



Semester 2 Alternative Assessment 2019/20

End-of-Module Assignment

MS5114 - Advanced Programming for Business Analytics

Student Name:	Jayakarthis Boovendran
Student ID:	19230487
Course:	MBY – MSc Business Analytics
Email:	J.Boovendran1@nuigalway.ie
Assignment Topic:	BIS Advanced Programming

Declaration

"In submitting this work I confirm that it is entirely my own. I acknowledge that I may be invited to undertake an online interview if there is any concern in relation to the integrity of my submission"

Signature:

Date: 11 May 2020

A handwritten signature in black ink, appearing to read "Jayakarthis".

Jayakarthis Boovendran

Question #2 : Data Storytelling for Improved Decision-making

I. Introduction

It is an understatement to declare the modern business environment as 'dynamic and uncertain'. With digitally empowered customers, there is a significant demand for businesses to adapt and respond to changing customer expectations continually. There is a compelling need for Businesses to understand user expectations, grasp the patterns, make data-driven decisions and execute action plans in as short a time as possible. Further, it is indisputable to state that the longevity and success of any business depend primarily on the decisions they make. Therefore, this study focuses primarily on the role of storytelling in effective decision-making and how data visualization enhances the art of storytelling.

II. Data Analysis and Data Storytelling

Data Analysis and Data Storytelling are different concepts; however, highly interrelated. I would define **Data analysis** as, a process or an experiment to derive solutions to a problem, research question or a hypothesis by analysing relevant data, applying appropriate algorithms on them and finally communicating the findings to the concerned for informed decision-making. Contrarily, **Data storytelling** is a procedure for communicating the results of the data analysis to the decision-makers. In order to expedite the decision-making process, it is essential to translate the data insights or any other technical aspects into layman's terms (Rouse 2015). Therefore, it is critical to **combine these concepts to improve decision-making**.

III. Is Data Storytelling necessary in Business Analytics?

"People hear statistics, but they feel stories" – Brent Dykes ([Forbes.com](https://www.forbes.com))

Data Storytelling has become a quintessential part of the decision-making process as it bridges the gap between complex information and the stakeholder's ability to interpret it correctly. Dykes (2016) claims that businesses can realise the full potential of data analytics only when they translate the derived knowledge into valuable business outcomes. Further, He argues that 'Data Storytelling' is the ultimate weapon that could facilitate this communication.

IV. Transforming Data into Visual Stories

a) Key Components

Dykes (2016) identifies the three critical elements of Data Storytelling as **data**, **visuals** and **narrative**. A typical Storytelling combines all these elements to convey the message effectively.

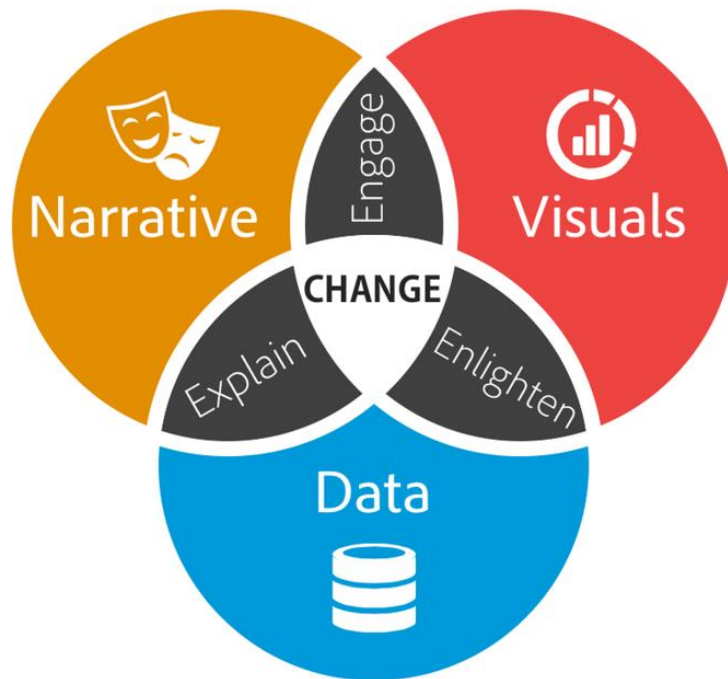


Figure 1 : Key Components of Data Storytelling (www.forbes.com)

b) Data Storytelling – The Process

Lee et al. (2015) illustrate the underlying process of storytelling with three significant steps; **Exploration of data**, **Creating a story** and **Telling the story**. It is critical to consider the target audience in all three steps.

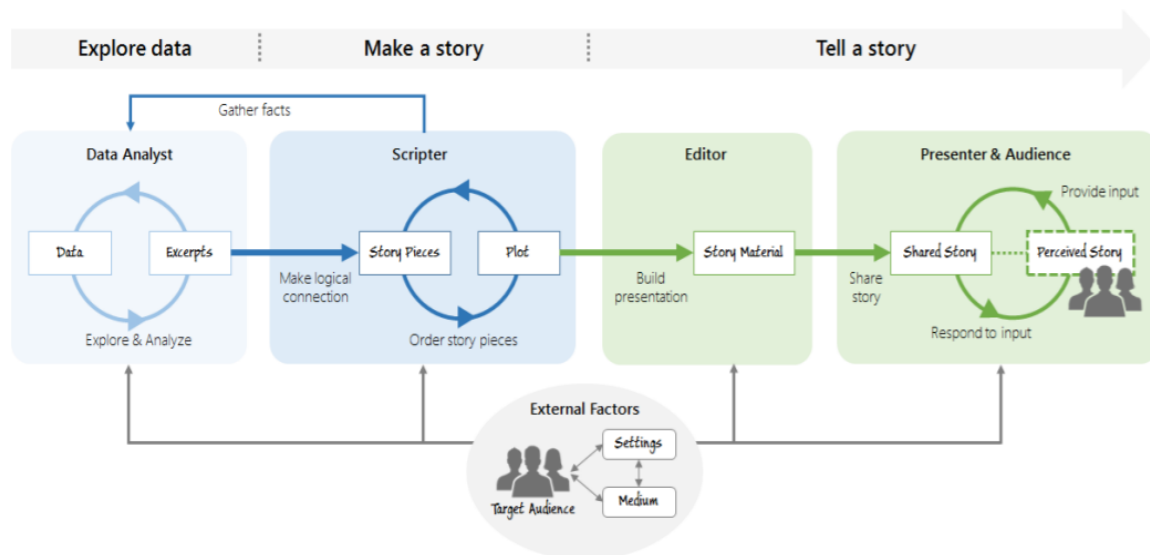


Figure 2: Visual Data Storytelling Process (Lee et al. 2015)

1. Data Exploration

This step concerns with the selection of essential data insights that are relevant for framing the story. Employing Data Visualizations help showcase the appealing aspects of the story. Besides, performing an exploratory data analysis will help identify the critical data points, and the KPIs required for the story. However, it is crucial to determine the inclusion and exclusion criteria in order to deliver the facts precisely.

2. Making a Story

The second step is to construct the storyline by assembling the selected data insights in a way that offers a logical flow and a better understanding of how the insights are connected. Furthermore, the storyline should be intriguing, compelling, and eventually should convey the message.

3. Telling a Story

The final step is to present the story to the target audience in a compelling and stimulating way. It comprises of steps like building a presentation material, sharing the story through the material, receiving feedback from the audience and finally accommodating the changes to the story.

V. Data Visualizations with Python

Data visualization play an incredible role in Storytelling by providing the building blocks to create engaging narratives (Kosara and Mackinlay 2013). Python allows users to create meaningful, expressive and aesthetically pleasing visuals using its libraries like **Matplotlib**, **Seaborn** and **Folium**. Moreover, Python offers a set of advanced libraries like **plotty**, **mpld3** and **pygal** to support interactive visual plots. In addition to that, these libraries are compatible with pandas dataframe and numpy, also, has in-built statistical functions to analyze data.

VI. Conclusion

An exceptional Data Storytelling should express complex ideas with clarity, precision and efficiency in order to influence the decision-making (Laskowski,2018). Leveraging the use of appropriate visualizations for Storytelling will help expedite the decision-making process. Therefore, ***“A picture is worth a thousand words....but only if that picture is designed well”***

Reference

- Bongshin Lee, Nathalie Henry Riche, Petra Isenberg, Sheelagh Carpendale. *“More than Telling a Story: A Closer Look at the Process of Transforming Data into Visually Shared Stories”*, IEEE Computer Graphics and Applications, Institute of Electrical and Electronics Engineers, 2015, 35 (5), pp.84–90. ff10.1109/MCG.2015.99ff. ffhal-01158445f
- Dykes, Brent (2016), *“Data Storytelling: The Essential Data Science Skill Everyone Needs”*, Forbes.com, Available at <https://www.forbes.com/sites/brentdykes> [Accessed 07 May 2020]
- Kosara, R. and Mackinlay, J., *“Storytelling: The Next Step for Visualization”*, in Computer, vol. 46, no. 5, pp. 44-50, May 2013, doi: 10.1109/MC.2013.36.
- Laskowski , N. (2015), *“Data storytelling is the next big thing in collaborative computing”*, Available at <https://searchcio.techtarget.com/podcast/Data-storytelling-is-the-next-big-thing-in-collaborative-computing> [Accessed 07 May 2020]
- Rouse,M. (2015), *“Data Storytelling”*, Available at <https://searchcio.techtarget.com/> [Accessed 07 May 2020]

Question #4 : Machine Learning – The Future of Businesses

I. Introduction

In the modern business environment, many companies are investing in technologies that could help better understand customer expectations, analyse and predict customer behaviour, explore new revenue opportunities and eventually facilitate competitive advantage in the market. Machine learning is one such technology that could effortlessly perform tasks that are beyond human capabilities (Shwartz and Shai 2014). Therefore, the study focuses primarily on machine learning and its application in Businesses.

II. Machine Learning

Samuel (1959) defines Machine learning as an ability of a computer to learn automatically without being programmed explicitly. From my point of view, Machine learning or automated learning is a process by which a machine responds automatically to a problem or a scenario based on the knowledge it obtained from learning an available set of data. The machine relies on patterns and inferences rather than explicitly defined instructions.

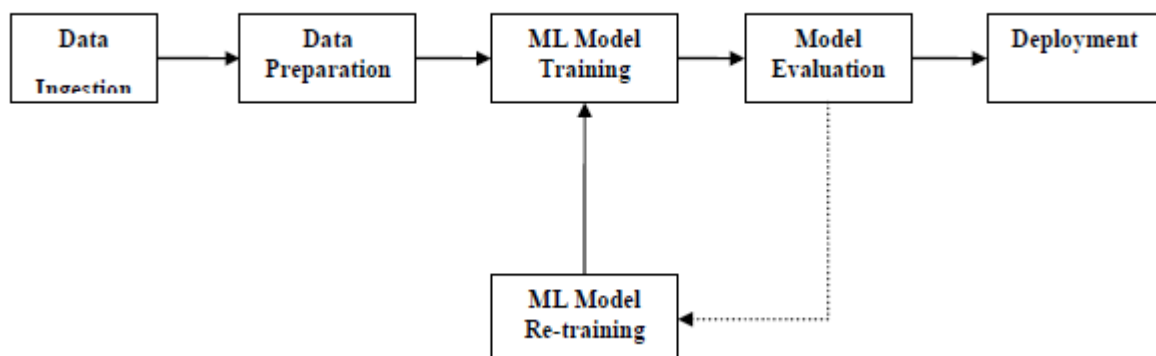


Figure 3: Machine learning Process (www.tutorialspoint.com)

III. Categories of Machine Learning

i) Supervised Learning

The model uses one or more independent variables (predictors) to predict a target variable. Here, the model is trained with labelled data until it attains a desired level of accuracy in predicting the target. In Supervised Learning, we can observe and direct the execution of a machine-learning model (Das 2017).

ii) Unsupervised Learning

The machine uses unlabeled training data. Hence, the system understands and learns by itself from the available data.

iii) Reinforcement Learning:

The primary objective is to train the machine to make accurate business decisions. Here, the machine trains itself using the trial and error method by interacting with the environment and learns continually.

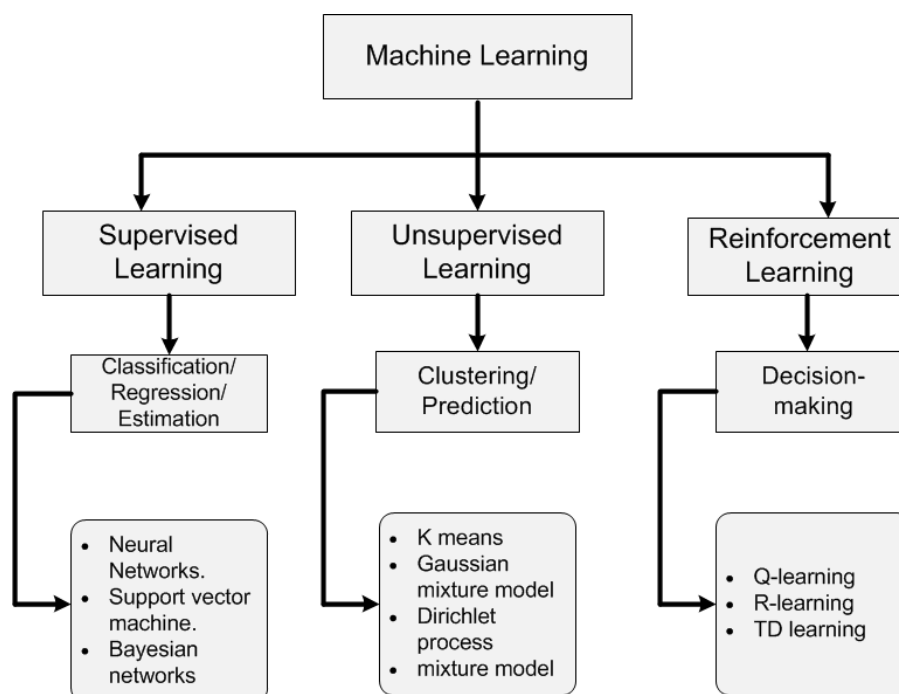


Figure 4 : Types of Machine-learning (www.researchgate.net)

Discussion:

However, in real-time, Supervised and Unsupervised algorithms are widely used in Business Applications than Reinforcement algorithms. Some frequently used machine learning according to Ray (2017) algorithms are;

- **Supervised Learning:** Regression algorithms (Linear, Logistic), KNN Classification, Decision Tree and Random Forest algorithms
- **Unsupervised Learning:** Apriori algorithm for Basket Analysis and K-means Clustering algorithm

IV. Business Applications of Machine Learning

The most common machine learning applications are churn modeling ,customer lifetime value modeling, recommendation engines, dynamic pricing, customer segmentation and image classification Castle (2017). Apart from the above, it is employed in fraud detection, Personalizing the customer experience and streamlining IT processes.

Case Study on Dell:

Dell implemented Machine learning to harness the power of their emails to attain better response from their customers. Eventually they observed an increase in CTR upto 50% and an increase in customer response upto 46%. Further, they also had an 77% increase in the number of add-to-carts.

V. Implications of Machine Learning on workforce

From being a novel initiative to an expected feature, Machine-learning has intruded into literally every industry. This transition towards automation often imposes severe consequences. According to Brynjolfsson and Mitchell (2017), the major implications are;

a) Technical Implications

- **Task Selection:**

Not all tasks can be automated! Only the tasks whose final results can be well-specified and evaluated automatically are appropriate for automation. Therefore, it requires care in selecting Machine-Learning tasks.

- **New framework:**

Machine-Learning tasks are different from routine tasks. Hence requires a framework more efficient than the conventional ones.

- **Well-defined Properties:**

It is crucial to analyze and define the Machine-Learning properties of the expected solution so that the system can automatically evaluate their results accurately.

b) Non-Technical Implications

- **Reduced demand for Labour**

In many cases, Machine-Learning directly substitutes for some tasks. Thus, replaces human and reduces labour demand eventually. However, over the years, the invention of any new product, service or tasks has always created new opportunities for labour.

- **Demand for skills**

ML tasks are highly creative and require a new set of skills. Demand for which will ultimately reflect on the wages.

- **A Continuous process**

Adoption of new technologies into an enterprise often take years as it requires alterations in business models, processes, organizational design, supply chains, and even cultural expectations.

VI. Python for Machine Learning

Implementing machine-learning models is as simple as a few lines of code in Python. Python has a set of dedicated libraries like Numpy, Scipy, Matplotlib, Pandas and, Scikit Learn to extend programming support to data analysis and machine-learning operations. (Santarcangelo J, 2017) However, **Scikit Learn** is the primary package that offers a collection of algorithms and tools specifically for machine-learning applications. These algorithms are compatible with other Python libraries like NumPy and SciPy. In addition to that, the library also offers support to the pipeline operations like data pre-processing, feature selection, feature extraction, splitting the data into training and test sets, defining machine-learning algorithms, fitting the models, tuning parameters, prediction, evaluation, and finally, exporting the model. For the reasons mentioned above, Scikit Learn stands as a popular one-stop solution for machine-learning.

VII. Conclusion

Machine learning is a power that is beyond imagination. In my perspective, we are only at the beginning stages of obtaining value from machine-learning. The full potential will only grow in the future. Therefore, investing in Machine-learning or AI technologies will genuinely benefit businesses.

References

- Samuel, A.L. "Some Studies in Machine Learning Using the Game of Checkers", IBM Journal of Research and Development, vol. 3, no. 3, pp. 210-229, July 1959, doi: 10.1147/rd.33.0210.
- Erik Brynjolfsson and Tom Mitchell (2017) "What can machine learning do? Workforce implications", DOI: 10.1126/science.aap8062
- Shai Shalev-Shwartz and Shai Ben-David (2014) , "Understanding Machine Learning: From Theory to Algorithms", Cambridge University Press, [Online] Available at <http://www.cs.huji.ac.il/~shais/UnderstandingMachineLearning>
- Santarcangelo, J. (2017), "Machine learning with Python", Available at <https://www.coursera.org/learn/machine-learning-with-python> [Accessed 08 May 2020]
- Ray, S. (2017),"Commonly used Machine Learning algorithms", Available at <https://www.analyticsvidhya.com/blog/2017/09/common-machine-learning-algorithms/> [Accessed 08 May 2020]
- Castle, N. (2017), "6 Common Machine Learning Applications for Business",Available at <https://blogs.oracle.com/datascience/6-common-machine-learning-applications-for-business> [Accessed 08 May 2020]
- Das, C. (2017), "What is machine learning and types of machine learning" towardsdatascience, Available at <https://towardsdatascience.com/what-is-machine-learning-and-types-of-machine-learning-andrews-machine-learning-part-1-9cd9755bc647> [Accessed 08 May 2019]