

ENPM703- Assignment-2

Part4: Dropout

Dropout

Dropout is a regularization method used to avoid overfitting, where the model fits the training data too well. By randomly turning off certain neurons during training, dropout helps the model learn better features that generalize well, improving its ability to handle unseen data.

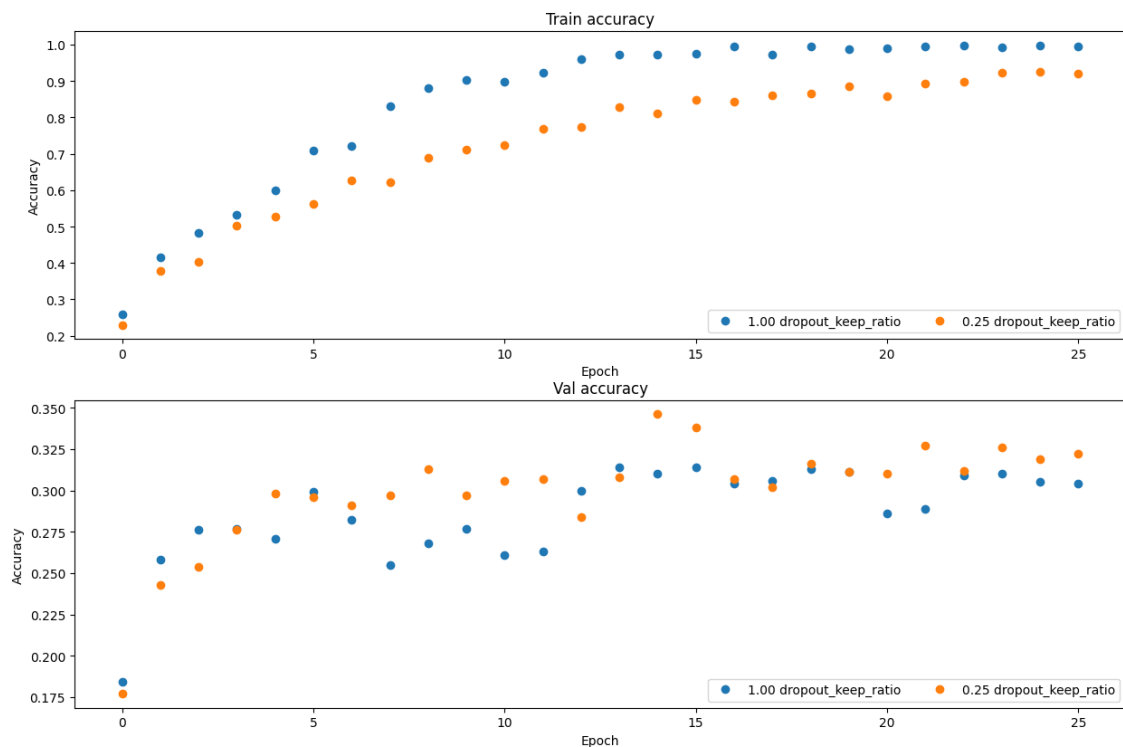
How Dropout Works

During training, dropout randomly deactivates some neurons in a layer with a certain probability $1-p$, where p is the keep probability. This means only part of the neurons contribute to the forward pass, making the network less dependent on specific neurons.

Vanila vs Inverted Dropout

- **Vanilla dropout** is the basic form of dropout where, during training, some neurons in a layer are randomly set to zero based on the dropout probability $1-p$. In vanilla dropout, the output values of the active neurons aren't changed during training. During prediction, the outputs need to be scaled by p to match what the model saw during training.
- **Inverted dropout** improves on vanilla dropout by scaling the active neurons' outputs by $1/p$ during training. This keeps the output level consistent whether or not dropout is applied. During prediction, no scaling or dropout is needed because the training already adjusted the outputs. This makes implementation easier and quicker during prediction.

Regularization Experiment



Without dropout, keeping the ratio at 1.0, the model quickly reaches high training accuracy, close to 1.0. However, the validation accuracy is relatively low and increases slowly, indicating overfitting. The model learns the training data well but struggles to generalize to new data.

With dropout, keeping the ratio at 0.25, training accuracy increases more slowly and remains lower, as dropout encourages the model to depend less on specific neurons, preventing overfitting. The validation accuracy is generally higher than without dropout, showing that the model generalizes better to new data when dropout is used, even if training accuracy is lower.

In summary, dropout enhances generalization by reducing overfitting, as shown by higher validation accuracy despite lower training accuracy.

Effects of Dropout

- Dropout reduces overfitting by forcing the network to learn using different paths, making the model more robust and less dependent on individual neurons. This encourages redundancy and prevents neurons from co-adapting.
- However, if the dropout rate is too high (low p), the model may underfit, meaning it won't learn effectively. It's important to tune p to find the right balance between reducing overfitting and keeping enough learning capacity.