



Tema 2.2

Proceso de análisis de datos

Elaborado por: María Guadalupe Chetla de Gante

Nivel: Alumnos de educación superior

Duración: 10 a 15 minutos

Tema: VII Ingeniería y Tecnología

GRADO DE DIFICULTAD: **M** Medio

Fecha de elaboración: Marzo 2019

2.2 Proceso de análisis de datos





Definición del problema

Un análisis de datos siempre comienza con un problema a resolver. Una vez que el problema ha sido definido y documentado, pasa a la planificación del proyecto de un análisis de datos.





Extracción

Los datos de muestra recopilados deben reflejar tanto como sea posible el mundo real.

Los datos se recopilan de diversas fuentes como dispositivos móviles, computadoras personales, cámaras, dispositivos de grabación; se obtienen de eventos o interacciones entre personas y personas con dispositivos.



Preparación

Se refiere a la obtención, limpieza, normalización y transformación de datos en un conjunto optimizado de datos en un formato preparado, normalmente tabular, adecuado para los métodos de análisis propuestos.





Exploración y visualización

Representa patrones, conexiones y relaciones en los datos de forma gráfica.



Modelado predictivo

Crea o elige un modelo estadístico para predecir la probabilidad de un resultado





Modelo de validación

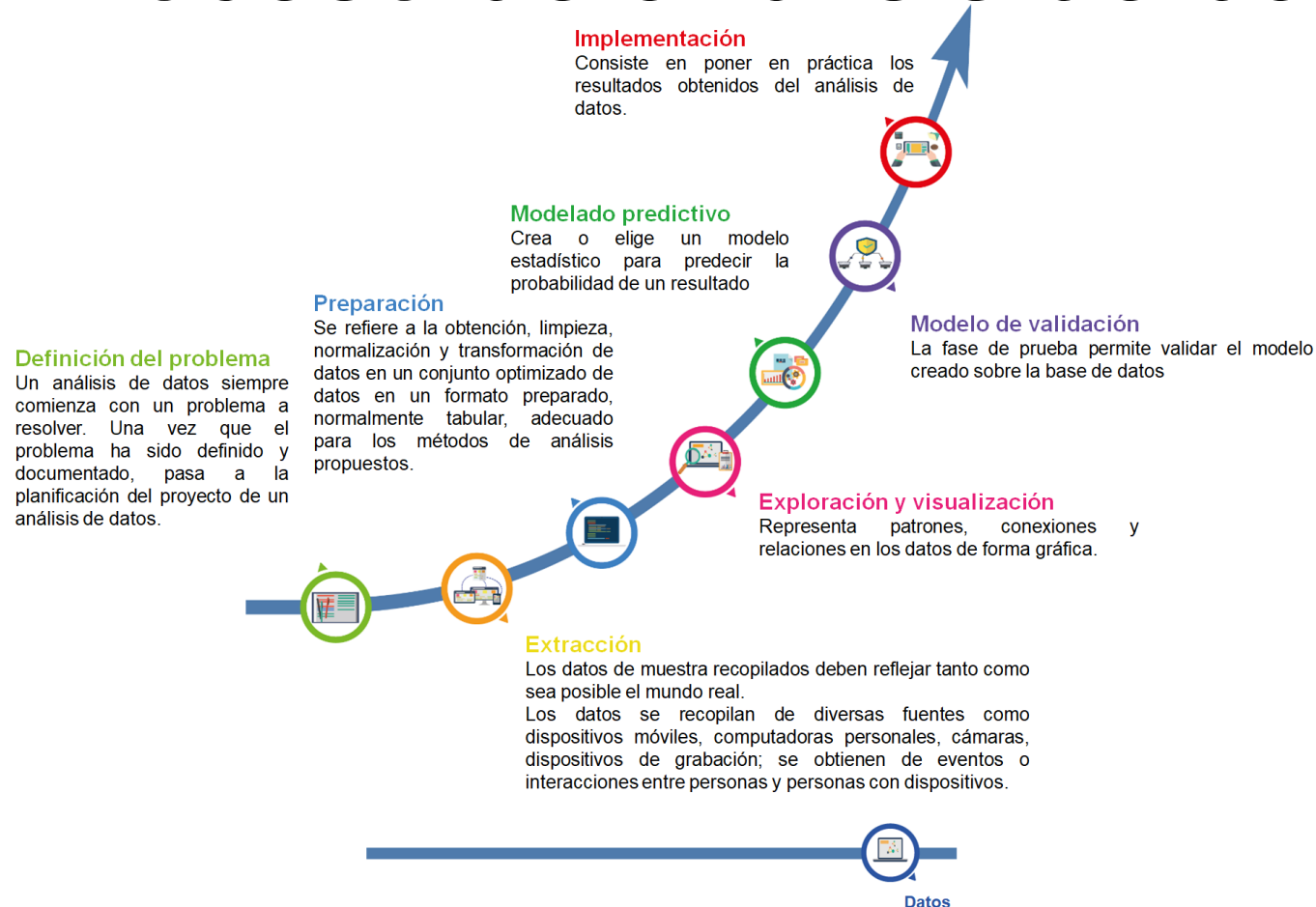
La fase de prueba permite validar el modelo creado sobre la base de datos

Implementación

Consiste en poner en práctica los resultados obtenidos del análisis de datos.



2.2 Proceso de análisis de datos





Referencias

1. Nelli, F. (2015). Python Data Analytics: Data Analysis and Science using pandas, matplotlib and the Python Programming Language. Apress
2. Katemangostar | Fotos y Vectores gratis. (2019). Freepik. Recuperado el 23 Enero del 2019, de <https://www.freepik.es>