Assignment 1

Certificates of Advanced Study (CAS) at the School of Information Studies:

At the School of Information Studies, along with the traditional Master's degree in Information Management, one can learn various specializations made available through Certifications of Advanced Study. The two Certifications provided are:

1. CAS in Data Science

A 15-credit, graduate-level program, CAS in Data Science helps students learn the many technologies under the umbrella of data analytics, such as Big Data management using IT tools, data mining, etc., enabling students to grow to become full-time data scientists [1].

2. CAS in Information Security Management

This certification enables students to learn aspects of high degree of technical security, risk management, etc. It a graduate-level program worth 15 credits, and is recognized by the National Security Agency (NSA) and the Department of Homeland Security (DHS) [1].

The iSchool also provides two other Certifications of Advanced Study, which are:

1. CAS in E-Government Management and Leadership

2. CAS in School Media

The curriculum of CAS in Data Science program is aligned with my interests. The various courses that are available through the certification program, such as Data Administration Concepts, Data Science, Data Mining, Metadata, etc., and the concepts of data such as its lifecycle of architecting, acquiring, analyzing and archiving, big data tools, etc., that we will learn through these courses, seem promising of instilling the required skills to become a successful data scientist.

Data Science is one of the most upcoming fields in Information Technology. Over the past decade, there has been an outburst of information-driven organizations. From the smallest of the lot to the whale of the profit-making realm of businesses, the one similarity that drives them is the information they possess. Information about their clientele, about themselves and their competition, and most importantly, information about the current trends in the market, all of it combined makes a knowledge base that decides whether a company will succeed or not. It is fascinating to see how data can enable a company to grow as one of the key market players, and all this can only be made possible if the company has the required personnel of intelligent data scientists who know the various ways in which data can be used to create the required information.

Career Analysis for Data Scientists:

Data Scientists can hold different positions in different organizations. They may have different titles, but the underlying set of responsibilities of a Data Scientist depends on various specializations like data analytics, data storage and management, data visualization, general systems management, etc., that they may have mastered. Data Scientists are crucial to solving big data problems in areas as diverse as clinical research, defense intelligence, customer behavior, medical diagnosis, and risk management.

As the field grows, Data Science graduates are shaping the first wave of data science practices and standards [1].

Top-ranked data science jobs include the following titles:

Data Scientist,

Data Engineer,

Analytics Consultant,

Business Analyst,

Data Analyst, etc.

The field of data science has grown enormously in the past few years, and it is believed to still have a much larger potential to grow. According to an article from Forbes, the data science and analytics market has been predicted to surpass a value of $200 billion [2]. International Data Corporation (IDC) says that the revenue for Big Data and Business Analytics will grow up to $203 billion by the year 2020 at a compound annual growth rate of 11.7%. By the end of 2017, revenue growth from information-based products will double the rest of the product/service portfolio for one third of Fortune 500 companies [3].

Data Scientist remains the top job in the market, not just because of the alarming rate at which data and information are driving businesses, but also because there is a shortage of talent is the field that has an even bigger scope to be explored and utilized. Not only are individuals with skills in statistics and analytics highly sought-after, but those with the soft skills to match are driving demand for data scientists. Business leaders are after professionals who can not only understand the numbers but also communicate their findings effectively. Because there is a still such a shortage of talent who can combine these two skillsets, salaries for data scientists are projected to grow over 6% this year alone [4].

Moreover, data science is not just restricted to companies that have a technological background. Every company, from Google to Walmart to the smallest grocery store around at the end of the street, uses data from all the available sources. Smaller organizations have realized that just like tech giants, they too can make better, more informed decisions with the help of data. While small-to-medium sized organizations are not churning out nearly as much data as larger enterprises, sifting through that data to extract meaningful insights into their businesses can be a powerful competitive advantage nonetheless [4].

Salary Ranges for Data Scientists:

As per recent records, the estimated national average salary for Data Scientists in United States is $123,732, calculated by over 2000 records submitted on Glassdoor [5].

Some examples of salary ranges in companies that hire data scientists are as follows [5] –

Facebook, $100,000 to $180,000, with an average of $134,715

Microsoft, $94,000 to $152,000, with an average of $123,011

PwC, $93,000 to $125,000, with an average of $104,624

KPMG, $80,000 to $122,000, with an average of $97,716

Google, $103,000 to $210,000, with an average of $152,856

Apple, $120,000 to $177,000, with an average of $145,974

As per the Bureau of Labor Statistics under the United States Department of Labor, Computer and Information Research Scientists, a title given to Data Scientists, had the following statistics in the year 2016 [6] –

Median Pay – $111,840 ($53.77 per hour)

Typical Entry-level Education – Doctoral or Professional degree

Internships in the field of data science are available throughout the year. Some internships currently available, and the qualifications required for them are as follows –

1. The Boeing Company, Bellevue, WA, US

Data Sciences Intern

Required Qualifications: Working knowledge of programming languages such R, Python, Java, Apache, C++, etc.

Job Description: Intern candidates will work effectively with cross-functional teams and technical leads to determine, define, and deploy analytic solutions to meet business goals. Evaluate business objectives, determine stakeholder needs, and identify requirements. Choose best fit methods, define algorithms, validate, and deploy models to achieve business results. Perform necessary data preparation and enhancements to models [7].

2. Gap Inc., US

Data Scientist Internship

Required Qualifications: Statistical and data analytical skills, languages like Python, R, SAS, SQL, knowledge of data visualization tools like Tableau or D3, statistical modeling skills.

Job Description: Work alongside our Customer Data & Analytics team and fellow interns on customer segmentation and descriptive analysis projects; use the latest technologies, tools, and technologies in the field- harnessing vast pools of information available in our brands' data systems; develop data product tools, such as dashboards and mapping tools; apply analytical and data science methods (regression, clustering, classification, text analysis) to real world datasets to solve ongoing business problems; present results of analysis to team members, business partners, and Gap executives; meet Sr. Data Scientists, advanced analytics experts, and Product Managers to understand the role that analytics play in a 16 billion-dollar global organization [8].

Success and Failure in IM Leadership and Management:

Jennifer Shin, Senior Principal Data Scientist at Nielsen, Founder and Chief Data Scientist at 8 Path Solutions, Director of Data Science at Comcast, Lecturer at UC Berkeley, is one of the most well recognized names in the field of data science. She is a graduate of Columbia University with a Master's degree in Mathematical Statistics. In an interview with James Kobielus, IBM Data Science Evangelist, Jennifer mentions her research in electronic health records and her consulting projects with some of the Fortune 500 companies as a data scientist.

On the topic of open-source data management tools like Hadoop, Spark, etc., she says that Hadoop has become a popular choice for many information management workers, whereas Spark is still a nascent technology that is growing to be liked and known by the data scientists' community, but most common platform, she says, has been R since it has been around longer and most people have been using it for some time now. When it comes to using open-source tools for advanced data visualization, she says that an open-source platform is very useful as it allows the data scientist to customize the algorithm to find the correct solution without investing huge amounts of money in another tool specific to just one project. Depending on what kind of visualizations the business requires, an open-source tool can be used with the extraction-transformation-loading (ETL) process.

When asked about qualifications required to become a successful data scientist, and whether one would need an advanced mathematics degree to succeed in this field or can anyone learn the required skills to be effective in a business context, she says that mathematics is a very big part of data science and analytics, but that said, it depends on what kind of data you will be working on or what kind of tools you would be using. Her honest opinion is that to be a successful data scientist who can work with all kinds of data and tools, a quantitative background is extremely important, because that is the only way to be able to understand what tool or technology to use in place of something else and to figure out what method is better than the other.

As James Kobielus mentions that to be a good data scientist as Jennifer, one needs to spend years getting the kind of knowledge and training that she did, and that someone needs to be a specific type of person to be able stand up to the full potential of a data scientist. To this, she says that is not necessary to be a polymath. She gives an example of how one can be in Biology and not necessarily work only on singular datasets. She does, although, believe that one must have a diverse background and have cross-industry knowledge. She also points out that data science is a practical field where one can only learn by actually working on datasets and not just reading about them, and as one practices the various tools, it gets easier to work with them. To be able to understand the concepts of data science better, one must, first and foremost, be willing to work and experiment with data and learn more about it. Secondly, Jennifer says that passion is what should drive a data scientist. This passion is not of being called a data scientist professional, but of looking at data in ways that no one else can and digging deeper to find more valuable information.

In the interview, Jennifer also talks about the art of science, which according to her is not having a great idea and wanting to explore it, but the art of science is being able to have certain skills and recognize things because of that skillset. She compares it to engineering, because you can never be sure of what the solution is going to look like and there is no blueprint, but you need to come up with solutions somehow to be able to get there. When talking about data, there exists some creativity which

makes it look like an art, but there is also a lot of skills that you need, creating a fine line between creativity and skillset.

The interview made visible a paramount skill that data and information professionals, such as Jennifer Shin, must possess. The skill of being eager to know what people want and understand the current trends. With multiple options available in the market for data analysis, it is very important to know what technology is upcoming and brings about a certain stability to what people need for their business needs.

The interview was presented by IBM Analytics and can be watched using the following link:

https://youtu.be/BRQhZFYDCiw

The interview gives its viewers insights into the world of information professionals like Jennifer Shin, and most importantly, shows what is really required to become successful in the field of Information Management.

Before talking about the success factors of IM professionals, it is important to understand what may cause failure in this field. Some practices that cause failures in this field are as follows: -

1. Resistance to Change:

Information professionals must be willing to learn about new tools and technologies and not show reluctance to shift from traditional style of working to an enhanced methodology.

2. Communicating Less:

Managers and leaders in the field of information management can only be deemed as successful if the information system that they implement works efficiently, which can only happen if the right system is built. To make sure that the right system is built, one must be open to communication on what is required of the system and what do they end users need.

3. Inability to Organize:

It is the responsibility of the Information Manager to have data that is organized and ready to be analyzed. Failure to transform data into a usable format and organize it according to the business process eventually causes failure of the system, and in turn, its manager.

Some factors that define success in IM Leadership and Management are as follows:

1. Eagerness to Learn:

An information professional must not be a polymath, but rather, must always be willing to learn new concepts and technologies. In the field of data science and analytics, it is not possible to learn everything by reading. It is only the amount of practical exposure that you get and that you learn from that decide how good a job you will do as a manger in this field.

2. Quantitative Background and Analytical Skills:

To understand the many possible outcomes of data analytics, one needs to have a quantitative background and the ability to think analytically. Understanding the methods and tools being used, what the visualizations mean, what is happening while the data is being analyzed and what information can

be used from it to make well-informed decisions for the organization, are all the skills that are required for in information professional, and need a quantitative and analytical background.

3. Understand the Needs of the Users and the Market:

One can analyze data in multiple ways, but a good information manager needs to know how exactly to work with the data to provide the solution that is required now. Secondly, it is very important for the information manager to understand the current trends in the market. There may be a thousand tools available to perform a specific task, but to know what people prefer and need is what makes a good manager and leader in the field of information management.

Looking at the above factors of success, is it clear that they are for an information management professional, but not for the information system as a whole. Factors responsible for success of an information system are in the implementation of its four components, which are: **People, Process, Technology and Structure** [9]**.**

People in an information system is an umbrella that shelters multiple individuals and groups. Managers, engineers, end users, all stakeholders who in any way are being affected by the information system make up the 'People' component of it. Understanding of the people involved makes a very big difference to the information system. From what the end user wants to the kind of work environment the engineer wants to work in to the acceptance level of the workers who will be using the information system, all aspects are parts of a stack of dominoes. Fall of one domino is the fall of the entire set, which in this case is the information system. It is also very important to realize that the information system is for the people, to make their lives easier and make business better in the process, so it is of utmost importance to manage the needs of all stakeholders [9].

Process in an information system has the same definition as process in the English language: the way things work. It is the series of steps necessary to complete a business activity. For example, consider the warehousing process of a grocery store – when there is a shortage of an item, the store manager indicates it to the inventory manager, who then checks the availability of the item in the storage. If available, the item will be restocked at the store, and if not, the warehouse manager will be notified to deliver the item to the store. The warehouse, in-turn, gets the items from the manufacturer with whom an order is placed. Thus, for the customer to get hold of an item, the item must come through a chain of manufacturer, warehouse, storage, and finally the stand at the grocery store from where the item is picked up. Even the organizational hierarchy can be thought of as a process, where an entry-level employee must go through the chain of command to reach to the top bosses. Making sure that the working of the process is maintained is very important for the success of the information system and its manager [9].

Technology encapsulates the IT aspect of an information system. It may enable or restrict activities through its operational design. The technology that is used in the information system has its benefits, but may also have its own constraints. Design of technology used in this system is based on the choices made by the engineers and programmers who developed it, or opted to use it over another technology. As an information systems manager, it is a pivotal decision to make when it comes to choosing the right technology and its ways of implementation, because failure of technology may cause the complete abandonment of the information system [9].

Structure is short for the organizational structure of the company, that is, the hierarchy, chain of command, and reporting structure within the information system. Understanding the structure component is crucial because user resistance, incentive systems, and relationships are often silent enemies of information system success that go undetected before, and sometimes even after, information system failure becomes apparent. A major characteristic of the structure component is the support of high-level management. With a great design for implementation of an information system, with the right people on board, the right processes being followed and the right technology to implement it, the project will still fail if it is not supported by all levels of the organizational structure, especially top-level executives [9].

These four component: People, Process, Technology, and Structure, are the four pillars of an Information System. They all must be adhered to and coordinate towards the success of the system.

References:

[1] School of Information Studies, Syracuse University. Retrieved from URL: https://ischool.syr.edu/academics/graduate/cas-overview/.

[2] Press, Gil (Jan 2017). 6 Predictions For The $203 Billion Big Data Analytics Market. Retrieved from URL: https://www.forbes.com/sites/gilpress/2017/01/20/6-predictions-for-the-203-billion-big-data-analytics-market/#428a816e2083.

[3] Framingham, Mass (Oct 2015). Double-Digit Growth Forecast for the Worldwide Big Data and Business Analytics Market Through 2020 Led by Banking and Manufacturing Investments, According to IDC – International Data Corporation. Retrieved from URL: https://www.idc.com/getdoc.jsp?containerId=prUS41826116.

[4] Zhang, Vivian, and, Neimeth, Chris (April 2017). 3 Reasons Why Data Scientist Remains the Top Job in America. Retrieved from URL: https://www.infoworld.com/article/3190008/big-data/3-reasons-why-data-scientist-remains-the-top-job-in-america.html.

[5] Glass Door Job Search. Retrieved from URL: https://www.glassdoor.com/Salaries/us-data-scientist-salary-SRCH_IL.0,2_IN1_KO3,17.htm.

[6] Bureau of Labor Statistics. Retrieved from URL: https://www.bls.gov/ooh/computer-and-information-technology/computer-and-information-research-scientists.htm.

[7] Boeing. Retrieved from URL: https://jobs.boeing.com/job/-/-/185/5506126?utm_source=Indeed&utm_medium=organic&utm_campaign=Indeed.

[8] Gap Inc. Retrieved from URL: https://jobs.gapinc.com/job/-/-/1649/5669183?src=JB-10324.

[9] Piccoli, Gabriele, and, Pigni, Federico (2016). Information Systems for Managers – Edition 3. Chapter One.