

مهندسی شبکه های عصبی توسط اتوماتانهای یادگیر: تعیین اندازه مطلوب برای شبکه های عصبی سه لایه

محمد رضا میبیدی

دانشیار دانشکده مهندسی کامپیوتر - دانشگاه صنعتی امیر کبیر

حمید بیگی

دانشجوی دکترای دانشکده مهندسی کامپیوتر - دانشگاه صنعتی امیر کبیر

(تاریخ دریافت ۷۹/۳/۲۵، تاریخ تصویب ۷۹/۱۰/۳)

چکیده

هدف از مهندسی شبکه های عصبی بررسی معایب و مزایای شبکه های عصبی مصنوعی و ارائه روشهایی برای بهبود کارایی آنهاست. یکی از موضوعات مورد بحث در مهندسی شبکه های عصبی چندلایه، یافتن ساختار مناسب (نزدیک به بهینه) برای حل مسئله میباشد. معیار و نحوه انتخاب اندازه شبکه عصبی برای یک مسئله خاص هنوز شناخته شده نیست. در روشهای کلاسیک، طراح شبکه در ابتدای آموزش ساختاری را برای شبکه تعیین و سپس شبکه را آموزش میدهد و ساختار تعیین شده در حین آموزش ثابت نگاه داشته میشود. تعیین اندازه شبکه عصبی (تعداد واحدها و لایه های مخفی) تاثیر عمده ای بر روی کارایی شبکه های عصبی خواهد داشت. طراحی یک شبکه با ساختار بهینه یک مسئله ذاتا مشکل است. بهمین جهت بیشتر الگوریتم های ارائه شده برای تعیین ساختار شبکه های عصبی، الگوریتم های تقریبی هستند. این الگوریتم ها را میتوان به پنج گروه عمده الگوریتم های هرس، الگوریتم های سازنده، الگوریتم های ترکیبی، الگوریتم های تکاملی و الگوریتم های بر اساس اتوماتانهای یادگیر تقسیم کرد. تنها الگوریتم گزارش شده بر اساس اتوماتان یادگیر، الگوریتم بقا نام دارد که توسط بیگی و میبیدی ارائه گردیده است. این الگوریتم از یک اتوماتان یادگیر مهاجرت اشیا (بمعنای یک ابزار جستجوی عمومی) و الگوریتم یادگیری انتشار خطا به عقب استفاده میکند و در ضمن آموزش، تعداد واحدهای مخفی یک شبکه سه لایه را تعیین مینماید. در این مقاله سه الگوریتم بر اساس اتوماتانهای یادگیر مهاجرت اشیا و الگوریتم یادگیری انتشار خطا به عقب ارائه شده است که ساختار مناسب شبکه عصبی سه لایه (ساختاریکه دارای اندازه کوچک، پیچیدگی آموزش کم و قدرت تعمیم بالا باشد) را برای یک مجموعه از الگوهای آموزشی تعیین می نماید. اغلب روشهای گزارش شده برای تعیین ساختار شبکه از الگوریتم های کوهنوردی استفاده میکنند و مشکل گرفتاری در حداقل محلی را دارند. در روش پیشنهادی در این مقاله بدلیل استفاده از روشهای جستجوی عمومی، امکان گرفتاری در حداقل های محلی فضای ساختارها کاهش می یابد. الگوریتم های ارائه شده روی مسائل متنوعی از جمله شناسایی اعداد فارسی و شناسایی اعداد انگلیسی آزمایش شده اند. از طریق شبیه سازی نشان داده شده است که شبکه های تولید شده توسط الگوریتم های پیشنهادی دارای کارایی بالایی هستند. همچنین الگوریتم های پیشنهادی با الگوریتم هرس کارنین مقایسه گردیده است. نتایج الگوریتم های پیشنهادی هم از لحاظ ساختار تولید شده، هم از نظر زمان مورد نیاز برای تعیین ساختار شبکه عصبی، بسیار بهتر از الگوریتم هرس کارنین عمل کرده اند.

واژه های کلیدی: مهندسی شبکه های عصبی، شبکه های عصبی چند لایه، ساختار شبکه های عصبی، اتوماتانهای یادگیر مهاجرت اشیا

مقدمه

شبکه را آموزش میدهد. ساختار تعیین شده در حین آموزش ثابت نگاه داشته میشود. تعیین تعداد واحدهای لایه مخفی و وزنهای آن تاثیر بسزایی بر روی کارایی شبکه دارد. یک شبکه با اندازه کوچک، خروجی دقیق تولید نمیکند و شبکه های با اندازه بزرگ، بسیار کند و پرهزینه خواهد بود و برای تعمیم مجموعه آموزشی، نیاز به مجموعه آموزشی بزرگی خواهد

هدف از مهندسی شبکه های عصبی بررسی معایب و مزایای شبکه های عصبی مصنوعی و ارائه روشهایی برای بهبود کارایی آنهاست. یکی از موضوعات مورد بحث در مهندسی شبکه های عصبی چندلایه یافتن ساختار مناسب (نزدیک به بهینه) برای حل مسئله میباشد. در روشهای کلاسیک، طراح شبکه در ابتدای آموزش ساختاری را برای شبکه تعیین و سپس