Hunting the Energy Optimal Fleet Composition

Sustainability and Emerging Transportation Technology (SETT) Conference

June 1, 2022

Qadeer Ahmed, Sharat Hegde, Manfredi Villani and Akarsh Konaje

Acknowledgment: "This material is based upon work supported by the U.S. Department of Energy's Office of Energy Efficiency and Renewable Energy (EERE) under the Advanced Vehicle Technologies Office Award Number DE-EE0009207."

THE OHIO STATE UNIVERSITY
CENTER FOR AUTOMOTIVE RESEARCH

VEHICLE SPECIFICATION OPTIMIZATION

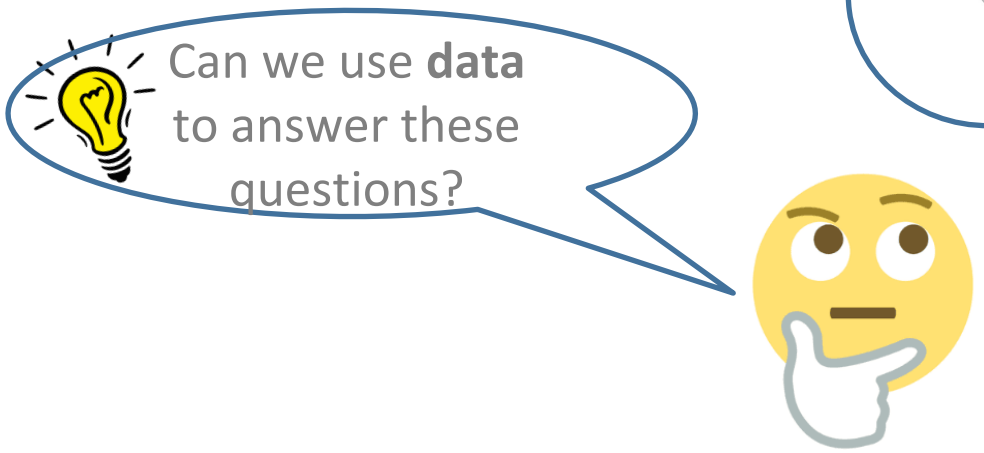
THE OHIO STATE UNIVERSITY
CENTER FOR AUTOMOTIVE RESEARCH

Vehicle Specification

the process of selecting a vehicle's components and sizes

How to select the right powertrain, and vehicle spec for target operation/route?

How to choose the right powertrain tech? (conventional or electric)

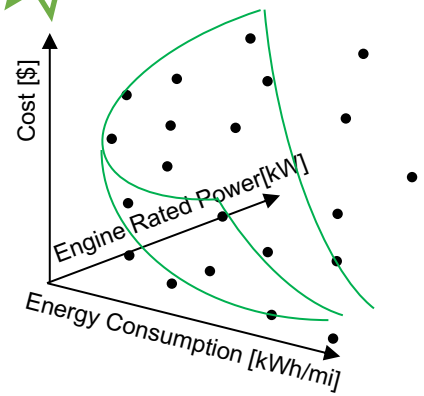


LET'S BUY A TRUCK!

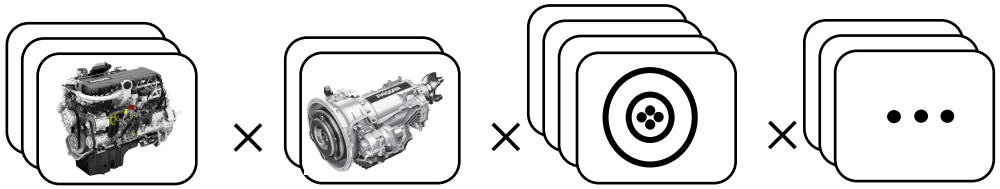
THE OHIO STATE UNIVERSITY
CENTER FOR AUTOMOTIVE RESEARCH

Powertrain Recommender System

PARETO OPTIMAL CONFIGURATIONS



Design Space



$70 \times 8765 \times 105 \times 150 \times 100 = 966,341,250,000$!!!
$3 \times 18 \times 40 \times 15 \times 3 = 97,200$

PERFORMANCE CHECK — Customer Requirements

Feasible Configurations

$3,600$

ESTIMATION MODEL — Drive Cycle

Energy Consumption
Fuel Economy
Carbon Emissions

Find non-dominated solutions

OPTIMIZATION MODEL

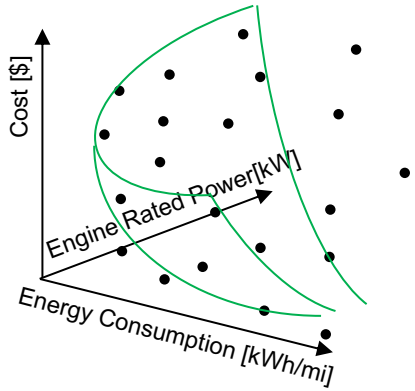
POWERTRAIN RECOMMENDER SYSTEM

Functionalities of the Powertrain Recommender System

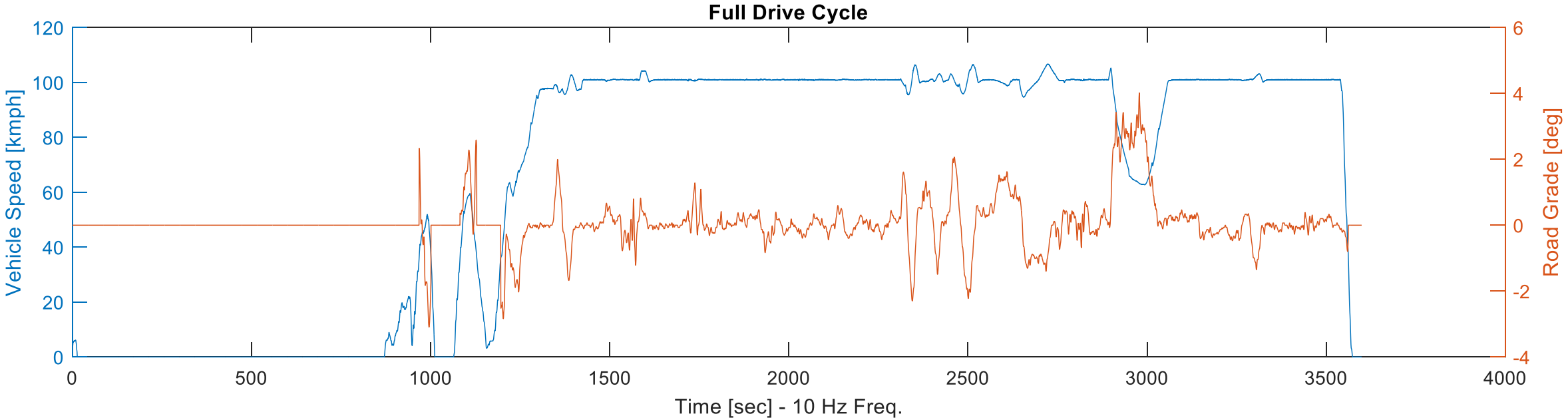
Evaluate Performance

Estimate Energy Consumption

Find the Pareto Optimal Configurations



DATA



DRIVE CYCLE

VEHICLE SPEED PROFILE
ROAD GRADE PROFILE
DISTANCE
TIME
WEATHER (TEMP, PRES.)
FUEL FLOW RATE

VEHICLE SPEC

ENGINE
TRANSMISSION
REAR AXLE
VEHICLE CHASSIS/BODY
AERODYNAMICS
MASS

WHAT IS THE GOAL?

Goal:

$$Estimate\ Energy\ Consumption = f(Drive\ Cycle, Vehicle\ Spec)$$

Large Scale Telematics Data

Expertise on Powertrain Dynamics & Simulation, Machine-Learning, Route-Optimization, UI-UX

A QUICK INTRO TO MACHINE LEARNING

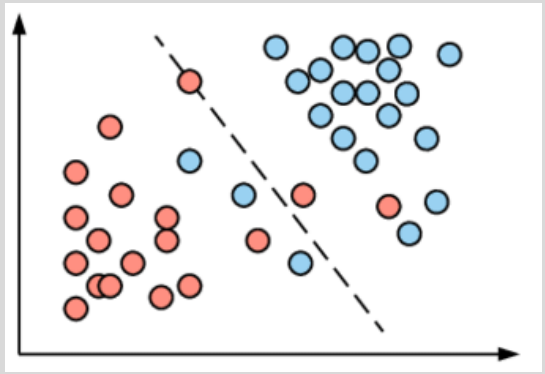
THE OHIO STATE UNIVERSITY
CENTER FOR AUTOMOTIVE RESEARCH

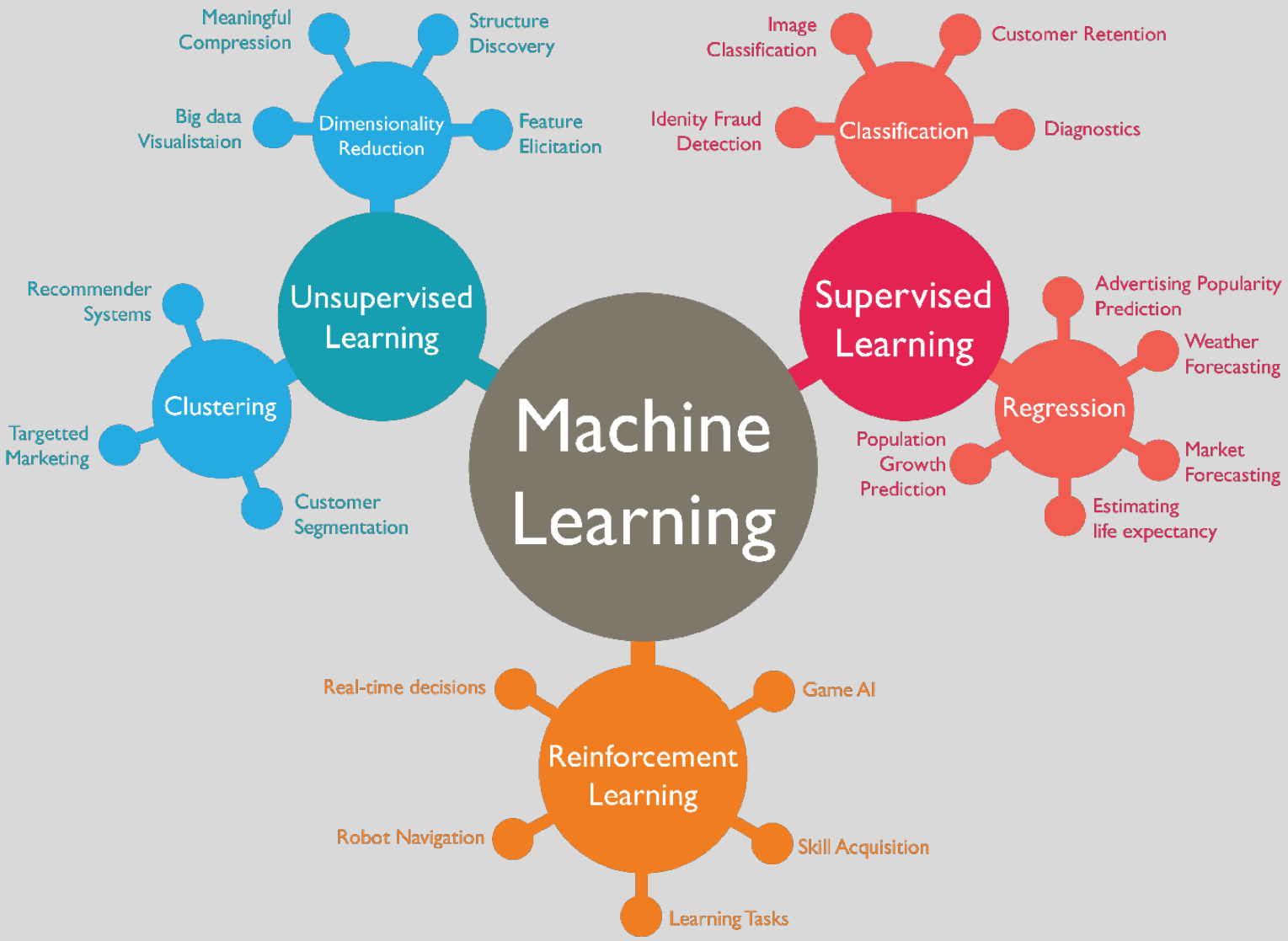
MACHINE LEARNING

ABILITY TO PERFORM TASKS WITHOUT EXPLICIT INSTRUCTIONS AND RELYING ON PATTERNS

SUPERVISED LEARNING

TASK OF LEARNING A FUNCTION THAT MAPS AN INPUT TO AN OUTPUT BASED ON EXAMPLE OF INPUT-OUTPUT PAIRS





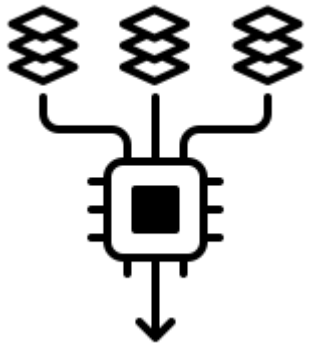
HOW TO BUILD ML MODELS

What do you want to know?

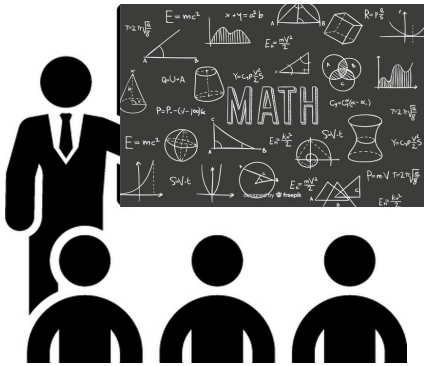
Collect Data

- Data Integration
- Data Cleaning
- Data Analysis
- Data Transformation (Normalization, Aggregation, Generalization)
- Data Visualization
- Data Reduction

Prepare Training, Validation, and Test Data



Train the model using mathematical algorithms



Use the model to predict future outcomes

WHAT DO WE WANT TO KNOW?

Goal:

$$Energy\ Consumption = f(Drive\ Cycle, Vehicle\ Spec)$$

$$where\ \ f = Machine\ Learning\ Model$$

Can we use Machine Learning?

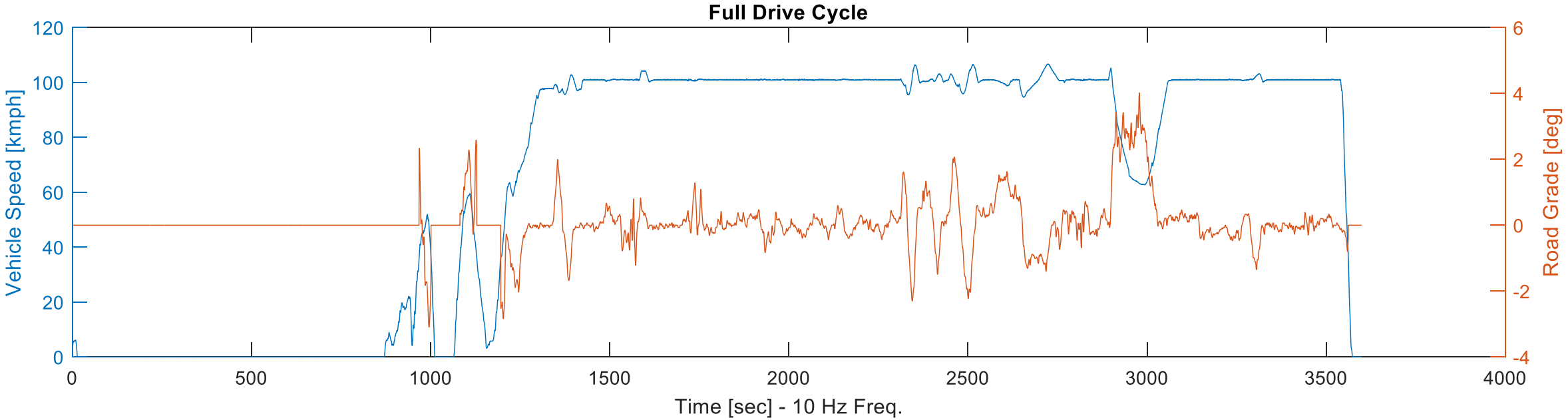
Is there a better alternative?
Computational Efficiency
Accuracy

Choosing the right ML Algorithm?
Type of Data? - Numeric Labelled Data
Supervised Learning Algorithm

Which Supervised Learning Algorithm ?
Classification/Regression? - Regression
Extrapolation/Interpolation? - Both

Algorithm
Random Forest
Neural Network

DATA COLLECTION



DRIVE CYCLE

VEHICLE SPEED PROFILE
ROAD GRADE PROFILE
DISTANCE
TIME
WEATHER (TEMP, PRES.)
FUEL FLOW RATE

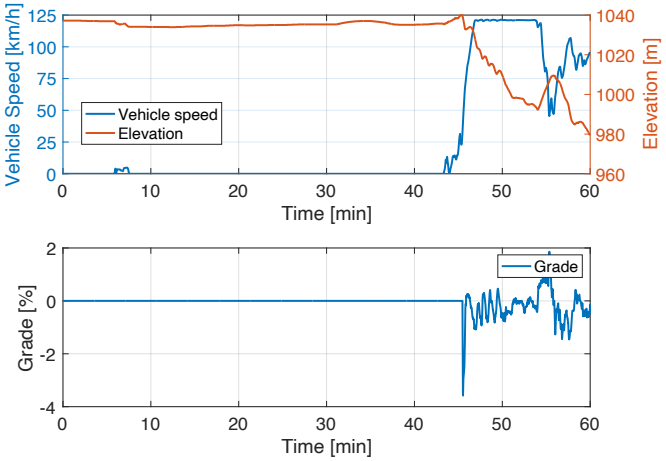
VEHICLE SPEC

ENGINE
TRANSMISSION
REAR AXLE
VEHICLE CHASSIS/BODY
AERODYNAMICS
MASS

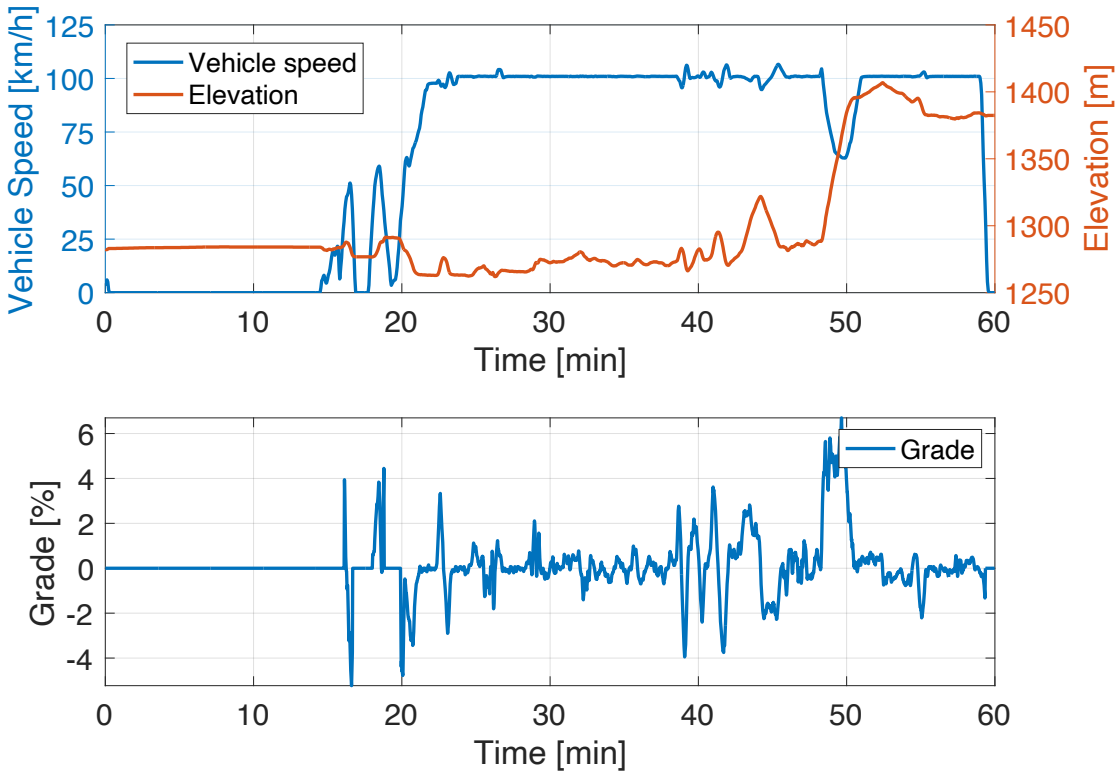
IDEAL DATA vs WHAT YOU GET

THE OHIO STATE UNIVERSITY
CENTER FOR AUTOMOTIVE RESEARCH

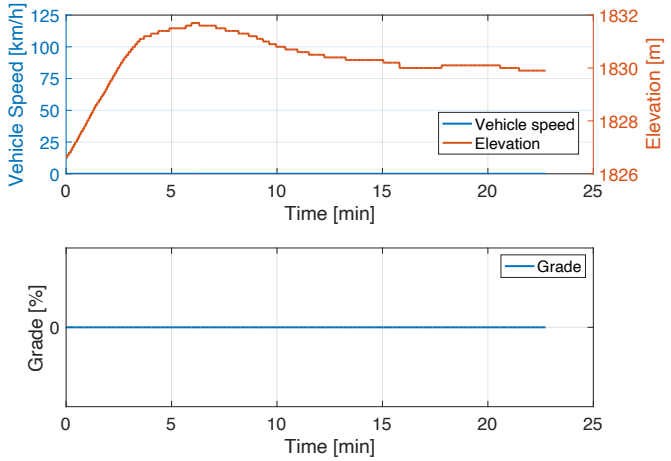
Long time, little data



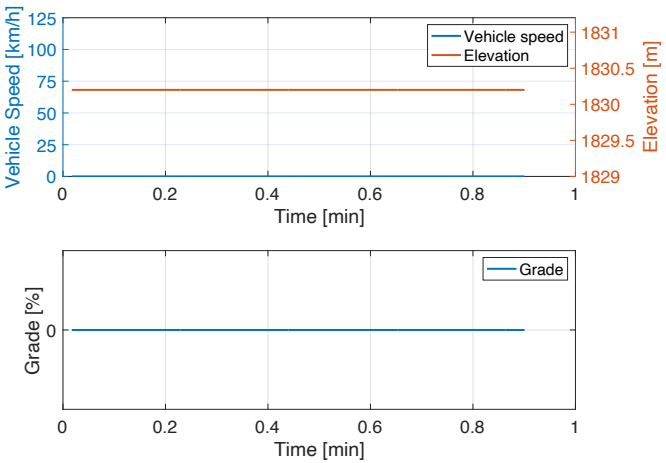
IDEAL



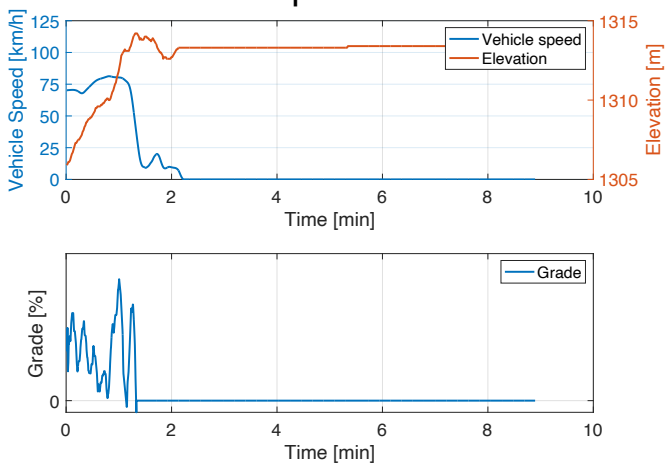
No actual variation in data



Just no data



Incomplete data



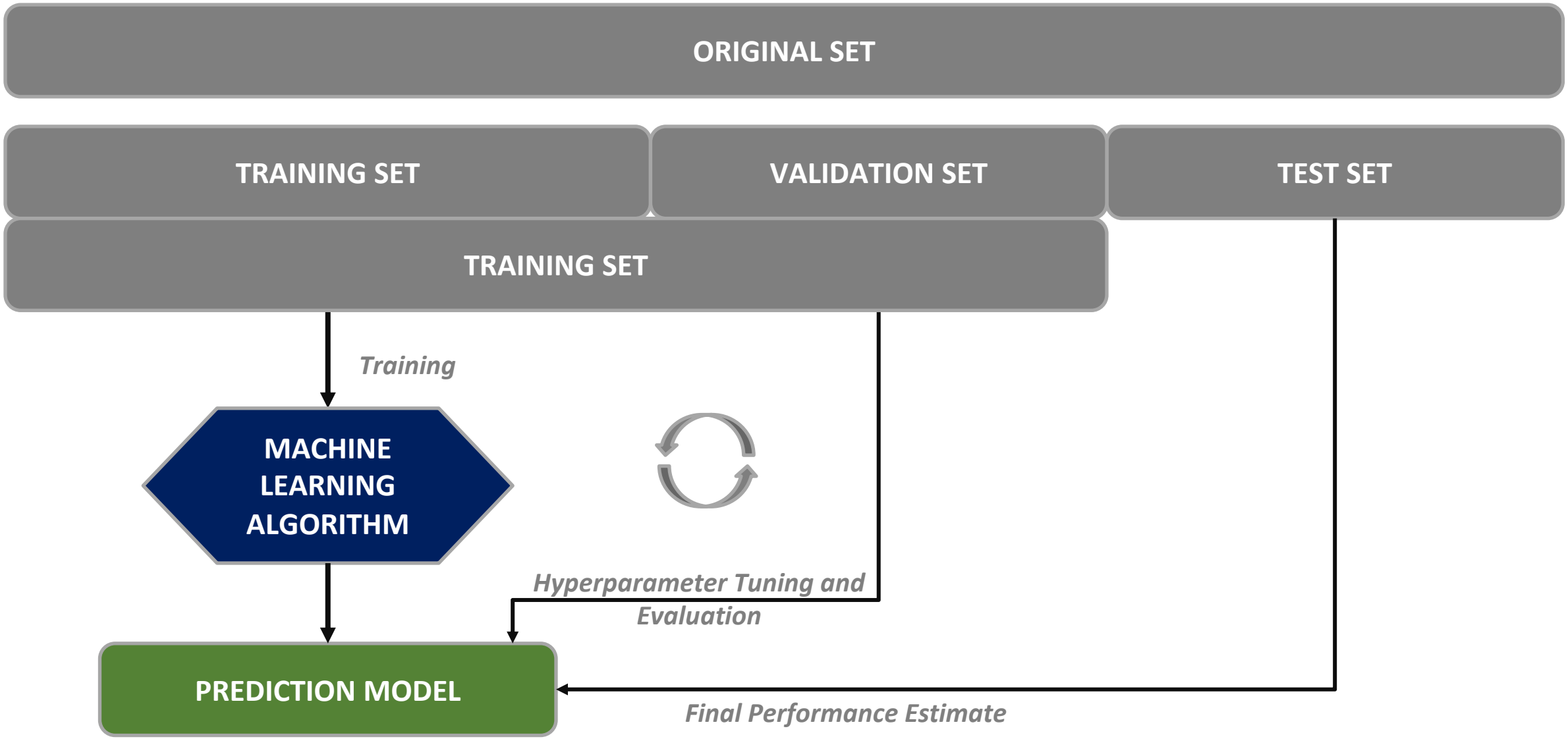
DATA PROCESSING

Data Integrate → Data Cleaning → Data Transformation → Data Reduction

Potential Issue	Source	Action
Trips < 1 Mi	Data logging errors or short trips	Remove
Non-zero Start/Stop	Data logging errors	Remove or Use Gap-Filling-Algos
Unrealistic Values	CAN bus signals error	Remove/Normalize
Time Gaps	Redundant information	Remove
First and Last days of data	Installation and removal of equipment	Remove
Missing data points	Data logging errors	Interpolation
Noise	Sensor noise	Smoothing

TRAINING, VALIDATION, & TEST

THE OHIO STATE UNIVERSITY
CENTER FOR AUTOMOTIVE RESEARCH



TRAINING DATA

Drive Cycle [x 400]	**Time**	Time_sec
	Distance	Distance_km
	Speed Profile	Max_Speed_kmph
		Min_Speed_kmph
		Avg_Speed_kmph
		Median_Speed_kmph
		StdDev_Speed_kmph
		Variance_Speed_kmph
		x25th_Speed_kmph
		x75th_Speed_kmph
	Road Grade Profile	Max_Grade_deg
		Min_Grade_deg
		Avg_Grade_deg
		Median_Grade_deg
		StdDev_Grade_deg
		Variance_Grade_deg
		x25th_Grade_deg
		x75th_Grade_deg
	Acceleration Profile	Max_Accln_mps2
		Min_Accln_mps2
		Avg_Accln_mps2
		Median_Accln_mps2
		StdDev_Accln_mps2
		Variance_Accln_mps2
		x25th_Accln_mps2
		x75th_Accln_mps2
	Idle Time	IdleTime_sec
	Weather	Avg_Pressure_kPa
		Avg_Temp_degC

INPUTS

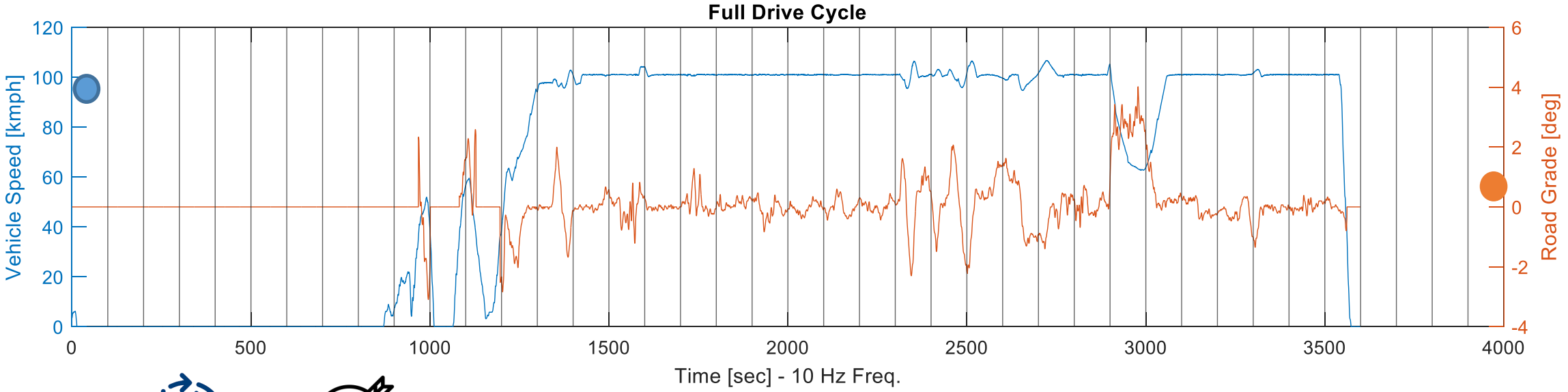
Vehicle Spec 19 Trucks	**Engine [x 3]**	EngineType
		EngineMaxSpeed_RPM
		EnginePeakPower_HP
		EnginePeakTorque_Nm
	Transmission [x 2]	Trans_Speed
		Trans_Deepest_Gear_Ratio
		Trans_Efficiency
	Rear Axle [x 12]	Rear_Axle_Ratio
		RearAxle_Efficiency
	Tire [x 10]	Tire_Radius_inch
		Road_Rolling_Resistance
	Body [x 12]	Vehicle_Frontal_Area_ft2
		Drag_Coefficient
		Curb_Weight_lbs
		GCVW_lbs

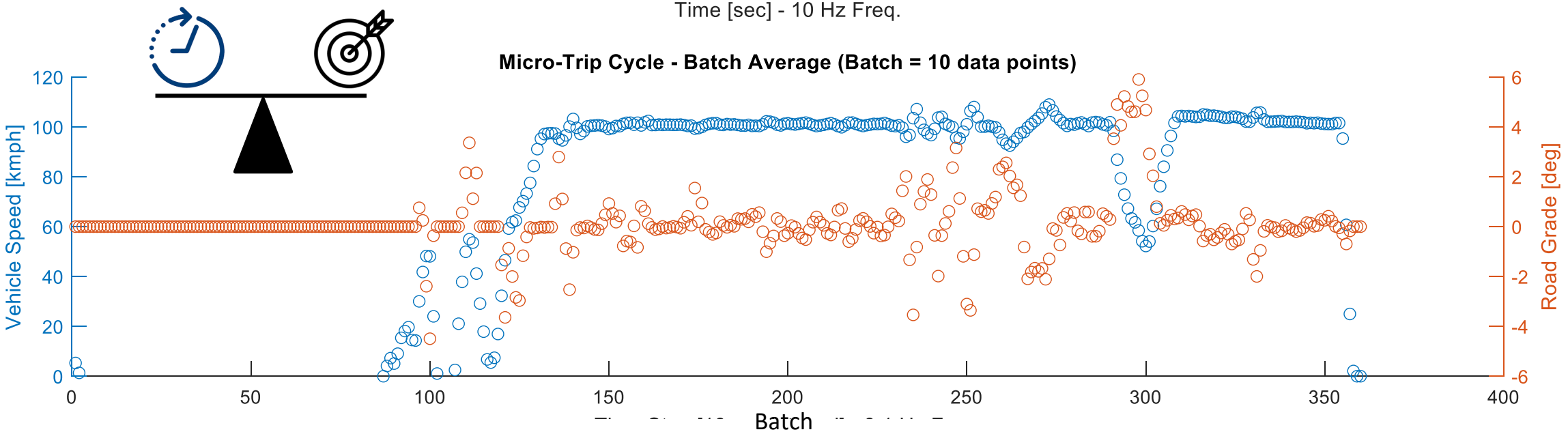
$$Estimate\ Energy\ Consumption = f(Drive\ Cycle, Vehicle\ Spec)$$

$$Supervised\ Learning\ (Input, Output)\ Pairs$$

OUTPUT
FUEL CONSUMPTION

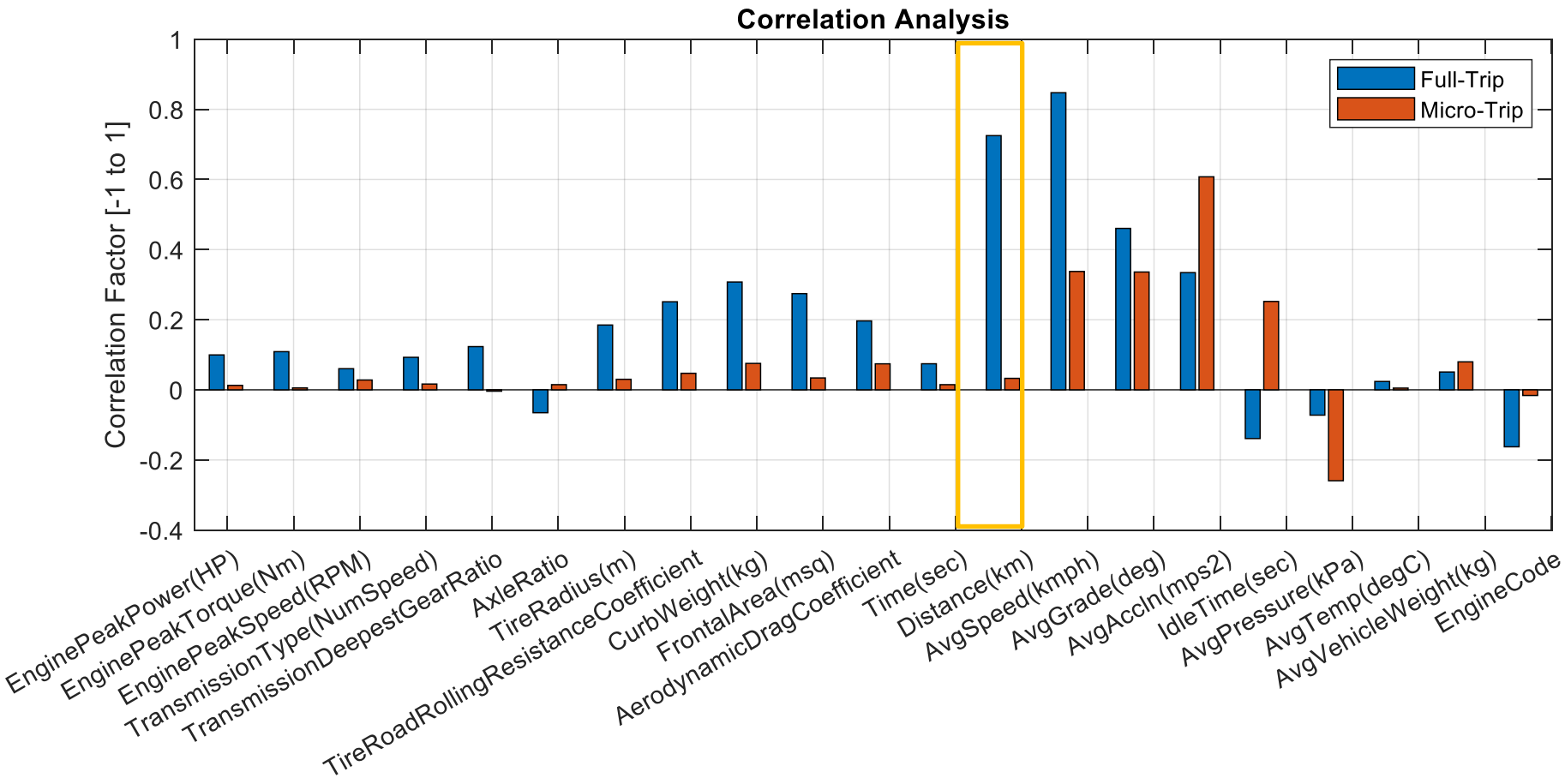
DRIVE CYCLE – FULL-TRIP vs. MICRO-TRIP





HOW DO WE KNOW WHAT'S IMPORTANT?

400 Drive Cycles with 19 Trucks



Correlation

Degree to which a pair of variables are linearly related

Range: $[-1,1]$

Correlation $\neq$ *Causation*

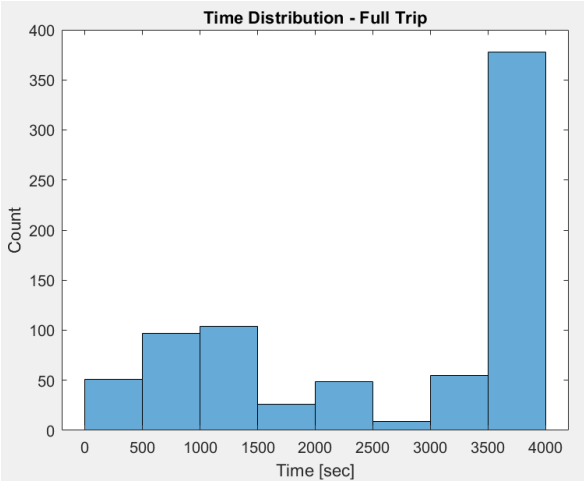
Correlation w.r.to Fuel Consumption

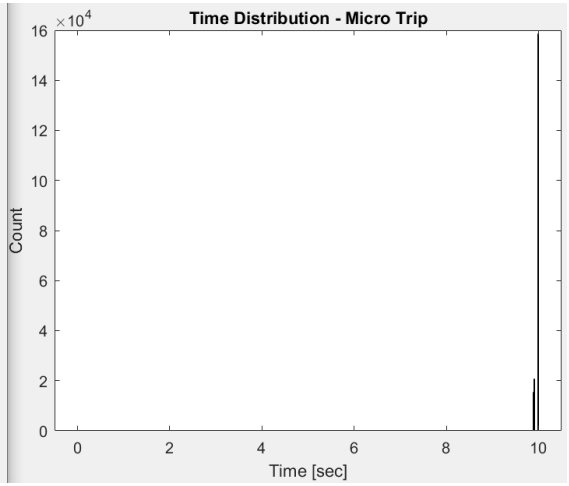
DATA DISTRIBUTION – FULL TRIP vs. MICRO TRIP

THE OHIO STATE UNIVERSITY
CENTER FOR AUTOMOTIVE RESEARCH

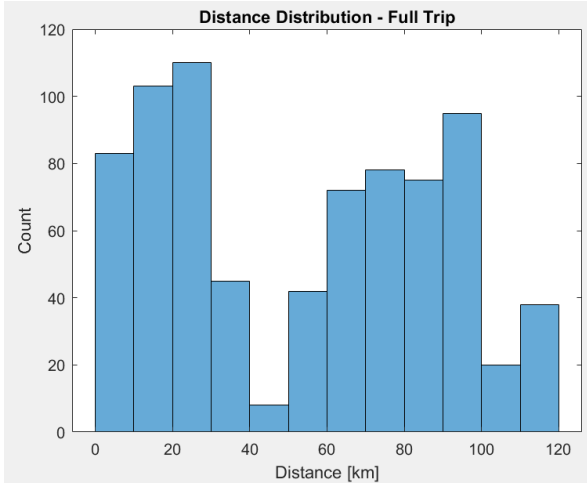
400 Drive Cycles with 19 Trucks

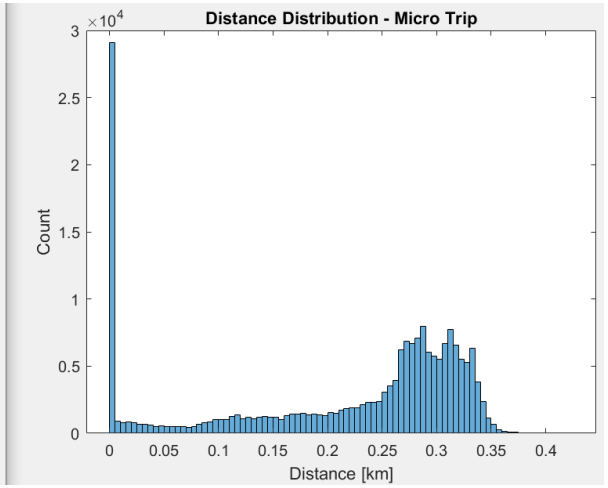
Time



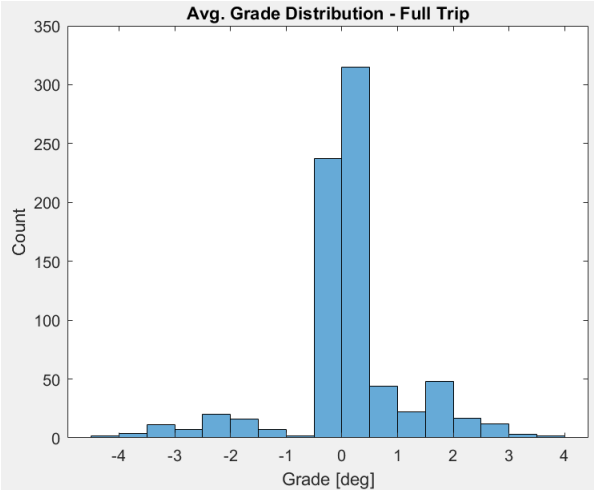


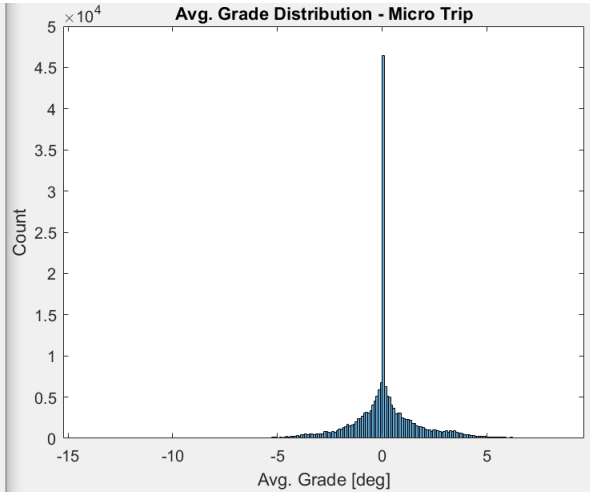
Distance



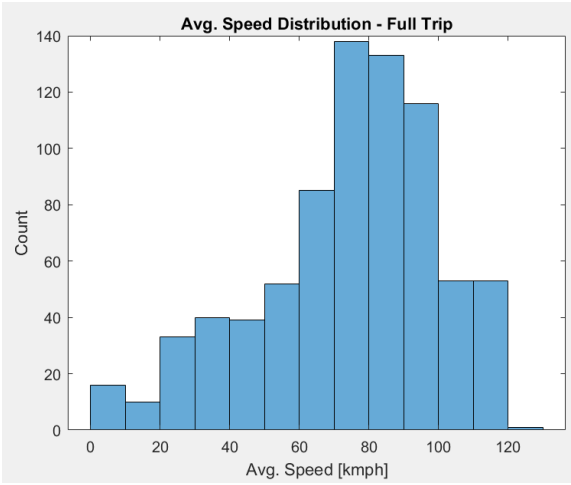


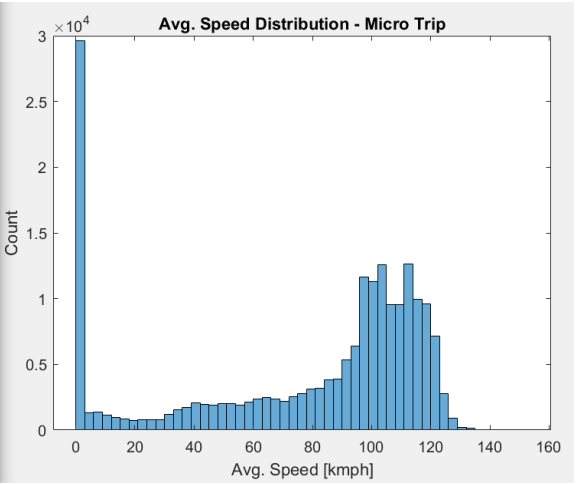
Avg. Grade





Avg. Speed

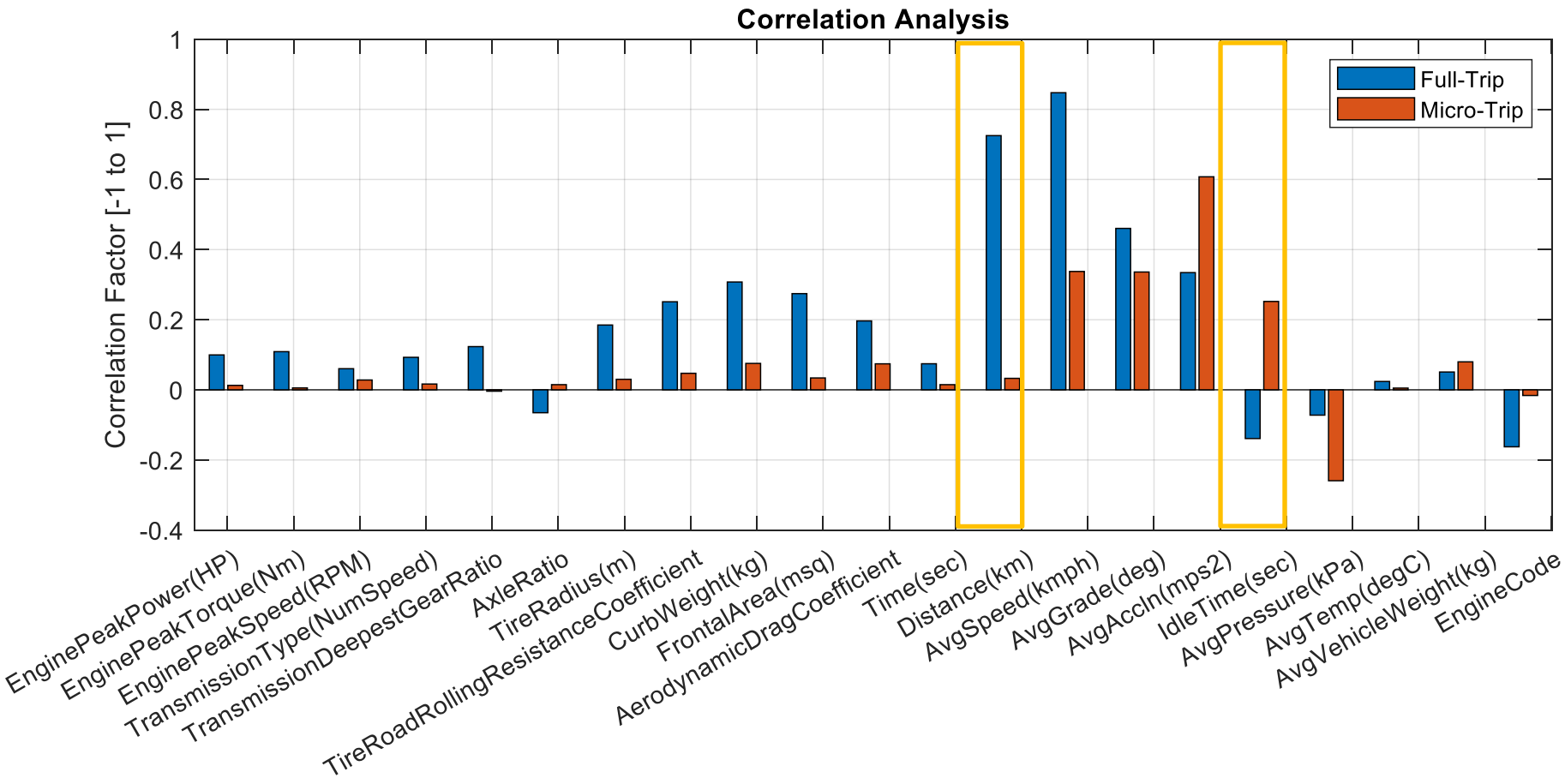




HOW DO WE KNOW WHAT'S IMPORTANT?

THE OHIO STATE UNIVERSITY
CENTER FOR AUTOMOTIVE RESEARCH

400 *Drive Cycles with* 19 *Trucks*



Correlation

Degree to which a pair of variables are linearly related

Range: $[-1,1]$

Correlation $\neq$ *Causation*

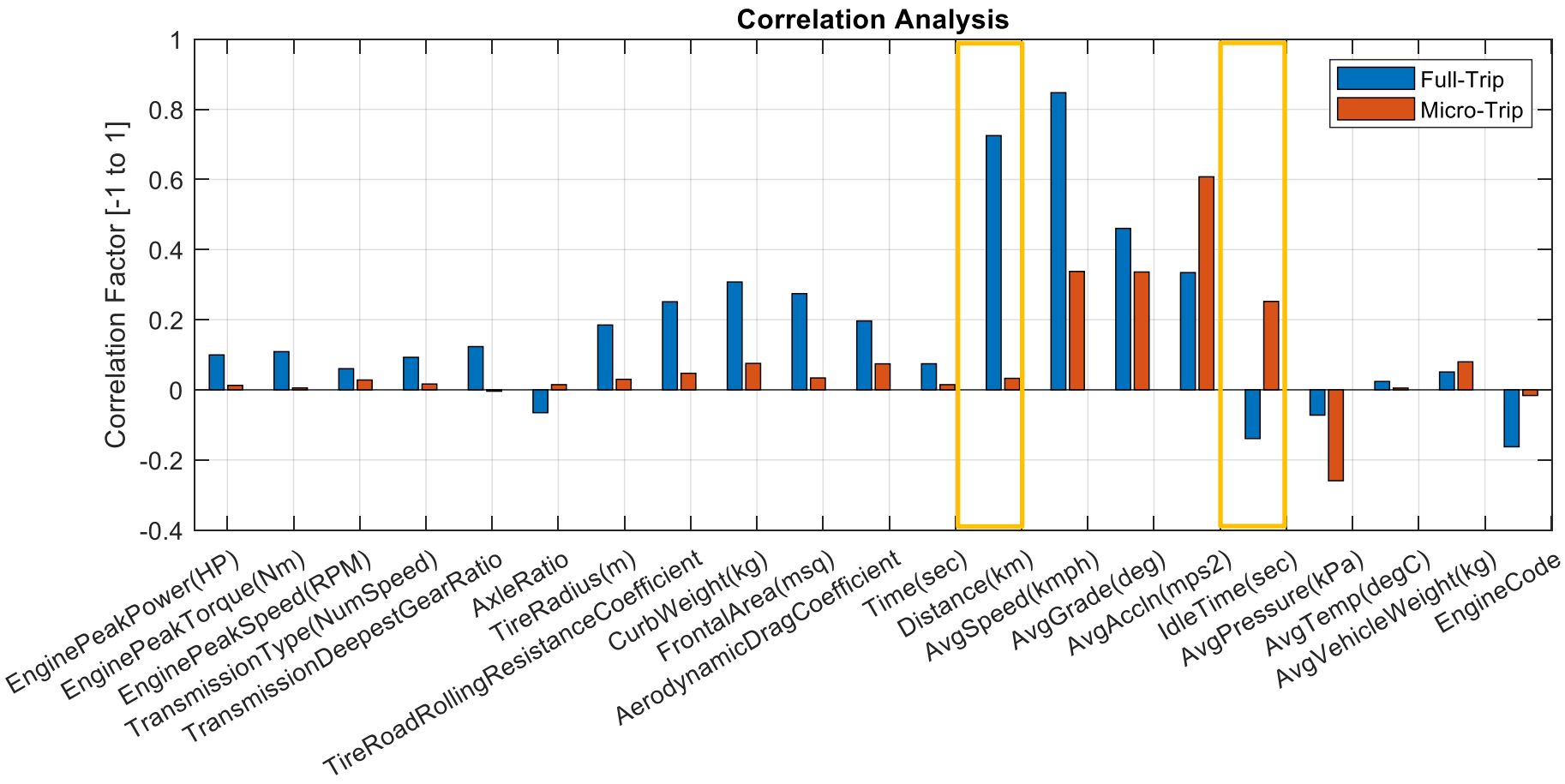
Bias

Correlation w.r.to Fuel Consumption

HOW DO WE KNOW WHAT'S IMPORTANT?

THE OHIO STATE UNIVERSITY
CENTER FOR AUTOMOTIVE RESEARCH

400 *Drive Cycles with* 19 *Trucks*



Correlation

Degree to which a pair of variables are linearly related

Range: $[-1,1]$

Correlation $\neq$ *Causation*

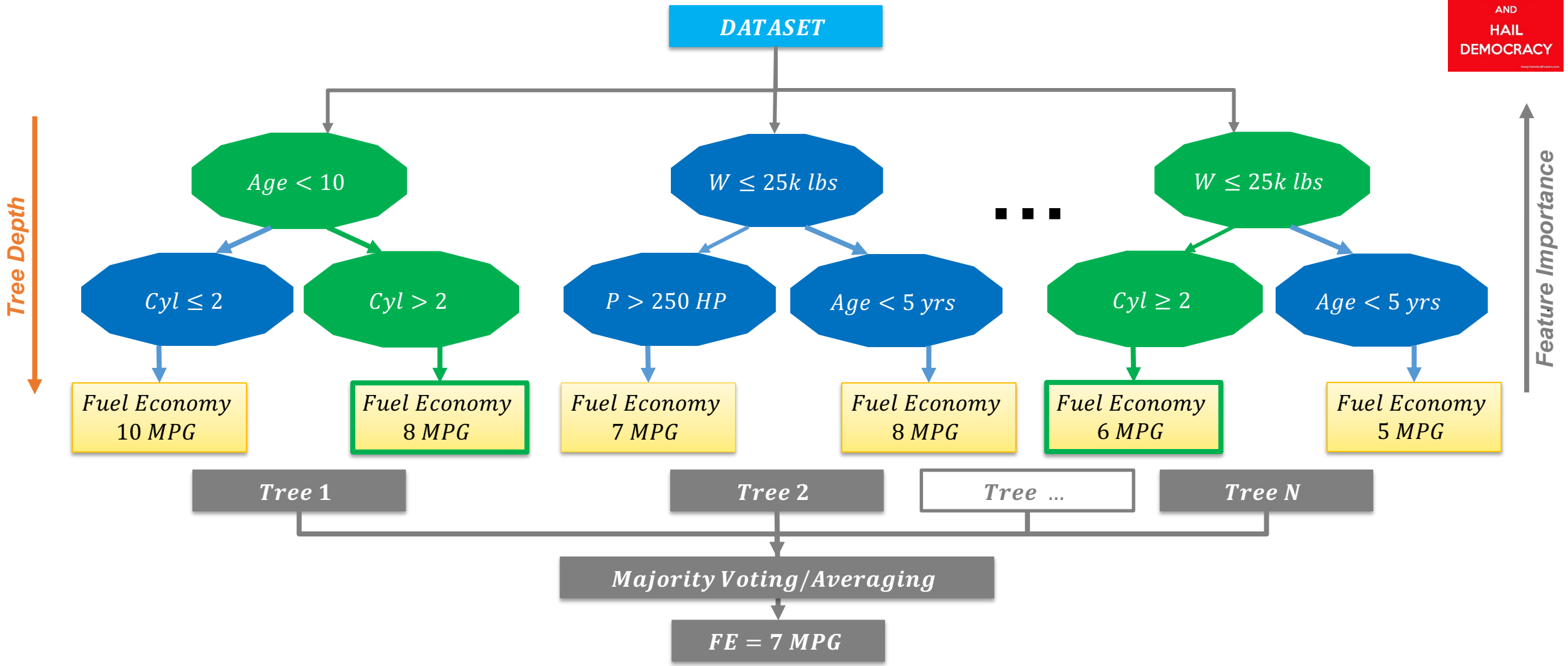
Bias

Correlation w.r.to Fuel Consumption

MACHINE LEARNING MODELS – RANDOM FOREST

THE OHIO STATE UNIVERSITY
CENTER FOR AUTOMOTIVE RESEARCH

$$Predict\ Fuel\ Economy = f(Weight, Engine\ Power, Num\ Cylinders, \& \ Age)$$



RANDOM FOREST - HYPERPARAMETER OPTIMIZATION

THE OHIO STATE UNIVERSITY
CENTER FOR AUTOMOTIVE RESEARCH

$$Energy\ Consumption = f(Drive\ Cycle, Vehicle\ Spec)$$

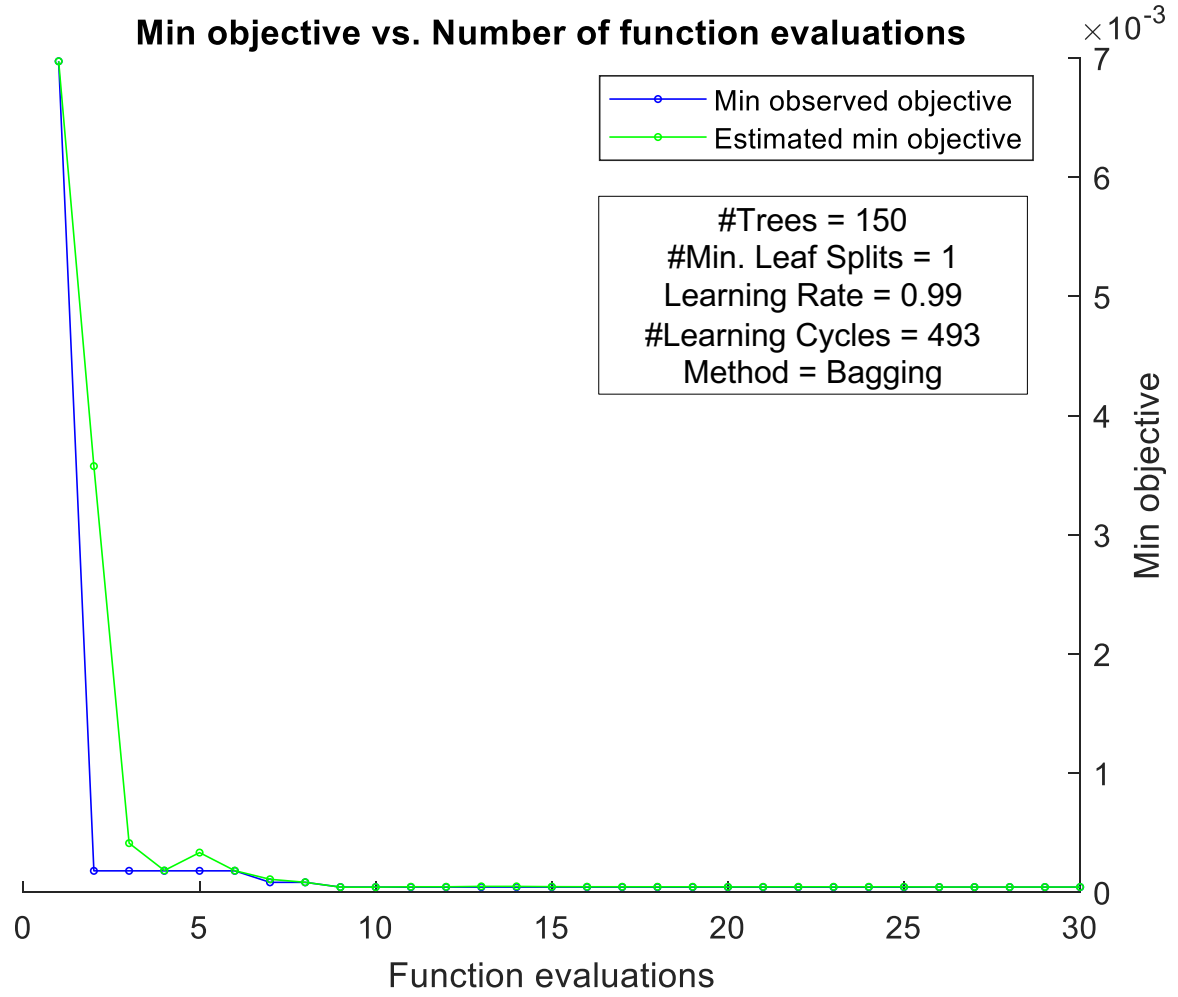
Random Forest Hyperparameters

Tree Depth
Learning Rate
Minimum Leaf Size
Number of Learning Cycles
Minimum Number of Splits
Training Method (Bagging vs. Boosting)

```
Best estimated feasible point (according to models):
    Method      NumLearningCycles     LearnRate     MinLeafSize
    _______     _________________     _________     ___________

    LSBoost           465              0.83153          1

Estimated objective function value = 0.00044326
Estimated function evaluation time = 324.094
```



RANDOM FOREST – PROS/CONS

PROS

Does not require large amount of data

Can handle linear and non-linear relationships well

Requires large number of trees to avoid overfitting

Provides good accuracy

Does not require validation data

Implicitly perform feature selection

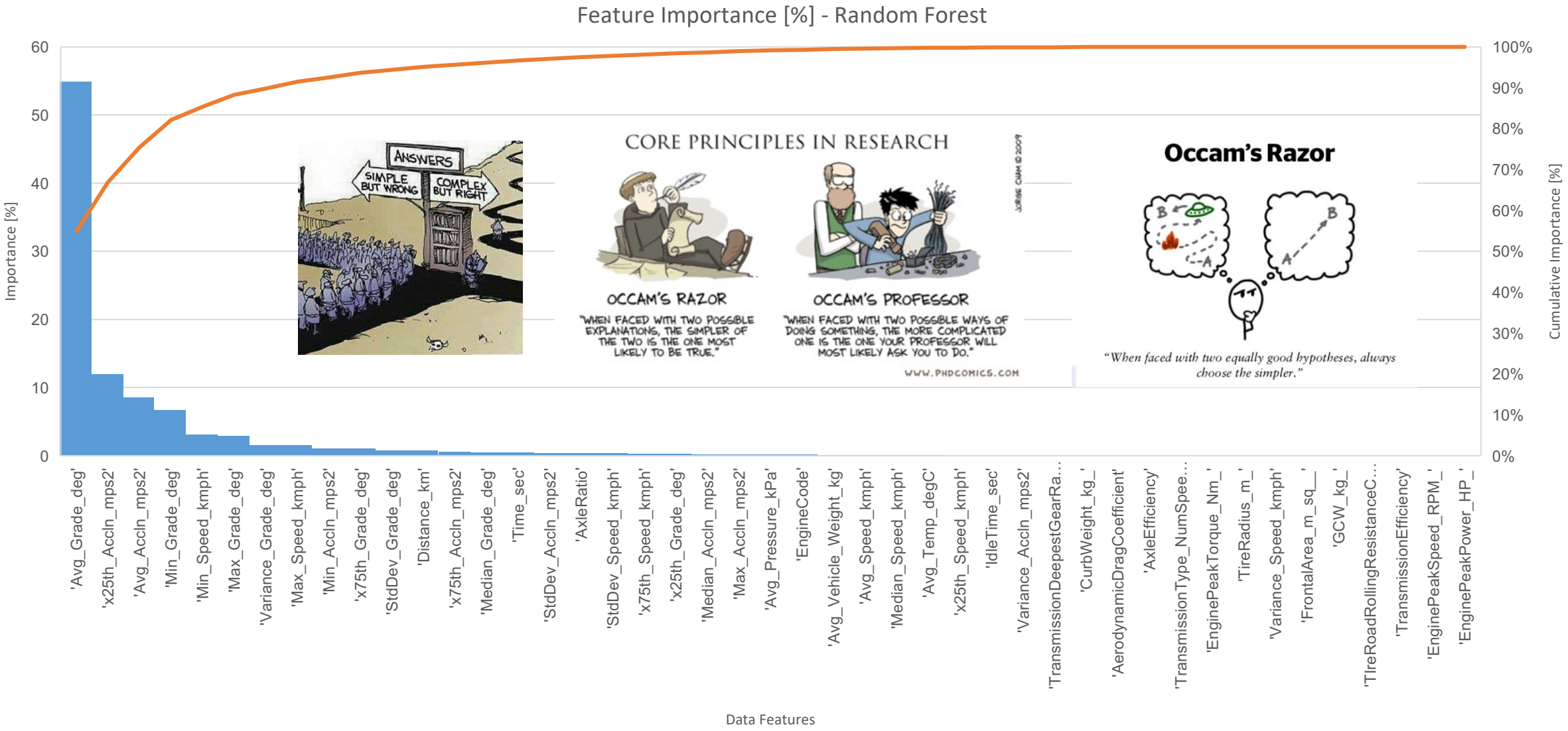
CONS

Cannot Extrapolate

"good for classification but not for regression"

Relatively Computationally Expensive

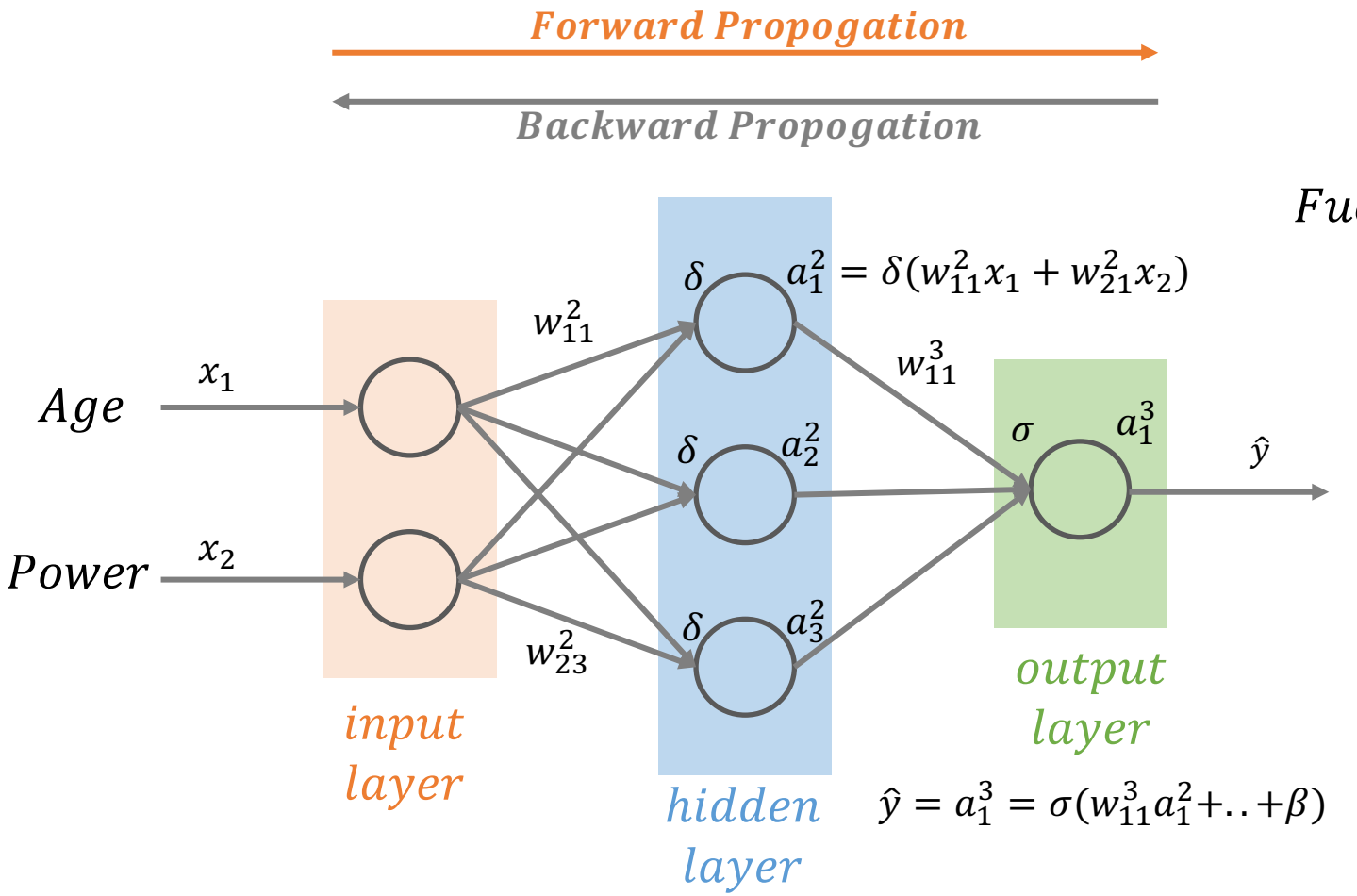
FEATURE IMPORTANCE – RANDOM FOREST



MACHINE LEARNING MODELS – NEURAL NETWORK

$$\boldsymbol{Predict\ Fuel\ Economy = f(Vehicle\ Age, Engine\ Power)}$$

Age	$Power$
1	250
2	250
3	150
…	…
7	250



Fuel Economy

$\hat{y}$	y	$e = y - \hat{y}$
8.5	9	0.5
8	8	0
10.8	10	-0.8
…	…	…
6.8	7	0.2

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2$$

NEURAL NETWORK - HYPERPARAMETER OPTIMIZATION

THE OHIO STATE UNIVERSITY
CENTER FOR AUTOMOTIVE RESEARCH

$$Energy\ Consumption = f(Drive\ Cycle, Vehicle\ Spec)$$

Neural Network Hyperparameters

Learning Rate
Number of hidden layers
Number of neurons in hidden layers
Activation function
Optimizer
Batch Size
Epochs

Results

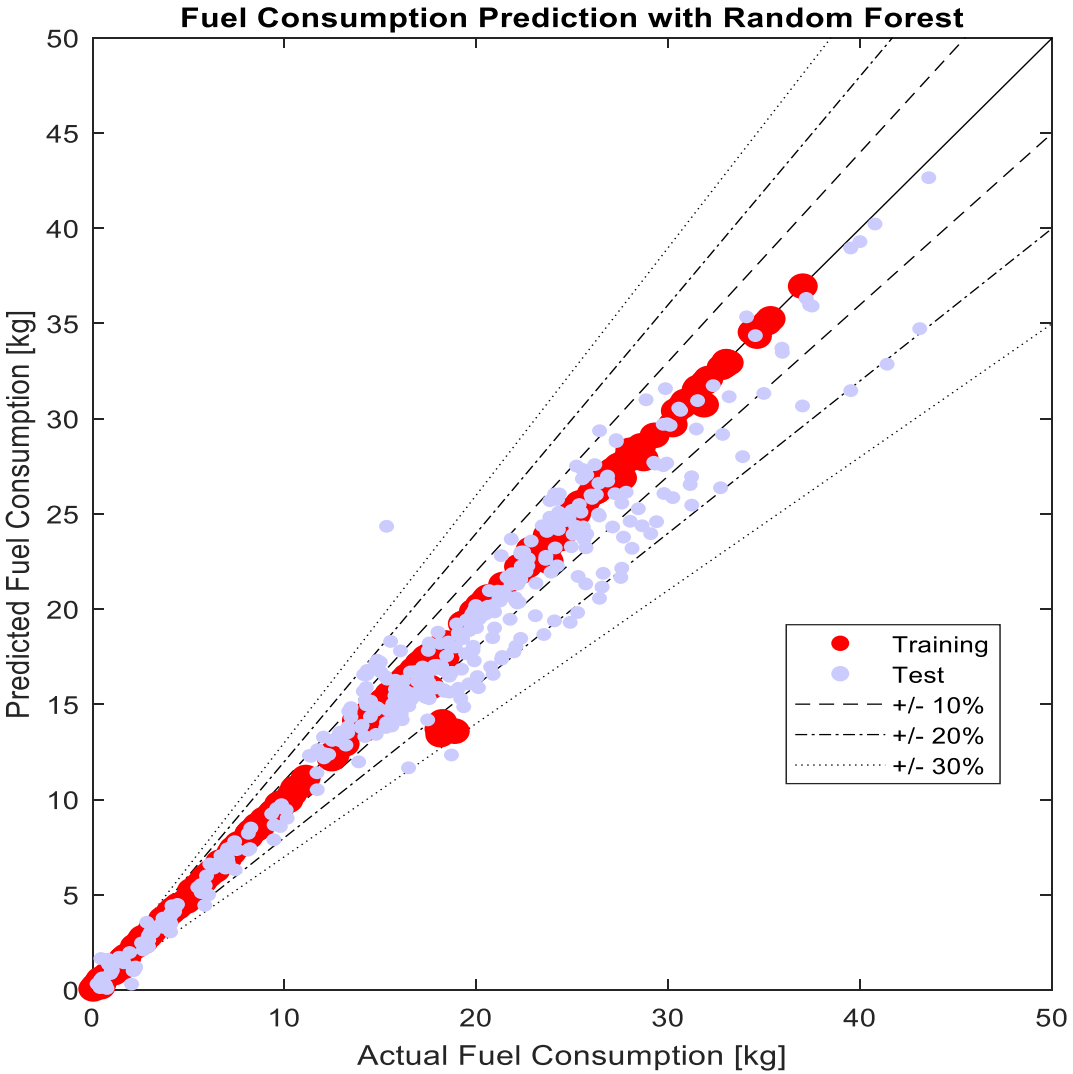
	Samples	MSE	R
Training:	75827	9.93451e-5	9.86243e-1
Validation:	17499	9.84205e-5	9.86383e-1
Testing:	23332	1.00411e-4	9.85941e-1

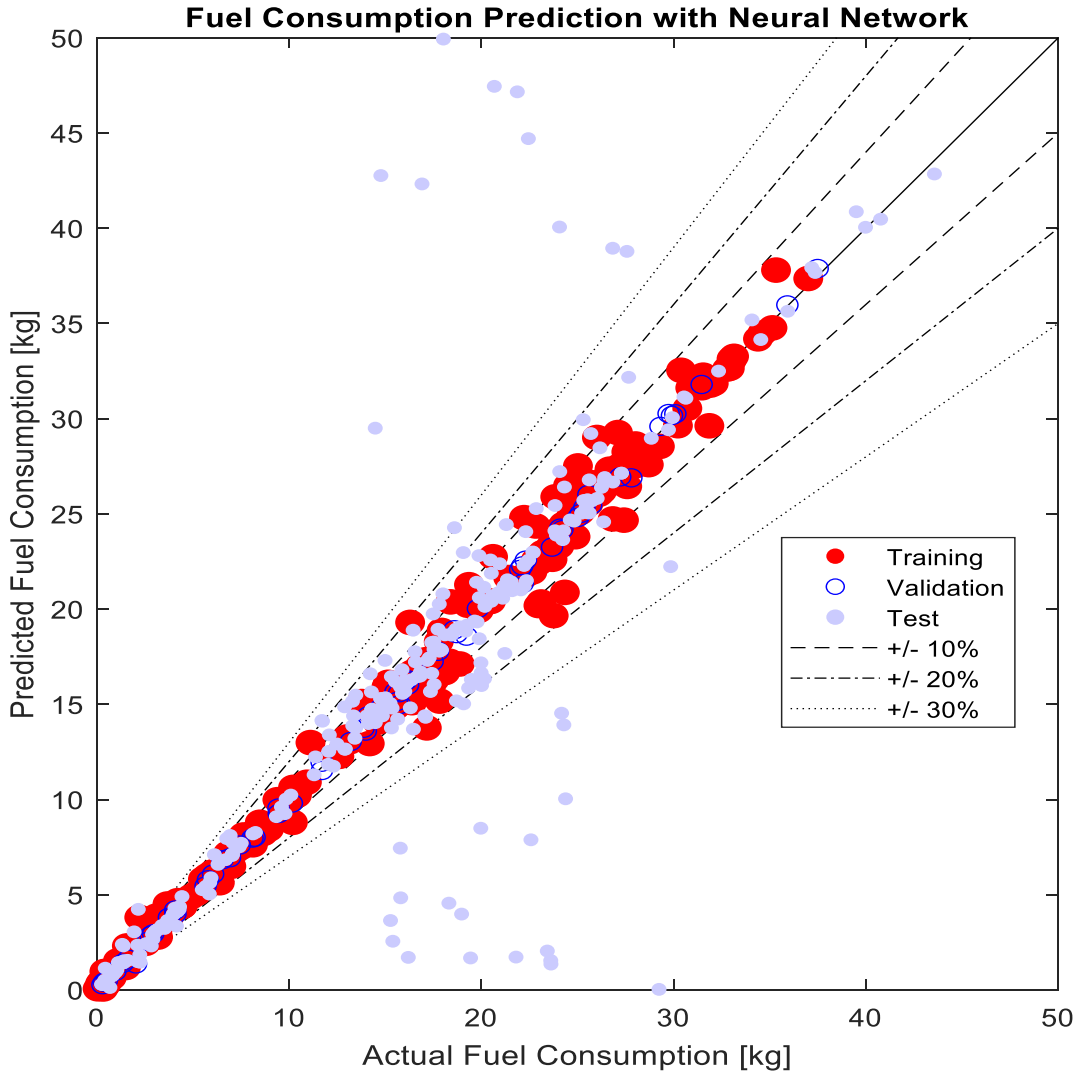
Results

	Samples	MSE	R
Training:	45	1.38169e-2	9.99973e-1
Validation:	10	14.20137e-0	9.11043e-1
Testing:	10	17.67531e-0	8.68514e-1

RANDOM FOREST vs. NEURAL NETWORK - MODEL ACCURACY

THE OHIO STATE UNIVERSITY
CENTER FOR AUTOMOTIVE RESEARCH





400 Drive Cycles with 19 Trucks

I/O OF POWERTRAIN RECOMMENDER SYSTEM

THE OHIO STATE UNIVERSITY
CENTER FOR AUTOMOTIVE RESEARCH

INPUTS

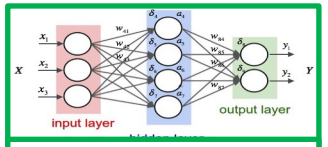
Dealer/Customer Input	Freight Capacity (lbs)
	Top Speed
	Max Road Grade
	Extreme Road Grade

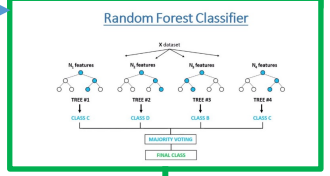
Design Space	EngineType
	EngineMaxSpeed_RPM
	EnginePeakPower_HP
	EnginePeakTorque_Nm
	Trans_Speed
	Trans_Deepest_Gear_Ratio
	Trans_Efficiency
Vehicle Spec	Rear_Axle_Ratio
	RearAxle_Efficiency
	Tire_Radius_inch
	Road_Rolling_Resistance
	Vehicle_Frontal_Area_ft2
	Drag_Coefficient
	Curb_Weight_lbs
	GCVW_lbs

	Time_sec
	Distance_km
	Max_Speed_kmph
	Min_Speed_kmph
	Avg_Speed_kmph
	Median_Speed_kmph
	StdDev_Speed_kmph
	Variance_Speed_kmph
	x25th_Speed_kmph
	x75th_Speed_kmph
	Max_Grade_deg
	Min_Grade_deg
Target	Avg_Grade_deg
	Median_Grade_deg
Drive Cycle	StdDev_Grade_deg
	Variance_Grade_deg
	x25th_Grade_deg
	x75th_Grade_deg
	Max_Accln_mps2
	Min_Accln_mps2
	Avg_Accln_mps2
	Median_Accln_mps2
	StdDev_Accln_mps2
	Variance_Accln_mps2
	x25th_Accln_mps2
	x75th_Accln_mps2
	IdleTime_sec
	Avg_Pressure_kPa
	Avg_Temp_degC

Design Space Filter

Supervised ML Models





Output	TopSpeed_FlatRoad_Eng_MPH
	TopSpeed_FlatRoad_Trans_MPH
	TopSpeed_On_Max_RoadGrade_MPH
	ExtremeRoadGrade_Startability_Accln_ftps2
	Accln0to30MPH_Time_sec
	Accln0to60MPH_Time_sec
	Fuel_Consumption_kg (ML)
	Fuel_Economy_MPG
	Freight_Efficiency_kWhpTonMi
	CO2_Emission_metTon
	Selling_Price_USD*
	Pareto_Optimal (MOOP)

OUTPUTS

WHAT DID WE LEARN TODAY?

How to buy a truck?

How to spec the truck that is best for your operation using a data driven approach?

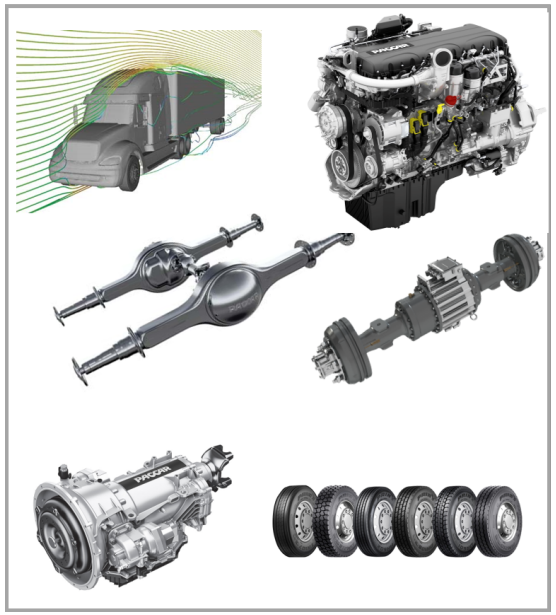
How to analyze telematics data?
Correlation, Bias

What is machine learning and how to build ML models?
Supervised learning algos to predict energy consumption
Random Forest, Neural Network

Fleet composition examples

SUMMARY

Design Space



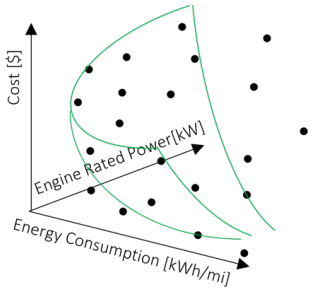
Candidate Powertrains

Customer's Requirements

Powertrain Recommender System

- *Evaluate Performance*
- *Predict (ML) Energy Consumption*
- *Find Non-Dominated Solutions*

BEV



GAS

Thank you

Qadeer Ahmed
ahmed.358@osu.edu

Full Legal Disclaimer: "This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof."