* "Understanding the Loss Surface of Neural Networks for Binary Classification" by Zhang et al. (2018)
  + This paper examines the loss surface of neural networks during training and analyzes how weights change during optimization.
* "The Marginal Value of Adaptive Gradient Methods in Machine Learning" by Reddi et al. (2019)
  + This paper investigates the dynamics of adaptive gradient-based optimization algorithms like Adam and discusses how they affect weight updates during training.
* "Characterizing signal propagation to close the performance gap in unnormalized ResNets" by Sedghi et al. (2019)
  + This paper focuses on the weight dynamics in residual networks (ResNets) and discusses the importance of signal propagation in deep neural networks.
* "On the Convergence and Robustness of Training Neural Networks" by Li et al. (2019)
  + This paper explores the convergence behavior and robustness of training deep neural networks, shedding light on weight-changing dynamics.
* "Measuring the Intrinsic Dimension of Objective Landscapes" by Li et al. (2018)
  + This paper proposes a method to measure the intrinsic dimension of the loss landscapes in neural networks, providing insights into the complexity of weight changes during training.
* "Understanding Training Dynamics of Deep Learning Systems via Transfer Matrix" by Yang et al. (2019)
  + This paper presents a theoretical framework for understanding the training dynamics of deep learning systems, including weight changes.
* "Weight Agnostic Neural Networks" by Gaier and Ha (2019)
  + While this paper primarily focuses on weight-agnostic networks, it discusses how weights evolve and affect model behavior during training.
* "On the Learning Dynamics of Deep Neural Networks" by Du et al. (2019)
  + This paper offers insights into the learning dynamics of deep neural networks, including weight updates and convergence properties.
* "Understanding Deep Learning Requires Rethinking Generalization" by Zhang et al. (2017)
  + This paper investigates the role of over-parameterization and optimization dynamics in deep learning, shedding light on how weights change during training.
* "The Marginal Value of Adaptive Gradient Methods in Machine Learning" by Reddi et al. (2019)
  + This paper explores the dynamics of adaptive gradient-based optimization algorithms like Adam, providing insights into how these methods affect weight updates during training.
* "A Gentle Introduction to Optimization" by Zhang (2019)
  + This paper provides an accessible introduction to optimization algorithms commonly used in deep learning and discusses their impact on weight dynamics.
* "The Lottery Ticket Hypothesis: Finding Sparse, Trainable Neural Networks" by Frankle and Carbin (2019)
  + This paper introduces the idea of "winning lottery tickets" in neural networks and explores the dynamics of weight pruning and training.
* "Characterizing signal propagation to close the performance gap in unnormalized ResNets" by Sedghi et al. (2019)
  + This paper focuses on the weight dynamics in residual networks (ResNets) and discusses the importance of signal propagation in deep neural networks.
* "Understanding Training Dynamics of Deep Neural Networks" by Du et al. (2019)
  + This paper offers insights into the learning dynamics of deep neural networks, including weight updates and convergence properties.
* "Measuring the Intrinsic Dimension of Objective Landscapes" by Li et al. (2018)
  + This paper proposes a method to measure the intrinsic dimension of the loss landscapes in neural networks, providing insights into the complexity of weight changes during training.
* "Weight Agnostic Neural Networks" by Gaier and Ha (2019)
  + While this paper primarily focuses on weight-agnostic networks, it discusses how weights evolve and affect model behavior during training.
* "The Information Bottleneck Theory of Deep Learning" by Tishby et al. (2015)
  + This influential paper explores the information bottleneck theory and how it relates to the dynamics of weight changes during deep learning.