

Tools

Hugo Villanueva

Ciencia de datos en Python.

Postgrado en análisis y Predicción de Datos

Universidad Galileo.

En el siguiente apartado se describirá brevemente lo entendido sobre las herramientas fundamentales que se van a utilizar en el curso.

Anaconda.

Es una distribución la cual permite utilizar paquetes en los lenguajes de programación R y Python, con un enfoque hacia la ciencia de datos, análisis de datos y machine learning, entre otros. Facilita la instalación de estos paquetes, creando ambientes virtuales para instalarlos precisamente en los proyectos que se necesitan, facilitando incluso utilizar versiones de Python o R distintas.

GIT

Es una herramienta de control de versiones, la cual ayuda a estructurar los distintos puntos de desarrollo de proyectos y mantener un historial de cada cambio de estos. Usa una ventana de terminal igual que los sistemas operativos en base unix. Del mismo modo registra quien hace estos cambios, fundamental en el seguimiento de trabajos en equipo. Se puede enlazar a una forma de trabajo remoto llamada Github, lo cual permite ver y modificar repositorios sin importar donde se esté.