Data Science Internship Journal - Ellyfe Ltd

The necessitude toolkit of work, I have adopted at this internship are as followed:

* Researching
* Trial and error
* Application to real dataset
* Researching: during my time at Ellyfe ltd, I have developed a habit of performing well-informed and thorough research, in order to improve my understanding of the data and industry, I am working with or in. Doing this allowed me to improve my domain knowledge of the company I work with, as a result of this, my proficiency in my technical skills were well supported and allowed me to make the right processes, analysis, visualisation to suite the company’s objective etc. The use of researching facilitated my ability to not only comprehend my own field of work but also gain knowledge in the processes that were used to extract the raw data, such as the health devices or machines that were employed to extract and record these data from end-user. As a result, I was able to recommend and contribute to changes or retesting in the data collection processes, when there were curiosities, inaccuracies or erros identified from the analysis or processing of these data.
* Trial and Error: Employing this trial-and-error method in my toolkit as a data scientist, allowed me to make iterative improvement, experiment, and discovery, allowing me to develop a more suitable and effective solution to objectives or goals that needed to be tackled. The use of trial-and-error method helped in further enhancing my knowledge of the data, I was or would be working with and possibly aim for the best possible outcome or result. In the field of working with Ellyfe ltd, the company was focused on measuring hydration levels in older age users, by means of trial-and-error, my work had involved making further exploration on the ideal parameters in order t detect the level of hydration in the older - age users. I also applied this toolkit in the process of visualising my data, by experimenting and choosing the most effective charts, graphs, or visual presentation to show my findings and insights. Guiding my manager, supervisor and stakeholder to have a better understanding of what they were looking at or working with. In the development stage of the predictive model, I had researched and made documentation about different machine learning algorithms that are in existent and also enhanced my understanding of the predictive models that were commonly used in the healthcare and fitness industries like deep learning model, artificial neural networks, convolutional neural networks, K – Nearest neighbour, support vector machine, etc. Through, trial-and-error method, I, along with my technical manager was able to assigned the predictive model that we believed was best fit for Ellyfe ltd.
* Application to real dataset: After performing the previous process, I was then well suited to begin applying my findings and knowledge gained, to the real datasets that I was working with. With the previous processes accomplished it made the work process on the real data, to be easy going and also reduced errors or gap in knowledge that may have affected my analysis or visualisation of the data being worked on. As I began to apply more of the previous processes, I began to get closer to the goals or objectives that needed to be tackled within Ellyfe ltd.

**Journal Entry:**

**Date:** 08/07/2024

**Summary of the Day:**

Today, I focused on the theme of hydration, what usually affects hydration and other relevant important information that would help improve my understanding of hydration. I also expanded my knowledge by learning about parameters that would be used to make an accurate prediction on the level of hydration, parameters like heart rate, skin temperature, ambient temperature, etc.

**Main activity:**

* Researching: making research on hydrations and factors that affected it or would be affected by different level of hydration. Ensured all findings and knowledge gained were from informative, reliable, and trustworthy sources, also documented and stored this research.

**Key learnings:**

* Improved my skill in industrial research to further enhance my understanding of the industry I am working in.
* Gained a deeper understanding of physiological data and its significance to hydration

Day 2: 09/07/2024

Summary of the day:

Today, I focused on researching about how hydration in older age adults differ from younger adults, and best parameters to detect hydration for these groups of adults, age 65 and older.

Main activity:

* Researching: advance research on hydration and its effect on older people, and how older age adults react differently to different level of hydration.

Key learnings:

* Further improvement of my skill in researching, being able to track down sources that would best fit the understanding I am looking for.
* Understanding on how older age adults react differently to hydration, e.g., older age adults receptors tend to become more dull or less responsive, so they are less aware if they are dehydrated or not.

Day 3:

Summary of the day:

Today, I research and enhanced my understanding of different medical devices that were used to measure parameters that had a relationship with hydration, such as heart rate, BIA, Spo2, skin temperature, etc. I also was able to use these sensors like electrocardiogram, ppg and see how they functions and also the result they output.

Main activity:

* Researching: further implementation of researching skills
* Real-life Demonstration and application of health devices.

Key learning:

* Further improvement of my skill in researching, being able to track down sources that would best fit the understanding I am looking for.
* Improved understanding of health devices and how they work and what data they produced.

Day 4:

Summary of the day:

Further completion of previous activity and development and using of python code to convert raw data from sensor readings to understandable data to be analysed.

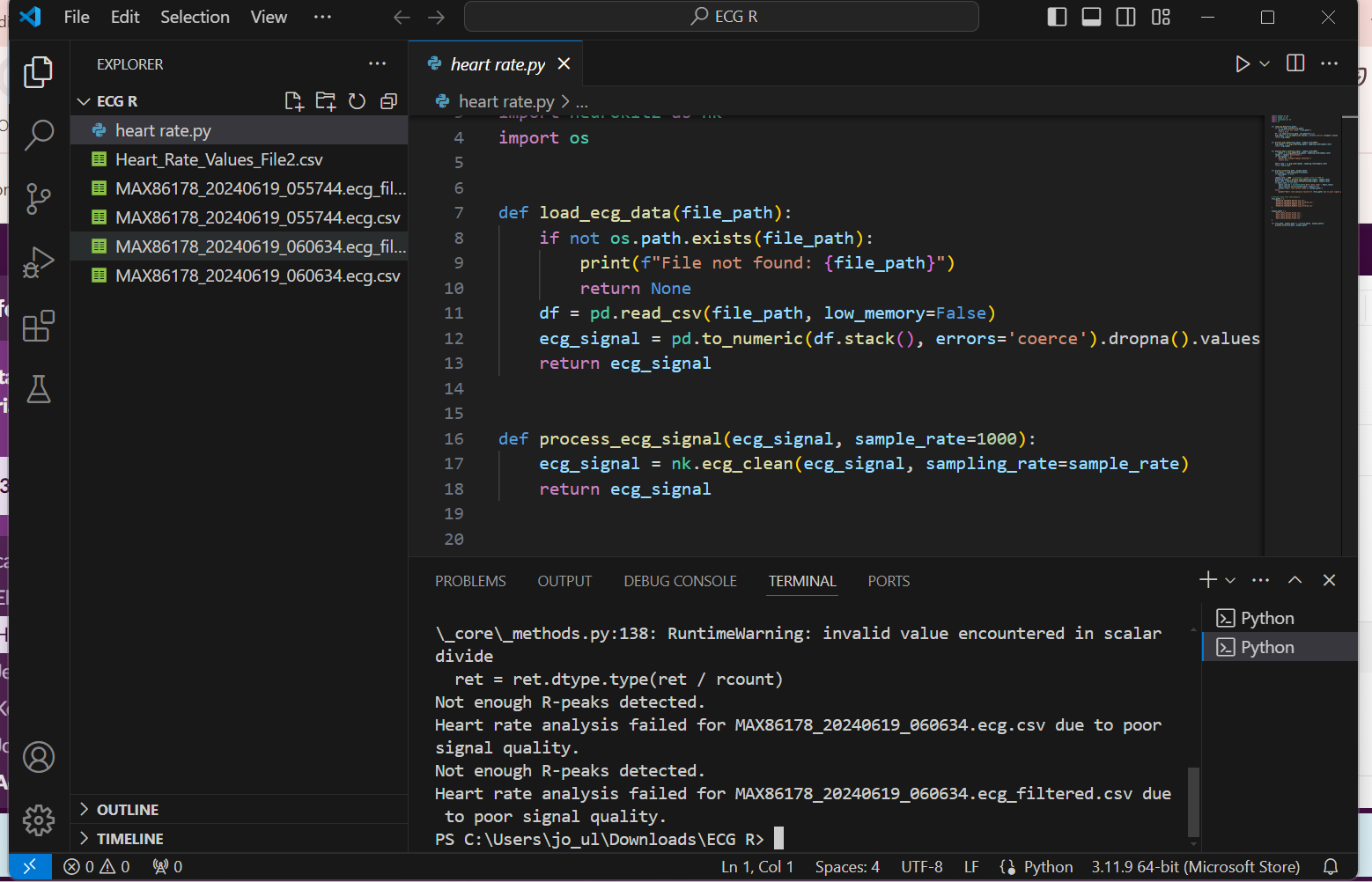
Main activity:

* Further completion of Previous report;
  + Uploaded information on different sensors used, e.g., PPG sensors, ECG sensors, etc.
  + Covering further information of components of sensors
* Utilised python code in order to convert CSV file containing raw readings from Electrocardiogram sensors to readable format that gave heart beats per minutes, enabling me to move on to analysis of the

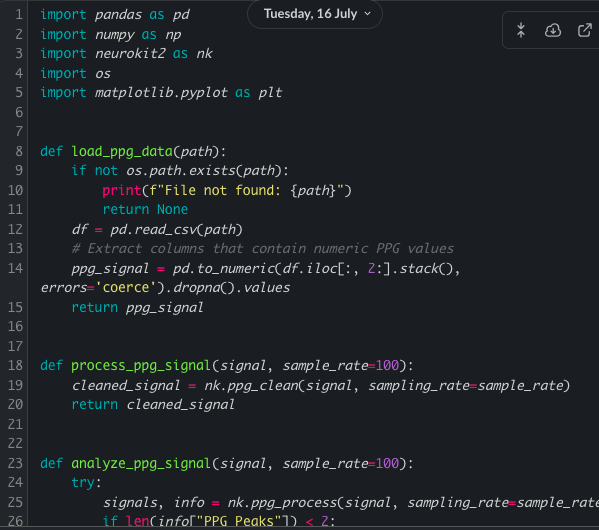
Key learning:

* Learnt how to develop a python code to convert ecg and ppg data from raw to processed.

ECG python code:



PPG python code:



Day 5:

Today, focused on data modelling of processed ECG and PPG dataset, to allow me to apply my understanding from researching done to the data and already form early insights to store for later use.

Main activity:

* Deployed data modelling on the ECG recorded data
* Deployed data modelling on the PPG recorded data

Key learning:

* Reverse engineering the custom python code use in processing of the data, helped in understanding the python code I was working with, its function and each part of the code.

Day 6:

Summary of the day:

Today, focused on learning about machine learning algorithms and the different predictive models used in the health and fitness technology industry.

Main activity:

* Researching on predictive models and their functionalities in the healthcare
* Learning predictive models and how they are used in the industry and how they are best used or what do they achieve or help predict

Learning keys:

* What machine learning algorithm is and how they make prediction
* Types of machine learning algorithms; k-Nearest Neighbour measure, CNNs, support vector machine, time series, supervised and unsupervised learning, etc.

Day 7:

Summary of the day:

I analysed competitor technology and algorithms for hydration prediction using PPG and ECG datasets, noting techniques like convolutional neural networks. I cleaned and processed these datasets, comparing results from different files and visualizing them in charts. I also evaluated the consistency and accuracy of the data, and researched the reliability of PPG and ECG for heart rate feedback.

Main activity:

* Competitor analysis
* Data cleansing
* Data visualisation
* Data processing
* data evaluation

Learning keys:

* Understanding how similar technologies and algorithms are used by competitors.
* Assessing the algorithms used in competitors' technologies, such as deep learning models, can guide the selection and refinement of your own predictive algorithms.
* How to evaluate the reliability of processed data, to ensure it is accurate and reliable.

Compilation of Excel Analysis of three dataset that gives final details of both PPG heart rate and ECG heart beats per minutes measures, with charts:

