

# **CPSC 322 Project Proposal:**

**Team Members:** Cole Elfstrom and Cam Chetcuti

**Project Name:** Car Sales Classifier

## **Dataset Information:**

<https://www.kaggle.com/datasets/doaaalsenani/usa-cers-dataset>

Our dataset is from kaggle and is already very clean. It contains 13 attributes in the table from brand of car to state sold in. We are going to try to predict the sale price of the vehicle based on the other attributes using our various classifiers.

## **Implementation/Technical Merit:**

As mentioned before the data is already very clean and should not require much preprocessing besides discretization on the mileage attribute and maybe regionizing the state it is sold in if thats possible (could be fun). We also will focus on which attributes truly determine the eventual sale price, as vin number and other factors probaby have no effect which we will find out with different feature selections.

## **Potential Impact of the Results:**

The results will provide a resource for determining how much your car is worth, and whether or not you are getting a fair price based on other cars similar to yours. Normally a knn classifier is used for this but it will be interesting to see if other classifiers are any better.

The stakeholders of this project are obviously dealers as they actively sell the cars, but more importantly the owners and buyers of the cars as they have a more transparent relationship now.