Discuss the importance of a postgraduate degree in the Computer Science field

Initial Post

The field of data science is vast and growing exponentially since the emergence of big data, making data science currently the "most trendy" job as stated by Englmeier and Murtagh (2017). Within this vast field of data science there are multiple roles and responsibilities that can be undertaken. A study by Baškarada and Koronios (2017) identified 6 key roles that are required for the data science process including domain expert, data engineer, statistician, computer scientist, communicator, and team leader. This is a range of roles that a very limited number of individuals possess if any, hence such an individual being deemed to be a “unicorn data scientist”. Therefore, to get the best possible output a team of data scientists with complementary skills may be required. On the other hand many employers look for workers with a vast range of skills and typically want to hire the least amount of staff possible.

Lo (2019) identified three types of data scientist, each taking on a range of skills and responsibilities. The types of data scientist are citizen data scientist, business data scientist and data science innovator. As AI and machine learning improves there may be less need for data scientists that only have a skill set that involves running processes, since these processes could be completed more efficiently and cost effectively by an AI. This type of data scientist is what Lo (2019) identifies as a citizen data scientist. Instead the role of a data scientist could potentially be to innovate and identify new problems. This role within data science requires a significantly higher skill level and a wider knowledge than simply running processes, which Lo (2019) identifies as a data science innovator.

References

Baškarada, S. and Koronios, A., 2017. Unicorn data scientist: the rarest of breeds. Program, [online] 51(1), pp.65-74. Available at: <https://www.emerald.com/insight/content/doi/10.1108/PROG-07-2016-0053/full/html> [Accessed 23 September 2022].

Englmeier, K. and Murtagh, F., 2017. Editorial: What Can We Expect from Data Scientists. Journal of theoretical and applied electronic commerce research, [online] 12(1), p.i-v. Available at: <https://www.scielo.cl/scielo.php?script=sci_arttext&pid=S0718-18762017000100001&lng=en&nrm=iso&tlng=en> [Accessed 23 September 2022].

Lo, V. (2019) Searching for the Perfect Unicorn. Available from: https://pubsonline.informs.org/do/10.1287/LYTX.2019.04.02/full/ [Accessed 23 September 2022].

Peer response- Guiseppe Raneil

Thank you for the great submission.

In addition to a high percentage of data science projects not not delivering the business outcomes, organisations also struggle to gain value from data potentially due to either a lack of skill, infrastructure or strategies (Kayser et al., 2018).

As AI and machine learning improves there may be less need for data scientists that only have a skill set that involves running processes, since these processes could be completed more efficiently and cost effective by an AI. Therefore, the role of a data scientist could potentially be channelled down the route of innovation and identifying new problems that the current level of AI can not solve. This role which Lo (2019) identifies as a data science innovator requires a significantly higher skill level and a wider knowledge than simply running pre-existing processes. For data scientist in an innovation role there is a great importance on having the correct set of people with the necessary skills that can integrate data science skills with the business need (Kayser et al., 2018).

As data science evolves through innovation certain skills and methods may become less important or obsolete therefore, data scientists and employers will have to adapt over time to be effective.

Kayser, V., Nehrke, B. and Zubovic, D., 2018. Data Science as an Innovation Challenge: From Big Data to Value Proposition. *Technology Innovation Management Review*, 8(3), pp.16-25.

Lo, V. (2019) Searching for the Perfect Unicorn. Available from: **https://pubsonline.informs.org/do/10.1287/LYTX.2019.04.02/full/** [Accessed 30 September 2022].

Peer Response- Tasweem Beelunkhan

Regarding your comment on the skills that a data scientist requires, a study by Baškarada and Koronios (2017) identified that there are 6 key roles required for the data science process including domain expert, data engineer, statistician, computer scientist, communicator, and team leader. This is a range of roles that require a vast array of skills. These skills are possessed by a very limited number of individuals if any, hence the term “unicorn data scientist” being bestowed on such an individual. On one hand many employers look for workers with a vast range of skills since they typically will want to hire the least amount of staff as possible to minimise expenditure. However, to get the best possible output, a team of data scientists with complimentary skills may be required.

With the increase in the level of sophistication of AI and machine learning, processes that can be automated will eventually no longer require data scientists to perform them. Since these processes could be completed more efficiently and cost effective by an AI, the role of the data scientist will have to adapt. Rather than employers seeing AI and machine learning as replacements for data scientists the two should be used co-efficiently to improve the field of data science. Data scientist implementing new application, procedures or solutions will allow machine intelligence to solve complex and challenging problems (Elshawi et al, 2018).

Baškarada, S. and Koronios, A., 2017. Unicorn data scientist: the rarest of breeds. Program, [online] 51(1), pp.65-74. Available at: <https://www.emerald.com/insight/content/doi/10.1108/PROG-07-2016-0053/full/html> [Accessed 23 September 2022].

Elshawi, R., Sakr, S., Talia, D. and Trunfio, P., 2018. Big Data Systems Meet Machine Learning

Challenges: Towards Big Data Science as a Service. *Big Data Research*, 14, pp.1-11.

Peer response- Abiodun Maborukoje

In response to the functions and duties that you mentioned, a study by Baškarada and Koronios (2017) mentions a “unicorn data scientist”. This is someone who has the complete set of skills to fulfil the 6 key roles that Baškarada and Koronios (2017) identified as being required to complete the data science process including domain expert, data engineer, statistician, computer scientist, communicator, and team leader. This is a range of roles that a very limited number of individuals possess if any, hence such an individual being likened to finding a unicorn. This seems to support your claim that data scientists will inevitably form into new more specialized roles, as no one person could have all the skills to cover all the requirements that is currently expected of a data scientist.

The next question that could be asked is how closely these new roles will work together and whether smaller companies will be able to fund the hiring of this many specialized roles? For smaller companies with smaller budgets it could be more practical to have a data scientist with a less specialized skill-set. With the steady improvement of AI and machine learning algorithms there could be less need for data scientists to have such a wide set of skill, instead exploitation of AI and machine learning can be implemented to solve problems reducing the workload of data scientists (Elshawi et al, 2018).

Baškarada, S. and Koronios, A., 2017. Unicorn data scientist: the rarest of breeds. Program, [online] 51(1), pp.65-74. Available at: <https://www.emerald.com/insight/content/doi/10.1108/PROG-07-2016-0053/full/html> [Accessed 30 September 2022].

Elshawi, R., Sakr, S., Talia, D. and Trunfio, P., 2018. Big Data Systems Meet Machine Learning Challenges: Towards Big Data Science as a Service. *Big Data Research*, 14, pp.1-11.

Summary Post

As Victoria Stapleton rightly points out citing Dietrich, et al. (2015) the role of a data scientist should be to answer the project requirements. Therefore many employers’ demands are outside the scope of the primary focus of a data scientist, vastly increasing the requirements and skills that a data scientist must have to levels unachievable by the majority.

In response to Lo (2019), Ting Wai Jeff Chau cited Berthold (2019) in which data scientists are classified into three categories based on experience rather than skills and responsibilities. Although Berthold (2019) is using a different factor to categorize data scientists he seems to come to a similar conclusion. A novice data scientist seems to have the same skill level and responsibilities as Lo's citizen data scientist and an expert seems to align themselves to data innovators. As Ting correctly points out it is difficult for a data scientist to have all the necessary skills and experience at the start of their careers. Therefore it is important for novice data scientists to be exposed to more experiences as a data scientist to gain more experience and widen their skill set they could slowly gain more roles and responsibilities until they are able to create new insights and think outside of the box.

Responses from my peers have altered my view slightly. I now see AI as a future means of assisting to fill the increasing number of jobs rather than as a direct replacement for data scientists.

Master data management (MDM) aims to bridge the gaps between organizational units and information systems (Vilminko-Heikkinen and Pekkola, 2017) to improve efficiency, accuracy and decision making (Das and Mishra, 2011). MDM is an important process to keep a "unique view of the truth as stated by Martins et al. (2022). However, there are still many issues that relate to MDM including data governance that need to be addressed.

Baškarada, S. and Koronios, A., 2017. Unicorn data scientist: the rarest of breeds. Program, [online] 51(1), pp.65-74. Available at: <https://www.emerald.com/insight/content/doi/10.1108/PROG-07-2016-0053/full/html> [Accessed 30 September 2022].

Berthold, M.R. (2019). What Does It Take to be a Successful Data Scientist? *Harvard Data Science Review 1(2)* doi:10.1162/99608f92.e0eaabfc. [Accessed 3 October 2022].

Das, T. and Mishra, M., 2011. A Study on Challenges and Opportunities in Master Data Management. *International Journal of Database Management Systems*, [online] 3(2), pp.129-139.

Dietrich, D., Heller, B. and Yang, B. (2015). *Data Science & Big Data Analytics Discovering, Analyzing, Visualizing, and Presenting Data*. Indianapolis Wiley.

Lo, V. (2019) Searching for the Perfect Unicorn. Available from: **https://pubsonline.informs.org/do/10.1287/LYTX.2019.04.02/full/** [Accessed 30 September 2022].

Martins, J., Mamede, H. and Correia, J., 2022. Risk compliance and master data management in banking – A novel BCBS 239 compliance action-plan proposal. *Heliyon*, 8(6), p.e09627.

Vilminko-Heikkinen, R. and Pekkola, S., 2017. Master data management and its organizational implementation. *Journal of Enterprise Information Management*, [online] 30(3), pp.454-475. Available at: <https://www-emerald-com.uniessexlib.idm.oclc.org/insight/content/doi/10.1108/JEIM-07-2015-0070/full/html> [Accessed 7 October 2022].