

Early Training Dynamics of MLPs with Varying Depth

