In [1]:
```
import pandas as pd
import numpy as np
from sklearn.model_selection import train_test_split, RepeatedStratifiedKFold, cross_
from sklearn.discriminant_analysis import LinearDiscriminantAnalysis
from sklearn import datasets
import matplotlib.pyplot as plt
```

In [2]:
```
df = pd.read_csv('Iris.csv')
df.head()
```

Out[2]:

	Id	SepalLengthCm	SepalWidthCm	PetalLengthCm	PetalWidthCm	Species
0	1	5.1	3.5	1.4	0.2	Iris-setosa
1	2	4.9	3.0	1.4	0.2	Iris-setosa
2	3	4.7	3.2	1.3	0.2	Iris-setosa
3	4	4.6	3.1	1.5	0.2	Iris-setosa
4	5	5.0	3.6	1.4	0.2	Iris-setosa

In [3]:
```
X = df[['SepalLengthCm', 'SepalWidthCm', 'PetalLengthCm', 'PetalWidthCm']]
y = df['Species']

model = LinearDiscriminantAnalysis()
model.fit(X, y)
```

Out[3]:
```
LinearDiscriminantAnalysis()
```

In [4]:
```
# evaluating the model
cv = RepeatedStratifiedKFold(n_splits=10, n_repeats=3, random_state=1)
scores = cross_val_score(model, X, y, scoring='accuracy', cv=cv, n_jobs=-1)
print(np.mean(scores))
```

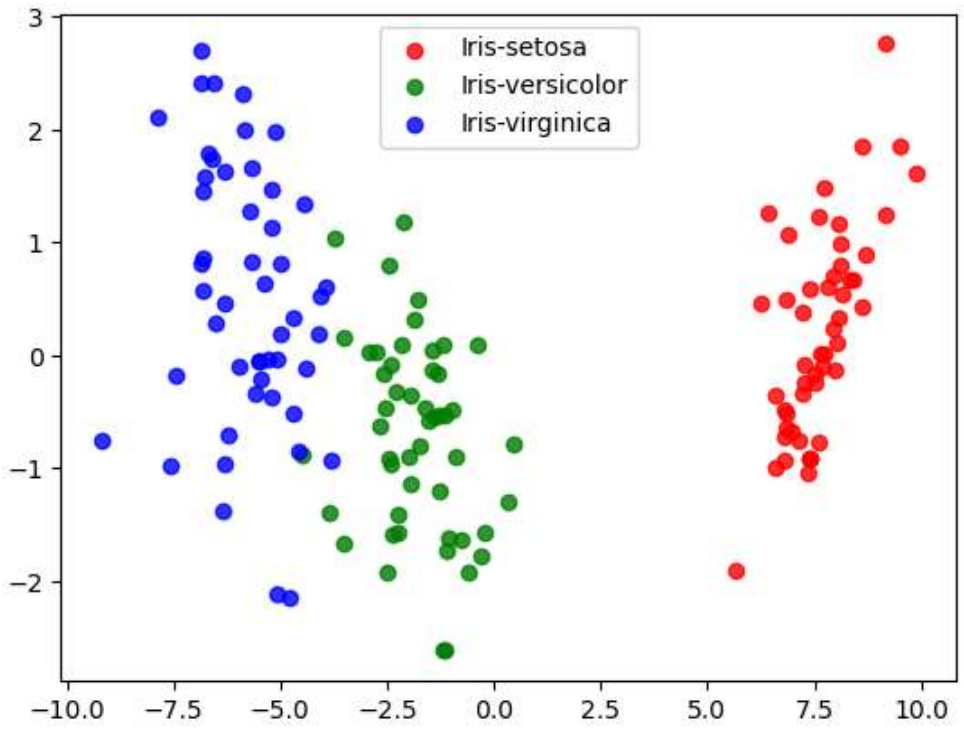
```
0.9800000000000001
```

In [5]:
```
data_plot = model.transform(X)

target_names = df['Species'].unique()

plt.figure()
colors = ['red', 'green', 'blue']
for color, i, target_name in zip(colors, range(len(target_names)), target_names):
    plt.scatter(data_plot[y == target_name, 0], data_plot[y == target_name, 1],
                alpha=.8, color=color, label=target_name)

plt.legend(loc='best', shadow=False, scatterpoints=1)
plt.show()
```



In [ ]:

In [ ]:

In [ ]: