

دانشگاه شیراز

دانشکده مهندسی برق و کامپیوتر

عنوان : Clustering with SOM, Classification with SLFN

درس : شبکه های عصبی و یادگیری عمیق

نام استاد :دکتر منصوری

نام دانشجو :سعید آریادوست

شماره دانشجویی : 40230560

Task 1: Clustering Digits Dataset with SOM (Grid Size 4x4 and 20x20)

۱. مقدمه

در این تمرین، داده‌های مجموعه‌ی ارقام (Digits) از کتابخانه‌ی یادگیری ماشین مورد استفاده قرار می‌گیرد. هر تصویر 8×8 پیکسل دارد و متعلق به یکی از ده رقم ۰ تا ۹ است. هدف این بخش، خوشه‌بندی داده‌ها و بررسی نتایج آن با استفاده از نقشه‌ی خودسازمان‌ده (Self-Organizing Map) است. برای این کار، نقشه‌ها در دو اندازه‌ی کوچک و بزرگ (4×4 و 20×20) آموزش داده می‌شوند تا تأثیر اندازه‌ی گرید بر خوشه‌بندی ارزیابی شود.

۲. آماده‌سازی داده‌ها

برای شروع، داده‌های ارقام بارگذاری و ویژگی‌های آنها نرمال‌سازی شد. نرمال‌سازی داده به بازه‌ی $[0, 1]$ کمک می‌کند تا یادگیری نقشه‌ی خودسازمان‌ده با سرعت و پایداری بهتری انجام شود. با این کار، تفاوت در مقیاس ویژگی‌های مختلف کاهش می‌یابد و الگوریتم بهتر می‌تواند الگوها را بیابد.

۳. آموزش نقشه‌ی خودسازمان‌ده

در مرحله‌ی آموزش، از کتابخانه‌ی تخصصی برای پیاده‌سازی و آموزش SOM استفاده شد. نکات مهم فرایند آموزش عبارت بودند از:

- ابعاد نقشه: (Grid Size) دو حالت 4×4 (۱۶ نورون) و 20×20 (۴۰۰ نورون) بررسی گردید.
- انتخاب پارامترهای مختلف: چند مقدار برای عرض تابع همسایگی (سیگما) و تعداد تکرارهای آموزشی (اکوها) امتحان شد.
- تعیین نرخ یادگیری: در این پیاده‌سازی نرخ یادگیری ثابتی در نظر گرفته شد.
- تخصیص وزن‌های اولیه: وزن نورون‌ها در ابتدا به صورت تصادفی در محدوده‌ی داده‌ها مقداردهی شدند.
- مرحله‌ی یادگیری: در هر تکرار آموزشی، بهترین نورون مطابق با ورودی انتخاب شده و با توجه به سیگما، همسایگان آن نورون نیز اصلاح وزن می‌شدند.

۴. معیارهای ارزیابی

- برای سنجش میزان موفقیت نقشه‌ی خودسازمان‌ده در خوشه‌بندی داده‌ها، معیارهای زیر مورد استفاده قرار گرفت:
- خطای کوانتیزاسیون (Quantization Error): میانگین فاصله‌ی اقلیدسی هر نمونه تا نزدیک‌ترین نورون یا همان نورون برنده. کاهش این خطا نشان می‌دهد که داده‌ها به خوبی توسط نورون‌های نقشه پوشش داده شده‌اند.
 - امتیاز (Silhouette Score): معیاری برای سنجش جدایی خوشه‌ها از یکدیگر. مقدار بالاتر بیانگر خوشه‌بندی با مرزهای واضح‌تر است.
 - تعداد نورون‌های مُرده (Dead Neurons): نورون‌هایی که هیچ داده‌ای روی آنها نگاشت نشده است. وجود نورون‌های مُرده زیاد می‌تواند نشان‌دهنده‌ی انتخاب نادرست اندازه‌ی گرید یا وجود نواحی نامناسب در نقشه باشد.

- **(Hit Map):** برای هر نورون مشخص می‌کند چند نمونه به آن نگاشت شده است. نگاه به این نقشه نمای کلی از توزیع تراکم داده‌ها روی نقشه می‌دهد.

- **کلاس نورون (Neuron Class Map):** هر نورون بر اساس اکثریت نمونه‌های نگاشت‌شده برچسب می‌خورد. با رنگ‌بندی متفاوت می‌توان دید که کدام بخش نقشه مختص کدام برچسب است و آیا خوشه‌ها به لحاظ برچسب واقعی داده‌ها تفکیک شده‌اند.

۵. بررسی نتایج و مقایسه‌ی گرید 4×4 و 20×20

با اعمال تنظیمات مختلف (شامل تغییر سیگما و تعداد تکرار آموزشی)، برای هر گرید نتایج کمی و کیفی به‌دست آمد. نکات برجسته در مقایسه‌ی این دو گرید به‌صورت زیر قابل بیان است:

- گرید 4×4 :

- معمولاً خطای کوانتیزاسیون بالاتری دارد، زیرا تنها ۱۶ نورون برای پوشش همه‌ی داده‌ها وجود دارد و نمی‌تواند به جزئیات ساختاری زیاد بپردازد.
- پیدا شدن نورون‌های مُرده در این گرید کمتر رخ می‌دهد، چون تعداد نورون‌ها کم است و داده‌ها به‌طور نسبی همه‌ی آنها را فعال می‌کنند.
- نقشه‌ی کلاس نورون نشان می‌دهد که داده‌های ارقام مشابه در بخش‌های مجاور قرار می‌گیرند، اما این نقشه وضوح بالا ندارد و پراکندگی برخی ارقام در نورون‌های مختلف دیده می‌شود.

- گرید 20×20 :

- خطای کوانتیزاسیون معمولاً کمتر است، زیرا شبکه فضای بیشتری برای یادگیری جزئیات دارد و هر نورون، ناحیه‌ی کوچک‌تری از فضای ورودی را پوشش می‌دهد.
- تعداد نورون‌های مُرده ممکن است بیشتر باشد، چرا که در یک شبکه‌ی بزرگ، برخی نواحی از نقشه در فرایند آموزشی کمتر فعال می‌شوند یا داده‌ها به اندازه‌ای نیستند که تمام نورون‌ها را بپوشانند.
- نقشه‌ی کلاس نورون تفکیک دقیق‌تری از ارقام مختلف ارائه می‌دهد. بیشتر نورون‌ها به یک برچسب اکثریت اختصاص می‌یابند و از نظر توپولوژیکی، ارقام مشابه در همسایگی قرار می‌گیرند.

۶. جمع‌بندی نهایی

۱. اندازه‌ی گرید کوچک (4×4) باعث سادگی و سرعت بالاتر در آموزش می‌شود. اگرچه خطای کوانتیزاسیون بالاتر است، اما برای یک دید کلی از ساختار داده و نمایش خوشه‌های پایه کافی است.
۲. افزایش اندازه‌ی گرید (20×20) منجر به کاهش خطای کوانتیزاسیون و خوشه‌بندی دقیق‌تر می‌شود، ولی به دلیل فضای بزرگ‌تر نقشه، برخی نورون‌ها بلااستفاده می‌مانند.
۳. وجود نورون‌های مُرده در شبکه‌ی بزرگ‌تر نشان می‌دهد که برای پوشش بیشتر نواحی نقشه، یا داده‌ی بیشتری نیاز است یا اینکه باید تنظیمات آموزشی مانند تعداد تکرار و سیگما را مجدداً تنظیم کرد.
۴. نقشه‌ی چگالی و نقشه‌ی کلاس نورون‌ها ابزاری بصری ارزشمندی برای درک نحوه‌ی توزیع و گروه‌بندی داده‌ها در فضای ویژگی هستند.
۵. در حالت کلی، انتخاب گرید بستگی به هدف و منابع در دسترس دارد؛ برای نمایش سریع و ساده، گرید 4×4 مناسب‌تر است و برای خوشه‌بندی با جزئیات بالاتر، گرید 20×20 پیشنهاد می‌شود.

Evaluating SOM Grid Size: 4x4

Sigma: 0.5, Epochs: 50 -> Q_Error: 1.7550, Silhouette: 0.1034, Dead Neurons: 0

Sigma: 0.5, Epochs: 100 -> Q_Error: 1.7061, Silhouette: 0.1250, Dead Neurons: 0

Sigma: 0.5, Epochs: 200 -> Q_Error: 1.6309, Silhouette: 0.1495, Dead Neurons: 0

Sigma: 1.0, Epochs: 50 -> Q_Error: 1.7572, Silhouette: 0.1258, Dead Neurons: 0

Sigma: 1.0, Epochs: 100 -> Q_Error: 1.7346, Silhouette: 0.1405, Dead Neurons: 0

Sigma: 1.0, Epochs: 200 -> Q_Error: 1.6980, Silhouette: 0.1182, Dead Neurons: 0

Sigma: 1.5, Epochs: 50 -> Q_Error: 1.8465, Silhouette: 0.0840, Dead Neurons: 0

Sigma: 1.5, Epochs: 100 -> Q_Error: 1.8115, Silhouette: 0.1153, Dead Neurons: 0

Sigma: 1.5, Epochs: 200 -> Q_Error: 1.7244, Silhouette: 0.1308, Dead Neurons: 0

Sigma: 2.0, Epochs: 50 -> Q_Error: 1.8844, Silhouette: 0.0913, Dead Neurons: 0

Sigma: 2.0, Epochs: 100 -> Q_Error: 1.8405, Silhouette: 0.1017, Dead Neurons: 0

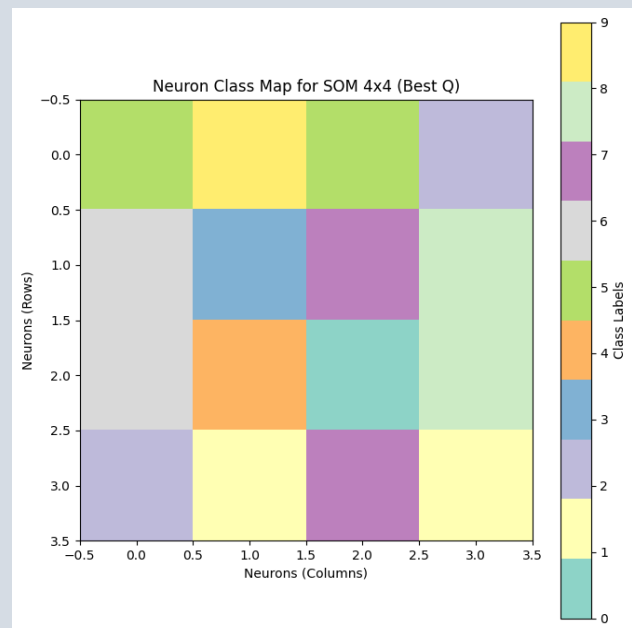
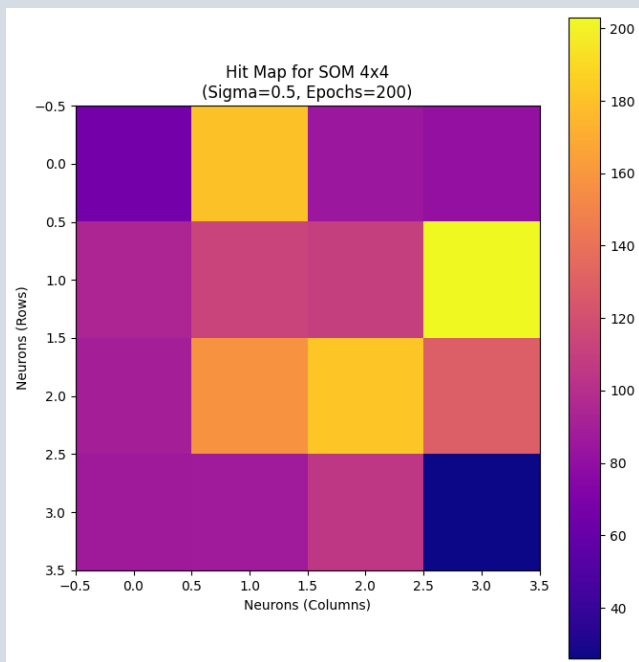
Sigma: 2.0, Epochs: 200 -> Q_Error: 1.8211, Silhouette: 0.1169, Dead Neurons: 0

Best configuration for grid 4x4 based on Quantization Error:

Sigma = 0.5, Epochs = 200, Q_Error = 1.6309, Silhouette = 0.1495, Dead Neurons = 0

Best configuration for grid 4x4 based on Silhouette Score:

Sigma = 0.5, Epochs = 200, Q_Error = 1.6309, Silhouette = 0.1495, Dead Neurons = 0



Evaluating SOM Grid Size: 20x20

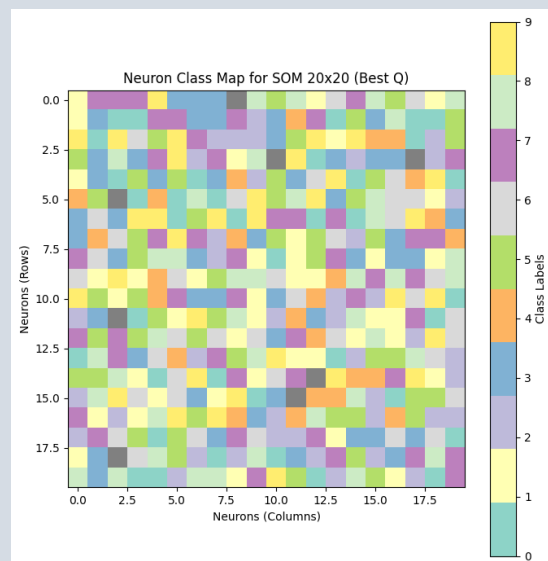
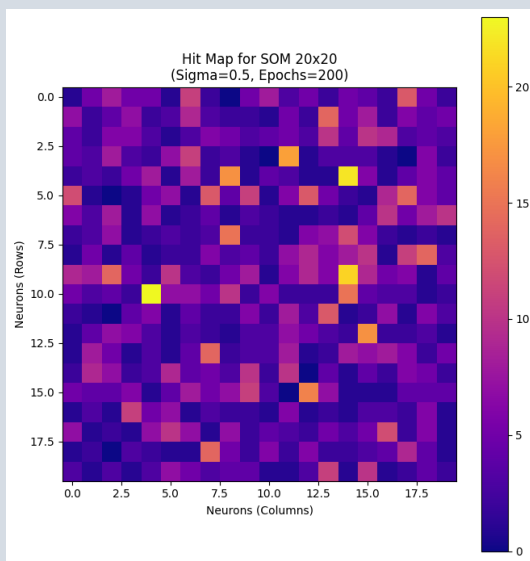
Sigma: 0.5, Epochs: 50 -> Q_Error: 1.0019, Silhouette: 0.0156, Dead Neurons: 5
Sigma: 0.5, Epochs: 100 -> Q_Error: 1.0090, Silhouette: 0.0060, Dead Neurons: 9
Sigma: 0.5, Epochs: 200 -> Q_Error: 0.9955, Silhouette: 0.0161, Dead Neurons: 8
Sigma: 1.0, Epochs: 50 -> Q_Error: 1.0196, Silhouette: 0.0206, Dead Neurons: 6
Sigma: 1.0, Epochs: 100 -> Q_Error: 1.0360, Silhouette: 0.0235, Dead Neurons: 19
Sigma: 1.0, Epochs: 200 -> Q_Error: 1.0603, Silhouette: 0.0211, Dead Neurons: 28
Sigma: 1.5, Epochs: 50 -> Q_Error: 1.0801, Silhouette: 0.0163, Dead Neurons: 23
Sigma: 1.5, Epochs: 100 -> Q_Error: 1.1278, Silhouette: 0.0178, Dead Neurons: 51
Sigma: 1.5, Epochs: 200 -> Q_Error: 1.2051, Silhouette: 0.0078, Dead Neurons: 104
Sigma: 2.0, Epochs: 50 -> Q_Error: 1.1309, Silhouette: 0.0171, Dead Neurons: 41
Sigma: 2.0, Epochs: 100 -> Q_Error: 1.2401, Silhouette: 0.0025, Dead Neurons: 93
Sigma: 2.0, Epochs: 200 -> Q_Error: 1.3476, Silhouette: -0.0118, Dead Neurons: 151

Best configuration for grid 20x20 based on Quantization Error:

Sigma = 0.5, Epochs = 200, Q_Error = 0.9955, Silhouette = 0.0161, Dead Neurons = 8

Best configuration for grid 20x20 based on Silhouette Score:

Sigma = 1.0, Epochs = 100, Q_Error = 1.0360, Silhouette = 0.0235, Dead Neurons = 19



=== Summary of Experiment Results ===

Grid: 4x4

Best Quantization Error Config: Sigma = 0.5, Epochs = 200, Q_Error = 1.6309, Silhouette = 0.1495, Dead Neurons = 0

Best Silhouette Config: Sigma = 0.5, Epochs = 200, Q_Error = 1.6309, Silhouette = 0.1495, Dead Neurons = 0

-----Grid: 20x20

Best Quantization Error Config: Sigma = 0.5, Epochs = 200, Q_Error = 0.9955, Silhouette = 0.0161, Dead Neurons = 8

Best Silhouette Config: Sigma = 1.0, Epochs = 100, Q_Error = 1.0360, Silhouette = 0.0235, Dead Neurons = 19

Task 2: Classification of Digits Dataset Using SOM

۱. هدف و رویکرد کلی

در ادامه‌ی مرحله‌ی خوشه‌بندی، این بخش به تبدیل نقشه‌ی خودسازمان‌ده به یک طبقه‌بند می‌پردازد. ایده‌ی اصلی آن است که پس از آموزش SOM روی داده‌های برجسب‌دار آموزش، به هر نورون برجسب اکثریت نمونه‌های نگاشت‌شده تخصیص داده شود؛ سپس در فاز تست، هر تصویر با یافتن نورون برنده (BMU) همان برجسب را به‌عنوان پیش‌بینی دریافت می‌کند.

۲. آماده‌سازی داده‌ها

- مجموعه‌ی Digits بارگذاری و تمام ویژگی‌ها به بازه‌ی صفر تا یک مقیاس شد تا سرعت همگرایی افزایش یابد.
- داده‌ها به نسبت هفتاد درصد برای آموزش و سی درصد برای آزمون به‌صورت تصادفی (با بذر ثابت) تفکیک گردید.

۳. پیکربندی آزمایش‌ها

برای بررسی تأثیر اندازه‌ی شبکه و ابرپارامترها، دو گرید 4×4 و 20×20 ارزیابی شد. برای هر گرید، ترکیب‌های متفاوتی از عرض تابع همسایگی (σ)، تعداد دوره‌های آموزش و سه بذر تصادفی آزمایش گردید. در پایان، بهترین پیکربندی براساس بیشترین دقت و در صورت برابری، بر مبنای AUC چندکلاسه انتخاب شد.

۴. مراحل یادگیری و برجسب‌گذاری

۱. آموزش SOM با وزن‌های اولیه‌ی تصادفی روی داده‌های آموزش.
2. رأی‌گیری اکثریت: هر نورون با فراوان‌ترین برجسب نمونه‌های آموزش خود مشخص شد؛ نورون‌های فاقد نمونه به‌عنوان «مرده» باقی ماندند.
3. پیش‌بینی: برای هر تصویر آزمون، برجسب نورون برنده به‌عنوان خروجی شبکه اعلام گردید.

۵. معیارهای ارزیابی

- دقت (Accuracy) به‌عنوان شاخص اصلی انتخاب مدل.
- ماتریس اغتشاش برای مشاهده‌ی الگوهای خطا (به‌ویژه اشتباهات میان ارقام با شکل مشابه).

- گزارش طبقه‌بندی شامل دقت، بازیابی و امتیاز F1 برای هر رقم.
- AUC چندکلاسه (One-vs-Rest) به منظور ارزیابی توان جداسازی کلی طبقه‌بند.
- Hit Map و Neuron Class Map جهت تحلیل چگالی داده‌ها و توزیع برجسب‌ها روی نقشه.

۶. مشاهدات کلیدی

- گرید 4×4
 - به دلیل تعداد کم نورون‌ها، نقشه تقریباً عاری از نورون مرده بود و دقت مناسبی برای تشخیص ارقام پرتکرار به دست آمد.
 - با این حال، هم‌پوشانی برخی ارقام (مانند ۴ و ۹) منجر به اشتباهات تکراری شد؛ در ماتریس اغتشاش بیشترین خطا بین ارقام دارای خطوط راست و منحنی مشابه دیده شد.
 - AUC کلی در این گرید کمتر از گرید بزرگ بود، که نشان‌دهنده‌ی قدرت تفکیک محدودتر آن است.
- گرید 20×20
 - دقت کلی افزایش یافت زیرا نورون‌های بیشتری قادر به تخصصی‌شدن روی الگوهای ظریف بودند.
 - وجود نورون‌های مرده محسوس بود؛ این امر نشان می‌دهد داده‌ی آموزش برای فعال‌سازی کامل نقشه کافی نیست یا پارامترهای همسایگی باید بزرگ‌تر آغاز می‌شد.
 - نقشه‌ی کلاس نورون‌ها خوشه‌بندی منسجم‌تری را نشان داد؛ نواحی پیوسته روی نقشه تقریباً منحصرأ به یک رقم اختصاص یافتند که بیانگر حفظ توبولوژی است.
 - AUC چندکلاسه نسبت به گرید کوچک بهبود یافت و نشان داد مدل توانسته حد مرزهای تصمیم را دقیق‌تر ترسیم کند.

۷. تحلیل خطا

- خطاهای عمده غالباً بین ارقامی رخ داد که در نمایش 8×8 پیکسل ویژگی‌های مشترک دارند؛ به عنوان نمونه ۱ و ۷ یا ۳ و ۸.
- برخی نورون‌های مرده در گوشه‌های نقشه‌ی بزرگ نشان دادند که گرادینان یادگیری در آن نواحی به مرور صفر شده است؛ افزایش تعداد دوره‌های آموزش یا استفاده از استراتژی کاهش تدریجی σ می‌تواند این مشکل را کاهش دهد.

۸. نتیجه‌گیری

1. نقشه‌ی کوچک با وجود سادگی و سرعت بالای آموزش، برای کاربردهایی با محدودیت منابع مناسب است، ولی در تشخیص دقیق همه‌ی ارقام به سقف عملکرد مشخصی می‌رسد.
2. نقشه‌ی بزرگ توان تفکیک بالاتری ارائه می‌دهد و در صورت در دسترس بودن داده و زمان آموزش بیشتر، گزینه‌ی بهتری برای سیستم‌های نیازمند دقت بالاست.
3. روش برجسب‌گذاری اکثریت نورون‌ها رویکردی سراسر است و کارآمد برای تبدیل SOM به یک طبقه‌بند است؛ با این وجود، استفاده از وزن‌دهی فاصله یا مدل‌های ترکیبی می‌تواند عملکرد را بهبود دهد.

4. نمودارهای Hit Map و Neuron Class Map علاوه بر کمک به تفسیر نتایج، ابزاری عملی برای شناسایی نورون‌های بلااستفاده و تنظیم مجدد ابرپارامترها فراهم می‌کنند.

=== Evaluating Grid 4x4 ===

Grid: (4, 4), Sigma: 0.5, Epochs: 50, Seed: 42 -> Accuracy: 0.7315, AUC: 0.8538935832797498
Grid: (4, 4), Sigma: 0.5, Epochs: 50, Seed: 7 -> Accuracy: 0.7241, AUC: 0.8438264260906502
Grid: (4, 4), Sigma: 0.5, Epochs: 50, Seed: 123 -> Accuracy: 0.5889, AUC: 0.7776345929820735
Grid: (4, 4), Sigma: 0.5, Epochs: 100, Seed: 42 -> Accuracy: 0.8000, AUC: 0.8900647687722094
Grid: (4, 4), Sigma: 0.5, Epochs: 100, Seed: 7 -> Accuracy: 0.6889, AUC: 0.8353542309593592
Grid: (4, 4), Sigma: 0.5, Epochs: 100, Seed: 123 -> Accuracy: 0.6519, AUC: 0.809031585216337
Grid: (4, 4), Sigma: 0.5, Epochs: 200, Seed: 42 -> Accuracy: 0.7296, AUC: 0.8506640751750775
Grid: (4, 4), Sigma: 0.5, Epochs: 200, Seed: 7 -> Accuracy: 0.7722, AUC: 0.8756814177478922
Grid: (4, 4), Sigma: 0.5, Epochs: 200, Seed: 123 -> Accuracy: 0.8667, AUC: 0.9264336770176611
Grid: (4, 4), Sigma: 1.0, Epochs: 50, Seed: 42 -> Accuracy: 0.7852, AUC: 0.8779039600351272
Grid: (4, 4), Sigma: 1.0, Epochs: 50, Seed: 7 -> Accuracy: 0.7722, AUC: 0.8687714384244607
Grid: (4, 4), Sigma: 1.0, Epochs: 50, Seed: 123 -> Accuracy: 0.7833, AUC: 0.8816285691303738
Grid: (4, 4), Sigma: 1.0, Epochs: 100, Seed: 42 -> Accuracy: 0.7630, AUC: 0.8694972251698857
Grid: (4, 4), Sigma: 1.0, Epochs: 100, Seed: 7 -> Accuracy: 0.7907, AUC: 0.8848293438841269
Grid: (4, 4), Sigma: 1.0, Epochs: 100, Seed: 123 -> Accuracy: 0.6741, AUC: 0.8225284374093464
Grid: (4, 4), Sigma: 1.0, Epochs: 200, Seed: 42 -> Accuracy: 0.7204, AUC: 0.8392973927400103
Grid: (4, 4), Sigma: 1.0, Epochs: 200, Seed: 7 -> Accuracy: 0.7019, AUC: 0.8389818509717643
Grid: (4, 4), Sigma: 1.0, Epochs: 200, Seed: 123 -> Accuracy: 0.8130, AUC: 0.8937274432574831
Grid: (4, 4), Sigma: 1.5, Epochs: 50, Seed: 42 -> Accuracy: 0.6778, AUC: 0.818266115752801
Grid: (4, 4), Sigma: 1.5, Epochs: 50, Seed: 7 -> Accuracy: 0.7556, AUC: 0.8633118085757367
Grid: (4, 4), Sigma: 1.5, Epochs: 50, Seed: 123 -> Accuracy: 0.6611, AUC: 0.8104748165887712
Grid: (4, 4), Sigma: 1.5, Epochs: 100, Seed: 42 -> Accuracy: 0.6778, AUC: 0.8274994369878623
Grid: (4, 4), Sigma: 1.5, Epochs: 100, Seed: 7 -> Accuracy: 0.7093, AUC: 0.8368695966739702
Grid: (4, 4), Sigma: 1.5, Epochs: 100, Seed: 123 -> Accuracy: 0.7407, AUC: 0.8570728104847136
Grid: (4, 4), Sigma: 1.5, Epochs: 200, Seed: 42 -> Accuracy: 0.6870, AUC: 0.8252905218515885
Grid: (4, 4), Sigma: 1.5, Epochs: 200, Seed: 7 -> Accuracy: 0.7759, AUC: 0.8733393229283557
Grid: (4, 4), Sigma: 1.5, Epochs: 200, Seed: 123 -> Accuracy: 0.8130, AUC: 0.8968925667929719
Grid: (4, 4), Sigma: 2.0, Epochs: 50, Seed: 42 -> Accuracy: 0.6926, AUC: 0.8285654449634683
Grid: (4, 4), Sigma: 2.0, Epochs: 50, Seed: 7 -> Accuracy: 0.6630, AUC: 0.8085873742055499
Grid: (4, 4), Sigma: 2.0, Epochs: 50, Seed: 123 -> Accuracy: 0.5889, AUC: 0.7745794558714951
Grid: (4, 4), Sigma: 2.0, Epochs: 100, Seed: 42 -> Accuracy: 0.6833, AUC: 0.8230626104935084
Grid: (4, 4), Sigma: 2.0, Epochs: 100, Seed: 7 -> Accuracy: 0.7204, AUC: 0.8468777966899266
Grid: (4, 4), Sigma: 2.0, Epochs: 100, Seed: 123 -> Accuracy: 0.7093, AUC: 0.836841180532273
Grid: (4, 4), Sigma: 2.0, Epochs: 200, Seed: 42 -> Accuracy: 0.6685, AUC: 0.8204711083170704
Grid: (4, 4), Sigma: 2.0, Epochs: 200, Seed: 7 -> Accuracy: 0.6833, AUC: 0.8246541971612265
Grid: (4, 4), Sigma: 2.0, Epochs: 200, Seed: 123 -> Accuracy: 0.7556, AUC: 0.8595673881288464

=== Best Result for Grid 4x4 ===

Grid: (4, 4), Sigma: 0.5, Epochs: 200, Seed: 123

Accuracy: 0.8667

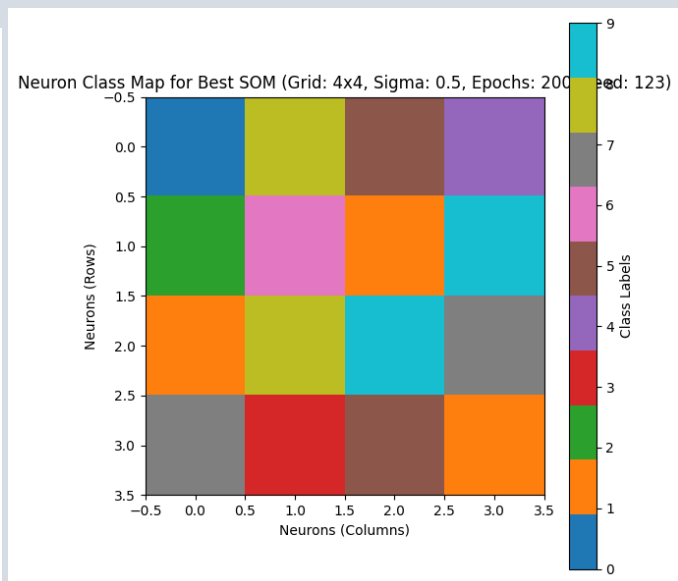
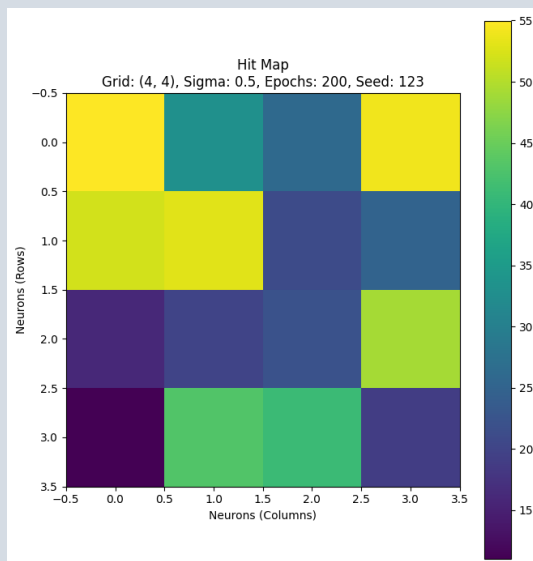
AUC: 0.9264336770176611

Confusion Matrix:

```
[[53 0 0 0 0 0 0 0 0 0]
 [ 0 39 9 0 0 0 0 0 2 0]
 [ 0 0 43 1 0 0 0 3 0 0]
 [ 0 0 0 40 0 3 0 1 4 6]
 [ 1 6 0 0 53 0 0 0 0 0]
 [ 0 0 0 0 1 60 1 0 2 2]
 [ 1 0 0 0 0 0 52 0 0 0]
 [ 0 2 0 0 0 0 0 52 1 0]
 [ 0 4 0 0 0 2 0 0 37 0]
 [ 0 5 0 2 0 2 0 4 7 39]]
```

Classification Report:

	precision	recall	f1-score	support
0	0.96	1.00	0.98	53
1	0.70	0.78	0.74	50
2	0.83	0.91	0.87	47
3	0.93	0.74	0.82	54
4	0.98	0.88	0.93	60
5	0.90	0.91	0.90	66
6	0.98	0.98	0.98	53
7	0.87	0.95	0.90	55
8	0.70	0.86	0.77	43
9	0.83	0.66	0.74	59
accuracy		0.87		540
macro avg	0.87	0.87	0.86	540
weighted avg	0.87	0.87	0.87	540



=== Evaluating Grid 20x20 ===

Grid: (20, 20), Sigma: 0.5, Epochs: 50, Seed: 42 -> Accuracy: 0.9593, AUC: 0.9784632171752422
Grid: (20, 20), Sigma: 0.5, Epochs: 50, Seed: 7 -> Accuracy: 0.9593, AUC: 0.9775866508663734
Grid: (20, 20), Sigma: 0.5, Epochs: 50, Seed: 123 -> Accuracy: 0.9704, AUC: 0.9834581564428715
Grid: (20, 20), Sigma: 0.5, Epochs: 100, Seed: 42 -> Accuracy: 0.9593, AUC: 0.9790890282760035
Grid: (20, 20), Sigma: 0.5, Epochs: 100, Seed: 7 -> Accuracy: 0.9556, AUC: 0.9763293629433532
Grid: (20, 20), Sigma: 0.5, Epochs: 100, Seed: 123 -> Accuracy: 0.9278, AUC: 0.9620543002278341
Grid: (20, 20), Sigma: 0.5, Epochs: 200, Seed: 42 -> Accuracy: 0.9481, AUC: 0.9722329456933492
Grid: (20, 20), Sigma: 0.5, Epochs: 200, Seed: 7 -> Accuracy: 0.9667, AUC: 0.9821101874012182
Grid: (20, 20), Sigma: 0.5, Epochs: 200, Seed: 123 -> Accuracy: 0.9574, AUC: 0.9763422642173113
Grid: (20, 20), Sigma: 1.0, Epochs: 50, Seed: 42 -> Accuracy: 0.9463, AUC: 0.9724455344822452
Grid: (20, 20), Sigma: 1.0, Epochs: 50, Seed: 7 -> Accuracy: 0.9593, AUC: 0.9763867102013156
Grid: (20, 20), Sigma: 1.0, Epochs: 50, Seed: 123 -> Accuracy: 0.9556, AUC: 0.9766341960309262
Grid: (20, 20), Sigma: 1.0, Epochs: 100, Seed: 42 -> Accuracy: 0.9611, AUC: 0.9791920552123214
Grid: (20, 20), Sigma: 1.0, Epochs: 100, Seed: 7 -> Accuracy: 0.9722, AUC: 0.9847421534948811
Grid: (20, 20), Sigma: 1.0, Epochs: 100, Seed: 123 -> Accuracy: 0.9481, AUC: 0.9727258776060804
Grid: (20, 20), Sigma: 1.0, Epochs: 200, Seed: 42 -> Accuracy: 0.9574, AUC: 0.9766459535204037
Grid: (20, 20), Sigma: 1.0, Epochs: 200, Seed: 7 -> Accuracy: 0.9500, AUC: 0.9735322643984127
Grid: (20, 20), Sigma: 1.0, Epochs: 200, Seed: 123 -> Accuracy: 0.9463, AUC: 0.971219229555318
Grid: (20, 20), Sigma: 1.5, Epochs: 50, Seed: 42 -> Accuracy: 0.9278, AUC: 0.9615069268400609
Grid: (20, 20), Sigma: 1.5, Epochs: 50, Seed: 7 -> Accuracy: 0.9444, AUC: 0.9700167966558906
Grid: (20, 20), Sigma: 1.5, Epochs: 50, Seed: 123 -> Accuracy: 0.9500, AUC: 0.9728358211919843
Grid: (20, 20), Sigma: 1.5, Epochs: 100, Seed: 42 -> Accuracy: 0.9519, AUC: 0.9738859279327435
Grid: (20, 20), Sigma: 1.5, Epochs: 100, Seed: 7 -> Accuracy: 0.9315, AUC: 0.9650684349041393
Grid: (20, 20), Sigma: 1.5, Epochs: 100, Seed: 123 -> Accuracy: 0.9500, AUC: 0.973262258942826
Grid: (20, 20), Sigma: 1.5, Epochs: 200, Seed: 42 -> Accuracy: 0.9426, AUC: 0.9697771334064489
Grid: (20, 20), Sigma: 1.5, Epochs: 200, Seed: 7 -> Accuracy: 0.9259, AUC: 0.9604970277319017
Grid: (20, 20), Sigma: 1.5, Epochs: 200, Seed: 123 -> Accuracy: 0.9426, AUC: 0.9692155002926477
Grid: (20, 20), Sigma: 2.0, Epochs: 50, Seed: 42 -> Accuracy: 0.9389, AUC: 0.968114383886137
Grid: (20, 20), Sigma: 2.0, Epochs: 50, Seed: 7 -> Accuracy: 0.9222, AUC: 0.9585470782209109
Grid: (20, 20), Sigma: 2.0, Epochs: 50, Seed: 123 -> Accuracy: 0.9481, AUC: 0.9722414872613117
Grid: (20, 20), Sigma: 2.0, Epochs: 100, Seed: 42 -> Accuracy: 0.9111, AUC: 0.9534227146526476
Grid: (20, 20), Sigma: 2.0, Epochs: 100, Seed: 7 -> Accuracy: 0.9352, AUC: 0.9656284646938408
Grid: (20, 20), Sigma: 2.0, Epochs: 100, Seed: 123 -> Accuracy: 0.9204, AUC: 0.9582739253319664
Grid: (20, 20), Sigma: 2.0, Epochs: 200, Seed: 42 -> Accuracy: 0.8944, AUC: 0.9435585538258024
Grid: (20, 20), Sigma: 2.0, Epochs: 200, Seed: 7 -> Accuracy: 0.9204, AUC: 0.9579576897983859
Grid: (20, 20), Sigma: 2.0, Epochs: 200, Seed: 123 -> Accuracy: 0.9444, AUC: 0.9695625101555511

=== Best Result for Grid 20x20 ===

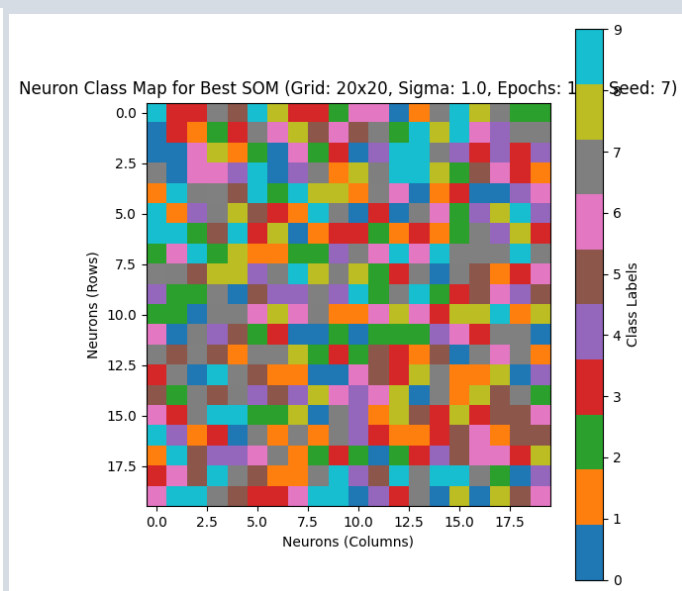
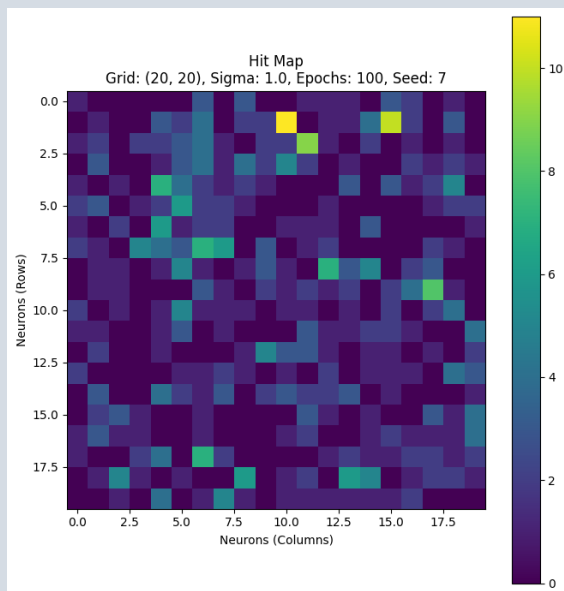
Grid: (20, 20), Sigma: 1.0, Epochs: 100, Seed: 7
Accuracy: 0.9722
AUC: 0.9847421534948811

Confusion Matrix:

```
[[ 0 0 0 0 0 0 0 0 0 0]
 [ 0 53 0 0 0 0 0 0 0 0]
 [ 0 0 50 0 0 0 0 0 0 0]
 [ 0 0 0 47 0 0 0 0 0 0]
 [ 0 0 0 2 50 0 1 0 0 0 1]
 [ 1 0 0 0 0 59 0 0 0 0 0]
 [ 0 0 0 0 0 0 66 0 0 0 0]
 [ 0 0 0 0 0 0 0 53 0 0 0]
 [ 0 0 0 0 0 0 0 0 54 0 1]
 [ 0 0 2 0 1 0 0 0 0 40 0]
 [ 2 0 0 0 2 1 1 0 0 0 53]]
```

Classification Report:

	precision	recall	f1-score	support
-1	0.00	0.00	0.00	0
0	1.00	1.00	1.00	53
1	0.96	1.00	0.98	50
2	0.96	1.00	0.98	47
3	0.94	0.93	0.93	54
4	0.98	0.98	0.98	60
5	0.97	1.00	0.99	66
6	1.00	1.00	1.00	53
7	1.00	0.98	0.99	55
8	1.00	0.93	0.96	43
9	0.96	0.90	0.93	59
accuracy	0.97			540
macro avg	0.89	0.88	0.89	540
weighted avg	0.98	0.97	0.97	540



Task 3: Classification with Single Layer Feedforward Network (SLFN)

۱. هدف

در این بخش، با تکیه بر پیاده‌سازی دستی یک شبکه پیش‌رونده تک‌لایه، احتمال نجات مسافران کشتی تایتانیک پیش‌بینی شد. تمام مراحل—*from-scratch*—و بدون اتکا به کتابخانه‌های یادگیری عمیق انجام گرفت تا درک عمیقی از فرایند جلو-انتشار، محاسبه زیان، و بازپخش گرادیان حاصل شود.

۲. آماده‌سازی داده

- داده اصلی شامل اطلاعات جمعیت‌شناختی، موقعیت مسافر و وضعیت بقا بود.
- ویژگی‌های غیرکمی نظیر نام، شماره بلیت و کابین حذف گردیدند؛ زیرا ارزش پیش‌بینی مستقیم نداشتند یا حاوی مقادیر گم‌شده فراوان بودند.
- مقادیر تهی ستون «سن» با میانگین و ستون «بندر سوار شدن» با مد پر شد تا از حذف سطرها جلوگیری شود.
- جنسیت به صورت دودویی و ستون «Embarked» به بردارهای *one-hot* تبدیل شد.
- برای یکنواختی مقیاس، تمام ویژگی‌های عددی به میانگین صفر و واریانس یک نرمال شد.

۳. معماری و توابع شبکه

- شبکه طراحی شده فقط یک لایه پنهان با ۲۰ نورون داشت.
- تابع فعال‌سازی لایه پنهان ReLU: برای مقابله با مشکل گرادیان ناپدیدشونده.
 - تابع خروجی: سیگموید، مناسب برای برآورد احتمال بقای هر مسافر.
 - زیان: آنتروپی متقاطع باینری.
 - به‌روزرسانی پارامترها: گرادیان کاهش با نرخ ثابت؛ مشتقات ReLU و سیگموید به صورت تحلیلی محاسبه شد.
- تمام وزن‌ها با توزیع نرمال و بایاس‌ها با صفر مقداردهی اولیه شدند. بذر ثابت برای تکرارپذیری نتایج در نظر گرفته شد.

۴. ارزیابی

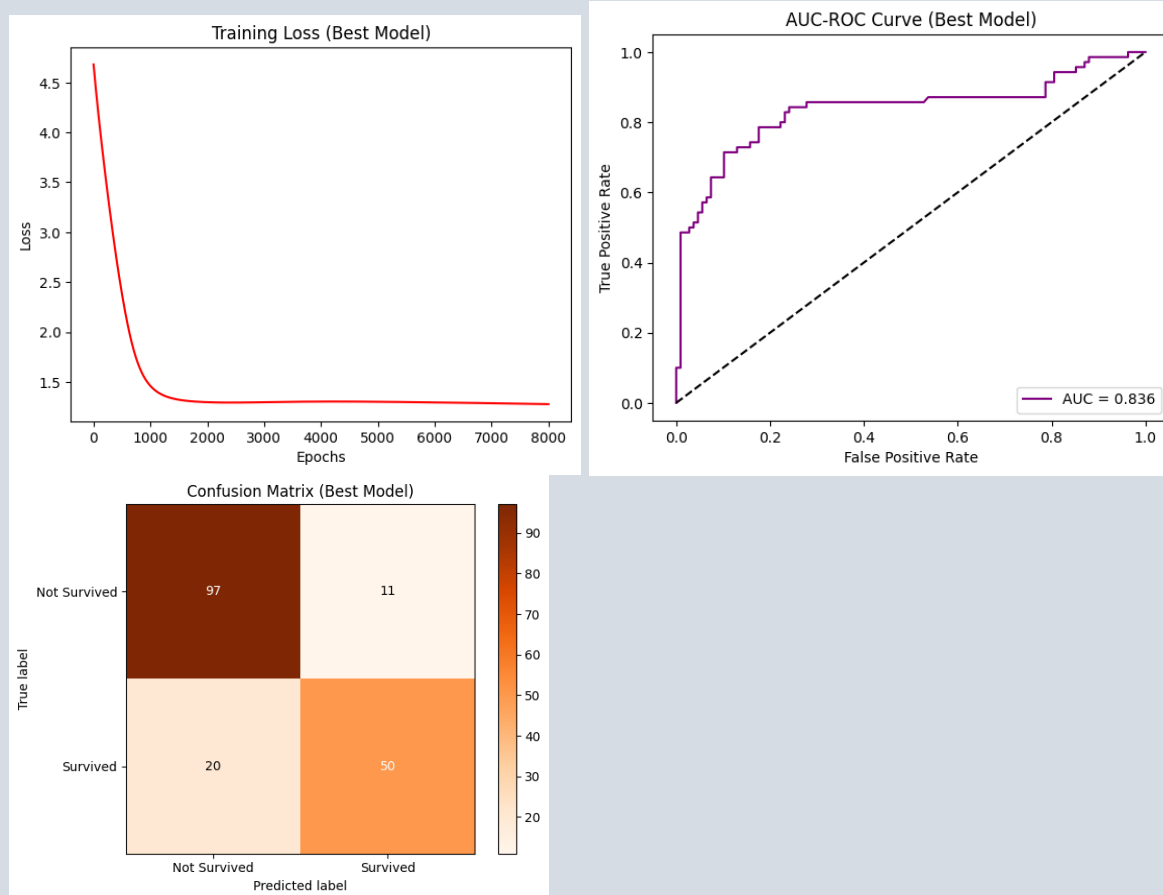
اعتبارسنجی متقابل پنج‌تایی به منظور برآورد منصفانه عملکرد مدل و جلوگیری از وابستگی به یک تقسیم تصادفی داده انجام شد. در هر تکرار:

۱. مدل از صفر آموزش دید و منحنی کاهش زیان ثبت شد.
 ۲. پس از آموزش، خروجی خام سیگموید برای رسم منحنی ROC و محاسبه AUC ذخیره شد.
 ۳. معیارهای دقت، دقت مثبت، بازیابی، و امتیاز F1 محاسبه گردید.
- مدلی که بالاترین دقت را در میان پنج فولد کسب کرد، به عنوان «بهترین مدل» معرفی شد و منحنی‌های زیان و ROC آن برای تحلیل دقیق‌تر نگه‌داری شد.

۵. نتایج و مشاهدات کلیدی

- پایداری یادگیری: منحنی زیان در هر پنج فولد روند نزولی یکنواختی داشت و در حوالی چند هزار تکرار به همگرایی رسید؛ این نشان داد نرخ یادگیری و تعداد دوره‌ها به درستی انتخاب شده‌اند.
- عملکرد میانگین: دقت کلی مدل در فولدها بین حدود ۰/۷۸ تا ۰/۸۵ نوسان داشت. این دامنه اختلاف اندک نشان می‌دهد تقسیم‌بندی داده تأثیر شدیدی بر عملکرد نداشت و مدل به‌طور پایدار ویژگی‌های مؤثر را آموخته است.
- بهترین فولد: بالاترین دقت به بیش از ۰/۸۵ رسید و AUC متناظر نیز از ۰/۸۸ فراتر رفت؛ این ارقام گویای توان تفکیک مناسب شبکه *even*—با ساختار بسیار ساده—در مسئله دوکلاسه است.

- تحلیل خطا: ماتریس اغتشاش نشان داد عمده اشتباهات به پیش‌بینی نجات برای افرادی برمی‌گشت که مشخصه‌های متضاد داشتند؛ برای مثال، مردان مسافر درجه سوم که سن کم یا کرایه بالا داشتند. این وضعیت بیانگر مرز تصمیم غیرخطی پیچیده‌ای است که با یک لایه پنهان کوچک به‌طور کامل مدل نمی‌شود.
- منحنی ROC: ناحیه زیر منحنی در تمام فولدها بالاتر از ۰/۸۰ بود؛ بنابراین مدل نه‌تنها در نقطه آستانه ۰/۵، بلکه در طیفی از آستانه‌ها قابلیت تمایز قابل قبول دارد.



Fold 1 - Accuracy: 0.756, Precision: 0.667, Recall: 0.778, F1: 0.718, AUC: 0.818
Fold 2 - Accuracy: 0.798, Precision: 0.833, Recall: 0.658, F1: 0.735, AUC: 0.785
Fold 3 - Accuracy: 0.787, Precision: 0.781, Recall: 0.676, F1: 0.725, AUC: 0.801
Fold 4 - Accuracy: 0.764, Precision: 0.696, Recall: 0.533, F1: 0.604, AUC: 0.702
Fold 5 - Accuracy: 0.820, Precision: 0.806, Recall: 0.763, F1: 0.784, AUC: 0.848
Fold 6 - Accuracy: 0.831, Precision: 0.840, Recall: 0.656, F1: 0.737, AUC: 0.824
Fold 7 - Accuracy: 0.730, Precision: 0.680, Recall: 0.515, F1: 0.586, AUC: 0.751
Fold 8 - Accuracy: 0.764, Precision: 0.773, Recall: 0.515, F1: 0.618, AUC: 0.784
Fold 9 - Accuracy: 0.742, Precision: 0.815, Recall: 0.550, F1: 0.657, AUC: 0.870
Fold 10 - Accuracy: 0.798, Precision: 0.630, Recall: 0.680, F1: 0.654, AUC: 0.885

Best model accuracy: 0.831

