Statistical and Predictive Modeling II (DATA 2204)
Assignment #1 – *k*-NN Regression (10% of Final Grade)
Professor: Ritwick Dutta

Mr. John Hughes has been reviewing the **EnergyUse-Cooling.csv** dataset and he would like you to create a ***standard and optimized k-NN regression model***.

The dataset contains the following variables:

Independent Variables:

X1 - Relative Compactness
X2 - Surface Area
X3 - Wall Area
X4 - Roof Area
X5 - Overall Height
X6 - Orientation
X7 - Glazing Area
X8 - Glazing Area Distribution

Dependent Variable:

Y- Cooling Load

The Ask:

1. Create a PowerPoint (PPT) presentation that includes the following:
   a. Cover Page (Title, Name (1st and last) and Student Number)
   b. Rational Statement (summary of the problem or problems to be addressed by the PPT) – *2%*
   c. Present the Correlation Heatmap and explain **two (2) insights** – *2%*
   d. Present the Learning Curve for the *k*-NN standard model and explain **two (2) insights** – *2%*
   e. Present and explain **two (2) insights** from the evaluation metrics (i.e. Adj. R^2, MAE, RMSE) for the Optimized *k* -NN Regression model – *2%*
   f. State and explain **two (2) recommendations** for Mr. John Hughes for next steps. – *2%*

Note: Please ensure that all key facts are in your slides and not in the notes section

Hint: Leverage the code from Wk2-kNNReg
Random State = 100 for all section

2. Provide an HTML copy of your python code

Please post your PowerPoint Document(.ppt) and HTML Python Code via assignments under Assignment #1 by 11:59 p.m. on Thursday, February 3rd, 2022

Grading Rubric

	Exemplary (9-10)	Proficient (7-8)	Incomplete (5-6)	Needs Improvement (0-4)
Analysis	Cover Page Complete Rational Statement is complete with supporting details Correlation Map presented with two (2) insights presented with explanation/ justification Learning Curve presented with two (2) insights presented with explanation/ justification Two (2) insights from the evaluation metrics presented and fully evaluated from the Optimized k-NN regression model	Cover Page Complete Rational Statement is complete with high-level supporting details Correlation Map presented with three (3) insights presented with high-level explanation/ justification Learning presented with three (3) insights presented with high-level explanation/ justification Three (3) insights from the evaluation metrics are presented with high-level evaluations for the Optimized k-NN regression model	Cover Page Incomplete Rational Statement is complete with missing supporting details Correlation Map presented with less than three (3) insights and/or Missing explanation/ justification Learning Curve presented with less than three (3) insights and/or Missing explanation/ justification Less than three (3) insights from the evaluation metrics are presented and evaluationed for the Optimized k-NN regression model	Cover Page missing Rational Statement missing Correlation Map and insights are missing or incorrect. Learning Curve and/or insights are missing or incorrect. Optimized k-NN regression model metric insights are missing or incorrect
Next Steps	Two (2) recommendations have been identified with detailed explanations.	Two (2) recommendations have been identified with only high-level explanations.	Less than Two (2) recommendations and incomplete explanations.	Recommendations are missing or incorrect.

Note: 50% Grade Penalty for missing Python HTML file