

## **Integración de Datos de Encuestas y Censos mediante Modelos de Aprendizaje Automático**

*La expansión de datos de censos utilizando datos de encuestas y técnicas de aprendizaje automático para generar nuevas variables y detalles sobre las bases de datos existentes. Este proyecto permite un análisis extremadamente detallado de temas como la pobreza y las condiciones laborales, superando notablemente los métodos de medición tradicionales.*

**Palabras clave:** Modelado predictivo, integración de datos, análisis de pobreza, condiciones laborales.

### Conocimientos deseables

*Programación en Python (numpy, pandas, scikit-learn, TensorFlow) Estadística aplicada Análisis y visualización de datos Conocimientos básicos de técnicas de machine learning*

### ¿Qué podría aprender quien realice esta tesis?

*Desarrollo y validación de modelos de aprendizaje automático Técnicas de integración de datos y manejo de grandes conjuntos de datos Herramientas avanzadas de análisis y visualización de datos Aplicación de métodos cuantitativos a problemas socioeconómicos*

---

### Dirección de la tesis

*Iglesias, Matias Nehuen  
DC*

Contacto: [matuteiglesias@gmail.com](mailto:matuteiglesias@gmail.com)

Más información en el pdf a continuación.

## Propuesta de tema de tesis

---

### **Integración de Datos de Encuestas y Censos mediante Modelos de Aprendizaje Automático**

**Resumen:** La expansión de datos de censos utilizando datos de encuestas y técnicas de aprendizaje automático para generar nuevas variables y detalles sobre las bases de datos existentes. Este proyecto permite un análisis extremadamente detallado de temas como la pobreza y las condiciones laborales.

**Temas Clave:** Modelado predictivo, integración de datos, análisis de pobreza, condiciones laborales.

---

**Propuesta:** Este proyecto tiene como objetivo principal mejorar la comprensión y análisis de temas socioeconómicos críticos mediante la integración de datos de censos y encuestas con técnicas avanzadas de aprendizaje automático. Utilizando variables compartidas entre ambos conjuntos de datos, se busca expandir las bases de datos censales con información detallada derivada de encuestas.

**Específicamente, la tesis se enfocará en:**

- 1. Desarrollo de Modelos Predictivos:**
  - Aplicar técnicas de aprendizaje automático para predecir variables adicionales en los datos censales a partir de datos de encuestas.
  - Evaluar la precisión y robustez de los modelos predictivos.
- 2. Integración de Datos:**
  - Identificar y utilizar variables comunes entre las encuestas y los censos para la integración de datos.
  - Desarrollar métodos para fusionar de manera efectiva estos conjuntos de datos, mejorando la calidad y la riqueza de la información disponible.
- 3. Análisis Socioeconómico:**
  - Realizar análisis detallados sobre temas como la pobreza y las condiciones laborales utilizando las nuevas variables generadas.
  - Evaluar el impacto de políticas públicas basándose en los datos enriquecidos.

**Conocimientos Deseables:**

- Programación en Python (numpy, pandas, scikit-learn, TensorFlow)
- Estadística aplicada
- Análisis y visualización de datos
- Conocimientos básicos de técnicas de machine learning

**¿Qué podría aprender quien realice esta tesis?**

- Desarrollo y validación de modelos de aprendizaje automático
- Técnicas de integración de datos y manejo de grandes conjuntos de datos
- Herramientas avanzadas de análisis y visualización de datos
- Aplicación de métodos cuantitativos a problemas socioeconómicos

**(co-) Dirección de la Tesis: Matías Iglesias**

**Contacto:** matuteiglesias@gmail.com