¿Cómo resolvemos desafíos con datos?

En el camino de desarrollo de la data te mostramos cuáles procesos intervienen y qué roles colaboran a llevar adelante un proyecto de datos

START

- Definir objetivos
- Necesidades a resolver
- Revisión de presupuesto de tiempo y recursos
- Qué tecnologías usaremos: NLP, Reinforcemnet Learning,



MLOps

- Data target y drift
- Calidad del modelo
- Calidad de la data
- Mantención y mejora continua
- Escalar y distribuir

Data Extraction

- Dónde adquiriremos la data
- Webscrapping
- Revisión de sesgos
- Datasets públicos
- Selección de data y filtrado



Data Preparation

- Limpieza de datos
- Normalización
- Casos anómalos
- Selección de variables
- Etiquetado de datos
- Análisis de datos



Data Modeling

- Machine Learning vs Deep Learning
- Volumen de datos
- Tracking de resultados
- Arquitecturas a utilizar
- Experimentación
- Resultados y evaluación

Detrás de un proyecto de ML

Detrás de un proyecto de Machine Learning se mueve la metodología ágil que respalda tareas secuenciales (o no) e iterativas. Nos permitió llevar adelante proyectos en ML, DL en más de 10 industrias y múltiples especialidades (NLP, VC, Modelos Predictivos, etc).



CIENCIA ÁGIL

En nuestro desarrollo de la ciencia ágil tomamos atributos de las metodologías ágiles de desarrollo y los vinculamos con la ciencia de datos **con la idea central de llegar a un MVP**. En caso de ser un cliente; obtener resultados. O si eres un estudiante; experimentar en la generación de un modelo de Machine Learning. O en caso de ser un profesional; ayudarte a mejorar KPIs.

- OBJETIVOS/SCOPE: Es el núcleo de la metodología. Se genera tareas precisas con límite de tiempo.
- SPRINT PLANNING: Priorizamos la creación de áreas y estructurar qué necesitamos hacer para llegar al objetivo.
- SPRINT REVIEW: Generamos demos, o entregables para demostrar conclusiones con los expertos/as.
- SPRINT RETRO: Obtenemos feedback y se retroalimenta el proceso en base a cuatro pilares que afectan el proyecto: presupuesto, tiempo, objetivo, y recursos.

Presupuesto

Tiempo

Till

Tiempo

En el inicio de cada sprints en ciencia de datos hay situaciones que no podemos controlar, y parte clave de este proceso es la experimentación y la apertura a la flexibilidad para hacer cambios y ajustarse a las nuevas tareas.





Periodo de tiempo en el que el DataTeam debe llegar a una nueva demo.

CONCEPTOS CLAVES

Demo

Objetivo

Entrega visual que resuma los resultados y conclusiones obtenidas en el sprint para discutir con el usuario real

4

Recursos

Data Team

Recursos involucrados en el proyecto, desde científicos de datos hasta infraestructura adquirida.

Retrospectiva

Tiempo de retrospección, qué obstáculos poder resolver mejor y cómo podemos llegar a nuestro objetivo.



4 pilares

Son conceptos claves que permiten el re-ajuste ágil de las actividades: Budget, Resources, Time, Scope.

