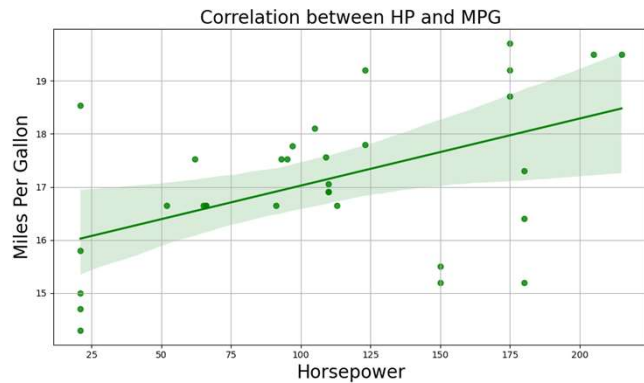


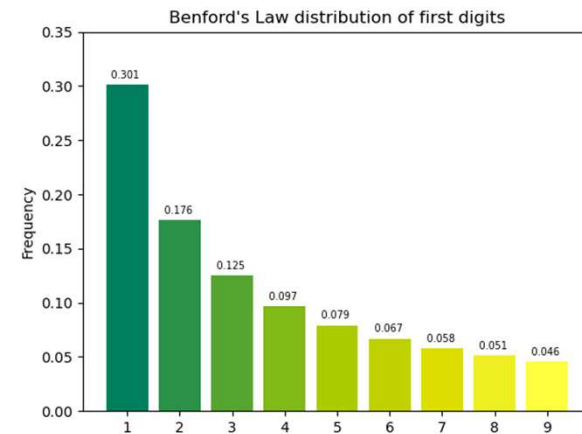
Detection of Correlation Reversal Manipulation via Benford's Law and Random Forest



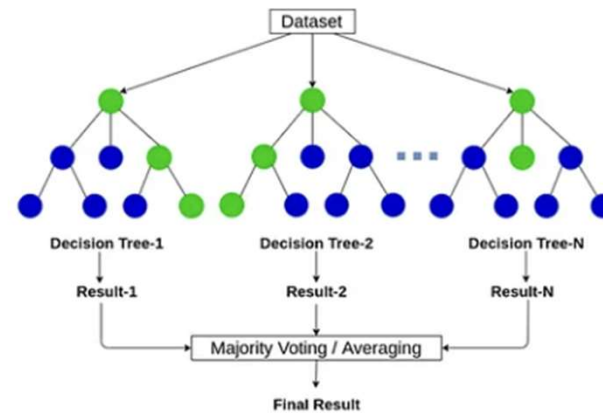
*“The more horsepower we have,
the more miles per gallon we get.”*

Detect Frauds

1. Data Manipulation
2. Correlation Reversal

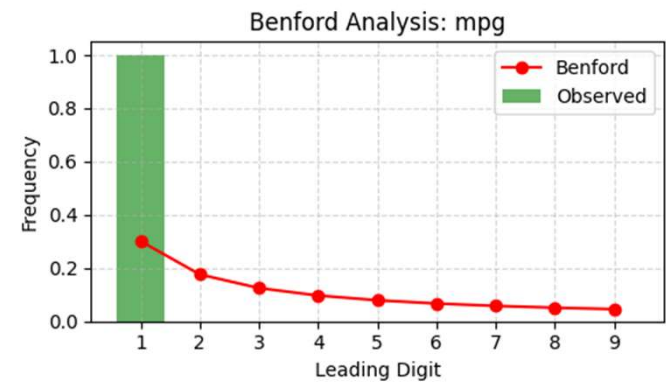
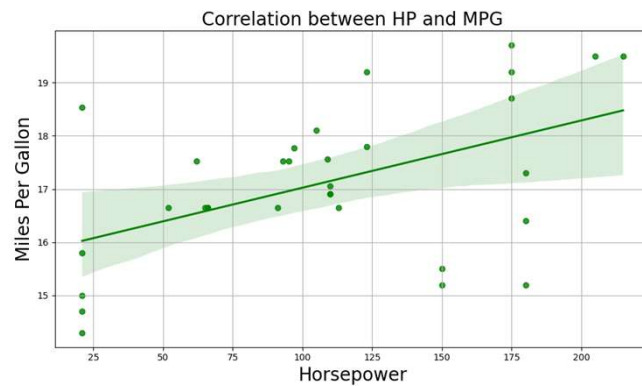


Benford's Law
(Statistical Analysis)



Random Forest
(Machine Learning)

Detect Manipulation using Benford's Law



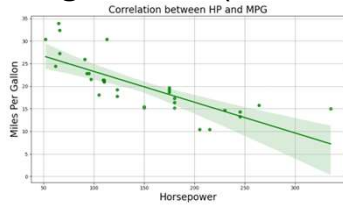
Chi-square = 74.3017, p-value = 0.00000000000006

Criteria: Chi-square > 30.58 and $p < 0.0001$

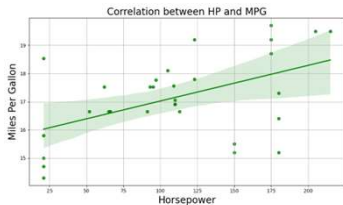
Classified as Benford Anomaly and extreme

Detect Manipulation using Random Forest

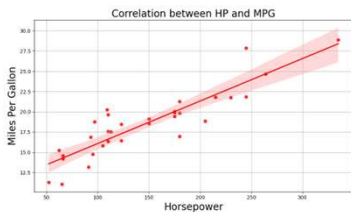
Original Data (Label=0)



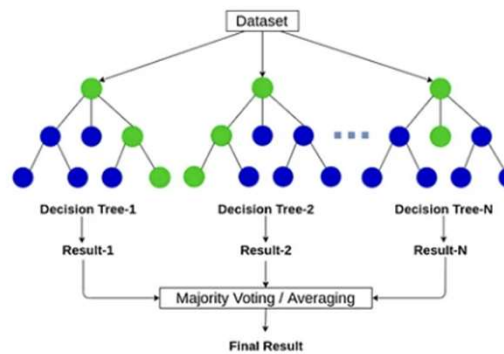
Manipulated Data (Label=1)



New Manipulated Data



Supervised Training



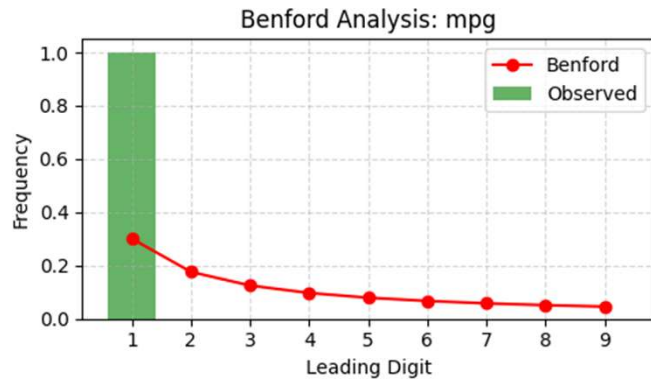
RF
Model

Classified

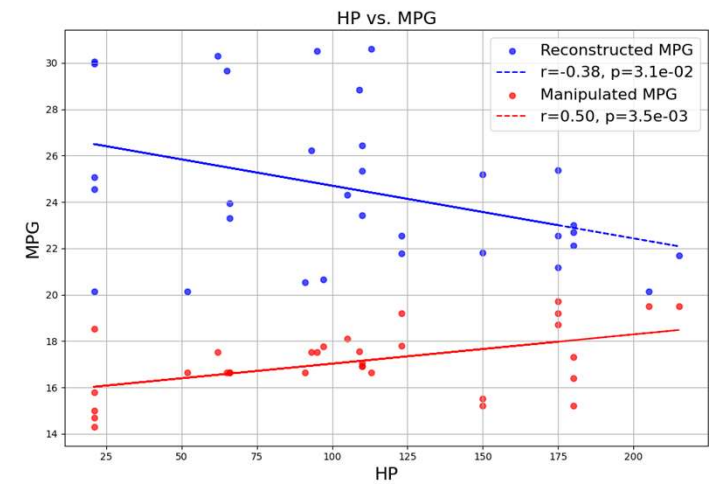
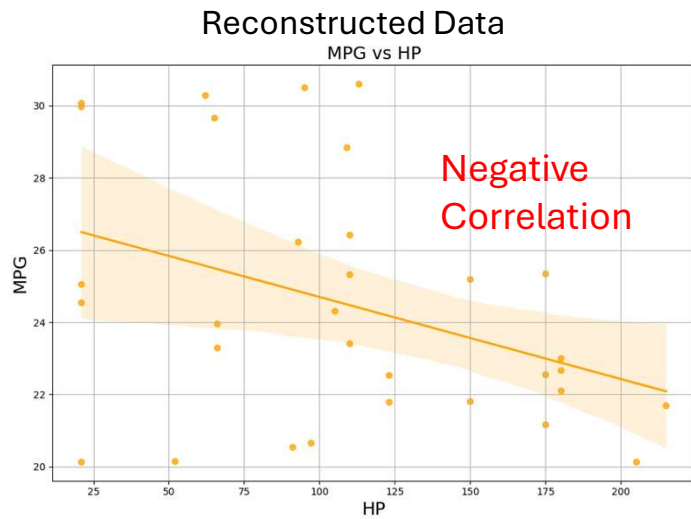
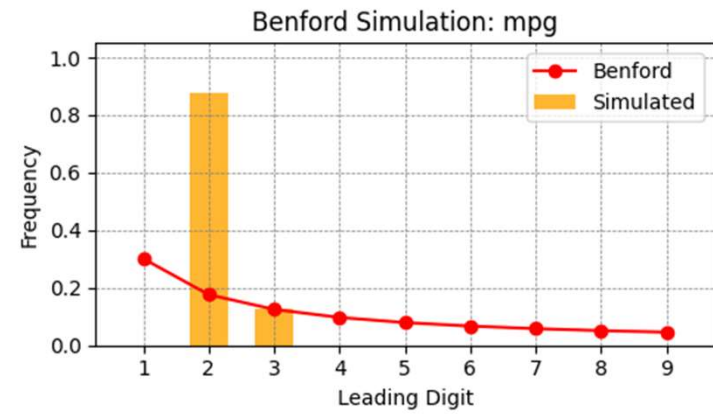
RF
Model

As **manipulated**

Detect Correlation Reversal using Benford's Law

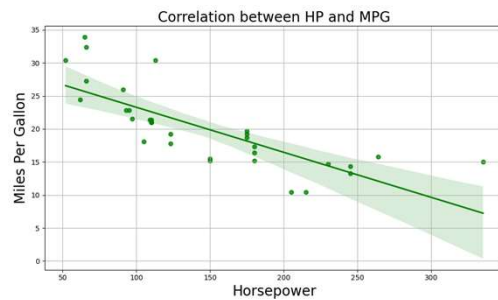


Stochastic
Simulation

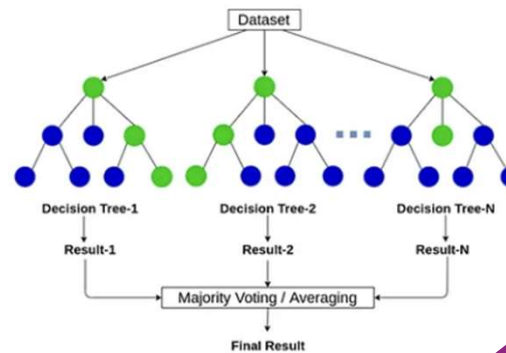


Detect Correlation Reversal using Random Forest

Original Data

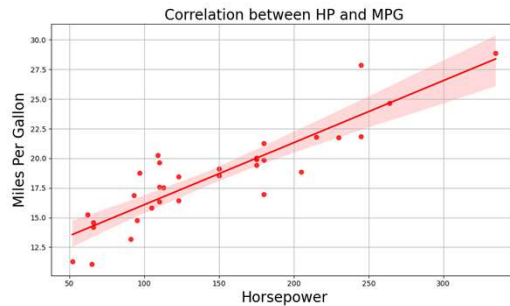


Training



RF
Model

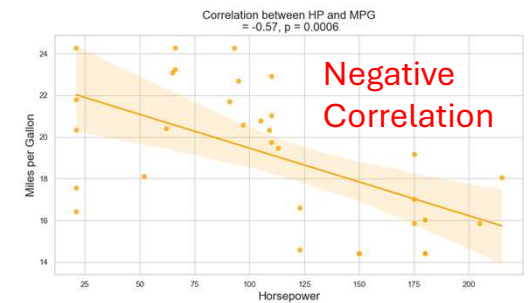
Manipulated Data



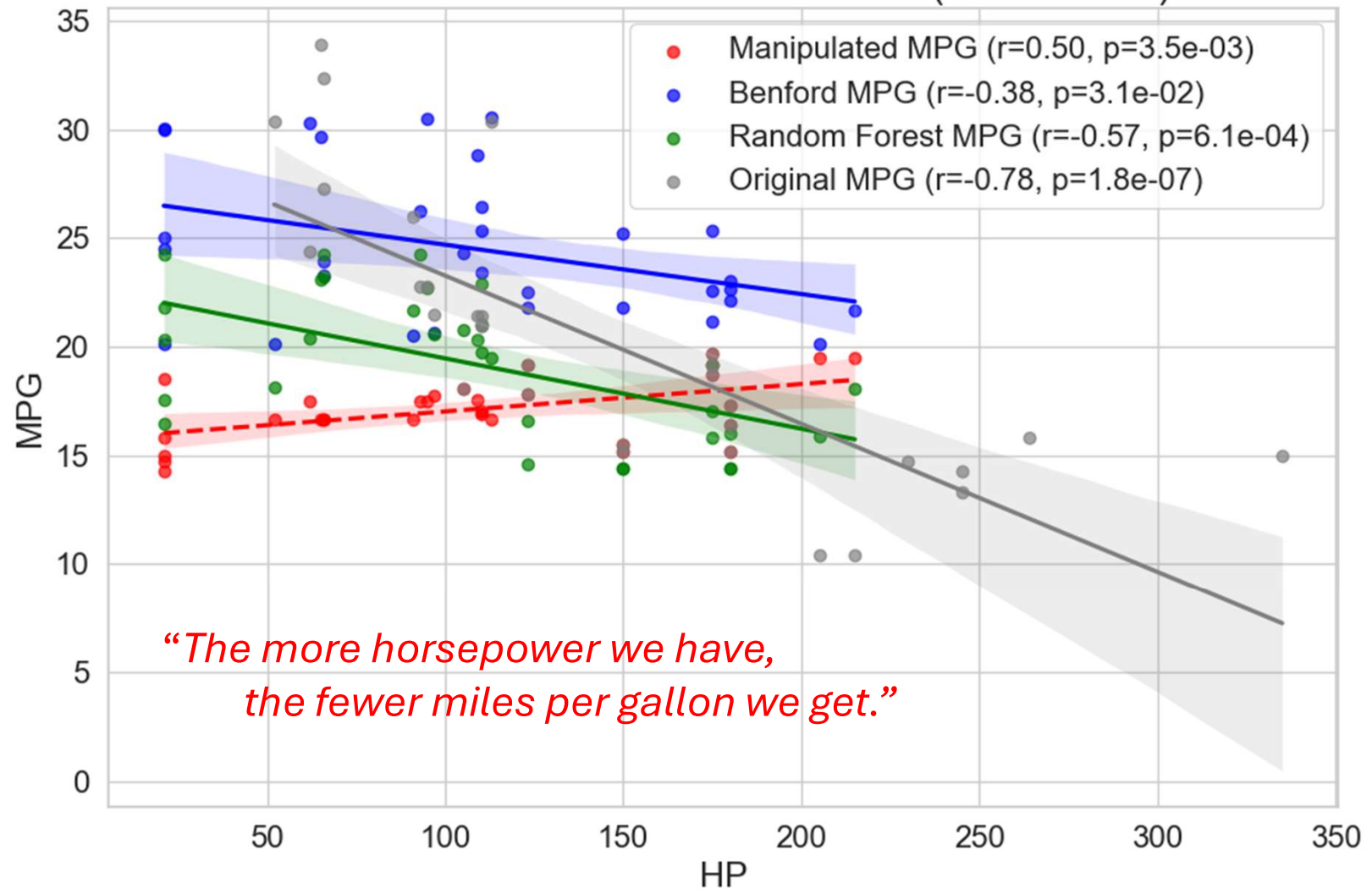
Regression

RF
Model

Reconstructed Data



Correlation between HP and MPG (Given Data)



Correlation between HP and MPG (My Data)

