

American University of Sharjah

College of Engineering

NGN 112 Introduction to Artificial Intelligence and Data Science

Spring 2024

Data Description:

Customer lifetime value abbreviated as CLV is defined as the projected revenue from a customer throughout the purchasing lifecycle. It represents the total amount of money a customer is expected to spend in business, or on products, during their lifetime. CLV is an important metric to know as it helps companies in make decisions about how much money to invest in acquiring new customers and retaining existing ones. The provided dataset represents an Auto Insurance company located in the United States. The company is facing issues in retaining its customers and wants to advertise promotional offers for its loyal ones. They are considering CLV as a parameter for this purpose. Full description of the provided data attributes is given the Table below. Your job is to help them in predicting their CLV accurately. Full list of attributes and their description is provided in the Table below.

Attribute	Description
State	Customer state
Coverage	Type of insurance coverage
Education	Customer education level
Emp_Status	Customer employment status
Gender	Customer gender
Income	Customer income
Loc_Code	Customer living area type
M_Status	Customer marital status
M_Prem	Monthly insurance premium
Mo_Claim	Months since last insurance claim
Mo_Policy	Customer months since policy inception
N_Complaints	Number of yet unresolved complaints
N_Policies	Number of insurance policies
P_Type	Type of insurance policy
S_Channel	Sales channel by which the customer was acquired
T_Claims	Total claims amount
V_Class	Vehicle type
V_Size	Vehicle size
CLV	Customer lifetime value

Your task is to build a:

- **Regression Model:** Build a regression model that predicts CLV based on the provided set of customers' attributes. Detailed project requirements are provided in the term project assignment document.