

	UNIVERSIDAD DE CALDAS	
FORMATO PARA CREACIÓN – MODIFICACIÓN DE ACTIVIDADES ACADÉMICAS		
CÓDIGO: R-1202-P-DC-503		VERSIÓN: 3

## PLAN INSTITUCIONAL DE ACTIVIDAD ACADÉMICA

### I. IDENTIFICACIÓN

Facultad que ofrece la Actividad Académica:	INTELIGENCIA ARTIFICIAL E INGENIERIAS		
Departamento que ofrece la Actividad Académica:	Sistemas		
Nombre de la Actividad Académica:	ANALÍTICA DE DATOS Y MACHINE LEARNING PARA LA OPTIMIZACIÓN INDUSTRIAL		
Código de la Actividad Académica:			
Versión del Programa Institucional de la Actividad Académica (PIAA):	1		
Acta y fecha del Consejo de Facultad para: aprobación _____ modificación _____	Acta No. _____ Fecha: _____		
Programas a los que se le ofrece la Actividad Académica (incluye el componente de formación al cual pertenece):	ESPECIALIZACIÓN EN TECNOLOGÍAS AVANZADAS PARA LA AUTOMATIZACIÓN INDUSTRIAL		
Actividad Académica abierta a la comunidad:	Si _____ No <u>X</u> _____		
Tipo de actividad: Teórica <u>X</u> _____ Teórico - Práctica _____ Práctica _____			
Horas teóricas:	36	Horas prácticas:	16
Horas presenciales:	48	Horas no presenciales:	64
Horas presenciales del docente:	48	Relación Presencial/No presencial:	1:1
Horas inasistencia con las que se repreuba:	5	Cupo máximo de estudiantes:	25
Habitable (Si o No):	SI	Nota aprobatoria:	3

Créditos que otorga:	3	Duración en semanas:	16
Requisitos (escribir los códigos y el nombre de las actividades académicas que son requisitos, diferenciados por programas para el caso de una actividad académica polivalente):			

1. **JUSTIFICACIÓN:** describe las razones por las cuales es importante la actividad académica desde la perspectiva del conocimiento, el objeto de formación del programa, el perfil profesional del egresado(s), y su lugar en el currículo.

En la era de la automatización industrial avanzada, la capacidad de recolectar, procesar y extraer valor de los datos generados por los sistemas de producción es un diferenciador competitivo fundamental. La Analítica de Datos y el Machine Learning (ML) proporcionan las metodologías y herramientas para transformar datos brutos en conocimiento accionable, permitiendo la optimización de procesos, el mantenimiento predictivo, la mejora de la calidad y la toma de decisiones estratégicas. Esta asignatura es esencial para el Especialista en Tecnologías Avanzadas para la Automatización Industrial, ya que le dota de las competencias para aplicar estas técnicas, pasar de un enfoque reactivo a uno proactivo y predictivo, y desbloquear nuevas oportunidades de eficiencia e innovación en los sistemas automatizados que diseña e implementa.

2. **OBJETIVOS:** describe en forma clara lo que se pretende con el desarrollo de la actividad académica.

Desarrollar en los estudiantes las competencias para aplicar principios y técnicas de analítica de datos y machine learning en la solución de problemas y la optimización de procesos en entornos industriales, utilizando datos provenientes de sistemas IIoT y de producción para mejorar la toma de decisiones y la eficiencia operativa.

2. Específicos:
1. Comprender los principios fundamentales de la analítica de datos y el machine learning en el contexto de la industria.
  2. Aplicar técnicas de machine learning para la optimización de procesos y la mejora de la eficiencia productiva.
  3. Utilizar herramientas de analítica de datos para identificar patrones, realizar predicciones y mejorar la toma de decisiones en tiempo real.
  4. Implementar soluciones de mantenimiento predictivo utilizando algoritmos de machine learning.
  5. Desarrollar proyectos que integren analítica de datos y machine learning para la resolución de problemas industriales.

NOTA: en el caso que el Programa Institucional de la Actividad Académica (PIAA) se desarrolle por competencias, es necesario completar los siguientes aspectos, en lugar de objetivos:

---

3. **COMPETENCIAS:** describe actuaciones integrales desde saber ser, el saber hacer y el saber conocer, para identificar, interpretar, argumentar y resolver problemas del contexto con idoneidad y ética.

1. Genéricas

- Pensamiento Analítico y Cuantitativo: Capacidad para abordar problemas de automatización desde una perspectiva basada en datos.
- Resolución de Problemas Basada en Evidencia: Habilidad para utilizar los resultados del análisis de datos para fundamentar decisiones y proponer soluciones de optimización.
- Manejo de Herramientas Tecnológicas: Destreza en el uso de software y lenguajes de programación para el análisis de datos y el machine learning.
- Interpretación y Comunicación de Resultados Complejos: Capacidad para traducir hallazgos técnicos en información comprensible y accionable.

2. Específicas

C4 (Corresponde al RA4 del programa): Utiliza herramientas de analítica de datos y sistemas de aprendizaje automático para la interpretación de información operativa y la toma de decisiones en tiempo real, orientadas a la mejora de procesos industriales. (Sub-competencias específicas de la asignatura):

- Aplicación de Metodologías de Analítica Industrial: Habilidad para aplicar el proceso KDD (Knowledge Discovery in Databases) o CRISP-DM en proyectos de analítica de automatización.
- Preparación de Datos Industriales: Competencia para realizar la limpieza, transformación e ingeniería de características de datos de sensores y procesos para el modelado.
- Modelado Predictivo y Descriptivo: Capacidad para seleccionar, entrenar y validar modelos de machine learning apropiados para problemas de predicción, clasificación o segmentación en la industria.
- Generación de Insights para la Optimización: Habilidad para visualizar e interpretar los resultados de los análisis para generar recomendaciones de optimización en procesos automatizados.

**COMPETENCIAS GENÉRICAS:** describen el conjunto de conocimientos, habilidades, destrezas y actitudes que le permiten al egresado del programa interactuar en diversos contextos de la vida profesional.

**COMPETENCIAS ESPECÍFICAS:** describen los comportamientos observables que se relacionan directamente con la utilización de conceptos, teorías o habilidades, logrados con el desarrollo del contenido de la Actividad Académica.

**4.RESULTADOS DE APRENDIZAJE (RA):** cada asignatura debe contener resultados de aprendizaje particulares, siempre articulados con los generales de cada

---

programa. Los RA de una asignatura pueden tributar a varios RA generales, y no necesariamente hay una relación uno a uno.

**Resultados de Aprendizaje (RA) (Alineados con RA4 del programa y adaptados):**

- RA.ADM.1. Aplicar el ciclo de vida de la analítica de datos (desde la adquisición y preprocesamiento hasta la visualización e interpretación) a conjuntos de datos provenientes de entornos de automatización industrial.
- RA.ADM.2. Seleccionar y aplicar técnicas de preprocesamiento y exploración de datos (limpieza, transformación, análisis exploratorio) para preparar datos industriales para el modelado.
- RA.ADM.3. Implementar y evaluar modelos básicos de machine learning (regresión, clasificación, clustering) utilizando software especializado para resolver problemas industriales como la predicción de fallos, la optimización de parámetros o la detección de anomalías.
- RA.ADM.4. Interpretar los resultados de los modelos de analítica y machine learning en el contexto de un problema de automatización, comunicando los hallazgos de manera efectiva para proponer acciones de mejora.

5. **CONTENIDO:** describe los temas y subtemas que se desarrollarán en la actividad académica. Estos deben estar en perfecta coherencia con los objetivos, método y evaluación de la asignatura y con los perfiles de formación de los programas a los que se ofrece la actividad académica.

**Módulo 1: Introducción a la Analítica y el ML en la Automatización Industrial**

- El valor de los datos en la automatización: De Big Data a Smart Data.
- Tipos de analítica: Descriptiva, Diagnóstica, Predictiva y Prescriptiva.
- Ciclo de vida de un proyecto de analítica de datos (CRISP-DM).
- Fuentes de datos en automatización: Sensores IIoT, SCADA, PLC, MES, CMMS.
- Conceptos fundamentales de Machine Learning:
  - Aprendizaje Supervisado (Regresión, Clasificación).
  - Aprendizaje No Supervisado (Clustering, Detección de Anomalías).
  - Introducción al Aprendizaje por Refuerzo.
- Ética y privacidad en el manejo de datos industriales.

**Módulo 2: Preprocesamiento y Exploración de Datos de Procesos Industriales**

- **Calidad y Preparación de Datos:**
  - Limpieza de datos: tratamiento de valores atípicos, datos faltantes, ruido en datos de series temporales.
- **Transformación de Datos:**
  - Normalización y estandarización.
  - Ingeniería de características (Feature Engineering).
  - Reducción de dimensionalidad (ej. PCA).
- **Análisis Exploratorio de Datos (EDA):**

- Estadística descriptiva para datos industriales.
- Visualización de datos (histogramas, box plots, scatter plots, series de tiempo).
- Herramientas para preprocesamiento y EDA (Python con Pandas, Matplotlib, Seaborn).

### **Módulo 3: Modelado con Machine Learning para la Optimización Industrial**

- **Modelos de Regresión para Predicción:**
  - Regresión Lineal, Regresión Polinómica, Random Forest Regressor.
  - Aplicaciones: predicción de consumo energético, estimación de vida útil restante (RUL).
- **Modelos de Clasificación para Detección y Diagnóstico:**
  - Regresión Logística, Árboles de Decisión, Random Forest, Support Vector Machines (SVM).
  - Aplicaciones: detección de anomalías en tiempo real, clasificación de calidad de producto, diagnóstico de fallos.
- **Modelos de Clustering para Segmentación:**
  - K-Means, Agrupamiento Jerárquico.
  - Aplicaciones: segmentación de regímenes de operación de máquinas.
- **Evaluación de Modelos de ML:** Métricas para regresión (RMSE, MAE, R<sup>2</sup>) y clasificación (Precisión, Recall, F1-Score, Curva ROC). Validación cruzada.
- Software y librerías para modelado (Python con Scikit-learn).

### **Módulo 4: Implementación, Interpretación y Casos de Uso en Automatización**

- Interpretación de los resultados de los modelos de ML en el contexto industrial.
- Comunicación de hallazgos a equipos técnicos y de gestión.
- Consideraciones para el despliegue de modelos (MLOps introductorio).
- **Casos de Uso Detallados:**
  - Mantenimiento Predictivo (PdM).
  - Optimización de Parámetros de Proceso.
  - Control de Calidad Inteligente.
  - Gestión Energética Eficiente.
  - Optimización de la Cadena de Suministro.
- Desafíos en la implementación de analítica y ML en la industria.
- El futuro del ML en la automatización industrial.

6. **METODOLOGÍA:** describe las estrategias educativas, métodos, técnicas, herramientas y medios utilizados para el desarrollo del contenido, en coherencia con los objetivos o competencias.

- **Clases Teórico-Prácticas (Virtual Sincrónico Viernes / Presencial Sábado):** Exposición de los conceptos de analítica y ML, y las metodologías de aplicación, complementada con demostraciones de software y ejemplos industriales.

- **Talleres Prácticos de Programación y Análisis (Presencial Sábado / Entornos Virtualizados):**
  - Ejercicios de preprocesamiento y exploración de datos utilizando Python (Pandas, Matplotlib).
  - Desarrollo y evaluación de modelos de ML utilizando Scikit-learn sobre conjuntos de datos industriales (reales anonimizados o simulados).
- **Aprendizaje Basado en Proyectos (Grupales):** Los estudiantes abordarán un problema industrial utilizando un ciclo completo de analítica, desde la comprensión del problema y los datos hasta la propuesta de un modelo y la interpretación de sus resultados.
- **Análisis de Casos de Estudio:** Estudio crítico de aplicaciones reales de analítica y ML en la industria, evaluando la metodología y el impacto.
- **Uso de Notebooks Interactivos (ej. Jupyter Notebooks, Google Colab):** Para facilitar la experimentación con código, datos y la documentación de los análisis.

7. **CRITERIOS GENERALES DE EVALUACIÓN:** describe las diferentes estrategias evaluativas, con valoraciones cuantitativas y reportes cualitativos, si son del caso, que se utilizarán para determinar si el estudiante ha cumplido con lo propuesto como objetivos o como competencias de la Actividad Académica. Ver reglamento estudiantil y política curricular.

- **Participación y Resolución de Ejercicios en Clase (Virtual y Presencial): 15%**
  - Evaluación de la comprensión de conceptos y la participación en discusiones sobre metodologías.
- **Entregables de Talleres Prácticos (Notebooks de Análisis y Modelado): 35%**
  - Calificación de los cuadernos de trabajo donde se evidencie el preprocesamiento de datos, la implementación de algoritmos de ML, la evaluación de modelos y la interpretación de resultados.
- **Análisis Crítico de Artículos o Casos de Estudio (Individual/Grupal): 20%**
  - Evaluación de la capacidad para analizar y sintetizar información de investigaciones o implementaciones de analítica/ML en la industria.
- **Proyecto Final Grupal (Aplicación de Analítica/ML a un Problema Industrial): 30%**
  - Desarrollo, documentación y presentación de un proyecto donde se aplique el ciclo de vida de la analítica de datos/ML para resolver un problema de automatización industrial.

---

8. **REFERENCIAS BIBLIOGRÁFICAS:** describe los textos guía, manuales, fuentes primarias, páginas de Internet, entre otras, que serán utilizadas para el desarrollo de la Actividad Académica.

- James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). An Introduction to Statistical Learning: with Applications in R. Springer. (También existen versiones y ejemplos en Python).
- Murphy, K. P. (2012). Machine Learning: A Probabilistic Perspective. MIT Press. (Más avanzado, para consulta).
- Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep Learning. MIT Press. (Para introducción a redes neuronales).
- McKinney, W. (2017). Python for Data Analysis: Data Wrangling with Pandas, NumPy, and IPython. O'Reilly Media. (2nd Edition o más reciente).
- VanderPlas, J. (2016). Python Data Science Handbook: Essential Tools for Working with Data. O'Reilly Media.
- Grus, J. (2019). Data Science from Scratch: First Principles with Python. O'Reilly Media. (2nd Edition).
- Aggarwal, C. C. (2015). Data Mining: The Textbook. Springer.
- Documentación oficial de librerías: Pandas, NumPy, Scikit-learn, Matplotlib, Seaborn, TensorFlow, Keras.
- Cursos online y tutoriales de plataformas como Coursera, edX, DataCamp, Kaggle.
- Artículos de conferencias y journals relevantes (ej. IEEE Transactions on Industrial Informatics, Computers in Industry, Journal of Manufacturing Systems)