User Guides

Group DS_06

Harry Lawson, Li Yunhao and Daniel Clark

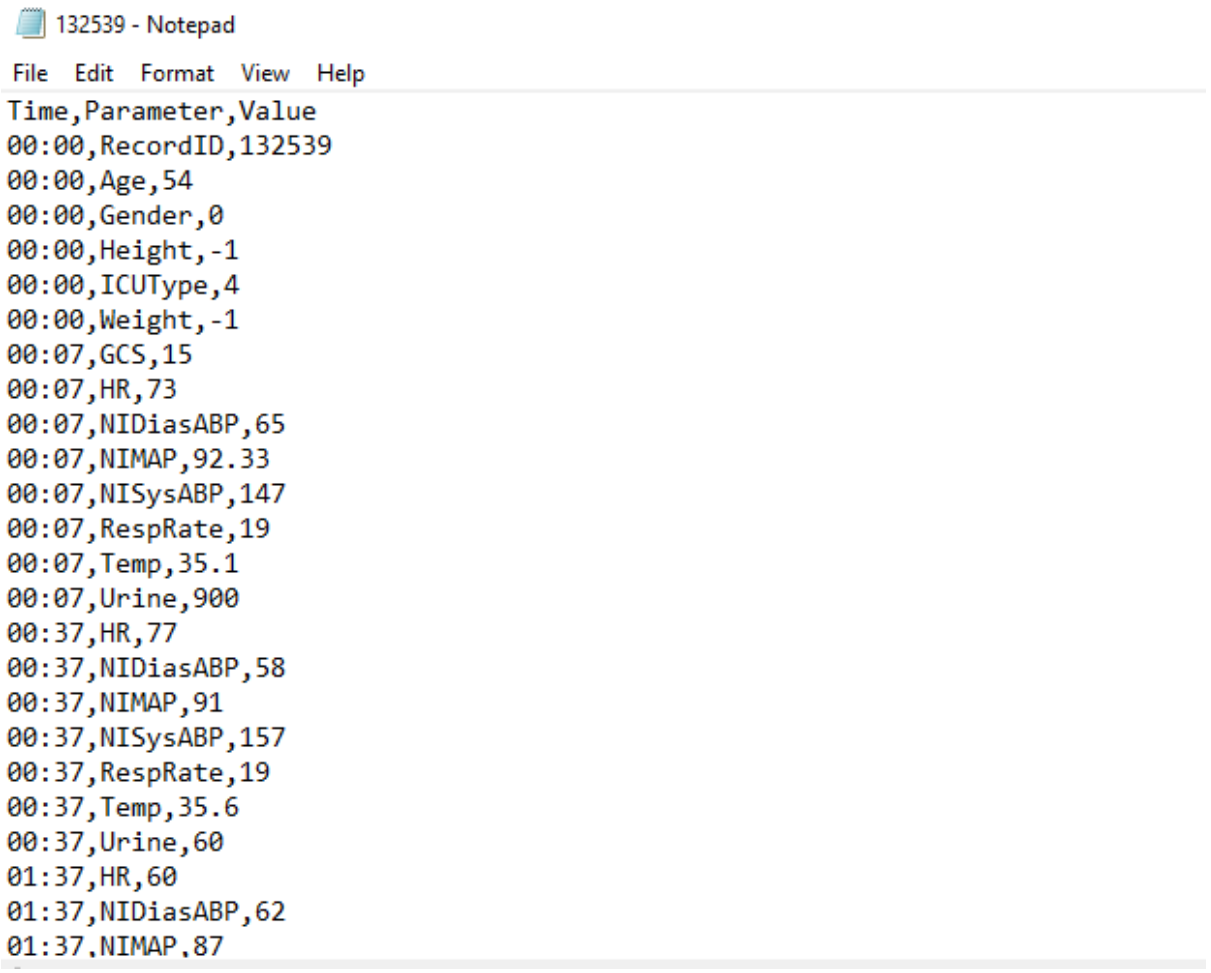
Word count: 1485

Table of Contents

End User Guide

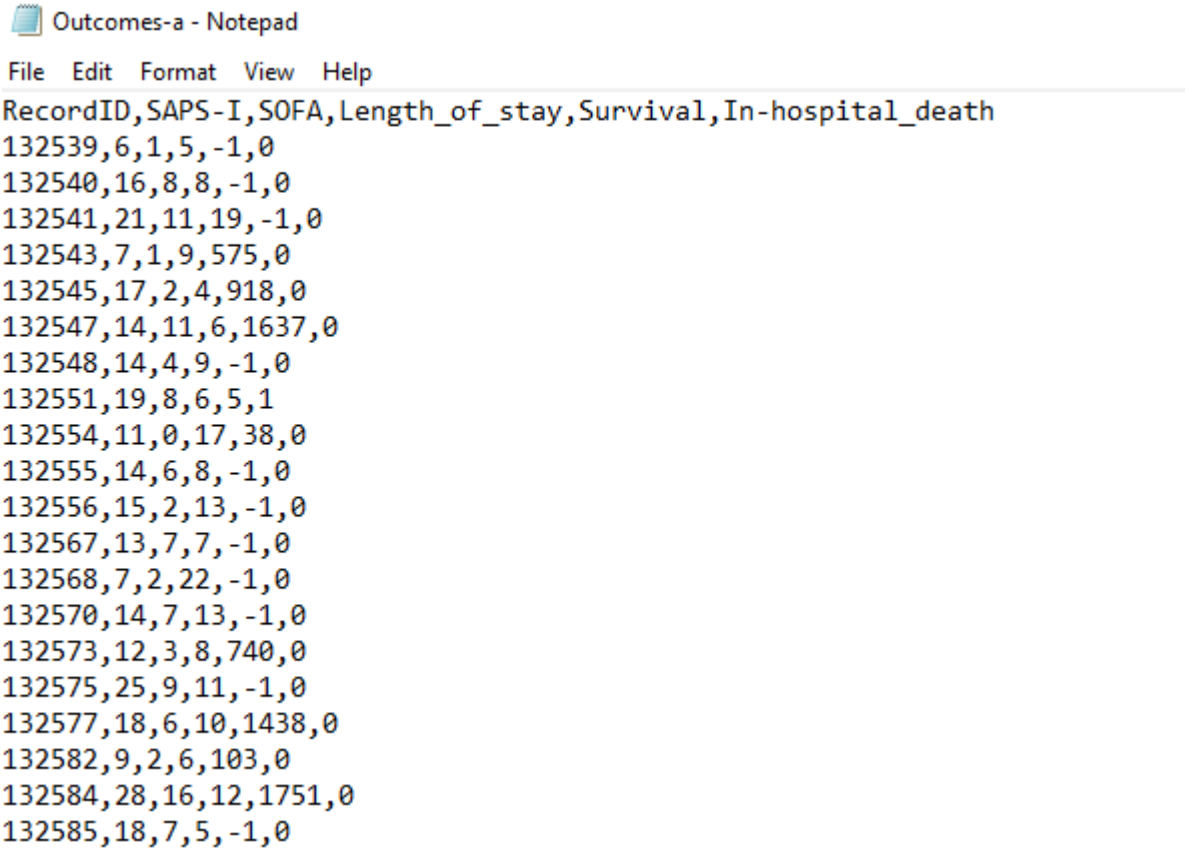
1.1 Data Preparation

Data is expected to be prepared in the same time series format and with the same variables as those in the PhysioNet Challenge. This means that the measurements for each patient at each time interval should be collected into a txt file as demonstrated in Figure 1 below. These files do not have any naming convention.



(Figure 1. Format for patient files)

The files for every patient must be put into a folder named ‘test_set’. Additionally, the outcomes for each patient should be uploaded into a single file named ‘test_set_outcomes’ that is formatted the same as the outcomes files in the PhysioNet Challenge. The format of this file is demonstrated in Figure 2 below.



(Figure 2. Format for outcome files)

Note that for accurate analysis, no missing values should exist in the ‘test_set_outcomes’ file. Missing values are accepted in the patient data but will reduce the accuracy of the analysis.

The folder ‘test_set’ (containing each patient's data) and the file ‘test_set_outcomes’ must then be placed in a new folder which must be compressed into a zip file. There are no required naming conventions for this zip file.

2.1 Sign up

The user will be directed to the home page of the web application when they access the local server. To sign up, the user will navigate to the sign up page by clicking the ‘Sign up’ tab on the home page as demonstrated below in Figure 3.

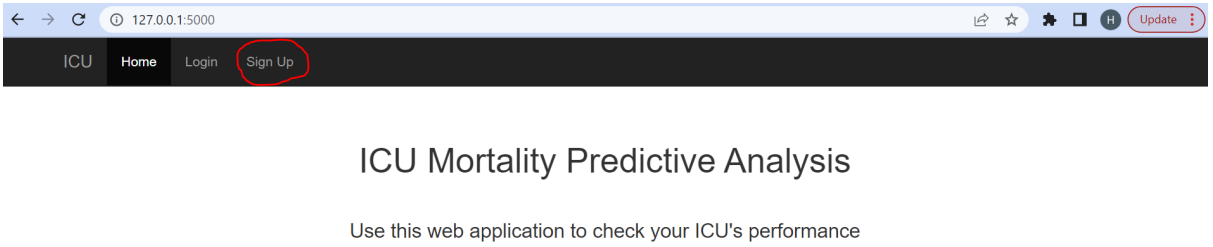


Figure 3. Sign up tab

The user will then be redirected to the sign up page and will be required to create a username (that is not already taken), enter a valid email address (that is not already taken), and enter a password that is at least 8 characters long (a possible signup is demonstrated in Figure 4 below).

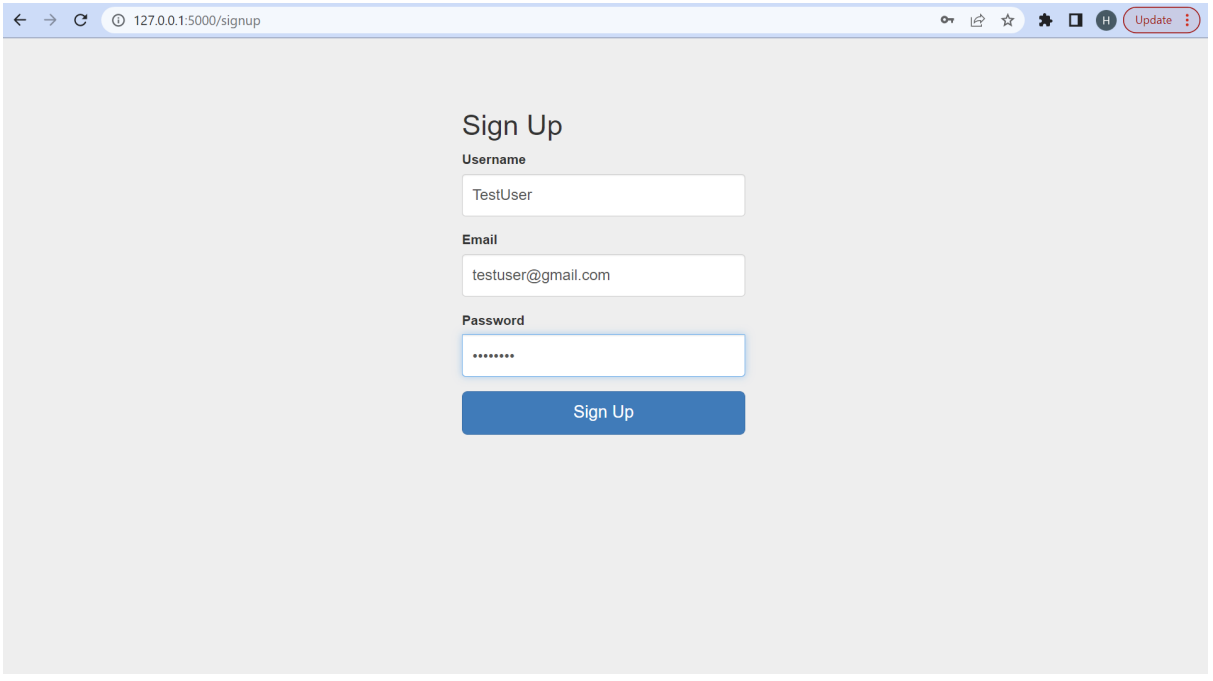


Figure 4. Sign up webpage

3.1 Log in

Once these requirements are met, the user will be redirected to the login page. The user must now enter their valid username and password to access the web application as demonstrated below in Figure 5.

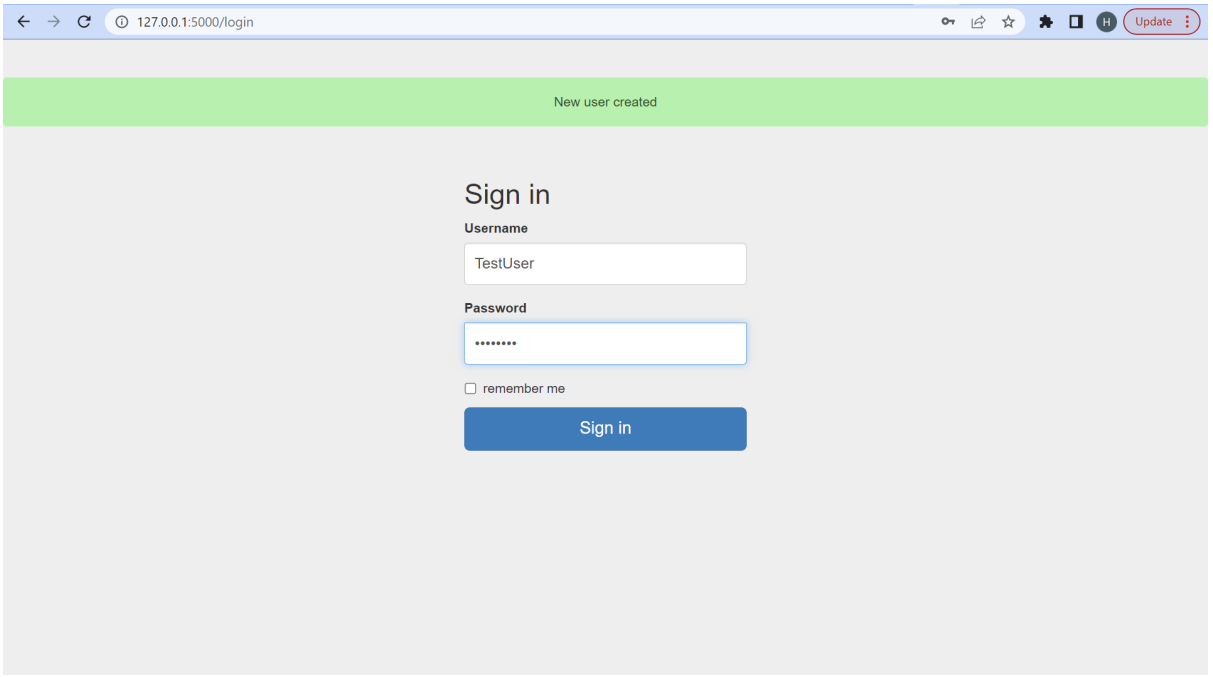


Figure 5. Sign in webpage

4.1 File upload

The user is then redirected to the file upload page of the web application. To upload the compressed file containing their patient data information (that is prepared correctly as per part 1 of the end user guide) the user must click the ‘Upload File’ button as demonstrated below in Figure 6 and select their compressed file as shown in Figure 7.

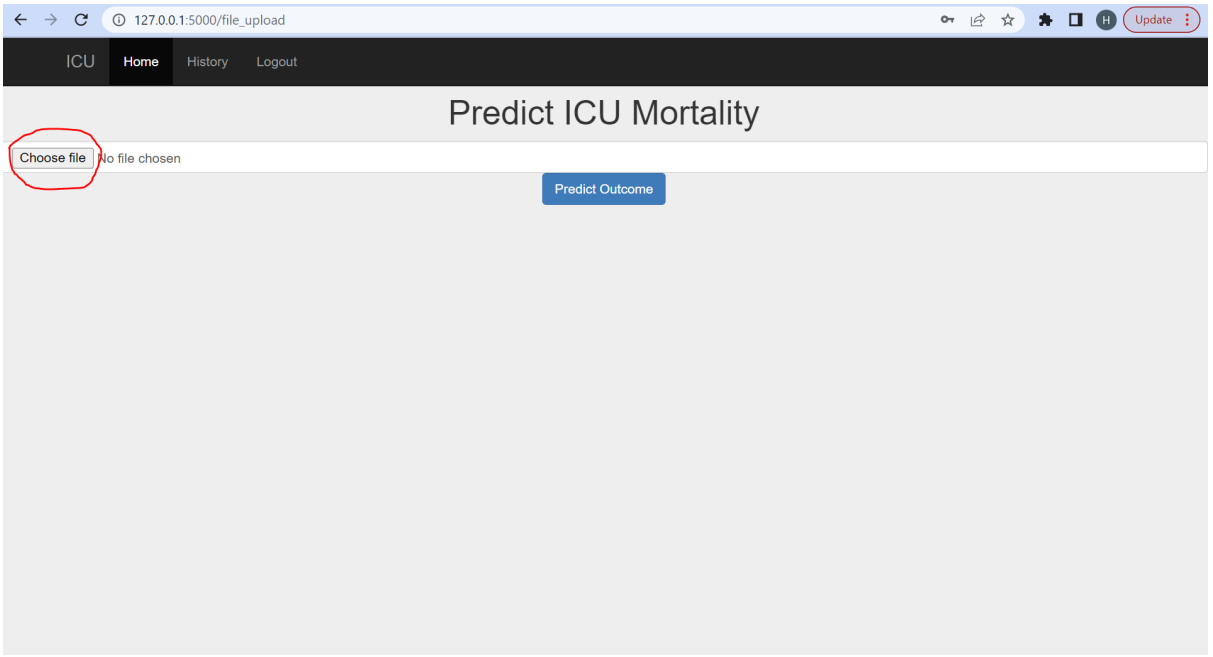


Figure 6. Choose file

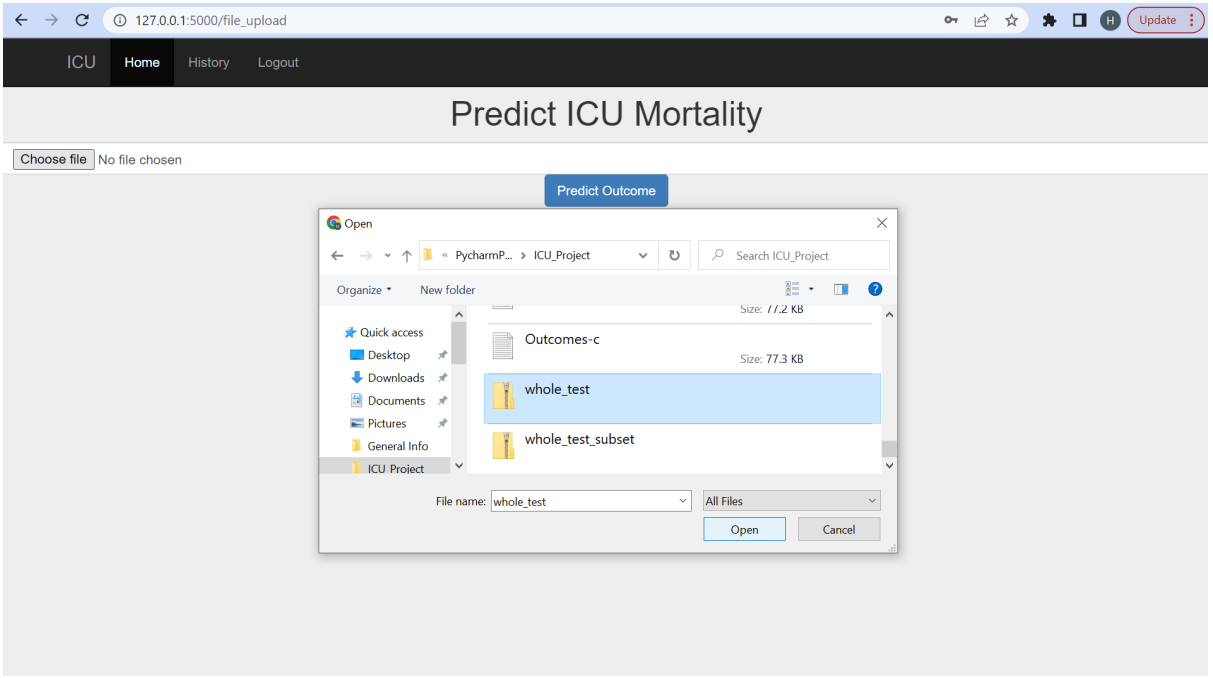


Figure 7. Open compressed file

Following this, the user must click the Predict Outcome button to run the analysis.

5.1 Analysis Page

Following the upload of the data, the web application will take time to cleanse, impute, and analyse the data. How long this will take depends on the amount of data uploaded and how many missing values are contained in the patient datasets. For a subset of the data provided by PhysioNet Challenge, the average time to compute analysis per patient is approximately 0.25 seconds.

Once this process is complete, the user will be redirected to the analysis page. The analysis page shows the number of patients uploaded, the number of patients that actually survived, the number of patients that were predicted to survive, and the overall score for the user's ICU unit for this sample of data compared to industry average. Additionally, this page includes a table for each Patient ID, Actual Outcome, Predicted Outcome, Probability of Outcome, and a link to a LIME analysis for that patient. This is demonstrated in Figure 7 below.

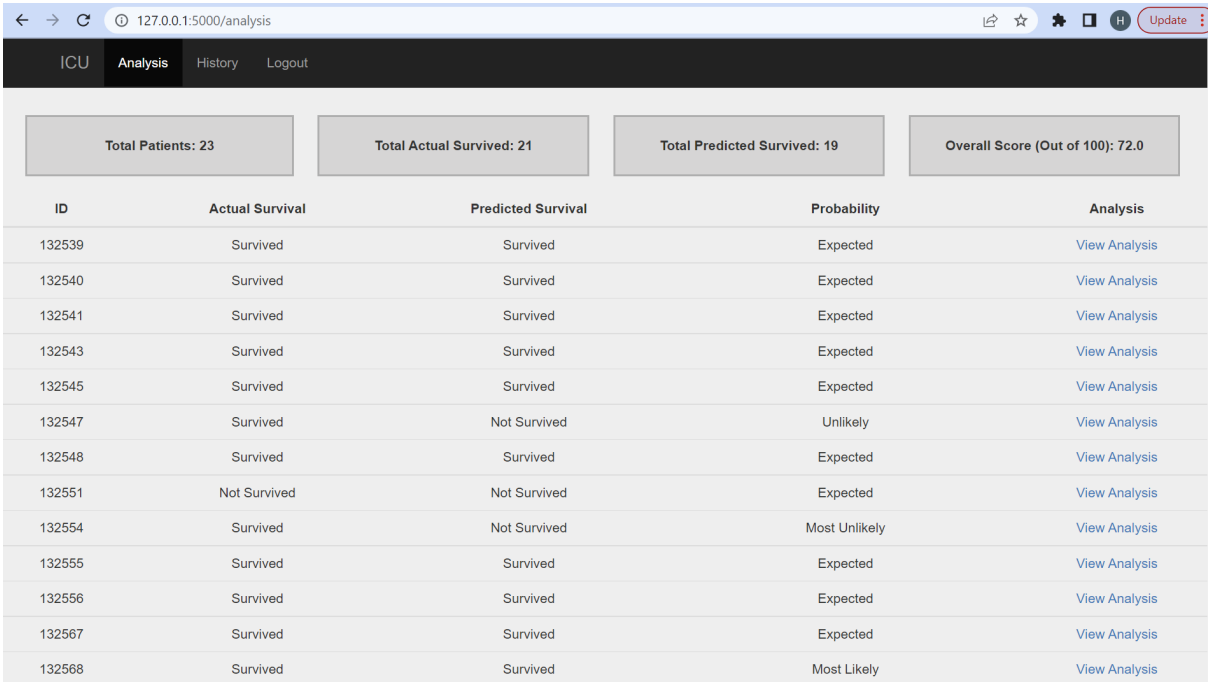


Figure 7. Analysis page

To access the LIME analysis for each patient, the user must click on that patient's analysis link as demonstrated in Figure 8. The time the LIME page takes to load will depend on the speed of the computer hosting the application but should be within 3 minutes.

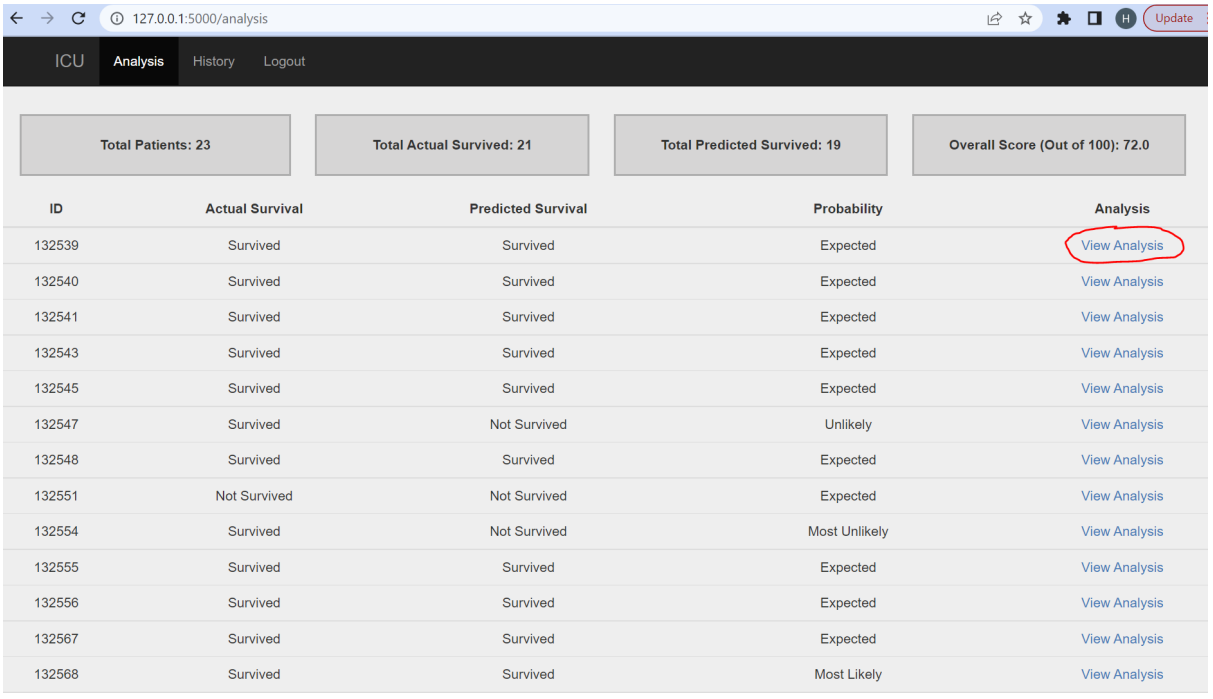


Figure 8. Access the LIME analysis.

This will open up a new tab on the user's web browser and demonstrate the LIME analysis. Every LIME analysis page will show the probability of the patient surviving (denoted as 0) and the probability of not surviving (denoted as 1) in the top left corner. Additionally, the LIME analysis shows which variables contributed to these probability predictions (shown in the colour coordinated

centre vertical bar chart) and what values each patient had for these variables (shown in the colour coordinated top right table). An example of patient 132582 is demonstrated below.

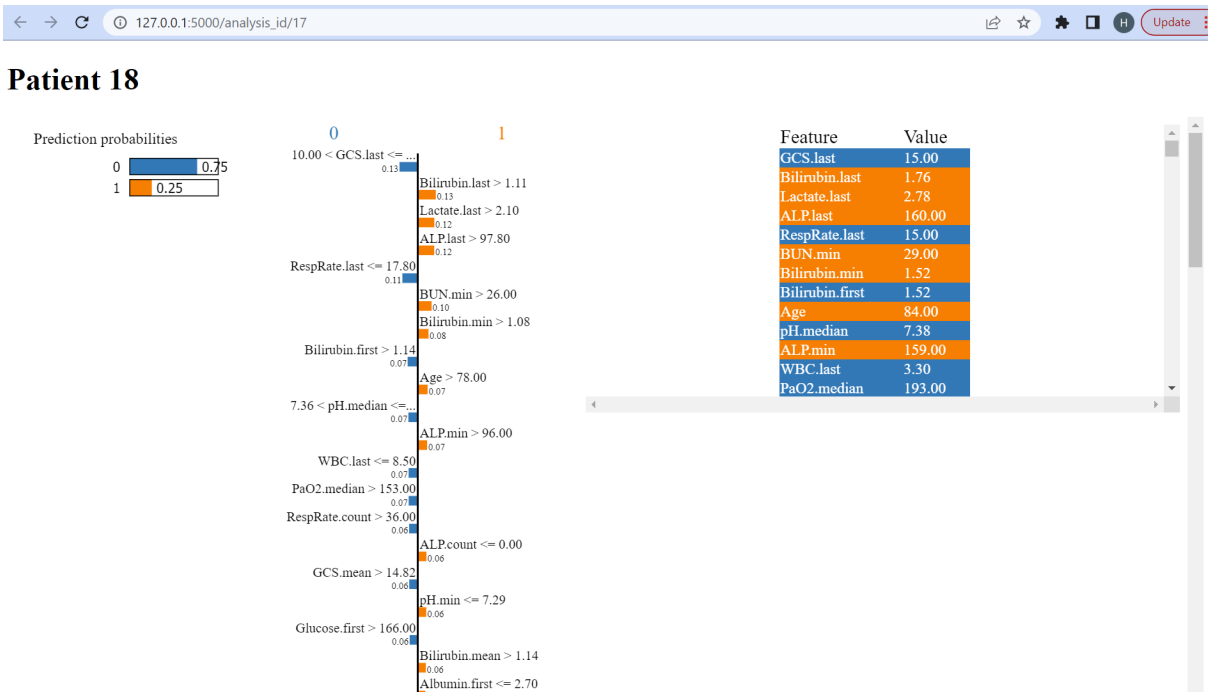


Figure 9. LIME analysis page.

6.1 History Page

To access the History page, the user must click on the history tab at the top of the webpage (shown in Figure 10). This will show the user any previous uploads they have performed as shown in Figure 11. The user can choose to re run previous uploads by selecting the analyse button.

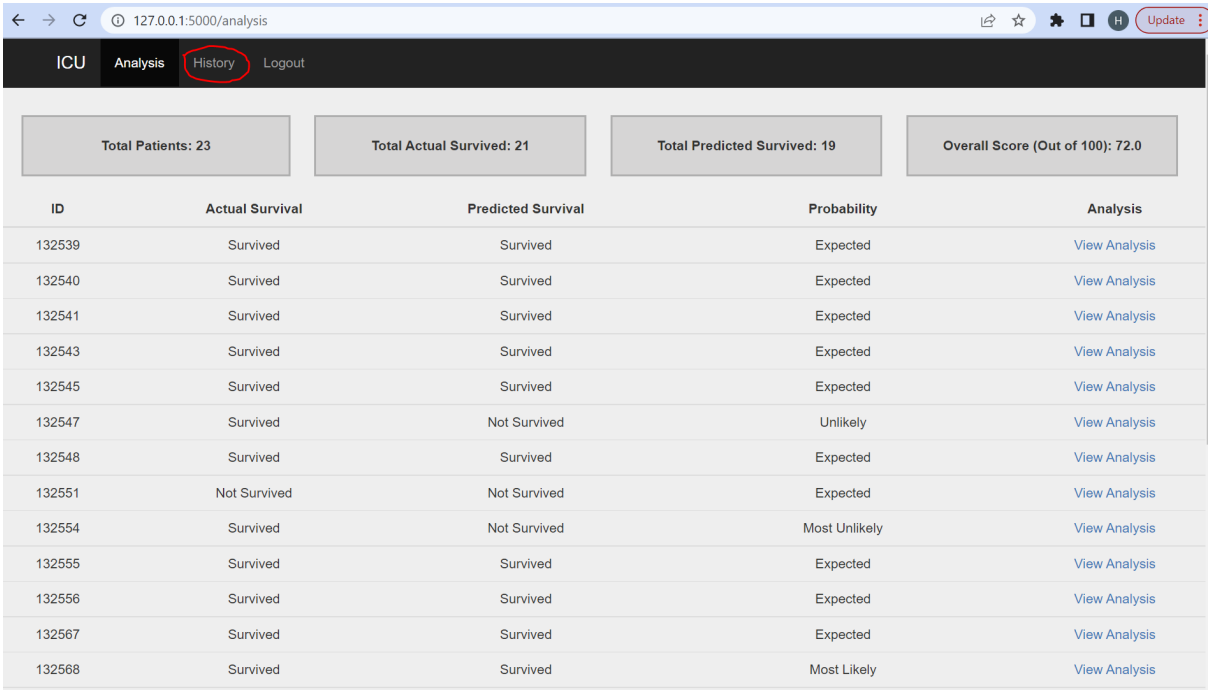


Figure 10. Select history page

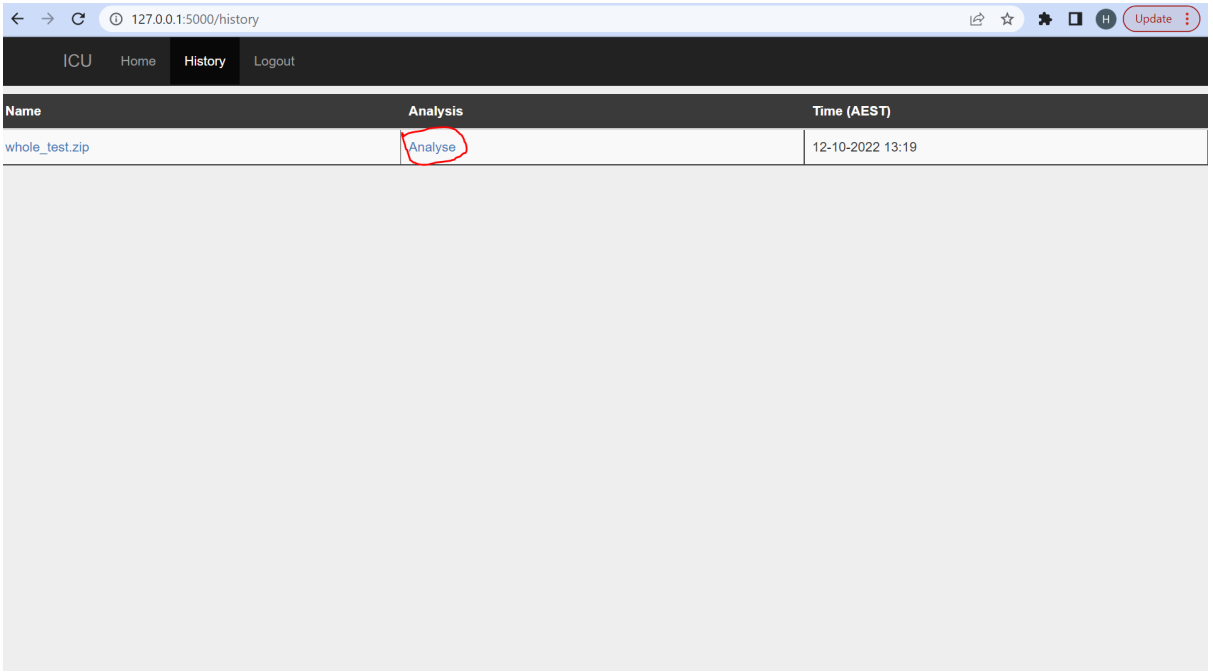


Figure 11. History page

To upload a new dataset, the user must click the home tab at the top of the webpage (shown in Figure 12) and then follow instructions for 4.1.

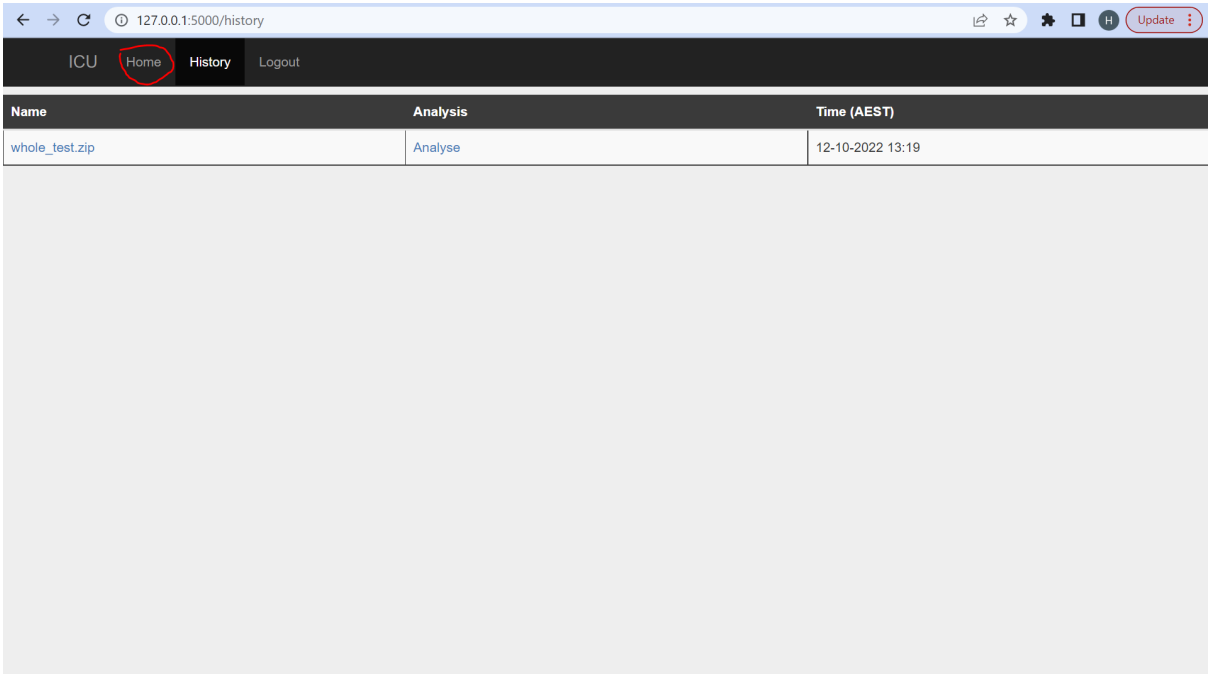


Figure 12. Home page for file upload

7.1 Log out

To log out, the user must click on the logout button as shown below in Figure 13 and will be redirected to the homepage for users not logged in.

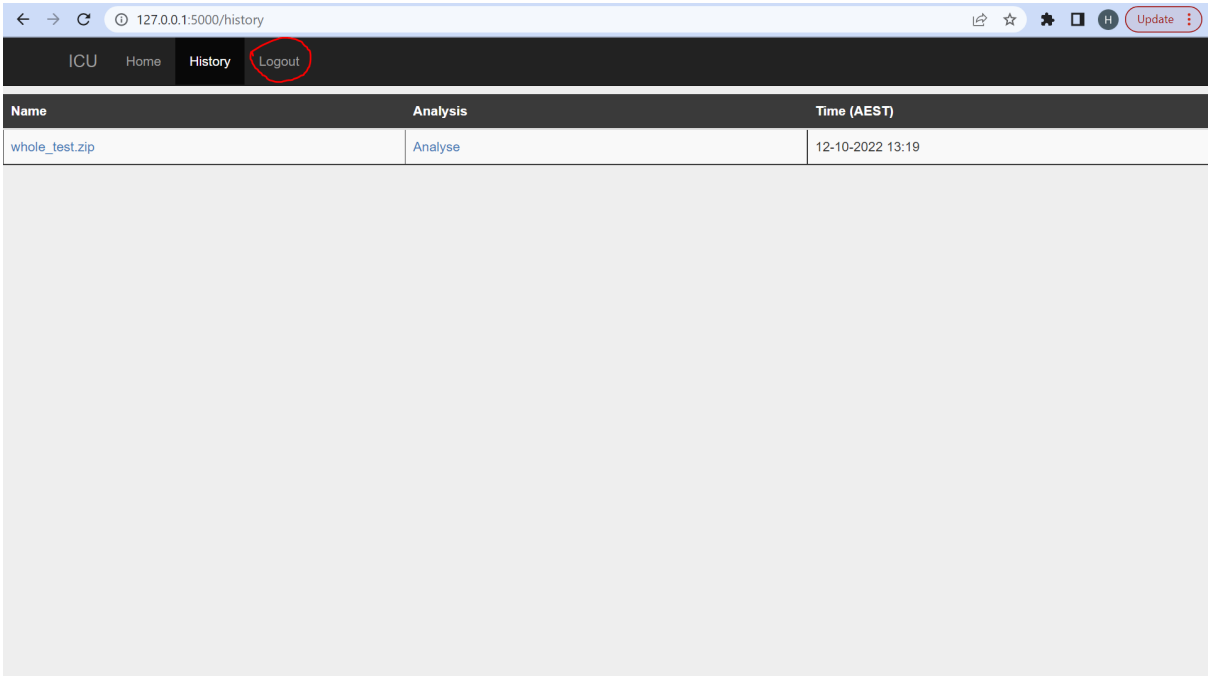


Figure 13. Select logout tab

Technical Guide

1.1 Setup development environment

To follow the technical guide it is recommended the admin uses an Anaconda Python environment and Visual Studio Code as the code editor. Download and install Anaconda here and Visual Studio Code here.

Multiple Python libraries will be installed by using the conda environment provided in the code folder called ICU_conda_environment.yml.

To import this into Anaconda, open up Anaconda and select the environments tab (Figure 14). Select the import button (Figure 15), select the open file explorer button (Figure 16) and choose the ICU_conda_environment.yml file (Figure 17).

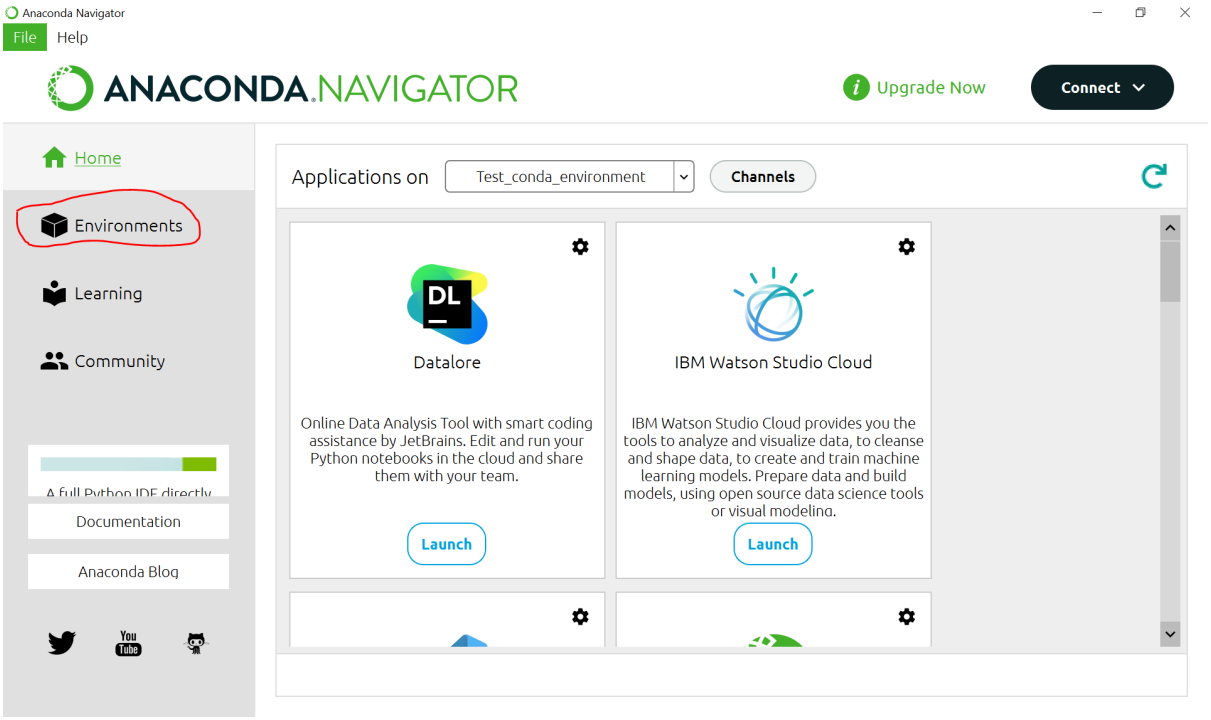


Figure 14. Select Environments Tab

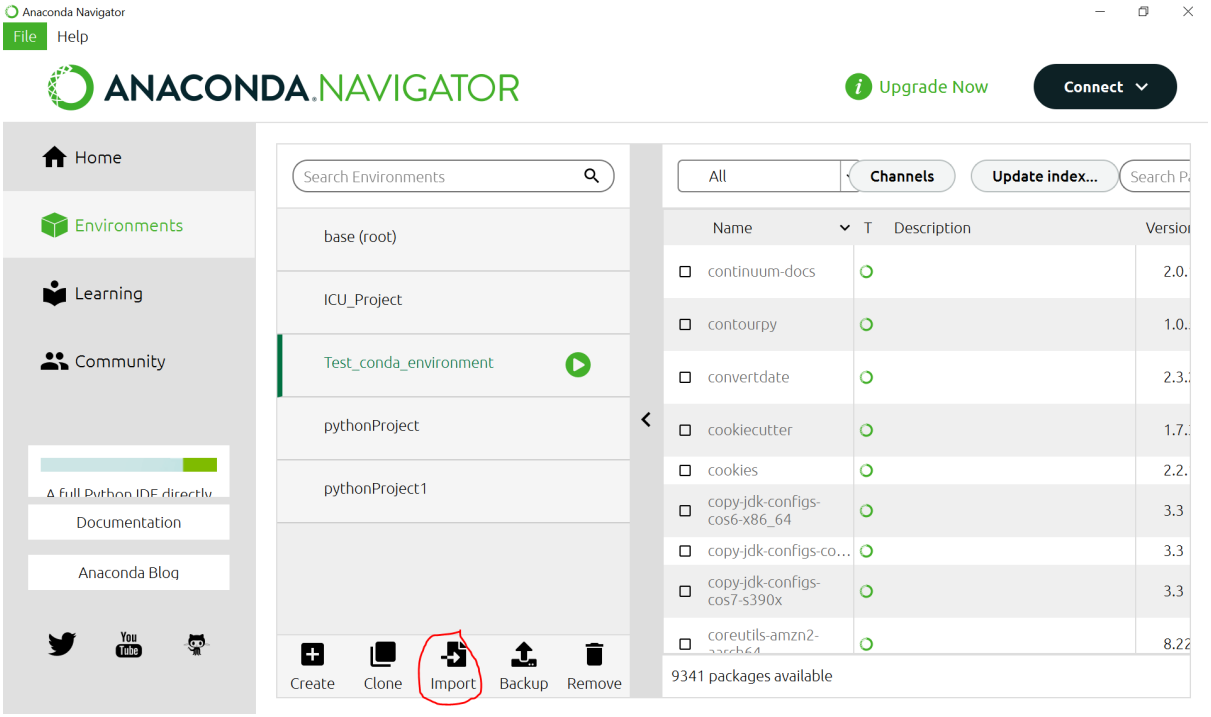


Figure 15. Select the import button

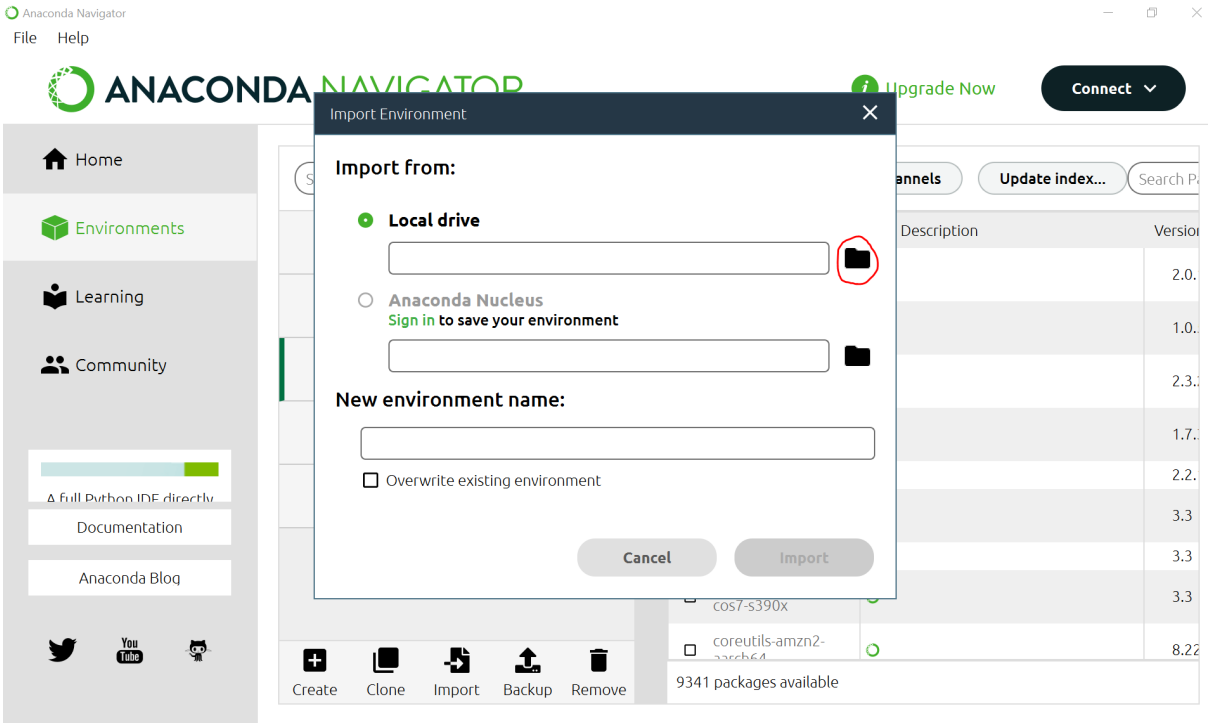


Figure 16. Select the open file explorer button

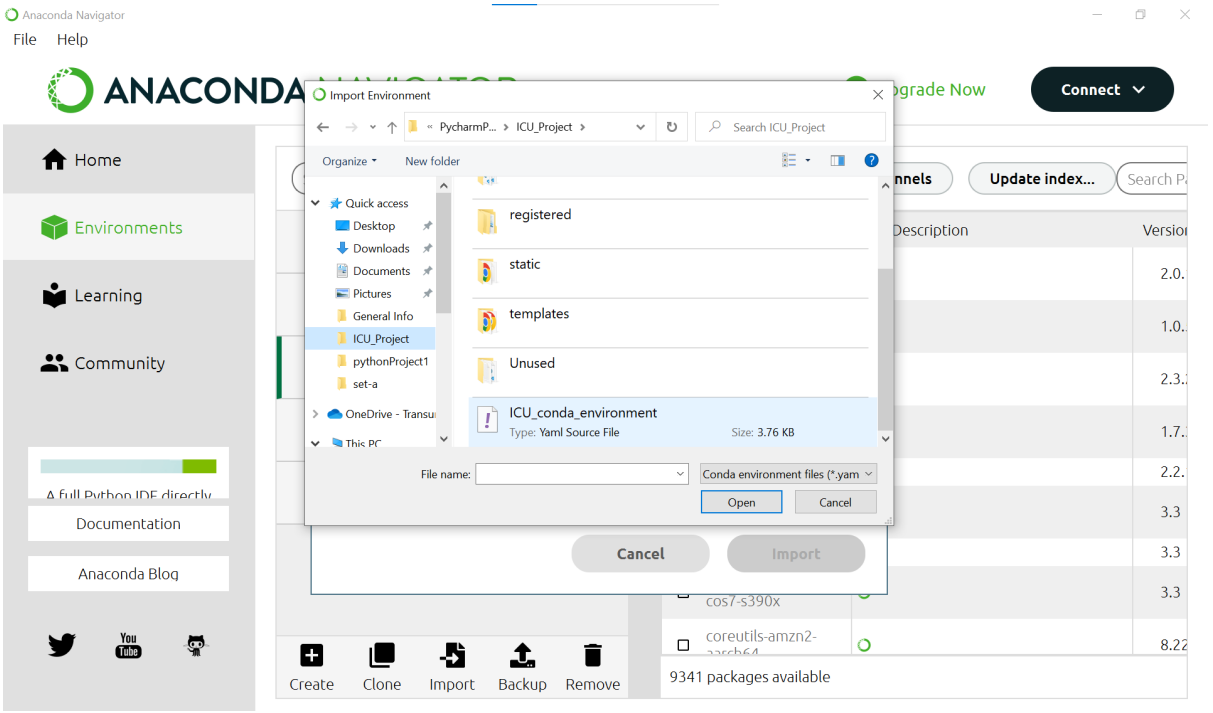


Figure 18. Choose ICU conda environment

Once the conda environment has been imported open Visual Studio Code and select the extension panel (Figure 19). Search Python in the search bar and click install (Figure 20).

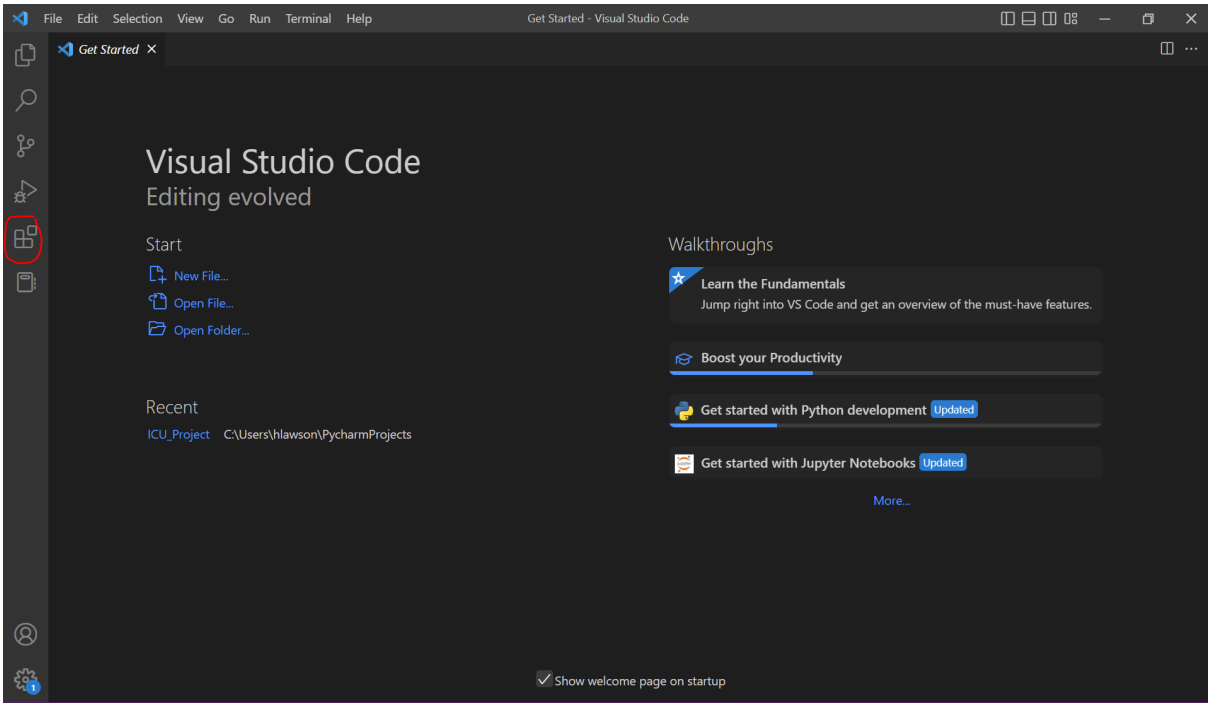


Figure 19. Select extensions

Figure 20. Install Python

Once Python is installed, select the explorer button (Figure 21) and close the extension window (Figure 22).

Figure 21. Select explorer button

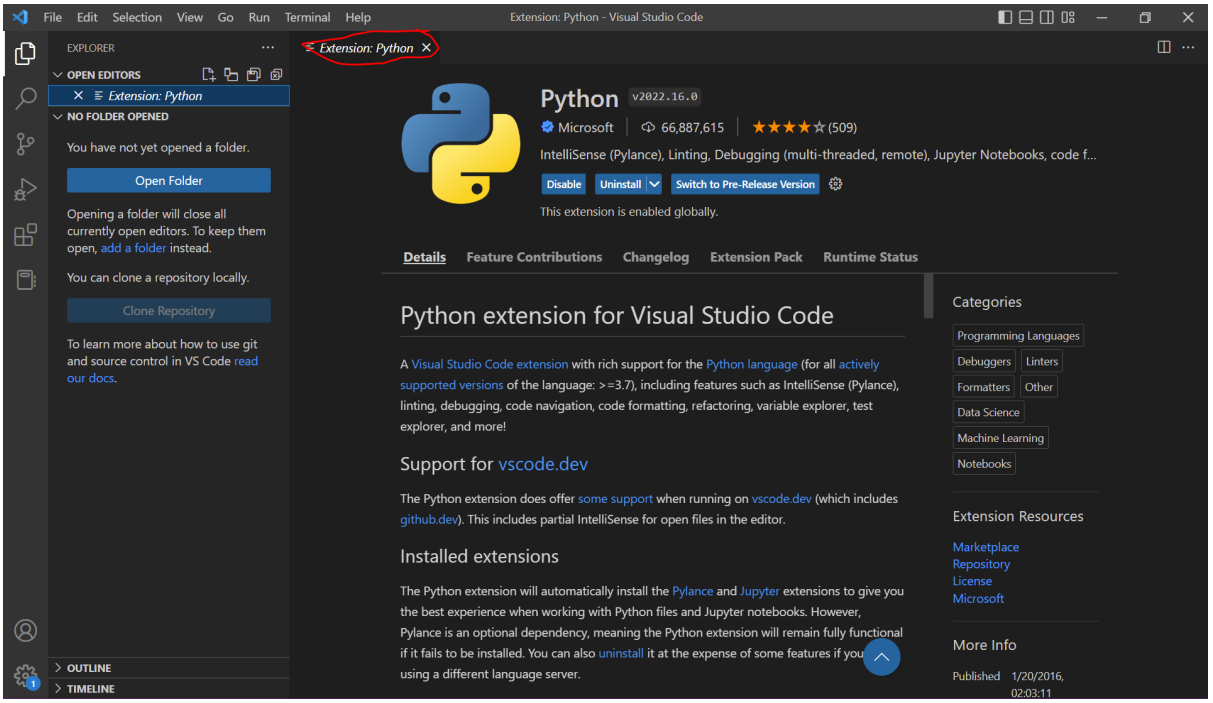


Figure 22. Close extension tab

Choose File > Preferences > Settings (Figure 23). Then type in 'Python path' in the search bar (Figure 24).

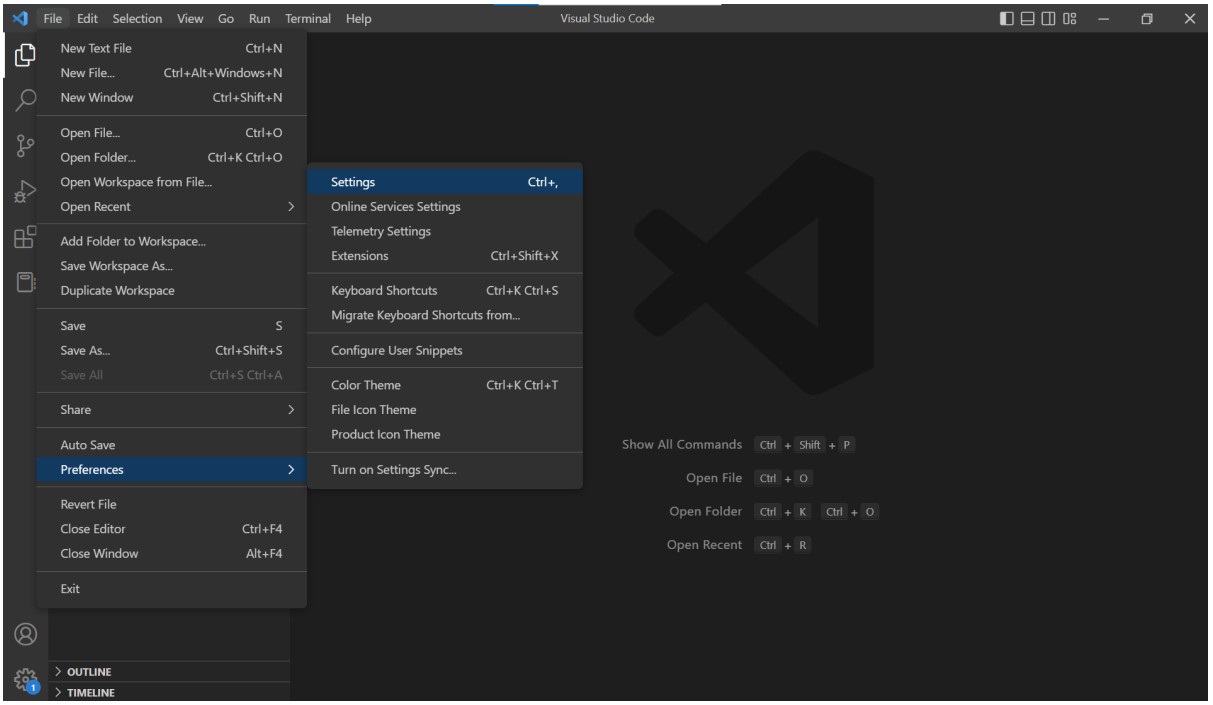


Figure 23. Select File > Preferences > Settings

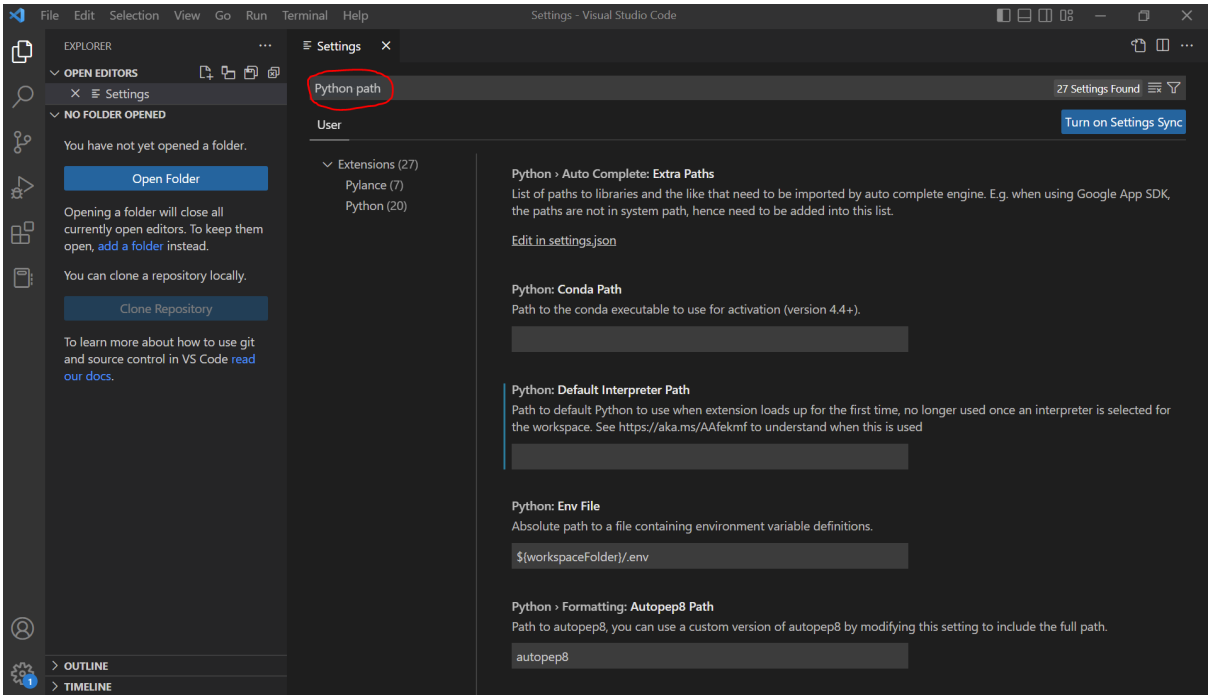


Figure 24. Search for Python path

Open up the File explorer app on the admins computer and search for and select the ‘Anaconda3’ folder in the search bar (Figure 25). Select the envs folder (Figure 26) and then select the ICU_conda_environment folder (Figure 27). Select the search bar and copy the current pathway (Figure 28).

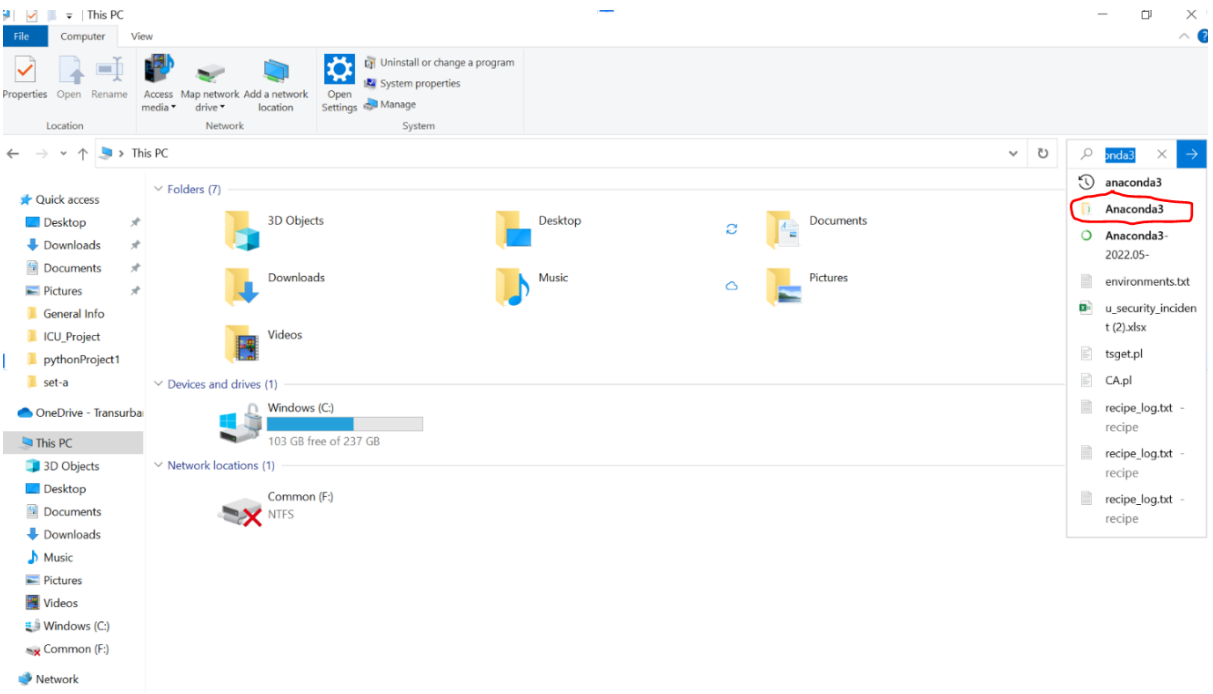


Figure 25. Find Anaconda3 folder

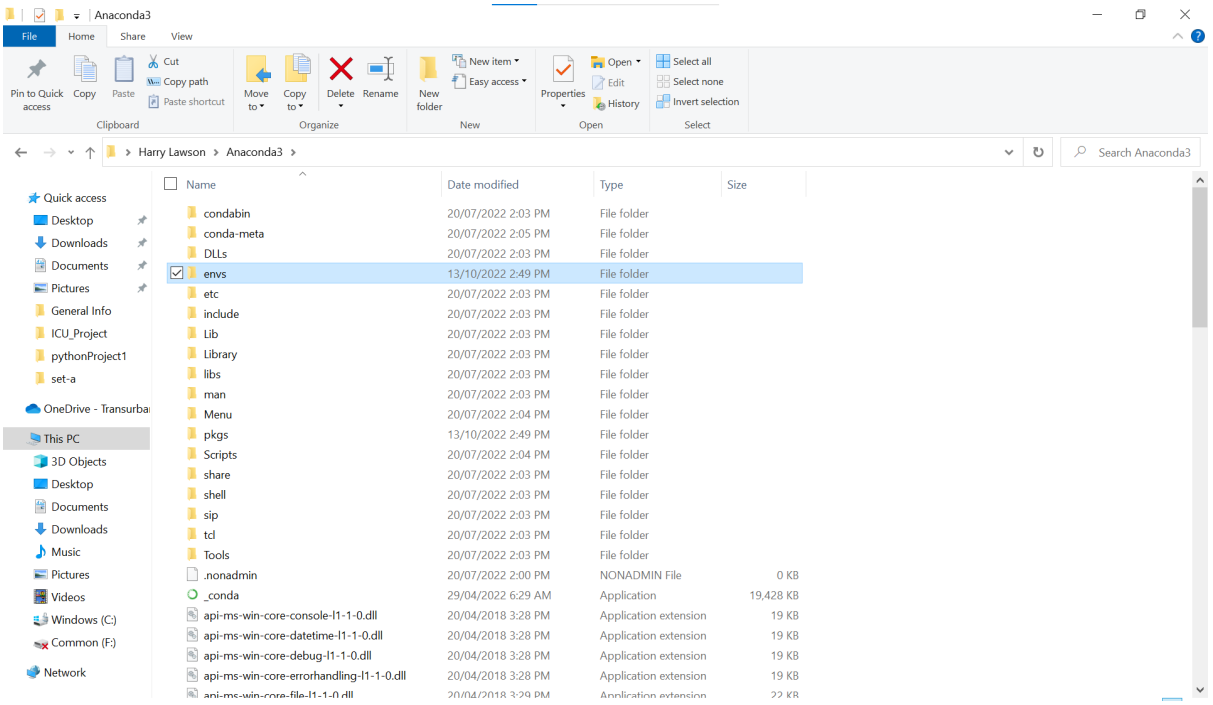


Figure 26. Select the envs folder

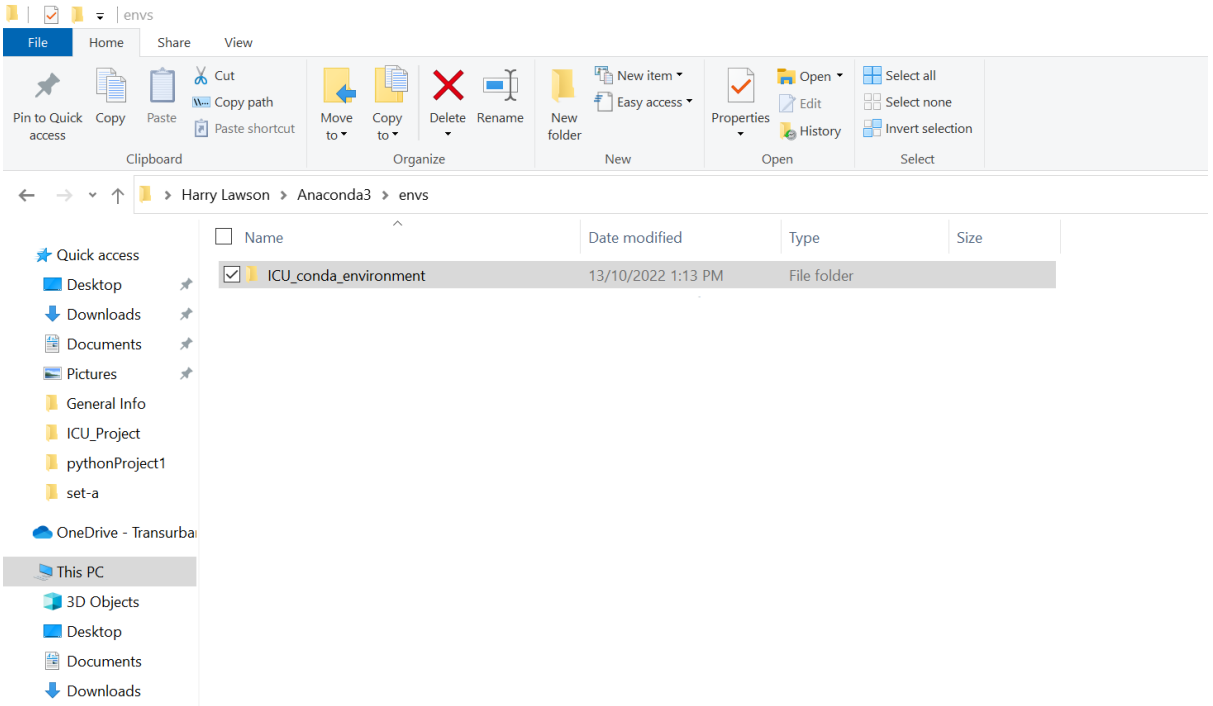


Figure 27. Select the ICU_conda_environment folder

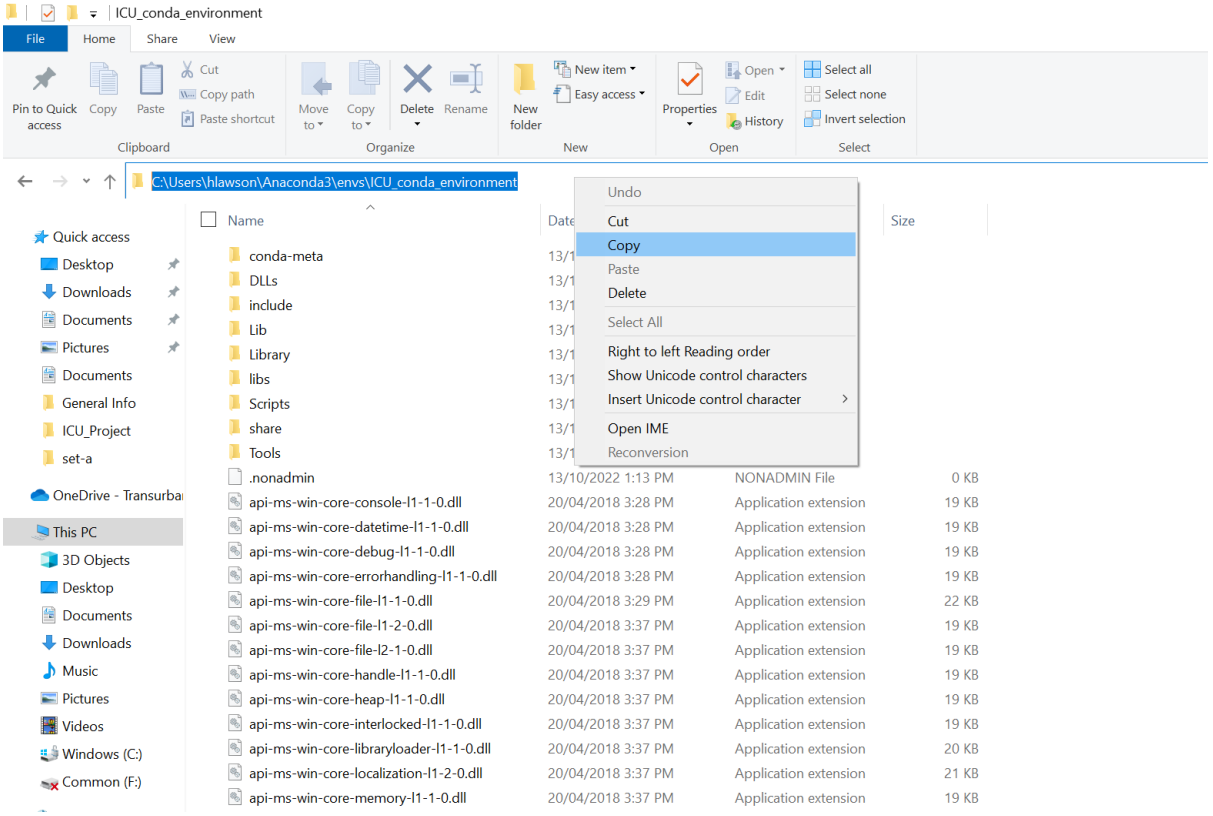


Figure 28. Copy path

Return to Visual Studio Code and place the pathway in the Python: Conda Path box and add '\python.exe' to the end of the path (Figure 29).

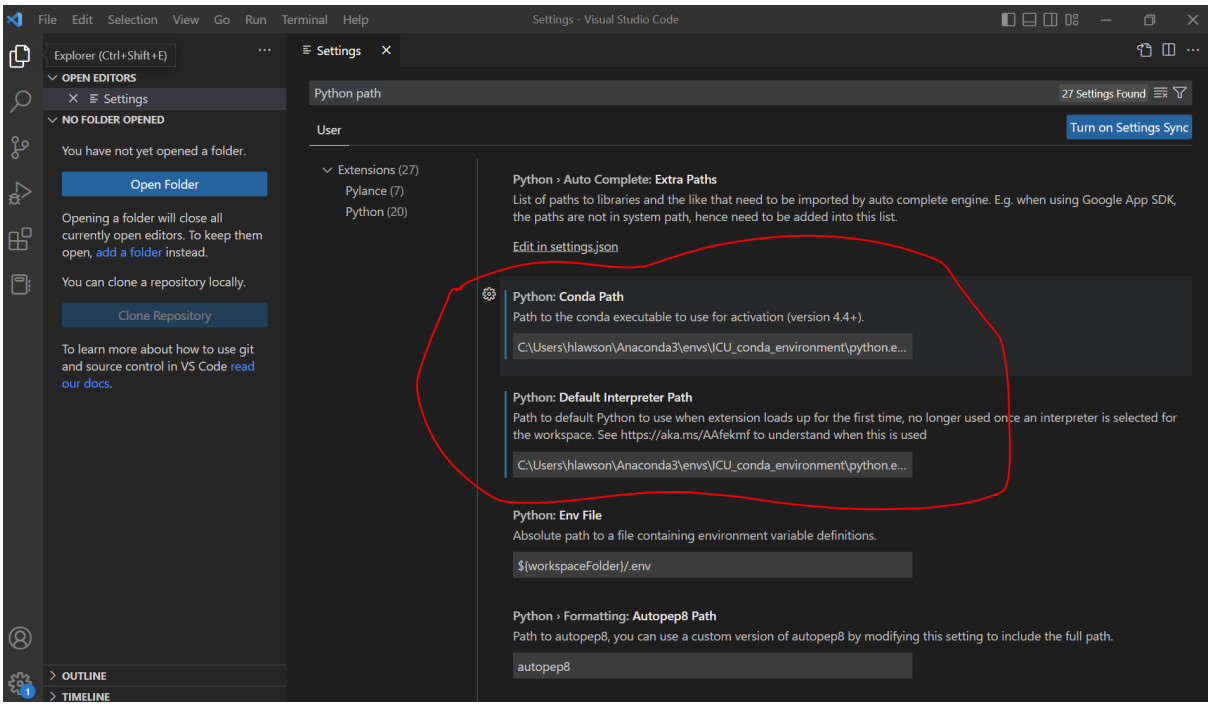


Figure 29. Add path + \python.exe

Open the code_folder file by selecting the ‘Open Folder’ button (Figure 30).

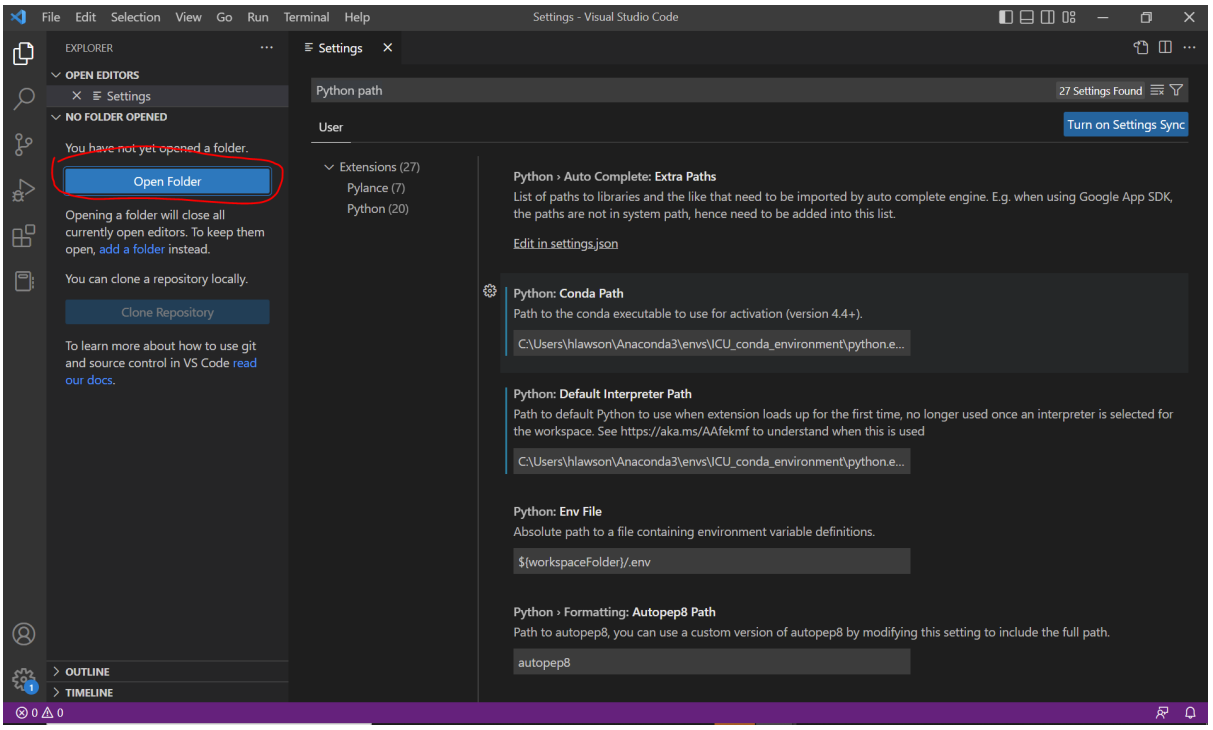


Figure 30. Open folder

2.1 Explanation of folders and files

The entire code folder contains python files, html files, css files, a compressed file containing the data used for testing, the PhysioNet dataset, and a machine learning model file.

2.2 Python files

All python files have documentation explaining what each file and function is doing.

The web application code is run using the example_login.py file.
The data wrangling code is run using the data_collection.py file.
The data imputation code is run using the knn_impute.py file.
The predictive analysis is run using the neural_network.py file.
The LIME analysis page is run using the lime_applied.py file..

2.3 HTML files and CSS files

All of the HTML files are found in the templates folder. All of the CSS files are found in the static folder.

The web applications home page uses index.html and index.css.
The web applications sign up page uses signup.html and signin.css.
The web applications login page uses login.html and signin.css.
The web applications file_upload page uses file_upload.html and file_upload.css.
The web applications analysis page uses analysis.html and analysis.css.
The web applications LIME page uses analysis_id.html.

The web applications history page uses history.html and history.css.

2.4 Remaining files

ICU_dataset_death_knnimputed.csv is the PhysioNet dataset and is used when the LIME analysis is run. The files whole_test.zip and whole_test_subset.zip contain data that can be used to upload during testing the web application. The model_file_name.pkl is the machine learning model.

3.1 Make application live through local host

Open up the ‘example_login.py’ file (Figure 31). Select the run button (Figure 32). Once the development server is ready, paste the given link (Figure 33) in any web browser.

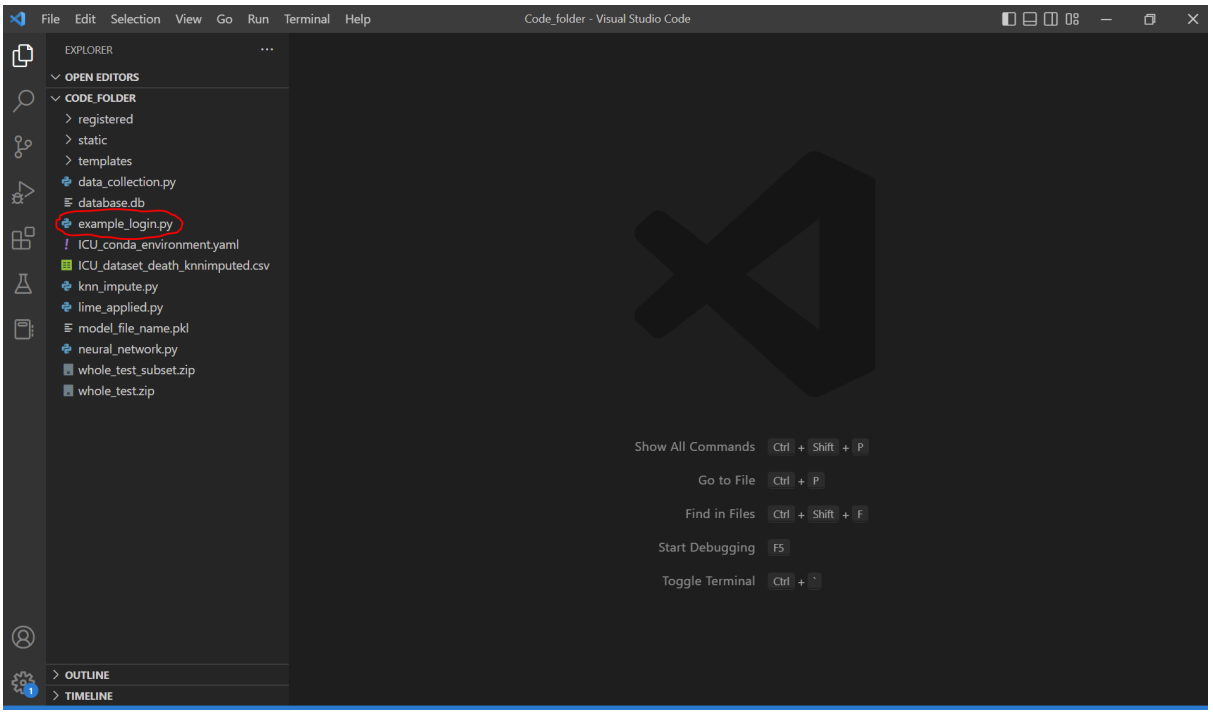


Figure 31. Select the example_login.py file

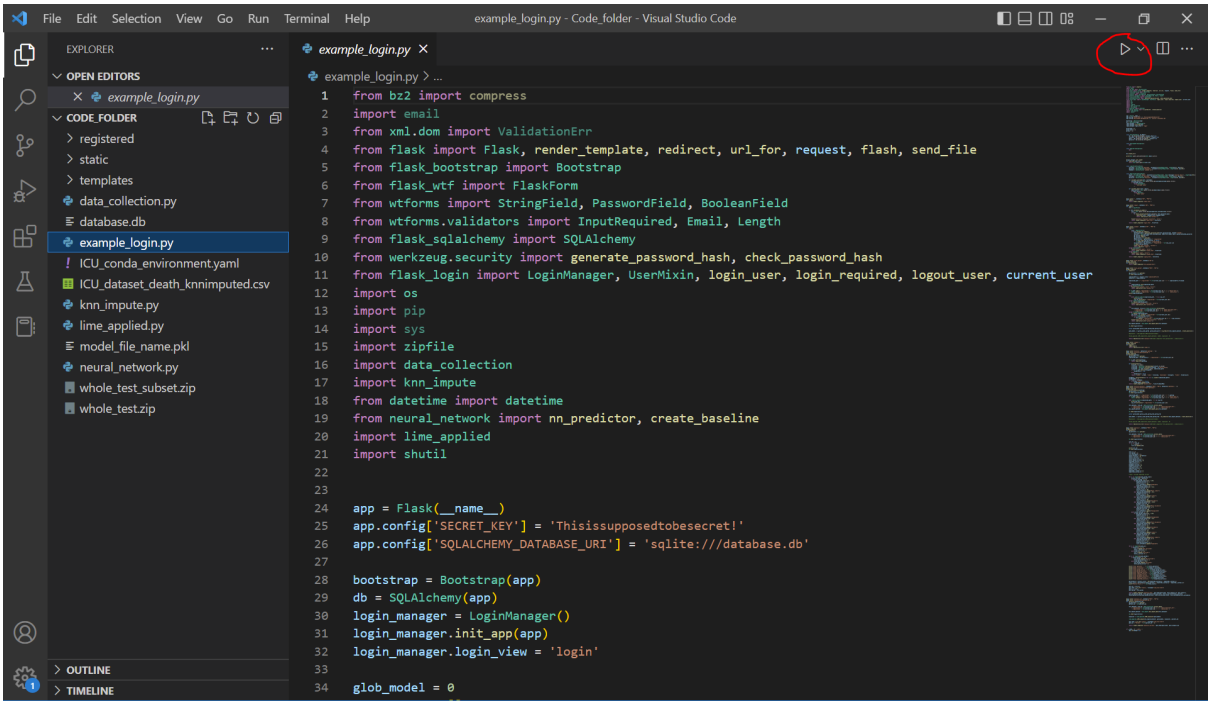


Figure 32. Select the run button

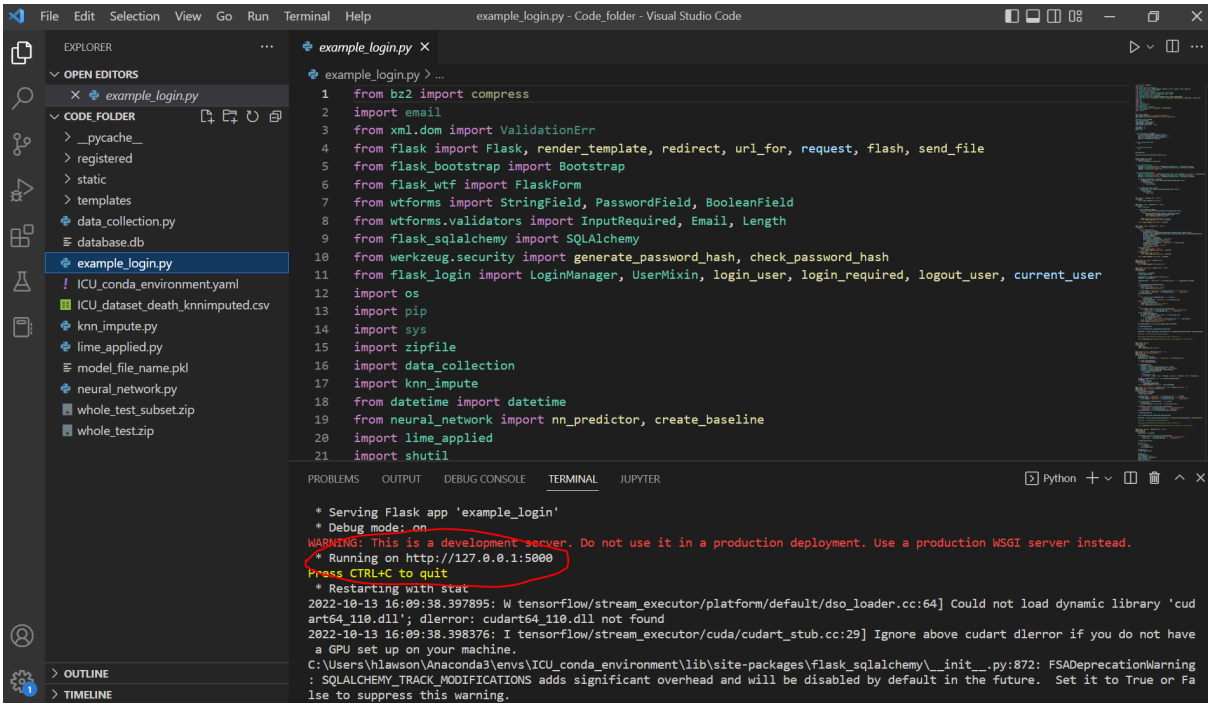


Figure 33. Copy link circled and paste in web browser

4.1 Manage database

The database is currently handled by manually deleting users and their information from the database. This is done by uncommenting out a block of code that is commented out in the example_login.py file (Figure 34) and choosing which user to delete by setting remove_user_id to the id of the user that is being deleted (Shown in Figure 35). This must be followed up by deleting that users folder in the registered folder (Figure 36). The folders in the registered folder are for each user in the system, and are denoted by that user's id.

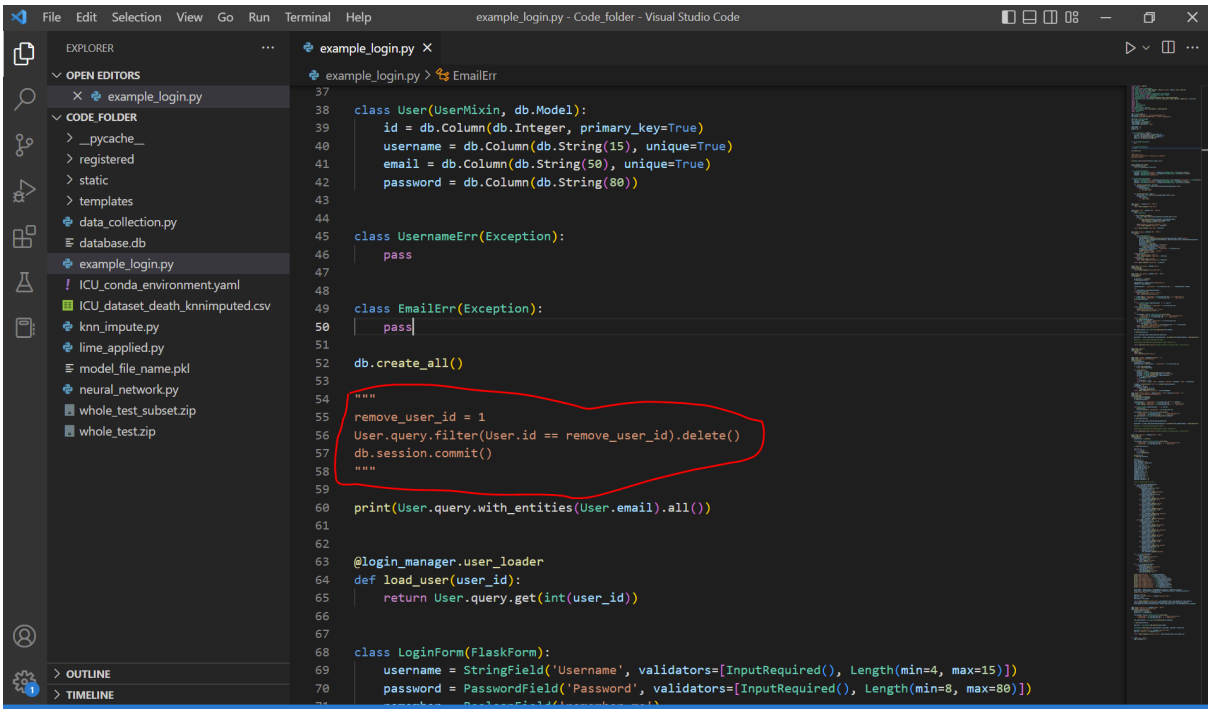


Figure 34. Delete user from database by uncommenting this code.

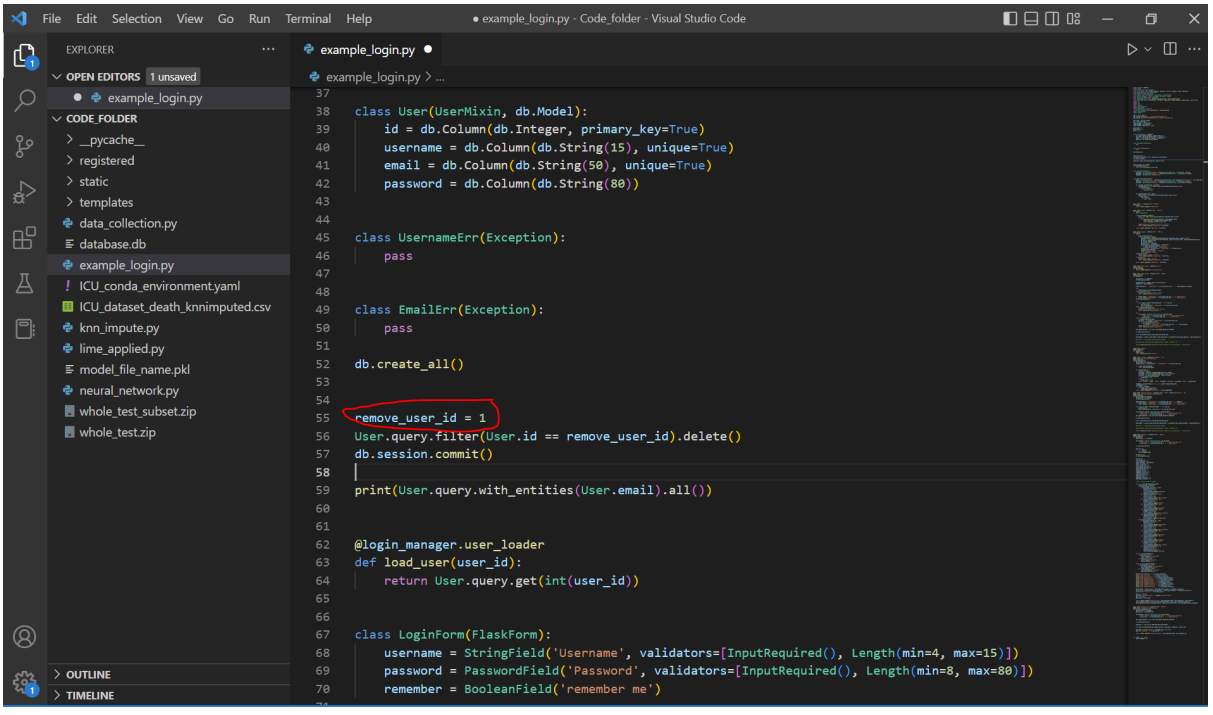


Figure 35. Choose the user to remove

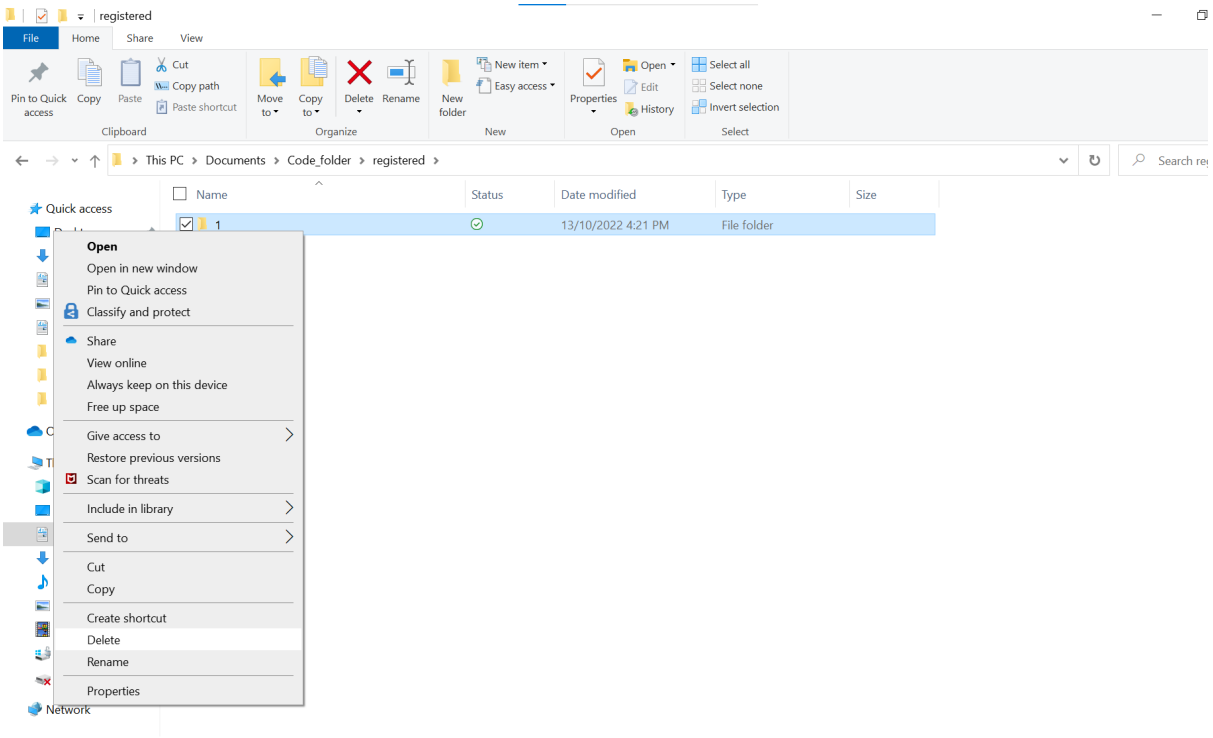


Figure 36. Delete user folder of documents