Anaconda es una distribución libre y abierta de los lenguajes Python y R, utilizada en ciencia de datos, y aprendizaje automático (machine learning). Esto incluye procesamiento de grandes volúmenes de Anaconda información, análisis predictivo y cómputos científicos. Está orientado a simplificar el despliegue y administración de los paquetes de **ANACONDA**° software. Creada en 2007. Scikit-Learn es una de estas librerías gratuitas para Python. Cuenta con algoritmos de clasificación, regresión, clustering y scikit-learn reducción de dimensionalidad. Scikit-Learn es una de estas librerías gratuitas para Python. Cuenta con algoritmos de clasificación, regresión, clustering y reducción de dimensionalidad. Publicada el 9 de noviembre de 2015. Se utiliza para realizar cálculos numéricos mediante diagramas de flujo de datos. Los nodos de los TensorFlow diagramas representan operaciones matemáticas y **TensorFlow** las aristas reflejan las matrices de datos multidimensionales. Se lanzó el 28 de marzo de 2015. El objetivo de la biblioteca es acelerar la creación de redes neuronales: para ello, Keras no funciona como unframework independiente, sino como una Keras interfaz de uso intuitivo (API) que permite acceder a varios frameworksde aprendizaje automático y desarrollarlos. Entre los frameworks compatibles con Keras, se incluyen Theano, Microsoft Cognitive Toolkit (anteriormente CNTK) y TensorFlow. Apache Spark es un framework de computación (entorno de trabajo) en clúster open-source. Apache Spark se puede considerar un sistema de computación en clúster de propósito general y Spark orientado a la velocidad. Proporciona APIs en Java, Scala, Python y R. También proporciona un motor optimizado que soporta la ejecución de graficos en general. También soporta un conjunto extenso y rico de herramientas de alto nivel Es una plataforma de transmisión de datos, capaz de manejar billones de eventos al día en tiempo Kafka - Python real. Utiliza un marco de ciencia de datos en proyectos que requieren acceder y manejar grandes cantidades de datos en tiempo real. Apache MXNet es un marco de software de aprendizaje profundo de código abierto, que se utiliza para entrenar e implementar redes neuronales profundas .Que cuenta con la mxnet capacidad de alternar entre programación simbólica y programación imperativa para lograr la máxima productividad, también escala y distribuye la formación. Herramientas para ciencia de datos en lenguaje Python Utilizada para trabajar con computación numérica y NumPy científica como un paquete de procesamiento de NumPy matrices. proporciona objetos de matriz que son 50 veces más rápidos que las listas de Python. bokeh Se utiliza para crear visualizaciones interactivas y escalables dentro de los navegadores. Bokeh se Bokeh construye en capas, la interactividad es lo fundamental de esta biblioteca. Chainer Es un marco Python de red neuronal de código abierto. Chainer es conocido por su velocidad, Chainer especialmente en comparación con otros marcos más sofisticados. Aprovecha las múltiples CPU para acelerar el entrenamiento. Es compatible con cualquier DeepLearning4j lenguaje. Puede crear redes neuronales profundas a partir de redes poco profunda. Es una biblioteca de aprendizaje profundo que proporciona a los profesionales componentes de alto nivel que brindan fácilmente resultados de Fats.ai vanguardia. También proporciona a los investigadores componentes de bajo nivel que pueden mezclarse y combinarse para crear nuevos enfogues. Es una interfaz de aprendizaje profundo de código abierto de Microsoft y Amazon. Permite el Gluon aprendizaje automático, desarrollar modelos rápidamente sin comprometer el rendimiento. Lab | pandas Es una herramienta de aprendizaje automático que se utiliza para explorar, limpiar, transformar y Pandas visualizar datos. Se utiliza para modelos y capacitación de aprendizaje automático.