

مروری بر روش های هرس در شبکه های عصبی عمیق با تاکید بر روش های هرس پیش از آموزش

عطیه فیروزه^{۱*}، مرتضی محجل^۲، محبوبه شمس^۳
^۱دانشجوی کارشناسی ارشد، دانشکده برق و کامپیوتر، صنعتی قم، قم، ایران.
^۲استادیار دانشکده برق و کامپیوتر، صنعتی قم، قم، ایران.

چکیده

با گسترش کاربرد شبکه های عصبی، عمیق شدن و افزایش پارامترهای شبکه، در عین حال محدودیت منابع محاسباتی، محدودیت در حافظه و تفسیرناپذیر شدن این شبکه ها، فشرده سازی شبکه های عصبی مورد توجه قرار گرفته است. فشرده سازی می بایست به صورت هوشمندانه باشد، به نحوی که ما را از مزایای بهره مندی از شبکه های عصبی عمیق جدا نکند. هرس به عنوان یکی از روش های فشرده سازی با حذف پارامترهای غیرضروری شبکه، همواره مورد اقبال پژوهشگران بوده است، به نحوی که در پژوهش های اخیر سعی شده است، مرحله ای با عنوان هرس پیش از آموزش شبکه، در مراحل قبل از راه اندازی شبکه گنجانده شود تا از مزایای فشرده سازی و هرس در مراحل آموزش و استنتاج شبکه بهره برده شود. در مقاله پیش رو سعی شده است، مروری بر روش های هرس با تاکید بر هرس های پیش از آموزش شبکه انجام شود. در ابتدا مبانی هرس مطرح شده، سپس انواع هرس به همراه تعریف ریاضی هریک مطرح و در نهایت بررسی دقیق تری روی هرس های پیش از آموزش شبکه انجام شده است.

کلمات کلیدی: شبکه عصبی عمیق، فشرده سازی شبکه عصبی، فرضیه بلیط بخت آزمایی، هرس پیش از راه اندازی شبکه.

A Review on Pruning Techniques in Deep Neural Networks with Emphasis on Prune at Initial

Atieh firoozeh*, Dr.M.Mohajel, Dr.M.Shamsi
Qom University of Technology, Qom, Iran.

Abstract

With the expansion of the use of neural networks, increase deep of NN, and increase of network parameters, as well as the limitation of computational resources, the limitation of memory, and the incomprehensibility of these networks, the compression of neural networks is necessary. Compression must be intelligent, so as not to deprive us of the benefits of deep neural networks. Pruning is one of the compression methods that eliminate unnecessary network parameters. in recent research, Pruning has always been favored by researchers as far as a step called pruning at the initial design that pruned the initial network to include the benefits compression and pruning in the training and inference. this article reviews pruning techniques in deep neural networks with emphasis on Prune at initializing. First, the basics of pruning are discussed, then the types of pruning with the mathematical definition of each discussed, and finally, a more detailed study of pruning before network training has been done.

Keywords: pruning NN, Sparsification NN, Compereson NN, LTH.

تاریخچه مقاله:

تاریخ ارسال: ۱۴۰۰/۰۲/۲۷

تاریخ اصلاحات: ۱۴۰۰/۰۳/۱۷

تاریخ پذیرش: ۱۴۰۰/۰۶/۰۲

تاریخ انتشار: ۱۴۰۰/۰۶/۳۱

Keywords:

Pruning NN,
Sparsification NN,
Compereson NN,
LTH

*نویسنده مسئول:

Firoozeh.a@qut.ac.ir