

# SIECI NEURONOWE - Ćwiczenie 4

W ćwiczeniu 4 odtworzymy zaimplementowaną już architekturę sieci w pełni połączonej korzystając z gotowego rozwiązania do budowania sieci neuronowych. Poniżej jest podany jest przykład uczenia w Pytorchu, inne rozwiązania są dopuszczalne (tensorflow, w jego obrębie keras który jest nawet prostszy), pod warunkiem że dopuszczają zdefiniowanie własnej architektury sieci.

Od framework do głębokiego uczenia oczekujemy w pierwszej kolejności:

- Implementacji operacji macierzowych/tensorowych w takim zakresie jakie będą nam potrzebne do budowy sieci
- Możliwości automatycznego wyliczania gradientów po zaimplementowanych operacjach
- Wydajnej implementacji dedykowanej GPU/TPU Oraz typowych funkcjonalności ułatwiających budowanie sieci neuronowych z wykorzystaniem powyższych (funkcje kosztu, standardowe warstwy, optimizery itp.).

Dla torcha, konwencja jest następująca: korzystamy z obiektów `torch.tensor` do przechowywania danych, natomiast sieć zbudowana jest z modułów dziedziczących po `torch.NN.module`. Przez `tensor` rozumiemy tutaj n-wymiarową tablicę liczb analogicznie do `numpy`. Dla tensorów nadpisane są podstawowe operatory matematyczne, dostępne są operacje na tablicach podobne do dostępnych w `numpy` (zmiany kształtu, transpozycje, agregacje takie jak średnia i suma etc.). W obiekcie tensora obok właściwej wartości może być zapisany również gradient. Gradient, jeżeli istnieje, zawsze będzie tensorem o tym samym wymiarze co właściwe dane tensora. Jest inicjalizowany przy pierwszym przejściu propagacji wstecz przez dany tensor.

Model powinien dziedziczyć po `torch.NN.module` i definiować `init()` – operacje wykonywane przy tworzeniu instancji klasy oraz `forward()` – funkcję która na podstawie danych wejściowych zwraca wyjście modelu.

Oficjalny tutorial odnośnie tego, jak uczyć model, opisuje pętlę uczącą w ten sposób:  
<https://pytorch.org/tutorials/beginner/introyt/trainingyt.html>

```
optimizer = torch.optim.SGD(model.parameters(), lr=0.001, momentum=0.9) #1
```

```
def train_one_epoch(epoch_index, tb_writer):
```

```
    running_loss = 0.  
    last_loss = 0.  
  
    # Here, we use enumerate(training_loader) instead of  
    # iter(training_loader) so that we can track the batch  
  
    # index and do some intra-epoch reporting  
    for i, data in enumerate(training_loader):  
  
        # Every data instance is an input + label pair
```

```

inputs, labels = data

# Zero your gradients for every batch!
optimizer.zero_grad() #2

# Make predictions for this batch
outputs = model(inputs) #3

# Compute the loss and its gradients
loss = loss_fn(outputs, labels) #4
loss.backward() #5

# Adjust learning weights
optimizer.step() #6

return loss

```

Jak i dlaczego to działa?

1. Obiekt optimizer służy do optymalizowania danego zbioru parametrów. Parametry są podklassą tensora specjalnie uwzględnianymi przez moduły.

`parameters()` dowolnego obiektu `torch.NN.module` zwraca parametry przynależące do tejże klasy oraz parametry każdej zmiennej przynależącej, która sama jest podklassą `torch.NN.module`. Uwaga: funkcja nie zadziała na zmiennych zagnieżdżonych jeśli jakikolwiek poziom nie jest podklassą `torch.NN.module!` Na przykład kod: `init(self): super(TinyModel, self).__init__()  
self.layer_list = [torch.nn.Linear(10,10) for i in range(123)]`

Stworzy listę 123 warstw liniowych  $10 \times 10$ , ale jako że jedyną zmienną przynależącą do klasy jest sama lista, model nie będzie poprawnie uwzględniał tych 123 warstw i ich macierzy wag w `parameters()`.

1. `zero_grad()` jest potrzebne, ponieważ moduł automatycznego różniczkowania akumuluje, a nie nadpisuje gradienty przy wielokrotnym wywołaniu.
2. Wywołanie instancji modelu jak funkcji (pythonowe **call**) jest równoważne `forward()`
3. Funkcje kosztu są dostępne jako parametryzowalne klasy w `torch.nn`. To znaczy, że najpierw instancjonujemy je jako obiekt, a potem wywołujemy `loss_fn()` na odpowiednich tensorach odpowiadających wejściom i pożdanym wyjściom – należy zwrócić uwagę na oczekiwany typ tensora w dokumentacji i pamiętać o domyślnych formatach! (W szczególności, domyślny float torcha NIE zgadza się z domyślnym w numpy – 32 vs 64bit!)
4. `backward()` to całe przejście wstecz w algorytmie propagacji wstecznej. Jak udało się to tak uprościć? Po pierwsze, każda operacja na tensorach dostępna w torchu, analogicznie do sposobu implementacji sugerowanego w poprzednim ćwiczeniu, ma zarówno funkcję `forward()`, jak i `backward()`. Po drugie, operacje w trakcie przechodzenia w przód są dodawane do grafu obliczeniowego zawierającego

informacje o tym skąd wejście do danej operacji. Istnienie grafu obliczeniowego pozwala przeiterować po wcześniej wywołanych operacjach wstecz i wyliczyć gradienty na każdym poziomie, oraz zsumować je w obiektach parametrów.

5. step() dokonuje zmiany parametrów zgodnie z wartością gradientu i przyjętą regułą uczenia. Dla SGD jest to klasyczne  $-\text{learning\_rate} * \text{gradient}$

W ćwiczeniu należy zbudować sieć o tych samych parametrach co w zadaniu poprzednim, i ocenić jej działanie na tych samych danych, tym razem sprawdzając wpływ:

- wybranego optimizera (SGD i dwa inne)
- rozmiaru batcha
- wartości współczynnika uczenia dla różnych optimizerów

Ćwiczenie oceniane jest w skali 0-10 pkt.

## Rozwiążanie Zadania

```
"""
train_heart_experiments.py

Założenia:
- DATA_PATH = '../heart-disease/processed.cleveland.data'
- COLUMN_NAMES zgodne z treścią zadania
- Binary classification: target = (num > 0)
- Eksperymenty: optymalizery = [SGD, Adam, RMSprop], kilka batch_size
i lr
- Wyniki zapisywane do results.csv
"""

import os
import random
import numpy as np
import pandas as pd
from sklearn.model_selection import train_test_split
from sklearn.preprocessing import StandardScaler
from sklearn.metrics import accuracy_score, roc_auc_score
import torch
import torch.nn as nn
import torch.nn.functional as F
from torch.utils.data import Dataset, DataLoader, TensorDataset
import matplotlib.pyplot as plt

# CONFIG

DATA_PATH = '../heart-disease/processed.cleveland.data'
COLUMN_NAMES = ['age', 'sex', 'cp', 'trestbps', 'chol', 'fbs',
'restecg',
'thalach', 'exang', 'oldpeak', 'slope', 'ca', 'thal',
'num']
```

```

SEED = 42

device = torch.device('cuda' if torch.cuda.is_available() else 'cpu')
np.random.seed(SEED)
random.seed(SEED)
torch.manual_seed(SEED)
if device.type == 'cuda':
    torch.cuda.manual_seed_all(SEED)

PLOTS_DIR = 'plots'
os.makedirs(PLOTS_DIR, exist_ok=True)
RESULTS_DIR = 'lab-4-results'
os.makedirs(RESULTS_DIR, exist_ok=True)
RESULTS_CSV = 'results_heart_experiments.csv'

# Hyperparam grid
optimizers_to_try = ['SGD', 'Adam', 'RMSprop']
batch_sizes = [16, 32, 64]
lrs = {
    'SGD': [0.1, 0.01, 0.001, 0.0001, 0.00001],
    'Adam': [0.1, 0.01, 0.001, 0.0001, 0.00001],
    'RMSprop': [0.1, 0.01, 0.001, 0.0001, 0.00001]
}
epochs = 200
weight_decay = 1e-4

# DATA LOADING & PREPROCESS

def load_and_preprocess(path):
    df = pd.read_csv(path, header=None, names=COLUMN_NAMES,
na_values='?')
    df = df.dropna(how='all')

    for col in df.columns:
        if df[col].isnull().any():
            df[col] = df[col].fillna(df[col].median())

    df['target'] = (df['num'] > 0).astype(int)
    X = df.drop(columns=['num', 'target']).values.astype(np.float32)
    y = df['target'].values.astype(np.int64)
    return X, y, df

X, y, df = load_and_preprocess(DATA_PATH)
print(f"Loaded X shape {X.shape}, y shape {y.shape}, device={device}")

X_temp, X_test, y_temp, y_test = train_test_split(X, y, test_size=0.2,
random_state=SEED, stratify=y)

```

```

X_train, X_val, y_train, y_val = train_test_split(X_temp, y_temp,
test_size=0.2, random_state=SEED, stratify=y_temp)

scaler = StandardScaler()
X_train = scaler.fit_transform(X_train)
X_val = scaler.transform(X_val)
X_test = scaler.transform(X_test)

def get_loaders(batch_size):
    train_ds = TensorDataset(torch.tensor(X_train),
    torch.tensor(y_train))
    val_ds = TensorDataset(torch.tensor(X_val), torch.tensor(y_val))
    test_ds = TensorDataset(torch.tensor(X_test),
    torch.tensor(y_test))
    train_loader = DataLoader(train_ds, batch_size=batch_size,
shuffled=True)
    val_loader = DataLoader(val_ds, batch_size=batch_size,
shuffled=False)
    test_loader = DataLoader(test_ds, batch_size=batch_size,
shuffled=False)
    return train_loader, val_loader, test_loader

# MODEL (fully-connected)

class NeuralNetwork(torch.nn.Module):
    def __init__(self, input_dim, hidden_dims=[32,16], dropout=0.2):
        super().__init__()

        layers = []
        last = input_dim

        for h in hidden_dims:
            layers.append(nn.Linear(last, h))
            layers.append(nn.ReLU())
            layers.append(nn.Dropout(dropout))
            last = h

        layers.append(nn.Linear(last, 1))
        self.net = nn.Sequential(*layers)

    def forward(self, x):
        return self.net(x).squeeze(1)

# TRAIN / EVAL helpers

def get_optimizer(name, params, lr):
    if name == 'SGD':
        return torch.optim.SGD(params, lr=lr, momentum=0.9,

```

```
weight_decay=weight_decay)
    elif name == 'Adam':
        return torch.optim.Adam(params, lr=lr,
weight_decay=weight_decay)
    elif name == 'RMSprop':
        return torch.optim.RMSprop(params, lr=lr,
weight_decay=weight_decay)
    else:
        raise ValueError(name)

bce_loss = nn.BCEWithLogitsLoss()

def train_one_epoch(model, loader, optimizer):
    model.train()

    running_loss = 0.0

    for inputs, labels in loader:
        inputs = inputs.to(device)
        labels = labels.float().to(device)

        optimizer.zero_grad()
        outputs = model(inputs)
        loss = bce_loss(outputs, labels)

        loss.backward()
        optimizer.step()

        running_loss += loss.item() * inputs.size(0)
    return running_loss / len(loader.dataset)

def evaluate(model, loader):
    model.eval()

    all_logits = []
    all_labels = []

    total_loss = 0.0

    with torch.no_grad():
        for inputs, labels in loader:
            inputs = inputs.to(device)
            labels = labels.float().to(device)

            logits = model(inputs)
            loss = bce_loss(logits, labels)
            total_loss += loss.item() * inputs.size(0)

            all_logits.append(logits)
            all_labels.append(labels)
```

```

logits = torch.cat(all_logits)
labels = torch.cat(all_labels)

probs = torch.sigmoid(logits)

acc = ((probs >= 0.5) == labels).float().mean().item()

try:
    auc = roc_auc_score(labels.cpu().numpy(), probs.cpu().numpy())
except Exception:
    auc = float('nan')

return total_loss / len(loader.dataset), acc, auc

# EXPERIMENT GRID

results_summary = []
results_detailed = []

for opt_name in optimizers_to_try:
    for batch_size in batch_sizes:
        train_loader, val_loader, test_loader =
get_loaders(batch_size)
        for lr in lrs[opt_name]:
            print(f"\n==== Running: optimizer={opt_name},
batch={batch_size}, lr={lr} ===")
            model = NeuralNetwork(input_dim=X_train.shape[1],
hidden_dims=[32, 16], dropout=0.2).to(device)
            optimizer = get_optimizer(opt_name, model.parameters(),
lr)
            best_val_auc = -1.0
            history = {'train_loss': [], 'val_loss': [], 'val_acc':
[], 'val_auc': []}

            for ep in range(1, epochs + 1):
                train_loss = train_one_epoch(model, train_loader,
optimizer)
                val_loss, val_acc, val_auc = evaluate(model,
val_loader)

                history['train_loss'].append(train_loss)
                history['val_loss'].append(val_loss)
                history['val_acc'].append(val_acc)
                history['val_auc'].append(val_auc)

            results_detailed.append({
                'optimizer': opt_name,
                'batch_size': batch_size,

```

```

        'lr': lr,
        'epoch': ep,
        'train_loss': train_loss,
        'val_loss': val_loss,
        'val_acc': val_acc,
        'val_auc': val_auc
    })

    if ep % 10 == 0 or ep == 1:
        print(f"ep {ep:02d}: train_loss={train_loss:.4f}
val_loss={val_loss:.4f} "
              f"val_acc={val_acc:.4f}
val_auc={val_auc:.4f}")

        if val_auc > best_val_auc:
            best_val_auc = val_auc
            best_state = {k: v.cpu() for k, v in
model.state_dict().items()}

            model.load_state_dict({k: v.to(device) for k, v in
best_state.items()})
            test_loss, test_acc, test_auc = evaluate(model,
test_loader)
            print(f"--> test_acc={test_acc:.4f},
test_auc={test_auc:.4f}")

            results_summary.append({
                'optimizer': opt_name,
                'batch_size': batch_size,
                'lr': lr,
                'test_loss': test_loss,
                'test_acc': test_acc,
                'test_auc': test_auc,
                'best_val_auc': best_val_auc
            })

```

**# ZAPIS WYNIKÓW**

```

summary_path = os.path.join(RESULTS_DIR, 'results_summary.csv')
detailed_path = os.path.join(RESULTS_DIR, 'results_detailed.csv')

pd.DataFrame(results_summary).to_csv(summary_path, index=False)
pd.DataFrame(results_detailed).to_csv(detailed_path, index=False)

print("\nAll experiments finished.")
print(f"Summary saved to {summary_path}")
print(f"Full epoch logs saved to {detailed_path}")

```

```

# ZBIORCZY WYKRES VAL AUC

import seaborn as sns

df_detailed = pd.DataFrame(results_detailed)

plt.figure(figsize=(10,6))
sns.set(style="whitegrid")

for opt_name, group in df_detailed.groupby('optimizer'):
    auc_mean = group.groupby('epoch')['val_auc'].mean()
    plt.plot(auc_mean.index, auc_mean.values, label=opt_name)

plt.title("Średnie val_auc w czasie dla optymalizatorów")
plt.xlabel("Epoch")
plt.ylabel("Validation AUC")
plt.legend(title="Optimizer")
plt.tight_layout()
plt.show()

# WYKRESY: val_auc vs epoch dla różnych lr w obrębie każdego optymalizatora

for opt_name, group in df_detailed.groupby('optimizer'):
    plt.figure(figsize=(10,6))
    sns.set(style="whitegrid")

    for lr_val, lr_group in group.groupby('lr'):
        auc_mean = lr_group.groupby('epoch')['val_auc'].mean()
        plt.plot(auc_mean.index, auc_mean.values,
label=f"lr={lr_val}")

    plt.title(f"Średnie val_auc w czasie - {opt_name}")
    plt.xlabel("Epoch")
    plt.ylabel("Validation AUC")
    plt.legend(title="Learning rate")
    plt.tight_layout()

    save_path = os.path.join(PLOTS_DIR, f"val_auc_{opt_name}.png")
    plt.savefig(save_path)
    print(f"Saved plot: {save_path}")
    plt.show()

# WYKRESY: val_auc vs epoch dla różnych batch_size (dla każdego optymalizatora)

for opt_name, group in df_detailed.groupby('optimizer'):
    plt.figure(figsize=(10,6))
    sns.set(style="whitegrid")

```

```

    for batch_val, batch_group in group.groupby('batch_size'):
        auc_mean = batch_group.groupby('epoch')['val_auc'].mean()
        plt.plot(auc_mean.index, auc_mean.values,
label=f"batch={batch_val}")

    plt.title(f"Średnie val_auc w czasie - {opt_name} (różne
batch_size)")
    plt.xlabel("Epoch")
    plt.ylabel("Validation AUC")
    plt.legend(title="Batch size")
    plt.tight_layout()

    save_path = os.path.join(PLOTS_DIR,
f"val_auc_batch_{opt_name}.png")
    plt.savefig(save_path)
    print(f"Saved plot: {save_path}")
    plt.show()

```

Loaded X shape (303, 13), y shape (303,), device=cpu

```

==== Running: optimizer=SGD, batch=16, lr=0.1 ====
ep 01: train_loss=0.6610 val_loss=0.5531 val_acc=0.8163 val_auc=0.8653
ep 10: train_loss=0.3272 val_loss=0.4107 val_acc=0.8367 val_auc=0.9007
ep 20: train_loss=0.3018 val_loss=0.5821 val_acc=0.8163 val_auc=0.8855
ep 30: train_loss=0.5319 val_loss=0.4468 val_acc=0.7755 val_auc=0.8535
ep 40: train_loss=0.4333 val_loss=0.5260 val_acc=0.6939 val_auc=0.7879
ep 50: train_loss=0.4432 val_loss=0.5552 val_acc=0.7959 val_auc=0.8519
ep 60: train_loss=0.5303 val_loss=0.7032 val_acc=0.6939 val_auc=0.7508
ep 70: train_loss=0.3596 val_loss=0.5196 val_acc=0.7959 val_auc=0.8990
ep 80: train_loss=0.3687 val_loss=0.7181 val_acc=0.7959 val_auc=0.8586
ep 90: train_loss=0.2953 val_loss=0.9380 val_acc=0.7551 val_auc=0.8519
ep 100: train_loss=0.3339 val_loss=0.8757 val_acc=0.7551
val_auc=0.7795
ep 110: train_loss=0.3249 val_loss=0.8356 val_acc=0.7551
val_auc=0.8157
ep 120: train_loss=0.2491 val_loss=1.0476 val_acc=0.7755
val_auc=0.8653
ep 130: train_loss=0.2712 val_loss=0.9893 val_acc=0.7755
val_auc=0.8603
ep 140: train_loss=0.3503 val_loss=1.2999 val_acc=0.7755
val_auc=0.8468
ep 150: train_loss=0.1741 val_loss=1.1442 val_acc=0.7755
val_auc=0.8788
ep 160: train_loss=0.2206 val_loss=1.1099 val_acc=0.7755
val_auc=0.8544
ep 170: train_loss=0.5970 val_loss=0.6910 val_acc=0.7347
val_auc=0.8047
ep 180: train_loss=0.3692 val_loss=0.5989 val_acc=0.8163
val_auc=0.8653
ep 190: train_loss=0.3463 val_loss=1.6577 val_acc=0.7959

```

```
val_auc=0.8114
ep 200: train_loss=0.3527 val_loss=1.0296 val_acc=0.7755
val_auc=0.8291
--> test_acc=0.7869, test_auc=0.8252

==== Running: optimizer=SGD, batch=16, lr=0.01 ====
ep 01: train_loss=0.6933 val_loss=0.6817 val_acc=0.6327 val_auc=0.7593
ep 10: train_loss=0.4245 val_loss=0.3306 val_acc=0.8776 val_auc=0.9495
ep 20: train_loss=0.3830 val_loss=0.3288 val_acc=0.8776 val_auc=0.9259
ep 30: train_loss=0.3357 val_loss=0.3327 val_acc=0.8367 val_auc=0.9141
ep 40: train_loss=0.2882 val_loss=0.3809 val_acc=0.8163 val_auc=0.8990
ep 50: train_loss=0.2801 val_loss=0.4479 val_acc=0.7551 val_auc=0.8737
ep 60: train_loss=0.2646 val_loss=0.4724 val_acc=0.7959 val_auc=0.8737
ep 70: train_loss=0.2812 val_loss=0.4866 val_acc=0.7551 val_auc=0.8535
ep 80: train_loss=0.2256 val_loss=0.5917 val_acc=0.7551 val_auc=0.8569
ep 90: train_loss=0.1624 val_loss=0.6513 val_acc=0.7347 val_auc=0.8384
ep 100: train_loss=0.1822 val_loss=0.5901 val_acc=0.7143
val_auc=0.8670
ep 110: train_loss=0.1365 val_loss=0.7099 val_acc=0.7551
val_auc=0.8434
ep 120: train_loss=0.2718 val_loss=0.8645 val_acc=0.7347
val_auc=0.8165
ep 130: train_loss=0.1832 val_loss=0.7369 val_acc=0.7959
val_auc=0.8653
ep 140: train_loss=0.1367 val_loss=0.8845 val_acc=0.7755
val_auc=0.8670
ep 150: train_loss=0.0954 val_loss=0.9051 val_acc=0.7755
val_auc=0.8704
ep 160: train_loss=0.1660 val_loss=0.9459 val_acc=0.7959
val_auc=0.8620
ep 170: train_loss=0.0880 val_loss=0.8724 val_acc=0.7755
val_auc=0.8721
ep 180: train_loss=0.1103 val_loss=0.9558 val_acc=0.7551
val_auc=0.8687
ep 190: train_loss=0.1634 val_loss=0.8563 val_acc=0.8163
val_auc=0.8805
ep 200: train_loss=0.0679 val_loss=0.9550 val_acc=0.8367
val_auc=0.8788
--> test_acc=0.7705, test_auc=0.9264

==== Running: optimizer=SGD, batch=16, lr=0.001 ====
ep 01: train_loss=0.7044 val_loss=0.6938 val_acc=0.4490 val_auc=0.5892
ep 10: train_loss=0.6627 val_loss=0.6537 val_acc=0.8571 val_auc=0.8552
ep 20: train_loss=0.6237 val_loss=0.6082 val_acc=0.8571 val_auc=0.8822
ep 30: train_loss=0.5821 val_loss=0.5534 val_acc=0.8571 val_auc=0.8923
ep 40: train_loss=0.5352 val_loss=0.4960 val_acc=0.8367 val_auc=0.8939
ep 50: train_loss=0.4867 val_loss=0.4487 val_acc=0.8163 val_auc=0.8939
ep 60: train_loss=0.4571 val_loss=0.4153 val_acc=0.8163 val_auc=0.8990
ep 70: train_loss=0.4215 val_loss=0.3913 val_acc=0.8163 val_auc=0.8990
```

```
ep 80: train_loss=0.4014 val_loss=0.3764 val_acc=0.8367 val_auc=0.8973
ep 90: train_loss=0.3982 val_loss=0.3754 val_acc=0.8163 val_auc=0.8973
ep 100: train_loss=0.3957 val_loss=0.3700 val_acc=0.8163
val_auc=0.8956
ep 110: train_loss=0.3779 val_loss=0.3669 val_acc=0.8367
val_auc=0.9007
ep 120: train_loss=0.3782 val_loss=0.3637 val_acc=0.8367
val_auc=0.9024
ep 130: train_loss=0.3647 val_loss=0.3613 val_acc=0.8367
val_auc=0.9024
ep 140: train_loss=0.3500 val_loss=0.3590 val_acc=0.8367
val_auc=0.9024
ep 150: train_loss=0.3627 val_loss=0.3613 val_acc=0.8367
val_auc=0.9024
ep 160: train_loss=0.3928 val_loss=0.3610 val_acc=0.8367
val_auc=0.9007
ep 170: train_loss=0.3427 val_loss=0.3614 val_acc=0.8367
val_auc=0.9057
ep 180: train_loss=0.3547 val_loss=0.3608 val_acc=0.8367
val_auc=0.9057
ep 190: train_loss=0.3801 val_loss=0.3658 val_acc=0.8571
val_auc=0.9074
ep 200: train_loss=0.3711 val_loss=0.3647 val_acc=0.8367
val_auc=0.9057
--> test_acc=0.8361, test_auc=0.9535

==== Running: optimizer=SGD, batch=16, lr=0.0001 ====
ep 01: train_loss=0.6852 val_loss=0.6883 val_acc=0.6735 val_auc=0.5960
ep 10: train_loss=0.6761 val_loss=0.6862 val_acc=0.6939 val_auc=0.6111
ep 20: train_loss=0.6789 val_loss=0.6842 val_acc=0.6939 val_auc=0.6145
ep 30: train_loss=0.6716 val_loss=0.6821 val_acc=0.6939 val_auc=0.6246
ep 40: train_loss=0.6746 val_loss=0.6800 val_acc=0.7143 val_auc=0.6481
ep 50: train_loss=0.6694 val_loss=0.6778 val_acc=0.7551 val_auc=0.6633
ep 60: train_loss=0.6654 val_loss=0.6756 val_acc=0.7347 val_auc=0.6700
ep 70: train_loss=0.6628 val_loss=0.6733 val_acc=0.7143 val_auc=0.6818
ep 80: train_loss=0.6634 val_loss=0.6708 val_acc=0.7143 val_auc=0.6902
ep 90: train_loss=0.6588 val_loss=0.6684 val_acc=0.6939 val_auc=0.7071
ep 100: train_loss=0.6583 val_loss=0.6659 val_acc=0.6939
val_auc=0.7256
ep 110: train_loss=0.6562 val_loss=0.6635 val_acc=0.7143
val_auc=0.7306
ep 120: train_loss=0.6523 val_loss=0.6607 val_acc=0.7143
val_auc=0.7424
ep 130: train_loss=0.6503 val_loss=0.6580 val_acc=0.7143
val_auc=0.7542
ep 140: train_loss=0.6458 val_loss=0.6552 val_acc=0.7143
val_auc=0.7576
ep 150: train_loss=0.6470 val_loss=0.6523 val_acc=0.7143
val_auc=0.7694
```

```
ep 160: train_loss=0.6338 val_loss=0.6491 val_acc=0.7143
val_auc=0.7828
ep 170: train_loss=0.6361 val_loss=0.6459 val_acc=0.7143
val_auc=0.7828
ep 180: train_loss=0.6340 val_loss=0.6427 val_acc=0.7347
val_auc=0.7845
ep 190: train_loss=0.6304 val_loss=0.6390 val_acc=0.7347
val_auc=0.7879
ep 200: train_loss=0.6251 val_loss=0.6353 val_acc=0.7347
val_auc=0.7997
--> test_acc=0.8033, test_auc=0.9123

==== Running: optimizer=SGD, batch=16, lr=1e-05 ====
ep 01: train_loss=0.7081 val_loss=0.7114 val_acc=0.3469 val_auc=0.2525
ep 10: train_loss=0.7037 val_loss=0.7112 val_acc=0.3469 val_auc=0.2542
ep 20: train_loss=0.6975 val_loss=0.7109 val_acc=0.3469 val_auc=0.2576
ep 30: train_loss=0.7021 val_loss=0.7107 val_acc=0.3469 val_auc=0.2609
ep 40: train_loss=0.7040 val_loss=0.7104 val_acc=0.3469 val_auc=0.2609
ep 50: train_loss=0.7053 val_loss=0.7101 val_acc=0.3469 val_auc=0.2660
ep 60: train_loss=0.7022 val_loss=0.7098 val_acc=0.3469 val_auc=0.2660
ep 70: train_loss=0.7061 val_loss=0.7096 val_acc=0.3469 val_auc=0.2694
ep 80: train_loss=0.7050 val_loss=0.7093 val_acc=0.3469 val_auc=0.2710
ep 90: train_loss=0.6970 val_loss=0.7090 val_acc=0.3469 val_auc=0.2744
ep 100: train_loss=0.7018 val_loss=0.7088 val_acc=0.3469
val_auc=0.2744
ep 110: train_loss=0.7006 val_loss=0.7085 val_acc=0.3469
val_auc=0.2795
ep 120: train_loss=0.7004 val_loss=0.7082 val_acc=0.3469
val_auc=0.2795
ep 130: train_loss=0.7010 val_loss=0.7079 val_acc=0.3673
val_auc=0.2828
ep 140: train_loss=0.7022 val_loss=0.7076 val_acc=0.3673
val_auc=0.2879
ep 150: train_loss=0.6975 val_loss=0.7074 val_acc=0.3673
val_auc=0.2912
ep 160: train_loss=0.7012 val_loss=0.7071 val_acc=0.3673
val_auc=0.2929
ep 170: train_loss=0.6980 val_loss=0.7068 val_acc=0.3673
val_auc=0.2963
ep 180: train_loss=0.6982 val_loss=0.7065 val_acc=0.3673
val_auc=0.2980
ep 190: train_loss=0.7043 val_loss=0.7062 val_acc=0.3673
val_auc=0.3047
ep 200: train_loss=0.6956 val_loss=0.7060 val_acc=0.3673
val_auc=0.3081
--> test_acc=0.4098, test_auc=0.3929

==== Running: optimizer=SGD, batch=32, lr=0.1 ====
ep 01: train_loss=0.6999 val_loss=0.6645 val_acc=0.5510 val_auc=0.8519
```

```
ep 10: train_loss=0.3854 val_loss=0.4427 val_acc=0.8163 val_auc=0.8906
ep 20: train_loss=0.3762 val_loss=0.4072 val_acc=0.8163 val_auc=0.8872
ep 30: train_loss=0.2754 val_loss=0.4131 val_acc=0.7551 val_auc=0.8889
ep 40: train_loss=0.3014 val_loss=0.3975 val_acc=0.7959 val_auc=0.8973
ep 50: train_loss=0.3123 val_loss=0.5314 val_acc=0.7755 val_auc=0.8771
ep 60: train_loss=0.3753 val_loss=0.6112 val_acc=0.7143 val_auc=0.8552
ep 70: train_loss=0.2973 val_loss=0.7923 val_acc=0.7551 val_auc=0.8645
ep 80: train_loss=0.1768 val_loss=0.8389 val_acc=0.7959 val_auc=0.8586
ep 90: train_loss=0.2366 val_loss=1.0484 val_acc=0.7959 val_auc=0.8822
ep 100: train_loss=0.2651 val_loss=0.7607 val_acc=0.7551
val_auc=0.8519
ep 110: train_loss=0.2115 val_loss=0.8136 val_acc=0.7551
val_auc=0.8678
ep 120: train_loss=0.2502 val_loss=0.7923 val_acc=0.6939
val_auc=0.8468
ep 130: train_loss=0.2281 val_loss=0.8234 val_acc=0.7755
val_auc=0.8552
ep 140: train_loss=0.8552 val_loss=0.8092 val_acc=0.6735
val_auc=0.7946
ep 150: train_loss=0.3571 val_loss=0.9436 val_acc=0.7347
val_auc=0.8300
ep 160: train_loss=0.5254 val_loss=1.8277 val_acc=0.6735
val_auc=0.8249
ep 170: train_loss=0.4010 val_loss=1.1483 val_acc=0.7347
val_auc=0.8586
ep 180: train_loss=0.2692 val_loss=1.3545 val_acc=0.6939
val_auc=0.8485
ep 190: train_loss=0.2976 val_loss=1.4895 val_acc=0.7551
val_auc=0.8426
ep 200: train_loss=0.2174 val_loss=1.9243 val_acc=0.7755
val_auc=0.8712
--> test_acc=0.8033, test_auc=0.8918
```

```
==== Running: optimizer=SGD, batch=32, lr=0.01 ====
ep 01: train_loss=0.6927 val_loss=0.6968 val_acc=0.5306 val_auc=0.3653
ep 10: train_loss=0.5946 val_loss=0.5645 val_acc=0.8163 val_auc=0.8855
ep 20: train_loss=0.4024 val_loss=0.3577 val_acc=0.8367 val_auc=0.9141
ep 30: train_loss=0.3816 val_loss=0.3659 val_acc=0.8367 val_auc=0.9074
ep 40: train_loss=0.3950 val_loss=0.3584 val_acc=0.8163 val_auc=0.9024
ep 50: train_loss=0.3548 val_loss=0.4062 val_acc=0.8163 val_auc=0.8923
ep 60: train_loss=0.3191 val_loss=0.3902 val_acc=0.8163 val_auc=0.8956
ep 70: train_loss=0.3851 val_loss=0.4815 val_acc=0.8163 val_auc=0.8535
ep 80: train_loss=0.3255 val_loss=0.3932 val_acc=0.8163 val_auc=0.8973
ep 90: train_loss=0.2755 val_loss=0.4221 val_acc=0.7755 val_auc=0.8889
ep 100: train_loss=0.2689 val_loss=0.4570 val_acc=0.7755
val_auc=0.8788
ep 110: train_loss=0.2546 val_loss=0.4551 val_acc=0.7959
val_auc=0.8737
ep 120: train_loss=0.2605 val_loss=0.5329 val_acc=0.7755
```

```
val_auc=0.8704
ep 130: train_loss=0.2290 val_loss=0.4490 val_acc=0.7551
val_auc=0.8939
ep 140: train_loss=0.1946 val_loss=0.5031 val_acc=0.7755
val_auc=0.8620
ep 150: train_loss=0.1730 val_loss=0.5277 val_acc=0.7551
val_auc=0.8737
ep 160: train_loss=0.2281 val_loss=0.5279 val_acc=0.7959
val_auc=0.8737
ep 170: train_loss=0.1865 val_loss=0.5891 val_acc=0.7755
val_auc=0.8552
ep 180: train_loss=0.1455 val_loss=0.5267 val_acc=0.8163
val_auc=0.8704
ep 190: train_loss=0.1909 val_loss=0.5837 val_acc=0.7959
val_auc=0.8687
ep 200: train_loss=0.1699 val_loss=0.5748 val_acc=0.7551
val_auc=0.8620
--> test_acc=0.8689, test_auc=0.9481

==== Running: optimizer=SGD, batch=32, lr=0.001 ===
ep 01: train_loss=0.6769 val_loss=0.6579 val_acc=0.6939 val_auc=0.7795
ep 10: train_loss=0.6692 val_loss=0.6393 val_acc=0.7347 val_auc=0.8367
ep 20: train_loss=0.6445 val_loss=0.6127 val_acc=0.7959 val_auc=0.8636
ep 30: train_loss=0.6092 val_loss=0.5838 val_acc=0.7959 val_auc=0.8805
ep 40: train_loss=0.5776 val_loss=0.5466 val_acc=0.7959 val_auc=0.8872
ep 50: train_loss=0.5650 val_loss=0.5130 val_acc=0.7959 val_auc=0.8923
ep 60: train_loss=0.5213 val_loss=0.4862 val_acc=0.8163 val_auc=0.8822
ep 70: train_loss=0.5007 val_loss=0.4545 val_acc=0.8163 val_auc=0.8855
ep 80: train_loss=0.4914 val_loss=0.4292 val_acc=0.8367 val_auc=0.8889
ep 90: train_loss=0.4764 val_loss=0.4117 val_acc=0.8367 val_auc=0.8923
ep 100: train_loss=0.4314 val_loss=0.3938 val_acc=0.8367
val_auc=0.8990
ep 110: train_loss=0.4230 val_loss=0.3798 val_acc=0.8367
val_auc=0.9040
ep 120: train_loss=0.4170 val_loss=0.3683 val_acc=0.8367
val_auc=0.9074
ep 130: train_loss=0.4070 val_loss=0.3572 val_acc=0.8367
val_auc=0.9125
ep 140: train_loss=0.3893 val_loss=0.3506 val_acc=0.8367
val_auc=0.9141
ep 150: train_loss=0.3936 val_loss=0.3454 val_acc=0.8367
val_auc=0.9158
ep 160: train_loss=0.4015 val_loss=0.3403 val_acc=0.8367
val_auc=0.9158
ep 170: train_loss=0.3919 val_loss=0.3372 val_acc=0.8367
val_auc=0.9209
ep 180: train_loss=0.3871 val_loss=0.3353 val_acc=0.8367
val_auc=0.9209
ep 190: train_loss=0.3877 val_loss=0.3298 val_acc=0.8571
```

```
val_auc=0.9192
ep 200: train_loss=0.3889 val_loss=0.3260 val_acc=0.8571
val_auc=0.9242
--> test_acc=0.8852, test_auc=0.9632

==== Running: optimizer=SGD, batch=32, lr=0.0001 ====
ep 01: train_loss=0.7105 val_loss=0.7112 val_acc=0.4490 val_auc=0.5219
ep 10: train_loss=0.7002 val_loss=0.7097 val_acc=0.4490 val_auc=0.5354
ep 20: train_loss=0.7035 val_loss=0.7075 val_acc=0.4490 val_auc=0.5572
ep 30: train_loss=0.7005 val_loss=0.7053 val_acc=0.4490 val_auc=0.5673
ep 40: train_loss=0.7035 val_loss=0.7032 val_acc=0.4490 val_auc=0.5825
ep 50: train_loss=0.6998 val_loss=0.7012 val_acc=0.4490 val_auc=0.5960
ep 60: train_loss=0.6964 val_loss=0.6990 val_acc=0.4490 val_auc=0.6195
ep 70: train_loss=0.6975 val_loss=0.6971 val_acc=0.4490 val_auc=0.6397
ep 80: train_loss=0.6898 val_loss=0.6954 val_acc=0.4490 val_auc=0.6515
ep 90: train_loss=0.6885 val_loss=0.6938 val_acc=0.4490 val_auc=0.6667
ep 100: train_loss=0.6883 val_loss=0.6919 val_acc=0.4490
val_auc=0.6785
ep 110: train_loss=0.6852 val_loss=0.6903 val_acc=0.4694
val_auc=0.6869
ep 120: train_loss=0.6840 val_loss=0.6885 val_acc=0.4898
val_auc=0.7020
ep 130: train_loss=0.6826 val_loss=0.6869 val_acc=0.5102
val_auc=0.7104
ep 140: train_loss=0.6831 val_loss=0.6848 val_acc=0.5102
val_auc=0.7189
ep 150: train_loss=0.6828 val_loss=0.6830 val_acc=0.5306
val_auc=0.7323
ep 160: train_loss=0.6731 val_loss=0.6810 val_acc=0.6122
val_auc=0.7441
ep 170: train_loss=0.6788 val_loss=0.6789 val_acc=0.6327
val_auc=0.7559
ep 180: train_loss=0.6701 val_loss=0.6773 val_acc=0.6122
val_auc=0.7626
ep 190: train_loss=0.6698 val_loss=0.6755 val_acc=0.6122
val_auc=0.7694
ep 200: train_loss=0.6690 val_loss=0.6739 val_acc=0.6122
val_auc=0.7778
--> test_acc=0.5574, test_auc=0.7911

==== Running: optimizer=SGD, batch=32, lr=1e-05 ====
ep 01: train_loss=0.6973 val_loss=0.6965 val_acc=0.5714 val_auc=0.4057
ep 10: train_loss=0.7007 val_loss=0.6963 val_acc=0.5714 val_auc=0.4091
ep 20: train_loss=0.6959 val_loss=0.6961 val_acc=0.5714 val_auc=0.4108
ep 30: train_loss=0.6963 val_loss=0.6959 val_acc=0.5714 val_auc=0.4091
ep 40: train_loss=0.6930 val_loss=0.6956 val_acc=0.5714 val_auc=0.4108
ep 50: train_loss=0.6952 val_loss=0.6954 val_acc=0.5714 val_auc=0.4158
ep 60: train_loss=0.6953 val_loss=0.6952 val_acc=0.5714 val_auc=0.4209
ep 70: train_loss=0.7001 val_loss=0.6950 val_acc=0.5714 val_auc=0.4259
```

```
ep 80: train_loss=0.6902 val_loss=0.6948 val_acc=0.5714 val_auc=0.4293
ep 90: train_loss=0.6924 val_loss=0.6946 val_acc=0.5714 val_auc=0.4310
ep 100: train_loss=0.6960 val_loss=0.6944 val_acc=0.5714
val_auc=0.4394
ep 110: train_loss=0.6975 val_loss=0.6941 val_acc=0.5714
val_auc=0.4411
ep 120: train_loss=0.6941 val_loss=0.6939 val_acc=0.5714
val_auc=0.4478
ep 130: train_loss=0.6943 val_loss=0.6937 val_acc=0.5714
val_auc=0.4495
ep 140: train_loss=0.6946 val_loss=0.6935 val_acc=0.5714
val_auc=0.4512
ep 150: train_loss=0.6950 val_loss=0.6933 val_acc=0.5714
val_auc=0.4529
ep 160: train_loss=0.6964 val_loss=0.6930 val_acc=0.5714
val_auc=0.4579
ep 170: train_loss=0.6883 val_loss=0.6928 val_acc=0.5714
val_auc=0.4613
ep 180: train_loss=0.6894 val_loss=0.6926 val_acc=0.5714
val_auc=0.4646
ep 190: train_loss=0.6919 val_loss=0.6924 val_acc=0.5714
val_auc=0.4646
ep 200: train_loss=0.6975 val_loss=0.6922 val_acc=0.5714
val_auc=0.4663
--> test_acc=0.5410, test_auc=0.4827

==== Running: optimizer=SGD, batch=64, lr=0.1 ====
ep 01: train_loss=0.6856 val_loss=0.6666 val_acc=0.6735 val_auc=0.8081
ep 10: train_loss=0.5463 val_loss=0.3416 val_acc=0.8571 val_auc=0.9259
ep 20: train_loss=0.6595 val_loss=0.6918 val_acc=0.7755 val_auc=0.8662
ep 30: train_loss=0.5655 val_loss=0.6812 val_acc=0.6327 val_auc=0.6902
ep 40: train_loss=0.6958 val_loss=0.4977 val_acc=0.7143 val_auc=0.9158
ep 50: train_loss=0.3659 val_loss=0.3995 val_acc=0.7959 val_auc=0.8965
ep 60: train_loss=0.4198 val_loss=0.5198 val_acc=0.7551 val_auc=0.9125
ep 70: train_loss=0.6191 val_loss=0.5945 val_acc=0.7143 val_auc=0.7828
ep 80: train_loss=0.7665 val_loss=0.7301 val_acc=0.7347 val_auc=0.7542
ep 90: train_loss=0.5245 val_loss=0.6578 val_acc=0.7143 val_auc=0.7609
ep 100: train_loss=0.4745 val_loss=0.5679 val_acc=0.8163
val_auc=0.8173
ep 110: train_loss=0.6389 val_loss=0.6079 val_acc=0.7347
val_auc=0.7584
ep 120: train_loss=0.5526 val_loss=0.7174 val_acc=0.5510
val_auc=0.7677
ep 130: train_loss=0.6374 val_loss=0.5486 val_acc=0.6531
val_auc=0.7407
ep 140: train_loss=1.6147 val_loss=1.1535 val_acc=0.7347
val_auc=0.7508
ep 150: train_loss=1.5396 val_loss=0.6726 val_acc=0.5510
val_auc=0.5370
```

```
ep 160: train_loss=0.6738 val_loss=0.6694 val_acc=0.4898  
val_auc=0.5370  
ep 170: train_loss=0.6775 val_loss=0.6633 val_acc=0.5510  
val_auc=0.5370  
ep 180: train_loss=0.6606 val_loss=0.6636 val_acc=0.5510  
val_auc=0.5370  
ep 190: train_loss=0.6563 val_loss=0.6746 val_acc=0.4898  
val_auc=0.5370  
ep 200: train_loss=0.6707 val_loss=0.6729 val_acc=0.5510  
val_auc=0.5556  
--> test_acc=0.5410, test_auc=0.5271

==== Running: optimizer=SGD, batch=64, lr=0.01 ====
ep 01: train_loss=0.6986 val_loss=0.6943 val_acc=0.5510 val_auc=0.3552  
ep 10: train_loss=0.6618 val_loss=0.6529 val_acc=0.7143 val_auc=0.8434  
ep 20: train_loss=0.5941 val_loss=0.5589 val_acc=0.8367 val_auc=0.8805  
ep 30: train_loss=0.4766 val_loss=0.4267 val_acc=0.7755 val_auc=0.9057  
ep 40: train_loss=0.4699 val_loss=0.4300 val_acc=0.7959 val_auc=0.9209  
ep 50: train_loss=0.4179 val_loss=0.3493 val_acc=0.8980 val_auc=0.9310  
ep 60: train_loss=0.4129 val_loss=0.3176 val_acc=0.8367 val_auc=0.9276  
ep 70: train_loss=0.3885 val_loss=0.3123 val_acc=0.8776 val_auc=0.9293  
ep 80: train_loss=0.3739 val_loss=0.3298 val_acc=0.8571 val_auc=0.9293  
ep 90: train_loss=0.3853 val_loss=0.3735 val_acc=0.8571 val_auc=0.9310  
ep 100: train_loss=0.3533 val_loss=0.3310 val_acc=0.8776  
val_auc=0.9276  
ep 110: train_loss=0.3073 val_loss=0.3259 val_acc=0.8163  
val_auc=0.9276  
ep 120: train_loss=0.3861 val_loss=0.3397 val_acc=0.8571  
val_auc=0.9276  
ep 130: train_loss=0.3238 val_loss=0.3433 val_acc=0.8367  
val_auc=0.9276  
ep 140: train_loss=0.3546 val_loss=0.4124 val_acc=0.8776  
val_auc=0.8956  
ep 150: train_loss=0.3243 val_loss=0.3504 val_acc=0.8163  
val_auc=0.9141  
ep 160: train_loss=0.3272 val_loss=0.3843 val_acc=0.8163  
val_auc=0.9007  
ep 170: train_loss=0.3137 val_loss=0.4089 val_acc=0.8571  
val_auc=0.8939  
ep 180: train_loss=0.2876 val_loss=0.3493 val_acc=0.8367  
val_auc=0.9074  
ep 190: train_loss=0.2666 val_loss=0.3700 val_acc=0.7959  
val_auc=0.9125  
ep 200: train_loss=0.2890 val_loss=0.4270 val_acc=0.8367  
val_auc=0.8822  
--> test_acc=0.7705, test_auc=0.9069

==== Running: optimizer=SGD, batch=64, lr=0.001 ====
ep 01: train_loss=0.7038 val_loss=0.7001 val_acc=0.4898 val_auc=0.4293
```

```
ep 10: train_loss=0.6948 val_loss=0.6917 val_acc=0.5102 val_auc=0.5438
ep 20: train_loss=0.6856 val_loss=0.6790 val_acc=0.6735 val_auc=0.7088
ep 30: train_loss=0.6747 val_loss=0.6657 val_acc=0.6735 val_auc=0.7710
ep 40: train_loss=0.6600 val_loss=0.6552 val_acc=0.7551 val_auc=0.7997
ep 50: train_loss=0.6525 val_loss=0.6455 val_acc=0.7959 val_auc=0.8182
ep 60: train_loss=0.6465 val_loss=0.6360 val_acc=0.7959 val_auc=0.8215
ep 70: train_loss=0.6390 val_loss=0.6235 val_acc=0.7959 val_auc=0.8283
ep 80: train_loss=0.6199 val_loss=0.6091 val_acc=0.7959 val_auc=0.8333
ep 90: train_loss=0.6077 val_loss=0.5943 val_acc=0.8163 val_auc=0.8367
ep 100: train_loss=0.5913 val_loss=0.5839 val_acc=0.8163
val_auc=0.8401
ep 110: train_loss=0.5803 val_loss=0.5696 val_acc=0.8163
val_auc=0.8418
ep 120: train_loss=0.5764 val_loss=0.5564 val_acc=0.8163
val_auc=0.8535
ep 130: train_loss=0.5480 val_loss=0.5449 val_acc=0.8163
val_auc=0.8620
ep 140: train_loss=0.5422 val_loss=0.5324 val_acc=0.8367
val_auc=0.8569
ep 150: train_loss=0.5274 val_loss=0.5184 val_acc=0.8367
val_auc=0.8653
ep 160: train_loss=0.5253 val_loss=0.5051 val_acc=0.8367
val_auc=0.8721
ep 170: train_loss=0.4934 val_loss=0.4933 val_acc=0.8367
val_auc=0.8737
ep 180: train_loss=0.4916 val_loss=0.4789 val_acc=0.8367
val_auc=0.8737
ep 190: train_loss=0.4800 val_loss=0.4661 val_acc=0.8367
val_auc=0.8788
ep 200: train_loss=0.4906 val_loss=0.4610 val_acc=0.8367
val_auc=0.8771
--> test_acc=0.9180, test_auc=0.9524
```

```
==== Running: optimizer=SGD, batch=64, lr=0.0001 ====
ep 01: train_loss=0.7096 val_loss=0.7098 val_acc=0.3878 val_auc=0.4175
ep 10: train_loss=0.7110 val_loss=0.7091 val_acc=0.3878 val_auc=0.4310
ep 20: train_loss=0.7087 val_loss=0.7082 val_acc=0.3878 val_auc=0.4394
ep 30: train_loss=0.7006 val_loss=0.7071 val_acc=0.3878 val_auc=0.4444
ep 40: train_loss=0.7066 val_loss=0.7058 val_acc=0.3878 val_auc=0.4562
ep 50: train_loss=0.7008 val_loss=0.7047 val_acc=0.3878 val_auc=0.4663
ep 60: train_loss=0.7022 val_loss=0.7036 val_acc=0.3878 val_auc=0.4714
ep 70: train_loss=0.6998 val_loss=0.7026 val_acc=0.4286 val_auc=0.4747
ep 80: train_loss=0.7011 val_loss=0.7020 val_acc=0.4490 val_auc=0.4865
ep 90: train_loss=0.7046 val_loss=0.7012 val_acc=0.4490 val_auc=0.4899
ep 100: train_loss=0.7011 val_loss=0.7004 val_acc=0.4694
val_auc=0.4966
ep 110: train_loss=0.6958 val_loss=0.6995 val_acc=0.4898
val_auc=0.5034
ep 120: train_loss=0.7001 val_loss=0.6984 val_acc=0.5102
```

```
val_auc=0.5118
ep 130: train_loss=0.6949 val_loss=0.6974 val_acc=0.5306
val_auc=0.5202
ep 140: train_loss=0.6938 val_loss=0.6967 val_acc=0.5306
val_auc=0.5286
ep 150: train_loss=0.6985 val_loss=0.6957 val_acc=0.5306
val_auc=0.5421
ep 160: train_loss=0.6924 val_loss=0.6949 val_acc=0.5306
val_auc=0.5438
ep 170: train_loss=0.6964 val_loss=0.6941 val_acc=0.5102
val_auc=0.5556
ep 180: train_loss=0.6915 val_loss=0.6930 val_acc=0.5306
val_auc=0.5657
ep 190: train_loss=0.6946 val_loss=0.6923 val_acc=0.5510
val_auc=0.5741
ep 200: train_loss=0.6946 val_loss=0.6917 val_acc=0.5510
val_auc=0.5825
--> test_acc=0.5246, test_auc=0.7532

==== Running: optimizer=SGD, batch=64, lr=1e-05 ===
ep 01: train_loss=0.7078 val_loss=0.7121 val_acc=0.4490 val_auc=0.5185
ep 10: train_loss=0.7118 val_loss=0.7121 val_acc=0.4490 val_auc=0.5185
ep 20: train_loss=0.7107 val_loss=0.7120 val_acc=0.4490 val_auc=0.5185
ep 30: train_loss=0.7095 val_loss=0.7119 val_acc=0.4490 val_auc=0.5202
ep 40: train_loss=0.7104 val_loss=0.7118 val_acc=0.4490 val_auc=0.5236
ep 50: train_loss=0.7098 val_loss=0.7117 val_acc=0.4490 val_auc=0.5269
ep 60: train_loss=0.7070 val_loss=0.7116 val_acc=0.4490 val_auc=0.5269
ep 70: train_loss=0.7059 val_loss=0.7115 val_acc=0.4490 val_auc=0.5303
ep 80: train_loss=0.7085 val_loss=0.7114 val_acc=0.4490 val_auc=0.5337
ep 90: train_loss=0.7081 val_loss=0.7113 val_acc=0.4490 val_auc=0.5354
ep 100: train_loss=0.7089 val_loss=0.7113 val_acc=0.4490
val_auc=0.5354
ep 110: train_loss=0.7065 val_loss=0.7112 val_acc=0.4490
val_auc=0.5354
ep 120: train_loss=0.7096 val_loss=0.7111 val_acc=0.4490
val_auc=0.5370
ep 130: train_loss=0.7083 val_loss=0.7110 val_acc=0.4490
val_auc=0.5370
ep 140: train_loss=0.7059 val_loss=0.7109 val_acc=0.4490
val_auc=0.5387
ep 150: train_loss=0.7070 val_loss=0.7107 val_acc=0.4490
val_auc=0.5387
ep 160: train_loss=0.7079 val_loss=0.7106 val_acc=0.4490
val_auc=0.5387
ep 170: train_loss=0.7061 val_loss=0.7106 val_acc=0.4490
val_auc=0.5387
ep 180: train_loss=0.7061 val_loss=0.7105 val_acc=0.4490
val_auc=0.5387
ep 190: train_loss=0.7101 val_loss=0.7103 val_acc=0.4490
val_auc=0.5404
```

```
ep 200: train_loss=0.7052 val_loss=0.7102 val_acc=0.4490
val_auc=0.5404
--> test_acc=0.4590, test_auc=0.4816

==== Running: optimizer=Adam, batch=16, lr=0.1 ====
ep 01: train_loss=0.5885 val_loss=0.4134 val_acc=0.8367 val_auc=0.9125
ep 10: train_loss=0.3987 val_loss=0.3979 val_acc=0.8163 val_auc=0.8737
ep 20: train_loss=0.4389 val_loss=0.3652 val_acc=0.8367 val_auc=0.8788
ep 30: train_loss=0.4839 val_loss=0.3983 val_acc=0.8163 val_auc=0.8249
ep 40: train_loss=0.5505 val_loss=0.4718 val_acc=0.8571 val_auc=0.8468
ep 50: train_loss=1.2601 val_loss=0.6326 val_acc=0.6122 val_auc=0.6061
ep 60: train_loss=0.4536 val_loss=0.8404 val_acc=0.7551 val_auc=0.7449
ep 70: train_loss=0.5690 val_loss=0.4553 val_acc=0.7755 val_auc=0.7685
ep 80: train_loss=0.4240 val_loss=0.4271 val_acc=0.8163 val_auc=0.7896
ep 90: train_loss=0.7722 val_loss=0.5608 val_acc=0.6939 val_auc=0.6591
ep 100: train_loss=0.4434 val_loss=0.3868 val_acc=0.8367
val_auc=0.8182
ep 110: train_loss=0.5432 val_loss=1.2148 val_acc=0.7551
val_auc=0.7929
ep 120: train_loss=0.5405 val_loss=2.0370 val_acc=0.6735
val_auc=0.6768
ep 130: train_loss=0.6324 val_loss=0.6562 val_acc=0.4694
val_auc=0.5842
ep 140: train_loss=0.5681 val_loss=0.8194 val_acc=0.6531
val_auc=0.6330
ep 150: train_loss=0.5977 val_loss=0.6254 val_acc=0.6122
val_auc=0.5842
ep 160: train_loss=0.5800 val_loss=0.9403 val_acc=0.6122
val_auc=0.5623
ep 170: train_loss=0.5304 val_loss=0.7756 val_acc=0.6531
val_auc=0.6195
ep 180: train_loss=0.5061 val_loss=2.0468 val_acc=0.6531
val_auc=0.6616
ep 190: train_loss=0.6667 val_loss=1.1994 val_acc=0.6735
val_auc=0.6389
ep 200: train_loss=0.4995 val_loss=1.0290 val_acc=0.6939
val_auc=0.7003
--> test_acc=0.7377, test_auc=0.7814

==== Running: optimizer=Adam, batch=16, lr=0.01 ====
ep 01: train_loss=0.6374 val_loss=0.4875 val_acc=0.7959 val_auc=0.8838
ep 10: train_loss=0.3055 val_loss=0.4078 val_acc=0.8163 val_auc=0.8771
ep 20: train_loss=0.2132 val_loss=0.4244 val_acc=0.8367 val_auc=0.9007
ep 30: train_loss=0.1361 val_loss=0.5347 val_acc=0.8163 val_auc=0.8855
ep 40: train_loss=0.1611 val_loss=0.7680 val_acc=0.8163 val_auc=0.8754
ep 50: train_loss=0.1865 val_loss=0.8774 val_acc=0.8367 val_auc=0.8805
ep 60: train_loss=0.1497 val_loss=1.1188 val_acc=0.7959 val_auc=0.8788
ep 70: train_loss=0.0952 val_loss=1.1166 val_acc=0.7551 val_auc=0.8704
ep 80: train_loss=0.1146 val_loss=1.2400 val_acc=0.7959 val_auc=0.8670
```

```
ep 90: train_loss=0.1113 val_loss=1.2713 val_acc=0.7755 val_auc=0.8653
ep 100: train_loss=0.0886 val_loss=1.1730 val_acc=0.7551
val_auc=0.8535
ep 110: train_loss=0.1278 val_loss=1.4131 val_acc=0.7347
val_auc=0.8687
ep 120: train_loss=0.0797 val_loss=1.2174 val_acc=0.7551
val_auc=0.8586
ep 130: train_loss=0.0883 val_loss=1.2853 val_acc=0.7551
val_auc=0.8670
ep 140: train_loss=0.0592 val_loss=1.1795 val_acc=0.7551
val_auc=0.8737
ep 150: train_loss=0.0649 val_loss=1.3008 val_acc=0.7551
val_auc=0.8788
ep 160: train_loss=0.1017 val_loss=1.3662 val_acc=0.7347
val_auc=0.8569
ep 170: train_loss=0.0604 val_loss=1.2260 val_acc=0.7143
val_auc=0.8636
ep 180: train_loss=0.0453 val_loss=1.4434 val_acc=0.7143
val_auc=0.8586
ep 190: train_loss=0.0411 val_loss=1.3773 val_acc=0.6939
val_auc=0.8485
ep 200: train_loss=0.0400 val_loss=1.7484 val_acc=0.6939
val_auc=0.8434
--> test_acc=0.7705, test_auc=0.9215

==== Running: optimizer=Adam, batch=16, lr=0.001 ====
ep 01: train_loss=0.6917 val_loss=0.6793 val_acc=0.5918 val_auc=0.7660
ep 10: train_loss=0.4734 val_loss=0.4629 val_acc=0.8571 val_auc=0.8822
ep 20: train_loss=0.4030 val_loss=0.3762 val_acc=0.8367 val_auc=0.9007
ep 30: train_loss=0.3631 val_loss=0.3605 val_acc=0.8571 val_auc=0.9125
ep 40: train_loss=0.3364 val_loss=0.3510 val_acc=0.8571 val_auc=0.9158
ep 50: train_loss=0.3062 val_loss=0.3552 val_acc=0.8367 val_auc=0.9209
ep 60: train_loss=0.3027 val_loss=0.3574 val_acc=0.8367 val_auc=0.9141
ep 70: train_loss=0.2838 val_loss=0.3663 val_acc=0.8163 val_auc=0.9125
ep 80: train_loss=0.2903 val_loss=0.3656 val_acc=0.8163 val_auc=0.9108
ep 90: train_loss=0.2454 val_loss=0.3680 val_acc=0.8571 val_auc=0.9125
ep 100: train_loss=0.2377 val_loss=0.3891 val_acc=0.8163
val_auc=0.9108
ep 110: train_loss=0.2447 val_loss=0.4092 val_acc=0.8163
val_auc=0.9007
ep 120: train_loss=0.2461 val_loss=0.4430 val_acc=0.7755
val_auc=0.8805
ep 130: train_loss=0.2186 val_loss=0.4700 val_acc=0.7959
val_auc=0.8838
ep 140: train_loss=0.1875 val_loss=0.4831 val_acc=0.7755
val_auc=0.8822
ep 150: train_loss=0.2023 val_loss=0.5034 val_acc=0.7755
val_auc=0.8788
ep 160: train_loss=0.2144 val_loss=0.5288 val_acc=0.7755
```

```
val_auc=0.8636
ep 170: train_loss=0.1727 val_loss=0.5758 val_acc=0.7755
val_auc=0.8687
ep 180: train_loss=0.1558 val_loss=0.5297 val_acc=0.7755
val_auc=0.8754
ep 190: train_loss=0.1622 val_loss=0.5811 val_acc=0.7755
val_auc=0.8687
ep 200: train_loss=0.1434 val_loss=0.6222 val_acc=0.7755
val_auc=0.8687
--> test_acc=0.8361, test_auc=0.9535

==== Running: optimizer=Adam, batch=16, lr=0.0001 ====
ep 01: train_loss=0.6870 val_loss=0.6848 val_acc=0.5306 val_auc=0.7593
ep 10: train_loss=0.6751 val_loss=0.6740 val_acc=0.7551 val_auc=0.8535
ep 20: train_loss=0.6572 val_loss=0.6598 val_acc=0.7347 val_auc=0.8973
ep 30: train_loss=0.6412 val_loss=0.6411 val_acc=0.7959 val_auc=0.9040
ep 40: train_loss=0.6297 val_loss=0.6194 val_acc=0.8163 val_auc=0.9158
ep 50: train_loss=0.6024 val_loss=0.5901 val_acc=0.8571 val_auc=0.9175
ep 60: train_loss=0.5706 val_loss=0.5575 val_acc=0.8776 val_auc=0.9175
ep 70: train_loss=0.5560 val_loss=0.5268 val_acc=0.8571 val_auc=0.9192
ep 80: train_loss=0.5152 val_loss=0.4946 val_acc=0.8571 val_auc=0.9209
ep 90: train_loss=0.4855 val_loss=0.4662 val_acc=0.8367 val_auc=0.9226
ep 100: train_loss=0.4647 val_loss=0.4418 val_acc=0.8367
val_auc=0.9226
ep 110: train_loss=0.4484 val_loss=0.4225 val_acc=0.8367
val_auc=0.9192
ep 120: train_loss=0.4506 val_loss=0.4079 val_acc=0.8367
val_auc=0.9192
ep 130: train_loss=0.4241 val_loss=0.3929 val_acc=0.8571
val_auc=0.9209
ep 140: train_loss=0.3943 val_loss=0.3804 val_acc=0.8571
val_auc=0.9209
ep 150: train_loss=0.4023 val_loss=0.3720 val_acc=0.8367
val_auc=0.9226
ep 160: train_loss=0.3962 val_loss=0.3643 val_acc=0.8367
val_auc=0.9242
ep 170: train_loss=0.3915 val_loss=0.3596 val_acc=0.8367
val_auc=0.9242
ep 180: train_loss=0.3800 val_loss=0.3561 val_acc=0.8367
val_auc=0.9226
ep 190: train_loss=0.3789 val_loss=0.3515 val_acc=0.8367
val_auc=0.9226
ep 200: train_loss=0.3665 val_loss=0.3476 val_acc=0.8367
val_auc=0.9242
--> test_acc=0.8689, test_auc=0.9567

==== Running: optimizer=Adam, batch=16, lr=1e-05 ====
ep 01: train_loss=0.6925 val_loss=0.6976 val_acc=0.4490 val_auc=0.6734
ep 10: train_loss=0.6901 val_loss=0.6954 val_acc=0.4490 val_auc=0.6936
```

```
ep 20: train_loss=0.6914 val_loss=0.6934 val_acc=0.4490 val_auc=0.7222
ep 30: train_loss=0.6863 val_loss=0.6913 val_acc=0.4490 val_auc=0.7576
ep 40: train_loss=0.6823 val_loss=0.6891 val_acc=0.4490 val_auc=0.7744
ep 50: train_loss=0.6783 val_loss=0.6869 val_acc=0.4490 val_auc=0.7929
ep 60: train_loss=0.6809 val_loss=0.6849 val_acc=0.4490 val_auc=0.8047
ep 70: train_loss=0.6792 val_loss=0.6826 val_acc=0.4490 val_auc=0.8249
ep 80: train_loss=0.6763 val_loss=0.6804 val_acc=0.4490 val_auc=0.8316
ep 90: train_loss=0.6750 val_loss=0.6782 val_acc=0.4490 val_auc=0.8451
ep 100: train_loss=0.6806 val_loss=0.6760 val_acc=0.4490
val_auc=0.8468
ep 110: train_loss=0.6743 val_loss=0.6736 val_acc=0.4694
val_auc=0.8502
ep 120: train_loss=0.6673 val_loss=0.6714 val_acc=0.4898
val_auc=0.8586
ep 130: train_loss=0.6695 val_loss=0.6690 val_acc=0.5102
val_auc=0.8670
ep 140: train_loss=0.6616 val_loss=0.6665 val_acc=0.5306
val_auc=0.8687
ep 150: train_loss=0.6614 val_loss=0.6641 val_acc=0.5714
val_auc=0.8704
ep 160: train_loss=0.6569 val_loss=0.6618 val_acc=0.5714
val_auc=0.8687
ep 170: train_loss=0.6588 val_loss=0.6593 val_acc=0.6122
val_auc=0.8687
ep 180: train_loss=0.6532 val_loss=0.6567 val_acc=0.6531
val_auc=0.8670
ep 190: train_loss=0.6505 val_loss=0.6541 val_acc=0.7143
val_auc=0.8687
ep 200: train_loss=0.6587 val_loss=0.6516 val_acc=0.7551
val_auc=0.8721
--> test_acc=0.7213, test_auc=0.9361
```

```
==== Running: optimizer=Adam, batch=32, lr=0.1 ====
ep 01: train_loss=0.6207 val_loss=0.5678 val_acc=0.5714 val_auc=0.8805
ep 10: train_loss=0.3926 val_loss=0.3424 val_acc=0.8367 val_auc=0.8923
ep 20: train_loss=0.3100 val_loss=0.4250 val_acc=0.8163 val_auc=0.8704
ep 30: train_loss=0.3572 val_loss=0.3847 val_acc=0.8367 val_auc=0.8771
ep 40: train_loss=0.3331 val_loss=0.4290 val_acc=0.8776 val_auc=0.8906
ep 50: train_loss=0.3570 val_loss=0.8358 val_acc=0.8163 val_auc=0.8333
ep 60: train_loss=0.3047 val_loss=0.5855 val_acc=0.8163 val_auc=0.8552
ep 70: train_loss=0.2831 val_loss=0.5414 val_acc=0.8163 val_auc=0.9040
ep 80: train_loss=0.5576 val_loss=0.4088 val_acc=0.8163 val_auc=0.8030
ep 90: train_loss=0.3585 val_loss=0.4308 val_acc=0.8163 val_auc=0.8653
ep 100: train_loss=0.3617 val_loss=0.4217 val_acc=0.7959
val_auc=0.8064
ep 110: train_loss=0.2600 val_loss=0.4191 val_acc=0.8163
val_auc=0.8409
ep 120: train_loss=0.5296 val_loss=0.9514 val_acc=0.8163
val_auc=0.7971
```

```
ep 130: train_loss=0.4483 val_loss=0.4696 val_acc=0.7959  
val_auc=0.7357  
ep 140: train_loss=0.4031 val_loss=0.4865 val_acc=0.7959  
val_auc=0.7879  
ep 150: train_loss=0.4098 val_loss=0.7277 val_acc=0.8163  
val_auc=0.8098  
ep 160: train_loss=0.3606 val_loss=0.3875 val_acc=0.8367  
val_auc=0.8182  
ep 170: train_loss=0.3377 val_loss=0.7946 val_acc=0.7959  
val_auc=0.7761  
ep 180: train_loss=0.3496 val_loss=0.6045 val_acc=0.8163  
val_auc=0.8098  
ep 190: train_loss=0.3386 val_loss=0.5881 val_acc=0.8163  
val_auc=0.7963  
ep 200: train_loss=0.2904 val_loss=0.5699 val_acc=0.8163  
val_auc=0.8182  
--> test_acc=0.8197, test_auc=0.8837

==== Running: optimizer=Adam, batch=32, lr=0.01 ====
ep 01: train_loss=0.6319 val_loss=0.5368 val_acc=0.8163 val_auc=0.8939  
ep 10: train_loss=0.3310 val_loss=0.3566 val_acc=0.8367 val_auc=0.9141  
ep 20: train_loss=0.2603 val_loss=0.3868 val_acc=0.8163 val_auc=0.8956  
ep 30: train_loss=0.2384 val_loss=0.5061 val_acc=0.8367 val_auc=0.8822  
ep 40: train_loss=0.1564 val_loss=0.6242 val_acc=0.7755 val_auc=0.8737  
ep 50: train_loss=0.2648 val_loss=0.5960 val_acc=0.7551 val_auc=0.8434  
ep 60: train_loss=0.1926 val_loss=0.6253 val_acc=0.7143 val_auc=0.8569  
ep 70: train_loss=0.2140 val_loss=0.5405 val_acc=0.7959 val_auc=0.8956  
ep 80: train_loss=0.1163 val_loss=0.6059 val_acc=0.7959 val_auc=0.8737  
ep 90: train_loss=0.1709 val_loss=0.6096 val_acc=0.7959 val_auc=0.8771  
ep 100: train_loss=0.1246 val_loss=0.6920 val_acc=0.7551  
val_auc=0.8636  
ep 110: train_loss=0.1293 val_loss=0.7145 val_acc=0.7755  
val_auc=0.8822  
ep 120: train_loss=0.1326 val_loss=0.8322 val_acc=0.7551  
val_auc=0.8535  
ep 130: train_loss=0.1133 val_loss=0.9223 val_acc=0.7551  
val_auc=0.8468  
ep 140: train_loss=0.1068 val_loss=0.7242 val_acc=0.7959  
val_auc=0.8603  
ep 150: train_loss=0.1061 val_loss=0.7236 val_acc=0.7755  
val_auc=0.8838  
ep 160: train_loss=0.0808 val_loss=0.8837 val_acc=0.7959  
val_auc=0.8620  
ep 170: train_loss=0.0889 val_loss=0.8990 val_acc=0.7143  
val_auc=0.8519  
ep 180: train_loss=0.0578 val_loss=0.9574 val_acc=0.7755  
val_auc=0.8771  
ep 190: train_loss=0.0804 val_loss=0.9305 val_acc=0.8367  
val_auc=0.8822
```

```
ep 200: train_loss=0.0762 val_loss=0.8351 val_acc=0.7551
val_auc=0.8838
--> test_acc=0.8197, test_auc=0.9416

==== Running: optimizer=Adam, batch=32, lr=0.001 ====
ep 01: train_loss=0.7011 val_loss=0.6915 val_acc=0.4490 val_auc=0.7710
ep 10: train_loss=0.5807 val_loss=0.5518 val_acc=0.8571 val_auc=0.9310
ep 20: train_loss=0.4462 val_loss=0.3917 val_acc=0.8571 val_auc=0.9175
ep 30: train_loss=0.3989 val_loss=0.3492 val_acc=0.8571 val_auc=0.9226
ep 40: train_loss=0.3645 val_loss=0.3222 val_acc=0.8571 val_auc=0.9310
ep 50: train_loss=0.3663 val_loss=0.3288 val_acc=0.8571 val_auc=0.9276
ep 60: train_loss=0.3420 val_loss=0.3250 val_acc=0.8571 val_auc=0.9276
ep 70: train_loss=0.3312 val_loss=0.3255 val_acc=0.8571 val_auc=0.9276
ep 80: train_loss=0.3131 val_loss=0.3219 val_acc=0.8571 val_auc=0.9310
ep 90: train_loss=0.3401 val_loss=0.3252 val_acc=0.8571 val_auc=0.9276
ep 100: train_loss=0.3271 val_loss=0.3239 val_acc=0.8571
val_auc=0.9327
ep 110: train_loss=0.2906 val_loss=0.3229 val_acc=0.8571
val_auc=0.9310
ep 120: train_loss=0.2879 val_loss=0.3276 val_acc=0.8571
val_auc=0.9276
ep 130: train_loss=0.2703 val_loss=0.3348 val_acc=0.8367
val_auc=0.9175
ep 140: train_loss=0.2777 val_loss=0.3428 val_acc=0.8163
val_auc=0.9209
ep 150: train_loss=0.2604 val_loss=0.3423 val_acc=0.8367
val_auc=0.9175
ep 160: train_loss=0.2817 val_loss=0.3459 val_acc=0.8367
val_auc=0.9209
ep 170: train_loss=0.2788 val_loss=0.3384 val_acc=0.8163
val_auc=0.9158
ep 180: train_loss=0.2786 val_loss=0.3407 val_acc=0.8163
val_auc=0.9141
ep 190: train_loss=0.2562 val_loss=0.3426 val_acc=0.8367
val_auc=0.9125
ep 200: train_loss=0.2490 val_loss=0.3778 val_acc=0.8163
val_auc=0.9057
--> test_acc=0.8852, test_auc=0.9632

==== Running: optimizer=Adam, batch=32, lr=0.0001 ====
ep 01: train_loss=0.6833 val_loss=0.6834 val_acc=0.5918 val_auc=0.7391
ep 10: train_loss=0.6728 val_loss=0.6733 val_acc=0.6531 val_auc=0.8232
ep 20: train_loss=0.6714 val_loss=0.6649 val_acc=0.7347 val_auc=0.8603
ep 30: train_loss=0.6524 val_loss=0.6560 val_acc=0.7959 val_auc=0.8822
ep 40: train_loss=0.6501 val_loss=0.6463 val_acc=0.8367 val_auc=0.8889
ep 50: train_loss=0.6413 val_loss=0.6360 val_acc=0.8367 val_auc=0.8889
ep 60: train_loss=0.6333 val_loss=0.6254 val_acc=0.8571 val_auc=0.8990
ep 70: train_loss=0.6238 val_loss=0.6142 val_acc=0.8571 val_auc=0.8990
ep 80: train_loss=0.6128 val_loss=0.6013 val_acc=0.8776 val_auc=0.9091
```

```
ep 90: train_loss=0.5914 val_loss=0.5864 val_acc=0.8571 val_auc=0.9091
ep 100: train_loss=0.5790 val_loss=0.5714 val_acc=0.8571
val_auc=0.9091
ep 110: train_loss=0.5716 val_loss=0.5558 val_acc=0.8367
val_auc=0.9108
ep 120: train_loss=0.5469 val_loss=0.5408 val_acc=0.8367
val_auc=0.9108
ep 130: train_loss=0.5436 val_loss=0.5246 val_acc=0.8367
val_auc=0.9141
ep 140: train_loss=0.5230 val_loss=0.5081 val_acc=0.8367
val_auc=0.9125
ep 150: train_loss=0.5144 val_loss=0.4914 val_acc=0.8571
val_auc=0.9108
ep 160: train_loss=0.5060 val_loss=0.4767 val_acc=0.8571
val_auc=0.9125
ep 170: train_loss=0.4894 val_loss=0.4634 val_acc=0.8571
val_auc=0.9091
ep 180: train_loss=0.4649 val_loss=0.4521 val_acc=0.8571
val_auc=0.9091
ep 190: train_loss=0.4657 val_loss=0.4399 val_acc=0.8571
val_auc=0.9108
ep 200: train_loss=0.4402 val_loss=0.4290 val_acc=0.8571
val_auc=0.9108
--> test_acc=0.8852, test_auc=0.9643

==== Running: optimizer=Adam, batch=32, lr=1e-05 ====
ep 01: train_loss=0.6916 val_loss=0.6905 val_acc=0.5510 val_auc=0.5118
ep 10: train_loss=0.6906 val_loss=0.6896 val_acc=0.5510 val_auc=0.5202
ep 20: train_loss=0.6876 val_loss=0.6887 val_acc=0.5510 val_auc=0.5404
ep 30: train_loss=0.6839 val_loss=0.6877 val_acc=0.5510 val_auc=0.5522
ep 40: train_loss=0.6852 val_loss=0.6869 val_acc=0.5510 val_auc=0.5657
ep 50: train_loss=0.6873 val_loss=0.6860 val_acc=0.5510 val_auc=0.5741
ep 60: train_loss=0.6856 val_loss=0.6851 val_acc=0.5510 val_auc=0.5909
ep 70: train_loss=0.6824 val_loss=0.6843 val_acc=0.5510 val_auc=0.6077
ep 80: train_loss=0.6882 val_loss=0.6836 val_acc=0.5510 val_auc=0.6195
ep 90: train_loss=0.6764 val_loss=0.6828 val_acc=0.5510 val_auc=0.6279
ep 100: train_loss=0.6769 val_loss=0.6820 val_acc=0.5510
val_auc=0.6330
ep 110: train_loss=0.6806 val_loss=0.6812 val_acc=0.5510
val_auc=0.6431
ep 120: train_loss=0.6788 val_loss=0.6803 val_acc=0.5510
val_auc=0.6515
ep 130: train_loss=0.6764 val_loss=0.6795 val_acc=0.5510
val_auc=0.6616
ep 140: train_loss=0.6801 val_loss=0.6787 val_acc=0.5510
val_auc=0.6700
ep 150: train_loss=0.6829 val_loss=0.6780 val_acc=0.5510
val_auc=0.6717
ep 160: train_loss=0.6845 val_loss=0.6772 val_acc=0.5510
```

```
val_auc=0.6768
ep 170: train_loss=0.6710 val_loss=0.6764 val_acc=0.5510
val_auc=0.6835
ep 180: train_loss=0.6647 val_loss=0.6756 val_acc=0.5510
val_auc=0.7037
ep 190: train_loss=0.6724 val_loss=0.6747 val_acc=0.5510
val_auc=0.7071
ep 200: train_loss=0.6753 val_loss=0.6739 val_acc=0.5510
val_auc=0.7121
--> test_acc=0.5738, test_auc=0.7327

==== Running: optimizer=Adam, batch=64, lr=0.1 ====
ep 01: train_loss=0.6791 val_loss=0.5085 val_acc=0.8163 val_auc=0.8906
ep 10: train_loss=0.3104 val_loss=0.4000 val_acc=0.8163 val_auc=0.9040
ep 20: train_loss=0.3295 val_loss=0.4286 val_acc=0.8367 val_auc=0.8813
ep 30: train_loss=0.6368 val_loss=0.3740 val_acc=0.8163 val_auc=0.8729
ep 40: train_loss=0.3224 val_loss=0.8733 val_acc=0.7959 val_auc=0.8384
ep 50: train_loss=0.3070 val_loss=0.4093 val_acc=0.7959 val_auc=0.8939
ep 60: train_loss=0.2766 val_loss=0.5367 val_acc=0.6939 val_auc=0.8687
ep 70: train_loss=0.2627 val_loss=0.4374 val_acc=0.7551 val_auc=0.8948
ep 80: train_loss=0.3709 val_loss=0.3823 val_acc=0.8163 val_auc=0.8855
ep 90: train_loss=0.4279 val_loss=0.3122 val_acc=0.8571 val_auc=0.9125
ep 100: train_loss=0.3506 val_loss=0.8975 val_acc=0.6939
val_auc=0.6978
ep 110: train_loss=0.4380 val_loss=0.6779 val_acc=0.8367
val_auc=0.8872
ep 120: train_loss=0.2636 val_loss=0.4681 val_acc=0.8367
val_auc=0.8805
ep 130: train_loss=0.2176 val_loss=0.9298 val_acc=0.8367
val_auc=0.8813
ep 140: train_loss=0.2258 val_loss=0.5432 val_acc=0.8367
val_auc=0.9099
ep 150: train_loss=0.2213 val_loss=0.5272 val_acc=0.8571
val_auc=0.9141
ep 160: train_loss=0.2627 val_loss=0.7980 val_acc=0.7959
val_auc=0.9015
ep 170: train_loss=0.2028 val_loss=0.8871 val_acc=0.8367
val_auc=0.8729
ep 180: train_loss=0.2418 val_loss=1.3205 val_acc=0.7959
val_auc=0.8577
ep 190: train_loss=0.2211 val_loss=1.4353 val_acc=0.8367
val_auc=0.8855
ep 200: train_loss=0.2529 val_loss=1.0488 val_acc=0.8163
val_auc=0.8384
--> test_acc=0.8689, test_auc=0.9324

==== Running: optimizer=Adam, batch=64, lr=0.01 ====
ep 01: train_loss=0.6771 val_loss=0.6116 val_acc=0.7959 val_auc=0.9040
ep 10: train_loss=0.3648 val_loss=0.3734 val_acc=0.8163 val_auc=0.9024
```

```
ep 20: train_loss=0.3178 val_loss=0.3298 val_acc=0.8571 val_auc=0.9276
ep 30: train_loss=0.3093 val_loss=0.3835 val_acc=0.7959 val_auc=0.9040
ep 40: train_loss=0.2556 val_loss=0.4758 val_acc=0.7959 val_auc=0.8737
ep 50: train_loss=0.2217 val_loss=0.4920 val_acc=0.7755 val_auc=0.8754
ep 60: train_loss=0.2322 val_loss=0.5204 val_acc=0.7755 val_auc=0.8721
ep 70: train_loss=0.2273 val_loss=0.6787 val_acc=0.7347 val_auc=0.8552
ep 80: train_loss=0.1561 val_loss=0.7381 val_acc=0.7143 val_auc=0.8552
ep 90: train_loss=0.1486 val_loss=0.7626 val_acc=0.7959 val_auc=0.8535
ep 100: train_loss=0.1632 val_loss=0.8157 val_acc=0.7959
val_auc=0.8468
ep 110: train_loss=0.1138 val_loss=0.9298 val_acc=0.7755
val_auc=0.8451
ep 120: train_loss=0.1660 val_loss=0.9610 val_acc=0.7755
val_auc=0.8552
ep 130: train_loss=0.2373 val_loss=0.8657 val_acc=0.7959
val_auc=0.8670
ep 140: train_loss=0.4324 val_loss=0.9082 val_acc=0.7959
val_auc=0.8586
ep 150: train_loss=0.2219 val_loss=0.5213 val_acc=0.7551
val_auc=0.8687
ep 160: train_loss=0.1595 val_loss=0.6147 val_acc=0.7551
val_auc=0.8586
ep 170: train_loss=0.2115 val_loss=0.7233 val_acc=0.7551
val_auc=0.8586
ep 180: train_loss=0.1384 val_loss=0.6787 val_acc=0.7551
val_auc=0.8737
ep 190: train_loss=0.1434 val_loss=0.7640 val_acc=0.7551
val_auc=0.8687
ep 200: train_loss=0.1098 val_loss=0.8488 val_acc=0.7551
val_auc=0.8535
--> test_acc=0.7869, test_auc=0.9221
```

```
==== Running: optimizer=Adam, batch=64, lr=0.001 ====
ep 01: train_loss=0.6931 val_loss=0.6867 val_acc=0.5510 val_auc=0.5741
ep 10: train_loss=0.6487 val_loss=0.6410 val_acc=0.7959 val_auc=0.8603
ep 20: train_loss=0.5873 val_loss=0.5825 val_acc=0.8163 val_auc=0.8771
ep 30: train_loss=0.5083 val_loss=0.5008 val_acc=0.8367 val_auc=0.8889
ep 40: train_loss=0.4692 val_loss=0.4361 val_acc=0.8571 val_auc=0.8973
ep 50: train_loss=0.3993 val_loss=0.3919 val_acc=0.8367 val_auc=0.9074
ep 60: train_loss=0.3760 val_loss=0.3745 val_acc=0.8571 val_auc=0.9040
ep 70: train_loss=0.3631 val_loss=0.3573 val_acc=0.8571 val_auc=0.9091
ep 80: train_loss=0.3744 val_loss=0.3495 val_acc=0.8571 val_auc=0.9141
ep 90: train_loss=0.3539 val_loss=0.3502 val_acc=0.8571 val_auc=0.9141
ep 100: train_loss=0.3413 val_loss=0.3488 val_acc=0.8367
val_auc=0.9158
ep 110: train_loss=0.3421 val_loss=0.3495 val_acc=0.8367
val_auc=0.9175
ep 120: train_loss=0.3641 val_loss=0.3522 val_acc=0.8367
val_auc=0.9125
```

```
ep 130: train_loss=0.3540 val_loss=0.3445 val_acc=0.8367  
val_auc=0.9175  
ep 140: train_loss=0.3320 val_loss=0.3356 val_acc=0.8776  
val_auc=0.9192  
ep 150: train_loss=0.3452 val_loss=0.3381 val_acc=0.8571  
val_auc=0.9209  
ep 160: train_loss=0.3257 val_loss=0.3410 val_acc=0.8367  
val_auc=0.9192  
ep 170: train_loss=0.3403 val_loss=0.3449 val_acc=0.8163  
val_auc=0.9158  
ep 180: train_loss=0.3500 val_loss=0.3496 val_acc=0.8367  
val_auc=0.9108  
ep 190: train_loss=0.3077 val_loss=0.3566 val_acc=0.8367  
val_auc=0.9057  
ep 200: train_loss=0.3167 val_loss=0.3564 val_acc=0.8163  
val_auc=0.9040  
--> test_acc=0.8689, test_auc=0.9535

==== Running: optimizer=Adam, batch=64, lr=0.0001 ====
ep 01: train_loss=0.6966 val_loss=0.6975 val_acc=0.4082 val_auc=0.4865  
ep 10: train_loss=0.6948 val_loss=0.6936 val_acc=0.4082 val_auc=0.5909  
ep 20: train_loss=0.6920 val_loss=0.6903 val_acc=0.4286 val_auc=0.6566  
ep 30: train_loss=0.6907 val_loss=0.6871 val_acc=0.4286 val_auc=0.7003  
ep 40: train_loss=0.6892 val_loss=0.6839 val_acc=0.4694 val_auc=0.7424  
ep 50: train_loss=0.6822 val_loss=0.6803 val_acc=0.4898 val_auc=0.7845  
ep 60: train_loss=0.6789 val_loss=0.6767 val_acc=0.5306 val_auc=0.8047  
ep 70: train_loss=0.6756 val_loss=0.6737 val_acc=0.5714 val_auc=0.8098  
ep 80: train_loss=0.6686 val_loss=0.6708 val_acc=0.5918 val_auc=0.8131  
ep 90: train_loss=0.6659 val_loss=0.6680 val_acc=0.6327 val_auc=0.8182  
ep 100: train_loss=0.6657 val_loss=0.6645 val_acc=0.7143  
val_auc=0.8199  
ep 110: train_loss=0.6619 val_loss=0.6602 val_acc=0.7347  
val_auc=0.8249  
ep 120: train_loss=0.6627 val_loss=0.6556 val_acc=0.7551  
val_auc=0.8215  
ep 130: train_loss=0.6528 val_loss=0.6506 val_acc=0.7551  
val_auc=0.8249  
ep 140: train_loss=0.6442 val_loss=0.6446 val_acc=0.7755  
val_auc=0.8350  
ep 150: train_loss=0.6434 val_loss=0.6397 val_acc=0.7755  
val_auc=0.8418  
ep 160: train_loss=0.6317 val_loss=0.6332 val_acc=0.7959  
val_auc=0.8502  
ep 170: train_loss=0.6311 val_loss=0.6274 val_acc=0.7755  
val_auc=0.8502  
ep 180: train_loss=0.6260 val_loss=0.6202 val_acc=0.7959  
val_auc=0.8535  
ep 190: train_loss=0.6157 val_loss=0.6144 val_acc=0.8163  
val_auc=0.8586
```

```
ep 200: train_loss=0.6150 val_loss=0.6076 val_acc=0.8163
val_auc=0.8620
--> test_acc=0.7869, test_auc=0.9307

==== Running: optimizer=Adam, batch=64, lr=1e-05 ====
ep 01: train_loss=0.6933 val_loss=0.6918 val_acc=0.5510 val_auc=0.4394
ep 10: train_loss=0.6935 val_loss=0.6915 val_acc=0.5510 val_auc=0.4478
ep 20: train_loss=0.6939 val_loss=0.6913 val_acc=0.5510 val_auc=0.4562
ep 30: train_loss=0.6921 val_loss=0.6911 val_acc=0.5510 val_auc=0.4663
ep 40: train_loss=0.6923 val_loss=0.6909 val_acc=0.5510 val_auc=0.4747
ep 50: train_loss=0.6921 val_loss=0.6907 val_acc=0.5510 val_auc=0.4798
ep 60: train_loss=0.6915 val_loss=0.6905 val_acc=0.5510 val_auc=0.4848
ep 70: train_loss=0.6943 val_loss=0.6902 val_acc=0.5510 val_auc=0.4949
ep 80: train_loss=0.6927 val_loss=0.6900 val_acc=0.5510 val_auc=0.5034
ep 90: train_loss=0.6911 val_loss=0.6898 val_acc=0.5510 val_auc=0.5118
ep 100: train_loss=0.6908 val_loss=0.6896 val_acc=0.5510
val_auc=0.5185
ep 110: train_loss=0.6920 val_loss=0.6894 val_acc=0.5510
val_auc=0.5303
ep 120: train_loss=0.6929 val_loss=0.6891 val_acc=0.5510
val_auc=0.5370
ep 130: train_loss=0.6884 val_loss=0.6890 val_acc=0.5510
val_auc=0.5354
ep 140: train_loss=0.6926 val_loss=0.6887 val_acc=0.5510
val_auc=0.5404
ep 150: train_loss=0.6898 val_loss=0.6886 val_acc=0.5510
val_auc=0.5471
ep 160: train_loss=0.6870 val_loss=0.6884 val_acc=0.5510
val_auc=0.5505
ep 170: train_loss=0.6922 val_loss=0.6881 val_acc=0.5510
val_auc=0.5657
ep 180: train_loss=0.6900 val_loss=0.6879 val_acc=0.5510
val_auc=0.5741
ep 190: train_loss=0.6898 val_loss=0.6877 val_acc=0.5510
val_auc=0.5791
ep 200: train_loss=0.6900 val_loss=0.6875 val_acc=0.5510
val_auc=0.5909
--> test_acc=0.5902, test_auc=0.5141

==== Running: optimizer=RMSprop, batch=16, lr=0.1 ====
ep 01: train_loss=36.3947 val_loss=0.8884 val_acc=0.7959
val_auc=0.8611
ep 10: train_loss=0.4426 val_loss=0.4200 val_acc=0.7347 val_auc=0.8889
ep 20: train_loss=0.4601 val_loss=0.4257 val_acc=0.7551 val_auc=0.8923
ep 30: train_loss=0.4466 val_loss=0.5401 val_acc=0.6735 val_auc=0.8401
ep 40: train_loss=0.4213 val_loss=0.4241 val_acc=0.7551 val_auc=0.8460
ep 50: train_loss=0.3597 val_loss=0.5971 val_acc=0.7143 val_auc=0.8367
ep 60: train_loss=0.5366 val_loss=0.9189 val_acc=0.7551 val_auc=0.8426
ep 70: train_loss=0.4033 val_loss=0.4478 val_acc=0.7959 val_auc=0.8805
ep 80: train_loss=0.6647 val_loss=0.5453 val_acc=0.7755 val_auc=0.8131
```

```
ep 90: train_loss=0.4854 val_loss=1.2348 val_acc=0.7755 val_auc=0.8165
ep 100: train_loss=0.3932 val_loss=0.9696 val_acc=0.8163
val_auc=0.8325
ep 110: train_loss=0.4126 val_loss=0.9920 val_acc=0.7959
val_auc=0.8064
ep 120: train_loss=0.3698 val_loss=2.7271 val_acc=0.7143
val_auc=0.7391
ep 130: train_loss=0.9519 val_loss=2.3812 val_acc=0.7551
val_auc=0.7778
ep 140: train_loss=2.3394 val_loss=0.4741 val_acc=0.7347
val_auc=0.8098
ep 150: train_loss=0.5346 val_loss=0.7368 val_acc=0.8163
val_auc=0.8384
ep 160: train_loss=0.3261 val_loss=1.0382 val_acc=0.7959
val_auc=0.8434
ep 170: train_loss=0.4791 val_loss=1.9770 val_acc=0.7959
val_auc=0.8350
ep 180: train_loss=0.4785 val_loss=0.7315 val_acc=0.7347
val_auc=0.7912
ep 190: train_loss=0.5680 val_loss=1.8766 val_acc=0.7551
val_auc=0.8098
ep 200: train_loss=0.3431 val_loss=2.8330 val_acc=0.6327
val_auc=0.5589
--> test_acc=0.6721, test_auc=0.6261

==== Running: optimizer=RMSprop, batch=16, lr=0.01 ====
ep 01: train_loss=0.5788 val_loss=0.3995 val_acc=0.8367 val_auc=0.9057
ep 10: train_loss=0.2723 val_loss=0.4328 val_acc=0.7755 val_auc=0.9024
ep 20: train_loss=0.2240 val_loss=0.5211 val_acc=0.7755 val_auc=0.8822
ep 30: train_loss=0.1888 val_loss=0.6142 val_acc=0.7755 val_auc=0.8636
ep 40: train_loss=0.1323 val_loss=0.7171 val_acc=0.8163 val_auc=0.8687
ep 50: train_loss=0.1301 val_loss=0.7720 val_acc=0.7959 val_auc=0.8838
ep 60: train_loss=0.1535 val_loss=0.8188 val_acc=0.8163 val_auc=0.8805
ep 70: train_loss=0.0691 val_loss=0.9488 val_acc=0.8163 val_auc=0.8923
ep 80: train_loss=0.1121 val_loss=0.9460 val_acc=0.8367 val_auc=0.8872
ep 90: train_loss=0.1524 val_loss=0.9806 val_acc=0.8163 val_auc=0.8822
ep 100: train_loss=0.0987 val_loss=1.0463 val_acc=0.7959
val_auc=0.8838
ep 110: train_loss=0.0803 val_loss=1.4277 val_acc=0.7755
val_auc=0.8653
ep 120: train_loss=0.0609 val_loss=1.4960 val_acc=0.7755
val_auc=0.8704
ep 130: train_loss=0.0578 val_loss=1.4584 val_acc=0.7959
val_auc=0.8687
ep 140: train_loss=0.1023 val_loss=1.5891 val_acc=0.7551
val_auc=0.8401
ep 150: train_loss=0.0461 val_loss=1.3604 val_acc=0.7959
val_auc=0.8721
ep 160: train_loss=0.0646 val_loss=1.3529 val_acc=0.8163
```

```
val_auc=0.8889
ep 170: train_loss=0.0554 val_loss=1.0996 val_acc=0.8163
val_auc=0.8771
ep 180: train_loss=0.0708 val_loss=1.5376 val_acc=0.7755
val_auc=0.8434
ep 190: train_loss=0.0470 val_loss=1.7256 val_acc=0.7551
val_auc=0.8687
ep 200: train_loss=0.0609 val_loss=1.9465 val_acc=0.7551
val_auc=0.8426
--> test_acc=0.8197, test_auc=0.9394

==== Running: optimizer=RMSprop, batch=16, lr=0.001 ====
ep 01: train_loss=0.6763 val_loss=0.6122 val_acc=0.8367 val_auc=0.9057
ep 10: train_loss=0.3883 val_loss=0.3452 val_acc=0.8367 val_auc=0.9158
ep 20: train_loss=0.3485 val_loss=0.3287 val_acc=0.8367 val_auc=0.9242
ep 30: train_loss=0.3481 val_loss=0.3287 val_acc=0.8571 val_auc=0.9226
ep 40: train_loss=0.3182 val_loss=0.3418 val_acc=0.8571 val_auc=0.9125
ep 50: train_loss=0.2876 val_loss=0.3414 val_acc=0.8367 val_auc=0.9125
ep 60: train_loss=0.2937 val_loss=0.3468 val_acc=0.8367 val_auc=0.9074
ep 70: train_loss=0.2935 val_loss=0.3648 val_acc=0.8163 val_auc=0.8973
ep 80: train_loss=0.2678 val_loss=0.3669 val_acc=0.8367 val_auc=0.9024
ep 90: train_loss=0.2389 val_loss=0.3754 val_acc=0.8163 val_auc=0.9024
ep 100: train_loss=0.2278 val_loss=0.3869 val_acc=0.7959
val_auc=0.8990
ep 110: train_loss=0.2178 val_loss=0.3929 val_acc=0.7755
val_auc=0.8923
ep 120: train_loss=0.2061 val_loss=0.4205 val_acc=0.7551
val_auc=0.8805
ep 130: train_loss=0.1656 val_loss=0.4531 val_acc=0.7755
val_auc=0.8788
ep 140: train_loss=0.1850 val_loss=0.4560 val_acc=0.7347
val_auc=0.8737
ep 150: train_loss=0.1708 val_loss=0.4991 val_acc=0.7755
val_auc=0.8670
ep 160: train_loss=0.1510 val_loss=0.5101 val_acc=0.7755
val_auc=0.8636
ep 170: train_loss=0.1487 val_loss=0.5377 val_acc=0.7755
val_auc=0.8603
ep 180: train_loss=0.1413 val_loss=0.5501 val_acc=0.7347
val_auc=0.8535
ep 190: train_loss=0.1129 val_loss=0.5756 val_acc=0.7755
val_auc=0.8603
ep 200: train_loss=0.1212 val_loss=0.5755 val_acc=0.7755
val_auc=0.8636
--> test_acc=0.8197, test_auc=0.9556

==== Running: optimizer=RMSprop, batch=16, lr=0.0001 ====
ep 01: train_loss=0.7050 val_loss=0.6915 val_acc=0.5510 val_auc=0.4747
ep 10: train_loss=0.6636 val_loss=0.6590 val_acc=0.5918 val_auc=0.8165
```

```
ep 20: train_loss=0.6404 val_loss=0.6336 val_acc=0.7959 val_auc=0.8586
ep 30: train_loss=0.6166 val_loss=0.6100 val_acc=0.8571 val_auc=0.8754
ep 40: train_loss=0.5975 val_loss=0.5868 val_acc=0.8571 val_auc=0.8838
ep 50: train_loss=0.5783 val_loss=0.5593 val_acc=0.8367 val_auc=0.8872
ep 60: train_loss=0.5596 val_loss=0.5323 val_acc=0.8367 val_auc=0.8872
ep 70: train_loss=0.5241 val_loss=0.5032 val_acc=0.8367 val_auc=0.8923
ep 80: train_loss=0.4844 val_loss=0.4748 val_acc=0.8367 val_auc=0.8956
ep 90: train_loss=0.4823 val_loss=0.4508 val_acc=0.8367 val_auc=0.9024
ep 100: train_loss=0.4695 val_loss=0.4305 val_acc=0.8367
val_auc=0.9024
ep 110: train_loss=0.4620 val_loss=0.4169 val_acc=0.8163
val_auc=0.9057
ep 120: train_loss=0.4290 val_loss=0.4044 val_acc=0.8163
val_auc=0.9091
ep 130: train_loss=0.4307 val_loss=0.3942 val_acc=0.8163
val_auc=0.9091
ep 140: train_loss=0.4109 val_loss=0.3877 val_acc=0.8163
val_auc=0.9141
ep 150: train_loss=0.4167 val_loss=0.3804 val_acc=0.8163
val_auc=0.9141
ep 160: train_loss=0.4009 val_loss=0.3730 val_acc=0.8163
val_auc=0.9209
ep 170: train_loss=0.3947 val_loss=0.3687 val_acc=0.8163
val_auc=0.9175
ep 180: train_loss=0.3966 val_loss=0.3638 val_acc=0.8163
val_auc=0.9175
ep 190: train_loss=0.3995 val_loss=0.3601 val_acc=0.8163
val_auc=0.9175
ep 200: train_loss=0.3774 val_loss=0.3569 val_acc=0.8163
val_auc=0.9175
--> test_acc=0.8689, test_auc=0.9632

==== Running: optimizer=RMSprop, batch=16, lr=1e-05 ====
ep 01: train_loss=0.7017 val_loss=0.7085 val_acc=0.4490 val_auc=0.5960
ep 10: train_loss=0.6994 val_loss=0.7063 val_acc=0.4490 val_auc=0.6347
ep 20: train_loss=0.7006 val_loss=0.7045 val_acc=0.4490 val_auc=0.6599
ep 30: train_loss=0.6973 val_loss=0.7030 val_acc=0.4490 val_auc=0.6785
ep 40: train_loss=0.6968 val_loss=0.7015 val_acc=0.4490 val_auc=0.6987
ep 50: train_loss=0.6923 val_loss=0.7000 val_acc=0.4490 val_auc=0.7155
ep 60: train_loss=0.6969 val_loss=0.6986 val_acc=0.4490 val_auc=0.7256
ep 70: train_loss=0.6902 val_loss=0.6971 val_acc=0.4490 val_auc=0.7458
ep 80: train_loss=0.6889 val_loss=0.6956 val_acc=0.4490 val_auc=0.7508
ep 90: train_loss=0.6896 val_loss=0.6942 val_acc=0.4490 val_auc=0.7626
ep 100: train_loss=0.6858 val_loss=0.6927 val_acc=0.4490
val_auc=0.7643
ep 110: train_loss=0.6850 val_loss=0.6911 val_acc=0.4490
val_auc=0.7727
ep 120: train_loss=0.6866 val_loss=0.6896 val_acc=0.4490
val_auc=0.7795
```

```
ep 130: train_loss=0.6876 val_loss=0.6881 val_acc=0.4490  
val_auc=0.7845  
ep 140: train_loss=0.6821 val_loss=0.6865 val_acc=0.4490  
val_auc=0.7963  
ep 150: train_loss=0.6839 val_loss=0.6849 val_acc=0.4490  
val_auc=0.7980  
ep 160: train_loss=0.6788 val_loss=0.6835 val_acc=0.4490  
val_auc=0.7997  
ep 170: train_loss=0.6783 val_loss=0.6820 val_acc=0.4490  
val_auc=0.8047  
ep 180: train_loss=0.6767 val_loss=0.6804 val_acc=0.4490  
val_auc=0.8081  
ep 190: train_loss=0.6742 val_loss=0.6788 val_acc=0.4490  
val_auc=0.8114  
ep 200: train_loss=0.6703 val_loss=0.6772 val_acc=0.4490  
val_auc=0.8131  
--> test_acc=0.4590, test_auc=0.8766

==== Running: optimizer=RMSprop, batch=32, lr=0.1 ====
ep 01: train_loss=79.4415 val_loss=1.9698 val_acc=0.8163  
val_auc=0.8805  
ep 10: train_loss=0.5003 val_loss=0.6158 val_acc=0.7959 val_auc=0.8308  
ep 20: train_loss=0.5550 val_loss=0.4388 val_acc=0.8367 val_auc=0.8561  
ep 30: train_loss=0.3607 val_loss=0.5486 val_acc=0.7959 val_auc=0.8443  
ep 40: train_loss=0.3678 val_loss=0.6088 val_acc=0.7959 val_auc=0.8367  
ep 50: train_loss=0.3490 val_loss=0.5091 val_acc=0.7959 val_auc=0.8005  
ep 60: train_loss=0.3121 val_loss=0.7220 val_acc=0.7755 val_auc=0.8418  
ep 70: train_loss=0.3727 val_loss=0.6600 val_acc=0.8367 val_auc=0.8687  
ep 80: train_loss=0.3984 val_loss=0.4856 val_acc=0.7959 val_auc=0.8224  
ep 90: train_loss=0.3644 val_loss=0.4967 val_acc=0.7959 val_auc=0.8114  
ep 100: train_loss=0.3826 val_loss=0.3811 val_acc=0.8367  
val_auc=0.8687  
ep 110: train_loss=0.2846 val_loss=0.7475 val_acc=0.7755  
val_auc=0.8266  
ep 120: train_loss=0.3128 val_loss=0.6814 val_acc=0.8367  
val_auc=0.8266  
ep 130: train_loss=0.4580 val_loss=0.3887 val_acc=0.8163  
val_auc=0.8316  
ep 140: train_loss=0.2748 val_loss=0.2984 val_acc=0.8980  
val_auc=0.8973  
ep 150: train_loss=0.3679 val_loss=0.4578 val_acc=0.7755  
val_auc=0.7811  
ep 160: train_loss=0.3215 val_loss=0.8238 val_acc=0.7551  
val_auc=0.7854  
ep 170: train_loss=0.4158 val_loss=0.5809 val_acc=0.7959  
val_auc=0.7980  
ep 180: train_loss=0.5836 val_loss=3.1412 val_acc=0.8163  
val_auc=0.8266  
ep 190: train_loss=0.3420 val_loss=0.4396 val_acc=0.7755
```

```
val_auc=0.8401
ep 200: train_loss=0.3502 val_loss=0.4887 val_acc=0.8163
val_auc=0.8594
--> test_acc=0.8197, test_auc=0.8891

==== Running: optimizer=RMSprop, batch=32, lr=0.01 ====
ep 01: train_loss=0.5821 val_loss=0.4808 val_acc=0.7959 val_auc=0.9040
ep 10: train_loss=0.3054 val_loss=0.3384 val_acc=0.8163 val_auc=0.9175
ep 20: train_loss=0.2865 val_loss=0.4682 val_acc=0.7959 val_auc=0.8805
ep 30: train_loss=0.1953 val_loss=0.4819 val_acc=0.7755 val_auc=0.8788
ep 40: train_loss=0.2240 val_loss=0.5301 val_acc=0.7755 val_auc=0.8670
ep 50: train_loss=0.1722 val_loss=0.6438 val_acc=0.7959 val_auc=0.8620
ep 60: train_loss=0.1642 val_loss=0.7705 val_acc=0.7959 val_auc=0.8737
ep 70: train_loss=0.1513 val_loss=0.8048 val_acc=0.7551 val_auc=0.8653
ep 80: train_loss=0.0939 val_loss=0.8243 val_acc=0.7755 val_auc=0.8552
ep 90: train_loss=0.1373 val_loss=0.8822 val_acc=0.7959 val_auc=0.8519
ep 100: train_loss=0.1307 val_loss=0.9131 val_acc=0.7755
val_auc=0.8636
ep 110: train_loss=0.0821 val_loss=0.8885 val_acc=0.7959
val_auc=0.8653
ep 120: train_loss=0.1042 val_loss=0.9769 val_acc=0.7551
val_auc=0.8519
ep 130: train_loss=0.0881 val_loss=1.0746 val_acc=0.7551
val_auc=0.8552
ep 140: train_loss=0.0945 val_loss=0.7580 val_acc=0.7755
val_auc=0.8653
ep 150: train_loss=0.0602 val_loss=0.9513 val_acc=0.7551
val_auc=0.8636
ep 160: train_loss=0.0628 val_loss=1.1510 val_acc=0.7959
val_auc=0.8653
ep 170: train_loss=0.0475 val_loss=1.1791 val_acc=0.7959
val_auc=0.8620
ep 180: train_loss=0.0776 val_loss=1.2674 val_acc=0.7959
val_auc=0.8620
ep 190: train_loss=0.0392 val_loss=1.4167 val_acc=0.8163
val_auc=0.8535
ep 200: train_loss=0.1016 val_loss=1.3684 val_acc=0.7959
val_auc=0.8603
--> test_acc=0.7869, test_auc=0.9134

==== Running: optimizer=RMSprop, batch=32, lr=0.001 ====
ep 01: train_loss=0.6315 val_loss=0.5835 val_acc=0.8776 val_auc=0.8939
ep 10: train_loss=0.3861 val_loss=0.3782 val_acc=0.8571 val_auc=0.9259
ep 20: train_loss=0.3657 val_loss=0.3535 val_acc=0.8776 val_auc=0.9226
ep 30: train_loss=0.3539 val_loss=0.3422 val_acc=0.8776 val_auc=0.9259
ep 40: train_loss=0.3299 val_loss=0.3392 val_acc=0.8776 val_auc=0.9259
ep 50: train_loss=0.3309 val_loss=0.3390 val_acc=0.8776 val_auc=0.9209
ep 60: train_loss=0.3150 val_loss=0.3388 val_acc=0.8571 val_auc=0.9242
ep 70: train_loss=0.3077 val_loss=0.3425 val_acc=0.8367 val_auc=0.9158
```

```
ep 80: train_loss=0.2853 val_loss=0.3455 val_acc=0.8367 val_auc=0.9192
ep 90: train_loss=0.3239 val_loss=0.3509 val_acc=0.8367 val_auc=0.9125
ep 100: train_loss=0.2685 val_loss=0.3552 val_acc=0.8367
val_auc=0.9158
ep 110: train_loss=0.2810 val_loss=0.3548 val_acc=0.8367
val_auc=0.9091
ep 120: train_loss=0.2819 val_loss=0.3524 val_acc=0.8367
val_auc=0.9125
ep 130: train_loss=0.2776 val_loss=0.3536 val_acc=0.8367
val_auc=0.9141
ep 140: train_loss=0.2615 val_loss=0.3652 val_acc=0.8367
val_auc=0.9074
ep 150: train_loss=0.2433 val_loss=0.3774 val_acc=0.8163
val_auc=0.9057
ep 160: train_loss=0.2481 val_loss=0.3701 val_acc=0.8163
val_auc=0.9040
ep 170: train_loss=0.2169 val_loss=0.3888 val_acc=0.8367
val_auc=0.9024
ep 180: train_loss=0.2397 val_loss=0.3995 val_acc=0.7959
val_auc=0.8939
ep 190: train_loss=0.2087 val_loss=0.4063 val_acc=0.7959
val_auc=0.8956
ep 200: train_loss=0.2186 val_loss=0.4362 val_acc=0.7959
val_auc=0.8889
--> test_acc=0.8689, test_auc=0.9524

==== Running: optimizer=RMSprop, batch=32, lr=0.0001 ====
ep 01: train_loss=0.6795 val_loss=0.6779 val_acc=0.6735 val_auc=0.7845
ep 10: train_loss=0.6609 val_loss=0.6575 val_acc=0.7755 val_auc=0.8468
ep 20: train_loss=0.6500 val_loss=0.6460 val_acc=0.8163 val_auc=0.8603
ep 30: train_loss=0.6355 val_loss=0.6365 val_acc=0.8163 val_auc=0.8653
ep 40: train_loss=0.6308 val_loss=0.6272 val_acc=0.7755 val_auc=0.8704
ep 50: train_loss=0.6137 val_loss=0.6177 val_acc=0.7755 val_auc=0.8721
ep 60: train_loss=0.6094 val_loss=0.6064 val_acc=0.7755 val_auc=0.8737
ep 70: train_loss=0.5968 val_loss=0.5963 val_acc=0.8163 val_auc=0.8721
ep 80: train_loss=0.5881 val_loss=0.5859 val_acc=0.8163 val_auc=0.8754
ep 90: train_loss=0.5798 val_loss=0.5747 val_acc=0.8163 val_auc=0.8754
ep 100: train_loss=0.5727 val_loss=0.5631 val_acc=0.8163
val_auc=0.8838
ep 110: train_loss=0.5566 val_loss=0.5526 val_acc=0.8163
val_auc=0.8838
ep 120: train_loss=0.5574 val_loss=0.5426 val_acc=0.8163
val_auc=0.8872
ep 130: train_loss=0.5317 val_loss=0.5320 val_acc=0.8163
val_auc=0.8872
ep 140: train_loss=0.5279 val_loss=0.5218 val_acc=0.8163
val_auc=0.8889
ep 150: train_loss=0.5223 val_loss=0.5111 val_acc=0.8163
val_auc=0.8956
```

```
ep 160: train_loss=0.4998 val_loss=0.5014 val_acc=0.8163  
val_auc=0.8973  
ep 170: train_loss=0.4862 val_loss=0.4926 val_acc=0.8163  
val_auc=0.8990  
ep 180: train_loss=0.4909 val_loss=0.4850 val_acc=0.8163  
val_auc=0.8990  
ep 190: train_loss=0.4863 val_loss=0.4768 val_acc=0.8367  
val_auc=0.8990  
ep 200: train_loss=0.4736 val_loss=0.4696 val_acc=0.8367  
val_auc=0.8990  
--> test_acc=0.8525, test_auc=0.9318

==== Running: optimizer=RMSprop, batch=32, lr=1e-05 ====
ep 01: train_loss=0.7029 val_loss=0.6942 val_acc=0.5510 val_auc=0.3805  
ep 10: train_loss=0.7007 val_loss=0.6927 val_acc=0.5510 val_auc=0.4125  
ep 20: train_loss=0.6982 val_loss=0.6917 val_acc=0.5510 val_auc=0.4276  
ep 30: train_loss=0.6990 val_loss=0.6910 val_acc=0.5510 val_auc=0.4461  
ep 40: train_loss=0.6976 val_loss=0.6902 val_acc=0.5510 val_auc=0.4630  
ep 50: train_loss=0.6974 val_loss=0.6894 val_acc=0.5510 val_auc=0.4764  
ep 60: train_loss=0.6963 val_loss=0.6887 val_acc=0.5510 val_auc=0.4815  
ep 70: train_loss=0.6890 val_loss=0.6880 val_acc=0.5510 val_auc=0.4882  
ep 80: train_loss=0.6928 val_loss=0.6873 val_acc=0.5510 val_auc=0.5051  
ep 90: train_loss=0.6900 val_loss=0.6866 val_acc=0.5510 val_auc=0.5219  
ep 100: train_loss=0.6892 val_loss=0.6859 val_acc=0.5510  
val_auc=0.5286  
ep 110: train_loss=0.6886 val_loss=0.6852 val_acc=0.5510  
val_auc=0.5421  
ep 120: train_loss=0.6926 val_loss=0.6845 val_acc=0.5510  
val_auc=0.5572  
ep 130: train_loss=0.6858 val_loss=0.6839 val_acc=0.5510  
val_auc=0.5690  
ep 140: train_loss=0.6887 val_loss=0.6833 val_acc=0.5510  
val_auc=0.5825  
ep 150: train_loss=0.6874 val_loss=0.6826 val_acc=0.5510  
val_auc=0.5976  
ep 160: train_loss=0.6921 val_loss=0.6820 val_acc=0.5510  
val_auc=0.6094  
ep 170: train_loss=0.6876 val_loss=0.6813 val_acc=0.5510  
val_auc=0.6145  
ep 180: train_loss=0.6834 val_loss=0.6806 val_acc=0.5510  
val_auc=0.6229  
ep 190: train_loss=0.6864 val_loss=0.6799 val_acc=0.5510  
val_auc=0.6347  
ep 200: train_loss=0.6882 val_loss=0.6792 val_acc=0.5510  
val_auc=0.6414  
--> test_acc=0.5410, test_auc=0.6981

==== Running: optimizer=RMSprop, batch=64, lr=0.1 ====
ep 01: train_loss=38.4028 val_loss=72.7942 val_acc=0.4490
```

```
val_auc=0.5000
ep 10: train_loss=0.6201 val_loss=0.3345 val_acc=0.8571 val_auc=0.8763
ep 20: train_loss=0.4820 val_loss=0.5385 val_acc=0.8571 val_auc=0.9024
ep 30: train_loss=0.4011 val_loss=0.3652 val_acc=0.8571 val_auc=0.8653
ep 40: train_loss=0.5810 val_loss=0.5092 val_acc=0.7551 val_auc=0.7643
ep 50: train_loss=0.4592 val_loss=0.3940 val_acc=0.8367 val_auc=0.8350
ep 60: train_loss=0.4678 val_loss=0.4010 val_acc=0.8367 val_auc=0.8182
ep 70: train_loss=0.3767 val_loss=0.3991 val_acc=0.8367 val_auc=0.8114
ep 80: train_loss=0.4276 val_loss=0.7511 val_acc=0.8367 val_auc=0.8721
ep 90: train_loss=0.5110 val_loss=0.4143 val_acc=0.8163 val_auc=0.7955
ep 100: train_loss=0.4201 val_loss=0.3966 val_acc=0.8367
val_auc=0.8333
ep 110: train_loss=0.3533 val_loss=0.3996 val_acc=0.8163
val_auc=0.8333
ep 120: train_loss=0.3595 val_loss=0.3749 val_acc=0.8367
val_auc=0.8502
ep 130: train_loss=0.4021 val_loss=0.3570 val_acc=0.8571
val_auc=0.8535
ep 140: train_loss=0.3525 val_loss=0.4188 val_acc=0.8163
val_auc=0.8316
ep 150: train_loss=0.3159 val_loss=0.3661 val_acc=0.8571
val_auc=0.8746
ep 160: train_loss=0.2807 val_loss=0.3980 val_acc=0.8367
val_auc=0.8502
ep 170: train_loss=0.3106 val_loss=0.3717 val_acc=0.8571
val_auc=0.8519
ep 180: train_loss=0.3930 val_loss=0.4327 val_acc=0.8163
val_auc=0.8114
ep 190: train_loss=0.4151 val_loss=0.5453 val_acc=0.8367
val_auc=0.8502
ep 200: train_loss=0.3699 val_loss=0.3877 val_acc=0.8367
val_auc=0.8350
--> test_acc=0.9016, test_auc=0.9329

==== Running: optimizer=RMSprop, batch=64, lr=0.01 ====
ep 01: train_loss=0.6469 val_loss=0.4074 val_acc=0.7551 val_auc=0.9007
ep 10: train_loss=0.3332 val_loss=0.4460 val_acc=0.7551 val_auc=0.8620
ep 20: train_loss=0.2320 val_loss=0.5996 val_acc=0.7755 val_auc=0.8586
ep 30: train_loss=0.1644 val_loss=0.7095 val_acc=0.7959 val_auc=0.8653
ep 40: train_loss=0.2008 val_loss=0.8513 val_acc=0.7755 val_auc=0.8552
ep 50: train_loss=0.1854 val_loss=0.8374 val_acc=0.7143 val_auc=0.8418
ep 60: train_loss=0.1867 val_loss=0.8388 val_acc=0.7551 val_auc=0.8620
ep 70: train_loss=0.1621 val_loss=0.9596 val_acc=0.7347 val_auc=0.8418
ep 80: train_loss=0.0928 val_loss=1.0022 val_acc=0.7143 val_auc=0.8485
ep 90: train_loss=0.1894 val_loss=1.1184 val_acc=0.7143 val_auc=0.8367
ep 100: train_loss=0.1549 val_loss=1.0242 val_acc=0.7551
val_auc=0.8552
ep 110: train_loss=0.1441 val_loss=0.9593 val_acc=0.7347
val_auc=0.8502
```

```
ep 120: train_loss=0.1190 val_loss=1.0157 val_acc=0.7755  
val_auc=0.8502  
ep 130: train_loss=0.1284 val_loss=1.1512 val_acc=0.7755  
val_auc=0.8535  
ep 140: train_loss=0.0877 val_loss=1.2002 val_acc=0.7755  
val_auc=0.8552  
ep 150: train_loss=0.1180 val_loss=1.1932 val_acc=0.7143  
val_auc=0.8451  
ep 160: train_loss=0.1338 val_loss=1.3647 val_acc=0.7143  
val_auc=0.8300  
ep 170: train_loss=0.0652 val_loss=1.4826 val_acc=0.7143  
val_auc=0.8283  
ep 180: train_loss=0.0601 val_loss=1.4136 val_acc=0.7143  
val_auc=0.8350  
ep 190: train_loss=0.0790 val_loss=1.5322 val_acc=0.7143  
val_auc=0.8502  
ep 200: train_loss=0.0445 val_loss=1.8122 val_acc=0.7347  
val_auc=0.8350  
--> test_acc=0.8197, test_auc=0.9399

==== Running: optimizer=RMSprop, batch=64, lr=0.001 ====
ep 01: train_loss=0.6953 val_loss=0.6518 val_acc=0.5918 val_auc=0.8939  
ep 10: train_loss=0.4353 val_loss=0.3910 val_acc=0.8776 val_auc=0.9175  
ep 20: train_loss=0.4118 val_loss=0.3539 val_acc=0.8571 val_auc=0.9192  
ep 30: train_loss=0.3835 val_loss=0.3373 val_acc=0.8776 val_auc=0.9259  
ep 40: train_loss=0.3741 val_loss=0.3377 val_acc=0.8571 val_auc=0.9259  
ep 50: train_loss=0.3633 val_loss=0.3292 val_acc=0.8571 val_auc=0.9259  
ep 60: train_loss=0.3552 val_loss=0.3317 val_acc=0.8571 val_auc=0.9293  
ep 70: train_loss=0.3747 val_loss=0.3483 val_acc=0.8367 val_auc=0.9293  
ep 80: train_loss=0.3362 val_loss=0.3564 val_acc=0.8367 val_auc=0.9141  
ep 90: train_loss=0.3700 val_loss=0.3471 val_acc=0.8367 val_auc=0.9209  
ep 100: train_loss=0.3592 val_loss=0.3377 val_acc=0.8571  
val_auc=0.9242  
ep 110: train_loss=0.3651 val_loss=0.3321 val_acc=0.8367  
val_auc=0.9242  
ep 120: train_loss=0.3319 val_loss=0.3230 val_acc=0.8367  
val_auc=0.9310  
ep 130: train_loss=0.3429 val_loss=0.3167 val_acc=0.8571  
val_auc=0.9343  
ep 140: train_loss=0.3276 val_loss=0.3204 val_acc=0.8571  
val_auc=0.9310  
ep 150: train_loss=0.3131 val_loss=0.3245 val_acc=0.8367  
val_auc=0.9276  
ep 160: train_loss=0.3279 val_loss=0.3231 val_acc=0.8367  
val_auc=0.9242  
ep 170: train_loss=0.3145 val_loss=0.3299 val_acc=0.8367  
val_auc=0.9226  
ep 180: train_loss=0.3373 val_loss=0.3218 val_acc=0.8367  
val_auc=0.9259
```

```
ep 190: train_loss=0.2935 val_loss=0.3250 val_acc=0.8367
val_auc=0.9259
ep 200: train_loss=0.3236 val_loss=0.3346 val_acc=0.8163
val_auc=0.9158
--> test_acc=0.8361, test_auc=0.9535

==== Running: optimizer=RMSprop, batch=64, lr=0.0001 ====
ep 01: train_loss=0.7103 val_loss=0.7004 val_acc=0.4490 val_auc=0.5791
ep 10: train_loss=0.6909 val_loss=0.6874 val_acc=0.4694 val_auc=0.7037
ep 20: train_loss=0.6876 val_loss=0.6810 val_acc=0.4898 val_auc=0.7239
ep 30: train_loss=0.6795 val_loss=0.6746 val_acc=0.5102 val_auc=0.7559
ep 40: train_loss=0.6742 val_loss=0.6698 val_acc=0.5306 val_auc=0.7677
ep 50: train_loss=0.6680 val_loss=0.6648 val_acc=0.5510 val_auc=0.7845
ep 60: train_loss=0.6673 val_loss=0.6608 val_acc=0.5918 val_auc=0.7946
ep 70: train_loss=0.6647 val_loss=0.6572 val_acc=0.6531 val_auc=0.8064
ep 80: train_loss=0.6545 val_loss=0.6533 val_acc=0.6735 val_auc=0.8148
ep 90: train_loss=0.6555 val_loss=0.6479 val_acc=0.6531 val_auc=0.8182
ep 100: train_loss=0.6473 val_loss=0.6427 val_acc=0.6531
val_auc=0.8199
ep 110: train_loss=0.6457 val_loss=0.6384 val_acc=0.6327
val_auc=0.8232
ep 120: train_loss=0.6462 val_loss=0.6338 val_acc=0.6531
val_auc=0.8266
ep 130: train_loss=0.6380 val_loss=0.6293 val_acc=0.6939
val_auc=0.8300
ep 140: train_loss=0.6292 val_loss=0.6240 val_acc=0.6735
val_auc=0.8350
ep 150: train_loss=0.6279 val_loss=0.6189 val_acc=0.6735
val_auc=0.8384
ep 160: train_loss=0.6158 val_loss=0.6138 val_acc=0.7143
val_auc=0.8418
ep 170: train_loss=0.6072 val_loss=0.6084 val_acc=0.7347
val_auc=0.8502
ep 180: train_loss=0.6058 val_loss=0.6035 val_acc=0.7143
val_auc=0.8552
ep 190: train_loss=0.6090 val_loss=0.5985 val_acc=0.7143
val_auc=0.8569
ep 200: train_loss=0.5986 val_loss=0.5924 val_acc=0.7143
val_auc=0.8620
--> test_acc=0.8525, test_auc=0.9383

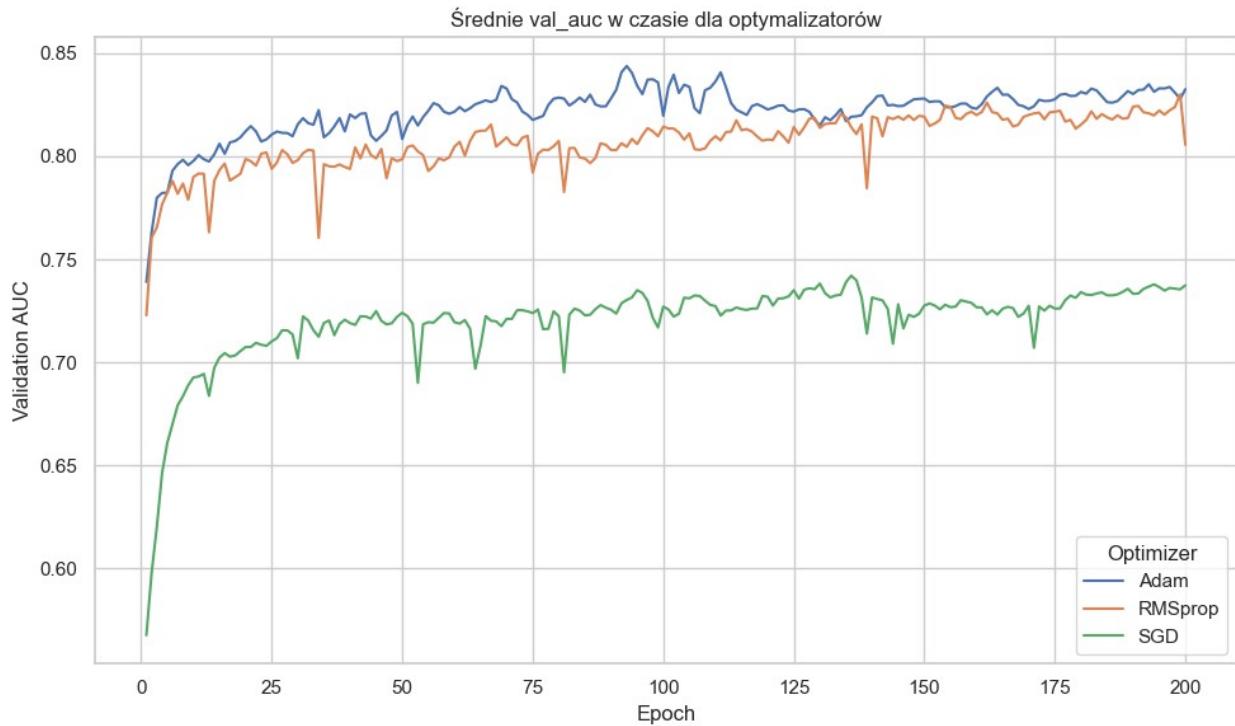
==== Running: optimizer=RMSprop, batch=64, lr=1e-05 ====
ep 01: train_loss=0.7155 val_loss=0.7101 val_acc=0.3878 val_auc=0.3822
ep 10: train_loss=0.7152 val_loss=0.7084 val_acc=0.4286 val_auc=0.3990
ep 20: train_loss=0.7097 val_loss=0.7074 val_acc=0.4286 val_auc=0.4125
ep 30: train_loss=0.7069 val_loss=0.7067 val_acc=0.4286 val_auc=0.4192
ep 40: train_loss=0.7062 val_loss=0.7060 val_acc=0.4694 val_auc=0.4226
ep 50: train_loss=0.7079 val_loss=0.7055 val_acc=0.4694 val_auc=0.4242
ep 60: train_loss=0.7115 val_loss=0.7049 val_acc=0.4694 val_auc=0.4276
```

```
ep 70: train_loss=0.7051 val_loss=0.7042 val_acc=0.4694 val_auc=0.4310
ep 80: train_loss=0.7028 val_loss=0.7037 val_acc=0.4694 val_auc=0.4360
ep 90: train_loss=0.7019 val_loss=0.7032 val_acc=0.4694 val_auc=0.4360
ep 100: train_loss=0.7065 val_loss=0.7027 val_acc=0.4694
val_auc=0.4411
ep 110: train_loss=0.7085 val_loss=0.7021 val_acc=0.4694
val_auc=0.4461
ep 120: train_loss=0.7021 val_loss=0.7016 val_acc=0.4694
val_auc=0.4512
ep 130: train_loss=0.7029 val_loss=0.7010 val_acc=0.4694
val_auc=0.4596
ep 140: train_loss=0.6982 val_loss=0.7005 val_acc=0.4694
val_auc=0.4613
ep 150: train_loss=0.7025 val_loss=0.7000 val_acc=0.4694
val_auc=0.4697
ep 160: train_loss=0.7017 val_loss=0.6994 val_acc=0.4898
val_auc=0.4764
ep 170: train_loss=0.6989 val_loss=0.6989 val_acc=0.5102
val_auc=0.4781
ep 180: train_loss=0.7018 val_loss=0.6983 val_acc=0.5102
val_auc=0.4848
ep 190: train_loss=0.7032 val_loss=0.6977 val_acc=0.5102
val_auc=0.4882
ep 200: train_loss=0.6988 val_loss=0.6972 val_acc=0.5306
val_auc=0.4916
--> test_acc=0.3934, test_auc=0.4848
```

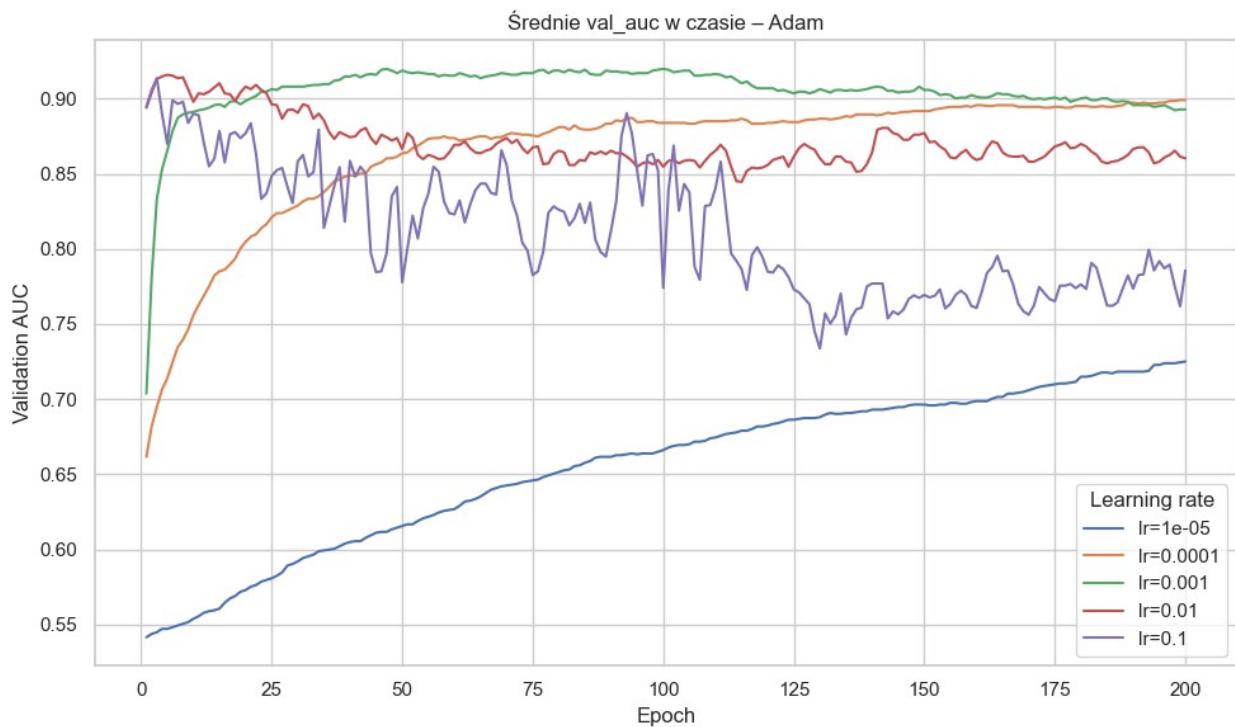
All experiments finished.

Summary saved to lab-4-results/results\_summary.csv

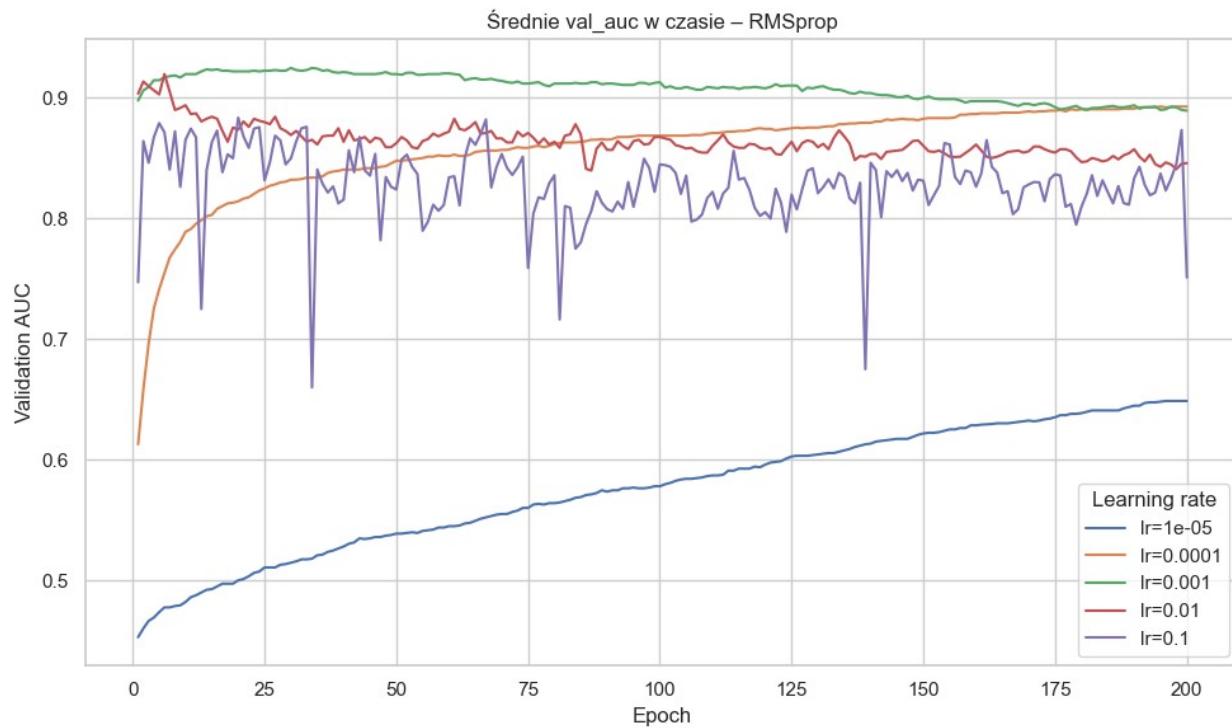
Full epoch logs saved to lab-4-results/results\_detailed.csv



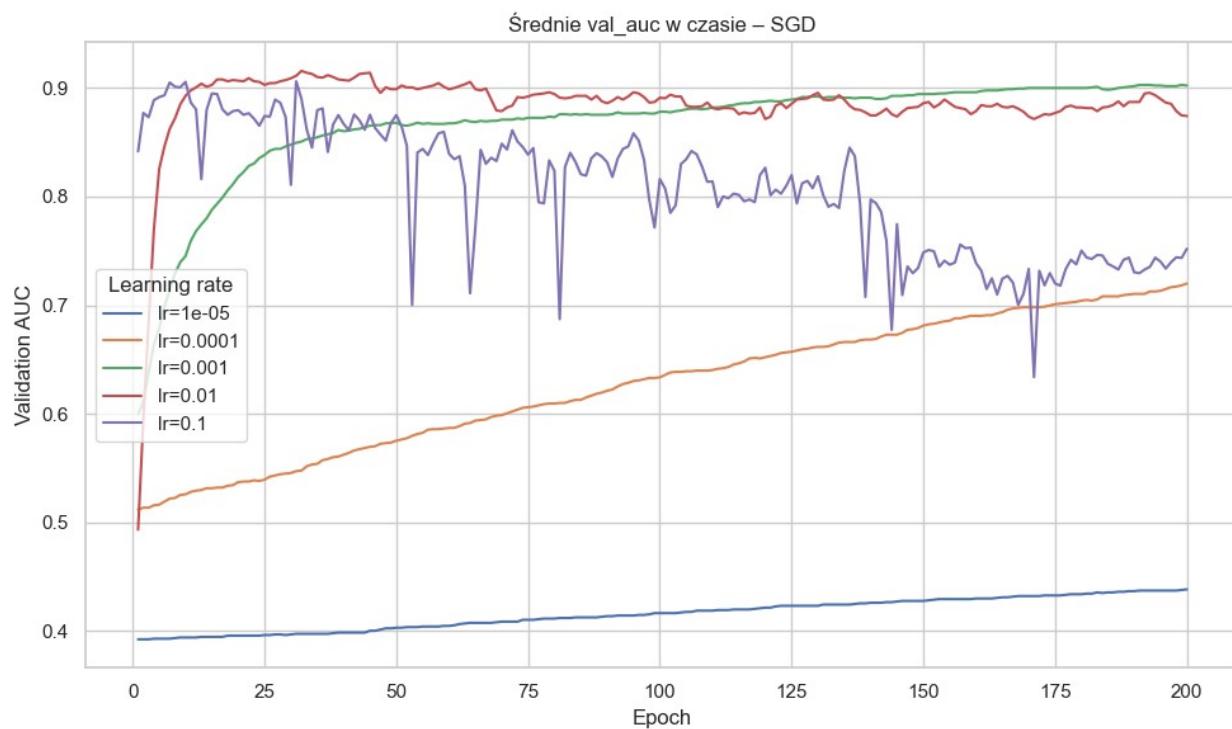
Saved plot: plots/val\_auc\_Adam.png



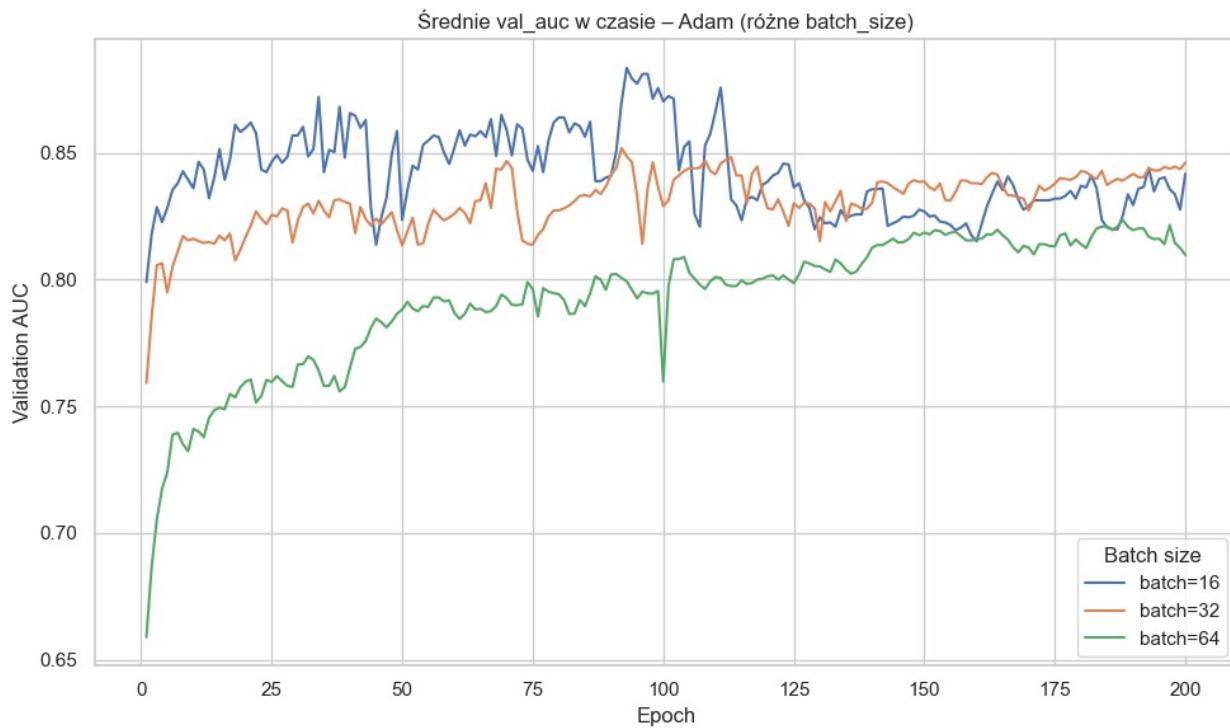
Saved plot: plots/val\_auc\_RMSprop.png



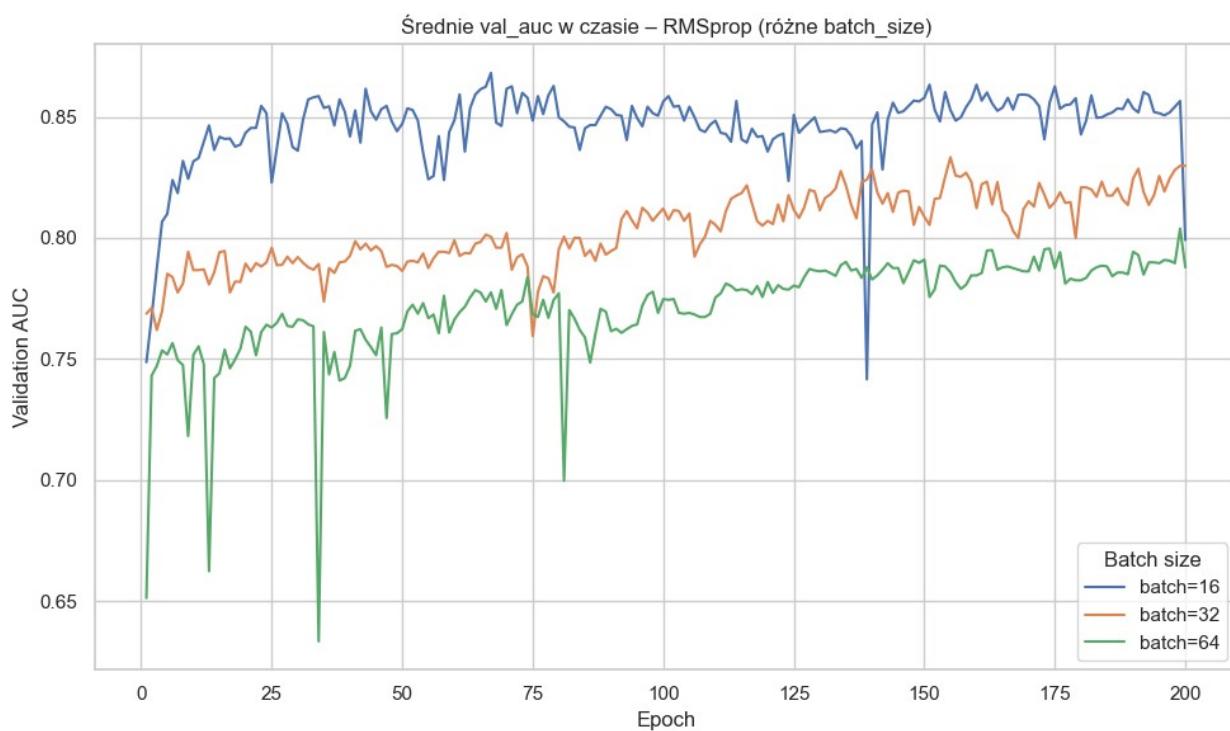
Saved plot: plots/val\_auc\_SGD.png



Saved plot: plots/val\_auc\_batch\_Adam.png



Saved plot: plots/val\_auc\_batch\_RMSprop.png



Saved plot: plots/val\_auc\_batch\_SGD.png

