INFO1111: Computing 1A Professionalism

2022 Semester 1

Team Project Report

Submission number: 1

Team Members:

Name	Student ID	Levels being attempted in this submission
Isaac Kim	510603294	Task 1, 2
Jack	??	??
Kyle	??	??

Contents

1. Level 1: Basic Skills

1.1. Developing industry skills

This section is completed as a team.
Throughout your Computing degree we will help you learn a range of new skills. Once you graduate however you will need to continue to learn new languages, new tools, new applications, etc. For this section you need to identify 5 approaches you can take to this continual learning. You should then put these in order from most effective to least effective, and then explain the circumstances in which each approach might be appropriate. (Target = ∼100 words per skill = ∼500 words total).

1.2. Skills: Isaac Kim : Data Science

1. Statistics

Statistics may look to first-year university students like myself to be a series of tedious calculations using fixed formulas. After all, that's what we've been doing in "Statistics" classes since high school. As a result, it's mind-boggling to grasp how statistical analysis and probability have such a profound impact on our daily lives. Statistics are used to forecast weather, restock store shelves, assess the state of the economy, and a variety of other things. Statistics on its own is a powerful tool for getting useful insights, but when combined with computer algorithms, machine learning, or AI, it can investigate a far larger pool of data and provide answers to business, science, and societal concerns that were previously deemed unsolvable. (Ram Dewani, 2021)

2. Programming Knowldege

Data science necessitates a mix of math and programming skills because it sits at the intersection of analytics and engineering. Data scientists will be unable to distinguish themselves from traditional statisticians if they are unable to use programming tools. Real data scientists will use sophisticated languages like Python or R to gain access to a wide range of useful functionality for dealing with Math, Statistics, and Scientific operations. Data scientists can utilise computer programmes to apply statistical expertise in a quick and secure manner, allowing them to uncover patterns in large amounts of disorganised data and make use of them.

3. Data Wrangling

Data wrangling is the process of organising and transforming new data into the appropriate format for later analysis. It's a crucial step that's well worth the investment because it helps analysts make better data-driven judgments in less time. Outlier treatment, missing value imputation, scaling, and correcting data types can help make a data collection more orderly and aid in the development of more accurate statistical results. Data wrangling, in my opinion, is just as significant as Oxford Dictionary's decision to structure its print dictionary alphabetically. It couldn't be worse if an English learner had to go through 1000 pages of paper simply to discover a definition of one word.

4. Data Visualisation

It might be difficult to grasp what a data collection has to offer as a story when looking at a series of statistics or numerical summaries. This is where Data Visualization, the more entertaining side of data science, may be able to assist with the solution. Data visualisation

enables data scientists to convey complex data sets to scholars and the general public in a more appealing and intelligible manner. Furthermore, because computer software can automate the process, creating complex histograms or pie charts takes only a few seconds and does not require the assistance of a skilled artist.

5. Machine Learning

Machine learning allows data scientists to easily uncover patterns in a variety of data sources and make continuous improvements to their models without the need for human involvement. Google and Facebook, for example, may develop highly targeted ad recommendation algorithms as more data from various people accumulates. Similarly, a skilled analyst could use data gathered by a Google search to find and predict popular software skills. On a larger scale, reliable prediction models, such as today's weather forecast system, can be constructed.

6. Deep learning

Deep learning is a step forwards in machine learning because it uses a revolutionary method to decision-making based on a learning model that analyses data with a logical structure in much the same way that a human would. Deep learning applications use a layered framework of algorithms known as an artificial neural network to do this analysis. Using this new, sophisticated technique, data scientists may construct software for self-driving cars that can distinguish between traffic signs, detect pedestrians, and identify automobiles. Deep learning can also be used to anticipate protein structures, create architectural plans, and design semiconductors. I am convinced that it has limitless potential for use in a variety of industries in the future.

7. Big Data Anaylsis

We produce 2.5 quintillions of data each day! The growth of the internet, social media networks, and the internet of things (IoT) has resulted in a rapid increase in the amount of data we generate. Volume, velocity, and veracity are all high in this data. If businesses do not properly analyse their acquired data in this data-rich environment, they will miss out on chances for wiser business decisions, more effective operations, bigger profitability, and happier consumers. This is why data analysts are so important in today's enterprises: they help organisations minimise expenses, make quicker, more rational choices, and catch client demand at the appropriate moment to generate niche goods.

8. Story-telling Skill

If it weren't for storytelling and visualisation, all of the data analysis and insights we create as data scientists would be pointless. Putting numbers and facts from your analysis on the table seldom gets you far. Imagine a weather forecaster walking in to tell people about an approaching blizzard. Their warning won't have any impact on the audience if they don't use appropriate visuals and storytelling techniques. Hence, forecasters use graphics and interactive methods to keep viewers hooked and informed.

Teamwork

I believe that the most significant element of data scientists is their ability to work in groups. Because data science is such a diverse area, it necessitates collaboration among

specialists from other professions. We can't be specialists in all of the subjects and abilities listed above as individuals. It is vital to concentrate on a few abilities in which we are interested and confident while also learning from our team to improve other skills.

1.3. Skills: add student 3 name here : Software Development

Your text goes here

1.4. Skills: add student 4 name here : Cyber Security

Your text goes here

2. Level 2: Basic Technology

Level 2 focuses on initial evaluation of the tech stack that is used by a selected company. All companies make use of a range of technologies, and these technologies need to work together. A tech stack is basically just this collection of technologies that collectively enable a company's systems. As an example, one of the most common technology stacks for supporting web servers is LAMP: Linux as the underlying operating system; Apache as a web server; MySQL as the supporting database; and Perl (or more recently PHP or Python) as the programming language.

Each student should choose a different tech stack and explain the role of each of the different technologies in that stack. Note that prior to researching your proposed tech stack and spending time writing about it, it might be a good idea to check with your tutor as to whether your chosen stack is suitable. (Target = ∼200-400 words per student).

2.1. Tech Stack: add student 1 name here

Your text goes here

2.2. Tech Stack: add student 2 name here

Your text goes here

2.3. Tech Stack: add student 3 name here

Your text goes here

2.4. Tech Stack: add student 4 name here

Your text goes here

3. Level 3: Advanced Skills

Level 3 focuses on more advanced technical skills (LaTeX and Git) and analysis of linkages and relationships between the items in the company tech stack.

The following is a list of advanced Git and LaTeX skills/features. Each student should select one pair of items from each list and demonstrate actual use of each item (either through activity in Git, or through including items in this report). (Target = ∼100 words per student for each feature).

- Git
  - Rebasing and Ignoring files
  - Forking and Special files
  - Resetting and Tags
  - Reverting and Automated merges
  - Hooks and Tags
- LaTeX
  - Cross-referencing and Custom commands
  - Footnotes/margin notes and creating new environments
  - Floating figures and editing style sheets
  - Graphics and advanced mathematical equations
  - Macros and hyperlinks

3.1. Advanced features: add student 1 name here

Explain your use of the advanced Git and LaTeX features.

3.2. Advanced features: add student 2 name here

Explain your use of the advanced Git and LaTeX features.

3.3. Advanced features: add student 3 name here

Explain your use of the advanced Git and LaTeX features.

3.4. Advanced features: add student 4 name here

Explain your use of the advanced Git and LaTeX features.

4. Level 4: Advanced Knowledge

Level 4 focuses on analysing your particular tech stack and considering alternatives. Each student should consider the tech stack they described for Level 2, and then discuss each of the following points:

- What are the strengths and limitations of this stack? (Target = ∼200 words).
- What alternatives exist, and under what situations might these alternatives be a better choice? (Target = ∼200 words).

4.1. Advanced Knowledge: add student 1 name here

Your text goes here

4.2. Advanced Knowledge: add student 2 name here

Your text goes here

4.3. Advanced Knowledge: add student 3 name here

Your text goes here

4.4. Advanced Knowledge: add student 4 name here

Your text goes here

References

Ram Dewani. (2021). *APA 14 must-have skills to become a data scientist (6th edition).* (See `https://www.analyticsvidhya.com/blog/2020/11/14-must-have-skills-to-become-a-data-scientist-with-resources/`)