Sprint 1 Retrospective/Review

Sprint Information

Start Date: 1/10/23

End Date: 25/20/23

Sprint Goals/objectives

From the beginning of this sprint, we had a clear vision in mind:

- Development of the frontend for the HealthAI – Web
- Development of the frontend for the HealthAI – App
- Breakthroughs in the Machine Learning Models and the API integration
- Strong clear communication and documentation
- Clear concise directions for all tasks during meetings

Project Supervisor – Gerard Mac Sweeney

Team Members:

- Luke Kenny R00212866 – Group Leader
- Alex Young R00219993
- Pierce Purcell R00221276
- Brian Sheridan R00238192

Agenda:

- During this review, there will be a number of topics discussed. From the entire lifespan of this first sprint. Highlighting key moments, roadblocks, what went well, what went wrong,

challenged and issues, improvement suggestions, lessons learned, summary and conclusion and more.

- We will each give our own personal insight into the project development lifetime and our own opinions on topics.

- We will also be providing information regarding the processes we undertook to ensure delivery of project requirements in a satisfactory timeframe.

- The Agile methodology we incorporated to ensure clear, concise communication between team members and maintaining a strong line of communication with out Project supervisor.

- The review will delve into project requirements which we have been working on from the sprint backlog, such as: Machine Learning Models and its Dataset practices, Flask API development. Web, and App development, as well as Requirements modelling and Sprint Project Management.

Alex Young – Team Member assigned to Machine Learning Models and Flask API

Sprint 1:

NB: Gathered Breast Cancer dataset, disregarded colon dataset.

Heart Disease Dataset:
- One-hot encoded "cp", "slope" and "ca" features - improved accuracy

Challenges:
- Imbalanced dataset: 526 instances of patients with identified heart disease and only 499 instances of patients without identified heart disease

Solution:
- Implemented SMOTE technique to synthetically generate more negative samples to match number of positive cases

- Dropping first category value with 'drop_first=True' lowered accuracy

Solution:
- Changing 'drop_first=False' to retain all columns

Challenges:
- One-hot encoding "restecg" and "thal" features lowered accuracy, possibly due to increased dimensionality (may lead to too many columns with too many unique values, or may not have enough unique values)

Solution:
- Removed one-hot encoding for "restecg" and "thal" features

Breast Cancer Dataset:
- Standardized all features - improved accuracy

Challenges:
- Imbalanced dataset: 357 instances of patients with identified breast cancer and only 212 instances of patients without identified breast cancer

Solution:

- Used SMOTE technique to synthetically generate more negative samples to match number of positive cases

Lung Cancer Dataset:

- Removed trailing spaces from the header names 'FATIGUE ' and 'ALLERGY '.
- Dropped ANXIETY and PEER_PRESSURE fields as I did not think they were directly correlated with lung cancer

Challenges:
- Imbalanced dataset: 270 instances of patients with identified lung cancer and only 39 instances of patients without identified lung cancer

Solution:
- Implemented SMOTE technique to synthetically generate more negative samples to match number of positive cases

Challenges:
- Attempted standardization but kept same accuracy levels
Solution:
- Removed standardization as did not harbour any value

Threshold: shifted to 0.4 threshold to favour higher recall. Recommended for healthcare scenarios to ensure that we can detect as many life-threatening diseases as possible, even if it means accepting some more false-positives.

MLAnalyser Class:
- Implemented train_logistic_regression(filename, target) method -> filename of csv, target feature of csv

- Implemented perform_analysis(self, type_analysis, feature_list) method -> type of analysis (heart/lung/breast), feature list passed from API

Flask API:
- Created @analysis route

Challenges:
- When receiving an analysis request for heart disease, the one-hot encoding in the originally trained model caused a 'missing columns' problem for the API when the user sent in the original columns

Solution:
- Built a one-hot encoder method which would fill in the missing columns if they were no longer present and set them to '0'

Personal Feedback:
Definitely a lot of knowledge gained from this sprint. My understanding of Logistic Regression and preprocessing techniques definitely improved and I can definitely say that I am very interested in the field of Machine Learning. A few different challenges were faced but managed to overcome them with a lot of testing as well as trial-and-error.

I also delved into Flask in order to learn how to build my own API. This was not too hard in comparison to Machine Learning. I found this whole process very fruitful for me and can definitely see the use of this in real-world applications.

Ideas and suggestions for making future sprints more efficient, effective, or smoother.

Feedback: While we did communicate on a weekly basis, more face-to-face interaction between team-members would be more beneficial.

END.

Pierce Purcell – Team Member assigned to frontend of HealthAI – Web

1. What Went Well:

Overall functionality and styling of website went very well. All fields, buttons, navigations, and necessary pages created to meet the requirements of the project. Links between different pages set up and ready for implementation of other aspects of the project. Design has a very user-friendly interface so that making it easier for users to interact and navigate the website.

A section where team members highlight and discuss the positive aspects of the sprint, such as achievements, successful tasks, and milestones.

2. Challenges and Issues:

• Challenges scaling project to fill web pages to give the website more modern look, researching html and css benefited to solve this issue.

• Managing css when project grew larger to correct any conflicting styles within the files.

• Uniqueness in the code so that it is easily read to keep track of progress.

Identification of any problems, roadblocks, or issues encountered during the sprint:

Action Items:

• Getting suitable images, logos, and other content to fit the overall design of the website.

• Adapting the websites layout and design for different screen sizes.

Specific actions that need to be taken to address challenges or to capitalize on what went well.

Improvement Suggestions:

• More JavaScript for more interactions throughout the website.

Ideas and suggestions for making future sprints more efficient, effective, or smoother.

Feedback:

• More in person meetings.

Luke Kenny – Project Leader assigned to Sprint Backlog, Requirements Setup and Documentation for submissions.

What went well:

- Efficiency of task completion
- Meeting schedule attendance
- Communication of team off-campus
- Strong delivery of deadlines

- Modelling of use cases, wireframes, test cases, backlog completion, and user stories.

- Was able to successfully draw up requirements models for the backlog of this sprint, demonstrating clear, concise instructions for users or stakeholders to be able to follow our progress and provide clear communication to our project supervisor on how our development timeline was going.

- Strong attendance of meetings with our Project Supervisor from all members whilst logging them in our group board.

- A milestone for me was when all features we set out to do were near completion which assisted me in drawing up the models for our submissions.

- This helped me better refine my models to ensure there is a clear guideline to our project deliverables.

Challenges:

- Creating use case diagrams, wireframes, user stories, documentation etc for rather complex backlog tasks which proved difficult and rather time consuming which was more than I expected at first.

- Setbacks included, adding more complexity to features which proved difficult to model, whilst also maintain a firm line of communication with other member to shape the requirements around their work.

Solution:

- Begun researching Agile Methodology and began deploying its best practices to the tasks at hand. Ensured all the modelling and diagrams were clear to users and easy to follow along. Whilst also maintaining the complexity of the requirements for our features.
- Maintained a line of communication with the individual team members and ensured that all members had a strong understanding of their tasks and maintained a positive environment between myself and the other members.

- Trusting in our team was crucial to overcoming any obstacles, as once the communication falls. It becomes incredibly difficult to maintain project deadlines and a clear-concise understanding of project requirements.

Improvement Suggestions:

- From having conversations with our team members, I have suggested we include more face-to-face meetings on campus. So that we may model up our development lifecycle for each sprint. To ensure that all members have a deep understanding of project roles and responsibilities. And so that we can improve our overall communication amongst ourselves. To ensure the team stays motivated and driven.
- A better balance between members for the tasks we have completed might be a better solution to workload. As I am planning to join team members to tackle large complex tasks.

Lessons Learned:

- Time management and communication are key to a successful project development team, without these the team will be spending more time fixing issues than delivering ideal timelines for project deliverables.
- I have learned a lot about organisational and team management skills. I hope to ensure the next sprint goes without any roadblocks and we deliver a desirable product.
- Task management: I wish to take on a more programming-oriented role, taking a keen interest in the LLM development for the next sprint.

Brian Sheridan – Team Member assigned to frontend of HealthAI- App

1. **What went well?**
    a. For my part of the sprint, I was working a lot with Android Studio. What went well was being able to layout and design the app pages using the Android studio tools. Android studio allows developers to drag and drop a lot of elements which can then be customized in the XML files. This makes it faster and easier to see the design of the pages in real time and be able to make changes quickly once you understand the XML files and how to edit them.

2. **Challenges and Issues**
    a. Some of the challenges I faced during this sprint was trying to connect the pages together and link everything together. Android Studio is new to me and trying to get everything linked together was quiet difficult at the beginning. Also trying to get all the views for each page to

be customized the way we want them and have this consistent across all pages took a good bit of time.

3. **Identification of any problems, roadblocks, or issues encountered during the sprint.**
   Action Items:
   a. One issue I ran into while creating the application frontend was that during the creation of the application the emulator that android studio allows developers to use to see their work on a virtual device stopped working. I couldn't figure out what the issue was for a while and it was only resolved once I changed the virtual device that I was working on. I am unsure as to what caused this issue but it is working on multiple other devices.

4. **Specific actions that need to be taken to address challenges or to capitalize on what went well.**
   Improvement Suggestions:

   a. In order to address challenges more efficiently and to get insight on the issue is to work as a team and to share all problems we encounter to get help from multiple sources to try and correct the problem and get the project on track. Sharing what went well with each of our sections of the project also allows others to see what we are aiming for and ensure that we all have the same ideas for how we want the project to develop.

5. **Ideas and suggestions for making future sprints more efficient, effective, or smoother. Feedback**:
    a. To further improve our sprints and make them more efficient and effective we have decided that we should try to meet more face to face as a group. Meeting for a few hours a week and sitting down together allows for more effective communication and will help us all with whatever problems we've encountered. It also allows us to be able to keep consistency across all of the sections of the project.

6. **Any feedback or comments from team members or stakeholders. Lessons Learned**:
    a. Android Studio is very useful tool for creating applications and as I become more familiar with it and learn more about what it can do and what it allows us to use in the application it will make developing the app a lot easier and should enable us to create a successful application.

7. **Insightful lessons or takeaways from the sprint that can inform future work**.
    a. Communication and constantly working on the project is vital to ensure we are all at the same place and are completing work as set out before each sprint. Communicating with each team member reduces to likelihood of running into any major road blocks.

Summary:

- Overall, I would say that the team members have learned a lot over this sprint. And it has had it's number of learning curves and problem solving. But the work we have completed has been satisfactory to what we set out to do, and the groundwork for the next sprint has been firmly established.

- We will incorporate all our members feedback and suggestions to ensure that Sprint 2 is delivered in a much more efficient fashion.

- That being said, the sprint was a great experience to learn more about Agile Methodology and how to deliver high quality software as a collaborative team.

Sprint 1 Contributors. 25/10/23

Alex Young

Brian Sheridan

Luke Kenny

Pierce Purcell