

Machine Learning Pipeline

Deborah Dormah Kanubala

February 16, 2020

What is a Machine Learning Pipeline ??

Data Acquisition



Source: <https://www.linkedin.com/pulse/business-intelligence-its-relationship-big-data-geekstyle>

Sources of Data

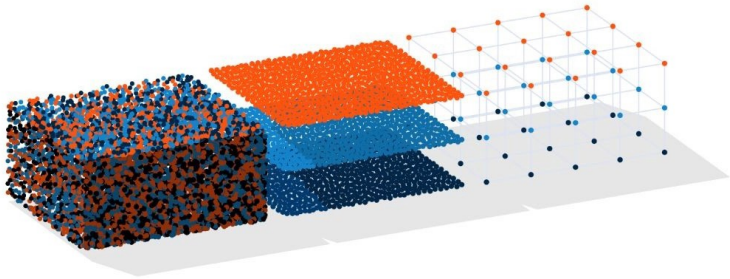
The Kaggle logo, featuring the word "kaggle" in a light blue, lowercase, sans-serif font with a small "TM" trademark symbol.The Yahoo! Finance logo, consisting of the word "YAHOO!" in a white, serif font above the word "FINANCE" in a smaller, white, sans-serif font, all on a purple background.The ZIND! logo, featuring the word "ZIND!" in a white, sans-serif font with colored dots above the 'i' and '!', set against a dark blue background.

Data Analysis



Source: <https://www.linkedin.com/pulse/>

Feature Engineering



Source: <https://www.linkedin.com/pulse/business-intelligence-its-relationship-big-data-geekstyle>

Feature Selection

All Features



Feature Selection



Final Features



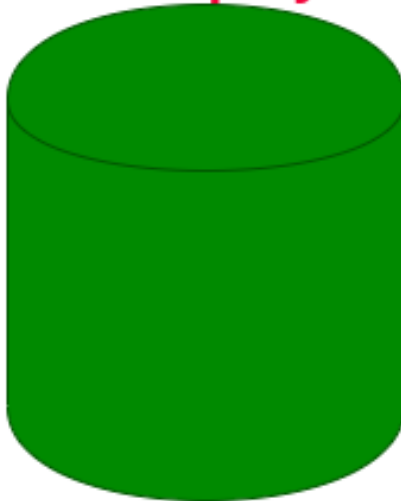
Source: <https://www.kdnuggets.com/2019/10/feature-selection-beyond-feature-importance.html>

Model Development



Source: <https://www.linkedin.com/pulse/business-intelligence-its-relationship-big-data-geekstyle>

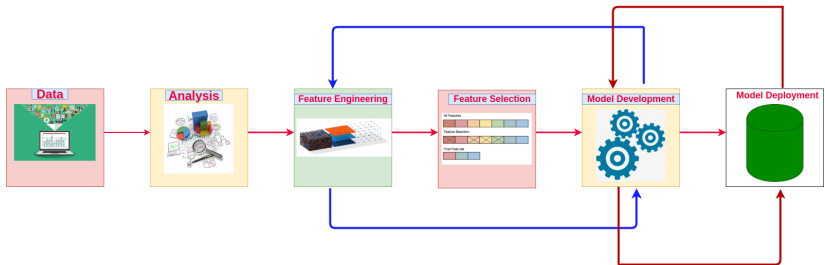
Model Deployment



Source: <https://www.linkedin.com/pulse/>

business-intelligence-its-relationship-big-data-geekstyle

Machine Learning Pipeline



Source: <https://www.linkedin.com/pulse/business-intelligence-its-relationship-big-data-geekstyle>