



Analista de datos Junior con Python

Universidad Nacional de Chilecito



UNIVERSIDAD
NACIONAL
de CHILECITO



Tramo 1: se tomará como equivalente de este tramo a los cursos introductorios dictados por otra unidad académica, en el marco de Argentina Programa 4.0.

Tramo 2: Programación Avanzada con Python

- **Módulo 1 - Introducción a la programación con Python:** Conceptos introductorios, preparación de entorno de trabajo.
- **Módulo 2 - Validación de Datos:** La importancia de la validación en programación. Validación de tipo y de rango. Captura de la reiteración de errores de entrada y su tratamiento.
- **Módulo 3 - Tipos de Datos Compuestos:** Cadena de caracteres y listas. Manipulación de tipos de datos compuestos: el caso de las cadenas de caracteres y las listas. Manipulación. Carga, recorrido, búsqueda. Algoritmos clásicos.
- **Módulo 4 - Numpy:** Librería Numpy. Array N Dimensionales. Funciones. Optimización.
- **Módulo 5 - Estadística Descriptiva, medidas de tendencia:** Media, Mediana, Moda. Medidas de Variabilidad: Rango, Varianza, Desvío estándar. Medidas asociadas entre variables: correlación.
- **Módulo 6 - Pandas:** estructuras de datos, funciones, operaciones de agrupación. Series, DataFrames.
- **Módulo 7 - Visualización de Datos:** Visualización de Datos. Tipos de gráficos: histogramas, barras, tortas, series, dispersión, diagramas de caja. Matplotlib.

Modalidad: virtual (sincrónica y asincrónica).

Horas cátedra: 100 hs. Duración del tramo: 2 meses.



Tramo 3: Análisis y Procesamiento de Datos

- **Módulo 1 - Fundamentos de programación con Python:** instalación del software Anaconda. Entorno de trabajo, ventaja de las librerías. Jupyter Notebooks. Documentación. Abrir y guardar proyecto.
- **Módulo 2 - Introducción a la Ciencia de Datos:** Introducción a la Ciencia de Datos. Obtención de conocimiento a partir de los datos. El concepto de patrón. El proceso KDD. Fases del proceso de extracción del conocimiento. La minería de datos como fase del proceso KDD. Relación con otras disciplinas..
- **Módulo 3 - Recuperación de información y recuperación de datos.** Proceso de recuperación de información. Preparación de Datos. Metadatos. Análisis de la información de entrada. Construcción y análisis de representaciones gráficas.
- **Módulo 4 - Manipulación de Datos:** Limpieza y transformación. Transformación y creación de atributos. Discretización y Numerización, Normalización de rango, escalado y centrado.
- **Módulo 5 - Visualización Avanzada:** Matplotlib. Exploración mediante Visualización y selección de datos.
- **Módulo 6 - Técnicas de Agrupamiento:** Técnicas de Agrupamiento. Métricas de calidad del agrupamiento. Tipos de agrupamiento: Jerárquico, partitivo y probabilista. Medidas de distancia y de conectividad. Proceso de agrupamiento. Clustering partitivo. Algoritmo k-means. Algoritmos de clustering jerárquicos aglomerativos y divisivos.
- **Módulo 7 - Árboles de Decisión:** Árboles de decisión. Métricas de selección de atributos. Entropía. Ganancia de Información. Tasa de Ganancia. Índice Gini. Poda y Sobreajuste. Algoritmos Id3, C4.5. Construcción de árboles para grandes volúmenes de datos.

Modalidad: virtual (sincrónica y asincrónica).

Horas cátedra: 100 hs. Duración del tramo: 2 meses.



Analista de Datos Junior con Python

EVALUACIÓN

Se realizará una evaluación formativa de carácter continuo a lo largo del trayecto, donde tendrás que ir realizando entregas de carácter práctico y teórico (múltiple choice).

Para aprobar el curso el participante deberá demostrar que comprende y maneja los conceptos básicos y demostrar habilidades para resolver los problemas que se planteen a lo largo de la cursada.

Considerando que la modalidad será virtual, la modalidad de evaluación será a través de:

- Asistencia al 70% de las clases sincrónicas.
- Participación activa en los foros propuestos como obligatorios.
- Aprobar el 80% de los trabajos prácticos y/o actividades propuestas.
- Al finalizar habrá un trabajo integrador final de carácter individual, el cual será entregado en la semana 6 de cada curso.

