



Nombre del curso: Python: una Introducción al mundo científico

Requerimientos: Cada alumno debe contar con computadora, al menos 4 Gb de RAM y 4 Gb de almacenamiento en disco duro disponible. Se trabajará con Windows 8 (se dará la explicación en este Sistema Operativo), aunque los programas pueden funcionar en Windows 7,8 10.

Dirigido a: Estudiantes, profesionistas y público en general de todas las áreas que estén interesados en adentrarse en el mundo de la investigación computacional; buscando «ideas» para proyectos de tesis o métodos para resolver problemas relacionados con el desarrollo de software de alto nivel empleando programas sencillos, pero de gran utilidad.

Consideraciones: El curso es ideal para profesionistas, investigadores, estudiantes con problemas de investigación y personas de todas las áreas del saber humano que deseen usar una computadora para resolver problemas propios de su actividad, pero que no tienen mucho tiempo para profundizar en el desarrollo de software, sino que quieren "aprender haciendo y resolviendo", usando herramientas ya probadas. El curso se empezará con una introducción básica del lenguaje, que es un lenguaje fácil de dominar. El curso no termina aquí, no es simplemente un aprender Python. A continuación, el curso expone a los asistentes a adentrarse en el mundo de la investigación computacional, dándoles las aplicaciones (herramientas de software) que resuelven variados problemas típicamente encontrados en proyectos de innovación. Las herramientas que se cubrirán se emplean en proyectos de investigación, de innovación, en tesis y desarrollos avanzados en ingeniería, economía, medicina, lingüística, sociología, etc. Las herramientas son paquetes de software de alto nivel, que cubren métodos para resolver problemas relacionados con el desarrollo de aplicaciones empleando programación sencilla, pero de gran utilidad para resolver problemas en variadas disciplinas con el respaldo de la inmensa comunidad Python cuyo objetivo primordial es simplemente, poner el software de alto nivel al alcance de quien lo necesite. Se explicarán las herramientas, demostrando su uso en clase, respondiendo preguntas generales sobre sus variantes, los parámetros que emplean y mencionando proyectos en los que se usan. Con esto se pretende fomentar la atracción de cualquier persona con deseo de aprendizaje mundo científico para que ayuden a mejorar el entorno en el que viven.

Contenido

1. Introducción a Python (1h:30m)
 - a. Introducción
 - b. Versiones de Python
 - c. Raíces (¿De qué lenguajes proviene Python?)
 - d. Características
 - e. Implementaciones
 - f. ¿Quién ocupa Python?
 - g. Instalación de Python
 - h. Instalación de Anaconda
 - i. Instalación de PyCharm
 - j. Primer programa

2. Conceptos básicos del lenguaje (30m)
 - a. Tipos de datos
 - i. Números
 - ii. Cadenas
 - iii. Booleanos
 - iv. Operadores aritméticos
 - v. Operadores a nivel de bit
 - vi. Solicitar datos al usuario

 - b. Colecciones (1h)
 - i. Listas
 - ii. Tuplas
 - iii. Diccionarios

 - c. Sentencias control de flujo (1h)
 - i. if
 - ii. switch ¿Por qué no existe?
 - iii. while
 - iv. for
 - v. range()

 - d. Programación orientada a objetos (1h)
 - i. Definición de función
 - ii. Funciones paso por valor y paso por referencia

- iii. Clases y objetos
 - iv. Herencia
 - v. Polimorfismo
 - vi. Encapsulación
 - vii. Getters y Setters
 - viii. Métodos especiales
- e. Conociendo más métodos (30m)
- i. Diccionarios
 - ii. Cadenas
 - iii. Listas
 - iv. Tuplas
- f. Módulos y paquetes (30m)
- i. Módulos
 - ii. Paquetes
- g. Entradas y salidas de archivos (30m)
- i. Creación de archivos
 - ii. Leer un archivo
 - iii. Escribir en un archivo
 - iv. Modos de acceso a los archivos
 - v. Lectura de archivos .csv
 - vi. Escritura de archivos .csv

3. NLTK (Natural Language ToolKit) (3h) [Es una introducción al análisis del lenguaje humano, específicamente a través de texto. Se muestran ejemplos para que el estudiante pueda adentrarse al mundo de la lingüística computacional].

- a. Instalación NLTK
- b. Tokens, palabras sin sentido, stemming
- c. Introducción a WordNet
 - i. Acceder a la taxonomía
 - ii. Obtener Synsets
 - iii. Sinónimos, hipónimos, hiperónimos de un Synset
 - iv. Calcular distancias entre Synsets (Wu & Palmer, Path)
- d. Acceder a los conjuntos de datos
- e. Clasificación de tweets (análisis de sentimientos positivos y negativos)
- f. Clasificación de texto

4. Clasificadores (3h) [Se muestra una introducción a un tema primordial en la inteligencia artificial que es el “reconocimiento de patrones”, específicamente el área de clasificación. La clasificación puede apoyar en el estudio de prácticamente cualquier campo del conocimiento humano. Por ejemplo, detección de terrorismo; apoyo en sistemas de protección civil; en descubrimiento de enfermedades, análisis del ADN, entre muchas].

- a. Introducción ¿Qué es un clasificador?, ¿Para qué se ocupa?
- b. Conjuntos de prueba particiones (conjuntos de entrenamiento y prueba estratificados)
- c. Construcción de clasificador KNN
- d. Aplicación de clasificadores (Clasificar correo como spam)

5. Manipulación de Base de datos (1h) [Se muestra de manera sencilla como poder manipular una base de datos usando Python, este bloque apoya a aquellos estudiantes que no tienen conocimientos profundos de bases de datos y busquen una implementación sencilla. Generalmente, las bases de datos, son el origen para cualquier proyecto que requiera una computadora para su implementación ya que facilitan en almacenamiento de la información].

- a. Introducción
- b. Instalación «Mysql connector»
- c. Conexión a una Base de datos
- d. Insertar registros en una base de datos
- e. Eliminar Registros

6. Graficar con Matplotlib (2h) [En este bloque se presenta una herramienta que apoya en la creación de gráficos convencionales que sirve para visualizar eficientemente los datos obtenidos de la resolución de algún problema. Se muestran ejemplos que pueden ser modificados sencillamente y que pueden ser adaptados sin complicaciones a todo tipo de datos que puedan representarse con variables numéricas. Este tipo de aplicaciones es útil en cualquier área en que se desee análisis de gráficas de barras, pastel, entre otras más específicas].

- a. ¿Qué es Matplotlib?
- b. Instalación
- c. Graficar arreglos
- d. Formato básico de las figuras
- e. Gráfica de pastel

- f. Histograma
- g. Gráfica de barras
- h. Múltiples figuras en una imagen
- i. Animación de la función seno
- j. Mostrando imágenes
- k. Gráfico en 3D

7. Graficar mediante Turtle. (2h) [Actualmente, el campo de Visualización de la información está adquiriendo mucha atención a nivel mundial. Aunado a los algoritmos complejos que analizan información, se requiere que los resultados representen datos de tal forma que un ser humano pueda comprenderla rápidamente independientemente de sus conocimientos. Turtle, es una herramienta que permite introducir a las personas al mundo de la creación de gráficas a través de una computadora para aquellos que requieran gráficos a la medida. Puede ser empleado por aplicaciones como juegos sencillos; aplicaciones relacionados con geometría analítica o visualización].

- a. Introducción a Turtle
- b. Línea, círculo, polígono
- c. Gráfica de barras
- d. Texto
- e. Eventos

8. Snap (análisis de grafos) (1h) [Analizar “redes” de información como redes sociales; el comportamiento de las publicaciones por científicos; el esparcimiento de un virus informático o alguna enfermedad; modelado de estaciones de suministro eléctrico o agua, es un campo de investigación con un gran alcance en ámbitos económicos y sociales (sólo por mencionar algunos) en los últimos años. La forma en cómo se produce la información permite que los grafos sean esenciales en la comprensión del funcionamiento de ciertos sistemas. Este bloque, muestra una herramienta útil para adentrarse en área de investigación de las redes complejas].

- a. Introducción
- b. Crear grafos
- c. Salvar grafos
- d. Manipular grafos

9. Análisis de Twitter (2h) [En este módulo se enseña a recuperar Tweets de la red social Twitter para poder analizarse posteriormente. Entre las aplicaciones que se pueden implementar al recuperar este tipo de mensajes destacan: ventajas competitivas; conocer la opinión de la población sobre ciertos eventos; publicidad dirigida, entre otras].

- a. Introducción
- b. Instalar Tweepy
- c. Obtener credenciales de acceso
- d. Recuperación de Tweets
- e. Tratamiento de Tweets

10. Extracción de PDF a texto (1h)

- a. Extracción de texto de documentos PDF
- b. Uso de indexador

11. Análisis de imágenes con OpenCV y Python. (2h). [Volúmenes gigantescos de información son generados a cada segundo en forma de imágenes. Debido a que para una persona es prácticamente imposible analizar tal flujo de información, resulta indispensable contar con herramientas que permitan simplificar al ser humano esta situación. Detección de tumores; análisis de delitos; seguimiento de manifestaciones; apoyo en desastres naturales; detección de zonas con incidencia delictiva; exploración del universo; ubicación de seguridad personal, son sólo algunos problemas latentes en los cuales se requiere un análisis de imágenes confiable para su resolución. El procesamiento digital de imágenes, es la rama de la ciencia que se encarga de trabajar con este tipo de problemas. En este bloque, se verá una herramienta que apoya en el análisis de imágenes de forma sencilla y de gran confianza].

- a. Cargar imágenes
- b. Desplegar imágenes
- c. Guardar imágenes
- d. Comparación entre imágenes
- e. Sumando imágenes
- f. Unión de imágenes

12. Diseño de figuras 3D con VTK y Python (3h) [Trabajar con gráficos de gran calidad apoya en labores como simulaciones o desarrollo de prototipos que pueden simplificar la labor de análisis por parte de los interesados. En este bloque, se presenta VTK, una herramienta de diseño gráfico sumamente poderosa que permite construir modelos para representar objetos en tres dimensiones, por ejemplo, el cerebro humano; piezas para sistemas (por ejemplo, autos); cadenas de ADN; información geo-espacial, entre muchas más. Este bloque presenta una breve introducción de la herramienta].

- a. Introducción
- b. Diseño de una figura geométrica
 - i. Mapper
 - ii. Actor
 - iii. Renderer
- c. Diseño de una flecha 3D
 - i. Mapper
 - ii. Actor
 - iii. Renderer
- d. Diseño de un cilindro 3D
 - i. Mapper
 - ii. Actor
 - iii. Renderer

13. Cifradores (Ciphers, en inglés) (4h) [La ciberseguridad es fundamental en la sociedad moderna, protección de datos; análisis de vulnerabilidades, encriptación, son sólo algunos usos de la seguridad informática. En este módulo se programan un par de Cifradores que ayudan proteger la información mediante codificaciones].

- a. ¿Qué es un Cifrador (Cipher)?
- b. Cifrador “César”
- c. Encriptar un archivo usando Cifrador “César”
- d. Desencriptar archivo usando Cifrador “César”
- e. Cifrador “Transportation”

14. Manipulación de mapas con Matplotlib y Python (2h) [Este módulo muestra la forma de usar herramientas de graficación en mapas con aplicaciones de propósito general sin usar herramientas profesionales GIS (Geographic Information System) complejas. Es muy útil en proyectos relacionados con los sistemas geo-espaciales].

- a. Introducción
- b. Desplegar mapa y océanos
- c. Mapa con relieve
- d. Mapa con proyección Mercator
- e. Mapa proyección Senoidal
- f. Desplegar datos en mapas

15. Geo-codificación (1h) [Se enseña a utilizar un módulo que permite encontrar direcciones (país, ciudad, calle, ...) a partir de coordenadas geográficas y lo contrario. Esta herramienta puede ser útil junto con herramientas GIS; análisis de formularios y detección de zonas específicas para envío de cuerpos de emergencia].

- a. Introducción
- b. Obtener coordenadas a partir de una dirección
- c. Obtener dirección a partir de coordenadas

16. Creación del juego Gato (Tic-tac-toe) (4h) [En este módulo se programa el bien conocido juego del gato empleando inteligencia artificial. Áreas como el modelado y la simulación matemática trabajan fuertemente para desarrollar teorías matemáticas que soporten y mejoren el desempeño de las máquinas con respecto a los “juegos”, y así conseguir desempeños óptimos. Los principales campos de acción de este tipo de enfoque las aplicaciones biológicas, psicológicas, médicas, computacionales (como el desarrollo de video juegos), entre otras].

- a. Introducción
- b. Explicación del problema
- c. Diseño del juego (Se usa inteligencia artificial)

17. Interfaz gráfica (3h) [El uso de interfaces gráficas para ejecutar nuestros programas facilita el uso del software. En este módulo se presentan de manera general los elementos para construir aplicaciones ejecutables sin necesidad de usar los programas a través de un IDE o Shell].

- a. Introducción
- b. Etiquetas
- c. Botones
- d. Canvas
- e. Entrada
- f. Frame
- g. ListBox
- h. Botón tipo menú
- i. Menú
- j. Mensaje
- k. Botones de radio
- l. Barra de escala
- m. ScrollBar
- n. Cajas de texto
- o. Ventanas por nivel
- p. SpinBox
- q. Paneles
- r. Marco de etiquetas
- s. Cajas de mensajes

18. Calcular similitud entre documentos de texto (3h) [En este módulo se presenta la base para el análisis de documentos de texto. Se aprende a convertir el texto en vectores para poder usar los algoritmos de inteligencia artificial. Este tipo de análisis puede ser usado para detectar plagio o extracción de características importantes sobre el texto como resúmenes; trabajos ampliamente abordados en la lingüística computacional].

- a. Introducción, ¿Qué es la similitud entre documentos?, ¿para qué calculamos la similitud entre documentos?
- b. Lectura de archivos
- c. Procesamiento de archivos (obtener tokens, remover stop-words)
- d. Obtener matriz término documento

e. Cálculo de similitud coseno

19. NumPy (2h) [En este módulo se estudia una herramienta que permite hacer cálculos de manera sencilla con gran precisión matemática. Permite el desarrollo de software científico de gran nivel con una implementación sencilla].

- a. ¿Qué es NumPy?
- b. Características de NumPy
- c. Uso de NumPy