

# Manual for Fiore\_v1.0

## Fiore\_v1.0

**Author:** Sk. Abdul Amin, M. Pharm., Ph.D.

**Tool Name:** Fiore\_v1.0 Tool

**Version:** 1.0 (Date: 18.04.2025)



### Introduction

**Fiore\_v1.0 Tool** is a complete package of comprehensive Python-based application designed to streamline the process of feature calculation, selection, and model development for machine learning tasks. It provides a user-friendly graphical interface built with customtkinter and integrates powerful libraries such as pandas, numpy, matplotlib, and scikit-learn.

It allows users to load training and testing sets, choose model parameters, tune the model using GridSearchCV or Optuna, and run 11 types of machine learning classification models such as Support Vector Classifier (SVC), Decision Tree Classifier, MLP Classifier, GaussianNB, Logistic Regression, KNeighbors Classifier, Bagging Classifier, Random Forest Classifier, Ada Boost Classifier, Gradient Boosting Classifier, XGB Classifier.

The tool also allows users to visualize results and save them to a specified directory.

### Fiore v1.0 Tool has three components

#### 1. **Fiore\_FC.py: Feature Calculation and Data Preparation**

**Purpose:** Calculate features, classify data based on median pIC50, and split into training and test sets.

#### 2. **Fiore\_FS.py: Feature Selection on Training Set**

**Purpose:** Perform feature selection using mutual information regression and remove highly correlated features.

#### 3. **Fiore\_MLC.py: Model Development and Validation**

**Purpose:** Develop and validate machine learning models with tuned hyperparameters using the selected features.

# Manual for Fiore\_v1.0

## Installation and Setup Guide: Fiore v1.0 Tool

### 1. Install Anaconda (run once only)

If you don't have Anaconda installed, follow these steps:

- Download Anaconda from the official website:  
<https://www.anaconda.com/products/distribution>
- Follow the on-screen instructions to complete the installation.

NOTE: *If Anaconda is already installed on your system, skip this step.*

---

### 2. Open Anaconda Prompt

- Open the **Anaconda Prompt** from the Start Menu or by searching for it in your system.

---

### 3. Create and Activate the Conda Environment

- **Step 3.1: Create the environment (run once only)**

`conda create -n Fiore_v1 python=3.9`

- **Step 3.2: Activate the environment**

`conda activate Fiore_v1`

---

### 4. Navigate to Your Working Directory

- Use the cd command to go to the directory where your project files are stored.

\*Example: `cd C:\Users\Amin\Downloads\Fiore_v1`

---

### 5. Check Required Files

Ensure that the following files are present in the **Fiore\_v1** folder:

- `requirements.txt` (Python libraries used in script)
- `Fiore_FC.py`, `Fiore_FS.py`, `Fiore_MLC.py` (the main application scripts)
- `Fiore_logo.jpg` (logo displayed in the GUI)

---

### 6. Install Required Packages (One-time setup only)

Install the required packages for running Fiore\_v1.0.

- **Option A – Using Pip (Recommended):**

`pip install -r requirements.txt`

---

### 7. Run the Tool

- Once everything is set up, run the tool with the following command:

`python Fiore_FC.py`

- This will launch the **Fiore\_v1.0 FC Tool** graphical user interface.

# Manual for Fiore\_v1.0

---

## Troubleshooting and Tips

### 1. Issue: conda or pip is not recognized

- Cause: The Anaconda installation path is not added to your system's environment variables.
- Solution: Reinstall Anaconda and ensure the checkbox "Add Anaconda to PATH" is selected during installation.
- Alternatively, use the Anaconda Prompt instead of your regular terminal or CMD.

### 2. Issue: ModuleNotFoundError when running Fiore\_v1.py

- Cause: Required packages are not installed in the active environment.
- Solution: Make sure you're in the correct environment:  
`conda activate Fiore_v1`
- Reinstall dependencies:

`pip install -r requirements.txt`

### 3. Issue: RDKit installation failed

- Solution: If pip installation fails, try installing RDKit via conda:

`conda install -c conda-forge rdkit`

### 4. Issue: GUI doesn't launch or closes instantly

- Cause: Errors in script or missing files.
- Solution: Make sure `Fiore_FC.py`, `Fiore_logo.jpg`, and `requirements.txt` are present in the same directory.
- Open a terminal, navigate to the project directory, and run:  
`python Fiore_v1.py`
- This way, you can read any error messages.

### 5. If the tool misbehaves or shows strange errors:

- Remove and recreate the environment:

`conda remove --name Fiore_v1 --all`

`conda create -n Fiore_v1 python=3.9`

`conda activate Fiore_v1`

`pip install -r requirements.txt`

### 7. Still having issues?

- Double-check Python version compatibility (`python --version`)
  - Restart the terminal or your system.
  - Ask for help or log errors to share with support.
-

# Manual for Fiore\_v1.0

## Details:

### 1. **Fiore\_FC.py: Feature Calculation and Data Preparation**

**Purpose:** Calculate features, classify data based on median pIC50, and split into training and test sets.

**Steps:**

1. **Load CSV File:** Click the "Load CSV File" button to select your dataset ([Example\\_dataset.csv](#)).
  2. **Select Columns:**
    - **ID Column:** Unique identifier for each sample.
    - **Smile Column:** Contains SMILES strings.
    - **Choose "pIC50" or "Class" from the "Select Column Type" dropdown**
  3. **Choose Split Method:**
    - **Random:** Random 80/20 split.
    - **KMCA:** Uses PCA + KMeans for cluster-aware stratification
  4. **Run Feature Calculation:** Click the "Run Feature Calculation" button to process the data.
  5. **Select Output Directory to save the results.**
- 

### 2. **Fiore\_FS.py: Feature Selection on Training Set**

**Purpose:** Perform feature selection using mutual information regression and remove highly correlated features.

**Process:**

1. **Load CSV File:** Click the "Load CSV File" (example: [train\\_data.csv](#)).
  2. **Select Columns:**
    - a. **ID Column:** Unique identifier for each sample.
    - b. **Target (Class) Column:** Dependent variable (e.g., [Class](#)).
    - c. **Start Feature Column:** First feature column to include.
    - d. **End Feature Column:** Last feature column to include (can select 'last', 'second last', or 'third last').
  3. **Set Thresholds:**
    - a. **Importance Threshold:** Minimum mutual information score to consider a feature important (default: 0.15).
    - b. **Correlation Threshold:** Maximum allowed correlation between features (default: 0.9).
  4. **Run Feature Selection:** Click the "Run Feature Selection" button to process the data.
  5. **Save the Train set file with selected features** (example: [train\\_data\\_FS.csv](#)).
- 

### 3. **Fiore\_MLC.py: Model Development and Validation**

**Purpose:** Develop and validate machine learning models with tuned hyperparameters using the selected features.

---



For further support or bug reports, please contact the author.

**DR. SK. ABDUL AMIN, M. Pharm., Ph. D.**

<https://orcid.org/0000-0003-4799-7322>

Dipartimento di Farmacia/DIFARMA

[Università degli Studi di Salerno, ITALY](#)

Email: [pharmacist.amin@gmail.com](mailto:pharmacist.amin@gmail.com)

