
EE/CprE/SE 491 BI-WEEKLY REPORT 04

03/18/2024 – 03/30/2024

Group number: 10

Project title: Accurate Cancer Prediction Using Artificial Intelligence

Client & Advisor: Professor Gaffar

Team Members:

Bishal Ghataney - Senior Engineer

Norfinn Norius - Client Communications

Thriambak Giriprakash - Minutes and Administration

Mark Hanson - Developer

Eric Schmitt - Developer

Chris Tague - Developer

- **Weekly Summary**

The week's main goal was improving the implementation and deployment of the application. We have set up individual pieces to make the application more realistic and real world ready. With the peer feedback, we are also looking into making our application more user friendly by altering how we display the prediction our AI model gives us.

- **Past week accomplishments**
  - Bishal:
    - Tried to tune the accuracy of the model using gradient boosting but kept having issues with the data that we have. I think the data is fine but the way I was incorporating it is maybe creating issues. I need to spend more time understanding the data.
    - Model calibration: tried if the current model accuracy is reflecting the prediction which I think is correct. I will be continuing on this model calibration more in the coming days by writing tests.

- Norfinn:
    - Helped Mark with implementation and debugging on the login
    - In preparation for expanding the Svelte UI next week I went through some tutorials and documentation as I do not have previous Svelte experience other than the very basic app setup we have so far.
    - We have discussed hosting another AI model in a different page on our website trained on open source data. In relation to this:
        - I researched open source data sets available related to cancer and the logistics of using them.
        - I came up with a number of options available to use that we will discuss this week. In particular there are several open source datasets available through the National Cancer Institute (NCI-DOE Collaboration AI/ML Resources) that look promising.
- Thriambak:
    - Researched cloud based databases which are quick and easy to deploy
    - Created a database on firebase for login functionality.
- Mark:
    - Worked on a login page for the web page.
    - Debugged the installation of a generic svelte login page with the current website.
- Eric:
    - Researched techniques for minimizing the error rate of our neural network.
- Chris:
    - Cleaned up the deployment of the application.
    - Began working on unsupervised learning as a precursor to finding better features than the whole dataset
    - Experimented with model more to improve accuracy

- **Pending issues**
    - The use of the login page somehow messes with the functionality of the AI model prediction. Mark has spent hours trying to resolve this to no avail.

- **Individual contributions**

Name	**Hours this week**	**Cumulative Hours**
Bishal Ghataney:	8	40
Norfinn Norius:	6	52
Thriambak Giriprakash:	6	36
Mark Hanson:	6	48
Eric Schmitt:	6	48
Chris Tague:	8	54

- **Plans for the upcoming week**
  - Bishal:
    - Try to help with the additional needed UX (user interface) web pages
    - Fix my current issue with gradient boosting and explore other ways to improve accuracy.
    - Understand the dataset more and write tests to verify the predicted accuracy is reflecting the dataset we have.
  - Norfinn:
    - Add multiple web pages to the Svelte application (it is currently just one).
    - Add an information webpage on the model and how to use it for the user to read
    - work on the UI of initial login page and existing page to make it more aesthetic and user friendly (it just has a button to load data at the moment and no formatting)
    - Continue to research possible alternative models that could be hosted in a different page and come to a consensus with teammates on implementation plans for that.
    - Help with any remaining debugging needed on the login database as needed
  - Thriambak:
    - Implement necessary changes to front end login screen to correctly store and retrieve data from the firebase db.
    - Look into other options if Firebase seems to not be the best solution.
  - Mark:
    - Get the login working with the current svelte setup.
    - Get the create account page made and working with the firebase database.

    - Get login working with the firebase database.
- Eric:
    - Work with Norfinn on finding possible alternative models for open source data we plan to include on a new tab
    - Help Norfinn with any UI debugging or implementation as needed
    - Help Mark with getting the login working
- Chris:
    - Improve model accuracy further by incorporating a combination of unsupervised learning
    - Aid in getting login page working and integrating UI with database

- **Summary of weekly advisor meeting**

Neither our TA nor our advisor attended our biweekly meeting.

Midterm Feedback

1. Summarize the feedback you received (both written and verbal).

They proposed enhancing the output of your model to include not only the raw prediction but also additional metrics such as confidence or standard error, while reminding users that predictions may not always be perfectly accurate. They also emphasized the importance of thorough documentation, particularly in relation to the improvements made to the model. The idea of refining the model's accuracy through the use of fewer data points or unsupervised learning was also put forward. In terms of resources, they advised ensuring the quality of any open source materials used. Finally, they recommended exploring various data augmentation techniques, such as cropping, resizing, flipping, and subsampling, but cautioned against overfitting and information loss.

2. Describe any new insights your team generated based on this feedback.

With the feedback we received, we definitely need to adjust our website to be more user friendly. It is a good idea to be more specific in our documentation, both for the advisor and client, but also for our group. It would also be beneficial when presenting our project to the panel at the end of the semester. Modifying the data through augmentation techniques was something we had not considered doing and will certainly look into it to improve our data set size and the accuracy of our AI model.

3. What steps are you taking based on the feedback?

The steps we will be taking based on the feedback from our peer review is to include the confidence percentage of our prediction and try refining the model's accuracy by selecting a few

data points instead of the whole set. We will also modify our UI to be more user friendly and we will be more specific with our documentation.