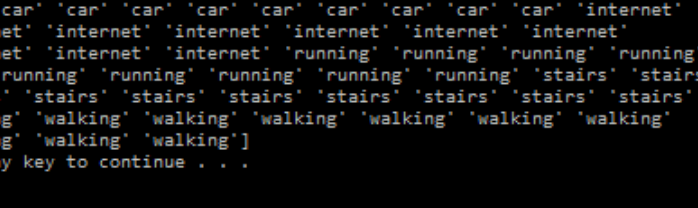


## Logistic Regression

This algorithm got 49/50 predictions right. I incorrectly classified one of the “running” test data points as a “car” data point.



A screenshot of a Windows command prompt window. The title bar at the top shows the file path 'C:\Program Files\Anaconda3\python.exe'. The command prompt displays a list of words: 'car', 'car', 'car', 'car', 'car', 'car', 'car', 'car', 'car', 'car', 'car', 'internet', 'internet', 'internet', 'internet', 'internet', 'internet', 'internet', 'internet', 'internet', 'internet', 'running', 'running', 'running', 'running', 'running', 'running', 'running', 'running', 'car', 'running', 'running', 'running', 'running', 'running', 'running', 'stairs', 'stairs', 'stairs', 'stairs', 'stairs', 'stairs', 'stairs', 'stairs', 'stairs', 'stairs', 'walking', 'walking', 'walking', 'walking', 'walking', 'walking', 'walking', 'walking', 'walking', 'walking', 'walking'. The word 'car' on the 18th line is highlighted with a red rectangular box. At the bottom of the window, the text 'Press any key to continue . . .' is visible.

## **K Neighbors and SVM**

Both of these algorithms successfully classified 50/50 test points. I think this worked better for this data in particular because there are clear clusters in the data—the boundary lines are not ambiguous.

Linear regression would probably work better in a case where the data is not in such clear clusters, and a regression line needs to be found that actually fits the data.