Importing Libraries

```
import pandas as pd
import numpy as np
from sklearn.datasets import load_iris
from sklearn.model_selection import train_test_split, cross_val_score, StratifiedKFold
from sklearn.linear_model import LogisticRegression
from sklearn.svm import SVC
from mlxtend.evaluate import paired_ttest_5x2cv
from sklearn.metrics import accuracy_score
```

Load IRIS dataset

```
data = load_iris()
X = pd.DataFrame(data.data, columns=data.feature_names)
y = pd.Series(data.target, name='Target')
```

Split Data, Train Model (Logistic Regression and Support Vector Machine) and Compare: accuracy

```
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.3, random_state=42, stratify=y)

# -------------------- Model 1: Logistic Regression --------------------
logreg_model = LogisticRegression(max_iter=500, random_state=42)
logreg_model.fit(X_train, y_train)
logreg_predictions = logreg_model.predict(X_test)
logreg_accuracy = accuracy_score(y_test, logreg_predictions)

# -------------------- Model 2: Support Vector Machine (SVM) --------------------
svm_model = SVC(random_state=42)
svm_model.fit(X_train, y_train)
svm_predictions = svm_model.predict(X_test)
svm_accuracy = accuracy_score(y_test, svm_predictions)

print("Model Accuracies")
print(f"Logistic Regression Accuracy: {logreg_accuracy:.4f}")
print(f"Support Vector Machine Accuracy: {svm_accuracy:.4f}")
```

```
Model Accuracies
Logistic Regression Accuracy: 0.9333
Support Vector Machine Accuracy: 0.9556
```

Paired T-Test

```
t_stat, p_value = paired_ttest_5x2cv(estimator1=logreg_model, estimator2=svm_model, X=X.values, y=y.values, random_seed=42)

print("Paired T-Test (5x2 Cross-Validation)")
print("\nNull Hypothesis (H0): There is no significant difference between the performances of Logistic Regression and SVM.")
print("Alternative Hypothesis (H1): There is a significant difference between the performances of Logistic Regression and SVM.")
print(f"T-Statistic: {t_stat:.4f}, P-value: {p_value:.4f}")

# Decision based on the p-value
if p_value <= 0.05:
    print("\nH₀ rejected: Logistic Regression and SVM have significantly different performances.")
    if logreg_accuracy > svm_accuracy:
        print("Recommendation: Use Logistic Regression as it performs better.")
    else:
        print("Recommendation: Use SVM as it performs better.")
else:
    print("\nH₀ accepted: No significant difference between Logistic Regression and SVM performances.")
    print("Recommendation: Both models are equally good.")
    if logreg_accuracy > svm_accuracy:
        recommended_model = "Logistic Regression"
    else:
        recommended_model = "SVM"
    print(f"Based on accuracy choose {recommended_model}")
```

```
Paired T-Test (5x2 Cross-Validation)

Null Hypothesis (H0): There is no significant difference between the performances of Logistic Regression and SVM.
Alternative Hypothesis (H1): There is a significant difference between the performances of Logistic Regression and SVM.
T-Statistic: 0.9129, P-value: 0.4032

H₀ accepted: No significant difference between Logistic Regression and SVM performances.
Recommendation: Both models are equally good.
Based on accuracy choose SVM
```