



## FORMATO PARA LA PRESENTACIÓN DE PROYECTOS DE INVESTIGACIÓN CON EL FINANCIAMIENTO DEL FEDU

### Título del proyecto

**Análisis comparativo de modelos de aprendizaje automático para la predicción de ingresos en adultos utilizando datos del censo de 1994**

### Resumen del Proyecto de Tesis

La investigación realizará el análisis del conjunto de datos de ingresos de adultos extraído del censo en USA, permitirá aplicar y comparar los modelos de aprendizaje automático y predecir personas que tienen mayores ingresos en un año.

### Palabras claves (Keywords)

Machine learning, modelos predictivo, censo, RFE

### Justificación del proyecto

La economía moderna y globalizada es cada vez más exigente, por ende las personas deben adquirir un conjunto de conocimiento y habilidades para obtener fuentes de ingreso por encima del promedio para satisfacer la economía familiar. Para satisfacer esta necesidad, se pretende identificar atributos que juegan un papel fundamental en la determinación del nivel de ingreso anual de una persona.

### Antecedentes del proyecto

Vidya Chockalingam, Sejal Shah y RonitShaw, desarrolla análisis exploratorio para conocer los datos y sus características útiles basándose en la varianza, realiza pre-procesado de datos y posteriormente aplicar diversos modelos de machine learning,  
Chet Lemon, Chris Zelazo y Kesav Mulakaluri, exploran los datos en su valor nominal para comprender las tendencias de los datos demográficos, aplicando modelos para predecir si una persona puede ganar más o menos de USD 50 000.  
Scaling Up the Accuracy of Naive-Bayes Classifiers:a Decision-Tree Hybrid, como parte del análisis de curvas que muestran cómo cambia la precisión a mayor cantidad de instancias para el algoritmo Naive-Bayes

### Hipótesis del trabajo (Es el aporte proyectado de la investigación en la solución del problema)

El análisis comparativo de modelos de aprendizaje automático permitirá diseñar





modelos para la predicción de ingresos en adultos utilizando datos del censo de 1994

### **Objetivo general**

Analizar y comparar modelos de aprendizaje automático para la predicción de ingresos en adultos utilizando datos del censo de 1994

### **Objetivos específicos**

Diseñar modelos predictivos de aprendizaje automático  
Seleccionar parámetros adecuados en el procesamiento del conjunto de datos

### **Referencias (Listar las citas bibliográficas con el estilo adecuado a su especialidad)**

V. Chockalingam, S. Shah, R. Shaw, Income Classification using Adult Census Data, [Online] disponible en:  
<https://cseweb.ucsd.edu/classes/wi17/cse258-a/reports/a120.pdf>.  
C, Lemon, C. Zelazo, K. Mulakaluri, Predicting if income exceeds USD50 000 per year based on 1994 US Census Data with Simple Classification Techniques, [Online] disponible

