



Improving Deep Neural Networks

Video 8: Model Saving using PyTorch & Hands-On

Recap: Techniques to improve performance

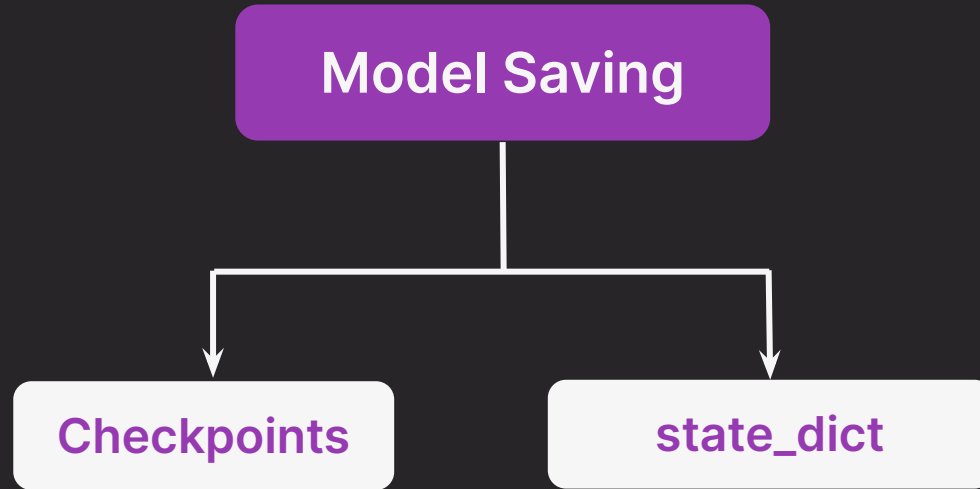
Up Next: Model Saving & Deployment

Model Saving in PyTorch

Model saving sometimes
takes weeks!!



Model Saving in PyTorch



Model Saving in PyTorch

1) Using Checkpoints

- Captures intermediates states during training
- **Components:** Parameters, No. of epochs, and optimizer parameters
- Facilitates continuous training from the specific epochs



Model Saving in PyTorch

1) Using Checkpoints

- Captures intermediates states during training
- **Components:** Parameters, No. of epochs, and optimizer parameters
- Facilitates continuous training from the specific epochs

`torch.save()`

To save checkpoints

`load_ckpt()`

To load checkpoints



Model Saving in PyTorch

2) Using state_dict

- Concise representation of the model parameters
- **Components:** Layer weights and biases, and learnable parameters



Choice of Model Saving

Factors to be considered:

- Complexity of the model
- Requirements for training resumption
- Need for Additional training metadata
- Deployment environment



Model Deployment in PyTorch

Tools to streamline model deployment:

- Torchserve
- **Gradio**
- Streamlit



Gradio

Key Features:

- Demo applications shared via URL
- **Pre-built input components:** Text boxes, sliders and image uploaders



Up-Next: Model Saving and Deployment Hands-On