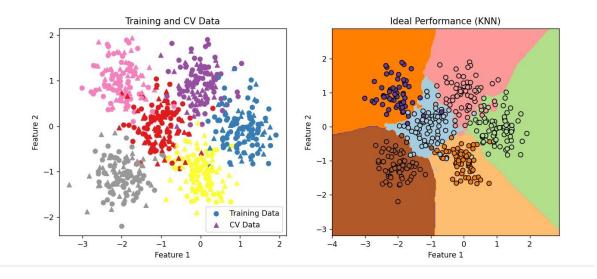
We compare, "training data", "cross-validation" and "ideal data"

```
from sklearn.datasets import make_blobs
from sklearn.model_selection import train_test_split
import numpy as np
import matplotlib.pyplot as plt
from sklearn.neighbors import KNeighborsClassifier
from sklearn.metrics import accuracy_score
classes = 6
centers = np.array([[-1, 0], [1, 0], [0, 1], [0, -1], [-2, 1], [-2, -1]])
X, y = make_blobs(n_samples=m, centers=centers, cluster_std=std,
random_state=2, n_features=2)
X_train, X_, y_train, y_ = train_test_split(X, y, test_size=0.50,
random_state=1)
X_cv, X_test, y_cv, y_test = train_test_split(X_, y_, test_size=0.20,
random_state=1)
def plot_decision_boundary(X, y, model, title):
    xx, yy = np.meshgrid(np.arange(x_min, x_max, h),
                            np.arange(y_min, y_max, h))
    Z = model.predict(np.c_[xx.ravel(), yy.ravel()])
    Z = Z.reshape(xx.shape)
   plt.contourf(xx, yy, Z, cmap=plt.cm.Paired)
    plt.scatter(X[:, 0], X[:, 1], c=y, edgecolors='k', cmap=plt.cm.Paired)
plt.figure(figsize=(12, 5))
plt.subplot(1, 2, 1)
plt.scatter(X_train[:, 0], X_train[:, 1], c=y_train, cmap=plt.cm.Set1, label='Training Data')
plt.scatter(X_cv[:, 0], X_cv[:, 1], c=y_cv, cmap=plt.cm.Set1, marker='^', label='CV Data')
plt.title('Training and CV Data')
plt.xlabel('Feature 1')
plt.ylabel('feature 2')
plt.legend()
plt.subplot(1, 2, 2)
model = KNeighborsClassifier(n_neighbors=5)
model.fit(X_train, y_train)
\verb|plot_decision_boundary(X_train, y_train, model, 'Ideal Performance (KNN)')| \\
cv_accuracy = accuracy_score(y_cv, model.predict(X_cv))
print(f'Cross-validation accuracy: {cv_accuracy:.2f}')
plt.tight_layout()
plt.show()
```

- Here is the training accuracy:

## - Here are the graphs:



We develop a complex model

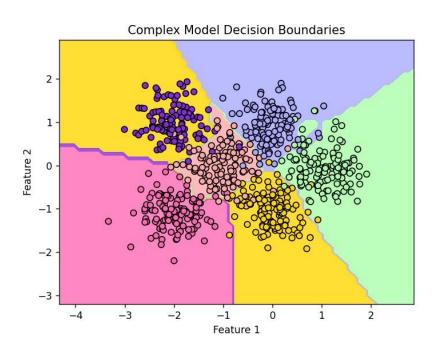
```
m - 888
std - 0.4
random_state=1)

X_cv, X_test, y_cv, y_test = train_test_split(X_tmp, y_tmp, test_size=0.20, random_state=1)
X_train = torch.FloatTensor(X_train)
y_train = torch.LongTensor(y_train)
X_cv = torch.FloatTensor(X_cv)
y_cv = torch.LongTensor(y_cv)
        ss ComplexModel(n.Andule):
    super(ComplexModel, self).__init__()
    self.fcl = nn.Linear(2, 110)
    self.fc2 = nn.Linear(128, 40)
    self.fc3 = nn.Linear(40, 6)
       def forward(self, x):
    x = torch.relu(self.fcl(x))
    x = torch.relu(self.fc2(x))
    x = self.fc3(x)
    return x
 criterion - mn.CrossEntropyLoss()
optimizer - optim.Adam(model_complex.parameters(), lr-6.881)
      cpocn in range(epocns):
    optimizer.zero_grad()
    outputs = model_complex(X_train)
    loss = criterion(outputs, y_train)
    loss.backward()
    optimizer.step()
      if (epoch:1) % 188 -- 8: print(f^*Epoch:\{epoch:1\}/\{epochs\}\}, \ Loss: \ (loss.item():.4f\}^*)
 with torch.no_grad():
    outputs_cv = model_complex(X_cv)
    _, predicted_cv = torch.max(outputs_cv, 1)
    cv_accuracy = accuracy_score(y_cv, predicted_cv)
 grid_tensor = torch.FloatTensor(np.c_[xx.ravel(), yy.ravel()])
Z = model(grid_tensor)
        _, Z = torch.max(Z, 1)
Z = Z.reshape(xx.shape)
        cmap = ListedColormap(['#FFAAAA', '#AAFFAA', '#AAAAFF', '#FF0788', '#F828E2', '#FF6984'])
plt.contourf(xx, yy, Z, cmap-cmap, alpha-8.8)
        plt.scatter(X[:,\;0],\;X[:,\;1],\;c-y,\;cmap-cmap,\;edgecolors-'k')
        plt.title(title)
        plt.xlabel('Feature 1')
plt.ylabel('Feature 2')
plt.show()
X_combined - torch.FloatTensor(np.vstack((X_train.numpy(), X_cv.numpy())))
y_combined - torch.longTensor(np.concatenate((y_train.numpy(), y_cv.numpy())))
plot_decision_boundary(model_complex, X_combined, y_combined, 'Complex Model Decision Boundaries')
```

- Here is the cross-validation accuracy

```
def plot_decision_boundary(model, X, y, title):
           x_{min}, x_{max} = X[:, 0].min() - 1, X[:, 0].max() + 1
           y_{min}, y_{max} = X[:, 1].min() - 1, <math>X[:, 1].max() + 1
           xx, yy = np.meshgrid(np.arange(x min, x max, 0.1),
                                 np.arange(y min, y max, 0.1))
PROBLEMS
           OUTPUT
                   DEBUG CONSOLE TERMINAL
PS C:\Users\ardah\UCM> python -u "c:\Users\ardah\UCM\p6 b\p6 b.py"
Cross-validation accuracy: 0.94
PS C:\Users\ardah\UCM> python -u "c:\Users\ardah\UCM\p6 b\p6 b2.py"
Epoch [100/1000], Loss: 0.3264
Epoch [200/1000], Loss: 0.2036
Epoch [300/1000], Loss: 0.1867
Epoch [400/1000], Loss: 0.1777
Epoch [500/1000], Loss: 0.1699
Epoch [600/1000], Loss: 0.1624
Epoch [700/1000], Loss: 0.1541
Epoch [800/1000], Loss: 0.1453
Epoch [900/1000], Loss: 0.1362
Epoch [1000/1000], Loss: 0.1266
Complex Model - Cross-validation accuracy: 90.62%
PS C:\Users\ardah\UCM>
```

Here is the plotted graph:



We will now try to develop a simple model

```
classes -
X_train, X_tmp, y_train, y_tmp - train_test_split(X, y, test_size-8.58,
X_cv, X_test, y_cv, y_test = train_test_split(X_tmp, y_tmp, test_size=0.20, random_state=1)

# Convert number arrays to PyTorch tensors
X_train = torch.FloatTensor(X_train)
y_train = torch.LongTensor(y_train)
X_cv = torch.FloatTensor(X_cv)
y_cv = torch.LongTensor(y_cv)
  lass SimpleModel(nn Module):

dof __init_(self):

super(SimpleModel, self).__init__()

self.fcl = nn.Lincar(2, 6)
       self.fc2 - mn.linear(2, 6)
self.fc2 - mn.linear(6, 6)
def forward(self, x):
x - torch.relu(self.fc1(x))
x - self.fc2(x)
return x
  model simple - SimpleModel()
criterion - nn.CrossEntropyLoss()
optimizer - optim.Adam(model_simple.parameters(), 1r-0.01)
      optimizer.zero_grad()
outputs = model_simple(X_train)
       optimizer.step()
        if (epoch+1) % 188 -- 8:
    print(f'Epoch [(epoch+1)/(epochs)], Loss: {loss.item():.4f}*)
 # Evaluate the rimple model on training and cross validation data with torch.no_grad():

outputs_train = model_simple(X_train)

_, predicted_train = torch.max(outputs_train, 1)

train_accuracy = accuracy_score(y_train, predicted_train)

outputs_cv = model_simple(X_cv)

_, predicted_cv = torch.max(outputs_cv, 1)

cv_accuracy = accuracy_score(y_cv, predicted_cv)

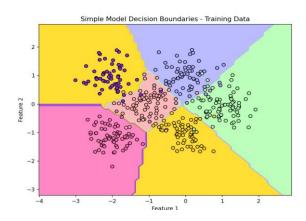
print(f'Simple Model = Training accuracy: (train_accuracy:.2%)')

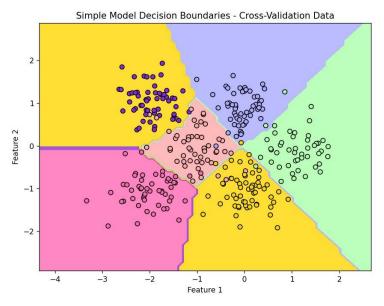
print(f'Simple Model = Cross-validation accuracy: (cv_accuracy:.2%)')

# America to also decision boundaries of the model an appetite data
  def plot_decision_boundary(model, X, y, title):
    x_min, x_max - X[:, 8].min() - 1, X[:, 8].max() + 1
    y_min, y_max - X[:, 1].min() - 1, X[:, 1].max() + 1
    xx, yy - np.meshgrid(np.arange(x_min, x_max, 8.1),
                                                np.arange(y_min, y_max, E.1))
        grid_tensor = torch.FloatTensor(np.c_[xx.ravel(), yy.ravel()])
        Z = model(grid_tensor)
       _, Z - torch.max(Z, 1)
Z - Z.reshape(xx.shape)
       cmap = ListedColormap(['#FFAAAA', '#AAFFAA', '#AAAAFF', '#FF0780', '#BAZ8E2', '#FF6984'])
plt.contourf(xx, yy, Z, cmap-cmap, alpha-0.8)
       \label{eq:plt.scatter} $$ plt.scatter(X[:, B], X[:, 1], c-y, cmap-cmap, edgecolors-'k') $$ plt.titlo(title) $$
       plt.xlabel('Feature 1')
plt.ylabel('Feature 2')
   odel_simple.eval()
plt.figure(figsize-($, 0))
plot_decision_boundary(model_simple, X_train.numpy(), y_train.numpy(), 'Simple Model Decision Boundaries - Training Data')
 plt.figure(figsize-(8, 6))
plot_decision_boundary(model_simple, X_cv.numpy(), y_cv.numpy(), 'Simple Model Decision Boundaries - Cross-Validation Data')
```

- Here is the output ad accuracy rate

```
print(f'Simple Model - Training accuracy: {train_accuracy:.2%}'
PROBLEMS
          OUTPUT
                   DEBUG CONSOLE
                                   TERMINAL
                                              PORTS
PS C:\Users\ardah\UCM> python -u "c:\Users\ardah\UCM\p6_b\p6_b3_1.py"
Epoch [100/1000], Loss: 0.4075
Epoch [200/1000], Loss: 0.2248
Epoch [300/1000], Loss: 0.2019
Epoch [400/1000], Loss: 0.1962
Epoch [500/1000], Loss: 0.1936
Epoch [600/1000], Loss: 0.1919
Epoch [700/1000], Loss: 0.1905
Epoch [800/1000], Loss: 0.1888
Epoch [900/1000], Loss: 0.1876
Epoch [1000/1000], Loss: 0.1866
Simple Model - Training accuracy: 92.50%
Simple Model - Cross-validation accuracy: 94.06%
PS C:\Users\ardah\UCM>
```





We will regularize the complex model

```
m - 888
std - 8.4
 X_train, X_tmp, y_train, y_tmp = train_test_split(X, y, test_size=0.50, random_state=1)
X_train = torch.FloatTensor(X_train)
y_train = torch.LongTensor(y_train)
X_cv = torch.FloatTensor(X_cv)
y_cv = torch.LongTensor(y_cv)
 class ComplexModelRegularized(nn.Module):
    def __init__(self):
         r__init__(self):
super(CosplexModelRegularized, self).__init__()
self.fcl = nn.Linear(2, 128)
self.fc2 = nn.Linear(128, 48)
self.fc3 = nn.Linear(48, 6)
         x = torch.relu(self.fc1(x))
x = torch.relu(self.fc2(x))
x = torch.relu(self.fc2(x))
return x
 model complex regularized - ComplexModelRegularized()
criterion - mn.CrossEntropyLoss()
optimizer - optim.Adam(model_complex_regularized.parameters(), lr-8.881, weight_decay-8.1)
 epochs - 1888
for epoch in range(epochs):
    outputs - model_complex_regularized(X_train)
loss - criterion(outputs, y_train)
    optimizer.step()
if (epoch+1) % 188 -- 8:
    print(f'Epoch [{epoch+1}/{epochs}], Loss: {loss.item():.4f}*)
 int torch.no_grad():
    outputs_cv = model_complex_regularized(X_cv)
    _, predicted_cv = torch.max(outputs_cv, 1)
    cv_accuracy = accuracy_score(y_cv, predicted_cv)
rimt(f'Complex Model with Regularization = Cross_validation accuracy: {cv_accuracy:.2%}')
  grid_tensor - torch.FloatTensor(np.c_[xx.ravel(), yy.ravel()])
Z - model(grid_tensor)
    _, Z - torch.max(Z, I)
Z - Z.reshape(xx.shape)
    cmap = ListedColormap(['#FFAAAA', '#AAFFAA', '#AAAAFF', '#FF0788', '#8A28E2', '#FF6984'])
plt.contourf(xx, yy, Z, cmap-cmap, alpha-8.8)
    plt.scatter(X[:, 0], X[:, 1], c-y, cmap-cmap, edgecolors-'k') plt.title(title)
    plt.xlabel('Feature 1')
plt.ylabel('Feature 2')
plt.figuro(figsize-(8, 6))
plot_decision_boundary∰model_complex_regularized, X_train.numpy(), y_train.numpy(), "Complex Model Decision Boundaries with Regularization - Training Data'∭
plt.figure(figize-(8, 6))
plot_decision_boundary(model_complex_regularized, X_cv.numpy(), y_cv.numpy(), 'Complex Model Decision Boundaries with Regularization - Cross-Validation Gata')
```

- Accuracy rate seems way lower than it should be... I wasnt able to solve this issue

```
Regularization Value: 0.3, Cross-validation Error: 0.8531, Training Error: 0.812: PS C:\Users\ardah\UCM\ppthon -u "c:\Users\ardah\UCM\p6_b\p6_b3_2.py"

Epoch [100/1000], Loss: 0.7264

Epoch [200/1000], Loss: 0.6461

Epoch [300/1000], Loss: 0.6676

Epoch [400/1000], Loss: 0.7018

Epoch [500/1000], Loss: 0.7309

Epoch [600/1000], Loss: 0.7505

Epoch [700/1000], Loss: 0.7655

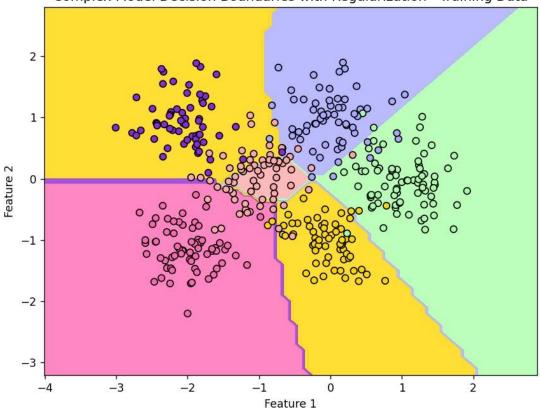
Epoch [800/1000], Loss: 0.7677

Epoch [1000/1000], Loss: 0.7687

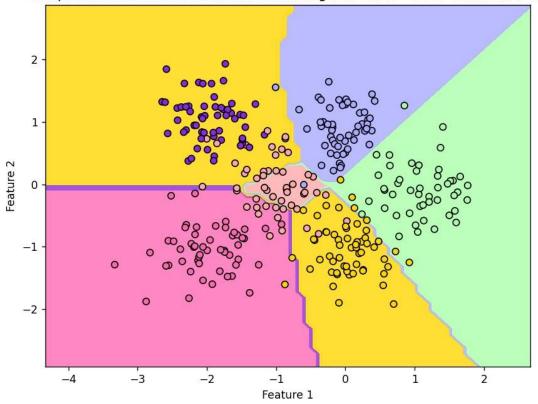
Complex Model with Regularization - Cross-validation accuracy: 85.62%

PS C:\Users\ardah\UCM\ []
```





Complex Model Decision Boundaries with Regularization - Cross-Validation Data



We need to find the most optimal lambda by trying different lambdas

```
X_train, X_tmp, y_train, y_tmp - train_test_split(X, y, test_size-0.50, random_state-1)
X_train = torch.FloatTensor(X_train)
y_train = torch.LongTensor(y_train)
       ss ComplexModel(nm.Module):
def __init__(self):
    super(ComplexModel, self).__init__()
    self.fc1 = nn.Linear(2, 120)
    self.fc2 = nn.Linear(120, 48)
    self.fc3 = nn.Linear(40, 6)
def forward(self, x):
    x = torch.relu(self.fc1(x))
    x = torch.relu(self.fc2(x))
    v = self.fc2(x)
               x = self.fc3(x)
 erunction to their and evaluate the model with specified regularization value

dof train_and_evaluate_model(regularization_value):

# Politicity to recommend
       model = ComplexModel[]
criterion = nn.CrossEntropyLoss()
optimizer = optim.Adam(model.parameters(), lr=0.801, weight_decay-regularization_value) = tlr=qualization_value
        epochs - 1000
       cv_errors - []
for epoch in range(epochs):
              optimizer.zero_grad()
outputs - model(X_train)
loss - criterion(outputs, y_train)
               optimizer.step()
                      h torch.no_grad():
    outputs_cv = model(X_cv)
    _, predicted_cv = torch.max(outputs_cv, 1)
    cv_accuracy = accuracy_score(y_cv, predicted_cv)
    cv_error = 1.8 = cv_accuracy = irror rate
    cv_errors.append(cv_error)
        final_cv_error - cv_errors[-1]
       with torch.no_grad():
    outputs_train = model(X_train)
    _, predicted_train = torch.max(outputs_train, 1)
    train_accuracy = accuracy_score(y_train, predicted_train)
    train_error = 1.8 = train_accuracy = line rate
return final_cv_error, train_error
 regularization_values = [0.0, 0.001, 0.01, 0.05, 0.1, 0.2, 0.3]
for reg_value in regularization_values:

cv_error, train_error = train_and_evaluate_model(reg_value)

cv_errors.append(cv_error)

train_errors.append(train_error)
        print(f'Regularization Value: (reg_value), Cross-validation Error: (cv_error: 4f), Training Error: (train_error: 4f)')
 plt.figure(figsize-(8, 6))
plt.flgure(figsize_(S, E))

plt.plot(regularization_values, train_errors, marker-'o', linestyle-'-', color-'b', label-'Training Error')

plt.plot(regularization_values, cv_errors, marker-'s', linestyle-'--', color-'r', label-'Cross-validation Error')

plt.title('Error vs Regularization Value')

plt.xlabel('Regularization Value (\lambda)')
 plt.grid(True)
```

- For me "0.1" seemed to be the most optimal, which contradicts with the actual results

```
PS C:\Users\ardah\UCM> python -u "c:\Users\ardah\UCM\p6_b\p6_b5.py"

Regularization Value: 0.0, Cross-validation Error: 0.0813, Training Error: 0.0550

Regularization Value: 0.001, Cross-validation Error: 0.0813, Training Error: 0.0575

Regularization Value: 0.01, Cross-validation Error: 0.0656, Training Error: 0.0750

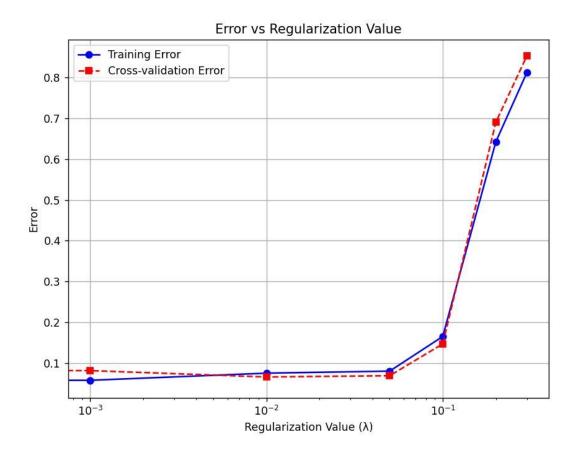
Regularization Value: 0.05, Cross-validation Error: 0.0687, Training Error: 0.0800

Regularization Value: 0.1, Cross-validation Error: 0.1469, Training Error: 0.1650

Regularization Value: 0.2, Cross-validation Error: 0.6906, Training Error: 0.6425

Regularization Value: 0.3, Cross-validation Error: 0.8531, Training Error: 0.8125
```

- This is the graph I obtained:



- We will use the X test to test the models...

```
import torch
import torch.nn as nn
import torch.optim as optim
from sklearn.datasets import make_blobs
from sklearn.model_selection import train_test_split
import numpy as np
import matplotlib.pyplot as plt
from sklearn.metrics import accuracy_score from matplotlib.colors import ListedColormap
from sklearn.neighbors import KNeighborsClassifier
m = 800
std = 0.4
X, y = make_blobs(n_samples=m, centers=centers, cluster_std=std,
                      random_state=2, n_features=2)
X_train, X_tmp, y_train, y_tmp = train_test_split(X, y, test_size=0.50,
                                                                 random state=1)
X_cv, X_test, y_cv, y_test = train_test_split(X_tmp, y_tmp, test_size=0.20, random_state=1)
X_train = torch.FloatTensor(X_train)
y_train = torch.LongTensor(y_train)
X_cv = torch.FloatTensor(X_cv)
y_cv = torch.LongTensor(y_cv)
X_test = torch.FloatTensor(X_test)
y_test = torch.LongTensor(y_test)
class SimpleModel(nn.Module):
     def __init__(self):
         self.fc1 = nn.Linear(2, 6)
self.fc2 = nn.Linear(6, 6)
     def forward(self, x):
    x = torch.relu(self.fc1(x))
     def __init__(self):
          self.fc1 = nn.Linear(2, 120)
self.fc2 = nn.Linear(120, 40)
     def forward(self, x):
    x = torch.relu(self.fc1(x))
    x = torch.relu(self.fc2(x))
    x = self.fc3(x)
def train_model(model, X_train, y_train):
     criterion = nn.CrossEntropyLoss()
optimizer = optim.Adam(model.parameters(), 1r=0.001)
     for epoch in range(epochs):
         optimizer.zero_grad()
          outputs = model(X_train)
          loss = criterion(outputs, y_train)
          optimizer.step()
model_simple = SimpleModel()
train model/model simple Y train v train)
```

```
optimizer.step()
model_simple = SimpleModel()
train_model(model_simple, X_train, y_train)
model_complex = ComplexModel()
train_model(model_complex, X_train, y_train)
def evaluate_model(model, X, y, dataset_name):
         outputs = model(X)
          _, predicted = torch.max(outputs, 1)
         accuracy = accuracy_score(y.numpy(), predicted.numpy())
print(f"{dataset_name} Accuracy: {accuracy:.2%}")
         return accuracy
test_accuracy_simple = evaluate_model(model_simple, X_test, y_test, "Simple Model (Test)")
test_accuracy_complex = evaluate_model(model_complex, X_test, y_test, "Complex Model (Test)")
model_ideal = KNeighborsClassifier(n_neighbors=5)
model_ideal.fit(X_train.numpy(), y_train.numpy())
test_accuracy_ideal = accuracy_score(y_test, model_ideal.predict(X_test.numpy()))
print(f"Ideal Model (Test) Accuracy: {test_accuracy_ideal:.2%}")
def plot_decision_boundary(model, X, y, title):
     x_min, x_max = X[:, 0].min() - 1, X[:, 0].max() + 1
y_min, y_max = X[:, 1].min() - 1, X[:, 1].max() + 1
xx, yy = np.meshgrid(np.arange(x_min, x_max, 0.1),
                             np.arange(y_min, y_max, 0.1))
     grid_tensor = torch.FloatTensor(np.c_[xx.ravel(), yy.ravel()])
     Z = model(grid_tensor)
     Z = Z.reshape(xx.shape)
     cmap = ListedColormap(['#FFAAAA', '#AAFFAA', '#AAAAFF', '#FFD700', '#8A2BE2', '#FF69B4'])
     plt.contourf(xx, yy, Z, cmap=cmap, alpha=0.8)
     plt.scatter(X[:,\;\theta],\;X[:,\;1],\;c=y,\;cmap=cmap,\;edgecolors='k')
     plt.title(title)
     plt.xlabel('Feature 1')
     plt.show()
plt.figure(figsize=(8, 6))
plot_decision_boundary(model_simple, X_test, y_test, 'Simple Model Decision Boundaries (Test)')
plt.figure(figsize=(8, 6))
plot_decision_boundary(model_complex, X_test, y_test, 'Complex Model Decision Boundaries (Test)')
plt.figure(figsize=(8, 6))
cmap = ListedColormap(['mFFAAAA', 'mAAFFAA', 'mAAAAFF', 'mFFD700', 'm8A28E2', 'mFF6984'])
x_min, x_max = X_test[:, 0].min() - 1, X_test[:, 0].max() + 1
y_min, y_max = X_test[:, 1].min() - 1, X_test[:, 1].max() + 1
xx, yy = np.meshgrid(np.arange(x_min, x_max, \theta.1),
                         np.arange(y_min, y_max, 0.1))
Z = model_ideal.predict(np.c_[xx.ravel(), yy.ravel()])
Z = Z.reshape(xx.shape)
\label{eq:plt.contourf} $$pt.contourf(xx, yy, Z, cmap=cmap, alpha=0.8)$ plt.scatter(X_test[:, 0], X_test[:, 1], c=y_test, cmap=cmap, edgecolors='k') plt.title('Ideal Model Decision Boundaries (KNN) (Test)')
plt.xlabel('Feature 1')
```

These are the accuracy rates...

```
PS C:\Users\ardah\UCM> python -u "c:\Users\ardah\UCM\p6_b\p6_b6.py"
Simple Model (Test) Accuracy: 82.50%
Complex Model (Test) Accuracy: 86.25%
Ideal Model (Test) Accuracy: 86.25%
```

