

# Vehicle Failure Prediction Using Machine Learning

Team Members: Madison Ashbach, Thomas Bye, Logan Chantry, Dylan Miska

Sponsor: Pedigree Technologies

Capstone 2022



Pedigree wants to predict vehicle failure types from sensor data.



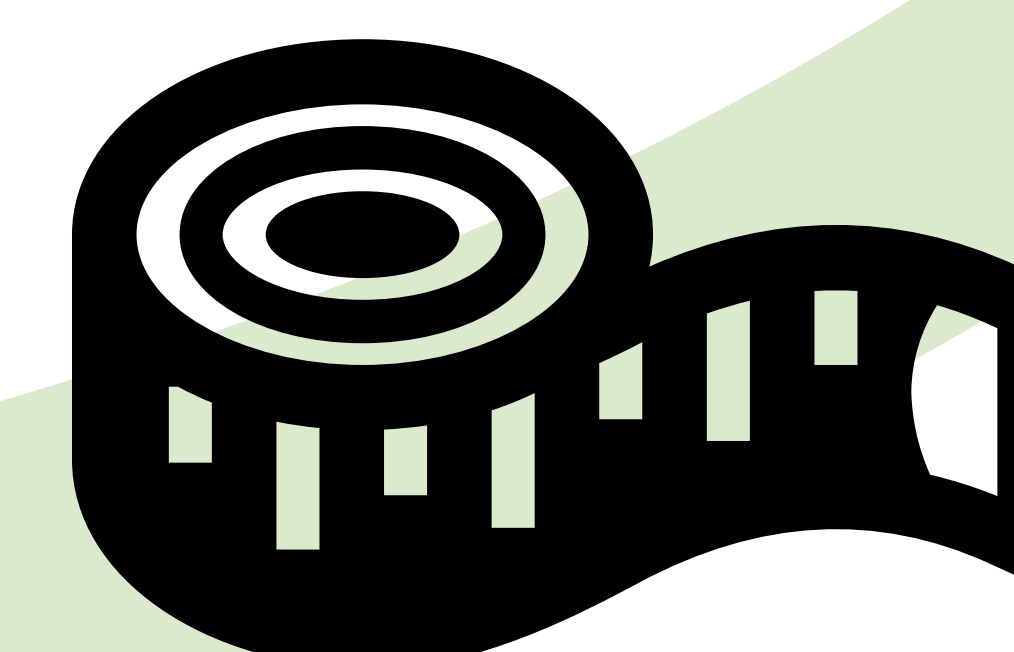
DTCs and failures were merged together and saved to a CSV file.

```
for i in range(length):
    fractionleft = i/length
    percentage = "{:.0%}".format(fractionleft)
    print(f'{i} of {length}. {percentage} complete.')
    row = dtcDf.loc[i]
    vehicle = row['Vehicle']

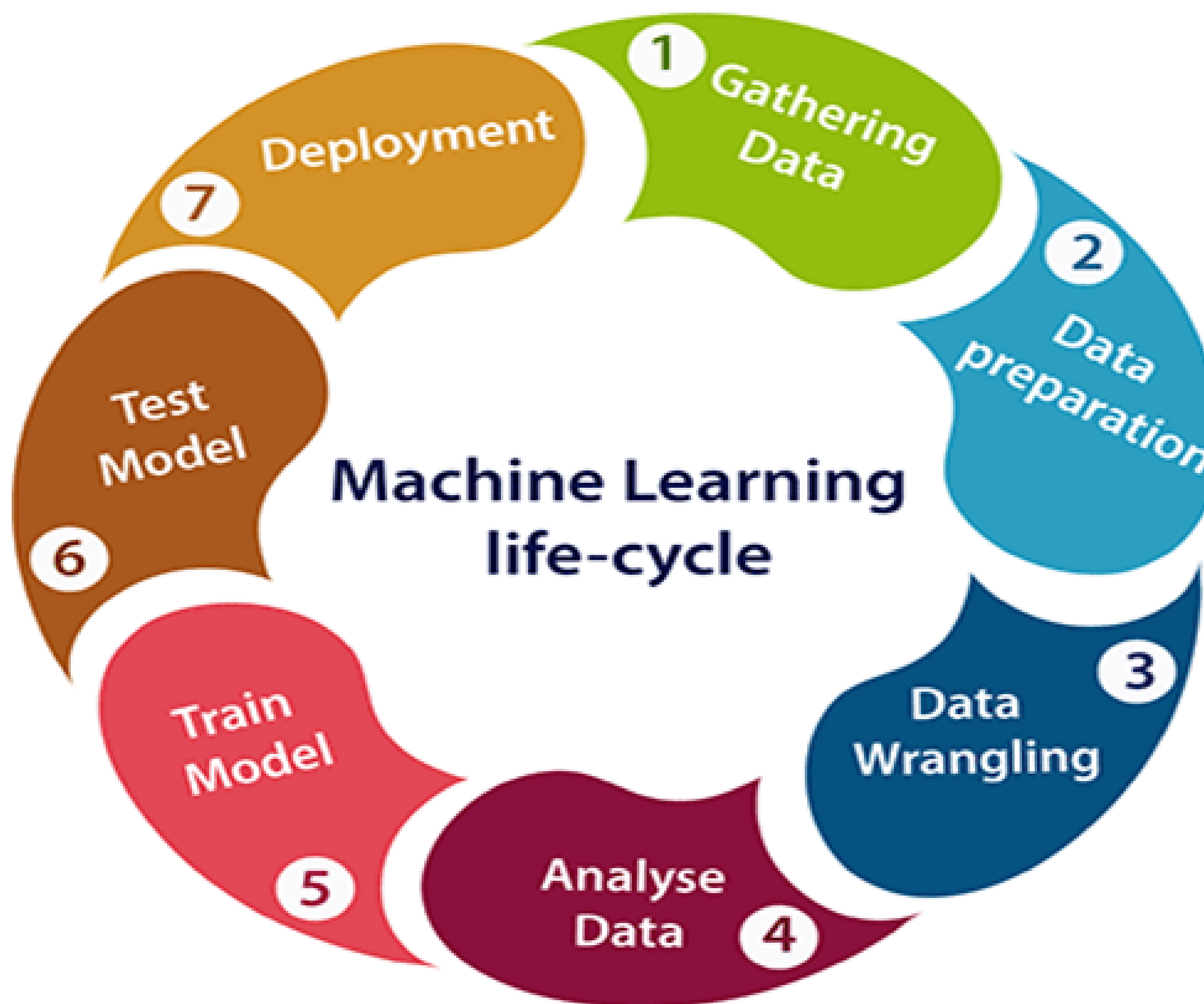
    time = row['Time']

    query = failuresDf[(failuresDf['Vehicle'] == vehicle) & (failuresDf['Time'] > time)]
    if(len(query.index) > 0):
        failureTime = query.iloc[-1]['Time']
        failureIssue = query.iloc[-1]['Issue']
        firstFailureTime.append(failureTime)
        firstFailureIssue.append(failureIssue)
    else:
        firstFailureTime.append('')
        firstFailureIssue.append('')
```

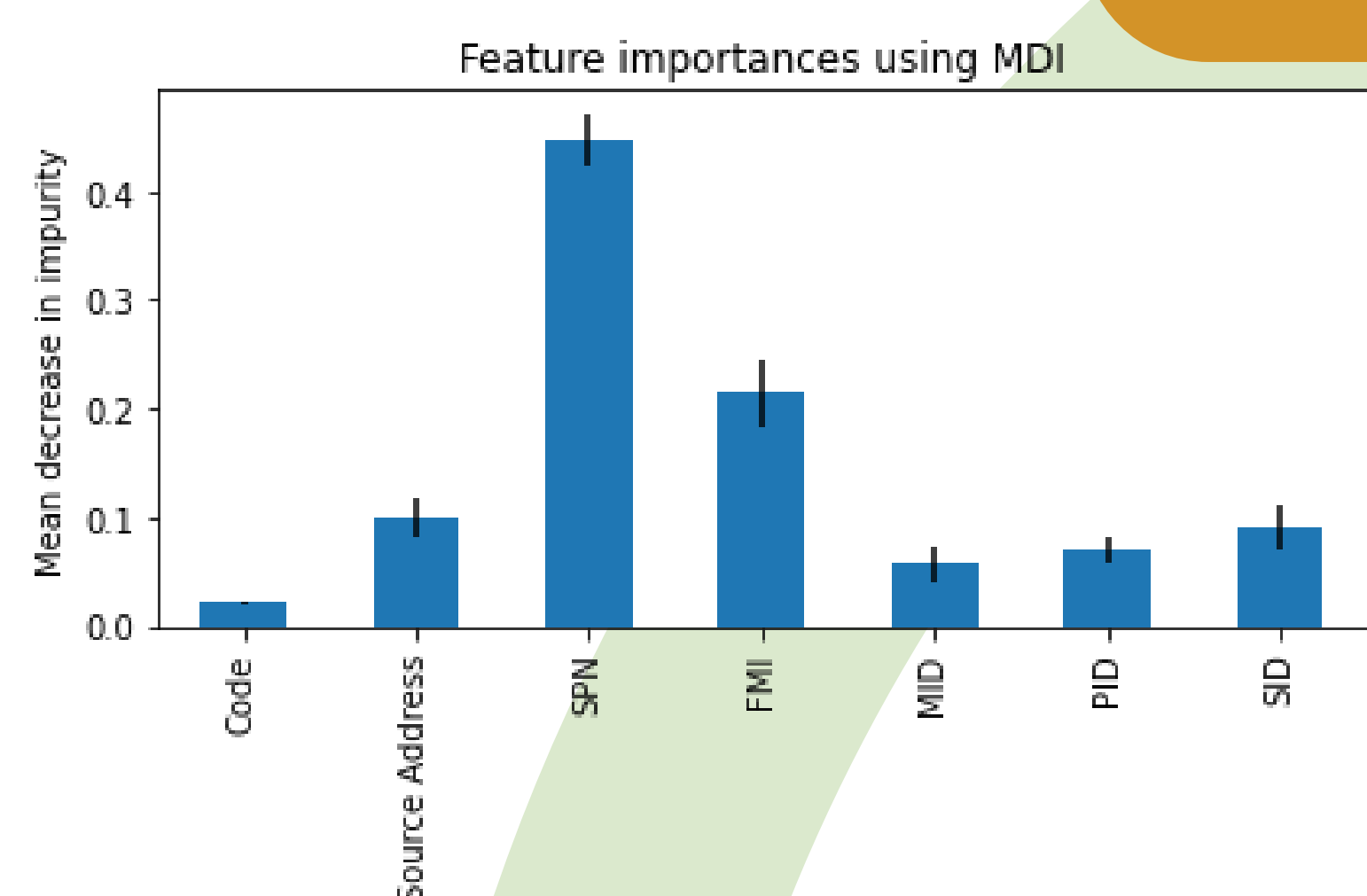
Vehicle without recorded failures were removed from the dataset.



Profiling reports were created using the Pandas Profiling library for data analysis.



Models trained using our methodology can be saved and loaded for production scenarios.



Random Forest had the highest accuracy against the remaining data at 58%.



A classification algorithm called Random Forest was trained using a subset of the data.