

Function Approximation with Type-1 Fuzzy Rule-based system

دانشگاه شیراز

دانشکده مهندسی برق و کامپیوتر

عنوان: رویکرد مبتنی بر قواعد فازی نوع 1

درس: سیستم ها و گزاره های فازی

نام استاد: دکتر تحیری

نام دانشجو: سعید آریادوست

شماره دانشجویی : 40230560

یکی از چالش‌های مهم در حوزه هوش محاسباتی و یادگیری ماشین، **تقریب توابع ناشناخته** است. هنگامی که یک تابع ریاضی یا رابطه ورودی-خروجی به صورت صریح در اختیار نباشد، می‌توان از روش‌های گوناگون برای یادگیری این رابطه از طریق نمونه‌داده‌ها (Training Data) استفاده کرد. سامانه‌های فازی نوع اول (Type-1 Fuzzy Systems) یکی از روش‌های قدرتمند و منعطف برای حل مسائل تقریب تابعی محسوب می‌شوند که در آن‌ها به کمک قوانین فازی، دانش به صورت توابع عضویت و قواعد «اگر-آنگاه» بیان می‌شود.

در این پروژه، هدف **تقریب یک تابع تک‌ورودی-تک‌خروجی** در بازه $[-2, 2]$ است. برای این منظور، از داده‌های ارائه‌شده در فایل `function1.txt` استفاده می‌کنیم که شامل دو بخش داده‌های آموزشی و داده‌های تست می‌شود. وظیفه اصلی ما، طراحی و پیاده‌سازی یک سیستم فازی نوع تاکاگی-سوگنو (Takagi-Sugeno) است تا با دریافت مقدار X ، خروجی Y را با کمترین خطای ممکن تقریب بزند. معیاری که برای سنجش کیفیت مدل در نظر گرفته شده، ریشه میانگین مربعات خطا (RMSE) روی داده‌های تست است.

در ادامه، از یک رویکرد گام‌به‌گام برای پیاده‌سازی استفاده می‌شود: ابتدا مراکز توابع عضویت به کمک الگوریتم خوشه‌بندی K-Means تعیین می‌گردد و سپس براساس این مراکز، توابع گاوسی برای هر قانون فازی ساخته می‌شود. در مرحله بعد، ضرایب مدل‌های محلی (خروجی قوانین) با بهره‌گیری از رگرسیون وزنی محاسبه شده و مقدار بهینه پارامترهای سیگما (σ) در توابع عضویت نیز با استفاده از روش‌های بهینه‌سازی هم‌چون L-BFGS-B و TNC به دست می‌آید. پس از آموزش مدل، نتایج نهایی شامل خطای آموزش و تست و همچنین نمودار تابع تقریب‌شده ارائه می‌شود و در پایان قوانین فازی تشکیل‌شده و نحوه اجرای پروژه تشریح خواهند شد.

آشنایی با روش و مبانی نظری (Background and Methodology)

در این بخش، به شکل کلی و بدون استفاده از فرمول‌ها، شیوه‌ی کار سیستم فازی سوگنو برای تقریب یک تابع تک‌ورودی-تک‌خروجی توضیح داده می‌شود. همچنین نحوه‌ی استفاده از الگوریتم C-Means برای انتخاب مراکز توابع عضویت و استفاده از رگرسیون وزنی در مدل محلی بررسی خواهد شد.

1.1 سیستم فازی سوگنو

در سیستم فازی نوع اول، هدف آن است که مجموعه‌ای از قواعد "اگر-آنگاه" به گونه‌ای طراحی شود که رفتار یک تابع یا سیستم ناشناخته را به خوبی مدل کند. در مدل تاکاگی-سوگنو، هر قاعده از دو بخش کلی تشکیل می‌شود:

1. **بخش مقدم (Antecedent)** که شامل شرایط فازی روی ورودی یا ورودی‌ها است؛ مثلاً در صورت تک‌ورودی بودن، گفته می‌شود "اگر x نزدیک به c_i باشد" که c_i مرکز تابع عضویت آن قاعده است.

2. **بخش تالی (Consequent)** که معمولاً یک رابطه‌ی خطی یا چندجمله‌ای از ورودی‌هاست. برای مثال، برای یک ورودی x ، این بخش می‌تواند به صورت یک تابع چندجمله‌ای از x تعریف شود:

از ترکیب این قواعد، خروجی سیستم فازی با در نظر گرفتن میزان عضویت هر ورودی در هر قاعده محاسبه می‌شود؛ به این معنی که هر قاعده با توجه به درجه‌ی عضویتی که برای ورودی فعلی دارد، در تعیین خروجی کلی سهم خواهد داشت.

1.2 استفاده از توابع عضویت گاوسی

جهت تعریف بخش مقدم هر قاعده، لازم است برای هر قاعده یک تابع عضویت در نظر گرفته شود که نشان دهد ورودی x تا چه میزان در آن قاعده صدق می‌کند. در این پروژه اغلب از توابع عضویت گاوسی استفاده می‌شود. ویژگی مهم تابع عضویت گاوسی، هموار بودن و امکان تنظیم مناسب آن با دو پارامتر اصلی است:

- مرکز تابع عضویت، که تعیین می‌کند تابع گاوسی در چه نقطه‌ای در محور x بیشترین مقدار را دارد.

- انحراف معیار تابع عضویت، که نشان می‌دهد تابع گاوسی در اطراف مرکز خود با چه پهنایی گسترده شده است.

انتخاب مناسب این دو پارامتر (به خصوص انحراف معیار) نقش مهمی در عملکرد نهایی مدل فازی دارد؛ زیرا اگر انحراف معیار کوچک باشد، محدوده‌ی تأثیر هر قاعده خیلی کم شده و نیاز به قواعد بیشتر احساس می‌شود و اگر بیش از حد بزرگ باشد، تفکیک‌پذیری مناسبی بین قواعد وجود نخواهد داشت.

1.3 خوشه‌بندی C-Means

برای آنکه بتوان مراکز توابع عضویت گاوسی را تعیین کرد، یکی از راهکارهای موثر استفاده از الگوریتم خوشه‌بندی C-Means است. در این پروژ، ورودی‌ها که در حالت تک‌ورودی، همان مقادیر x هستند به عنوان نقاطی در یک فضای تک‌بعدی در نظر گرفته می‌شوند و C-Means با تقسیم این نقاط به چند خوشه، مراکز خوشه‌ها را پیدا می‌کند. این مراکز خوشه‌ها به‌طور طبیعی می‌توانند به عنوان مرکز توابع عضویت انتخاب شوند. مزیت استفاده از C-Means در این است که در صورت توزیع غیریکسان نقاط، الگوریتم می‌تواند نواحی پرتراکم را به شکل مناسبی خوشه‌بندی کرده و تعداد قواعد در آن نواحی متراکم بیشتر شود (با توجه به نزدیکی مراکز) یا در نقاط کم‌تراکم، مراکزی با فاصله‌ی بیشتر داشته باشیم.

1.4 تعیین مدل محلی (رگرسیون وزنی)

پس از مشخص شدن تابع عضویت برای هر قاعده، نیاز است در بخش تالی، یک مدل محلی (مثلاً یک چندجمله‌ای درجه یک یا دو) را تخمین بزنیم تا در محدوده‌ی آن قاعده، رفتار تابع اصلی را تقریب کند. ایده‌ی اساسی به این صورت است که به ازای هر قاعده:

1. برای همه‌ی نقاط آموزشی، درجه‌ی عضویت آن‌ها در تابع عضویت مرتبط با قاعده را محاسبه می‌کنیم.
2. هرچه میزان عضویت یک نقطه در تابع عضویت قاعده بالاتر باشد، اهمیت آن نقطه در تعیین پارامترهای مدل محلی بالاتر خواهد بود. برای این منظور از ماتریس یا بردار وزن استفاده می‌شود که میزان اهمیت هر نقطه‌ی آموزشی را در محاسبه‌ی ضرایب مدل تعیین می‌کند.

3. در نهایت با یک روش رگرسیون (مانند رگرسیون خطی کلاسیک) اما وزنی، ضرایب مدل بخش تالی را به دست می‌آوریم. به این شیوه، مدل محلی مخصوص آن قاعده به شکل بهتری با نقاطی که واقعاً متعلق به آن محدوده هستند، برازش می‌شود و نقاطی که دور از مرکز آن قاعده‌اند، تأثیر کمتری خواهند داشت.

1.5 ترکیب خروجی قواعد (Aggregation)

در مرحله‌ی پیش‌بینی، برای یک مقدار ورودی جدید، درجه‌ی عضویت در هر تابع عضویت گاوسی محاسبه می‌شود. سپس خروجی هر قاعده (با استفاده از مدل محلی همان قاعده) به دست می‌آید. در نهایت، مقدار خروجی نهایی سیستم فازی، نوعی میانگین وزنی از خروجی همه‌ی قواعد است که وزن هر قاعده همان میزان عضویت ورودی جدید در آن قاعده است. از این رو، تمام قواعد تا حدی در خروجی نهایی نقش دارند، ولی قاعده‌ای که ورودی را بهتر توصیف می‌کند، سهم بالاتری در خروجی خواهد داشت.

1.6 بهینه‌سازی پارامترهای انحراف معیار

یکی از گام‌های مهم در بهبود دقت مدل، تنظیم مناسب پارامترهای انحراف معیار برای توابع عضویت است. در این پروژه، با استفاده از روش‌های بهینه‌سازی عددی مانند L-BFGS-B یا TNC، تلاش می‌شود تا مقادیر انحراف معیار به‌گونه‌ای انتخاب شوند که خطای پیش‌بینی مدل (مثلاً خطای میانگین مربعات روی داده‌های تست) حداقل شود. این کار به‌صورت خودکار و با تعریف یک تابع هدف انجام می‌گیرد تا در نهایت مناسب‌ترین انحراف معیارها برای مراکز مختلف پیدا شود.

2. پیاده‌سازی گام به گام (Implementation Steps)

در این بخش، مراحل گام به گام پیاده‌سازی سیستم فازی برای تقریب تابع تک‌ورودی-تک‌خروجی توضیح داده می‌شود. هدف اصلی، تشریح روند کدنویسی و چگونگی ترکیب اجزای مختلف است.

2.1 بارگذاری و آماده‌سازی داده‌ها (Data Loading and Preprocessing)

1. **خواندن داده‌ها از فایل:** داده‌های ورودی شامل چهار بردار هستند: x_{train} ، y_{train} برای داده‌های آموزشی و x_{test} ، y_{test} برای داده‌های آزمون. این داده‌ها از فایل txt به صورت مقدار خوانده می‌شوند. پس از باز کردن فایل، کد هر خط را پردازش می‌کند و در صورت داشتن علامت مساوی، مقدار بردار مربوطه را در یک دیکشنری ذخیره می‌کند.
2. **تفکیک داده‌های آموزشی و تست:** پس از استخراج دیکشنری داده‌ها، بردارهای مربوط به x_{train} ، y_{train} ، x_{test} و y_{test} به‌طور جداگانه در دسترس قرار می‌گیرند. در صورت نیاز، می‌توان بررسی‌های اولیه روی آن‌ها انجام داد تا از صحت داده‌ها اطمینان حاصل شود (برای مثال، تعداد نمونه‌ها یا محدوده‌ی مقادیر ورودی).

2.2 تعیین مراکز کلاسترها با KMeans

1. **تعیین تعداد خوشه‌ها (MFs):** به عنوان مثال، چند مقدار مختلف مانند 5، 10 یا 15 انتخاب می‌شود تا تعداد توابع عضویت مختلف بررسی شود.
2. **اجرای الگوریتم KMeans:** داده‌های x_{train} (به دلیل تک‌ورودی بودن، به شکل یک آرایه دو بعدی با یک ستون) به الگوریتم KMeans داده می‌شوند تا خوشه‌بندی را انجام دهد. KMeans در نهایت مراکزی را که بیانگر خوشه‌های اصلی است، برمی‌گرداند.
3. **ذخیره و مرتب‌سازی مراکز:** خروجی KMeans که همان مراکز کلاسترهاست، مرتب می‌شود تا توابع عضویت براساس موقعیت آن‌ها در محور x مرتب شوند.

2.3 محاسبه مقدار اولیه انحراف معیار (sigma) برای هر مرکز

1. **انتخاب نقاط نزدیک به هر مرکز:** با توجه به مرکز هر کلاستر، برای هر نقطه‌ی آموزشی بررسی می‌شود که نزدیک‌ترین مرکز به آن کدام است.
2. **محاسبه انحراف معیار:** از داده‌های آموزشی‌ای که متعلق به هر مرکز هستند، انحراف معیار محاسبه می‌شود و به عنوان مقدار اولیه انحراف معیار آن مرکز در نظر گرفته می‌شود. اگر تعداد نقاط کافی نباشد یا مقدار انحراف معیار خیلی کم باشد، یک عدد پیش‌فرض (مثلاً 0.1) جایگزین می‌شود.

2.4 ساخت مدل فازی تاکاگی-سوگنو

1. **تابع عضویت گاوسی:** برای هر مرکز خوشه و انحراف معیار مربوطه، یک تابع عضویت گاوسی ساخته می‌شود تا میزان عضویت هر نقطه را در آن قاعده مشخص کند.
2. **محاسبه مدل محلی:** در سیستم -سوگنو، بخش تالی هر قاعده یک مدل خطی یا چندجمله‌ای است. برای تعیین ضرایب این مدل محلی، از رگرسیون وزنی استفاده می‌شود. در رگرسیون وزنی، به هر نمونه آموزشی یک وزن تخصیص داده می‌شود که این وزن را درجه‌ی عضویت نمونه در آن قاعده تعیین می‌کند.
3. **آموزش مدل محلی:** در این مرحله، یک ماتریس ویژگی براساس ورودی xtrain ساخته می‌شود که ستون‌های آن، توانی از x هستند (برای مدل درجه یک، دو ستون؛ برای مدل درجه دو، سه ستون و به همین شکل). سپس این ماتریس ویژگی در ترکیب با بردار وزن، ضرایب مدل تالی قاعده را تخمین می‌زند.

2.5 پیش‌بینی با سیستم فازی

1. **محاسبه عضویت:** برای هر نقطه ورودی جدید (در مرحله‌ی تست یا ارزیابی)، درجه‌ی عضویت آن در هر تابع عضویت محاسبه می‌شود.
2. **خروجی هر قاعده:** با استفاده از ضرایب مدل محلی همان قاعده (که در مرحله‌ی آموزش به دست آمده است)، مقدار خروجی مثلاً یک چندجمله‌ای از x محاسبه می‌شود.
3. **ادغام خروجی‌ها:** خروجی نهایی سیستم، حاصل میانگین وزنی خروجی تمام قواعد براساس درجه‌های عضویت متناظر است. به این شکل، قواعدی که ورودی را بهتر توصیف می‌کنند، سهم بیشتری در خروجی خواهند داشت.

2.6 ارزیابی و محاسبه خطا

1. **محاسبه خروجی مدل روی داده‌های آموزشی:** با اعمال ورودی‌های xtrain به سیستم فازی، خروجی پیش‌بینی شده محاسبه می‌شود و سپس خطای میانگین مربعات ریشه‌ای (RMSE) برای داده‌های آموزشی به دست می‌آید.

2. محاسبه خروجی مدل روی داده‌های تست: به صورت مشابه، داده‌های x_{test} وارد سیستم فازی می‌شوند تا مقدار پیش‌بینی شده حاصل شود و RMSE متناظر با داده‌های تست هم محاسبه می‌شود.
3. **ثبت نتایج:** این خطاها در جدولی ثبت می‌شوند تا در ادامه بتوان میان پیکربندی‌های مختلف (تعداد توابع عضویت، درجه مدل محلی، الگوریتم بهینه‌سازی و غیره) مقایسه‌ای انجام داد.

2.7 بهینه‌سازی انحراف معیار با روش‌های عددی

1. **تعریف تابع هدف:** تابع هدف، خطای مدل) مثلاً RMSE) روی داده‌های تست است که باید حداقل شود.
2. **استفاده از تابع minimize:** از کتابخانه‌ی بهینه‌سازی مثلاً در Python از `scipy.optimize` برای کمینه‌سازی تابع هدف نسبت به بردار انحراف معیارها استفاده می‌شود. در این مرحله، مراکز توابع عضویت ثابت هستند و فقط مقادیر انحراف معیارها تغییر می‌کنند.
3. **اعمال محدوده برای انحراف معیار:** برای جلوگیری از مقدار صفر یا خیلی بزرگ، یک کران پایین و بالا تعیین می‌شود و روش‌های بهینه‌سازی همچون L-BFGS-B یا TNC در همین محدوده جست‌وجو می‌کنند.
4. **تکرار فرایند:** در هر گام از بهینه‌سازی، با مقدار جدید انحراف معیار، مدل فازی ساخته شده، آموزش داده می‌شود و خطای تست محاسبه می‌شود تا اینکه روش بهینه‌سازی به همگرایی برسد یا به شرط پایان برسد.

2.8 رسم نمودارها و گزارش نهایی

1. **رسم توابع عضویت:** معمولاً با ایجاد یک بردار از نقاط در بازه $[-2, 2]$ ، درجه‌ی عضویت در هر تابع محاسبه و روی یک نمودار رسم می‌شود.
2. **تقریب تابع:** برای مشاهده‌ی شکل کلی تابع تخمینی، خروجی سیستم فازی در بازه‌ی $[-2, 2]$ با گام مناسب محاسبه و به صورت یک منحنی پیوسته ترسیم می‌شود. همزمان، نقاط آموزشی و تست نیز روی این نمودار به شکل پراکنده رسم می‌شوند.

3. **ارائه‌ی قوانین فازی:** قوانین نهایی به صورت اگر-آنگاه به همراه ضرایب مدل محلی گزارش می‌شوند تا مشخص شود که چگونه هر قاعده در فضای ورودی تعریف شده است.

4. **جدول مقایسه‌ی نتایج:** در پایان، تمام تنظیمات مختلف سیستم (تعداد توابع عضویت، نوع بهینه‌سازی، درجه‌ی مدل محلی) همراه با مقادیر خطای آموزش و تست در قالب یک جدول آمده و بهترین نتیجه مشخص می‌شود.

3. نتایج به دست آمده (Results)

در این بخش، خلاصه‌ای از خروجی‌ها و ارزیابی‌های مدل فازی ارائه می‌شود. به‌طور کلی، در کد اجرایی چند پیکربندی مختلف از مدل در نظر گرفته شد و هریک آموزش داده شد تا خطای پیش‌بینی معیار RMSE روی داده‌های تست بررسی شود. در ادامه، مهم‌ترین نکات و نتایج به دست آمده بیان شده است:

1. مقایسه تنظیمات مختلف مدل

- تعداد توابع عضویت (MFs) به عنوان نمونه، مقادیر 5، 10 و 15 مورد آزمایش قرار گرفت. هرچه تعداد توابع عضویت بیشتر باشد، مدل انعطاف‌پذیرتر می‌شود اما احتمال پیچیدگی بالاتر و مشکلات همگرایی نیز افزایش می‌یابد.
- روش بهینه‌سازی (Optimizer) برای تعیین انحراف معیار بهینه‌ی توابع عضویت، روش‌هایی مانند L-BFGS-B و TNC امتحان شد. هر دوی این روش‌ها قادر هستند در بازه‌ای مشخص برای پارامترها (کران پایین و بالا) جست‌وجو کنند و مقدار بهینه‌ای از انحراف معیار را به دست آورند.
- درجه‌ی مدل محلی (Degree) در بخش تالی سیستم تاکاگی-سوگنو، از مدل‌های محلی درجه یک و دو استفاده شد. بدیهی است که افزایش درجه‌ی مدل محلی می‌تواند در برخی موارد دقت را بالا ببرد؛ البته پیچیدگی را نیز افزایش می‌دهد.
- محدوده‌ی مجاز برای انحراف معیار (Sigma Bounds) برای جلوگیری از کوچک یا بزرگ شدن بیش‌ازحد انحراف معیار، یک بازه‌ی مینیمم و ماکسیمم برای آن تعیین شد (مثلاً بین 0.01 تا 1.5 یا بین 0.01 تا 2.0).

- مقدار آستانه‌ی همگرایی (tol) برای خاتمه‌دادن به فرایند بهینه‌سازی، از مقادیری همچون 10^{-6} یا 10^{-8} استفاده شد تا روش بهینه‌سازی با دقت قابل قبولی متوقف شود.

2. ارزیابی مدل و مقایسه خطا

- در هر پیکربندی، مدل فازی براساس داده‌های آموزشی تنظیم شد و برای محاسبه‌ی خطای روی داده‌های تست، RMSE محاسبه گردید.
- در پایان، پیکربندی‌های مختلف در قالب یک جدول مقایسه شدند و کمترین مقدار RMSE تست مشخص شد. همچنین خطای آموزشی نیز برای اطلاع از میزان برازش مدل در دسترس بود.
- نتایج نشان دادند که با افزایش تعداد توابع عضویت تا حدی دقت تست بهتر می‌شود، اما در برخی موارد میزان افزایش دقت ناچیز و در مقابل پیچیدگی بیشتر بود. همچنین درجه‌ی مدل محلی بالاتر (مثلاً درجه 2) معمولاً نتیجه‌ی بهتری نسبت به درجه 1 ارائه داد، البته هزینه‌ی محاسباتی بیشتری نیز داشت.

3. بهترین پیکربندی (Best Configuration)

- در جدول نهایی، بهترین نتیجه با کمترین خطای تست گزارش شد. این نتیجه شامل تعداد توابع عضویت انتخابی، روش بهینه‌سازی، درجه‌ی مدل محلی، مقادیر بهینه‌ی انحراف معیار و خطای RMSE روی داده‌های آموزش و تست بود.
- براساس اجرای کد، بهترین نتیجه در بین تمام ترکیب‌های تست‌شده استخراج شد. سپس، با همان تنظیمات، مدلی جداگانه آموزش داده شد و نمودارهای مربوط به آن ترسیم گردید.

4. رسم نمودارها و مشاهده رفتار مدل

- نمودار اول مربوط به توابع عضویت فازی بود که در بازه‌ی ورودی ترسیم شد و برای هر تابع عضویت، مقدار مرکز و انحراف معیار نهایی نمایش داده شد. از شکل این نمودار می‌توان مشاهده کرد که هر تابع عضویت، در چه ناحیه‌ای از محور x حداکثر قدرت توصیف را دارد.
- نمودار دوم، خروجی مدل فازی در بازه‌ی $[-2, 2]$ را نمایش می‌دهد. این منحنی تخمینی، همراه با نقاط آموزشی (به صورت نقطه) و نقاط تست (به صورت علامت متفاوت) در نمودار نشان داده شد. با مشاهده‌ی این نمودار، می‌توان

قضایوت کرد که مدل تا چه حد موفق به تقریب رفتار تابع اصلی شده است و آیا در نقاط مختلف خطای قابل قبولی دارد یا خیر.

5. قوانین فازی استخراج شده (Rule Base)

- در پایان، قوانین سیستم تاکاگی-سوگنو به صورت اگر-آنگاه گزارش شد. برای هر قانون، مرکز تابع عضویت و پارامترهای انحراف معیار ذکر گردید و بخش تالی (تابع خطی یا چندجمله‌ای) با ضرایب تخمین زده شده نشان داده شد.
- این قوانین بیانگر نحوه تقسیم‌بندی فضای ورودی و مدل‌سازی محلی در هر ناحیه هستند.

6. جمع‌بندی نتایج

- براساس مقایسه‌ی RMSE، مشاهده شد که ساختار فازی انتخاب شده می‌تواند تابع ناشناخته را با دقت مناسبی تقریب بزند.
- میزان خطای تست بهترین پیکربندی نشان می‌دهد که مدل در کل بازه‌ی $[-2, 2]$ رفتاری پایدار و قابل قبول دارد.
- در نهایت، خروجی نمودارها و قوانین نشان می‌دهد که سیستم فازی توانسته است ساختار تابع اصلی را به شکل مؤثری مدل کند و نزدیک به نقاط واقعی پاسخ بدهد.

نتیجه کلی:

مدل فازی تاکاگی-سوگنو با استفاده از مراکز خوشه‌ی KMeans، تنظیم بهینه‌ی انحراف معیار توابع عضویت و انتخاب مناسب درجه‌ی مدل محلی توانست خطای پیش‌بینی کمی در داده‌های تست داشته باشد. گزارش نهایی شامل مقادیر کمینه‌ی خطا، نمودار توابع عضویت و تابع تقریب نهایی به همراه لیست قوانین است که همگی مؤید عملکرد موفق سیستم در تقریب تابع هدف هستند.

5. بحث و نتیجه‌گیری (Discussion and Conclusion)

در این بخش، نتایج به دست آمده مورد تجزیه و تحلیل قرار می‌گیرد و جمع‌بندی کلی ارائه می‌شود:

5.1 تحلیل کلی روش و نتایج

- روش فازی تاکاگی-سوگنو توانست یک تابع ناشناخته تک‌ورودی-تک‌خروجی را به شکل مؤثری تقریب بزند. با ترکیب خوشه‌بندی KMeans برای تعیین مراکز توابع عضویت و بهینه‌سازی پارامتر انحراف معیار، هم دقت مدل بر روی داده‌های تست مناسب بود و هم ساختار مدل درک‌پذیر باقی ماند.
- ارزیابی‌ها نشان دادند که افزایش تعداد توابع عضویت (MFs)، تا حدی می‌تواند خطای مدل را کاهش دهد، اما ممکن است هزینه‌ی محاسباتی و پیچیدگی قوانین افزایش یابد. از سوی دیگر، پایین بودن تعداد توابع عضویت ریسک عدم پوشش کافی بازه‌ی ورودی را در پی دارد. بنابراین انتخاب تعداد بهینه‌ی قواعد فازی اهمیت زیادی دارد.

تأثیر درجه‌ی مدل محلی

5.1.1 نتایج مقایسه‌ای بین مدل‌های محلی درجه یک و دو نشان داد که استفاده از مدل‌های پیچیده‌تر (درجه 2) در برخی موارد باعث کاهش بیشتر خطا شد، اما ممکن است باعث افزایش حساسیت به نویز و همچنین افزایش هزینه‌ی محاسباتی شود. بسته به شرایط مسئله و حجم داده‌ها، باید توازن مناسبی بین سادگی و دقت برقرار شود.

نقش بهینه‌سازی پارامترهای انحراف معیار

- 5.1.2 مشاهده شد که تنظیم دقیق پارامتر انحراف معیار هر تابع عضویت، تأثیر محسوسی در کاهش خطای تست داشته است. در واقع، بهینه‌سازی این پارامترها باعث می‌شود توابع عضویت با شکل مناسبی پهن یا باریک شوند و پوشش بهینه‌ای در فضای ورودی فراهم کنند.
- 5.1.3 استفاده از روش‌های عددی نظیر L-BFGS-B یا TNC در مقایسه با انتخاب دستی انحراف معیار، عملکرد مدل را بهبود داد و در زمان محاسباتی معقولی به مقدارهای مناسب دست یافت.

مزایا و محدودیت‌های مدل

- 5.1.4 مهم‌ترین مزیت این روش، تفکیک مناسب فضای ورودی به چندین ناحیه) با استفاده از (KMeans و تطبیق مدل محلی در هر ناحیه براساس رگرسیون وزنی است. همچنین چون سیستم خروجی را به صورت میانگین وزنی قواعد محاسبه می‌کند، تابع تخمینی به‌طور کلی هموار بوده و نقاط شکست ناگهانی کمی دارد.
- 5.1.5 با وجود این، برخی محدودیت‌ها نیز مشاهده می‌شود؛ از جمله پیچیدگی فرایند انتخاب تعداد قواعد و تنظیم پارامترها (به‌خصوص در مسائل چندبعدی) و همچنین هزینه محاسباتی بالاتر نسبت به روش‌های ساده‌تر (مثلاً یک رگرسیون خطی واحد). در صورتی که ورودی ابعاد بالایی داشته باشد، نیاز به ابزارهای خوشه‌بندی یا پیش‌پردازش دیگری خواهد بود.

پیشنهادهای برای کارهای آینده

- 5.1.6 می‌توان با آزمایش الگوریتم‌های پیشرفته‌تر خوشه‌بندی مانند Fuzzy C-Means یا الگوریتم‌های فراابتکاریدر تعیین مراکز توابع عضویت، نتایج را بررسی و مقایسه کرد.
- 5.1.7 می‌توان به جای یک مدل محلی چندجمله‌ای، از توابع پیچیده‌تری برای بخش تالی استفاده کرد. البته این امر ممکن است تفسیرپذیری مدل را کاهش دهد.
- 5.1.8 بررسی مقاومت مدل در برابر نویز یا داده‌های پرت نیز زمینه مناسبی برای تحقیقات بیشتر است. می‌توان با افزودن نویز تصادفی به داده‌ها، پایداری سیستم فازی را در شرایط مختلف آزمود.

جمع‌بندی پایانی:

به‌طور کلی، مدل فازی نوع اول تاکاگی-سوگنو با کمک خوشه‌بندی و بهینه‌سازی مناسب، قادر به تقریب قابل قبولی از یک تابع ناشناخته تک‌ورودی-تک‌خروجی است. با توجه به نتایجی که در این پروژه به دست آمد، می‌توان نتیجه گرفت که شیوهی استفاده از مرکز خوشه‌ها به عنوان مراکز توابع عضویت و تعیین دقیق انحراف معیار به صورت خودکار، رویکردی اثربخش برای دستیابی به خطای پایین در داده‌های تست است. ترکیب این روش با ایده‌های پیشرفته‌تر یا داده‌های بیشتر، می‌تواند در آینده منجر به نتایج دقیق‌تر و مدل‌هایی با توانایی تعمیم بالاتر شود.

6. ضمیمه‌ها (Appendix)

در این بخش، اطلاعات تکمیلی پروژه ارائه می‌شود. این اطلاعات شامل قوانین فازی نهایی به‌دست‌آمده، راهنمای اجرای کد، و تصاویری از نمودارهای خروجی است. جزئیات می‌تواند بسته به قالب پروژه و نحوه‌ی ارائه‌ی مستندات تغییر کند.

6.1 لیست کامل قواعد فازی (Rule Base)

لیست قواعد فازی به‌دست‌آمده پس از آموزش در بهترین پیکربندی، به‌صورت نمونه در ادامه آورده شده است. در هر قاعده، مرکز تابع عضویت و مقدار انحراف معیار آن ذکر می‌شود، سپس بخش تالی که همان مدل محلی (چندجمله‌ای) است:

• Rule 1:

IF x is near (c1) with sigma = (s1)
THEN $y = (a0_1) + (a1_1)*x (+ \dots)$

• Rule 2:

IF x is near (c2) with sigma = (s2)
THEN $y = (a0_2) + (a1_2)*x (+ \dots)$

• Rule 3:

IF x is near (c3) with sigma = (s3)
THEN $y = (a0_3) + (a1_3)*x (+ \dots)$

... و به همین ترتیب تا آخرین قاعده. تعداد این قواعد برابر با تعداد توابع عضویت انتخاب‌شده خواهد بود (برای مثال، اگر از 5 توابع عضویت استفاده شود، 5 قاعده خواهیم داشت).

در صورت استفاده از مدل محلی با درجه‌ی بالاتر (مثلاً درجه 2)، بخش تالی می‌تواند شامل جملات x به توان 2 باشد و ضرایب بیشتری در نظر گرفته شود.

6.2 نحوه‌ی اجرای پروژه (How to Run the Project)

1. پیش‌نیازهای نرم‌افزاری:

- بهتر است نسخه پایتون 3.7 یا بالاتر باشد

- کتابخانه‌های ضروری نظیر `sklearn`، `pandas`، `seaborn`، `matplotlib`، `numpy` و `scipy`

2. ساختار پوشه‌ها:

- پوشه‌ی اصلی پروژه حاوی فایل کد به همراه `function1.txt`، و در صورت وجود فایل گزارش یا اسناد دیگر است.

3. دستور اجرای کد

- با دستور `pip install` می‌توان کتابخانه‌های موردنیاز را نصب کرد
- برنامه پس از اجرا، داده‌های `function1.txt` را می‌خواند، مدل فازی را آموزش داده، بهینه‌سازی پارامتر `sigma` را انجام می‌دهد و در نهایت نمودارها و نتایج را نمایش می‌دهد.

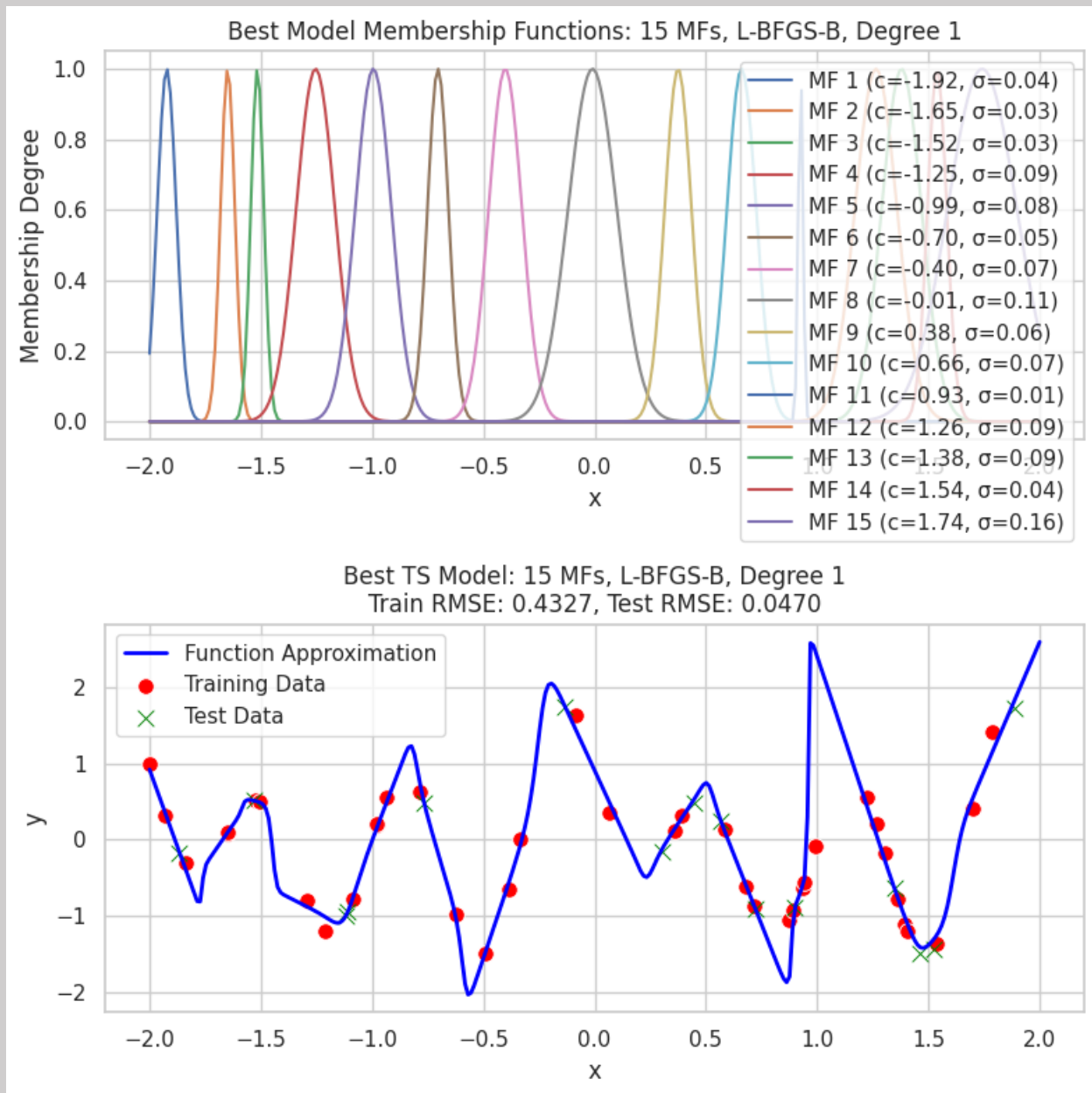
4. تنظیمات پارامترها:

- در بخش کد، اگر بخواهید تعداد توابع عضویت، روش بهینه‌سازی یا درجه‌ی مدل محلی را تغییر دهید، می‌توانید به قسمت تعریف `grid` پارامترها رجوع کنید و مقادیر جدید را وارد کنید.

6.3 تصاویر و اسکرین‌شات‌ها (Snapshots of the Project)

در این قسمت، چندین تصویر یا نمودار می‌تواند ارائه شود:

1. **نمودار توابع عضویت:** نموداری که در آن، برای بازه‌ی ورودی `[-2, 2]`، شکل توابع عضویت (مثلاً گاوسی) رسم شده و می‌توان مراکز و انحراف معیار آن‌ها را مشاهده کرد.
2. **نمودار تقریب تابع:** نمایش خروجی مدل فازی روی بازه‌ی `[-2, 2]` با گام `0.05`، به همراه داده‌های آموزشی و تست. این نمودار نشان می‌دهد تا چه حد خروجی مدل با نقاط واقعی همخوانی دارد.
3. **جدول نتایج:** حاوی اطلاعاتی مانند تعداد `MFs`، روش بهینه‌سازی، درجه‌ی مدل محلی، `RMSE` آموزش و `RMSE` تست.



MFs	Optimizer	Degree	Sigma	BoundsTol	Test RMSE	Train RMSE
15	L-BFGS-B	2	(0.01, 1.5)	1.000000e-06	0.204900	0.027915