روشی جدیدجهت آموزش شبکه عصبیMLP با استفاده ازقوانین فازی

محمد کمـالی
گروه کامپیوتر – هوش مصنوعی
دانشگاه آزاد اسلامی – واحد مشهد
kamali622@yahoo.com

سعید راحتی قوچانی
گروه کامپیوتر – هوش مصنوعی
دانشگاه آزاد اسلامی – واحد مشهد
rahati@mshdiau.ac.ir

واژه‌های کلیدی: قوانین فازی – آموزش پرسپترون – MLP – همگرایی – پس‌انتشار خطا – شبکه عصبی

چکیده

استفاده از منطق فازی در آموزش شبکه‌های عصبی به دلیل نزدیکی آن به زبان انسانی و نیز به جهت همگرایی سریعتر در آموزش و نیاز به عصبهای کمتر، آن را نسبت به دیگر روشها مانند آموزش مبتنی بر الگو یا تکرار، متمایز کرده است.

در این مقاله به کمک یک تدوین جدید برای قوانین متناسب با ساختار شبکه‌های عصبی پرسپترون چندلایه، روشی برای آموزش آنها بر اساس قوانین فازی ارائه می‌شود.

مدل پیشنهادی خود را با یک مدل آموزش دیده بر اساس روش پس‌انتشار خطا از نظر سرعت همگرایی و دقت در یادگیری و تعداد عصبهای موجود در شبکه‌ی عصبی مقایسه می‌کنیم.

برای این کار آزمایشی برای درون‌یابی یک تابع پیچیده ترتیب داده شد. متوسط خطا برای یک مجموعه آموزش ۲۱۶ عضوی، ۰/۱۹ در ۲۰۰ بار تکرار و برای آموزش به روش کلاسیک برای همین مجموعه، متوسط خطا ۰/۴ در ۶۰۰ بار تکرار آموزش می‌باشد.

New Method for learning MLP Neural Network using Fuzzy Rules

Keywords: Fuzzy rules, Training, Neural networks, train perceptrons , back propagation , convergence

Abstract:

Since the fuzzy logic is more consistent with human languages and have more speed in convergence of learning in neural networks, and require fewer neurons, it become more prominent than pattern-based or repeated algorithms.

In this paper, using new formulation for rule construction that is appropriate for MLP neural networks structures, we present an approach to learning them according to fuzzy rules.

We compare our neural network that was learned upon proposed model with a learned network that was learned by back propagation methods in point of view of accuracy and speed of convergence.

We applied our method on interpolation of functions task. Using proposed method, Mean of error was 0.19 for 216 member training set in 200 epochs, versus 0.4 in 600 epochs.

روشی جدیدجهت آموزش شبکه عصبیMLP با استفاده ازقوانین فازی

واژه‌های کلیدی: قوانین فازی - آموزش پرسپترون - MLP – همگرایی - پس‌انتشار خطا – شبکه عصبی

۱. مقدمه

شبکه‌های عصبی دارای ساختارهای قدرتمندی برای دسته‌بندی[1] داده‌ها و یادگیری الگوها می‌باشند. در چند جمله می‌توانیم اهم مزایای شبکه‌های عصبی را به صورت ذیل بیان کنیم : مقاوم بودن نسبت به نویز، شناسایی الگوها (صرفا با بخشی از الگو) و یادگیری الگوهای جدید، اجرا به طور موازی، ارائه نگاشتهای غیرخطی برای مسائل پیچیده، عدم نیاز به دانستن ارتباط بین داده‌ها، قابلیت خود گردانی وخود تنظیمی و خود آموزشی، قابلیت مدل‌سازی سیستمهای مختلف با توابع غیر خطی.

یک شبکه‌های عصبی مرحله به مرحله بر اساس اطلاعاتی که در یافت می‌کند تغییر کرده و هر چه این اطلاعات بیشتر شود شبکه خطای خودش را کاهش می‌دهد. این امر می‌تواند با رفتار یک انسان در یک محیط ناشناخته مقایسه گردد. با وجود الهام تفکر شبکه‌های عصبی از مغز انسان، تفاوتهای عمده‌ای بین این دو وجود دارد. پویا بودن مغز، توانایی ایجاد لینکهای جدید در مغز انسان، درگیر بودن شبکه‌های عصبی مصنوعی با مسأله نفرین بعدیت[2]، زمان‌بر بودن آموزش شبکه‌های عصبی مصنوعی و مشکل همگرایی نسبت به نقطه بهینه تفاوتهای اساسی بین این دو هستند. ساز و کارهای شناختی مغز در سطوح بالاتر از اعداد است. در صورتی که عموماً شبکه‌های عصبی مصنوعی به ارائه یک ساختار سطح پایین از داده‌ها می‌پردازند. با وجود سرعت بسیار پایینترحرکت سیگنالهای عصبی در دستگاه عصبی طبیعی انسان که ناشی از تفاوت غلظت مواد شیمیایی در عصب است و بالا بودن این سرعت در لینکهای شبکه عصبی مصنوعی به علت سرعت بالای سیگنالهای الکتریکی و تفاوت انتقال اطلاعات در محیطهای مختلف از لحاظ سرما و گرما و بیماری در عصبهای طبیعی، سرعت درک و فراگیری مغز یک انسان قابل مقایسه با شبکه عصبی مصنوعی نبوده و به مراتب اختلافات زیادی بین این دو وجود دارد.

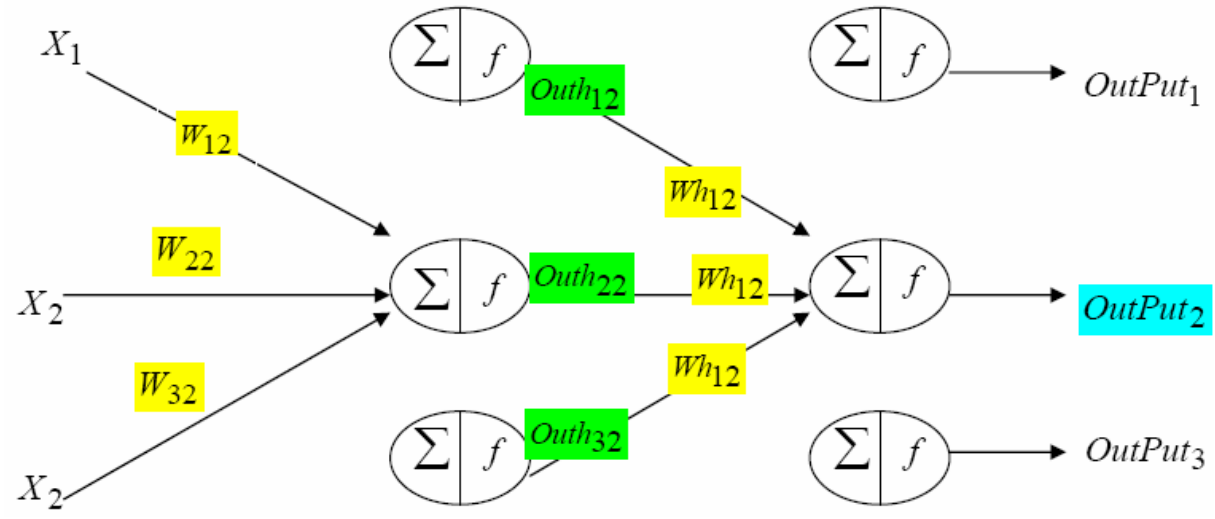
دو مشکل دیگر به طور اساسی در شبکه‌های عصبی دیده می‌شود: عدم توانایی در درگیر شدن با اطلاعات زبانی و عدم توانایی ترکیب داده‌های عددی با داده‌های زبانی یا منطقی. می‌توان با استفاده از آموزش شبکه‌های عصبی به روش فازی دو مشکل اساسی شبکه‌های عصبی را بهبود داد. سیستمهای عصبی فازی برای فراگیری دانش انسانی[3] مناسب می‌باشند. که قابلیت به‌کارگیری هر نوع داده‌ای (عددی، زبانی، منطقی) را دارند. مهمترازهمه اینکه یک سیستم فازی بر پایه تقلید از نحوه تصمیم‌گیری انسان در موارد مبهم و نامعلوم وهمچنین قابلیت درگیری با اطلاعات نادرست و ناقص پیاده سازی می شود .

در ادامه‌ی این مقاله ابتدا یک الگوریتم بر پایه قوانین فازی که می‌تواند یک شبکه عصبیMLP را قدرتمند کند تا برخی از مشکلات ذکر شده در بالا را حل نماید را ارائه دهیم ودربخش سوم ساختار تشخیص متغیرهای زبانی مورد استفاده در الگوریتم را تشریح می‌کنیم و سپس در بخش آخر چگونگی به کار بردن توابع عضویت در این سیستم را مورد توجه قرار خواهیم داد. ودرادامه با مقایسه نتایج ملاحظه خواهیم کرد که چقدر می‌تواند استفاده از این الگوریتم که بر پایه قوانین فازی ارائه شده بر روی شبکه عصبی مورد استفاده مفید واقع گردد.

[1] Classification
[2] Curse Of Dimensionality
[3] Rule Generation

۲. آموزش فازی یک شبکه‌ی پرسپترون با یک لایه پنهان:

ابتدا یک شبکه ساده که از بردار ورودی، یک لایه پنهان و یک لایه خروجی تشکیل شده است را در نظر می‌گیریم و با استفاده از روش پس‌انتشار خطا[4] برای به‌روزآوری وزنهای لایه پنهان اقدام می‌کنیم. شکل ۲ یک شبکه با ترکیب ۳ ورودی، ۳ خروجی و یک لایه مخفی با ۳ عصب را نشان می‌دهد. البته معمولاً در یک عصب علاوه بر ورودیهای متداول، ورودی قابل تنظیم دیگری به نام بایاس نیز وجود دارد.



الف)

$$OutPut_i = f\left(\sum_{i=1}^{3} Outh_j X_i Wh_{ji}\right)$$

ب)

$$Outh_j = \left(\sum_{i=1}^{3} X_i Wij\right)$$

ج)

شکل ۱- الف) شبکه عصبی پرسپترون با ۳ ورودی و ۳ خروجی و یک لایه پنهان. ب) خروجی عصب لایه آخر. ج)خروجی عصب لایه پنهان.
در این اشکال W_{ji}، وزن ورودی به عصب j لایه پنهان از ورودی i، X_i، بردار ورودی، $Outh_j$، خروجی عصب j از لایه پنهان، Wh_{ji} وزن ورودی به عصب i ام لایه خروجی از عصب j ام لایه پنهان می‌باشد.

در یک شبکه عصبی با یک لایه پنهان دو نوع اتصال داریم: اتصال ورودی به لایه پنهان از لایه ورودی و اتصال خروجی از لایه پنهان به لایه خروجی. پس دو مجموعه بردار وزن داریم. V_{ij} وزن ورودی به عصب j از عصب i است و W_{ji} وزن ورودی به عصب i ام لایه خروجی از عصب j ام لایه پنهان است. در آموزش شبکه عصبی با روش پس‌انتشار خطا این وزنها طوری تنظیم می‌شوند که خطا به حداقل میزان خود برسد. باید توجه کرد که شبکه به صورت با سرپرستی[5] آموزش داده می‌شود. در آموزش به شیوه فازی این اصلاح وزنها بر اساس متغیرهای زبانی صورت می‌گیرند.

اکنون ساز و کار آموزش و اصلاح وزنها در یک شبکه عصبی MLP با یک لایه پنهان بر اساس قوانین فازی را شرح می‌دهیم. ابتدا ورودیها به شبکه اعمال می‌شوند و بر طبق روابط موجود در شکل ۲-ب خروجی هر عصب لایه پنهان محاسبه می‌گردد. سپس این خروجیها با توجه به رابطه ۲-ج در وزنهای مشخص شده ضرب شده و روی آنها مجموع گرفته می‌شود که پس از گذشتن از تابع F خروجی نهایی به‌دست می‌آید. در اینجا یک سطح تصمیم‌گیری وجود دارد. باتوجه به تحت ناظربودن آموزش شبکه اگر خطای بین خروجی واقعی شبکه با خروجی مطلوب کمتر از مقدار تعیین شده باشد آموزش انجام شده است و شبکه خروجیهای مورد نظر را تولید کرده است. ولی اگر این خروجیها با خروجیهای مطلوب یکی نباشد و اختلاف زیادی داشته

[4] BackPropagation
[5] Supervised

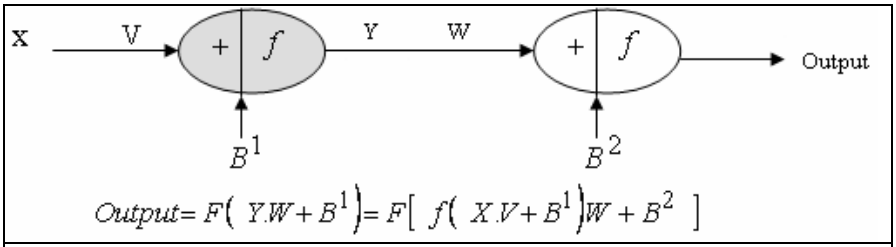
باشند باید خطا را محاسبه و آنرا به لایه ماقبل یعنی همان لایه پنهان برگشت دهیم تا خطا کم شود. این کاهش خطا بر اساس قوانین فازی انجام خواهد گرفت. پس هدف به‌روزآوری وزنهای وابسته به این اتصال می‌باشد.

وقتی یک عصب حاصل ضرب درونی بین بردار ورودی به اتصال و بردار وزنهای وابسته به این اتصال را محاسبه می‌کند در واقع میزان شباهت بین این دو بردار را محاسبه می‌کند. بنابراین هر چه شباهت بین بردار X و W زیادتر باشد مقدار خروجی این عصب بزرگتر خواهد بود. این درجه شباهت بین بردارهای موجود در یک عصب را می‌توان به وسیله قوانین فازی بیان نمود.

الف - وقتی خروجی مطلوب عصب پایین باشد مثل این است که درجه شباهت بین بردار X و W وارد شده به آن باید کم باشد، یعنی این دو بردار با هم اختلاف زیاد دارند.

ب - اگر خروجی مطلوب عصب بالا باشد یعنی اینکه شباهت بین بردار X و W زیاد بوده و این دو بردار اختلاف کمتری نسبت به هم دارند.

اکنون می‌توانیم با توجه به مطالب فوق برای اصلاح وزنها، قواعد زبانی بنویسیم. اکنون یک ساختار عصبی دولایه مطابق شکل ۲ را در نظر بگیرید:



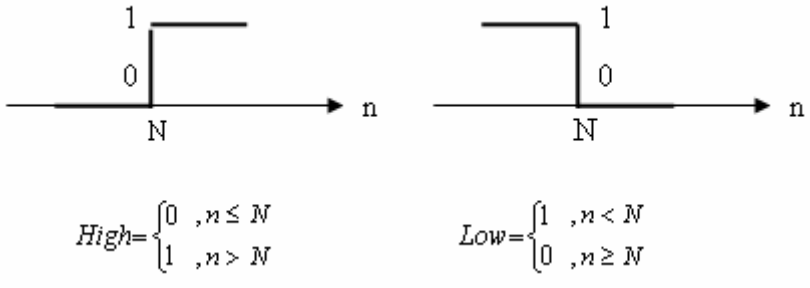
شکل ۲) نمایش دو عصب پشت سرهم در حالت کلی همراه با محاسبه خروجی آن.

اگر مقدار موجود در اتصال (Y) برابر A و وزن مربوط به اتصال، B و خروجی واقعی عصبی که اتصال به آن تمام شده است C و خروجی مطلوب D باشد آنگاه وزن اصلاح شده، E است. به عبارت دیگر:

If Y=A , W=B , Output = C, out = D Then W' = E

در اینجا A، B، C، D، E برچسبهای زبانی هستند. اکنون مجموعه‌ی قوانین فازی برطبق رابطه‌ی بالا به هر اتصال لایه خروجی مطابق شکل ۲ اعمال می‌شود. این قوانین برای لایه خروجی اعمال می‌شوند و به طریق مشابهی که در ادامه خواهد آمد، مجموعه قوانین فازی مشابهی برای به‌روزآوری وزنهای مربوط به لایه پنهان طراحی می‌گردند. چون هر نرون بیش از یک ورودی را از چندین اتصال نرونهای لایه ماقبل می گیرد بنا بر این اطلاعات مربوط به خروجی مطلوب آن با استفاده از عملگر های مجتمع دسته بندی می شوند.

حال با توجه به اینکه تابع فعالیت *Hard Limiting* است، می‌توانیم متغیرهای زبانی *Low* و *High* را به هر یک از وزن مطلوب لایه‌ی خروجی عصب و خروجی واقعی به‌دست آمده از عصب اختصاص دهیم.



شکل ۳) تابع فعالیت عصب.

برچسبهای وابسته به متغیرهای زبانی وزن و اصلاح وزن عبارتند از: Small، Medium، و Big.

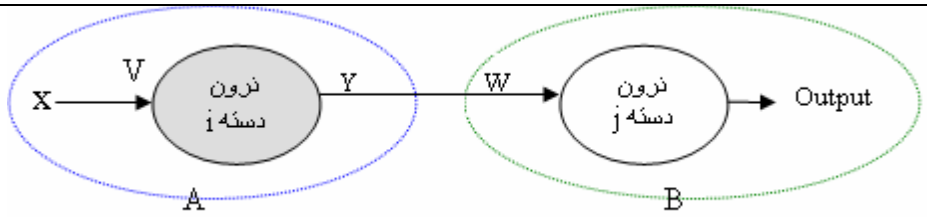
اگر وزنهای وابسته به اتصالات ورودی یک عصب که خروجی آن منطبق بر خروجی مطلوب نباشد، باید به وسیله الگوریتم آموزش،

اصلاح شوند. با توجه به مطالب ذکر شده می‌توان دو دسته قوانین فازی برای اصلاح وزن که هدف اصلی در آموزش است، طراحی کنیم:

الف) قوانین دسته اول: این قوانین برای اصلاح وزنهای اتصالات بین عصبهای لایه پنهان و لایه خروجی می‌باشند. بر اثر اعمال این قوانین، وزن این اتصالات، یعنی w_{ij} و مقدار Y_i یعنی خروجی عصبهای لایه پنهان اصلاح می‌شوند، که این امر خود با اصلاح وزنهای لایه اول یعنی V_{ij} امکان‌پذیر می‌باشد. بنابراین قوانین این دسته برای رسیدن به هدف، می‌توانند هر دو نوع وزن را در اتصالات مختلف اصلاح کنند.

ب) قوانین دسته دوم: که فقط روی وزنهای عصبهای لایه پنهان از طرف ورودیها اثر کنند و این وزنها را به‌روزآوری می‌کنند.

شکل ۴ مجموعه‌ی قوانین برای آموزش اتصالات لایه‌ی خروجی و پنهان را نشان می‌دهد. راهبرد الگوریتم کلی آموزش شبکه بر پایه قوانین فازی به شرح ذیل است: ابتدا مقدار وزن وابسته به اتصال خروجی اصلاح می‌شود:



الف)

ب)	ج)
High =مقدار موجود در اتصال(Y)	A = High
High = خروجی واقعی عصب (Output)	Out =High
Low = خروجی مطلوب (Out)	if Output = Low
Medium = وزن وابسته به اتصال (W)	then اصلاح وزن (W') = Low
	W = Medium

شکل ۴) تقسیم شبکه به دو قسمت برای اعمال یک دسته قانون به هر قسمت. ب) قوانین قسمت A (قوانین دسته‌ی دوم). ج) قوانین قسمت B (قوانین دسته‌ی اول).

شکل کلی قانونهای اعمال شونده به لایه خروجی به صورت ذیل می‌باشد:

1) **If** Y=A, W=B, Out = C, Output = D **Then** وزن اصلاح شده = E
2) **If** Y=A, W=B, Out = C, Output = D **Then** (V)وزن اصلاح شده = E, Y0=F (۱)

و شکل کلی قانونهای اعمال شونده به لایه‌ی پنهان به صورت ذیل تعیین می‌شود:

If X = A, V = B, Y = C, Yo = D **Then** V = E (۲)

چنانچه خروجی شبکه بر خروجی مطلوب، منطبق نباشد بر طبق رابطه‌ی (۱) یا باید مقدار وزنها عوض شود یا مقدار موجود در اتصال عصب خروجی. ابتدا مقدار وزن وابسته به اتصال تغییر کرده و به سمت Small گرایش پیدا می‌کند و اصلاح وزن در اینجا مقدار Small تعیین می‌گردد. با توجه به خطا، اگر نتیجه مطلوب به دست نیامد یعنی خروجی مطلوب با توجه به خطا درحد Medium نشده است و مقدار زیاد نبوده و در مرحله بعد باید مقدار موجود در اتصال تغییر کند و این جز با اصلاح وزن لایه قبل امکان ندارد بنابراین وزن لایه ماقبل متناسب با این مقدار تغییر می‌کند.

در مورد پارامتر های بدست آمده از اولین لایه ملاحظه می گردد که اگرخروجی مطلوب به خروجی واقعی بدست آمده منطبق نباشد دیگرمقدار موجود در اتصال که همان ورودیها باشند تغییر نمی کنند و فقط قوانین دسته دوم می‌توانند مقدار وزن را تغییر دهند که اصلاح وزن در اینجا ازHigh به سمت Medium و سپس Low گرایش پیدا می کند. با توجه به خروجی مطلوب و خروجی واقعی بدست آمده.

۳. نتایج تجربی

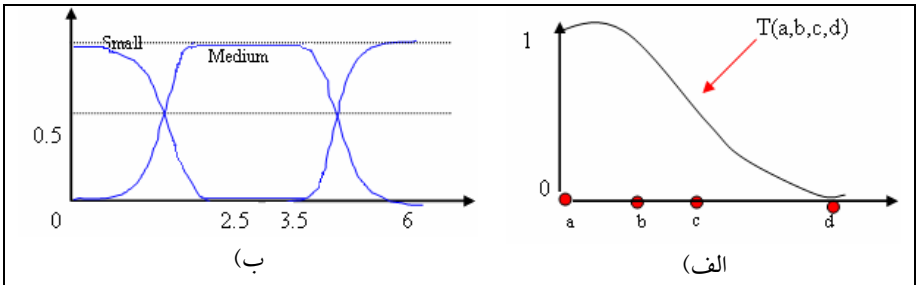
برای مقایسه‌ی دو روش مختلف آموزش شبکه عصبی MLP آزمایشی برای درونیابی تابع سه متغیره مطابق رابطه (۳) ترتیب می‌دهیم.

$$f = (x^{-1/5} + y^{1/5} + z^{-1/5})^{1.5} \qquad (3)$$

برای این کار یک مجموعه‌ی آموزشی با ۲۱۶ عضو برای آموزش شبکه‌ی عصبی و یک مجموعه‌ی آزمایشی با ۱۲۵ نمونه برای محاسبه‌ی میزان صحت یادگیری شبکه، تشکیل می‌دهیم.

برای آموزش شبکه عصبی به روش پس‌انتشار خطا از یک شبکه‌ی عصبی پرسپترون با ۶۷ عصب و ۶۰۰ بار تکرار استفاده کردیم که خطا مجموعه‌ی آموزشی به ۰,۴ کاهش یافت. خطای شبکه برای مجموعه‌ی آزمایشی که تاکنون آن را ندیده است، به ۰,۵ می‌رسد.

برای آموزش شبکه‌ی عصبی به روش فازی از یک شبکه با ۸ عصب و مجموعه‌های فازی با توابع عضویت gbellmf استفاده کردیم (شکل (۵-الف)). برای هر متغیر سه برچسب زبانی Small، Medium و Big مطابق شکل (۵-ب) وجود دارد. اشکال (۵-ج) و (۵-د) نیز محدوده‌های مقادیر زبانی و شکل (۵- ه)، قوانین فازی مورد استفاده برای حل مسأله درون‌یابی با توجه به توابع عضویت متغیرهای ورودی و خروجی را نشان می‌دهد.



الف) ب)

Range of linguistic values: [0 1].
Small: T (0, 0.3, 0.3, 0.3).
Medium: T (0, 0.3, 0.3, 1).
Big: T (0.6, 1, ∞ , ∞).

ج)

Range of linguistic values: [0 6].
Small: T (0, 2.5, 2.5, 2.5).
Medium: T (0, 2.5, 2.5, 6).
Big: T (3.5, 6, 6, 6).

د)

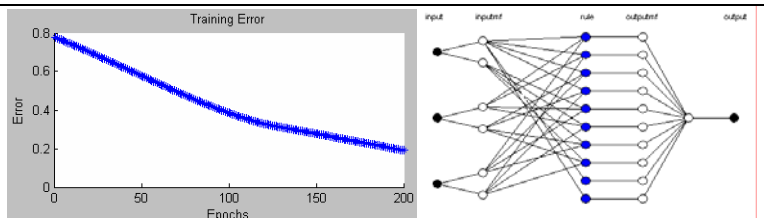
1. If [input1 is in1mf1] and [input2 is in2mf1] and and [input3 is in3mf1] then [output is out1mf1]
2. If [input1 is in1mf1] and [input2 is in2mf1] and and [input3 is in3mf2] then [output is out1mf2]
3. If [input1 is in1mf1] and [input2 is in2mf2] and and [input3 is in3mf1] then [output is out1mf3]
4. If [input1 is in1mf1] and [input2 is in2mf2] and and [input3 is in3mf2] then [output is out1mf4]
5. If [input1 is in1mf2] and [input2 is in2mf1] and and [input3 is in3mf1] then [output is out1mf5]
6. If [input1 is in1mf2] and [input2 is in2mf1] and and [input3 is in3mf2] then [output is out1mf6]
7. If [input1 is in1mf2] and [input2 is in2mf2] and and [input3 is in3mf1] then [output is out1mf7]
8. If [input1 is in1mf2] and [input2 is in2mf2] and and [input3 is in3mf2] then [output is out1mf8]

ه)

شکل ۵- الف) تابع عضویت gbellmf. ب) برچسب های متغیر زبانی Weigh ج) دامنه تغییرات متغیر های زبانی وزنها. د) دامنه تغییرات متغیر های زبانی تغییرات وزنها. ه) قوانین فازی مورد استفاده برای حل مسأله درون‌یابی با توجه به توابع عضویت متغیرهای ورودی و خروجی.

ساختار این شبکه و منحنی کاهش خطا در حین آموزش به ترتیب در اشکال (۶-الف) و (۶-ب) دیده می‌شوند. خطای این شبکه پس از ۲۰۰ بار تکرار به ۰,۱۹ می‌رسد.

همچنین خطای مجموعه‌ی آزمایشی در این شبکه ۰,۴ می‌باشد و این در حالی است که خطای شبکه‌ای که به روش پس‌انتشار خطا آموزش دیده است، ۰,۵ می‌باشد. خلاصه‌ی این نتایج در جدول ۱ آورده شده است.



ب الف

شکل ۶- الف)ساختارفازی مورد استفاده برای مدل. ب) منحنی کاهش خطا در جریان آموزش

همان‌گونه که در این جدول نیز مشاهده می‌شود، علی‌رغم استفاده از ابعاد شبکه‌ی کوچکتر در روش فازی، هم برای مجموعه‌ی آموزشی و هم برای مجموعه‌ی آزمایشی نتایج بهتری به‌دست آمده است.

روش آموزش	خطای آموزش	خطای آزمایش	اندازه‌ی مجموعه‌ی آموزشی	اندازه‌ی مجموعه‌ی آزمایشی	تعداد تکرار	تعداد عصب
مرسوم	0.4	0.5	216	125	600	67
فازی	0.19	0.4	216	125	200	8

جدول ۱) نتایج بدست آمده از دو روش فازی و فازی پیشنهادی برای شبکه.

٤. نتیجه‌گیری:

استفاده از روشهای فازی در آموزش شبکه‌های عصبی موجب استفاده از ابعاد کوچکتر و نتایج بهتری می‌شود. در این مقاله این کار برای شبکه‌های پرسپترون با یک لایه‌ی پنهان و کاربرد درون‌یابی، مورد مطالعه قرار گرفت، لیکن می‌توان آن را به‌سادگی به شبکه‌های چند لایه (و اصولاً شبکه‌های پیش‌رو) تعمیم داد. علاوه برفوق، به دلیل نزدیکی مفاهیم (مقادیر و قوانین) فازی به زبان انسانی به کارگیری این روشها برای کاربردهای دیگری به جز درون‌یابی (که صرفاً یک کاربرد محاسباتی محسوب می‌گردد)، نتایج خوبی به‌همراه داشته باشد. یکی از مسائل بسیار مهم استخراج خودکار قوانین فازی است که می‌تواند در مطالعات بعدی مورد توجه قرار گیرد.

٥. مراجع :

[1] David S. Touretzky, Artificial Neural Networks ; Spring 2004

[2] F.L. Chung, Wang Shitong, Deng Zhao Hong, Hu DeWine, Fuzzy kernel hyper ball ; Soft computing, 2004

[3] Simone Marinai, Marco Gori, and Giovanni Soda, Artificial Neural Networks for Document Analysis and Recognition,

[4] J.L. Castro, M. Delgado, C.J. Mantas, A fuzzy rule-based algorithm to train perceptrons.

[5] Phillip E Mitchell, Hong Yan ; Newspaper Document Analysis featuring Connected Line Segmentation, 2001 IEEE

[6] Connectionist Models: Basics, Srini Narayanan, CS182/CogSci110/Ling109 ; Spring 2005

[7] Jiu–Lun Fan, Wen-Zhi Zhen, Wei–Xin xie; Suppressed fuzzy c-means clustering algorithm; pattern Recognition Letters, 2003.

[۸] لی وانگ , ترجمه محمد تشنه لب،نیما صفاپور، داریوش افیونی؛ سیستمهای فازی و کنترل فازی ، ۱۳۸۰

[۹] محمد باقر منهاج، مبانی شبکه‌های عصبی(هوش محاسباتی)، ۱۳۸۱

[10] Pier Paolo D'Urso, Paolo Giordanib; A weighted fuzzy *c*-means clustering model for fuzzy data ; Computational Statistics & Data Analysis , 2004