**Meeting Time:** 9:00 am – 10:38 am 2022.12.12

**Attendees:** Xia Jiang, Zhen Yang

**Meeting agenda**

1. Tested iMedbot: tested both prediction and modeling training services. Predict service now is doing better and most of the issues commented by Dr. Jiang was resolved. Zhen did a good job about the explaining the values of features when hovering around the input buttons based on Dr. Jiang’s suggestion from last meeting. But there are still major problems with the model training services (See Dr. Jiang’s comments below).
2. Made new comments and suggestions based on testing results today: See below.
3. Evaluated the work done based on the comments made by Jiang previously.
4. Work assignment.

**Issues/Questions and Comments**

Jiang’s comments and suggesting based on the new testing of iMedbot today.

Regarding Model Prediction:

1. Extend the “hover around button” mechanism to other features that contain values that are not common, such as the values of histology2 (see screenshot below)

Graphical user interface, text, application, chat or text message

Description automatically generated

1. When an error message shows up, and the user clicked on OK, the prompt should go back to the spot where the error occurs. For example, when error message is generated due to the fact that the user-selected dataset does not match the required dataset for 10-year model training (see below), then we should go back to the dataset selection page rather than the “prediction or model training” page as it showed during the testing. This was from previous meeting, but not yet resolved.

Jiang’s comments: This issue is still not resolved completely, for example, we had errors when testing modeling training service today, but no error handling, and we (users) sit there but have no idea what to do. The only clickable button is “submit”, but when submitted, error occurs and went back to submit … This is bad, and need to be resolved immediately!

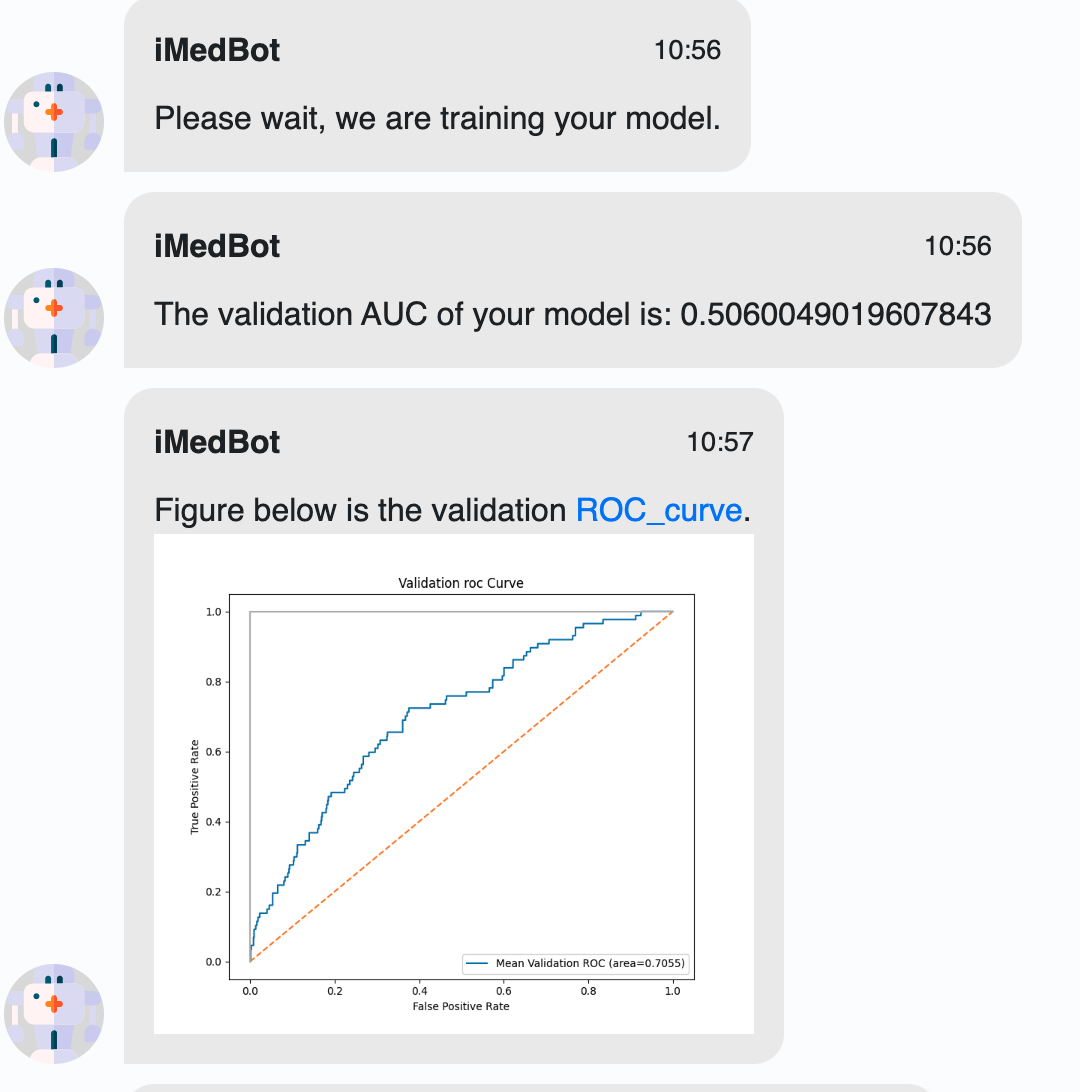
About Model Training

1. When select “Shap”, no results, no error handling.
2. When choose not using default hyperparameters, no results, no error handling.
3. For testing new patients, just use all default values (all 0’s ), the AUC is very high, this is unexpected, because 0’s usually represent all negative values.
4. Even the target class contained in the dataset is “survival”, the result assume it was “breast cancer metastasis”
5. User data statistics and results don’t say what the target class is.
6. The “Prediction New Patients” don’t explain where are the values are coming from.
7. There were no error handling for model training (discussed previously).
8. The list of buttons are not organized. For example, end task is not the last on the list. Another example, run example dataset did not following show example dataset.

**Graphical user interface, text, application, chat or text message

Description automatically generated**

1. Testing with Jiang local datasets did not work at all. Jiang used 00.200.0.txt, data-wisconsin-breast-cancer.csv, and data-00.200.0-2.txt. Did not get any results.
2. Resulted AUC value and ROC\_curve don’t match (see below).



1. The issues mentioned above are obvious ones, which should be found by Zhen and Fixed by Zhen before the meeting. Apparently Zhen did not do a test.

**Ongoing tasks that cover more than a week**

Revise and Improve IMedBot

Tasks will include but are not limited to the following:

1. Revised the current version. Many things, and I will write about them in the specific task for the coming week. Please get all done before our next meeting, and get the errors fixed ASAP and deploy error-free version immediately.
2. Resolve the “deployment” crisis. Currently, we all work on the main branch. When we make a change and push to github, it will trigger an automatic deployment on the AWS site, in which case AWS will charge us. Another problem is when there is a crash in the development work, the main branch will also be affected. Potential solutions: 1. Look into writing our own deploy pipeline without using the paid service (Conder doing this eventually perhaps next year, when you get really familiar with the system). 2. Looking into established a developmental branch, which will not be deployed automatically, but with which we can do development and testing work and conveniently merge it to the main branch for deployment once the new features are confirmed.
3. We will incorporate google analytics to the iMedBot.
4. We will develop a user online survey for the model training service. We currently have a simple online survey for the prediction service, but we don’t have one developed for the model training service call. We plan to further enhance the current survey and develop a new one that is tailored to the model training service
5. We will develop a user registration system that is currently missing;
6. We will develop a backend database during the expansion project. The iMedBot currently does not have a backend storage which can be used to store proper information such as user registration information and user feedback collected via online survey results. The information stored in such a database can be very useful to further improve the quality of the serviced provided by the iMedBot;
7. We will develop an online user manual during the expansion award;
8. We will develop online videos for further user guidance;
9. We will develop a Trello board that would be connected to our current github repository for iMedBot. The Trello board will further promote user-developer interactions and encourage the user involvement in the development work such as testing and providing feedback in real time. It will automatically update the users with the newest development of the iMedBot and inform the developers the user feedback.

**Specific tasks before next meeting.**

Zhen Yang did pretty good work this week to fix most of the problems of iMedbot, discovered last meeting (see Jiang’s comments and suggestion for the previously suggested changes below). There are a few remaining tasks from last week, which are either hard to resolve according to Zhen or not as urgent. Zhen should keep in mind of them until they are resolved.

Based on Dr. Jiang’s tests conducted today during the meeting, we found more severe and urgent issues, especially for the model training service (see the new comments 1 through 11 under **Issues/Questions and Comments).** These issues must be fixed before the next meeting (Monday 2022.12.18).

**Comments from previously (a week or more than a week) and evaluation today.**

1. Need to fix the “logic” of the program, perhaps in many spots. For example, during the test, my dataset was rejected, but I can still click on the “view your dataset”, then once a user clicks on the button, the program paused. This comment is from previous week. Some might have been fix, but more discovered during the text today, such as the choice of hyperparameter value input. Continue to fix the logic flow everywhere, think from the standpoint view of a user. The bottom line is not to confuse user and always user friendly (intuitively).

Jiang’s evaluation: not completely

1. I notice that we need to reword/revise our sentences in many places. For example (See below) the “patient risk of breast cancer recurrence” misses the information of the “year“, and note breast metastasis is not exactly the same as the “breast cancer recurrence”. We should just say the “patient risk of 5-year (or 10-year …) breast cancer metastasis”. Another example is “we are calculating the recurrence” is not accurate. I wonder what is the purpose of the whole sentence. If it is absolutely necessary to have this sentence, we could just say “Thank you, your patient information entry is complete and will be used by our program.” A third example, the whole sentence “What is the following thing you would like to do” does not sound smooth to me. We could just say “Which task would you like to do next?”

Jiang’s comment: Changed to breast cancer metastasis of n-year. Perhaps n-year breast cancer metastasis sounds better.

Further more, rather than repeat the same sentence “Please select one choice according to your situation” in many places, you can separate the difference choices (buttons) more obviously if possible. Currently, all choices appear to be a one big button. One possible solution is adding an outline in different color to each of the choice button.

Jiang’s comments: repeating sentence of “Please select …” is now gone. “One big button” issue is not resolved. Zhen said it was hard to resolve this issue because it was like that as default.

Finally, when a “Send” button at the bottom, which is not used. This is the legacy when we allow a user to enter random questions, which did not seem to work at the time. It should be removed now and can be readded in the right spot in the future if we decide to go back the random conversation mode. This can be considered as one of the “logic” issues as mentioned previously.

Jiang’s evaluation: This is still not done.

Graphical user interface, text, application, chat or text message

Description automatically generated

1. Another rewording example below. “Could you tell me your ER?” is can be replaced by “Could you select the ER status of your patient”. Again, we assume that the iMedBot is talking with a researcher or a medical professional instead of with a patient directly. So the questions should be asked by keeping in mind whom you are talking to. Please check on this throughout the program.

Jiang’s comments: Zhen made changes about these. Looked much better now.

1. Order of the predictors also matters, see blow. I would say it would be more intuitive if ER\_percent follows the ER, rather than the current situation in which PR\_percent follows the ER. Please ensure the predictors appear in the same order as they appear in the original dataset for all three years. I wonder why the order has been changed, for example, in original dataset, the PR\_percent follows PR, but now PR appears way behind PR\_percent, which should not happen even if some of the features were removed.

Graphical user interface, text, application, chat or text message

Description automatically generated

Graphical user interface, text, application, chat or text message

Description automatically generated

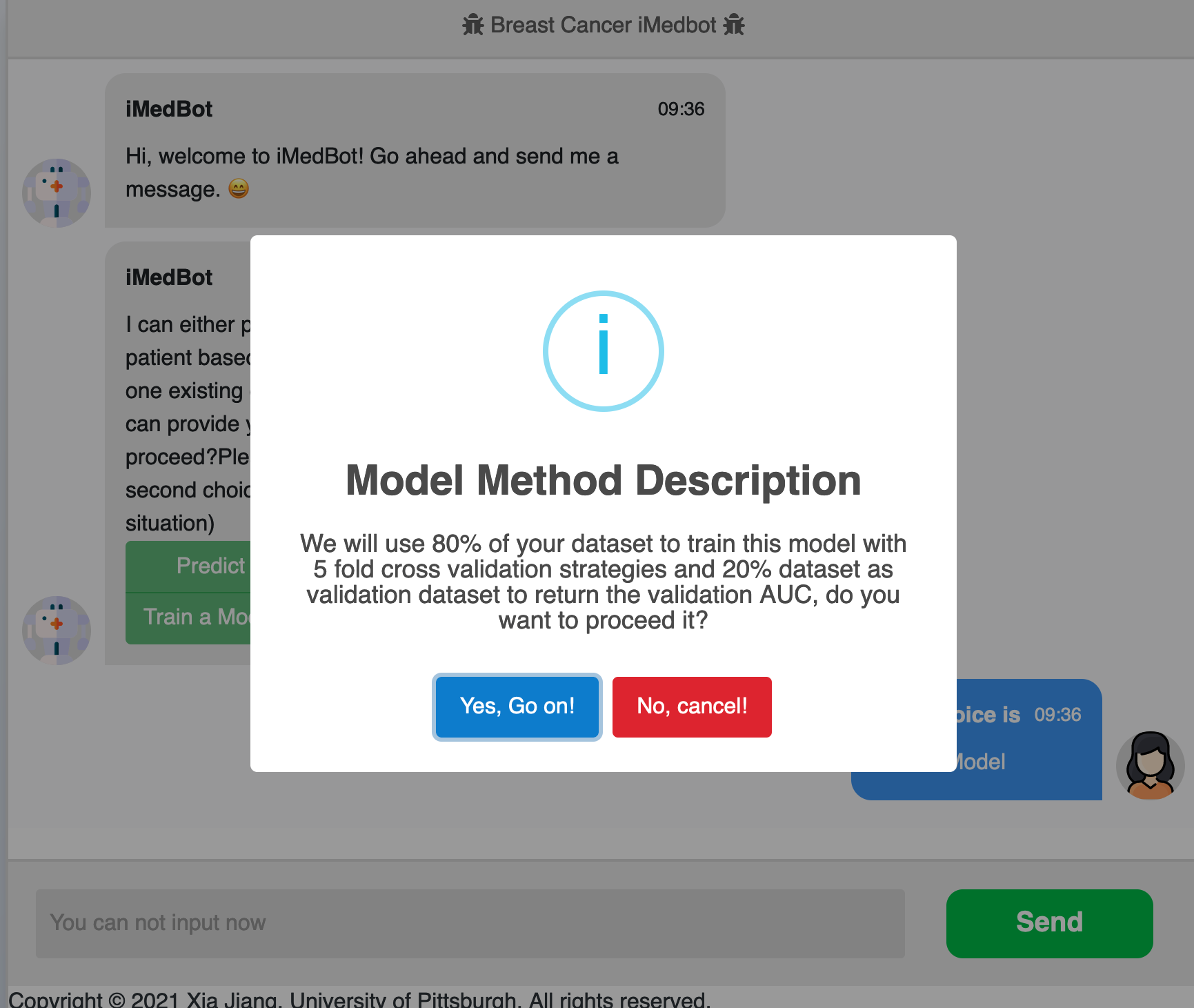
Graphical user interface, text, application, chat or text message

Description automatically generated

Jiang: Zhen said he changed the orders according the order of the original data.

Previous comments about model training

1. Again, please pay attention to wording. For example, what is “Model method description”? Could it be better to say “About model training” …. Or “About ‘Train a Model’ service”. Consider changing the service name from “Train a model” to “Model training”. The paragraph about the model training can be further refined.



1. Need to change the current logic in term of the user choice of the hyperparameter values input. Once the user choose default, then should go with default, rather than re-prompt for the same choice gain.

Jiang’s comments. This was done. But I think now we should add a button to allow a user to change his/her mind. The button could be called “No, changed my mind”. Buttons should be kept but with different wording.

**Graphical user interface, text, application, chat or text message

Description automatically generated**

**Graphical user interface, text, application

Description automatically generated**

1. Besides, the default should be appeared and explained better before the user makes selection. See the last screenshot above.

Jiang’s comments: This was done. It is better to have the hyperparameters explained and be consistent between default and no default. I would suggest show all hyperparameters default and no default.

1. For the hyperparameters that we decide to set as default such as the current mstruct, there is be a good reason, and then it should not show to confuse user. See the last screenshot.
2. Hyperparameters should be better explained somewhere (see the last screenshot above). What is mstruct, what is drate …. We should not assume that user knows the meaning of these terms.

Jiang’s comments: Notice the changes have been made regarding this.

1. Please better explain your results, such as what is AUC, and how a user should interpret the result.

Jiang ‘s comments: Yes, changes were made. Nice! But in term of the meaning AUC, suggest that add a bit more about the interpretation of different values (range of values), for example, 50% is worst … normally between 0.1 and 1, ….

1. Error testing “Predicting for new patients”. See blow.

Graphical user interface, application

Description automatically generated

1. The mechanism for user selecting values did not go smoothly during test. See blow. Need to resolve.

Graphical user interface, application

Description automatically generated

Why learning rate and batch size have different size of font? Again, Learning rate and Batch size and other user input values can be better explained.

Jiang’s comments: Zhen said it was fixed, and added hover around style of explanations.

Jiang’s comments: This is resolved! Good job.

1. When a user dataset is rejected (see example below), please add a mechanism so that a user can go back to the previous page rather restart. A similar problem was already mentioned previously. Need to resolve all similar issues.

Jiang’s comments: This was done in two occasions as indicated in 3)

Graphical user interface, text, application, chat or text message

Description automatically generated

Jiang comments: Resolved.

1. The large user dataset issue. During the test today, the 52 MB dataset that Dr. Jiang chose did not work and the program just paused, and not further hints for user. Suggestions: 1) Do a thorough test on your computer and give hint to user based on your test, such as given a computer hardware information, what is the maximum size of user dataset expected; 2) If user selected an very large dataset that is not expected not working with our current system, we should reject it and give a reason; 3) In case the training service is working properly as happened today, we should give a message explaining why (such as slow internet at the user site), and give prompt for further action.

Jiang’s evaluation: The current maximum was set as 500K by Zhen because he said this was already set in the current system. I suggest that Zhen do some tests and research (for example from the AWS EC2 capacity) and find out the absolute maximum. This instruction was given multiple times including last meeting (See above).

Jiang: Not done yet.

**Jiang’s new comments based on today’s testing.**

1. During the testing (15 year prediction), the program stopped working after making some value selections. See screenshot below, after selected the duct as the input value of Histology, everything is greyed, and can go anywhere further.

Graphical user interface, text, application, chat or text message

Description automatically generated

Jiang: solved!

2)

The instructions for Uploading Local Dataset (see below) should appear after a user clicks on “Uploading Local Dataset”, rather that appear only when a user clicks on “View Example Dataset”.

A picture containing table

Description automatically generated

Jiang: changed according to my instruction.

3)

The page for the meanings of the features values should be further formatted: For examples (see screenshot below), there should be some space to the right of the page; could user bold font for each of the terms; be consistent in terms of the format of the terms.

Text

Description automatically generated

Jiang: Changed and kept both. Nice work!

4)

In terms of explaining the meaning of the feature values, the current solution (as shown in 3)) works, but to make it more user-friendly, can you use the “hovering around” mechanism instead of opening up a separate page?

Jiang: Zhen did a good on this!

**Less urgent tasks**