**CSE-6363-002 Machine Learning Programming Assignment-1**

Name: - Pinky Patel

UTA\_ID: - 1002251887

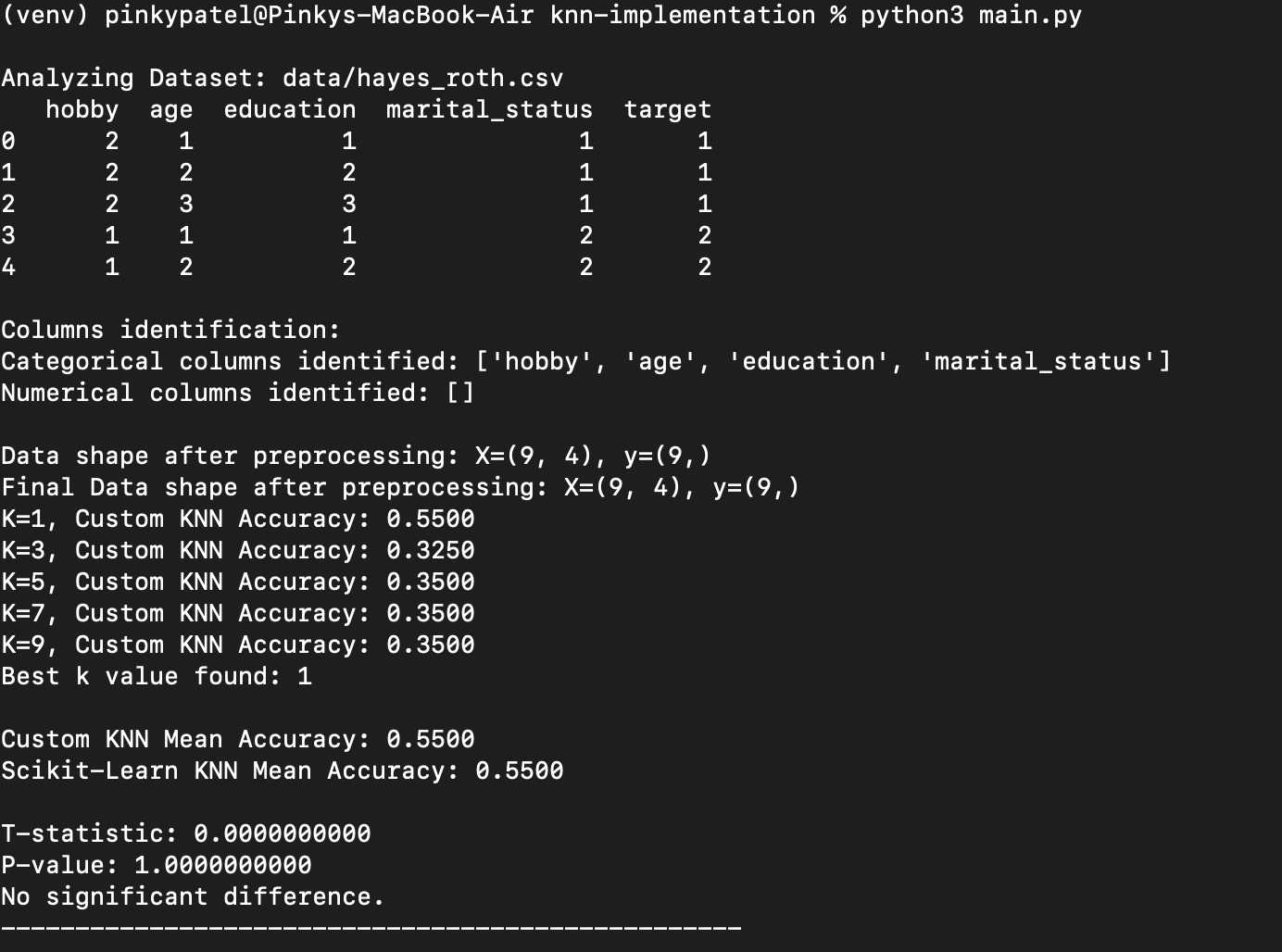
**Comparison between implementation of a Custom KNN algorithm and a sklearn KNN algorithm**

**1.Hayes Roth Dataset.**

While implementing this dataset, I achieved an accuracy of 55.00% using KNN from scratch, whereas Sklearn resulted in an accuracy of 55.00%.

Hypothesis Testing:

Null Hypothesis: The difference between the accuracies of the two tests is insignificant.



Custom KNN Mean Accuracy: 0.5500

Scikit-Learn KNN Mean Accuracy: 0.5500

T-statistic: 0.0000000000

P-value: 1.0000000000

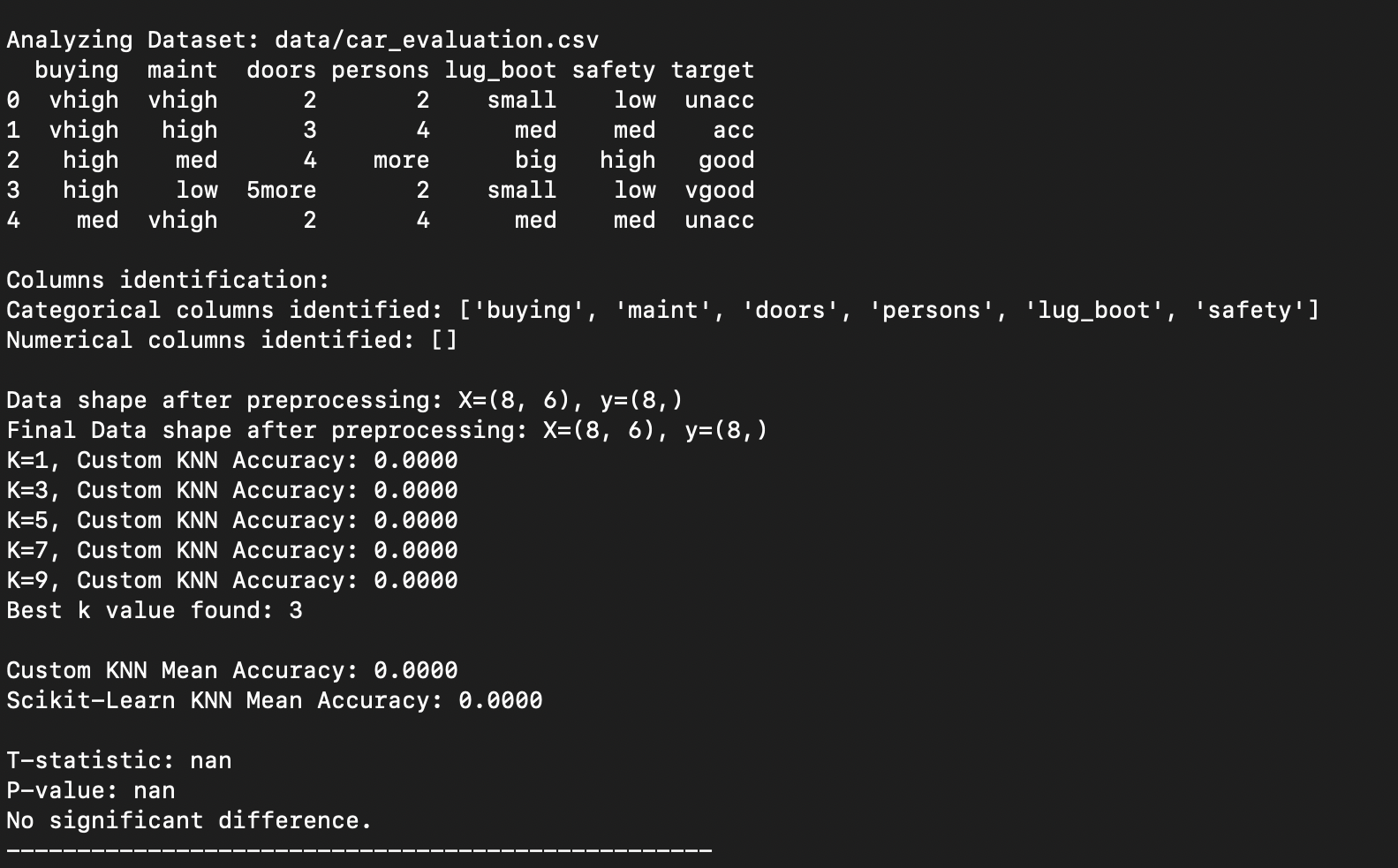
No significant difference.

**2.Car-Evaluation Dataset.**

While implementing this dataset, I achieved an accuracy of 0.00% using KNN from scratch, whereas using Sklearn also resulted in an accuracy of 0.00%.

Hypothesis Testing:

Null Hypothesis: The difference between the accuracies of the two tests is insignificant.



Custom KNN Mean Accuracy: 0.0000

Scikit-Learn KNN Mean Accuracy: 0.0000

T-statistic: nan

P-value: nan

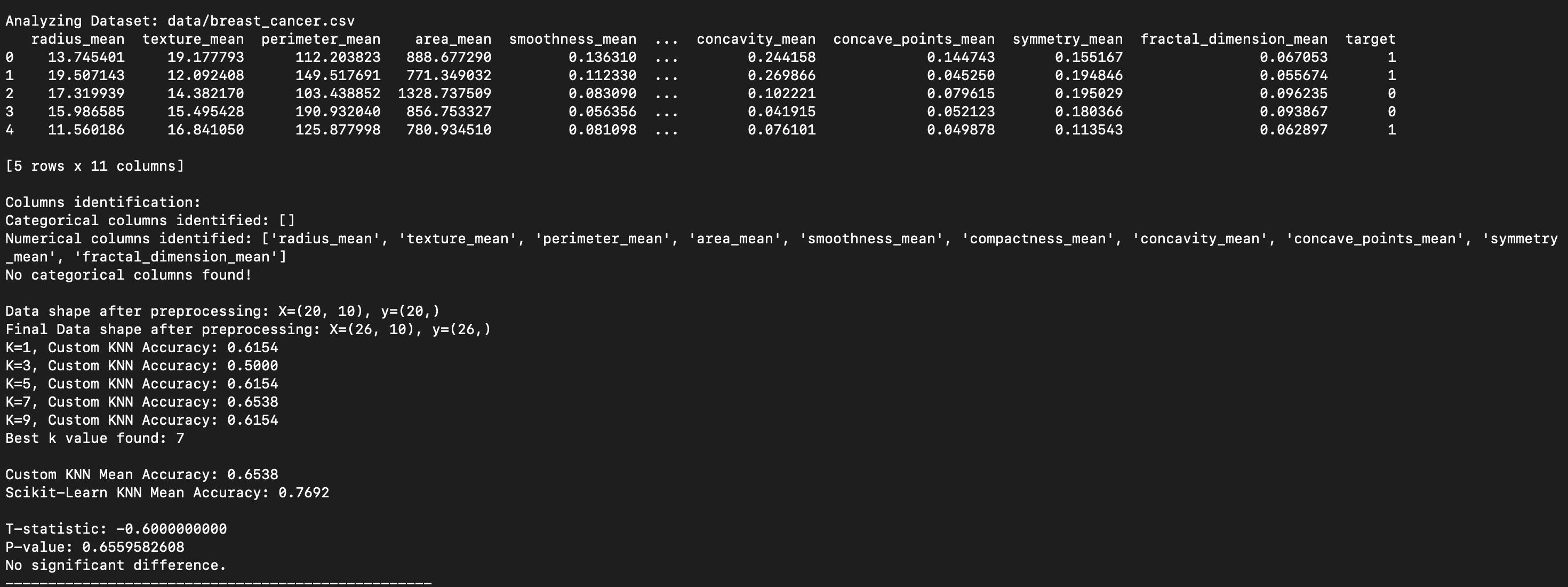
No significant difference.

**3.Breast Cancer Dataset.**

While implementing this dataset, I achieved an accuracy of 65.38% using KNN from scratch, whereas using Sklearn resulted in an accuracy of 76.92%.

Hypothesis Testing:

Null Hypothesis: The difference between the accuracies of the two tests is insignificant.

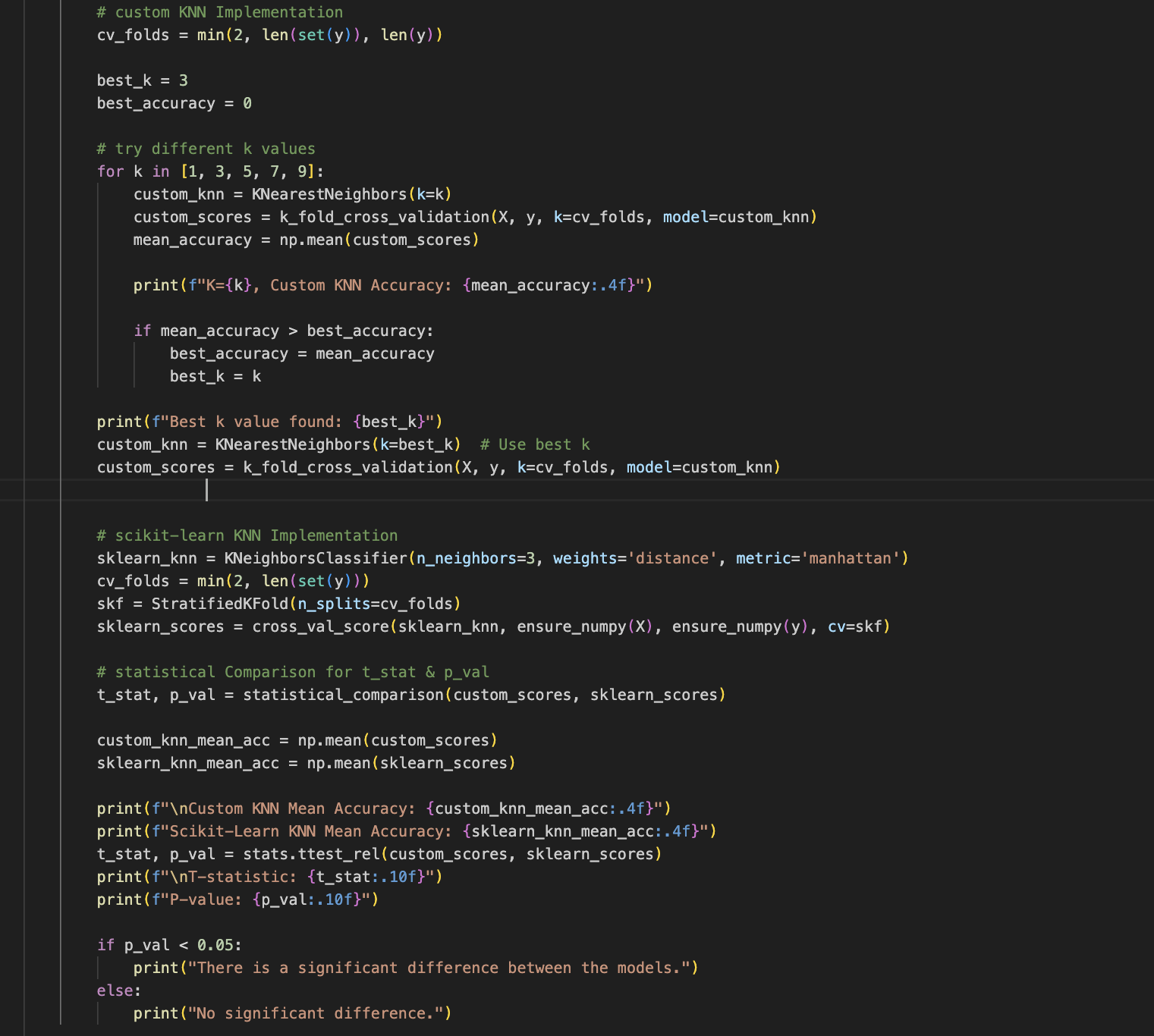


Custom KNN Mean Accuracy: 0.6538   
Scikit-Learn KNN Mean Accuracy: 0.7692

T-statistic: -0.60  
 P-value: 0.6559   
No significant difference.

**Hypothesis Testing: (For all 3 datasets)**

**Null Hypothesis:** The difference between the accuracies of the two tests is insignificant.



**References:-**

* <https://machinelearningmastery.com/tutorial-to-implement-k-nearest-neighbors-in-python-from-scratch/>
* <https://machinelearningmastery.com/k-fold-cross-validation/>
* <https://archive.ics.uci.edu/dataset/44/hayes+roth>
* <https://archive.ics.uci.edu/dataset/19/car+evaluation>
* <https://archive.ics.uci.edu/dataset/14/breast+cancer>
* <https://www.geeksforgeeks.org/k-nearest-neighbors-with-python-ml/>
* <https://scikit-learn.org/stable/modules/generated/sklearn.neighbors.KNeighborsClassifier.html>
* <https://www.geeksforgeeks.org/how-to-conduct-a-paired-samples-t-test-in-python/>