

BOOTCAMP EN DATA SCIENCE

Híbrido

(9 meses)

DIPLOMADO EJECUTIVO

¡Transforma tu perfil profesional y destaca en el mercado laboral!

- Obtén tres insignias digitales que validan tus habilidades.
- Credenciales verificables con tecnología blockchain, que fortalecen tu perfil profesional.
- Accede a +20 programas en diversas áreas de estudio, diseñados para responder a las demandas actuales de los empleadores



Reconocimiento Diplomado Ejecutivo

Obtén un Constancia digital UVM del programa con tecnología blockchain y código de verificación



Reconocimiento Curso IA

Adquiere nuevos conocimientos y desarrolla tus habilidades para el uso práctico y estratégico de las herramientas de IA en tu entorno profesional.



Constancia DC-3 por STPS

Documento avalado por la Secretaría de Trabajo y Previsión Social (STPS) que acredita las habilidades y competencias adquiridas.

Objetivo:

- Conocer el lenguaje técnico de la ciencia de datos para el desarrollo de soluciones y toma de decisiones basadas en el análisis de los modelos matemáticos.

Dirigido a:

- Gerentes, Directores, Empresarios, Coordinadores de Planeación Estratégica y Tomadores de decisión y público en general interesado en el tema.

Reconocimiento:

- Al finalizar tu programa recibirás:
 - **Diploma Digital** con **validez curricular** y **tecnología Blockchain** con código QR y de verificación.
 - Certificado Digital Internacional Blockchain de **Embiz Foundation de *Data Science***.
 - Certificado **Ibaktor DC-3** de Competencias Laborales de ***Data Science***.

¿Por qué UVM?

60 años de experiencia académica, más de 150 programas educativos y más de 180 programas de excelencia a nivel nacional.

Adquieres conocimientos y habilidades esenciales aplicables de manera inmediata a tu actividad profesional.

Los profesores que imparten las Certificaciones y Diplomados siguen un modelo de enseñanza con ejemplos reales, pues cada uno de ellos es experto y reconocido en su campo.

Flexibilidad educativa que te permite estudiar a tu ritmo, a cualquier hora y en cualquier lugar.

Los Diplomados y Certificaciones de UVM enriquecen tu CV y te posicionan como el mejor candidato.

Al estudiar el programa podrás:

Comprender los fundamentos y conceptos de la ciencia de datos como herramienta para la gestión.



Utilizar modelos matemáticos y estadísticos para desarrollo de soluciones de *data science*.



Aplicar herramientas para la toma de decisiones en la organización en la que se encuentre basada en la ciencia de datos, mediante la clasificación, análisis y consolidación de datos.



Interpretar datos para una efectiva presentación y visualización gráfica de datos masivos.



BLOQUES DE APRENDIZAJE

Comprender los fundamentos y el alcance de la ciencia de datos, identificando sus principales campos de aplicación y los tipos de modelos utilizados para resolver problemas en contextos reales.

Introdutorio a la ciencia de datos, programación Python y R

Comprender los principios fundamentales de *big data*, incluyendo sus características clave (las 5 V), los modelos de almacenamiento distribuido, las principales arquitecturas y herramientas tecnológicas asociadas.

Big data y almacenamiento a gran escala

Comprender los principios, estructuras y tipos de redes neuronales utilizadas en la ciencia de datos, incluyendo arquitecturas como redes *feedforward*, convolucionales y recurrentes. Implementar modelos básicos y avanzados mediante herramientas como TensorFlow o PyTorch.

Introducción a redes neuronales

Dominar los principios teóricos y prácticos de la visualización avanzada de datos, integrando conocimientos de precepción visual, narrativa con datos y diseño centrado en el usuario.

Visualización de datos: Looker, Tableau y herramientas avanzadas

Aplicar técnicas estadísticas avanzadas para el análisis, modelado y validación de datos complejos. Comprender y utilizar métodos de inferencia, análisis multivariado, series temporales, estadística bayesiana y diseño de experimentos.

Estadística avanzada y métodos cuantitativos

BLOQUES DE APRENDIZAJE

Utilizar de forma eficiente los lenguajes Python y R en entornos especializados para la manipulación, análisis y visualización de datos. Dominar diferentes bibliotecas como Pandas, Matplotlib, etc y herramientas estadísticas.

Programación: Python y R

Aplicar técnicas de aprendizaje automático supervisado, no supervisado y profundo (*deep learning*) para resolver problemas reales de ciencia de datos. Conocer la importancia de la Inteligencia Artificial en la ciencia de datos.

Machine Learning e Inteligencia Artificial

Aplicar principios de ética y responsabilidad en el desarrollo de modelos de Inteligencia Artificial, incluyendo sesgos, y cumplimiento de regulaciones. Comprender y desarrollar sistemas de recomendación avanzados y técnicas de *forecasting* con redes neuronales recurrentes.

Data science y storytelling avanzado

BLOQUES DE APRENDIZAJE

Introductorio a la ciencia de datos, programación Python y R

1. Introducción a la ciencia de los datos
 - a. Campos de aplicación del *big data*
 - b. Modelos
 - c. Discusión sobre cómo se utilizan los modelos para resolver problemas en diferentes campos
 - d. Aplicaciones específicas de la ciencia de datos
 - i. Detección de fraudes en transacciones financieras
 - ii. Análisis de sentimientos en redes sociales
 - iii. Predicción de demanda y optimización de la cadena de suministro
 - iv. Personalización de la experiencia del cliente
 - v. Diagnóstico y predicción de enfermedades en la industria de la salud
 - vi. Programación Python y R

Big data y almacenamiento a gran escala

1. Fundamentos de *big data*
 - a. Las 5 V del *big data* (Volumen, Velocidad, Variedad, Veracidad, Valor)
 - b. Arquitecturas de almacenamiento distribuido
 - c. CAP Theorem y bases de datos NoSQL
 - d. Data Lakes vs Data Warehouses vs Data Lakehouses
2. Sistemas de almacenamiento distribuido
 - a. Hadoop HDFS: arquitectura y casos de uso
 - b. Apache Cassandra: bases de datos columnares distribuidas
 - c. MongoDB: almacenamiento de documentos a escala
 - d. Apache HBase: bases de datos NoSQL sobre Hadoop
3. Formatos de datos para *big data*
 - a. Apache Parquet: almacenamiento columnar optimizado
 - b. Apache ORC: optimización para *analytics*
 - c. Delta Lake: versionado y ACID en *data lakes*
 - d. Apache Iceberg: evolución de esquemas en *petabytes*
4. Cloud Storage y Data Lakes
 - a. AWS: S3, Redshift, EMR para almacenamiento escalable
 - b. Google Cloud: BigQuery, Cloud Storage, Dataflow
 - c. Azure: Data Lake Storage, Synapse Analytics
 - d. Estrategias de particionado y organización de datos
5. Conceptos de procesamiento en tiempo real
 - a. ¿Qué es *stream processing*? (conceptos básicos)
 - b. Diferencias entre *batch* y *streaming*
 - c. Casos de uso: *analytics* en tiempo real
 - d. Introducción a Apache Kafka y Spark Streaming (solo conceptual)

Introducción a redes neuronales

1. Comprensión de la arquitectura de las redes neuronales
 - a. Introducción a los conceptos fundamentales de las redes neuronales
 - b. Estructura básica de una red neuronal: capas, nodos, conexiones
 - c. Diferentes tipos de arquitecturas de redes neuronales: *feedforward*, recurrentes, convolucionales
 - d. Funciones de activación y cómo se utilizan en las redes neuronales
 - e. Aprendizaje supervisado y no supervisado en las redes neuronales
2. Construcción de redes neuronales y cómo se utilizan en la ciencia de datos
 - a. Implementación de redes neuronales utilizando herramientas de aprendizaje profundo, como TensorFlow o PyTorch
 - b. Preprocesamiento de datos para el entrenamiento de redes neuronales
 - c. Entrenamiento de redes neuronales utilizando diferentes técnicas de aprendizaje automático, como retropropagación o descenso de gradiente estocástico
 - d. Validación y evaluación de redes neuronales
 - e. Optimización de hiperparámetros de las redes neuronales
3. Uso de codificadores automáticos
 - a. Concepto de codificador automático y su papel en la reducción de dimensionalidad
 - b. Tipos de codificadores automáticos: *denoising*, *variational*, *contractive*
 - c. Implementación de codificadores automáticos utilizando herramientas de aprendizaje profundo
4. Modelos de perceptrón y redes neuronales artificiales
 - a. Perceptrón y perceptrón multicapa: estructura y funcionamiento
 - b. Redes neuronales artificiales y su relación con el cerebro humano
 - c. Uso de redes neuronales artificiales para clasificación, regresión y predicción
5. Aprendizaje supervisado y no supervisado
 - a. Concepto de aprendizaje supervisado y no supervisado en las redes neuronales
 - b. Diferencias entre el aprendizaje supervisado y no supervisado
 - c. Ejemplos de aplicaciones de aprendizaje supervisado y no supervisado
6. Algoritmos de aprendizaje y arquitecturas de red
 - a. Algoritmos de aprendizaje automático utilizados en las redes neuronales, como retropropagación, descenso de gradiente estocástico y *backpropagation through time*
 - b. Arquitecturas de red populares, como las redes neuronales convolucionales, recurrentes y *autoencoder*
7. Funcionamiento de las redes neuronales convolucionales y recurrentes
 - a. Concepto de redes neuronales convolucionales y su papel en la clasificación de imágenes
 - b. Funcionamiento de las redes neuronales recurrentes y su papel en el procesamiento de secuencias de datos, como texto o audio
 - c. Ejemplos de aplicaciones prácticas de las redes neuronales convolucionales y recurrentes

Visualización de datos: Looker, Tableau y herramientas avanzadas

1. Teoría de visualización avanzada
 - a. Principios de percepción visual y cognición
 - b. *Grammar of graphics* y teoría de Wilkinson
 - c. *Storytelling* con datos y narrativa visual
 - d. *Design thinking* aplicado a visualización
2. Tableau para *data science*
 - a. Tableau Desktop: desde básico hasta avanzado
 - b. *Calculated fields* y *table calculations* complejas
 - c. *Statistical functions* en Tableau
 - d. Integración con modelos de Python/R
 - e. *Dashboards* para presentar resultados de ML
3. Looker Studio y Google Cloud
 - a. Looker Studio: visualización y modelado de datos
 - b. Conexión con BigQuery y fuentes de Google Cloud
 - c. LookML para modelado semántico
 - d. *Dashboards* colaborativos y embebidos
4. Power BI para científicos de datos
 - a. DAX avanzado para cálculos complejos
 - b. Power BI con Python/R: integración de *scripts*
 - c. *Custom visuals* para resultados de ML
 - d. Power BI Service y colaboración
5. Visualización programática
 - a. D3.js para visualizaciones *web* personalizadas
 - b. Streamlit para aplicaciones interactivas de ML
 - c. Dash de Plotly para *dashboards* científicos
 - d. Shiny para aplicaciones analíticas en R
6. Proyectos de visualización científica
 - a. *Dashboard* de resultados de modelos ML
 - b. Visualización de datos multidimensionales
 - c. Interfaces para exploración de *datasets*
 - d. Comunicación de hallazgos científicos

Estadística avanzada y métodos cuantitativos

1. Estadística inferencial avanzada
 - a. Distribuciones de probabilidad complejas
 - b. Teorema central del límite y aplicaciones
 - c. Intervalos de confianza y pruebas de hipótesis avanzadas
 - d. ANOVA multifactorial y ANCOVA
2. Métodos estadísticos multivariados
 - a. Análisis de componentes principales (PCA)
 - b. Factor Analysis y análisis de correspondencias
 - c. Análisis discriminante y clasificación estadística
 - d. *Clustering* estadístico y métodos de agrupación
3. Series temporales y *forecasting*
 - a. Componentes de series temporales y descomposición
 - b. Modelos ARIMA y SARIMA avanzados
 - c. Modelos de suavizado exponencial (Holt-Winters)
 - d. Prophet de Facebook para *forecasting* automático
 - e. Evaluación de modelos de pronóstico
4. Métodos bayesianos
 - a. Teorema de Bayes e inferencia bayesiana
 - b. MCMC y métodos de simulación
 - c. Modelos bayesianos con PyMC3 y Stan
 - d. Comparación bayesiana vs frecuentista
5. Diseño de experimentos y causalidad
 - a. *A/B testing* avanzado y diseños experimentales
 - b. Análisis de poder estadístico
 - c. Inferencia causal y métodos de identificación
 - d. *Propensity score matching*

Programación: Python y R

1. Python para *data science*
 - a. Configuración de entornos (Anaconda, Jupyter, Google Colab)
 - b. NumPy y manipulación de *arrays* multidimensionales
 - c. Pandas avanzado para análisis de datos
 - d. Matplotlib, Seaborn y Plotly para visualización
2. R para análisis estadístico
 - a. RStudio y configuración del entorno R
 - b. Manipulación de datos con *dplyr* y *tidyr*
 - c. Visualización con *ggplot2*
 - d. Análisis estadístico con R base y *packages* especializados
3. Integración y casos prácticos
 - a. Comparación Python vs R: cuándo usar cada uno
 - b. *Notebooks* reproducibles y documentación
 - c. APIs y conexión a bases de datos desde ambos lenguajes
 - d. Proyectos integradores combinando ambas herramientas

Machine Learning e Inteligencia Artificial

1. Fundamentos de *Machine Learning*
 - a. Tipos de aprendizaje: supervisado, no supervisado, reforzado
 - b. Evaluación de modelos y métricas de performance
 - c. *Cross-validation* y técnicas de validación
 - d. *Feature engineering* y selección de variables
2. Algoritmos de *Machine Learning*

Supervisado

 - a. Regresión: lineal, logística, Ridge, Lasso, Elastic Net
 - b. Árboles de decisión y Random Forest
 - c. Support Vector Machines (SVM)
 - d. *Gradient Boosting*: XGBoost, LightGBM, CatBoost
 - e. *Ensemble methods* y *stacking*

Machine Learning no supervisado

 - a. Clustering: K-means, jerárquico, DBSCAN, Gaussian Mixture
 - b. Reducción de dimensionalidad: PCA, t-SNE, UMAP
 - c. *Association rules* y *market basket analysis*
 - d. *Anomaly detection* y detección de outliers
4. *Deep Learning* y redes neuronales
 - a. Redes neuronales básicas y perceptrón multicapa
 - b. TensorFlow y Keras para *deep learning*
 - c. Convolutional Neural Networks (CNN) para imágenes
 - d. Recurrent Neural Networks (RNN/LSTM) para secuencias
 - e. / y modelos pre-entrenados
5. Natural Language Processing (NLP)
 - a. *Preprocessing* de texto: tokenización, *stemming*, lematización
 - b. *Sentiment analysis* y clasificación de texto
 - c. *Topic modeling*: LDA, BERTopic
 - d. Named Entity Recognition (NER)
 - e. Transformers básicos: BERT, GPT
6. *Computer vision*
 - a. Procesamiento de imágenes con OpenCV
 - b. Clasificación de imágenes con CNN
 - c. *Object detection* básico
 - d. *Image segmentation* conceptos
7. MLOps y productivización
 - a. Versionado de modelos con MLflow
 - b. *Hyperparameter tuning* automatizado
 - c. *Model deployment* básico con Flask/FastAPI
 - d. *Model monitoring* y *drift detection*
 - e. A/B testing para modelos ML
8. AutoML y herramientas avanzadas
 - a. AutoML con H2O, Auto-sklearn
 - b. Google AutoML y Azure AutoML
 - c. *Feature engineering* automatizado
 - d. Neural Architecture Search básico

Data science y storytelling avanzado

1. IA ética y responsable
 - a. Bias y *fairness* en modelos de ML
 - b. *Explainable AI*: SHAP, LIME, interpretabilidad
 - c. Privacidad diferencial y técnicas de anonimización
 - d. Regulaciones de IA: GDPR, AI Act y consideraciones legales
2. Sistemas de recomendación
 - a. *Collaborative filtering* y *content-based filtering*
 - b. *Matrix factorization* y métodos híbridos
 - c. *Deep learning* para recomendaciones
 - d. Evaluación de sistemas de recomendación
3. *Time Series Forecasting* Avanzado
 - a. LSTM y GRU para series temporales complejas
 - b. Prophet con componentes personalizados
 - c. *Forecasting ensembles* y combinación de modelos
 - d. *Forecasting* con variables exógenas
4. *Storytelling* y comunicación científica
 - a. Narrativa con datos para audiencias técnicas y ejecutivas
 - b. Presentación de resultados de modelos complejos
 - c. Documentación científica y reproducibilidad
 - d. *Portfolio* profesional en *data science*
5. Casos de uso empresariales
 - a. *Customer analytics*: segmentación, *lifetime value*, churn
 - b. *Marketing analytics*: *attribution*, *campaign optimization*
 - c. *Financial analytics*: *credit scoring*, *fraud detection*
 - d. *Operations analytics*: *supply chain*, *demand forecasting*
6. Proyectos Capstone
 - a. Proyecto integral de *data science* *end-to-end*
 - b. Desde *problem definition* hasta *deployment*
 - c. Presentación a *stakeholders* técnicos y de negocio
 - d. Metodología de proyectos de ciencia de datos

Beneficios de la modalidad

Clases en vivo, actividades interactivas y casos prácticos. Puedes interactuar con profesores y otros alumnos para tener una experiencia más enriquecedora.

Networking. Tienes la oportunidad de construir una red de contactos profesionales con otras personas que tienen intereses similares o se desempeñan en el mismo ámbito.

Asesoría y acompañamiento. Cuentas con un facilitador por módulo para guiarte durante tu curso.

Aplica lo que aprendas de forma inmediata.

SÉ PARTE DE LA UVM



@uvmmx



uvm



@uvmmx



uvm.mx