**FACULTAD DE INGENIERÍA**

**ESCUELA DE INGENIERIA DE COMPUTACIÓN Y SISTEMAS**



**PROYECTO FINAL**

**MainFrame 1**

**INTEGRANTES:**

Azabache Medina, Jean Pierre

Patiño Hermoza, Ze Carlos

**DOCENTE:**

Cueva Chávez, Walter Manuel

**Trujillo, 2020**

**Descripción del caso de estudio**

En este proyecto, analizamos un conjunto de datos de transacciones con tarjetas de crédito realizadas durante un período de dos días. El conjunto de datos contiene 284.807 transacciones, de las cuales 492 (0,17%) son fraudulentas.

Cada transacción tiene 30 características, todas ellas numéricas. Las características V1, V2, ..., V 28 son el resultado de una transformación PCA, para poder proteger la confidencialidad por ello la información básica sobre estas funciones no está disponible. La función Tiempo contiene el tiempo transcurrido desde la primera transacción, y la función Monto contiene el monto de la transacción. La variable de respuesta, *Clase*, es 1 en el caso de fraude y 0 en caso contrario.

Objetivo en este proyecto es construir modelos para predecir si una transacción con tarjeta de crédito es fraudulenta. Intentaremos un enfoque de aprendizaje supervisado. También crearemos visualizaciones para ayudarnos a comprender la estructura de los datos y descubrir patrones interesantes.

**Procedimiento**

Lograr el procedimiento de identificación de un fraude electrónico, primero se realiza la preparación de los datos (muestreo, normalización e imputación) que luego ingresan en el modelo de aprendizaje automático, la cual analizará los datos etiquetados que previamente fueron muestreados para su entrenamiento, segundo creará un modelo de detección basado en su aprendizaje, tercero ingresan los datos de validación, para la comprobación de sus predicciones con los que ya fueron identificados en el dataset.

Finalmente, el modelo esta apto para cumplir su objetivo en el negocio, no obstante, el modelo deberá tener una constante mejora y actualización para aumentar su precisión en las nuevas modalidades de evasión hacia el sistema.

**Conclusiones**

Pudimos identificar con precisión las transacciones fraudulentas de tarjetas de crédito utilizando un modelo de bosque aleatorio, regresión logística, redes neuronales, gracias a estos modelos podemos predecir las transacciones fraudulentas que ocurrirán en los bancos y estos podrán tomar medidas para evitar esto (planes de contingencia).