

## ارزیابی عملکرد ماشین یادگیر نهایی در قطعه بندی تصاویر

محمود امین طوسی، سکینه خورسندی، مهدی زعفرانی

دانشکده ریاضی و علوم کامپیوتر. دانشگاه حکیم سبزواری. سبزوار. ایران

### چکیده

ماشین یادگیر نهایی (ELM) از جمله جدیدترین ابزارهای یادگیری ماشین است که استفاده‌های روزافزونی در کاربردهای متعدد حوزه یادگیری ماشین پیدا کرده است. مزیت اصلی این روش نسبت به شبکه عصبی پرسپترون چند لایه (MLP)، کم بودن زمان آموزش آن نسبت به این شبکه عصبی معروف است. در بسیاری از مقالات اصلی ماشین یادگیر نهایی، دقت این سیستم بیشتر از MLP گزارش شده است. در این مقاله برای ارزیابی این ادعا، دو نسخه از ماشین یادگیر نهایی در کاربرد قطعه‌بندی تصاویر با شبکه عصبی پرسپترون چند لایه مورد مقایسه قرار گرفته‌اند. آزمایشات مختلف بر روی پایگاه داده BSD و با در نظر گرفتن تعداد نورون‌های متفاوت در لایه پنهان انجام شده است. نتایج آزمایشات با معیارهای مختلف حاکی از آن بوده است که در حالت میانگین، خطای طبقه‌بندی شبکه عصبی پرسپترون چند لایه از ماشین یادگیر نهایی کمتر است؛ با این حال آموزش ماشین یادگیر نهایی بسیار سریعتر از MLP می‌باشد. با تعداد ۱۰۰ نورون در لایه پنهان، MLP، 2.5 درصد «صحت» بیشتری از ELM دارد، در حالیکه زمان آموزش MLP، 155 برابر زمان آموزش ELM است.

واژه های کلیدی: قطعه‌بندی تصویر، ماشین یادگیر نهایی، ماشین یادگیر نهایی دو لایه پنهان، پرسپترون چند لایه

### مقدمه یا بیان مسئله

از آنجا که هدف مقاله مقایسه شبکه عصبی پرسپترون چندلایه با ماشین یادگیر نهایی و ماشین یادگیر نهایی دو لایه است، ابتدا بصورت مختصر مروری

### پرسپترون چندلایه

در آموزش شبکه‌های عصبی پرسپترون چندلایه (MLP) با "فازده پس انتشار خطا" نمونه‌ای به شبکه عرضه می‌شود و خروجی آن محاسبه می‌گردد، با مقایسه خروجی شبکه و خروجی واقعی، ضرایب وزنی شبکه بین لایه ورودی و لایه پنهان و ضرایب وزنی بین لایه پنهان و لایه خروجی به نحوی تغییر می‌یابند که در دفعات بعد خروجی درست‌تری حاصل شود.

### ماشین یادگیر نهایی

ماشین یادگیر نهایی (ELM) برای آموزش شبکه‌های عصبی پیشخور تک لایه پنهان با اجتناب از یک روش آموزش تکراری وقت‌گیر و هدف کاهش زمان آموزش و بهبود عملکرد تعمیم پیشنهاد شده است که در آن پارامترهای لایه پنهان، یعنی وزن‌های ورودی بین لایه ورودی و لایه پنهان و همچنین بایاس‌ها، همزمان به طور تصادفی تولید می‌شوند و در طول فاز آموزش ثابت می‌مانند و فقط وزن‌های خروجی بین لایه پنهان و لایه خروجی با روش حداقل مربعات منظم بصورت زیر قابل تنظیم‌اند:

$$\beta = H^{\dagger} T$$

که در آن منظور از  $H^{\dagger}$  معکوس تعمیم یافته مور-پنروز ماتریس خروجی لایه پنهان است. زمانیکه  $H^T H$  نامفرد است  $H^{\dagger} = H^T (H H^T)^{-1} H$  و زمانیکه  $H H^T$  نامفرد است  $H^{\dagger} = H (H^T H)^{-1}$  بصورت  $H^{\dagger} = H^T (H H^T)^{-1} H$  محاسبه می‌شود. منظور از  $T$  هم ماتریس خروجی واقعی نمونه‌های آموزشی است.

### ماشین یادگیر نهایی دولایه

ماشین یادگیر نهایی دولایه (TELM) یکی از توسعه‌های (ELM) است که یک لایه پنهان به آن اضافه نموده که پارامترهای لایه پنهان اول بطور تصادفی مقداردهی اولیه می‌شوند ولی پارامترهای لایه پنهان دوم بدست می‌آیند. ابتدا وزن‌های ورودی و بایاس اولین لایه پنهان به طور تصادفی مقداردهی اولیه می‌شوند و خروجی مورد انتظار لایه پنهان دوم طبق رابطه  $H_1 = T \beta^{\dagger}$  محاسبه می‌شود. سپس ماتریس  $W_{HE} = [B_1 W_H]$  که  $B_1$  بایاس و  $W_H$  وزن ورودی لایه پنهان دوم هستند طبق رابطه  $W_{HE} = g^{-1}(H_1^T)(H_E^T)^{\dagger}$  به دست می‌آید که  $H_E = [1 H]$  و "1" برداری با عناصر اسکالر ۱ است.  $g^{-1}(x)$  معکوس تابع خروجی  $g(x)$  را نشان می‌دهد. خروجی واقعی لایه پنهان دوم با  $H_2 = g(W_{HE} H_E^T)$  به دست می‌آید. در نهایت

$$\beta = H_2^{\dagger} T$$

### روش پژوهش

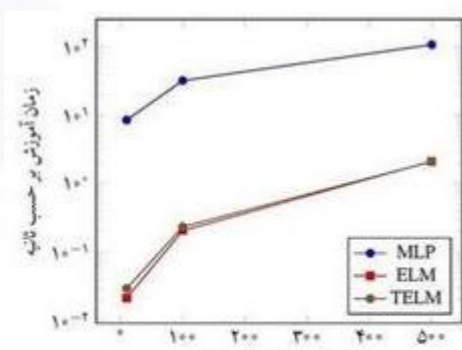
ستون دوم با عنوان «تصویر اصلی» در جدول زیر - که شامل نتایج بصری است - این تصاویر را نشان می‌دهند. پیکسل‌های آموزشی مشخص شده توسط کاربر از ناحیه زمینه با رنگ سیاه و از ناحیه پیش زمینه با رنگ سفید مشخص شده‌اند. ستون سوم از جدول نمایش دهنده «نقاط آموزشی» در تصویر مربوطه است. داده‌های درست قطعه‌بندی شده توسط افراد مختلف تحت عنوان «جواب واقعی» در ستون چهارم از جدول آورده شده است. متناظر با هر پیکسل، مقادیر شدت رنگ RGB تمامی پیکسل‌ها در یک همسایگی ۳\*۳ آن، واریانس همسایگی و گرادیان‌های افقی و عمودی آن به عنوان ویژگی‌های هر پیکسل مورد استفاده قرار گرفته است. آزمایشات روی شبکه‌های عصبی با تعداد ۱۰، ۱۰۰ و ۵۰۰ نورون پنهان انجام شده است. برای هر یک از تصاویر تمام الگوریتم‌های ۱۰ مرتبه اجرا شده‌اند و تصویر متناظر با اجرای دارای بیشترین «صحت» برای تعداد ۱۰۰ نورون پنهان در جدول آمده است. از معیارهای صحت، حساسیت، نرخ مثبت کاذب و نرخ منفی کاذب به منظور ارزیابی دقت روش‌ها استفاده شده و زمان آموزش آن‌ها نیز مورد سنجش قرار گرفته است.

ردیف	تصویر اصلی	نقاط آموزشی	جواب واقعی	MLP	ELM	TELM
۱						
۲						
۳						
۴						
۵						
۶						
۷						
۸						
۹						
۱۰						
۱۱						
۱۲						
۱۳						
۱۴						
۱۵						

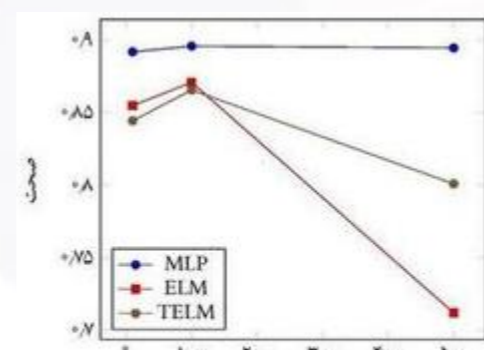
جدول ۱: مقایسه خروجی قطعه‌بندی روش‌های مختلف با ۱۰۰ نورون در لایه مخفی. خطوط سفید و سیاه در تصاویر ستون دوم نمایش دهنده نقاط آموزشی هستند. ستون سوم، قطعه‌بندی واقعی و سه ستون آخر خروجی سه شیوه مورد نظر را نشان می‌دهند.

### یافته‌های پژوهش

بر اساس نتایج بدست آمده، در حالت ۱۰۰ نورون در لایه پنهان، هر سه شبکه کارایی نسبتاً خوبی داشتند. در این وضعیت، در حالیکه صحت MLP فقط حدود ۰/۰۲۵ از ELM بیشتر است، اما زمان آموزش آن ۱۵۵ برابر زمان آموزش ELM است. وضعیت مشابهی در خصوص TELM برقرار است. لذا اگر در کاربرد مدنظر قطعه‌بندی، زمان آموزش از اهمیت بالایی برخوردار باشد و بتوان از مختصر کاهش دقت چشم‌پوشی نمود، ماشین یادگیر نهایی ELM انتخاب بهتری نسبت به MLP می‌باشد.



شکل ۲: مقایسه میانگین «زمان آموزش» روش‌های مختلف در تعداد نورون‌های لایه پنهان مشخص شده.



شکل ۱: مقایسه میانگین «صحت» روش‌های مختلف در تعداد نورون‌های لایه پنهان مشخص شده.

### منابع

- D.E. Rumelhart, G.E. Hinton, and R.J. Williams, "Neurocomputing: Foundations of research," chap. Learning Representations by Back-propagating Errors, pp.696-699, Cambridge, MA, USA: MIT Press, 1988.  
G.Huang, G.-B. Huang, S.Song, and K.You, "Trends in extreme learning machines: A review," Neural Networks, vol.61, pp.32 -- 48, 2015.  
G.-B. Huang and H.A. Babri, "Upper bounds on the number of hidden neurons in feedforward networks with arbitrary bounded nonlinear activation functions," IEEE Transactions on Neural Networks, vol.9, no.1, pp.224-229, 1998.