
EE/CprE/SE 491 BI-WEEKLY REPORT 03

02/11/2024 – 02/24/2024

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

Bi-Weekly Summary

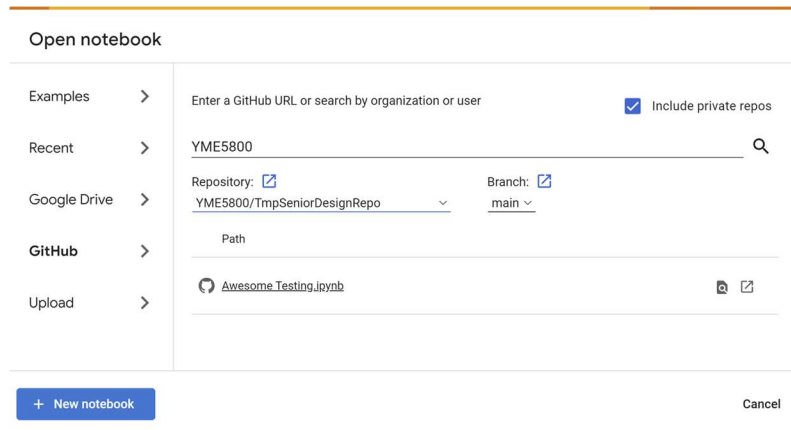
The UI has been modified to be more visually appealing. In addition to this, the UI is now capable of interacting with our model. The model upload process is not automatic at the moment which is one improvement for the future. There have also been improvements to the model; the data has been normalized, which dropped the MAE by about 10%, from 50 months to 45 months.

Past week accomplishments

- Bishal Ghataney:
 - Researched and explored advanced machine learning techniques to improve the accuracy of the model that we currently have by watching videos and searches on how others have tuned their model to improve accuracy

- Explored techniques such as ensemble methods, gradient boosting or neural architecture search to enhance accuracy
 - Now that I have learned these methods this week, I will try to investigate more and try to implement the Ensemble Learning Method to combine the predictions of multiple models for improved accuracy. This includes techniques such as bagging, boosting or stacking.
- I tried incorporating a transfer learning technique this week by using the pre-trained models developed for related tasks but I couldn't succeed but I will try this technique again next week.
- I also worked on investigating the model calibration this week a little bit for our model to make sure that the output accuracy properly reflects the predictions and I will be continuing on this next week.
- Norfinn Norius:
 - Researched how to use TensorflowJS to load our model into the browser
 - Set up a Python Linux environment, downloaded TensorflowJS and converted our trained Keras model to TF layers.
 - Implemented a basic Javascript/HTML webpage that loads a local csv file of data and converts it to a tensor that inputs into the model which then returns a prediction.
 - Implemented a basic UI for the website that includes a button to run the AI and shows the outputted result on the page.
- Thriambak Giriprakash:
 - Assisted Norfinn with integrating the trained model into the website
 - Assisted Norfinn with debugging the website
 - Looked into API changes which need to be made to fit the new front end webpage
- Mark Hanson:
 - Assisted Norfinn with integrating the trained model into the website
 - Assisted Norfinn with debugging the website
 - Researched cancer and different cancer cells
 - A cancerous cell can be any cell of the human body, meaning different cancer cells can look different

- Eric Schmitt:
 - Researched Further into Colab-Gitlab versioning. Findings below:
 - You can clone a repo in the colab then from there you will be able to use git commands in the notebook in order to push/pull the updates through the colab notebook.
 - Unfortunately, in order to clone the repo you would have to save your gitlab password/tokens in colab which we won't do for security purposes.
 - I looked for many workarounds saving passwords/tokens in colab and couldn't find any that would work while using gitlab. There used to be a way to launch a GUI in colab to enter account info instead of having to save passwords but that is no longer supported as of 2021.
 - Possible solution for Colab versioning: If we use github instead of gitlab you can easily and securely connect your github account and open the colab notebook through your github. I tried this with my personal github account for proof of concept and it works smoothly as seen below in the screenshot.



- Chris Tague:
 - Normalized data
 - Collected data into compressed CSV to eliminate some steps of the model creation process
 - Tweaked model to reduce MAE by 10%

Pending issues (If applicable: Were there any unexpected complications? Please elaborate.)

- Improve model metrics more

- Deploy UI
- Sort out remaining issues with frontend and model communication
- Solve Gitlab/Colab version control issue
- Create sign in and account system for front end for user security + HIPAA compliance

Individual contributions

<u>Name</u>	<u>Hours this week</u>	<u>Cumulative Hours</u>
Bishal Ghataney:	7	24
Norfinn Norius:	8	40
Thriambak Giriprakash:	6	36
Mark Hanson:	6	36
Eric Schmitt:	6	36
Chris Tague:	7	38

Comments and extended discussion (Optional)

Some of our members are beginning to think that there may be little correlation with the given pathology data and the cancer recurrence rate

Plans for the upcoming week

Improve UI, model accuracy, and deployments.

Add a backend on the website to upload a non-local file and implement a secure login page.

Summary of weekly advisor meeting (If applicable/optional)

Advisor did not show up. TA's told us to continue the work we are doing; other teams who have duplicates of this project seem to be having worse MAE than we are thus far.

Broader Context

Individual reviews:

- Chris
 - The vision of the project has changed, so many of the impacted communities have changed. With these new insights, I would make lots of changes to our broader context.
 - The positive effects of this is that it provides a means of prediction for cancer recurrence, which doesn't exist in many forms, and not at all in this form, today.
 - A potential negative effect could be a broad range of time for when the recurrence will happen; however having some knowledge leaves you at worst no worse than prior.
- Bishal:
 - The main intent of this project has not changed which is to serve as a support tool for the medical professional on cancer identification and recurrence. Clients also will have access to this tool but they should highly consult the medical professional.
 - One thing that clients must know is that we will have addressed any concerns related to clients or patients privacy, informed consent, and potential biases in the model's prediction to ensure the ethical and responsible use of the technology before the model is in service.
 - The heavy investigation of model calibration must take before this model ever gets used just to be extra sure that the output of the model is properly reflecting the prediction of the model
 - It is very necessary to develop a plan for continuous improvement and refinement of the model that'd be created based on ongoing feedback, data updates and technological advancement.
- Mark
 - The people who this project is meant to help has not changed. Cancer patients and medical professionals will still rely upon this tool for diagnoses.
 - With the model deployed, cancer diagnoses will become more accurate and with less time to verify diagnoses. With cancer, the sooner it is identified, the more options medical professionals have and the higher the survival chance of the patient.

- Saving lives and giving care to those who need it are some of the most important moral obligations to society. So extra power or training is minimal in comparison. In addition, users of the tool should be informed of drawbacks and proper usage of the website to minimize any errors. Doctors should also be informed that this is a tool, not a miracle cancer predictor. They should still make their own assessments of the patient's conditions.
- Norfinn
 - Everyday more people and their loved ones are affected by cancer. These people are becoming more aware of the large impact cancer has. Due to this the project's scope has grown but the effects have not changed from our broader context.
 - Once our program is deployed into the world it will be able to detect cancer occurrence and recurrence in patients. This would positively affect the patients by confirming medical professionals' diagnosis, giving the patients more confidence in the diagnosis.
 - A potential negative effect would be false positives or negatives. The goal of this project is to make the AI as accurate as we can in order to minimize incorrect outputs as much as possible. On top of that this AI should be used as a tool so results should already be verified by a medical professional.
- Eric
 - We have not identified any new effects that are different from our broader context which was to provide patients and medical professionals with reliable and accurate information.
 - We can provide evidence of positive effects by showing how our program can increase the confidence and satisfaction of patients and medical professionals by providing them with clear and trustworthy results. We could support our arguments with statistics, testimonials and case studies once we deploy the program.
 - We can justify negative effects by acknowledging the limitations and challenges of our program and by explaining how we are minimizing them. We can admit that our program is not perfect and there is a possibility of false positives or negatives, but we can also emphasize that our program will be constantly

improving from feedback and data. We can also stress that our program is not intended to replace human judgment or expertise, but to complement and assist them.

- Thriambak
 - The effects that we had at the start of the year are the same now, that is, to provide accurate predictions on the likelihood of cancer recurrence based on an input of medical data of a given patient.
 - Our output will be based on trends noted in existing cancer patients' recurrence rate. Showing accurate results based on this data which is fixed in reality will allow us to provide evidence for positive effects.
 - We can justify negative effects by referring to the general limitations of predictive AI, the quality of the data being input; and the data that we built our algorithm on. But the strength in AI lies in its ability to learn with proper guidance. As more people use the algorithm, correcting its mistakes, and affirming its correctness, the algorithm will improve in its ability to function properly.

Updates:

Communities affected

- Patients who are in remission from cancer
- Families of patients, healthcare policy makers, and the general public.

Societal Needs Addressed

– The project addresses the need for cancer recurrence predictions, and will help those struggling with cancer and those close to them, as well as medical professionals planning for the future

Public Health, Safety, and Welfare

- Positive impacts include, early and accurate cancer prediction could lead to improved patient survival rates, and may reduce the number of unnecessary procedures.
- Negative impacts include, potential false negatives or false positives. This could lead to missed treatments or unnecessary stress and medical interventions.

– To address this negative effect, doctors and other medical professionals should continue with their own diagnosis independent from this tool, and use this program as a “second opinion” and not take it as fact.

Global, Cultural, and Social

– Positive impacts: Universal access to such a tool could lead to standardized cancer care across different regions, which would benefit areas with less access to expert diagnostics. The system might also overcome bias in human-based diagnoses.

– Challenges: Cultural skepticism towards AI may make people hesitant to trust the diagnoses. Additionally, if we are too reliant on the technology it could lead to medical professionals to lose a bit of their skill.

– To address these challenges, medical professionals should only use this program as a tool to help confirm their diagnosis. If the program is used as a tool the patients will still be able to trust the doctor's opinion.

Environmental

– Positive impact: If the system reduces the need for other means of testing due to improved accuracy, it could lead to decreased usage of medical resources and chemicals related to cancer screenings.

– Negative impact: The energy consumption of training and deploying the neural network could be significant.

– Justification: If this model does take the place of other medical tests, then the power that originally went to those other tests could be used in the deployment and usage of this model. While even more power may be needed, it is still preferable to give patients the best treatment and care the medical industry can give.

Economic

– Positive impact: Faster and more accurate diagnosis could reduce the overall treatment cost by catching cancers early, leading to less aggressive treatments and shorter hospital stays. For healthcare providers the efficiency of an automated system could lead to reduced labor costs.

– Challenges: There might be potential job displacements if AI takes over roles traditionally held by medical professionals.

– This challenge isn't very concerning because this tool will most likely be used as an aid rather than a replacement so there is very little risk to it taking medical professional's jobs.