How Do Companies Use Big Data Analytics in Real World?

The most valuable item for any company in modern times is data! Companies can work much more efficiently by analyzing large amounts of data and making business decisions on that basis. This means that **Big Data Analytics** is the current path to profit!



[**Big Data Analytics**](https://www.geeksforgeeks.org/data-analytics-and-its-type/) is much more objective than the older methods and companies can make the correct business decisions using data insights. There was a time when companies could only interact with their customers on one in stores. And there was no way to know what individual customers wanted on a large scale. But that has all changed with the coming of Big Data Analytics. Now companies can directly engage with each customer online personally and know what they want!

So let’s see the different ways companies can use Big Data Analytics in the real world to improve their performance and become even more successful (and rich!) with time.

### 1. Companies use Big Data Analytics to Increase Customer Retention

Big Data Analytics allows a company to observe customer trends and then market their products specifically keeping their customers in mind.

An example of a company that uses Big Data Analytics to Increase Customer Retention is **Amazon**

### 2. Companies use Big Data Analytics to create Marketing Campaigns

Big Data Analytics is necessary to analyze the customer base and understand what people want so that the marketing campaign is successful in converting more people. This can be done by monitoring the current online trends, understanding customer behavior in the market and then cashing on that to create a successful marketing campaign.

An example of a company that uses Big Data Analytics to create Marketing Campaigns is **Netflix**.

### 3. Companies use Big Data Analytics for Risk Management

It can be used to collect and analyze the vast internal data available in the company archives that can help in developing both short term and long term risk management models.

An example of a company that uses Big Data Analytics for Risk Management is **Starbucks**

### 4. Companies use Big Data Analytics for Supply Chain Handling

Big Data Analytics analyze their raw materials, products in their warehouse inventories and their retailer details to understand their production and shipment needs. This will make Supply Chain Handling much easier which will lead to fewer errors and consequently fewer losses for the company.

An example of a company that uses Big Data Analytics for Supply Chain Handling is **PepsiCo**.

### 5. Companies use Big Data Analytics for Product Creation

Big Data Analytics aims to do for Product Creation. Companies can use data like previous product response, customer feedback forms, competitor product successes, etc. to understand what types of products customers want and then work on that. In this way, companies can create new products as well as improve their previous products according to market demand and become much more successful and popular.

An example of a company that uses Big Data Analytics for Product Creation is **Burberry**, a British luxury fashion house.

## [Announcing TensorFlow Quantum: An Open Source Library for Quantum Machine Learning](http://ai.googleblog.com/2020/03/announcing-tensorflow-quantum-open.html)

“Nature isn’t classical, damnit, so if you want to make a simulation of nature, you’d better make it quantum mechanical.” — Physicist [Richard Feynman](https://www.nobelprize.org/prizes/physics/1965/feynman/biographical/)

Today, in collaboration with the [University of Waterloo](https://uwaterloo.ca/), [X](https://x.company/), and [Volkswagen](https://www.volkswagenag.com/en/news/2018/06/volkswagen-tests-quantum-computing-in-battery-research.html), release of [TensorFlow Quantum](https://www.tensorflow.org/quantum) (TFQ), an open-source library for the rapid prototyping of quantum ML models. TFQ provides the tools necessary for bringing the quantum computing and machine learning research communities together to control and model natural or artificial quantum systems; e.g. [Noisy Intermediate Scale Quantum](https://arxiv.org/abs/1801.00862) (NISQ) processors with ~50 - 100 qubits.

[](https://1.bp.blogspot.com/-Fh22J53FPQs/XmKe2PJG78I/AAAAAAAAFZs/YOD35UMCxcwQbPQaRM8BSWU7niY2I5LoQCLcBGAsYHQ/s1600/image1.jpg)

Under the hood, TFQ integrates [Cirq](https://ai.googleblog.com/2018/07/announcing-cirq-open-source-framework.html) with [TensorFlow](https://www.tensorflow.org/), and offers high-level abstractions for the design and implementation of both discriminative and generative quantum-classical models by providing quantum computing primitives compatible with existing TensorFlow APIs, along with high-performance quantum circuit simulators.  
  
**What is a Quantum ML Model?**  
A quantum model has the ability to represent and generalize data with a quantum mechanical origin. However, to understand quantum models, two concepts must be introduced - quantum data and hybrid quantum-classical models.  
  
**Quantum data** exhibits [superposition](https://en.wikipedia.org/wiki/Quantum_superposition) and [entanglement](https://en.wikipedia.org/wiki/Quantum_entanglement), leading to joint probability distributions that could require an exponential amount of classical computational resources to represent or store.  
The second concept to introduce is**hybrid quantum-classical models**. Because near-term quantum processors are still fairly small and noisy, quantum models cannot use quantum processors alone — NISQ processors will need to work in concert with classical processors to become effective. As TensorFlow already supports heterogeneous computing across CPUs, GPUs, and TPUs, it is a natural platform for experimenting with hybrid quantum-classical algorithms.  
  
TFQ contains the basic structures, such as qubits, gates, circuits, and measurement operators that are required for specifying quantum computations.

**How TFQ works**  
TFQ allows researchers to construct quantum datasets, quantum models, and classical control parameters as tensors in a single computational graph. The outcome of quantum measurements, leading to classical probabilistic events, is obtained by [TensorFlow Ops](https://www.tensorflow.org/guide/create_op). Training can be done using standard Keras functions.  
  
To provide some intuition on how to use quantum data, one may consider a supervised classification of quantum statesusing a quantum neural network. Just like classical ML, a key challenge of quantum ML is to classify “noisy data”. To build and train such a model, the researcher can do the following:

1. **Prepare a quantum dataset**
2. **Evaluate a quantum neural network model**
3. **Sample or Average**
4. **Evaluate a classical neural networks model** .
5. **Evaluate Cost Function**
6. **Evaluate Gradients & Update Parameters**

|  |
| --- |
|  |

A key feature of TensorFlow Quantum is the ability to simultaneously train and execute many quantum circuits. This is achieved by TensorFlow’s ability to parallelize computation across a cluster of computers, and the ability to simulate relatively large quantum circuits on multi-core computers.

# Dimensionality Reduction in Data Mining

Big data is the large scale of data sets that have multi-level variables and that grow really fast. Volume is the most important aspect of big data. Extremely big size of data in big data forms multidimensional datasets. Having multiple dimensions for the in a large data set makes the job of analyzing those or looking for any kind of patterns in the data really hard.

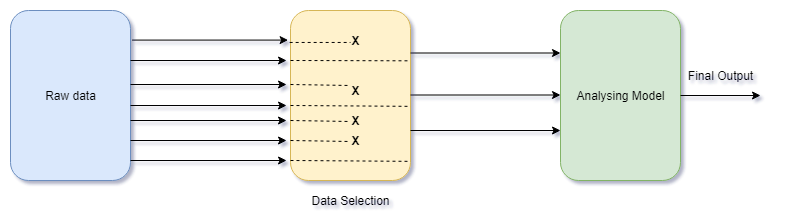
Dimensionality reduction is the process of reducing the number of random variables or attributes under consideration

For an example you may have a dataset with hundreds of features (columns in your database). Then dimensionality reduction is that you reduce those features of attributes of data by combining or merging them in such a way that it will not loose much of the significant characteristics of the original dataset. One of the major problem that occurs with high dimensional data is widely known as the “Curse of Dimensionality”. This pushes us to reduce the dimensions of our data if we want to use them for analysis.

## Techniques of dimensionality reduction

Dimensionality reduction is accomplished based on either **feature selection** or **feature extraction**. Feature selection is based on omitting those features from the available measurements which do not contribute to class separability. In other words, redundant and irrelevant features are ignored. Feature extraction, on the other hand, considers the whole information content and maps the useful information content into a lower dimensional feature space.

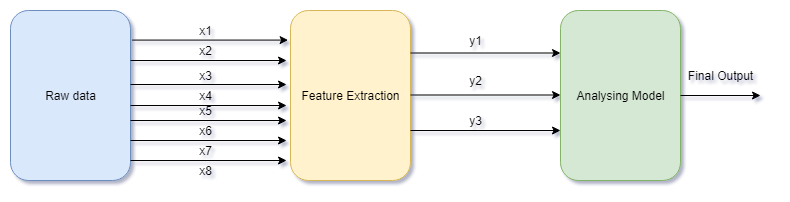
## Feature selection techniques



Feature selection process

As a stand-alone task, feature selection can be unsupervised (e.g. Variance Thresholds) or supervised (e.g. Genetic Algorithms). You can also combine multiple methods if needed.

## Feature extraction techniques



Feature extraction process

Feature extraction is for creating a new, smaller set of features that still captures most of the useful information. This can come as supervised(e.g. LDA) and unsupervised(e.g. PCA) methods.