Imagen que contiene dibujo

El contenido generado por IA puede ser incorrecto.

UNIVERSIDAD PERUANA DE CIENCIAS APLICADAS

FACULTAD DE INGENIERÍA

TRABAJO PARCIAL

Ciclo:

2025-1

Curso:

Programación Concurrente y Distribuida (CC65)

Clasificación de Condición de Autos Usados Utilizando Programación Concurrente en Go

Integrantes:

Rodriguez Calla, Jair Stephano

Niño Suarez, David Joaquín

Llanos, Carlos

Profesor:

Jara García, Carlos Alberto

Lima, 13 de mayo de 2025

# Planteamiento del Problema

**Objetivo General**

Desarrollar una aplicación concurrente en el lenguaje Go que clasifique el nivel de riesgo de inversión en vehículos usados (Bajo, Moderado o Alto) utilizando un algoritmo de Árbol de Decisión, sin depender de librerías externas. Esta clasificación se realizará en base a características relevantes de cada vehículo, tales como la marca, el modelo, el año de fabricación, el precio, el kilometraje y su condición actual.

**Problema que se Quiere Resolver**

El mercado de autos usados es altamente competitivo y genera enormes volúmenes de datos. En este contexto, concesionarios, plataformas de compraventa y empresas de renting necesitan herramientas que les permitan evaluar rápidamente si un vehículo representa una buena oportunidad de inversión.

Sin embargo, el proceso de análisis tradicional es manual, lento y propenso a errores subjetivos, lo que puede conducir a decisiones de inversión poco rentables o riesgosas.

Desde el punto de vista computacional, procesar grandes cantidades de información de forma secuencial es ineficiente. Por ello, se plantea resolver este problema aplicando técnicas de programación concurrente en Go, lo que permitirá acelerar la evaluación de vehículos a gran escala, aprovechando múltiples núcleos de procesamiento.

**Utilidad Práctica**

La solución automatiza la evaluación del riesgo de inversión de vehículos, lo cual es de gran utilidad para:

* Plataformas de compraventa de autos online.
* Concesionarios que buscan maximizar la rentabilidad de sus inventarios.
* Empresas de logística o renting que deben tomar decisiones informadas sobre la compra de flotas de vehículos.

Al clasificar de forma automática y eficiente el nivel de riesgo de inversión, estas organizaciones podrán:

* Tomar decisiones financieras más acertadas.
* Reducir el riesgo de adquirir vehículos con baja rentabilidad.
* Mejorar la planificación de inventarios y las estrategias de reventa.
* Decisión Automatizada

La aplicación clasifica cada vehículo en uno de los siguientes niveles de riesgo de inversión:

* Bajo Riesgo: Alta probabilidad de rentabilidad en la reventa o uso.
* Moderado Riesgo: Rentabilidad incierta, sujeta a factores adicionales.
* Alto Riesgo: Alta probabilidad de pérdida económica si se adquiere.

Esta clasificación se realiza mediante un Árbol de Decisión implementado en Go, utilizando reglas de negocio derivadas de fuentes reconocidas como Kelley Blue Book, Edmunds y MMR, y optimizadas para ejecución concurrente.

# Explicación del dataset Seleccionado

Para el desarrollo de este trabajo se seleccionó el dataset “Synthetic Car Sales Dataset (Over Million Records)”, disponible en la plataforma Kaggle (<https://www.kaggle.com/datasets/jayavarman/synthetic-car-sales-dataset-over-million-records>). Este conjunto de datos fue generado de manera sintética y contiene más de un millón de registros, cumpliendo con el requisito mínimo establecido en la rúbrica del curso.

El dataset simula información transaccional de ventas de autos a nivel mundial, con un enfoque en datos realistas generados para análisis de datos, modelado predictivo y pruebas de algoritmos de aprendizaje automático. Su estructura diversa lo convierte en una base ideal para la implementación de algoritmos de clasificación y la evaluación de estrategias de inversión en vehículos.

**Justificación de la Elección**

Este dataset fue elegido debido a:

* Su gran volumen de datos, permitiendo probar la eficiencia de la aplicación concurrente.
* La diversidad de variables disponibles, ideales para evaluar el riesgo de inversión de vehículos.
* Su carácter sintético, que evita problemas de privacidad y permite su uso libre en entornos académicos.

En el contexto de este trabajo, se utilizan principalmente las variables:

* Brand, Model, Year, Price, Mileage, Country como variables predictivas.
* Riesgo\_Inversion como la etiqueta de clasificación (Bajo, Moderado, Alto).

El resto de las columnas fue descartado durante la fase de preprocesamiento por no aportar valor predictivo directo.

**Descripción de las columnas**

|  |  |  |
| --- | --- | --- |
| Columna | Tipo de dato | Descripción |
| Brand | String | Marca del vehículo (ej. Toyota, Ford, BMW). |
| Model | String | Modelo específico del vehículo (ej. Camry, Mustang, 3 Series). |
| Year | Integer | Año de fabricación del vehículo (rango: 2000–2024). |
| Price | Float | Precio del vehículo en USD (entre $5,000 y $80,000). |
| Mileage | Integer | Kilometraje total del vehículo (entre 0 y 200,000 millas). |
| Color | String | Color exterior del vehículo (ej. Red, Blue, Black). |
| Condition | String | Estado general del vehículo: “New”, “Used” o “Certified Pre-Owned”. Esta es la variable objetivo (target) que se desea predecir. |
| First Name | String | Nombre del comprador. |
| Last Name | String | Apellido del comprador. |
| Address | String | Dirección del comprador. |
| Country | String | País donde se encuentra el comprador (ej. USA, Canadá, Alemania). |