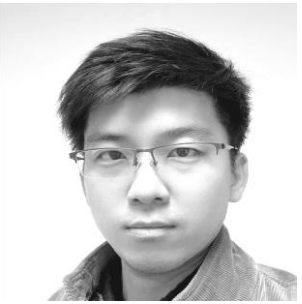
95-791 DATA MINING
FALL 2022 | 10/12/2022

Building Energy and Carbon Emissions Benchmarking Prediction

A Data Mining Study for Infrastructure Energy Management

Devashri Karve
Graduate Student
dkarve@andrew.cmu.edu
CEE Department
Carnegie Mellon University

Tanay Kulkarni
Graduate Student
tskulkar@andrew.cmu.edu
CEE Department
Carnegie Mellon University



Tian Li
Ph.D. Student
tili@cmu.edu
School of Architecture
Carnegie Mellon University

Mentor
Prof. Dr. Gabriela Gongora-Svartzman

Prompt

We catered 30% of city's energy demands in 2022!
Do you have an estimate for 2023?

Private Players

ENERGY

Private Energy Supplier

Private Developer

We must plan energy consumption for a new institution in the city.
Can we get an estimate for the first year?

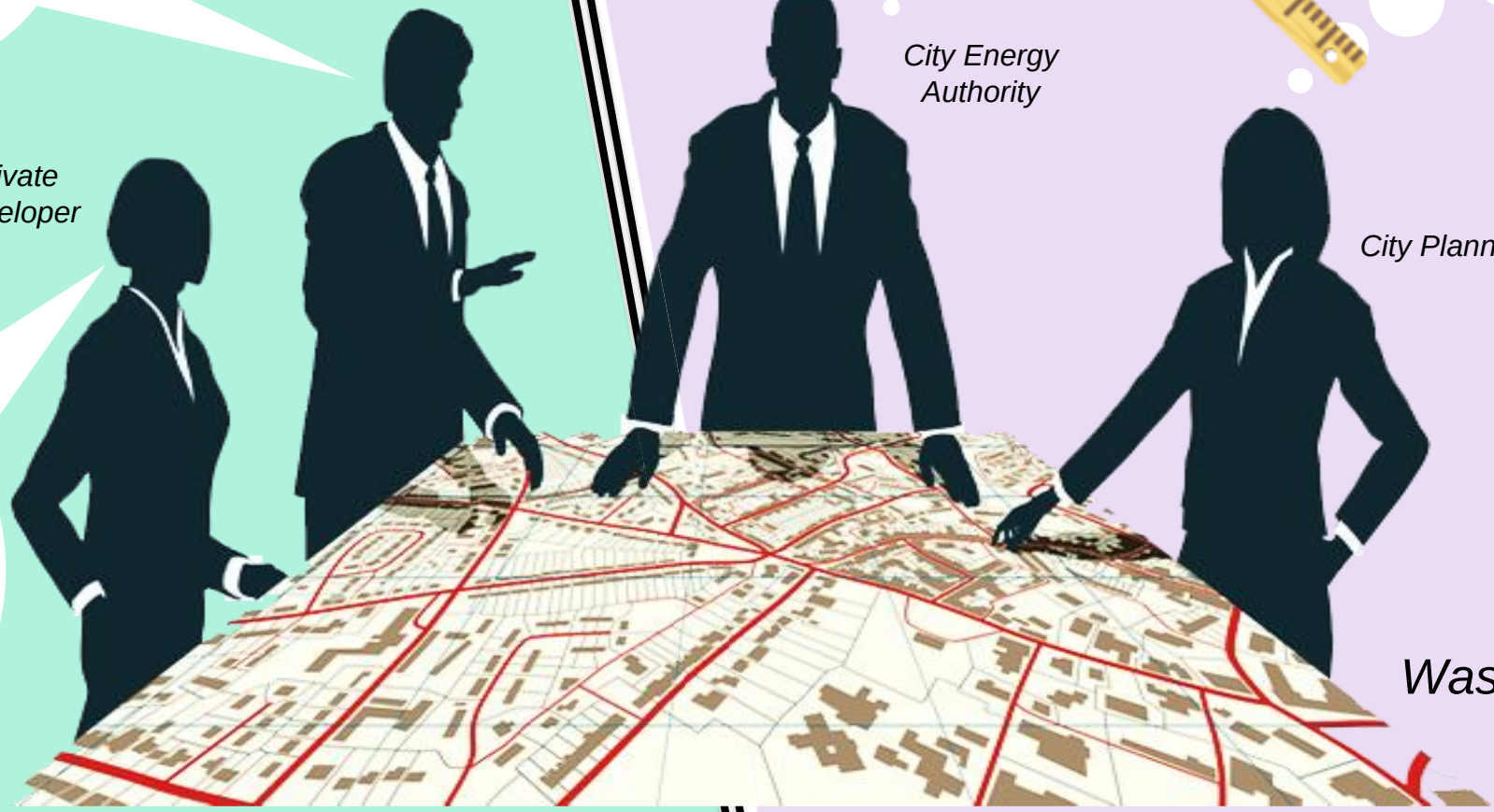
Government

Jeez! Do we have the answers?

We've been capturing all this DATA! It's time we USE IT!
Can we use MACHINE LEARNING?

City Energy Authority

City Planner



Washington, DC

Data availability

year built

primary use

water use

hot water

property status

building area

water score rating

chilled water

monthly gas consumption

steam

carbon intensity

source energy use intensity

site energy use intensity

natural gas

renewable energy

oil use

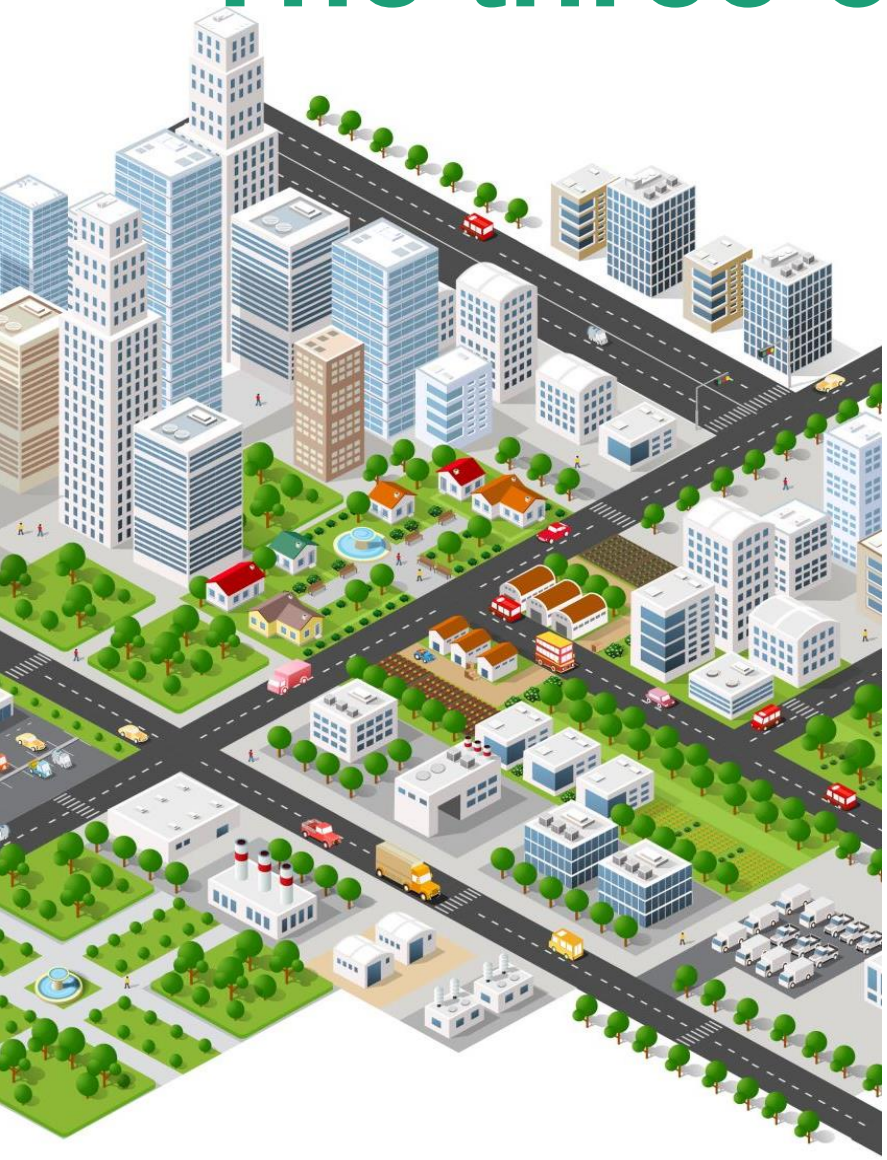
monthly electricity consumption

energy star rating

electricity grid

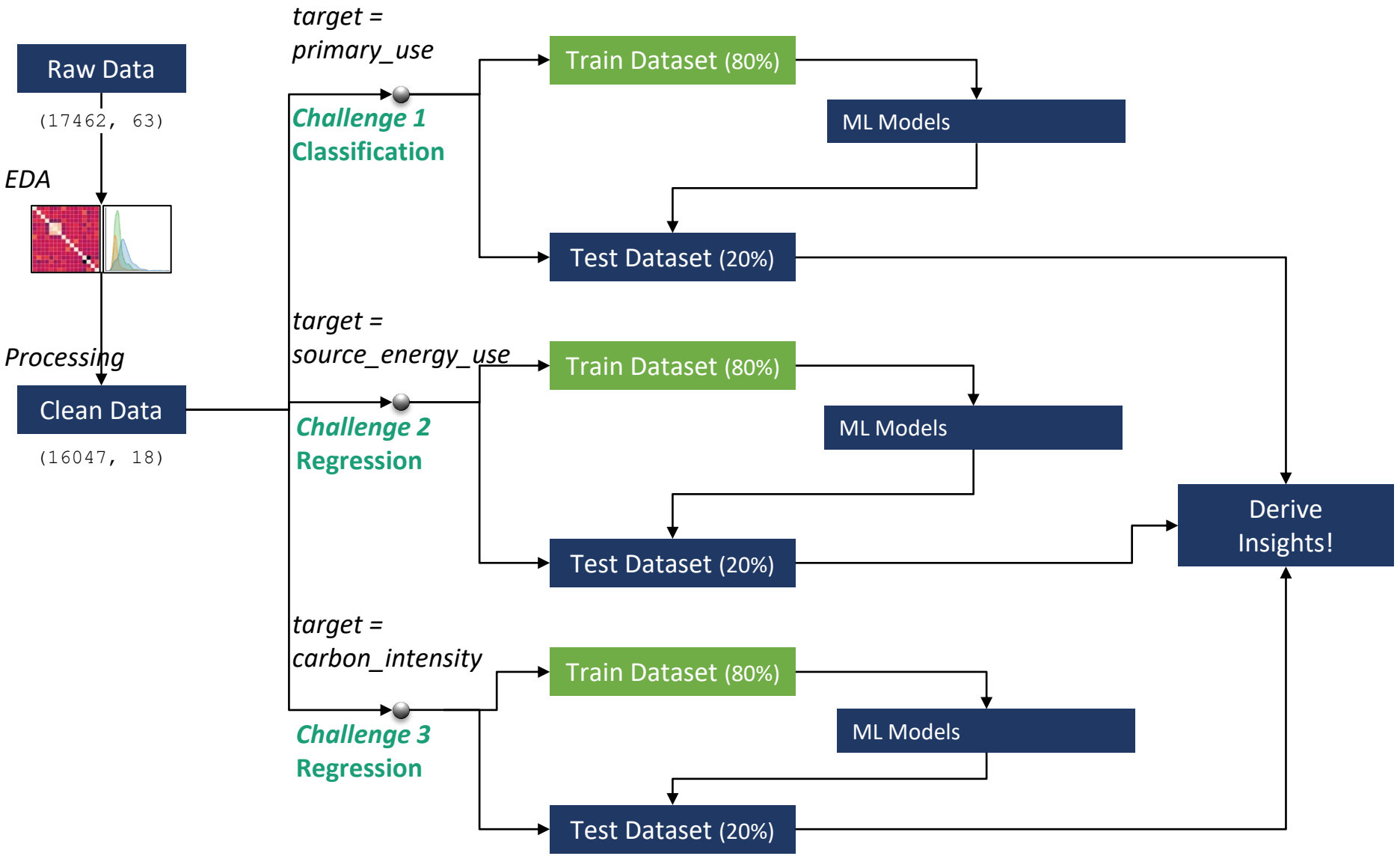
carbon emission

The three challenges

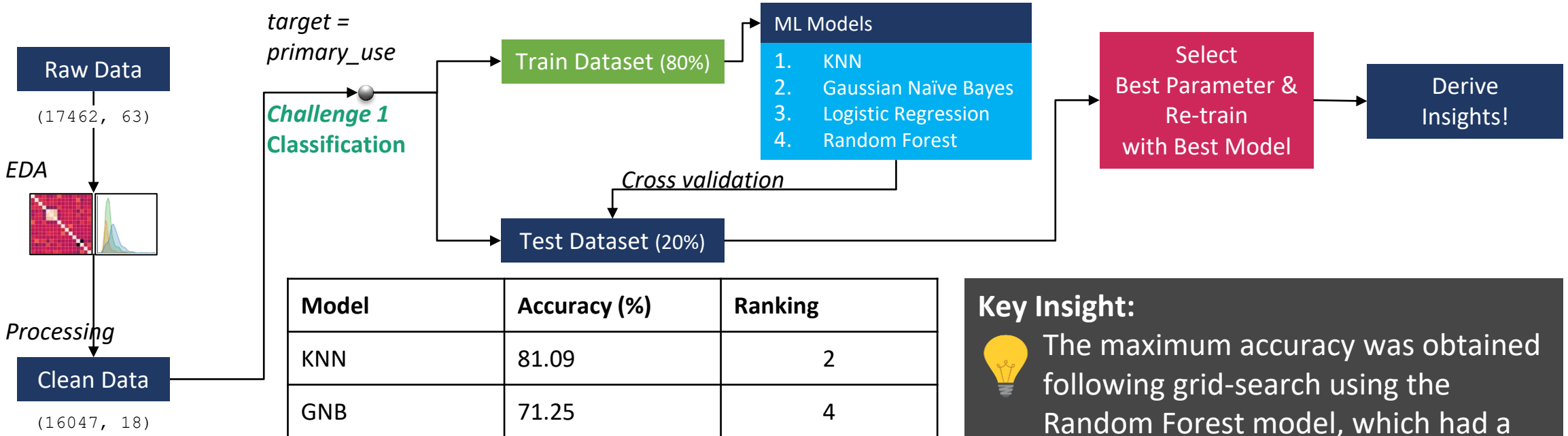


1. Predict primary use of the building
   - 'Residential', 'Commercial', or 'Institutional’
2. Forecasting the building's EUI to describe the building energy use
   - Better prescribe predicted energy consumption.
   - Help the city authorities to plan the city's energy distribution for future scenarios.
3. Predict Carbon Intensity
   - Based on predicted primary use and EUI along with other features

Data mining process adopted:



Challenge 1: Predict ‘primary use’



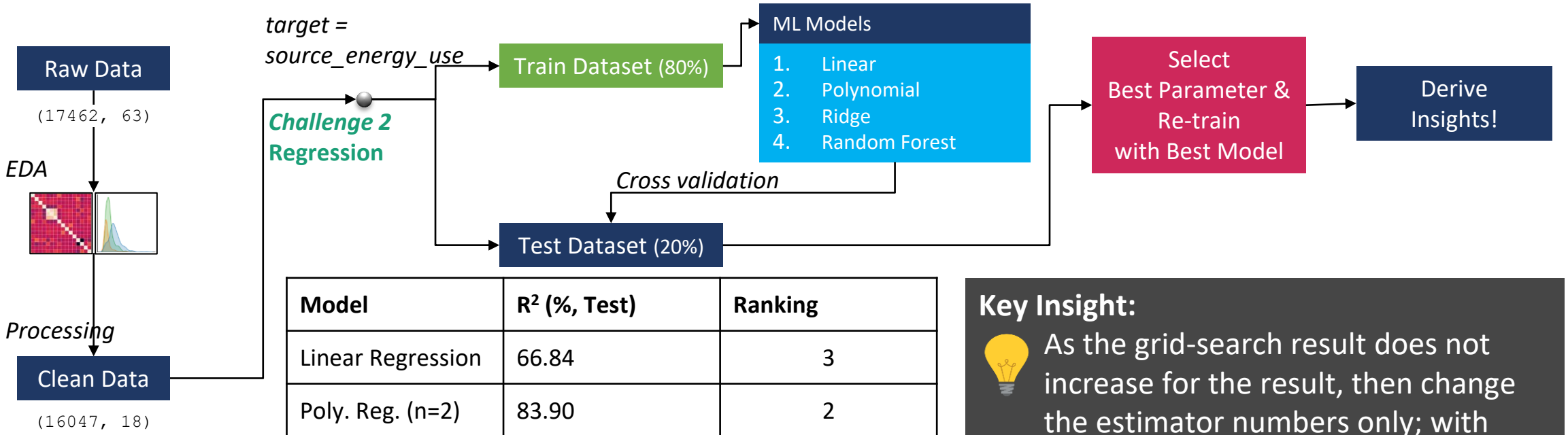
Model	Accuracy (%)	Ranking
KNN	81.09	2
GNB	71.25	4
Logistic Reg	74.77	3
Random Forest	**92.83**	**1**

Hyper-parameters tuned, Refitted on best Model	Accuracy (%)
Random Forest	**92.99**

Key Insight:

The maximum accuracy was obtained following grid-search using the Random Forest model, which had a training accuracy of 92.99%, surpassing that of the Random Forest with default value!

Challenge 2: Predict ‘energy use intensity’



Model	R^2 (%, Test)	Ranking
Linear Regression	66.84	3
Poly. Reg. (n=2)	83.90	2
Ridge	66.84	3
Random Forest	**93.86**	**1**

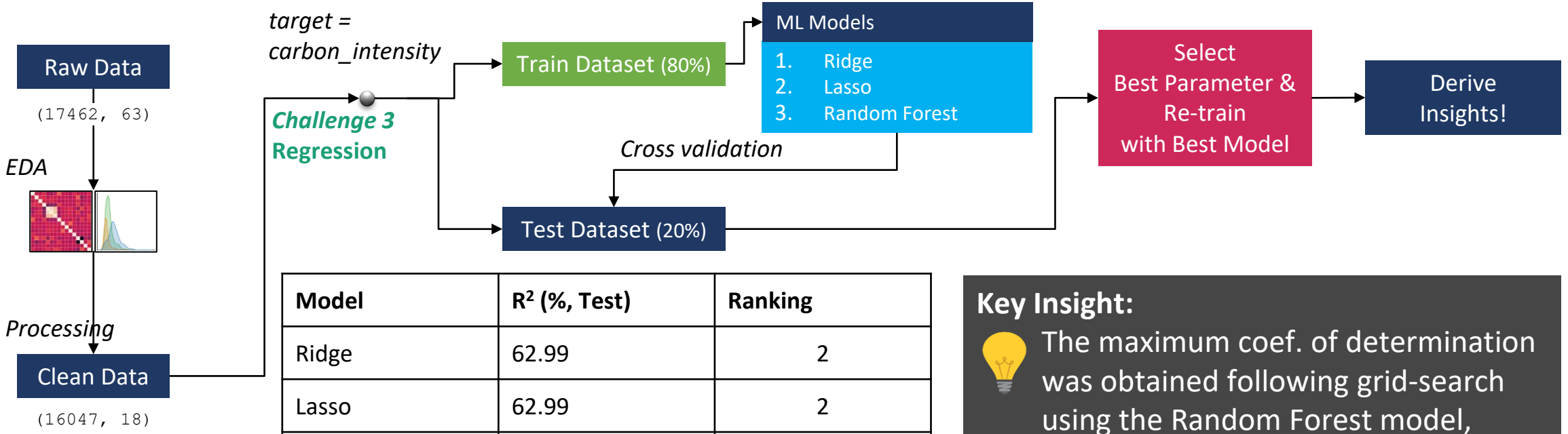
Hyper-parameter tuned, Refitted on best Model	R^2 (%, Test)
Random Forest	**93.33**

Changed estimator number	R^2 (%, Test)
Random Forest	**93.95**

Key Insight:

As the grid-search result does not increase for the result, then change the estimator numbers only; with other default parameters, select the best estimators! Finally, got an improved coefficient of determination of 93.95%.

Challenge 3: Predict ‘carbon intensity’



Model	R^2 (%, Test)	Ranking
Ridge	62.99	2
Lasso	62.99	2
Random Forest	**96.92**	**1**

Hyper-parameter tuned, Refitted on best Model	R^2 (%, Test)
Random Forest	**97.28**

Key Insight:

The maximum coef. of determination was obtained following grid-search using the Random Forest model, which had a training r-squared of 97.28%, surpassing that of the Random Forest with default value!

Conclusions

- Random Forest classification predict building primary use type efficiently.
- Random Forest Regression performs the best compared to Linear, Ridge and Polynomial models to predict the source energy use intensity.
- Random Forest Regression with Grid-search can achieve the best result compared to Lasso, Ridge models to predict carbon intensity

Future work

- ‘Repeatability’ in other cities!
- Demographic and meteorological data can also be used in this study.

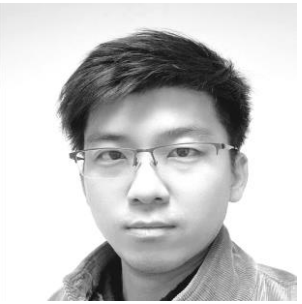
95-791 DATA MINING
FALL 2022 | 10/12/2022

Building Energy and Carbon Emissions Benchmarking Prediction

A Data Mining Study for Infrastructure Management

Devashri Karve
Graduate Student
dkarve@andrew.cmu.edu
CEE Department
Carnegie Mellon University

Tanay Kulkarni
Graduate Student
tskulkar@andrew.cmu.edu
CEE Department
Carnegie Mellon University



Tian Li
Doctoral Student
tili@cmu.edu
School of Architecture
Carnegie Mellon University

Mentor
Prof. Dr. Gabriela Gongora-Svartzman