

الگوریتمی بر اساس اتوماتان‌های یادگیر برای تعیین حداقل تعداد واحدهای مخفی در یک شبکه عصبی سه لایه

محمدرضا میبیدی
دانشیار

حمید بیگی
دانشجوی دکترا

دانشکده مهندسی کامپیوتر، دانشگاه صنعتی امیرکبیر

چکیده

یکی از مسائل حل نشده در شبکه‌های عصبی سه لایه انتخاب ساختار بهینه (ساختار با حداقل تعداد واحدهای لایه میانی) برای شبکه می باشد. در شبکه‌های عصبی سه لایه، در ابتدای آموزش، یک ساختار مشخص برای شبکه انتخاب و سپس شبکه آموزش داده می شود. کارایی الگوریتم آموزش برای این شبکه‌ها، به میزان زیادی به ساختار انتخاب شده دارد. برای مشخص نمودن ساختار بهینه باید یک جستجوی کامل در فضای ساختارها انجام گیرد که این خود دارای هزینه بالایی است. مسئله پیدا کردن ساختار بهینه برای شبکه‌های عصبی، در گروه مسائل NP-Hard قرار دارد و به همین جهت برای کاهش زمان تعیین ساختار مناسب برای شبکه، روش‌های تقریبی از جمله الگوریتم‌های سازنده، الگوریتم‌های هرس، الگوریتم‌های ترکیبی و الگوریتم‌های تکاملی پیشنهاد شده اند که ساختار نزدیک به بهینه را تولید می کنند. در اغلب این روش‌ها، از الگوریتم‌های گوهنوردی به عنوان ابزار جستجو استفاده می شود که مشکل گرفتاری در حداقل‌های محلی را دارند. در این مقاله الگوریتمی به نام الگوریتم بقا، برای تعیین ساختار شبکه عصبی سه لایه معرفی شده است که از یک اتوماتان یادگیر (به عنوان یک ابزار جستجوی عمومی) و الگوریتم انتشار خطا به عقب استفاده می کند و در ضمن آموزش، ساختار نزدیک به بهینه را برای شبکه تعیین می نماید. استفاده از اتوماتان‌های یادگیر به عنوان یک ابزار جستجوی عمومی مشکل گرفتاری در حداقل‌های محلی را تا حدودی حل می نماید و احتمال رسیدن به بهترین ساختار را افزایش می دهد. در الگوریتم بقا، آموزش از یک شبکه بزرگ شروع شده و اتومان یادگیر با افزودن و کاستن واحدهای مخفی، حداقل تعداد واحدهای مورد نیاز لایه مخفی را تعیین می کند.

A Learning Automata Based Algorithm for Determination of Minimum Number of Hidden Units for Three Layers Neural Networks

H. Beigy
Ph.D. Student

M. R. Meybodi
Associate Professor

Computer Engineering Department,
Amirkabir University of Technology

Abstract

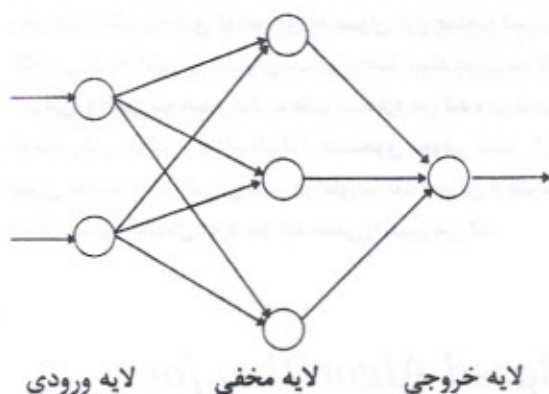
One of the unsolved problems in multi-layer neural networks is the problem of determination of optimal topology: a topology with minimum number of hidden units. There is no method to determine the optimal topology for multi-layer neural networks for a given problem. Usually the designer selects a topology for the network and then trains it. The selected topology remains fixed during the training period. Since the problem of determination of optimal topology for neural networks belongs to class of NP-Hard problems, most of the existing algorithms for determination of the topology are approximate algorithms. These algorithms could be classified into four main groups: pruning algorithms, constructive algorithms, hybrid algorithms, and evolutionary algorithms. These algorithms can produce near optimal solutions. Most of these algorithms use hill-climbing method and may stuck at local minima. In this paper a learning automata based algorithm called survival algorithm, for determination of the number of hidden units of three layers neural networks is proposed. The proposed algorithm uses learning automaton as a global search method in order to increase the probability of escaping from local minima and hence to increase the probability of obtaining the optimal topology. In survival algorithm, the training begins with a large network, and then by adding and deleting hidden units the optimal topology is obtained.

کلمات کلیدی

ساختار شبکه عصبی، آموزش شبکه عصبی، اتوماتان یادگیر، الگوریتم انتشار خطا به عقب، شبکه های سه لایه.

مقدمه

شبکه های عصبی دارای کاربردهای زیادی می باشند [۲۹] [۴۱]. بیشتر این کاربردها از شبکه های عصبی چند لایه پرسپترون^۱ استفاده می کنند. شبکه های عصبی چند لایه پرسپترون^۱ (که به اختصار شبکه های عصبی چند لایه نامیده می شوند)، شبکه های پیش خوری^۲ هستند که در آنها یک یا بیشتر لایه مخفی بین لایه های ورودی و خروجی وجود دارد. در این شبکه ها معمولاً اتصال های بین لایه ها کامل در نظر گرفته می شود و ورودی های واحدهای هر لایه فقط به خروجی لایه قبل از خود متصل هستند. شکل ۱ یک شبکه عصبی ۳ - لایه پرسپترون^۱ را که در این مقاله مدنظر می باشد، نشان می دهد.



شکل (۱) شبکه عصبی ۳ - لایه پرسپترون^۱.

شبکه عصبی چند لایه پرسپترون^۱ توسط رزنبلت^۳ پیشنهاد شد [۴۱] که در آن از تابع پله^۴ برای تابع فعالیت^۵ واحدها استفاده گردیده است. فقدان یک الگوریتم یادگیری مناسب یکی از محدودیت های این شبکه های چند لایه پرسپترون^۱ است. یک روش معمول برای آموزش چنین شبکه های، استفاده از روش تکراری نزول در امتداد گرادیان می باشد که به دلیل عدم مشتق پذیری تابع پله این روش قابل استفاده نیست. استفاده از یک تابع غیرخطی مشتق پذیر برای تابع