

# Web Application User's Manual

## (Version 1.0)

Web Application Documentation (last updated: July 18, 2025)

Washington State University, Richland, WA 99354, USA

For more information, please contact Suraiya Akhter (Email: [suraiya.akhter@wsu.edu](mailto:suraiya.akhter@wsu.edu))

Please follow the user manual for using our web application. If you find our web application useful, please cite the following papers.

1. Akhter, S. and Miller, J.H., 2025. Evaluating Feature Selection Methods and Feature Contributions for Cardiovascular Disease Risk Prediction. medRxiv, pp.2025-07.

Our application will automatically generate the required features of the testing data. Users can test multiple data simultaneously and augment new sequences to the training data to boost the predictive ability of the machine learning models through the web application. The flowchart of our web application is given below.

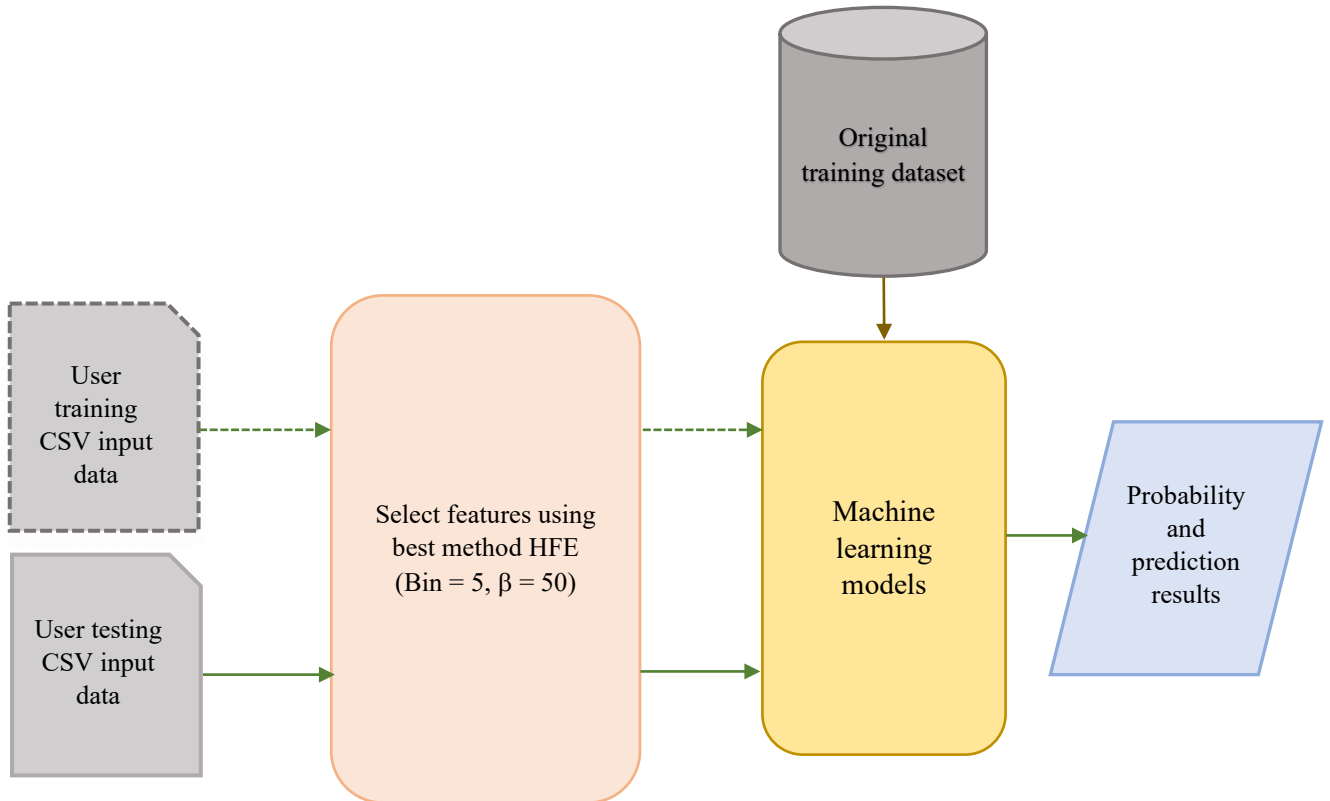


Fig 1: Workflow.

### Home Page:

Fig 2. shows the home page of our web application.

## CVDRP: A Web Application for Cardiovascular Disease Risk Prediction

Welcome Predicted CVDR Probability Scores Help Data

### Upload Files

This web application predicts cardiovascular disease risk from an input CSV file. The CSV file must include the following columns: SEQN, BPQ101D, RIDAGEYR, BPQ020, BPQ080, RXQ033, SMQ020, LBXTC, DIQ010, RIAGENDR, DMDEDUC2, LBXRDI, BPKOSY1, INDFMPIR, LBXP1TS1, LBDHDD, LBXGH, and CVD\_risk. If any of these columns are missing, the application will return an error. A sample input file is provided below. The user manual can be found under the Help menu. To predict cardiovascular disease risk, please upload a CSV file using the data fields listed above. If you wish to add new data to the model (training set), please use the 'Add new cardiovascular diseases risk data to training (CSV)' upload box for cardiovascular and the 'Add new non-cardiovascular diseases risk data to training (CSV)' upload box for non-cardiovascular, respectively. After uploading the necessary file, click the 'CVDR PREDICTION' button. Once the classification results are generated, you will be automatically redirected to the 'Predicted CVDR' page. Then, click the 'PROBABILITY ESTIMATION' button to view the probability scores on the 'Probability Score' page.

### Download Results

The prediction and probability results can be downloaded by clicking the 'PREDICTION RESULTS' and 'PROBABILITY RESULTS' buttons, respectively. Additionally, the training/testing datasets are available in the 'Data' menu.

### CSV Formatting

An example CSV file can be obtained by clicking on the 'Download Input Samples' button. To predict new data, the example CSV file should be in the form shown below:

DOWNLOAD INPUT SAMPLES

If you find our web application useful, please cite the following paper.  
Akhter, S. and Miller, J.H., 2025. Evaluating Feature Selection Methods and Feature Contributions for Cardiovascular Disease Risk Prediction. medRxiv, pp.2025-07.

Fig 2. home page

### Cardiovascular Disease Risk Detection:

To identify cardiovascular and non-cardiovascular disease risk, users need to follow below steps:

1. First, users can upload the input test data file (CSV file) by clicking on the Browse button under "Choose an input CSV file." For example, in Fig. 3, the "input\_sample.fasta" file is uploaded. During the upload of the input CSV file, all action buttons (CVDR PREDICTION, PROBABILITY ESTIMATION, PREDICTION RESULTS, and PROBABILITY RESULTS) are inactive. After completing the upload of data, they will become active and ready to use.

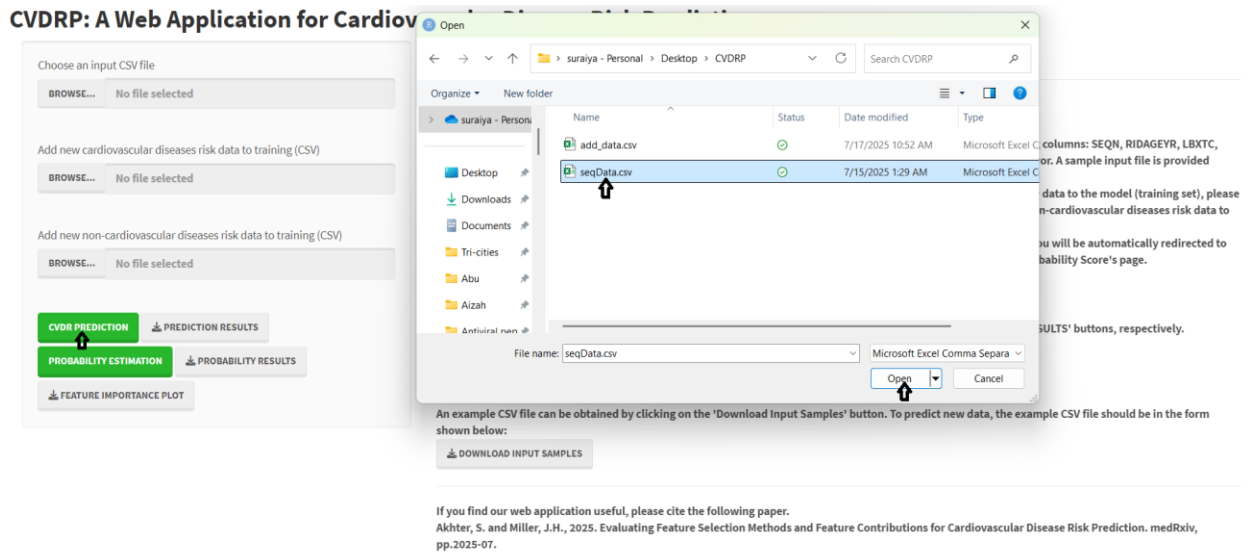


Fig 3: File upload

- Once the necessary files are uploaded, please click the ' CVDR PREDICTION ' button first. This will automatically navigate you to the 'Predicted CVDR' page once the classification results are generated. Refer to the example in Fig. 4 below.

### CVDRP: A Web Application for Cardiovascular Disease Risk Prediction

Choose an input CSV file

BROWSE... seqData.csv

**Upload complete**

Add new cardiovascular diseases risk data to training (CSV)

BROWSE... No file selected

Add new non-cardiovascular diseases risk data to training (CSV)

BROWSE... No file selected

**CVDR PREDICTION**    PREDICTION RESULTS

**PROBABILITY ESTIMATION**    PROBABILITY RESULTS

FEATURE IMPORTANCE PLOT

Welcome	Predicted CVDR	Probability Scores	Help	Data
Total predicted data...60				
1	130970			
2	131452			
3	132095			
4	132150			
5	132152			
6	132928			
7	133123			
8	133167			
9	133192			
10	133299			
11	133315			
12	133690			
13	133843			
14	134073			
15	134241			
16	134528			
17	134822			
18	135022			
19	135087			
20	135130			
21	135536			
22	135582			
23	135807			
24	136072			
25	136107			
26	136113			
27	136129			
28	136200			
29	136321			
30	136334			
31	136724			
32	136897			
33	136968			
34	137022			

Fig 4: CVDR prediction

- Next, click the PROBABILITY ESTIMATION button to navigate to the 'Probability Scores' page containing the probability values of the sequences. Refer to the example in Fig. 5 below.

### CVDRP: A Web Application for Cardiovascular Disease Risk Prediction

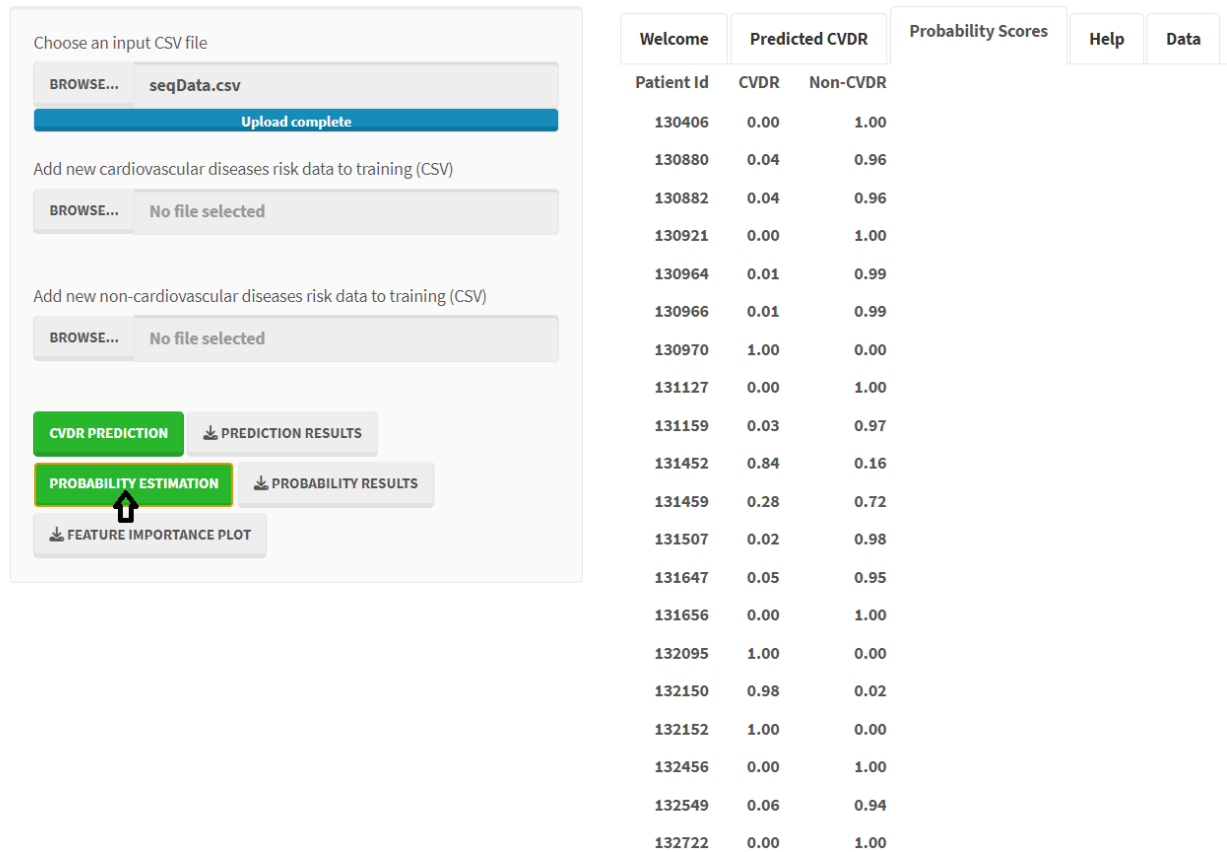


Fig 5: Probability estimation

- SHAP Analysis:  
Our web application generates box plot based on the best method (bin=10 and  $\beta = 25$ ). See the following plot for the best model .

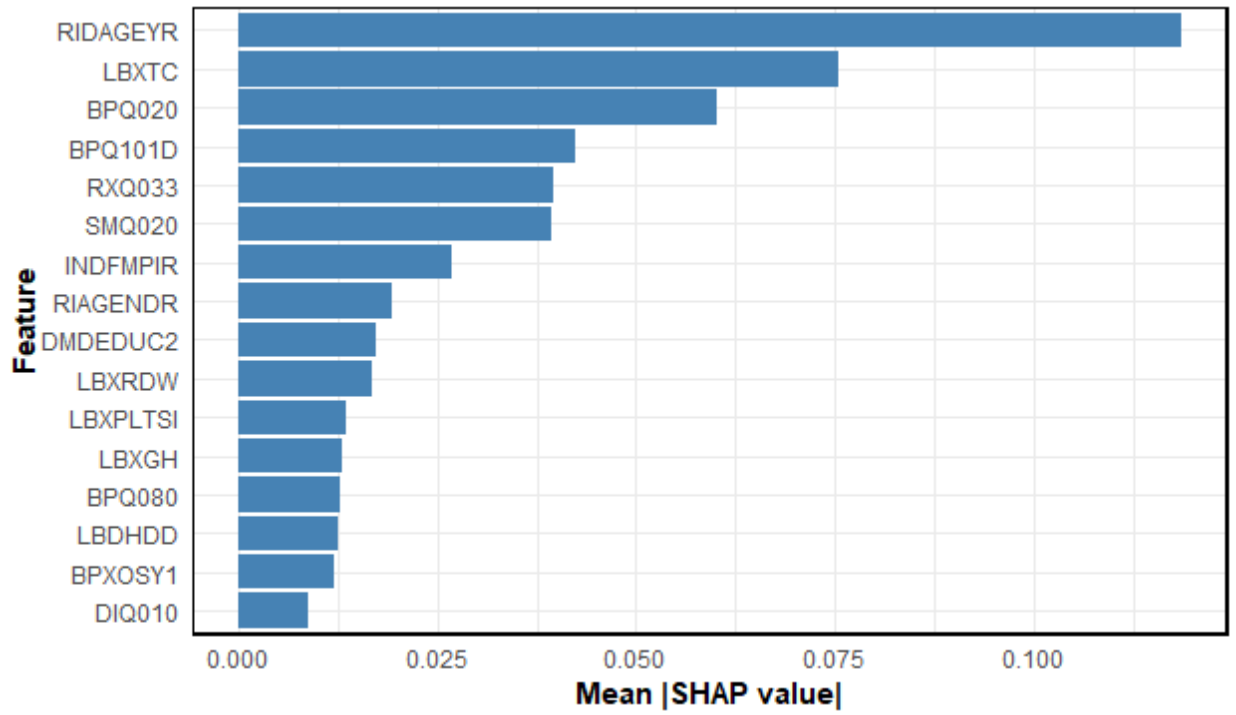


Fig 6: Features importance plot