SDF Merger: A Desktop and Web Tool for Merging Chemical Structure Data Files

Md. Ismiel Hossen Abir
Independent Researcher
ismielabir286@gmail.com

Summary

The SDF Merger is a cross-platform, simple tool designed to efficiently merge Structure-Data File (SDF) files. SDF files are widely used in cheminformatics, computational chemistry, and drug discovery research. SDF files store molecular structures with their properties. Researchers used SDF files for diverse workflows such as virtual screening, molecular docking, and high-throughput compound testing. For molecular docking needs, used large number of ligand files which can be SDF files. Managing this large number of SDF files manually is time-consuming and chance to errors. Therefore, creating an SDF merger tool is required for a reliable and accessible use. This tool provides two interfaces such as a desktop application and a webpage interface. The desktop application, built using Python, Tkinter, and RDKit. These support research-intensive workflows, multiple processing of SDF files, and detailed logging. The web application, implemented in pure JavaScript, HTML, and CSS that requires no installation and processes files entirely client-side, ensuring data privacy while providing immediate, drag-and-drop accessibility. Both the desktop and web interface required no installation. The web interface is already live on the GitHub page, and the desktop application as an .exe file is available for download from the SourceForge site. Users can use any of the platforms as per their needs and requirements. The aim of this tool is to solve the complexity of merging SDF file for research purposes. By bridging the gap, SDF Merger uses SDF file management, making it accessible to researchers without programming expertise. Its cross-platform availability, such as a Windows executable and a browser-based web interface, ensures broad usability in academic and educational settings. The tool tried to compute research workflows and reduce technical barriers for users across diverse research environments.

Statement of Need

SDF files are the standard for exchanging chemical structure information in computational chemistry and drug discovery. Typical research workflows involve processing hundreds or thousands of SDF files from various sources, such as compound libraries, virtual screening results, or molecular dynamics simulations. Therefore, merging these files is essential to efficient the workflow in these fields education and research perspective. Existing solutions either require programming expertise or are complex to make a merged SDF file. There is a clear need for a simple, accessible tool that bridges this

gap. SDF Merger addresses these challenges by providing accessibility, open source, and cross-platform availability. Here, no programming skills are required for basic merging operations. The tool is available as both Windows executable and a browser-based web interface. Web version processes files entirely client-side.

Features and Tool Description

This tool is available in both a webpage interface and a desktop application. Desktop application features, including graphical user interface which build using rdkit toolkit for cheminformatics [1], several python libraries such as tkinter, threading, os, etc. [2]. Here, graphical user interface built with Python's Tkinter, offering intuitive file selection and processing controls. This tool also uses batch processing, which can automatically detect and process all .sdf files in a selected directory. Then, for process monitoring, it uses real-time progress bar and detailed logging of merge operations. After that, for quick assess the tool build webpage of this too where user can directly merge multiple SDF file into one SDF file without download and installations. In the web implementation, all file operations occur locally in the user's browser. Users can drag and drop or browse files to upload, where users can upload multiple files at once or a zip file. It is suitable for use both desktops and mobile devices. One of the benefit of this tool is no data uploads to external servers. The web application is implemented in vanilla JavaScript, HTML, and CSS, utilizing the File API for client-side file operations. It performs basic SDF parsing and validation without external dependencies. After that, user can immediate download the one merge file and that can use for research purpose.

Availability and Use Case

All source code and webpage are available in Github.
Source Repository: https://github.com/IsmielAbir/SDF-File-Merge
Webpage: https://ismielabir.github.io/SDF-File-Merge/
For desktop application, it is available in SourceForge and code available open source both available in Github repository and SourceForge site.
Desktop Application: https://sourceforge.net/projects/sdf-merger/ For installations and use of desktop application, users need to directly download the .exe file from SourceForge site. After that, they can directly use the tool by double click to open the tool. Then select input folder path and output location. After that, click merge button and it will automatically merge and save in the output location. For the webpage need to go Github repository and open the webpage. Then, simply upload files, merge, and download. The use case of these tools is for both educational and research perspectives.

License

SDF Merger is distributed under the MIT license, allowing free use, modification, and distribution for both academic and commercial purposes.

References

[1] Greg Landrum et al. Rdkit: Open-source cheminformatics, 2006.

[2] Python Software Foundation. Python (version 3.9.6) [programming language]. `https://www.python.org/`, 2021.