



<b>Materia:</b> Minería de Datos I	<b>Código:</b>
<b>Créditos:</b> 3	<b>Carga horaria total:</b> 48 hs
<b>Departamento:</b>	<b>Modalidad:</b> Virtual <b>Año:</b> 2023
<b>Carrera de:</b> Tecnicatura Superior en Ciencia de Datos	
<b>Plan de estudios:</b>	

### 1. Presentación:

En el contexto actual, la Minería de Datos ha emergido como una disciplina esencial en el mundo de la informática y la toma de decisiones. Con la creciente digitalización de la sociedad y la acumulación masiva de datos en prácticamente todos los sectores, la habilidad para extraer conocimientos valiosos y tomar decisiones basadas en datos se ha convertido en un activo estratégico para empresas, instituciones y profesionales en diversos campos.

El flujo constante de información, desde redes sociales y sistemas de salud hasta transacciones financieras y dispositivos conectados, ha creado una enorme reserva de datos. Sin embargo, estos datos por sí solos son inabarcables y a menudo caóticos. Aquí es donde la Minería de Datos desempeña un papel fundamental al convertir estos datos en información significativa y accionable.

A continuación se mencionan algunas de las muchas utilidades de la Minería de Datos:

- **Descubrimiento de Patrones y Tendencias:** La Minería de Datos permite descubrir patrones ocultos y tendencias en grandes conjuntos de datos. Por ejemplo, en el comercio electrónico, se pueden identificar patrones de compra de los clientes para personalizar recomendaciones y mejorar la retención.
- **Toma de Decisiones Informada:** Los conocimientos extraídos de la Minería de Datos brindan una base sólida para la toma de decisiones informadas. En medicina, por ejemplo, los análisis de datos pueden ayudar a diagnosticar enfermedades de manera más precisa y anticipar brotes de enfermedades.

- **Optimización de Procesos:** Las organizaciones pueden optimizar sus procesos y operaciones mediante el análisis de datos históricos y en tiempo real. En la industria manufacturera, se pueden identificar ineficiencias y realizar ajustes para mejorar la producción.
- **Personalización y Experiencia del Usuario:** La Minería de Datos permite crear experiencias personalizadas para los usuarios. En el ámbito de los servicios en línea, como la música en streaming, los algoritmos de recomendación utilizan datos de preferencias anteriores para ofrecer contenido relevante.
- **Innovación y Desarrollo:** La comprensión de los patrones de los datos puede inspirar nuevas ideas y oportunidades de negocio. Las startups pueden usar la Minería de Datos para identificar nichos no atendidos en el mercado.
- **Predicción y Planificación:** Con modelos de pronóstico basados en datos históricos, es posible predecir eventos futuros con cierto grado de precisión. En la logística, por ejemplo, se pueden anticipar demandas de envío y planificar rutas eficientes.
- **Investigación Científica:** En áreas como la genómica y la climatología, los análisis de datos a gran escala permiten realizar investigaciones más exhaustivas y descubrir relaciones complejas.

Este curso proporciona una base sólida en conceptos teóricos y habilidades prácticas para permitir a los estudiantes abordar problemas del mundo real utilizando técnicas de Minería de Datos. Además, fomenta la comprensión de la ética y responsabilidad en la aplicación de la Minería de Datos en diversos contextos. Los estudiantes adquirirán las herramientas necesarias para explorar datos, tomar decisiones informadas y utilizar conocimientos ocultos para resolver problemas complejos en una variedad de campos.

En resumen, los conocimientos en Minería de Datos se han convertido en una herramienta indispensable para comprender y aprovechar el vasto flujo de información en la sociedad moderna. Capacitar a los estudiantes en esta disciplina no solo les brinda habilidades valiosas para el mercado laboral, sino que también contribuye al avance tecnológico, la toma de decisiones informadas y la mejora de procesos en una amplia gama de industrias y disciplinas.

## **2. Objetivos de aprendizaje:**

Al finalizar el curso, los estudiantes serán capaces de:

- Comprender los conceptos fundamentales de la Minería de Datos y ser capaz de identificar y aplicar las etapas esenciales del proceso de Minería de Datos en conjuntos de datos reales.
- Realizar un análisis exploratorio de datos efectivo, utilizando estadísticas descriptivas y técnicas de visualización para identificar patrones, relaciones y valores atípicos en conjuntos de datos.
- Familiarizarse con algoritmos de aprendizaje supervisado, aplicar técnicas de clasificación y ser capaz de evaluar y comparar el rendimiento de diferentes modelos utilizando métricas relevantes.
- Comprender los principios del aprendizaje no supervisado y aplicar algoritmos de agrupamiento para segmentar datos, además de realizar reducción de dimensionalidad y poder interpretar los resultados obtenidos.
- Aplicar técnicas de procesamiento de lenguaje natural para analizar y clasificar textos, realizar preprocesamiento de texto y extracción de características, y explorar las aplicaciones de Minería de Datos en el procesamiento de texto.
- Reconocer la importancia del procesamiento de imágenes en la Minería de Datos, aplicar técnicas de preprocesamiento y extracción de características en imágenes, y explorar aplicaciones de Minería de Datos en imágenes y visión por computadora.
- Identificar aplicaciones prácticas de la Minería de Datos en diversos campos, evaluar las implicaciones éticas y de privacidad, y utilizar los resultados de manera responsable y consciente.

## **3. Contenidos:**

### **Unidad 1: Introducción a la Minería de Datos y Conceptos Fundamentales**

- Introducción a la Minería de Datos: definición, objetivos y aplicaciones.

- Proceso de Minería de Datos: desde la selección de datos hasta la evaluación de modelos.
- Conceptos de conjunto de datos, atributo, instancia y clases.
- Tipos de datos: numéricos, categóricos y mixtos.
- Preprocesamiento de datos: limpieza, transformación y reducción de dimensiones.

Competencias teóricas y prácticas:

- Comprender los fundamentos de la Minería de Datos y sus aplicaciones.
- Identificar las etapas del proceso de Minería de Datos.
- Realizar preprocesamiento básico en un conjunto de datos.
- Diferenciar entre diferentes tipos de datos y atributos.

## **Unidad 2: Exploración de Datos y Visualización**

- Análisis exploratorio de datos: estadísticas descriptivas y gráficos.
- Visualización de datos: diagramas de dispersión, histogramas, gráficos de caja, etc.
- Identificación de patrones y relaciones en los datos.
- Detección de valores atípicos y valores faltantes.

Competencias teóricas y prácticas:

- Realizar análisis exploratorio de datos utilizando herramientas estadísticas y visuales.
- Interpretar patrones y relaciones en los datos.
- Identificar y manejar valores atípicos y valores faltantes.

## **Unidad 3: Modelado Predictivo y Aprendizaje Supervisado**

- Conceptos de aprendizaje supervisado.
- Algoritmos de clasificación: árboles de decisión, regresión logística, k-vecinos más cercanos, etc.
- Evaluación de modelos: precisión, recall, F1-score, matriz de confusión, curvas ROC, etc.
- Técnicas de validación cruzada.

Competencias teóricas y prácticas:

- Comprender el concepto de aprendizaje supervisado.
- Aplicar algoritmos de clasificación en la resolución de problemas.

- Evaluar y comparar el rendimiento de diferentes modelos.

#### **Unidad 4: Agrupamiento y Aprendizaje No Supervisado**

- Introducción al aprendizaje no supervisado.
- Algoritmos de agrupamiento (clustering): k-medias, jerárquico, DBSCAN, etc.
- Reducción de dimensionalidad: Análisis de Componentes Principales (PCA) y t-SNE.
- Interpretación de resultados en el aprendizaje no supervisado.

Competencias teóricas y prácticas:

- Comprender el concepto de aprendizaje no supervisado.
- Aplicar algoritmos de agrupamiento en la segmentación de datos.
- Realizar reducción de dimensionalidad para visualización y mejora del rendimiento.

#### **Unidad 5: Minería de Texto y Procesamiento de Lenguaje Natural**

- Extracción de características de texto.
- Técnicas de preprocesamiento de texto: tokenización, lematización, eliminación de stopwords, etc.
- Análisis de sentimiento y clasificación de textos.
- Aplicaciones de procesamiento de lenguaje natural (NLP).

Competencias teóricas y prácticas:

- Entender cómo se aplica la Minería de Datos en el procesamiento de texto.
- Realizar preprocesamiento de texto y extracción de características.
- Clasificar textos y analizar sentimientos.

#### **Unidad 6: Minería de Datos en Imágenes y Procesamiento de Imágenes**

- Introducción al procesamiento de imágenes y su importancia en la Minería de Datos.
- Representación de imágenes: píxeles, canales de color, formatos de imagen, etc.
- Preprocesamiento de imágenes: filtrado, segmentación y eliminación de ruido.
- Extracción de características visuales: descriptores de textura, forma y color.
- Aplicaciones de la Minería de Datos en imágenes: reconocimiento de objetos, detección de rostros, diagnóstico médico, etc.

Competencias teóricas y prácticas:

- Comprender la importancia del procesamiento de imágenes en la Minería de Datos.
- Familiarizarse con la representación y preprocesamiento de imágenes.
- Aplicar técnicas de extracción de características en imágenes.
- Explorar aplicaciones prácticas de Minería de Datos en imágenes.

#### **Unidad 7: Aplicaciones y Ética en Minería de Datos**

- Aplicaciones prácticas de la Minería de Datos en diferentes campos (salud, comercio, redes sociales, etc.).
- Consideraciones éticas y privacidad en la Minería de Datos.
- Uso responsable de los resultados de la Minería de Datos.
- Tendencias y avances en el campo de la Minería de Datos.

Competencias teóricas y prácticas:

- Identificar aplicaciones prácticas de la Minería de Datos en diversos campos.
- Comprender los aspectos éticos y de privacidad en la Minería de Datos.
- Evaluar las implicaciones éticas y sociales de los resultados obtenidos.

#### **4. Trabajos de laboratorio:**

**No aplica.**

#### **5. Referencias:**

##### **Libros de Texto:**

- "Introducción a la Minería de Datos" por Pang-Ning Tan, Michael Steinbach, Vipin Kumar.
- "Data Science for Business" por Foster Provost, Tom Fawcett.
- "Pattern Recognition and Machine Learning" por Christopher M. Bishop.
- "Data Mining: Practical Machine Learning Tools and Techniques" por Ian H. Witten, Eibe Frank, Mark A. Hall.
- "Introduction to Data Mining" por Pang-Ning Tan, Michael Steinbach, Vipin Kumar.

### **Recursos en Línea:**

- Coursera: Plataforma que ofrece cursos en línea sobre Minería de Datos y temas relacionados. Cursos como "Machine Learning" de Stanford University o "Applied Data Science with Python" de University of Michigan pueden ser útiles para complementar el conocimiento.
- Kaggle: Plataforma que ofrece conjuntos de datos y competencias en Minería de Datos. Los estudiantes pueden practicar y aplicar lo aprendido en proyectos reales.
- UCI Machine Learning Repository: Un repositorio en línea de conjuntos de datos utilizados para la investigación en aprendizaje automático y Minería de Datos.
- Libros y Tutoriales en línea: Sitios como "Towards Data Science", "Medium", y "DataCamp" ofrecen tutoriales y artículos sobre temas actuales en Minería de Datos y Aprendizaje Automático.
- Blogs de Expertos: Explorar blogs y sitios web de expertos en el campo, como el blog de Andrew Ng y el blog de Sebastian Raschka, puede proporcionar información actualizada y enfoques prácticos.

### **7. Metodología de enseñanza:**

Este curso se imparte en formato virtual con clases sincrónicas.

### **Carga Horaria**

- 30hs teóricas
- 18hs prácticas
- No aplica hs de laboratorio

### **8. Recursos didácticos:**

En las clases sincrónicas se trabajará con material en diapositivas donde se expondrán los conceptos teóricos. Para los laboratorios se trabajará con plataformas virtuales donde se pueden preparar y resolver los mismos.

Por otra parte, en el aula virtual (campus) se pondrán a disposición los siguientes materiales:

- Actividades de clases y demostraciones
- Ejercicios de laboratorio
- Artículos, conferencias y videos relevantes
- Soluciones de ejemplo

## **9. Modalidad de evaluación y aprobación:**

### **- *Modalidad de evaluación***

Se evaluará la participación de los alumnos en las actividades propuestas. Se espera que interactúen en forma activa y comprometidamente con el/la docente y sus compañeros, y que cumplan con las tareas en tiempo y forma. Además, los alumnos deben aplicar a las siguientes evaluaciones, para que los alumnos puedan demostrar la apropiación de herramientas teóricas/prácticas trabajadas en clase:

- Dos Exámenes Parciales (EP).
- Dos Trabajos Prácticos (TP).

La entrega de los TPs se realizará vía aula virtual en fechas estipuladas. Con posterioridad, los alumnos recibirán una devolución individual o grupal acerca de su desempeño.

### **- *Requisitos de aprobación de la cursada:***

#### **Son requisitos de aprobación del curso:**

- Haber participado en clases.
- Aprobar los dos exámenes parciales con una nota superior a 4 (cuatro).
- Aprobar los dos trabajos prácticos con una nota superior a 4 (cuatro) en función de los criterios mínimos descritos y comunicados en el campus.
- Aprobar examen final con una nota superior a 4 (cuatro).