1)}; \theta^{(i)})]$$
$$v_z^{(i)} = V_{q(z^{(i-1)})}[f^{(i)}(z^{(i-1)}; \theta^{(i)})]$$

Max-Pooling:

Use the formulas to derive the formulae.

For

$$C=\max(A,B)$$

the formula comes out to be

$$\mu_C = \sqrt{\sigma_A^2 + \sigma_B^2} \cdot \phi(\alpha) + (\mu_A - \mu_B) \cdot \Phi(\alpha) + \mu_B$$

$$v_c = (\mu_A + \mu_B)\sqrt{\sigma_A^2 + \sigma_B^2} \cdot \phi(\alpha) + (\mu_A^2 + v_A) \cdot \Phi(\alpha) + (\mu_B^2 + v_B) \cdot (1 - \Phi(\alpha)) - \mu_C^2$$

Where

$$\alpha = \frac{(\mu_A - \mu_B)}{\sqrt{\sigma_A^2 + \sigma_B^2}}$$

Relu:

Y= max(0,X)

Given X is a Gaussian random variable with mean μ and standard deviation σ, the mean and variance of Y are given as follows:

$$\mu_{relu} = \mu . \Phi(\mu/\sigma) + \sigma . \phi(\mu/\sigma)$$
$$\nu_{relu} = (\mu * \mu + \nu) . \Phi(\mu/\sigma) + \mu \sigma . \phi(\mu/\sigma) - \mu_{relu} * \mu_{relu}$$

Φ , ϕ represents cdf and pdf of standard normal distribution respectively.

This is an activation layer used in the convolution neural nets and this formula was used for passing Relu nonlinearity in the means and variances of next layer based on previous layers.

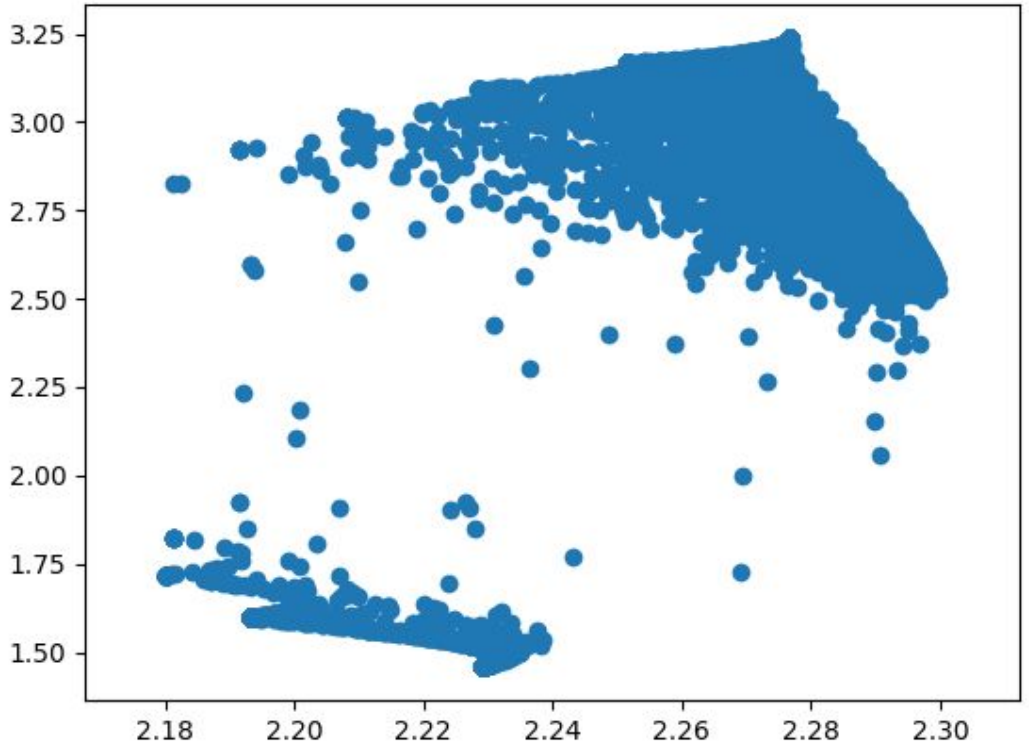

Figure 1: ADF- Cross-entropy vs Entropy of MNIST data using Dirichlet Distribution and ADF CNN

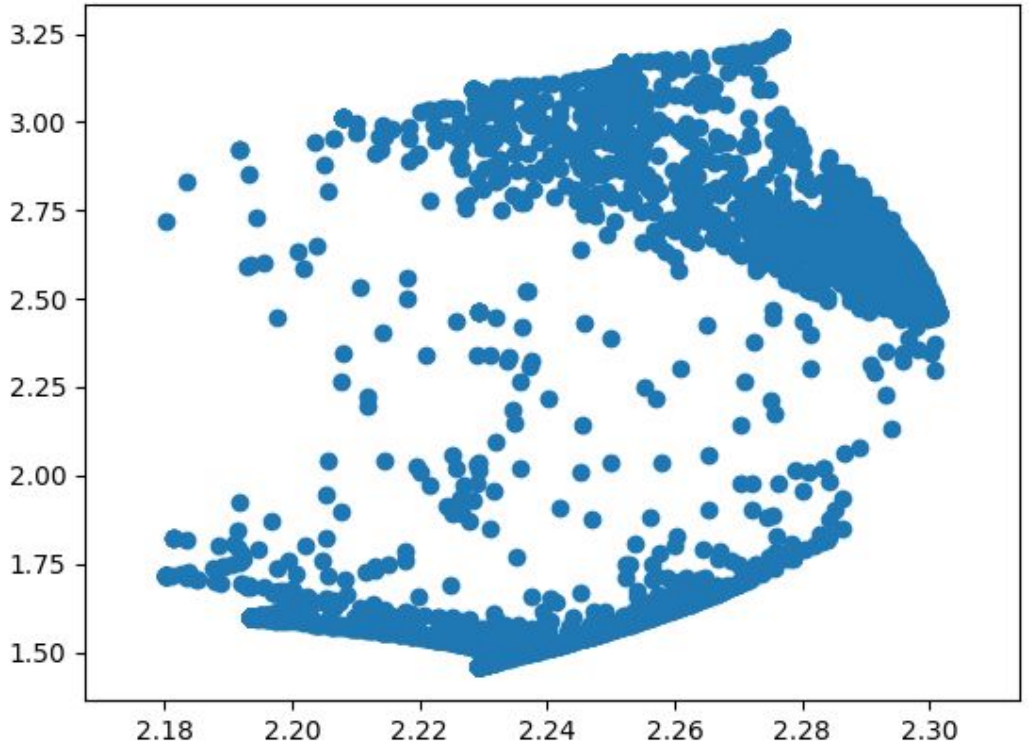

Figure 2: ProbOut- Cross-entropy vs Entropy of MNIST data using Dirichlet Distribution and Simple CNN outputs

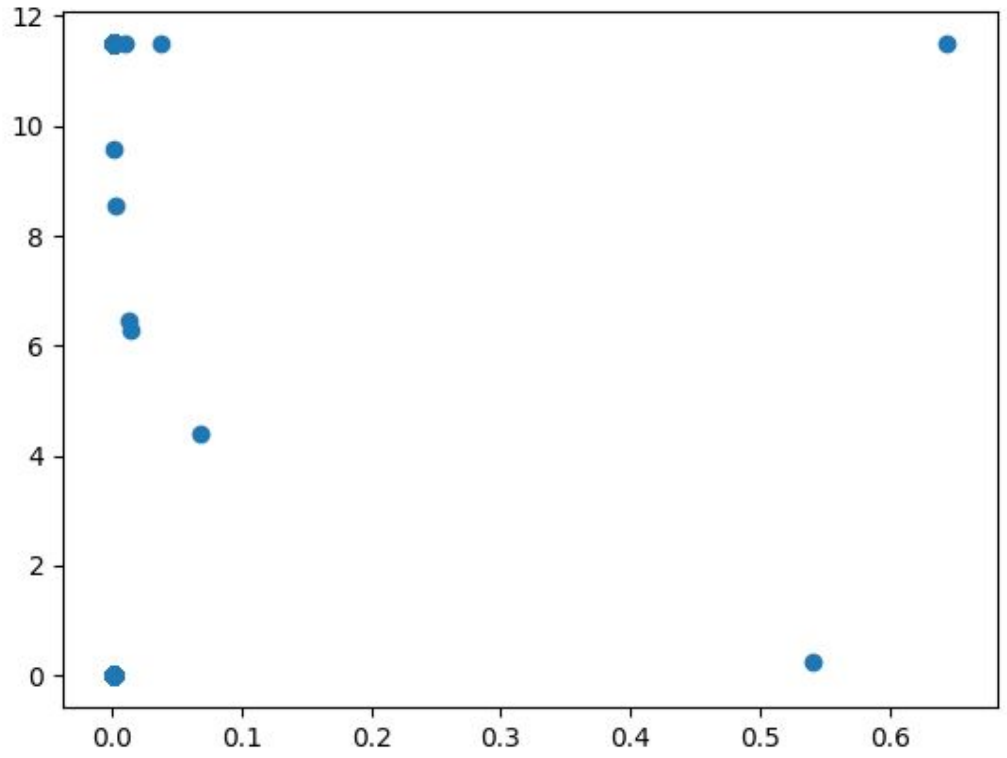

Figure 3: Softmax- Cross-entropy vs Entropy of MNIST data using Softmax categorical classification

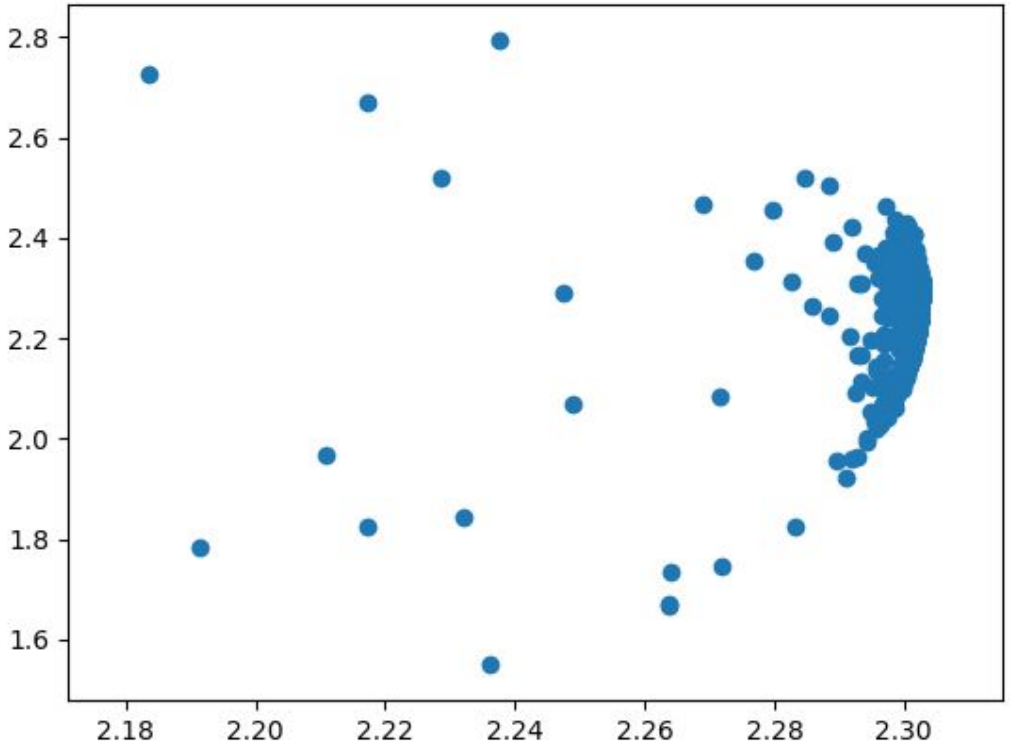

Figure 4: ADF- Cross-entropy vs Entropy of CIFAR10 data using Dirichlet Distribution and ADF CNN

```
2018-11-23 04:52:13.618627: E tensorflow/core/grappler/optimizers/dependency_optimizer.cc:666] Iteration = 1, topological sort failed with message: The graph couldn't be sort
32768/32768 [==============================] - 380s 12ms/step - loss: -72.3974 - custom_metric: 0.7429 - val_loss: -74.3977 - val_custom_metric: 0.9521
Epoch 2/10
32768/32768 [==============================] - 385s 12ms/step - loss: -73.6696 - custom_metric: 0.9266 - val_loss: -74.7012 - val_custom_metric: 0.9647
Epoch 3/10
32768/32768 [==============================] - 360s 11ms/step - loss: -73.8849 - custom_metric: 0.9448 - val_loss: -74.8565 - val_custom_metric: 0.9697
Epoch 4/10
32768/32768 [==============================] - 361s 11ms/step - loss: -73.9631 - custom_metric: 0.9515 - val_loss: -74.9011 - val_custom_metric: 0.9702
Epoch 5/10
32768/32768 [==============================] - 367s 11ms/step - loss: -74.0324 - custom_metric: 0.9561 - val_loss: -74.9053 - val_custom_metric: 0.9727
Epoch 6/10
32768/32768 [==============================] - 374s 11ms/step - loss: -74.0784 - custom_metric: 0.9608 - val_loss: -74.9583 - val_custom_metric: 0.9740
Epoch 7/10
32768/32768 [==============================] - 384s 12ms/step - loss: -74.1211 - custom_metric: 0.9615 - val_loss: -74.9691 - val_custom_metric: 0.9751
Epoch 8/10
32768/32768 [==============================] - 362s 11ms/step - loss: -74.1422 - custom_metric: 0.9637 - val_loss: -75.0001 - val_custom_metric: 0.9744
Epoch 9/10
32768/32768 [==============================] - 361s 11ms/step - loss: -74.1478 - custom_metric: 0.9634 - val_loss: -74.9406 - val_custom_metric: 0.9751
Epoch 10/10
32768/32768 [==============================] - 358s 11ms/step - loss: -74.1776 - custom_metric: 0.9645 - val_loss: -75.0233 - val_custom_metric: 0.9758
```

Figure 5: Output for ADF for MNIST database using epoch=10

```
2018-11-23 06:57:58.466898: E tensorflow/core/grappler/optimizers/dependency_optimizer.cc:666] Iteration = 1, topological sort failed with message: The graph couldn't be sort
16384/16384 [==============================] - 248s 15ms/step - loss: -70.0563 - custom_metric: 0.1639 - val_loss: -70.4023 - val_custom_metric: 0.2979
Epoch 2/10
16384/16384 [==============================] - 266s 16ms/step - loss: -70.5706 - custom_metric: 0.2397 - val_loss: -71.0727 - val_custom_metric: 0.3018
Epoch 3/10
16384/16384 [==============================] - 241s 15ms/step - loss: -71.4947 - custom_metric: 0.2522 - val_loss: -72.4384 - val_custom_metric: 0.3105
Epoch 4/10
16384/16384 [==============================] - 235s 14ms/step - loss: -72.8315 - custom_metric: 0.1971 - val_loss: -73.0903 - val_custom_metric: 0.1982
Epoch 5/10
16384/16384 [==============================] - 236s 14ms/step - loss: -73.0894 - custom_metric: 0.2083 - val_loss: -73.1254 - val_custom_metric: 0.2539
Epoch 6/10
16384/16384 [==============================] - 235s 14ms/step - loss: -73.1136 - custom_metric: 0.2327 - val_loss: -73.1388 - val_custom_metric: 0.3115
Epoch 7/10
16384/16384 [==============================] - 240s 15ms/step - loss: -73.1274 - custom_metric: 0.2516 - val_loss: -73.1545 - val_custom_metric: 0.3105
Epoch 8/10
16384/16384 [==============================] - 246s 15ms/step - loss: -73.1402 - custom_metric: 0.2580 - val_loss: -73.1530 - val_custom_metric: 0.3428
Epoch 9/10
16384/16384 [==============================] - 240s 15ms/step - loss: -73.1437 - custom_metric: 0.2617 - val_loss: -73.1557 - val_custom_metric: 0.3145
Epoch 10/10
16384/16384 [==============================] - 239s 15ms/step - loss: -73.1497 - custom_metric: 0.2700 - val_loss: -73.1580 - val_custom_metric: 0.3389
>>>
```

Figure 6: ADF output for CIFAR10 database using epoch=10