CSO 211 PROJECT

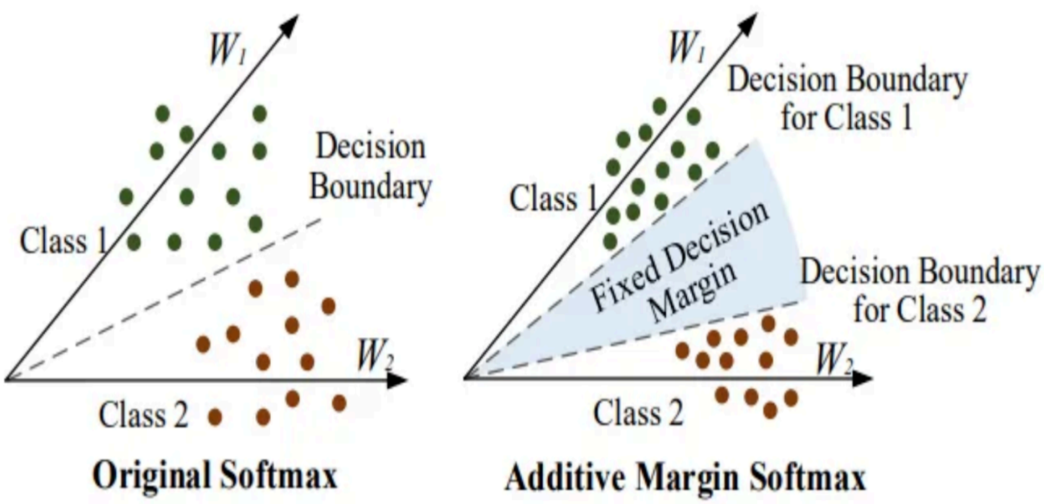
Anshika Agarwal
22124009

Enhanced AM softmax loss

AM SoftMax loss

- The Additive Margin Softmax (AM-Softmax) loss is an extension of the traditional softmax function used in neural networks, primarily in tasks like face recognition, where creating well-separated feature representations is crucial.
- AM-Softmax introduces an additional margin to the standard softmax function, aiming to enhance the angular margin between classes in the learned feature space.
- The AM-Softmax loss formula includes an angular margin term, which adds a specific margin angle to the standard softmax loss.
- This margin aims to increase the separation between classes, improving the network's ability to distinguish between different categories, especially in scenarios where inter-class variations are significant, such as facial recognition systems.

☐ By employing the AM-Softmax loss during training, neural networks not only learn to classify inputs but also create feature representations that are more robust and distinctive, enabling better performance in tasks with high inter-class variability.



Original AM Softmax Function:

$$\mathcal{L}_{AMS} = -\frac{1}{n}\sum_{i=1}^{n} log \frac{e^{s\cdot\left(cos\theta_{y_i} - m\right)}}{e^{s\cdot\left(cos\theta_{y_i} - m\right)} + \sum_{j=1, j\neq y_i}^{c} e^{s\cdot cos\theta_j}}$$

$$= -\frac{1}{n}\sum_{i=1}^{n} log \frac{e^{s\cdot\left(W_{y_i}^T \boldsymbol{f}_i - m\right)}}{e^{s\cdot\left(W_{y_i}^T \boldsymbol{f}_i - m\right)} + \sum_{j=1, j\neq y_i}^{c} e^{s W_j^T \boldsymbol{f}_i}}$$

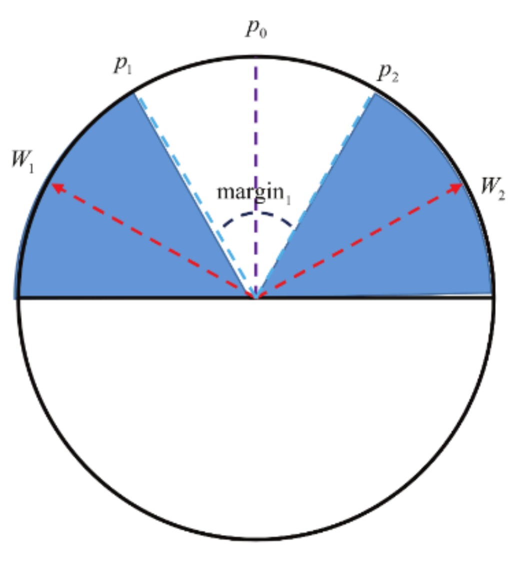
Changes and their advantage over loss function:

1.Additional margin:

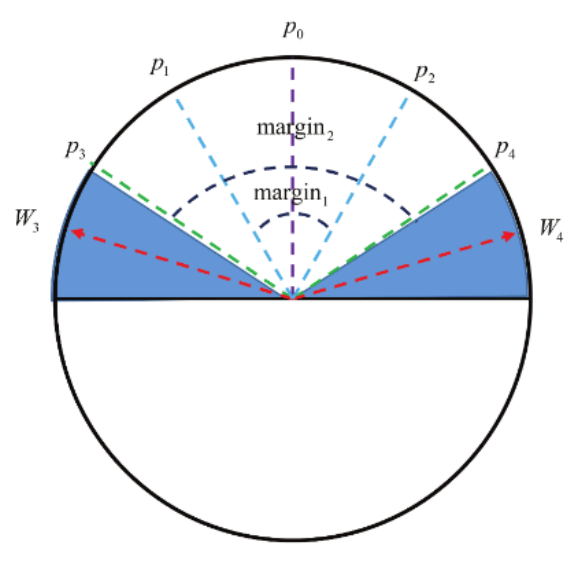
Change 1:
y_true=y_true_margin

Advantage:

- ☐ Simultaneous Improvement: It enhances both intra-class compactness and inter-class separability concurrently.
- ☐ Better Feature Separation: By introducing an additive margin to both intra-class and inter-class angular variations, promotes increased separation between different classes.



AM SoftMax Loss



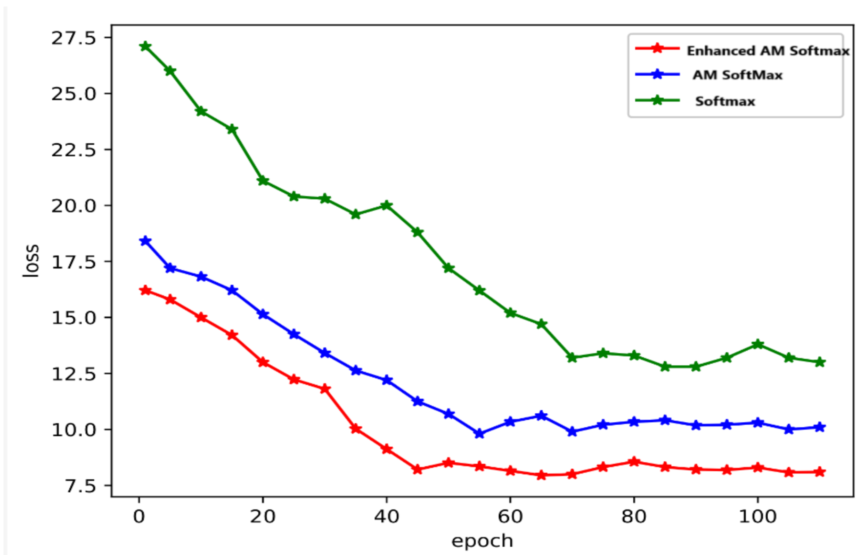
Enhanced AM SoftMax Loss

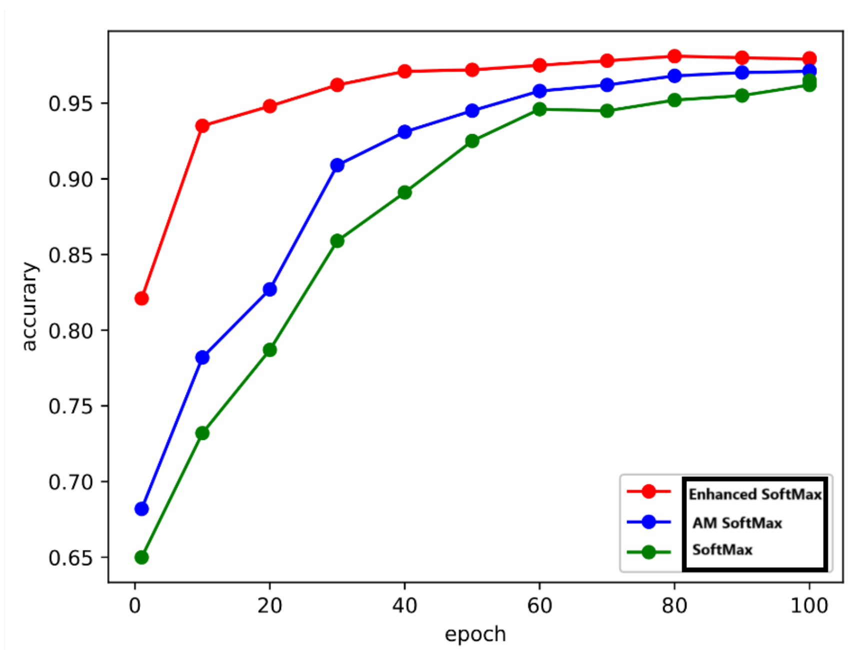
2.Exponential scaling factor:

Advantage:

Exponential scaling factor to manage loss:

- ☐ Loss reduction: It Reduces the cost function resulting in decreased loss.
- ☐ Enhances Compactness: reduces gap between interclass elements resulting in compact class behavior. And increased accuracy.





Change 2:

Introducing a new factor M,

M=e^m

```
y_pred = y_true * (y_pred - margin)/M + (1 - y_true) * y_pred
```

New function:

$$L_{a??} = -\frac{1}{n}\sum_{i=1}^{n} \frac{e^{\frac{s.(cos\theta yi - m)}{M}}}{e^{\frac{s.(cos\theta yi - m)}{M}} + \sum_{j=1,\, j\neq y_i}^{n} e^{s.(cos\theta_{Y} + m)}}$$

Conclusion

In conclusion,Enhanced AM Softmax is a better version of AM Softmax since by introducing the additive margin the inter-class separation is increased ,also the exponential factor helped to reduce loss and hence it's easy to study classes which proved to give the best increase in the model performance.

The Enhanced AM Softmax has a clearer geometrical explanation and can obtain highly discriminative features for face recognition. Extensive experimental evaluation of several recent state-of-the-art softmax loss functions are conducted on the relevant face recognition benchmarks, CASIA-Webface, LFW, CALFW, CPLFW, and CFP-FP. We show that the proposed loss function consistently outperforms the state-of-the-art.