**LaRFScv 0.1**

LaRFScv (*Lasso Regression-based Feature Selection with Cross-Validation*) is a Shiny web application designed for performing Lasso Regression with Cross-validation (Lasso-CV). Lasso regression is a regression analysis method that performs both variable selection and regularization to improve the predictive accuracy and interpretability of the statistical model.

This application allows users to upload a CSV dataset, choose a dependent variable, and perform Lasso regression with cross-validation to identify the optimal regularization parameter (lambda) and select significant features. It provides visualizations and a downloadable CSV file of the selected features for further analysis.

**How to Run the Software**

1. **Download LaRFScv**:
   * Visit the LaRFScv GitHub repository.
   * Click on the **Releases** tab to find the latest release version.
   * Download the **LaRFScv 0.1** release package (usually a ZIP file).
2. **Extract the Files**:
   * Extract the downloaded ZIP file to a location of your choice.
3. **Run LaRFScv**:
   * Open the extracted folder.
   * Locate the **LaRFScv.bat** file.
   * Double-click on **LaRFScv.bat** to launch the portable standalone app.
   * Alternatively, if you prefer to use the default browser, run **LaRFScv\_browser.bat**.

**Features**

* **Upload CSV File**: Users can upload a CSV file containing their dataset.
* **Select Dependent Variable**: Users can choose the dependent variable from the dataset.
* **Cross-validation**: Users can specify the number of folds for cross-validation.
* **Coefficient Cut-off**: Users can set a cut-off value to select significant features based on the coefficient magnitude.
* **Seed Value**: Users can set a seed value for reproducibility.
* **Analyze Button**: Triggers the Lasso regression analysis.
* **Download CSV**: Allows users to download a CSV file containing the selected features.
* **MSE vs Log Lambda Plot**: Displays the Mean Squared Error (MSE) against Log Lambda values.
* **Selected Features vs Coefficient Plot**: Presents the selected features and their coefficients.
* **Selected Features Table**: Shows the list of selected features and their corresponding coefficients.
* **Error Message Box**: Displays an error message if there is an issue with reading the CSV file.

**Dependencies**

* [shiny](https://shiny.posit.co/): R package for web application development.
* [glmnet](https://cran.r-project.org/web/packages/glmnet/index.html): R package for Lasso and elastic-net regularized generalized linear models.
* [ggplot2](https://ggplot2.tidyverse.org/): R package for data visualization.
* [data.table](https://cran.r-project.org/web/packages/data.table/): R package for data manipulation
* [dplyr](https://dplyr.tidyverse.org/): R package for grammar of data manipulation

**Author**

SUVANKAR BANERJEE

**Version**

0.1

**License**

This project is licensed under the GPL v3.0 License.

**Compatibility**

Windows systems (Windows 7 and above)

**Note**

If you use this tool in your work, cite the following article:

*Banerjee, S., Jana, S., Jha, T., Ghosh, B., & Adhikari, N. (2024). An assessment of crucial structural contributors of HDAC6 inhibitors through fragment-based non-linear pattern recognition and molecular dynamics simulation approaches.* ***Computational biology and chemistry****, 110, 108051.* <https://doi.org/10.1016/j.compbiolchem.2024.108051>.

**Acknowledgments**

A special thanks to all the supporters who have made this tool possible. Your inspiration and support are invaluable.