```
import pandas as pd
import numpy as np
from scipy.stats import chi2_contingency, norm
from mlxtend.evaluate import paired_ttest_5x2cv
from sklearn.model_selection import train_test_split, cross_val_score, StratifiedKFold
from sklearn.tree import DecisionTreeClassifier
from sklearn.ensemble import RandomForestClassifier
from sklearn.metrics import accuracy_score
from sklearn.datasets import load_breast_cancer
```

Load Dataset (breast cancer dataset)

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

X['Target'] = y

X.head(1)
```

	mean radius	mean texture	mean perimeter	mean area	mean smoothness	mean compactness	mean concavity	mean concave points	mean symmetry	mean fractal dimension	...	worst texture	worst perimeter	worst area	worst smoothness
0	17.99	10.38	122.8	1001.0	0.1184	0.2776	0.3001	0.1471	0.2419	0.07871	...	17.33	184.6	2019.0	0.1622

1 rows × 31 columns

Chi-Squared Test (Testing the relationship between 'mean radius' and 'Target')

```
X['mean radius (binned)'] = pd.cut(X['mean radius'], bins=5, labels=False)
contingency_table = pd.crosstab(X['mean radius (binned)'], X['Target'])
chi2_stat, p_value_chi2, dof, _ = chi2_contingency(contingency_table)

print("Chi-Squared Test")
print("\nNull Hypothesis (H0): The feature 'mean radius (binned)' and the target variable are independent.")
print("Alternative Hypothesis (H1): The feature 'mean radius (binned)' and the target variable are associated.")
print(f"\nChi2 Statistic: {chi2_stat}, P-value: {p_value_chi2}, Degrees of Freedom: {dof}")
if p_value_chi2 <= 0.05:
    print("\nH₀ rejected: The feature 'mean radius (binned)' is significantly associated with the target.")
else:
    print("\nH₀ accepted: No significant association between 'mean radius (binned)' and the target.")
    X = X.drop(columns=['mean radius (binned)'])
```

```
Chi-Squared Test

Null Hypothesis (H0): The feature 'mean radius (binned)' and the target variable are independent.
Alternative Hypothesis (H1): The feature 'mean radius (binned)' and the target variable are associated.

Chi2 Statistic: 318.75152913941974, P-value: 9.752672443260828e-68, Degrees of Freedom: 4

H₀ rejected: The feature 'mean radius (binned)' is significantly associated with the target.
```

Split Data, Train Model (Decision Tree and Random Forest) and Compare: accuracy

```
X = X.drop(columns=['Target'])

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

dt_model = DecisionTreeClassifier(random_state=42)
dt_model.fit(X_train, y_train)
dt_predictions = dt_model.predict(X_test)

rf_model = RandomForestClassifier(random_state=42)
rf_model.fit(X_train, y_train)
rf_predictions = rf_model.predict(X_test)

dt_accuracy = accuracy_score(y_test, dt_predictions)
rf_accuracy = accuracy_score(y_test, rf_predictions)

print("Decision Tree Accuracy:", dt_accuracy)
print("Random Forest Accuracy:", rf_accuracy)
```

```
Decision Tree Accuracy: 0.9298245614035088
Random Forest Accuracy: 0.9766081871345029
```

Paired T-Test

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

print("Paired T-Test (5x2 Cross-Validation)")
print("\nNull Hypothesis (H0): The performance of the Decision Tree and Random Forest models is the same.")
print("Alternative Hypothesis (H1): The performance of the Decision Tree and Random Forest models is different.")
print(f"\nT-Statistic: {t_stat}, P-value: {p_value}")
if p_value <= 0.05:
    print("\nH₀ rejected: The Random Forest and Decision Tree models have significantly different performances.")
else:
    print("\nH₀ accepted: No significant difference between Random Forest and Decision Tree performances.")
```

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

Null Hypothesis (H0): The performance of the Decision Tree and Random Forest models is the same.
Alternative Hypothesis (H1): The performance of the Decision Tree and Random Forest models is different.

T-Statistic: -2.690573848818753, P-value: 0.043268965333263756

H₀ rejected: The Random Forest and Decision Tree models have significantly different performances.
```

Z-Test

```
cv_strategy = StratifiedKFold(n_splits=10, shuffle=True, random_state=42)
dt_scores = cross_val_score(dt_model, X, y, cv=cv_strategy)
rf_scores = cross_val_score(rf_model, X, y, cv=cv_strategy)

differences = rf_scores - dt_scores
mean_diff = np.mean(differences)
std_diff = np.std(differences, ddof=1)
std_error = std_diff / np.sqrt(len(differences))

if std_error == 0:
    print("\nStandard error is zero, Z-Test cannot be performed.")
    z_stat = np.nan
    p_value_z = np.nan
else:
    z_stat = mean_diff / std_error
    p_value_z = 2 * (1 - norm.cdf(abs(z_stat)))

print("Z-Test (10-Fold Cross-Validation)")
print("\nNull Hypothesis (H0): There is no difference in mean performance between the Decision Tree and Random Forest models.")
print("Alternative Hypothesis (H1): There is a difference in mean performance between the two models.")
print(f"\nZ-Statistic: {z_stat}, P-value: {p_value_z}")
if p_value_z <= 0.05:
    print("\nH₀ rejected: The difference in performance between Random Forest and Decision Tree is statistically significant.")
else:
    print("\nH₀ accepted: No significant difference between Random Forest and Decision Tree performances.")
```

```
Z-Test (10-Fold Cross-Validation)

Null Hypothesis (H0): There is no difference in mean performance between the Decision Tree and Random Forest models.
Alternative Hypothesis (H1): There is a difference in mean performance between the two models.

Z-Statistic: 2.968489416634384, P-value: 0.0029926740776020644

H₀ rejected: The difference in performance between Random Forest and Decision Tree is statistically significant.
```