

OUTLINE

- Introduction: Era of Data
 - DA, ML & AI: What are they?
 - Data Analytics (DA)
 - Machine Learning (ML)
 - Artificial Intelligence (AI)
 - DA, ML & AI: Knowledge Base
 - Statistics & Mathematics
 - Computer Programming
 - Data Visualization
 - DA, ML & AI: Oil & Gas Applications
 - Oil and Gas Data
 - Applications
 - Company Examples
 - DA, ML & AI: How to Acquire the skills?
 - Working Platforms
 - Learning Resources
 - Conclusions



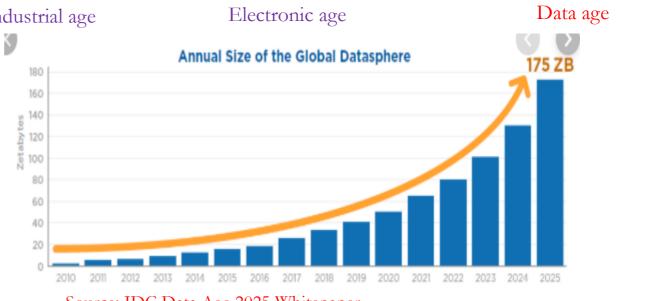
INTRODUCTION

Era of Data

It is the age of data. Data is everywhere!!!



- Technology makes it possible to capture and store vast quantities of data
- Amount of data is growing. 2.5 billion gigabytes of data generated everyday in 2012. [1]



Source: IDC Data Age 2025 Whitepaper

INTRODUCTION

Era of Data

Data versus Information

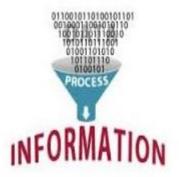






Information: patterns, trends, insights and relationships that underlie data

- Information is what is needed for decision-making and problem-solving processes.
- The process of extracting information from data is generally known as Data Mining (DM). The required body of knowledge is Data Science (DS).
- Data Analytics (DA), Machine Learning (ML) and Artificial Intelligence (AI) are all parts of Data Science



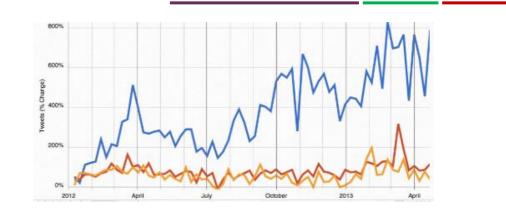
Data Analytics

- Data Analytics: a set of tasks performed on data with the aid of specialized systems and softwares in order to describe or infer the information contained in such data.
- DA tasks include:
 - Data collection identifying sources, subsetting, assembling
 - Data integration combining data from different sources into a common format
 - Data preparation manipulating and organizing data to conform it to analytics requirements
 - Data cleaning fixing quality problems: errors, suspicious, omitted and duplicated data
 - Data modeling fitting data into conventional models: linear regression etc.
 - Data visualization presenting data with charts, graphs to aid information mining
 - Data interpretation

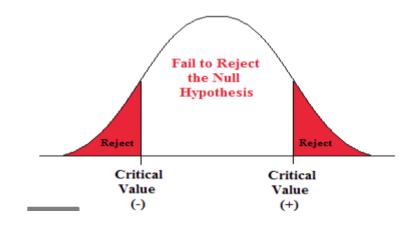


Data Analytics

- Data analytics could be exploratory or confirmatory
- Exploratory Data Analysis (EDA): finds patterns and relationship
 - more like the work of a crime investigator [2]



Confirmatory Data Analysis (CDA): determines if hypothesis are
true or false – more like the work of a judge [2]



Machine Learning

- Machine learning refers to techniques by which machines (computers) are made to analyze data and recognize (learn) patterns and trends in data without relying on standard rule-based programming practices.
- Core of ML: making machines perform tasks based on past experiences passed to the machines as training data

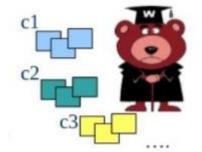


- ML could handle larger and more complicated data and could find patterns more quickly than conventional data analytics tools.
- ML requires less human effort and less assumptions.
- A subfield of computer science and AI

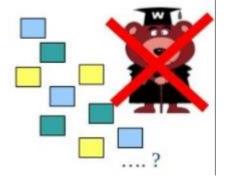
Machine Learning

- Machine learning types/algorithms includes:
 - Regression
 - Classification
 - Clustering
 - Dimensionality reduction
 - Feature extraction
 - Decision tree
 - Random forest
 - K-nearest neigbors.

- Two major approaches to machine learning:
 - Supervised learning train with labeled (input/output) data – to predict output for new inputs. Examples – regression, classification.

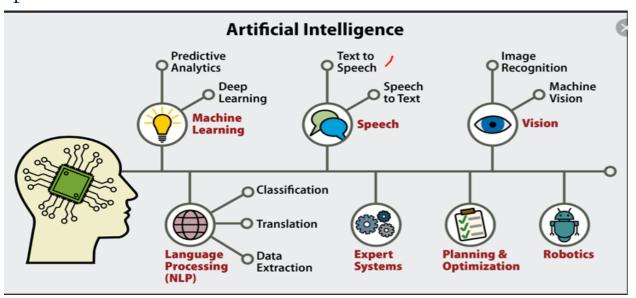


 Unsupervised learning – train with unlabeled data – to detect patterns.
Examples – clustering analysis.



Artificial Intelligence

- Artificial intelligence (AI) is a sub-discipline of computer science with focus on developing machines (computers) capable of performing tasks requiring human intelligence.
- AI systems work by learning (from data) and improving performance over time.



Source: datamation.com



Major application areas of AI includes: machine learning (ML), natural language processing (NLP), computer vision (CV), speech recognition, virtual assistants etc.

DA, ML & AI: KNOWLEDE BASE

Certain domains of knowledge are necessary to attain competence in DA, ML & AI

Mathematics and Statistics

Matrix (linear) algebra: forms the basis of [1 3 5 7] • [2 4 6 8] many ML techniques



- Multivariable calculus
- Data distributions, statistical estimators, hypothesis tests



Computer Programming

- Often, there will be the need to write codes to implement workflows for specific datasets and objectives.
- For automation and flexibility
- Statistical programming language like R and data querying language (SQL)

Data Visualization

- Turning numeric data to visual objects aids communication and decision making process.
- Understanding the principles of visual encoding of data.



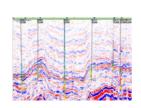
Creativity, critical and intuitive thinking, and problem solving skills also needed

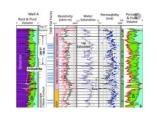
The oil and gas sector is prolific in data generation and measurements.

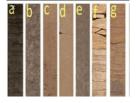
Data items include:

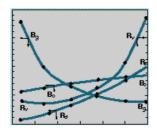
- Seismic surveys
- Well logs
- Core data
- Fluid data
- Pressure and temperature data
- Production test data
- Production and injection data volumes, rates
- Drilling performance data
- Pipeline inspection data
- Equipment maintenance and failure data
- Crude trading data

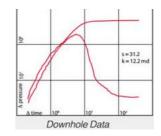
Oil and Gas Data

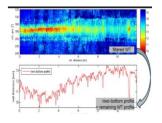
















Applications

- Resource estimation using AI to predict reserves
- Subsurface modeling and analysis using ML to find correlations among geological variables
- Wells analysis historical and real-time well data analysis
- In-fill drilling using AI to optimize number and locations of wells
- Well trajectory optimization sensors and AI tools



Applications

- Orill bit performance monitoring
- Fluid properties using ML algorithms to develop correlations
- Production optimization using computer vision to analyse
- Equipment monitoring and maintenance pumps, compressors; using AI to predict unexpected failures
- Crude logistics and trading oil grade data, price and market data



Applications

- Investment portfolio management using AI/ML to evaluate opportunities
- Mealth, Safety and Environment using AI/ML to perform root cause analysis
- Inventory management using AI/ML to predict demands, stock levels, warehouse utilization etc
- Finance using AL/ML for cost allocations



Company Examples

Baker Hughes & NVIDIA – to use AI-based to help clients gain insights from data





Total and Google – AI systems to help geologists interpret subsurface images





ExxonMobil and MIT – autonomous underwater exploration robots





BP invested in start-up Belmont Tech – developing a geoscience platform - "Sandy"

Working Platforms

Here are some platforms on which DA/ML/AI workflows can be implemented. Most of them are open source platforms



Python – a programming environment – popular among data science practitioner



SAS



SQL – database management



IBM Watson – AI platform



Working Platforms

Scikit-learn – ML platform on Python





ggplot2 – visualization in R



Matplotlib – visualization in Python matplotlib



Apache Hadoop – Big data platform





Tableau – analytics and data visualization



Working Platforms

Google Analytics



MSExcel



Knime



Weka



CONCLUSION

Learning Resources

There is abundance of learning resources from which DA, ML & AI skills could be learnt. Many of these resources are open access. You could earn recognized certificates and badges on some of these platforms

Data Camp



Coursera



IBM – Digital Nation Africa







Learning Resources

Big Data University



Cognitive Courses







MIT OCW



Kaggle – a community of data scientists – contain lots of data



DA, ML and AI are technologies that would certainly impact the oil and gas value chain in the emerging future. In order to leverage on this, young graduates should acquire these skills.

THANK YOU

FOR

LISTENING!!!

BIBLIOGRAPY

- Ralph Jacobson (2013). IBM Consumer Products Industry Blog, https://www.ibm.com/blogs/insights-on-business/consumer-products/2-5-quintillion-bytes-of-data-created-every-day-how-does-cpg-retail-manage-it/
- 2. Tukey, J. W. (1977). Exploratory data analysis. Reading, MA: Addison-Wesley.