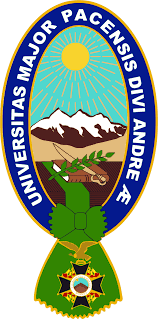
**UNIVERSIDAD MAYOR DE SAN ANDRES**

**FACULTAD DE CIENCIAS PURAS Y NATURALES**

**CARRERA DE INFORMATICA**



**EXAMEN FINAL (PROYECTO)**

**MATERIA:**  INTELIGENCIA ARTIFICIAL

**SIGLA**: INF – 354

**NOMBRE**: UNIV. ITATI CARLA TORREZ MENDEZ

**DOCENTE**: Ph.D. MOISES MARTIN SILVA CHOQUE

**GESTION**: II / 2024

1. **Del dataset trabajado en los exámenes, describa el mismo a detalle.**

**Descripción del Dataset: Breast Cancer Data Set**

* **Contexto del Dataset**

Este dataset contiene características extraídas de imágenes digitalizadas de células mamarias de pacientes diagnosticados con cáncer. Fue diseñado para ayudar en la detección y diagnóstico de cáncer de mama mediante la identificación de patrones en los datos que permitan distinguir entre tipos de tumores.

* **Características principales del Dataset**
* **ID**: Identificador único de cada paciente.
* **Diagnosis**: Etiqueta que indica el tipo de cáncer:
  + **M**: Maligno (tumor que tiende a crecer rápidamente y puede invadir tejidos cercanos).
  + **B**: Benigno (tumor no cancerígeno, crecimiento lento y sin invasión de otros tejidos).
* **Características visuales promedio del tumor** (variables numéricas):
  + radius\_mean: Promedio del radio de las células.
  + texture\_mean: Uniformidad de las células.
  + perimeter\_mean: Promedio del perímetro de las células.
  + area\_mean: Promedio del área de las células.
  + smoothness\_mean: Regularidad de las células.
  + compactness\_mean: Grado de compactación de las células.
  + concavity\_mean: Grado de concavidad de las células.
  + concave\_points\_mean: Cantidad promedio de puntos cóncavos en el contorno de las células.
* **Propósito del Dataset**

El conjunto de datos puede ser utilizado para:

* Entrenar modelos de aprendizaje supervisado que clasifiquen automáticamente tumores en malignos o benignos.
* Evaluar y validar algoritmos predictivos para el diagnóstico temprano de cáncer de mama.
* Analizar los factores visuales que influyen en el diagnóstico, contribuyendo a la comprensión del comportamiento de diferentes tipos de tumores.

Además de los valores promedio mencionados, el dataset también puede incluir métricas relacionadas con rangos de valores, como desviaciones estándar o máximos/mínimos de las mismas características. Esto enriquece la capacidad del dataset para capturar detalles relevantes.

1. **Objetivo de Investigación**

El objetivo principal de este estudio es **desarrollar y evaluar un modelo de clasificación supervisado capaz de predecir el diagnóstico (maligno o benigno) de pacientes con base en características visuales de sus tumores**. A partir de este análisis, se busca:

1. Mejorar la precisión del diagnóstico médico mediante herramientas automáticas.
2. Identificar qué características visuales tienen mayor impacto en la clasificación.
3. Reducir el margen de error humano en diagnósticos médicos críticos.