Anti-Distillation or Teacher Learning: Knowledge Transfer from Simple Model to a Complex One

A Preprint

Kseniia Petrushina
Moscow Institute of Physics and Technology
Dolgoprudny, Russia
petrushina.ke@phystech.edu

Andreii Hraboviy
Moscow Institute of Physics and Technology
Dolgoprudny, Russia
grabovoy.av@phystech.edu

Oleg Bakhteev
Moscow Institute of Physics and Technology
Dolgoprudny, Russia
bakhteev@phystech.edu

Abstract

This paper considers the problem of adapting the model to a new data with a large amount of information. We propose to build a model of greater complexity with further knowledge transfer from a simple model to it. It is necessary to take into account not only the quality of the prediction on the original samples, but also the adaptability to novel data and the robustness of the obtained solution. The novelty of the work lies in the fact that our method allows to adapt the model to the increase of complexity of the training sample without losing quality. This study considers both probabilistic and algebraic methods for obtaining a new model.

Keywords Anti-Distillation · Distillation · Knowledge Transfer · Weight Initialization

1 Introduction