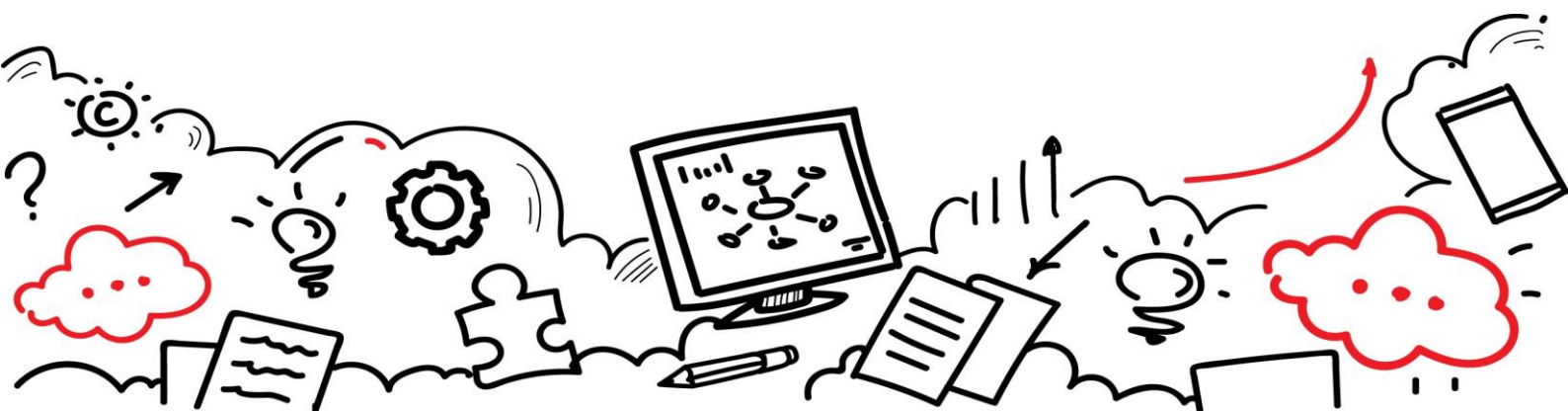


HOW DATA SCIENCE IS CHANGING THE BUSINESS LANDSCAPE

BRIGHTRED RESOURCING LIMITED
NOVEMBER 2019

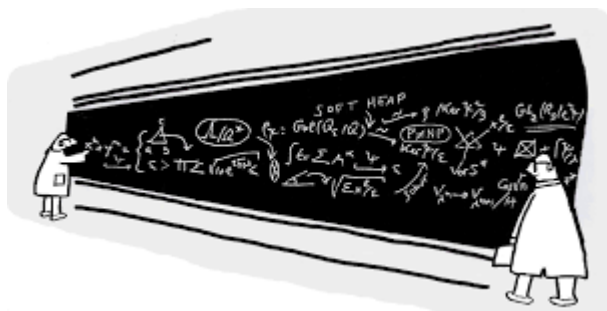


HOW DATA SCIENCE IS CHANGING THE BUSINESS LANDSCAPE

With Artificial Intelligence (AI) and Machine Learning (ML) taking centre stage of the IT space, Data Science has become the buzz word in recent times. Since William S Cleveland first coined the term 'Data Science' in 2001, this technology has been rapidly innovating. AI, Smart Apps, Edge Computing, ML based security solutions are some of the key Data Science solutions that are making waves in recent times.

AN OVERVIEW OF DATA SCIENCE

Data Science is a multi-disciplinary field that extracts useful and actionable insights from both structured and unstructured data using various scientific algorithms, processes and methods. It is actually a unified system that combines data mining, data analytics, statistics and machine learning to discover hidden patterns in raw data.



Data Science is an umbrella term that covers various types of data analytics solutions. Simply put, it is the study of data. Here are some of the key components of Data Science.

A) PREDICTIVE ANALYTICS

This data model algorithm analyses current and historic data to predict future events. For instance, lending institutions should know if a customer will be able to repay a loan on time. Using predictive analytics, it is

possible to assess behavioural patterns of the customer and make informed decisions.

B) PREDICTIVE CASUAL ANALYTICS

Casual analytics is a method wherein the algorithm tries to find the reason for the hidden patterns of the data instead of just proving the cause-effect relationship.

C) PRESCRIPTIVE ANALYTICS

This data model algorithm has the ability to not only predict future events but also has the ability to dynamically make decisions and apply wherever applicable. Google self-driving vehicle technology is a notable example of this algorithm.



D) SUPERVISED ML ALGORITHMS

When you already have data to build models, it is called supervised machine learning algorithms. This algorithm is best suited for fraud prevention programs.

E) UNSUPERVISED ML ALGORITHMS

Here, there is no existing data to build models. So, you need to find hidden patterns and make predictions.

THE NEED FOR DATA SCIENCE

The advent of Internet technology has brought a digital revolution that resulted in a massive influx of data. As more and more products and services entered the network, even more data was generated. As such, the first job for businesses was to store and manage big data in an organised format. Data management frameworks such as Hadoop have solved this problem. Using these frameworks, businesses were able to store and manage big data. The next big task was to derive









business value from this data. What is the use of storing huge volumes of data that are of no business value? To derive business value from this data, it should be processed. Data Science does this.

In every business organisation, 70%-80% of the data is in unstructured format while the remaining 20-30% accounts for structured data. Structured data can be processed using regular Business Intelligence (BI) tools. However, processing unstructured data is a challenge. This data contains information generated from system logs, audio/video streaming, IoT and other analog instruments data. The sudden increase in mobile and IoT networks adds to this challenge. As such, businesses need robust and complex algorithms and analytics tools to process this raw data.



Data Science helps businesses to analyse raw data and discover hidden patterns to derive business insights while predicting future events. By analysing data models, data science enables businesses to understand customer behavioural patterns, purchase history and browsing history, so that they can offer customised services to each customer. In addition, Data Science enables businesses to predict customer shopping requirements and offer products and services with better precision.

Here are a few examples of applications, already using the data science models.

-  Email Spam Filter
-  Abusive Content Filter
-  Recommendations Engine
-  Robotics
-  Automated Piracy Detection
-  Digital Assistants
-  Self-driving Cars
-  Search Engine Results

DATA SCIENCE Vs BI

A Data Scientist is not to be confused with a Data Analyst. Similarly, BI is different from Data Science. BI deals with studying previous data to understand business progress. Using a dashboard, you can generate customised reports to understand critical business insights. Data Science takes this concept to a new level. It discovers hidden patterns in data and predicts future events. While BI primarily uses visualisation and statistics, Data Science uses complex algorithms such as ML, neuro-linguistics programming (NLP), graph analysis, etc.

BigML, Weka, RapidMiner are some of the tools used for Data Science processes.

Parameters	Data Analyst	Data Science
Overview	Processes data and offers clear insights into business processes.	In addition to extracting meaningful insights from data, predicts future events
Skills	Statistics, Visualizations, Data Mining, Data Wrangling	Statistics, Visualizations, Data Mining, Data Wrangling, machine learning algorithms, software engineering
Knowledge	Knowledge of Hadoop-based analytics	Knowledge of analytic functions, median, rank etc.
Technologies	Tableau, Java, Hadoop, AWS, Hive etc.	R, Python, AWS, MySQL, Hive, Machine Learning libraries
Jobs	Answers questions formulated by the business team	Creates questions and answers them

THE STATE OF DATA SCIENCE MARKET IN THE UK

- BUSINESS GROWTH -



The Data Science market is rapidly evolving. According to [MHR Analytics](#), **80% of companies in the UK are planning to hire a data scientist or seek data consultancy in 2019.** 47% of companies believe that their workforce lacks the digital skills to compete with others in the next decade.

Similarly the [Royal Society](#) reports that the number of **Data Science jobs in the UK tripled in the last 5 years.** In 2013, British employers posted 8,157 Data Science jobs and in 2018, they posted 27,033 jobs, showing an increase of 231%. The demand for Data Science jobs increased by 563% in Northern Ireland followed by 269% in North West UK and 112% in South West UK.

The **average salary for data scientists is £64,376**, which is an increase of 22% in this period. Similarly, ML, scripting languages, SQL and Big Data are the most frequently required skills for British employers.

Europe needs 346,000 data scientists by 2020, according to IBM's [The Quant Crunch](#) report. 28% of digital jobs will be data science related by 2020. However, most of the jobs lie vacant for more than 45 days owing to shortage of Data Science skills. Almost half of European companies

are now struggling to fill Data Science positions in their companies.

These numbers show a change in the nature of work and the dominance of Data Science jobs in the coming years.

HOW DATA SCIENCE ADDS VALUE TO YOUR BUSINESS?

Today, Data Science has become an essential component of business processes. It adds great value to business operations, including the following key benefits of Data Science for businesses.

Say No to Fraud Incidents

With prescriptive and casual analytics, businesses are now able to identify future fraud events and risks factors to proactively resolve them with ease. Using machine learning algorithms such as predictive fraud propensity models, businesses are able to identify unusual data patterns and create alerts to mitigate risks and frauds.



Say No Resource Wastage

Data Science enables organisations to identify resource utilisation at different times and thereby allows them to optimise resource usage. With dynamic provisioning of resources that are augmented by predictive outage incidents, businesses can optimise resources while minimising outages for a better ROI. An example of this is, the airline industry which has been experiencing hard times owing to slow down, increased fuel price and discounted offers. They were unable to manage occupancy ratio to establish operating profits. With data science, airlines are now able to optimise occupancy, choose the shortest routes and predict flight delays to achieve operational profits.

Increased sales with relevant offers

Data Science enables businesses to understand customer shopping patterns at the granular level. With predictive analysis, businesses are able to predict what customers would buy in the coming days and offer them relevant products at customised prices at the right time to



increase sales. In addition, it identifies new markets and opportunities for businesses.

Improved Customer Service



Using AI-powered chat bots, businesses now offer a higher level of customer service to customers 24/7. While

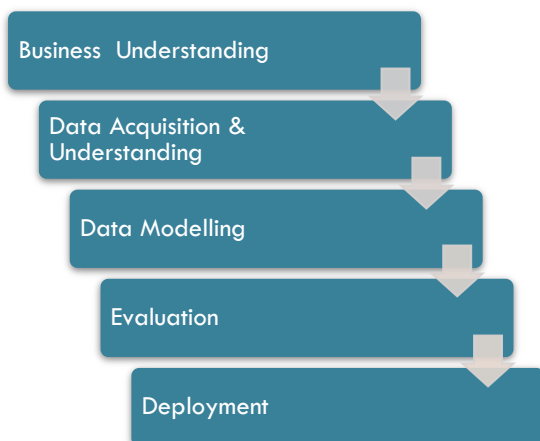
it optimises resources, customer service is always available and is customised to each user. For instance, [eBay chatbot](#) is an advanced solution that enables customers to purchase any product at the best price. Chatbot will search for that product on all channels and give you the best price. Similarly, [niki.ai](#) is India's first fully-automated AI App that enables users to make recharges, payments and seamlessly perform all day-to-day transactions with ease.

Better Decision Making

Data Science helps the management to get clear insights into business processes, identify new opportunities, create goals, test decisions, and recruit the right talent and help in adapting best practices. Simply put, Data Science offers quantifiable and data driven evidence for better decision making.

HOW DATA SCIENCE LIFECYCLE WORKS?

Each business has different data and business requirements. As such, Data Science and models vary with each business vertical. However, the entire Data Science lifecycle can be categorised into five standard phases.



1) BUSINESS UNDERSTANDING

The first phase of the Data Science lifecycle is understanding business processes. Acquiring huge volumes of data doesn't offer any value but deriving clear insights does. It all boils down to asking the right questions. To do so, define your business objectives and goals that the Data Science model should achieve. Then, you can frame the right questions.

For instance,

- How much AWS computing power is being utilised during peak hours and how much computing power lies idle in lean hours?
- What information in your marketing campaigns is more relevant to your customers?
- What is the churn rate of the company?
- Is the company offering discounted prices to customers who would normally choose a brand even with higher prices?
- What is the average room occupancy of hotels in a City?

By asking the right questions, you will know what type of data is required. It is essential to determine the budget as well as the availability of resources; people, technology and time.

2) DATA ACQUISITION AND UNDERSTANDING

Now that you have determined business requirements, you need to break it into smaller components. Then, it becomes easy to source and collect data for each component. Data should be collected from internal databases as well as external sources such as social media, web scraping and third-party solutions. So, ingest the data from the source into the target environment and explore it to check if it serves the analytics goal.

As real-world data comes with several discrepancies, you need to first visualise data, cleanse it and prepare reports before building data models. It is important to setup a data pipeline for streamlined and automated flow of data to refresh data and score new data.

3) DATA MODELLING

Based on your business objectives, you need to choose the modelling algorithms so you can draw relationships between variables using statistical

formulas, visualisation tools, Exploratory Data Analytics (EDA), together with model planning tools such as R, SAS/ACCESS, SQL Analysis Services etc. As a best practise, split the data randomly while building models and evaluate and test the data sets and compare results from alternative models to determine the best solution.

There are three modelling algorithms:

- a) **Supervised ML algorithms:** Prediction based on existing and historic data
- b) **Unsupervised ML algorithms:** The goal is to organize data into clusters to derive a structure.
- c) **Learning reinforcement:** Here, the algorithm chooses one action for each data point.

Accuracy, linearity, training time, parameters and features are some of the aspects to consider while choosing a model algorithm, using model building tools such as WEKA, Statistica, Matlab, Alpine Miner, SAS Enterprise Miner etc.

4) EVALUATION

For every model that is used, a standard report template is prepared. The next phase is to evaluate the outcome using predefined metrics. These metrics depend on the business objective of the data science project. For instance, if you are trying to provide automated spam protection, metrics to look at are accuracy, log loss, Area Under the Curve (AUC) etc. Comparing results of alternate methods will allow you to determine the one that provides more precision.

In addition, check if the test model is ready to be deployed. Check if the test data gives you the confidence that it answers all the questions; or should you do feature engineering or try an alternate algorithm. Based on your evaluation; repeat the steps from 1-4.

5) DEPLOYMENT

Once the performing model data sets are determined, the next phase is to deploy them to your Production Environment for application consumption. These models are exposed via an open API interface to applications such as online websites, social media programs, backend databases, dashboard, spreadsheets, business

apps etc. As a best practice, you should create a status dashboard, deployment modelling final report and the final solution document for further reference.

Based on the production environment, you can recode ML models. For instance, if the models are written in Java and the production environment favours Python, you can recode them. Finally, confirm if the model pipeline and results meet customer needs with acceptable accuracy. At this stage, the data model is handed over to the operations entity and documentation is reviewed and finalised.

HOW DIFFERENT INDUSTRIES ARE LEVERAGING DATA SCIENCE

Data Science is quickly turning into necessity for every industry in recent times. Here are five examples of industries that are leveraging Data Science.

RETAIL



The retail sector deals with products and customers. The primary focus of a retail business is to deliver higher customer satisfaction and retain customers while gaining new ones. Using Data Science, the retail sector is now able to study and analyse product reviews, analyse call center records to discover hidden patterns and scour through the social media stream to get the pulse of customers. Using this data, different data science models are prepared with regard to current and future business happenings. By understanding current processes, business get clear insights into business processes. At the same time, predictive and prescriptive analytics enable them to identify customers that are about to leave their business and proactively take measures to retain them. With churn prediction solutions, you can forecast churn rates and reduce them.

Similarly, a recommendations engine helps businesses to suggest relevant products to customers at the right time. E.g.: Amazon recommendations engine.

amazon.com

Recommended for You

Amazon.com has new recommendations for you based on [items](#) you purchased or told us you own.

MANUFACTURING

The manufacturing sector needs Data Science the most as it deals with various types of data sets. Right from IoT sensors, analog devices, computing devices and multimedia data, the industry deals with a great deal of unstructured data. As such, Data Science enables them to efficiently process it and gain useful insights. Firstly, predictive maintenance is a big advantage to the industry. Using predictive analysis, data scientists study the performance of each machine, warranty claims and forecast problems that arise in future. Using this data, companies perform predictive maintenance to eliminate outages while efficiently managing overproduction, logistics, and inventory. It also reduces idle time. Predictive maintenance is done on time-based or usage based.



Robotics is a big asset for the manufacturing industry. With Data Science that offers Artificial Intelligence, organisations are able to use robots to perform manual and repetitive jobs with ease. In addition, companies can efficiently manage supply chain relations while offering dynamic and optimised pricing with regard to global market pricing. With inputs for new products and development techniques and energy efficient sustainable models, data science helps manufacturing industries to improve production at optimised costs.

For instance, General Electric (GE) produces high pressure nozzles and shrouds that are fixed at the hottest portion of aircraft engines. As such, GE created an Asset Performance Management (APM) program using data science for predictive and conditional maintenance of aircrafts.

BANKING & FINANCIAL

Banks and financial institutions consistently use Data Science to study credit risk management. To offer credit to customers, banks analyse customer transactions, customer ratings, social media profiles, data from



payment gateways, and ecommerce shopping trends to identify the trustworthiness and creditworthiness a customer. Based on this data, financial institutions are creating customer credit scores and credit profiles to offer instant and paperless loans. Personalised marketing with recommendation engines bring more sales and increased customer satisfaction. Another striking advantage of data science is determining the lifetime value of customers. Banks are now able to create beneficial relationships with valued customers to increase revenues and sustain growth.

Moreover, by analysing business processes, banks are able to detect fraud and take action accordingly. Banks are also able to do customer segmentation to create customised products for customers based on the location, income, business etc. Investments banks operate risk modelling to regulate financial activities, manage acquisitions and mergers, provide capital for financing corporates, restructuring organisations etc. With churn prediction solutions, banks are able to identify dissatisfied customers and are able to act quickly to retain them.

TRAVEL



Data Science has transformed the travel industry from customer segmentation methods to getting a 360° view of a customer. By tracking the online footprint of a customer and combining it with itineraries, location, CRM data etc., travel companies create a 360° profile of each customer.

Using data modelling and predictive analysis, companies offer customised offers to each customer. Similarly, companies predict demand and realign their marketing strategies accordingly via recommendation engines.

For instance, when a customer checks exotic places from London, travel agencies offer best hotels in that destination area, cheap flight tickets and discount vouchers, etc. Travel agencies can also

optimise travel routes to save time and costs. Using predictive models, the companies increase prices on weekends and special occasions while decreasing prices and offering great discounts in off season.

Using sentiment analysis where companies read customer reviews and determine customer emotions, organisations can check how satisfied the customer is with the company and proactively manage campaigns. Customer service becomes easy and effective via AI-powered travel bots.

HEALTHCARE

The healthcare industry significantly benefits from data science as it deals with data-heavy images and different types of data sets. MRI, CT, X-ray etc., produce data-heavy images while IoT devices constantly generate data. With deep learning machine algorithms, healthcare institutions analyse this data to come up with accurate diagnosis of diseases. For instance, healthcare institutions use Hadoop's MapReduce to determine optimal parameters of various diagnosis tasks. Other tools include iDASH, MPI etc. Besides, data science can offer predictive medicine by predicting future diseases based on patient history and current parameters and lifestyle.



Medical researchers that study DNA are using Data Science techniques to identify the connection between genetics and diseases as well as how the body responds to various drugs. This data can be used to predict the chances of future diseases. Drug discovery is another area that needs data science. A normal drug after thorough testing requires 10-12 years to becoming authentic. ML algorithms shorten this time. Besides, healthcare institutions can use AI-powered assistants to offer basic support to patients. In addition, they can provide tips for a healthy lifestyle. Managing patient records is also made easy.

DATA SCIENCE AS A CAREER

Data Science is a multi-disciplinary entity. Having a certificate in statistics doesn't serve the purpose. You need to have soft skills and hard skills. Having a thorough knowledge of statistics and mathematics to

analyse and visualise data, you will also need to have a good understanding of ML algorithms and domain expertise to implement them.

Data scientist jobs are in great demand in recent times. According to a [LinkedIn Emerging Jobs 2019 Report](#), **'Data Scientist' is one of the fastest growing jobs in 2019.**

Secondly, the market is not saturated yet as there are no complete data scientists in the field. Besides, Data Science is now being implemented across all business verticals, and therefore brings greater opportunities for data scientists. Moreover, it is the most lucrative job as it highly paid. In addition to bringing a higher pay package, Data Science challenges you every day to make data better and products smarter. It is a constantly evolving industry.

THE Bottom Line

Being relatively new, Data Science seems to be a complex concept for many people. However, it doesn't push non-data scientists away from understanding its potential and the business value that it offers to organisations. The key requirement for data science is intense intellectual curiosity and deep thinking. You should always be keen to learn new things by asking questions and finding ways to get answers to those questions. The idea of discovery hidden patterns and secrets from data is what motivates data scientists in their job.

Brightred is one of the few specialist recruitment solutions companies that has a thorough understanding of Data Science. BRIGHTRED provides highly experienced data scientists for businesses of all verticals.

Contact us right now to enrich your data journey and make better business decisions!

CALL US FOR A CHAT!

0203 8000 555

www.brightred.com