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)

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.

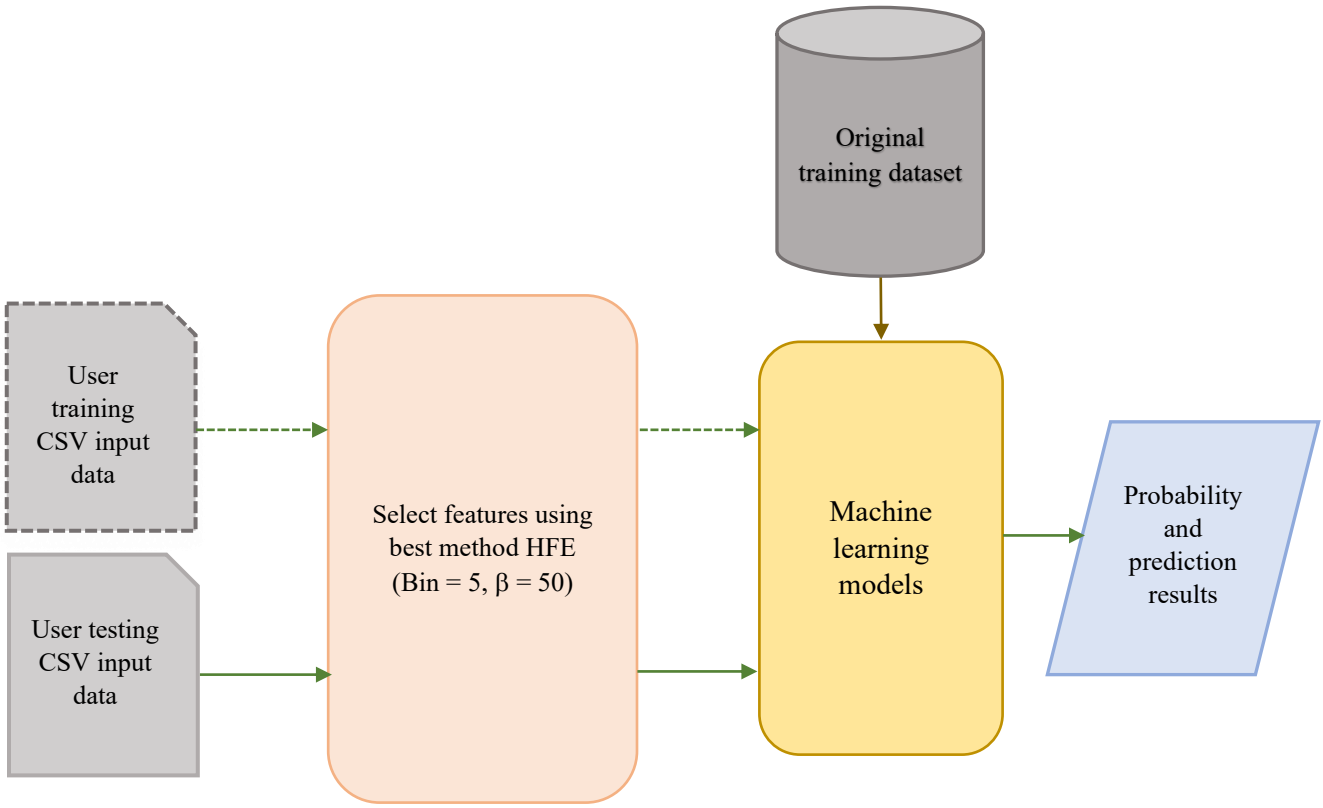

Fig 1: Workflow.

Home Page:

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

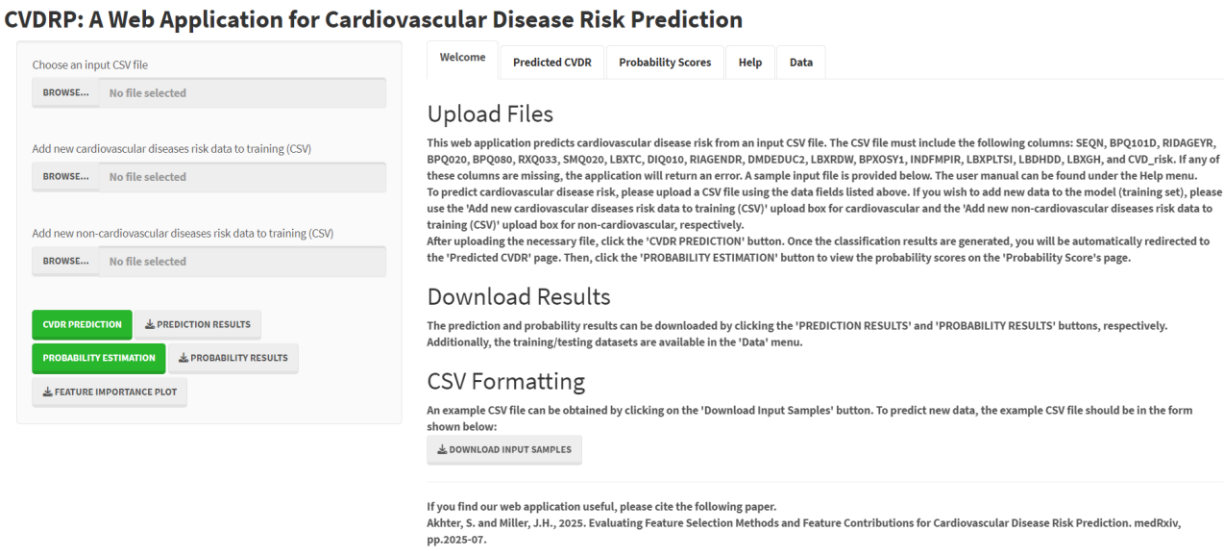

Fig 2. home page

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.

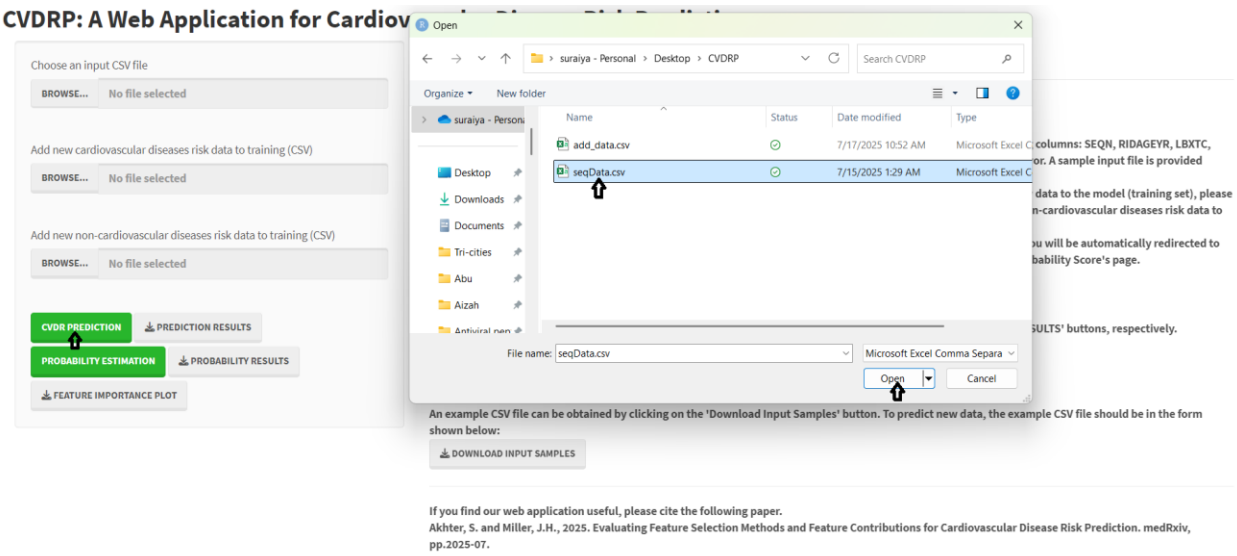

Fig 3: File upload

2. 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.

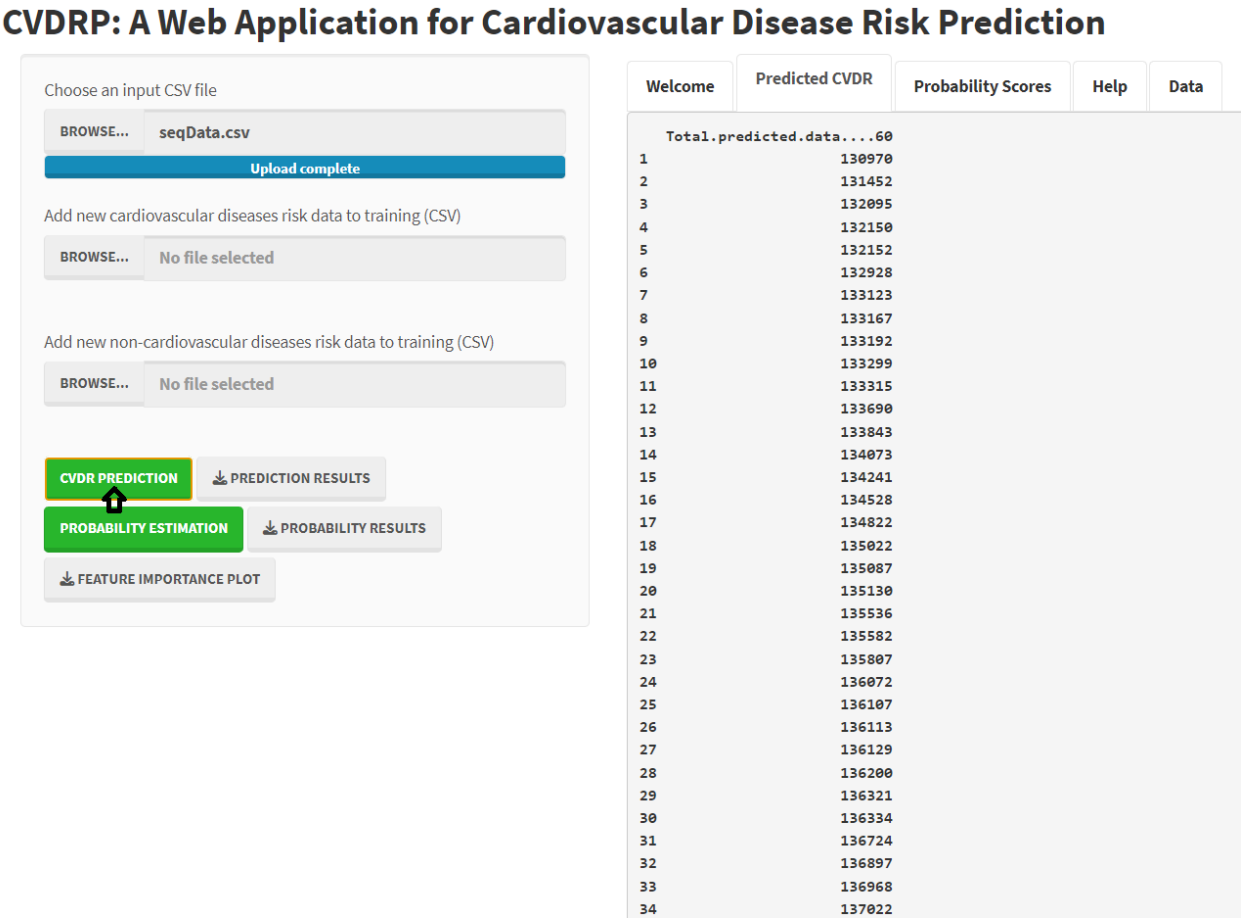

Fig 4: CVDR prediction

3. 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.

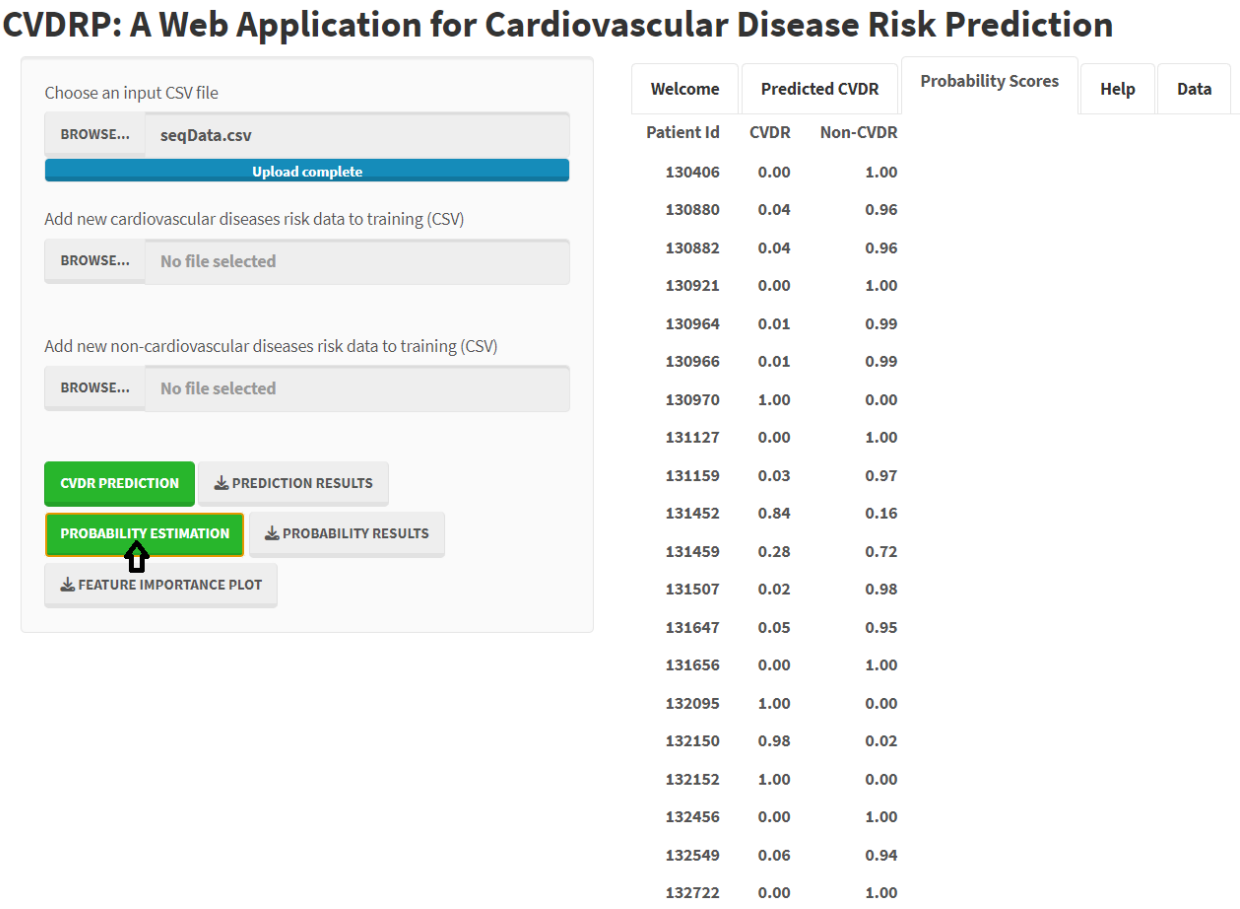

Fig 5: Probability estimation

4. 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 .

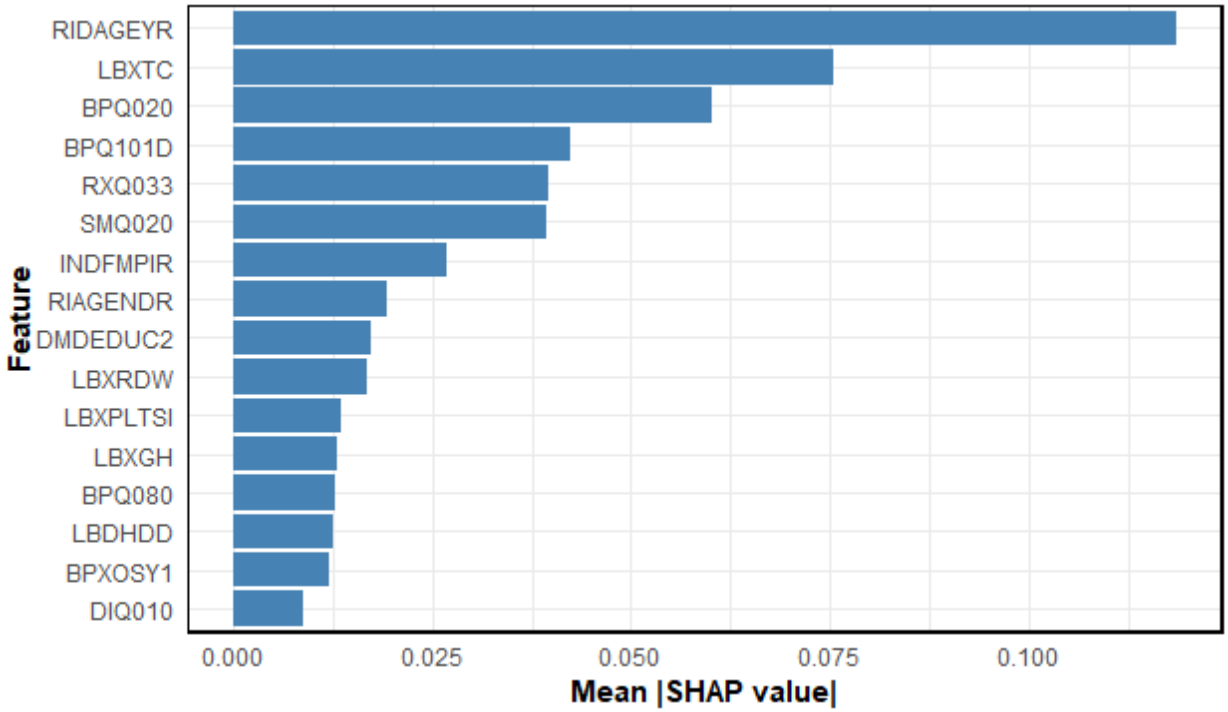

Fig 6: Features importance plot