

Introducción a la Epidemiología con Oracle Analytics/Data Visualization y Técnicas de Machine Learning

1. Descripción general

Esta microcredencial introduce los fundamentos de la epidemiología moderna y su aplicación práctica mediante Oracle Analytics/Data Visualization, con un enfoque especial en la visualización y el uso de técnicas de Machine Learning. El estudiante aprenderá a gestionar, analizar y visualizar datos epidemiológicos, aplicar técnicas estadísticas y de aprendizaje automático para la exploración, predicción y clasificación de datos, y construir cuadros de mando (dashboards) orientados a la vigilancia y toma de decisiones en salud pública.

2. Perfil de participantes

Estudiantes de doctorado, profesionales sanitarios, analistas de datos o personal técnico de salud pública interesados en aplicar técnicas de análisis de datos y Machine Learning en el ámbito de la epidemiología. No se requieren conocimientos previos de epidemiología ni de Machine Learning, pero sí manejo básico de herramientas estadísticas. Se requiere ordenador portátil con acceso a la versión Desktop, Cloud o Server de Oracle Analytics/Data Visualization.

3. Resultados de aprendizaje

Al finalizar la microcredencial, el estudiante será capaz de:

- Diseñar análisis descriptivos y analíticos básicos.
- Importar y transformar datos epidemiológicos en Oracle Analytics.
- Construir visualizaciones epidemiológicas (curvas epidémicas, mapas, diagramas de cajas).
- Aplicar modelos de Machine Learning (regresión, clasificación, clustering) a problemas de salud pública.
- Comunicar hallazgos mediante informes y cuadros de mando (dashboards) interactivos

4. Estructura y contenidos

Módulo 1. Introducción a Oracle Analytics:

- Conceptos básicos.
- Preparación de datos: “datasets”
- Flujos de Datos: “dataflows”
- Visualización de datos: “workbooks”

Módulo 2. Gestión y preparación de datos epidemiológicos

- Tipos de datos en epidemiología.
- Importación de datasets en Oracle Analytics.
- Limpieza, unión, filtrado y creación de variables.
- Preprocesamiento: detección de valores atípicos, imputación, normalización.

Módulo 3. Análisis exploratorio y visualización

- Curvas epidémicas y series temporales.
- Mapas de distribución geográfica.
- Análisis multivariante.
- Técnicas de clustering y segmentación con Machine Learning para identificar patrones epidemiológicos.

Módulo 4. Modelos predictivos y de clasificación con Machine Learning

- Regresión lineal y logística en epidemiología.
- Modelos multivariantes adaptados a datos de salud.
- Algoritmos de clasificación: árboles de decisión, regresión logística, redes neuronales básicas.
- Evaluación de modelos: precisión, sensibilidad, especificidad, curvas ROC.
- Validación interna y externa de modelos de Machine Learning.

Módulo 5. Proyecto final aplicado

- Análisis de un problema epidemiológico real.
- Aplicación de técnicas de Machine Learning para predicción o clasificación.
- Construcción de un dashboard en Oracle Analytics.

- Presentación de resultados por escrito y coevaluación entre pares.

5. Metodología

La microcredencial combina sesiones teóricas breves con prácticas intensivas. Se trabajará con datasets de vigilancia epidemiológica aportados por el alumno. Se fomentará el aprendizaje activo mediante retos analíticos, trabajo en equipo y desarrollo de proyectos aplicados. El uso de Oracle Analytics/Data Visualization permitirá aplicar técnicas de Machine Learning de forma visual e interactiva.

6. Evaluación

- Actividades prácticas semanales.
- Participación, ejercicios y resolución de problemas.
- Proyecto final con elaboración de un informe visual y evaluación entre pares.

7. Recursos:

- Se ofrece acceso gratuito a Oracle Analytics Cloud y Oracle Analytics Desktop gracias a la afiliación de la EHU al programa Oracle Academy.
- Plataforma EHUON con datos epidemiológicos agregados y anonimizados (p. ej., brotes, series temporales, cohortes).
- Video tutoriales originales creados con Avatares
- Galería de visualizaciones y dashboards de ejemplo.
- Manuales de referencia sobre Visualización, Oracle Analytics y Machine Learning aplicado a la salud pública.