

BOOTCAMP EN DATA SCIENCE

Híbrido

(224 horas)

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 UVM con validez curricular y tecnología Blockchain con código QR y de verificación.**
 - **Certificado Internacional de Embiz Foundation.**
 - **Certificado de competencias laborales DC-3 de la STPS.**

¿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.



MÓDULOS

01 Introducción a la ciencia de los datos y Herramientas estadísticas.

1. Campos de aplicación del big data
 - a. Discusión sobre diferentes campos y sectores que utilizan big data y la ciencia de datos.
 - b. Ejemplos de aplicaciones en la industria, gobierno, investigación, marketing, etc.
2. Modelos
 - a. Definición de modelos y su importancia en la ciencia de datos.
 - b. Presentación de los diferentes tipos de modelos utilizados en la ciencia de datos, como modelos estadísticos, modelos de aprendizaje automático y modelos de redes neuronales.
3. Discusión sobre cómo se utilizan los modelos para resolver problemas en diferentes campos.
4. Descriptiva y de estimación
 - a. Tablas de frecuencia: presentación de datos en forma de tabla y cómo interpretarlas.
 - b. Probabilidades: introducción a la teoría de la probabilidad y cómo se aplica en la ciencia de datos.
 - c. Inferencial: cómo realizar inferencias a partir de datos de muestra y cómo interpretar los resultados.
 - d. Modelo lineal: explicación de los modelos lineales y cómo se ajustan a los datos.
 - e. Forecast: concepto de pronóstico y técnicas para hacer pronósticos.
5. Implementación de usos gráficos y clasificación
 - f. Visualización de datos: diferentes tipos de gráficos y técnicas de visualización utilizados en la ciencia de datos.
 - g. Clasificación: técnicas para clasificar datos en diferentes categorías.
6. Experimentación de modelos estadísticos
 - h. Cómo probar diferentes modelos y técnicas estadísticas.
 - i. Análisis de resultados y selección de modelos adecuados.

02 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.

MÓDULOS

03 Machine Learning.

1. Introducción al aprendizaje automático y cómo se utiliza en la ciencia de datos:

a. Concepto de aprendizaje automático y su papel en la ciencia de datos.

b. Diferentes tipos de aprendizaje automático: supervisado, no supervisado, por refuerzo.

c. Ejemplos de aplicaciones de aprendizaje automático en la industria, el gobierno, la investigación, la salud y otros campos.

2. Funcionalidades y uso de software para la implementación de modelos de machine learning:

a. Descripción de las bibliotecas y herramientas populares de aprendizaje automático, como Scikit-Learn, TensorFlow, PyTorch y Keras.

b. Uso de estas bibliotecas para construir modelos de aprendizaje automático.

c. Preprocesamiento de datos para el aprendizaje automático.

3. Programación y utilización de funciones para el análisis de datos y la implementación de modelos estadísticos:

a. Concepto de funciones y su papel en el análisis de datos.

b. Implementación de funciones en lenguajes de programación populares, como Python o R.

c. Uso de funciones para implementar modelos de aprendizaje automático.

4. Identificación y análisis de grandes volúmenes de datos:

a. Concepto de big data y su impacto en la ciencia de datos.

b. Métodos para almacenar y procesar grandes volúmenes de datos, como Hadoop o Spark.

c. Técnicas para analizar grandes conjuntos de datos, como el aprendizaje por lotes o el aprendizaje incremental.

5. Reducción de la dimensionalidad y técnicas de clasificación:

a. Técnicas para reducir la dimensionalidad de los datos, como el análisis de componentes principales o la selección de características.

b. Métodos de clasificación, como los árboles de decisión, las redes neuronales y los algoritmos de vecinos más cercanos.

c. Evaluación y comparación de modelos de clasificación.

6. Procesos de máquina y ajuste de modelos:

a. Concepto de proceso de máquina y su papel en el ajuste de modelos.

b. Métodos de ajuste de modelos, como la validación cruzada y la selección de modelos.

c. Métodos para evitar el sobreajuste y el subajuste.

7. Modelos de regresión y clasificación:

a. Regresión lineal y polinómica, y su uso en la predicción de valores numéricos.

b. Modelos de clasificación, como los árboles de decisión, las redes neuronales y los algoritmos de vecinos más cercanos.

c. Ejemplos de aplicaciones de modelos de regresión y clasificación.

MÓDULOS

03 Machine Learning.

8. Modelos de árboles de decisión y ensamblaje de modelos:

- a. Concepto de modelos de árboles de decisión y su uso en la clasificación y regresión.
- b. Ensamblaje de modelos, como el bagging, el boosting y el apilamiento.
- c. Ejemplos de aplicaciones de modelos de árboles de decisión y ensamblaje de modelos.

9. Algoritmos de clustering y redes neuronales:

- a. Algoritmos de clustering, como k-means clustering jerárquico.
- b. Redes neuronales y su uso en el clustering y la clasificación no supervisada.
- c. Ejemplos de aplicaciones de algoritmos de clustering y redes neuronales.

10. Procesamiento del lenguaje natural y minería de textos:

- a. Concepto de procesamiento del lenguaje natural y su papel en la minería de textos.
- b. Métodos para el preprocesamiento y análisis de texto, como la tokenización, la eliminación de stop words y la lematización.
- c. Modelos de aprendizaje automático utilizados en la minería de textos, como el modelo de bolsa de palabras y el modelo de lenguaje.
- d. Ejemplos de aplicaciones de procesamiento del lenguaje natural y minería de textos.

11. Aprendizaje profundo:

- a. Concepto de aprendizaje profundo y su relación con las redes neuronales.
- b. Diferentes arquitecturas de aprendizaje profundo, como las redes neuronales convolucionales y recurrentes.

c. Ejemplos de aplicaciones de aprendizaje profundo, como el reconocimiento de imágenes y el procesamiento de lenguaje natural.

- d. Uso de modelos pre-entrenados y transferencia de aprendizaje:
- e. Concepto de modelos pre-entrenados y transferencia de aprendizaje.
- f. Uso de modelos pre-entrenados para acelerar el proceso de entrenamiento y mejorar la precisión.

g. Transferencia de aprendizaje para aplicaciones en diferentes dominios y conjuntos de datos.

12. Problemas comunes en el aprendizaje automático y cómo solucionarlos:

- a. Sobreajuste y subajuste de modelos.
- b. Falta de representatividad y desequilibrio de los datos.
- c. Métodos para evitar estos problemas, como la selección de modelos y la validación cruzada.

13. Ética y privacidad en el aprendizaje automático:

- a. Concepto de ética y privacidad en el aprendizaje automático.
- b. Problemas éticos y de privacidad comunes en el aprendizaje automático, como el sesgo algorítmico y la recolección de datos personales.
- c. Métodos para abordar estos problemas, como la transparencia algorítmica y la protección de datos personales.

04 Introducción a Python.

1. Instalación y configuración de Python.
 - a. Sintaxis y estructuras de datos básicas en Python.
 - b. Uso de Python para manipular y analizar datos.
2. Números:
 - a. Operaciones aritméticas básicas en Python.
 - b. Funciones matemáticas y estadísticas en Python.
 - c. Uso de Python para realizar cálculos y análisis numérico.
3. Strings:
 - a. Manipulación de cadenas de caracteres en Python.
 - b. Funciones y métodos para manipular strings en Python.
 - c. Uso de Python para procesamiento de texto y análisis de lenguaje natural.
4. Condicionales:
 - a. Uso de condicionales en Python, como if-else y switch-case.
 - b. Uso de condicionales para realizar selecciones y filtros en datos.
5. Bucles:
 - a. Uso de bucles en Python, como for y while.
 - b. Uso de bucles para iterar sobre datos y realizar operaciones repetitivas.
6. Programación orientada a objetos:
 - a. Concepto de programación orientada a objetos y su uso en Python.
 - b. Clases, objetos, atributos y métodos en Python.
 - c. Uso de la programación orientada a objetos para modelar datos y solucionar problemas.
7. Uso de Scikit-Learn:
 - a. Descripción de la biblioteca Scikit-Learn y sus principales características.
 - b. Uso de Scikit-Learn para implementar modelos de aprendizaje automático, como regresión lineal y clasificación.
8. Visualización de datos con SQL y Google Big Query:
 - a. Uso de SQL para realizar consultas y manipulaciones de datos en bases de datos relacionales.
 - b. Introducción a Google Big Query y su uso para analizar grandes conjuntos de datos.
 - c. Visualización de datos utilizando herramientas como Matplotlib y Seaborn.
9. Descripción de datos y usos de filtros:
 - a. Descripción de los datos y su importancia en el análisis de datos.
 - b. Uso de filtros para reducir el conjunto de datos y facilitar su análisis.
 - c. Ejemplos de aplicaciones de filtros en diferentes tipos de datos.
10. VS Code:
 - a. Descripción del editor de código VS Code y sus características.
 - b. Configuración y personalización de VS Code.
 - c. Uso de VS Code para desarrollar proyectos de ciencia de datos.
11. Virtual machines:
 - a. Concepto de máquinas virtuales y su uso en la ciencia de datos.
 - b. Creación y configuración de máquinas virtuales para el análisis de datos.
 - c. Ejemplos de aplicaciones de máquinas virtuales en la ciencia de datos.
12. Big data:
 - a. Descripción de big data y su impacto en la ciencia de datos.
 - b. Técnicas y herramientas para procesar grandes volúmenes de datos, como Hadoop y Spark.
 - c. Ejemplos de aplicaciones de big data en diferentes industrias.
13. Rendimiento de modelos:
 - a. Medición del rendimiento de los modelos de aprendizaje automático, como la precisión, la sensibilidad y la especificidad.
 - b. Técnicas para mejorar el rendimiento de los modelos, como el ajuste de hiperparámetros y la selección de modelos.
 - c. Ejemplos de aplicaciones de medición y mejora del rendimiento de los modelos.
14. Predicciones mediante API:
 - a. Concepto de API y su uso en la ciencia de datos.
 - b. Uso de API para realizar predicciones basadas en modelos de aprendizaje automático.
 - c. Ejemplos de aplicaciones de API en diferentes industrias.

05 Visualización de datos y Herramientas para la visualización de datos

1. ¿Qué es la visualización de datos?
2. La importancia de la visualización de datos en la ciencia de datos y en la toma de decisiones en general.
3. Tipos de visualizaciones de datos y cuándo utilizar cada una.
4. Diseño de visualizaciones efectivas:
 - a. Principios básicos de diseño de visualizaciones.
 - b. Cómo seleccionar el tipo de visualización adecuado para los datos.
 - c. Cómo optimizar el formato y estilo de la visualización para una comunicación efectiva.
 - d. Ejemplos de mejores prácticas y errores comunes en el diseño de visualizaciones.
5. Herramientas de visualización de datos:
 - a. Herramientas de visualización de datos populares, como Tableau, Power BI, D3.js y Matplotlib.
 - b. Cómo utilizar estas herramientas para crear visualizaciones de datos interactivas y personalizadas.
 - c. Ejemplos de aplicaciones de visualización de datos en diferentes industrias.
6. Visualización de datos avanzada:
 - a. Técnicas avanzadas de visualización de datos, como la visualización de redes y la visualización 3D.
 - b. Cómo utilizar técnicas avanzadas de visualización para comunicar patrones y tendencias complejos en los datos.
 - c. Ejemplos de aplicaciones de visualización de datos avanzada en diferentes industrias.
7. Introducción a Power BI:
 - a. Descripción de la herramienta de visualización de datos Power BI.
 - b. Instalación y configuración de Power BI.
 - c. Uso de Power BI para crear informes y paneles interactivos.
8. Manipulación de objetos en Power BI:
 - a. Descripción de los objetos en Power BI, como gráficos, tablas y matrices.
 - b. Manipulación de objetos en Power BI, como la creación de campos calculados y la adición de filtros.
9. Generación de gráficos en Power BI:
 - a. Uso de Power BI para crear diferentes tipos de gráficos, como gráficos de barras, gráficos de líneas y gráficos circulares.
 - b. Personalización de gráficos en Power BI, como la modificación de etiquetas y colores.
10. Cuadros interactivos en Power BI:
 - a. Uso de cuadros interactivos en Power BI, como los filtros y los selectores.
 - b. Creación de cuadros interactivos en Power BI para permitir a los usuarios interactuar con los datos.
11. Introducción a Tableau:
 - a. Descripción de la herramienta de visualización de datos Tableau.
 - b. Instalación y configuración de Tableau.
 - c. Uso de Tableau para crear informes y paneles interactivos.
12. Manipulación de objetos en Tableau:
 - a. Descripción de los objetos en Tableau, como hojas de trabajo, paneles y dashboards.
 - b. Manipulación de objetos en Tableau, como la creación de campos calculados y la adición de filtros.
13. Generación de gráficos en Tableau:
 - a. Uso de Tableau para crear diferentes tipos de gráficos, como gráficos de barras, gráficos de líneas y gráficos circulares.
 - b. Personalización de gráficos en Tableau, como la modificación de etiquetas y colores.
14. Análisis de tablas en Tableau:
 - a. Uso de Tableau para analizar datos en tablas.
 - b. Creación de tablas dinámicas en Tableau para permitir a los usuarios explorar los datos.
15. Dashboards en Tableau:
 - a. Creación de dashboards interactivos en Tableau.
 - b. Uso de dashboards en Tableau para resumir y presentar los datos de manera clara y concisa.

06 Aplicaciones específicas de la ciencia de datos y StoryTelling

1. Detección de fraudes en transacciones financieras:
 - a. Análisis de patrones y comportamientos sospechosos.
 - b. Técnicas de machine learning y minería de datos para detectar fraudes.
 - c. Ejemplos de aplicaciones en la industria financiera.
2. Análisis de sentimientos en redes sociales:
 - a. Procesamiento de lenguaje natural y análisis de texto.
 - b. Técnicas de clasificación y agrupamiento para el análisis de sentimientos.
 - c. Ejemplos de aplicaciones en la industria del marketing y publicidad.
3. Predicción de demanda y optimización de la cadena de suministro:
 - a. Análisis de datos históricos y de tendencias.
 - b. Modelos de pronóstico y optimización de inventario.
 - c. Ejemplos de aplicaciones en la industria de la logística y el transporte.
4. Personalización de la experiencia del cliente:
 - a. Análisis de datos de comportamiento del cliente.
 - b. Modelos de recomendación y personalización.
 - c. Ejemplos de aplicaciones en la industria del comercio electrónico y el entretenimiento.
5. Diagnóstico y predicción de enfermedades en la industria de la salud:
 - a. Análisis de datos médicos y registros de salud electrónicos.
 - b. Modelos de predicción y diagnóstico de enfermedades.
6. ¿Qué es el storytelling?
7. ¿La importancia del storytelling en la ciencia de datos y en la comunicación en general.
8. Elementos clave de una historia efectiva.
9. Técnicas y herramientas para contar historias con datos.
10. Pasos para contar historias con datos:
 - a. Definir el objetivo de la historia.
 - b. Identificar la audiencia y adaptar el mensaje.
 - c. Seleccionar los datos adecuados para contar la historia.
 - d. Diseñar la estructura de la historia y establecer una narrativa clara.
 - e. Seleccionar las visualizaciones adecuadas para respaldar la historia.
 - f. Presentar la historia de manera efectiva.
11. Tipos de historias de datos:
 - a. Historias de datos exploratorias:
 - i. Exploración y análisis de datos para descubrir patrones y tendencias.
 - ii. Ejemplos de aplicaciones en la industria del marketing y publicidad.
 - b. Historias de datos explicativas:
 - i. Explicación de datos y análisis para respaldar una conclusión o recomendación.
 - ii. Ejemplos de aplicaciones en la industria de la toma de decisiones empresariales.
 - c. Historias de datos persuasivas:
 - i. Presentación de datos y análisis para persuadir a una audiencia de una idea o acción.
 - ii. Ejemplos de aplicaciones en la industria de la política y el activismo.
11. Ejemplos de aplicaciones en la industria de la salud.

07 Proyecto Final

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.

Estudia a tu ritmo. Consulta todas las sesiones grabadas en el horario que más te convenga.

Soporte técnico. Cuentas con atención técnica en todo momento para ayudarte a solucionar cualquier problema que se presente.

Nota: Si no asistes a las sesiones en vivo con el profesor en las fechas y horarios establecidos, tendrás 30 días naturales para ver completa la grabación de la clase en Teams® y realizar la actividad asignada para que acredites el módulo.

SÉ PARTE DE LA UVM



@uvmmx



uvm



@uvmmx



uvm.mx