

Introduccion a data science

Data Science

Es el proceso de recopilar los datos para obtener informacion importante por un lado para crear estrategias de negocio y por otro creacion de software con la ayuda de la inteligencia artificial

Proceso a Realizar

- Obtencion de Datos: Fuentes, Encuestas
- Limpiar los Datos
- Explorar los datos, analizar y crear reportes
- Uso de modelos de machine learning para saber informacion
- Integrar IA, Inteligencia Artificial

En este se usa el modelo del metodo cientifico

Las ciencias de Datos son:

- Matematicas & Estadistica
 - Ciencias de la computacion
 - Conocimiento del Dominio
 - Aprender a Programar
 - Aprender Estadistica
 - Conocer las Industrias
-

Que es la Inteligencia Artificial

Data science no es lo mismo que Inteligencia Artificial

- **Algoritmos para emular la inteligencia Humana, el como sentimos, el como aprendemos etc:** reconociendo patrones en muchos datos
- **IA:** Empresas que se dedican a crear inteligencia Artificial
- No tiene conciencia ni mente propia, si un robot es hecho para leer libros solo eso sabra hacer y nada mas

Machine Learning (Proceso)

- Datos de entrada: Aqui aprende y va entendiendo gracias a muchos datos
- Analiza los datos
- Predecir gracias al analisis anterior

Ejemplo: El maching learning puede servir para predecir las compras y ventas de una empresa a futuro

Proceso de Ciencia de Datos: Integrar ML a Producto

**Ingesta de Datos ---> Validacion de los Datos ---> Preparacion de los Datos
---> Evaluacion de MOdelo ---> Validacion de Modelo ---> Despliegue de
Modelo ---> Interfaz de Usuario**

- Ingesta significa que introducimos los datos y los comience a leer

Diferencias entre data science

| **Data science | IA Inteligencia Artificial |
|----------------------------------------------------------------------------------|----------------------------------------------------------------------|
| Proceso de analizar datos para generar predicciones o generar toma de desiciones | Algoritmos para predecir eventos futuros, emulan inteligencia humana |

En el proceso de data science podemos crear herramientas para crear las predicciones

Que es Big data

- Grandes volumenenes de datos, escala grande
- No se puede trabajar con ellos en computadoras normales

Para que un proceso sea considerado "Big Data" tiene que cumplir con lo siguiente

| 5V de Big Data | Definicion |
|-----------------------|---------------------------------------------------|
| Volumen | Almacenamiento masivo de datos, en muchas fuentes |

| 5V de Big Data | Definicion |
|------------------|--------------------------------------------------------|
| Velocidad | Los datos se generan en tiempo real |
| Variedad | Ejemplos: Tablas, videos, Imagenes, Bases de datos etc |
| Veracidad | Calidad de los datos, sean confiables |
| Valor | Los datos deben de proporcionar valor a la empresa |

Para realizar big data necesitamos dividir los datos en muchas maquinas, en varias parte para poder ser procesados, usaremos Tecnologias de AWS servicios en linea

Se relaciona

- Es la materia prima donde podremos realizar analisis de datos
- Crear machine learning para automatizar todo

Que no es Data Science y ¿Porque no Aprender?

- No es magia, no solucionara todo automaticamente
- No es Inteligencia Artificial
- No es tener datos sin valor
- No son puras mates
- No solo se trabaja con Big Data

Porque no aprender

- No gustar del aprendizaje continuo, las tecnologias actualizan
- No amar el trabajo "sucio" limpiar los datos, eso sera complicado
- No te gusta comunicar o negociar lo que encuentras
- No te sientes motivado a ayudar a otras personas
- No te gusta hacer que las cosas pasen, osea hacer que las cosas se hagan

Porque aprender

- Me gusta tomar desiciones serteras
- Si me interesa saber como se aplican los datos
- Ver como las organizaciones crecen gracias a los datos
- Crear productos con la IA para revolucionar al mundo