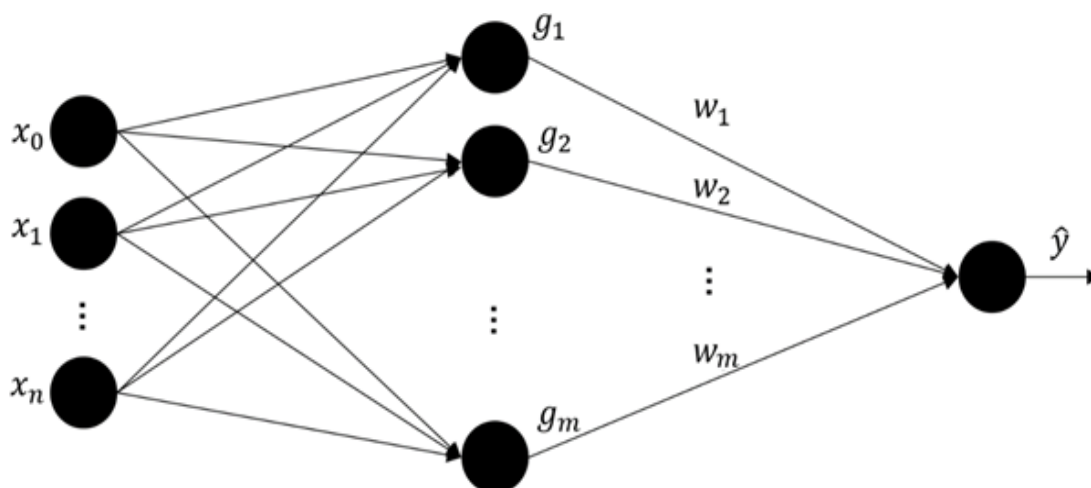


هدف پروژه:

پیاده سازی شبکه عصبی مصنوعی RBF و آموزش آن با استفاده از fcm برای طبقه بندی چند کلاسه

تعریف پروژه:

در این پروژه قصد داریم شبکه عصبی rbf را با استفاده الگوریتم خوشه بندی فازی fcm آموزش دهیم و از آن برای طبقه بندی چند کلاسه استفاده نماییم.



شکل بالا، شبکه عصبی rbf را نشان میدهد که در آن x_0 تا x_n داده‌های اولیه هستند.

لایه اول شبکه عصبی به صورت یک تابع radial Basis عمل میکند و فضای x ها را به فضای G تبدیل میکند. با فرض اینکه مجموعه داده n تایی و داده‌ها d بعدی باشند و m مرکز خوشه با الگوریتم fcm به دست آمده باشند (در قدم اول قبل از آموزش شبکه عصبی، باید با استفاده از خوشه بندی فازی، fcm مرکز خوشه‌ها را به دست آورید)، ماتریس G به این صورت خروجی‌های لایه اول را نشان می‌دهد:

$$G = \begin{bmatrix} g_1(X^1) & \cdots & g_m(X^1) \\ \vdots & \ddots & \vdots \\ g_1(X^n) & \cdots & g_m(X^n) \end{bmatrix}$$

درایه سطر k م و ستون i م از ماتریس G بدین صورت تعریف می‌شود:

$$g_i(X_k) = e^{-\gamma (X_k - V_i)^T C_i^{-1} (X_k - V_i)}$$

که در آن γ شعاع نظیر مرکز خوشه، V_i مرکز خوشه i م که توسط fcm به دست آمده است و ماتریس کواریانس C یک ماتریس مربعی $d \times d$ است بدین صورت محاسبه می‌شود:

$$C_i = \frac{\sum_{k=1}^{Ndata} (u_{ki})^m (X_k - V_i)(X_k - V_i)^T}{\sum_{k=1}^{Ndata} u_{ki}^m}$$

که در آن u_{ki} تعلق داده‌ی X_k به خوشه‌ی V_i می‌باشد. $(X_k - V_i)$ یک بردار $d \times 1$ می‌باشد. همچنین

توجه نمایید که به تعداد خوشه‌ها ماتریس C داریم، یعنی C_1, C_2, \dots, C_M .

حال در ادامه با تبدیل خطی می‌توان خروجی \hat{Y} را به صورت زیر محاسبه کرد:

$$\hat{Y} = \text{argmax}(G \times W)$$

ماتریس وزن‌ها، W را به این صورت محاسبه نماییم:

$$W = (G^T G)^{-1} G^T Y$$

که در آن ماتریس Y یک ماتریس $n \times c$ است که n تعداد داده‌ها و c تعداد کلاس‌هاست و به صورت زیر تعریف می‌شود:

$$\text{label}(x_i) = y \Rightarrow Y[i][y] = 1, \quad Y[i][j \neq y] = 0$$

در واقع اگر داده x_i مثلاً متعلق به کلاس ۳ باشد، سومین خانه از سطر i م (با فرض شروع از درایه صفر، درایه ۲) ماتریس Y ، یک و بقیه خانه‌های سطر صفر هستند.

پس از به دست آوردن ماتریس G و W ، با توجه به اینکه G یک ماتریس $n \times m$ و W یک ماتریس $m \times c$ می‌باشد، ماتریس حاصل از $G \times W$ یک ماتریس $n \times c$ می‌باشد.

حال بردار \hat{Y} که همان بردار برچسب‌های داده‌هاست، ایندکس‌های بزرگترین مقدار هر سطر از ماتریس $G \times W$ است.

برای مثال اگر دو داده ی سه کلاسه را در نظر بگیریم، اگر ماتریس حاصل ضرب گفته شده به صورت زیر باشد:

$$G \times W = \begin{bmatrix} 1.5 & 2.5 & 3.5 \\ 7.5 & 2.1 & 0.4 \end{bmatrix}$$

در سطر اول 3.5 و در سطر دوم 7.5 بیشترین مقادیر هستند که به ترتیب ایندکس 3 و 1 را دارند. بنابراین:

$$\hat{y} = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$$

یعنی شبکه عصبی به داده اول برچسب 3 و به داده دوم برچسب 1 داده است.

محاسبه دقت خروجی:

دقت خروجی شبکه عصبی، نسبت تعداد برچسب‌های درست به کل داده‌هاست.

به عبارت دیگر:

$$accuracy(y, \hat{y}) = 1 - \frac{\sum (abs(sign(y - \hat{y})))}{n}$$

(sign، تابع علامت است که مقادیر مثبت را به یک و مقادیر منفی را به منفی یک و صفر را به صفر نگاشت می‌کند)

پارامترهای مسئله

همانطور که در طول تعریف پروژه بیان شد، پارامترهای مسئله عبارتند از:

- **m: تعداد خوشه‌ها** برای خوشه‌بندی فازی.

تعداد کلاسترهای بهینه لزوماً با تعداد کلاس‌های داده‌ها برابر نیست. یعنی داخل مجموعه داده‌های

دو کلاسه لزوماً جواب بهینه‌ی مسئله با دو مرکز به دست نمی‌آید.

برای به دست آوردن تعداد مراکز بهینه، می‌توان بعد از مرحله‌ی خوشه‌بندی و با توجه به دقت شبکه

عصبی عمل کرد. مثلاً داخل یک حلقه تعداد کلاسترها را با افزایش گام یکی یکی یا دوتا دوتا از ۲ تا

۳۰ تغییر بدهید (این اعداد صرفاً مثالند هر حدی پایین و بالایی و گامی که بنظرتان منطقی هست رو می‌توانید در نظر بگیرید) و از این طریق تعداد خوشه‌های بهینه را پیدا کنید.

- γ : شعاع خوشه‌ها. این پارامتر را می‌توانید برای بار اول 0.1 در نظر بگیرید و با افزایش یا کاهش گاما تاثیر آن را روی دقت به دست آمده ببینید.

نحوه آموزش شبکه عصبی:

دیتاست را به دو بخش داده‌های آموزشی و داده‌های تست تقسیم نمایید. به این منظور 70% از دیتاست را به داده‌های آموزش و 30% را به داده‌های تست اختصاص دهید. بدیهی است که تقسیم بندی باید به صورت رندوم انجام شود. همانطور که کامل‌تر بیان شد، داده‌های آموزشی را با توجه به پارامتر m خوشه‌بندی کرده و سپس با استفاده از مراکز خوشه‌ها و محاسبه ماتریس G و W شبکه عصبی خود را آموزش دهید. سعی کنید پارامترهای مسئله را طوری تغییر دهید که بیشترین درصد دقت را از داده‌های آموزشی بگیرید.

تست شبکه عصبی:

حال با m مرکز خوشه به دست آمده از مرحله آموزش، ماتریس G' را برای داده‌های تست حساب به دست آورده و با محاسبه‌ی $\text{argmax}(G' \times W)$ که برچسب‌های داده تست می‌باشد، دقت را برای داده‌های تست به دست آورید. دقت کنید که ماتریس W را از مرحله آموزش داریم. علاوه بر این، لازم است که پس از برچسب‌زنی داده‌های تست، آنها را نمایش دهید؛ به این صورت که داده‌هایی که در یک کلاس هستند و درست برچسب‌گذاری شده‌اند با یک رنگ نمایش داده شوند (و طبیعتاً رنگ کلاس‌ها با هم تفاوت دارد) و تمام داده‌هایی که برچسب نادرست گرفته‌اند با یک رنگ (مثلاً قرمز) رسم شوند. علاوه بر برچسب گذاری‌ها، **مرکزهای خوشه‌ها** هم رسم شوند.

گزارش پروژه

شامل موارد زیر است:

- پارامترهای مسئله و دقت‌های به دست آمده:

تعداد خوشه‌ها (m) و اندازه شعاع مرکزها (γ). مسئله را با مقادیر مختلف پارامترها حل کنید تا به بهترین دقت آموزش و تست برسید. دقت آموزش و تست را گزارش نمایید.

- مسئله را برای حالتی حل کنید که با شعاع مرکز 0.1، تعداد خوشه‌ها را زیاد در نظر بگیریم (برای مثال 40 خوشه). دقت برای داده‌های آموزش و تست را گزارش کرده و تحلیل کنید.

- مسئله را برای حالتی حل کنید که تعداد خوشه‌ها را 3 و شعاع مرکزها را 1 در نظر بگیریم. دقت برای داده‌های آموزش و تست را گزارش کرده و تحلیل کنید.

- رسم نتایج:

نتایج مسئله (طبقه‌بندی داده‌ها و مراکز خوشه‌ها و مرزهای خوشه‌ها) مشابه آنچه در بخش تست شبکه عصبی توضیح داده شد، رسم شوند.

مرزهای خوشه‌ها:

بعد از اینکه تعداد خوشه‌های بهینه به دست آمدند، برای رسم مرزهای خوشه‌ها ساده ترین کاری که می‌توان انجام داد به صورت زیر است:

فرض کنید که داده‌ها در راستای مولفه‌ی اول (محور x) بین ۲,۵ تا ۳,۷۵ پخش شده‌اند و در راستای مولفه‌ی دوم (محور y) بین ۱ تا ۵,۵؛ برای دیدن مرزهای خوشه‌ها کافیست که تعداد زیادی داده به صورت رندوم $uniform$ ، بین این مرزها تولید کنید و سپس این داده‌ها را به FCM بدهید. برای هر داده یک خروجی

فازی داده می‌شود. (با توجه به اینکه مراکز خوشه‌ها از قبل به دست آمده‌اند، برای به دست آوردن تعلق داده های جدید به خوشه‌ها از همان رابطه u_{ij} در $fcmeans$ استفاده نمایید.) خوشه‌ی هر داده را خوشه‌ای که داده بیش‌ترین تعلق به آن را دارد در نظر بگیرید و هر خوشه را با یک رنگ مجزا معین کنید. با این رویکرد و از آنجایی که این تعداد زیاد داده‌ها عملاً کل فضا را می‌پوشانند، مرزبندی خوشه‌ها مشخص می‌شود.

برای تولید داده‌ها هم می‌توان به صورت رندوم عمل کرد یا اینکه فضا را به صورت بهینه‌تر پیمایش کنیم که توضیحات تکمیلی در این زمینه در کانال تلگرام درس ارائه شده است.

نکات تکمیلی

- برای آموزش و تست شبکه عصبی خود می‌توانید از دیتاست پیوست استفاده نمایید.
- پیاده‌سازی پروژه به صورت تک نفره می‌باشد و زبان برنامه نویسی آزاد است.
- تحویل به صورت حضوری خواهد بود و تسلط روی کد و مفاهیم مطرح شده در پروژه بخش مهمی از معیارهای ارزیابی است.
- ددلاین پروژه تا ساعت ۲۳:۵۵ شنبه ۵ بهمن می‌باشد.
- برای مطرح کردن سوالات و اشکالات خود می‌توانید با ایمیل ci.fall2019@gmail.com در ارتباط باشید.
- کدهای پیاده‌سازی شده و گزارش خود را در قالب یک فایل زیپ و با فرمت ID_FName LName.zip در مودل درس بارگزاری نمایید.